Algebra pre informatikov

Martin Sleziak

13. februára 2024

Obsah

Kapitola 1

Úvod

Verzia: 13. februára 2024

Tento text je zamýšľaný ako pomocný text k predmetom Algebra (1), Algebra (2) a Algebra (3) na odbore informatika. (Samozrejme, možno môže byť – či už celý alebo niektoré časti – zaujímavý aj pre človeka, ktorý nemá zapísané tieto predmety.) Určite to nie je tak, že by text bol úplne totožný s tým, čo odzneje na prednáške – ale snažíme sa dodržiavať podobné označenie v texte aj na prednáškach a cvičeniach. Na rozdiel od prednášky, písaný text v podstate nie je limitovaný rozsahom na počet hodín – takže sa sem dali zaradiť aj niektoré veci navyše.

Treba určite počítať s tým, že text je vo vývoji a bude sa postupne upravovať.

TODO prvá časť – hlavne lineárna algebra

TODO vymenovať aplikácie

TODO prečo všeobecne (nad ľubovoľným poľom)

1.1 Sylaby

1.1.1 1-INF-115 Algebra (1)

Základné pojmy potrebné k abstraktnému vybudovaniu vektorových priestorov (grupy, polia, vektorové priestory). Podpriestory, lineárna závislosť a nezávislosť vektorov, Steinitzova veta, báza vektorového priestoru. Matice. Lineárne zobrazenia. Kompozícia lineárnych zobrazení, inverzné matice. Riešenia homogénnych a nehomogénnych systémov lineárnych rovníc. Determinanty, základné vlastnosti a aplikácie.

1.1.2 1-INF-156 Algebra (2)

Skalárny súčin, ortonormálna báza a ortogonálna projekcia na podpriestor. Kvadratické formy a ich kanonické tvary. Pozitívna (semi)definitnosť matice a kvadratickej formy a kritériá na overenie pozitívnej definitnosti. Zmena bázy, podobné matice. Podobnosť matice s diagonálnou maticou. Vlastné čísla a vlastné vektory, charakteristický polynóm. Ortogonálne matice, ortogonálna podobnosť, Schurova veta a veta o hlavných osiach. Okruhy polynómov, rozklad polynómov na ireducibilné polynómy, (viacnásobné) korene polynómov, derivácia a Taylorov rozvoj polynómov.

1.1.3 2-INF-182 Algebra (3)

Grupy, podgrupy, homomorfizmy, faktorové grupy. Okruhy, ideály, maximálne ideály a prvoideály, vzťah k poliam a oborom integrity pri faktorizácii. Euklidovské okruhy, okruhy hlavných ideálov, gaussovské okruhy. Teória deliteľnosti a veta o rozklade na ireducibilné prvky. Rozšírenia polí. Konečné polia, klasifikácia konečných polí. Niektoré aplikácie rozšírení polí a konečných polí. Použitie rýchlej Fourierovej transformácie pri násobení veľkých čísel. (Výber tém v danom semestri sa môže upraviť na základe záujmu študentov.)

1.1.4 1-INF-156 Algebra (2) – v akreditácii do školského roku 2021/22

Grupy, podgrupy, homomorfizmy, faktorové grupy. Okruhy, ideály, maximálne ideály a prvoideály, vzťah k poliam a oborom integrity pri faktorizácii. Euklidovské okruhy, okruhy hlavných ideálov, gausovské okruhy. Teória deliteľnosti a veta o rozklade na ireducibilné prvky. Okruhy polynómov, rozklad polynómov na ireducibilné polynómy, (viacnásobné) korene polynómov, derivácia a Taylorov rozvoj polynómov. Rozšírenia polí. Riešenie antických problémov (duplicita kocky, trisekcia uhla, kvadratúra kruhu). Konečné polia, klasifikácia konečných polí, šifrovanie RSA.

1.1.5 2-INF-182 Algebra (3) – v akreditácii do školského roku 2021/22

Skalárny súčin, ortonormálna báza a ortogonálna projekcia na podpriestor. Kvadratické formy a ich kanonické tvary. Pozitívna (semi)definitnosť matice a kvadratickej formy a kritériá na overenie pozitívnej definitnosti. Zmena bázy, podobné matice. Podobnosť matice s diagonálnou maticou. Vlastné čísla a vlastné vektory, charakteristický polynóm. Ortogonálne matice, ortogonálna podobnosť, Schurova veta a veta o hlavných osiach. Symetrické polynómy. Použitie rýchlej Fourierovej transformácie pri násobení veľkých čísel. PageRank algoritmus.

Kapitola 2

Množiny a zobrazenia

Na vysokej škole sa stretnete s trochu iným prístupom k matematike ako doteraz. Pre modernú matematiku je typický axiomatický prístup, ktorý spočíva v tom, že vychádzame z niektorých pojmov, ktoré nedefinujeme (chápeme ich ako základné, zvyknú sa nazývať *primitívne pojmy*). Okrem nich zavedieme niekoľko axióm, ktoré hovoria o ich základných vlastnostiach. Všetky ďalšie výsledky musia byť odvodené len z týchto axióm, všetky ďalšie pojmy sú definované len pomocou primitívnych pojmov.

Prvou matematickou knihou, v ktorej sa používali axiómy, boli Euklidove Základy.[1] Priekopníkom používania axiomatickej metódy v modernej matematike bol David Hilbert.[2] V súčasnosti sa za základ matematiky, na základe ktorého sa dajú sformalizovať všetky študované disciplíny, považuje teória množín.

Jedným z cieľov snahy zachytiť matematiku ako celok pomocou axióm (tzv. Hilbertov program) bola snaha o dôkaz bezospornosti matematiky. Dnes je známe, že tento cieľ sa nedá naplniť v takom rozsahu, ako si to predstavoval Hilbert. (Viac sa o tom dá dozvedieť napríklad v knihe [Z2], ktorá však vyžaduje aspoň základné znalosti z teórie množín.) Napriek tomu mal však Hilbertov program obrovský vplyv na podobu modernej matematiky.

Axiomatický prístup má svoje výhody aj nevýhody. Formalizácia prináša výhodu v tom, že sa dôkazy dajú ľahšie skontrolovať (dokonca sa dajú, hoci pomerne pracne, prepísať do takej podoby, aby boli skontrolovateľné počítačovým programom). Takisto ak zadefinujeme nejaký pojem pomocou nejakého systému axióm a z týchto axióm dokážeme nejaké tvrdenia, vieme, že tieto tvrdenia platia pre každý matematický objekt, ktorý spĺňa tieto axiómy. Tento princíp bude v rámci tejto prednášky často používaný.

Za nevýhodu možno považovať to, že formalizáciou sa môže do istej miery stratiť intuitívne porozumenie pojmom s ktorými pracujeme. Preto je dôležité dbať na obe stránky – nielen naučiť sa pracovať s formalizmom, ktorý budeme používať, ale tvrdeniam a dôkazom aj rozumieť.

Zápis tvrdení i definícií nových pojmov bude o čosi formálnejší, než ste boli zvyknutí doteraz. Hoci matematický jazyk, ktorý sa používa v dôkazoch, môže byť pre vás nový a trochu neobvyklý, je veľmi dôležité, aby ste sa ho naučili používať, a ešte dôležitejšie aby ste tomuto jazyku a hlavne dôkazom, ktoré budeme robiť, aj porozumeli.

[1] Euklides (3.–2. stor. pr.n.l) bol grécky matematik a geometer. Jeho najvýznamnejšie dielo *Základy* sa počtom vydaní radí na druhé miesto medzi všetkými knihami v histórii, hneď po Biblii.

[2] David Hilbert (1862–1943) bol nemecký matematik. Je považovaný za jedného z najvýznamnejších matematikov v svojom období ale aj v celej histórii matematiky.

2.1 Dôkazy

y:SECTDOKAZY}

Z toho, čo sme povedali o axiomatickom prístupe je zrejmé, že všetky tvrdenia, ktoré budeme používať, sa budeme snažiť aj dokázať. Preto je užitočné povedať si pár slov o dôkazoch. Veľmi dobrý úvodný text o tejto problematike je [KGGS, Kapitola 1.1]. Knihy [L] a [HS] sú venované rôznym postupom používaným pri dôkazoch matematických tvrdení, môžete tam nájsť (okrem iného) aj samostatnú kapitolu venovanú matematickej indukcii.

Ešte prv ako sa začneme venovať formálnej stránke dôkazov, mohli by sme sa zamyslieť nad tým, načo vlastne dokazujeme. Dôkaz poskytuje overenie správnosti tvrdenia – hoci sa vám niektoré tvrdenia môžu zdať intuitívne zrejmé, až podrobný dôkaz nám dá istotu, že je skutočne správne. Pre človeka, ktorý sa venuje matematike, je teda prirodzená potreba vidieť aj dôkaz vysloveného tvrdenia. Považoval by som za úspech tejto prednášky, keby na konci semestra boli pre vás matematické dôkazy väčšmi prostriedkom na overenie, či platí matematická veta, ktorá vás zaujíma, než nutným zlom, ktoré musíte zvládnuť kvôli úspešnému absolvovaniu skúšky.

2.1.1 Základné typy dôkazov

Hoci spôsob, ako dokazovať matematické tvrdenia, sa najlepšie naučíte na konkrétnych príkladoch, predsa len na tomto mieste zhrnieme terminológiu a ukážeme si na veľmi jednoduchých príkladoch niektoré základné typy dôkazov.

Najjednoduchším typom dôkazu je *priamy dôkaz.* Jednoducho postupujeme z predpokladov tvrdenia, až kým sa nám nepodarí dostať dokazovaný výrok. Vyskúšajme si to na príklade nasledujúceho tvrdenia.

{dokazy:TVRZVYSOK}

Tvrdenie 2.1.1. *Nech n je celé číslo. Potom zvyšok čísla n^2 po delení 4 je 0 alebo 1. Pritom ak n je párne, tak n^2 je deliteľné 4 a ak n je nepárne tak n^2 má zvyšok 1.*

Dôkaz. Sú 2 možnosti. Buď n je párne, alebo n je nepárne.

Ak n je párne, tak $n = 2k$ (pre nejaké celé číslo k) a $n^2 = 4k^2$, čiže n^2 má zvyšok 0 po delení 4.

Ak n je nepárne, tak $n = 2k + 1$, čiže $n^2 = 4k^2 + 4k + 1 = 4(k^2 + k) + 1$ má zvyšok 1 po delení 4. □

Iný typ dôkazu je *dôkaz sporom.* Pri tomto type dôkazu začneme s predpokladom, že dokazované tvrdenie neplatí. Ak sa nám z tohoto tvrdenia podarí odvodiť niečo, čo určite platiť nemôže, musí byť predpoklad o neplatnosti dokazovaného tvrdenia chybný – a tým sme tvrdenie vlastne dokázali.

Tvrdenie 2.1.2. *Ak pre celé čísla a, b, c platí rovnosť $a^2 + b^2 = c^2$, tak aspoň jedno z čísel a, b je párne.*

Dôkaz. Sporom. Predpokladajme, že by a aj b boli nepárne. Potom podľa tvrdenia 2.1.1 majú a^2 aj b^2 zvyšok 1 po delení 4. Ich súčet $c^2 = a^2 + b^2$ má potom zvyšok 2. Podľa tvrdenia 2.1.1 sú však druhá mocnina môže mať ako zvyšok po delení 4 iba číslo 0 alebo 1 – dostali sme spor. □

Nepriamy dôkaz sa dosť podobá na dôkaz sporom. Väčšina tvrdení, ktoré dokazujeme majú tvar implikácie: snažíme sa dokázať, že ak platia predpoklady P, tak platí aj záver Z. Nepriamy dôkaz spočíva v tom, že namiesto toho dokazujeme: Ak neplatí záver Z, tak neplatí ani predpoklad P. (Skúste si rozmyslieť, prečo je to to isté ako dokazovať pôvodnú implikáciu. K implikáciam aj k princípu nepriameho dôkazu sa ešte vrátime v časti 2.1.4 venovanej tautológiám.)

Tvrdenie 2.1.3. *Nech n je prirodzené číslo. Ak n^2 je deliteľné štyrmi, tak n je párne.*

Dôkaz. Nepriamo. Nech n je nepárne, teda má tvar $n = 2k + 1$. Potom $n^2 = (2k + 1)^2 = 4k^2 + 4k + 1 = 4(k^2 + k) + 1$ má zvyšok 1 po delení štyrmi, čiže nie je deliteľné štyrmi. □

Všimnite si, že v predchádzajúcom tvrdení sme namiesto výroku z tvrdenia dokazovali: Ak n nie je párne, tak n^2 nie je deliteľné štyrmi.

2.1.2 Matematická indukcia

Často používaný, a preto aj dosť dôležitý spôsob dôkazu, je dôkaz pomocou matematickej indukcie. Povedzme si teda o ňom pár slov a ilustrujeme si ho na niekoľkých príkladoch. (S matematickou indukciou ste sa už stretli na strednej škole, táto podkapitolka slúži len na pripomenutie.)

Dôkaz *matematickou indukciou* spočíva v tom, že ak chceme dokázať nejaký výrok $V(n)$ pre všetky prirodzené čísla $n \in \mathbb{N}$, dokážeme ho najprv pre $n = 0$ a ďalej dokážeme, že ak platí $V(n)$, tak tento výrok platí pre nasledujúce číslo, teda platí $V(n + 1)$. (Toto treba dokázať pre všetky prirodzené čísla n.)

Dôkaz, že z $V(n)$ vyplýva $V(n+1)$ nazývame *indukčný krok* a výrok $V(n)$ použitý v tomto kroku sa volá *indukčný predpoklad.*

Dokazovaný výrok môže obsahovať viacero premenných, preto pri dôkaze matematickou indukciou vždy treba uviesť, vzhľadom na ktorú premennú sa indukcia robí.

Niekedy sa indukčného kroku v tvare $V(n) \Rightarrow V(n+1)$ (z $V(n)$ vyplýva $V(n+1)$) používa tvar $V(n-1) \Rightarrow V(n)$ (z $V(n-1)$ vyplýva $V(n)$). Oba prístupy sú rovnocenné, v druhom prípade samozrejme dokazujeme indukčný krok len pre $n \geq 1$. (Pre $n = 0$ by ani nedával zmysel, lebo $V(-1)$ vôbec nemusí byť zadefinované.) My budeme používať prvý z týchto dvoch prístupov.

Niekedy (keď výrok dokazujeme pre všetky čísla počnúc od nejakého daného čísla n_0) nerobíme prvý krok indukcie pre $n = 0$ ale pre $n = n_0$. Aj v indukčnom kroku potom môžeme použiť predpoklad $n \geq n_0$ namiesto $n \geq 0$.

{dokazy:PRMIGEOM}

Príklad 2.1.4. Uvažujme geometrickú postupnosť určenú predpisom $a_0 = 1$ a $a_{n+1} = 2a_n$. (Teda členy tejto postupnosti sú $a_0 = 1, a_1 = 2, a_2 = 4, \ldots, a_n = 2^n, \ldots$) Dokážeme, že pre súčet prvých $n + 1$ členov tejto postupnosti platí

$$1 + 2 + \cdots + 2^n = \sum_{k=0}^{n} 2^k = 2^{n+1} - 1.$$

Budeme postupovať matematickou indukciou vzhľadom na n.

1° Pre $n = 0$ dostaneme rovnosť $2^0 = 2^1 - 1$, teda tvrdenie platí.

2° Indukčný krok. Predpokladajme, že rovnosť platí pre n. Potom pre $n + 1$ dostaneme

$$1+2+\cdots+2^n+2^{n+1} = (1+2+\cdots+2^n)+2^{n+1} \overset{\text{IP}}{=} 2^{n+1}-1+2^{n+1} = 2.2^{n+1}-1 = 2^{n+2}-1,$$

čo môžeme stručnejšie zapísať ako

$$\sum_{k=0}^{n+1} 2^k = 2^{n+1} + \sum_{k=0}^{n} 2^k \overset{\text{IP}}{=} 2^{n+1} + 2^{n+1} - 1 = 2.2^{n+1} - 1 = 2^{n+2} - 1.$$

Indukčný predpoklad sme použili na mieste označenom IP.

Takmer rovnakým spôsobom by ste mohli odvodiť vzorec pre súčet prvých $n + 1$ členov ľubovoľnej geometrickej postupnosti. (Teda a_0 bude ľubovoľné a aj dvojku nahradíme ľubovoľným kvocientom $q \neq 1$.) Vyskúšajte si to!

zy:POZNDEFMI}

Poznámka 2.1.5. Okrem dôkazov sa matematická indukcia používa aj pri definíciach. Ak chceme zadefinovať nejaký matematický objekt pre ľubovoľné prirodzené číslo n, môžeme postupovať tak, že ho zadefinujeme pre $n = 0$ a ďalej zavedieme $(n + 1)$-vý objekt pomocou n-tého. V takomto prípade hovoríme o *definícii matematickou indukciou.*

Definíciu matematickou indukciou sme už použili v predchádzajúcom príklade – postupnosť (a_n) bola určená tým, že $a_0 = 1$ a $a_{n+1} = 2a_n$.

Veľmi dôležitým variantom matematickej indukcie je *úplná indukcia.* V tomto prípade v indukčnom kroku dokazujeme platnosť nového tvrdenia nie pomocou platnosti pre predchádzajúce číslo, ale v dôkaze využijeme platnosť tvrdenia pre všetky (prípadne viaceré) menšie čísla.

Zatiaľ čo indukčný krok matematickej indukcie sme mohli schematicky zapísať ako

$$V(n) \Rightarrow V(n+1)$$

pri úplnej indukcii v indukčnom kroku dokazujeme

$$V(0), V(1), V(2), \ldots, V(n) \Rightarrow V(n+1);$$

inak povedané, treba dokázať že ak výrok $V(k)$ platí pre všetky $k < n$, tak platí aj pre číslo n. (Samozrejme, aj tu by sme mohli úplne rovnocenne použiť indukčný krok kde by sme dokazovali $V(n)$ z platnosti dokazovaného výroku pre $k = 0, 1, \ldots, n-1$. Niekedy budeme používať úplnú indukciu aj v takejto podobe.) Aj tu platí, že indukciu môžeme začať aj od iného prvku namiesto nuly.

Príklad 2.1.6. Uvažujme postupnosť určenú predpisom $a_0 = 1$ a $a_{n+1} = a_0 + a_1 + \cdots + a_n = \sum_{k=0}^{n} a_k$. (Všimnite si, že sme túto postupnosť definovali pomocou úplnej indukcie – definícia $(n+1)$-vého prvku využíva všetky menšie prvky.) Dokážeme úplnou indukciou vzhľadom na n, že pre $n \geq 1$ platí $a_n = 2^{n-1}$.

1° Pre $n = 0$ máme $a_1 = 1 = 2^{1-1}$, teda v tomto prípade dokazovaná rovnosť platí.

2° Indukčný krok: Predpokladajme, že $a_k = 2^k$ pre $k = 0, 1, \ldots, n$. Potom

$$a_{n+1} = \sum_{k=0}^{n} a_k = a_0 + \sum_{k=1}^{n} a_k \overset{\text{IP}}{=} 1 + \sum_{k=1}^{n} 2^{k-1} \overset{(1)}{=} 1 + \sum_{j=0}^{n-1} 2^j \overset{(2)}{=} 1 + 2^n - 1 = 2^n.$$

(V rovnosti (1) sme zaviedli novú sumačnú premennú $j = k - 1$ a v rovnosti (2) sme využili výsledok dokázaný v príklade 2.1.4.)

Poznámka 2.1.7. Určite ste si všimli, že v príkladoch dôkazov matematickou indukciou sme používali zápisy typu $\sum_{k=1}^{n} a_k = a_1 + \ldots + a_n$. (Pokiaľ nie ste na používanie znaku $\sum$ zvyknutí zo stednej školy, vo vysokoškolskej matematike sa s ním budete stretávať veľmi často, takže si naň treba čím skôr zvyknúť.) Zjednodušene sa dá povedať, že ak sa pri úprave výrazov vystkytne výraz obsahujúci tri bodky (...), v skutočnosti je za touto úpravou skrytý dôkaz matematickou indukciou. (Väčšinou natoľko jednoduchý, že dôkaz správnosti tejto úpravy samostatne nedokazujeme.)

2.1.3 Drobné rady ako dokazovať

Hoci nasledujúca rada na prvý pohľad znie veľmi naivne, pri hľadaní dôkazu nejakého tvrdenia je užitočné uvedomiť si, čo všetko máme zadané a čo potrebujeme dokázať. Pri jednoduchších

dôkazoch sa často stane, že keď si poriadne zapíšeme vlastnosť, ktorú chceme dokázať a takisto všetky predpoklady, takmer okamžite zbadáme, ako postupovať. Samozrejme, aj pri zložitejších dôkazoch je veľmi dobre nestrácať zo zreteľa aké predpoklady môžeme použiť a k čomu vlastne chceme dospieť.

Niekedy môže byť užitočné aj hľadanie protipríkladu, prípadne overenie tvrdenia na konkrétnych príkladoch. Môže sa stať, že si uvedomíte, prečo sa vám kontrapríklad nedarí zostrojiť alebo pri overovaní, či tvrdenie platí pre konkrétny príklad prídete na nejakú zákonitosť, ktorá vám nakoniec pomôže dané tvrdenie dokázať.

Ďalšia rada je, že občas nie je zlé uvedomiť si, či ste skutočne v dôkaze použili všetky predpoklady. Tvrdenia a úlohy v učebniciach bývajú často formulované tak, že tam nie sú uvedené žiadne zbytočné predpoklady navyše. Preto dôkaz, kde ste nepoužili všetky predpoklady, je trochu podozrivý a treba ho skontrolovať. (Aj keď sa samozrejme môže stať, že zákerný autor úlohy či skúšajúci mohol pridať nejaké predpoklady navyše.)

2.1.4 Výroky, logické spojky, tautológie

{dokazy:SSECTTAUT}

Toto je ďalšia téma, ktorú by ste mali ovládať už zo strednej školy – napriek tomu však pripomenieme niektoré základné fakty.

Definícia 2.1.8. *Negáciou* výroku P rozumieme výrok „neplatí P“. Označujeme ju $\neg P$.

Pre dva výroky P a Q nazývame ich *konjunkciou* výrok „P a Q“, označujeme $P \wedge Q$.

Disjunkcia je výrok „P alebo Q“, označujeme $P \vee Q$.

Pod *implikáciou* rozumieme výrok „ak platí P, tak platí Q“, označujeme $P \Rightarrow Q$.

Ekvivalencia výrokov P a Q je výrok „P platí práve vtedy, keď platí Q“, označujeme $P \Leftrightarrow Q$.

Tieto definície logických spojok sú zhrnuté v nasledujúcich pravdivostných tabuľkách.[3]

P	$\neg P$
1	0
0	1

P	Q	$P \wedge Q$
1	1	1
1	0	0
0	1	0
0	0	0

P	Q	$P \vee Q$
1	1	1
1	0	1
0	1	1
0	0	0

P	Q	$P \Rightarrow Q$
1	1	1
1	0	0
0	1	1
0	0	1

P	Q	$P \Leftrightarrow Q$
1	1	1
1	0	0
0	1	0
0	0	1

Tautológie môžeme overovať jednoducho metódou, ktorú poznáte zo strednej školy.

Príklad 2.1.9. Overme napríklad tautológiu $P \vee (\neg P)$ (princíp vylúčenia tretieho).

P	$\neg P$	$P \vee \neg P$
1	0	1
0	1	1

Ako ďalší príklad si ukážeme overenie jedného z de Morganových pravidiel.

{dokazy:PRDEMORGAN}

Príklad 2.1.10. *De Morganove pravidlá* sú pravidlá ako negovať konjunkciu a disjunkciu.

$$\neg(P \wedge Q) \quad \Leftrightarrow \quad \neg P \vee \neg Q$$
$$\neg(P \vee Q) \quad \Leftrightarrow \quad \neg P \wedge \neg Q$$

Samozrejme, pretože teraz vo výroku vystupuje viacero premenných, budeme potrebovať viac riadkov tabuľky na to, aby sme vyčerpali všetky možnosti.

[3] Na označovanie pravdivosti a nepravdivosti budeme v tabuľke používať symboly 1 a 0. Niekedy sa zvyknú používať aj T a F, ako skratky pre anglické true a false.

P	Q	$P \vee Q$	$\neg(P \vee Q)$	$\neg P \wedge \neg Q$	$\neg(P \vee Q) \Leftrightarrow (\neg P \wedge \neg Q)$
1	1	1	0	0	1
1	0	1	0	0	1
0	1	1	0	0	1
0	0	0	1	1	1

Niekedy si môžeme pri overovaní platnosti tautológie použiť aj jednoduchší postup. V predchádzajúcom príklade sme napríklad mohli na základe symetrie overovať o jeden riadok menej. Inú možnosť zjednodušenia ilustruje nasledujúci príklad.

{dokazy:PRTAUTNEPR}

Príklad 2.1.11. Dokážeme tautológiu $(P \Rightarrow Q) \Leftrightarrow (\neg Q \Rightarrow \neg P)$. (Táto tautológia súvisí s princípom nepriameho dôkazu. Implikácia $\neg Q \Rightarrow \neg P$ sa zvykne nazývať *obmena implikácie* $P \Rightarrow Q$.)

Aby sme dokázali ekvivalenciu dvoch výrokov, stačí ukázať, že výrok na ľavej strane je nepravdivý práve v tých prípadoch, kedy je nepravdivý výrok na pravej strane.

Implikácia je nepravdivá jedine v prípade, že ľavý výrok je pravdivý a pravý je nepravdivý (prípad $1 \Rightarrow 0$). Teda výrok $P \Rightarrow Q$ je nepravdivý práve vtedy, keď $P = 1$ a $Q = 0$. Podobne, aby bol výrok $\neg Q \Rightarrow \neg P$ nepravdivý, musí byť $\neg Q = 1$ a $\neg P = 0$, čo je presne ten istý prípad $P = 1$ a $Q = 0$. Vidíme, že obe strany ekvivalencie majú vždy tú istú pravdivostnú hodnotu.

(Tento spôsob overenia tautológie sa až tak veľmi nelíši od tabuľkovej metódy – vlastne sme si len rozmysleli, v ktorých riadkoch tabuľky sa na oboch stranách uvedenej ekvivalencie vyskytne 0 – zdá sa mi byť bližší ku spôsobu, ako prirodzene uvažujeme o výrokoch.)

V cvičení 2.1.1 nájdete viacero tautológií. Je dobré si uvedomiť ako súvisia tautológie s niektorými typmi dôkazov. Tautológia z príkladu 2.1.11 je presne princíp nepriameho dôkazu, ktorý sme už spomínali. Tautológiu z cvičenia 2.1.1b) budeme tiež často používať pri dokazovaní – namiesto výroku tvaru $P \Leftrightarrow Q$ dokážeme zvlášť jednotlivé implikácie $P \Rightarrow Q$ a $Q \Rightarrow P$.

2.1.5 Negácia výrokov s kvantifikátormi

{dokazy:SSECTNEGKVANT}

Okrem logických spojok budeme na zápis tvrdení používať aj kvantifikátory. Budeme používať všeobecný (univerzálny) kvantifikátor

$$(\forall x \in A)P(x)$$

vo význame „pre každý prvok x množiny A platí $P(x)$“ a existenčný kvanfitifkátor

$$(\exists x \in A)P(x),$$

ktorý znamená „existuje prvok x množiny A, pre ktorý platí $P(x)$“. (Tu $P(x)$ predstavuje *výrokovú funkciu* – výrok, v ktorom vystupuje „premenná“ x.)

Veľmi dôležité budú nasledovné pravidlá pre negáciu výrokov s kvantifikátormi.

$$\begin{aligned} \neg[(\forall x)P(x)] \quad &\Leftrightarrow \quad (\exists x)(\neg P(x)), \\ \neg[(\exists x)P(x)] \quad &\Leftrightarrow \quad (\forall x)(\neg P(x)). \end{aligned}$$

(Teda existenčný kvantifikátor sa mení na všeobecný a obrátene a výrok pod kvantifikátorom sa zneguje.)

Cvičenia

{dokazy:CVTAU

Úloha 2.1.1. Dokážte, že nasledujúce výroky sú tautológie:
a) $(\neg P \vee Q) \Leftrightarrow (P \Rightarrow Q)$
b) $(P \Leftrightarrow Q) \Leftrightarrow [(P \Rightarrow Q) \wedge (Q \Rightarrow P)]$
c) $\neg(P \wedge Q) \Leftrightarrow (\neg P \vee \neg Q)$
d) $((P \wedge Q) \Rightarrow R) \Leftrightarrow (P \Rightarrow (Q \Rightarrow R))$

2.2 Množiny a zobrazenia

2.2.1 Množiny

Už sme spomenuli, že základným stavebným kameňom súčasnej matematiky je teória množín. Viac sa o nej dozviete v 4. ročníku (viď. tiež [ŠS]). V tejto prednáške úplne vystačíme so základnými predstavami o množinách. (Takýto intuitívny prístup k teórii množín sa zvykne nazývať naivná teória množín – na rozdiel od axiomatickej teórie množín, ktorá systematicky definuje pojem množiny pomocou niekoľkých základných axióm.) Navyše, všetky množiny, s ktorými sa budeme stretávať budú množina komplexných čísel a rôzne jej podmnožiny (ako napríklad $\mathbb{R}$, $\mathbb{Z}$, $\mathbb{N}$). Tieto množiny poznáte zo strednej školy a mali by ste o nich mať dobrú intuitívnu predstavu, takže si až tak veľmi nemusíme robiť starosti s tým, že si celkom presne nevysvetlím, čo sa rozumie pod pojmom množina.

Pre naše účely stačí množinu chápať ako súhrn nejakých prvkov. Pričom prvky budú najčastejšie čísla, niekedy aj množiny alebo zobrazenia. A takisto sa budeme stretávať s množinami, ktoré sa dajú z množín čísel vytvoriť pomocou množinových operácií ako sú napríklad zjednotenie alebo karteziánsky súčin.

Každá množina je určená svojimi prvkami. Inak povedané, dve množiny sa rovnajú, ak majú rovnaké prvky.

Definícia 2.2.1. Vzťah byť prvok množiny zapisujeme ako $x \in A$, čítame „x patrí A.“

Hovoríme, že množiny A a B sa *rovnajú* (označujeme $A = B$), ak platí

$$(x \in A) \qquad \Leftrightarrow \qquad (x \in B)$$

pre ľubovoľný prvok x.

Množiny, ktorá nemá nijaké prvky nazývame *prázdna množina* a označujeme $\emptyset$.

To znamená, že pre nijaké x neplatí $x \in \emptyset$.

Ďalej pripomenieme niektoré základné vzťahy a operácie s množinami. (Opäť ide o opakovanie – pravdepodobne ste o nich už počuli.)

Definícia 2.2.2. Hovoríme, že A je *podmnožinou* B, ak každý prvok množiny A patrí aj do B, označujeme $A \subseteq B$. Inak povedané, $A \subseteq B$ ak pre každé x platí

$$(x \in A) \qquad \Rightarrow \qquad (x \in B).$$

Vzťah množín A a B, pre ktoré platí $A \subseteq B$ sa tiež zvykne nazývať *inklúzia.*

Inklúziu budeme často používať pri dôkaze rovnosti nejakých množín. Platí totiž

$$A = B \qquad \Leftrightarrow \qquad (A \subseteq B) \wedge (B \subseteq A).$$

(Vyplýva to z tautológie uvedenej v cvičení 2.1.1b) v časti 2.1. Ak totiž za výroky vystupujúce v tejto tautológii dosadíme $x \in A$ a $x \in B$, tak dostaneme, že tvrdenie $(x \in A) \Leftrightarrow (x \in B)$,

čo je presne rovnosť množín A a B, je ekvivalentné s tvrdením $[(x \in A) \Rightarrow (x \in B)] \wedge [(x \in B) \Rightarrow (x \in A)]$, čo môžeme skrátene zapísať $(A \subseteq B) \wedge (B \subseteq A)$.)

Pripomenieme si teraz niekoľko spôsobov, ako môžeme z daných množín vytvárať nové množiny.

Často budeme definovať množiny zápisom tvaru

$$\{x \in A; P(x)\},$$

ktorý znamená množinu všetkých tých prvkov z A, pre ktoré platí výrok $P(x)$. (Pričom A je nejaká vopred daná množina.) Napríklad z množiny prirodzených čísel $\mathbb{N}$ môžeme vybrať párne čísla $E = \{n \in \mathbb{N}; n \text{ je deliteľné číslom } 2\}$.

Zo strednej školy poznáte základné operácie s množinami – prienik, zjednotenie a rozdiel množín – pripomeňme si však ich definície. Použijeme pri nich práve typ zápisu, ktorý sme pred chvíľou spomínali.

Definícia 2.2.3. *Zjednotenie* dvoch množín A a B je množina

$$A \cup B = \{x; x \in A \vee x \in B\}.$$

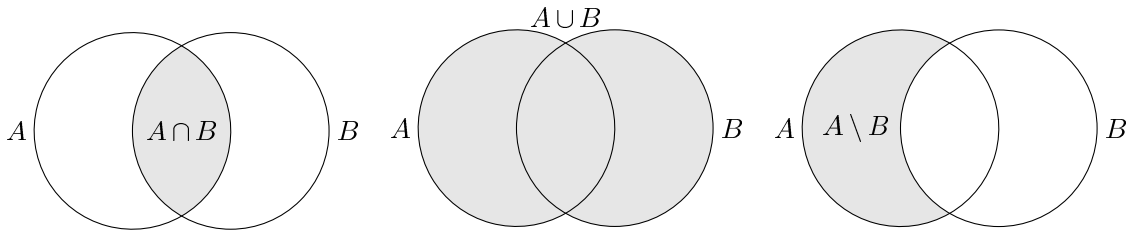
Prienik dvoch množín A a B je množina

$$A \cap B = \{x \in A; x \in B\}.$$

Rozdiel dvoch množín A a B je množina

$$A \setminus B = \{x \in A; x \notin B\}$$

Operácie s množinami znázorňujeme pomocou *Vennových diagramov* (obr. 2.1).

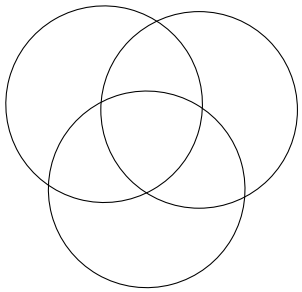


Obr. 2.1: Operácie s množinami {zobrazenia:FIGMNCUP}

Keďže množiny často definujeme pomocou nejakého výroku, ktorý majú spĺňať všetky prvky množiny, dajú sa operácie s množinami chápať aj ako iný spôsob vyjadrovania o výrokoch. Aj identity s množinami môžeme overovať tak, že pracujeme s výrokmi typu $x \in A$. Iný spôsob je použiť Vennove diagramy – v tom prípade je potrebné nakresliť si množiny vo všeobecnej (alebo tiež generickej polohe) – tak, aby sme na diagrame mali body pre každú možnú kombináciu pravdivostných hodnôt výrokov $x \in A$, $x \in B$ atď.

Často budeme používať aj zjednotenie a prienik nekonečného počtu množín. Ak $\{A_i; i \in I\}$ je nejaký systém množín (oindexovaný prvkami množiny I), tak používame označenia

$$\bigcup_{i \in I} A_i = \{x; (\exists i \in I) x \in A_i\};$$
$$\bigcap_{i \in I} A_i = \{x; (\forall i \in I) x \in A_i\}.$$



Obr. 2.2: Generická poloha {zobrazenia:M

(Pri prieniku navyše požadujeme, aby $I \neq \emptyset$; prienik prázdneho systému množín nedefinujeme.) V prípade, že indexová množina je $I = \{1, 2, \ldots\}$, budeme používať označenie $\bigcup_{n=1}^{\infty} A_n$ resp. $\bigcap_{n=1}^{\infty} A_n$.

Bude pre nás dôležité ešte jedna operácia s množinami – karteziánsky súčin.

Definícia 2.2.4. Ak A, B sú množiny, tak ich *karteziánsky súčin* je množina všetkých usporiadaných dvojíc (a, b) takých, že $a \in A$ a $b \in B$. Označujeme ho

$$A \times B = \{(a, b); a \in A, b \in B\}.$$

Hoci pojem *usporiadaná dvojica* sme nedefinovali, malo by byť intuitívne jasné, čo sa ním myslí. Slovíčko usporiadaná je v názve preto, že záleží na poradí. Usporiadanú dvojicu prvkov a a b budeme označovať (a, b).

Napríklad

$$\{0, 1\} \times \{0, 1\} = \{(0, 0), (0, 1), (1, 0), (1, 1)\}$$

a dvojice $(0, 1)$ a $(1, 0)$ považujeme za rôzne. (Oproti tomu, množiny $\{0, 1\}$ a $\{1, 0\}$ sú rovnaké – pretože pri množinách záleží len na tom, ktoré prvky tam patria.)

2.2.2 Zobrazenia

{zobrazenia:SSECTZOBR}

Pod zobrazením z množiny X do množiny Y rozumieme akýkoľvek predpis, ktorý každému prvku množiny X priradí jediný prvok množiny Y. Formálne môžeme zobrazenie zadefinovať nasledovne:

Definícia 2.2.5. *Zobrazenie* $f\colon X \to Y$ z množiny X do množiny Y je podmnožina f množiny $X \times Y$ taká, že ku každému $x \in X$ existuje práve jedno $y \in Y$ s vlastnosťou $(x, y) \in f$.

Množinu X budeme tiež nazývať *definičný obor* zobrazenia f a množina Y je *obor hodnôt* zobrazenia f.

Namiesto zápisu $(x, y) \in f$ budeme používať zápis $y = f(x)$.

Podmienka, že ku každému $x \in X$ existuje práve jedno $y \in Y$, je presne formalizáciou toho, čo chápeme pod priradením (jediného) prvku $y = f(x)$ každému prvku $x \in X$.

Poznamenajme, že (podobne ako pri mnohých ďalších označeniach) pri použití inej literatúry je často nutné skontrolovať, či sa zhodujú použité definície. (V prípadoch, keď sa

budeme hovoriť o pojmoch a označeniach, ktoré sa často zvyknú v matematickej literatúre líšiť, na to vždy upozorníme.)

V tomto prípade je vhodné spomenúť, že [KGGS] používa namiesto $y = f(x)$ zápis $y = xf$. (V súvislosti so zobrazeniami sa v [KGGS] vyskytujú aj ďalšie odlišnosti, ku ktorým sa o chvíľu dostaneme.)

{zobrazenia:PR1}

Príklad 2.2.6. Uvedieme niekoľko príkladov zobrazení.
$f_1 \colon \mathbb{N} \to \mathbb{N}$, $f_1(n) = 2n + 1$
$f_2 \colon \mathbb{N} \to \mathbb{N}$, $f_2(n) = 2n$
$f_3 \colon \mathbb{N} \to \mathbb{N}$, $f_3(n) = \begin{cases} n+1, & \text{ak } n \text{ je párne} \\ n-1, & \text{ak } n \text{ je nepárne} \end{cases}$
$f \colon \mathbb{R} \to \mathbb{R}$, $f(x) = x \cdot \sin x$
$g \colon \mathbb{R} \to \mathbb{R}$, $g(x) = \sin x$
$h \colon \mathbb{R} \to \mathbb{R}$, $h(x) = x^2$

Definícia 2.2.7. Hovoríme, že dve zobrazenia $f \colon X \to Y$ a $g \colon Z \to W$ sa *rovnajú,* ak $X = Z$, $Y = W$ a $f(x) = g(x)$ pre každé $x \in X$. (Inými slovami, ak sa rovnajú ich definičné obory, obory hodnôt a obe zobrazenia nadobúdajú v každom bode rovnakú hodnotu.) Rovnosť zobrazení označujeme $f = g$.

{zobrazenia:PRROVNOST}

Príklad 2.2.8. Dve zobrazenia sa môžu rovnať aj keď sú zapísané iným zápisom. Napríklad zobrazenia $k, l \colon \mathbb{R} \to \mathbb{R}$ určené predpismi $k(x) = \sin^2 x + \cos^2 x$ a $l(x) = 1$ sa rovnajú.

O celkom zaujímavom vývoji pojmu zobrazenia v histórii teórie množín sa viac môžete dozvedieť napríklad v úvodnej kapitole knihy [BŠ].

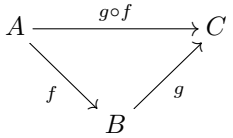
V tejto kapitole ešte zavedieme ďalšie dôležité pojmy, ako sú skladanie zobrazení, bijektívne, injektívne a surjektívne zobrazenia.

{zobrazenia:DEFSKLAD}

Definícia 2.2.9 (Skladanie zobrazení)**.** Ak $f \colon X \to Y$, $g \colon Y \to Z$ sú zobrazenia, tak *zložením zobrazení* f *a* g nazývame zobrazenie $g \circ f \colon X \to Z$ také, že pre každé $x \in X$ platí

$$g \circ f(x) = g(f(x)).$$

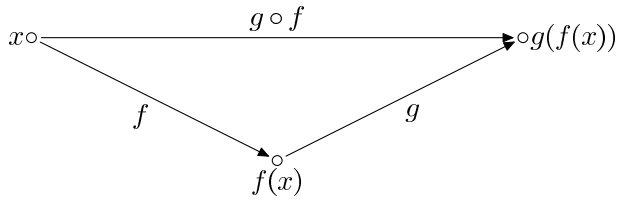
Zobrazenia môžeme skladať iba vtedy, keď obor hodnôt prvého zobrazenia je rovnaký ako definičný obor druhého zobrazenia.



Ak si predstavíme zobrazenia ako šípku smerujúcu od x ku $f(x)$, tak skladanie zobrazení znamená, že z x prejdeme najprv po šípke f a potom po šípke g, čiže sa dostaneme do $g(f(x))$. Táto predstava o skladaní zobrazení je znázornená na obrázku 2.3.

Napríklad pre funkcie g a h z príkladu 2.2.6 dostaneme $h \circ g(x) = \sin^2 x$ a $g \circ h(x) = \sin x^2$. Vidíme, že pre zobrazenia $g, h \colon A \to A$ vo všeobecnosti neplatí $g \circ h = h \circ g$.

POZOR!!! V niektorej literatúre (napríklad v [KGGS]) nájdete skladanie zobrazení definované v opačnom poradí (t.j. $g \circ f(x) = f(g(x))$), my ho budeme používať tak, ako sme ho zaviedli v definícii 2.2.9. (Na prvý pohľad vyzerá zápis $g \circ f(x) = f(g(x))$ nelogicky; môžeme ho chápať tak, že píšeme zobrazenia v takom poradí, ako sa skladajú. Ak by ste sa pozreli



Obr. 2.3: Skladanie zobrazení {zobrazenia:F

do [KGGS], zistili by ste, že namiesto $f(x)$ používajú autori zápis xf. Pri takomto označení je skutočne logickejší zápis $x(g \circ f) = (xg)f$.)[4]

Teraz dokážeme veľmi dôležitú vlastnosť skladania zobrazení.

{zobrazenia:TVRASOC}

Tvrdenie 2.2.10 (Asociatívnosť skladania zobrazení)**.** *Nech $f\colon X \to Y$, $g\colon Y \to Z$, $h\colon Z \to W$ sú zobrazenia, potom*

$$(h \circ g) \circ f = h \circ (g \circ f).$$

Dôkaz. Obe zobrazenia, ktoré porovnávame, sú zobrazenia z množiny X do množiny W. Majú teda rovnaké definičné obory i obory hodnôt.

Zostáva nám teda overiť, či nadobúdajú rovnaké hodnoty. To zistíme jednoduchým výpočtom:

$$((h \circ g) \circ f)(x) = (h \circ g)(f(x)) = h(g(f(x)))$$
$$(h \circ (g \circ f))(x) = h((g \circ f)(x)) = h(g(f(x)))$$

□

V ďalších častiach tejto prednášky budú dosť dôležité typy zobrazení, ktoré teraz zadefinujeme.

Definícia 2.2.11. Nech $f\colon X \to Y$ je zobrazenie. Hovoríme, že f je *injektívne* (*prosté*) *zobrazenie* (alebo tiež injekcia), ak pre všetky $x, y \in X$ také, že $x \neq y$ platí $f(x) \neq f(y)$.

Hovoríme, že f je *surjekcia* (*surjektívne zobrazenie, zobrazenie na*), ak pre každé $y \in Y$ existuje také, $x \in X$, že $f(x) = y$.

Hovoríme, že f je *bijekcia* (*bijektívne zobrazenie*), ak f je súčasne injekcia aj surjekcia.

Ekvivalentná definícia injekcie by bola, ak by sme namiesto

$$x \neq y \ \Rightarrow f(x) \neq f(y)$$

uvažovali podmienku

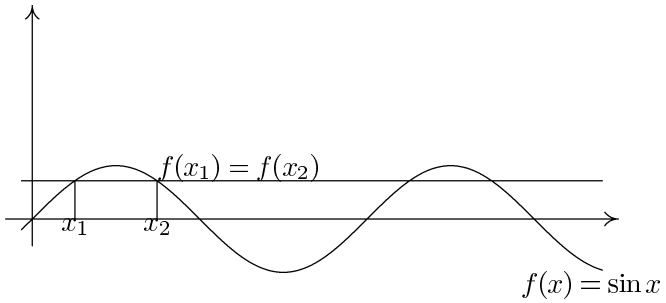
$$f(x) = f(y) \ \Rightarrow x = y.$$

(Lebo výroky $P \Rightarrow Q$ a $\neg Q \Rightarrow \neg P$ sú ekvivalentné, pozri príklad 2.1.11.)

Definíciu surjektívneho zobrazenia by sme mohli preformulovať tak, že každá hodnota y z oboru hodnôt Y sa nadobúda v aspoň jednom bode definičného oboru.

[4] Pravdepodobne si budete musieť zvyknúť na to, že na rôznych prednáškach sa stretnete s rozličnými spôsobmi označenia; napríklad pri skladaní zobrazení ale aj pri ďalších iných veciach. Ja som zvolil používanie poradia skladania tak, aby bolo rovnaké s definíciou skladania zobrazení, ktorú používate na prednáške z matematickej analýzy v prvom ročníku – teda aby ste nemali problém s používaním 2 rozličných označení počas toho istého ročníka.

Z grafu zobrazenia $f\colon \mathbb{R} \to \mathbb{R}$ môžeme jednoducho vyčítať, či ide o injekciu alebo surjekciu. Stačí sa pozrieť na vodorovné priamky – rovnobežné s osou x. Zobrazenie je injektívne, ak každá takáto priamka pretína graf funkcie najviac raz (pozri Obr. 2.4). Zobrazenie je surjektívne, ak každá takáto priamka pretína graf funkcie aspoň raz. Zobrazenie je bijektívne, ak každá takáto priamka pretína graf funkcie práve raz.



azenia:FIG03} Obr. 2.4: Ak vodorovná priamka pretne graf v 2 bodoch, zobrazenie nie je injektívne

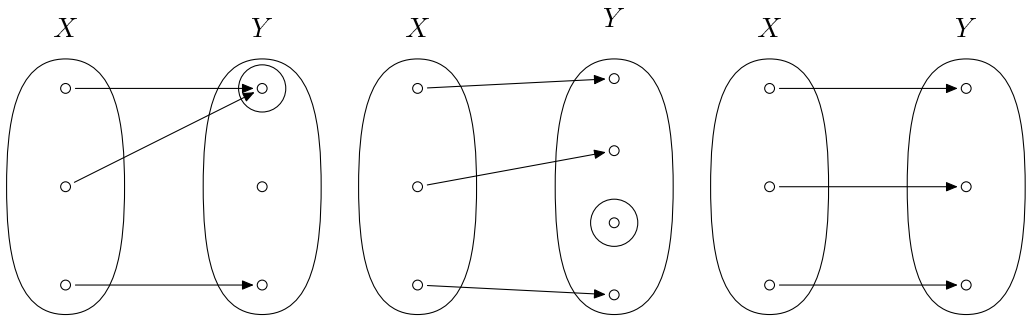
Príklad 2.2.12. Zobrazenie $f(x) = \sin x$, $f\colon \mathbb{R} \to \mathbb{R}$ nie je ani surjektívne ani injektívne. (Napríklad $f(0) = f(\pi) = f(2\pi) = 0$, preto f nie je injektívne. Hodnota $2 \in \mathbb{R}$ sa nenadobúda v žiadnom bode.)

Ak však zmeníme obor hodnôt $g(x) = \sin x$, $g\colon \mathbb{R} \to \langle -1, 1\rangle$, dostaneme už surjektívne zobrazenie.

Zobrazenie $h(x) = \sqrt{x}$, $h\colon \langle 0, \infty) \to \mathbb{R}$ je injektívne. Opäť, ak by sme zmenili obor hodnôt, dostaneme surjektívne zobrazenie $j(x) = \sqrt{x}$, $j\colon \langle 0, \infty) \to \langle 0, \infty)$. Zobrazenie j je bijekcia (je injektívne aj surjektívne.)

Iným príkladom bijektívneho zobrazenia je $k\colon (-\frac{\pi}{2}, \frac{\pi}{2}) \to \mathbb{R}$, $k(x) = \operatorname{tg} x$.

Opäť si môžeme pomôcť predstavou zobrazení ako šípok. Zobrazenie je injekcia, ak nenastane situácia znázornená na ľavom obrázku v 2.5, keď sa viaceré šípky stretnú v jednom bode. O surjekciu ide vtedy, ak do každého bodu ide aspoň jedna šípka, čiže nenastane situácia znázornená na prostrednom obrázku. V prípade bijekcie ide z každého prvku x jediná šípka do jediného bodu y, teda bijekcia určuje jedno-jednoznačné priradenie medzi prvkami.



Obr. 2.5: Ilustrácia injekcie, surjekcie a bijekcie {zobrazenia:FIGBIJ}

{zobrazenia:TVRSKLINJ}

Tvrdenie 2.2.13. *Zloženie dvoch injekcií je injekcia, zloženie dvoch surjekcií je surjekcia, zloženie dvoch bijekcií je bijekcia.*

Dôkaz. Dané zobrazenia označme $f\colon X \to Y$, $g\colon Y \to Z$.

Najprv predpokladajme, že f aj g sú injekcie. Nech ďalej $x, y \in X$ majú tú vlastnosť, že

$$(g \circ f)(x) = (g \circ f)(y).$$

Túto rovnosť môžeme prepísať ako

$$g(f(x)) = g(f(y)).$$

Pretože g je injekcia, vyplýva z tejto rovnosti

$$f(x) = f(y).$$

Opäť použitím definície injekcie, ale tentokrát pre zobrazenie f, dostaneme

$$x = y.$$

Dokázali sme implikáciu $(g \circ f)(x) = (g \circ f)(y) \Rightarrow x = y$, teda $g \circ f$ je skutočne injekcia.

Teraz nech f aj g sú surjekcie. Máme dokázať, že ku každému $z \in Z$ existuje $x_0 \in X$ také, že $g(f(x_0)) = z$. Z toho, že g je surjekcia, vieme, že existuje $y_0 \in Y$ také, že $g(y_0) = z$. Podobne (pretože f je surjekcia), k tomuto y_0 vieme nájsť x_0 také, že $f(x_0) = y_0$. Spojením týchto dvoch rovností ale dostaneme

$$g(f(x_0)) = g(y_0) = z,$$

teda sme skutočne našli x_0, ktoré sa zobrazením $g \circ f$ zobrazí na z.

Posledná časť tvrdenia (ktorá hovorí o skladaní bijekcií) ľahko vyplýva z prvých dvoch častí. □

Ak chcete lepšie porozumieť predchádzajúcemu dôkazu (alebo sa pokúšate dokázať si toto tvrdenie samostatne), opäť môže byť užitočné pomôcť si kreslením šípok.

Definícia 2.2.14. Zobrazenie $id_X\colon X \to X$ také, že $id_X(x) = x$ pre každé $x \in X$ sa nazýva *identické zobrazenie* (*identita*).

Všimnime si, že pre ľubovoľné zobrazenie $f\colon X \to Y$ platí

$$f \circ id_X = f \qquad \text{a} \qquad id_Y \circ f = f.$$

{zobrazenia:DEFINV}

Definícia 2.2.15. Nech $f\colon X \to Y$ a $g\colon Y \to X$ sú zobrazenia. Ak platí

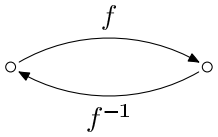
$$g \circ f = id_X$$
$$f \circ g = id_Y$$

tak hovoríme, že zobrazenie g je *inverzné zobrazenie k* f. Inverzné zobrazenie k zobrazeniu f označujeme f^{-1}.

Ak k nejakému zobrazeniu existuje inverzné zobrazenie, tak takéto zobrazenie je jediné (úloha 2.2.5). Vďaka tejto jednoznačnosti má zmysel zaviesť označenie pre inverzné zobrazenie – znak f^{-1} skutočne označuje len jediné konkrétne zobrazenie.

Vzťahy definujúce inverzné zobrazenie môžeme ekvivalentne prepísať aj tak, že pre každé $x \in X$ a pre každé $y \in Y$ platí

$$f^{-1}(f(x)) = x,$$
$$f(f^{-1}(y)) = y.$$



zenia:FIGINV}

Obr. 2.6: Inverzné zobrazenie

V zmysle našej intuície o zobrazení ako systéme šípok vychádzajúcich z každého prvku x, zodpovedá inverzné zobrazenie šípkam idúcim opačným smerom (obr. 2.6).

Aby takéto zobrazenie existovalo, do každého prvku y musí ísť aspoň jedna šípka (aby sme sa mali kam vrátiť) a do žiadneho prvku y nesmie ísť viac ako jedna šípka (aby sme dostali zobrazenie, t.j. len jediný prvok, do ktorého sa vraciame.) Tento fakt je sformulovaný v nasledujúcom tvrdení.

{zobrazenia:TVRBIJINV}

Tvrdenie 2.2.16. *Inverzné zobrazenie k f existuje práve vtedy, keď f je bijekcia.*

Dôkaz. $\boxed{\Rightarrow}$ Predpokladajme, že existuje inverzné zobrazenie k zobrazeniu $f\colon X \to Y$, označme ho g. Z definície inverzného zobrazenia dostaneme

$$f(x_1) = f(x_2) \;\Rightarrow\; g(f(x_1)) = g(f(x_2)) \;\Rightarrow\; g \circ f(x_1) = g \circ f(x_2) \;\Rightarrow\; x_1 = x_2,$$

čo znamená, že f je prosté.

Ešte potrebujeme ukázať, že ku každému $y \in Y$ existuje $x_0 \in X$ s vlastnosťou $f(x_0) = y$. Stačí zvoliť $x_0 = g(y)$, pretože (opäť podľa definície inverzného zobrazenia)

$$f(x_0) = f(g(y)) = y.$$

$\boxed{\Leftarrow}$ Predpokladajme, že $f\colon X \to Y$ je bijekcia, chceme dokázať, že existuje inverzné zobrazenie $g\colon Y \to X$. Majme ľubovoľné $y \in Y$. Pretože f je surjekcia, existuje aspoň jedno $x \in X$ také, že $f(x) = y$. Pretože f je injekcia, existuje najviac jedno také x – celkove teda dostávame, že existuje práve jedno x s touto vlastnosťou. Zobrazenie g teda môžeme definovať tak, že $g(y)$ je práve ten prvok x, pre ktorý platí $f(x) = y$. Inak povedané, g je určené podmienkou

$$g(y) = x \;\Leftrightarrow\; f(x) = y.$$

Ak zobrazenie g definujeme takýmto spôsobom dostaneme

$$g(f(x)) = x$$

(lebo $g(f(x))$ je práve ten prvok, ktorý sa zobrazí na $f(x)$, čo je presne prvok x) a

$$f(g(y)) = y$$

(lebo $g(y)$ sa, podľa definície zobrazenia g, zobrazí na y.) □

Ukážeme si ešte jeden dôkaz jednej implikácie z predchádzajúceho tvrdenia. (Je to v podstate ten istý dôkaz, inak formulovaný.) Inverzné zobrazenie tu dostaneme tak, že vymeníme poradie vo všetkých usporiadaných dvojiciach patriacich do f. Toto zodpovedá predstave o tom, že otáčame jednotlivé šípky (zodpovedá to výmene poradia) a takisto geometrickej predstave, ktorú poznáte ešte zo strednej školy, keď ste inverzné zobrazenie získali symetrickým zobrazením grafu funkcie podľa osi $y = x$.

Iný dôkaz implikácie $\boxed{\Leftarrow}$. Predpokladáme, že $f\colon X \to Y$ je bijekcia, chceme ukázať, že existuje inverzné zobrazenie $g\colon Y \to X$.

Položme

$$g := \{(y,x); (x,y) \in f\} = \{(y,x); y = f(x), x \in X\} = \{(f(x),x); x \in X\}.$$

Očividne je g podmnožinou karteziánskeho súčinu $Y \times X$. Najprv si uvedomme, že g je skutočne zobrazenie z Y do X.

Pretože f je súčasne injekcia a surjekcia, pre každé $y \in Y$ existuje práve jedno $x \in X$ také, že $y = f(x)$, resp. $(x,y) \in f$.

Máme teda zobrazenie $g\colon Y \to X$. Priamo z jeho definície vyplýva platnosť ekvivalencie

$$g(y) = x \qquad \Leftrightarrow \qquad y = f(x),$$

z ktorej rovnakým spôsobom ako v predchádzajúcom dôkaze dostaneme, že g je inverzné zobrazenie k zobrazeniu f. □

{zobrazenia:TVRINFGF}

Tvrdenie 2.2.17. *Nech $f\colon X \to Y$ a $g\colon Y \to Z$ sú bijekcie. Potom*

$$(f^{-1})^{-1} = f$$
$$(g \circ f)^{-1} = f^{-1} \circ g^{-1}$$

Všimnime si, že v predpokladoch máme, že f aj g sú bijekcie. Teda z tvrdenia 2.2.16 vieme, že existujú inverzné zobrazenia f^{-1} a g^{-1}. (Vďaka tomu môžeme o nich hovoriť; ak by neexistovali, tak uvedené tvrdenie by vôbec nemalo zmysel.)

Dôkaz. Použitím definície inverzného zobrazenia pre zobrazenie f máme $f^{-1} \circ f = id_X$ a $f \circ f^{-1} = id_Y$. To ale presne hovorí, že f je inverzné zobrazenie k f^{-1}.

Podobne vidíme, že

$$(f^{-1} \circ g^{-1}) \circ (g \circ f) = f^{-1} \circ (g^{-1} \circ g) \circ f = f^{-1} \circ id_Y \circ f = f^{-1} \circ f = id_X.$$

Analogickým spôsobom overíme, že $(g \circ f) \circ (f^{-1} \circ g^{-1}) = id_Z$. Tým sme vlastne overili obe podmienky z definície inverzného zobrazenie k $g \circ f$. □

Dôsledok 2.2.18. *Ak f je bijekcia, tak aj f^{-1} je bijekcia.*

2.2.3 Vzor a obraz množiny*

Definícia 2.2.19. Nech $f\colon X \to Y$ je zobrazenie, $A \subseteq X$, $B \subseteq Y$. Množinu

$$f[A] = \{f(a); a \in A\}$$

nazývame *obrazom* množiny A v zobrazení f. Množinu

$$f^{-1}(B) = \{x \in X; f(x) \in B\}$$

nazývame *vzorom* množiny B v zobrazení f.

Cvičenia

{zobrcvic:ULOZIMPSUR}

Úloha 2.2.1. Dokážte: Ak $g \circ f$ je surjekcia, tak aj g je surjekcia. Platí aj opačná implikácia? Musí byť f surjekcia?

{zobrcvic:ULOZIMPINJ}

Úloha 2.2.2. Dokážte: Ak $g \circ f$ je injekcia, tak f je injekcia.

Úloha 2.2.3. Dokážte: Ak $g \circ f$ je bijekcia, tak f je injekcia a g je surjekcia.

{zobrcvic:ULO6b}

Úloha 2.2.4. Nech $f\colon X \to Y$ je zobrazenie a $X \neq \emptyset$ (t.j. X je neprázdna množina). Potom:
a) f je injekcia práve vtedy, keď existuje g také, že $g \circ f = id_X$.
b) f je surjekcia práve vtedy, keď existuje h také, že $f \circ h = id_Y$.
c) K zobrazeniu f existuje inverzné zobrazenie práve vtedy, keď f je bijekcia. (Tým sme znovu dokázali tvrdenie 2.2.16.)

{zobrcvic:ULO6}

Úloha 2.2.5. Nech $f\colon X \to Y$, $g\colon Y \to X$, $h\colon Y \to X$ sú zobrazenia. Ak g aj h sú inverzné zobrazenia k f, tak $g = h$.

Úloha 2.2.6. Nech M, N sú konečné množiny, M má m prvkov a N má n prvkov. Koľko existuje zobrazení množiny M do množiny N?

Úloha 2.2.7. Nech A je konečná množina a $f\colon A \to A$ je zobrazenie. Dokážte:
a) Ak f je injekcia, tak f je bijekcia.
b) Ak f je surjekcia, tak f je bijekcia.

{zobrcvic:ULO10}

Úloha 2.2.8. Dokážte: Zobrazenie $f\colon X \to Y$ je surjekcia práve vtedy, keď pre každú množinu Z a všetky zobrazenia $g, h\colon Y \to Z$ platí: Ak $g \circ f = h \circ f$, tak $g = h$.

Úloha 2.2.9. Dokážte: Zobrazenie $f\colon X \to Y$ je injekcia práve vtedy, keď pre každú množinu Z a všetky zobrazenia $g, h\colon Z \to X$ platí: Ak $f \circ g = f \circ h$, tak $g = h$.

Úloha 2.2.10$^+$. Dokážte: $f[A \cup B] = f[A] \cup f[B]$, $f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B)$.

{zobrcvic:ULO13}

Úloha 2.2.11$^+$. Ktoré z nasledujúcich tvrdení platia a ktoré neplatia? Zdôvodnite.
a) $f[A \cap B] = f[A] \cap f[B]$
b) $f[A \cap B] \subset f[A] \cap f[B]$
c) $f[A \cap B] \supset f[A] \cap f[B]$
d) $f^{-1}(A \cap B) = f^{-1}(A) \cap f^{-1}(B)$
e) $f^{-1}(A \cap B) \subset f^{-1}(A) \cap f^{-1}(B)$
f) $f^{-1}(A \cap B) \supset f^{-1}(A) \cap f^{-1}(B)$
g) $f[f^{-1}(B)] = B$
h) $f[f^{-1}(B)] \subset B$
i) $f^{-1}(f[A]) = A$
j) $f^{-1}(f[A]) \subset A$
k) $g \circ f[A] = g[f[A]]$

Úloha 2.2.12$^+$. Ak X je množina, tak $P(X)$ budeme označovať množinu všetkých jej podmnožín. Nech $f\colon X \to Y$ je zobrazenie a $g\colon P(X) \to P(Y)$ je zobrazenie definované tak, že $g(A) = f[A]$ pre ľubovoľnú podmnožinu $A \subseteq X$. Dokážte, že f je prosté práve vtedy, keď g je prosté.

2.3 Permutácie

V tejto podkapitole sa budeme zaoberať istým typom zobrazení, ktorý budeme v ďalších častiach využívať.

Definícia 2.3.1. Ak M je konečná množina, tak bijekciu $\varphi\colon M \to M$ budeme nazývať *permutáciou* množiny M.

My sa budeme zaoberať (pre zjednodušenie) len permutáciami množiny $\{1, 2 \ldots, n\}$, kde n je prirodzené číslo. (Každú konečnú množinu M vieme očíslovať číslami od 1 po n, kde n je počet prvkov množiny M.)

Zo strednej školy (z kombinatoriky) poznáte termín permutácia v zmysle preusporiadanie nejakej (konečnej) množiny. Je to v istom zmysle, to isté, čo definujeme tu – zobrazenie z množiny $\{1, 2, \ldots, n\}$ vlastne určuje nejaké poradie prvkov tejto množiny, čiže skutočne zodpovedá nejakému jej preusporiadaniu.

Nie je ťažké si uvedomiť, že počet permutácií množiny $\{1, \ldots, n\}$ je práve $n! = n \cdot (n-1) \cdot (n-2) \cdots 2 \cdot 1$. Máme totiž práve n možností, ako môžeme vybrať prvok $\varphi(1)$. Pri výbere prvku $\varphi(2)$ však už nemôžeme použiť ten istý prvok ako v prvom prípade (inak by zobrazenie φ nebolo injektívne), teda nám zostáva práve $(n-1)$ možností. Pre výber obrazu ďalšieho prvku je už iba $(n-2)$ možností, atď.

Permutáciu $\varphi\colon \{1, \ldots, n\} \to \{1, \ldots, n\}$ budeme zapisovať v tvare

$$\begin{pmatrix} 1 & 2 & \ldots & n \\ \varphi(1) & \varphi(2) & \ldots & \varphi(n) \end{pmatrix},$$

čiže pod každé číslo $1, 2, \ldots, n$ zapíšeme jeho obraz v permutácii φ.

Napríklad permutáciu na množine $\{1, 2, 3\}$, pre ktorú $\varphi(1) = 1$, $\varphi(2) = 3$ a $\varphi(3) = 2$ zapíšeme ako $\left(\begin{smallmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{smallmatrix}\right)$, identickú permutáciu zapíšeme ako $\left(\begin{smallmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{smallmatrix}\right)$. (Budeme ju tiež označovať id, keďže ide o identické zobrazenie.)

Niekedy sa kvôli stručnosti používa označenie, kde sa vynechá prvý riadok obsahujúci čísla $1, 2, \ldots, n$. (Napríklad [Bó] používa iba jednoriadkové označenie. Väčšina kníh s touto tématikou používa jednoriadkový aj dvojriadkový zápis – podľa toho, ktorý sa práve hodí.) My sa budeme pridržiavať označenia, ktoré sme zaviedli, z toho dôvodu, že spomínané zjednodušenie by sa mohlo pliesť s označením pre cykly, o ktorých sa viac dozviete neskôr.

Ďalšou výhodou dvojriadkového zápisu je, že ho môžeme použiť pre ľubovoľnú množinu – jednoriadkový zápis môžeme použiť ak máme na množine M nejaké prirodzené usporiadanie, vďaka ktorému vieme prvý riadok jednoznačne doplniť.

Z tvrdenia 2.2.13 vyplýva, že zložením dvoch permutácií opäť dostaneme permutáciu. Pri skladaní permutácií (a takisto pri skladaní zobrazení) budeme často vynechávať znak $\circ$, teda píšeme $\varphi\tau$ namiesto $\varphi \circ \tau$.

Ak napríklad $\varphi = \left(\begin{smallmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{smallmatrix}\right)$, $\tau = \left(\begin{smallmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{smallmatrix}\right)$, tak $\varphi\tau = \left(\begin{smallmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{smallmatrix}\right)$, $\tau\varphi = \left(\begin{smallmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{smallmatrix}\right)$. Vidíme, že skladanie permutácií vo všeobecnosti nie je komutatívne.

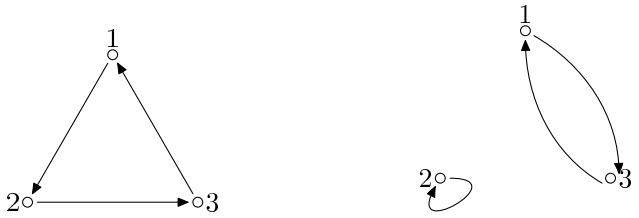
Skladanie permutácií $\varphi\tau$ si môžeme predstaviť tak, že ľavú permutáciu napíšeme pod pravú a preusporiadame stĺpce dolnej permutácie tak, aby jednotlivé čísla súhlasili.

$$\begin{matrix} \tau = \left(\begin{smallmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{smallmatrix}\right) \\ \varphi = \left(\begin{smallmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{smallmatrix}\right) \end{matrix} \quad \longrightarrow \quad \begin{matrix} \tau = \left(\begin{smallmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{smallmatrix}\right) \\ \varphi = \left(\begin{smallmatrix} 3 & 2 & 1 \\ 1 & 3 & 2 \end{smallmatrix}\right) \end{matrix} \quad \longrightarrow \quad \varphi\tau = \left(\begin{smallmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{smallmatrix}\right)$$

Permutácie φ a τ sú znázornené na obrázku 2.7.

To isté môžeme inak vyjadriť tak, že čísla, ktoré sú v dolnom riadku v zápise ľavej permutácie, napíšeme do dolného riadku výslednej permutácie v takom poradí, aké udáva pravá permutácia (tretie, druhé, prvé). Čiže číslo 3, ktoré je na prvom mieste v permutácii τ udáva, že na prvom mieste vo $\varphi\tau$ permutácii bude tretie číslo z φ, čiže 1, 2 na druhom mieste v τ určuje, že na druhom mieste bude to, čo je na druhom mieste vo φ, t.j. 2 a posledná jednotka určuje, že na treťom mieste má byť 2.

Pozor!!! V [KGGS] je skladanie zobrazení (a teda aj skladanie permutácií) definované v opačnom poradí, ako sme ho definovali my.



{perm:FIG3}

Obr. 2.7: Permutácie τ a φ

Inverzné zobrazenie k permutácii je permutácia. Vypočítať ju môžeme jednoducho takým spôsobom, že vymeníme riadky a preusporiadame stĺpce do požadovaného tvaru. Z $\varphi = \left(\begin{smallmatrix}1&2&3\\2&3&1\end{smallmatrix}\right)$ tak dostaneme

$$\varphi^{-1} = \left(\begin{smallmatrix}2&3&1\\1&2&3\end{smallmatrix}\right) = \left(\begin{smallmatrix}1&2&3\\3&1&2\end{smallmatrix}\right).$$

Ak φ je permutácia, tak namiesto $\varphi \circ \varphi$ budeme písať φ^2 a podobne namiesto $\underbrace{\varphi \circ \ldots \circ \varphi}_{n\text{-krát}}$ používame φ^n.

Matematicky správnejšie by bolo povedať, že φ^n definujeme matematickou indukciou:
1° $\varphi^0 = id$, $\varphi^1 = \varphi$
2° $\varphi^{n+1} = \varphi \circ \varphi^n$.

{perm:PRMOCN}

Príklad 2.3.2. Uvažujme permutáciu $\varphi = \left(\begin{smallmatrix}1&2&3&4\\1&3&4&2\end{smallmatrix}\right)$ množiny $\{1,2,3,4\}$. Pokúsme sa vypočítať permutáciu φ^{50}.

Ľahko zistíme, že

$$\varphi^1 = \varphi = \left(\begin{smallmatrix}1&2&3&4\\1&3&4&2\end{smallmatrix}\right)$$
$$\varphi^2 = \varphi \circ \varphi = \left(\begin{smallmatrix}1&2&3&4\\1&4&2&3\end{smallmatrix}\right)$$
$$\varphi^3 = \varphi \circ \varphi^2 = \left(\begin{smallmatrix}1&2&3&4\\1&2&3&4\end{smallmatrix}\right)$$

Ďalšími výpočtami by sme zistili, že sa stále budú opakovať tieto 3 permutácie. (Môžeme to dokázať matematickou indukciou.) Z toho dostaneme $\varphi^{50} = \varphi^{3\cdot 16+2} = \varphi^2 = \left(\begin{smallmatrix}1&2&3&4\\1&4&2&3\end{smallmatrix}\right)$. (Vlastne sme využili vzťah $\varphi^{m.n} = (\varphi^m)^n$, ktorého dôkaz sme ponechali ako cvičenie v úlohe 2.3.4.)

Cvičenia

Úloha 2.3.1. Uvažujme o permutáciach na množine $M = \{1,2,3,4,5\}$. Aká je inverzná permutácia ku: $\left(\begin{smallmatrix}1&2&3&4&5\\1&3&4&2&5\end{smallmatrix}\right)$, $\left(\begin{smallmatrix}1&2&3&4&5\\1&2&3&5&4\end{smallmatrix}\right)$, $\left(\begin{smallmatrix}1&2&3&4&5\\3&1&4&5&2\end{smallmatrix}\right)$? Urobte aj skúšku správnosti, t.j. po vypočítaní φ^{-1} overte, či $\varphi^{-1}\varphi = \varphi\varphi^{-1} = id$. $[\left(\begin{smallmatrix}1&2&3&4&5\\1&4&2&3&5\end{smallmatrix}\right), \left(\begin{smallmatrix}1&2&3&4&5\\1&2&3&5&4\end{smallmatrix}\right), \left(\begin{smallmatrix}1&2&3&4&5\\2&5&1&3&4\end{smallmatrix}\right)]$

Úloha 2.3.2. $M = \{1,2,3,4\}$. Vypočítajte: $\left(\begin{smallmatrix}1&2&3&4\\1&3&4&2\end{smallmatrix}\right)\left(\begin{smallmatrix}1&2&3&4\\1&2&4&3\end{smallmatrix}\right)\left(\begin{smallmatrix}1&2&3&4\\4&1&3&2\end{smallmatrix}\right)$. Určte inverznú permutáciu k výsledku.

{perm:ULO3}

Úloha 2.3.3. Čomu sa rovná φ^{120}, ak $\varphi = \left(\begin{smallmatrix}1&2&3&4\\1&4&2&3\end{smallmatrix}\right)$?

{perm:ULOMI}

Úloha 2.3.4. Matematickou indukciou dokážte, že $\varphi^{n+m} = \varphi^n \circ \varphi^m$, $\varphi^{nm} = (\varphi^n)^m$.

Kapitola 3

Grupy a polia

Naším cieľom je dostať sa k definícii vektorového priestoru a lineárneho zobrazenia. Ich dôležitosť je v tom, že sú vhodným aparátom na popis lineárnych javov.

Predtým nás ale čaká ešte veľmi dlhá cesta, na ktorej sa však naučíme veľmi veľa užitočných vecí. Na definíciu vektorových priestorov použijeme pojem poľa. Na definíciu poľa použijeme pojem grupy. A navyše, predtým než sa dostaneme k definícii grupy, potrebujeme uviesť niektoré základné poznatky o binárnych operáciach.

Táto dlhá cesta k definícii vektorového priestoru je cenou za to, že chceme vektorové priestory definovať v čo najväčšej obecnosti – budeme pracovať s vektorovými priestormi nad ľubovoľným poľom. Keby sme sa rozhodli pracovať iba s vektorovými priestormi nad poľom reálnych čísel, mohli by sme si túto námahu ušetriť – len by sme stručne zopakovali vlastnosti reálnych čísel a hneď by sme zadefinovali vektorový priestor. (V niektorých učebniciach sa takto aj naozaj postupuje.)

Výhoda tohoto postupu je v tom, že dostaneme všeobecnejšie výsledky. Ďalší, vôbec nie zanedbateľný, prínos je, že sa naučíte mnohé užitočné veci a zvyknete si na axiomatický prístup k definovaniu nových pojmov a dokazovaniu ich vlastností.

3.1 Binárne operácie

Hlavným cieľom tejto kapitoly je zadefinovať grupy a polia. Hoci najčastejšie budeme pracovať s reálnymi a komplexnými číslami, je užitočné uvedomiť si, že tvrdenia, ktoré dokážeme budú platiť pre ľubovoľné pole. Aby sme však vôbec mohli zaviesť pojmy grupa a pole, potrebujeme najprv vedieť, čo sú to binárne operácie.

{binop:DEFBINOP}

Definícia 3.1.1. *Binárna operácia* $*$ na množine A je zobrazenie z množiny $A \times A$ do A.

Namiesto $*(a, b)$ budeme používať označenie $a * b$, tento zápis budeme niekedy skracovať ako ab.

Opäť, podobne ako pri zobrazeniach, môžeme binárnu operáciu chápať ako predpis, ktorý však v tomto prípade priradí dvom prvkom z množiny A priradí nejaký prvok z tej istej množiny. Takisto zápis $a * b$ resp. ab zodpovedá tomu, na čo sme zvyknutí pri binárnych operáciach, s ktorými sme sa stretli doteraz, ako je sčitovanie a násobenie.

Tiež si všimnime, že v definícii vystupujú *usporiadané* dvojice (prvky množiny $X \times Y$), teda výsledok operácie $a * b$ a $b * a$ nemusí predstavovať ten istý prvok.

{binop:PR1}

Príklad 3.1.2. Operácie $+$ a $\cdot$ sú binárne operácie na množine $\mathbb{R}$ reálnych čísel.

Operácia $\cdot$ je binárna operácia na množine $\mathbb{R} \setminus \{0\}$, pretože nulu nemôžeme dostať ako súčin dvoch nenulových čísel. Naopak, $+$ nie je binárna operácia na $\mathbb{R} \setminus \{0\}$, lebo $-1+1=0$, teda dvom číslam z množiny $\mathbb{R} \setminus \{0\}$ operácia $+$ priradí číslo, ktoré do tejto množiny nepatrí.

Operácie $+$ aj $\cdot$ sú binárne operácie na množine $\mathbb{R}^+$. Na množine $\mathbb{R}^-$ predstavuje $+$ binárnu operáciu, $\cdot$ však už nie je binárna operácia na tejto množine.

{binop:PRZ5}

Príklad 3.1.3. Ako ďalší príklad definujme operáciu $\oplus$ na množine $\mathbb{Z}_5 = \{0,1,2,3,4\}$ takto: $a \oplus b$ takto: $a \oplus b = (a+b) \bmod 5$. Operácia mod tu predstavuje zvyšok (v tomto prípade po delení piatimi). Napríklad $1 \oplus 2 = 3$, $2 \oplus 3 = 0$, $3 \oplus 3 = 1$ (čísla najprv sčítame a potom urobíme zvyšok po delení 5).

Podobne môžeme definovať binárnu operáciu $\odot$ ako $a \odot b = (a.b) \bmod 5$.

Binárnu operáciu na konečnej množine môžeme tiež určiť tabuľkou: (do riadku a a stĺpca b píšeme výsledok operácie $a \oplus b$.)

$\oplus$	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

$\odot$	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

Podobná operácia by sa dala definovať aj pre ľubovoľné prirodzené číslo $n \geq 2$. S množinami $\mathbb{Z}_n = \{0,1,\ldots,n-1\}$ a binárnymi operáciami $\oplus$ a $\odot$ (čiže sčitovanie a násobenie modulo n) sa v priebehu tejto prednášky stretnete ešte dosť často.

{binop:PRTRIANG}

Príklad 3.1.4. Binárna operácia nemusí byť vždy daná predpisom – ak ide o konečnú množinu môže byť zadaná tabuľkou. (Dalo by sa povedať, že tabuľka binárnej operácie vlastne do istej miery zodpovedá zobrazeniu definovanému rôznymi predpismi pre rôzne prípady. Tu však sú jednotlivé časti definičného oboru len jednoprvkové.)

Napríklad môžeme definovať takúto binárnu operáciu $\triangle$ na množine $\{0,1,2\}$:

$\triangle$	0	1	2
0	0	1	2
1	0	1	2
2	0	2	1

{binop:PRLEFT}

Príklad 3.1.5. Ako posledný príklad binárnej operácie, ktorej vlastnosti budeme vyšetrovať, definujme binárnu operáciu

$$a \lhd b = a$$

na množine $\mathbb{N}$.

Teraz sa budeme zaoberať niektorými základnými vlastnosťami binárnych operácií. Aby sme si ich lepšie ozrejmili, pre každú z nich overíme, či túto vlastnosť majú binárne operácia $+$ a $\cdot$ na množine $\mathbb{R}$, $\oplus$ na množine $\mathbb{Z}_5$, operácia $\triangle$ z príkladu 3.1.4 a operácia $\lhd$ z príkladu 3.1.5.

Vlastnosti, ktoré budeme definovať budú komutatívnosť, asociatívnosť, existencia (ľavého a pravého) neutrálneho prvku a existencia inverzného prvku. Pri overovaní, či uvedené binárne operácie majú túto vlastnosť, postupne vyplníme nasledujúcu tabuľku

	LN	PN	K	A	IP
$(\mathbb{R}, +)$					
$(\mathbb{R}, \cdot)$					
$(\mathbb{Z}_5, \oplus)$					
$(\mathbb{Z}_3, \triangle)$					
$(\mathbb{N}, \lhd)$					

Definícia 3.1.6. Nech $*$ je binárna operácia na množine M. Hovoríme, že $e \in M$ je *ľavý neutrálny prvok* operácie $*$, ak pre všetky $m \in M$ platí

$$e * m = m.$$

Podobne, e je *pravý neutrálny prvok,* ak

$$m * e = m$$

pre každé $m \in M$.

Ak e je súčasne ľavý aj pravý neutrálny prvok operácie $*$, hovoríme, že e je *neutrálny prvok.*

Zo strednej školy vieme, že $0 + m = m + 0 = m$, $1.m = m.1 = m$. To znamená, že 0 je neutrálny prvok pre operáciu $+$ a 1 je neutrálny prvok pre operáciu $\cdot$ na množine $\mathbb{R}$. Pretože rovnosť $0 + m = m + 0 = m$ zostane v platnosti aj keď zo všetkých prvkov urobíme zvyšok po delení 5, 0 je aj neutrálnym prvkom operácie $\oplus$ na množine $\mathbb{Z}_5$.

Prv než sa pozrieme na operáciu $\triangle$, skúsme si ozrejmiť ako sa prejaví existencia neutrálneho prvku na tabuľke binárnej operácie. Rovnosť $e * m = m$ znamená, že v riadku e sa vyskytnú rovnakú prvky ako v hlavičke tabuľky. To isté, ale pre stĺpce, vyplýva z rovnosti $m * e = m$. Skutočne, môžeme si všimnúť v tabuľke operácie $\oplus$ na $\mathbb{Z}_5$, že v riadku 0 a v stĺpci 0 sa opakujú prvky 0,1,2,3,4. Ak sa v prípade operácie $\triangle$ pozrieme na riadky, vidíme, že 0 a 1 sú ľavé neutrálne prvky. Keď skontrolujeme stĺpce, zistíme, že táto operácia nemá pravý neutrálny prvok. Teda táto operácia nemá neutrálny prvok. Priamo z definície operácie $\lhd$ vidno, že v tomto prípade je každý prvok pravým neutrálnym prvkom, ale ľavý neutrálny prvok nemáme ani jeden.

	LN	PN	K	A	IP
$(\mathbb{R}, +)$	0	0			
$(\mathbb{R}, \cdot)$	1	1			
$(\mathbb{Z}_5, \oplus)$	0	0			
$(\mathbb{Z}_3, \triangle)$	0,1	x			
$(\mathbb{N}, \lhd)$	x	✓			

Teraz si dokážeme, že binárna operácia nemôže mať viac ako jeden neutrálny prvok. Spôsob dôkazu je typický pre prípad, keď chceme dokázať, že nejaký objekt je danou vlastnosťou jednoznačne určený. Pri dôkaze tvrdení takéhoto typu sa veľmi často postupuje tak, že uvažujeme dva objekty, ktoré majú túto vlastnosť a snažíme sa dokázať, že sa rovnajú.

{binop:JEDNNEUTR}

Tvrdenie 3.1.7. *Nech $*$ je binárna operácia na množine M. Ak e_1 je jej ľavý neutrálny a e_2 je jej pravý neutrálny prvok, tak $e_1 = e_2$.*

Špeciálne, ak má binárna operácia $*$ na množine M neutrálny prvok, tak tento neutrálny prvok je jediný.

Dôkaz. Ak e_1, e_2 sú ľavý a pravý neutrálny prvok operácie $*$, tak

$$e_1 \overset{(1)}{=} e_1 * e_2 \overset{(2)}{=} e_2,$$

pričom rovnosť (1) platí vďaka tomu, že e_2 je pravý neutrálny prvok a rovnosť (2) platí vďaka tomu, že e_1 je ľavý neutrálny prvok. □

Z toho špeciálne vyplýva, že ak máme viacero ľavých neutrálnych prvkov, žiadny prvok nie je pravým neutrálnym prvkom. Na príklade operácie $\lhd$ sme videli, že jednostranných neutrálnych prvkov môže byť aj nekonečne veľa.

Definícia 3.1.8. Binárna operácia $*$ na množine M je *komutatívna,* ak pre všetky $x, y \in M$ platí

$$x * y = y * x.$$

Komutatívnosť vlastne znamená, že môžeme dvojice prvkov vymieňať. Väčšina operácií, s ktorými budeme pracovať, je komutatívna.

Opäť operácia $+$ a $\cdot$ na $\mathbb{R}$ sú komutatívne. V prípade operácie $\oplus$ si stačí uvedomiť, že rovnosť $x + y = y + x$ sa zachová, aj keď urobíme zvyšok modulo 5, teda dostaneme

$$(x + y) \bmod 5 = (y + x) \bmod 5,$$
$$x \oplus y = y \oplus x.$$

Operácia $\triangle$ nie je komutatívna. Stačí si všimnúť, že

$$0 \triangle 2 \neq 2 \triangle 0,$$
$$2 \neq 0.$$

Podobne sa dá overiť, že operácia $\lhd$ na množine $\mathbb{N}$ nie je komutatívna.

	LN	PN	K	A	IP
$(\mathbb{R}, +)$	0	0	✓		
$(\mathbb{R}, \cdot)$	1	1	✓		
$(\mathbb{Z}_5, \oplus)$	0	0	✓		
$(\mathbb{Z}_3, \triangle)$	0,1	x	x		
$(\mathbb{N}, \lhd)$	x	✓	x		

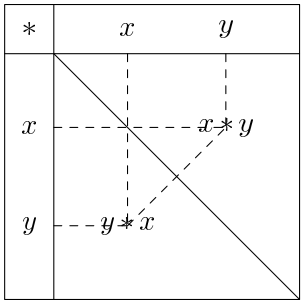
Opäť je užitočné si všimnúť, ako sa komutatívnosť prejaví na tabuľke binárnej operácie. Ak je binárna operácia komutatívna, musí byť tabuľka symetrická podľa hlavnej diagonály (pretože prvky $x * y$ a $y * x$ sú na diagonálne symetrických pozíciach).

Definícia 3.1.9. Binárna operácia $*$ na množine M je *asociatívna,* ak pre všetky $x, y, z \in M$ platí

$$(x * y) * z = x * (y * z).$$

Uzátvorkovanie v predchádzajúcej rovnosti znamená, ktorú operáciu robíme ako prvú. Formálne sa môžeme na asociatívny zákon pozerať ako na „prehadzovanie zátvoriek“.

$$\begin{array}{ccccccc} (& x & * & y &) & * & z \\ & \searrow & & & & \searrow & \\ x & * & (& y & * & z &) \end{array}$$



Obr. 3.1: Komutatívnosť a tabuľka binárnej operácie

Niekedy, v zložitejších výrazoch, aby sme znázornili, kde presne sme použili asociatívny zákon, podčiarkneme na ľavej strane rovnosti tie zátvorky, ktoré „prehadzujeme“.

$$\underline{(}x * y\underline{)} * z = x * (y * z).$$

Vlastnosť asociatívnosti vlastne hovorí to, že nezáleží na uzátvorkovaní, inak povedané zápis $x * y * z$ predstavuje ten istý prvok, bez ohľadu na to, aké uzátvorkovanie zvolíme. Preto zátvorky môžeme vynechávať.

Hoci vlastnosť asociatívnosti tak, ako sme ju definovali, hovorí len o tom, že zátvorky môžeme vynechať, ak ide o súčin 3 prvkov, nie je ťažké si uvedomiť, že to platí aj pre viac prvkov. K tomuto faktu sa ešte vrátime na konci tejto podkapitoly.

O operáciach $+$ a $\cdot$ vieme, že sú asociatívne. Pre operáciu $\oplus$ môžeme použiť podobnú úvahu ako pri komutatívnosti. Keď si všimneme, že

$$2\triangle(0\triangle 2) = 2\triangle 2 = 1,$$
$$(2\triangle 0)\triangle 2 = 0\triangle 2 = 2,$$

vidíme, že operácia $\triangle$ nie je asociatívna

Operácia $\lhd$ je asociatívna, lebo pre ľubovoľné $a, b, c, \in \mathbb{N}$ platí

$$a \lhd (b \lhd c) = a \lhd b = a$$
$$(a \lhd b) \lhd c = a \lhd c = a$$

Môžeme teda doplniť ďalší stĺpec našej tabuľky.

	LN	PN	K	A	IP
$(\mathbb{R}, +)$	0	0	✓	✓	
$(\mathbb{R}, \cdot)$	1	1	✓	✓	
$(\mathbb{Z}_5, \oplus)$	0	0	✓	✓	
$(\mathbb{Z}_3, \triangle)$	0,1	x	x	x	
$(\mathbb{N}, \lhd)$	x	✓	x	✓	

Teraz nasleduje definícia inverzného prvku. O inverznom prvku má zmysel hovoriť iba vtedy, ak má binárna operácia neutrálny prvok. Opäť má zmysel hovoriť o ľavom a pravom inverznom prvku.

Definícia 3.1.10. Nech $*$ je binárna operácia na množine M. Nech $a \in M$ a nech e je neutrálny prvok operácie $*$. Prvok $b \in M$ je *inverzný* k prvku a, ak platí

$$a * b = b * a = e.$$

V prípade, že platí $a * b = e$, b nazývame *pravý inverzný prvok* k a. Ak $b * a = e$, tak b je *ľavý inverzný prvok* k b.

{binop:PRASOC}

Príklad 3.1.11. Pre operáciu + platí

$$a + (-a) = 0 = (-a) + a,$$

teda ľubovoľný prvok $a \in \mathbb{R}$ má inverzný prvok a je to prvok $-a$. (Ako sme spomenuli pred chvíľou, pretože táto operácia je komutatívna, stačí vlastne overovať len jednu z uvedených rovností.)

Podobne pre komutatívnu operáciu $\cdot$ bude inverzným prvkom ku $a \neq 0$ prvok $\frac{1}{a}$, lebo

$$a \cdot \frac{1}{a} = 1.$$

Číslo 0 nemá v tejto operácii inverzný prvok, lebo $0.b = 0$ pre ľubovoľné $b \in \mathbb{R}$.

V prípade množiny $\mathbb{Z}_5$ a operácie $\oplus$ platí $0 \oplus 0 = 1 \oplus 4 = 2 \oplus 3 = 0$, teda ku každému prvku sme našli inverzný prvok. (Všimnime si, že inverzný prvok k a je práve zvyšok čísla $-a$ po delení 5. Vedeli by ste tento fakt zdôvodniť?)

Pre operáciu $\triangle$ nemá zmysel hovoriť o inverznom prvku, lebo táto operácia nemá ani neutrálny prvok. Teraz už teda môžeme konečne vyplniť celú našu tabuľku.

	LN	PN	K	A	IP
$(\mathbb{R}, +)$	0	0	✓	✓	✓
$(\mathbb{R}, \cdot)$	1	1	✓	✓	x
$(\mathbb{Z}_5, \oplus)$	0	0	✓	✓	✓
$(\mathbb{Z}_3, \triangle)$	0,1	x	x	x	x
$(\mathbb{N}, \lhd)$	x	✓	x	✓	x

{binop:JEDNINV}

Tvrdenie 3.1.12. *Nech $*$ je asociatívna operácia na množine M a $*$ má neutrálny prvok e. Ak existuje inverzný prvok k a, tak je jednoznačne určený.*

Dôkaz. Predpokladajme, že b_1 a b_2 sú inverzné prvky ku a. Postupnými úpravami nasledujúcej rovnosti (ktorá vyplýva z asociatívnosti) dostaneme

$$\begin{aligned} b_1 * (a * b_2) &= (b_1 * a) * b_2 \\ b_1 * e &= e * b_2 \\ b_1 &= b_2 \end{aligned}$$

□

Pretože pre asociatívnu operáciu je inverzný prvok jednoznačne určený prvkom a, môžeme preň zaviesť označenie. Budeme ho označovať a^{-1}.

V tabuľke binárnej operácie vieme inverzný prvok nájsť tak, že v riadku a vyhľadáme stĺpec, kde sa vyskytuje neutrálny prvok. To isté urobíme v stĺpci a – nájdeme riadok, v ktorom je neutrálny prvok. Ak sa nájdený riadok a stĺpec zhodujú, tak sme našli inverzný prvok k a.

3.1.1 Zovšeobecnený asociatívny zákon*

Ako sme sľúbili, vrátime sa na chvíľu ešte k asociatívnemu zákonu. Pôjde nám o to, aby sme si uvedomili, že pre asociatívnu operáciu môžeme skutočne vynechávať zátvorky, aj keď vo výraze vystupuje viac prvkov ako 3.

Najprv si to ukážeme pre 4 prvky a potom podáme aj formálny dôkaz, že to platí pre ľubovoľný počet prvkov.

Na začiatok sa pokúste nájsť všetky možné uzátvorkovania výrazu $a \circ b \circ c \circ d$. V nasledujúcej úlohe si môžete overiť, či ste skutočne našli všetky.

Pokúste sa potom pomocou asociatívneho zákona odvodiť, že všetky tieto vyjadrenia sa rovnajú. (Možností ako to dokazovať je veľmi veľa, jednu nájdete v nasledujúcej úlohe.)

{binop:PRASOC4}

Príklad 3.1.13. Dokážte, že ak $\circ$ je binárna operácia na množine A a $\circ$ je asociatívna, tak ľubovoľné uzátvorkovanie výrazu $a \circ b \circ c \circ d$ predstavuje ten istý prvok.

Všetky možné uzátvorkovania sú[1]
(1) $a \circ ((b \circ c) \circ d)$
(2) $a \circ (b \circ (c \circ d))$
(3) $(a \circ b) \circ (c \circ d)$
(4) $((a \circ b) \circ c) \circ d$
(5) $(a \circ (b \circ c)) \circ d$
Sú to skutočne všetky možné uzátvorkovania – najprv sme uviedli tie, kde sa ako posledná operácia vykoná vynásobenie prvkom a zľava, výraz (3) predstavuje jediné uzátvorkovanie, kde sú prvky rozdelené na dve dvojice a výrazy (4) a (5) sú tie uzátvorkovania, kde sa ako posledné vykoná vynásobenie prvkom d.

Podľa asociatívneho zákona platí

$$(b \circ c) \circ d = b \circ (c \circ d).$$

Ak vynásobíme túto rovnosť zľava prvkom a, dostaneme

$$a \circ ((b \circ c) \circ d) = a \circ (b \circ (c \circ d)),$$

čo je vlastne rovnosť (1) = (2).

Podobne, ak použijeme asociatívnosť pre prvky a, b, c a vzniknutú rovnosť vynásobíme sprava prvkom d, dostaneme (4) = (5).

Dvojnásobným použitím asociatívneho zákona (podčiarknutím sú zvýraznené zátvorky, ktoré sme v príslušnej rovnosti „prehodili“) dostaneme

$$a \circ \underline{(b \circ (c \circ d))} = (a \circ b) \circ \underline{(c \circ d)} = ((a \circ b) \circ c) \circ d,$$

čo je vlastne rovnosť (2) = (3) = (4).

Spojením všetkých týchto rovností dostaneme, že všetky uvedené výrazy sa rovnajú.

Predchádzajúci príklad by vás snáď mohol presvedčiť o tom, že niečo podobné platí aj pre uzátvorkovanie ľubovoľného počtu prvkov. Ale vôbec nie je jasné, ako by sa takéto niečo dalo dokázať. To si ukážeme v nasledujúcom (nepovinnom) dôkaze.

{binop:TVRZOVSASOC}

Tvrdenie 3.1.14 (Zovšeobecnený asociatívny zákon)**.** *Nech $\cdot$ je asociatívna binárna operácia na množine A. Potom súčin $a_1 * a_2 * \cdots * a_n$ nezávisí od spôsobu uzátvorkovania.*

Dôkaz. Tvrdenie budeme dokazovať úplnou indukciou vzhľadom na n. Pri dôkaze budeme postupovať tak, že si vyberieme jedno uzátvorkovanie, konkrétne $a_1 * (a_2 * (a_3 * \ldots (a_{n-1} * a_n)))$, a budeme sa snažiť dokázať, že všetky ostatné možné uzátvorkovania predstavujú ten istý prvok.

[1] V prípade, že vám vyšiel iný počet alebo iné uzátvorkovania, skúste si ozrejmiť čo presne rozumieme pod uzátvorkovaním – je to taký zápis, ktorý jednoznačne určuje poradie vykonaných operácií. Pritom poradie prvkov a, b, c, d nesmieme prehadzovať.

1° Pre $n = 2$ máme jediné možné uzátvorkovanie, preto tvrdenie platí.

2° Predpokladajme teraz, že tvrdenie platí pre ľubovoľný počet prvkov menší ako n.

Ak máme nejako uzátvorkovaný výraz $a_1 * a_2 * \ldots * a_n$, sú dve možnosti ako môže vyzerať. Buď má tvar

$$a_1 * \underbrace{(a_2 * a_3 * \ldots * a_n)}_{\text{nejako uzátvorkované}}$$

alebo

$$(a_1 * \ldots * a_k) * (a_{k+1} * \ldots * a_n),$$

kde opäť prvá a druhá zátvorka môžu byť uzátvorkované ľubovoľným spôsobom.

V prvom prípade, podľa indukčného predpokladu

$$a_2 * a_3 * \cdots * a_n = a_2 * (a_3 * \cdots (a_{n-1} * a_n)).$$

(Inými slovami, ľubovoľné uzátvorkovanie $n-1$ prvkov $a_2, a_3, \ldots, a_n$ sa rovná prvku určenému nami zvoleným usporiadaním.) Vynásobením tejto rovnosti zľava prvkom a_1 dostaneme

$$a_1 * (a_2 * a_3 * \cdots * a_n) = a_1 * (a_2 * (a_3 * \cdots * (a_{n-1} * a_n))).$$

Zostáva nám ešte druhý prípad. V tomto prípade najprv použijeme indukčný predpoklad pre obe zátvorky

$$(a_1 * \cdots * a_k) * (a_{k+1} * \cdots * a_n) = (a_1 * (a_2 * \cdots * a_k)) * (a_{k+1} * (a_{k+2} * \cdots * a_n)).$$

Využitím asociatívnosti dostaneme

$$\underline{(}a_1 * (a_2 * \cdots * a_k)\underline{)} * (a_{k+1} * (a_{k+2} * \cdots * a_n)) = a_1 * ((a_2 * \cdots * a_k) * (a_{k+1} * (a_{k+2} * \cdots * a_n))),$$

čím sme upravili daný výraz na tvar súčinu prvku a_1 a nejako uzátvorkovaných ostatných prvkov (čo je presne prvý prípad, ktorý sme už vyriešili).

Teraz teda opäť stačí použiť indukčný predpoklad na prvky v zátvorke a dostaneme požadovaný tvar

$$a_1 * ((a_2 * \cdots * a_k) * (a_{k+1} * (a_{k+2} * \cdots * a_n))) = a_1 * (a_2 * (a_3 * \cdots * (a_{n-1} * a_n))).$$

□

Poznámka* **3.1.15.** Ako zaujímavosť môžeme spomenúť, že počet uzátvorkovaní $n + 1$ symbolov určuje n-té *Catalanove číslo*

$$C_n = \frac{1}{n+1}\binom{2n}{n}.$$

Dôkaz tohoto faktu nie je jednoduchý. Pri niektorých odvodeniach uvedenej formuly sa používa rovnosť

$$C_{n+1} = \sum_{i=0}^{n} C_i C_{n-i},$$

môžete sa skúsiť zamyslieť nad tým, či by ste ju vedeli odvodiť. Súvisí táto rovnosť s nejakým spôsobom (algoritmom) na vymenovanie všetkých možných uzátvorkovaní?

Cvičenia

Úloha 3.1.1. Vypíšte všetky možné binárne operácie na množine $\{0,1\}$. Ktoré z nich sú asociatívne, komutatívne, majú neutrálny prvok? Pre ktoré existuje ku každému prvku aj inverzný?

Úloha 3.1.2. Na $\mathbb{Z}_7 = \{0,1,2,3,4,5,6\}$ definujme operácie $\oplus$ a $\odot$ podobne ako pre $\mathbb{Z}_5$ v príklade 3.1.3. (Teda ako obvyklé sčitovanie a násobenie, ibaže po urobení tejto operácie urobíme zvyšok po delení 7, čím dostaneme prvok zo $\mathbb{Z}_7$.) Zistite, či sú tieto operácie asociatívne, komutatívne, či existuje neutrálny prvok a či má každý prvok inverzný. Vedeli by ste to v prípade operácie $\oplus$ nájsť inverzný prvok aj bez toho, že by ste skúšali jednotlivé prvky?

Úloha 3.1.3. Nájdite binárnu operáciu,
a) pre ktorú aspoň jeden prvok má viacero ľavých inverzných prvkov,
b) ktorá je asociatívna a aspoň jeden prvok má viacero ľavých inverzných prvkov.

Úloha 3.1.4. Na $\mathbb{R}$ definujme operáciu $x * y = x + y + x^2y$. Overte, že každé $x \in \mathbb{R}$ má vzhľadom na túto binárnu operáciu jediný pravý, ale existujú reálne čísla, ktoré nemajú ľavý inverzný prvok.

Úloha 3.1.5*. Ak viete, že ide o tabuľku asociatívnej binárnej operácie, doplňte chýbajúce výsledky (ak sa to dá).

	a	b	c
a	b	a	c
b			
c			

3.2 Grupy

V tejto časti sa konečne dostávame k definícii grupy. Pre nás síce grupy poslúžia len ako pomocný aparát – na to, aby sme mohli jednoduchšie zadefinovať pojem poľa a dokázať niektoré vlastnosti polí – od svojho vzniku sa teória grúp stala samostatnou, veľmi rozsiahlou matematickou disciplínou, zasahujúcou do najrozličnejších oblastí matematiky. Viac sa o grupách dozviete neskôr. Podobne aj o poliach sa v priebehu vášho štúdia dozviete aj ďalšie zaujímavé fakty, na tejto prednáške uvedieme len tie, ktoré budeme potrebovať pri práci s vektorovými priestormi.

Definícia 3.2.1. Dvojica $(G, *)$, kde G je množina a $*$ je binárna operácia na G, sa nazýva *grupa,* ak

(i) operácia $*$ je asociatívna,

(ii) operácia $*$ má neutrálny prvok, (neutrálny prvok budeme spravidla označovať e)

(iii) ku každému prvku $g \in G$ existuje inverzný prvok vzhľadom na operáciu $*$. (Tento inverzný prvok budeme označovať g^{-1}.)

Poznámka 3.2.2. Pretože požadujeme existenciu neutrálneho prvku $e \in G$, z definície grupy automaticky vyplýva, že množina G je neprázdna, $G \neq \emptyset$.

Príklad 3.2.3. Na základe tabuľky, ktorú sme vyplnili v príklade 3.1.11 vidíme, že $(\mathbb{R}, +)$ aj $(\mathbb{Z}_5, \oplus)$ sú grupy.

Naopak, $(\mathbb{R}, \cdot)$ nie je grupa, pretože 0 nemá inverzný prvok. V tomto prípade môžeme situáciu zachrániť, ak zmeníme množinu, na ktorej budeme túto binárnu operáciu uvažovať. Tvrdíme, že $(\mathbb{R} \setminus \{0\}, \cdot)$ je grupa.

O tom, že $\cdot$ je binárna operácia aj na množine $\mathbb{R} \setminus \{0\}$ sme už hovorili v príklade 3.1.2. Neutrálny prvok je 1, toto číslo patrí do množiny $\mathbb{R} \setminus \{0\}$. Takisto asociatívnosť sa neporuší pri prechode ku menšej množine. Pretože sme vynechali nulu, každý prvok $a \in \mathbb{R} \setminus \{0\}$, už teraz má inverzný prvok $\frac{1}{a}$, ktorý tiež patrí do množiny $\mathbb{R} \setminus \{0\}$.

Definícia 3.2.4. Grupa $(G, *)$ sa nazýva *komutatívna,* ak operácia $*$ na G je komutatívna. (Tiež sa používa termín *abelovská grupa.*)

Nie každá grupa je komutatívna. Príkladom nekomutatívnej grupy je grupa S_n všetkých permutácií n-prvkovej množiny pre $n \geq 3$ (úloha 3.2.2).

Veta 3.2.5 (Zákony o krátení)**.** *Ak $(G, *)$ je grupa, tak pre ľubovoľné $a, b, c \in G$ platí*

$$\begin{aligned} a * b = a * c \quad &\Rightarrow \quad b = c \\ b * a = c * a \quad &\Rightarrow \quad b = c \end{aligned}$$

Inak povedané, zákony o krátení hovoria, že v grupe môžeme krátiť ľubovoľným prvkom zľava aj sprava.

Dôkaz. Z rovnosti $a * b = a * c$ dostaneme vynásobením prvkom a^{-1} zľava

$$\begin{aligned} a^{-1} * (a * b) &= a^{-1} * (a * c) \\ (a^{-1} * a) * b &= (a^{-1} * a) * c \\ e * b &= e * c \\ b &= c \end{aligned}$$

(V jednotlivých krokoch sme využili asociatívnosť, definíciu inverzného a neutrálneho prvku.)

Implikácia $b * a = c * a \Rightarrow b = c$ sa dokáže analogicky, ibaže prvkom a^{-1} budeme násobiť sprava. □

Zákony o krátení môžeme zapísať aj v ekvivalentnej podobe (obmenená implikácia, pozri príklad 2.1.11):

$$\begin{aligned} b \neq c \quad &\Rightarrow \quad a * b \neq a * c \\ b \neq c \quad &\Rightarrow \quad b * a \neq c * a \end{aligned}$$

Skúsme si premyslieť, ako sa prejavia zákony o krátení v tabuľke binárnej operácie. Prvky tvaru $a * b$, kde $b \in G$, sú práve prvky v riadku a. Zákon o krátení v tvare $b \neq c \Rightarrow a * b \neq a * c$ teda znamená, že v rôznych stĺpcoch riadku a sa objavia rôzne výsledky operácie $*$. Inak povedané, v žiadnom riadku sa nesmie viackrát vyskytnúť ten istý prvok. Podobne, zákon o krátení zľava hovorí, že v žiadnom stĺpci sa nezopakuje nijaký prvok viackrát.

{grp:VTINVINV}

Veta 3.2.6. *Nech $(G, *)$ je grupa. Potom pre ľubovoľné $a, b \in G$ platí*

$$\begin{aligned} (a^{-1})^{-1} &= a \\ (a * b)^{-1} &= b^{-1} * a^{-1} \end{aligned}$$

Dôkaz. Najprv použijeme dvakrát definíciu inverzného prvku. Inverzný prvok a^{-1} musí spĺňať rovnosť

$$(a^{-1})^{-1} * a^{-1} = e.$$

Definícia inverzného prvku pre a nám dáva

$$a * a^{-1} = e.$$

Porovnaním týchto 2 rovností dostaneme

$$(a^{-1})^{-1} * a^{-1} = a * a^{-1}$$

a zo zákona o krátení vyplýva

$$(a^{-1})^{-1} = a.$$

Aj na dôkaz druhej rovnosti chceme využiť zákony o krátení. Z definície inverzného prvku máme

$$(a * b)^{-1} * (a * b) = e.$$

Aby sme mohli využiť zákon o krátení bolo by dobre, keby sme nejako upravili výraz $(b^{-1} * a^{-1}) * (a * b)$, tak sa pokúsme upraviť ho (budeme používať asociatívny zákon a definíciu inverzného a neutrálneho prvku).

$$\underline{(b^{-1} * a^{-1})} * (a * b) = b^{-1} * (a^{-1} * \underline{(a * b)}) = b^{-1} * ((a^{-1} * a) * b) = b^{-1} * (e * b) = b^{-1} * b = e$$

Zistili sme teda, že $(a * b)^{-1} * (a * b) = (b^{-1} * a^{-1}) * (a * b)$ a zo zákona o krátení potom vyplýva

$$(a * b)^{-1} = b^{-1} * a^{-1}.$$

□

V predchádzajúcom dôkaze sme podrobne rozpisovali každé použitie asociatívneho zákona. Pretože sme už v používaní asociatívneho zákona ostrieľaní, môžeme si prácu zjednodušiť vynechávaním zátvoriek (ktoré si vieme na patričných miestach domyslieť) a uvedenú úpravu zapísať stručnejšie ako

$$b^{-1} * a^{-1} * a * b = b^{-1} * e * b = b^{-1} * b = e.$$

{grp:POZNADIT}

Poznámka 3.2.7. Grupy, s ktorými ste sa doteraz najčastejšie stretli, sú asi $(\mathbb{R}, +)$ a $(\mathbb{R} \setminus \{0\}, \cdot)$. Aj pre grupy vo všeobecnosti sa veľmi často zvykne grupová operácia často označovať $+$ alebo $\cdot$. Vtedy hovoríme o *aditívnom* (pomocou znamienka $+$) alebo *multiplikatívnom* (pomocou $\cdot$) zápise grupovej operácie.

Pri aditívnom zápise sa zvykne inverzný prvok zapisovať ako $-a$, pri multiplikatívnom ako a^{-1} alebo aj $\frac{1}{a}$. Takisto neutrálny prvok sa v závislosti od použitej symboliky niekedy označuje ako 0 alebo 1.

Rozdiel medzi týmito dvoma druhmi zápisu sa napríklad prejaví aj vtedy, ak chceme zapísať n-násobné použitie operácie na prvok a (pozri úlohu 3.2.16). Pri aditívnom zápise sa spravidla používa symbol $n \times a$, pri multiplikatívnom a^n.

My budeme používať označenia tak, ako sme ich zaviedli v tejto kapitole. Teda keď budeme pracovať s grupou vo všeobecnosti, budeme neutrálny prvok označovať e a inverzný prvok a^{-1}. V prípade konkrétnych grúp a hlavne v prípade polí a vektorových priestorov, kde sa priamo v definícii vyskytne binárna operácia označovaná ako $+$ však uprednostníme aditívny zápis. (V oboch spomínaných definíciach na to výslovne upozorníme). Takisto budeme používať aditívny zápis pre grupu $(\mathbb{Z}_n, \oplus)$ a mnohé ďalšie grupy, kde je prirodzené označiť grupovú operáciu znakom $+$. (Napríklad v úlohe 3.2.8, kde sa zaoberáme sčitovaním funkcií, je prirodzené označovať neutrálny prvok znakom 0.)

Cvičenia

{grpcvic:U3}

Úloha 3.2.1. Ktoré z uvedených množín tvoria vzhľadom na dané operácie grupu? V ktorých prípadoch je táto grupa komutatívna?
a) $(\mathbb{Z}, \cdot)$ (celé čísla s obvyklým násobením)
b) $(\mathbb{R}, \cdot)$ (reálne čísla s obvyklým násobením)
c) $(\mathbb{R} \setminus \{0\}, \cdot)$, d) $(\mathbb{C}, +)$, e) $(\mathbb{C}, \cdot)$, f) $(\mathbb{C} \setminus \{0\}, \cdot)$
g) $(\mathbb{R}^2, +)$ (so sčitovaním definovaným po zložkách)
h) $\mathbb{R}$ s operáciou $*$, $a * b = a + b - 1$
i) $\mathbb{R} \setminus \{-1\}$ s operáciou $*$, $a * b = ab + a + b$
j) Množina všetkých párnych celých čísel vzhľadom na sčitovanie.
k) Množina všetkých nepárnych celých čísel vzhľadom na sčitovanie.
l) $(\mathbb{Z}_5, \oplus)$

{grpcvic:SN}

Úloha 3.2.2. Tvoria všetky permutácie na konečnej množine M s operáciou skladania zobrazení grupu? Je táto grupa komutatívna? Urobte tabuľku grupovej operácie v prípade $M = \{1, 2, 3\}$.

Úloha 3.2.3. Je $(\mathbb{R}, *)$, kde $a * b = ab + a + b$, grupa? Ak nie, vedeli by ste vynechať niektorý prvok a z množiny $\mathbb{R}$ tak, aby $(\mathbb{R} \setminus \{a\}, *)$ bola grupa?

Úloha 3.2.4. Nech G je množina všetkých funkcií $f_{a,b} \colon \mathbb{R} \to \mathbb{R}$, ktoré sú tvaru $f_{a,b}(x) = ax + b$ pre nejaké reálne čísla $a, b \in \mathbb{R}$. Tvorí táto množina funkcií s operáciou skladania zobrazení grupu? Je množina $\{f_{a,b}; a, b \in \mathbb{R}, a \neq 0\}$ s operáciou skladania zobrazení grupa? Dostaneme grupu, ak vezmeme len také $a, b \in \mathbb{R}$, že $a = 1$? V tých prípadoch, keď dostaneme grupu, je táto grupa komutatívna?

Úloha 3.2.5. Nech $G = \{z \in \mathbb{C} : |z| = 1\}$. Je G s operáciou $\cdot$ (násobenie komplexných čísel) grupa? Označme $C_n = \{z \in \mathbb{C} : z^n = 1\}$. Je $(C_n, \cdot)$ grupa?

{grpcvic:U4}

Úloha 3.2.6*. Budeme uvažovať o nasledujúcich operáciach s množinami:
$A \cup B = \{x; x \in A \vee x \in B\}$ (zjednotenie)
$A \cap B = \{x; x \in A \wedge x \in B\}$ (prienik)
$A \setminus B = \{x; x \in A \wedge x \notin B\}$ (rozdiel)
$A \div B = \{x; x \in A \Leftrightarrow x \in B\}$ (symetrická diferencia - ekvivalentne ju môžeme definovať ako $A \div B = (A \setminus B) \cup (B \setminus A)$)
Ak X je ľubovoľná množina, $P(X)$ označíme množinu všetkých jej podmnožín. Potom $\cup$, $\cap$, $\setminus$, $\div$ sú binárne operácie na $P(X)$. Je $P(X)$ s niektorou z týchto operácií grupa?

{grpcvic:U10}

Úloha 3.2.7. Označme:
$M_1 = \{f \colon \mathbb{Z} \to \mathbb{Z}; f \text{ je bijekcia}\}$
$M_2 = \{f \in M_1; f(n) = n \text{ pre všetky celé čísla } n \text{ až na konečný počet}\}$
$M_3 = \{f \in M_1; f(n) = n \text{ len pre konečný počet } n\}$.
Ktoré z množín M_1, M_2, M_3 tvoria grupu spolu s operáciou skladania zobrazení?

{grpcvic:FCIESCIT}

Úloha 3.2.8. Nech G je množina všetkých zobrazení $f \colon \mathbb{R} \to \mathbb{R}$. Na tejto množine definujeme operáciu $\oplus$ tak, že $(f \oplus g)(x) = f(x) + g(x)$. Je G s touto operáciou grupa? Ak definujeme $(f \odot g)(x) = f(x) \cdot g(x)$, bude $(G, \odot)$ grupa? Ktoré funkcie treba vynechať, aby sme dostali grupu?

Úloha 3.2.9. Nech $M \neq \emptyset$ je množina a $(G, \circ)$ je grupa. Nech H je množina všetkých zobrazení $f \colon M \to G$. Definujme na H binárnu operáciu $*$ tak, že $(f * g)(x) = f(x) \circ g(x)$. Je $(H, *)$ grupa?

{grpcvic:RN}

Úloha 3.2.10. Na množine $\mathbb{R}^n$ (teda na množine všetkých usporiadaných n-tíc reálnych čísel) definujeme binárnu operáciu + ako $(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n)$. Je $\mathbb{R}^n$ s touto operáciou grupa? (Použili sme symbol + v dvoch rôznych významoch – raz ako operáciu na $\mathbb{R}^n$, ktorú definujeme, a raz ako dobre známe sčitovanie na množine $\mathbb{R}$. Keby sme chceli byť dôslední, zaviedli by sme nový symbol pre operáciu na $\mathbb{R}^n$, napríklad $(x_1, \ldots, x_n) \oplus (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n)$. K tomuto problému – používanie rovnakého symbolu v rôznych významoch – sa ešte vrátime.)

{grpcvic:BIJ1}

Úloha 3.2.11. Ak $(G, \circ)$ je grupa a $a \in G$ je nejaký jej prvok, tak zobrazenie $f_a \colon G \to G$ definované ako $f_a(x) = a \circ x$ je bijekcia.

{grpcvic:BIJ2}

Úloha 3.2.12. Nech $(G, \circ)$ je grupa. Dokážte, že zobrazenie $f \colon G \to G$ definované ako $f(x) = x^{-1}$ je bijekcia.

Úloha 3.2.13*. Nech G je neprázdna množina a $\circ$ je asociatívna binárna operácia na G. Potom G je grupa práve vtedy, keď pre ľubovoľné $a, b \in G$ majú rovnice

$$a \circ x = b$$
$$y \circ a = b$$

riešenie v G (inými slovami, pre ľubovoľné $a, b \in G$ existujú $x, y \in G$, ktoré spĺňajú tieto dve rovnosti.)

{grpcvic:KONECKRAT}

Úloha 3.2.14*. Nech G je konečná množina a $\circ$ je binárna operácia na G taká, že platí asociatívny zákon a zákony o krátení. Dokážte, že G je grupa.

Úloha 3.2.15*. Dokážte, že v konečnej grupe, ktorá má párny počet prvkov, existuje prvok rôzny od neutrálneho prvku taký, že $a \circ a = e$.

{grpcvic:ANAN}

Úloha 3.2.16. Nech $(G, *)$ je grupa a $a \in G$. Potom pre ľubovoľné $n \in \mathbb{N}$ definujeme matematickou indukciou prvok a^n nasledovne:
$a^0 = e$
$a^{n+1} = a^n * a$.
(Je to presne to, čo by sme intuitívne chápali pod zápisom $\underbrace{a * a * \ldots * a}_{n\text{-krát}}$.)

Túto definíciu môžeme rozšíriť aj na záporné čísla tak, že pre $n \in \mathbb{N}$ položíme $a^{-n} = (a^{-1})^n$. Tým je a^n definované pre ľubovoľné $a \in G$ a $n \in \mathbb{Z}$. (Všimnite si, že to korešponduje s označením a^{-1}, ktoré používame pre inverzný prvok.)

Dokážte, že pre ľubovoľné $a, b \in G$ a $m, n \in \mathbb{Z}$ platí:
a) $a^{m+n} = a^m * a^n$,
b) $(a^m)^n = a^{mn}$,
c) ak $a * b = b * a$, tak $a^n * b^n = (a * b)^n$,

grpcvic:ULOSUCINNADRUHU}

Úloha 3.2.17. Nech konečná množina $G = \{e, a_1, \ldots, a_n\}$ tvorí s operáciou $*$ komutatívnu grupu a e je jej neutrálny prvok. Dokážte, že $(a_1 * a_2 * \cdots * a_n)^2 = e$.

Úloha 3.2.18. Nech $*$ je binárna operácia na množine A, taká, že pre každé $a, b, c \in A$ platí $a * (b * c) = (a * c) * b$ a $*$ má neutrálny prvok. Dokážte, že operácia $*$ je komutatívna a asociatívna.

Úloha 3.2.19. Nech $(G, \circ)$ je grupa. Dokážte, že ak $x \circ x = x$, tak $x = e$.

Úloha 3.2.20. Zistite, či $(\mathbb{R}^+ \times \mathbb{R}, \square)$, kde pre každé $(a,b),(c,d) \in \mathbb{R}^+ \times \mathbb{R}$ definujeme $(a,b)\square(c,d) = (2ac, b+d)$, je grupa.

Úloha 3.2.21. Nech $G = \mathbb{R} \times (\mathbb{R} \setminus \{0\})$. Definujme na tejto množine binárnu operáciu $*$ predpisom $(a,b) * (c,d) = (a+bc, bd)$. Je to skutočne binárna operácia? Je $(G,*)$ grupa? Je to komutatívna grupa?

Úloha 3.2.22. Doplňte nasledujúcu tabuľku

	a	b	c	d
a				
b				d
c			d	
d				

tak aby ste dostali grupu.

Úloha 3.2.23. Ak pre každý prvok x grupy $(G,\circ)$ platí $x \circ x = e$, tak táto grupa je komutatívna.

Úloha 3.2.24. Je množina $\mathbb{Q}$ s operáciou $\triangleleft$ definovanou ako $a \triangleleft b = ab - a$ grupa? Je táto operácia komutatívna? Má ľavý neutrálny prvok? Má pravý neutrálny prvok?

Úloha 3.2.25. Nech $*$ je asociatívna binárna operácia na množine M, ktorá má neutrálny prvok e. Ak pre nejaké $x \in M$ platí $x * x = x$ a ku x existuje ľavý inverzný prvok, tak $x = e$.

{grpcvic:ULOHLEFTNPIP}

Úloha 3.2.26*. Nech $*$ je binárna operácia na množine G, ktorá
a) je asociatívna,
b) má ľavý neutrálny prvok t.j. existuje prvok $e \in G$ taký, že $(\forall x \in G) e * x = x$
c) pre každý prvok $x \in G$ existuje $y \in G$ také, že $y * x = e$ (kde e označuje prvok z časti b) t.j.

$$(\forall x \in G)(\exists y \in G) y * x = e$$

(stručne môžeme povedať, že ku každému prvku existuje ľavý inverzný prvok vzhľadom na e).
Dokážte, že potom $(G,*)$ je grupa.

{grpcvic:ULOKONECKOMUTRA

Úloha 3.2.27. Nech G je konečná grupa, $|G| = n$. Neutrálny prvok tejto grupy označme e a jej prvky označme ako $a_1, \ldots, a_n$ (t.j. $G = \{a_1, \ldots, a_n\}$).
a) Ukážte, že pre ľubovoľné $a \in G$ platí $G = \{aa_1, \ldots, aa_n\}$.
b) Ukážte, že pre ľubovoľné $a \in G$ platí $a^n = e$.
(Poznámka: Takéto tvrdenie platí aj pre nekomutatívnej grupy – v tom prípade ale treba použiť iný argument. Dá sa to odvodiť napríklad ako dôsledok Lagrangeovej vety.)

3.3 Polia

{pole:SECT}

V tejto podkapitole zadefinujeme polia. Základnými príkladmi polí sú reálne čísla, racionálne čísla a komplexné čísla. Väčšina vlastností, o ktorých budeme hovoriť, vám preto bude dobre známa.

Výhoda toho, že sa budeme mnohé užitočné vlastnosti týchto známych číselných oborov dokázať pomocou niekoľkých jednoduchých základných axióm spočíva v tom, že aj pre iné číselné množiny, ktoré spĺňajú axiómy z definície poľa, budeme môcť automaticky použiť všetky výsledky, ktoré si dokážeme o poliach vo všeobecnosti.

{pole:DEF}

Definícia 3.3.1. Nech F je množina, $+$ a $\cdot$ sú binárne operácie na F. Hovoríme, že trojica $(F,+,\cdot)$ je *pole,* ak

{pole:1}

(i) $(F, +)$ je komutatívna grupa, jej neutrálny prvok budeme označovať 0;

{pole:2}

(ii) $(F \setminus \{0\}, \cdot)$ je komutatívna grupa, jej neutrálny prvok budeme označovať 1;

{pole:3}

(iii) pre ľubovoľné $a, b, c \in F$ platí

$$a(b + c) = ab + ac,$$
$$(a + b)c = ac + bc.$$

(Túto vlastnosť nazývame *distributívnosť*.)

Pre inverzný prvok v grupe $(F, +)$ budeme používať označenie $-a$, t.j. pre túto grupu používame aditívny zápis. Prvok $-a$ nazývame *opačný prvok* k prvku a.

Pre grupu $(F \setminus \{0\}, \cdot)$ budeme používať multiplikatívny zápis, teda inverzný prvok k prvku $a \neq 0$ poľa F vzhľadom na operáciu $\cdot$ budeme značiť a^{-1}. Ak použijeme termín *inverzný prvok* v súvislosti s poľom a nešpecifikujeme binárnu operáciu, myslí sa tým práve prvok a^{-1}.

Namiesto $b + (-c)$ budeme používať stručnejší zápis $b - c$.

O operáciach + a $\cdot$ v poli F budeme niekedy hovoriť ako o sčitovaní a násobení (súčte a súčine), presne tak ako je to v najzákladnejších príkladoch polí.

Aby bolo jasné, ktorá operácia sa vykoná najskôr, mali by sme používať zápis ako napríklad $(a \cdot b) + (c \cdot d)$. Budeme používať rovnakú konvenciu, aká je zaužívaná pre reálne čísla – operácia $\cdot$ má vyššiu prioritu ako operácia +, teda predchádzajúci zápis môžeme stručnejšie zapísať ako $ab + cd$.

Príklad 3.3.2. $(\mathbb{Q}, +, \cdot)$, $(\mathbb{R}, +, \cdot)$ a $(\mathbb{C}, +, \cdot)$ sú polia. Vlastnosti (i) a (ii) sme overili v časti o grupách (resp. v cvičeniach za ňou), zo strednej školy vieme, že pre tieto číselné obory platí aj distributívnosť.

Pole by sme mohli ekvivalentne definovať aj nasledujúcim spôsobom. (Všimnite si, že pojem grupy nám umožnil túto definíciu zapísať oveľa stručnejšie.)

{pole:DEF2}

Definícia 3.3.3. Pole je množina F, na ktorej sú definované 2 binárne operácie + a $\cdot$ spĺňajúce:

(i) pre všetky $a, b, c \in F$ platí $a + (b + c) = (a + b) + c$,

(ii) pre všetky $a, b \in F$ platí $a + b = b + a$,

(iii) existuje prvok $0 \in F$ taký, že pre každé $a \in F$ sa $a + 0 = a$,

(iv) ku každému $a \in F$ existuje $b \in F$ tak, že $a + b = 0$,

(v) pre všetky $a, b, c \in F$ platí $a \cdot (b \cdot c) = (a \cdot b) \cdot c$,

(vi) pre všetky $a, b \in F$ platí $a \cdot b = b \cdot a$,

(vii) existuje prvok $1 \in F$ taký, že $1 \neq 0$ a pre každé $a \in F$ sa $a \cdot 1 = a$,

(viii) ku každému $a \in F$, $a \neq 0$ existuje $b \in F$ tak, že $a \cdot b = 1$,

(ix) pre všetky $a, b, c \in F$ sa $a \cdot (b + c) = a \cdot b + a \cdot c$.

Overenie ekvivalentnosti týchto 2 definícií ponechávame ako cvičenie (úloha 3.3.1). Možno vám pri tom pomôžu niektoré zo základných vlastností poľa, ktoré odvodíme v nasledujúcom tvrdení. (My budeme používať definíciu 3.3.1. Samozrejme, akonáhle viete dokázať ekvivalentnosť oboch definícií, môžete používať ktorúkoľvek z nich.)

{pole:TVR1}

Tvrdenie 3.3.4. *Nech $(F, +, \cdot)$ je pole. Potom pre $a, b, c \in F$ platí*

{pole:1.1}

(i) $a \cdot 0 = 0,\ 0 \cdot a = 0,$ {pole:1.1a}

(ii) $a \cdot b = b \cdot a,$ {pole:itONE}

(iii) $1 \cdot a = a \cdot 1 = a,$ {pole:1.2}

(iv) $(-a) \cdot b = -a \cdot b,$ {pole:1.3}

(v) $(-a) \cdot (-b) = a \cdot b,$ {pole:1.4}

(vi) $a \cdot b = 0 \Rightarrow a = 0 \vee b = 0,$ {pole:1.5}

(vii) $a \cdot b = a \cdot c \wedge a \neq 0 \Rightarrow b = c,$ {pole:1.6}

(viii) $a \cdot a = a \Rightarrow a = 0 \vee a = 1.$

Dôkaz. (i) Rovnosť

$$0 + 0 = 0$$

vynásobíme zľava prvkom a. Po použití distributívneho zákona dostaneme

$$a.0 = a.(0 + 0) = a.0 + a.0$$
$$a.0 = 0$$

(V poslednej úprave sme využili zákon o krátení – môžeme ho použiť vďaka tomu, že $(F, +)$ je grupa.)

Rovnosť $0.a = 0$ sa ukáže takmer rovnako. (Prvkom a budeme násobiť sprava.)

(ii) Ak a aj b sú rôzne od nuly, tak tvrdenie vyplýva z komutatívnosti grupy $(F \setminus \{0\}, \cdot)$. Prípad, že niektoré z nich je nulové, je vyriešený v (i).

(iii) Ak $a \neq 0$, tak táto rovnosť vyplýva z toho, že 1 je neutrálny prvok grupy $(F \setminus \{0\}, \cdot)$. Pre $a = 0$ máme $1.0 = 0 = 0.1$ z (i).

(iv) Použitím distributívnosti a definície opačného prvku dostaneme

$$a.b + (-a).b = (a + (-a)).b = 0.b \overset{(i)}{=} 0.$$

Teda $(-a).b$ je skutočne opačný prvok k $a.b$, čiže naozaj platí

$$(-a).b = -(a.b).$$

(v) Dvojnásobným použitím rovnosti (iv) dostaneme

$$(-a).(-b) = -a.(-b) = -(-a.b) = a.b.$$

(V poslednej rovnosti sme použili fakt, že $-(-a) = a$, čo je vlastne tvrdenie $(a^{-1})^{-1} = a$ z vety 3.2.6 prepísané do aditívneho zápisu.)

(vi) Fakt, že $\cdot$ je binárna operácia na množine $F \setminus \{0\}$ (ktorý je v definícii 3.3.1 ukrytý v tom, že $(F \setminus \{0\}, \cdot)$ je grupa) vlastne hovorí, že

$$a \neq 0 \wedge b \neq 0 \Rightarrow a.b \neq 0.$$

Z tejto implikácie dostaneme tvrdenie (vi) ako obmenenú implikáciu.

(vii) Z rovnosti $ab = ac$ dostaneme

$$a(b - c) = ab - ac = 0.$$

Pretože $a \neq 0$, z (vi) vyplýva $b - c = 0$ a $b = c$.

(viii) Rovnosť $a.a = a$ môžeme upraviť na tvar

$$a \cdot a - a \cdot 1 = 0$$
$$a(a - 1) = 0$$

Na základe (vi) potom dostaneme $a = 0$ alebo $a = 1$. □

Zistiť, že racionálne, reálne a komplexné čísla spĺňajú vlastnosti poľa bolo pomerne jednoduché. Ďalším základným príkladom poľa bude pre nás pole $(\mathbb{Z}_p, \oplus, \odot)$, kde p je prvočíslo. Teraz sa preto budeme venovať definícii operácií $\oplus$ a $\odot$ na množine $\mathbb{Z}_n$ a dokážeme, že ak n je prvočíslo, tak to je skutočne pole.

Definícia 3.3.5. Nech $n \in \mathbb{N}$, $n \geq 2$. Množinu $\mathbb{Z}_n$ definujeme ako $\mathbb{Z}_n := \{0, 1, \ldots, n - 1\}$. (Teda množina $\mathbb{Z}_n$ obsahuje všetky možné zvyšky po delení číslom n.)

Na množine $\mathbb{Z}_n$ zavedieme operácie $\oplus$ a $\odot$ predpisom

$$a \oplus b = (a + b) \bmod n,$$
$$a \odot b = (ab) \bmod n,$$

kde operácia mod označuje zvyšok po delení číslom n (pozri dodatok **??**).

Delenie so zvyškom poznáte zo strednej školy, na pripomenutie si uveďme zopár príkladov.

Príklad 3.3.6. V tomto príklade budeme počítať v $\mathbb{Z}_7$.

$2 \oplus 4 = 6$

$3 \oplus 6 = 2$, pretože $3 + 6 = 9$ a zvyšok 9 po delení 7 je 2 ($9 = 1 \cdot 7 + 2$).

$2 \odot 3 = 6$

$2 \odot 4 = 1$, lebo $2 \cdot 4 = 8$ a $8 = 1 \cdot 7 + 1$ (Tým sme vlastne zistili, že 4 je inverzný prvok k 2 v poli $\mathbb{Z}_7$; čiže $2^{-1} = 4$ a $4^{-1} = 2$.)

$3 \odot 6 = 4$, pretože $3 \cdot 6 = 18$ a $18 = 2 \cdot 7 + 4$

Pokúsme sa vyrátať aj nejaký zložitejší výraz ako napríklad
$3 \odot (4 \oplus 2) \oplus 5 \odot (3 \oplus 6) = 3 \odot 6 \oplus 5 \odot 2 = 4 \oplus 3 = 0.$

Všimnime si, že ten istý výsledok by sme dostali, keby sme najprv použili obvyklé sčitovanie aj násobenie a až na záver urobili zvyšok modulo 7.
$3 \cdot (4 + 2) + 5 \cdot (3 + 6) = 3 \cdot 6 + 5 \cdot 9 = 18 + 45 = 63 = 9 \cdot 7 + 0$
Nie je to náhoda – takto to funguje vždy. Viac sa o tom môžete dočítať v dodatku **??**.

Ak budete mať na pamäti túto zákonitosť, môže vám to niekedy pomôcť pri výpočtoch operácií v $\mathbb{Z}_n$. Využijeme ju aj v nasledujúcom dôkaze (povinnom pre tých, ktorí si trúfajú na A-čko; jediná náročnejšia časť je tam dôkaz existencie inverzného prvku – overenie ostatných podmienok by mal zvládnuť každý).

Budeme potrebovať pripomenúť aj pojem prvočísla, ktorý poznáte zo strednej školy.

Definícia 3.3.7. Číslo $n \in \mathbb{N}$, $n > 1$, nazývame *zloženým číslom,* ak $n = m.k$ pre nejaké $m, k \in \mathbb{N}$ také, že $1 < m, k < n$.

Ak $n \in \mathbb{N}$, $n > 1$, nie je zložené, tak ho nazývame *prvočíslo.*

Číslo 1 nepovažujeme ani za prvočíslo ani za zložené číslo.

Inými slovami, číslo n je prvočíslo, ak nie je deliteľné žiadnymi inými prirodzenými číslami okrem 1 a n. Napríklad 9 je zložené číslo, lebo 9 je násobkom 3, ale 7 je prvočíslo (číslo 7 nie je deliteľné nijakým menším číslom okrem 1).

V dôkaze nasledujúcej vety budeme potrebovať túto dôležitú vlastnosť prvočísel:

$$p \mid mn \Rightarrow p \mid m \vee p \mid n,$$

t.j. ak je súčin dvoch čísel deliteľný prvočíslom p, musí p deliť jedno z týchto čísel. (Túto vlastnosť nebudeme dokazovať, mali by ste ju poznať zo strednej školy.)

Veta 3.3.8. *Ak p je prvočíslo, tak $(\mathbb{Z}_p, \oplus, \odot)$ je pole.*

Dôkaz. Overíme podmienky (i,ii,iii) z definície 3.3.1.

$(\mathbb{Z}_p, \oplus)$ je komutatívna grupa

Ľahko vidno, že $\oplus$ je binárna operácia na množine $\mathbb{Z}_p$.

Asociatívnosť: Pre ľubovoľné $a, b, c \in \mathbb{Z}_p$ platí

$$\begin{gathered}(a + b) + c = a + (b + c)\\ ((a + b) + c) \bmod p = (a + (b + c)) \bmod p\\ (((a + b) \bmod p) + c) \bmod p = (a + ((b + c) \bmod p)) \bmod p\\ (a \oplus b) \oplus c = a \oplus (b \oplus c)\end{gathered}$$

(Pri poslednej úprave sme využili, že nezáleží na tom, že zvyšok po delení číslom p urobíme po každej operácii, alebo najprv urobíme obvyklé sčitovanie/násobenie a až z takto získaného výsledku urobíme zvyšok po delení číslom p.)

Komutatívnosť:

$$\begin{gathered}a + b = b + a\\ (a + b) \bmod p = (b + a) \bmod p\\ a \oplus b = b \oplus a\end{gathered}$$

Neutrálny prvok je 0.

$$\begin{gathered}a + 0 = a\\ (a + 0) \bmod p = a \bmod p = a\\ a \oplus 0 = a\end{gathered}$$

Inverzný prvok k a je $(-a) \bmod p$, čiže zvyšok čísla $-a$ po delení p. Skutočne

$$a \oplus (-a) \bmod p = (a + ((-a) \bmod p)) \bmod p = (a + (-a)) \bmod p = 0 \bmod p = 0.$$

Všimnite si, že sme zatiaľ nikde nevyužili predpoklad, že p je prvočíslo. Táto prvá časť tvrdenia teda platí aj pre zložené čísla. Spomínaný predpoklad však budeme potrebovať v nasledujúcej časti dôkazu.

$(\mathbb{Z}_p \setminus \{0\}, \odot)$ je komutatívna grupa

Pri dôkaze asociatívnosti, komutatívnosti a existencie neutrálneho prvku budeme postupovať takmer rovnako ako v predchádzajúcej časti. V tomto prípade však budeme musieť dať pozor aj na to, či $\odot$ je skutočne binárna operácia na množine $\mathbb{Z}_p \setminus \{0\}$.

Binárna operácia: Chceme ukázať, že ak $a, b \in \mathbb{Z}_p \setminus \{0\}$, tak $a \odot b \neq 0$. Ak by platilo $a \odot b = (a.b) \bmod p = 0$, znamená to, že $p \mid a.b$ (zvyšok čísla $a.b$ po delení p je 0). Keďže p je prvočíslo, tak musí platiť $p \mid a$ alebo $p \mid b$. Lenže v množine $\mathbb{Z}_p \setminus \{0\} = \{1, 2, \ldots, p-1\}$ nie je žiadny násobok p.

Asociatívnosť:

$$\begin{aligned} a.(b.c) &= (a.b).c \\ (a.(b.c)) \bmod p &= ((a.b).c) \bmod p \\ (a.(b.c) \bmod p) \bmod p &= ((a.b \bmod p).c) \bmod p \\ a \odot (b \odot c) &= (a \odot b) \odot c \end{aligned}$$

Komutatívnosť:

$$\begin{aligned} a.b &= b.a \\ (a.b) \bmod p &= (b.a) \bmod p \\ a \odot b &= b \odot a \end{aligned}$$

Neutrálny prvok je 1:

$$\begin{aligned} a \cdot 1 &= a \\ (a \cdot 1) \bmod p &= a \bmod p = a \\ a \odot 1 &= a \end{aligned}$$

Existencia inverzného prvku. Nech $a \neq 0$. Chceme vedieť, či existuje $b \in \mathbb{Z}_p$ také, že $a \odot b = 1$. Najprv si všimnime, že pre $k, l \in \mathbb{Z}_p$ platí implikácia

$$a \odot k = a \odot l \qquad \Rightarrow \qquad k = l$$

(inak povedané, môžeme krátiť nenulovým číslom.)

Ak $(a \cdot k) \bmod p = (a \cdot l) \bmod p$, čiže ak aj al dávajú rovnaký zvyšok po delení p, tak platí $p \mid a \cdot (k - l)$. Pretože p je prvočíslo, nemá vlastných deliteľov, a teda musí deliť buď a alebo $k - l$. Pritom $a \neq 0$ a iné násobky čísla p v $\mathbb{Z}_p$ už nie sú. Teda $p \mid k - l$. Pretože $k, l \in \mathbb{Z}_p$, ich rozdiel je z množiny $\{-(p-1), -(p-2), \ldots, 0, 1, \ldots, p-2, p-1\}$. V tejto množine je jediným násobkom čísla p opäť nula, preto $k - l = 0$ a $k = l$.

Implikácia, ktorú sme práve dokázali, ale znamená, že $a \odot k$, kde $k \in \mathbb{Z}_p$, nadobúda p rôznych hodnôt. (Nemôže nadobudnúť rovnakú hodnotu pre dve rôzne čísla $k \neq k'$.) Pre vhodné číslo $k \in \mathbb{Z}_p$ sa teda objaví každá hodnota zo $\mathbb{Z}_p$, špeciálne aj 1, čo sme chceli dokázať.

Distributívnosť:

$$\begin{aligned} a.(b + c) &= ab + ac, \\ (a.(b + c)) \bmod p &= (ab + ac) \bmod p, \\ a \odot (b \oplus c) &= a \odot b \oplus a \odot c. \end{aligned}$$

□

Môžeme si všimnúť, že prvočíselnosť čísla p sme využívali iba v dvoch častiach dôkazu: $\odot$ je binárna operácia na $\mathbb{Z}_p \setminus \{0\}$ a ku každému nenulovému prvku existuje inverzný.

Všimnite si tiež, že sme v najdôležitejšej časti dôkazu najprv dokázali zákon o krátení a z neho sme odvodili existenciu inverzného prvku. Tento postup súvisí s postupom, ktorý sa dá použiť v úlohe 3.2.14 *.

Tiež si môžeme všimnúť, že dôkaz je len existenčný – ukázali sme existenciu inverzných prvkov, ale v dôkaze sme neuviedli nijaký algoritmus, ako ich hľadať. Zatiaľ teda budeme postupovať tak, že jednoducho vyskúšame všetky možnosti (príklad 3.3.11), čo nie je až také hrozné, keď má pole, ktoré skúmame, málo prvkov. Iná možnosť, ako hľadať inverzný prvok, by bolo použitie malej Fermatovej vety (príklad 3.3.13). Neskôr, na predmete Algebra 2, sa dozviete o Euklidovom algoritme na hľadanie najväčšieho spoločného deliteľa. Ten sa takisto dá použiť na tento účel.

Ak n je zložené, $(\mathbb{Z}_n, \oplus, \odot)$ nie je pole.

Pozrime sa najprv na $\mathbb{Z}_4$ s násobením modulo 4.

Príklad 3.3.9. $(\mathbb{Z}_4, \oplus, \odot)$ nie je pole.

$\odot$	0	1	2	3
0	0	0	0	0
1	0	1	2	3
2	0	2	0	2
3	0	3	2	1

Z tabuľky si môžeme všimnúť, že neplatí viacero vlastností poľa:

- ku prvku 2 neexistuje inverzný prvok vzhľadom na operáciu $\odot$;
- platí $2 \odot 1 = 2 \odot 3$, teda v $\mathbb{Z}_4 \setminus \{0\}$ neplatí zákon o krátení, čiže $(\mathbb{Z}_4 \setminus \{0\}, \odot)$ nie je grupa;
- rovnosť $2 \odot 2 = 0$ ukazuje, že v $\mathbb{Z}_4$ neplatí tvrdenie 3.3.4(vii).

Práve posledná zo spomenutých vlastností sa dá pomerne jednoducho zovšeobecniť na ľubovoľné zložené číslo.

{pole:PRIKLZMN}

Príklad 3.3.10. Ak n je zložené číslo, tak $(\mathbb{Z}_n, \oplus, \odot)$ nie je pole.

Ak n je zložené, znamená to, že $n = m.k$ pre nejaké celé čísla m, k s vlastnosťou $1 < m, k < n$. Špeciálne to znamená, že $m, k \in \mathbb{Z}_n \setminus \{0\}$. Ak na obe strany rovnosti $n = m.k$ použijeme operáciu zvyšok po delení n, dostaneme

$$0 = m \odot k,$$

pričom $m \neq 0$ a $k \neq 0$. Teda aj v tomto prípade sme zistili, že neplatí tvrdenie 3.3.4(vii) a $(\mathbb{Z}_n, \oplus, \odot)$ nemôže byť pole.

Označenie $\oplus$ a $\odot$ korešponduje s našou dohodou, že v poli budeme používať aditívny zápis pre operáciu $+$ a pre operáciu $\cdot$ multiplikatívny zápis. (Jediný rozdiel je, že kvôli odlíšeniu týchto operácií ich dávame do krúžku.) V súlade s touto dohodou budeme označovať inverzný prvok k prvku a vzhľadom na operáciu $\oplus$ ako $-a$ a inverzný prvok vzhľadom na $\odot$ ako a^{-1}.

V nasledujúcom príklade sa budeme zaoberať práve inverznými prvkami vzhľadom na operácie $\oplus$ a $\odot$ v poli $\mathbb{Z}_p$. (Pre jednoduchosť si vyberieme $p = 7$.)

{pole:PRINVZ}

Príklad 3.3.11. Na základe dohody o označovaní opačného prvku v $\mathbb{Z}_7$ platí $-1 = 6$, $-2 = 5$, $-3 = 4$, $-4 = 3$, $-5 = 2$, $-6 = 1$, $-0 = 0$. (Teda inverzný prvok k $a \in \mathbb{Z}_7$ je $7 - a$.) V predchádzajúcom zápise -1 neznamená celé číslo -1 ale opačný prvok k prvku $1 \in \mathbb{Z}_7$.

Využívanie opačných prvkov môže niekedy zjednodušiť výpočty s operáciami $\oplus$ a $\odot$. Napríklad v $\mathbb{Z}_7$ máme
$3 \odot 6 = 3 \odot (-1) = -3 = 4$.
(Súčin $3.(-1)$ sa vyráta ľahšie ako 3.6. Je to len ilustračný príklad – výraznejšie zjednodušenie to prinesie až vtedy, keď počítame viacero operácií a vychádzajú tam väčšie čísla.) Iný príklad: $(2 \oplus 3) \odot (2 \odot 3) = 5 \odot 6 = (-2) \odot (-1) = 2 \odot 1 = 2$. (Využili sme Tvrdenie 3.3.4(v). Pretože sme už dokázali, že $\mathbb{Z}_7$ je pole, môžeme pri výpočtoch používať čokoľvek, čo sme dokázali o poliach vo všeobecnosti.)

Videli sme, že nájsť opačný prvok v $\mathbb{Z}_7$ je jednoduché. S inverzným prvkom je to o niečo komplikovanejšie. Zatiaľ jediný spôsob, ako to môžeme urobiť je vyskúšať všetky možnosti. Skúsme napríklad vypočítať 3^{-1} v $\mathbb{Z}_7$. Kandidáti na inverzný prvok sú 1,2,3,4,5,6. Vypočítame:
$1 \odot 3 = 3$
$2 \odot 3 = 6$
$3 \odot 3 = 2$
$4 \odot 3 = 5$
$5 \odot 3 = 1$
Na piaty pokus sa nám podarilo nájsť prvok, ktorý dáva v súčin s trojkou rovný 1. Teda práve tento prvok je inverzný k 3:
$3^{-1} = 5$.

Ak by sme boli o niečo pozornejší, mohli sme prestať už po druhom kroku. V ňom sme totiž dostali:
$2 \odot 3 = 6 = -1$
z čoho vyplýva
$1 = -(2 \odot 3) = (-2) \odot 3$, čiže $3^{-1} = -2 = 5$. (Takto to funguje aj vo všeobecnosti – vždy nám stačí vyskúšať iba prvú polovicu možností, určite sa tam vyskytne buď 1 alebo -1.)

{pole:DEFAN}

Definícia 3.3.12. Ak n je celé číslo a a, b sú prvky poľa F, tak definujeme $n \times a$ takto:
$0 \times a = 0$,
$(n+1) \times a = n \times a + a$ (zatiaľ sme to indukciou definovali pre prirodzené čísla),
Ak $n > 0$ tak definujeme $(-n) \times a = -(n \times a)$ (tým sme rozšírili definíciu aj na záporné čísla).
Podobne definujeme pre $a \neq 0$:
$a^0 = 1$,
$a^{n+1} = a^n.a$,
$a^{-n} = (a^n)^{-1}$ $(n > 0)$.

Teda tvrdenie 3.3.4(viii) by sme mohli zapísať aj v tvare $a^2 = a \Rightarrow a = 1 \vee a = 0$. (Namiesto $a.a$ budeme stručnejšie písať a^2, pozri definíciu 3.3.12.)

Predchádzajúca definícia je príkladom definície matematickou indukciou (poznámka 2.1.5). Menej formálne to môžeme vyjadriť ako $n \times a = \underbrace{a + a + \cdots + a}_{n\text{-krát}}$ a $a^n = \underbrace{a \cdot a \cdot \cdots \cdot a}_{n\text{-krát}}$.

Nulu sme z definície a^n vynechali preto, že by bol problém definovať 0^z pre $z \leq 0$. Pre prirodzené čísla je výraz 0^n zmysluplný.

Niektoré základné vlastnosti týchto dvoch operácií nájdete v úlohe 3.3.5. (Viaceré z nich použijeme v nasledujúcom príklade.)

{pole:PRMALFERM}

Príklad 3.3.13. Vypočítajme a^6 pre prvky $a \in \mathbb{Z}_7$.
$0^6 = 0$
$1^6 = 1$
$2^6 = (2^3)^2 = 1^2 = 1$

$3^6 = (3^2)^3 = 2^3 = 1$
$4^6 = (-3)^6 = 3^6 = 1$
$5^6 = (-2)^6 = 2^6 = 1$
$6^6 = (-1)^6 = 1^6 = 1$

To, že pre všetky $a \neq 0$ sme dostali $a^6 = 1$ nie je náhoda. Pre ľubovoľné prvočíslo v poli $\mathbb{Z}_p$ platí $a^{p-1} = 1$ (pre nenulové $a \in \mathbb{Z}_p$). Toto tvrdenie je známe ako *malá Fermatova veta,* stretnete sa s ním ešte viackrát. Teoreticko-číselný dôkaz môžete nájsť napríklad v [Č, Sl3]. Návod na iný dôkaz tejto vety (využívajúci algebraické idey) nájdete v úlohe 3.3.14 alebo [KGGS, 174/9*].

Z rovnosti $a^{p-1} = 1$ špeciálne vyplýva, že v $\mathbb{Z}_p$ platí $a^{-1} = a^{p-2}$.

Cvičenia

{pole:CVIC1}

Úloha 3.3.1. Dokážte ekvivalenciu definície 3.3.1 a 3.3.3.

{pole:U6}

Úloha 3.3.2. Ktoré z uvedených množín tvoria spolu s obvyklým sčitovaním a násobením pole?
a) $F = \{a + ib; a \in \mathbb{R}, b \in \mathbb{R}, b \geq 0\}$
b) $F = \{a + ib; a \in \mathbb{Q}, b \in \mathbb{Q}\}$
c) $F = \{a + ib; a \in \mathbb{Z}, b \in \mathbb{Z}\}$
d) $F = \{a + b\sqrt{5}; a \in \mathbb{Q}, b \in \mathbb{Q}\}$
e) $F = \{a + \sqrt{3}ib; a \in \mathbb{Q}, b \in \mathbb{Q}\}$
f) $F = \{a + \frac{b}{\sqrt{2}}; a, b, c \in \mathbb{Q}\}$
g*) $F = \{a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6}; a, b, c, d \in \mathbb{Q}\}$ (Hint: Možno pomôže prepísať si túto množinu do tvaru $F = \{a + b\sqrt{3}; a, b \in F'\}$, kde $F' = \{a + b\sqrt{2}; a, b \in \mathbb{Q}\}$.)
h*) $F = \{a + b\sqrt{2} + c\sqrt{3}; a, b, c \in \mathbb{Q}\}$
i*) $F = \{a + b\sqrt[3]{5}; a \in \mathbb{Q}, b \in \mathbb{Q}\}$

Úloha 3.3.3. V poli $\mathbb{Z}_5$ vyrátajte $2^{-1} \oplus 4$, $(-2) \oplus 4$, $2^{-1} \odot 3$ a $-4 \odot 3^{-1}$.

Úloha 3.3.4. V $\mathbb{Z}_5$ vyrátajte 2^3, $(2^{-1})^4$, $2 \odot (4^{-1})^3$, $(4 \odot 2^{-1})^3$, $(-1)^5 \odot (4 \odot 3^{-1})^2$.

{pole:UL2}

Úloha 3.3.5. Nech m, n sú celé čísla, a, b, $b_1, \ldots, b_n$ sú prvky poľa F. V úlohách f) až j) predpokladáme, že $a \neq 0$. Dokážte:[2]
a) $m \times a + n \times a = (m + n) \times a$
b) $m \times a + m \times b = m \times (a + b)$
c) $m \times (n \times a) = (mn) \times a$
d) $a \cdot (n \times b) = n \times (a \cdot b)$
e) $(m \times a)(n \times b) = (mn) \times (a \cdot b)$
f) $m \times (m \times a)^{-1} = a^{-1}$
g) $a^m \cdot a^n = a^{m+n}$
h) $a^m \cdot b^m = (a \cdot b)^m$
i) $(a^m)^n = a^{mn}$
j) $a^{2k} = (-a)^{2k}$
k) $n \times 0 = 0$
l) $1^n = 1$

[2]Podúlohy by mali byť usporiadané tak, že ak v dôkaze niektorej z nich potrebujeme nejaké pomocné tvrdenie, máme ho už dokázané v niektorej z predchádzajúcich častí tejto úlohy. Ak by sa Vám zdalo, že poradie nie je správne, ozvite sa mi. Môžeme sa spolu pozrieť na to, či som sa pomýlil alebo či je dôvodom odlišného poradia to, že sa to dá dokazovať aj inak.

Úloha 3.3.6. V ľubovoľnom poli F platí:

$$a+b=a+c \Rightarrow b=c$$
$$(a+b)(c+d)=ac+ad+bc+bd$$
$$-(-a)=a$$
$$-0=0$$
$$-(a+b)=(-a)+(-b)$$
$$(a-b)c=ac-bc$$
$$1\neq 0$$
$$a\cdot a=1 \Leftrightarrow a=1 \vee a=-1$$
$$a^2=b^2 \Leftrightarrow a=b \vee a=-b$$
$$a\cdot(b_1+\ldots+b_n)=a\cdot b_1+\ldots+a\cdot b_n$$

Úloha 3.3.7. Na množine $\mathbb{R}^+$ všetkých kladných reálnych čísel zadefinujme operácie $\oplus$ a $\odot$ tak, že $x\oplus y=x.y$ a $x\odot y=x^y$. Ktoré z axióm poľa spĺňa $(\mathbb{R}^+,\oplus,\odot)$?

{pole:UL3}

Úloha 3.3.8. Nech F je pole a $a\in F$. Definujeme zobrazenie $f_a\colon F\to F$ tak, že $f_a(b)=a+b$. Je f_a bijekcia? Ak áno, ako vyzerá zobrazenie f_a^{-1}? Čomu sa rovná $f_a\circ f_b$?

Ďalej definujeme $g_a\colon F\to F$ pre $a\neq 0$ tak, že $g_a(b)=a\cdot b$. Je to bijekcia?

{pole:ATA374}

Úloha 3.3.9. Nech na množine $M=\{0,1\}$ sú operácie + a $\cdot$ dané tabuľkami

+	0	1
0	0	1
1	1	0

$\cdot$	0	1
0	0	0
1	1	1

Ukážte, že $(M,+)$ a $(M\setminus\{0\},\cdot)$ sú komutatívne grupy a že platí distributívny zákon $(a+b)c=ac+bc$. Je $(M,+,\cdot)$ pole?

Úloha 3.3.10. Zistite, či $(\mathbb{R},+,*)$, kde + je obvyklé sčitovanie reálnych čísel a pre každé $a,b\in\mathbb{R}$ $a*b=-2ab$, je pole.

{polecvic:KOMPL}

Úloha 3.3.11. Na $\mathbb{R}\times\mathbb{R}$ definujeme operácie + a $\cdot$ takto:
a) $(a,b)+(c,d)=(a+c,b+d)$ a $(a,b)\cdot(c,d)=(ac,bd)$,
b) $(a,b)+(c,d)=(a+c,b+d)$ a $(a,b)\cdot(c,d)=(ac-bd,ad+bc)$.
c) $(a,b)+(c,d)=(a+c,b+d)$ a $(a,b)\cdot(c,d)=(ac-bd,ad+bc-bd)$
d) $(a,b)+(c,d)=(a+c,b+d)$ a $(a,b)\cdot(c,d)=(bd-ac,ad+bc)$
Je potom $(\mathbb{R}\times\mathbb{R},+,\cdot)$ pole?

Úloha 3.3.12*. Pre ktoré prvky a poľa $\mathbb{Z}_7$ má riešenie rovnica $x^2=a$? Koľko je takých prvkov v poli $\mathbb{Z}_{109}$?

{polecvic:BINOM}

Úloha 3.3.13*. Dokážte, že:
a) V ľubovoľnom poli platí $(a+b)^m=a^m+\binom{m}{1}\times a^{m-1}b+\binom{m}{2}\times a^{m-2}b^2+\ldots+\binom{m}{m-1}ab^{m-1}+b^m$. (Súčet na pravej strane sa zvykne označovať takto: $\sum_{k=0}^{m}\binom{m}{k}\times a^{m-k}b^k$.)
b) V poli $\mathbb{Z}_p$ platí: $(a\oplus b)^p=a^p\oplus b^p$.

{polecvic:FERM}

Úloha 3.3.14*. Pomocou úlohy 3.3.13 dokážte matematickou indukciou vzhľadom na a, že v $\mathbb{Z}_p$ platí rovnosť $a^p=a$ (pre ľubovoľné $a\in\mathbb{Z}_p$). (Toto je vlastne iná formulácia malej Fermatovej vety.)

{polecvic:POLERAD}

Úloha 3.3.15. Nech F je konečné pole a platí $|F|=n$. Ukážte, že potom pre každý prvok $a\in F\setminus\{0\}$ platí $a^{n-1}=1$. (Týmto sme súčasne získali aj iné odvodenie malej Fermatovej vety.)

Kapitola 4

Vektorové priestory

4.1 Vektorový priestor

Na strednej škole ste sa už stretli s pojmom vektoru. Pracovali ste hlavne s vektormi v rovine a v trojrozmernom priestore. Tieto vektory budú tvoriť špeciálny prípad toho, čo budeme nazývať vektorový priestor.

Náš prístup bude opäť axiomatický, čo nám umožní používať dokázané výsledky okrem týchto vektorov aj na mnohé iné prípady. Okrem toho všeobecnosť našich úvah bude väčšia i vďaka tomu, že budeme pracovať nad ľubovoľným poľom.

{vpr:DEF}

Definícia 4.1.1. Nech F je pole a $V \neq \emptyset$ je množina. Nech $+$ je binárna operácia na V a každej dvojici $c \in F$, $\vec{\alpha} \in V$ je priradený prvok $c \cdot \vec{\alpha} \in V$, pričom platí pre ľubovoľné $c, d \in F$ a $\vec{\alpha}, \vec{\beta} \in V$:

{vpr:1}

(i) $(V, +)$ je komutatívna grupa,

{vpr:2}

(ii) $c \cdot (\vec{\alpha} + \vec{\beta}) = c \cdot \vec{\alpha} + c \cdot \vec{\beta}$,

{vpr:3}

(iii) $(c + d) \cdot \vec{\alpha} = c \cdot \vec{\alpha} + d \cdot \vec{\alpha}$,

{vpr:4}

(iv) $(c \cdot d) \cdot \vec{\alpha} = c \cdot (d \cdot \vec{\alpha})$,

{vpr:5}

(v) $1 \cdot \vec{\alpha} = \vec{\alpha}$.

Potom hovoríme, že V je *vektorový priestor* nad poľom F.

Namiesto $c \cdot \vec{\alpha}$ budeme často používať stručnejší zápis $c\vec{\alpha}$.

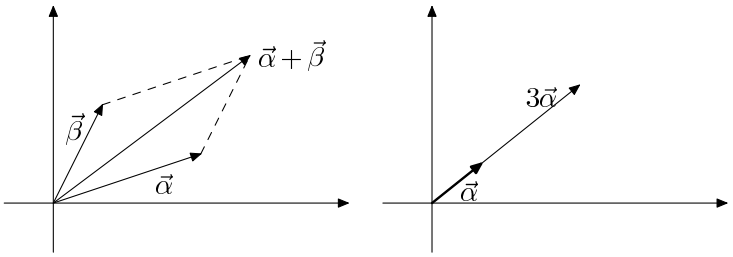
Prvky množiny V budeme nazývať *vektory* a spravidla ich budeme označovať gréckymi písmenami a šípkou. Pre prvky poľa F budeme niekedy používať termín *skaláry.*

Všimnite si, že hoci pre násobenie v poli F aj pre násobenie vektoru skalárom používame rovnaký symbol, z toho, medzi akými objektami sa tento symbol vyskytuje je jasné, ktorú z týchto dvoch možností máme na mysli. (Dalo by sa povedať, že vlastnosť (iv) z definície 4.1.1 hovorí o kompatibilite týchto dvoch operácií.)

Neutrálny prvok komutatívnej grupy $(V, +)$ budeme označovať $\vec{0}$ a nazývať *nulový vektor.*

Inverzný prvok v grupe $(V, +)$ budeme označovať $-\vec{\alpha}$ a nazývame *opačný vektor* k vektoru $\vec{\alpha}$. Vektor $\vec{\alpha} - \vec{\beta} := \vec{\alpha} + (-\vec{\beta})$ sa nazýva *rozdiel* vektorov $\vec{\alpha}$ a $\vec{\beta}$.

Príklad 4.1.2. Vektory v rovine so sčitovaním a násobením ako ho poznáte zo strednej školy, tvoria vektorový priestor nad poľom $\mathbb{R}$ (obr. 4.1).



Obr. 4.1: Operácie s vektormi v rovine {vpr:VECT1}

{vpr:PRRN}

Príklad 4.1.3. Nech $n \in \mathbb{N}$ a $V = \mathbb{R}^n$, teda V pozostáva z usporiadaných n-tíc reálnych čísel, kde n je nejaké prirodzené číslo. Potom $\mathbb{R}^n$ je vektorový priestor nad poľom $\mathbb{R}$.

Sčitovanie vektorov a násobenie skalárom definujeme po zložkách, čím sa myslí

$$(x_1, x_2, \ldots, x_n) + (y_1, y_2 \ldots, y_n) = (x_1 + y_1, x_2 + y_2 \ldots, x_n + y_n)$$

(teda sčítame príslušné súradnice oboch n-tíc)

$$c(x_1, x_2, \ldots, x_n) = (cx_1, cx_2 \ldots, cx_n)$$

(každú súradnicu vynásobíme skalárom c).

Sčitovanie a násobenie použité na jednotlivých súradniciach je už obvyklé sčitovanie a násobenie reálnych čísel (c, x_k aj y_k sú reálne čísla). [1]

Aby sme ukázali, že takto skutočne dostaneme vektorový priestor, treba overiť podmienky z definície 4.1.1. Princíp overenia je v podstate rovnaký u všetkých podmienok z tejto definície: aby sme overili rovnosť dvoch n-tíc, stačí overiť rovnosť na ľubovoľnej súradnici. Keď už pracujeme s niektorou konkrétnou súradnicou, dostaneme rovnosť v poli $\mathbb{R}$, ktorej platnosť vyplýva z toho, že $\mathbb{R}$ spĺňa definíciu poľa.

Vlastnosť (i) sme už overovali v úlohe 3.2.10. Pretože dôkazy ostatných vlastností sú skutočne veľmi podobné, overme pre ilustráciu len vlastnosť (iii) z definície vektorového priestoru. Majme teda ľubovoľný vektor $\vec{\alpha} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ a skaláry $c, d \in F$. Potom

$$(c + d)\vec{\alpha} = ((c + d)x_1, \ldots, (c + d)x_n) = (cx_1 + dx_1, \ldots, cx_n + dx_n) = c\vec{\alpha} + d\vec{\alpha}.$$

(Rovnosť medzi súradnicovými vyjadreniami platí vďaka tomu, že c, d, x_i sú prvky poľa $\mathbb{R}$, teda rovnosť $(c + d)x_i = cx_i + dx_i$ vyplýva z distributívneho zákona.)

V prípade $n = 2$ dostaneme vektorový priestor $\mathbb{R}^2$, čo je vlastne vektorový priestor z predchádzajúceho príkladu v prípade, že vektory v rovine zapíšeme pomocou súradníc.

{vpr:PRRR}

Príklad 4.1.4. Nech $V = \{f \colon \mathbb{R} \to \mathbb{R}\}$, teda V je množina všetkých zobrazení z $\mathbb{R}$ do $\mathbb{R}$. Túto množinu budeme obvykle označovať ako $\mathbb{R}^{\mathbb{R}}$.

Pre $f, g \in V$ a $c \in \mathbb{R}$ zadefinujeme sčitovanie a násobenie nasledovne:

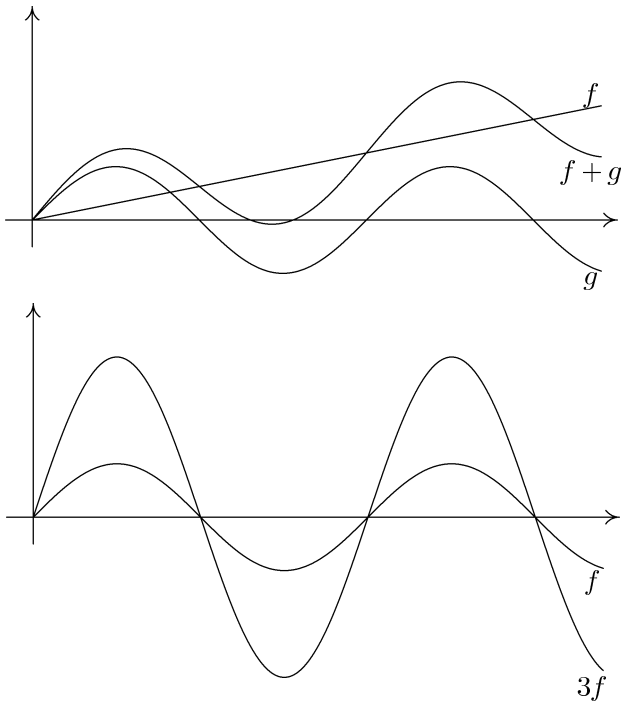
$$(f + g)(x) := f(x) + g(x),$$
$$(c \cdot f)(x) := c \cdot f(x),$$

[1] Všimnite si, že symbol + používame v dvoch významoch: na ľavej strane rovnosti označuje operáciu na množine $\mathbb{R}^n$ a na pravej strane rovnosti operáciu na $\mathbb{R}$. Podobne aj $\cdot$ sa tu vyskytuje v dvoch rozličných významoch. Keby sme chceli byť veľmi dôslední, mali by sme pre operácie s n-ticami zaviesť iné označenie. S podobnou situáciou sme sa už stretli aj v prípade násobenia v poli a vo vektorovom priestore nad týmto poľom. Kvôli stručnosti a jednoduchosti označenia budeme často používať zápisy takéhoto typu. Treba si na to zvyknúť.

kde $x \in \mathbb{R}$. (Na vysvetlenie: v uvedených rovnostiach sú na ľavej strane operácie, ktoré definujeme. Na pravej strane rovnosti ide už o obvyklé sčitovanie a násobenie reálnych čísel – vieme, že $f(x), g(x) \in \mathbb{R}$. Tým, že sme zadefinovali funkčnú hodnotu v každom bode $x \in \mathbb{R}$, sú zobrazenia $f + g, c \cdot f \colon \mathbb{R} \to \mathbb{R}$ jednoznačne určené. Symboly $+$ a $\cdot$ tu vystupujú opäť v dvoch rôznych významoch – podobne ako v predchádzajúcom príklade.)

Inak môžeme predchádzajúcu definíciu preformulovať tak, že v každom bode sčítame funkčné hodnoty resp. prenásobíme funkčnú hodnotu konštantou.

S týmito operáciami tvorí množina $\mathbb{R}^{\mathbb{R}}$ vektorový priestor. Dôkaz tohoto faktu je do istej miery podobný ako pre priestor $\mathbb{R}^n$. V tomto prípade pri dôkaze vlastností vektorového priestoru overujeme rovnosť funkcií, čím sa dostaneme k rovnosti funkcií po dosadení ľubovoľného $x \in \mathbb{R}$ a keď už pracujeme s funkčnými hodnotami, sú to prvky poľa $\mathbb{R}$, čiže môžeme využiť vlastnosti poľa (úloha 4.1.4).



Obr. 4.2: Operácie v priestore $\mathbb{R}^{\mathbb{R}}$ {vpr:VECTRR}

(V tomto prípade sme nepoužili označenie pomocou gréckych písmen, ale označenie, ktoré obvykle používame pre funkcie. Tento príklad by mal ilustrovať, že prvkami vektorového priestoru skutočne môžu byť najrozličnejšie objekty.)

{vpr:POZNFN}

Poznámka 4.1.5. Podobným spôsobom ako pre reálne čísla by sme mohli definovať vektorové priestory F^n a F^M nad ľubovoľným poľom F (pre akékoľvek prirodzené číslo $n \in \mathbb{N}$; resp. pre akúkoľvek množinu M). Overenie, že sú to naozaj vektorové priestory by bolo takmer rovnaké ako v prípade $F = \mathbb{R}$. (Všimnite si, že sme nepoužili žiadnu vlastnosť, ktorá by bola špecifická pre $\mathbb{R}$ a neplatila v ľubovoľnom poli.) S priestorom F^n sa ešte stretneme neskôr.

Podobne ako v prípade polí budú nasledovať niektoré základné vlastnosti, ktoré sa dajú ľahko odvodiť priamo z definície vektorového priestoru.

{vpr:VT1}

{vpr:1.1}
Veta 4.1.6. *Nech V je vektorový priestor nad poľom F, $c \in F$ a $\vec{\alpha} \in V$.*

{vpr:1.2}
(a) $0 \cdot \vec{\alpha} = \vec{0}$,

{vpr:1.3}
(b) $c \cdot \vec{0} = \vec{0}$,

{vpr:1.4}
(c) $c \cdot \vec{\alpha} = \vec{0}$ *práve vtedy, keď* $c = 0$ *alebo* $\vec{\alpha} = \vec{0}$,

(d) $(-c) \cdot \vec{\alpha} = -c \cdot \vec{\alpha}$.

Dôkaz. (a) Keď rovnosť $0 = 0 + 0$ vynásobíme vektorom $\vec{\alpha}$, dostaneme

$$0 \cdot \vec{\alpha} = (0+0)\vec{\alpha} \overset{(iii)}{=} 0 \cdot \vec{\alpha} + 0 \cdot \vec{\alpha}$$

(Využili sme aj vlastnosť (iii) z definície vektorového priestoru.) Zo zákona o krátení (v grupe $(V, +)$) dostaneme $0 \cdot \vec{\alpha} = \vec{0}$.

(b) Budeme postupovať veľmi podobne, tentokrát skalárom $c \in F$ vynásobíme rovnosť $\vec{0} = \vec{0} + \vec{0}$ a použijeme vlastnosť (ii) z definície vektorového priestoru. Dostaneme

$$c \cdot \vec{0} = c \cdot (\vec{0} + \vec{0}) \overset{(ii)}{=} c \cdot \vec{0} + c \cdot \vec{0},$$

z čoho vyplýva (opäť na základe zákona o krátení v grupe $(V, +)$), že $c \cdot \vec{0} = \vec{0}$.

(c) Nech $c \cdot \vec{\alpha} = \vec{0}$ a $c \neq 0$. Potom existuje k prvku c inverzný prvok c^{-1}. Vynásobením uvedenej rovnosti prvkom c^{-1} zľava dostaneme

$$c^{-1}(c \cdot \vec{\alpha}) = c^{-1} \cdot \vec{0}.$$

Ľavú stranu môžeme upraviť ako

$$c^{-1}(c \cdot \vec{\alpha}) \overset{(iv)}{=} (c^{-1} \cdot c)\vec{\alpha} = 1 \cdot \vec{\alpha} \overset{(v)}{=} \vec{\alpha}.$$

Pre pravú stranu máme

$$c^{-1} \cdot \vec{0} \overset{(b)}{=} \vec{0}$$

podľa prvej časti tejto vety. Dostali sme teda rovnosť $\vec{\alpha} = \vec{0}$.

(d) Jednoduchou úpravou dostaneme

$$c \cdot \vec{\alpha} + (-c).\vec{\alpha} \overset{(iii)}{=} (c - c)\vec{\alpha} = 0 \cdot \vec{\alpha} \overset{(a)}{=} \vec{0}.$$

□

Cvičenia

Úloha 4.1.1. Nech $\vec{\alpha} = (1, 3, 6)$, $\vec{\beta} = (2, 1, 5)$, $\vec{\gamma} = (4, -3, 3)$. Vypočítajte $7\vec{\alpha} - 3\vec{\beta} - 2\vec{\gamma}$, $2\vec{\alpha} - 3\vec{\beta} + \vec{\gamma}$ vo vektorovom priestore $\mathbb{R}^3$. [(-7,24,21), (0,0,0)]

{vprcvic:POSTUP}
Úloha 4.1.2. Ukážte, že F je vektorový priestor nad F.

Úloha 4.1.3. Nech V je množina všetkých postupností reálnych čísel. Pre postupnosti $a = (a_n)_{n=1}^{\infty}$ a $b = (b_n)_{n=1}^{\infty}$ definujeme $a + b = (a_n + b_n)_{n=1}^{\infty}$ a $c.a = (c.a_n)_{n=1}^{\infty}$. Overte, že V s týmito operáciami tvorí vektorový priestor nad poľom $\mathbb{R}$.

cvic:FUNKCIE}

Úloha 4.1.4. Nech M je neprázdna množina, F je pole. Potom množina všetkých zobrazení $f\colon M \to F$ so sčitovaním a násobením definovaným po bodoch (pozri príklad 4.1.4) tvorí vektorový priestor nad poľom F. (Ak sa Vám zdá táto úloha príliš zložitá, riešte ju iba pre $F = M = \mathbb{R}$.)

Skúste si tiež uvedomiť, že týmto spôsobom sme súčasne overili, že priestory F^n (príklad 4.1.3 a poznámka 4.1.5), F^F (príklad 4.1.4 a poznámka 4.1.5) a postupnosti prvkov z F (úloha 4.1.3) tvoria vektorové priestory. (Postupnosti môžeme chápať ako zobrazenia z $\mathbb{N}$ do F. Usporiadané n-tice môžeme chápať ako zobrazenia z $\{1, 2, \ldots, n\}$ do F.)

{vprcvic:FNADF}

Úloha 4.1.5. Nech F je ľubovoľné pole a nech $\vec{\alpha}$ je ľubovoľný prvok. Nech $V = \{\vec{\alpha}\}$. Na V zavedieme operáciu sčitovania ako $\vec{\alpha} + \vec{\alpha} = \vec{\alpha}$ a násobenie skalárom $c.\vec{\alpha} = \vec{\alpha}$ (pre každé $c \in F$). Dokážte, že V je vektorový priestor nad poľom F.

{vprcvic:Z2NA2}

Úloha 4.1.6. Overte, že $\mathbb{Z}_2 \times \mathbb{Z}_2$ so sčitovaním a násobením skalárom definovaným po zložkách tvorí vektorový priestor nad poľom $\mathbb{Z}_2$.

Úloha 4.1.7. Nech F je pole, $V = F^n$. Definujeme $(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n)$, $c(x_1, \ldots, x_n) = (cx_1, \ldots, cx_n)$ pre $c, x_1, \ldots, x_n, y_1, \ldots, y_n \in F$. Potom V je vektorový priestor nad poľom F.

Úloha 4.1.8. Koľko prvkov má vektorový priestor $(\mathbb{Z}_3)^n$? Čomu sa v tomto priestore rovná $\vec{\alpha} + \vec{\alpha} + \vec{\alpha}$?

{vprcvic:INTRV}

Úloha 4.1.9. Overte, že všetky zobrazenia $f\colon \langle 0, 1\rangle \to \mathbb{R}$ so sčitovaním a násobením skalárom definovaným po bodoch tvoria vektorový priestor nad poľom $\mathbb{R}$.

{vprcvic:QNADR}

Úloha 4.1.10. Overte, že $\mathbb{R}$ je vektorový priestor nad $\mathbb{Q}$, $\mathbb{C}$ je vektorový priestor nad $\mathbb{R}$, $\mathbb{C}$ je vektorový priestor nad $\mathbb{Q}$. Je $\mathbb{C}$ vektorový priestor nad $\mathbb{Z}$?

Úloha 4.1.11. Nech V je vektorový priestor nad poľom F, $c, c_1 \ldots c_k \in F$, $\vec{\alpha}, \vec{\alpha}_1, \ldots, \vec{\alpha}_n \in V$. Dokážte, že potom platí $c(\vec{\alpha}_1 + \ldots \vec{\alpha}_n) = c\vec{\alpha}_1 + \ldots + c\vec{\alpha}_n$, $(c_1 + \ldots + c_k)\vec{\alpha} = c_1\vec{\alpha} + \ldots + c_k\vec{\alpha}$. Čomu sa rovná $(c_1 + \ldots + c_k)(\vec{\alpha}_1 + \ldots \vec{\alpha}_n)$?

Úloha 4.1.12. Dokážte, že vo vektorovom priestore V nad poľom F pre každé $\vec{\alpha}, \vec{\beta} \in V$, $c \in F$ platí:
a) $c(\vec{\alpha} - \vec{\beta}) = c\vec{\alpha} - c\vec{\beta}$
b) $c(-\vec{\alpha}) = -c\vec{\alpha}$
c) $(c - d)\vec{\alpha} = c\vec{\alpha} - d\vec{\alpha}$
d) $(-c)(-\vec{\alpha}) = c\vec{\alpha}$
e) $-(\vec{\alpha} + \vec{\beta}) = (-\vec{\alpha}) + (-\vec{\beta})$
f) $\vec{\alpha} - (\vec{\beta} + \vec{\gamma}) = (\vec{\alpha} - \vec{\beta}) - \vec{\gamma}$

Úloha 4.1.13. Pre celé číslo n a vektor $\vec{\alpha}$ definujeme $n \times \vec{\alpha}$ podobným spôsobom, ako sme definovali $n \times a$ pre prvok a nejakého poľa F. Dokážte, že potom platí $n \times (c.\vec{\alpha}) = c.(n \times \vec{\alpha})$.

Úloha 4.1.14. Zistite, či $\mathbb{R} \times \mathbb{R}$ s operáciami $+$ a $\cdot$ definovanými tak, že $(a, b) + (c, d) = (a + c, b + d)$ pre ľubovoľné $(a, b), (c, d) \in \mathbb{R} \times \mathbb{R}$ a $r \cdot (a, b) = (ra, 2rb)$ pre ľubovoľné $r \in \mathbb{R}$, je vektorový priestor nad $\mathbb{R}$.

Úloha 4.1.15. Zistite, či $(\mathbb{R}^+, \oplus, \odot)$ je vektorový priestor nad $\mathbb{R}$, ak definujeme $x \oplus y = xy$, $c \odot x = x^c$ pre $x, y \in \mathbb{R}^+$, $c \in \mathbb{R}$.

4.2 Podpriestory

{ppr:DEF}

Definícia 4.2.1. Ak V je vektorový priestor nad poľom F, $S \neq \emptyset$ a $S \subseteq V$, tak S nazveme *podpriestorom* (alebo tiež *vektorovým podpriestorom*) priestoru V, ak

{ppr:1}

(i) pre ľubovoľné $\vec{\alpha}, \vec{\beta} \in S$ platí $\vec{\alpha} + \vec{\beta} \in S$,

{ppr:2}

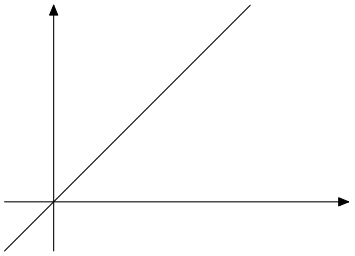
(ii) pre ľubovoľné $\vec{\alpha} \in S$ a $c \in F$ platí $c\vec{\alpha} \in S$.

Inými slovami, podpriestor vektorového priestoru V je taká podmnožina S, ktorá je uzavretá vzhľadom sčitovanie aj vzhľadom na násobenie skalárom.

{ppr:POZNULA}

Poznámka 4.2.2. Všimnime si, že každý podpriestor S priestoru V musí obsahovať nulový vektor $\vec{0}$. Vyplýva to z toho, že $S \neq \emptyset$, teda obsahuje aspoň jeden vektor $\vec{\alpha}$. Z uzavretosti na násobenie skalárom vyplýva, že musí obsahovať aj vektor $\vec{0} = 0.\vec{\alpha}$.

Príklad 4.2.3. Priamku v rovine, ktorá prechádza počiatkom súradnicovej sústavy, môžeme chápať ako množinu vektorov (obrázok 4.3). Táto množina tvorí jej vektorový podpriestor.



{ppr:FIGVECT3}

Obr. 4.3: Priamka v rovine ako príklad vektorového podpriestoru

Príklad 4.2.4. Nech $V = \mathbb{R}^3$ a [2]

$$S = \{(x, y, z) \in \mathbb{R}^3; x + y + z = 0\}.$$

Potom S je podpriestor priestoru V. Overíme podmienky z definície podpriestoru.

Ak $\vec{\alpha} = (x, y, z) \in S$ a $\vec{\beta} = (x', y', z')$, znamená to, že

$$\begin{aligned} x + y + z &= 0, \\ x' + y' + z' &= 0. \end{aligned}$$

Sčítaním týchto 2 rovníc dostaneme

$$(x + x') + (y + y') + (z + z') = 0,$$

preto aj vektor $\vec{\alpha} + \vec{\beta} = (x + x', y + y', z + z')$ spĺňa podmienku, pomocou ktorej sme definovali podmnožinu S.

[2] Hoci v tomto príklade nám stačí overiť definíciu, nie je na škodu uvedomiť si aj to, čo predstavujú vektory zo zadaného podpriestoru geometricky. Zo strednej školy by ste mali vedieť, že pre ľubovoľné reálne čísla $a, b, c, d \in \mathbb{R}$ také, že $(a, b, c) \neq 0$, množina bodov z $\mathbb{R}^3$ vyhovujúca rovnici $ax + by + cz = d$ je rovina. Navyše viete, že (a, b, c) je normálový vektor tejto roviny. Teda v našom prípade máme rovinu kolmú na vektor $(1, 1, 1)$ prechádzajúcu počiatkom súradnicovej sústavy.
Podobne ak by sme pracovali v $\mathbb{R}^2$, tak množina bodov v rovine vyhovujúcich rovnici $ax + by = c$ (kde $a, b \in \mathbb{R}$ sú nejaké reálne čísla a $(a, b) \neq 0$), je priamka, ktorej normálový vektor je (a, b).

Ak $x + y + z = 0$, tak prenásobením konštantou c dostaneme

$$cx + cy + cz = 0,$$

teda $c\vec{\alpha} = (cx, cy, cz) \in S$.

Príklad 4.2.5. Ak V je ľubovoľný vektorový priestor, tak $S = \{\vec{0}\}$ je podpriestor priestoru V.

Skutočne, jediný možný súčet, ktorý môžeme dostať z prvkov S je $\vec{0} + \vec{0} = \vec{0}$, a táto množina je uzavretá aj vzhľadom na násobenie skalárom, $c\vec{0} = \vec{0}$.

Ľahko sa dá overiť aj to, že V je podpriestor V, t.j. každý vektorový priestor je podpriestorom samého seba.

{ppr:POZNKVANT}

Poznámka 4.2.6. Ak S je podpriestor vektorového priestoru V nad poľom F, tak aj S je vektorový priestor nad F (s operáciami rovnako definovanými ako v priestore V, inak povedané, „zdedenými“ z V).

To nám dáva ďalšiu možnosť, ako overiť, že nejaká množina je vektorový priestor nad F. V prípade, že ide o podmnožinu nejakého iného vektorového priestoru (pre ktorý sme už overili všetky vlastnosti), stačí nám len overiť podmienky z definície podpriestoru.

Pri vysvetlení, prečo to platí, si môžeme uvedomiť aj o niečo všeobecnejšie pravidlá, ktoré fungujú pri overovaní axióm nejakého tvaru. Chceme overiť, či S je vektorový priestor – pri tom máme overiť viacero vlastností.

Ako prvé vlastnosti máme podmienky, že $+$ je binárna operácia na S a násobenie má prvku $c \in F$ a vektoru $\vec{\alpha} \in S$ priradiť opäť prvok z S. To zabezpečia podmienky (i), (ii) z definície vektorového podpriestoru.

Ďalšou podmienkou je, že $(V, +)$ je komutatívna grupa. Požiadavku asociatívnosti môžeme zapísať v tvare

$$(\forall \vec{\alpha}, \vec{\beta}, \vec{\gamma} \in S)(\vec{\alpha} + \vec{\beta}) + \vec{\gamma} = \vec{\alpha} + (\vec{\beta} + \vec{\gamma}).$$

Teda je to výrok, ktorý hovorí, že pre všetky prvky z S má platiť nejaká rovnosť. Pretože ale už vieme, že táto vlastnosť platí pre ľubovoľné prvky z väčšej množiny V, tým skôr musí platiť pre ľubovoľné prvky z jej podmnožiny S.

Analogický argument samozrejme funguje aj pre každú vlastnosť, ktorú môžeme zapísať len pomocou všeobecného kvantifikátora a nejakej rovnosti. Vďaka tomu pre S nemusíme overovať ani vlastnosti (ii,iii,iv,v) z definície vektorového priestoru (definícia 4.1.1) a komutatívnosť operácie $+$.

Keď overujeme existenciu neutrálneho a inverzného prvku v $(S, +)$, sme v trochu inej situácii. Existenciu neutrálneho prvku môžeme zapísať v tvare

$$(\exists \vec{\varepsilon} \in S)\ (\forall \vec{\alpha} \in S)\ \vec{\varepsilon} + \vec{\alpha} = \vec{\alpha},$$

tentokrát v našej podmienke vystupuje okrem všeobecného kvantifikátora aj existenčný kvantifikátor. Z poznámky 4.2.2 vieme, že $\vec{0}$ patrí do S. Vieme, že $\vec{0}$ je neutrálny prvok vo $(V, +)$. Teda spĺňa podmienku

$$(\forall \vec{\alpha} \in S)\ \vec{0} + \vec{\alpha} = \vec{\alpha}.$$

O podmienkach takéhoto tvaru sme sa pred chvíľou už presvedčili, že sa dedia na podmnožiny. Súčasne vieme, že $\vec{0} \in S$. Preto $\vec{0}$ je neutrálny prvok aj v $(S, +)$.

Existenciu inverzného prvku môžeme zapísať podmienkou

$$(\forall \vec{\alpha} \in S)\ (\exists \vec{\beta} \in S)\vec{\alpha} + \vec{\beta} = \vec{0},$$

ktorá hovorí, že ku každému vektoru $\vec{\alpha}$ má existovať inverzný prvok na sčitovanie. Vieme však, že $\vec{\alpha}$ má inverzný prvok $-\vec{\alpha}$ vo $(V,+)$. Podobným spôsobom, akým sme dokázali jednoznačnosť neutrálneho prvku v tvrdení 3.1.12, by sme vedeli dokázať, že aj inverzný prvok v S musí byť ten istý, ako inverzný prvok vo V. Teda jediné, čo potrebujeme zistiť, je či aj vektor $-\vec{\alpha}$ patrí do S. To však vyplýva z toho, že $-\vec{\alpha} = (-1).\vec{\alpha}$, teda podľa podmienky (i) patrí do S.

{ppr:TVRKRIT}

Tvrdenie 4.2.7 (Kritérium vektorového podpriestoru)**.** *Nech V je vektorový priestor nad poľom F a $S \subseteq V$, $S \neq \emptyset$. Potom S je podpriestor V práve vtedy, keď pre ľubovoľné $c, d \in F$ a $\vec{\alpha}, \vec{\beta} \in V$ platí*

{ppr:EQKRIT}
$$\vec{\alpha}, \vec{\beta} \in S \qquad \Rightarrow \qquad c\vec{\alpha} + d\vec{\beta} \in S. \tag{4.1}$$

Dôkaz. Ako sme už niekoľkokrát spomenuli, ekvivalenciu 2 výrokov môžeme dokázať tak, že dokážeme implikácie oboma smermi.

To, že S je neprázdna, overovať nemusíme, pretože táto podmienka sa vyskytuje v oboch prípadoch.

$\boxed{\Rightarrow}$ Ak $\vec{\alpha}, \vec{\beta} \in S$, tak podľa (ii) platí $c\vec{\alpha} \in S$ a $d\vec{\beta} \in S$. Z toho na základe (i) dostaneme $c\vec{\alpha} + d\vec{\beta} \in S$.

$\boxed{\Leftarrow}$ Ak množina S spĺňa podmienku (4.1), tak dosadením $c = d = 1$ dostaneme, že S spĺňa (i). Ak zvolíme $d = 0$, dostaneme podmienku (ii). □

Ďalšia vlastnosť, ktorá bude pre nás užitočná, je fakt, že prienik dvoch podpriestorov vektorového priestoru je opäť podpriestor.

{ppr:VTPRIEN}

Veta 4.2.8. *Ak S a T sú podpriestory vektorového priestoru V, tak aj $S \cap T$ je podpriestor V.*

Dôkaz. Pretože $\vec{0} \in S$ aj $\vec{0} \in T$, platí $\vec{0} \in S \cap T$, čiže $S \cap T \neq \emptyset$.

Overíme podmienku (4.1). Ak $\vec{\alpha}, \vec{\beta} \in S \cap T$, tak platí $c\vec{\alpha} + d\vec{\beta} \in S$ (lebo S je podpriestor V) a súčasne $c\vec{\alpha} + d\vec{\beta} \in T$ (lebo T je podpriestor V). To znamená, že $c\vec{\alpha} + d\vec{\beta} \in S \cap T$. □

Tvrdenia takéhoto typu sa jednoduchým spôsobom dajú rozšíriť z dvoch objektov na ľubovoľný konečný počet. (Pokúste sa to overiť podrobne.)

{ppr:DOSPRIEN}

Dôsledok 4.2.9. *Nech $n \in \mathbb{N}$. Ak $S_1, S_2, \ldots, S_n$ sú podpriestory priestoru V, tak aj $\bigcap_{i=1}^{n} S_i$ je podpriestor priestoru V.*

Veľmi podobným spôsobom, ako sme dokázali vetu 4.2.8, sa dá overiť, že podobné tvrdenie platí aj pre nekonečne veľa podpriestorov.

{ppr:VTPRIENINF}

Veta 4.2.10. *Nech $I \neq \emptyset$ je ľubovoľná neprázdna množina a S_i je podpriestor priestoru V pre každé $i \in I$. Potom aj $\bigcap_{i \in I} S_i$ je podpriestor priestoru V.*

Dôkaz$^{\triangle}$. Označme $S = \bigcap_{i \in I} S_i$. Stačí si uvedomiť, že vektor $\vec{\gamma} \in S$ práve vtedy, keď pre všetky $i \in I$ platí $\vec{\gamma} \in S_i$. Na základe toho dostaneme

$$\vec{\alpha}, \vec{\beta} \in S \;\Rightarrow\; (\forall i \in I)\vec{\alpha}, \vec{\beta} \in S_i \;\Rightarrow\; (\forall i \in I)c\vec{\alpha} + d\vec{\beta} \in S_i \;\Rightarrow\; c\vec{\alpha} + d\vec{\beta} \in S.$$

Podľa tvrdenia 4.2.7 je teda S vektorový podpriestor priestoru V. □

Cvičenia

Úloha 4.2.1. Podrobne dokážte dôsledok 4.2.9.

Úloha 4.2.2. Dokážte, že množina všetkých funkcií $f\colon \mathbb{R} \to \mathbb{R}$, ktoré sú tvaru $a + b\cos x + c\sin x$ pre nejaké $a, b, c \in \mathbb{R}$ tvoria vektorový podpriestor priestoru všetkých reálnych funkcií $\mathbb{R}^{\mathbb{R}}$.
rcvic:PODPRY}

Úloha 4.2.3. Ktoré z týchto množín tvoria vektorový podpriestor priestoru $\mathbb{R}^3$?
a) $M = \{(x_1, x_2, x_3) \in \mathbb{R}^3; x_1 \in \mathbb{Z}\}$
b) $M = \{(x_1, x_2, x_3) \in \mathbb{R}^3; x_1 = 0\}$
c) $M = \{(x_1, x_2, x_3) \in \mathbb{R}^3; x_1 = 0 \vee x_2 = 0\}$
d) $M = \{(x_1, x_2, x_3) \in \mathbb{R}^3; 3x_1 + 4x_2 = 1\}$
e) $M = \{(x_1, x_2, x_3) \in \mathbb{R}^3; 7x_1 - x_2 = 0\}$
f) $M = \{(x_1, x_2, x_3) \in \mathbb{R}^3; x_1 + x_2 = x_3\}$
g) $M = \{(x_1, x_2, x_3) \in \mathbb{R}^3; |x_1| = |x_2|\}$
h) $M = \{(x_1, x_2, x_3) \in \mathbb{R}^3; x_1 + x_2 + x_3 \geq 0\}$
i) $M = \{(x_1, x_2, x_3) \in \mathbb{R}^3; 2x_1 = -x_2 = x_3\}$
j) $M = \{(x_1, x_2, x_3) \in \mathbb{R}^3; x_1 + x_2 + x_3 = 0\}$.
{pprcvic:PPRFCIE}

Úloha 4.2.4. Ktoré z týchto podmnožín tvoria vektorový podpriestor priestoru reálnych funkcií $\mathbb{R}^{\mathbb{R}}$?
a) funkcie $f\colon \mathbb{R} \to \mathbb{R}$ s vlastnosťou $2f(0) = f(1)$
b) nezáporné funkcie
c) funkcie $f\colon \mathbb{R} \to \mathbb{R}$ s vlastnosťou $f(1) = 1 + f(0)$
d) funkcie $f\colon \mathbb{R} \to \mathbb{R}$ s vlastnosťou $(\forall x \in \langle 0, 1\rangle) f(x) = f(1 - x)$
e) ohraničené funkcie $f\colon \mathbb{R} \to \mathbb{R}$
f) spojité funkcie $f\colon \mathbb{R} \to \mathbb{R}$
h) funkcie $f\colon \mathbb{R} \to \mathbb{R}$ také, že existuje konečná $\lim\limits_{x\to\infty} f(x)$
i*) funkcie $f\colon \mathbb{R} \to \mathbb{R}$ také, že existuje konečná alebo nekonečná $\lim\limits_{x\to\infty} f(x)$.
{pprcvic:POLYN}

Úloha 4.2.5. Overte, či
a) množina všetkých polynómov s reálnymi koeficientami,
b) množina všetkých polynómov s reálnymi koeficientami stupňa najviac n,
c) množina všetkých polynómov párneho stupňa,
d) množina všetkých polynómov stupňa práve n
sú vektorové priestory. Sčitovanie a násobenie skalárom definujeme rovnako ako pre reálne funkcie.

Úloha 4.2.6. Nech S, T sú podpriestory vektorového priestoru V nad poľom F. Ukážte, že $S \cup T$ je podpriestor priestoru V práve vtedy, keď $S \subseteq T$ alebo $T \subseteq S$.
{pprcvic:ULOEKVIVPPR}

Úloha 4.2.7. Nech V je vektorový priestor nad poľom F a $S \neq \emptyset$ je podmnožina V. Ukážte, že S je podpriestor V práve vtedy, keď pre ľubovoľné $c \in F$ a $\vec{\alpha}, \vec{\beta} \in S$ platí $c\vec{\alpha} + \vec{\beta} \in S$.

4.3 Lineárna kombinácia, lineárna nezávislosť

4.3.1 Lineárna kombinácia a lineárny obal

{lnze:DEFLK}

Definícia 4.3.1. Nech V je vektorový priestor nad poľom F. Hovoríme, že vektor $\vec{\alpha}$ je *lineárnou kombináciou* vektorov $\vec{\alpha}_1, \vec{\alpha}_2, \ldots, \vec{\alpha}_n$, ak existujú skaláry $c_1, c_2, \ldots, c_n \in F$ také, že

$$\vec{\alpha} = c_1\vec{\alpha}_1 + c_2\vec{\alpha}_2 + \cdots + c_n\vec{\alpha}_n.$$

Skaláry $c_1, c_2, \ldots, c_n$ nazývame *koeficienty lineárnej kombinácie.*

Príklad 4.3.2. $(1,0) + (0,1) = (1,1)$, teda vektor $(1,1)$ je lineárna kombinácia vektorov $(1,0)$ a $(0,1)$ v $\mathbb{R}^2$.

$2 \cdot (1,0,0) + 3 \cdot (0,1,0) = (2,3,0)$, teda vektor $(2,3,0)$ je lineárna kombinácia vektorov $(1,0,0)$ a $(0,1,0)$ v $\mathbb{R}^3$.

Tvrdenie 4.3.3. *Nech V je vektorový priestor nad poľom F. Ak $\vec{\alpha}_1, \vec{\alpha}_2, \ldots, \vec{\alpha}_n \in V$, tak množina*

$$M = \{c_1\vec{\alpha}_1 + c_2\vec{\alpha}_2 + \cdots + c_n\vec{\alpha}_n; c_i \in F \text{ pre } i = 1, 2 \ldots, n\}$$

je podpriestor vektorového priestoru V.

Tento podpriestor nazývame lineárny obal *vektorov $\vec{\alpha}_1, \vec{\alpha}_2, \ldots, \vec{\alpha}_n$ alebo* podpriestor generovaný vektormi $\vec{\alpha}_1, \vec{\alpha}_2, \ldots, \vec{\alpha}_n$. *Označujeme ho*

$$M =: [\vec{\alpha}_1, \vec{\alpha}_2, \ldots, \vec{\alpha}_n].$$

Ak platí $[\vec{\alpha}_1, \vec{\alpha}_2, \ldots, \vec{\alpha}_n] = V$, hovoríme, že vektory $\vec{\alpha}_1, \vec{\alpha}_2, \ldots, \vec{\alpha}_n$ generujú *vektorový priestor V.*

Definícia množiny $[\vec{\alpha}_1, \vec{\alpha}_2, \ldots, \vec{\alpha}_n]$ vlastne hovorí, že $[\vec{\alpha}_1, \vec{\alpha}_2, \ldots, \vec{\alpha}_n]$ je množina všetkých lineárnych kombinácií vektorov $\vec{\alpha}_1, \vec{\alpha}_2, \ldots, \vec{\alpha}_n$.

Dôkaz. Aby sme dokázali, že M je podpriestor vektorového priestoru V, stačí nám overiť, že táto množina je uzavretá na sčitovanie a skalárne násobky.

Ak máme dva vektory

$$\vec{\alpha} = c_1\vec{\alpha}_1 + c_2\vec{\alpha}_2 + \cdots + c_n\vec{\alpha}_n$$
$$\vec{\beta} = d_1\vec{\alpha}_1 + d_2\vec{\alpha}_2 + \cdots + d_n\vec{\alpha}_n$$

tak aj vektor $\vec{\alpha} + \vec{\beta} = c_1\vec{\alpha}_1 + c_2\vec{\alpha}_2 + \cdots + c_n\vec{\alpha}_n + d_1\vec{\alpha}_1 + d_2\vec{\alpha}_2 + \cdots + d_n\vec{\alpha}_n = (c_1 + d_1)\vec{\alpha}_1 + (c_2 + d_2)\vec{\alpha}_1 + \cdots + (c_n + d_n)\vec{\alpha}_n$ má tvar, aký požadujeme v definícii množiny M. Takisto pre $c \in F$ dostaneme

$$c\vec{\alpha} = c(c_1\vec{\alpha}_1 + c_2\vec{\alpha}_2 + \cdots + c_n\vec{\alpha}_n) = cc_1\vec{\alpha}_1 + cc_2\vec{\alpha}_2 + \cdots + cc_n\vec{\alpha}_n,$$

čiže aj vektor $c\vec{\alpha}$ patrí do M. □

{lnze:PRSTANDBAZA}

Príklad 4.3.4. Pre vektorový priestor $\mathbb{R}^3$ platí $\mathbb{R}^3 = [(1,0,0), (0,1,0), (0,0,1)]$.

Skutočne, ľubovoľný vektor $(x,y,z) \in \mathbb{R}^3$ sa dá vyjadriť ako lineárna kombinácia $x(1,0,0) + y(0,1,0) + z(0,0,1)$.

Podobne sa dá dokázať, že $[(1,0,\ldots,0), (0,1,0,\ldots,0), \ldots, (0,\ldots,0,1)] = \mathbb{R}^n$.

Na vygenerovanie podpriestoru $S = \{(x,y,z) \in \mathbb{R}^3; x+y+z = 0\}$ dokonca stačia 2 vektory. Všimnime si, že rovnica $x+y+z = 0$ je ekvivalentná s rovnicou $z = -x-y$, teda podpriestor S môžeme zapísať aj v tvare $S = \{(x,y,-x-y); x,y \in \mathbb{R}\}$. Teraz už vidíme, že každý vektor z S sa dá zapísať ako lineárna kombinácia $(x,y,-x-y) = x(1,0,-1) + y(0,1,-1)$ a platí $S = [(1,0,-1), (0,1,-1)]$.

Všimnime si, že každý podpriestor, ktorý obsahuje vektory $\vec{\alpha}_1, \vec{\alpha}_2, \ldots, \vec{\alpha}_n$ musí obsahovať aj všetky ich lineárne kombinácie (pretože ich vieme dostať opakovaním sčitovania a násobenia skalárom a na tieto 2 operácie sú podpriestory uzavreté). Teda podpriestory sú uzavreté vzhľadom na lineárne kombinácie vektorov. Toto tvrdenie sformalizujeme a dokážeme v nasledujúcej leme.

Lema 4.3.5. *Ak $\vec{\alpha}_1, \vec{\alpha}_2, \ldots, \vec{\alpha}_n \in S$, kde S je podpriestor vektorového priestoru V nad poľom F, aj ich ľubovoľná lineárna kombinácia $c_1\vec{\alpha}_1 + c_2\vec{\alpha}_2 + \cdots + c_n\vec{\alpha}_n$ patrí do podpriestoru S.*

Dôkaz. Chceme ukázať, že ľubovoľné lineárna kombinácia $c_1\vec{\alpha}_1 + c_2\vec{\alpha}_2 + \cdots + c_k\vec{\alpha}_k$ patrí do S. Budeme postupovať indukciou vzhľadom na k.

1° Pre $k = 1$ prakticky niet čo dokazovať. (Dokazovaný výrok pre $k = 1$ je $\vec{\alpha} \in S \Rightarrow c\vec{\alpha} \in S$.)

Pre $k = 2$ dostaneme tvrdenie $\vec{\alpha}_1, \vec{\alpha}_2 \in S \Rightarrow c_1\vec{\alpha}_1 + c_2\vec{\alpha}_2 \in S$, ktoré vyplýva z kritéria vektorového podpriestoru (tvrdenie 4.2.7).

2° Predpokladajme, že tvrdenie platí pre lineárnu kombináciu k vektorov. Dokážeme, že platí aj pre $(k+1)$ vektorov.

$$c_1\vec{\alpha}_1 + c_2\vec{\alpha}_2 + \cdots + c_{k+1}\vec{\alpha}_{k+1} = (\underbrace{c_1\vec{\alpha}_1 + c_2\vec{\alpha}_2 + \cdots + c_k\vec{\alpha}_k}_{\in S}) + c_{k+1}\vec{\alpha}_{k+1} \in S$$

Podľa indukčného predpokladu lineárna kombinácia k vektorov (prvá zátvorka) patrí do S a po pripočítaní vektora $c_{k+1}\vec{\alpha}_{k+1}$ (ktorý tiež patrí do S) dostaneme opäť vektor z S. □

Táto lema vlastne hovorí, že podpriestor je uzavretý vzhľadom na tvorbu lineárnych kombinácií. Túto podmienku by sme mohli pridať ako ďalšiu ekvivalentnú podmienku hovoriacu, kedy je neprázdna podmnožina vektorového priestoru podpriestorom. (Zatiaľ máme dve takéto podmienky – definíciu podpriestoru a kritérium vektorového podpriestoru.)

Veta 4.3.6. *Ak $\vec{\alpha}_1, \vec{\alpha}_2, \ldots, \vec{\alpha}_n \in S$, kde S je podpriestor vektorového priestoru V nad poľom F, tak $[\vec{\alpha}_1, \vec{\alpha}_2, \ldots, \vec{\alpha}_n] \subseteq S$.*

Dôkaz. Vyplýva z predchádzajúcej lemy. □

Poznámka 4.3.7. Predchádzajúca veta hovorí, že podpriestor $[\vec{\alpha}_1, \vec{\alpha}_2, \ldots, \vec{\alpha}_n]$ je najmenší podpriestor priestoru V, ktorý obsahuje vektory $\vec{\alpha}_1, \vec{\alpha}_2, \ldots, \vec{\alpha}_n$.

Pod slovom *najmenší* tu rozumieme, že ak S je taký podpriestor V, že $\vec{\alpha}_1, \vec{\alpha}_2, \ldots, \vec{\alpha}_n \in S$, tak $[\vec{\alpha}_1, \vec{\alpha}_2, \ldots, \vec{\alpha}_n] \subseteq S$. (Často sa používa aj termín *najmenší vzhľadom na inklúziu.*)

Tento podpriestor je prienikom všetkých podpriestorov V, ktoré obsahujú vektory $\vec{\alpha}_1, \vec{\alpha}_2, \ldots, \vec{\alpha}_n$. Pretože prienik podpriestorov je opäť podpriestor (veta 4.2.10) dostaneme takto podpriestor priestoru V. Pretože sme urobili prienik všetkých podpriestorov, je takto získaný prienik najmenší podpriestor vzhľadom na inklúziu, ktorý obsahuje $\vec{\alpha}_1, \vec{\alpha}_2, \ldots, \vec{\alpha}_n$.

{lnze:VTOBAL}

Veta 4.3.8. *Nech $\vec{\alpha}_1, \vec{\alpha}_2, \ldots, \vec{\alpha}_n \in V$, $\vec{\beta} \in V$, kde V je vektorový priestor nad poľom F. Potom $\vec{\beta}$ je lineárnou kombináciou vektorov $\vec{\alpha}_1, \vec{\alpha}_2, \ldots, \vec{\alpha}_n$ práve vtedy, keď*

$$[\vec{\alpha}_1, \vec{\alpha}_2, \ldots, \vec{\alpha}_n] = [\vec{\alpha}_1, \vec{\alpha}_2, \ldots, \vec{\alpha}_n, \vec{\beta}].$$

Dôkaz. $\boxed{\Rightarrow}$ Chceme ukázať rovnosť 2 množín – to môžeme dokazovať tak, že dokážeme obe inklúzie. Pritom inklúzia $[\vec{\alpha}_1, \vec{\alpha}_2, \ldots, \vec{\alpha}_n] \subseteq [\vec{\alpha}_1, \vec{\alpha}_2, \ldots, \vec{\alpha}_n, \vec{\beta}]$ je zrejmá. Opačná inklúzia vyplýva z toho, že ak máme nejaký vektor $\vec{\gamma} \in [\vec{\alpha}_1, \vec{\alpha}_2, \ldots, \vec{\alpha}_n, \vec{\beta}]$, čo znamená, že

$$\vec{\gamma} = c_1\vec{\alpha}_1 + c_2\vec{\alpha}_2 + \cdots + c_n\vec{\alpha}_n + c\vec{\beta}$$

pre nejaké $c_1, c_2, \ldots, c_n, c \in F$ a ak vieme, že β je lineárna kombinácia vektorov $\vec{\alpha}_1, \vec{\alpha}_2, \ldots, \vec{\alpha}_n$, čiže

$$\vec{\beta} = d_1\vec{\alpha}_1 + d_2\vec{\alpha}_2 + \cdots + d_n\vec{\alpha}_n$$

pre nejaké $d_1, d_2, \ldots, d_n$, tak úpravou dostaneme

$$\vec{\gamma} = c_1\vec{\alpha}_1+c_2\vec{\alpha}_2+\cdots+c_n\vec{\alpha}_n+cd_1\vec{\alpha}_1+cd_2\vec{\alpha}_2+\cdots+cd_n\vec{\alpha}_n = (c_1+cd_1)\vec{\alpha}_1+(c_2+cd_2)\vec{\alpha}_1+\ldots+(c_n+cd_n)\vec{\alpha}_n,$$

čo znamená, že

$$\vec{\gamma} \in [\vec{\alpha}_1, \vec{\alpha}_2, \ldots, \vec{\alpha}_n].$$

(Stručne: Lineárna kombinácia lineárnych kombinácií je opäť lineárna kombinácia.)

$\boxed{\Leftarrow}$ Podľa predpokladu $\vec{\beta} \in [\vec{\alpha}_1, \vec{\alpha}_2, \ldots, \vec{\alpha}_n]$, teda $\vec{\beta}$ je lineárna kombinácia vektorov $\vec{\alpha}_1, \vec{\alpha}_2, \ldots, \vec{\alpha}_n$. □

4.3.2 Lineárna nezávislosť

V tejto podkapitole zadefinujeme pojem, ktorý bude pre nás v ďalšom štúdiu veľmi dôležitý.

Definícia 4.3.9. Nech V je vektorový priestor nad poľom F. Vektory $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ sú *lineárne závislé,* ak existujú $c_1, \ldots, c_n \in F$, ktoré nie sú všetky nulové a platí

$$c_1\vec{\alpha}_1 + \cdots + c_n\vec{\alpha}_n = \vec{0}.$$

(Stručne: $\vec{0}$ je nenulovou lineárnou kombináciou vektorov $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$.)

V opačnom prípade hovoríme, že vektory $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ *lineárne nezávislé.*

Príklad 4.3.10. Najprv sa pozrime na niektoré špeciálne prípady. Ak $n = 1$, teda ak máme len jediný vektor $\vec{\alpha}$, tento vektor tvorí lineárne závislú množinu práve vtedy, keď $\vec{\alpha} = \vec{0}$ (vyplýva to z vety 4.1.6 (c)).

Ak $n = 2$, čiže máme 2 vektory $\vec{\alpha}$ a $\vec{\beta}$, tak sú lineárne závislé práve vtedy, keď jeden z nich je násobkom druhého, čiže $\vec{\alpha} = c\vec{\beta}$ alebo $\vec{\beta} = c\vec{\alpha}$ pre nejaké $c \in F$ (úloha 4.3.8).

Vektory $(0,1), (1,0), (1,1) \in \mathbb{R}^2$ sú lineárne závislé, lebo $1\cdot(0,1)+1\cdot(0,1)-1\cdot(1,1) = (0,0)$. (Nulový vektor v $\mathbb{R}^2$ je $(0,0)$.)

Ekvivalentne môžeme lineárnu nezávislosť definovať tak, že vektory $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ sú lineárne nezávislé práve vtedy, keď platí implikácia

{lnze:EQLNZE}
$$c_1\vec{\alpha}_1 + c_2\vec{\alpha}_2 + \cdots + c_n\vec{\alpha}_n = \vec{0} \qquad \Rightarrow \qquad c_1 = c_2 = \ldots = c_n = 0. \tag{4.2}$$

Táto formulácia lineárnej nezávislosti bude pre nás často výhodnejšia pri overovaní, či nejaké vektory sú lineárne nezávislé.

Príklad 4.3.11. Vektory $(1,0), (0,1)$ vo vektorovom priestore $\mathbb{R}^2$ sú lineárne nezávislé. Skutočne, z rovnosti $c_1(1,0) + c_2(0,1) = (c_1, c_2) = (0,0)$ vyplýva $c_1 = c_2 = 0$.

Poznámka 4.3.12. Aby sme si ozrejmili, že uvedené dve definície lineárnej nezávislosti sú skutočne ekvivalentné, potrebujeme si najprv pripomenúť, ako sa negujú výroky s kvantifikátormi.[3]

Pre negácie výrokov s kvantifikátormi platia dve jednoduché pravidlá (pozri 2.1.5):

$$\neg[(\forall x)P(x)] \quad \Leftrightarrow \quad (\exists x)(\neg P(x)),$$
$$\neg[(\exists x)P(x)] \quad \Leftrightarrow \quad (\forall x)(\neg P(x)).$$

[3]Ako sme si už kedysi povedali, kvantifikátory sú len spôsobom na zápis istého druhu výrokov. Tu síce odvodíme ekvivalenciu týchto 2 definícií pomocou formálnych pravidiel pre prácu s kvantifikátormi, je to však presne to isté, čo dostaneme aj logickou úvahou, kvantifikátory nám poslúžia len na stručnejší, jednoduchší a prehľadnejší zápis týchto úvah.

(Teda existenčný kvantifikátor sa mení na všeobecný a obrátene a výrok pod kvantifikátorom sa zneguje.)

My by sme radi overili, či (4.2) je skutočne negáciou definície lineárne závislých vektorov. Pokúsme sa teda najprv prepísať definíciu lineárne závislých vektorov. (Snáď jediným drobným problémom je, ako zapísať, že aspoň jeden zo skalárov $c_1, \ldots, c_n \in F$ je nenulový.)

Spomínanú definíciu by sme mohli zapísať takto

$$(\exists c_1, \ldots, c_n \in F)[c_1\vec{\alpha}_1 + \cdots + c_n\vec{\alpha}_n = 0 \wedge (c_1 \neq 0 \vee c_2 \neq 0 \vee \ldots \vee c_n \neq 0)].$$

Teraz už použitím pravidiel o negovaní výrokov s kvantifikátormi a de Morganových zákonov dostaneme

$$(\forall c_1, \ldots, c_n \in F)[c_1\vec{\alpha}_1 + \cdots + c_n\vec{\alpha}_n \neq 0 \vee (c_1 = 0 \wedge c_2 = 0 \wedge \ldots \wedge c_n = 0)].$$

Keď si uvedomíme, že $\neg P \vee Q$ je vlastne iný zápis implikácie $P \Rightarrow Q$ (inak povedané, $(\neg P \vee Q) \Leftrightarrow (P \Rightarrow Q)$ je tautológia, pozri úlohu 2.1.1) vidíme, že sme dostali

$$(\forall n \in \mathbb{N})(\forall c_1, \ldots, c_n \in F)(c_1\vec{\alpha}_1 + \cdots + c_n\vec{\alpha}_n = 0 \Rightarrow c_1 = c_2 = \ldots = c_n = 0),$$

čiže implikáciu (4.2).

Nasledujúce výsledky budú pre nás veľmi užitočné v nasledujúcej podkapitole.

{lnze:VTLZOST}

Veta 4.3.13. *Nech V je vektorový priestor nad poľom F. Nech n je prirodzené číslo, $n \geq 2$ a $\vec{\alpha}_1, \ldots, \vec{\alpha}_n \in V$. Vektory $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ sú lineárne závislé práve vtedy, keď niektorý z nich je lineárnou kombináciou ostatných.*

Dôkaz. $\boxed{\Rightarrow}$ Ak sú vektory $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ lineárne závislé, znamená to, že platí rovnosť

$$c_1\vec{\alpha}_1 + \cdots + c_n\vec{\alpha}_n = \vec{0}$$

pre nejaké $c_1, \ldots, c_n \in F$, ktoré nie sú všetky nulové. Zvoľme si niektorý nenulový index $c_i \neq 0$. Pretože $c_i \neq 0$, existuje inverzný prvok c_i^{-1}. Úpravou predchádzajúcej rovnosti dostaneme

$$\begin{aligned} -c_i\vec{\alpha}_i &= c_1\vec{\alpha}_1 + \cdots + c_{i-1}\vec{\alpha}_{i-1} + c_{i+1}\vec{\alpha}_{i+1} + \cdots + c_n\vec{\alpha}_n \\ -\vec{\alpha}_i &= c_i^{-1}c_1\vec{\alpha}_1 + \cdots + c_i^{-1}c_{i-1}\vec{\alpha}_{i-1} + c_i^{-1}c_{i+1}\vec{\alpha}_{i+1} + \cdots + c_i^{-1}c_n\vec{\alpha}_n \\ \vec{\alpha}_i &= -c_i^{-1}c_1\vec{\alpha}_1 - \ldots - c_i^{-1}c_{i-1}\vec{\alpha}_{i-1} - c_i^{-1}c_{i+1}\vec{\alpha}_{i+1} - \ldots - c_i^{-1}c_n\vec{\alpha}_n \end{aligned}$$

Teda $\vec{\alpha}_i$ je lineárna kombinácia ostatných vektorov.

$\boxed{\Leftarrow}$ Bez ujmy na všeobecnosti,[4] nech vektor, ktorý je lineárnou kombináciou ostatných, je vektor $\vec{\alpha}_1$. To znamená, že

$$\vec{\alpha}_1 = c_2\vec{\alpha}_2 + \cdots + c_n\vec{\alpha}_n,$$

čiže

$$-1 \cdot \vec{\alpha}_1 + c_2\vec{\alpha}_2 + \cdots + c_n\vec{\alpha}_n = \vec{0}.$$

Zistili sme, že $\vec{0}$ sa dá získať ako lineárna kombinácia vektorov $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$, pričom hneď prvý koeficient -1 je nenulový. □

[4] Frázu „bez ujmy na všeobecnosti“ nájdete v matematických textoch dosť často. Myslí sa tým, že použijeme dodatočný argument, ktorý môže o niečo zjednodušiť zápis dôkazu alebo dôkaz, ale je zrejmé, že analogický dôkaz by platil aj bez tohoto predpokladu. Napríklad v tomto prípade nám výber vektora $\vec{\alpha}_1$ umožní jednoduchší zápis a navyše všeobecnú situáciu vieme previesť na tento prípad vhodným prečíslovaním vektorov. Iná možnosť by bola postupovať podobným postupom ako v predchádzajúcej časti dôkazu, znamenalo by to však o niečo komplikovanejší zápis.

{lnze:VTLZPRE

Veta 4.3.14. *Nech V je vektorový priestor nad poľom F. Nech $\vec{\alpha}_1, \ldots, \vec{\alpha}_n \in V$ sú vektory také, že $\vec{\alpha}_1 \neq \vec{0}$. Vektory $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ sú lineárne závislé práve vtedy, keď niektorý z nich je lineárnou kombináciou predchádzajúcich.*

Dôkaz. Implikácia $\boxed{\Leftarrow}$ vyplýva z predchádzajúcej vety.

Implikáciu $\boxed{\Rightarrow}$ dokážeme veľmi podobným spôsobom ako v predchádzajúcom dôkaze.

Ak vektory $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ sú lineárne závislé, znamená to podľa (4.2), že

$$c_1\vec{\alpha}_1 + \cdots + c_n\vec{\alpha}_n = \vec{0}$$

pre nejaké $c_1, \ldots, c_n \in F$, pričom aspoň jedno c_i, $i \in \{1, 2, \ldots, n\}$ je nenulové.

Nech $k \in \{1, 2, \ldots, n\}$ je posledný index z tejto množiny, pre ktorý je c_k nenulové. (Taký index existuje, pretože množina $\{1, 2, \ldots, n\}$ je konečná. Navyše, platí $k \geq 2$, pretože $\vec{\alpha}_1 \neq \vec{0}$.)

Potom predchádzajúcu rovnicu môžeme prepísať do tvaru

$$c_1\vec{\alpha}_1 + c_2\vec{\alpha}_2 + \cdots + c_k\vec{\alpha}_k = \vec{0}$$

a úpravou dostaneme

$$c_k\vec{\alpha}_k = -c_1\vec{\alpha}_1 - c_2\vec{\alpha}_2 - \cdots - c_{k-1}\vec{\alpha}_{k-1}.$$

Pretože $c_k \neq 0$, existuje inverzný prvok c_k^{-1}. Keď predchádzajúcu rovnosť prenásobíme c_k^{-1} dostaneme

$$\vec{\alpha}_k = -c_k^{-1}c_1\vec{\alpha}_1 - c_k^{-1}c_2\vec{\alpha}_2 - \cdots - c_k^{-1}c_{k-1}\vec{\alpha}_{k-1},$$

teda $\vec{\alpha}_k$ je skutočne lineárnou kombináciou predchádzajúcich vektorov. □

Nasledujúca veta bude kľúčová pri definovaní dimenzie vektorového priestoru v nasledujúcej kapitole.

{lnze:VTSTEINITZ}

Veta 4.3.15 (Steinitzova veta o výmene)**.** *Nech V je vektorový priestor nad poľom F. Ak $V = [\vec{\alpha}_1, \ldots, \vec{\alpha}_n]$ (vektorový priestor V je generovaný vektormi $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$) a $\vec{\beta}_1, \ldots, \vec{\beta}_s \in V$ sú lineárne nezávislé vektory, tak*

(i) *$s \leq n$,*

(ii) *z vektorov $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ sa dá vybrať $n - s$ vektorov, ktoré spolu s vektormi $\vec{\beta}_1, \ldots, \vec{\beta}_s$ generujú V.*

V prípade, že vám nie je úplne jasné, čo hovorí táto veta, môžete sa pozrieť na príklad 4.3.16, prípadne si vyskúšať urobiť ďalšie podobné príklady sami.

Dôkaz. Matematickou indukciou vzhľadom na s.

1° Najprv uvažujme prípad, že $s = 1$. Vektor $\vec{\beta}_1$ je lineárne nezávislý, teda nenulový. Pretože $\vec{\beta}_1 \in V$, je vektor $\vec{\beta}_1$ lineárna kombinácia vektorov $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$. To znamená, že vektory $\vec{\beta}_1, \vec{\alpha}_1, \ldots, \vec{\alpha}_n$ sú lineárne závislé. Preto niektorý z nich je lineárnou kombináciou predchádzajúcich. Pritom to nemôže byť vektor $\vec{\beta}_1$, lebo $\vec{\beta}_1 \neq \vec{0}$.

Ak $\vec{\alpha}_i$ je lineárna kombinácia predchádzajúcich vektorov, tak podľa vety 4.3.8 $V = [\vec{\alpha}_1, \ldots, \vec{\alpha}_n] = [\vec{\beta}_1, \vec{\alpha}_1, \ldots, \vec{\alpha}_n] = [\vec{\beta}_1, \vec{\alpha}_1, \ldots, \vec{\alpha}_{i-1}, \vec{\alpha}_{i+1}, \ldots, \vec{\alpha}_n]$.

Pretože $V = [\vec{\alpha}_1, \ldots, \vec{\alpha}_n]$ obsahuje aspoň jeden nenulový vektor $\vec{\beta}_1$, platí $V \neq \{\vec{0}\}$ a $n \geq 1$.

2° Predpokladajme, že tvrdenie platí pre číslo s. Budeme sa snažiť dokázať, že platí aj pre $s + 1$.

Máme teda daných $s+1$ lineárne nezávislých vektorov $\vec{\beta}_1, \ldots, \vec{\beta}_{s+1} \in V$. Podľa indukčného predpokladu vieme vektory $\vec{\beta}_1, \ldots, \vec{\beta}_s$ doplniť $n-s$ vektormi spomedzi vektorov $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ tak, aby generovali celý priestor. Ďalej platí $s \leq n$.

Predpokladajme, že by platilo $s = n$. To by znamenalo, že (podľa indukčného predpokladu) sa dajú vektory $\vec{\beta}_1, \ldots, \vec{\beta}_s$ doplniť $n - s = 0$ vektormi, čiže $V = [\vec{\beta}_1, \ldots, \vec{\beta}_s]$. Pretože $\vec{\beta}_{s+1} \in [\vec{\beta}_1, \ldots, \vec{\beta}_s]$, vektor $\vec{\beta}_{s+1}$ je lineárna kombinácia vektorov $\vec{\beta}_1, \ldots, \vec{\beta}_s$, čo je spor s tým, že vektory $\vec{\beta}_1, \ldots, \vec{\beta}_{s+1}$ sú lineárne nezávislé. Musí teda platiť $s < n$, čiže

$$s + 1 \leq n.$$

Bez ujmy na všeobecnosti môžeme predpokladať, že vektory, ktorými môžeme doplniť $\vec{\beta}_1, \ldots, \vec{\beta}_s$ sú vektory $\vec{\alpha}_1, \ldots, \vec{\alpha}_{n-s}$. (Takúto situáciu vieme dosiahnuť vhodným prečíslovaním vektorov $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$.) Platí teda $\vec{\beta}_{s+1} \in V = [\vec{\beta}_1, \ldots, \vec{\beta}_s, \vec{\alpha}_1, \ldots, \vec{\alpha}_{n-s}]$.

Z toho vyplýva, že vektory $\vec{\beta}_1, \ldots, \vec{\beta}_{s+1}, \vec{\alpha}_1, \ldots, \vec{\alpha}_{n-s}$ sú lineárne závislé, čiže niektorý z nich je lineárnou kombináciou predchádzajúcich vektorov. Nemôže to však byť žiadny z vektorov $\vec{\beta}_1, \ldots, \vec{\beta}_{s+1}$, lebo tieto vektory sú lineárne nezávislé. Musí to byť niektorý z $\vec{\alpha}_1, \ldots, \vec{\alpha}_{n-s}$, bez ujmy na všeobecnosti nech je to $\vec{\alpha}_{n-s}$. Z vety 4.3.8 potom dostaneme

$$V = [\vec{\beta}_1, \ldots, \vec{\beta}_{s+1}, \vec{\alpha}_1, \ldots, \vec{\alpha}_{n-s}] = [\vec{\beta}_1, \ldots, \vec{\beta}_{s+1}, \vec{\alpha}_1, \ldots, \vec{\alpha}_{n-(s+1)}].$$

□

Z dôkazu môžeme vidieť, prečo sa predchádzajúca veta nazýva veta o výmene. V indukčnom kroku sme vymenili jeden z vektorov $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ za vektor $\vec{\beta}_{s+1}$.

Všimnime si, že na konci predchádzajúceho dôkazu môže nastať aj situácia $n = s + 1$, vtedy zápis $\vec{\alpha}_1, \ldots, \vec{\alpha}_{n-(s+1)}$ predstavuje 0 vektorov. Možno trochu neobvyklý zápis – ale dá sa ľahko si uvedomiť, že prípad $n = s + 1$ by fungoval v podstate rovnako. (Nie sú tam problémy s tým, že by sme uvažovali lineárny obal prázdnej množiny vektorov – sú tam totiž aj vektory $\vec{\beta}_1, \ldots, \vec{\beta}_{s+1}$.)

Podobne podmienka $n \geq 2$ vo vete 4.3.13 a podmienka $\vec{\alpha}_1 \neq \vec{0}$ vo vete 4.3.14 slúžia práve nato, aby sme sa vyhli prípadu, že v dôkaze (alebo už priamo v tvrdení vety) sa vyskytne lineárny obal prázdnej množiny vektorov (ten sme totiž nedefinovali). Môžete si rozmyslieť, že keby sme definitoricky položili $[\emptyset] = \{\vec{0}\}$, teda lineárny obal prázdnej množiny by bol nulový vektorový priestor, prešli by dôkazy týchto viet aj po vynechaní spomínaných podmienok.

Cieľom nasledujúceho príkladu, ktorým uzavrieme túto podkapitolu, je ilustrovať (na konkrétnych príkladoch), čo hovorí Steinitzova veta o výmene.

{lnze:PRSTEINITZ}

Príklad 4.3.16. Budeme pracovať vo vektorovom priestore $V = \mathbb{R}^3$ nad poľom $\mathbb{R}$.

a) Zvoľme $\vec{\alpha}_1 = (1, 0, 0)$, $\vec{\alpha}_2 = (0, 1, 0)$ a $\vec{\alpha}_3 = (0, 0, 1)$. Platí $V = [\vec{\alpha}_1, \vec{\alpha}_2, \vec{\alpha}_3]$ (pozri príklad 4.3.4). Ďalej nech $\vec{\beta}_1 = (1, 1, 0)$ a $\vec{\beta}_2 = (1, 0, 1)$. Tieto dva vektory sú lineárne nezávislé, lebo ani jeden z nich nie je násobok toho druhého.

Podľa vety 4.3.15 môžeme k vektorom $\vec{\beta}_1$, $\vec{\beta}_2$ pridať niektorý z vektorov $\vec{\alpha}_1$, $\vec{\alpha}_2$, $\vec{\alpha}_3$ tak, aby sme dostali trojicu vektorov, ktorá generuje V. Vyskúšajme pridať napríklad $\vec{\alpha}_1$. Chceme overiť, či platí $V = [\vec{\beta}_1, \vec{\beta}_2, \vec{\alpha}_1]$. Na to nám stačí ukázať, že $\vec{\alpha}_1, \vec{\alpha}_2, \vec{\alpha}_3 \in [\vec{\beta}_1, \vec{\beta}_2, \vec{\alpha}_1]$.

Skutočne máme

$$\begin{aligned}
\vec{\alpha}_1 &= \vec{\alpha}_1, \\
\vec{\alpha}_2 &= (1, 1, 0) - (1, 0, 0) = \vec{\beta}_1 - \vec{\alpha}_1, \\
\vec{\alpha}_3 &= (1, 0, 1) - (1, 0, 0) = \vec{\beta}_2 - \vec{\alpha}_1,
\end{aligned}$$

čiže všetky tieto vektory vieme získať ako lineárne kombinácie $\vec{\beta}_1$, $\vec{\beta}_2$ a $\vec{\alpha}_1$. To znamená, že $V = [\vec{\alpha}_1, \vec{\alpha}_2, \vec{\alpha}_3] \subseteq [\vec{\beta}_1, \vec{\beta}_2, \vec{\alpha}_1]$.

Môžete sa presvedčiť o tom, že v tomto prípade by sme dokonca dostali celý priestor aj vtedy, ak by sme namiesto $\vec{\alpha}_1$ použili vektor $\vec{\alpha}_2$ či vektor $\vec{\alpha}_3$. Takéto niečo však Steinitzova veta netvrdí – tá zaručuje len to, že aspoň jeden z týchto troch vektorov musí fungovať. Skúsme si teda ukázať ešte nejaký príklad, kde nie je úplne jedno, ktorý vektor si vyberieme na doplnenie.

b) Nech opäť me $\vec{\alpha}_1 = (1,0,0)$, $\vec{\alpha}_2 = (0,1,0)$ a $\vec{\alpha}_3 = (0,0,1)$, ale tentokrát $\vec{\beta}_1 = (1,1,0)$, $\vec{\beta}_2 = (0,1,0)$. Vektory $\vec{\beta}_1$ a $\vec{\beta}_2$ sú lineárne nezávislé.

Ak k vektorom $\vec{\beta}_{1,2}$ pridáme jeden z vektorov $\vec{\alpha}_{1,2}$, tak nedostaneme celý priestor V. Z rovností

$$\vec{\alpha}_1 = \vec{\beta}_1 - \vec{\beta}_2$$
$$\vec{\alpha}_2 = \vec{\beta}_2$$

totiž vidíme, že vektory $\vec{\alpha}_{1,2}$ sú lineárne kombinácie vektorov $\vec{\beta}_1$ a $\vec{\beta}_2$. Teda podľa vety 4.3.8 $[\vec{\beta}_1, \vec{\beta}_2] = [\vec{\beta}_1, \vec{\beta}_2, \vec{\alpha}_1] = [\vec{\beta}_1, \vec{\beta}_2, \vec{\alpha}_2]$ a ľahko sa dá presvedčiť o tom, že tento podpriestor neobsahuje vektor $\vec{\alpha}_3$. (Všetky vektory z tohoto podpriestoru majú tretiu súradnicu nulovú.)

Ak však pridáme vektor $\vec{\alpha}_3$, tak už dostaneme celý priestor $V = [\vec{\beta}_1, \vec{\beta}_2, \vec{\alpha}_3]$. Platí totiž

$$\vec{\alpha}_1 = \vec{\beta}_1 - \vec{\beta}_2$$
$$\vec{\alpha}_2 = \vec{\beta}_2$$
$$\vec{\alpha}_3 = \vec{\alpha}_3$$

čiže každý z vektorov $\vec{\alpha}_{1,2,3}$ vieme dostať ako lineárnu kombináciu $\vec{\beta}_1$, $\vec{\beta}_2$ a $\vec{\alpha}_3$.

Cvičenia

Úloha 4.3.1. Dokážte, že vektory $\vec{\alpha}_1, \ldots, \vec{\alpha}_n \in V$, kde $n \geq 2$ a $\vec{\alpha}_n \neq \vec{0}$, sú lineárne závislé práve vtedy, keď niektorý z nich je lineárnou kombináciou nasledujúcich.

Úloha 4.3.2. Nech $\vec{\alpha}, \vec{\beta}, \vec{\gamma}$ sú ľubovoľné vektory z vektorového priestoru V nad poľom $\mathbb{R}$. Potom $[\vec{\alpha}, \vec{\beta}, \vec{\gamma}] = [\vec{\alpha} + \vec{\beta}, \vec{\alpha} - \vec{\beta}, \vec{\gamma}]$.

Úloha 4.3.3. Nech $M = \{(x,y,z) \in \mathbb{R}^3; 2x + 3y + 5z = 0\}$. Ukážte, že M je vektorový podpriestor $\mathbb{R}^3$ a nájdite vektory, ktoré ho generujú.

Úloha 4.3.4. P_n označme množinu všetkých polynómov stupňa najviac n s reálnymi koeficientami. P_n je podpriestor vektorového priestoru všetkých zobrazení $f\colon \mathbb{R} \to \mathbb{R}$. Platí $P_n = [1, x, \ldots, x^n]$?

{lnzecvic:LZVZ5}

Úloha 4.3.5. Zistite, či dané vektory sú lineárne závislé v príslušnom vektorovom priestore:
a) (1,2,3), (1,3,2), (2,1,5) v $\mathbb{R}^3$,
b) (1,2,3), (1,3,2), (2,1,5), (1,127,3) v $\mathbb{R}^3$,
c) (1,3,4), (2,1,3), (3,1,4) v $\mathbb{Z}_5^3$
d) (1,3,4), (2,1,3), (3,1,4) v $\mathbb{Z}_7^3$.

{lnzecvic:LZFCIE}

Úloha 4.3.6. Zistite, či sú nasledujúce funkcie lineárne závislé vo vektorovom priestore všetkých funkcií z $\mathbb{R}$ do $\mathbb{R}$:
a) $x + 1$, x^2, x^3,
b) 1, $x + a$, $x^2 + bx + c$ (a, b, c môžu byť ľubovoľné reálne čísla),

c*) 1, $\cos x$, $\cos^2(\frac{x}{2})$,
d) x, $x(x-1)$, $x(x-1)(x-2)$,
e) 1, $\cos x$, $\cos 2x$.

Úloha 4.3.7. Ak $\vec{\alpha}$, $\vec{\beta}$, $\vec{\gamma}$ sú lineárne nezávislé vo vektorovom priestore V nad poľom $\mathbb{R}$, tak aj $\vec{\alpha}+\vec{\beta}$, $\vec{\alpha}+\vec{\gamma}$, $\vec{\beta}+\vec{\gamma}$ sú lineárne nezávislé. (Platilo by to aj vo vektorovom priestore nad poľom $\mathbb{Z}_2$?)

{lnze:CV1}

Úloha 4.3.8. Množina $\{\vec{\alpha}\}$ je lineárne nezávislá práve vtedy, keď $\vec{\alpha} \neq \vec{0}$. Dva vektory $\vec{\alpha}, \vec{\beta}$ sú lineárne závislé práve vtedy, keď jeden z nich je násobkom druhého (t.j. existuje $c \in F$ tak, že $c.\vec{\alpha} = \vec{\beta}$), alebo jeden z nich je $\vec{0}$.

Ak vektory $\vec{\alpha}, \vec{\beta}$ sú lineárne nezávislé, tak $\vec{\alpha}, \vec{\beta}, \vec{\gamma}$ sú lineárne závislé práve vtedy, keď $\vec{\gamma}$ je lineárna kombinácia vektorov $\vec{\alpha}, \vec{\beta}$.

{lnzecvic:ULOHRNADQ}

Úloha 4.3.9*. Overte, že $\mathbb{R}$ je vektorový priestor nad poľom $\mathbb{Q}$. Dokážte, že v tomto priestore sú 1, $\sqrt{2}$ a $\sqrt{3}$ lineárne nezávislé.

Úloha 4.3.10. Ukážte, že vo vektorovom priestore $\mathbb{R}$ nad $\mathbb{Q}$ (z predošlej úlohy) sú lineárne nezávislé vektory $1+3\sqrt{2}$ a $2-\sqrt{2}$.

Úloha 4.3.11*. Sú 1, $\sqrt{2}$, $\sqrt{3}$, $\sqrt{6}$ lineárne nezávislé vo vektorovom priestore $\mathbb{R}$ nad poľom $\mathbb{Q}$? (Hint: Úlohu môže o niečo zjednodušiť, ak sa pozriete na 1 a $\sqrt{3}$ ako prvky priestoru $\mathbb{R}$ nad poľom $\mathbb{Q}(\sqrt{2}) = \{a+b\sqrt{3}; a, b \in \mathbb{Q}\}$.)

Úloha 4.3.12. Nech $\vec{\alpha}, \vec{\beta}, \vec{\gamma}$ sú ľubovoľné vektory. Zistite, či sú tieto systémy vektorov lineárne závislé:
a) $\vec{\alpha}, \vec{\beta}, \vec{\alpha}+\vec{\beta}, \vec{\gamma}$, b) $\vec{\alpha}, \vec{\beta}, 0$, c) $\vec{\alpha}, \vec{\alpha}, \vec{\beta}, \vec{\gamma}$, d) $\vec{\alpha}+\vec{\beta}+\vec{\gamma}, \vec{\alpha}+\vec{\beta}, \vec{\alpha}+\vec{\gamma}, \vec{\beta}+\vec{\gamma}$.

Úloha 4.3.13. Nájdite 4 vektory v $\mathbb{R}^2$ tak, aby každé dva z nich boli lineárne nezávislé.

Úloha 4.3.14. Nech vektory $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ sú lineárne nezávislé vektory v nejakom vektorovom priestore nad poľom $\mathbb{R}$. Sú aj vektory $\vec{\alpha}_1, \vec{\alpha}_1+2\vec{\alpha}_2, \ldots, \vec{\alpha}_1+2\vec{\alpha}_2+\ldots+n\vec{\alpha}_n$ lineárne nezávislé?

4.4 Báza a dimenzia

V tejto podkapitole zadefinujeme pojmy báza a dimenzia vektorového priestoru. Pri dôkazoch základných výsledkoch o nich bude pre nás základným prostriedkom Steinitzova veta o výmene.

{baza:DEFFG}

Definícia 4.4.1. Nech V je vektorový priestor. Hovoríme, že V je *konečnorozmerný* ak existuje taká konečná množina vektorov $\{\vec{\alpha}_1, \ldots, \vec{\alpha}_n\}$, že platí $[\vec{\alpha}_1, \ldots, \vec{\alpha}_n] = V$.

Inými slovami: konečnorozmerný vektorový priestor je priestor, ktorý je generovaný nejakou konečnou množinou vektorov.

{baza:DEF}

Definícia 4.4.2. Nech V je vektorový priestor nad poľom F. Množinu vektorov $\{\vec{\alpha}_1, \ldots, \vec{\alpha}_n\}$ nazývame *bázou* priestoru V, ak

(i) vektory $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ sú lineárne nezávislé,

(ii) $V = [\vec{\alpha}_1, \ldots, \vec{\alpha}_n]$.

(Stručne: Báza je taká množina lineárne nezávislých vektorov, ktorá generuje celý priestor.)

Príklad 4.4.3. Priestor $V = \{\vec{0}\}$ nemá bázu (pretože v ňom neexistujú žiadne lineárne nezávislé vektory). Je však konečnorozmerný, keďže $V = [\vec{0}]$.

Príklad 4.4.4. Nech F je pole. Ako F^n budeme označovať vektorový priestor všetkých usporiadaných n-tíc prvkov poľa F. Sčitovanie a násobenie skalárom definujeme po súradniciach (podobne ako v príklade 4.1.4 pre $F = \mathbb{R}$).

Ako $\vec{\varepsilon}_i$ označíme vektor, ktorý má na všetkých súradniciach 0, iba na i-tej súradnici 1, teda

$$\begin{aligned}\vec{\varepsilon}_1 &= (1, 0, \ldots, 0),\\ \vec{\varepsilon}_2 &= (0, 1, \ldots, 0),\\ &\ldots\\ \vec{\varepsilon}_n &= (0, \ldots, 0, 1).\end{aligned}$$

Vektory $\vec{\varepsilon}_1, \vec{\varepsilon}_2, \ldots, \vec{\varepsilon}_n$ tvoria bázu vektorového priestoru F^n. Túto bázu nazývame *štandardná báza* F^n.

Overme, že táto množina vektorov spĺňa podmienky z definície 4.4.2.

Ak $c_1(1, 0, \ldots, 0) + c_2(0, 1, \ldots, 0) + \ldots + c_n(0, \ldots, 0, 1) = (c_1, c_2, \ldots, c_n) = (0, 0, \ldots, 0)$, tak platí $c_1 = c_2 = \ldots = c_n = 0$, teda vektory $\vec{\varepsilon}_1, \vec{\varepsilon}_2, \ldots, \vec{\varepsilon}_n$ sú naozaj lineárne nezávislé.

Ak máme ľubovoľný vektor $(x_1, x_2, \ldots, x_n) \in F^n$, dá sa získať ako lineárna kombinácia $(x_1, x_2, \ldots, x_n) = x_1(1, 0, \ldots, 0) + x_2(0, 1, \ldots, 0) + \ldots + x_n(0, \ldots, 0, 1) = x_1\vec{\varepsilon}_1 + x_2\vec{\varepsilon}_2 + \cdots + x_n\vec{\varepsilon}_n$. Teda vektory $\vec{\varepsilon}_1, \vec{\varepsilon}_2, \ldots, \vec{\varepsilon}_n$ skutočne generujú celý priestor F^n.

Veľmi podobným postupom ako sme použili v tomto príklade by sme vedeli zdôvodniť, že F^M je nekonečnorozmerný priestor pre ľubovoľné pole F a *nekonečnú* množinu M.

Ako neskôr ukážeme, všetky konečnorozmerné vektorové priestory nad poľom F sú v istom zmysle podobné ako priestory F^n.

{baza:VTDVEBAZY}

Veta 4.4.5. *Ľubovoľné dve bázy konečnorozmerného vektorového priestoru V majú rovnaký počet prvkov.*

Dôkaz. Nech $\{\vec{\alpha}_1, \ldots, \vec{\alpha}_n\}$ a $\{\vec{\beta}_1, \ldots, \vec{\beta}_s\}$ sú dve bázy toho istého vektorového priestoru V. Pretože $V = [\vec{\alpha}_1, \ldots, \vec{\alpha}_n]$ a vektory $\vec{\beta}_1, \ldots, \vec{\beta}_s$ sú lineárne nezávislé, podľa Steinitzovej vety o výmene platí

$$s \leq n.$$

Analogicky môžeme dokázať opačnú nerovnosť $n \leq s$. Tieto dve nerovnosti spolu dávajú rovnosť $n = s$. □

Veta 4.4.6. *Nech V je konečnorozmerný vektorový priestor. Ak $\vec{\beta}_1, \ldots, \vec{\beta}_s \in V$ sú lineárne nezávislé, tak sa dajú doplniť na bázu priestoru V.*

Dôkaz. Ak V je konečnorozmerný, tak podľa definície existujú vektory $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ také, že $V = [\vec{\alpha}_1, \ldots, \vec{\alpha}_n]$. Na základe Steinitzovej vety môžeme vektory $\vec{\beta}_1, \ldots, \vec{\beta}_s$ doplniť niektorými z týchto vektorov tak, aby generovali celý priestor. Nech k je najmenší možný počet vektorov, ktorými ich môžeme takto doplniť. (Steinitzova veta hovorí, že sa to určite dá $n - s$ vektormi, nevylučuje však, že niekedy môže stačiť aj menší počet.) Bez ujmy na všeobecnosti, nech $V = [\vec{\beta}_1, \ldots, \vec{\beta}_s, \vec{\alpha}_1, \ldots, \vec{\alpha}_k]$.

Chceme dokázať, že vektory $\vec{\beta}_1, \ldots, \vec{\beta}_s, \vec{\alpha}_1, \ldots, \vec{\alpha}_k$ tvoria bázu priestoru V. Pretože sme ich vybrali tak, že generujú celý priestor, zostáva nám dokázať, že sú lineárne nezávislé.

Postupujme sporom – predpokladajme, že by boli lineárne závislé. Potom je niektorý z nich lineárnou kombináciou predchádzajúcich vektorov. Bez ujmy na všeobecnosti, nech je to $\vec{\alpha}_k$. (Nemôže to byť žiadny z vektorov $\vec{\beta}_1, \ldots, \vec{\beta}_s$, pretože tieto vektory sú lineárne nezávislé.) Potom ale platí

$$V = [\vec{\beta}_1, \ldots, \vec{\beta}_s, \vec{\alpha}_1, \ldots, \vec{\alpha}_k] = [\vec{\beta}_1, \ldots, \vec{\beta}_s, \vec{\alpha}_1, \ldots, \vec{\alpha}_{k-1}],$$

čo je v spore s tým, že k je najmenší možný počet vektorov, ktorými sa vektory $\vec{\beta}_1, \ldots, \vec{\beta}_s$ dajú doplniť tak, aby generovali celý priestor V. □

Dôsledok 4.4.7. *Každý konečnorozmerný vektorový priestor $V \neq \{\vec{0}\}$ má bázu.*

Poznámka 4.4.8. Ak sme nejaký pojem definovali, vôbec to nemusí znamenať, že taký objekt aj naozaj existuje. Preto je predchádzajúci dôsledok dôležitý. (Hoci táto poznámka znie nesmierne naivne, skutočne sa možno často stretnúť s chybami takéhoto typu.)

Definícia 4.4.9. *Dimenziou* konečnorozmerného vektorového priestoru V nazývame počet prvkov ľubovoľnej jeho bázy. (Pre nulový priestor dodefinujeme $d(\{\vec{0}\}) = 0$.) Toto číslo označujeme $d(V)$.

Poznámka 4.4.10. Aby mala predchádzajúca definícia zmysel, museli sme najprv dokázať, že v konečnorozmernom vektorovom priestore existuje báza a že ľubovoľné dve bázy musia mať rovnaký počet prvkov; teda naša definícia nezávisí od voľby bázy.

S podobnou situáciou – že sa nejaký objekt zadefinuje, ale bude potrebné overiť správnosť definície – sa v matematike stretnete ešte veľakrát.

Príklad 4.4.11. Pretože vektory $\vec{\varepsilon}_1, \vec{\varepsilon}_2, \ldots, \vec{\varepsilon}_n$ tvoria bázu vektorového priestoru F^n, platí $d(F^n) = n$.

Dôsledok 4.4.12. *Ak V je konečnorozmený vektorový priestor a $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ sú lineárne nezávislé vo V, tak $n \leq d(V)$.*

Príklad 4.4.13. Vektory $(1,2,3), (2,3,4), (3,4,5), (4,5,6)$ sú lineárne závislé v $\mathbb{R}^3$.

Pretože vieme, že $d(\mathbb{R}^3) = 3$, nemôžu byť podľa predchádzajúcej vety v tomto priestore viac ako 3 lineárne nezávislé vektory.

Bázu sme definovali pomocou dvoch podmienok. Nasledujúca, veľmi užitočná veta hovorí, že ak už vieme, že nejaká množina vektorov má „správny“ počet prvkov, môžeme jednu z týchto podmienok vynechať.

{baza:VTEKV}

Veta 4.4.14. *Nech V je konečnorozmerný vektorový priestor a $d(V) = n$. Nasledujúce podmienky sú ekvivalentné:*

{baza:E1}

(i) *$\{\vec{\alpha}_1, \ldots, \vec{\alpha}_n\}$ je báza priestoru V,*

{baza:E2}

(ii) *vektory $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ sú lineárne nezávislé,*

{baza:E3}

(iii) $V = [\vec{\alpha}_1, \ldots, \vec{\alpha}_n]$.

Dôkaz. Implikácie (i) ⇒ (ii), (i) ⇒ (iii) vyplývajú priamo z definície.

(ii) ⇒ (i): Ak máme n lineárne nezávislých vektorov $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$, podľa Steinitzovej vety ich môžeme doplniť $n - n = 0$ vektormi na množinu generujúcu celý priestor V. Teda nemusíme pridávať žiadne vektory a už množina $\{\vec{\alpha}_1, \ldots, \vec{\alpha}_n\}$ je báza (generuje V a je aj lineárne nezávislá).

(iii)⇒ (i): Sporom. Ak by boli vektory $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ lineárne závislé, dali by sa niektoré z nich vynechať tak, aby stále tieto vektory generovali celý priestor V. Dostali by sme k vektorov, ktoré generujú V, pričom $k < n$. Súčasne by priestor V mal n-prvkovú bázu, ktorá je tvorená lineárne nezávislými vektormi. Zo Steinitzovej vety potom vyplýva $n < k$. Odvodili sme súčasnú platnosť nerovností $k < n$ aj $n < k$ – spor. □

Príklad 4.4.15. S použitím predchádzajúcej vety by sme mohli overiť, že vektory $\vec{\varepsilon}_1, \ldots, \vec{\varepsilon}_n$ tvoria bázu priestoru F^n. Prvý spôsob: Overili by sme, že generujú celý priestor. Druhý spôsob: Sú lineárne nezávislé. (Kým sme nevedeli, že $d(V) = n$, potrebovali sme overiť obe tieto vlastnosti.)

{baza:VTBAZAJEDNOZ}

Veta 4.4.16. *Nech V je vektorový priestor. Vektory $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ tvoria bázu priestoru V práve vtedy, keď každý vektor $\vec{\beta}$ sa dá jednoznačne vyjadriť ako*

$$\vec{\beta} = c_1\vec{\alpha}_1 + \cdots + c_n\vec{\alpha}_n.$$

Dôkaz. $\boxed{\Rightarrow}$ Pretože $V = [\vec{\alpha}_1, \ldots, \vec{\alpha}_n]$, každý vektor sa dá vyjadriť ako lineárna kombinácia vektorov $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$. Ešte treba overiť jednoznačnosť takéhoto vyjadrenia. Nech $\vec{\beta} = c_1\vec{\alpha}_1 + \cdots + c_n\vec{\alpha}_n = d_1\vec{\alpha}_1 + \cdots + d_n\vec{\alpha}_n$ sú dve vyjadrenia vektoru $\vec{\beta}$. Úpravou tejto rovnosti dostaneme

$$(c_1 - d_1)\vec{\alpha}_1 + \ldots + (c_n - d_n)\vec{\alpha}_n = \vec{0}.$$

Z lineárnej nezávislosti vektorov $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ vyplýva $c_i - d_i = 0$, čiže $c_i = d_i$ pre $i = 1, 2, \ldots, n$.

$\boxed{\Leftarrow}$ Pretože každý vektor z V sa dá vyjadriť pomocou vektorov $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$, tieto vektory generujú priestor V, čiže $V = [\vec{\alpha}_1, \ldots, \vec{\alpha}_n]$.

Ďalej vieme, že $\vec{0}$ sa dá vyjadriť ako lineárna kombinácia vektorov $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ jediným spôsobom. Z rovnosti $c_1\vec{\alpha}_1 + \cdots + c_n\vec{\alpha}_n = \vec{0} = 0\cdot\vec{\alpha}_1 + \ldots + 0\cdot\vec{\alpha}_n$ teda vyplýva $c_1 = \ldots = c_n = 0$. Zistili sme, že vektory $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ sú lineárne nezávislé. □

Ešte si ukážeme pár užitočných tvrdení o podpriestoroch konečnorozmerných priestorov.

{baza:VTPPRKON}

Veta 4.4.17. *Ľubovoľný podpriestor S konečnorozmerného priestoru V je konečnorozmerný. Navyše, $d(S) \leq d(V)$.*

Dôkaz. Pretože $S \subseteq V$ a $d(V) = n$, číslo n udáva horné ohraničenie pre počet lineárne nezávislých vektorov z S. Nech $\vec{\alpha}_1, \ldots, \vec{\alpha}_k$ je najväčší systém lineárne nezávislých vektorov z S. Platí $k \leq n$. Stačí nám dokázať, že $\vec{\alpha}_1, \ldots, \vec{\alpha}_k$ tvorí bázu priestoru S, čiže $S = [\vec{\alpha}_1, \ldots, \vec{\alpha}_k]$.

Predpokladajme, že by existoval vektor $\vec{\alpha} \in S$, ktorý nepatrí do $[\vec{\alpha}_1, \ldots, \vec{\alpha}_k]$. Teda $\vec{\alpha}$ sa nedá vyjadriť ako lineárna kombinácia vektorov $\vec{\alpha}_1, \ldots, \vec{\alpha}_k$, čo znamená, že $\vec{\alpha}_1, \ldots, \vec{\alpha}_k, \vec{\alpha}$ sú lineárne nezávislé. To však je spor s predpokladom, že $\vec{\alpha}_1, \ldots, \vec{\alpha}_k$ je najväčší systém lineárne nezávislých vektorov v S. □

Úloha 4.4.1. Viete povedať, na ktorom mieste predchádzajúceho dôkazu sme využili, že V je konečnorozmerný?

{baza:TVRPPRDIM}

Tvrdenie 4.4.18. *Ak S je podpriestor konečnorozmerného vektorového priestoru V a $d(S) = d(V)$, tak $S = V$.*

Dôkaz. Označme $n := d(S) = d(V)$. Nech $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ je báza S. Keďže je to n vektorov vo V, ktoré sú lineárne nezávislé, podľa vety 4.4.14 je to súčasne báza V. Teda $S = [\vec{\alpha}_1, \ldots, \vec{\alpha}_n] = V$. □

Pozrime sa na konkrétnom príklade na to, aká je báza niektorých jednoduchých podpriestorov priestoru $\mathbb{R}^4$.

Príklad 4.4.19. Pripomeňme, že v priestore $V = \mathbb{R}^4$ máme nulový podpriestor $\{\vec{0}\}$, o ktorom vieme že $d(\{\vec{0}\}) = 0$. Takisto vieme, že V je podpriestor V a v tomto prípade $d(V) = 4$. Pozrieme sa na nejaké ďalšie, o niečo menej triviálne podpriestory.

Začnime s podpriestorom

$$S_1 = \{(x, y, w, z) \in \mathbb{R}^4; x + y + w + z = 0\}.$$

(Nebudeme overovať, že je to podpriestor – nemalo by byť pre vás ťažké overiť definíciu alebo kritérium vektorového podpriestoru.)

Ak si zvolíme ľubovoľné $y, w, z \in \mathbb{R}$, tak aby vektor (x, y, w, z) patril do S_1, tak musí platiť $x = -y - w - z$. Priamo dosadením do rovnice $x + y + z + w = 0$ zistíme, že každý vektor tvaru $(-y - w - z, y, w, z)$ patrí so S_1. Teda tento podpriestor môžeme inak vyjadriť ako

$$S_1 = \{(-y - w - z, y, w, z); y, w, z \in \mathbb{R}\}.$$

Po úprave máme $(-y - w - z, y, w, z) = y(-1, 1, 0, 0) + w(-1, 0, 1, 0) + z(-1, 0, 0, 1)$, z čoho vidíme, že

$$S_1 = [(-1, 1, 0, 0), (-1, 0, 1, 0), (-1, 0, 0, 1)].$$

Je ľahké overiť, že vektory $(-1, 1, 0, 0)$, $(-1, 0, 1, 0)$, $(-1, 0, 0, 1)$ sú aj lineárne nezávislé, takže tieto tri vektory tvoria bázu podpriestoru S_1. Dostávame teda, že $d(S_1) = 3$.

Podpriestor S_1 sme dostali ako množinu riešení jedinej rovnice $x + y + z + w = 0$. Aby sme mali o trošičku komplikovanejší príklad, skúsme pridať ešte jednu rovnicu a pozrieť sa na podpriestor

$$S_2 = \{(x, y, w, z) \in \mathbb{R}^4; x + y + w + z = 0, x - y + w - z = 0\}.$$

Opäť môžeme použiť podobnú úvahu ako pre podpriestor S_1. Predpokladajme, že máme zvolené $w, z \in \mathbb{R}$ a pozrime sa na to, či sú jednoznačne určené x aj y.

Sčítaním zadaných rovníc dostaneme $2x + 2w = 0$, z čoho vyplýva $x = -w$. Keď ich odčítame, tak dostaneme $2y + 2z = 0$, čo znamená, že $y = -z$. Teda všetky vektory z S_2 musia mať tvar $(-w, -z, w, z) = w(-1, 0, 1, 0) + z(0, -1, 0, 1)$ a každý vektor takéhoto tvaru patrí do S_2, o čom sa dá ľahko presvedčiť dosadením. Zistili sme, že

$$S_2 = \{(-w, -z, w, z); w, z \in \mathbb{R}\} = [(-1, 0, 1, 0), (0, -1, 0, 1)].$$

Vektory $(-1, 0, 1, 0)$ a $(0, -1, 0, 1)$ sú lineárne nezávislé, teda sme našli dvojprvkoú bázu pre S_2. Znamená to, že $d(S_2) = 2$.

Skúsme pridať ešte jednu rovnicu.

$$S_3 = \{(x, y, w, z) \in \mathbb{R}^4; x + y + w + z = 0, x - y + w - z = 0, x - y - w + z = 0\}.$$

Z výpočtov, ktoré sme robili pre S_2, vieme, že ak nejaký vektor vyhovuje prvým dvom rovniciam, tak musí spĺňať $x = -w$ a $y = -z$. Keď dosadíme do tretej rovnice, dostaneme

$$x - y - w + z = -w + z - w + z = -2w + 2z = 0.$$

Z tejto rovnice vyplýva, že $w = z$. Teda akýkoľvek vektor z S_3 musí mať tvar $(-z, -z, z, z)$. Dosadením sa presvedčíme, že takéto vektory vyhovujú zadaným rovniciam, a teda

$$S_3 = \{(-z, -z, z, z); z \in \mathbb{R}\} = [(-1, -1, 1, 1)].$$

Vidíme, že $d(S_3) = 1$.

Opäť pridajme ešte jednu rovnicu, napríklad sa skúsme pozrieť na podpriestor

$$S_4 = \{(x,y,w,z) \in \mathbb{R}^4; x+y+w+z=0, x-y+w-z=0, x-y-w+z=0, x+y+w-z=0\}.$$

Ak dosadíme do poslednej rovnice $x=-z$, $y=-z$, $w=z$ (už vieme, že tieto rovnosti spĺňajú riešenia prvých troch rovníc),tak dostaneme

$$x+y+w-z = -z-z+z-z = -2z = 0,$$

z čoho vyplýva, že $z=0$. Preto

$$S_4 = \{(0,0,0,0)\}$$

a $d(S_4)=0$.

Neskôr, v časti 5.7, sa naučíme ako efektívnejšie a jednoduchšie rátať bázu a dimenziu podpriestorov podobných ako v predošlej úlohe.

Poznámka 4.4.20. Na predošlej úlohe si môžeme všimnúť istú zákonitosť. Začali sme s priestorom dimenzie 4 a po pridaní novej rovnice dimenzia vždy klesla o jedna. To nemusí platiť vždy – dôležité je, že sme pridávavli rovnicu, ktorá nevyplýva z predošlých rovníc (nebola ich lineárnou kombináciou.)

Niečo podobné si vieme geometricky predstaviť v $\mathbb{R}^3$. Jedna rovnica nám určí rovinu, teda podpriestor dimenzie 2. (Budeme teraz uvažovať len lineárne rovnice s nulovou pravou stranou, teda vždy dostaneme rovinu prechádzajúcu nulou.) Ak pridáme ďalšiu rovnicu, dostaneme ďalšiu rovinu. Ak je to iná rovina ako tá, čo vyšla z prvej rovnice, tak ako prienik dostaneme priamku – podpriestor dimenzie 1. Ak túto priamku pretneme s ďalšou rovinou, tak by sme mali dostať jediný bod – podpriestor dimenzie 0. Nestalo by sa to iba v prípade, že nová rovina by prechádzala touto priamkou. To by zodpovedalo lineárnej závislosti rovníc, s ktorými pracujeme.

Takisto si to môžete rozmyslieť v rovine, kde takýmto rovniciam zodpovedajú priamky.

Keď sa budeme zaoberať riešeniami sústav lineárnych rovníc, tak túto geometrickú predstavu budeme vedieť presnejšie vyjadriť a aj dokázať – vlastne to vyjadruje dôsledok 5.7.5.

Ukážme si ešte jeden príklad, kde by sme na prvý pohľad neočakávali, že nám vedomosti o dimenzii vektorových priestorov môžu pomôcť.

{baza:PRIKLPOLESQRT32}

Príklad 4.4.21. Pokúsme sa ukázať, že

$$F = \{a + b\sqrt[3]{2} + c\sqrt[3]{2^2}; a,b,c \in \mathbb{Q}\}$$

je pole.

Úlohy podobného typu sme riešili v úlohe 3.3.2. Veľa vlastností sa ukáže veľmi jednoducho (mnohé sa priamo zdedia z poľa reálnych čísel.) Jediná vlastnosť, ktorej dôkaz nie je úplne jednoduchý, je existencia inverzného prvku pre každé $\alpha \in F \setminus \{0\}$. Takže sa poďme pozrieť na túto vlastnosť.

Skúsme si ale ešte predtým uvedomiť to, že F je vektorový priestor nad poľom $\mathbb{Q}$. (To sa overí veľmi ľahko – na základe podobných argumentov ako v úlohe 4.1.10.) Navyše tiež ľahko vidieť, že tento vektorový priestor je vygenerovaný prvkami 1, $\sqrt[3]{2}$ a $\sqrt[3]{2^2}$, teda má dimenziu najviac 3.

Nech teda $\alpha = a+b\sqrt[3]{2}+c\sqrt[3]{2^2}$, $\alpha \neq 0$. Čísla 1, α, α^2, α^3 všetky patria do F. Sú to teda 4 vektory vo vektorovom priestore dimenzie najviac 3, musia teda byť lineárne závislé. Z toho dostávame, že existujú racionálne čísla $a,b,c,d \in \mathbb{Q}$, ktoré nie sú všetky nulové a platí pre ne

$$a + b\alpha + c\alpha^2 + d\alpha^3 = 0.$$

Ak $a \neq 0$, tak túto rovnosť môžeme upraviť na tvar

$$-\alpha\left(\frac{b}{a}+\frac{c}{a}\alpha^2+\frac{d}{a}\alpha^3\right)=1,$$

čím je dokázané, že pre α existuje inverzný prvok.

Ak by platilo $a=0$, tak máme rovnosť $\alpha(b+c\alpha+d\alpha^2)=0$, z ktorej na základe nenulovosti prvku α dostaneme

$$b+c\alpha+d\alpha^2=0.$$

Opäť, ak $b\neq 0$, vieme dostať inverzný prvok podobný spôsobom ako v predošlom prípade. Ak $b=0$, tak dostaneme $c\alpha+d=0$.

Teraz už dostávame buď $\alpha=\frac{d}{c}$, čo znamená, že α je racionálne a má inverzný prvok. Alebo dostaneme $a=b=c=d=0$, lenže o týchto číslach sme predpokladali, že aspoň jedno z nich je nenulové.

Môžete vyskúšať, či by ste vedeli priamo vypočítať inverzný prvok k danému prvku poľa F. Zistíte, že je to pomerne zdĺhavý výpočet. V príklade 6.5.6 si ukážeme, ako sa tento výpočet dá o niečo urýchliť. V budúcom semestri na predmete Algebra 2 sa budeme učiť aj o konečných rozšíreniach polí. V súvislosti s touto témou budú často užitočné úvahy takého typu, ako sme robili v tomto príklade.

V tejto časti sme sa zaoberali iba konečnorozmernými vektorovými priestormi. (A takisto aj v nasledujúcich častiach nájdete veľa výsledkov, ktoré dokážeme iba pre konečnorozmerné vektorové priestory.) Azda by bolo užitočné vidieť aj príklad vektorového priestoru, ktorý nie je konečnorozmerný.

Príklad 4.4.22. Vektorový priestor $\mathbb{R}^{\mathbb{R}}$ všetkých zobrazení z $\mathbb{R}$ do $\mathbb{R}$ (príklad 4.1.4) nie je konečnorozmerný.

Predpokladajme, že by bol konečnorozmerný. Potom by existoval konečný počet funkcií $g_1,\dots,g_n\colon\mathbb{R}\to\mathbb{R}$ tak, že $[g_1,\dots,g_n]=\mathbb{R}^{\mathbb{R}}$. Ak sa nám podarí zostrojiť $n+1$ funkcií, ktoré sú v $\mathbb{R}^{\mathbb{R}}$ lineárne nezávislé, tak pomocou Steinitzovej vety ľahko dostaneme spor ($n+1\leq n$).

Pokúsme sa teda definovať takéto funkcie. Pre $k=0,1,\dots,n$ definujme zobrazenie $f_k\colon\mathbb{R}\to\mathbb{R}$ ako

$$f_k(x)=\begin{cases}1, & \text{ak } x=k\\ 0, & \text{ak } x\neq k\end{cases}$$

Tvrdíme, že $f_0,\dots,f_n$ sú lineárne nezávislé. Skutočne, ak platí rovnosť $c_0f_0+c_1f_1+\dots+c_nf_n=0$ (kde 0 označuje nulovú funkciu), tak pre každé $x\in\mathbb{R}$ máme

$$c_0f_0(x)+c_1f_1(x)+\dots+c_nf_n(x)=0.$$

Špeciálne, musí to platiť aj keď za x dosadíme $k=0,1,\dots,n$. V takom prípade však dostávame $f_k(k)=1$ a $f_j(k)=0$, teda z predchádzajúce rovnosti priamo dostávame

$$c_k=0.$$

Poznámka 4.4.23. Možno vám napadla otázka, či sa dá definovať báza aj pre nekonečnorozmerné vektorové priestory. Dá sa to, v tomto prípade sa zvykne nazývať *Hamelova báza.* Na jej zavedenie by sme však potrebovali podstatne väčšie vedomosti z teórie množín. Dokonca platí aj analógia vety 4.4.5, čiže aj ľubovoľné 2 Hamelove bázy majú rovnaký „počet“ prvkov – s tým rozdielom, že pre nekonečné množiny najprv treba definovať nový pojem, ktorý by zodpovedal počtu prvkov konečných množín (nazýva sa kardinalita množiny, viac

sa o nej dozviete na iných predmetoch). Pre prípad, že by vás to zaujímalo a chceli by ste sa k tomuto problému časom vrátiť uvediem aj niekoľko odkazov na literatúru. V [NS] je pekným spôsobom dokázané, že ľubovoľné dve Hamelove bázy toho istého priestoru musí mať rovnakú „veľkosť“ (kardinalitu). V [ŠS, Kapitola 10.3] autori definujú Hamelovu bázu v špeciálnom prípade – pre reálne čísla ako vektorový priestor nad poľom $\mathbb{Q}$ (úloha* 4.3.9). Niečo o Hamelovej báze (a aj nejakých je aplikáciách) si môžete prečítať aj v [Sl2].

Ešte raz zdôrazňujem, že túto poznámku som sem vložil len kvôli tomu, aby ste vedeli, kde môžete hľadať v prípade, že by ste sa k takémuto niečomu chceli neskôr vrátiť. (Zatiaľ by to pre vás bolo pomerne ťažké, potrebujete na to najprv poznať základné fakty o kardinalite množín a na dôkaz existencie bázy aj niečo o Zornovej leme.)

Cvičenia

Úloha 4.4.2. Zistite, či dané vektory tvoria bázu v $\mathbb{R}^3$:
a) (1,2,3), (1,-2,3), (1,2,-3)
b) (1,1,1), (1,1,0), (1,0,1)
c) (1,0,0), (0,1,0), (0,0,1), (1,1,1).

Úloha 4.4.3. Zistite, či dané vektory tvoria bázu v $\mathbb{Z}_5^3$:
a) (1,2,3), (2,3,4), (0,3,1)
b) (1,0,0), (0,1,2), (2,1,3)
c) (0,1,2), (3,0,1), (1,0,2).

Úloha 4.4.4. P_n označme priestor všetkých polynómov stupňa najviac n. Overte, že $d(P_n) = n+1$ a že $1, x-1, \dots (x-1)^n$ je báza tohoto priestoru.

Úloha 4.4.5. Určte dimenziu podpriestoru $[\vec{\alpha}, \vec{\beta}, \vec{\gamma}]$, ak $\vec{\alpha} = (1,3,2,1)$, $\vec{\beta} = (4,9,5,4)$ a $\vec{\gamma} = (3,7,4,3)$ v $\mathbb{R}^4$.

{bazacvic:DOPLNBAZ}

Úloha 4.4.6. Ak sa to dá, doplňte dané vektory na bázu príslušného vektorového priestoru:
a) (1,1,2), (2,1,3) v $\mathbb{R}^3$,
b) x^2-1, x^2+1 v priestore polynómov stupňa najviac 3,
c) (1,2,3,0), (3,4,1,2) v $\mathbb{Z}_5^4$.

Úloha 4.4.7. Máme dané vektory $\vec{\alpha}_1 = (1,1,2,0)$, $\vec{\alpha}_2 = (0,0,3,1)$ v priestore $\mathbb{Z}_5^4$. Koľko existuje možností na výber vektorov $\vec{\alpha}_3, \vec{\alpha}_4 \in \mathbb{Z}_5^4$ tak, aby tieto štyri vektory tvorili bázu?

Úloha 4.4.8. Ak každý z vektorov $\vec{\beta}_1, \dots, \vec{\beta}_k$ je lineárnou kombináciou vektorov $\vec{\alpha}_1, \dots, \vec{\alpha}_m$, tak $d([\vec{\beta}_1, \dots, \vec{\beta}_k]) \leq d([\vec{\alpha}_1, \dots, \vec{\alpha}_m])$.

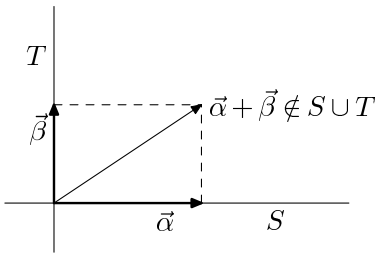
Úloha 4.4.9. Overte, že množina $S = \{f\colon \mathbb{R} \to \mathbb{R} : (\exists a, b \in \mathbb{R})(\forall x \in \mathbb{R}) f(x) = ax + b\}$ je podpriestor priestoru všetkých funkcií z $\mathbb{R}$ do $\mathbb{R}$. Nájdite funkcie $g, h \in S$ také, že $S = [g, h]$.

Úloha 4.4.10. Zistite, či $S = \{f\colon \mathbb{R} \to \mathbb{R}; f(x) = ax^2 + bx + c, a, b, c \in \mathbb{R}\}$ je vektorový podpriestor priestoru reálnych funkcií. Ak áno, nájdite, $g_1, g_2, g_3 \in S$ také, že $S = [g_1, g_2, g_3]$.

Úloha 4.4.11. Nájdite bázu pre každý vektorový podpriestor z úlohy 4.2.3.

4.5 Lineárne a direktné súčty podpriestorov

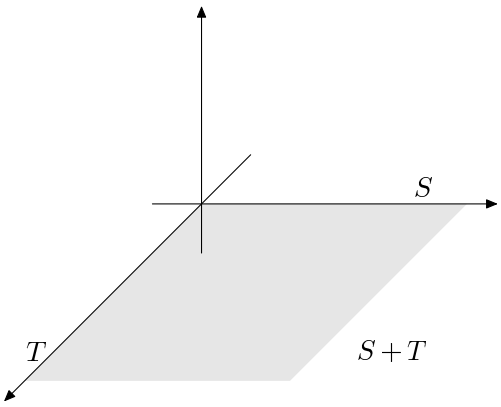
Už vieme, že prienik podpriestorov vektorového priestoru je tiež podpriestor (vety 4.2.8 a 4.2.10). Ako je to so zjednotením? Ak si zvolíme podpriestory $S = [(1,0,0)]$ a $T = [(0,1,0)]$



cty:FIGVECT4} Obr. 4.4: Zjednotenie 2 podpriestorov nemusí byť podpriestor

priestoru R^3, tak vidíme, že $S \cup T$ nie je vektorový podpriestor, lebo $(1,0,0) \in S \cup T$, $(0,1,0) \in S \cup T$, ale $(1,0,0)+(0,1,0) = (1,1,0) \notin S \cup T$.

Zaujímalo by nás, ako vyzerá najmenší podpriestor, ktorý obsahuje S aj T. Z obrázku 4.5 môžeme zistiť, že v tomto prípade je to podpriestor $[(1,0,0),(0,1,0)]$.



cty:FIGVECT5} Obr. 4.5: Najmenší podpriestor obsahujúci S aj T

Ukážeme si, ako možno nájsť takýto podpriestor vo všeobecnosti, pre ľubovoľné dva podpriestory daného vektorového priestoru V.

Veta 4.5.1. *Nech S, T sú vektorové podpriestory vektorového priestoru V nad poľom F. Potom*

$$S+T = \{\vec{\alpha}+\vec{\beta}; \vec{\alpha} \in S, \vec{\beta} \in T\}$$

je podpriestorom vektorového priestoru V.

Táto veta vlastne hovorí, že množina všetkých vektorov, ktoré sa dajú získať ako súčty vektorov z S a z T, tvorí vektorový podpriestor. Všimnite si, že v predchádzajúcom príklade bolo $S+T = [(1,0,0),(0,1,0)]$.

Dôkaz. Overíme podmienky z definície vektorového podpriestoru. Množina $S+T$ je *neprázdna,* lebo $\vec{0} \in S$, $\vec{0} \in T$, čiže $\vec{0} = \vec{0}+\vec{0} \in S+T$.

$S+T$ je uzavretá na súčty: Ak $\vec{\gamma}_1, \vec{\gamma}_2 \in S+T$, tak vektory $\vec{\gamma}_1$, $\vec{\gamma}_2$ sa dajú napísať v tvare $\vec{\gamma}_1 = \vec{\alpha}_1+\vec{\beta}_1$, $\vec{\gamma}_2 = \vec{\alpha}_2+\vec{\beta}_2$, kde $\vec{\alpha}_1, \vec{\alpha}_2 \in S$ a $\vec{\beta}_1, \vec{\beta}_2 \in T$. Potom $\vec{\gamma}_1+\vec{\gamma}_2 = (\vec{\alpha}_1+\vec{\beta}_1)+(\vec{\alpha}_2+\vec{\beta}_2) = (\vec{\alpha}_1+\vec{\alpha}_2)+(\vec{\beta}_1+\vec{\beta}_2)$. (Využili sme komutatívnosť a asociatívnosť sčitovania.) Pretože S je

vektorový podpriestor vektor $\vec{\alpha}_1 + \vec{\alpha}_2$ patrí do S, podobne $\vec{\beta}_1 + \vec{\beta}_2 \in T$. Ukázali sme, že vektor $\vec{\gamma}_1 + \vec{\gamma}_2$ sa dá napísať ako súčet vektora z S a vektora z T, teda $\vec{\gamma}_1 + \vec{\gamma}_2 \in S + T$.

$S + T$ je uzavretá na násobenie skalárom: Ak $\vec{\gamma} \in S + T$, tak $\vec{\gamma} = \vec{\alpha} + \vec{\beta}$ pre nejaké $\vec{\alpha} \in S$ a $\vec{\beta} \in T$. Nech $c \in F$ je ľubovoľný skalár. Potom $c\vec{\gamma} = c\vec{\alpha} + c\vec{\beta}$. Pritom $c\vec{\alpha} \in S$, $c\vec{\beta} \in T$, čiže $c\vec{\gamma} \in S + T$. □

Definícia 4.5.2. Ak S, T sú podpriestory vektorového podpriestoru V, tak vektorový podpriestor $S + T$ sa nazýva *lineárny súčet* podpriestorov S a T.

Vidno, že S aj T sú podmnožiny $S + T$, čiže $S + T$ obsahuje oba podpriestory S aj T. ($\vec{\alpha} \in S \Rightarrow \vec{\alpha} = \vec{\alpha} + \vec{0} \in S + T$, podobne pre T.) Priestor $S + T$ je skutočne najmenší vektorový podpriestor priestoru V, ktorý obsahuje S aj T. Ak totiž $S, T \subseteq U$ a U je vektorový podpriestor V, tak U musí obsahovať všetky súčty tvaru $\vec{\alpha} + \vec{\beta}$, pretože $\vec{\alpha} \in S \subseteq S + T$ a $\vec{\beta} \in T \subseteq S + T$.

{sucty:VT1}

Veta 4.5.3. *Nech S a T sú podpriestory vektorového priestoru V nad poľom F. Nech $S = [\vec{\alpha}_1, \ldots, \vec{\alpha}_n]$, $T = [\vec{\beta}_1, \ldots, \vec{\beta}_m]$. Potom $S + T = [\vec{\alpha}_1, \ldots, \vec{\alpha}_n, \vec{\beta}_1, \ldots, \vec{\beta}_m]$.*

Dôkaz. Je zrejmé, že vektory $\vec{\alpha}_1, \ldots, \vec{\alpha}_n, \vec{\beta}_1, \ldots, \vec{\beta}_m$ patria do $S+T$. Keďže $S+T$ je vektorový podpriestor, musí potom platiť $[\vec{\alpha}_1, \ldots, \vec{\alpha}_n, \vec{\beta}_1, \ldots, \vec{\beta}_m] \subseteq S + T$.

Ešte treba dokázať opačnú inklúziu, čiže chceme ukázať, že

$$\vec{\gamma} \in S + T \;\Rightarrow\; \vec{\gamma} \in [\vec{\alpha}_1, \ldots, \vec{\alpha}_n, \vec{\beta}_1, \ldots, \vec{\beta}_m].$$

Ak $\vec{\gamma} \in S + T$, tak $\vec{\gamma} = \vec{\alpha} + \vec{\beta}$, kde $\vec{\alpha} \in [\vec{\alpha}_1, \ldots, \vec{\alpha}_n]$ a $\vec{\beta} \in [\vec{\beta}_1, \ldots, \vec{\beta}_m]$. To znamená, že existujú $c_1, \ldots, c_n, d_1, \ldots, d_m \in F$ tak, že $\vec{\alpha} = c_1\vec{\alpha}_1 + \ldots + c_n\vec{\alpha}_n$ a $\vec{\beta} = d_1\vec{\beta}_1 + \cdots + d_m\vec{\beta}_m$. Potom $\vec{\gamma} = c_1\vec{\alpha}_1 + \cdots + c_n\vec{\alpha}_n + d_1\vec{\beta}_1 + \cdots + d_m\vec{\beta}_m$, čiže $\vec{\gamma} \in [\vec{\alpha}_1, \ldots, \vec{\alpha}_n, \vec{\beta}_1, \ldots, \vec{\beta}_m]$. □

{sucty:VT2}

Veta 4.5.4. *Nech S, T sú podpriestory konečnorozmerného priestoru V. Potom*[5]

$$d(S) + d(T) = d(S + T) + d(S \cap T).$$

Dôkaz. Podľa vety 4.4.17 každý podpriestor konečnorozmerného priestoru je tiež konečnorozmerný, teda všetky dimenzie, ktoré vystupujú vo vete, sú skutočne definované.

V prípade, že $S \subseteq T$ máme $S + T = T$ a $S \cap T = S$, z čoho je zrejmé, že tvrdenie vety platí. Prípad $T \subseteq S$ je symetrický.

Zostáva teda prípad, že $S \not\subseteq T$ a $T \not\subseteq S$. Môžeme potom predpokladať, že $S \cap T$ má bázu $\vec{\gamma}_1, \ldots, \vec{\gamma}_r$. (V prípade, že $S \cap T = \{\vec{0}\}$, tak tento podpriestor nemá bázu – vtedy stačí zobrať $r = 0$ a vo zvyšku dôkazu môžeme postupovať úplne rovnako.) Tieto vektory patria do priestoru S a sú lineárne nezávislé, preto ich možno doplniť na bázu priestoru S, čiže $S = [\vec{\gamma}_1, \ldots, \vec{\gamma}_r, \vec{\alpha}_1, \ldots, \vec{\alpha}_s]$. Podobne môžeme v T zvoliť bázu $\vec{\gamma}_1, \ldots, \vec{\gamma}_r, \vec{\beta}_1, \ldots, \vec{\beta}_t$. Podľa vety 4.5.3 dostaneme

$$S + T = [\vec{\gamma}_1, \ldots, \vec{\gamma}_r, \vec{\alpha}_1, \ldots, \vec{\alpha}_s, \vec{\beta}_1, \ldots, \vec{\beta}_t].$$

Stačí, ak dokážeme, že tieto vektory sú lineárne nezávislé, lebo potom tvoria bázu v $S+T$ a máme $d(S + T) + d(S \cap T) = r + s + t + r = (r + s) + (r + t) = d(S) + d(T)$.

Nech $c_1\vec{\gamma}_1 + \cdots + c_r\vec{\gamma}_r + d_1\vec{\alpha}_1 + \cdots + d_s\vec{\alpha}_s + e_1\vec{\beta}_1 + \cdots + e_t\vec{\beta}_t = \vec{0}$. Potom $\vec{\delta} = c_1\vec{\gamma}_1 + \cdots + c_r\vec{\gamma}_r + d_1\vec{\alpha}_1 + \cdots + d_s\vec{\alpha}_s = -e_1\vec{\beta}_1 + \cdots + -e_t\vec{\beta}_t$ patrí do podpriestoru $S \cap T$. (Patrí do S, lebo je lineárnou kombináciou vektorov $\vec{\gamma}_1, \ldots, \vec{\gamma}_r, \vec{\alpha}_1, \ldots, \vec{\alpha}_s$. Do T patrí preto, že je lineárnou kombináciou vektorov $\vec{\beta}_1, \ldots, \vec{\beta}_t$.) Teda $\vec{\delta} = c'_1\vec{\gamma}_1 + \ldots + c'_2\vec{\gamma}_r$. Vďaka tomu, že

[5] Tento vzorec pripomína vzorec pre počet prvkov zjednotenia dvoch množín $|S \cup T| = |S| + |T| - |S \cap T|$.

vyjadrenie vektora pomocou prvkov bázy je jednoznačné, dostaneme, že $d_1 = \ldots = d_s = 0$ a $e_1 = \ldots = e_t = 0$. Potom máme $c_1\vec{\gamma}_1 + \cdots + c_r\vec{\gamma}_r = \vec{0}$ a (pretože $\vec{\gamma}_1, \ldots, \vec{\gamma}_r$ je báza) $c_1 = \ldots = c_r = 0$. □

Definícia 4.5.5. Nech S, T sú podpriestory vektorového priestoru V nad poľom F a nech $S \cap T = \{\vec{0}\}$. Potom podpriestor $S + T$ nazývame *direktný* (*priamy*) *súčet* podpriestorov S a T a označujeme ho $S \oplus T$.

{sucty:VTSUCTYEKVIV}

Veta 4.5.6. *Nech S, T, P sú podpriestory konečnorozmerného vektorového priestoru V nad poľom F. Tieto podmienky sú potom ekvivalentné:*

(i) $P = S \oplus T$

(ii) $P = S + T$ *a* $d(P) = d(S) + d(T)$

(iii) Ak $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ *je báza podpriestoru* S *a* $\vec{\beta}_1, \ldots, \vec{\beta}_m$ *je báza podpriestoru* T*, tak* $\vec{\alpha}_1, \ldots, \vec{\alpha}_n, \vec{\beta}_1, \ldots, \vec{\beta}_m$ *je báza podpriestoru* P*.*

(iv) $P = S + T$ *a každý vektor* $\vec{\gamma} \in P$ *sa dá jediným spôsobom vyjadriť v tvare* $\vec{\alpha} + \vec{\beta}$*, kde* $\vec{\alpha} \in S$ *a* $\vec{\beta} \in T$*. (T.j. ak* $\vec{\gamma} = \vec{\alpha}_1 + \vec{\beta}_1 = \vec{\alpha}_2 + \vec{\beta}_2$*, pričom* $\vec{\alpha}_1, \vec{\alpha}_2 \in S$ *a* $\vec{\beta}_1, \vec{\beta}_2 \in T$*, tak* $\vec{\alpha}_1 = \vec{\alpha}_2$ *a* $\vec{\beta}_1 = \vec{\beta}_2$*.)*

Dôkaz. $(i) \Rightarrow (ii)$ Podľa vety 4.5.4 dostaneme $d(S) + d(T) = d(S + T) + d(S \cap T) = d(P) + d(\{\vec{0}\}) = d(P)$.

$(ii) \Rightarrow (iii)$ Podľa vety 4.5.3 je $S + T = [\vec{\alpha}_1, \ldots, \vec{\alpha}_n, \vec{\beta}_1, \ldots, \vec{\beta}_m]$. Pretože $d(P) = n + m$ a našli sme $n + m$ vektorov, ktoré sú ho generujú, musia byť tieto vektory lineárne nezávislé a tvoria bázu.

$(iii) \Rightarrow (iv)$ Z vety 4.5.3 vyplýva, že $S+T = [\vec{\alpha}_1, \ldots, \vec{\alpha}_n, \vec{\beta}_1, \ldots, \vec{\beta}_m]$, teda $P = S+T$. Nech $\vec{\gamma} = \vec{\alpha}+\vec{\beta} = \vec{\alpha}'+\vec{\beta}'$. Potom $\vec{\alpha}-\vec{\alpha}'+\vec{\beta}-\vec{\beta}' = \vec{0}$. Ak vyjadríme vektory $\vec{\alpha}-\vec{\alpha}' \in S$ a $\vec{\beta}-\vec{\beta}' \in T$ pomocou báz týchto podpriestorov, čiže $\vec{\alpha}-\vec{\alpha}' = c_1\vec{\alpha}_1+\cdots+c_n\vec{\alpha}_n$ a $\vec{\beta}-\vec{\beta}' = d_1\vec{\beta}_1+\cdots+d_m\vec{\beta}_m$ dostaneme

$$c_1\vec{\alpha}_1 + \cdots + c_n\vec{\alpha}_n + d_1\vec{\beta}_1 + \cdots + d_m\vec{\beta}_m = \vec{0}.$$

Pretože vektory $\vec{\alpha}_1, \ldots, \vec{\alpha}_n, \vec{\beta}_1, \ldots, \vec{\beta}_m$ tvoria bázu v P, sú lineárne nezávislé a $c_1 = \ldots = c_n = d_1 = \ldots = d_m = 0$. Z toho vyplýva, že $\vec{\alpha} - \vec{\alpha}' = \vec{0}$ a $\vec{\beta} - \vec{\beta}' = \vec{0}$, čiže $\vec{\alpha} = \vec{\alpha}'$ a $\vec{\beta} = \vec{\beta}'$.

$(iv) \Rightarrow (i)$ Potrebujeme ukázať iba že $S \cap T = \{\vec{0}\}$. Ak $\vec{\gamma} \in S \cap T$, tak $\vec{\gamma}$ môžeme vyjadriť ako súčet vektora z S a vektora z T týmito dvoma spôsobmi: $\vec{\gamma} = \vec{\gamma} + \vec{0} = \vec{0} + \vec{\gamma}$. Z jednoznačnosti potom vyplýva, že $\vec{\gamma} = \vec{0}$. □

Cvičenia

{suctycvic:SUCTY}

Úloha 4.5.1. Zistite[6] $d(U)$, $d(V)$, $d(U + V)$, $d(U \cap V)$, bázu $U + V$ a bázu $U \cap V$
a) v $\mathbb{R}^2$ pre $U = [(2, 5)]$, $V = [(1, 3)]$
b) v $\mathbb{R}^3$ pre $U = [(1, 2, 3), (-1, 2, 3)]$, $V = [(2, 1, 4), (-2, 1, 4)]$
c) v $\mathbb{R}^4$ pre $U = [(1, 0, 1, 0), (1, 0, 0, 1)]$, $V = [(1, 1, 1, 0), (1, 0, 1, 1)]$
d) v $\mathbb{R}^4$ pre $U = [(1, 2, 3, 4), (1, 1, 1, 1), (4, 3, 2, 1)]$, $V = [(1, 1, 0, 0), (0, 0, 1, 1), (1, 0, 0, 0)]$.
[a)1,1,2,0; b)2,2,3,1; c)2,2,4,0; d)2,3,4,1]

Úloha 4.5.2. Nech $T = [(1, 3, 2), (2, 1, 3), (3, 4, 0)]$ je podpriestor $(\mathbb{Z}_5)^3$. Existuje podpriestor S taký, že $(\mathbb{Z}_5)^3 = T \oplus S$? Ak áno, nájdite ho! Je tento podpriestor jednoznačne určený?

[6] Túto úlohu budeme riešiť neskôr, keď sa (v časti 5.2) naučíme jednoduchý spôsob ako nájsť dimenziu a bázu daného podpriestoru $\mathbb{R}^n$. Zaradil som ju však sem, pretože súvisí s témou tejto podkapitoly.

Úloha 4.5.3. Nech $S \neq T$ sú dva podpriestory vektorového priestoru F^3 nad poľom F a $d(S) = 2$, $d(T) = 2$. Dokážte, že $d(S \cap T) \geq 1$.

Úloha 4.5.4. Ak máme zadané podpriestory

$$W_1 = \{(x, y, z) \in \mathbb{R}^3; x + y - z = 0\},$$
$$W_2 = \{(x, y, z) \in \mathbb{R}^3; 3x + y - 2z = 0\},$$
$$W_3 = \{(x, y, z) \in \mathbb{R}^3; x - 7y + 3z = 0\},$$

nájdite $\dim(W_1 \cap W_2 \cap W_3)$ a $\dim(W_1 + W_2)$, $\dim(W_1 + W_3)$, $\dim(W_2 + W_3)$.

{suctycvic:ULOSUCETN}

Úloha 4.5.5*. Nech $S_1, \ldots, S_n$ a P sú podpriestory vektorového priestoru V. Ukážte, že nasledujúce podmienky sú ekvivalentné:

(i) $P = S_1 + \cdots + S_n$ a pre každé $i = 1, \ldots, n$ platí $S_i \cap (S_1 + \cdots + S_{i-1} + S_{i+1} + \cdots + S_n) = \{\vec{0}\}$;

(ii) Každý vektor $\vec{\gamma} \in P$ sa dá práve jedným spôsobom vyjadriť v tvare $\vec{\gamma} = \vec{\alpha}_1 + \cdots + \vec{\alpha}_n$, kde $\vec{\alpha}_i \in S_i$ pre $i = 1, \ldots, n$.

Ktorúkoľvek z týchto ekvivalentných podmienok môžeme zobrať za definíciu priameho súčtu $S_1 \oplus \cdots \oplus S_n$. (Pri definícii lineárneho súčtu n podpriestorov nijaké komplikácie nenastanú a môžeme ho definovať jednoducho ako $S_1 + \cdots + S_n = \{\vec{\alpha}_1 + \cdots + \vec{\alpha}_n; \vec{\alpha}_i \in S_i \text{ pre } i = 1, \ldots, n\}$.)

Úloha 4.5.6. Dokážte, že ak $\vec{e_1}, \ldots, \vec{e_k}$ je báza vektorového priestoru V, tak $V = [\vec{e_1}] \oplus \ldots \oplus [\vec{e_k}]$. (Ako definíciu priameho súčtu viacerých podpriestorov môžete brať ktorúkoľvek z ekvivalentných podmienok z úlohy 4.5.5*.)

Kapitola 5

Lineárne zobrazenia a matice

5.1 Matice

Definícia 5.1.1. *Maticou* typu $m \times n$ nad poľom F nazývame ľubovoľnú tabuľku pozostávajúcu z prvkov poľa F, ktorá má m riadkov a n stĺpcov.

Matice zapisujeme v tvare

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

pričom a_{ij} označuje prvok v i-tom riadku a j-tom stĺpci.

Niekedy bude výhodné použiť stručnejší zápis $||a_{ij}||$, čím myslíme, že pre stručnosť niekedy len uvedieme predpis pre prvok i-teho riadku a j-teho stĺpca.

Príklad 5.1.2. $\left(\begin{smallmatrix} 2 & -1 & 3 \\ 4 & -2 & 5 \end{smallmatrix}\right)$ je matica typu 2×3 nad $\mathbb{R}$.

Definícia 5.1.3. Nech A, B sú matice typu $m \times n$ nad poľom F a $c \in F$.

(a) Súčet matíc $A = ||a_{ij}||$ a $B = ||b_{ij}||$ je matica $A + B = ||a_{ij} + b_{ij}||$.

(b) Matica $c \cdot A = ||ca_{ij}||$ sa nazýva *c-násobok* matice A.

(Teda sčitovanie matíc a násobenie matice skalárom definujeme po súradniciach.)

Všimnime si, že súčet matíc definujeme len pre matice rovnakého typu.

Príklad 5.1.4. Uvažujme matice typu 2×2 nad $\mathbb{R}$.
$\left(\begin{smallmatrix} 1 & -1 \\ 0 & 1 \end{smallmatrix}\right) + \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 2 & 0 \\ 0 & 2 \end{smallmatrix}\right)$
2. $\left(\begin{smallmatrix} 1 & -1 \\ 0 & 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 2 & -2 \\ 0 & 2 \end{smallmatrix}\right)$

Veta 5.1.5. *Matice typu $m \times n$ nad poľom F s takto definovaným sčitovaním a násobením skalármi tvoria vektorový priestor nad poľom F.*

Dôkaz. Úloha 5.1.1. (Vlastne si stačí uvedomiť, že je to to isté ako priestor F^{mn} – matice typu $m \times n$ tvoria len inak zapísané mn-tice prvkov z F, operácie sú definované tak, že korešpondujú s priestorom F^{mn}.) □

Vďaka tomu, že matice tvoria vektorový priestor nad F môžeme využívať všetky vlastnosti, ktoré poznáme z vektorových priestorov, ako napríklad identitu $c(A + B) = cA + cB$.

Budeme používať označenie $-A = ||-a_{ij}||$ pre *opačnú maticu* k matici A a $0 = ||0||$ pre *nulovú maticu.*

Definícia 5.1.6. Maticu typu $n \times n$ (teda takú, ktorá má rovnaký počet riadkov a stĺpcov) nazývame *štvorcová matica.*

Maticu

$$I = I_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix}$$

typu $n \times n$, ktorá má na diagonále jednotky a mimo diagonály nuly, nazývame *jednotková matica.*

Štvorcová matica, ktorá má mimo diagonály iba nuly (t.j. $a_{ij} = 0$ pre $i \neq j$) sa nazýva *diagonálna matica.* (Príkladom diagonálnej matice je jednotková matica.)

{mat:POZKRON}

Poznámka* **5.1.7.** Jednotkovú maticu by sme mohli definovať ako $I = ||\delta_{ij}||$, kde

$$\delta_{ij} = \begin{cases} 1 & \text{ak } i = j, \\ 0 & \text{ak } i \neq j. \end{cases}$$

Takto definovaný symbol sa v matematike často používa, nazýva sa *Kroneckerov symbol.*

Často budeme používať aj pojem transponovanej matice.

Definícia 5.1.8. *Transponovaná matica* k matici A typu $m \times n$ je matica A^T typu $n \times m$ určená ako

$$A^T = ||a_{ji}||.$$

Štvorcová matica A sa nazýva *symetrická,* ak $A = A^T$ a *antisymetrická,* ak $A = -A^T$.

Teda A^T je vlastne matica A prevrátená symetricky podľa hlavnej diagonály.

Môžeme si všimnúť, že platí $I^T = I$, $(A^T)^T = A$, $(A + B)^T = A^T + B^T$ a $(cA)^T = cA^T$ (úloha 5.1.3).

Príklad 5.1.9. Ak $A = \left(\begin{smallmatrix} 2 & -1 & 3 \\ 4 & -2 & 5 \end{smallmatrix}\right)$, tak $A^T = \left(\begin{smallmatrix} 2 & 4 \\ -1 & -2 \\ 3 & 5 \end{smallmatrix}\right)$

Matice $\left(\begin{smallmatrix} 1 & 2 \\ 2 & 3 \end{smallmatrix}\right)$ a $\left(\begin{smallmatrix} 1 & 2 & 1 \\ 2 & 1 & 0 \\ 1 & 0 & 1 \end{smallmatrix}\right)$ sú symetrické, matice $\left(\begin{smallmatrix} 0 & 2 \\ -2 & 0 \end{smallmatrix}\right)$ a $\left(\begin{smallmatrix} 0 & 2 & 1 \\ -2 & 0 & -1 \\ -1 & 1 & 0 \end{smallmatrix}\right)$ sú antisymetrické.

Cvičenia

{mat:ULOVPR}

Úloha 5.1.1. Overte, že matice typu $m \times n$ nad poľom F (spolu so sčitovaním matíc a násobením matice skalárom) tvoria vektorový priestor nad F.

Úloha 5.1.2. Dokážte, že diagonálne matice tvoria podpriestor vektorového priestoru všetkých matíc typu $n \times n$.

{matcvic:symaasym}

Úloha 5.1.3. Nech matice A a B sú rovnakého typu. Dokážte, že potom $(A+B)^T = A^T+B^T$ a $(A^T)^T = A$. Čomu sa rovná $(c_1A + c_2B)^T$?

{matcvic:SYMAASYM}

Úloha 5.1.4. Dokážte, že
a) množina všetkých symetrických matíc typu $n \times n$ a
b) množina všetkých antisymetrických matíc typu $n \times n$
tvoria podpriestory vektorového priestoru všetkých matíc typu $n \times n$. Je vektorový priestor matíc typu $n \times n$ direktný súčet týchto podpriestorov?

5.2 Riadková ekvivalencia a hodnosť matice

{rtm:SECTRTM}

Definícia 5.2.1. *Podpriestorom prislúchajúcim matici* A typu $m \times n$ nad poľom F nazývame podpriestor priestoru F^n generovaný riadkami matice A. Označujeme ho V_A.

Aby sme rozumeli predchádzajúcej definícii, uvedomme si, že každý riadok matice je vlastne n-tica prvkov z F, čiže ho môžeme chápať ako vektor z F^n.

Ak matica A má riadky $\vec{\alpha}_1, \ldots, \vec{\alpha}_m$, tak podpriestor prislúchajúci tejto matici je vlastne $V_A = [\vec{\alpha}_1, \ldots, \vec{\alpha}_m]$.

Príklad 5.2.2. Pozrime sa na pár príkladov nad poľom $\mathbb{R}$.

$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \quad V_A = [(1,0,1),(0,1,1)]$

$I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad V_I = [(1,0,0),(0,1,0),(0,0,1)] = \mathbb{R}^3$

Vidíme, že jednotkovej matici I zodpovedá štandardná báza priestoru $\mathbb{R}^3$, preto jej prislúcha celý priestor $\mathbb{R}^3$.

Zavedieme teraz úpravy, ktoré nám umožnia jednoduchšie popísať priestor prislúchajúci danej matici.

Definícia 5.2.3. *Elementárne riadkové operácie* na matici A nad poľom F sú:

1. výmena 2 riadkov matice,
2. vynásobenie niektorého riadku matice nenulovým prvkom c poľa F,
3. pripočítanie násobku niektorého riadku k inému riadku.

Hovoríme, že matice A a B sú *riadkovo ekvivalentné* ak maticu B možno z A dostať pomocou konečnej postupnosti elementárnych riadkových operácií. Ak matice A a B sú riadkovo ekvivalentné, zapisujeme to ako $A \sim B$.

{hodnost:PR1}

Príklad 5.2.4. Nasledujúce matice sme dostali z prvej pomocou elementárnych riadkových operácií. Sú to teda riadkovo ekvivalentné matice.

$$A = \begin{pmatrix} 1 & 2 & 5 \\ 2 & -2 & 1 \\ 3 & -2 & 3 \end{pmatrix} \overset{(1)}{\sim} \begin{pmatrix} 1 & 2 & 5 \\ 0 & -6 & -9 \\ 3 & -2 & 3 \end{pmatrix} \overset{(2)}{\sim} \begin{pmatrix} 1 & 2 & 5 \\ 0 & -6 & -9 \\ 0 & -8 & -12 \end{pmatrix} \overset{(3)}{\sim} \begin{pmatrix} 1 & 2 & 5 \\ 0 & 2 & 3 \\ 0 & 2 & 3 \end{pmatrix} \overset{(4)}{\sim} \begin{pmatrix} 1 & 2 & 5 \\ 0 & 2 & 3 \\ 0 & 0 & 0 \end{pmatrix} \overset{(5)}{\sim} \begin{pmatrix} 1 & 0 & 2 \\ 0 & 2 & 3 \\ 0 & 0 & 0 \end{pmatrix} \overset{(6)}{\sim} \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & \frac{3}{2} \\ 0 & 0 & 0 \end{pmatrix}$$

Elementárne riadkové operácie, ktoré sme použili sú:
(1) k 2.riadku sme pripočítali (-2)-násobok prvého (inak povedané, odčítali sme dvojnásobok),
(2) od 3.riadku sme odčítali 3-násobok prvého,
(3) 2.riadok sme vynásobili $-\frac{1}{3}$, 3.riadok sme vynásobili $-\frac{1}{4}$,
(4) od 2.riadku sme odpočítali tretí,
(5) od prvého riadku sme odpočítali druhý,
(6) druhý riadok sme vynásobili $\frac{1}{2}$ (čiže sme vlastne ho vydelili 2).

Poznámka 5.2.5. Podobným spôsobom sa dajú definovať aj elementárne stĺpcové operácie.

{hodnost:POZNEKV}

Poznámka 5.2.6. Je zrejmé, že ak $A \sim B$ a $B \sim C$, tak platí aj $A \sim C$. (Postupnosťou elementárnych riadkových operácií vieme z A dostať najprv B a potom z B dostať C.)

Okrem toho elementárne riadkové operácie možno obrátiť, preto ak $A \sim B$, tak platí aj $B \sim A$.

(Nie je ťažké si uvedomiť, že ak B dostaneme s A výmenou 2 riadkov, tak výmenou tých istých riadkov dostaneme z matice B pôvodnú maticu A. Takisto, ak použijeme vynásobenie niektorého riadku prvkom $c \neq 0$ poľa F, tak pôvodnú maticu dostaneme tak, že tento riadok vynásobíme c^{-1}. Ak sme v matici B získali k-ty riadok pripočítaním c-násobku j-teho riadku,

čiže k-ty riadok novej matice pomocou riadkov pôvodnej matice vieme vyjadriť ako $\vec{\alpha}_k + c\vec{\alpha}_j$, tak $(\vec{\alpha}_k + c\vec{\alpha}_j) - c\vec{\alpha}_j = \vec{\alpha}_k$, čiže pripočítaním $(-c)$-násobku j-teho riadku ku k-temu dostaneme z B pôvodnú maticu.)

{hodnost:VTVA

Veta 5.2.7. *Elementárne riadkové operácie nemenia podpriestor prislúchajúci danej matici.* (*Teda riadkovo ekvivalentným maticiam zodpovedá rovnaký podpriestor.*)

Dôkaz. Chceme ukázať, že ak na matici A, ktorej riadky sú $\vec{\alpha}_1, \ldots, \vec{\alpha}_m$, urobíme ktorúkoľvek z 3 elementárnych riadkových operácií, podpriestor prislúchajúci novej matici bude rovnaký ako podpriestor $V_A = [\vec{\alpha}_1, \ldots, \vec{\alpha}_m]$.

V prípade výmeny 2 riadkov je to jasné – ak vektory napíšeme v inom poradí, tak vygenerujú ten istý podpriestor.

Ďalšou elementárnou operáciou je vynásobenie niektorého riadku skalárom $c \neq 0$. Potom priestor prislúchajúci novej matici je $V_B = [\vec{\alpha}_1, \ldots, \vec{\alpha}_{i-1}, c\vec{\alpha}_i, \vec{\alpha}_{i+1}, \ldots, \vec{\alpha}_n]$. Pretože $c\vec{\alpha}_i \in V_A$, všetky vektory generujúce priestor V_B patria do V_A, preto tam patria aj všetky ich lineárne kombinácie, čo znamená $V_B \subseteq V_A$. Obrátenú inklúziu $V_A \subseteq V_B$ dostaneme rovnakým spôsobom: vektor $\vec{\alpha} = c^{-1}.(c\vec{\alpha})$ totiž patrí do V_B.

Zostáva nám posledná operácia – pripočítanie násobku niektorého riadku k inému riadku. Bez ujmy na všeobecnosti predpokladajme, že sme pripočítavali c-násobok druhého riadku k prvému (vektory môžeme ľubovoľne preusporiadať bez toho, aby sme zmenili podpriestor, ktorý generujú). Chceme teda ukázať, že $V_A = [\vec{\alpha}_1, \ldots, \vec{\alpha}_m] = [\vec{\alpha}_1 + c\vec{\alpha}_2, \vec{\alpha}_2, \ldots, \vec{\alpha}_m] = V_B$. Každý z vektorov generujúcich V_B je lineárna kombinácia vektorov $\vec{\alpha}_1, \ldots, \vec{\alpha}_m$, preto platí $V_B \subseteq V_A = [\vec{\alpha}_1, \ldots, \vec{\alpha}_m]$. Obrátene, vektor $\vec{\alpha}_1 = (\vec{\alpha}_1 + c\vec{\alpha}_2) - c\vec{\alpha}_2$ je lineárna kombinácia vektorov, ktoré generujú V_B, preto platí aj $V_A \subseteq V_B$. □

Definícia 5.2.8. Matica A je *redukovaná trojuholníková matica,* ak:

(i) Vedúci (=prvý nenulový) prvok každého riadku matice je 1.

(ii) Každý stĺpec obsahujúci vedúci prvok niektorého riadku má prvky v ostatných riadkoch nulové.

(iii) Nulové riadky ležia pod nenulovými riadkami. (Presnejšie povedané: Akýkoľvek nulový riadok musí byť nižšie ako akýkoľvek nenulový riadok.)

(iv) Vedúci prvok ľubovoľného nenulového riadku je napravo od vedúcich prvkov všetkých nenulových riadkov nad ním a naľavo od vedúcich prvkov riadkov pod ním (t.j. vedúce riadky sú usporiadané zľava doprava).

Napríklad matica $\begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & \frac{3}{2} \\ 0 & 0 & 0 \end{pmatrix}$, ktorú sme dostali v príklade 5.2.4 je redukovaná trojuholníková matica.

Ľubovoľná redukovaná trojuholníková matica vyzerá zhruba takto:

$$\begin{pmatrix} 0 & \ldots & 0 & \boxed{1} & * & 0 & * & 0 & * & \ldots & * \\ 0 & \ldots & 0 & 0 & 0 & \boxed{1} & * & 0 & * & \ldots & * \\ 0 & \ldots & 0 & 0 & 0 & 0 & 0 & \boxed{1} & * & \ldots & * \\ 0 & \ldots & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots & 0 \\ 0 & \ldots & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots & 0 \end{pmatrix}$$

V predchádzajúcej schéme $*$ označuje miesta, kde môže byť ľubovoľný prvok (nulový alebo nenulový). Vidíme, že vedúce jednotky (vyznačené štvorčekom) idú zľava doprava

{hodnost:VTEKVRTM}

Veta 5.2.9. *Každá matica nad poľom F je riadkovo ekvivalentná s nejakou redukovanou trojuholníkovou maticou.*

Dôkaz. Nech A je matica typu $m \times n$.

Tvrdenie vety dokážeme indukciou vzhľadom na m – počet riadkov matice A.

Ak $m = 1$, tak máme jediný riadok. V prípade, že je tento riadok nulový, matica už je v redukovanom trojuholníkovom tvare. Ak nie, tak tento riadok má tvar $(0, \ldots, 0, a_{1s}, a_{1,s+1}, \ldots, a_{1n})$, kde a_{1s} je prvý nenulový prvok v danom riadku. Vynásobením a_{1s}^{-1} dostaneme maticu $(0, \ldots, 0, 1, a_{1s}^{-1}a_{1,s+1}, \ldots, a_{1s}^{-1}a_{1n})$, čiže vedúci prvok jej jediného riadku je 1 a táto matica je v redukovanom trojuholníkovom tvare.

Indukčný krok: predpokladáme, že tvrdenie vety platí pre každú maticu, ktorá má m riadkov, chceme dokázať, že platí aj pre ľubovoľnú maticu A typu $(m+1) \times n$.

Ak A je nulová matica, tak je v redukovanom trojuholníkovom tvare. V opačnom prípade, nech s je prvý stĺpec, ktorý je nenulový. Teda tento stĺpec obsahuje aspoň jeden nenulový prvok.

$$\begin{pmatrix} 0 & \ldots & 0 & a_{1s} & * & * & * & a_{1n} \\ 0 & \ldots & 0 & * & * & * & * & * \\ 0 & \ldots & 0 & a_{ks} \neq 0 & * & * & * & a_{ks} \\ 0 & \ldots & 0 & * & * & * & * & * \\ 0 & \ldots & 0 & a_{m+1,s} & * & * & * & a_{ns} \end{pmatrix}$$

Výmenou riadkov vieme dostať maticu, ktorá má v s-tom stĺpci nenulový prvok už v prvom riadku. (Ak to spĺňa už pôvodná matica, nie je potrebné vymieňať riadky.)

$$\begin{pmatrix} 0 & \ldots & 0 & b_{1s} \neq 0 & * & * & * & * \\ 0 & \ldots & 0 & b_{2s} & * & * & * & * \\ & & & \vdots & & & & \\ 0 & \ldots & 0 & b_{ks} & * & * & * & * \\ & & & \vdots & & & & \\ 0 & \ldots & 0 & b_{m+1,s} & * & * & * & * \end{pmatrix}$$

Aby sme v prvom riadku dostali vedúcu jednotku, vynásobíme ho b_{1s}^{-1}.

$$\begin{pmatrix} 0 & \ldots & 0 & 1 & * & * & * & * \\ 0 & \ldots & 0 & b_{2s} & * & * & * & * \\ & & & \vdots & & & & \\ 0 & \ldots & 0 & b_{ks} & * & * & * & * \\ & & & \vdots & & & & \\ 0 & \ldots & 0 & b_{m+1,s} & * & * & * & * \end{pmatrix}$$

Teraz vieme vynulovať všetky ostatné prvky v s-tom stĺpci – na to stačí od k-teho riadku (pre $k = 2, 3, \ldots, m+1$) odpočítať b_{ks}-násobok prvého riadku.

$$\begin{pmatrix} 0 & \ldots & 0 & 1 & * & * & * & * \\ 0 & \ldots & 0 & 0 & c_{2,s+1} & \ldots & \ldots & c_{2,n} \\ 0 & \ldots & 0 & \vdots & * & * & * & * \\ 0 & \ldots & 0 & 0 & c_{k,s+1} & \ldots & \ldots & c_{k,n} \\ 0 & \ldots & 0 & \vdots & * & * & * & * \\ 0 & \ldots & 0 & 0 & c_{m+1,s+1} & \ldots & \ldots & c_{m,n} \end{pmatrix}$$

Teraz nastala správna chvíľa použiť indukčný predpoklad – podľa neho vieme podmaticu pozostávajúcu zo všetkých riadkov predchádzajúcej matice okrem prvého upraviť na redukovanú trojuholníkovú maticu. Takto dostaneme maticu tvaru, ktorý schematicky znázorníme takto:

$$\begin{pmatrix} 0 & \dots & 0 & 1 & * & * & * & * \\ 0 & \dots & 0 & 0 & \boxed{1} & 0 & * & 0 \\ 0 & \dots & 0 & 0 & 0 & \boxed{1} & * & 0 \\ 0 & \dots & 0 & 0 & 0 & 0 & 0 & \boxed{1} \\ 0 & \dots & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \dots & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Inak povedané, podmatica pozostávajúca z riadkov 2 až $m+1$ spĺňa definíciu redukovanej trojuholníkovej matice. Jediná podmienka, z definície redukovanej trojuholníkovej matice, ktorá môže byť narušená v celej matici, je, že v niektorom zo stĺpcov obsahujúcich vedúcu jednotku môže táto matica obsahovať nenulový prvok. Pripočítaním vhodného násobku týchto riadkov vieme aj tieto prvky matice vynulovať. (Presnejšie to môžeme zapísať takto: označme prvky prvého riadku v $(s+1)$-vom až n-tom stĺpci. Nech stĺpce obsahujúce vedúce jednotky sú $i_2, i_3, \dots, i_k$. Potom od prvého riadku odpočítame c_{1,i_2}-násobok 2.riadku, c_{1,i_3}-násobok 3.riadku, atď.)

$$\begin{pmatrix} 0 & \dots & 0 & 1 & 0 & 0 & * & 0 \\ 0 & \dots & 0 & 0 & \boxed{1} & 0 & * & 0 \\ 0 & \dots & 0 & 0 & 0 & \boxed{1} & * & 0 \\ 0 & \dots & 0 & 0 & 0 & 0 & 0 & \boxed{1} \\ 0 & \dots & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \dots & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Z naznačených úprav je vidno, že výsledná matica je skutočne redukovaná trojuholníková matica. □

Predchádzajúci dôkaz vlastne súčasne popisuje aj algoritmus, ako môžeme upraviť ľubovoľnú maticu na redukovaný trojuholníkový tvar.

Ako príklad úpravy na redukovanú trojuholníkovú maticu nám môže opäť poslúžiť príklad 5.2.4.

Zatiaľ sme si povedali, čo sú redukované trojuholníkové matice a vysvetlili sme si postup, akým môžeme z ľubovoľnej matice pomocou elementárnych riadkových úprav dostať redukovanú trojuholníkovú maticu. Teraz by sme chceli ukázať, prečo sú redukované trojuholníkové matice užitočné.

{hodnost:VTRTMLN}

Veta 5.2.10. *Nenulové riadky redukovanej trojuholníkovej matice sú lineárne nezávislé.*

Najprv ilustrujme túto vetu na príklade. Opäť použijeme redukovanú trojuholníkovú maticu z príkladu 5.2.4.

Príklad 5.2.11. Riadky redukovanej trojuholníkovej matice z príkladu 5.2.4 sú $\vec{\alpha} = (1,0,2)$ a $\vec{\beta} = (0,1,\frac{3}{2})$. Rovnosť $c\vec{\alpha} + d\vec{\beta} = \vec{0}$ znamená, že

$$c(1,0,2) + d(0,1,\frac{3}{2}) = (c, d, 2c + \frac{3}{2}d) = (0,0,0),$$

z čoho dostaneme (porovnaním prvých 2 súradníc) $c = 0$ a $d = 0$. Platí teda implikácia $c\vec{\alpha} + d\vec{\beta} = \vec{0} \Rightarrow c = d = 0$, čiže vektory $\vec{\alpha}$ a $\vec{\beta}$ sú lineárne nezávislé.

Videli sme, že dôležité boli tie súradnice, na ktorých sa nachádzajú vedúce jednotky. V nasledujúcom dôkaze postupujeme takmer identicky ako v príklade, ktorý sme práve uviedli.

Dôkaz. Nech $\vec{\alpha}_1, \dots, \vec{\alpha}_k$ sú nenulové riadky redukovanej trojuholníkovej matice A. Nech $i_1, \dots, i_k$ sú stĺpce s vedúcimi jednotkami.

Vektor $c_1\vec{\alpha}_1 + \cdots + c_k\vec{\alpha}_k$ má na mieste i_j prvok c_j (pre $j = 1, 2, \ldots, k$), takže rovnosť $c_1\vec{\alpha}_1 + \cdots + c_k\vec{\alpha}_k = \vec{0}$ implikuje $c_j = 0$ (pretože nulový vektor má na tomto mieste nulu). Zistili sme, že platí $c_1\vec{\alpha}_1 + \cdots + c_k\vec{\alpha}_k = \vec{0} \Rightarrow c_1 = c_2 = \ldots = c_k = 0$, čiže vektory $\vec{\alpha}_1, \ldots, \vec{\alpha}_k$ sú skutočne lineárne nezávislé. □

Definícia 5.2.12. *Hodnosť matice* A je dimenzia podpriestoru V_A prislúchajúceho tejto matici. Označujeme ju $h(A)$.

Z tejto definície máme rovnosť $h(A) = d(V_A)$ a pretože elementárne riadkové operácie nemenia priestor prislúchajúci danej matici (veta 5.2.7), nemenia ani hodnosť matice.

Z vety 5.2.10 vyplýva, že hodnosť redukovanej trojuholníkovej matice je počet jej nenulových riadkov. Teda hodnosť matice môžeme rátať pomocou úpravy na redukovanú trojuholníkovú maticu. Pre maticu z príkladu 5.2.4 dostaneme $h(A) = 2$.

Okrem toho, že redukované trojuholníkové matice sú užitočné na výpočet hodnosti (a tým aj výpočet dimenzie podpriestoru generovaného nejakými zadanými vektormi), dajú sa využiť aj na to, aby sme zistil, či nejaký vektor patrí do podpriestoru V_A.

{hodnost:PR2}

Príklad 5.2.13. V príklade 5.2.4 sme ukázali, že matice $A = \begin{pmatrix} 1 & 2 & 5 \\ 2 & -2 & 1 \\ 3 & -2 & 3 \end{pmatrix}$ a $B = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & \frac{3}{2} \\ 0 & 0 & 0 \end{pmatrix}$ sú riadkovo ekvivalentné čo znamená, že $V_A = V_B = [(1, 0, 2), (0, 1, \frac{3}{2})]$.

Pokúsme sa zistiť, či vektor $\vec{\alpha} = (1, 4, 4)$ patrí do V_A. Ak to platí, musí byť tento vektor lineárnou kombináciou vektorov $(1, 0, 2)$ a $(0, 1, \frac{3}{2})$. Dostávame teda rovnosť

$$(1, 4, 4) = c_1(1, 0, 2) + c_2(0, 1, \frac{3}{2}) = (c_1, c_2, 2c_1 + \frac{3}{2}c_2).$$

Porovnaním prvých dvoch súradníc dostaneme $c_1 = 1$ a $c_2 = 4$. Vektor na pravej strane poslednej rovnice má potom ale tretiu súradnicu rovnú $2.1 + \frac{3}{2}.4 = 8 \neq 4$, čiže vektor $\vec{\alpha}$ nepatrí do V_A.

Postup z predchádzajúceho príkladu sa dá použiť na dôkaz tohoto tvrdenia:

{hodnost:LMAB}

Lema 5.2.14. *Nech A je redukovaná trojuholníková matica typu $m \times n$ nad poľom F. Označme jej nenulové riadky $\vec{\alpha}_1, \ldots, \vec{\alpha}_k$ a ako $i_1, \ldots, i_k$ označme čísla stĺpcov, v ktorých sú vedúce jednotky. Potom $\vec{\alpha} = (c_1, \ldots, c_n) \in V_A$ práve vtedy, keď $\vec{\alpha} = c_{i_1}\vec{\alpha}_1 + c_{i_2}\vec{\alpha}_2 + \ldots + c_{i_k}\vec{\alpha}_k$.*

Dôkaz. Ak $\vec{\alpha} \in V_A$, tak $\vec{\alpha}$ je lineárnou kombináciou vektorov $\vec{\alpha}_1, \ldots, \vec{\alpha}_k$, čiže $\vec{\alpha} = d_1\vec{\alpha}_1 + \cdots + d_k\vec{\alpha}_k$

Pozrime sa, aká je i_j-ta súradnica vektoru $\vec{\alpha} = d_1\vec{\alpha}_1 + \cdots + d_k\vec{\alpha}_k$, pre ľubovoľné $j = 1, 2, \ldots, k$. Na tejto zložke majú vektory $\vec{\alpha}_1, \ldots, \vec{\alpha}_k$ nulu s výnimkou vektora α_j, ktorý tam má 1. Preto na i_j-tej súradnici vektora $d_1\vec{\alpha}_1 + \cdots + d_k\vec{\alpha}_k$ je d_j, dostávame rovnosť $d_j = c_{i_j}$. Zistili sme, že koeficienty lineárnej kombinácie sú skutočne rovné $c_{i_1}, c_{i_2}, \ldots, c_{i_k}$.

Obrátená implikácia je zrejmá. □

{hodnost:VTRTMVAVB}

Veta 5.2.15. *Ak A a B sú redukované trojuholníkové matice rovnakého typu $m \times n$ nad poľom F a $V_A = V_B$, tak $A = B$.*

Dôkaz. Aby sme ukázali, že 2 redukované trojuholníkové matice sa rovnajú, stačí dokázať, že vedúce jednotky sú v rovnakých stĺpcoch a v ostatných stĺpcoch obsahujú matice rovnaké prvky.

Pretože $h(A) = h(B)$, tieto matice majú rovnaký počet nenulových riadkov. Označme ho k. Nenulové riadky matice A označme $\vec{\alpha}_1, \ldots, \vec{\alpha}_k$ a $i_1 < i_2 < \ldots < i_k$ nech sú čísla stĺpcov, ktoré obsahujú vedúce jednotky. Podobne nenulové riadky matice B označme $\vec{\beta}_1, \ldots, \vec{\beta}_k$ a nech vedúce jednotky tejto matice sú v stĺpcoch $j_1 < j_2 < \ldots < j_k$.

Postupujme sporom. Predpokladajme, že by vedúce jednotky neboli v rovnakých stĺpcoch. Nech t je prvý index, kde $i_t \neq j_t$. Bez ujmy na všeobecnosti môžeme predpokladať $i_t < j_t$. Vektor $\vec{\alpha}_t$ má na miestach $i_1, \ldots, i_{t-1}$ nuly. Preto

$$\vec{\alpha}_t = 0\vec{\beta}_1 + \ldots + 0\vec{\beta}_{t-1} + c_t\vec{\beta}_t + c_{t+1}\vec{\beta}_{t+1} + \ldots + c_k\vec{\beta}_k.$$

Vektor na pravej strane tejto rovnosti má na prvých $j_t - 1$ miestach 0, čiže ju má aj na mieste i_t, čo je spor (keďže sa má rovnať vektoru $\vec{\alpha}_t$, ktorý tam má prvok 1).

Zistili sme, že predpoklad $i_t \neq j_t$ vedie k sporu, preto $i_t = j_t$.

Teraz stačí ukázať, že pre všetky $t = 1, 2, \ldots, k$ platí $\vec{\alpha}_t = \vec{\beta}_t$ – to znamená, že aj ostatné prvky matíc A a B sa rovnajú.

Vektor $\vec{\alpha}_t$ má 0 na miestach $i_1, \ldots, i_{t-1}$ aj $i_{t+1}, \ldots, i_k$ a na mieste i_t má jednotku. Podľa lemy 5.2.14 teda $\vec{\alpha}_t = 0\vec{\beta}_1 + \ldots + 0\vec{\beta}_{t-1} + 1\vec{\beta}_t + 0\vec{\beta}_{t+1} + 0\vec{\beta}_k$, čiže $\vec{\alpha}_t = \vec{\beta}_t$. □

Na získanie lepšieho porozumenia predchádzajúceho dôkazu je možno dobre si ukázať jeho najdôležitejší krok na konkrétnom príklade.

Príklad 5.2.16. Majme redukovanú trojuholníkovú maticu

$$A = \left(\begin{smallmatrix} 1 & 1 & 0 & 2 & 1 \\ 0 & 0 & 1 & 1 & 2 \end{smallmatrix}\right)$$

ktorá má vedúce jednotky v prvom a treťom stĺpci.

Predpokladajme, že by existovala redukovaná trojuholníková matica taká, že $V_A = V_B$ ale táto matica by mala vedúce jednotky v prvom a štvrtom stĺpci, čiže by mala tvar

$$B = \left(\begin{smallmatrix} 1 & b_{12} & b_{13} & 0 & b_{15} \\ 0 & 0 & 0 & 1 & b_{25} \end{smallmatrix}\right)$$

To by znamenalo, že $(0, 0, 1, 1, 2) \in V_B$ a podľa lemy 5.2.14 z toho dostaneme $(0, 0, 1, 1, 2) = 0.(1, b_{12}, b_{13}, 0, b_{15}) + 1.(0, 0, 0, 1, b_{25})$, čiže $(0, 0, 1, 1, 2) = (0, 0, 0, 1, b_{25})$; ale tieto vektory sa líšia na tretej súradnici, teda uvedená rovnosť nemôže platiť.

Iný spôsob, ako môžeme ukázať, že $V_A \neq V_B$: Ak by sa tieto podpriestory rovnali, tak by platilo $(0, 0, 0, 1, b_{25}) \in V_A = [(1, 1, 0, 2, 1), (0, 0, 1, 1, 2)]$. Z lemy 5.2.14 potom dostaneme $(0, 0, 0, 1, b_{25}) = 0.(1, 1, 0, 2, 1) + 0.(0, 0, 1, 1, 2) = (0, 0, 0, 0, 0)$. Dostali sme teda rovnosť nulového a nenulového vektoru, čo je spor.

Predchádzajúce vety môžeme zhrnúť nasledovne.

{hodnost:DOSVAVB}

Dôsledok 5.2.17. *Nech A a B sú matice typu $m \times n$ nad poľom F. Nasledovné podmienky sú ekvivalentné:*

{hodnost:it1} (i) *A a B sú riadkovo ekvivalentné,*

{hodnost:it2} (ii) *$V_A = V_B$,*

{hodnost:it3} (iii) *A a B sú riadkovo ekvivalentné s tou istou redukovanou trojuholníkovou maticou.*

Dôkaz. Veta 5.2.7 vlastne znamená implikáciu (i) ⇒ (ii).

Podľa vety 5.2.9 je A riadkovo ekvivalentná s nejakou redukovanou trojuholníkovou maticou A' a B je ekvivalentná s redukovanou trojuholníkovou maticou B'. Pretože $V_{B'} = V_B = V_A = V_{A'}$, podľa vety 5.2.15 $A' = B'$. Tým je dokázaná implikácia (ii) ⇒ (iii).

Podľa poznámky 5.2.6, ak $A \sim T$ a $B \sim T$ (T je redukovaná trojuholníková matica), tak aj $A \sim B$. Teda platí aj implikácia (iii) ⇒ (i). □

Poznámka 5.2.18. Postup z príkladu 5.2.13 a lemy 5.2.14 môžeme použiť na akúsi „polovičnú skúšky správnosti“ pri počítaní redukovanej trojuholníkovej matice. Pretože podobný postup sa dá použiť aj v iných situáciach, vysvetlime si ho trochu podrobnejšie.

Z dôsledku 5.2.17 vieme, že ak matica A je podobná redukovanej trojuholníkovej matici B, tak $V_A = V_B$. Túto rovnosť síce nevieme priamo overiť, vieme však ľahko zistiť, či $V_A \subseteq V_B$ (tak, že postupne overíme, či jednotlivé riadky matice A patria do V_B; použijeme na to rovnaký postup ako v príklade 5.2.13).

Ak nám takáto „poloskúška“ nevyjde vieme dokonca pomerne jednoducho nájsť poslednú úpravu od konca, v ktorej sme spravili chybu. Skúsme urobiť v úprave nejakej matice na redukovaný trojuholníkový tvar náročky chybu a ilustrovať si, ako ju vieme nájsť.

$\begin{pmatrix} 1 & -2 & -2 & 2 \\ 2 & 0 & -1 & -1 \\ 3 & 0 & -4 & -4 \end{pmatrix} \overset{(1)}{\sim} \begin{pmatrix} 1 & -2 & -2 & 2 \\ 0 & 4 & 3 & -5 \\ 0 & 0 & -\frac{5}{2} & -\frac{5}{2} \end{pmatrix} \overset{(2)}{\sim} \begin{pmatrix} 1 & -2 & -2 & 2 \\ 0 & 4 & 3 & -5 \\ 0 & 0 & 1 & 1 \end{pmatrix} \overset{(3)}{\sim} \begin{pmatrix} 1 & -2 & -2 & 2 \\ 0 & 4 & 0 & -2 \\ 0 & 0 & 1 & 1 \end{pmatrix} \overset{(4)}{\sim} \begin{pmatrix} 1 & -2 & -2 & 2 \\ 0 & 1 & 0 & -\frac{1}{2} \\ 0 & 0 & 1 & 1 \end{pmatrix} \overset{(5)}{\sim}$
$\begin{pmatrix} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & -\frac{1}{2} \\ 0 & 0 & 1 & 1 \end{pmatrix}$

Keď urobíme skúšku pre prvú maticu, nevyjde. (Konkrétne pre druhý riadok dostaneme $2(1,0,0,3) - (0,0,1,1) = (2,0,-1,5)$.)

Aby sme našli chybu, robíme skúšku pre matice, ktoré sme dostali ako medzivýsledky, až kým nenarazíme na situáciu, že pred niektorou úpravou skúška nesedí a po nej už áno. To znamená, že v tejto úprave sa zmenil podpriestor prislúchajúci matici, a teda táto úprava nemôže byť správna.

Napríklad pre maticu po úprave (2) skúška nesedí ($4.(0,1,0,-\frac{1}{2})+3.(0,0,1,1) = (0,4,3,1)$). Musíme teda chybu hľadať napravo od tejto matice.

Vyskúšame maticu po úprave (4) – skúška vyjde. Vyskúšame maticu po úprave (3) – skúška opäť vyjde. Keďže pre maticu pred úpravou (3) nám skúška vyšla, ale po nej nie, museli sme spraviť v tejto úprave chybu.

Samozrejme, môže sa stať, že urobíme chybu takého typu, ktorú takáto poloskúška neodhalí.

Správny postup je

$\begin{pmatrix} 1 & -2 & -2 & 2 \\ 2 & 0 & -1 & -1 \\ 3 & 0 & -4 & -4 \end{pmatrix} \overset{(1)}{\sim} \begin{pmatrix} 1 & -2 & -2 & 2 \\ 0 & 4 & 3 & -5 \\ 0 & 0 & -\frac{5}{2} & -\frac{5}{2} \end{pmatrix} \overset{(2)}{\sim} \begin{pmatrix} 1 & -2 & -2 & 2 \\ 0 & 4 & 3 & -5 \\ 0 & 0 & 1 & 1 \end{pmatrix} \overset{(3)}{\sim} \begin{pmatrix} 1 & -2 & -2 & 2 \\ 0 & 4 & 0 & -8 \\ 0 & 0 & 1 & 1 \end{pmatrix} \overset{(4)}{\sim} \begin{pmatrix} 1 & -2 & -2 & 2 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 1 \end{pmatrix} \overset{(5)}{\sim}$
$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 1 \end{pmatrix}$

(1) `3.r-=(3/2)*2.r; 2.r-=2*1.r` (Týmto zápisom sa myslí to, že od tretieho riadku sa odpočíta (3/2)-násobok druhého a od druhého dvojnásobok prvého) (2) `3.r*=-2/5` (3) `2.r-=3*3.r` (4) `2.r*=1/4` (5) `1.r+=2*2.r+2*3.r`

Poznámka 5.2.19. Ešte na tomto mieste spomeniem jednu často sa vyskytujúcu chybu.

Pomerne často sa študenti robia pri úpravách matíc niečo také, že namiesto pripočítania c-násobku niektoré riadku k inému urobia c-násobok niektorého riadku a pripočítajú iný riadok. (Alebo dokonca násobok iného riadku.) Napríklad sa vyskytuje

$$\begin{pmatrix} 1 & 2 & -1 \\ 1 & -2 & -2 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & -1 \\ 0 & 4 & 1 \end{pmatrix}$$

pričom použitá úprava je `2r:=1r-2r`, t.j. druhý riadok je nahradený rozdielom prvý riadok mínus druhý riadok.

Takáto úprava by bola (v podstate) v poriadku, pretože sme sa od matice $\begin{pmatrix} \vec{\alpha}_1 \\ \vec{\alpha}_2 \end{pmatrix}$ dostali k matici $\begin{pmatrix} \vec{\alpha}_1 \\ \vec{\alpha}_1 - \vec{\alpha}_2 \end{pmatrix}$, ktorá je s pôvodnou maticou riadkovo ekvivalentná. Vieme ju totiž dostať tak, že postupne robíme tieto úpravy: vynásobíme druhý riadok číslom (-1) a potom

pripočítame prvý riadok (1-násobok prvého riadku).

$$\begin{pmatrix}1 & 2 & -1\\ 1 & -2 & -2\end{pmatrix} \sim \begin{pmatrix}1 & 2 & -1\\ -1 & 2 & 2\end{pmatrix} \sim \begin{pmatrix}1 & 2 & -1\\ 0 & 4 & 1\end{pmatrix}$$

Teda sme vlastne iba naraz zapísali viacero úprav naraz, čo sme robili už aj predtým.

Napriek tomu, by som považoval za rozumné sa takémuto typu úprav vyhnúť. (Alebo si aspoň vždy poriadne premyslieť, aké úpravy ste urobili, ak ich takto skombinujete viacero naraz.) Sú na to dva dôvody.

Môže sa potom stať, že spravíte takúto chybu: Urobíte úpravu, že prvý riadok nahradíte rozdielom druhý riadok mínus prvý riadok (čo je v poriadku – len ste spojili viac úprav do jednej) a druhý riadok upravíte podobne ako v predošlom príklade. T.j. spravili sme takéto veci: `1r:=2r-1r`, `2r:=1r-2r`. Pozrime sa, čo dostaneme takouto úpravou (a čo vyjde ak pokračujeme ďalej).

$$\begin{pmatrix}1 & 2 & -1\\ 1 & -2 & -2\end{pmatrix} \sim \begin{pmatrix}0 & -4 & -1\\ 0 & 4 & 1\end{pmatrix} \sim \begin{pmatrix}0 & -4 & -1\\ 0 & 0 & 0\end{pmatrix}$$

Takto sme dostali z matice hodnosti 2 maticu hodnosti 1, čiže je tam zjavne niekde chyba. (A určite ľahko prídete na to kde.)

Predošlému problému by sme sa vedeli vyhnúť, ak by sme neurobili viac takýchto „zlých“ úprav naraz. Aj tak sa mi zdá rozumnejšie nezvykať si príliš na používanie úprav takéhoto typu. Dôvod je ten, že aj keď by to bolo v poriadku teraz (kým nás zaujíma len riadková ekvivalencia, hodnosť, podpriestore prislúchajúci matici a podobné veci), mohlo by nám to narobiť problémy pri počítaní determinantov. Neskôr – v časti 6.3.2 – sa naučíme, ako sa dajú používať elementárne riadkové operácie na počítanie determinantov. Budeme využívať fakt, že výmena riadkov mení determinant na opačný, pripočítanie násobku niektorého riadku k inému determinant nezmení, a vynásobenie konštantou c zmení determinant c-krát. Ak by sme použili úpravu typu `2r:=1r-2r`, tak veľmi ľahko prehliadneme, že táto úprava je zložená z viacerých, kde jedna z nich je vynásobenie matice číslom -1 (a teda aj determinant by sa mal zmeniť na (-1)-násobok). Veľmi sa to totiž podobá na typ operácie „pripočítanie násobku riadku k inému“ (ako `2r:=2r-1r`), ktoré determinant nemení.

Stručne zhrnuté: Ak používate viacero úprav naraz, treba si vždy poriadne uvedomiť, aké úpravy ste použili. A možno je lepšie nepoužívať naraz viac úprav rôznych typov. (Ak pripočítam násobok prvého riadku k druhému i k tretiemu v tom istom kroku, veľmi sa nemám kde pomýliť. Ak pripočítam k druhému riadku prvý i tretí súčasne, tiež sa nie je veľmi kde pomýliť. Z príkladov, ktoré som uviedol, by mohlo byť vidieť, že ak kombinujem viacero úprav rôznych typov, alebo počítam s už upraveným riadkom, dá sa vcelku ľahko spraviť chyba.)

Cvičenia V nasledujúcich úlohách, ak nie je uvedené inak, uvažujeme matice nad poľom $\mathbb{R}$.

`{rtmcvic:UL04B}`

Úloha 5.2.1. Nájdite redukované trojuholníkové matice riadkovo ekvivalentné s nasledujúcimi maticami a) nad poľom $\mathbb{R}$ b) nad poľom $\mathbb{Z}_5$

$\begin{pmatrix}2&3&1\\1&2&1\\3&4&1\end{pmatrix}$ $\begin{pmatrix}1&1&3\\3&2&2\\0&4&3\\4&1&1\end{pmatrix}$ $\begin{pmatrix}2&3&4&1\\3&2&3&2\\1&4&0&0\\4&1&2&3\end{pmatrix}$

Úloha 5.2.2. Ak sa to dá, doplňte dané vektory na bázu vektorového priestoru $(\mathbb{Z}_5)^4$:
a) $(1,2,0,0)$, $(3,4,0,1)$
b) $(1,2,3,4)$, $(1,1,1,1)$, $(3,2,1,0)$
c) $(2,3,4,1)$, $(3,2,4,1)$, $(0,2,3,2)$
d) $(1,3,1,4)$, $(3,0,4,3)$, $(2,3,1,1)$

Úloha 5.2.3. Ak sa to dá, doplňte dané vektory na bázu priestoru
$S = [(1,0,0,0,1),(0,1,0,0,2),(0,0,1,0,1),(0,0,0,1,3)] \subseteq (\mathbb{Z}_5)^5$.
a) $\vec{\alpha}_1 = (1,2,1,1,4)$, $\vec{\alpha}_2 = (2,4,2,1,0)$
b) $\vec{\alpha}_1 = (1,1,2,4,2)$, $\vec{\alpha}_2 = (3,3,2,2,0)$
c) $\vec{\alpha}_1 = (1,3,2,0,4)$, $\vec{\alpha}_2 = (2,1,4,0,3)$
d) $\vec{\alpha}_1 = (1,2,3,3,1)$, $\vec{\alpha}_2 = (2,3,0,0,3)$

Úloha 5.2.4. Zistite, či nasledujúce matice tvoria bázu vektorového priestoru všetkých matíc typu 2×2 nad poľom $\mathbb{R}$:
a) $\left(\begin{smallmatrix}1&2\\0&4\end{smallmatrix}\right), \left(\begin{smallmatrix}2&3\\5&0\end{smallmatrix}\right), \left(\begin{smallmatrix}3&0\\1&2\end{smallmatrix}\right), \left(\begin{smallmatrix}0&5\\4&2\end{smallmatrix}\right)$ b) $\left(\begin{smallmatrix}1&2\\3&4\end{smallmatrix}\right), \left(\begin{smallmatrix}2&3\\1&2\end{smallmatrix}\right), \left(\begin{smallmatrix}1&1\\4&1\end{smallmatrix}\right), \left(\begin{smallmatrix}1&4\\0&5\end{smallmatrix}\right)$

Úloha 5.2.5. Zistite, ktoré z daných vektorov patria do podpriestoru $[(1,4,1,0),(2,3,-2,-3),(0,2,-5,-6)]$ priestoru $\mathbb{R}^4$: a) $(4,11,-3,-3)$, b) $(1,0,11,12)$, c) $(3,0,4,1)$, d) $(1,-1,2,-2)$.

Úloha 5.2.6. Zistite, či $[\vec{\beta}_1, \vec{\beta}_2] \subseteq [\vec{\gamma}_1, \vec{\gamma}_2, \vec{\gamma}_3]$ vo vektorovom priestore $\mathbb{R}^4$ nad poľom $\mathbb{R}$, ak $\vec{\gamma}_1 = (1,1,5,1)$, $\vec{\gamma}_2 = (1,0,2,1)$, $\vec{\gamma}_3 = (2,1,0,1)$, $\vec{\beta}_1 = (1,1,5,1)$ a $\vec{\beta}_2 = (-1,1,-6,-2)$.

Úloha 5.2.7. Zistite hodnosti matíc
$\left(\begin{smallmatrix}1&2&3&4&5\\0&-1&3&8&0\\0&0&2&3&0\end{smallmatrix}\right)$ $\left(\begin{smallmatrix}1&0&-1&2&3\\0&0&5&0&1\\0&0&0&1&-1\\0&0&0&0&7\end{smallmatrix}\right)$ $\left(\begin{smallmatrix}1&0&\dots&0\\0&1&\dots&0\\\dots&\dots&\dots&\dots\\0&0&\dots&1\\1&1&\dots&1\end{smallmatrix}\right)$ $\left(\begin{smallmatrix}2&-1&3&-2&4\\4&-2&5&1&7\\2&-1&1&8&2\end{smallmatrix}\right)$ $\left(\begin{smallmatrix}1&3&5&-1\\2&-1&-3&4\\5&1&-1&7\\7&7&9&1\end{smallmatrix}\right)$

{rtmcvic:HODN}

Úloha 5.2.8. Upravte danú maticu nad poľom $\mathbb{R}$ na redukovaný trojuholníkový tvar a určte hodnosť matice
$\left(\begin{smallmatrix}1&-2&-2&2\\2&2&-1&-1\\3&3&-4&-4\end{smallmatrix}\right)$ $\left(\begin{smallmatrix}3&-1&3&2&5\\5&-3&2&3&4\\1&-3&-5&0&-7\\7&-5&1&4&1\end{smallmatrix}\right)$ $\left(\begin{smallmatrix}1&3&0&5&0&-1\\2&6&1&10&0&0\\5&15&2&25&-1&-4\\3&9&1&15&0&-1\end{smallmatrix}\right)$ $\left(\begin{smallmatrix}4&3&-5&2&3\\8&6&-7&4&2\\4&3&-8&2&7\\4&3&1&2&-5\\8&6&-1&4&-6\end{smallmatrix}\right)$

{rtmcvic:HODNPAR}

Úloha 5.2.9. Určte hodnosť danej matice v závislosti od parametra $c \in \mathbb{R}$
$A = \left(\begin{smallmatrix}1&c&-1&2\\2&-1&c&5\\1&10&-6&1\end{smallmatrix}\right)$ $\left(\begin{smallmatrix}3&2&c&2c\\1&-1&3&-c\\2&3&0&1\end{smallmatrix}\right)$ $\left(\begin{smallmatrix}2&c{+}1&0\\2&c{-}1&2c\\c&c&c\end{smallmatrix}\right)$ $\left(\begin{smallmatrix}2&c{+}1&0\\4&c{-}1&2c\\c&c&c\end{smallmatrix}\right)$

Úloha 5.2.10. Zistite, či priestor [(2,4,4,2,4),(3,1,1,2,2),(4,3,3,2,0)] je podpriestor priestoru [(1,1,0,1,4),(2,1,3,3,1),(3,2,1,1,3)] a) nad $\mathbb{Q}$, b) nad $\mathbb{Z}_5$, c) nad $\mathbb{Z}_7$.

Úloha 5.2.11. Zistite, ktoré z daných matíc sú navzájom riadkovo ekvivalentné:
$\left(\begin{smallmatrix}2&3&1\\4&3&3\\1&2&4\end{smallmatrix}\right)$ $\left(\begin{smallmatrix}1&3&1\\2&4&3\\3&1&2\end{smallmatrix}\right)$ $\left(\begin{smallmatrix}0&3&1\\2&0&3\\1&1&2\end{smallmatrix}\right)$ $\left(\begin{smallmatrix}3&0&1\\0&3&2\\1&0&3\end{smallmatrix}\right)$ $\left(\begin{smallmatrix}1&0&1\\0&1&1\\2&1&3\end{smallmatrix}\right)$ $\left(\begin{smallmatrix}3&2&1\\4&2&1\\1&0&1\end{smallmatrix}\right)$

Úloha 5.2.12. Nájdite bázu daného podpriestoru a určite jeho dimenziu:
a) $[(1,1,0,-1),(0,1,2,1),(1,0,1,-1),(1,1,-6,-3),(-1,-5,1,0)]$ v $\mathbb{R}^4$;
b) $[(1,2,2,0,1),(1,2,0,1,2),(1,2,-2,1,0)]$ v $\mathbb{R}^5$
c) $[(1,2,2,0,1),(1,2,0,1,2),(1,2,3,1,0)]$ v $\mathbb{Z}_5^5$
d) $[(1,2,2,0,1),(1,2,0,1,2),(1,2,3,1,0)]$ v $\mathbb{Z}_7^5$.

Úloha 5.2.13. Zistite, pre aké hodnoty parametra c sú dané vektory lineárne závislé
a) $(-1,0,-1)$, $(2,1,2)$, $(1,1,c)$ v $\mathbb{R}^3$;
b) $(1,1,3)$, $(2,1,2)$, $(c,0,-c)$ v $\mathbb{R}^3$;
c) $(2,0,-1)$, $(3,2,0)$, $(1,-2,c)$ v $\mathbb{R}^3$.

Úloha 5.2.14*. Určite hodnosť matice:

$$\begin{pmatrix} 1 & 1 & \dots & 1 & 1 \\ a_1 & a_2 & \dots & a_n & a_{n+1} \\ a_1^2 & a_2^2 & \dots & a_n^2 & a_{n+1}^2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_1^n & a_2^n & \dots & a_n^n & a_{n+1}^n \end{pmatrix}$$

ak viete, že $a_1, \ldots, a_{n+1}$ sú navzájom rôzne reálne čísla (t.j. $a_i \neq a_j$ pre všetky $i \neq j$).

Pri riešení tejto úlohy môžete použiť fakt, že elementárne stĺpcové operácie nemenia hodnosť, resp. to, že $h(A) = h(A^T)$. (Tento fakt dokážeme neskôr.) Ale mala by sa dať vyriešiť aj bez použitia tejto veci.

Všetky príklady, v ktorých vystupujú len celé čísla, si môžete upraviť tak, že jednotlivé členy matice nahradíte ich zvyškami po delení 3 (5 , 7) a riešite rovnakú úlohu nad $\mathbb{Z}_3$ ($\mathbb{Z}_5$, $\mathbb{Z}_7$).

5.3 Lineárne zobrazenia

V tejto časti zavedieme vlastne najdôležitejší koncept tejto prednášky – lineárne zobrazenia. Ak by sme povedali, že všetko čo sme robili doteraz sme robili iba s cieľom, aby sme si pripravili vhodné prostriedky na popis lineárnych zobrazení, neboli by sme ďaleko od pravdy. Táto prednáška totiž do veľkej miery smeruje k tomu, aby sme pochopili lineárne javy, ktoré sa v matematike (ale aj vo fyzike a ďalších aplikáciach) popisujú práve pomocou lineárnych zobrazení.

{lzob:DEFLZ}

Definícia 5.3.1. Ak V a W sú vektorové priestory nad poľom F a $f\colon V \to W$ je zobrazenie z V do W, tak hovoríme, že f je *lineárne zobrazenie,* ak pre ľubovoľné $\vec{\alpha}, \vec{\beta} \in V$ a ľubovoľné $c \in F$ platí

{lzob:1}

(i) $f(\vec{\alpha} + \vec{\beta}) = f(\vec{\alpha}) + f(\vec{\beta})$,

{lzob:2}

(ii) $f(c\vec{\alpha}) = cf(\vec{\alpha})$.

Inými slovami, lineárne zobrazenia sú tie zobrazenia, ktoré zachovávajú základné operácie popisujúce vektorový priestor.

Asi sa hodí ilustrovať nový pojem na niekoľkých príkladoch. Začnime najprv nejakými veľmi triviálnymi:

Príklad 5.3.2. Uvažujeme identické zobrazenie $id_V\colon V \to V$, t.j. $id_V(\vec{\alpha}) = \vec{\alpha}$. (Pričom V je ľubovoľný vektorový priestor nad nejakým poľom F.) Prakticky ihneď vidno, že ide o lineárne zobrazenie, lebo podmienky z definície 5.3.1 aplikované na toto konkrétne zobrazenie sú rovnosti

$$\vec{\alpha} + \vec{\beta} = \vec{\alpha} + \vec{\beta},$$
$$c\vec{\alpha} = c\vec{\alpha},$$

ktoré očividne platia pre ľubovoľné $c \in F$, $\vec{\alpha}, \vec{\beta} \in V$.

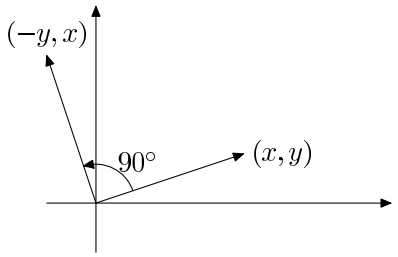
Iný triviálny príklad dostaneme, ak vezmeme nejaké dva vektorové priestory V, W nad tým istým poľom F a definujeme

$$f(\vec{\alpha}) = \vec{0}$$

pre $\vec{\alpha} \in V$. (T.j. každému vektoru sme priradili nulový vektor. Môžeme toto zobrazenie nazvať *nulové zobrazenie.*) Opäť je veľmi ľahké skontrolovať, že ide o lineárne zobrazenie, vlastne si stačí všimnúť, že platia rovnosti

$$\vec{0} = \vec{0} + \vec{0},$$
$$\vec{0} = c\vec{0}.$$

Azda sa teraz môžeme skúsiť posunúť k zaujímavejším (menej triviálnym) príkladom.



zob:FIGVECT6} Obr. 5.1: Otočenie v rovine o 90°

{lzob:PROTOC}

Príklad 5.3.3. Otočenie v rovine o 90° je lineárne zobrazenie.

Otočenie v rovine o 90° má vyjadrenie v súradniciach[1]

$$f(x,y) = (-y, x).$$

Overme podmienky z definície lineárneho zobrazenia:

$$f(x,y) + f(x',y') = (-y,x) + (-y',x') = (-(y+y'), x+x') = f(x+x', y+y'),$$
$$f(cx,cy) = (-cy, cx) = c(-y,x) = cf(x,y).$$

{lzob:PR2}

Príklad 5.3.4. Zobrazenie $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ určené predpisom $f(x,y) = (2x+y, x+3y)$ je lineárne zobrazenie. (V istom zmysle je typickým príkladom lineárneho zobrazenia, ako uvidíme neskôr.)

$$f(x,y) + f(x',y') = (2x+y, x+3y) + (2x'+y', x'+3y') = \\ (2(x+x') + y + y', x + x' + 3(y+y')) = f(x+x', y+y'),$$

$$f(cx,cy) = (2cx+cy, cx+3cy) = c(2x+y, x+3y) = cf(x,y).$$

Veľmi podobne by sme vedeli ukázať, že pre ľubovoľné reálne čísla a, b, c, d predpis

$$f(x,y) = (ax+by, cx+dy)$$

určuje lineárne zobrazenie. Neskôr uvidíme, ako sa to dá zovšeobecniť a tiež to, že každé lineárne zobrazenie $\mathbb{R}^2 \to \mathbb{R}^2$ vyzerá takto; pozri poznámku 5.4.11.

{lzob:VTCHARLZ}

Veta 5.3.5. *Nech V, W sú vektorové priestory nad poľom F a $f\colon V \to W$ je zobrazenie. Nasledujúce podmienky sú ekvivalentné:*

{lzob:1.1}

(a) *zobrazenie f je lineárne,*

{lzob:1.2}

(b) $f(c\vec{\alpha} + d\vec{\beta}) = cf(\vec{\alpha}) + df(\vec{\beta})$ *pre ľubovoľné $c, d \in F$ a ľubovoľné $\vec{\alpha}, \vec{\beta} \in V$,*

{lzob:1.3}

(c) $f(c_1\vec{\alpha}_1 + \cdots + c_n\vec{\alpha}_n) = c_1 f(\vec{\alpha}_1) + \ldots + c_n f(\vec{\alpha}_n)$ *pre ľubovoľné $c_1, \ldots, c_n \in F$, $\vec{\alpha}_1, \ldots, \vec{\alpha}_n \in V$.*

{lzob:1.4}

(d) $f(c\vec{\alpha} + \vec{\beta}) = cf(\vec{\alpha}) + f(\vec{\beta})$ *pre ľubovoľné $c \in F$ a $\vec{\alpha}, \vec{\beta} \in V$.*

[1] Keby sme boli presní, mali by sme písať $f((x,y))$ – jednu zátvorku kvôli zobrazeniu a druhú z označenia vektora. Rozhodli sme sa, že si zápis trochu zjednodušíme.

Tretia podmienka v predchádzajúcej vete hovorí, že lineárne zobrazenia sú práve zobrazenia zachovávajúce lineárne kombinácie.

Dôkaz. (a) $\Rightarrow$ (b) Vyplýva priamo z definície lineárneho zobrazenia.

$$f(c\vec{\alpha} + d\vec{\beta}) \overset{(i)}{=} f(c\vec{\alpha}) + f(d\vec{\beta}) \overset{(ii)}{=} cf(\vec{\alpha}) + df(\vec{\beta})$$

(b) $\Rightarrow$ (c) Dostaneme opakovaným použitím (b). (Formálny dôkaz by sme urobili pomocou matematickej indukcie.)

(c) $\Rightarrow$ (a) Ak dosadíme $n = 1$ dostaneme podmienku (ii) z definície lineárneho zobrazenia. Pre $n = 2$ a $c_1 = c_2 = 1$ máme podmienku (i).

Ak už máme overenú ekvivalenciu prvých troch podmienok, tak ukázať, že aj (d) je s nimi ekvivalentná je pomerne jednoduché. Túto časť dôkazu som ponechal ako cvičenie – úloha 5.3.2. □

Tvrdenie 5.3.6. *Ak f je lineárne zobrazenie, tak $f(\vec{0}) = \vec{0}$.*

Dôkaz.

$$f(\vec{0}) = f(\vec{0} + \vec{0}) = f(\vec{0}) + f(\vec{0})$$

Vykrátením $f(\vec{0})$ dostaneme $f(\vec{0}) = \vec{0}$. □

Nasledujúcu vetu niektorí autori nazývajú *základná veta o lineárnych zobrazeniach.*

{lzob:VTOBRBAZY}

Veta 5.3.7. *Nech V, W sú vektorové priestory. Nech $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ je báza priestoru V a nech $\vec{\beta}_1, \ldots, \vec{\beta}_n \in W$. Potom existuje práve jedno lineárne zobrazenie $f\colon V \to W$ také, že*

$$f(\vec{\alpha}_i) = \vec{\beta}_i$$

pre $i = 1, 2, \ldots, n$.

Dôkaz. Nech $\vec{\alpha} \in V$. Pretože $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ tvorí bázu priestoru V, existujú jednoznačne určené skaláry $c_1, \ldots, c_n \in F$ také, že

$$\vec{\alpha} = c_1\vec{\alpha}_1 + \cdots + c_n\vec{\alpha}_n.$$

Potom $f(\vec{\alpha})$ definujeme ako

$$f(\vec{\alpha}) = c_1\vec{\beta}_1 + \cdots + c_n\vec{\beta}_n.$$

(Pretože $c_1, \ldots, c_n$ sú jednoznačne určené, je f dobre definované, t.j., nemôže sa stať, že by sme takto tomu istému $\vec{\alpha}$ priradili dve rôzne hodnoty. Súčasne $f(\vec{\alpha})$ nemôže mať inú hodnotu, ak to má byť lineárne zobrazenie – vyplýva to z toho, že lineárne zobrazenia zachovávajú lineárne kombinácie. Z toho vyplýva jednoznačnosť zobrazenia f.)

Ukážeme, že takto definované zobrazenie je skutočne lineárne. Uvažujme dva vektory

$$\vec{\alpha} = c_1\vec{\alpha}_1 + \cdots + c_n\vec{\alpha}_n$$
$$\vec{\alpha'} = d_1\vec{\alpha}_1 + \cdots + d_n\vec{\alpha}_n$$

Potom platí

$$\vec{\alpha} + \vec{\alpha'} = (c_1 + d_1)\vec{\alpha}_1 + \ldots + (c_n + d_n)\vec{\alpha}_n$$
$$f(\vec{\alpha} + \vec{\alpha'}) = (c_1 + d_1)\vec{\beta}_1 + \ldots + (c_n + d_n)\vec{\beta}_n = c_1\vec{\beta}_1 + \cdots + c_n\vec{\beta}_n + d_1\vec{\beta}_1 + \cdots + d_n\vec{\beta}_n = f(\vec{\alpha}) + f(\vec{\alpha'})$$

Podobne dostaneme

$$c.\vec{\alpha} = cc_1\vec{\alpha}_1 + \cdots + cc_n\vec{\alpha}_n$$
$$f(c.\vec{\alpha}) = cc_1\vec{\beta}_1 + \cdots + cc_n\vec{\beta}_n = c(c_1\vec{\beta}_1 + \cdots + c_n\vec{\beta}_n) = c.f(\vec{\alpha})$$

□

Lineárne zobrazenie je jednozačne určené obrazmi prvkov (ľubovoľnej) bázy.
V priestore F^n máme štandardnú bázu

$$\vec{\varepsilon}_1 = (1, 0, \ldots, 0), \vec{\varepsilon}_2 = (0, 1, \ldots, 0), \ldots, \vec{\varepsilon}_n = (0, \ldots, 0, 1).$$

Táto báza nám umožní popísať ľubovoľné zobrazenie z F^n do F^k spôsobom, ktorý je v istom zmysle kanonický.

{lzob:DEFMATICAZOBRAZENI

Definícia 5.3.8. Nech F je pole. *Matica lineárneho zobrazenia* $f\colon F^m \to F^n$ je matica typu $m \times n$ ktorej k-ty riadok je vektor $f(\vec{\varepsilon}_k)$.

Maticu zobrazenia f budeme označovať A_f.
Každému lineárnemu zobrazeniu $f\colon F^m \to F^n$ sme takto priradili nejakú maticu A_f typu $m \times n$.
Obrátene, ľubovoľnou maticou typu $m \times n$ je jednoznačne určené lineárne zobrazenie $f\colon F^m \to F^n$. (Riadky matice určujú obrazy bázových vektorov, jednoznačnosť a existencia takéhoto zobrazenia vyplývajú z vety 5.3.7.) Lineárne zobrazenie prislúchajúce matici A budeme označovať f_A.

Príklad 5.3.9. Uvažujme lineárne zobrazenie $f\colon \mathbb{R}^2 \to \mathbb{R}^3$ dané predpisom $f(x, y) = (2x + y, x + y, x + 2y)$. Dosadením zistíme, že platí

$$f(\vec{\varepsilon}_1) = f(1, 0) = (2, 1, 1)$$
$$f(\vec{\varepsilon}_2) = f(0, 1) = (1, 1, 2)$$

Teda matica tohoto zobrazenia je

$$\begin{pmatrix} 2 & 1 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

Veta 5.3.10. *Nech U, V, W sú vektorové priestory nad tým istým poľom F. Ak $f\colon U \to V$ a $g\colon V \to W$ sú lineárne zobrazenia, tak aj $g \circ f$ je lineárne zobrazenie.*

Dôkaz. Na overenie použijeme podmienku (b) z vety 5.3.5. Nech $\vec{\alpha}, \vec{\beta} \in U$ a $c, d \in F$. Potom dostaneme

$$g(f(c\vec{\alpha} + d\vec{\beta})) = g(cf(\vec{\alpha}) + df(\vec{\beta})) = cg(f(\vec{\alpha})) + dg(f(\vec{\beta})).$$

(Využili sme najprv linearitu zobrazenia f a potom linearitu zobrazenia g.) □

Poznámka 5.3.11. Ľahko sa overí, že ak $f, g\colon V \to W$ sú lineárne zobrazenia, tak aj zobrazenia $f + g$ a $c.f$ sú lineárne.

Teraz si ukážeme na konkrétnom príklade, ako vieme nájsť maticu lineárneho zobrazenia, ak máme dané obrazy niektorých vektorov. (V prípade, že tieto vektory tvoria bázu, také zobrazenie existuje podľa vety 5.3.7.)

{lzobcvic:ULOZOBR}

Úloha 5.3.1. Nájdite maticu lineárneho zobrazenia $f\colon \mathbb{R}^3 \to \mathbb{R}^4$, pre ktoré platí:
a) $f(2, 0, 3) = (1, 2, -1, 1)$, $f(4, 1, 5) = (4, 5, -2, 1)$, $f(3, 1, 2) = (1, -1, 1, -1)$,
b) $f(2, 0, 3) = (1, 2, -1, 1)$, $f(4, 1, 5) = (4, 5, -2, 1)$, $f(2, -1, 4) = (-1, 1, -1, 2)$,
c) $f(2, 0, 3) = (1, 2, -1, 1)$, $f(4, 1, 5) = (4, 5, -2, 1)$, $f(2, -1, 4) = (1, -1, 1, -1)$.

Postupujeme tak, že si napíšeme do matice vektory a ich obrazy a ľavú časť sa snažíme upraviť riadkovými úpravami na jednotkovú maticu (aby sme našli obrazy vektorov $\vec{\varepsilon}_i$)
$\left(\begin{array}{ccc|cccc} 2&0&3&1&2&-1&1\\ 4&1&5&4&5&-2&1\\ 3&1&2&1&-1&1&-1 \end{array}\right) \sim \left(\begin{array}{ccc|cccc} 1&0&\frac{3}{2}&\frac{1}{2}&1&-\frac{1}{2}&\frac{1}{2}\\ 0&1&-1&2&1&0&-1\\ 0&1&-\frac{5}{2}&-\frac{1}{2}&4&\frac{5}{2}&-1 \end{array}\right) \sim \left(\begin{array}{ccc|cccc} 1&0&\frac{3}{2}&\frac{1}{2}&1&-\frac{1}{2}&\frac{1}{2}\\ 0&1&-1&2&1&0&-1\\ 0&0&-\frac{3}{2}&-\frac{5}{2}&-5&\frac{5}{2}&-\frac{3}{2} \end{array}\right) \sim \left(\begin{array}{ccc|cccc} 1&0&0&-2&-4&2&-1\\ 0&1&0&\frac{11}{3}&\frac{13}{3}&-\frac{5}{3}&0\\ 0&0&1&\frac{5}{3}&\frac{10}{3}&-\frac{5}{3}&1 \end{array}\right)$
Hľadaná matica je $\begin{pmatrix} -2&-4&2&-1\\ \frac{11}{3}&\frac{13}{3}&-\frac{5}{3}&0\\ \frac{5}{3}&\frac{10}{3}&-\frac{5}{3}&1 \end{pmatrix}$.

Základná myšlienka algoritmu, ktorý sme práve popísali, je v tom, že po každom kroku platí, že vektor na ľavej strane sa zobrazením s danými vlastnosťami zobrazí na vektor napravo od neho. Skutočne, ak máme v nejako kroku $f(\vec{\alpha}_i) = \vec{\beta}_i$ a $f(\vec{\alpha}_j) = \vec{\beta}_j$ (kde $\vec{\alpha}_k$ a $\vec{\beta}_k$ označuje k-ty riadok ľavej resp. pravej časti matice) a pripočítame k i-temu riadku c-násobok j-teho riadku, platí tento vzťah aj pre riadky novej matice:

$$f(\vec{\alpha}_i + c\vec{\alpha}_j) = f(\vec{\alpha}_i) + cf(\vec{\alpha}_j) = \vec{\beta}_i + c\vec{\beta}_j.$$

Podobne sa to dá overiť pre ostatné elementárne riadkové operácie.[2]

Skúšku správnosti môžeme urobiť tak, že overíme, či f naozaj nadobúda zadané hodnoty, napríklad $f(3,1,2) = 3(-2,-4,2,-1) + 1(\frac{11}{3},\frac{13}{3},-\frac{5}{3},0) + 2(\frac{5}{3},\frac{10}{3},-\frac{5}{3},1) = (1,-1,1,-1)$.

V prípade, že skúška nevyjde, chybu môžeme hľadať tak, že skúšame pre medzivýsledky, či sa vektor na ľavej strane zobrazí na vektor ležiaci od neho napravo v zobrazení určenom maticou, ktorá nám vyšla. Samozrejme chybu môžeme hľadať aj tak, že kontrolujeme jednotlivé úpravy.

b) $\left(\begin{array}{ccc|cccc} 2&0&3&1&2&-1&1\\ 4&1&5&4&5&-2&1\\ 2&-1&4&-1&1&-1&2 \end{array}\right) \sim \left(\begin{array}{ccc|cccc} 2&0&3&1&2&-1&1\\ 0&1&-1&2&1&0&-1\\ 0&-1&1&-2&-1&0&1 \end{array}\right) \sim \left(\begin{array}{ccc|cccc} 2&0&3&1&2&-1&1\\ 0&1&-1&2&1&0&-1\\ 0&0&0&0&0&0&0 \end{array}\right)$

Vidíme, že môžeme $f(0,0,1)$ zvoliť ľubovoľne. Označme $f(0,0,1) = (a,b,c,d)$. Potom dostaneme
$\left(\begin{array}{ccc|cccc} 2&0&3&1&2&-1&1\\ 0&1&-1&2&1&0&-1\\ 0&0&1&a&b&c&d \end{array}\right) \sim \left(\begin{array}{ccc|cccc} 1&0&0&\frac{1-3a}{2}&1-\frac{3}{2}b&\frac{1-3c}{2}&1-\frac{3}{2}d\\ 0&1&0&2+a&1+b&c&-1+d\\ 0&0&1&a&b&c&d \end{array}\right)$
Pre každé $a,b,c,d \in \mathbb{R}$ je táto matica maticou zobrazenia f s požadovanými vlastnosťami.

c) $\left(\begin{array}{ccc|cccc} 2&0&3&1&2&-1&1\\ 4&1&5&4&5&-2&1\\ 2&-1&4&1&-1&1&-1 \end{array}\right) \sim \left(\begin{array}{ccc|cccc} 2&0&3&1&2&-1&1\\ 0&1&-1&2&1&0&-1\\ 0&-1&1&0&-3&-2&-2 \end{array}\right) \sim \left(\begin{array}{ccc|cccc} 2&0&3&1&2&-1&1\\ 0&1&-1&2&1&0&-1\\ 0&0&0&2&-2&-2&-2 \end{array}\right)$

Pre lineárne zobrazenie f, ktoré by spĺňalo podmienky zo zadania by muselo platiť $f(0,0,0) = (2,-2,-2,-2)$, ale také lineárne zobrazenie neexistuje. (Lineárne zobrazenie vždy zobrazuje nulový vektor na nulový vektor.)

Cvičenia

{lzobcvic:ULOCHARLZ}

Úloha 5.3.2. Ukážte, že $f\colon V \to W$ je lineárne práve vtedy, keď pre ľubovoľné $c \in F$ a $\vec{\alpha}, \vec{\beta} \in V$ platí $f(c\vec{\alpha} + \vec{\beta}) = cf(\vec{\alpha}) + f(\vec{\beta})$.

Úloha 5.3.3. Nájdite maticu lineárneho zobrazenia $f\colon (\mathbb{Z}_7)^2 \to (\mathbb{Z}_7)^2$ a napíšte jeho predpis.
a) $f(1,1) = (0,1)$, $f(6,1) = (3,2)$
b) $f(2,3) = (1,0)$, $f(3,2) = (6,1)$

Úloha 5.3.4. Nájdite maticu lineárneho zobrazenia $f\colon \mathbb{R}^4 \to \mathbb{R}^4$ takého, že:
a) $f(1,2,3,1) = (1,3,1,0)$, $f(2,1,3,0) = (0,1,3,1)$, $f(3,2,1,0) = (1,0,3,0)$, $f(2,2,3,4) = (3,1,0,4)$
b) $f(1,2,3,4) = (0,0,0,0)$, $f(2,1,3,1) = (1,0,3,1)$, $f(0,1,2,0) = (2,0,1,0)$, $f(1,0,3,1) =$

[2] Takýto postup je typický pri dokazovaní správnosti algoritmov. Našli sme tvrdenie, ktoré platí po každom kroku algoritmu. (Nazývame ho *invariant.*) O tomto invariante treba dokázať, že: a) platí na začiatku výpočtu; b) vykonanie jedného kroku algoritmu nezmení platnosť invariantu; c) ak platí invariant na konci výpočtu, tak algoritmus skutočne robí, to čo má.

$(2,1,3,1)$
c) $f(0,1,1,1)=(1,0,0,0)$, $f(1,0,1,1)=(0,1,0,0)$, $f(1,1,0,1)=(0,0,1,0)$, $f(1,1,1,0)=(0,0,0,1)$

Úloha 5.3.5. Nech V a W sú vektorové priestory nad poľom F a $f\colon V\to W$ je lineárne zobrazenie. Ak $\vec{\alpha}_1,\dots,\vec{\alpha}_n$ sú lineárne závislé vektory, tak aj $f(\vec{\alpha}_1),\dots,f(\vec{\alpha}_n)$ sú lineárne závislé vektory.

{lzobcvic:KERIMPPR}

Úloha 5.3.6. Nech $f\colon V\to W$ je lineárne zobrazenie z vektorového priestoru V do vektorového priestoru W nad poľom F. Dokážte:
Ak S je podpriestor vektorového priestoru V, tak $f[S]=\{f(\vec{\alpha});\vec{\alpha}\in S\}$ je podpriestor vektorového priestoru W.
Ak T je podpriestor vektorového priestoru W, tak $f^{-1}(T)=\{\vec{\alpha}\in V : f(\vec{\alpha})\in T\}$ je podpriestor vektorového priestoru V.

5.4 Súčin matíc

V predchádzajúcej časti sme sa naučili, že lineárne zobrazenie $F^m\to F^n$ je jednoznačne určené maticou typu $m\times n$ a obrátene, každému takémuto lineárnemu zobrazeniu prislúcha jeho matica. Tiež sme sa dozvedeli, že zložením lineárnych zobrazení opäť vznikne lineárne zobrazenie. Preto je prirodzená otázka ako vyzerá matica prislúchajúca zloženému zobrazeniu.

{suc:PRSUCMAT}

Príklad 5.4.1. Uvažujme zobrazenie $f\colon\mathbb{R}^2\to\mathbb{R}^3$ určené maticou $\left(\begin{smallmatrix}1&0&2\\2&1&1\end{smallmatrix}\right)$ a zobrazenie $g\colon\mathbb{R}^3\to\mathbb{R}^2$ určené maticou $\left(\begin{smallmatrix}3&1\\1&1\\0&-1\end{smallmatrix}\right)$. Pokúsme sa vypočítať maticu zloženého zobrazenia $A_{g\circ f}$.

Budeme označovať štandardnú bázu v $\mathbb{R}^2$ ako $\vec{\delta}_1,\vec{\delta}_2$ a štandardnú bázu v $\mathbb{R}^3$ ako $\vec{\varepsilon}_1,\vec{\varepsilon}_2,\vec{\varepsilon}_3$.

Vypočítajme obrazy vektorov štandardnej bázy:

$$g(f(\vec{\delta}_1))=g(1,0,2)=g(\vec{\varepsilon}_1+2\vec{\varepsilon}_3)=g(\vec{\varepsilon}_1)+2g(\vec{\varepsilon}_3)=(3,1)+2.(0,-1)=(3,-1)$$
$$g(f(\vec{\delta}_2))=g(2,1,1)=g(2\vec{\varepsilon}_1+\vec{\varepsilon}_2+\vec{\varepsilon}_3)=2g(\vec{\varepsilon}_1)+g(\vec{\varepsilon}_2)+g(\vec{\varepsilon}_3)=2.(3,1)+(1,1)+(0,-1)=(7,2)$$

Z toho dostávame

$$A_{g\circ f}=\left(\begin{smallmatrix}3&-1\\7&2\end{smallmatrix}\right).$$

Pokúsme sa zopakovať tento výpočet vo všeobecnosti. Máme teda dve lineárne zobrazenia a im prislúchajúce matice:

$$f\colon F^m\to F^n\qquad A_f=\begin{pmatrix}a_{11}&\dots&a_{1n}\\\vdots&\ddots&\vdots\\a_{m1}&\dots&a_{mn}\end{pmatrix}\text{ typu } m\times n$$
$$g\colon F^n\to F^k\qquad A_g=\begin{pmatrix}b_{11}&\dots&b_{1k}\\\vdots&\ddots&\vdots\\b_{n1}&\dots&b_{nk}\end{pmatrix}\text{ typu } n\times k$$

Opäť nech $\vec{\delta}_1,\dots,\vec{\delta}_m$ je štandardná báza F^m a $\vec{\varepsilon}_1,\dots,\vec{\varepsilon}_n$ je štandardná báza F^n.

Pretože $g\circ f\colon F^m\to F^k$ je matica zloženého zobrazenia matica typu $m\times k$. Riadky matice $A_{g\circ f}$ sú vektory $g(f(\vec{\delta}_1)),g(f(\vec{\delta}_2)),\dots,g(f(\vec{\delta}_m))$.

Nech $i \in \{1, 2, \dots m\}$. Vypočítajme príslušný riadok matice $A_{g\circ f}$.

$$g(f(\vec{\delta}_i)) = g(a_{i1}, a_{i2}, \dots, a_{in}) =$$
$$g(a_{i1}\vec{\varepsilon}_1 + a_{i2}\vec{\varepsilon}_2 + \dots + a_{in}\vec{\varepsilon}_n) = g(a_{i1}\vec{\varepsilon}_1) + g(a_{i2}\vec{\varepsilon}_2) + \dots + g(a_{in}\vec{\varepsilon}_n) =$$
$$a_{i1}(b_{11}, b_{12}, \dots, b_{1k}) +$$
$$a_{i2}(b_{21}, b_{22}, \dots, b_{2k}) +$$
$$\vdots$$
$$a_{in}(b_{n1}, b_{n2}, \dots, b_{nk}) =$$
$$(a_{i1}b_{11} + a_{i2}b_{21} + \dots + a_{in}b_{n1}, a_{i1}b_{12} + a_{i2}b_{22} + \dots + a_{in}b_{n2}, \dots, a_{ik}b_{1k} + a_{i2}b_{2k} + \dots + a_{in}b_{nk})$$

Vektor $g(f(\vec{\delta}_i))$ má na j-tej súradnici hodnotu

$$c_{ij} := a_{i1}b_{1j} + a_{i2}b_{2j} + \dots + a_{in}b_{nj} = \sum_{t=1}^{n} a_{it}b_{tj}.$$

Matica zloženého zobrazenia je matica $A_{g\circ f} = ||c_{ij}||$ typu $m \times k$.

Všimnime si, že prvok c_{ij} vlastne získame tak, že vezmeme i-ty riadok matice A a j-ty stĺpec matice B (dostaneme tak 2 vektory rovnakej dĺžky n), vynásobíme hodnoty na rovnakých súradniciach a takto získané hodnoty sčítame. (Je to vlastne skalárny súčin i-teho riadku matice A a j-teho stĺpca a matice B – o skalárnom súčine ešte budeme hovoriť neskôr.)

Definícia 5.4.2. Ak A je matica typu $m \times n$ a B je matica typu $n \times k$ nad poľom F, tak maticu $C = ||c_{ij}||$ typu $m \times k$, kde

$$c_{ij} = \sum_{t=1}^{n} a_{it}b_{tj}$$

pre $i = 1, 2, \dots, m$ a $j = 1, 2, \dots, k$, nazývame *súčin matíc* A a B. Označujeme ju AB alebo $A \cdot B$.

Dôležité je si všimnúť, že súčin matíc definujeme iba v prípade, že počet stĺpcov prvej matice sa rovná počtu riadkov druhej matice.

$$m \times \boxed{n \quad n} \times k$$

Výsledok je matica typu $m \times k$.

Príklad 5.4.3. $\begin{pmatrix} 1 & 0 & 2 \\ 2 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 3 & 1 \\ 1 & 1 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 3 & -1 \\ 7 & 2 \end{pmatrix}$
$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 2 \\ 2 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 2 \\ 2 & 1 & 1 \end{pmatrix}$

Veta 5.4.4. *Nech F je pole, $f\colon F^m \to F^n$ a $g\colon F^n \to F^k$ sú lineárne zobrazenia. Potom platí*

$$A_{g\circ f} = A_f \cdot A_g$$

Dôkaz. Súčin matíc sme definovali práve tak, aby platil predchádzajúci vzťah. (Dôkaz tejto vety spočíva vlastne v odvodení vzťahu pre $A_{g\circ f}$, ktoré sme uviedli za príkladom 5.4.1.) □

POZOR na zmenu poradia v predchádzajúcej vete. T.j. matice na pravej strane rovnosti sú zapísané v inom poradí ako je zápis skladania zobrazení. (Opäť platí, že niektoré knihy, ako napríklad [KGGS], definujú poradie skladania zobrazení inak, čo samozrejme ovplyvní aj poradie v predchádzajúcej rovnosti.)

c:PRZLOZOTOC}

Príklad 5.4.5. Nie je ťažké uvedomiť si, že otočenie v rovine o uhol α (proti smeru pohybu hodinových ručičiek) je lineárne zobrazenie a jeho matica je

$$\begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}$$

(Pre $\alpha = 90°$ sme sa na toto zobrazenie pozreli v príklade 5.3.3.)

Ak zložíme otočenie o uhol α a otočenie o uhol β, mali by sme dostať otočenie o uhol $\alpha + \beta$. Už vieme, že zloženému zobrazeniu zodpovedá súčin matíc. Pozrime sa, čo dostaneme ako súčin matíc týchto dvoch zobrazení.

$$\begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} \cos\beta & \sin\beta \\ -\sin\beta & \cos\beta \end{pmatrix} = \begin{pmatrix} \cos\alpha\cos\beta - \sin\alpha\sin\beta & \cos\alpha\sin\beta + \sin\alpha\cos\beta \\ -\cos\alpha\sin\beta - \sin\alpha\cos\beta & \cos\alpha\cos\beta - \sin\alpha\sin\beta \end{pmatrix}$$

Ak si spomenieme na vzorce pre trigonometrické funkcie súčtu uhlov

$$\begin{aligned} \cos(\alpha+\beta) &= \cos\alpha\cos\beta - \sin\alpha\sin\beta, \\ \sin(\alpha+\beta) &= \cos\alpha\sin\beta + \sin\alpha\cos\beta, \end{aligned}$$

tak vidíme, že matica na pravej strane rovnosti sa skutočne rovná $\left(\begin{smallmatrix} \cos(\alpha+\beta) & \sin(\alpha+\beta) \\ -\sin(\alpha+\beta) & \cos(\alpha+\beta) \end{smallmatrix}\right)$.

Navyše ak tieto vzorce zabudneme, tak násobenie matíc nám dáva možnosť, ako si ich ľahko odvodiť. (Iná možnosť, ako si ich zapamätať, je násobenie komplexných čísel – pozri dôkaz vety C.2.3.)

Dôsledok 5.4.6. *Násobenie matíc je asociatívne, teda*

$$A \cdot (B \cdot C) = (A \cdot B) \cdot C$$

pre ľubovoľné matice také, že ich možno násobiť v uvedenom poradí.

Dôkaz. Ľubovoľná matica je matica nejakého lineárneho zobrazenia. Označme zobrazenia prislúchajúce daným maticiam f, g a h. Dostaneme

$$A_f \cdot (A_g \cdot A_h) = A_f \cdot (A_{h\circ g}) = A_{(h\circ g)\circ f} = A_{h\circ(g\circ f)} = A_{g\circ f} \cdot A_h = (A_f \cdot A_g) \cdot A_h.$$

(Inak povedané, vďaka tomu, že poznáme vzťah medzi násobením matíc a skladaním zobrazení, asociatívnosť násobenia matíc ľahko vyplýva z asociatívnosti skladania zobrazení.) □

To isté tvrdenie môžeme dokázať aj priamo z definície súčinu.

Dôkaz. Majme matice A, B, C typov $m \times n$, $n \times k$, $k \times l$. Vyjadrime prvok v i-tom riadku a j-tom stĺpci matice $A(BC)$. Dostaneme

$$\sum_{t=1}^{n} a_{it} \sum_{u=1}^{k} b_{tu}c_{uj} = \sum_{t=1}^{n}\sum_{u=1}^{k} a_{it}(b_{tu}c_{uj}).$$

Keď vypočítame prvok v i-tom riadku a j-tom stĺpci matice $(AB)C$ dostaneme

$$\sum_{u=1}^{k} \left(\sum_{t=1}^{n} a_{it}b_{tu} \right) c_{uj} = \sum_{u=1}^{k}\sum_{t=1}^{n} (a_{it}b_{tu})c_{uj},$$

čiže presne tú istú sumu, len s iným poradím sčitovania. □

Príklad 5.4.7. Násobenie matíc nie je komutatívne. (Vyplýva to aj z toho, že nie vždy, keď je definovaný súčin AB je definovaný aj súčin BA. Ukážeme si však aj príklad, kde sú súčiny v oboch poradiach definované, ale rôzne.)

$$\begin{pmatrix}1 & 1\\ 0 & 0\end{pmatrix}\begin{pmatrix}1 & 0\\ -1 & 0\end{pmatrix}=\begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix}$$
$$\begin{pmatrix}1 & 0\\ -1 & 0\end{pmatrix}\begin{pmatrix}1 & 1\\ 0 & 0\end{pmatrix}=\begin{pmatrix}1 & 1\\ -1 & -1\end{pmatrix}$$

Veta 5.4.8. *Nech matice A, B, C nad poľom F majú také rozmery, že uvedené súčty a súčiny majú zmysel.*

$$I_m A = A = A I_n$$
$$A(B+C) = AB + AC$$
$$(B+C)D = BD + CD$$

Dôkaz. Rovnosť dvoch výrazov obsahujúcich matice môžeme overiť tak, že vypočítame prvok v i-tom riadku a j-tom stĺpci matice na ľavej a pravej strane rovnosti a výsledky porovnáme.

Prvok v i-tom riadku a j-tom stĺpci matice $I_m A$ má tvar

$$c_{ij} = \delta_{i1}a_{1j} + \delta_{i2}a_{2j} + \ldots + \delta_{im}a_{mj},$$

kde ako δ_{ij} sme označili prvok i-teho riadku a j-teho stĺpca jednotkovej matice I_m (pozri tiež poznámku 5.1.7). Pretože z čísel δ_{ij} je len $\delta_{ii} = 1$ a ostatné sú nulové, dostávame z predchádzajúcej rovnosti priamo

$$c_{ij} = a_{ij}.$$

Vzťah pre násobenie jednotkovou maticou sprava sa overí rovnako.

Označme $D := A(B+C)$. Potom

$$d_{ij} = a_{i1}(b_{1j}+c_{1j}) + a_{i2}(b_{2j}+c_{2j}) + \ldots + a_{in}(b_{nj}+c_{nj}) =$$
$$a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{in}b_{nj} + a_{i1}c_{1j} + a_{i2}c_{2j} + \ldots + a_{in}c_{nj},$$

čo je presne súčet prvkov i-teho riadku a j-teho stĺpca matíc AB a AC. Teda skutočne platí

$$A(B+C) = AB + AC.$$

Predchádzajúce odvodenie by sme mohli stručnejšie a prehľadnejšie zapísať ako

$$d_{ij} = \sum_{t=1}^{n} a_{it}(b_{tj}+c_{tj}) = \sum_{t=1}^{n} a_{it}b_{tj} + \sum_{t=1}^{n} a_{it}c_{tj}.$$

Vzťah $(B+C)D = BD + CD$ sa overí úplne analogicky. □

Vidíme, že matica I má podobnú vlastnosť ako neutrálny prvok nejakej binárnej operácie. Mohla by nám napadnúť otázka, či štvorcové matice náhodou netvoria grupu – už vieme, že násobenia matíc je asociatívne, chýba nám teda ešte inverzný prvok. K otázke existencie inverznej matice sa dostaneme v ďalšej podkapitole.

Poznámka* 5.4.9. Aj tu by sme mohli postupovať tak, že by sme namiesto matíc porovnávali zobrazenia, ktoré zodpovedajú maticiam vystupujúcim v uvedených rovnostiach. (Môžete si to vyskúšať.) Treba si pritom uvedomiť, že zobrazenie zodpovedajúce súčtu matíc je súčet zobrazení a jednotkovej matici zodpovedá identické zobrazenie.

Pri overovaní distributívnosti sa takýmto spôsobom dostanete ku vzťahom medzi skladaním zobrazení a súčtom zobrazení, ktoré neplatia všeobecne, platia však pre lineárne zobrazenia (čo nám úplne stačí).

uc:POZNfX=XA}

Poznámka 5.4.10. Vektor môžeme chápať ako maticu typu $1 \times m$. Vďaka tomu môže mať zmysel aj násobenie matice a vektora.

Nech $f\colon F^m \to F^n$ je lineárne zobrazenie a $\vec{\alpha} \in F^m$. Potom platí nasledujúca veľmi užitočná rovnosť

$$f(\vec{\alpha}) = \vec{\alpha}A_f.$$

Ak $\vec{\alpha}$ chápeme ako maticu typu $1 \times m$ nad poľom F, tak uvedená rovnosť skutočne má zmysel – vynásobením matíc typu $1 \times m$ a $m \times n$ dostaneme maticu typu $1 \times n$. Túto maticu môžeme chápať ako vektor z F^n.

Platnosť uvedenej rovnosti vyplýva priamo z definície násobenia matíc. Stačí si uvedomiť, že ak riadky matice A označíme ako $\vec{\alpha}_1, \ldots, \vec{\alpha}_m$, tak

$$\vec{\alpha}A = (a_1, \ldots, a_m)\begin{pmatrix}\vec{\alpha}_1\\ \vdots \\ \vec{\alpha}_m\end{pmatrix} = a_1\vec{\alpha}_1 + \cdots + a_m\vec{\alpha}_m =$$
$$a_1 f(\vec{\varepsilon}_1) + \cdots + a_n f(\vec{\varepsilon}_n) = f(a_1\vec{\varepsilon}_1 + \cdots + a_n\vec{\varepsilon}_n) = f(a_1, \ldots, a_n) = f(\vec{\alpha}).$$

Všimnime si, že pomocou predchádzajúceho zápisu dostaneme

$$g(f(\vec{\alpha})) = g(\vec{\alpha}A_f) = \vec{\alpha}(A_f A_g),$$

čiže

$$A_{g\circ f} = A_f A_g.$$

Ak si teda zapamätáme, že lineárne zobrazenie je vlastne násobenie maticou *sprava,* tak si ľahko zapamätáme aj to, že sa poradie skladania zobrazení pri súčine matíc musí meniť.

{suc:POZNXAJELIN}

Poznámka 5.4.11. Z výsledkov v tejto kapitole tiež vidíme, že *každé* lineárne zobrazenie $f\colon F^m \to F^n$ má tvar $f(\vec{x}) = \vec{x}A$ pre nejakú maticu $A \in M_{m,n}(F)$. (Konkrétne to je matica $A = A_f$.)

Napríklad pre lineárne zobrazenia $f\colon F^2 \to F^2$ dostávame, že každé takéto zobrazenie bude mať tvar $f(\vec{\alpha}) = \vec{\alpha}A$ pre nejakú maticu

$$A = \begin{pmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix} \in M_{2,2}(F).$$

To nám dáva predpis

$$f(x, y) = (x, y)\begin{pmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix} = (a_{11}x + a_{21}y, a_{12}x + a_{22}y).$$

Toto je vlastne zovšeobecnenie príkladu 5.3.4.

Súčasne si môžeme všimnúť, že linearitu zobrazenia zadaného takýmto predpisom vieme ľahko overiť aj priamo z definície s využitím vlastností súčinu matíc:

$$f(\vec{\alpha} + \vec{\beta}) = (\vec{\alpha} + \vec{\beta})A = \vec{\alpha}A + \vec{\beta}A = f(\vec{\alpha}) + f(\vec{\beta})$$
$$f(c\vec{\alpha}) = (c\vec{\alpha})A = c(\vec{\alpha}A) = cf(\vec{\alpha})$$

Ak porovnáte, ako sme zapísali dôkaz linearity zobrazenia f teraz (keď už poznáme základné vlastnosti násobenia matíc) a ako sme ho zapisovali v príklade 5.3.4, tak by malo byť jasné, že použitie matíc nám umožnilo zapísať tento dôkaz stručnejšie a tiež je oveľa prehľadnejší.

V súvislosti so súčinom matíc bude pre nás často užitočná aj rovnosť

$$(AB)^T = B^T A^T, \tag{5.1}$$ {suc:EQTRANSP}

ktorej dôkaz je ponechaný ako cvičenie (úloha 5.4.1).

Cvičenia

{succvic:TRANSP}

Úloha 5.4.1. Dokážte:
a) $(AB)^T = B^T A^T$
b) Ak A je symetrická matica, tak aj A^n pre každé $n \in \mathbb{N}$ je symetrická matica.

Úloha 5.4.2. Vypočítajte $A^2+2AB+B^2$, $A^2+2BA+B^2$, $A^2+AB+BA+B^2$, $(A+B)^2$, ak $A = \left(\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right)$ $B = \left(\begin{smallmatrix}1&3\\2&1\end{smallmatrix}\right)$

Úloha 5.4.3. Vyrátajte EA a AE pre $A = \left(\begin{smallmatrix}1&2&3\\1&0&-1\\-1&-2&1\end{smallmatrix}\right)$ a a) $E = \left(\begin{smallmatrix}1&0&1\\0&1&0\\0&0&1\end{smallmatrix}\right)$ b) $E = \left(\begin{smallmatrix}1&0&0\\0&3&0\\0&0&1\end{smallmatrix}\right)$ c) $E = \left(\begin{smallmatrix}0&0&1\\0&1&0\\1&0&0\end{smallmatrix}\right)$ d) $E = \left(\begin{smallmatrix}1&0&-3\\0&1&0\\0&0&1\end{smallmatrix}\right)$. Vedeli by ste nájsť riadkovú/stĺpcovú operáciu, pomocou ktorej dostaneme z matice A maticu EA resp. AE? (Viac sa o súvise násobenia matíc a elementárnych riadkových/stĺpcových operácií môžete dozvedieť v podkapitole 5.6).

{succvic:ULOSTOPA}

Úloha 5.4.4. Pre štvorcovú maticu C typu $n \times n$ budeme výraz $\mathrm{Tr}(C) = \sum_{k=1}^{n} c_{kk} = c_{11} + c_{22} + \cdots + c_{nn}$ nazývať *stopa matice* C.

Ukážte, že ak A, B sú matice typu $n \times n$ nad poľom F, tak platia rovnosti $\mathrm{Tr}(A) = \mathrm{Tr}(A^T)$ a $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$.

Zistite, či pre ľubovoľné matice A, B, C typu $n \times n$ platia vzťahy $\mathrm{Tr}(ABC) = \mathrm{Tr}(CBA)$ a $\mathrm{Tr}(ABC) = \mathrm{Tr}(ACB)$. (Svoje tvrdenie zdôvodnite!) Ak niektorý z týchto vzťahov neplatí, bude platiť za dodatočného predpokladu, matica A je symetrická?

Úloha 5.4.5. Dokážte, alebo vyvráťte nasledujúce tvrdenie: Ak A, B sú štvorcové matice typu $n \times n$ a $A^2 = B^2$, tak $A = B$ alebo $A = -B$.

Úloha 5.4.6. Nech $C = AB$, kde A, B sú matice. Musí potom platiť $V_C \subseteq V_A$? Musí platiť $V_C \subseteq V_B$? Musí platiť $V_A \subseteq V_C$, $V_B \subseteq V_C$? (Svoje tvrdenie zdôvodnite, t.j. dokážte, alebo nájdite kontrapríklad.)

{succvic:ULOHABHA}

Úloha 5.4.7. Nech A, B sú matice nad poľom F typu $m \times n$ resp. $n \times k$. Dokážte, že $h(AB) \leq h(A)$. Dokážte, že ak $n = k$ a B je regulárna, tak $h(AB) = h(A)$.

{succvic:ULOHABHB}

Úloha 5.4.8. Nech A, B sú matice nad poľom F typu $m \times n$ resp. $n \times k$. Dokážte, že $h(AB) \leq h(B)$. Dokážte, že ak $m = n$ a A je regulárna, tak $h(AB) = h(B)$.

5.5 Inverzná matica

Pripomeňme, že zobrazenie $g\colon Y \to X$ nazývame inverzným zobrazením k zobrazeniu $f\colon X \to Y$, ak

$$g \circ f = id_X$$
$$f \circ g = id_Y$$

(definícia 2.2.15) a označujeme ho f^{-1}. Ďalej vieme, že inverzné zobrazenie k zobrazeniu f existuje práve vtedy, keď f je bijekcia (tvrdenie 2.2.16).

{inv:VTINVLZ}

Veta 5.5.1. *Ak $f\colon V \to W$ je lineárne zobrazenie a existuje inverzné zobrazenie $f^{-1}\colon W \to V$, tak f^{-1} je lineárne zobrazenie.*

Dôkaz. Nech $\vec{\alpha}, \vec{\beta} \in W$. Označme $\vec{\alpha}_1 = f^{-1}(\vec{\alpha})$ a $\vec{\beta}_1 = f^{-1}(\vec{\beta})$. To znamená, že $\vec{\alpha}_1$, $\vec{\beta}_1$ sú (jednoznačne určené) vektory z V, pre ktoré platí $f(\vec{\alpha}_1) = \vec{\alpha}$ a $f(\vec{\beta}_1) = \vec{\beta}$. Potom (pre ľubovoľné $c, d \in F$) platí

$$f(c\vec{\alpha}_1 + d\vec{\beta}_1) = cf(\vec{\alpha}_1) + df(\vec{\beta}_1) = c\vec{\alpha} + d\vec{\beta}.$$

Zistili sme, že vektor $c\vec{\alpha}_1 + d\vec{\beta}_1$ sa zobrazením f zobrazí na vektor $c\vec{\alpha} + d\vec{\beta}$, čo znamená

$$f^{-1}(c\vec{\alpha} + d\vec{\beta}) = c\vec{\alpha}_1 + d\vec{\beta}_1.$$

Pretože táto rovnosť platí pre ľubovoľné $c, d \in F$, podľa vety 5.3.5 je zobrazenie f^{-1} lineárne. □

Vieme, že inverzné zobrazenie existuje iba k bijektívnym zobrazeniam. V prípade, že je zobrazenie f lineárne, vieme odvodiť pomerne jednoduché kritérium na zistenie či je to bijekcia. Najprv dokážeme lemu, ktorá charakterizuje injektívne a surjektívne lineárne zobrazenia.

{inv:LMINJSURJ}

Lema 5.5.2. *Nech $f\colon V \to W$ je lineárne zobrazenie a $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ je báza priestoru V.*

{IS.1}

(i) *Zobrazenie f je injekcia práve vtedy, keď vektory $f(\vec{\alpha}_1), \ldots, f(\vec{\alpha}_n)$ sú lineárne nezávislé.*

{IS.2}

(ii) *Zobrazenie f je surjekcia práve vtedy, keď $[f(\vec{\alpha}_1), \ldots, f(\vec{\alpha}_n)] = W$ (teda ak vektory $f(\vec{\alpha}_1), \ldots, f(\vec{\alpha}_n)$ generujú celý priestor W).*

Dôkaz časti (i). $\boxed{\Rightarrow}$ Nech

$$c_1 f(\vec{\alpha}_1) + \ldots + c_n f(\vec{\alpha}_n) = \vec{0}.$$

Z linearity zobrazenia f dostaneme

$$f(c_1\vec{\alpha}_1 + \cdots + c_n\vec{\alpha}_n) = f(\vec{0}).$$

Pretože f je prosté, vyplýva z tejto rovnosti

$$c_1\vec{\alpha}_1 + \cdots + c_n\vec{\alpha}_n = \vec{0}$$

a keďže vektory $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ sú lineárne nezávislé, dostávame $c_1 = \ldots = c_n = 0$. Z rovnosti $c_1 f(\vec{\alpha}_1) + \ldots + c_n f(\vec{\alpha}_n) = \vec{0}$ sme odvodili, že všetky koeficienty vystupujúce v tejto lineárnej kombinácii sú nulové, teda vektory $f(\vec{\alpha}_1), \ldots, f(\vec{\alpha}_n)$ sú skutočne lineárne nezávislé.

$\boxed{\Leftarrow}$ Nech pre nejaké vektory $\vec{\alpha}, \vec{\beta} \in V$ platí $f(\vec{\alpha}) = f(\vec{\beta})$. Vektory $\vec{\alpha}$ a $\vec{\beta}$ vieme vyjadriť pomocou bázových vektorov

$$\vec{\alpha} = c_1\vec{\alpha}_1 + \cdots + c_n\vec{\alpha}_n$$
$$\vec{\beta} = d_1\vec{\alpha}_1 + \cdots + d_n\vec{\alpha}_n$$

Odčítaním týchto 2 rovností dostaneme $\vec{\alpha} - \vec{\beta} = (c_1 - d_1)\vec{\alpha}_1 \ldots + (c_n - d_n)\vec{\alpha}_n$. Ak zobrazíme obe strany tejto rovnosti lineárnym zobrazením f, dostaneme

$$\vec{0} = f(\vec{\alpha}) - f(\vec{\beta}) = f(\vec{\alpha} - \vec{\beta}) = (c_1 - d_1)f(\vec{\alpha}_1) + \ldots + (c_n - d_n)f(\vec{\alpha}_n).$$

Pretože podľa predpokladu sú $f(\vec{\alpha}_1), \ldots, f(\vec{\alpha}_n)$ lineárne nezávislé, vyplýva z toho $c_i - d_i = 0$, čo znamená, že $c_i = d_i$ (pre $i = 1, 2, \ldots, n$) a $\vec{\alpha} = \vec{\beta}$.

Dôkaz časti (ii). $\boxed{\Rightarrow}$ Inklúzia $[f(\vec{\alpha}_1), \ldots, f(\vec{\alpha}_n)] \subseteq W$ je zrejmá, potrebujeme dokázať obrátenú inklúziu. Nech $\vec{\alpha}$ je ľubovoľný vektor z W. Pretože zobrazenie f je surjektívne, existuje vektor $\vec{\beta} \in V$ taký, že $f(\vec{\beta}) = \vec{\alpha}$. Vektor $\vec{\beta}$ sa dá vyjadriť ako lineárna kombinácia bázových vektorov

$$\vec{\beta} = c_1\vec{\alpha}_1 + \cdots + c_n\vec{\alpha}_n.$$

Z toho dostaneme

$$\vec{\alpha} = f(\vec{\beta}) = c_1 f(\vec{\alpha}_1) + \ldots + c_n f(\vec{\alpha}_n),$$

čo znamená, že $\vec{\alpha} \in [f(\vec{\alpha}_1), \ldots, f(\vec{\alpha}_n)]$.

$\boxed{\Leftarrow}$ Nech $\vec{\gamma} \in W = [f(\vec{\alpha}_1), \ldots, f(\vec{\alpha}_n)]$. Potom existujú skaláry $c_1, \ldots, c_n \in F$ také, že

$$\gamma = c_1 f(\vec{\alpha}_1) + \ldots + c_n f(\vec{\alpha}_n) = f(c_1\vec{\alpha}_1 + \cdots + c_n\vec{\alpha}_n).$$

Našli sme vzor pre ľubovoľný vektor $\vec{\gamma} \in W$, čo znamená, že zobrazenie f je surjektívne. □

Z predchádzajúcej lemy priamo vyplýva

{inv:VTBIJLZ}

Veta 5.5.3. *Nech $f\colon V \to W$ je lineárne zobrazenie a $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ je báza priestoru V. Zobrazenie f je bijekcia práve vtedy, keď vektory $f(\vec{\alpha}_1), \ldots, f(\vec{\alpha}_n)$ tvoria bázu vektorového priestoru W.*

{inv:DOSINJ}

Dôsledok 5.5.4. *Nech $f\colon F^n \to F^n$ je lineárne zobrazenie. Nasledujúce podmienky sú ekvivalentné:*

(i) *f je bijekcia,*

(ii) *f je prosté,*

(iii) *f je surjektívne.*

Dôkaz. V priestore s dimenziou n je $f(\vec{\alpha}_1), \ldots, f(\vec{\alpha}_n)$ báza $\Leftrightarrow$ tieto vektory sú lineárne nezávislé $\Leftrightarrow$ $[f(\vec{\alpha}_1), \ldots, f(\vec{\alpha}_n)] = F^n$ (veta 4.4.14). □

{inv:VTDOSBIJ}

Dôsledok 5.5.5. *Nech $f\colon F^n \to F^n$ je lineárne zobrazenie. Nasledujúce podmienky sú ekvivalentné:*

(a) *zobrazenie f je bijekcia,*

(b) *existuje inverzné zobrazenie f^{-1},*

(c) *$h(A_f) = n$.*

Dôkaz. Stačí si uvedomiť, že hodnosť matice je dimenzia priestoru prislúchajúceho tejto matici, čo je v našom prípade celý priestor F^n.

Obrátene, ak je hodnosť matice A_f rovná n, tak jej riadky tvoria n lineárne nezávislých vektorov v priestore dimenzie n, sú teda bázou celého priestoru. Riadky matice A_f sú však práve obrazy vektorov štandardnej bázy. □

Definícia 5.5.6. Nech A je matica typu $n \times n$. Hovoríme, že matica B je *inverzná* k matici A, ak platí

$$AB = BA = I_n.$$

Označujeme ju $B =: A^{-1}$.

Poznámka 5.5.7. Predchádzajúca definícia vlastne hovorí, že $f_A \circ f_B = f_B \circ f_A = id$, čiže inverzná matica je práve matica zodpovedajúca inverznému zobrazeniu. Práve tento fakt využijeme pri výpočte inverznej matice – budeme postupovať takým spôsobom, že vypočítame maticu inverzného zobrazenia postup z úlohy 5.3.1. (Čiže začneme s maticou $(A|I)$ a upravujeme ju kým nedostaneme maticu $(I|A^{-1})$.)

NVJEDNASTACI}

Poznámka 5.5.8. Je užitočné si uvedomiť, že ak pre *štvorcové* matice A, B rovnakých rozmerov platí niektorá z rovností $AB = I$ alebo $BA = I$, tak už B musí byť inverzná matica k A.

Aby sme to overili, preložme tieto rovnosti do rečí skladania zobrazení. Dostaneme

$$f_B \circ f_A = id$$
$$f_A \circ f_B = id$$

Využijeme úlohy 2.2.1 a 2.2.2. V prvom prípade vidíme, že f_A je injekcia (lebo zložené zobrazenie je injekcia) a podľa dôsledku 5.5.4 je potom f_A bijekcia. Podobne, v druhom prípade dostaneme, že f_A je surjekcia, čiže aj bijekcia.

Preto (v oboch prípadoch) má matica A inverznú maticu A^{-1}. Ak ňou vynásobíme rovnosť $AB = I$ zľava, dostaneme $B = A^{-1}$. Ak predpokladáme platnosť rovnosti $BA = I$, môžeme ju vynásobiť A^{-1} sprava a opäť máme $B = A^{-1}$.

V oboch prípadoch sme dostali rovnosť $B = A^{-1}$, čiže pre matice $n \times n$ stačí overiť jednu z uvedených rovností. (Môže však existovať matica typu $m \times n$ pre $m \neq n$ taká, že $BA = I$, tá samozrejme nie je inverznou maticou k A.)

Definícia 5.5.9. Štvorcová matica typu $n \times n$ sa nazýva *regulárna,* ak $h(A) = n$.

Z dôsledku 5.5.5 vyplýva nasledujúca veta.

{inv:VTREG}

Veta 5.5.10. *Nech A je matica typu $n \times n$. K matici A existuje inverzná matica práve vtedy, keď A je regulárna.*

Môžeme si všimnúť, že inverznú maticu k súčinu vypočítame ako súčin inverzných matíc – treba si však dať pozor na poradie. (Podobne to bolo s transponovanými maticami – úloha 5.4.1.) Toto tvrdenie je analogické k výsledkom o $(g \circ f)^{-1}$ pri skladaní zobrazení (tvrdenie 5.4.1) a tiež o $(x * y)^{-1}$ v grupe (veta 3.2.6).

{inv:VTINVAB}

Tvrdenie 5.5.11. *Nech $A, B \in M_{n,n}(F)$ sú regulárne matice. Potom platí:*

(i) $(A^{-1})^{-1} = A$,

(ii) $(A \cdot B)^{-1} = B^{-1} \cdot A^{-1}$

(iii) $(A^T)^{-1} = (A^{-1})^T$

Dôkaz. Úloha 5.5.1. □

Definícia 5.5.12. Bijektívne lineárne zobrazenie $f\colon V \to W$ nazývame *izomorfimus vektorových priestorov* V a W (alebo tiež *lineárny izomorfizmus*).

Ak existuje bijektívne zobrazenie $f\colon V \to W$, hovoríme, že vektorové priestory V a W sú izomorfné. Fakt, že V a W sú izomorfné označujeme $V \cong W$.

Dôsledok 5.5.13. *Ak V, W sú konečnorozmerné vektorové priestory a $V \cong W$, tak $d(V) = d(W)$.*

Dôkaz. Ak V a W sú izomorfné, tak existuje bijekcia $f\colon V \to W$, ktorá je súčasne lineárnym zobrazením.

Označme $d(V) = n$. Potom V má n-prvkovú bázu $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$. Podľa vety 5.5.3 je potom $f(\vec{\alpha}_1), \ldots, f(\vec{\alpha}_n)$ bázou vo W, čo znamená, že $d(W) = n$. □

Poznámka 5.5.14. Vieme, že bijektívne zobrazenie je jedno-jednoznačné priradenie medzi prvkami množiny V a prvkami množiny W. Ak je toto zobrazeníe navyše lineárne, znamená to, že táto jedno-jednoznačná korešpondencia navyše rešpektuje operácie definované na vektorových priestoroch V a W.

Fakt, že 2 vektorové priestory sú izomorfné, teda znamená, že sú v podstate rovnaké, len ich prvky sú inak označené (pomenované). Izomorfizmus poskytuje „preklad“ medzi týmito dvoma pomenovaniami.

Nasledujúca veta hovorí, že každý priestor dimenzie n je izomorfný priestoru F^n. (A teda každý konečnorozmerný priestor je izomorfný s F^n pre niektoré n.)

{inv:VTKANONFN}

Veta 5.5.15. *Nech V je vektorový priestor nad poľom F a $d(V) = n$. Potom V je izomorfný s priestorom F^n.*

Dôkaz. Ak $d(V) = n$, znamená to, že V má n-prvkovú bázu $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$. Podľa vety 5.3.7 existuje jediné lineárne zobrazenie $f\colon V \to F^n$ s vlastnosťou $f(\vec{\alpha}_1) = \vec{\varepsilon}_1, \ldots, f(\vec{\alpha}_n) = \vec{\varepsilon}_n$. Podľa vety 5.5.3 je toto zobrazenie bijekcia, čiže je to izomorfizmus medzi V a F^n. □

Cvičenia

{invcvic:ULOINVAB}

Úloha 5.5.1. Dokážte, že ak A, B sú regulárne štvorcové matice rovnakých rozmerov nad tým istým poľom F, tak platí:
a) $(A^{-1})^{-1} = A$,
b) $(A \cdot B)^{-1} = B^{-1} \cdot A^{-1}$ c) $(A^T)^{-1} = (A^{-1})^T$

{invcvic:INVERZ}

Úloha 5.5.2. Nájdite inverznú maticu k daným maticiam nad $\mathbb{R}$:
$\begin{pmatrix} 2 & 3 & 1 \\ 4 & 3 & 3 \\ 1 & 2 & 1 \end{pmatrix}$ $\begin{pmatrix} 1 & 3 & 1 \\ 2 & 4 & 3 \\ 3 & 1 & 2 \end{pmatrix}$ $\begin{pmatrix} 0 & 3 & 1 \\ 2 & 0 & 3 \\ 1 & 1 & 2 \end{pmatrix}$ $\begin{pmatrix} 3 & 0 & 1 \\ 0 & 3 & 2 \\ 1 & 0 & 1 \end{pmatrix}$ $\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 2 & -1 & 3 \end{pmatrix}$ $\begin{pmatrix} 3 & 2 & 1 \\ 4 & 2 & 1 \\ 1 & 0 & 1 \end{pmatrix}$ $\begin{pmatrix} 1 & \sqrt{2} & \sqrt{6} \\ 0 & 1 & \sqrt{3} \\ 0 & 0 & 1 \end{pmatrix}$ Výsledky:
$\begin{pmatrix} \frac{3}{4} & \frac{1}{4} & -\frac{3}{2} \\ \frac{1}{4} & -\frac{1}{4} & \frac{1}{2} \\ -\frac{5}{4} & \frac{1}{4} & \frac{3}{2} \end{pmatrix}$ $\begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{10} & -\frac{1}{10} \\ -1 & \frac{4}{5} & -\frac{1}{5} \end{pmatrix}$ $\begin{pmatrix} 3 & 5 & -9 \\ 1 & 1 & -2 \\ -2 & -3 & 6 \end{pmatrix}$ $\begin{pmatrix} \frac{1}{2} & 0 & -\frac{1}{2} \\ \frac{1}{3} & \frac{1}{3} & -1 \\ -\frac{1}{2} & 0 & \frac{3}{2} \end{pmatrix}$ $\begin{pmatrix} 2 & -\frac{1}{2} & -\frac{1}{2} \\ 1 & \frac{1}{2} & -\frac{1}{2} \\ -1 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$ $\begin{pmatrix} -1 & 1 & 0 \\ \frac{3}{2} & -1 & -\frac{1}{2} \\ 1 & -1 & -1 \end{pmatrix}$ $\begin{pmatrix} 1 & -\sqrt{2} & 0 \\ 0 & 1 & -\sqrt{3} \\ 0 & 0 & 1 \end{pmatrix}$

Úloha 5.5.3. Zistite, či je zadaná matica nad poľom $\mathbb{Z}_5$ regulárna; ak áno, nájdite inverznú:
$\begin{pmatrix} 3 & 1 & 1 & 3 \\ 1 & 4 & 0 & 1 \\ 2 & 1 & 1 & 4 \\ 3 & 2 & 2 & 1 \end{pmatrix}$, $\begin{pmatrix} 2 & 1 & 3 & 1 \\ 0 & 1 & 3 & 4 \\ 3 & 1 & 2 & 1 \\ 1 & 1 & 4 & 2 \end{pmatrix}$, $\begin{pmatrix} 2 & 1 & 3 & 1 \\ 1 & 2 & 4 & 3 \\ 3 & 1 & 2 & 3 \\ 1 & 3 & 1 & 3 \end{pmatrix}$, $\begin{pmatrix} 2 & 1 & 3 & 1 \\ 1 & 2 & 3 & 3 \\ 3 & 1 & 4 & 3 \\ 1 & 3 & 2 & 4 \end{pmatrix}$,
Výsledky: $\begin{pmatrix} 4 & 0 & 0 & 3 \\ 2 & 4 & 1 & 1 \\ 3 & 1 & 1 & 1 \\ 3 & 0 & 1 & 3 \end{pmatrix}$, $\begin{pmatrix} 0 & 2 & 4 & 4 \\ 0 & 0 & 2 & 4 \\ 1 & 2 & 3 & 4 \\ 3 & 0 & 1 & 1 \end{pmatrix}$, $\begin{pmatrix} 0 & 4 & 3 & 3 \\ 3 & 4 & 0 & 0 \\ 1 & 0 & 0 & 3 \\ 0 & 3 & 4 & 0 \end{pmatrix}$, $\begin{pmatrix} 0 & 0 & 1 & 3 \\ 4 & 0 & 3 & 3 \\ 4 & 4 & 2 & 2 \\ 0 & 3 & 4 & 0 \end{pmatrix}$

Úloha 5.5.4. Nech $f\colon (\mathbb{Z}_5)^4 \to (\mathbb{Z}_5)^4$ je lineárne zobrazenie také, že $f(1,2,3,1) = (2,0,1,0)$, $f(0,2,3,1) = (1,2,0,3)$, $f(1,0,3,4) = (3,2,1,0)$, $f(4,1,3,2) = (2,3,1,1)$. Nájdite maticu zobrazenia f^{-1}.

Úloha 5.5.5. Zistite, či $\begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$ je regulárna a) nad $\mathbb{Z}_2$ b) nad $\mathbb{Z}_3$, ak áno, nájdite inverznú.

Úloha 5.5.6. Koľko existuje lineárnych zobrazení spĺňajúcich zadané podmienky? Koľko z nich je injektívnych? Koľko je surjektívnych?
a) $f\colon \mathbb{Z}_5^3 \to \mathbb{Z}_5^4$, $f(1,3,1) = (1,1,1,3)$, $f(2,1,3) = (0,1,3,4)$, $f(2,1,0) = (1,4,0,0)$;
b) $f\colon \mathbb{Z}_5^4 \to \mathbb{Z}_5^3$, $f(1,0,3,1) = (0,1,3)$, $f(2,1,3,1) = (1,1,3)$, $f(1,1,4,1) = (2,2,1)$;
c) $f\colon \mathbb{Z}_5^4 \to \mathbb{Z}_5^3$, $f(1,0,3,1) = (0,1,3)$, $f(2,1,3,1) = (1,1,3)$, $f(1,1,0,0) = (1,0,0)$;

Úloha 5.5.7. Dokážte, alebo nájdite kontrapríklad:
a) Nech A, B, C sú štvorcové matice typu $n \times n$ a platí $AB = AC$. Potom $B = C$.
b) Nech A, B, C sú štvorcové matice typu $n \times n$ a platí $AB = AC$. Nech navyše A je regulárna matica. Potom $B = C$.

Úloha 5.5.8. Zistite, či pre dané matice A, B nad poľom $\mathbb{Z}_5$ existuje matica X nad tým istým poľom taká, že $AX = B$. Zistite, či je X maticami A, B jednoznačne určená. Ak taká matica existuje, tak aspoň jednu takú maticu nájdite.

a) $A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 3 & 1 \\ 3 & 2 & 3 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 1 & 3 \\ 2 & 0 & 1 \\ 0 & 2 & 3 \end{pmatrix}$

b) $A = \begin{pmatrix} 0 & 1 & 3 \\ 2 & 1 & 0 \\ 1 & 3 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 3 & 1 \\ 1 & 3 & 4 \end{pmatrix}$

c) $A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 2 & 3 \\ 2 & 3 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 3 & 3 \\ 1 & 4 & 0 \\ 3 & 0 & 3 \end{pmatrix}$

d) $A = \begin{pmatrix} 1 & 2 & 0 \\ 3 & 1 & 0 \\ 1 & 4 & 4 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 4 & 2 \\ 0 & 2 & 1 \\ 4 & 1 & 3 \end{pmatrix}$.

e) $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 3 \\ 2 & 4 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 3 & 1 \\ 3 & 2 & 3 \end{pmatrix}$

f) $A = \begin{pmatrix} 1 & 1 & 3 \\ 0 & 1 & 1 \\ 2 & 3 & 2 \end{pmatrix}$, $B = \begin{pmatrix} 2 & 1 & 2 \\ 1 & 1 & 4 \\ 3 & 1 & 0 \end{pmatrix}$

Úloha 5.5.9. Vypočítajte $A^{-1}B$ a $B^{-1}A$. Skúste to urobiť bez výpočtu A^{-1} resp. B^{-1}.
$A = \left(\begin{smallmatrix} 0 & 3 & 1 \\ 2 & 0 & 3 \\ 1 & 1 & 2 \end{smallmatrix}\right)$ $B = \left(\begin{smallmatrix} 1 & 2 & 1 \\ 0 & 1 & -1 \\ 0 & -1 & 2 \end{smallmatrix}\right)$
Ako skúšku správnosti môžete vyskúšať, či po vynásobení výsledku zľava maticou A (resp. B) dostanete maticu B (resp. A).

5.6 Elementárne riadkové operácie a súčin matíc

{rtm2:SECT}

V tejto časti si povieme, ako môžeme elementárne riadkové (stĺpcové) operácie vyjadriť pomocou násobenia matíc.

Definícia 5.6.1. Pre ľubovoľnú elementárnu riadkovú operáciu na matici typu $m \times n$ nazveme *maticou elementárnej riadkovej operácie* maticu typu $m \times m$, ktorá vznikne vykonaním tejto operácie na jednotkovej matici I_m.

Podobne môžeme definovať maticu sĺpcovej operácie.

Príklad 5.6.2. Uvažujme matice s 3 riadkami. Potom výmene prvého a tretieho riadku zodpovedá matica E_1, pripočítaniu dvojnásobku druhého riadku k prvému matica E_2 a vynásobeniu prvého riadku číslom 3 matica E_3.

$$E_1 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \qquad E_2 = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad E_3 = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Vedeli by ste nájsť stĺpcové operácie, ktorým by zodpovedali tieto 3 matice? (Odpoveď: Výmena prvého a tretieho stĺpca, pripočítanie 2-násobku prvého stĺpca k druhému a vynásobenie prvého stĺpca číslom 3.)

Tvrdenie 5.6.3. *Ak matica B vznikne z matice A vykonaním nejakej elementárnej riadkovej operácie a E je matica tejto riadkovej operácie, tak $B = EA$.*

Príklad 5.6.4. Ak $A = \begin{pmatrix} 1 & 0 \\ 1 & 2 \\ -1 & -1 \end{pmatrix}$ a E_2 je matica z predchádzajúceho príkladu, tak $E_2A = \begin{pmatrix} 3 & 4 \\ 1 & 2 \\ -1 & -1 \end{pmatrix}$ je skutočne matica, ktorá vznikne z A pripočítaním dvojnásobku druhého riadku k prvému.

Dôkaz. Tvrdenie overíme jednoducho priamym výpočtom. Budeme postupovať pre každý typ elementárnej riadkovej operácie zvlášť.

Nech $A = ||a_{ij}||$ je matica typu $m \times n$.

Výmene k-teho a l-teho riadku zodpovedá matica $E = ||e_{ij}||$, ktorej prvky sú len nuly a jednotky, pričom jednotky sú iba na pozíciach e_{ii} pre $i \neq k, l$ a tiež e_{kl} a e_{lk}. Vidíme, že v súčine EA budú všetky riadky okrem k-teho a l-teho rovnaké ako v matici A. (Vo výraze $\sum_{t=1}^{m} e_{it}a_{tj}$ jediný nenulový sčítanec je $e_{ii}a_{ij} = a_{ij}$.) Podobne v k-tom riadku dostaneme prvky l-teho riadku a obrátene.

$$\sum_{t=1}^{m} e_{kt}a_{tj} = e_{kl}a_{lj} = a_{lj},$$
$$\sum_{t=1}^{m} e_{lt}a_{tj} = e_{lk}a_{kj} = a_{kj}.$$

Matica EA je teda skutočne matica, ktorá vznikne výmenou týchto 2 riadkov.

Vynásobeniu k-teho riadku skalárom c zodpovedá matica E, ktorá má mimo diagonály nuly a na diagonále jednotky s výnimkou prvkov $e_{kk} = c$. Opäť sa nezmenia ostatné riadky a v k-tom riadku dostávame

$$\sum_{t=1}^{m} e_{kt}a_{tj} = e_{kk}a_{kj} = ca_{kj}.$$

Teda k-ty riadok novej matice je skutočne c-násobok k-teho riadku pôvodnej matice.

Teraz uvažujme pripočítanie c-násobku l-teho riadku ku k-temu. V tomto prípade má matica E na diagonále jednotky a mimo diagonály je jediný nenulový prvok $e_{kl} = c$. Opäť vidno, že mimo k-teho riadku sa prvky nezmenia. V k-tom riadku dostaneme

$$\sum_{t=1}^{m} e_{kt}a_{tj} = e_{kk}a_{kj} + e_{kl}a_{lj} = a_{kj} + ca_{lj}.$$

Teda k-ty riadok matice EA je skutočne súčet k-teho riadku matice A a c-násobku l-tého riadku matice A. □

Úplne analogicky sa dá dokázať podobné tvrdenie pre stĺpcové operácie.

Tvrdenie 5.6.5. *Ak matica B vznikne z matice A vykonaním nejakej elementárnej stĺpcovej operácie a E je matica tejto stĺpcovej operácie, tak $B = AE$.*

Užitočné je si všimnúť, že matica ľubovoľnej elementárnej riadkovej operácie je regulárna a inverzná matica k nej je tiež matica elementárnej riadkovej operácie. (Dalo by sa povedať, že je to matica „inverznej“ riadkovej operácie – pozri poznámku 5.2.6. Práve fakt, že elementárne riadkové operácie sú invertovateľné môžeme použiť ako jednu z možností na zdôvodnenie, že matice elementárnych riadkových operácií sú regulárne.)

{rtm2:ROZKLADELEM}

Poznámka 5.6.6. Ak A je ľubovoľná matica, tak pomocou elementárnych riadkových operácií z nej vieme dostať redukovanú trojuholníkovú maticu

$$R = E_1 E_2 \dots E_k A.$$

Potom platí

$$A = E_k^{-1} \dots E_2^{-1} E_1^{-1} R.$$

Takto sme maticu A vyjadrili ako súčin pomerne jednoduchých matíc (jedna z nich je v redukovanom trojuholníkovom tvare, ostatné sú matice elementárnych riadkových operácií). To môže byť užitočné v dôkazoch niektorých tvrdení – hlavne v prípade, že dokazované tvrdenie vieme ľahko dokázať pre matice takéhoto tvaru a tiež vieme dokázať, že platnosť tvrdenia sa zachová ak prejdeme k súčinu matíc.

Ilustráciou tohoto prístupu je napríklad alternatívny dôkaz vety 5.7.2.

Pomocou tohoto vzťahu medzi násobením matíc a môžeme lepšie porozumieť spôsobu, akým sme počítali inverzné matice.

Začali sme s maticou

$$(A|I)$$

a v každom kroku sme urobili nejakú riadkovú operáciu, ktorá zodpovedá vynásobeniu oboch častí matice zľava nejakou maticou riadkovej operácie.

$$(A|I) \sim (E_1 A|E_1 I) \sim (E_2 E_1 A|E_2 E_1 I) \sim \dots \sim (E_n \dots E_2 E_1 A|E_n \dots E_2 E_1 I) = (I|E),$$

kde ako E sme označili maticu $E := E_n \dots E_2 E_1$. Pretože táto matica spĺňa rovnosť $EA = I$ (túto rovnosť vidíme z ľavej časti matice), je to inverzná matica k A. Tiež si môžeme všimnúť, že v každom kroku dostaneme maticu tvaru $(DA|D)$, t.j. ak vynásobíme pravú časť sprava maticou A, dostaneme ľavú časť matice.

Podobne sa dá pozerať aj na riešenie sústav lineárnych rovníc, ktorými sa budeme zaoberať v nasledujúcej kapitole.

Takisto pri výpočte matice zobrazenia sme mali zadaných viacero podmienok tvaru $\vec{\beta}_i A = \vec{\gamma}_i$. Ak poukladáme vektory $\vec{\beta}_i$ do matice B ako riadky a podobne z vektorov $\vec{\gamma}_i$ vytvoríme maticu C, tieto podmienky môžeme zapísať ako jedinú maticovú rovnosť

$$BA = C.$$

Pri výpočte sme postupovali tak, že sme maticu C zľava násobili nejakými maticami riadkových operácií, a to konkrétne takými, že z matice B vytvoria jednotkovú maticu, čiže ich súčin je matica I.

To znamená, že $A = B^{-1}C$ (ak B je regulárna) alebo presnejšie, $A = B'C$, kde B' je taká matica, že $B'B = I$.

Poznámka 5.6.7. Ďalší užitočný fakt, ktorý by nám mal byť po prečítaní tejto kapitoly jasný, je, že násobenie maticou A zľava zodpovedá vytvoreniu lineárnych kombinácii riadkov v matici B (riadky matice A určujú koeficienty).

Podobne, ak maticu B násobíme maticou A sprava, tak stĺpce v BA budú lineárne kombinácie stĺpcov B s koeficientmi určenými stĺpcami matice A.

(Videli sme, že niečo takéto platí pre matice riadkových operácií. Priamo z definície násobenia matíc sa dá overiť, že to platí aj pre ľubovoľné matice.)

{rtm2cvic:ULOEROREG}

Cvičenia

Úloha 5.6.1*. Nech $A, B \in M_{m,n}(F)$. Dokážte: Matice A, B sú riadkovo ekvivalentné práve vtedy, keď existuje regulárna matica $R \in M_{m,m}(F)$ taká, že $B = RA$.

tm2cvic:ULOREGSUCINELEM}

Úloha 5.6.2*. Nech A je štvorcová matica (nad nejakým poľom F). Dokážte: Matica A je regulárna práve vtedy, keď A sa dá dostať ako súčin nejakých matíc elementárnych riadkových operácií. (T.j. $A = E_1 E_2 \dots E_k$, kde každá z matíc $E_1, E_2, \dots, E_k$ zodpovedá nejakej ERO.)

5.7 Sústavy lineárnych rovníc

{suct:SECTSUST}

Definícia 5.7.1. *Sústavou lineárnych rovníc* rozumieme systém rovníc tvaru

{sust:EQSUST}

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= c_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= c_2 \\ &\dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &= c_m \end{aligned} \tag{5.2}$$

kde $a_{ij}, c_i \in F$ pre všetky prípustné hodnoty indexov i a j.

Riešenie sústavy lineárnych rovníc je n-tica $(x_1, \dots, x_n)$ ktorá spĺňa všetky uvedené rovnice. Ak existuje aspoň jedno riešenie sústavy lineárnych rovníc, hovoríme, že táto sústava je *riešiteľná.* Skaláry $c_1, \dots, c_n$ nazývame *pravé strany,* a_{ij} sú *koeficienty* a x_i sú neznáme.

Maticu

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

nazývame *matica sústavy* (5.2).

Maticu

$$A' = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & c_1 \\ a_{21} & a_{22} & \dots & a_{2n} & c_2 \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} & c_m \end{pmatrix}$$

nazývame *rozšírená matica sústavy* (5.2).

Pomocou matice sústavy môžeme zadefinovať *maticový zápis* sústavy

$$A\vec{x}^T = \vec{c}^T$$

alebo

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{pmatrix}$$

Skutočne, $(x_1, \ldots, x_n)$ je riešením sústavy (5.2) práve vtedy, keď platí uvedená maticová rovnosť.

{sust:VTRIADEKV}

Veta 5.7.2. *Ak rozšírené matice dvoch sústav lineárnych rovníc sú riadkovo ekvivalentné, tak tieto dve sústavy majú rovnakú množinu riešení.*

Dôkaz. Predpokladajme, že z rozšírenej matice sústavy $(A|c)$ sme dostali nejakou elementárnou riadkovou operáciou maticu $(B|d)$. Vďaka tomu, že elementárne riadkové operácie sú invertibilné (t.j. možno ich obrátiť, pozri poznámku 5.2.6) stačí nám dokázať, že každé riešenie sústavy $(A|c)$ je aj riešením sústavy $(B|d)$.

Výmena riadkov je vlastne výmena 2 rovníc, čo samozrejme neovplyvní, či $(x_1, \ldots, x_n)$ je riešením sústavy.

Ak prenásobíme i-ty riadok skalárom $c \neq 0$, znamená to, že z rovnice $a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n = c_i$ dostaneme rovnicu $ca_{i1}x_1 + ca_{i2}x_2 + \ldots + ca_{in}x_n = cc_i$. Pretože druhá rovnica je c-násobkom prvej, je jasné, že ak $x_1, x_2, \ldots, x_n$ spĺňa prvú uvedenú rovnicu, musí spĺňať aj druhú z nich.

Zostáva nám posledný typ elementárnych riadkových úprav. Predpokladajme, že sme novú maticu získali pripočítanim c-násobku j-teho riadka k i-temu riadku. To znamená, že i-ta rovnica sústavy sa zmenila na rovnicu

$$(a_{i1} + ca_{j1})x_1 + (a_{i2} + ca_{j2})x_2 + \ldots + (a_{in} + ca_{jn})x_n = c_i + cc_j.$$

Ak však $x_1, \ldots, x_n$ spĺňa rovnosť $a_{j1}x_1 + a_{j2}x_2 + \ldots + a_{jn}x_n = c_j$, tak spĺňa aj jej c-násobok $ca_{j1}x_1 + ca_{j2}x_2 + \ldots + ca_{jn}x_n = cc_j$. Sčítaním tejto rovnice a rovnice $a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n = c_i$ dostaneme práve i-ty riadok novej sústavy. Pretože $x_1, \ldots, x_n$ spĺňa obe uvedené rovnice, musí spĺňať aj rovnicu, ktorú dostaneme ich sčítaním. □

Pomocou toho, čo sme sa dozvedeli o súvise elementárnych riadkových úprav s násobením matíc v časti 5.6 môžeme dokázať tú istú vetu aj iným spôsobom, ktorý je snáď do istej miery jasnejší (alebo prinajmenšom stručnejšie zapísaný – lebo hovorí to isté, čo predchádzajúci dôkaz, iba používa trochu iný pohľad na elementárne riadkové operácie).

Dôkaz. Označme maticu uvažovanej elementárnej riadkovej operácie E. Uvedomme si najprv, že urobiť elementárnu riadkovú úpravu na rozšírenej matici sústavy je presne to isté, ako keby sme túto úpravu urobili zvlášť na matici A a zvlášť na stĺpcovom vektore $\vec{c}^T$. To znamená, že ak pôvodná sústava bola (pri zápise v maticovom tvare) $A\vec{x}^T = \vec{c}^T$, tak po úprave dostaneme.

$$EA\vec{x}^T = E\vec{c}^T.$$

Z toho vidíme, že každé riešenie pôvodnej sústavy je aj riešením sústavy s upravenou maticou (obe strany rovnosti sme vynásobili zľava tou istou maticou E, tým sme nezmenili platnosť rovnosti).

Obrátene, ak pre $\vec{x}$ platí $EA\vec{x}^T = E\vec{c}^T$, tak prenásobením tejto rovnosti maticou E^{-1} zľava dostaneme, že $\vec{x}$ je aj riešením pôvodnej sústavy. □

5.7.1 Homogénne sústavy lineárnych rovníc

V prípade, že pravé strany sú nulové ($c_1 = c_2 = \ldots = c_n = 0$), nazývame sústavu (5.2) *homogénna* sústava lineárnych rovníc. Ľahko si môžeme všimnúť, že v prípade homogénnej sústavy je nulový vektor $(0, 0, \ldots, 0)$ riešením sústavy. Toto riešenie nazývame *triviálne riešenie.*

{sust:VTHOMPPR}

Veta 5.7.3. *Množina všetkých riešení homogénnej sústavy lineárnych rovníc tvorí podpriestor priestoru F^n.*

Dôkaz. Stačí overiť vlastnosti z definície podpriestoru.

Ak $\vec{\alpha}$ a $\vec{\beta}$ sú riešeniami homogénnej sústavy s maticou A, znamená to, že $A.\vec{\alpha}^T = \vec{0}^T$ a $A.\vec{\beta}^T = \vec{0}^T$.

Sčítaním týchto rovností dostaneme $A.(\vec{\alpha} + \vec{\beta})^T = \vec{0}^T$, teda aj $\vec{\alpha} + \vec{\beta}$ je riešením tejto sústavy. Ak prvú rovnosť vynásobíme skalárom F, máme $A.(c\vec{\alpha})^T = \vec{0}^T$, čo znamená, že aj $c.\vec{\alpha}$ je riešením sústavy. □

Rozšírenú maticu sústavy lineárnych rovníc môžeme teda upraviť na redukovanú trojuholníkovú maticu. Predpokladajme, že sme navyše preusporiadali premenné (čo vlastne zodpovedá permutácii niektorých stĺpcov) tak, aby vo výslednej matici boli ako prvé tie stĺpce, kde vystupujú vedúce jednotky. Navyše môžeme vynechať všetky nulové riadky bez toho, aby sme nejako ovplyvnili množinu riešení. Dostaneme takto maticu, ktorej zodpovedá sústava

{sust:EQHOM}

$$\begin{aligned} x_1 + c_{1,r+1}x_{r+1} + c_{1,r+2}x_{r+2} + \ldots + c_{1,n}x_n &= 0 \\ x_2 + c_{2,r+1}x_{r+1} + c_{2,r+2}x_{r+2} + \ldots + c_{2,n}x_n &= 0 \\ &\ldots \\ x_r + c_{r,r+1}x_{r+1} + c_{r,r+2}x_{r+2} + \ldots + c_{r,n}x_n &= 0 \end{aligned} \tag{5.3}$$

pričom r označuje hodnosť pôvodnej matice (a teda aj matice C).

Vidíme, že ak si zvolíme hodnotu neznámych $x_{r+1}, x_{r+2} \ldots, x_n$, dá sa z týchto rovníc dorátať hodnota neznámych $x_1, x_2, \ldots, x_n$. Ak postupne dosadíme 1 za x_{r+k} a 0 za ostatné neznáme, ktoré si môžeme voliť (pre $k = 1, 2, \ldots, n - r$) dostaneme tieto riešenia sústavy (5.3):

{sust:EQBAZARIES}

$$\begin{aligned} \vec{\gamma}_{r+1} &= (-c_{1,r+1}, -c_{2,r+1}, \ldots, -c_{r,r+1}, 1, 0, \ldots, 0), \\ \vec{\gamma}_{r+2} &= (-c_{1,r+2}, -c_{2,r+2}, \ldots, -c_{r,r+2}, 0, 1, \ldots, 0), \\ &\ldots \\ \vec{\gamma}_n &= (-c_{1,n}, -c_{2,n}, \ldots, -c_{r,n}, 0, \ldots, 0, 1). \end{aligned} \tag{5.4}$$

{sust:VTBAZARIES}

Veta 5.7.4. *Vektory $\vec{\gamma}_{r+1}, \vec{\gamma}_{r+2}, \ldots, \vec{\gamma}_n$ tvoria bázu priestoru riešení homogénnej sústavy* (5.3).

Dôkaz. Z toho, ako sme ich získali, vieme, že tieto vektory sú riešenia (5.3).

Ich lineárnu nezávislosť overíme tiež pomerne jednoducho: ak

$$\vec{\alpha} = d_{r+1}\vec{\gamma}_{r+1} + d_{r+2}\vec{\gamma}_{r+2} + \cdots + d_n\vec{\gamma}_n = \vec{0}$$

tak vektor $\vec{\alpha}$ má na $(r + k)$-tej súradnici hodnotu d_{r+k}, z čoho dostaneme $d_{r+k} = 0$ pre $k = 1, 2, \ldots, n - r$.

Zostáva dokázať, že vektory $\vec{\gamma}_{r+1}, \vec{\gamma}_{r+2}, \ldots, \vec{\gamma}_n$ generujú celý priestor riešení, teda že každé riešenie homogénnej sústavy (5.3) možno získať ako ich lineárnu kombináciu.

Nech teda $b_1, b_2, \ldots, b_n$ je riešením (5.3). Z toho dostaneme

$$b_i = -c_{i,r+1}b_{r+1} - c_{i,r+2}b_{r+2} - \ldots - c_{i,n}b_n$$

pre $i = 1, 2, \ldots, r$. Z týchto rovností priamo vyplýva

$$\vec{\beta} = b_{r+1}\vec{\gamma}_{r+1} + b_{r+2}\vec{\gamma}_{r+2} + \ldots + b_n\vec{\gamma}_n,$$

teda vektor $\vec{\beta}$ je lineárnou kombináciou vektorov $\vec{\gamma}_{r+1}, \vec{\gamma}_{r+2}, \ldots, \vec{\gamma}_n$. □

t:DOSDIMRIES}

Dôsledok 5.7.5. *Nech A je matica typu $m \times n$ a S je priestor riešení homogénnej sústavy lineárnych rovníc s maticou A. Potom*

$$d(S) = n - h(A).$$

ZNVYMENASTLP}

Poznámka 5.7.6. V skutočnosti sme vlastne dôsledok 5.7.5 zatiaľ dokázali iba v špeciálnom prípade, že sú premenné (resp. stĺpce našej matice) vhodne usporiadané. V celom dôkaze vety 5.7.4 sme totiž predpokladali, že premenné sú usporiadané tak, aby výsledná redukovaná trojuholníková matica mala vedúce jednotky na začiatku.

Môžeme si však uvedomiť, že presne rovnaký dôkaz by fungoval aj bez výmeny premenných; len by ho bolo o dosť ťažšie formálne zapísať. Všetky výpočty v tomto dôkaze totiž robíme po súradniciach.

Iný možný pohľad je uvedomiť si, že permutácia súradníc predstavuje izomorfizmus $F^n \to F^n$. Aj ak zúžime toto zobrazenie na nejaký podpriestor, tak dostaneme opäť izomorfizmus medzi dvoma podpriestormi. (V tomto prípade podpriestorom riešení pôvodnej sústavy a podpriestorom riešení sústavy s vymenenými premennými.) Izomorfizmus zachováva dimenziu priestoru.

Dôsledok 5.7.7. *Homogénna sústava lineárnych rovníc s n neznámymi, ktorej matica má hodnosť n, má len triviálne riešenie.*

Ilustrujme si predchádzajúci postup na dvoch veľmi jednoduchých príkladoch homogénnych sústav lineárnych rovníc nad $\mathbb{R}$.

Príklad 5.7.8. $\left(\begin{smallmatrix}1&1&1&|&0\\1&0&1&|&0\\0&1&1&|&0\end{smallmatrix}\right) \sim \left(\begin{smallmatrix}1&0&1&|&0\\0&1&0&|&0\\0&1&1&|&0\end{smallmatrix}\right) \sim \left(\begin{smallmatrix}1&0&1&|&0\\0&1&0&|&0\\0&0&1&|&0\end{smallmatrix}\right) \sim \left(\begin{smallmatrix}1&0&0&|&0\\0&1&0&|&0\\0&0&1&|&0\end{smallmatrix}\right)$

V tomto prípade nám nezostali žiadne premenné, ktoré by sme mohli voliť – vo všetkých stĺpcoch máme vedúce jednotky. Prepísaním sústavy z maticového zápisu priamo dostaneme (ako jediné možné riešenie) triviálne riešenie $x_1 = 0$, $x_2 = 0$, $x_3 = 0$.

Poznámka 5.7.9. Všimnime si, že na pravé strany pri všetkých elementárnych riadkových úpravách zostávajú nulové. Kvôli stručnejšiemu zápisu, pri riešení homogénnych sústav budeme vynechávať pravé strany a budeme namiesto rozšírenej matice sústavy písať len maticu sústavy. (Budeme si pamätať, že pravé strany sú nuly, ale nebudeme ich po každom kroku znovu písať.)

Príklad 5.7.10. Vynechajme teraz z predchádzajúcej sústavy jednu rovnicu. Úpravou na redukovanú trojuholníkovú maticu dostaneme

$\left(\begin{smallmatrix}1&1&1\\1&0&1\end{smallmatrix}\right) \sim \left(\begin{smallmatrix}1&0&1\\0&1&0\end{smallmatrix}\right)$

Vidíme, že sústava je ekvivalentná so sústavou $x_1 + x_3 = 0$ a $x_2 = 0$. Premennú x_3 môžeme voliť (v treťom stĺpci nie je vedúca jednotka). Nech teda $x_3 = t$, kde $t \in \mathbb{R}$ je parameter. Keď z predchádzajúcich rovníc vyjadríme x_1 a x_2, máme $x_1 = -t$ a $x_2 = 0$. Množina riešení tejto sústavy je teda $\{(t, 0, -t); t \in \mathbb{R}\}$.

Ľahko môžeme overiť, že množina riešení je skutočne podpriestor priestoru $\mathbb{R}^3$. Pretože každý vektor z množiny riešení má tvar $t.(1, 0, -1)$, môžeme ju zapísať aj ako $\{(t, 0, -t); t \in \mathbb{R}\} = [(1, 0, -1)]$.

Aj postup riešenia nehomogénnych sústav je veľmi podobný – podrobnejšie ho rozoberieme v nasledujúcej podkapitole. Predtým však ešte dokážeme vetu, ktorú môžeme chápať ako obrátenie vety 5.7.3.

{sust:VTPPRHOM}

Veta 5.7.11. *Každý podpriestor priestoru F^n je množinou riešení nejakého homogénneho systému lineárnych rovníc.*

Dôkaz. Ak S je podpriestor F^n, tak S je konečnorozmerný (veta 4.4.17). Má teda konečnú bázu $\vec{\alpha}_1, \ldots, \vec{\alpha}_r$.

Nech B je matica, ktorej riadky tvoria vektory $\vec{\alpha}_1, \ldots, \vec{\alpha}_r$,

$$B = \begin{pmatrix} \vec{\alpha}_1 \\ \vdots \\ \vec{\alpha}_r \end{pmatrix}.$$

Podľa predchádzajúcej vety má podpriestor riešení homogénnej sústavy

$$B \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \vec{0}^T$$

bázu $\vec{\gamma}_{r+1}, \ldots, \vec{\gamma}_n$.

Označme ako A maticu, ktorej riadkami sú vektory $\vec{\gamma}_{r+1}, \ldots, \vec{\gamma}_n$,

$$A = \begin{pmatrix} \vec{\gamma}_{r+1} \\ \vdots \\ \vec{\gamma}_n \end{pmatrix}.$$

Pretože každý vektor $\vec{\gamma}_i$ je riešením sústavy s maticou B, platí $B.\vec{\gamma}_i^T = \vec{0}^T$. Z toho dostaneme

$$B.A^T = 0$$

(treba si uvedomiť, že i-ty stĺpec matice A je $\vec{\gamma}_i^T$, z čoho vyplýva, že stĺpce matice $B.A^T$ môžeme vypočítať ako $B.\vec{\gamma}_i^T$). Transponovaním predchádzajúceho vzťahu dostaneme (na základe (5.1))

$$A.B^T = 0.$$

Keď porovnáme i-ty stĺpec matice na ľavej a pravej strane predchádzajúcej rovnosti, dostaneme

$$A\vec{\alpha}_i^T = \vec{0}^T,$$

teda vektory $\vec{\alpha}_1, \ldots, \vec{\alpha}_r$ sú riešeniami homogénnej sústavy $A\vec{x}^T = \vec{0}^T$.

Označme ako M priestor riešení tejto sústavy. Jeho dimenzia je

$$d(M) = n - h(A) = n - (n - r) = r.$$

Súčasne platí $S \subseteq M$ (pretože všetky vektory $\vec{\alpha}_1, \ldots, \vec{\alpha}_r$ patria do M) a $d(S) = d(M)$, teda podľa tvrdenia 4.4.18 platí $S = M$. □

5.7.2 Gaussova eliminačná metóda

Gaussovou eliminačnou metódou nazývame algoritmus na riešenie sústav lineárnych rovníc, o ktorom sme hovorili v predchádzajúcej kapitole. Ide teda o postup, pri ktorom rozšírenú maticu sústavy najprv upravíme na redukovanú trojuholníkovú maticu a z nej už potom vieme zistiť riešenie pôvodnej sústavy.

V prípade, že počas úprav dostaneme riadok tvaru $(0 \ldots 0|c)$, kde $c \neq 0$, sústava nemá riešenie. (Takýto riadok zodpovedá rovnici $0x_1 + \ldots + 0x_n = c$.) V takomto prípade samozrejme nemusíme ďalej pokračovať v upravovaní na RTM.

Ak niektoré stĺpce (v upravenej matici) neobsahujú vedúcu jednotku, tak im prislúchajúce premenné zvolíme za parametre.

Ukážeme si tento postupu na niekoľkých jednoduchých príkladoch. V prípade homogénnych sústav sme mali dve možnosti – buď existovalo jediné riešenie (pri homogénnej sústave

to bolo triviálne riešenie) alebo riešení bolo viac (tvorili podpriestor). Pri nehomogénnej sústave lineárnych rovníc už množina riešení netvorí vektorový podpriestor a navyše pribudne ešte ďalšia možnosť – môže sa stať, že sústava nemá nijaké riešenie.

Príklad 5.7.12. Riešme sústavu

$$\begin{array}{rrrrl} x_1 & -2x_2 & +3x_3 & -4x_4 & = 4 \\ & x_2 & -x_3 & +x_4 & = -3 \\ x_1 & +3x_2 & & -3x_4 & = 1 \\ & -7x_2 & +3x_3 & +x_4 & = -3 \end{array}$$

nad poľom $\mathbb{R}$.

Danú sústavu najprv prepíšeme do matice a potom upravujeme rozšírenú maticu sústavy až kým nedostaneme redukovaný trojuholníkový tvar.

$$\left(\begin{array}{cccc|c} 1 & -2 & 3 & -4 & 4 \\ 0 & 1 & -1 & 1 & -3 \\ 1 & 3 & 0 & -3 & 1 \\ 0 & -7 & 3 & 1 & -3 \end{array}\right) \overset{(1)}{\sim} \left(\begin{array}{cccc|c} 1 & -2 & 3 & -4 & 4 \\ 0 & 1 & -1 & 1 & -3 \\ 0 & 5 & -3 & 1 & -3 \\ 0 & -7 & 3 & 1 & -3 \end{array}\right) \overset{(2)}{\sim} \left(\begin{array}{cccc|c} 1 & -2 & 3 & -4 & 4 \\ 0 & 1 & -1 & 1 & -3 \\ 0 & 0 & 2 & -4 & 12 \\ 0 & 0 & -4 & 8 & -24 \end{array}\right) \overset{(3)}{\sim} \left(\begin{array}{cccc|c} 1 & -2 & 3 & -4 & 4 \\ 0 & 1 & -1 & 1 & -3 \\ 0 & 0 & 1 & -2 & 6 \\ 0 & 0 & 1 & -2 & 6 \end{array}\right)$$

$$\overset{(4)}{\sim} \left(\begin{array}{cccc|c} 1 & -2 & 3 & -4 & 4 \\ 0 & 1 & -1 & 1 & -3 \\ 0 & 0 & 1 & -2 & 6 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \overset{(5)}{\sim} \left(\begin{array}{cccc|c} 1 & -2 & 3 & -4 & 4 \\ 0 & 1 & 0 & -1 & -3 \\ 0 & 0 & 1 & -2 & 6 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \overset{(6)}{\sim} \left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & -8 \\ 0 & 1 & 0 & -1 & 3 \\ 0 & 0 & 1 & -2 & 6 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

(1) `3.r-=1.r` (Týmto zápisom myslím to, že od tretieho riadku sa odčíta prvý.)
(2) `3.r-=5*2.r; 4.r+=7*1.r`
(3) `3.r*=1/2; 4.r*=-1/4`
(4) `3.r-=4.r`
(5) `2.r+=3.r`
(6) `1.r+=2*2.r-3*3.r`

Štvrtý stĺpec neobsahuje vedúcu jednotku. Preto x_4 zvolíme za parameter - položíme $x_4 = t$. Dostaneme potom $x_1 = -8$, $x_2 = t + 3$, $x_3 = 2t + 6$. Množina všetkých riešení je teda $\{(-8, 3 + t, 6 + 2t, t); t \in \mathbb{R}\}$.

(Ak by bola vedúca jednotka v každom stĺpci redukovanej trojuholníkovej matice, ktorú sme dostali z matice sústavy, mali by sme situáciu ešte jednoduchšiu – dostali by sme jediné riešenie. Jedine v prípade, že by bola vedúca jednotka aj v stĺpci pravých strán, čo zodpovedá rovnici $0 = 1$, by sústava nemala žiadne riešenie.)

Skúšku správnosti urobíme tak, že dosadíme výsledok do pôvodnej sústavy. V prípade, že v riešení vystupuje parameter, buď dosadíme výsledok aj s parametrom, alebo to vyskúšame pre nejaké dve hodnoty parametra (také, aby sa nám dobre rátalo). Ak je parametrov viac, môžeme napríklad zvoliť najprv všetky parametre za nulu (tým skontrolujeme riešenie nehomogénneho systému) a potom vždy jeden z parametrov položíme rovný 1 a ostatné 0.

Úpravu, v ktorej sme spravili chybu, môžeme nájsť tak, že skúšame, pre ktoré z matíc získaných počas upravovania náš výsledok ešte vyhovuje a pre ktoré už nie. (Ak nejaká n-tica $(x_1, \ldots, x_n)$ vyhovuje sústave, ktorú sme dostali v jednom kroku alebo sústave v nasledujúcom kroku už nevyhovuje (alebo je to obrátene), táto úprava musí byť chybná. Vyplýva to z toho, že elementárne riadkové operácie nemenia množinu riešení.)

Príklad 5.7.13. Riešme v $\mathbb{Z}_5$ určenú maticou

$$\left(\begin{array}{cccc|c} 1 & 1 & 0 & 0 & 1 \\ 1 & 2 & 4 & 0 & 2 \\ 0 & 1 & 3 & 4 & 3 \\ 0 & 0 & 4 & 4 & 4 \end{array}\right)$$

$$\left(\begin{array}{cccc|c} 1&1&0&0&1\\ 1&2&4&0&2\\ 0&1&3&4&3\\ 0&0&4&4&4 \end{array}\right) \overset{(1)}{\sim} \left(\begin{array}{cccc|c} 1&1&&&1\\ &1&4&&1\\ &1&3&4&3\\ &&4&4&4 \end{array}\right) \overset{(2)}{\sim} \left(\begin{array}{cccc|c} 1&1&&&1\\ &1&4&&1\\ &&4&4&2\\ &&4&4&4 \end{array}\right) \overset{(3)}{\sim} \left(\begin{array}{cccc|c} 1&1&&&1\\ &1&4&&1\\ &&4&4&2\\ 0&0&0&0&2 \end{array}\right)$$

(1) `2.r+=4*1.r`(2) `3.r+=4*2.r` (3) `4.r+=4*3.r` (Kvôli stručnosti a prehľadnosti som v matici vynechával nulové koeficienty.)

Pretože sme dostali riadok zodpovedajúci rovnici $0x_1+0x_2+0x_3+0x_4 = 2$, sústava nemá riešenie.

5.7.3 Frobeniova veta

V tejto časti dokážeme vetu, ktorá poskytuje kritérium na riešiteľnosť nehomogénnych sústav lineárnych rovníc. Predtým však potrebujeme ukázať, že hodnosť matice je rovnaká ako hodnosť transponovanej matice. (Túto vetu neskôr ešte dokážeme dvoma odlišnými spôsobmi v časti 5.9.)

{sust:VTHAT}

Veta 5.7.14. *Pre každú maticu A nad poľom F platí $h(A) = h(A^T)$.*

Dôkaz. Nech A je matica typu $m \times n$.

Ak označíme i-ty stĺpec matice A ako $\vec{\alpha}_i$, tak platí

$$h(A^T) = d[\vec{\alpha}_1, \ldots, \vec{\alpha}_n].$$

Bez toho, aby sme zmenili hodnosť, môžeme preusporiadať stĺpce matice tak, aby po úprave na redukovanú trojuholníkovú maticu boli vedúce jednotky v prvých r stĺpcoch, kde $r = h(A)$.

Budeme sa zaoberať riešeniami homogénnej sústavy

$$A\vec{x}^T = \vec{0}^T,$$

ktorú môžeme ekvivalentne prepísať ako

{sust:EQSTLPCE}

$$x_1\vec{\alpha}_1 + \cdots + x_n\vec{\alpha}_n = \vec{0}. \tag{5.5}$$

(Rozmyslite si prečo – vyplýva to priamo z definície súčinu matíc.)

Všetky riešenia tejto sústavy sú určené bázou (5.4). Špeciálne z toho, že

$$\vec{\gamma}_i = (-c_{1,i}, -c_{2,i}, \ldots, -c_{r,i}, 0, \ldots, 0, 1, 0, \ldots, 0)$$

(kde $i \in \{r+1, r+2, \ldots, n\}$) je riešením (5.5), vyplýva

$$-c_{1,i}\vec{\alpha}_1 - c_{2,i}\vec{\alpha}_2 - \ldots - c_{r,i}\vec{\alpha}_r + \vec{\alpha}_i = \vec{0},$$
$$\vec{\alpha}_i = c_{1,i}\vec{\alpha}_1 + c_{2,i}\vec{\alpha}_2 + \ldots + c_{r,i}\vec{\alpha}_r,$$

a teda $\vec{\alpha}_i$ je lineárna kombinácia vektorov $\vec{\alpha}_1, \ldots, \vec{\alpha}_r$ pre všetky $i = r+1, r+2, \ldots, n$.

Teda

$$[\vec{\alpha}_1, \ldots, \vec{\alpha}_n] = [\vec{\alpha}_1, \ldots, \vec{\alpha}_r]$$

a

$$h(A^T) = d([\vec{\alpha}_1, \ldots, \vec{\alpha}_n]) \leq r,$$

čo znamená, že

$$h(A) \geq h(A^T).$$

Použitím tejto nerovnosti pre maticu A^T však dostaneme

$$h(A^T) \geq h((A^T)^T) = h(A),$$

a teda $h(A^T) = h(A)$. □

Poznámka 5.7.15. Pretože vykonanie riadkovej operácie na transponovanej matici A^T zodpovedá stĺpcovej operácii na matici A, z práve dokázanej vety vyplýva, že pri výpočte hodnosti matice môžeme ľubovoľne kombinovať riadkové a stĺpcové operácie.

Veta 5.7.16 (Frobeniova)**.** *Nehomogénna sústava lineárnych rovníc* (5.2) *je riešiteľná práve vtedy, keď matica sústavy a rozšírená matica sústavy majú rovnakú hodnosť, t.j.*

$$h(A) = h(A').$$

Dôkaz. Označme $\vec{\gamma} = (c_1, \ldots, c_m)$ vektor pozostávajúci z pravých strán, čiže naša sústava má tvar $A\vec{x}^T = \vec{\gamma}^T$. Ďalej označme stĺpce matice A ako $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$.

$\boxed{\Rightarrow}$ Ak $x_1, \ldots, x_n$ je riešením tejto sústavy, znamená to, že

$$\vec{\gamma} = x_1\vec{\alpha}_1 + \cdots + x_n\vec{\alpha}_n.$$

Z toho vyplýva

$$[\vec{\alpha}_1, \ldots, \vec{\alpha}_n] = [\vec{\alpha}_1, \ldots, \vec{\alpha}_n, \vec{\gamma}]$$
$$h(A^T) = d([\vec{\alpha}_1, \ldots, \vec{\alpha}_n]) = d([\vec{\alpha}_1, \ldots, \vec{\alpha}_n, \vec{\gamma}]) = h(A'^T)$$

Pretože podľa predchádzajúcej vety má každá matica rovnakú hodnosť ako jej transponovaná matica, dostali sme

$$h(A) = h(A').$$

$\boxed{\Leftarrow}$ Predpokladajme teraz, že $h(A) = h(A')$. To znamená, že podpriestory $[\vec{\alpha}_1, \ldots, \vec{\alpha}_n]$ a $[\vec{\alpha}_1, \ldots, \vec{\alpha}_n, \vec{\gamma}]$ majú rovnakú dimenziu. Pretože jeden z nich je navyše podpriestorom druhého, podľa tvrdenia 4.4.18 z toho vyplýva rovnosť

$$[\vec{\alpha}_1, \ldots, \vec{\alpha}_n] = [\vec{\alpha}_1, \ldots, \vec{\alpha}_n, \vec{\gamma}].$$

To znamená, že $\vec{\gamma}$ je lineárnou kombináciou vektorov $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$, teda existujú $x_1, \ldots, x_n \in F$ také, že

$$\vec{\gamma} = x_1\vec{\alpha}_1 + \cdots + x_n\vec{\alpha}_n.$$

Ako sme si už uvedomili v predchádzajúcej časti dôkazu, táto rovnosť je ekvivalentná s tým, že $x_1, \ldots, x_n$ je riešenie sústavy (5.2). □

Nasledujúca veta hovorí, že ak máme jedno konkrétne (partikulárne) riešenie nehomogénnej lineárnej sústavy, tak všetky ostatné riešenia nehomogénnej sústavy môžeme získať ako súčet tohoto riešenia a ľubovoľného riešenia príslušnej homogénnej sústavy.

Oveľa dôležitejší ako samotné znenie vety je veľmi dôležitý koncept rozdelenia riešenia na homogénnu a nehomogénnu časť. S týmto prístupom sa stretnete ešte veľakrát – dá sa využiť v podstate všade, kde sa vyskytuje linearita, napríklad pri lineárnych diferenciálnych rovniciach, pri lineárnych rekurentných rovniciach alebo tiež pri afinných priestoroch a afinných zobrazeniach – čo je niečo podobné ako vektorové priestory a lineárne zobrazenia, len sú doplnené o „nehomogénnu“ zložku (viac sa o afinných priestoroch môžete dočítať napríklad v [HZK]).

Veta 5.7.17. *Nech $\vec{\alpha}$ je riešenie sústavy lineárnych rovníc*

$$A\vec{\alpha}^T = \vec{\gamma}^T \qquad \text{(N)} \quad \texttt{\{sust:EQN\}}$$

a S je podpriestor pozostávajúci zo všetkých riešení homogénneho systému

$$A\vec{\alpha}^T = \vec{0}^T. \qquad \text{(H)} \quad \texttt{\{sust:EQH\}}$$

Potom $T = \{\vec{\alpha} + \vec{\beta}; \vec{\beta} \in S\}$ je množina všetkých riešení (N).

Inak povedané, ľubovoľné riešenie (N) sa dá získať ako súčet vektora $\vec{\gamma}$ (partikulárneho riešenia (N)) a nejakého riešenia homogénnej sústavy (H).

Dôkaz. Rovnosť dvoch množín (množiny T a množiny všetkých riešení) budeme dokazovať tak, že dokážeme obe inklúzie.

Najprv ukážeme, že každý prvok množiny T je riešením sústavy (N). Skutočne, pre prvok tvaru $\vec{\alpha}+\vec{\beta}$ platí

$$A(\vec{\alpha}+\vec{\beta})^T = A\vec{\alpha}^T + A\vec{\beta}^T = \vec{\gamma}^T + \vec{0}^T = \vec{\gamma}^T.$$

Zostáva ukázať, že každé riešenie (N) má uvedený tvar. Nech teda $\vec{\delta}$ je ľubovoľné riešenie (N), čiže platí $A\vec{\delta}^T = \vec{\gamma}^T$. Potom platí

$$A(\vec{\delta}-\vec{\alpha})^T = A\vec{\delta}^T - A\vec{\alpha}^T = \vec{\gamma}^T - \vec{\gamma}^T = \vec{0}^T.$$

Potom

$$\vec{\delta} = \vec{\alpha} + \underbrace{(\vec{\delta}-\vec{\alpha})}_{\in S},$$

teda je to súčet vektora $\vec{\alpha}$ a prvku z S. Ukázali sme, že ľubovoľné riešenie (N) patrí do T. □

Cvičenia

{sustcvic:SYS1}

Úloha 5.7.1. Nájdite všetky riešenia daných sústav rovníc nad poľom $\mathbb{R}$:

$$\begin{array}{rrrrl} x_1 & -x_2 & +2x_3 & -3x_4 & =1 \\ & x_2 & -x_3 & +x_4 & =-3 \\ x_1 & +3x_2 & & -3x_4 & =1 \\ & -7x_2 & +3x_3 & +x_4 & =3 \end{array} \qquad \begin{array}{rl} x_1+x_2 & =1 \\ x_1+x_2+x_3 & =4 \\ x_2+x_3+x_4 & =-3 \\ x_3+x_4+x_5 & =2 \\ x_4+x_5 & =-1 \end{array}$$

$$\begin{array}{rrrrl} 2x & -5y & +3z & +t & =5 \\ 3x & -7y & +3z & -t & =-1 \\ 5x & -9y & +6z & +2t & =7 \\ 4x & -6y & +3z & +t & =8 \end{array} \qquad \begin{array}{rrrrl} x & +2y & +4z & -3t & =0 \\ 3x & +5y & +6z & -4t & =0 \\ 4x & +5y & -2z & +3t & =0 \\ 3x & +8y & +24z & -19t & =0 \end{array}$$

$$\begin{array}{rrrrl} x & +4y & -2z & +8t & =12 \\ & y & -7z & +2t & =-4 \\ & & 5z & -t & =7 \\ & & z & +3t & =-5 \end{array}$$

{sustcvic:SUS2}

Úloha 5.7.2. Riešte v $\mathbb{Z}_5$ sústavu určenú maticou:

$$\left(\begin{array}{cccc|c} 1&1&0&3&1\\1&2&4&0&2\\2&1&3&4&3\\3&0&4&4&4\end{array}\right) \left(\begin{array}{cccc|c} 3&1&2&2&1\\4&4&2&1&0\\0&1&2&4&1\\2&1&1&2&3\end{array}\right) \left(\begin{array}{cccc|c} 2&4&1&1&2\\3&3&3&2&1\\1&4&2&1&1\\4&2&0&3&2\end{array}\right) \left(\begin{array}{cccc|c} 1&2&3&2&4\\2&3&1&1&3\\4&3&1&3&2\\3&4&3&2&1\end{array}\right)$$

Úloha 5.7.3. Riešte v $\mathbb{R}$ sústavu určenú maticou:

$$\left(\begin{array}{ccc|c} 3&-2&1&11\\1&1&-3&7\\11&-4&-3&10\end{array}\right) \left(\begin{array}{ccc|c} 1&2&-1&2\\3&-1&2&7\\1&0&-1&-2\\2&1&1&7\end{array}\right) \left(\begin{array}{ccc|c} 1&2&-3&1\\-1&3&-2&3\\0&5&-5&4\end{array}\right) \left(\begin{array}{ccc|c} 1&-2&1&0\\4&1&-1&2\\1&2&4&0\end{array}\right) \left(\begin{array}{ccc|c} 1&4&-3&0\\1&-3&-1&0\\2&1&-4&0\end{array}\right)$$

Riešenie: a) nemá riešenie, b) (1,2,3) c) $(t-\frac{3}{5}, t+\frac{4}{5}, t)$, d) $(\frac{20}{47}, \frac{6}{47}, -\frac{8}{47})$, e) $(\frac{13}{7}t, \frac{2}{7}t, t)$

Úloha 5.7.4. Riešte v $\mathbb{Z}_7$ sústavu určenú maticou:

$$\left(\begin{array}{ccc|c} 1&0&1&1&5\\0&1&1&1&6\\3&1&2&3&0\\0&3&6&1&4\end{array}\right) \left(\begin{array}{ccc|c} 1&1&0&1&0\\2&1&0&3&1\\3&1&1&1&5\\0&1&2&3&6\end{array}\right) \left(\begin{array}{ccc|c} 2&1&4&1&4\\1&3&3&1&5\\4&1&5&1&6\\2&3&1&4&2\end{array}\right) \left(\begin{array}{ccc|c} 1&2&1&3&1\\2&1&2&3&2\\3&1&1&1&1\\0&6&5&3&3\end{array}\right) \left(\begin{array}{ccc|c} 1&2&1&3&1\\2&3&1&4&2\\1&1&0&1&1\\0&4&4&1&0\end{array}\right)$$

Úloha 5.7.5. Môžete si vymyslieť kopec vlastných sústav. Stačí najprv zvoliť riešenie, koeficienty a dorátať pravé strany. Skúste vymyslieť aj také sústavy, ktoré nemajú riešenie alebo majú viac než jedno riešenie.

Úloha 5.7.6. Nájdite reálne čísla a, b, c tak, aby graf funkcie $f(x) = ax^2 + bx + c$ prechádzal bodmi (1,2), (-1,6) a (2,3).

Úloha 5.7.7. Nájdite nejakú homogénnu sústavu rovníc so 4 neznámymi nad $\mathbb{R}$, ktorej riešením je daný podpriestor:
a) $S = [(1, 4, 0, 1), (1, 0, 3, -3), (0, 2, 0, 1)]$
b) $S = [(1, -1, 1, -2), (1, 1, 0, -1), (3, 1, 1, -4)]$

Úloha 5.7.8. Nájdite hodnotu parametra $b \in \mathbb{R}$, pre ktorú má daná sústava riešenie. Pre túto hodnotu aj vyjadrite množinu riešení.

$$\begin{aligned} x_1 + 4x_2 - 3x_3 + 2x_4 &= 2 \\ 2x_1 + 7x_2 - 4x_3 + 4x_4 &= 3 \\ -x_1 - 5x_2 + 5x_3 - 2x_4 &= b \\ 3x_1 + 10x_2 - 5x_3 + 6x_4 &= 4 \end{aligned}$$

Úloha 5.7.9. Zistite pre aké hodnoty parametra $a \in \mathbb{R}$ systém

$$\begin{aligned} x + y + 7z &= -7 \\ 2x + 3y + 17z &= -16 \\ x + 2y + (a^2 + 1)z &= 3a \end{aligned}$$

a) má práve jedno riešenie; b) má nekonečne veľa riešení; c) nemá žiadne riešenie.

Úloha 5.7.10$^+$. V závislosti od parametra $a \in \mathbb{R}$ riešte systém daný maticou:
a) $\left(\begin{array}{cc|c} a & 1 & a^2 \\ 1 & a & 1 \end{array}\right)$ b) $\left(\begin{array}{cc|c} a & 1 & a^3 \\ 1 & a & 1 \end{array}\right)$

Úloha 5.7.11$^+$. Ako vyzerajú, v závislosti od parametra p, riešenia sústavy danej maticou:
$\left(\begin{array}{ccc|c} p & 1 & 1 & 1 \\ 1 & p & 1 & 1 \\ 1 & 1 & p & 1 \\ 1 & 1 & 1 & p \end{array}\right)$

Úloha 5.7.12*. O sústave n rovníc o n neznámych nad poľom $\mathbb{R}$ vieme, že jej koeficienty tvoria aritmetickú postupnosť (ako napríklad pre maticu $\left(\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{array}\right)$), že $n \geq 2$ a že táto sústava má jediné riešenie. Nájdite riešenie sústavy.

5.8 Jadro a obraz lineárneho zobrazenia

Definícia 5.8.1. Nech V a W sú vektorové priestory nad poľom F a $f\colon V \to W$ je lineárne zobrazenie. Potom *jadrom lineárneho zobrazenia* f nazývame množinu

$$\operatorname{Ker} f = \{\vec{\alpha} \in V; f(\vec{\alpha}) = \vec{0}\}$$

a *obrazom lineárneho zobrazenia* f nazývame množinu

$$\operatorname{Im} f = \{f(\vec{\alpha}); \vec{\alpha} \in V\}.$$

Inými slovami, $\operatorname{Ker} f$ obsahuje práve tie vektory z V, ktoré sa zobrazia na nulový vektor a $\operatorname{Im} f$ obsahuje obrazy všetkých vektorov z V. Ľahko sa overí, že $\operatorname{Ker} f$ aj $\operatorname{Im} f$ sú vektorové podpriestory. (Môžete si všimnúť, že ide o špeciálny prípad úlohy 5.3.6.)

Tvrdenie 5.8.2. *Nech V a W sú vektorové priestory nad poľom F a $f\colon V \to W$ je lineárne zobrazenie. Potom* $\operatorname{Ker} f$ *je vektorový podpriestor priestoru V a* $\operatorname{Im} f$ *je vektorový podpriestor priestoru W.*

Dôkaz. Pretože $f(\vec{0}) = \vec{0}$, platí $\vec{0} \in \operatorname{Ker} f$, teda $\operatorname{Ker} f \neq \emptyset$.

Ak $\vec{\alpha}, \vec{\beta} \in \operatorname{Ker} f$, znamená to, že $f(\vec{\alpha}) = f(\vec{\beta}) = \vec{0}$. Z linearity potom dostaneme

$$f(\vec{\alpha} + \vec{\beta}) = f(\vec{\alpha}) + f(\vec{\beta}) = \vec{0} + \vec{0} = \vec{0},$$

čiže aj $\vec{\alpha} + \vec{\beta} \in \operatorname{Ker} f$.

Podobne, ak $c \in F$ a $\vec{\alpha} \in \operatorname{Ker} f$, dostaneme

$$f(c.\vec{\alpha}) = c.f(\vec{\alpha}) = c.\vec{0} = \vec{0}$$

a $c.\vec{\alpha} \in \operatorname{Ker} f$.

Pretože $f(\vec{0}) = \vec{0}$, platí $\vec{0} \in \operatorname{Im} f$, teda $\operatorname{Im} f \neq \emptyset$.

Ak $\vec{\alpha}, \vec{\beta} \in \operatorname{Im} f$, znamená to, že tieto vektory sú obrazmi nejakých vektorov z V, označme ich $\vec{\alpha}_1$ a $\vec{\beta}_1$. Máme teda

$$f(\vec{\alpha}_1) = \vec{\alpha}$$
$$f(\vec{\beta}_1) = \vec{\beta}$$
$$f(\vec{\alpha}_1 + \vec{\beta}_1) = \vec{\alpha} + \vec{\beta}$$

Teda vektor $\vec{\alpha} + \vec{\beta}$ je obrazom vektora $\vec{\alpha}_1 + \vec{\beta}_1$, čiže patrí do $\operatorname{Im} f$.

Podobne sa ukáže

$$c\vec{\alpha} = cf(\vec{\alpha}_1) = f(c\vec{\alpha}_1),$$

teda aj $c\vec{\alpha} \in \operatorname{Im} f$. □

Všimnime si, že $\operatorname{Im} f$ vieme vygenerovať obrazmi generátorov priestoru V:

{kerim:TVRIMAGENER}

Tvrdenie 5.8.3. *Nech $f\colon V \to W$ je lineárne zobrazenie a $V = [\vec{\alpha}_1, \ldots, \vec{\alpha}_n]$. Potom* $\operatorname{Im} f = [f(\vec{\alpha}_1), \ldots, f(\vec{\alpha}_n)]$.

Dôkaz. $\boxed{\subseteq}$ Ak $\vec{\beta} \in \operatorname{Im} f$, znamená to, že $\vec{\beta} = f(\vec{\alpha})$ pre nejaké $\vec{\alpha} \in V$. Z toho, že $\vec{\alpha} \in V = [\vec{\alpha}_1, \ldots, \vec{\alpha}_n]$ máme, že $\vec{\alpha}$ je lineárna kombinácia vektorov $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$, teda sa dá vyjadriť ako $\vec{\alpha} = c_1\vec{\alpha}_1 + \cdots + c_n\vec{\alpha}_n$. Z linearity zobrazenia f potom máme

$$\vec{\beta} = f(\vec{\alpha}) = f(c_1\vec{\alpha}_1 + \cdots + c_n\vec{\alpha}_n) = c_1 f(\vec{\alpha}_1) + \cdots + c_n f(\vec{\alpha}_n).$$

Zistili sme, že $\vec{\beta} \in [f(\vec{\alpha}_1), \ldots, f(\vec{\alpha}_n)]$.

$\boxed{\supseteq}$ Predpokladajme teraz, že $\vec{\beta} \in [f(\vec{\alpha}_1), \ldots, f(\vec{\alpha}_n)]$. To znamená, že sa dá napísať ako lineárna kombinácia týchto vektorov, teda

$$\vec{\beta} = c_1 f(\vec{\alpha}_1) + \cdots + c_n f(\vec{\alpha}_n)$$

pre nejaké $c_1, \ldots, c_n \in F$. Ak opäť použijeme to, že f je lineárne zobrazenie, tak dostaneme

$$\vec{\beta} = f(c_1\vec{\alpha}_1 + \cdots + c_n\vec{\alpha}_n).$$

Ukázali sme, že $\vec{\beta}$ je obrazom vektora $\vec{\alpha} = c_1\vec{\alpha}_1 + \cdots + c_n\vec{\alpha}_n \in V$, teda $\vec{\beta} \in \operatorname{Im} f$. □

Teraz si povieme, ako súvisí jadro a obraz lineárneho zobrazenia s tým, či je toto zobrazenie surjektívne alebo injektívne.

Tvrdenie 5.8.4. *Nech V s W sú vektorové priestory nad poľom F a $f\colon V \to W$ je lineárne zobrazenie.*

Zobrazenie f je injektívne práve vtedy, keď $\operatorname{Ker} f = \{\vec{0}\}$.

Dôkaz. $\boxed{\Rightarrow}$ Predpokladajme, že f je injektívne. Vieme, že $f(\vec{0}) = \vec{0}$. Z injektívnosti vyplýva, že iný vektor sa už na nulový vektor nemôže zobraziť, preto $\operatorname{Ker} f = \{\vec{0}\}$.

$\boxed{\Leftarrow}$ Nech $\operatorname{Ker} f = \{\vec{0}\}$. Ak $f(\vec{\alpha}) = f(\vec{\beta})$, tak $f(\vec{\alpha} - \vec{\beta}) = \vec{0}$, čiže $\vec{\alpha} - \vec{\beta} \in \operatorname{Ker} f$. To ale znamená, že $\vec{\alpha} - \vec{\beta} = \vec{0}$, a teda $\vec{\alpha} = \vec{\beta}$. □

Dôkaz nasledujúceho tvrdenia vynecháme, ide vlastne len o inak prepísanú definíciu surjektívnosti.

Tvrdenie 5.8.5. *Nech V a W sú vektorové priestory nad poľom F a $f\colon V \to W$ je lineárne zobrazenie.*

Zobrazenie f je surjektívne práve vtedy, keď $\operatorname{Im} f = W$.

Dôsledok 5.8.6. *Lineárne zobrazenie $f\colon V \to W$ je izomorfizmus práve vtedy, keď* $\operatorname{Im} f = W$ *a* $\operatorname{Ker} f = \{\vec{0}\}$.

{kerim:VTDIKER}

Veta 5.8.7. *Nech V a W sú konečnorozmerné vektorové priestory a $f\colon V \to W$ je lineárne zobrazenie. Potom*

$$d(V) = d(\operatorname{Ker} f) + d(\operatorname{Im} f).$$

Dôkaz. Nech $\vec{\alpha}_1, \ldots, \vec{\alpha}_k$ je báza $\operatorname{Ker} f$ (teda $d(\operatorname{Ker} f) = k$). Bázu $\operatorname{Ker} f$ vieme doplniť na bázu celého priestoru V vektormi $\vec{\beta}_1, \ldots, \vec{\beta}_l$. (Teda $d(V) = k + l$.) Označme $S = [\vec{\beta}_1, \ldots, \vec{\beta}_l]$.

Definujme zobrazenie $g\colon S \to \operatorname{Im} f$ ako zúženie zobrazenia f, t.j. $g(\vec{\alpha}) = f(\vec{\alpha})$ pre všetky $\vec{\alpha} \in S$. Ukážeme, že toto zobrazenie je izomorfizmus.

Je zrejmé, že ide o lineárne zobrazenie.

Ďalej máme $\operatorname{Im} f = [f(\vec{\alpha}_1), \ldots, f(\vec{\alpha}_k), f(\vec{\beta}_1), \ldots, f(\vec{\beta}_l)] = [\vec{0}, \ldots, \vec{0}, f(\vec{\beta}_1), \ldots, f(\vec{\beta}_l)] = [f(\vec{\beta}_1), \ldots, f(\vec{\beta}_l)] = [g(\vec{\beta}_1), \ldots, g(\vec{\beta}_l)] = \operatorname{Im} g$. Z toho vyplýva, že toto zobrazenie je surjektívne.

Ešte ukážeme, že g je injektívne. Na to nám stačí ukázať, že $\operatorname{Ker} g = \{\vec{0}\}$. Priamo z definície zobrazenia g vidíme, že $\operatorname{Ker} g = \operatorname{Ker} f \cap S$. Tento prienik obsahuje iba nulový vektor. (Ak nejaký vektor $\vec{\alpha}$ patrí do $\operatorname{Ker} g \cap S$, tak sa dá vyjadriť súčasne ako lineárna kombinácia vektorov $\vec{\alpha}_1, \ldots, \vec{\alpha}_k$ a aj vektorov $\vec{\beta}_1, \ldots, \vec{\beta}_l$. Keďže $\vec{\alpha}_1, \ldots, \vec{\alpha}_k, \vec{\beta}_1, \ldots, \vec{\beta}_l$ tvoria bázu, každý vektor má jednoznačné vyjadrenie ako lineárna kombinácia týchto vektorov. Porovnaním vyjadrení

$$\vec{\alpha} = c_1\vec{\alpha}_1 + \cdots + c_k\vec{\alpha}_k + 0\vec{\beta}_1 + \cdots + 0\vec{\beta}_l = 0\vec{\alpha}_1 + \cdots + 0\vec{\alpha}_k + d_1\vec{\beta}_1 + \cdots + d_l\vec{\beta}_l$$

dostaneme – na základe jednoznačnosti – že všetky koeficienty sú nulové a $\vec{\alpha} = \vec{0}$.)

Zistili sme teda, že f je izomorfizmus medzi priestormi S a $\operatorname{Im} f$. Izomorfizmus zachováva dimenziu (pretože zobrazuje bázu na bázu), teda máme

$$l = d(S) = d(\operatorname{Im} f)$$

a z toho dostaneme

$$d(V) = k + l = d(\operatorname{Ker} f) + d(\operatorname{Im} f).$$

□

Cvičenia

{kerimcvic:JA

Úloha 5.8.1. Nájdite bázu obrazu a bázu jadra lineárneho zobrazenia $f\colon (\mathbb{Z}_5)^4 \to (\mathbb{Z}_5)^4$ s danou maticou. V ktorých prípadoch je toto zobrazenie surjektívne a v ktorých injektívne?
$\begin{pmatrix} 3&1&2&2\\4&3&2&1\\0&1&2&4\\2&0&1&2 \end{pmatrix}$ $\begin{pmatrix} 2&4&1&1\\3&3&3&2\\1&4&2&1\\4&2&0&3 \end{pmatrix}$ $\begin{pmatrix} 1&2&3&4\\2&3&1&4\\4&3&2&1\\3&4&1&2 \end{pmatrix}$

Úloha 5.8.2. Nájdite bázu a dimenziu $\operatorname{Ker} f$ aj $\operatorname{Im} f$ pre dané lineárne zobrazenie. Rozhodnite, či toto zobrazenie je injektívne, surjektívne, bijektívne.
a) $f\colon \mathbb{R}^4 \to \mathbb{R}^3$, $f(x_1,x_2,x_3,x_4) = (x_1+x_2+x_3+x_4, 2x_2+x_3+x_4, 4x_2+2x_3+2x_4)$.
b) $f\colon M_{2,2}(\mathbb{R}) \to M_{2,2}(\mathbb{R})$, $f\colon \left(\begin{smallmatrix} a&b\\c&d \end{smallmatrix}\right) \mapsto \left(\begin{smallmatrix} a+b&b+c\\c+d&d+a \end{smallmatrix}\right)$
c) $f\colon V \to V$, kde $V = \{ax^2+bx+c; a,b,c \in \mathbb{R}\}$ je podpriestor $\mathbb{R}^{\mathbb{R}}$ a $f\colon p(x) \mapsto p'(x)$. (T.j. f priradí polynómu $p(x)$ jeho deriváciu $p'(x)$.)

{kerimcvic:PROSTELIN}

Úloha 5.8.3. Nájdite lineárne zobrazenie (ak také existuje), ktoré je prosté a spĺňa podmienky:
a) $f(1,0,1) = (2,2,1)$, $f(1,-1,1) = (1,2,-2)$, $f(0,1,-2) = (0,-1,2)$,
b) $f(1,0,1) = (2,2,1)$, $f(1,-1,1) = (1,2,-2)$, $f(1,1,1) = (3,2,4)$,
c) $f(1,0,1) = (2,2,1)$, $f(0,-1,2) = (0,1,1)$, $f(1,1,-1) = (2,3,2)$.

Úloha 5.8.4. Nájdite lineárne zobrazenie $f\colon \mathbb{R}^3 \to \mathbb{R}^3$ (ak také existuje), pre ktoré: $f(3,2,3) = (5,-3,-2)$, $f(0,2,1) = (2,0,-2)$, $f(3,0,3) = (3,-3,0)$. Určte bázu a dimenziu jeho jadra a obrazu.

Úloha 5.8.5. Definujme lineárne zobrazenie $f\colon \mathbb{R}^4 \to \mathbb{R}^2$ ako $f(x_1,x_2,x_3,x_4) = (3x_1+x_2+2x_3-x_4, 2x_1+4x_2+x_3-x_4)$ a označme $U_1 = \operatorname{Ker} f$.

Ďalej definujme lineárne zobrazenie $g\colon \mathbb{R}^2 \to \mathbb{R}^4$ ako $g(y_1,y_2) = (y_1-y_2, y_1-3y_2, 2y_1-8y_2, 3y_1-27y_2)$ a označme $U_2 = \operatorname{Im} g$.

Vidíme, že U_1 aj U_2 sú podpriestory $\mathbb{R}^4$.

Nájdite bázy priestorov U_1, U_2, $U_1 \cap U_2$ a $U_1 + U_2$.

Úloha 5.8.6. Nech $f\colon V \to V$ je lineárne zobrazenie. Ako f^2 budeme označovať $f \circ f$. Dokážte

(a) $\operatorname{Ker} f^2 \supseteq \operatorname{Ker} f$,

(b) $\operatorname{Im} f^2 \subseteq \operatorname{Im} f$,

(c) $f^2 = 0 \Leftrightarrow \operatorname{Ker} f \supseteq \operatorname{Im} f$.

Úloha 5.8.7. Nech $f\colon V \to V$ je lineárne zobrazenie a $g\colon V \to V$ je určené predpisom $g(\vec{\alpha}) = \vec{\alpha} - f(\vec{\alpha})$ (inak povedané, $g = id_V - f$). Dokážte:
a) Zobrazenie g je lineárne.
b) Ak $\operatorname{Ker} f = \operatorname{Im} g$, tak $f \circ f = f$. (Zobrazenie s takouto vlastnosťou sa zvykne nazývať *projekcia.*)

Úloha 5.8.8. Nech $f\colon V \to V$ je lineárne zobrazenie a $g\colon V \to V$ je určené predpisom $g(\vec{\alpha}) = \vec{\alpha} - f(\vec{\alpha})$ (inak povedané, $g = id_V - f$). Dokážte:
a) Zobrazenie g je lineárne.
b) Ak $\operatorname{Ker} f = \operatorname{Im} g$, tak $f \circ f = f$. (Zobrazenie s takouto vlastnosťou sa zvykne nazývať *projekcia.*)

Úloha 5.8.9. Dokážte, že $h(A^TA) = h(A)$ pre ľubovoľnú maticu typu $n \times n$ nad $\mathbb{R}$. (Hint 1: Možno pomôže, ak si uvedomíte, že v $\mathbb{R}^n$ platí $\vec{\alpha}\vec{\alpha}^T = 0 \Leftrightarrow \vec{\alpha} = \vec{0}$. Hint 2: Iná možnosť je skúsiť to najprv dokázať pre RTM. Hint 3: Ak vymyslíte úplne iné riešenie, nedajte sa zviesť zo stopy predchádzajúcimi dvoma hintami.)

Úloha 5.8.10*. Nech A, B sú štvorcové matice typu $n \times n$ nad poľom F. Dokážte, že ak $I - AB$ je regulárna, tak aj $I - BA$ je regulárna.

5.9 Hodnosť transponovanej matice

{kerim:SECTTRANS}

Už sme jedným spôsobom ukázali, že

$$h(A) = h(A^T)$$

(veta 5.7.14). Tu uvedieme dva ďalšie spôsoby. Prvý z nich bude využívať práve dokázanú vetu 5.8.7.

Dôkaz vety 5.7.14. Ku matici A typu $m \times n$ prislúcha lineárne zobrazenie $f\colon F^m \to F^n$. Vieme, že toto zobrazenie je určené predpisom

$$f(\vec{\alpha}) = \vec{\alpha}A$$

(poznámka 5.4.10).

Súčasne vieme, že podpriestor $\operatorname{Im} f$ je generovaný riadkami tejto matice. Preto $h(A) = d(\operatorname{Im} f)$.

Do $\operatorname{Ker} f$ patria práve vektory, pre ktoré platí

$$\vec{\alpha}A = \vec{0},$$

z čoho transponovaním dostávame

$$A^T\vec{\alpha}^T = \vec{0}^T,$$

teda sú to práve riešenia homogénneho systému s maticou A^T. Podľa dôsledku 5.7.5 je dimenzia množiny riešení takéhoto systému rovná $m - h(A^T)$ (pretože počet stĺpcov matice A je m).

Z vety 5.8.7 potom dostaneme

$$m = m - h(A^T) + h(A),$$

z čoho vyplýva $h(A^T) = h(A)$. □

Uvedieme ešte jeden, pomerne jednoduchý dôkaz tejto vety.

Dôkaz vety 5.7.14. Dôkaz bude pozostávať z 2 častí: Najprv ukážeme, že táto veta platí pre redukované trojuholníkové matice. Ďalej si uvedomíme, že stĺpcové operácie nemenia hodnosť matice. Z toho už potom vyplynie tvrdenie vety.

Ak B je redukovaná trojuholníková matica typu $m \times n$ a $h(B) = k$, znamená to, že B má k nenulových riadkov a navyše, má k stĺpcov, ktoré obsahujú jedinú (vedúcu) jednotku. Potom B^T je matica typu $n \times m$, v ktorej sú nenulové prvky iba v prvých k stĺpcoch a navyše obsahuje ako svoje riadky vektory $\vec{e}_1, \ldots, \vec{e}_k$ štandardnej bázy priestoru F^m (tieto riadky zodpovedajú tým stĺpcom pôvodnej matice, v ktorých boli vedúce jednotky). Z toho je zrejmé, že priestor V_{B^T} prislúchajúci tejto matici je generovaný vektormi $\vec{e}_1, \ldots, \vec{e}_k$ a teda $h(B^T) = d(V_{B^T}) = k = h(B)$.

Skúsme sa teraz zamyslieť nad tým ako menia dimenziu stĺpcové operácie. Každá stĺpcová operácia zodpovedá nejakému lineárnemu zobrazeniu $F^m \to F^m$ (vykonanie stĺpcovej operácie znamená zobrazenia každého riadku týmto zobrazením). Konkrétne, výmene dvoch stĺpcov i-teho a j-teho stĺpca zodpovedá zobrazenie (pre $i < j$)

$$(x_1, \ldots, x_n) \mapsto (x_1, \ldots, x_{i-1}, x_j, x_{i+1}, \ldots, x_{j-1}, x_i, x_{j+1}, \ldots, x_n)$$

pripočítaniu c-násobku i-teho stĺpca k j-temu zodpovedá

$$(x_1, \ldots, x_n) \mapsto (x_1, \ldots, x_{j-1}, x_j + cx_i, x_{j+1}, \ldots, x_n)$$

a vynásobeniu j-teho riadku konštantou $c \neq 0$ zodpovedá zobrazenie

$$(x_1, \ldots, x_n) \mapsto (x_1, \ldots, x_{j-1}, cx_j, x_{j+1}, \ldots, x_n).$$

Každé z týchto zobrazení je lineárne a navyše k nemu existuje inverzné (to vyplýva napríklad z toho, že stĺpcové operácie sú invertovateľné, ale dá sa to ľahko overiť aj priamo). Všetky takéto zobrazenia sú teda izomorfizmy.

Pretože stĺpcová úprava zodpovedá zobrazeniu podpriestoru prislúchajúceho danej matici nejakým izomorfizmom a izomorfizmus nemení dimenziu, je zrejmé, že stĺpcové operácie nemenia hodnosť matice.

Majme teraz maticu A. Matica A je riadkovo ekvivaletná s nejakou redukovanou trojuholníkovou maticou B. Platí $h(A) = h(B) = h(B^T)$. Z matice B^T však vieme dostať maticu A^T pomocou elementárnych stĺpcových operácií. (Riadkové operácie na pôvodnej matici totiž zodpovedajú stĺpcovým operáciam na transponovanej matici.) Preto máme aj rovnosť $h(B^T) = h(A^T)$ a spojením týchto dvoch rovností dostaneme

$$h(A) = h(A^T).$$

□

Cvičenia

Úloha 5.9.1. Zistite hodnosť matice

$$A = \begin{pmatrix} b & b & b-a \\ a-b & -b & a \\ a+b & b & 0 \end{pmatrix}$$

v závislosti od hodnôt parametrov $a, b \in \mathbb{R}$.

5.10 Násobenie blokových matíc*

{blok:SECT}

Ešte si stručne povedzme o inom pohľade na násobenie matíc, ktorý sa môže v niektorých situáciách hodiť.

Niekedy je vhodné rozdeliť maticu na menšie podmatice, napríklad takýmto spôsobom:

$$M = \left(\begin{array}{cc|cc} 0 & 1 & 3 & 1 \\ 1 & 2 & 4 & 1 \\ \hline 1 & 1 & 3 & 1 \\ 2 & -1 & 0 & 3 \end{array}\right).$$

T.j. zobrali sme nejakú maticu $m \times n$ a nejako sme rozdelili počty riadkov a počty stĺpcov; označme jednotlivé rozmery $m = m_1 + \cdots + m_k$ a $n = n_1 + \cdots + n_l$. Takýmto spôsobom sme maticu rozdelili na $k \times l$ menších matíc, pričom matica na pozícii (i, j) má rozmery $m_i \times n_j$. Takto rozloženej matici zvykneme hovoriť *bloková matica.*

Teda maticu uvedenú v našom príklade by sme mohli zapísať ako

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

pre

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix}, \qquad B = \begin{pmatrix} 3 & 1 \\ 4 & 1 \end{pmatrix}, \qquad C = \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix}, \qquad D = \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}.$$

V tomto prípade mali všetky bloky rovnaké rozmery. Nemusí to tak byť vždy, tú istú maticu by sme mohli rozdeliť aj takto:

$$M = \left(\begin{array}{ccc|c} 0 & 1 & 3 & 1 \\ 1 & 2 & 4 & 1 \\ 1 & 1 & 3 & 1 \\ \hline 2 & -1 & 0 & 3 \end{array}\right).$$

Iným, veľmi jednoduchým, príkladom zápisu matice ako blokovej matice by bolo rozdelenie matici na bloky veľkosti 1×1.

Otázka je, či nám takéto rozdelenie môže nejako pomôcť pri výpočte súčinu matíc. Skúsme sa pozrieť najprv na nejaký jednoduchý prípad – opäť si zoberme nejaké dve matice veľkosti 4×4 rozdelené na bloky 2×2. Pýtame sa, či platí:

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} E & F \\ G & H \end{pmatrix} \stackrel{?}{=} \begin{pmatrix} AE + BG & AF + BH \\ CE + DG & CF + DH \end{pmatrix}$$

Ako $A, B, \ldots, H$ sme označili matice rozmerov 2×2. Vďaka tomu všetky maticové súčiny na pravej strane majú zmysel a dostaneme tam maticu zloženú zo štyroch blokov veľkosti 2×2, teda maticu správnych rozmerov.

Ak chceme overiť uvedenú rovnosť, mohli by sme to urobiť tak, že porovnáme jednotlivé prvky v oboch maticiach. Pozrime sa na prvok na pozícii $(1, 1)$ v súčine. Tento prvok dostaneme pomocou prvkov v prvom riadku prvej matice, čo sú konkrétne $a_{11}, a_{12}, b_{11}, b_{12}$; a prvkov v prvom stĺpci druhej matice, teda $e_{11}, e_{21}, f_{11}, f_{21}$. Na pozícii $(1, 1)$ potom dostaneme

$$a_{11}e_{11} + a_{12}e_{21} + b_{11}f_{11} + b_{12}f_{21}.$$

Všimnime si, že tento výraz môžeme prepísať ako $(a_{11}e_{11}+a_{12}e_{21})+(b_{11}f_{11}+b_{12}f_{21})$. V tomto zápise je prvý sčítanec presne ľavý horný prvok matice AE a druhý sčítanec je ľavý horný prvok matice BG.

Teda naozaj v ľavom hornom rohu dostávame to isté, ak urobíme súčin matíc aj ak sa pozrieme na maticu $AE + BG$.

Nemalo by nás veľmi prekvapiť, že to funguje (a že to takto bude fungovať aj všeobecne). V skutočnosti sme vlastne len rozdelili prvky v prvom riadku na prvky patriace matici A a prvky patriace matici B. To isté sme spravili v prvom stĺpci druhej matice.

Všeobecne by sme mohli sformulovať to, že matice sa dajú násobiť po blokoch, asi takto:

Tvrdenie 5.10.1. *Nech*

$$A = \begin{pmatrix} A_{11} & \ldots & A_{1n} \\ \ldots\ldots\ldots\ldots\ldots \\ A_{m1} & \ldots & A_{mn} \end{pmatrix} B = \begin{pmatrix} B_{11} & \ldots & B_{1k} \\ \ldots\ldots\ldots\ldots\ldots \\ A_{n1} & \ldots & A_{nk} \end{pmatrix}$$

sú blokové matice a navyše pre ľubovoľné prípustné i, j, k majú príslušné bloky A_{ij} a B_{jk} také rozmery, že sa tieto matice dajú násobiť. Potom ich súčin $C = AB$ sa dá zapísať ako bloková matica pozostávajúca z $m \times k$ blokov

$$C = \begin{pmatrix} A_{11} & \ldots & A_{1k} \\ \ldots\ldots\ldots\ldots\ldots \\ A_{m1} & \ldots & A_{mk} \end{pmatrix}$$

pričom

$$C_{ij} = \sum_{s=1}^{n} A_{is} B_{sj}$$

pre $i = 1, \ldots, m$, $j = 1, \ldots, k$.

Toto tvrdenie nebudeme dokazovať – dôkaz by bol len cvičením na zápis výrazov s viacerými indexami a asi by iba zahmlil základnú myšlienku, ktorú sme si už vysvetlili. Môžeme si všimnúť, že špeciálne prípady výsledku, ktorý sme sa tu snažili sformulovať všeobecne, sme už na niektorých miestach použili.

Napríklad výpočet v poznámke 5.4.10 je vlastne súčin blokových matíc (aj keď bloky v prvej matici sú veľkosti 1×1.)

Podobne sa dá pozerať na to, čo sme robili v časti 5.6 týkajúcej sa súvisu elementárnych riadkových operácií so súčinom matíc. Konkrétne sa tam vyskytol krok, ktorý môžeme zapísať aj ako $E(A|I) = (EA|EI)$. (To, že urobiť riadkovú operáciu na celej matici, alebo zvlášť na oboch častiach matice rozdelenej uvedeným spôsobom, nie je nič prekvapivé. Tu to uvádzame len ako ďalšiu ilustráciu súčinu blokových matíc.)

Takisto niektoré argumenty použité v dôkaze vety 5.7.11 predstavujú vlastne násobenie blokových matíc. (Opäť ide o pomerne jednoduchý prípad, keďže jednu z matíc sme rozdelili iba na jediný blok a druhú na stĺpce.)

Podobné argumenty sa vyskytnú napríklad aj v dôkaze vety 6.4.1 o determinante súčinu matíc.

Kapitola 6

Determinanty

{CHDETER}

6.1 Motivácia

Na začiatku tejto kapitoly uvedieme dva motivačné príklady. Ako neskôr uvidíme, v oboch z nich určitým spôsobom vystupujú determinanty, jeden z nich nám ponúka aj istú geometrickú predstavu o pojme determinantu.

{det:PRSUST}

Príklad 6.1.1. Pokúsme sa nájsť všeobecné riešenie sústavy

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= c_1 \\ a_{21}x_1 + a_{22}x_2 &= c_2 \end{aligned}$$

Prvú rovnicu vynásobíme a_{22} a odčítame od nej a_{12}-násobok druhej rovnice. Dostaneme:

$$\begin{aligned} (a_{11}a_{22} - a_{12}a_{21})x_1 &= c_1a_{22} - c_2a_{12} \\ x_1 &= \frac{c_1a_{22} - c_2a_{12}}{a_{11}a_{22} - a_{12}a_{21}} \end{aligned}$$

v prípade, že $a_{11}a_{22} - a_{12}a_{21} \neq 0$.

Ak zavedieme označenie

$$\begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{vmatrix} = b_{11}b_{22} - b_{12}b_{21},$$

tak predchádzajúcu rovnosť môžeme vyjadriť ako

$$x_1 = \frac{\begin{vmatrix} c_1 & a_{12} \\ c_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}.$$

Podobným spôsobom by sme mohli odvodiť, že platí

$$x_2 = \frac{\begin{vmatrix} a_{11} & c_1 \\ a_{21} & c_2 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}.$$

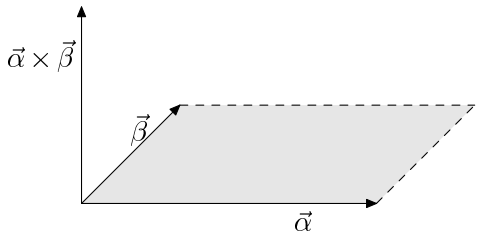
Definíciu, ktorú sme zaviedli v predchádzajúcom príklade, neskôr rozšírime aj na štvorcové matice väčších rozmerov ako 2×2 a práve tento výraz budeme nazývať determinant.

{det:PROBJEM}

Príklad 6.1.2. Dva vektory v rovine určujú rovnobežník. Zo strednej školy viete, že jeho obsah možno vypočítať pomocou vektorového súčinu (obrázok 6.1). Konkrétne, ak je rovnobežník určený vektormi $\vec{\alpha} = (a_{11}, a_{12})$ a $\vec{\beta} = (a_{21}, a_{22})$, tak tieto vektory najprv doplníme treťou súradnicou 0 na vektory $(a_{11}, a_{12}, 0)$ a $(a_{21}, a_{22}, 0)$ a potom vypočítame ich vektorový súčin $(0, 0, a_{11}a_{22} - a_{12}a_{21})$. Obsah rovnobežníka je veľkosť vektora, ktorý sme vyrátali, čiže

$$S = |a_{11}a_{22} - a_{12}a_{21}|.$$

Až na znamienko sme opäť dostali výraz z predchádzajúceho príkladu (čiže determinant). (Pričom znamienko má tiež svoj význam – určuje orientáciu vektorov.)



{det:FIGVEKTSUC} Obr. 6.1: Vektorový súčin

Pokúsme sa ešte pokúsiť o riešenie analogickej úlohy v trojrozmernom priestore. V tomto prípade 3 vektory $\vec{\alpha}$, $\vec{\beta}$, $\vec{\gamma}$ určia rovnobežnosten. Jeho objem by sme vedeli vyrátať ako súčin obsahu podstavy a jeho výšky. Pritom výšku môžeme určiť ako priemet vektora $\vec{\gamma}$ do smeru vektora $\vec{\alpha} \times \vec{\beta}$. Tento priemet sa dá vyrátať ako $|\vec{\gamma}| \cos \alpha$, kde α je uhol, ktorý zvierajú vektory $\vec{\alpha} \times \vec{\beta}$ a $\vec{\gamma}$.

Teda až na znamienko je jeho objem určený výrazom

$$|\vec{\alpha} \times \vec{\beta}||\vec{\gamma}| \cos \alpha = (\vec{\alpha} \times \vec{\beta}).\vec{\gamma},$$

kde $\cdot$ označuje skalárny súčin vektorov. (Budeme sa ním zaoberať neskôr, ale už ste sa s ním stretli aj na strednej škole a poznáte niektoré jeho základné vlastnosti.)

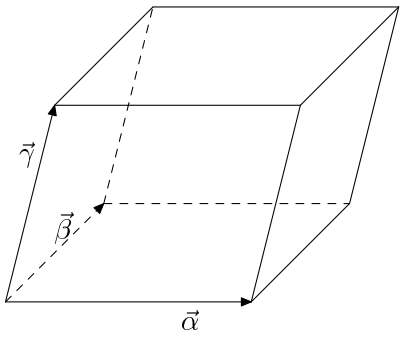
Pokúsme sa vyčísliť tento výraz pre

$$\vec{\alpha} = (a_{11}, a_{12}, a_{13})$$
$$\vec{\beta} = (a_{21}, a_{22}, a_{23})$$
$$\vec{\gamma} = (a_{31}, a_{32}, a_{33})$$

Vieme, že $\vec{\alpha} \times \vec{\beta} = (\left|\begin{smallmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{smallmatrix}\right|, -\left|\begin{smallmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{smallmatrix}\right|, \left|\begin{smallmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{smallmatrix}\right|)$.

Dostaneme teda

$$a_{31}\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} - a_{32}\begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix} + a_{33}\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} =$$
$$a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}$$



VNOBEZNOSTEN}

Obr. 6.2: Rovnobežnosten určený 3 vektormi v $\mathbb{R}^3$

Odvodili sme výrazy, ktoré predstavujú determinant štvorcovej matice 2×2 a 3×3. Ako by sme mohli túto definíciu rozšíriť na vyššie rozmery?

Mohli by sme postupovať analogicky ako v predchádzajúcich príkladoch. Jedna možnosť, ako by sme definovali determinanty, by bolo všeobecné riešenie rovníc vyšších stupňov. Vychádzali by nám síce čoraz komplikovanejšie výrazy, ale snáď by sme v nich časom objavili nejakú zákonitosť.

Príklad 6.1.2 nám dáva dobrú geometrickú predstavu determinantu – ako objem telesa určeného danými vektormi v $\mathbb{R}^n$. Nie je však úplne jasné, čo chápať pod objemom v n-rozmernom priestore. Ale snáď by sme na to vedeli prísť. Objem jednotkovej kocky – čiže n-rozmerného rovnobežnostena určeného vektormi $\vec{\varepsilon}_1, \ldots, \vec{\varepsilon}_n$ – je zrejme 1. A vieme celkom dobre popísať (na základe analógie s dvojrozmerným a trojrozmerným prípadom), ako sa tento objem zmení pri transformáciach ako je skosenie, natiahnutie v smere niektorej z jeho strán alebo súmernosť podľa roviny (či skôr „nadroviny“, ako sa zvykne nazývať $(n-1)$-rozmerný podpriestor v $\mathbb{R}^n$). A pomocou týchto transformácií by sme vedeli dostať z jednotkovej kocky ľubovoľný n-rozmerný rovnobežnosten (prinajmenšom v trojrozmere máme o tom celkom dobrú geometrickú predstavu; neskôr si niečo povieme aj o tom ako súvisia tieto transformácie s riadkovými operáciami na matici), takže takýmto spôsobom by sme tiež boli schopný vyrátať objem každého rovnobežnostena – a teda ľubovoľný determinant.

Nebudeme postupovať ani jedným z naznačených spôsobov – naša definícia determinantu bude celkom iná a na prvý pohľad veľmi zvláštna. Neskôr však uvidíme, že pri našej definícií budú platiť pre riešenie sústavy n rovníc o n neznámych analogické vzťahy ako sme dostali v príklade 6.1.1 a takisto sa determinant správa vzhľadom na niektoré transformácie spôsobom, ktorý sme pred chvíľou spomenuli.

6.2 Definícia determinantu

Definícia 6.2.1. V tejto kapitole budeme označovať ako S_n množinu všetkých permutácií množiny $\{1, 2, \ldots, n\}$.

Dvojica $(\varphi(k), \varphi(s))$ sa volá *inverzia* permutácie φ, ak $k < s$ ale $\varphi(k) > \varphi(s)$. Počet inverzií permutácie φ budeme označovať $i(\varphi)$.

Príklad 6.2.2. Permutácia $\left(\begin{smallmatrix}1&2&3&4\\4&1&3&2\end{smallmatrix}\right)$ má 4 inverzie (4,1), (4,3), (4,2), (3,2):

$$\begin{pmatrix}1&2&3&4\\4&1&3&2\end{pmatrix}$$
$$\begin{array}{cccc}4&1&&\\4&&3&\\4&&&2\\&&3&2\end{array}$$

{det:DEF}

Definícia 6.2.3. Nech A je matica typu $n\times n$ nad poľom F, $A=||a_{ij}||$. *Determinant matice* A je

{det:EQDEF}
$$|A|=\sum_{\varphi\in S_n}(-1)^{i(\varphi)}a_{1\varphi(1)}a_{2\varphi(2)}\dots a_{n\varphi(n)}. \tag{6.1}$$

Symbolom $\sum_{\varphi\in S_n}$ rozumieme, že sčitujeme cez celú množinu S_n, teda pre každú permutáciu $\varphi\in S_n$ pripočítame jeden sčítanec uvedeného tvaru. (Množina S_n je konečná, teda takýto súčet je jednoznačne definovaný.)

Príklad 6.2.4. Ak máme maticu typu 1×1 tak dostaneme $|A|=|a_{11}|=a_{11}$.

Pre maticu typu 2×2 tiež dostaneme jednoduchý vzorec. S_2 obsahuje len 2 permutácie $\left(\begin{smallmatrix}1&2\\1&2\end{smallmatrix}\right)a\left(\begin{smallmatrix}1&2\\2&1\end{smallmatrix}\right)$. Prvá z nich nemá inverzie, druhá má jednu inverziu. Z (6.1) dostaneme v tomto prípade

$$\begin{vmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{vmatrix}=a_{11}a_{22}-a_{12}a_{21}.$$

V prípade matice typu 3×3 obsahuje množina S_3 už 6 prvkov. V nasledujúcej tabuľke sú vypísané všetky permutácie z S_3, ich inverzie a počet inverzií.

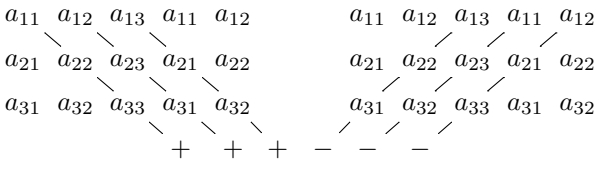
φ	$i(\varphi)$	inverzie
$\left(\begin{smallmatrix}1&2&3\\1&2&3\end{smallmatrix}\right)$	0	
$\left(\begin{smallmatrix}1&2&3\\1&3&2\end{smallmatrix}\right)$	1	$(3,2)$
$\left(\begin{smallmatrix}1&2&3\\2&1&3\end{smallmatrix}\right)$	1	$(2,1)$
$\left(\begin{smallmatrix}1&2&3\\2&3&1\end{smallmatrix}\right)$	2	$(2,1)$ $(3,1)$
$\left(\begin{smallmatrix}1&2&3\\3&1&2\end{smallmatrix}\right)$	2	$(3,1)$ $(3,2)$
$\left(\begin{smallmatrix}1&2&3\\3&2&1\end{smallmatrix}\right)$	3	$(3,2)$ $(3,1)$ $(2,1)$

Z (6.1) teda dostaneme (permutácie uvedieme v takom poradí, aby išli najprv tie, ktoré majú kladné znamienko, čiže párny počet inverzií)

$$\begin{vmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{vmatrix}=a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{11}a_{23}a_{32}-a_{12}a_{21}a_{33}-a_{13}a_{22}a_{31}.$$

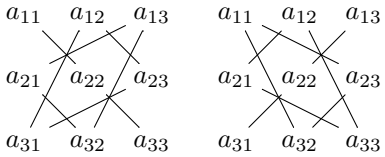
Tento vzorec pre výpočet determinantu matice 3×3 nazývame *Sarrusovo pravidlo.*

Spôsob, ako si ho možno ľahko zapamätať, je pridať ešte raz k matici A je prvé 2 stĺpce. Súčiny na hlavných diagonálach dávame do vzorca s kladným znamienkom, súčiny na vedľajších diagonálach so záporným znamienkom.



Ten istý výsledok dostaneme ako pod maticu podpíšeme ešte raz jej prvé 2 riadky.

Inou mnemotechnickou pomôckou je vyznačiť si „kladné“ a „záporné“ diagonály v pôvodnej matici – bez pripisovania prvkov matice.



Príklad 6.2.5. $\begin{vmatrix} 1 & 3 & 2 \\ 1 & 3 & 1 \\ 0 & 1 & 2 \end{vmatrix} = 1.3.2 + 3.1.0 + 2.1.1 - 2.3.0 - 1.1.1 - 2.3.1 = 1$

Priamo z definície sa dá dokázať užitočná vlastnosť determinantu: determinant transponovanej matice je rovnaký ako determinant pôvodnej matice.

{det:VTTRANS}

Veta 6.2.6. *Nech A je matica typu $n \times n$. Potom*

$$|A| = |A^T|.$$

Dôkaz. V definícii determinantu (6.1) vystupujú súčiny tvaru $a_{1\varphi(1)}a_{2\varphi(2)}\dots a_{n\varphi(n)}$. Okamžite vidíme, že v takomto súčine sa objaví práve raz prvok prvého riadku matice A (konkrétne na prvom mieste), práve raz prvok druhého riadku, atď.

Ako je to so stĺpcami? Druhé súradnice, ktoré predstavujú stĺpce, sú $\varphi(1), \varphi(2), \dots, \varphi(n)$. Vďaka tomu, že φ je bijekcia, objaví sa každý stĺpec práve raz. Konkrétne prvok j-teho stĺpca sa vyskytne v činiteli $a_{\varphi^{-1}(j)j}$ (lebo $i = \varphi^{-1}(j)$ je presne to číslo, ktoré sa zobrazí na j, čiže spĺňa $\varphi(i) = j$).

Ak teda usporiadame činitele vystupujúce v takomto súčine nie podľa prvých, ale podľa druhých súradnic, dostaneme iný zápis pre ten istý súčin.

$$a_{1\varphi(1)}a_{2\varphi(2)}\dots a_{n\varphi(n)} = a_{\varphi^{-1}(1)1}a_{\varphi^{-1}(2)2}\dots a_{\varphi^{-1}(n)n}.$$

Sčítaním takýchto rovností cez všetky permutácie $\varphi \in S_n$ dostaneme

$$|A| = \sum_{\varphi\in S_n} (-1)^{i(\varphi)} a_{1\varphi(1)}a_{2\varphi(2)}\dots a_{n\varphi(n)} = \sum_{\varphi\in S_n} (-1)^{i(\varphi)} a_{\varphi^{-1}(1)1}a_{\varphi^{-1}(2)2}\dots a_{\varphi^{-1}(n)n}.$$

Označme prvok v i-tom riadku a j-tom stĺpci transponovanej matice ako a'_{ij}, t.j. $a'_{ij} = a_{ji}$.

Potom poslednú rovnosť môžeme prepísať ako

$$|A| = \sum_{\varphi\in S_n} (-1)^{i(\varphi)} a'_{1\varphi^{-1}(1)}a'_{2\varphi^{-1}(2)}\dots a'_{n\varphi^{-1}(n)}.$$

Ďalej si uvedomme, že priradenie $\varphi \mapsto \varphi^{-1}$ je bijekcia z S_n do S_n. (Ľahko to môžete overiť na základe vlastností inverzného zobrazenia. Vyplýva to napríklad aj z toho, že $(S_n, \circ)$ je grupa (úloha 3.2.2) a z úlohy 3.2.12.) Teda, ak v predchádzajúcej sume namiesto φ^{-1} použijeme všade φ, znamená to len preusporiadanie sčítancov, ale hodnotu súčtu to neovplyvní.

$$|A| = \sum_{\varphi\in S_n} (-1)^{i(\varphi^{-1})} a'_{1\varphi(1)}a'_{2\varphi(2)}\dots a'_{n\varphi(n)}.$$

Na to, aby sme vpravo dostali determinant matice A^T, stačilo by dokázať, že $i(\varphi) = i(\varphi^{-1})$. To skutočne platí. Inverzie permutácie φ sú totiž určené takými dvojicami indexov, pre ktoré platí

$$i < j \qquad \wedge \qquad \varphi(i) > \varphi(j).$$

Ak označíme $i' = \varphi(i)$ a $j' = \varphi(j)$, tak predchádzajúca podmienka je ekvivalentná podmienke

$$\varphi^{-1}(i') < \varphi^{-1}(j') \qquad \wedge \qquad i' > j'.$$

Našli sme teda jedno-jednoznačné priradenie medzi dvojicami, ktoré určujú inverzie permutácií φ a φ^{-1}. Teda skutočne platí $i(\varphi) = i(\varphi^{-1})$ a

$$|A| = |A^T|.$$

□

6.3 Výpočet determinantov

Doteraz sme uviedli ako počítať determinanty rozmeru maximálne 3×3, pričom sme vlastne postupovali priamo z definície. Na výpočet determinantov vyššieho stupňa sa naučíme dva postupy. Prvý z nich je Laplaceov rozvoj a druhý je použitie elementárnych riadkových a stĺpcových operácií.

6.3.1 Laplaceov rozvoj

Nech A je štvorcová matica typu $n \times n$. Zvoľme si (pevne) nejaké $i \in \{1, 2, \ldots, n\}$. Potom determinant matice A sa dá upraviť na tvar

$$|A| = a_{i1}A_{i1} + a_{i2}A_{i2} + \ldots + a_{in}A_{in}.$$

Vyplýva to z toho, že v každom sčítanci v sume (6.1) vystupuje práve jeden prvok tvaru a_{ik} (konkrétne je to $a_{i\varphi(i)}$). Aby sme získali uvedenú rovnosť, stačí vyňať a_{ij} z tých sčítancov v ktorých sa vyskytuje.

Podobne by sme mohli postupovať aj pre prvky niektorého stĺpca $a_{1j}, a_{2j}, \ldots, a_{nj}$. Dostali by sme

$$|A| = a_{1j}A_{1j} + a_{2j}A_{2j} + \ldots + a_{nj}A_{nj}.$$

Výraz A_{ij} nazývame *algebraický doplnok prvku* a_{ij}.

Naším najbližším cieľom bude zistiť, čomu sa rovná A_{ij}.

Pokúste sa sami si vyskúšať zistiť všetky možné hodnoty A_{ij} pre maticu 3×3, výsledky si môžete skontrolovať v nasledujúcom príklade.

Príklad 6.3.1. Pre maticu typu 3×3 sme odvodili

$$|A| = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}.$$

Skúsme v tomto konkrétnom prípade urobiť spomínaný rozvoj pre druhý riadok a druhý stĺpec.

Pre druhý riadok dostaneme:

$$|A| = a_{21}(a_{13}a_{32} - a_{12}a_{33}) + a_{22}(a_{11}a_{33} - a_{13}a_{31}) + a_{23}(a_{12}a_{31} - a_{11}a_{32})$$

Pre druhý stĺpec dostaneme:

$$|A| = a_{12}(a_{23}a_{31} - a_{21}a_{33}) + a_{22}(a_{11}a_{33} - a_{13}a_{31}) + a_{32}(a_{13}a_{21} - a_{11}a_{23})$$

Z uvedených výrazov vieme vyčítať hodnoty A_{12}, A_{21}, A_{22}, A_{23}, A_{32}. Keby sme urobili ešte 2 ďalšie rozvoje (povedzme podľa prvého a tretieho riadku), zistili by sme aj ostatné hodnoty algebraických doplnkov.

$$A_{11} = a_{22}a_{33} - a_{23}a_{32}$$
$$A_{12} = a_{23}a_{31} - a_{21}a_{33}$$
$$A_{13} = a_{21}a_{32} - a_{22}a_{31}$$
$$A_{21} = a_{13}a_{32} - a_{12}a_{33}$$
$$A_{22} = a_{11}a_{33} - a_{13}a_{31}$$
$$A_{23} = a_{12}a_{31} - a_{11}a_{32}$$
$$A_{31} = a_{12}a_{23} - a_{13}a_{22}$$
$$A_{32} = a_{13}a_{21} - a_{11}a_{23}$$
$$A_{33} = a_{11}a_{22} - a_{12}a_{21}$$

Možno by ste z hodnôt vyrátaných v predchádzajúcom príklade vedeli uhádnuť nejakú všeobecnú zákonitosť. Váš tip sa potvrdí dokázaním nasledujúcej vety.

Pre maticu A typu $n \times n$ a $i, j \in \{1, 2, \ldots, n\}$ označíme ako M_{ij} maticu, ktorá vznikne z matice A vynechaním i-teho riadku a j-teho stĺpca.

Veta 6.3.2. *Pre algebraický doplnok prvku a_{rs} štvorcovej matice A platí*

$$A_{rs} = (-1)^{r+s}|M_{rs}|$$

Dôkaz. Priamo z definície vyplýva, že

$$A_{rs} = \sum_{\varphi' \in S_n^{rs}} (-1)^{i(\varphi')} a_{1,\varphi'(1)} a_{2\varphi'(2)} \ldots a_{r-1,\varphi'(r-1)} a_{r+1,\varphi'(r+1)} \ldots a_{n,\varphi'(n)},$$

kde ako S_n^{rs} sme označili množinu tých permutácií z S_n, pre ktoré $\varphi'(r) = s$.

Pre maticu M_{rs} tejto permutácii zodpovedá permutácia φ, ktorá je určená predpisom

$$\varphi(k) = \begin{cases} \varphi'(k), & \text{ak } \varphi'(k) < s \\ \varphi'(k) - 1, & \text{ak } \varphi'(k) > s \end{cases} \quad \text{pre } k < r$$

$$\varphi(k) = \begin{cases} \varphi'(k+1), & \text{ak } \varphi'(k+1) < s \\ \varphi'(k+1) - 1, & \text{ak } \varphi'(k+1) > s \end{cases} \quad \text{pre } k \geq r.$$

Chceli by sme zistiť, aký je vzťah medzi $i(\varphi')$ a $i(\varphi)$. Každá inverzia permutácie φ zodpovedá nejakej inverzii pôvodnej permutácie φ'. Niektoré inverzie sme však stratili:
a) Ak platí $\varphi'(j) > s$ pre $j < r$, tak dvojica $(\varphi'(j), s)$ tvorí inverziu pôvodnej permutácie, ale v permutácií φ nemáme inverziu, ktorá by jej zodpovedala. Označme počet takýchto inverzií k. Znamená to, že medzi prvkami $j \in \{1, 2, \ldots, r-1\}$ je k prvkov takých, že $\varphi'(j) > s$. Pre zostávajúcich $(r-1) - k$ prvkov teda platí $\varphi'(j) < s$. (Nemôže platiť $\varphi'(j) = s$, lebo na s sa zobrazí jedine r.)
b) Ak $\varphi'(j) < s$ pre $j > r$, tak opäť dostaneme inverziu, ku ktorej nemáme zodpovedajúcu inverziu v permutácii φ. Prvkov s vlastnosťou $\varphi'(j) < s$ je práve $s - 1$. Pritom, ako sme videli v prípade a), z nich práve $(r-1) - k$ spĺňa nerovnosť $j < r$. Inverzií typu b) je teda $(s-1) - [(r-1) - k] = s - r + k$.

Spolu sme „stratili“ $s - r + 2k$ inverzií. Teda $i(\varphi') = i(\varphi) + s - r + 2k$ a

$$(-1)^{i(\varphi')} = (-1)^{i(\varphi)} + (-1)^{s-r} + (-1)^{2k} = (-1)^{i(\varphi)} + (-1)^{s+r}.$$

Dosadením do vyjadrenia pre A_{rs}, ak navyše zavedieme označenie $M_{rs} = ||b_{ij}||$, získame dokazovanú rovnosť

$$A_{rs} = (-1)^{r+s} \sum_{\varphi \in S_{n-1}} (-1)^{i(\varphi)} b_{1\varphi(1)} \ldots b_{n\varphi(n)} = (-1)^{r+s}|M_{ij}|.$$

□

Pretože predchádzajúci dôkaz nie je úplne jednoduchý, uvedieme ešte iný, do istej miery podobný (v nádeji, že keď uvidíte viacero pohľadov na dôkaz tej istej vety, bude to o čosi jasnejšie).

Dôkaz. Kvôli jednoduchosti začnime s prípadom $r = s = n$. Priamo z definície determinantu vidíme, že a_{nn} sa vyskytne v tých sčítancoch, kde $\varphi(n) = n$. To ale znamená, že permutácia φ zobrazí $\{1, 2, \ldots, n-1\}$ na $\{1, 2, \ldots, n-1\}$, čiže takéto permutácie sú v jednojednoznačnej korešpondencii s permutáciami množiny $\{1, 2, \ldots, n-1\}$. Navyše počet inverzií zodpovedajúcich si permutácií je očividne rovnaký. (Vynechali sme len $\varphi(n) = n$, tento prvok nevystupoval v žiadnej inverzii.) Máme teda

$$A_{nn} = \sum_{\varphi \in S_{n-1}} (-1)^{i(\varphi)} a_{1\varphi(1)} a_{2\varphi(2)} \ldots a_{n-1,\varphi(n-1)} = |M_{nn}|.$$

(Opäť, priamo z definície dostaneme, že výraz na pravej strane rovnosti je determinant matice, ktorá vznikne vynechaním posledného riadku a posledného stĺpca.)

Ďalej sa pozrime na to, ako sa zmení algebraický doplnok A_{rs} ak vymeníme r-tý riadok matice A s nasledujúcim. Nech teda matica B vznikne z matice A tak, že vymeníme r-tý a $(r+1)$-vý riadok. Chceme porovnať sčítance vystupujúce v determinante matice A obsahujúce prvok a_{rs} s tými sčítancami v determinante matice B, ktoré obsahujú $b_{r+1,s} = a_{r,s}$. V prvom prípade sčitujeme cez všetky permutácie také, že $\varphi(r) = s$ (množinu týchto permutácií označíme opäť S_n^{rs}).

$$A_{rs} = \sum_{\varphi \in S_n^{rs}} (-1)^{i(\varphi)} a_{1,\varphi(1)} \ldots a_{r-1,\varphi(r-1)} a_{r+1,\varphi(r+1)} \ldots a_{n,\varphi(n)}.$$

Pri výpočte $B_{r+1,s}$ sčitujeme cez všetky permutácie také, že $\varphi(r+1) = s$:

$$B_{r+1,s} = \sum_{\varphi \in S_n^{r+1,s}} (-1)^{i(\varphi)} b_{1,\varphi(1)} \ldots b_{r-1,\varphi(r-1)} b_{r,\varphi(r)} b_{r+2,\varphi(r+2)} \ldots b_{n,\varphi(n)}.$$

Súčasne, z definície matice B máme $b_{ij} = a_{ij}$ pre $i \neq r, r+1$ a $b_{rs} = a_{r+1,s}$, teda

$$B_{r+1,s} = \sum_{\varphi \in S_n^{r+1,s}} (-1)^{i(\varphi)} a_{1,\varphi(1)} \ldots a_{r-1,\varphi(r-1)} a_{r+1,\varphi(r)} a_{r+2,\varphi(r+2)} \ldots a_{n,\varphi(n)}.$$

Uvedomme si ďalej, že výmena prvkov na r-tej a $(r+1)$-vej pozícii dáva korešpondenciu medzi množinami S_n^{rs} a $S_n^{r+1,s}$. Konkrétne, z permutácie

$$\varphi = \begin{pmatrix} 1 & \ldots & r & r+1 & \ldots & n \\ \varphi(1) & \ldots & s & \varphi(r+1) & \ldots & \varphi(n) \end{pmatrix}$$

patriacej do S_n^{rs} dostaneme permutáciu

$$\varphi' = \begin{pmatrix} 1 & \ldots & r & r+1 & \ldots & n \\ \varphi(1) & \ldots & \varphi(r+1) & s & \ldots & \varphi(n) \end{pmatrix}$$

patriacu do $S_n^{r+1,s}$ a obrátene. Vďaka tejto bijektívnej korešpondencii môžeme predchádzajúcu sumu prepísať ako

$$B_{r+1,s} = \sum_{\varphi \in S_n^{r,s}} (-1)^{i(\varphi')} a_{1,\varphi(1)} \ldots a_{r-1,\varphi(r-1)} a_{r+1,\varphi(r+1)} a_{r+2,\varphi(r+2)} \ldots a_{n,\varphi(n)}.$$

Vidíme, že v oboch sumách sa vyskytujú presne tie isté členy, zostáva len zistiť aký je vzťah medzi $i(\varphi)$ a $i(\varphi')$. Tieto permutácie sa líšia len na r-tom a $(r+1)$-mieste, čiže jediná inverzia, ktorou sa môžu líšiť, je $(\varphi(r), \varphi(r+1))$. Skutočne, ak $(s, \varphi(r+1))$ tvoria inverziu permutácie φ, tak v novej permutácii nebudeme mať na tomto mieste inverziu a obrátene, ak tu φ nemá inverziu, vo φ' dostaneme inverziu. Počet permutácií sa teda líši o túto jednu inverziu, čiže $i(\varphi') = i(\varphi) \pm 1$, a teda

$$B_{r+1,s} = -\sum_{\varphi \in S_n^{r,s}} (-1)^{i(\varphi)} a_{1,\varphi(1)} \dots a_{r-1,\varphi(r-1)} a_{r+1,\varphi(r+1)} a_{r+2,\varphi(r+2)} \dots a_{n,\varphi(n)} = -A_{rs}.$$

Vďaka vete 6.2.6 vieme, že to isté sa stane pri výmene susedných stĺpcov.

Posledné pozorovanie potrebné na dokončenie dôkazu je, že podmatica M_{rs} matice A je rovnaká ako podmatica $M_{r+1,s}$ matice B. (Z matice B vynechávame $(r+1)$-vý riadok, čo je presne r-tý riadok pôvodnej matice.)

Ak teraz chceme zistiť algebraický doplnok A_{rs} prvku a_{rs} matice A, môžeme postupovať tak, že $n-r-1$ výmenami susedných riadkov presunieme prvok a_{rs} do n-tého riadku, a potom urobíme ešte $n-s-1$ výmen stĺpcov, po ktorých prvok a_{rs} pôvodnej matice bude prvkom c_{nn} matice C, ktorú takto dostaneme. Už vieme, že algebraický doplnok tohoto prvku je

$$A_{nn} = |M_{rs}|$$

(podmatica, ktorá vznikne z C vynechaním posledného riadku a posledného stĺpca je presne tá podmatica pôvodnej matice, ktorú sme dostali vynechaním r-tého riadku a s-tého stĺpca.) Súčasne sme pri každej výmene riadku/stĺpca zmenili znamienko, preto

$$A_{rs} = (-1)^{n-r+n-s} A_{nn} = (-1)^{2(n-r-s)+r+s} A_{nn} = (-1)^{r+s} A_{nn} = (-1)^{r+s} |M_{rs}|.$$

□

Dôsledok 6.3.3 (Laplaceov rozvoj determinantu)**.** *Nech A je matica typu $n \times n$. Potom*

$$|A| = (-1)^{i+1} a_{i1} |M_{i1}| + (-1)^{i+2} a_{i2} |M_{i2}| + \dots + (-1)^{i+n} a_{in} |M_{in}| \quad (6.2)$$ {det:EQLAPR}

$$|A| = (-1)^{j+1} a_{1j} |M_{1j}| + (-1)^{j+2} a_{2j} |M_{2j}| + \dots + (-1)^{j+n} a_{nj} |M_{nj}| \quad (6.3)$$ {det:EQLAPS}

Prvú rovnosť uvedenú v predchádzajúcom dôsledku nazývame *Laplaceov rozvoj determinantu matice A podľa i-teho riadku,* druhú *Laplaceov rozvoj podľa j-teho stĺpca.*

{det:PRLAPL}

Príklad 6.3.4. Nasledujúci determinant vypočítame Laplaceovým rozvojom podľa druhého stĺpca.

$$\begin{vmatrix} 2 & 2 & 0 & 1 \\ 1 & 0 & -1 & 1 \\ 2 & 3 & 1 & 1 \\ 2 & 0 & -1 & 2 \end{vmatrix} = -2 \begin{vmatrix} 1 & -1 & 1 \\ 2 & 1 & 1 \\ 2 & -1 & 2 \end{vmatrix} - 3 \begin{vmatrix} 2 & 0 & 1 \\ 1 & -1 & 1 \\ 2 & -1 & 2 \end{vmatrix} = (-2).1 - 3.(-1) = 1$$

6.3.2 Výpočet pomocou riadkových a stĺpcových operácií

{det:SSECTERO}

V časti 5.2 sme si ukázali, ako možno pomocou elementárnych riadkových úprav upraviť ľubovoľnú maticu na redukovanú trojuholníkovú maticu. Ak by sme vedeli, ako elementárne riadkové úpravy ovplyvňujú hodnotu determinantu a ak by sme vedeli vypočítať determinant redukovanej trojuholníkovej matice, tak by sme získali ďalšiu metódu na výpočet determinantov. Práve to je naším najbližším cieľom.

Začneme s tým, že overíme, ako menia hodnotu determinantu jednotlivé elementárne riadkové operácie.

{det:VTCNAS}

Veta 6.3.5. *Ak maticu B získame z A vynásobením k-teho riadku skalárom $c \in F$, tak*

$$|B| = c|A|.$$

Dôkaz. Označme $B = ||b_{ij}||$ a $A = ||a_{ij}||$. Potom platí $b_{ij} = a_{ij}$ pre $i \neq k$ a $b_{kj} = ca_{kj}$. Priamo z definície determinantu potom dostaneme

$$\begin{aligned}|B| &= \sum_{\varphi \in S_n} (-1)^{i(\varphi)} b_{1,\varphi(1)} b_{2,\varphi(2)} \dots b_{k-1,\varphi(k-1)} b_{k,\varphi(k)} b_{k+1,\varphi(k+1)} \dots b_{n,\varphi(n)} = \\ &\sum_{\varphi \in S_n} (-1)^{i(\varphi)} a_{1,\varphi(1)} a_{2,\varphi(2)} \dots a_{k-1,\varphi(k-1)} c a_{k,\varphi(k)} a_{k+1,\varphi(k+1)} \dots a_{n,\varphi(n)} = \\ &c \sum_{\varphi \in S_n} (-1)^{i(\varphi)} a_{1,\varphi(1)} a_{2,\varphi(2)} \dots a_{k-1,\varphi(k-1)} a_{k,\varphi(k)} a_{k+1,\varphi(k+1)} \dots a_{n,\varphi(n)} = c|A|.\end{aligned}$$

□

{det:DOSNULRIAD}

Dôsledok 6.3.6. *Ak matica A má nulový riadok, tak $|A| = 0$.*

Dôkaz. Stačí v predchádzajúcej vete dosadiť $c = 0$. □

{det:VTROVRIAD}

Veta 6.3.7. *Ak má matica A dva rovnaké riadky, tak $|A| = 0$.*

Dôkaz. Matematickou indukciou vzhľadom na n.

1° Pre $n = 2$ tvrdenie platí:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{11} & a_{12} \end{vmatrix} = a_{11}a_{12} - a_{11}a_{12} = 0.$$

2° Predpokladajme, že tvrdenia platí pre ľubovoľnú maticu typu $n \times n$. Nech A je matica typu $(n+1) \times (n+1)$, ktorej i-ty a j-ty riadok sú rovnaký. Ak urobíme rozvoj determinantu matice A podľa niektorého iného riadku (okrem i-teho a j-teho), tak všetky matice M_{rs} vystupujúce v Laplaceovom rozvoji sú matice typu $n \times n$ a majú dva rovnaké riadky. Podľa indukčného predpokladu $|M_{rs}| = 0$ pre všetky prípustné hodnoty r, s, a teda z Laplaceovho rozvoja (6.2) vidíme, že aj $|A| = 0$. □

{det:VTSUCRIAD}

Veta 6.3.8. *Nech matice A a B sú matice typu $n \times n$, ktoré sa líšia len v k-tom riadku. Potom $|A| + |B| = |C|$, kde $c_{ij} = a_{ij} = b_{ij}$ pre $i \neq k$ a $c_{kj} = a_{kj} + b_{kj}$.*

Dôkaz. Urobme rozvoj matice C podľa k-teho riadku. Dostaneme

$$|C| = (-1)^{k+1}(a_{k1} + b_{k1})|M_{k1}| + (-1)^{k+2}(a_{k2} + b_{k2})|M_{k2}| + \dots + (-1)^{k+n}(a_{kn} + b_{kn})|M_{kn}|.$$

Pritom podmatice $M_{k1}, M_{k2}, \dots, M_{kn}$ už pozostávajú len z tých prvkov, ktoré sú vo všetkých troch maticiach rovnaké. Teda tie istú podmatice budú vystupovať v rozvojoch matíc A a B podľa k-teho riadku. Pomocou nich dostaneme:

$$\begin{aligned}|A| &= (-1)^{k+1}a_{k1}|M_{k1}| + (-1)^{k+2}a_{k2}|M_{k2}| + \dots + (-1)^{k+n}a_{kn}|M_{kn}|, \\ |B| &= (-1)^{k+1}b_{k1}|M_{k1}| + (-1)^{k+2}b_{k2}|M_{k2}| + \dots + (-1)^{k+n}b_{kn}|M_{kn}|.\end{aligned}$$

Sčítaním týchto dvoch rovností a porovnaním s rozvojom determinantu matice C dostaneme

$$|A| + |B| = |C|.$$

□

Predchádzajúcu vetu by sme mohli ľahko overiť aj priamo na základe definície determinantu.

Dôkaz.

$$|C| = \sum_{\varphi\in S_n} (-1)^{i(\varphi)} c_{1\varphi(1)} \dots c_{k-1,\varphi(k-1)} c_{k\varphi(k)} c_{k+1,\varphi(k+1)} \dots c_{n\varphi(n)} =$$
$$\sum_{\varphi\in S_n} (-1)^{i(\varphi)} a_{1\varphi(1)} \dots a_{k-1,\varphi(k-1)} (a_{k\varphi(k)} + b_{k\varphi(k)}) a_{k+1,\varphi(k+1)} \dots a_{n\varphi(n)} =$$
$$\sum_{\varphi\in S_n} (-1)^{i(\varphi)} a_{1\varphi(1)} \dots a_{k-1,\varphi(k-1)} a_{k\varphi(k)} a_{k+1,\varphi(k+1)} \dots a_{n\varphi(n)} +$$
$$\sum_{\varphi\in S_n} (-1)^{i(\varphi)} a_{1\varphi(1)} \dots a_{k-1,\varphi(k-1)} b_{k\varphi(k)} a_{k+1,\varphi(k+1)} \dots a_{n\varphi(n)} = |A| + |B|$$

□

{det:VTPLUSCNAS}

Veta 6.3.9. *Ak matica B vznikne z A pripočítaním c-násobku niektorého riadku k inému (pričom $c \in F$), tak $|B| = |A|$.*

Dôkaz. Nech B vznikne z A pripočítaním c-násobku k-teho riadku k l-temu riadku, pričom $k \neq l$. Teda B má všetky riadky rovnaké ako matica A, len prvky l-teho riadku majú tvar $b_{lj} = ca_{kj} + a_{lj}$.

Nech A' je matica, ktorá má všetky riadky rovnaké ako matica A len l-ty riadok matice A' sa rovná k-temu riadku matice A. (Teda $a'_{ij} = a_{ij}$ pre $i \neq l$ a $a'_{lj} = a_{kj}$.) Táto matica má dva rovnaké riadky, teda podľa vety 6.3.7 je $|A'| = 0$.

Uvažujme ďalej maticu A'', ktorá vznikne z A' vynásobením k-teho riadku skalárom c. (Teda $a''_{ij} = a_{ij}$ pre $i \neq l$ a $a''_{lj} = ca_{kj}$.) Z vety 6.3.5 máme $|A''| = c|A'| = 0$.

Teraz si stačí všimnúť, že maticu B dostaneme z matíc A a A'' spôsobom popísaným vo vete 6.3.8. Teda

$$|B| = |A| + |A''| = |A|.$$

□

Veta 6.3.10. *Ak matica B vznikne z A vzájomnou výmenou dvoch riadkov, tak $|B| = -|A|$. (Výmena 2 riadkov matice mení znamienko determinantu.)*

Dôkaz. Označme $\vec{\alpha}_1, \dots, \vec{\alpha}_n$ riadky matice A. Uvažujme maticu, ktorá má rovnaké riadky ako A, len namiesto i-teho aj namiesto j-teho riadku má $\vec{\alpha}_i + \vec{\alpha}_j$. Pretože táto matica má rovnaký i-ty a j-ty riadok, podľa vety 6.3.7

$$\begin{vmatrix} \vec{\alpha}_1 \\ \vec{\alpha}_i + \vec{\alpha}_j \\ \vec{\alpha}_i + \vec{\alpha}_j \\ \vec{\alpha}_n \end{vmatrix} = 0$$

(Kvôli stručnosti označenia sme vynechali sme všetky riadky, ktoré sú rovnaké ako v matici A, okrem prvého a posledného. Toto označenie rozhodne nie je korektné, snáď je však dostatočne zrozumiteľné.)

Tento determinant súčasne vieme prepísať pomocou vety 6.3.8 a vety 6.3.7

$$\begin{vmatrix} \vec{\alpha}_1 \\ \vec{\alpha}_i + \vec{\alpha}_j \\ \vec{\alpha}_i + \vec{\alpha}_j \\ \vec{\alpha}_n \end{vmatrix} = \begin{vmatrix} \vec{\alpha}_1 \\ \vec{\alpha}_i + \vec{\alpha}_j \\ \vec{\alpha}_i \\ \vec{\alpha}_n \end{vmatrix} + \begin{vmatrix} \vec{\alpha}_1 \\ \vec{\alpha}_i + \vec{\alpha}_j \\ \vec{\alpha}_j \\ \vec{\alpha}_n \end{vmatrix} = \begin{vmatrix} \vec{\alpha}_1 \\ \vec{\alpha}_i \\ \vec{\alpha}_i \\ \vec{\alpha}_n \end{vmatrix} + \begin{vmatrix} \vec{\alpha}_1 \\ \vec{\alpha}_j \\ \vec{\alpha}_i \\ \vec{\alpha}_n \end{vmatrix} + \begin{vmatrix} \vec{\alpha}_1 \\ \vec{\alpha}_i \\ \vec{\alpha}_j \\ \vec{\alpha}_n \end{vmatrix} + \begin{vmatrix} \vec{\alpha}_1 \\ \vec{\alpha}_j \\ \vec{\alpha}_j \\ \vec{\alpha}_n \end{vmatrix} = \begin{vmatrix} \vec{\alpha}_1 \\ \vec{\alpha}_j \\ \vec{\alpha}_i \\ \vec{\alpha}_n \end{vmatrix} + \begin{vmatrix} \vec{\alpha}_1 \\ \vec{\alpha}_i \\ \vec{\alpha}_j \\ \vec{\alpha}_n \end{vmatrix}$$

Porovnaním týchto 2 vzťahov dostaneme

$$0 = |A| + |B|,$$

čiže skutočne $|B| = -|A|$. □

Teraz už vieme, ako ovplyvňujú hodnotu determinantu jednotlivé elementárne riadkové operácie. Podľa vety 6.2.6 platí $|A| = |A^T|$. Pretože riadkové operácie použité na transponovanú maticu A^T zodpovedajú stĺpcovým operáciam na pôvodnej matici A, všetky dokázané tvrdenia platia aj pre stĺpcové operácie. (Pri výpočte determinantov môžeme teda kombinovať riadkové aj stĺpcové operácie.)

Doteraz dokázané vety nám však len umožňujú porovnať determinant danej matice s determinantom redukovanej trojuholníkovej matice, ktorú dostaneme. Aby sme mohli túto metódu naozaj použiť na výpočet determinantu, potrebujeme ešte vedieť určiť determinant matice, ktorá je v redukovanom trojuholníkovom tvare. Na to nám poslúžia nasledujúce dva výsledky.

{det:VTRTM}

Veta 6.3.11. *Ak A je horná trojuholníková matica (pod hlavnou diagonálou má nuly), tak determinant matice A sa rovná súčinu prvkov na diagonále.*

$$|A| = a_{11}a_{22}\dots a_{nn}$$

Dôkaz. Stačí ukázať, že pre každú permutáciu $\varphi \in S_n$ okrem identickej permutácie je súčin $a_{1\varphi(1)}a_{2\varphi(2)}\dots a_{n\varphi(n)}$ nulový. Na to stačí, aby bol nulový niektorý činiteľ $a_{i\varphi(i)}$. Pretože predpokladáme, že A je horná trojuholníková matica, určite platí $a_{i\varphi(i)} = 0$ pre $i > \varphi(i)$. Zostáva nám teda ukázať, že aspoň jedno také i existuje.

Ak $\varphi \in S_n$ a $\varphi \neq id$, tak existuje $i \in \{1, 2, \dots, n\}$ také, že $i \neq \varphi(i)$. Nech i je najväčšie také číslo, čiže $i = \max\{k; \varphi(k) \neq k\}$. Označme $j = \varphi(i)$. Nemôže platiť $\varphi(j) = j$, lebo potom by sa na j zobrazili 2 rôzne prvky, čo je v spore s predpokladom, že φ je bijekcia. Teda platí $i > j = \varphi(i)$. □

{det:DOSRTM}

Dôsledok 6.3.12. *Determinant diagonálnej matice sa rovná súčinu diagonálnych prvkov.*

$$\begin{vmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{vmatrix} = d_1 d_2 \dots d_n$$

Príklad 6.3.13. Vypočítajme determinant z príkladu 6.3.4 tentokrát použitím riadkových a stĺpcových úprav.

$$\begin{vmatrix} 2&2&0&1\\1&0&-1&1\\2&3&1&1\\2&0&-1&2 \end{vmatrix} \overset{(1)}{=} \begin{vmatrix} 2&2&0&1\\0&0&-1&0\\3&3&1&2\\1&0&-1&1 \end{vmatrix} \overset{(2)}{=} - \begin{vmatrix} 2&2&0&1\\0&0&1&0\\3&3&0&2\\1&0&0&1 \end{vmatrix} \overset{(3)}{=} - \begin{vmatrix} 2&2&0&1\\0&0&1&0\\3&3&0&2\\1&0&0&1 \end{vmatrix} \overset{(4)}{=} - \begin{vmatrix} 2&2&0&1\\0&0&1&0\\1&1&0&1\\1&0&0&1 \end{vmatrix} \overset{(5)}{=} - \begin{vmatrix} 2&2&0&1\\0&0&1&0\\0&1&0&0\\1&0&0&1 \end{vmatrix} \overset{(6)}{=} \begin{vmatrix} 2&2&0&1\\0&1&0&0\\0&0&1&0\\1&0&0&1 \end{vmatrix} \overset{(7)}{=}$$
$$\begin{vmatrix} 2&0&0&1\\0&1&0&0\\0&0&1&0\\1&0&0&1 \end{vmatrix} \overset{(8)}{=} \begin{vmatrix} 1&0&0&0\\0&1&0&0\\0&0&1&0\\1&0&0&1 \end{vmatrix} \overset{(9)}{=} \begin{vmatrix} 1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1 \end{vmatrix} = 1$$

Elementárne riadkové a stĺpcové operácie, ktoré sme použili sú:
(1) 3. stĺpec sme pripočítali k prvému a štvrtému stĺpcu
(2) 2. riadok sme vynásobili -1
(3) odčítali sme druhý riadok od tretieho a pripočítali sme ho k štvrtému
(4) odpočítali sme 1. riadok od tretieho
(5) odpočítali sme 4. riadok od tretieho
(6) výmena druhého a tretieho riadku
(7) odpočítali sme 2-násobok 2. riadku od prvého

(8) odpočítali sme 4. riadok od prvého
(9) odčítali sme 1. riadok od štvrtého

Asi najviac si zjednodušíme prácu, ak budeme kombinovať oba postupy – riadkové a stĺpcové úpravy aj Laplaceov rozvoj. Napríklad namiesto kroku (2) alebo (3) v predchádzajúcom postupe sme mohli použiť Laplaceov rozvoj podľa 2. riadku a dostali by sme sa tak k determinantu 3×3.

{det:POZNOBJEM}

Poznámka 6.3.14. Všetky výsledky, ktoré sme odvodili pre zmeny determinantu pri elementárnych riadkových úpravách zodpovedajú geometrickej intuícii, ktorú sme spomínali – že determinant môžeme chápať (až na znamienko) ako objem.

Konkrétne vynásobenie niektorého riadku konštantou znamená c-násobné natiahnutie rovnobežnostena v smere niektorej z jeho hrán, pričom sa aj objem zväčší c-krát. Podobne pripočítanie násobku i-teho riadku k j-temu predstavuje vlastne skosenie rovnobežnostena v smere i-tej hrany. Pri ňom sa nemení objem. (Podobne – snáď ešte jednoduchšie – si môžete rozmyslieť, že to funguje pre dvojrozmerný rovnobežník.)

Veta 6.3.15. *Nech A je štvorcová matica typu $n \times n$. Matica A je regulárna práve vtedy, keď $|A| \neq 0$.*

Dôkaz. Pripomeňme, že matica A je regulárna práve vtedy, keď hodnosť matice je n. Vieme, že hodnosť matice sa rovná počtu nenulových riadkov v redukovanej trojuholníkovej matici M, ktorá je riadkovo ekvivalentná s A.

Ak matica A nie je regulárna, tak príslušná redukovaná trojuholníková matica má aspoň jeden nulový riadok. Podľa dôsledku 6.3.6 má teda nulový determinant. Ako sme dokázali v predchádzajúcich vetách, žiadna z elementárnych riadkových úprav nemení nulovosť a nenulovosť determinantu. Preto aj determinant matice A je nulový.

Ak A je regulárna, tak príslušná redukovaná trojuholníková matica má n nenulových riadkov. Pretože ide o maticu typu $n \times n$, priamo z definície redukovanej trojuholníkovej matice vyplýva, že to je jednotková matica I_n, ktorá má nenulový determinant $|I_n| = 1$. □

Teraz, keď už vieme, že sú pre nás podstatné len regulárne matice, mohli by sme vetu 6.3.11 a dôsledok 6.3.12 odvodiť v opačnom poradí. Priamy dôkaz dôsledku 6.3.12 by sa podobal na dôkaz vety 6.3.11, bol by však jednoduchší v tom, že teraz už máme nuly aj nad diagonálou a nepotrebujeme hľadať i také, že $i > \varphi(i)$. (Stačí nám, že $i \neq \varphi(i)$.) Akonáhle už máme dokázaný dôsledok 6.3.12 a chceme overiť vetu 6.3.11 pre regulárnu hornú trojuholníkovú maticu, stačí si uvedomiť, že takúto maticu vieme upraviť na diagonálnu už len používaním pripočítavania niektorého riadku k inému a pri tejto úprave sa hodnota determinantu nemení (veta 6.3.9).

6.4 Determinant súčinu matíc

Teraz dokážeme ešte jeden dôležitý výsledok týkajúci sa determinantov. Stručne povedané, tento výsledok hovorí, že determinant súčinu matíc sa rovná súčinu determinantov.

{det:VTDETSUC}

Veta 6.4.1. *Nech A, B sú dve matice typu $n \times n$ nad poľom F. Potom platí*

$$|AB| = |A| \cdot |B|.$$

Dôkaz. Označme riadky matice A ako $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$. Teda

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix} = \begin{pmatrix} \vec{\alpha}_1 \\ \vec{\alpha}_2 \\ \vdots \\ \vec{\alpha}_n \end{pmatrix}$$

Priamo z definície súčinu matíc (rozmyslite si to!) sa dá zistiť, že riadky matice AB majú tvar $\vec{\alpha}_k B$, čiže

$$AB = \begin{pmatrix} \vec{\alpha}_1 B \\ \vec{\alpha}_2 B \\ \vdots \\ \vec{\alpha}_n B \end{pmatrix}$$

Pretože $\vec{\alpha}_k = \sum_{i=1}^{n} a_{ki}\vec{e}_i$, môžeme túto rovnosť upraviť na tvar

$$AB = \begin{pmatrix} \sum_{i_1=1}^{n} a_{1i_1}\vec{e}_{i_1} B \\ \vdots \\ \sum_{i_n=1}^{n} a_{ni_n}\vec{e}_{i_n} B \end{pmatrix}$$

Pomocou viacnásobného použitia vety 6.3.8 postupne dostaneme

$$|AB| = \sum_{i_1=1}^{n} a_{1i_1} \begin{vmatrix} \vec{e}_{i_1} B \\ \sum_{i_2=1}^{n} a_{2i_2}\vec{e}_{i_2} B \\ \vdots \\ \sum_{i_n=1}^{n} a_{ni_n}\vec{e}_{i_n} B \end{vmatrix} = \ldots = \sum_{i_1=1}^{n} \ldots \sum_{i_n=1}^{n} a_{1i_1} a_{2i_2} \ldots a_{ni_n} \begin{vmatrix} \vec{e}_{i_1} B \\ \vec{e}_{i_2} B \\ \vdots \\ \vec{e}_{i_n} B \end{vmatrix}$$

Teraz si uvedomme, že ak $i_j = i_k$ pre nejaké $j \neq k$, tak príslušný determinant v predchádzajúcej sume je nulový (lebo má 2 rovnaké riadky). Teda nenulové sčítance budú iba tie, kde n-tica $(i_1, i_2, \ldots, i_n)$ predstavuje permutáciu čísel $1, \ldots, n$. Dostávame teda

$$|AB| = \sum_{\varphi \in S_n} a_{1\varphi(1)} a_{2\varphi(2)} \ldots a_{n\varphi(n)} \begin{vmatrix} \vec{e}_{\varphi(1)} B \\ \vec{e}_{\varphi(2)} B \\ \vdots \\ \vec{e}_{\varphi(n)} B \end{vmatrix}$$

Matica vystupujúca v predchádzajúcej rovnosti je vlastne matica B s poprehadzovanými riadkami. Pomocou výmen niektorých riadkov z nej vieme dostať maticu B. O chvíľu si ukážeme, že sa to dá urobiť pomocou $i(\varphi)$ výmen. Na základe toho prejde predchádzajúca rovnosť na tvar

$$|AB| = \sum_{\varphi \in S_n} a_{1\varphi(1)} a_{2\varphi(2)} \ldots a_{n\varphi(n)} (-1)^{i(\varphi)} |B| = |A| \cdot |B|.$$

Zostáva nám teda overiť, že potrebný počet výmen je skutočne $i(\varphi)$. Pri upravovaní „poprehadzovanej matice na maticu“ B môžeme postupovať tak, že najprv premiestníme prvý riadok na prvé miesto a to tak, že ho vždy vymieňame s predchádzajúcim, až kým sa nedostane na správnu pozíciu. (Ak už je na prvom mieste, nerobíme žiadne výmeny.) Takto sme urobili toľko výmen riadkov, koľko má permutácia φ inverzií obsahujúcich číslo 1. Teraz môžeme podobným spôsobom presunúť druhý riadok na druhé miesto. Počet výmen je rovnaký ako počet tých inverzií, ktoré obsahujú 2 ale nie 1 (pretože prvý riadok sme už presunuli). Takto postupujeme ďalej. Vidíme, že pri takomto algoritme dostaneme maticu B pomocou presne $i(\varphi)$ výmen riadkov. □

Dôkaz, ktorý sme uviedli, je v princípe rovnaký ako v [Ax, Theorem 10.31]. V [K, Veta 6.2.18] môžete nájsť dôkaz, ktorý využíva rozloženie matice na súčin redukovanej trojuholníkovej matice a matíc elementárnych riadkových operácií. Ešte jeden, úplne iný dôkaz, môžete nájsť v [KGGS, Veta 2.14.7].

Hovorili sme o tom, že ako geometrický význam determinantu si môžeme predstaviť objem rovnobežnostena určeného riadkami matice. V prípade, že danú maticu chápeme ako maticu lineárneho zobrazenia, je to presne objem rovnobežnostena na ktorý sa zobrazí jednotková kocka (určená vektormi štandardnej bázy). Determinant nám teda hovorí, koľkokrát sa pri zobrazení daným lineárnym zobrazením zväčší objem jednotkovej kocky. Pretože ide o lineárne zobrazenie, aj objem ľubovoľného rovnobežnostena sa zväčší v rovnakom pomere. A presne toto tvrdenie vlastne hovorí veta, ktorú sme práve dokázali. (V poznámke 6.3.14 sme hovorili o súvise medzi touto geometrickou predstavou a elementárnymi riadkovými operáciami. Spomínaný dôkaz z [K] teda vlastne zodpovedá tejto geometrickej intuícii – ľubovoľné lineárne zobrazenie sme najprv rozložili na jednoduchšie zobrazenia, o ktorých vieme ukázať koľkokrát zväčšujú objem. Pomocou toho vieme určiť, koľkokrát sa zväčší objem pri použití pôvodného zobrazenia.)

6.5 Využitie determinantov

6.5.1 Výpočet inverznej matice

{det:VTINVERZNA}

Veta 6.5.1. *Ak A je regulárna matica typu $n \times n$, tak*

$$A^{-1} = \frac{1}{|A|} \begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \dots & \dots & \dots & \dots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{pmatrix}$$

kde A_{ij} označuje algebraický doplnok prvku a_{ij}.

Poznámka 6.5.2. Maticu

$$\begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \dots & \dots & \dots & \dots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{pmatrix}$$

nazývame *adjungovaná matica* k matici A a označuje adj A. Teda vyjadrenie inverznej matice z predchádzajúcej vety môžeme zapísať aj v tvare

$$A^{-1} = \frac{\operatorname{adj} A}{|A|}.$$

POZOR na výmenu poradia indexovania – algebraické doplnky v adjungovanej matici nie sú indexované tak, ako v pôvodnej matici ale podobným spôsobom ako v transponovanej matici.

Dôkaz. Treba dokázať, že $A.\frac{\operatorname{adj}(A)}{|A|} = I$. Označme maticu $A.\frac{\operatorname{adj}(A)}{|A|}$ ako C. Pre jej prvky platí

$$c_{ij} = \frac{1}{|A|} \sum_{k=1}^{n} a_{ik} A_{jk}.$$

Potrebujeme vlastne ukázať, že $c_{ii} = 1$ a $c_{ij} = 0$ pre $i \neq j$.

Pre $i = j$ dostávame

$$c_{ii} = \frac{1}{|A|} \sum_{k=1}^{n} a_{ik} A_{ik}.$$

Suma v predchádzajúcej rovnosti je presne rozvoj matice A podľa i-teho riadku, čiže sme dostali $c_{ii} = \frac{|A|}{|A|} = 1$.

Pre $i \neq j$ si všimneme, že suma vo výraze

$$c_{ij} = \frac{1}{|A|} \sum_{k=1}^{n} a_{ik} A_{jk}.$$

predstavuje Laplaceov rozvoj matice, ktorá vznikne z A nahradením j-teho riadku i-tym, podľa (nového) j-teho riadku. Pretože táto matica má dva riadky rovnaké, jej determinant je nulový, z čoho $c_{ij} = 0$. □

V predchádzajúcom dôkaze sme overili pre maticu $B = \frac{\operatorname{adj}(A)}{|A|}$ rovnosť $AB = I$. V definícii inverznej matice však vystupuje aj rovnosť $BA = I$. Nie je to chyba? Zabudli sme ju overiť?

Nie, nie je to chyba. V predpokladoch vety totiž máme, že A je regulárna, teda A^{-1} existuje. Vďaka tomu z rovnosti $AB = I$ po vynásobení A^{-1} dostaneme $B = A^{-1}$. (A navyše z poznámky 5.5.8 vieme, že nám stačí overiť iba jednu z týchto dvoch rovností.)

{det:PRINV}

Príklad 6.5.3. Nech $A = \begin{pmatrix} 1 & -1 & 1 \\ 2 & 1 & 1 \\ 2 & -1 & 2 \end{pmatrix}$. Priamym výpočtom dostaneme $|A| = 1$ a $\operatorname{adj} A = \begin{pmatrix} 3 & 1 & -2 \\ -2 & 0 & 1 \\ -4 & -1 & 3 \end{pmatrix}$.

Z toho dostávame, že

$$A^{-1} = \frac{1}{|A|} \operatorname{adj} A = \begin{pmatrix} 3 & 1 & -2 \\ -2 & 0 & 1 \\ -4 & -1 & 3 \end{pmatrix}.$$

Vynásobením matíc sa môžeme presvedčiť, že skutočne platí $A.A^{-1} = A^{-1}.A = I$.

{det:POZN2x2}

Poznámka 6.5.4. Práve odvodený vzorec pre inverznú maticu bude veľmi jednoduchý v prípade, že pracujeme s maticou veľkosti 2×2. Vtedy totiž všetky podmatice vystupujúce v algebraických doplnkoch sú typu 1×1, čiže jediná hodnota v takejto matici je priamo rovná determinantu.

Konkrétne pre maticu

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

dostaneme $A_{11} = a_{22}$, $A_{12} = -a_{21}$, $A_{21} = -a_{12}$ a $A_{22} = a_{11}$. Ak to dosadíme do vzorca pre inverznú maticu, máme

$$A^{-1} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix} = \frac{1}{|A|} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}.$$

Ľahko môžete priamy výpočtom skontrolovať, že súčin týchto dvoch matíc skutočne dáva jednotkovú maticu. A je to vzorec, ktorý sa vcelku ľahko zapamätá. (Vlastne sme len prvky na diagonále vymenili a pri prvkoch mimo diagonály zmenili znamienko. Takto upravenú maticu sme vydelili determinantom.)

6.5.2 Cramerovo pravidlo

Na začiatku kapitoly sme si ukázali, ako sa dá pomocou determinantu vyjadriť riešenie sústavy 2 lineárne nezávislých rovníc o 2 neznámych. Teraz odvodíme analogický výsledok pre sústavy n rovníc s n neznámymi.

Majme sústavu

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & c_1 \\ a_{21} & a_{22} & \dots & a_{2n} & c_2 \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} & c_n \end{array}\right)$$

Označme maticu sústavy ako A a $C := (c_1, c_2, \dots, c_n)^T$ maticu, v ktorej sú do stĺpca zapísané pravé strany. Hľadáme vlastne takú maticu X, pre ktorú platí

$$AX = C.$$

Ak je matica A regulárna, vynásobením A^{-1} zľava dostaneme

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = A^{-1}C = \frac{1}{|A|} \begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \dots & \dots & \dots & \dots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$$

Z predchádzajúcej rovnosti dostaneme

$$x_i = \frac{1}{|A|} \sum_{j=1}^{n} A_{ji} c_j.$$

Nech A_i označuje maticu, ktorú dostaneme ak v matici A nahradíme i-ty stĺpec stĺpcom $(c_1, c_2, \dots, c_n)^T$ (čiže pravými stranami). Potom si môžeme všimnúť, že výraz vystupujúci v predchádzajúcej rovnosti je presne Laplaceov rozvoj matice A_i podľa i-teho stĺpca. (Iné stĺpce sme nemenili, preto algebraické doplnky vystupujúce v Laplaceovom rozvoji sú rovnaké ako pre maticu A.) Platí teda

$$|A_i| = \sum_{j=1}^{n} A_{ji} c_j,$$

z čoho dostaneme

$$x_i = \frac{|A_i|}{|A|}.$$

Tým sme odvodili vzorec pre riešenia sústavy lineárnych rovníc, ktorá má regulárnu maticu. Tento vzorec sa nazýva *Cramerovo pravidlo.*

Príklad 6.5.5. Majme sústavu $\left(\begin{array}{ccc|c} 1 & -1 & 1 & 1 \\ 2 & 1 & 1 & 2 \\ 2 & -1 & 2 & 3 \end{array}\right)$. V príklade 6.5.3 sme vypočítali determinant $|A| = 1$.

Ostatné determinanty, ktoré potrebujeme na použitie Cramerovho pravidla sú

$\begin{vmatrix} 1 & -1 & 1 \\ 2 & 1 & 1 \\ 3 & -1 & 2 \end{vmatrix} = \begin{vmatrix} 1 & -1 & 1 \\ 2 & 1 & 1 \\ 0 & -1 & 0 \end{vmatrix} = -2 + 1 = -1$

(Úprava, ktorú sme použili, bolo odčítanie prvých dvoch riadkov od tretieho.)

$\begin{vmatrix} 1 & 1 & 1 \\ 2 & 2 & 1 \\ 2 & 3 & 2 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 1 \\ 2 & 0 & 1 \\ 2 & 1 & 2 \end{vmatrix} = 2 - 1 = 1$

$\begin{vmatrix} 1 & -1 & 1 \\ 2 & 1 & 2 \\ 2 & -1 & 3 \end{vmatrix} = \begin{vmatrix} 1 & -1 & 0 \\ 2 & 1 & 0 \\ 2 & -1 & 1 \end{vmatrix} = 1 + 2 = 3$

Z toho dostaneme riešenie sústavy $(-1, 1, 3)$.

{det:PRIKLPOLI

Príklad 6.5.6. V príklade 4.4.21 sme dokázali, že $F = \{a+b\sqrt[3]{2}+c\sqrt[3]{2^2}; a, b, c \in \mathbb{Q}\}$ je pole. Poďme sa pozrieť na to, či by sme vedeli nájsť vzorec, ktorý určuje inverzný prvok k danému prvku z F. Chceme teda pre dané $a, b, c \in \mathbb{Q}$ nájsť také racionálne čísla x, y, z, aby platilo

$$(a + b\sqrt[3]{2} + c\sqrt[3]{2^2})(x + y\sqrt[3]{2} + z\sqrt[3]{2^2}) = 1.$$

Najprv, ako prípravu na samotný výpočet, si všimnime, že $a + b\sqrt[3]{2} + c\sqrt[3]{2^2} = 0$ platí práve vtedy, keď $a = b = c = 0$. Môžeme to ukázať sporom.

Predpokladajme, že by platilo $a + b\sqrt[3]{2} + c\sqrt[3]{2^2} = 0$ a aspoň jedno z čísel a, b, c by bolo nenulové. Ak $c = 0$ tak vieme z tejto rovnosti ukázať, že buď $a = b = 0$ (čiže všetky tri čísla by boli nulové), alebo $\sqrt[3]{2} = -a/b$. Čiže by sme dostali, že $\sqrt[3]{2}$ je racionálne číslo, čo je spor.

Ak $c \neq 0$, tak vieme dostať rovnosť tvaru

$$\sqrt[3]{2^2} = e + f\sqrt[3]{2},$$

kde e aj f sú racionálne čísla. Vynásobením tejto rovnice číslom $\sqrt[3]{2}$ dostaneme

$$8 = e\sqrt[3]{2} + f\sqrt[3]{2^2} = e\sqrt[3]{2} + f(e + f\sqrt[3]{2}) = (e + f^2)\sqrt[3]{2} + ef$$
$$8 - ef = (e + f^2)\sqrt[3]{2}$$

Ak je ľavá strana nenulová, tak opäť vieme ukázať, že $\sqrt[3]{2}$ je racionálne číslo a dostaneme spor. Ak je nulová, tak máme $8 = ef$ a $e + f^2 = 0$. Z toho dostaneme $f^3 = -8$, jediné racionálne riešenie je $f = -2$. Súčasne máme $e = -f^2 = -4$. Teda v rovnosti $\sqrt[3]{2^2} = e + f\sqrt[3]{2}$ máme na ľavej strane kladné číslo a na pravej strane záporné číslo, čiže sme opäť dostali spor.

Z toho, čo sme si práve ukázali, vyplýva aj implikácia $a+b\sqrt[3]{2}+c\sqrt[3]{2^2} = a'+b'\sqrt[3]{2}+c'\sqrt[3]{2^2}$ $\Rightarrow a = a'$, $b = b'$, $c = c'$.

Teraz už máme pripravené všetko na to, aby sme sa pustili do hľadanie inverzného prvku.

Chceli by sme, aby platilo $(a + b\sqrt[3]{2} + c\sqrt[3]{2^2})(x + y\sqrt[3]{2} + z\sqrt[3]{2^2}) = 1$. Ľavú stranu tejto rovnosti môžeme upraviť ako

$$(a+b\sqrt[3]{2}+c\sqrt[3]{2^2})(x+y\sqrt[3]{2}+z\sqrt[3]{2^2}) = (ax+2cy+2bz)+(bx+ay+2cz)\sqrt[3]{2}+(cx+by+az)\sqrt[3]{2^2}.$$

Dostávame teda, že musia platiť tieto rovnosti:

$$\begin{aligned} ax + 2cy + 2bz &= 0 \\ bx + ay + 2cz &= 0 \\ cx + by + az &= 0 \end{aligned}$$

Toto je vlastne sústava lineárnych rovníc s neznámymi x, y, z, v prípade, že determinant matice sústavy je nenulový, tak bude mať jediné riešenie, ktoré budeme vedieť vyjadriť na základe Cramerovho pravidla. Vyrátajme teda najprv determinant matice sústavy:

$$\begin{vmatrix} a & 2c & 2b \\ b & a & 2c \\ c & b & a \end{vmatrix} = a^3 + 2b^3 + 4c^3 - 6abc.$$

Ak chceme použiť Cramerovo pravidlo mali by sme sa najprv presvedčiť o tom, že tento determinant je nenulový. Riešime teda vlastne otázku, či pre $a, b, c \in \mathbb{Q}$ môže nastať rovnosť $a^3 + 2b^3 + 4c^3 - 6abc = 0$. Chceme ukázať, že to môže nastať iba pre $a = b = c = 0$.

Zdôvodnenie tohoto faktu nie je úplne jednoduché. Môžeme to urobiť napríklad takto: Označme $u = a$, $v = b\sqrt[3]{2}$, $w = c\sqrt[3]{2^2}$. Po tejto substitúcii prejde naša rovnosť do tvaru

$$u^3 + v^3 + w^3 - 3uvw = 0.$$

Ďalej použijeme rozklad

$$u^3+v^3+w^3-3uvw = (u+v+w)(u^2+v^2+w^2-uv-uw-vw) = \frac{1}{2}(u+v+w)((u-v)^2+(v-w)^2+(w-u)^2).$$

(Táto rovnosť sa ľahko overí priamym výpočtom, objaviť ju ale nie je úplne jednoduché.)

Ak sme overili predošlú rovnosť, tak už vieme, že náš výraz sa bude rovnať nule iba ak $u + v + w = 0$ alebo $(u - v)^2 + (v - w)^2 + (w - u)^2 = 0$. V prvom prípade dostávame $u + v + w = a + b\sqrt[3]{2} + c\sqrt[3]{2^2} = 0$. Už sme overili, že to nastane iba ak $a = b = c = 0$. V druhom prípade by sa mal rovnať nule súčet troch nezáporných reálnych čísel, čo znamená, že $u - v = v - w = w - u = 0$, a teda $u = v = w$. Potom dostávame $a = \sqrt[3]{2}b = c\sqrt[3]{2^2}$ a ak sú a, b, c nenulové tak dostaneme, že $\sqrt[3]{2}$ je racionálne, čo je spor.

Podarilo sa nám teda ukázať, že uvedený determinant je pre racionálne čísla a, b, c nenulový. Ak teraz použijeme Cramerovo pravidlo, dostaneme

$$x = \frac{\begin{vmatrix} 1 & 2c & 2b \\ 0 & a & 2c \\ 0 & b & a \end{vmatrix}}{\begin{vmatrix} a & 2c & 2b \\ b & a & 2c \\ c & b & a \end{vmatrix}} = \frac{a^2 - 2bc}{a^3 + 2b^3 + 4c^3 - 6abc}$$

$$y = \frac{\begin{vmatrix} a & 1 & 2b \\ b & 0 & 2c \\ c & 0 & a \end{vmatrix}}{\begin{vmatrix} a & 2c & 2b \\ b & a & 2c \\ c & b & a \end{vmatrix}} = \frac{2c^2 - ab}{a^3 + 2b^3 + 4c^3 - 6abc}$$

$$z = \frac{\begin{vmatrix} a & 2c & 1 \\ b & a & 0 \\ c & b & 0 \end{vmatrix}}{\begin{vmatrix} a & 2c & 2b \\ b & a & 2c \\ c & b & a \end{vmatrix}} = \frac{b^2 - ac}{a^3 + 2b^3 + 4c^3 - 6abc}$$

Priamym výpočtom sa dá skontrolovať, že $x+y\sqrt[3]{2}+z\sqrt[3]{2^2}$ je skutočne inverzný prvok k prvku $a + b\sqrt[3]{2} + c\sqrt[3]{2^2}$.

Príklad 6.5.7. Ešte sa pozrime na to, že rovnosť

$$u^3 + v^3 + w^3 - 3uvw = (u + v + w)(u^2 + v^2 + w^2 - uv - uw - vw),$$

ktorú sme použili v predošlom príklade, sa dá odvodiť pomocou determinantov. Platí:

$$u^3 + v^3 + w^3 - 3uvw = \begin{vmatrix} u & v & w \\ w & u & v \\ v & w & u \end{vmatrix} = \begin{vmatrix} u & v & w \\ w & u & v \\ u+v+w & u+v+w & u+v+w \end{vmatrix} =$$

$$= (u + v + w)\begin{vmatrix} u & v & w \\ w & u & v \\ 1 & 1 & 1 \end{vmatrix} = (u + v + w)(u^2 + v^2 + w^2 - uv - uw - vw)$$

Viaceré ďalšie aplikácie tejto faktorizácie sú spomenuté napríklad v článku [Ma].

Cvičenia

Úloha 6.5.1. Vypočítajte determinanty: $\begin{vmatrix} -2 & 3 & -3 & -1 \\ 1 & -2 & 3 & 2 \\ 0 & 1 & 1 & 1 \\ 1 & -1 & -1 & -2 \end{vmatrix}$ $\begin{vmatrix} -2 & 3 & -2 & -1 \\ 1 & -2 & 1 & 2 \\ 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & -1 \end{vmatrix}$ $\begin{vmatrix} -2 & 3 & -1 & -1 \\ 1 & -2 & 1 & 2 \\ 1 & 1 & 2 & -1 \\ 0 & 1 & -1 & -1 \end{vmatrix}$
Ak existuje inverzná matica, aký bude jej determinant. Výsledky (bez záruky): $0, -8, 8$.

Úloha 6.5.2. Vyriešte v $\mathbb{Z}_5$ pomocou Cramerovho pravidla: $\left(\begin{array}{ccc|c} 3 & 4 & 0 & 1 \\ 1 & 1 & 2 & 1 \\ 3 & 4 & 1 & 0 \end{array}\right)$ $\left(\begin{array}{ccc|c} 1 & 1 & 4 & 1 \\ 0 & 1 & 2 & 2 \\ 1 & 0 & 3 & 3 \end{array}\right)$ $\left(\begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 2 & 1 & 3 & 2 \end{array}\right)$

Úloha 6.5.3. Pomocou Cramerovho pravidla riešte:

$$\begin{array}{rrrrl} x_1 & +5x_2 & +4x_3 & +3x_4 & = 1 \\ 2x_1 & -x_2 & +2x_3 & -x_4 & = 0 \end{array} \qquad \begin{array}{rrrl} x_1 & +2x_2 & +x_3 & = 1 \\ 2x_1 & +x_2 & -x_3 & = 0 \end{array}$$

(Návod: Skúste zvoliť x_3, x_4 za parametre.)

Úloha 6.5.4. Určte determinanty daných matíc. Viete na základe výsledku určiť ich hodnosť pre niektoré hodnoty $c \in \mathbb{R}$?
$\begin{vmatrix} 1 & 2 & c-1 \\ c-2 & 1 & 0 \\ c & 1 & 0 \end{vmatrix}$ $\begin{vmatrix} 1 & 1 & c-1 \\ c-2 & 1 & 0 \\ 0 & 1 & c \end{vmatrix}$ $\begin{vmatrix} 2 & c+1 & 0 \\ 2 & c-1 & 2c \\ 1 & 1 & 1 \end{vmatrix}$

Úloha 6.5.5. Nájdite inverznú maticu k maticiam z úlohy 5.5.2 pomocou determinantu.

Úloha 6.5.6. Vypočítajte inverznú maticu:
$\begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{pmatrix}$ $\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & -1 \\ 1 & 4 & 9 & 1 \\ 1 & 8 & 27 & -1 \end{pmatrix}$

Úloha 6.5.7. Ukážte, že v ľubovoľnom poli platí $x + y + z = 0 \Rightarrow x^3 + y^3 + z^3 - 3xyz = 0$.
Skúste sa zamyslieť nad tým, či to viete odvodiť použitím vhodného determinantu.

{detcvic:VANDERMOND}

Úloha 6.5.8*. $\begin{vmatrix} 1 & a_1 & a_1^2 & \dots & a_1^n \\ 1 & a_2 & a_2^2 & \dots & a_2^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & a_n & a_n^2 & \dots & a_n^n \\ 1 & a_{n+1} & a_{n+1}^2 & \dots & a_{n+1}^n \end{vmatrix} = ?$
[Výsledok by mal byť $\prod_{1 \le i < j \le n+1} (a_j - a_i)$, t.j. súčin výrazov tvaru $a_j - a_i$ pre všetky $i < j$.]

{detcvic:DET1}

Úloha 6.5.9*. $D_n = \begin{vmatrix} 2 & 1 & 0 & 0 & \dots & 0 \\ 1 & 2 & 1 & 0 & \dots & 0 \\ 0 & 1 & 2 & 1 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & 2 & 1 \\ 0 & \dots & 0 & 0 & 1 & 2 \end{vmatrix} = ?$

{detcvic:DET2}

Úloha 6.5.10*. $D_n = \begin{vmatrix} a+b & ab & 0 & 0 & \dots & 0 \\ 1 & a+b & ab & 0 & \dots & 0 \\ 0 & 1 & a+b & ab & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & a+b & ab \\ 0 & \dots & 0 & 0 & 1 & a+b \end{vmatrix} = ?$

{detcvic:DET3}

Úloha 6.5.11*. $D_n = \begin{vmatrix} n & 1 & 1 & \dots & 1 \\ 1 & n & 1 & \dots & 1 \\ \dots & \dots & \dots & \dots & \\ 1 & \dots & 1 & 1 & n \end{vmatrix} = ?$

Úloha 6.5.12*. Vypočítajte determinant matice typu $n \times n$

$$D_n = \begin{vmatrix} x & a & a & \dots & a \\ a & x & a & \dots & a \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ a & \dots & a & x & a \\ a & \dots & a & a & x \end{vmatrix} = ?$$

(Teda ide o maticu, kde diagonálne prvky sú rovné x a všetky prvky mimo diagonály sú rovné a.)

detcvic:DET4}

Úloha 6.5.13*. $D_n = \begin{vmatrix} 1 & 1 & \dots & \dots & 1 \\ 1 & 2 & 1 & \dots & 1 \\ 1 & 1 & 3 & \dots & 1 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 1 & 1 & \dots & 1 & n \end{vmatrix} = ?$

Úloha 6.5.14*. Nájdite determinant matice A_n typu $n \times n$, ktorej prvky sú určené predpisom $a_{ij} = \min\{i, j\}$, t.j. napríklad

$$A_2 = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \qquad A_3 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{pmatrix} \qquad A_4 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 \\ 1 & 2 & 3 & 3 \\ 1 & 2 & 3 & 4 \end{pmatrix}$$

Úloha 6.5.15. Nech A je matica typu 7×7, ktorej prvky sú nepárne celé čísla. Ukážte, že $|A|$ je celé číslo a ďalej, že $|A|$ je celočíselný násobok čísla 64.

Úloha 6.5.16. Ak A, B sú štvorcové matice a matica AB je regulárna, tak obe matice A, B sú regulárne. (Môžete sa zamyslieť nad tým, či to viete zdôvodniť s pomocou determinantov a aj nad riešením bez nich. Takisto sa môžete skúsiť zamyslieť nad tým, čo sa stane ak matice nie sú štvorcové.)

{detcvic:ULOHBLOK}

Úloha 6.5.17*. Ukážte, že pre determinant blokovej matice platí

$$\det \begin{pmatrix} A & B \\ 0 & D \end{pmatrix} = \det \begin{pmatrix} A & 0 \\ C & D \end{pmatrix} = \det(A) \cdot \det(D)$$

(Hint: Možno sa oplatí začať tým, že si rozmyslíte, ako to je s determinantami nejakých jednoduchších matíc – napríklad $\det \begin{pmatrix} A & B \\ 0 & I \end{pmatrix}$ alebo $\det \begin{pmatrix} A & 0 \\ 0 & D \end{pmatrix}$.) [1]

Úloha 6.5.18. Ak viete, že 195, 403 a 247 sú násobky čísla 13, viete ukázať (bez toho, aby ste ho museli vyrátať), že aj $\begin{vmatrix} 1 & 9 & 5 \\ 4 & 0 & 3 \\ 2 & 4 & 7 \end{vmatrix}$ je celočíselný násobok 13?

[1] Azda sa oplatí poznamenať, že vo všeobecnosti sa determinant $\begin{vmatrix} A & C \\ D & B \end{vmatrix}$ nemusí rovnať $|A||B| - |C||D|$ ani $|AB - CD|$. Nejaké podmienky, kedy takéto niečo funguje, sa dajú nájsť v článku [Si].

Kapitola 7

Euklidovské vektorové priestory

Táto kapitola je spracovaná prevažne na základe [KGGS, 1.16,1.17].

7.1 Skalárny súčin

Skalárny súčin vektorov patriacich do $\mathbb{R}^2$ alebo $\mathbb{R}^3$ poznáte zo strednej školy. Tam ste skalárny súčin vektorov $\vec{\alpha} = (a_1, a_2)$ a $\vec{\beta} = (b_1, b_2)$ definovali ako

$$\langle \vec{\alpha}, \vec{\beta} \rangle = a_1 b_1 + a_2 b_2.$$

(Používali ste iné označenie pre skalárny súčin, ako budeme používať my.) Takisto ste sa na strednej škole naučili, ako súvisí skalárny súčin s veľkosťou vektora a uhlom, ktorý zvierajú 2 vektory:

$$\langle \vec{\alpha}, \vec{\beta} \rangle = |\vec{\alpha}||\vec{\beta}| \cos \varphi,$$
$$|\vec{\alpha}| = \sqrt{a_1^2 + a_2^2} = \sqrt{\langle \vec{\alpha}, \vec{\alpha} \rangle}.$$

My by sme teraz chceli zaviesť definíciu skalárneho súčinu o niečo všeobecnejšie – chceli by sme popísať, aké vlastnosti by mal mať skalárny súčin, aby sme pomocou neho mohli zmysluplne hovoriť o veľkosti alebo uhle vektorov z daného vektorového priestoru. Budeme opäť postupovať axiomaticky – zavedieme si niekoľko základných vlastností skalárneho súčinu, z ktorých sa budú dať odvodiť ostatné.

Definícia 7.1.1. Nech $(V, +, \cdot)$ je vektorový priestor nad poľom $\mathbb{R}$. Potom zobrazenie $g\colon V \times V \to \mathbb{R}$ sa nazýva *skalárny súčin* na V, ak pre ľubovoľné $\vec{\alpha}, \vec{\beta} \in V$ a $c \in \mathbb{R}$ platí

{SK1} *(i)* $g(\vec{\alpha}, \vec{\beta}) = g(\vec{\beta}, \vec{\alpha})$,

{SK2} *(ii)* $g(\vec{\alpha} + \vec{\beta}, \vec{\gamma}) = g(\vec{\alpha}, \vec{\gamma}) + g(\vec{\beta}, \vec{\gamma})$,

{SK3} *(iii)* $g(c\vec{\alpha}, \vec{\beta}) = cg(\vec{\alpha}, \vec{\beta})$,

{SK4} *(iv)* ak $\vec{\alpha} \neq \vec{0}$, tak $g(\vec{\alpha}, \vec{\alpha}) > 0$.

Vektorový priestor V spolu so skalárnym súčinom g nazývame *euklidovským vektorovým priestorom.*

Predchádzajúcu definíciu môžeme stručne preformulovať tak, že zobrazenie g je symetrické (i), kladne definitné (iv), a bilineárne (ii) a (iii). S pojmom kladnej definitnosti sa ešte stretneme, budeme sa ním zaoberať podrobnejšie v časti 8.3. Tento pojem by ste mohli poznať aj z matematickej analýzy, kde ste sa s ním mohli stretnúť v súvislosti s hľadaním extrémov viac premenných. Pod bilinearitou rozumieme to, že zobrazenie je lineárne v oboch premenných – ak zvolím pevne vektor $\vec{\alpha}$ a mením $\vec{\beta}$, môžeme ho chápať ako zobrazenie, ktoré vektoru $\vec{\beta}$ priradí reálne číslo. Z (ii) a (iii) vidíme, že toto zobrazenie je lineárne. Rovnako je to aj v prípade, že fixujeme $\vec{\beta}$.

Všimnite si, že skalárny súčin sme definovali iba pre vektorové priestory nad poľom $\mathbb{R}$.

Namiesto $g(\vec{\alpha},\vec{\beta})$ budeme používať označenie $\langle\vec{\alpha},\vec{\beta}\rangle$. Pri tomto označení uvedené vlastnosti môžeme prepísať nasledovne:

{SC1}

(i) $\langle\vec{\alpha},\vec{\beta}\rangle = \langle\vec{\beta},\vec{\alpha}\rangle$,

{SC2}

(ii) $\langle\vec{\alpha}+\vec{\beta},\vec{\gamma}\rangle = \langle\vec{\alpha},\vec{\gamma}\rangle + \langle\vec{\beta},\vec{\gamma}\rangle$,

{SC3}

(iii) $\langle c\vec{\alpha},\vec{\beta}\rangle = c\langle\vec{\alpha},\vec{\beta}\rangle$,

{SC4}

(iv) ak $\vec{\alpha}\neq\vec{0}$, tak $\langle\vec{\alpha},\vec{\alpha}\rangle > 0$.

Podmienku (iv) možno ekvivalentne vyjadriť aj tak, že pre každý vektor $\vec{\alpha}$ platí $\langle\vec{\alpha},\vec{\alpha}\rangle) \geq 0$ a rovnosť nastáva jedine pre $\vec{\alpha}=\vec{0}$. (Skúste si rozmyslieť, ako z prvých troch podmienok v definícii vyplýva, že $\langle\vec{0},\vec{\alpha}\rangle = 0$ pre ľubovoľné $\vec{\alpha}$.)

Ak V je euklidovský vektorový priestor, tak aj každý jeho podpriestor je euklidovský priestor (s rovnako definovaným skalárnym súčinom).

Poznámka 7.1.2. Niekedy sa skalárny súčin definuje aj pre vektorové priestory nad poľom $\mathbb{C}$. V tomto prípade sa podmienka (i) zmení na

$$\langle\vec{\alpha},\vec{\beta}\rangle = \overline{\langle\vec{\beta},\vec{\alpha}\rangle}.$$

Všimnime si, že táto podmienka implikuje, že $\langle\vec{\alpha},\vec{\alpha}\rangle = \overline{\langle\vec{\alpha},\vec{\alpha}\rangle}$, a teda $\langle\vec{\alpha},\vec{\alpha}\rangle\in\mathbb{R}$. Vďaka tomu má zmysel aj podmienka (iv).

My sa však budeme zaoberať iba reálnymi euklidovskými priestormi.

{skal:PRIKLRN}

Príklad 7.1.3. Zoberme si vektorový priestor $\mathbb{R}^n$ s obvyklým sčitovaním a skalárnym násobkom (po zložkách). Potom pre vektory $\vec{\alpha}=(a_1,\ldots,a_n)$ a $\vec{\beta}=(b_1,\ldots,b_n)$ definujeme

$$\langle\vec{\alpha},\vec{\beta}\rangle = \sum_{k=1}^{n} a_k b_k.$$

V prípade $\mathbb{R}^2$ alebo $\mathbb{R}^3$ dostávame skalárny súčin ako ho poznáte zo strednej školy.

Vlastnosti z definície skalárneho súčinu sa overia pomerne jednoducho. Vlastnosť (i) je zrejmá. Vlastnosti (ii) a (iii) sa overia jednoduchou úpravou:

$$\sum_{k=1}^{n}(a_k+b_k)c_k = \sum_{k=1}^{n}(a_kc_k+b_kc_k) = \sum_{k=1}^{n}a_kc_k + \sum_{k=1}^{n}b_kc_k,$$
$$\sum_{k=1}^{n}ca_k = c\sum_{k=1}^{n}a_k.$$

Aby sme overili vlastnosť (iii), stačí si všimnúť, že

$$\langle \vec{\alpha}, \vec{\alpha} \rangle = \sum_{k=1}^{n} a_k^2.$$

Pretože $a_k^2 \geq 0$, aj skalárny súčin $\langle \vec{\alpha}, \vec{\alpha} \rangle \geq 0$ a rovný 0 bude iba v prípade, že všetky sčítance sú nulové, t.j. $a_k = 0$ pre každé k a $\vec{\alpha} = 0$.

Príklad 7.1.4. Definujme na $\mathbb{R}^2$ skalárny súčin nasledovne:

$$\langle \vec{\alpha}, \vec{\beta} \rangle = a_1 b_1 + a_1 b_2 + a_2 b_1 + 2 a_2 b_2,$$

pre $\vec{\alpha} = (a_1, a_2)$ a $\vec{\beta} = (b_1, b_2)$. Vlastnosti (i)–(iii) sa overia ľahko. Vlastnosť (iv) vyplýva z toho, že

$$\langle \vec{\alpha}, \vec{\alpha} \rangle = a_1^2 + 2 a_1 a_2 + 2 a_2^2 = (a_1 + a_2)^2 + a_2^2,$$

čiže $\langle \vec{\alpha}, \vec{\alpha} \rangle = 0$ platí práve vtedy, keď $a_1 = 0$ a súčasne $a_1 + a_2 = 0$, teda $a_1 = a_2 = 0$, čiže $\vec{\alpha} = \vec{0}$.

{skal:PRIKLMATICA}

Príklad 7.1.5. Postup z predchádzajúceho príkladu sa dá zovšeobecniť. Všimnime si, že je to špeciálny prípad nasledujúceho zápisu:

$$g(\vec{\alpha}, \vec{\beta}) = (a_1 \dots a_n) C \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \vec{\alpha} C \vec{\beta}^T,$$

kde C je reálna matica typu $n \times n$. Takýto predpis priradí dvom vektorom $\vec{\alpha} = (a_1, \dots, a_n)$, $\vec{\beta} = (b_1, \dots, b_n) \in \mathbb{R}^n$ nejaké reálne číslo. Nie vždy to však bude skalárny súčin.

Všimnime si, že tento predpis môžeme zapísať aj takto:

$$g(\vec{\alpha}, \vec{\beta}) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_i c_{ij} b_j.$$

V prípade, že ide o symetrickú maticu, čiže $c_{ij} = c_{ji}$, ľahko zistíme, že je splnená podmienka (i). Podmienky (ii) a (iii) sú splnené pre ľubovoľnú maticu.

S podmienkou (iv) je to o niečo komplikovanejšie. Budeme sa ňou zaoberať neskôr.

{PRIKLINT}

Príklad 7.1.6. Ako $C(a, b)$ označíme množinu všetkých spojitých funkcií $f\colon \langle a, b \rangle \to \mathbb{R}$. Tieto funkcie tvoria vektorový podpriestor priestoru všetkých funkcií z $\langle a, b \rangle$ do $\mathbb{R}$ a predpis

$$\langle f, g \rangle = \int_a^b f(x) g(x) dx$$

definuje skalárny súčin na tomto priestore. (Takto by sme vedeli definovať skalárny súčin aj na podstatne väčšom priestore funkcií – stačilo by zvoliť nejaké podmienky, ktoré zaručia, že súčin $f(x)g(x)$ bude mať konečný integrál. Keď však použijeme aj nespojité funkcie, budeme mať problémy pri overovaní podmienky (iv). Pravdepodobne v niektorom z vyšších ročníkov sa na analýze stretnete s Fourierovými radmi, kde sa objaví tento istý skalárny súčin a dozviete sa tam aj ako sa to dá definovať tak, aby to fungovalo aj pre iné funkcie, nielen spojité.)

Ak máme euklidovský vektorový priestor, tak môžeme prirodzeným spôsobom zadefinovať veľkosť vektora a uhol dvoch vektorov. Uvidíme, že takto zavedená veľkosť vektora spĺňa viaceré vlastnosti, ktoré platia pre veľkosť vektora v $\mathbb{R}^2$ a $\mathbb{R}^3$.

Definícia 7.1.7. Nech V je euklidovský vektorový priestor. Potom pre $\vec{\alpha} \in V$ definujeme veľkosť vektora $\vec{\alpha}$ ako

$$|\vec{\alpha}| = \sqrt{\langle \vec{\alpha}, \vec{\alpha} \rangle}.$$

Všimnite si, že podmienka (iv) z definície skalárneho súčinu zaručí, že veľkosť je definovaná pre ľubovoľný vektor (nikdy v predpise pre $|\vec{\alpha}|$ nedostaneme odmocninu zo záporného čísla.)

Niekedy sa používa aj označenie $\|\vec{\alpha}\|$ (napríklad v [KGGS]).

{skal:TVRVELK}

Tvrdenie 7.1.8. *Nech V je euklidovský vektorový priestor. Pre ľubovoľné $\vec{\alpha}, \vec{\beta} \in V$ a $c \in \mathbb{R}$ platí:*

{TVSK1} (i) $|\vec{\alpha}| \geq 0$

{TVSK2} (ii) $|\vec{\alpha}| = 0 \Leftrightarrow \vec{\alpha} = \vec{0}$

{TVSK3} (iii) $|c\vec{\alpha}| = |c||\vec{\alpha}|$

{TVSK4} (iv) $|\langle \vec{\alpha}, \vec{\beta} \rangle| \leq |\vec{\alpha}||\vec{\beta}|$ (*Schwarzova nerovnosť*)

{TVSK5} (v) $|\vec{\alpha} + \vec{\beta}| \leq |\vec{\alpha}| + |\vec{\beta}|$ (*trojuholníková nerovnosť*)

V (iv) *nastáva rovnosť práve vtedy, keď vektor $\vec{\alpha}$ je násobkom vektora $\vec{\beta}$.*
V (v) *nastane rovnosť, ak $\vec{\alpha}$ je nezáporným násobkom $\vec{\beta}$.*

Nerovnosť z (iv) môžete stretnúť aj pod názvom *Cauchy-Schwarzova* alebo *Cauchy-Buňakovského nerovnosť.*

Dôkaz. Vlastnosti (i), (ii), (iii) sa overia ľahko priamo z definície.

(iv) Dokážeme použitím vlastnosti (i) pre vektor $\vec{\alpha} + c\vec{\beta}$, kde c je ľubovoľné reálne číslo. Na základe vlastností skalárneho súčinu môžeme urobiť tieto úpravy:

$$|\vec{\alpha} + c\vec{\beta}|^2 = \langle \vec{\alpha} + c\vec{\beta}, \vec{\alpha} + c\vec{\beta} \rangle = \langle \vec{\alpha}, \vec{\alpha} \rangle + 2c\langle \vec{\alpha}, \vec{\beta} \rangle + c^2 \langle \vec{\beta}, \vec{\beta} \rangle = |\vec{\alpha}|^2 + 2c\langle \vec{\alpha}, \vec{\beta} \rangle + c^2 |\vec{\beta}|^2 \geq 0.$$

Pretože uvedená nerovnosť má platiť pre každé reálne číslo c a môžeme ju chápať ako kvadratickú nerovnicu s neznámou c, diskriminant tejto nerovnice nesmie byť kladný (aby príslušná kvadratická rovnica nemala nenulové reálne korene)

$$D = 4\langle \vec{\alpha}, \vec{\beta} \rangle^2 - 4|\vec{\alpha}|^2 |\vec{\beta}|^2 \leq 0.$$

Z tejto nerovnosti dostaneme

$$\langle \vec{\alpha}, \vec{\beta} \rangle^2 \leq |\vec{\alpha}|^2 |\vec{\beta}|^2$$
$$|\langle \vec{\alpha}, \vec{\beta} \rangle| \leq |\vec{\alpha}||\vec{\beta}|$$

Tým je dokázaná platnosť nerovnosti (iv) pre ľubovoľné vektory $\vec{\alpha}$, $\vec{\beta}$. Ešte sa pozrime na otázku, kedy nastáva rovnosť. Z rovnosti $|\langle \vec{\alpha}, \vec{\beta} \rangle| = |\vec{\alpha}||\vec{\beta}|$ vyplýva $D = 0$. Potom existuje také $c \in \mathbb{R}$, že platí $|\vec{\alpha} + c\vec{\beta}| = 0$. To znamená, že $\vec{\alpha} + c\vec{\beta} = 0$, čiže $\alpha = -c\vec{\beta}$. Zistili sme teda, že ak nastane rovnosť, tak $\vec{\alpha}$ musí nutne byť násobkom $\vec{\beta}$. Ľahko sa overí, že ak vektor $\vec{\alpha}$ je násobkom vektora $\vec{\beta}$, tak rovnosť skutočne nastane.

(v) Pokúsme sa upraviť výraz $|\vec{\alpha}+\vec{\beta}|^2$. Platí

$$|\vec{\alpha}+\vec{\beta}|^2 = \langle \vec{\alpha}+\vec{\beta}, \vec{\alpha}+\vec{\beta}\rangle = \langle \vec{\alpha},\vec{\alpha}\rangle + 2\langle \vec{\alpha},\vec{\beta}\rangle + \langle \vec{\beta},\vec{\beta}\rangle = \\ |\vec{\alpha}|^2 + 2\langle \vec{\alpha},\vec{\beta}\rangle + |\vec{\beta}|^2 \overset{(1)}{\leq} |\vec{\alpha}|^2 + 2|\vec{\alpha}||\vec{\beta}| + |\vec{\beta}|^2 = (|\vec{\alpha}|+|\vec{\beta}|)^2$$

(v nerovnosti (1) sme použili Schwarzovu nerovnosť (iv)). Z poslednej nerovnosti už vyplýva (v).

Aby platila rovnosť, musí platiť rovnosť v Schwarzovej nerovnosti použitej v (1). Teda $\vec{\alpha} = k\vec{\beta}$ pre nejaké $k \in \mathbb{R}$. Ľahko sa overí, že k rovnosti dôjde iba v prípade, že $k \geq 0$. (Stačí si uvedomiť, že $|\vec{\alpha}+k\vec{\alpha}| = |1+k|\cdot|\vec{\alpha}|$ a $|\vec{\alpha}|+|k\vec{\alpha}| = (1+|k|)\cdot|\vec{\alpha}|$.) □

Schwarzova nerovnosť pre priestor $\mathbb{R}^n$ s obvyklým skalárnym súčinom sa často používa pri dôkaze rôznych nerovností.

{skal:EQCBRN}
$$\left|\sum_{k=1}^{n} x_k y_k\right| \leq \sqrt{\sum_{k=1}^{n} x_k^2 \sum_{k=1}^{n} y_k^2} \tag{7.1}$$

Definícia 7.1.9. Nech V je euklidovský vektorový priestor.

Uhol dvoch nenulových vektorov definujeme ako taký uhol, pre ktorý platí

$$\cos\varphi = \frac{\langle \vec{\alpha},\vec{\beta}\rangle}{|\vec{\alpha}||\vec{\beta}|}.$$

V prípade, že niektorý z vektorov je nulový, položíme $\varphi = 0$.

Všimnite si, že vďaka Schwarzovej nerovnosti je výraz vystupujúci v definícii uhla dvoch vektorov z intervalu $\langle -1, 1\rangle$, teda takýto uhol skutočne existuje.

Definícia 7.1.10. Vektory $\vec{\alpha},\vec{\beta} \in V$ nazveme *kolmé* (*ortogonálne*), ak $\langle \vec{\alpha},\vec{\beta}\rangle = 0$.

O k-tici vektorov $\vec{\alpha}_1,\ldots,\vec{\alpha}_k$ hovoríme, že tieto vektory sú ortogonálne, ak ľubovoľné 2 z nich sú ortogonálne, t.j. $\langle \vec{\alpha}_i,\vec{\alpha}_j\rangle = 0$ pre každé $i \neq j$.

{TVRORTOGLNZ}
Tvrdenie 7.1.11. *Nech V je euklidovský vektorový priestor. Ak nenulové vektory $\vec{\alpha}_1,\ldots,\vec{\alpha}_k$ sú ortogonálne, tak sú lineárne nezávislé.*

Dôkaz. Nech $c_1,\ldots,c_k \in \mathbb{R}$ sú také, že $c_1\vec{\alpha}_1 + \cdots + c_k\vec{\alpha}_k = \vec{0}$. Zoberme ľubovoľné $i \in \{1,2,\ldots,k\}$. Potom dostaneme

$$0 = \langle \vec{\alpha}_i,\vec{0}\rangle = \langle \vec{\alpha}_i, c_1\vec{\alpha}_1 + \cdots + c_k\vec{\alpha}_k\rangle = c_i|\vec{\alpha}_i|^2.$$

Táto rovnosť môže platiť jedine ak $\vec{\alpha}_i = \vec{0}$ alebo $c_i = 0$. Pretože $\vec{\alpha}_i \neq \vec{0}$, platí $c_i = 0$. Použitím rovnakej úvahy pre všetky $i = 1,2,\ldots,k$ dostaneme $c_1 = c_2 = \ldots = c_k = 0$. Teda dané vektory sú lineárne nezávislé. □

Definícia 7.1.12. Nech V je euklidovský priestor a $M \subseteq V$. Potom

$$M^\perp = \{\vec{\alpha} \in V; \langle \vec{\alpha},\vec{\beta}\rangle = 0 \text{ pre všetky } \vec{\beta} \in M\}$$

sa nazýva *ortogonálny doplnok* množiny M.

Tvrdenie 7.1.13. *Nech V je euklidovský priestor a $M \subseteq V$. Potom $M^\perp$ je vektorový podpriestor priestoru V.*

Dôkaz. Zrejme $\vec{0} \in M^\perp$, preto $M^\perp$ je neprázdna množina.

Treba ešte overiť, že pre všetky $\vec{\alpha}_1, \vec{\alpha}_2 \in M^\perp$ a $c, d \in \mathbb{R}$ aj $c\vec{\alpha}_1 + d\vec{\alpha}_2 \in M^\perp$. Ak pre všetky $\vec{\beta} \in M$ platí $\langle \vec{\alpha}_1, \vec{\beta} \rangle = 0$ a $\langle \vec{\alpha}_2, \vec{\beta} \rangle = 0$, tak aj

$$\langle c\vec{\alpha}_1 + d\vec{\alpha}_2, \vec{\beta} \rangle = c\langle \vec{\alpha}_1, \vec{\beta} \rangle + d\langle \vec{\alpha}_2, \vec{\beta} \rangle = 0,$$

teda $c\vec{\alpha}_1 + d\vec{\alpha}_2 \in M^\perp$. □

Tvrdenie 7.1.14. *Ak V je euklidovský priestor a $M \subseteq N \subseteq V$, tak*

$$N^\perp \subseteq M^\perp.$$

Dôkaz. Ak $\alpha \in N^\perp$, tak $\langle \vec{\alpha}, \vec{\beta} \rangle = 0$ pre všetky vektory $\vec{\beta} \in N$. To ale znamená, že $\langle \vec{\alpha}, \vec{\beta} \rangle = 0$ platí aj pre všetky vektory $\vec{\beta} \in M$ (pretože $M \subseteq N$), a teda $N^\perp \subseteq M^\perp$. □

Lema 7.1.15. *Nech V je euklidovský priestor a $\vec{\alpha}_1, \ldots, \vec{\alpha}_k \in V$. Nech $S = [\vec{\alpha}_1, \ldots, \vec{\alpha}_k]$ je podpriestor vygenerovaný týmito vektormi. Potom $S^\perp = \{\vec{\alpha}_1, \ldots, \vec{\alpha}_k\}^\perp$.*

Dôkaz. Z toho, že $\{\vec{\alpha}_1, \ldots, \vec{\alpha}_k\} \subseteq S$ vyplýva inklúzia $S^\perp \subseteq \{\vec{\alpha}_1, \ldots, \vec{\alpha}_k\}^\perp$.

Naopak, nech $\vec{\beta} \in \{\vec{\alpha}_1, \ldots, \vec{\alpha}_k\}^\perp$. To znamená, že $\langle \vec{\beta}, \vec{\alpha}_i \rangle = 0$ pre $i = 1, 2, \ldots, k$. Potom $\langle \vec{\beta}, \vec{\alpha} \rangle = 0$ aj pre ľubovoľný vektor $\vec{\alpha} \in [\vec{\alpha}_1, \ldots, \vec{\alpha}_k]$, pretože každý vektor z tohoto podpriestoru má tvar $\alpha = c_1\vec{\alpha}_1 + \cdots + c_k\vec{\alpha}_k$ a

$$\langle \vec{\beta}, c_1\vec{\alpha}_1 + \cdots + c_k\vec{\alpha}_k \rangle = c_1\langle \vec{\beta}, \vec{\alpha}_1 \rangle + \ldots + c_k\langle \vec{\beta}, \vec{\alpha}_k \rangle = 0.$$

□

{skal:TVRCAPBOT}

Tvrdenie 7.1.16. *Ak V je euklidovský priestor a S, T sú podpriestory V, tak*

$$(S + T)^\perp = S^\perp \cap T^\perp.$$

Dôkaz. Pretože $S \subseteq S + T$ aj $T \subseteq S + T$ máme $(S + T)^\perp \subseteq S^\perp$ a súčasne $(S + T)^\perp \subseteq T^\perp$, z čoho vyplýva

$$(S + T)^\perp \subseteq S^\perp \cap T^\perp.$$

Naopak, ak $\vec{\alpha} \in S^\perp \cap T^\perp$, tak vektor $\vec{\alpha}$ je kolmý na ľubovoľný vektor z S aj na ľubovoľný vektor z T. Každý vektor z $S + T$ sa dá zapísať v tvare $\vec{\beta} + \vec{\gamma}$, kde $\vec{\beta} \in S$ a $\vec{\gamma} \in T$, takže potom platí

$$\langle \vec{\alpha}, \vec{\beta} + \vec{\gamma} \rangle = \langle \vec{\alpha}, \vec{\beta} \rangle + \langle \vec{\alpha}, \vec{\gamma} \rangle = 0,$$

teda α je kolmý na každý vektor z $S + T$, čiže patrí do $(S + T)^\perp$. Ukázali sme, že platí aj opačná inklúzia

$$S^\perp \cap T^\perp \subseteq (S + T)^\perp.$$

□

7.2 Gram-Schmidtov ortogonalizačný proces

Definícia 7.2.1. Vektory $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ sa nazývajú *ortonormálne,* ak pre všetky i platí $|\vec{\alpha}_i| = 1$ a pre $i \neq j$ platí

$$\langle \vec{\alpha}_i, \vec{\alpha}_j \rangle = 0.$$

Stručne povedané, sú to ortogonálne normované vektory (pod slovom „normované“ rozumieme, že ich veľkosť je 1).

Z tvrdenia 7.1.11 vyplýva, že ortonormálne vektory sú lineárne nezávislé. Ak ich teda bude dosť veľa (v prípade konečnorozmerného priestoru toľko, koľko je dimenzia priestoru), môžu tvoriť bázu.

Definícia 7.2.2. Ak vektory $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ sú ortonormálne a tvoria bázu vektorového priestoru V, tak túto bázu nazývame *ortonormálna báza.*

Príklad 7.2.3. Najjednoduchší príklad je štandardná báza $\vec{e}_1, \ldots, \vec{e}_n$ v priestore $\mathbb{R}^n$ so štandardným skalárnym súčinom

$$\langle \vec{\alpha}, \vec{\beta} \rangle = \sum_{k=1}^{n} a_k b_k.$$

V takomto euklidovskom priestore majú všetky vektory $\vec{e}_1, \ldots, \vec{e}_n$ veľkosť 1 a každý z nich je kolmý na všetky ostatné.

Výhoda ortonormálnej bázy spočíva v tom, že ak máme 2 vektory vyjadrené pomocou ortonormálnej bázy veľmi ľahko vypočítame ich skalárny súčin – v podstate rovnako ako v predchádzajúcom príklade.

Majme $\vec{\alpha} = c_1\vec{\alpha}_1 + \cdots + c_n\vec{\alpha}_n$ a $\vec{\beta} = d_1\vec{\alpha}_1 + \cdots + d_n\vec{\alpha}_n$, kde vektory $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ tvoria ortonormálnu bázu. Potom

$$\langle \vec{\alpha}, \vec{\beta} \rangle = \langle c_1\vec{\alpha}_1 + \cdots + c_n\vec{\alpha}_n, d_1\vec{\alpha}_1 + \cdots + d_n\vec{\alpha}_n \rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} c_i d_j \langle \vec{\alpha}_i, \vec{\alpha}_j \rangle.$$

Jediné nenulové členy v predchádzajúcej sume sú tie, kde $i = j$. Navyše vieme, že $\langle \vec{\alpha}_i, \vec{\alpha}_i \rangle = 1$. Dostaneme teda

$$\langle \vec{\alpha}, \vec{\beta} \rangle = \sum_{i=1}^{n} c_i d_i.$$

Naším najbližším cieľom je ukázať ako z ľubovoľnej bázy v euklidovskom vektorovom priestore vieme dostať ortonormálnu bázu. Dôkaz nasledujúcej vety poskytuje jej konštrukciu, ktorá sa zvykne nazývať Gram-Schmidtov ortogonalizačný proces.

{skal:VTGRAM}

Veta 7.2.4. *Nech V je euklidovský vektorový priestor a $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ je báza priestoru V. Potom existuje ortonormálna báza $\vec{\beta}_1, \ldots, \vec{\beta}_n$ priestoru V.*

Dôkaz. Nech $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ je báza priestoru V. Najprv sa pokúsime nájsť takú bázu $\vec{\gamma}_1, \ldots, \vec{\gamma}_n$ priestoru V, ktorej vektory sú ortogonálne.

$$\begin{aligned}
\vec{\gamma}_1 &= \vec{\alpha}_1 \\
\vec{\gamma}_2 &= \vec{\alpha}_2 + c_{21}\vec{\gamma}_1 \\
\vec{\gamma}_3 &= \vec{\alpha}_3 + c_{31}\vec{\gamma}_1 + c_{32}\vec{\gamma}_2 \\
&\vdots \\
\vec{\gamma}_n &= \vec{\alpha}_n + c_{n1}\vec{\gamma}_1 + c_{n2}\vec{\gamma}_2 + \ldots + c_{n,n-1}\vec{\gamma}_{n-1}
\end{aligned}$$

Budeme postupovať indukciou. Prvý krok indukcie je jasný - stačí položiť $\vec{\gamma}_1 = \vec{\alpha}_1$.

Predpokladajme teraz, že už sme našli k ortogonálnych vektorov $\vec{\gamma}_1, \ldots, \vec{\gamma}_k$, ktoré majú uvedený tvar. Navyše platí

$$[\vec{\alpha}_1, \ldots, \vec{\alpha}_k] = [\vec{\gamma}_1, \ldots, \vec{\gamma}_k].$$

Chceli by sme nájsť vektor $\vec{\gamma}_{k+1}$ kolmý na všetky predchádzajúce, ktorý by mal navyše tvar

$$\vec{\gamma}_{k+1} = \vec{\alpha}_{k+1} + c_{k+1,1}\vec{\gamma}_1 + c_{k+1,2}\vec{\gamma}_2 + \ldots + c_{k+1,k}\vec{\gamma}_k$$

a súčasne taký, aby platilo

$$[\vec{\alpha}_1, \ldots, \vec{\alpha}_{k+1}] = [\vec{\gamma}_1, \ldots, \vec{\gamma}_{k+1}].$$

Ak má byť tento vektor kolmý na predchádzajúce, musia platiť rovnosti

$$\begin{aligned} 0 = \langle \vec{\gamma}_{k+1}, \vec{\gamma}_1 \rangle &= \langle \vec{\alpha}_{k+1}, \vec{\gamma}_1 \rangle + c_{k+1,1}\langle \vec{\gamma}_1, \vec{\gamma}_1 \rangle \\ 0 = \langle \vec{\gamma}_{k+1}, \vec{\gamma}_2 \rangle &= \langle \vec{\alpha}_{k+1}, \vec{\gamma}_2 \rangle + c_{k+1,2}\langle \vec{\gamma}_2, \vec{\gamma}_2 \rangle \\ 0 = \langle \vec{\gamma}_{k+1}, \vec{\gamma}_3 \rangle &= \langle \vec{\alpha}_{k+1}, \vec{\gamma}_3 \rangle + c_{k+1,3}\langle \vec{\gamma}_3, \vec{\gamma}_3 \rangle \\ &\vdots \\ 0 = \langle \vec{\gamma}_{k+1}, \vec{\gamma}_k \rangle &= \langle \vec{\alpha}_{k+1}, \vec{\gamma}_k \rangle + c_{k+1,k}\langle \vec{\gamma}_k, \vec{\gamma}_k \rangle \end{aligned} \tag{7.2}$$

{EQGAMMAK}

(V každej rovnici sme vynechali všetky členy obsahujúce $\langle \vec{\gamma}_i, \vec{\gamma}_j \rangle$ pre $i \neq j$, $i, j \in \{1, 2, \ldots, k\}$, pretože podľa indukčného predpokladu sú tieto hodnoty nulové.) Z predchádzajúcich rovníc môžeme vyjadriť všetky koeficienty $c_{k+1,i}$:

$$c_{k+1,i} = -\frac{\langle \vec{\alpha}_{k+1}, \vec{\gamma}_i \rangle}{\langle \vec{\gamma}_i, \vec{\gamma}_i \rangle}$$

pre každé $i = 1, 2, \ldots, k$.

Z rovníc (7.2) vidno, že pre takéto hodnoty $c_{k+1,i}$ bude vektor $\vec{\gamma}_{k+1}$ skutočne kolmý na všetky predchádzajúce vektory.

Ďalej vieme, že $\vec{\alpha}_{k+1} \notin [\vec{\alpha}_1, \ldots, \vec{\alpha}_k] = [\vec{\gamma}_1, \ldots, \vec{\gamma}_k]$ (lebo vektory $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ sú lineárne nezávislé). Teda aj $\vec{\alpha}_{k+1}, \vec{\gamma}_1, \ldots, \vec{\gamma}_k$ sú lineárne nezávislé, čiže ich lineárnou kombináciou nemôžeme dostať $\vec{0}$. Pretože $\vec{\gamma}_{k+1} = \vec{\alpha}_{k+1} + c_{k+1,1}\vec{\gamma}_1 + c_{k+1,2}\vec{\gamma}_2 + \ldots + c_{k+1,k}\vec{\gamma}_k$ je lineárna kombinácia týchto vektorov a koeficient pri $\vec{\alpha}_{k+1}$ je $1 \neq 0$. Z toho vyplýva, že $\vec{\gamma}_{k+1} \neq \vec{0}$.

Súčasne platí $\vec{\gamma}_{k+1} \in [\vec{\alpha}_{k+1}, \vec{\gamma}_1, \ldots, \vec{\gamma}_k] = [\vec{\alpha}_1, \ldots, \vec{\alpha}_{k+1}]$. Teda $[\vec{\gamma}_1, \ldots, \vec{\gamma}_{k+1}] \subseteq [\vec{\alpha}_1, \ldots, \vec{\alpha}_{k+1}]$.

Ďalej $\vec{\alpha}_{k+1} = \vec{\gamma}_{k+1} - (c_{k+1,1}\vec{\gamma}_1 + c_{k+1,2}\vec{\gamma}_2 + \ldots + c_{k+1,k}\vec{\gamma}_k)$ je lineárna kombinácia vektorov $\vec{\gamma}_1, \ldots, \vec{\gamma}_{k+1}$, čiže platí aj obrátená inklúzia $[\vec{\alpha}_1, \ldots, \vec{\alpha}_{k+1}] \subseteq [\vec{\gamma}_1, \ldots, \vec{\gamma}_{k+1}]$.

Takto sme dostali bázu priestoru V, ktorej vektory sú ortogonálne. Aby boli ortonormálne, potrebujeme, každý z nich vydeliť jeho veľkosťou, čiže ortonormálnu bázu dostaneme tak, že položíme

$$\vec{\beta}_i = \frac{\vec{\gamma}_i}{|\vec{\gamma}_i|}.$$

□

Príklad 7.2.5. Zoberme si priestor $V = [(1, 0, 1, 0), (0, 2, -1, 1), (0, 2, 1, 3)]$. Ľahko sa overí, že tieto vektory sú lineárne nezávislé, teda tvoria bázu priestoru V. Pomocou Gram-Schmidtovho procesu nájdeme ortogonálnu bázu pre V. Položíme

$$\vec{\gamma}_1 = \vec{\alpha}_1 = (1, 0, 1, 0).$$

Ďalej chceme nájsť vektor $\vec{\gamma}_2 = \vec{\alpha}_2 + c\vec{\gamma}_1 = (0, 2, -1, 1) + c(1, 0, 1, 0) = (c, 2, c-1, 1)$ tak, aby bol kolmý na $\vec{\gamma}_1 = (1, 0, 1, 0)$. Dostávame teda rovnosť

$$\langle (c, 2, c-1, c+1), (1, 0, 1, 0) \rangle = c + c - 1 = 2c - 1 = 0,$$

z ktorej vyplýva $c = \frac{1}{2}$ a $\vec{\gamma}_2 = (\frac{1}{2}, 2, -\frac{1}{2}, 1)$.

Tretí vektor $\vec{\gamma}_3$ hľadáme v tvare $\vec{\gamma}_3 = \vec{\alpha}_3 + d\vec{\gamma}_1 + e\vec{\gamma}_2$ tak, aby

$$\langle \vec{\gamma}_3, \vec{\gamma}_1 \rangle = \langle \vec{\alpha}_3, \vec{\gamma}_1 \rangle + d\langle \vec{\gamma}_1, \vec{\gamma}_1 \rangle = 0$$
$$\langle \vec{\gamma}_3, \vec{\gamma}_2 \rangle = \langle \vec{\alpha}_3, \vec{\gamma}_2 \rangle + e\langle \vec{\gamma}_2, \vec{\gamma}_2 \rangle = 0$$

z čoho

$$d = -\frac{\langle \vec{\alpha}_3, \vec{\gamma}_1 \rangle}{\langle \vec{\gamma}_1, \vec{\gamma}_1 \rangle}$$
$$e = -\frac{\langle \vec{\alpha}_3, \vec{\gamma}_2 \rangle}{\langle \vec{\gamma}_2, \vec{\gamma}_2 \rangle}$$

Keď vypočítame $\langle \vec{\alpha}_3, \vec{\gamma}_1 \rangle = 1$, $\langle \vec{\gamma}_1, \vec{\gamma}_1 \rangle = 2$, $\langle \vec{\alpha}_3, \vec{\gamma}_2 \rangle = \frac{13}{2}$ a $\langle \vec{\gamma}_2, \vec{\gamma}_2 \rangle = \frac{11}{2}$, dostaneme

$$d = -\frac{1}{2}$$
$$e = -\frac{13}{11}$$
$$\vec{\gamma}_3 = \vec{\alpha}_3 - \frac{1}{2}\vec{\gamma}_1 - \frac{13}{11}\vec{\gamma}_2$$
$$\vec{\gamma}_3 = \left(-\frac{12}{11}, -\frac{4}{11}, \frac{12}{11}, \frac{20}{11}\right)$$

Zatiaľ sme teda dostali ortogonálne vektory, ktoré generujú V. Aby sme z nich dostali ortonormálne, musíme ich predeliť veľkosťou. Platí

$$|\vec{\gamma}_1| = \sqrt{2}$$
$$|\vec{\gamma}_2| = \frac{\sqrt{11}}{\sqrt{2}}$$
$$|\vec{\gamma}_3| = \frac{\sqrt{704}}{11} = \frac{8\sqrt{11}}{11} = \frac{8}{\sqrt{11}}$$

a teda ortonormálna báza priestoru V je

$$\vec{\beta}_1 = \frac{1}{\sqrt{2}}(1, 0, 1, 0)$$
$$\vec{\beta}_2 = \frac{\sqrt{2}}{\sqrt{11}}(\frac{1}{2}, 2, -\frac{1}{2}, 1)$$
$$\vec{\beta}_3 = \frac{\sqrt{11}}{8}\left(-\frac{12}{11}, -\frac{4}{11}, \frac{12}{11}, \frac{20}{11}\right) = \frac{1}{8\sqrt{11}}(-12, -4, 12, 20) = \frac{2}{\sqrt{11}}(-3, -1, 3, 5).$$

Vidíme, že vektory, ktoré sme dostali vyzerajú pomerne zložito. Nevedeli by sme si nejako zjednodušiť tieto výpočty? Možnože keby sme mali bázové vektory pôvodnej bázy o niečo jednoduchšie, aj ortonormálna báza by vyšla jednoduchšia. Ale dostať „peknú“ bázu vieme – to sa dá urobiť pomocou elementárnych riadkových operácií. Takže skúsme ešte takýto postup – vypočítajme najprv jednoduchšiu bázu pre priestor V.

$\begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 2 & -1 & 1 \\ 0 & 2 & 1 & 3 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 2 & -1 & 1 \\ 0 & 0 & 2 & 2 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 2 & -1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 2 & 0 & 2 \\ 0 & 0 & 1 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$

Vidíme teda, že $V = [(1, 0, 0, -1), (0, 1, 0, 1), (0, 0, 1, 1)]$, čiže tentokrát ako štartovací bod pre Gram-Schmidtovu ortogonalizáciu použijeme bázové vektory $\vec{\alpha}_1 = (1, 0, 0, -1)$, $\vec{\alpha}_2 =$

$(0,1,0,1)$, $\vec{\alpha}_3 = (0,0,1,1)$. (Dúfam, že Vás nebude príliš pliesť, že tentokrát sme ako $\vec{\alpha}_1$, $\vec{\alpha}_2$ a $\vec{\alpha}_3$ označili úplne iné vektory ako v prvej časti príkladu. Dôvod nie je ten, že by sme mali príliš málo gréckych písmeniek, ale ten, že som chcel aby sa označenie zhodovalo s označením použitým v predchádzajúcom dôkaze.)

Opäť dostaneme:

$$\vec{\gamma}_1 = \vec{\alpha}_1 = (1,0,0,-1).$$

Vektor $\vec{\gamma}_2$ hľadáme v tvare $\vec{\alpha}_2 + c\vec{\gamma}_1$ a z podmienky, že $\langle \vec{\gamma}_2, \vec{\gamma}_1 \rangle = 0$ nám vyjde, že

$$c = -\frac{\langle \vec{\gamma}_1, \vec{\alpha}_2 \rangle}{\langle \vec{\gamma}_1, \vec{\gamma}_1 \rangle} = -\frac{-1}{2} = \frac{1}{2}$$

$$\vec{\gamma}_2 = (0,1,0,1) + \frac{1}{2}(1,0,0,-1) = \left(\frac{1}{2}, 1, 0, \frac{1}{2}\right)$$

Ďalej hľadáme $\vec{\gamma}_3$ v tvare $\vec{\gamma}_3 = \vec{\alpha}_3 + d\vec{\gamma}_1 + e\vec{\gamma}_2$. Koeficienty e a f opäť určíme z podmienok ortogonality.

$$d = -\frac{\langle \vec{\alpha}_3, \vec{\gamma}_1 \rangle}{\langle \vec{\gamma}_1, \vec{\gamma}_1 \rangle} = \frac{1}{2}$$

$$e = -\frac{\langle \vec{\alpha}_3, \vec{\gamma}_2 \rangle}{\langle \vec{\gamma}_2, \vec{\gamma}_2 \rangle} = -\frac{1}{3}$$

$$\vec{\gamma}_3 = (0,0,1,1) + \frac{1}{2}(1,0,0,-1) - \frac{1}{3}\left(\frac{1}{2}, 1, 0, \frac{1}{2}\right) = \left(\frac{1}{3}, -\frac{1}{3}, 1, \frac{1}{3}\right)$$

Teraz už zostáva len každý vektor predeliť jeho veľkosťou.

$$\vec{\beta}_1 = \frac{\vec{\gamma}_1}{|\vec{\gamma}_1|} = \frac{1}{\sqrt{2}}(1,0,0,-1)$$

$$\vec{\beta}_2 = \frac{\vec{\gamma}_2}{|\vec{\gamma}_2|} = \sqrt{\frac{2}{3}}\left(\frac{1}{2}, 1, 0, \frac{1}{2}\right)$$

$$\vec{\beta}_3 = \frac{\vec{\gamma}_3}{|\vec{\gamma}_3|} = \frac{\sqrt{3}}{2}\left(\frac{1}{3}, -\frac{1}{3}, 1, \frac{1}{3}\right)$$

V predošlom príklad sme použili presne postup z dôkazu. Poďme si ešte ukázať inú možnosť ako by sme mohli nájsť ortogonálnu bázu podpriestoru V z predošlej úlohy.

Príklad 7.2.6. Chceme hľadať bázu podpriestoru $V = [(1,0,0,-1),(0,1,0,1),(0,0,1,1)]$. Začnime tým, že si uvedomíme, že tento podpriestor je presne množina vektorov kolmých na $(1,-1,-1,1)$. Inak povedané, našli sme $V^\perp = [(1,-1,-1,1)]$. (Ten vlastne vieme získať jednoducho riešením homogénnej sústavy, ktorej riadky sú bázové vektory podpriestoru V.)

Toto sa dá spraviť pre ľubovoľný podpriestor. Môžeme si tiež uvedomiť, že inak sa dá na to pozrieť tak, že sme vlastne vyjadrili tento podpriestor v tvare $V = \{(x_1,x_2,x_3,x_4) \in \mathbb{R}^4; x_1 - x_2 - x_3 + x_4 = 1\}$. Teda vlastne je to množina riešení homogénnej sústavy. V tomto prípade ide o veľmi jednoduchú sústavu pozostávajúcu iba z jednej rovnice. Ale z minulého semestra vieme, že každý podpriestor je množinou riešení nejakej sústavy (veta 5.7.11).

Vyjadrili sme podpriestor V trochu iným spôsobom, ktorý by nám mal pomôcť pri hľadaní ortogonálnej bázy. Máme teda pôvodnú bázu $\vec{\alpha}_1$, $\vec{\alpha}_2$, $\vec{\alpha}_3$, ktorú by sme chceli trochu pomeniť.

Začnime tým, že $\vec{\gamma}_1 = \vec{\alpha}_1 = (1,0,0,-1)$. Teraz by sme chceli nájsť ďalší vektor. Od neho chceme aby bol kolmý na $\vec{\gamma}_1$. Navyše chceme aby patril do V, čo je ekvivalentné s tým, že je kolmý na $(1,-1,-1,1)$. Tieto dve podmienky nám dajú sústavu dvoch homogénnych rovníc:

$$\begin{pmatrix} 1 & 0 & 0 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & -1 & -1 & 2 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & -2 \end{pmatrix}$$

Za $\vec{\gamma}_2$ môžeme zobrať ľubovoľné nenulové riešenie tejto sústavy – priestor riešení je dvojrozmerný, máme teda veľa možností. Zoberme napríklad $\vec{\gamma}_2 = (1, 2, 0, 1)$.

Teraz by sme chceli nájsť vektor $\vec{\gamma}_3$, ktorý má spĺňať obe podmienky určené predošlou sústavou (má byť kolmý na $\vec{\gamma}_1$ a patriť) do V. Ale pribudne nám aj nová podmienka, že má byť kolmý na $\vec{\gamma}_2$, a tým aj nová rovnica.

$$\begin{pmatrix} 1 & 0 & 0 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & 2 & 0 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & -2 \\ 0 & 2 & 0 & 2 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & -2 \\ 0 & 1 & 0 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -3 \end{pmatrix}$$

Opäť za $\vec{\gamma}_3$ môžeme voliť ľubovoľné nenulové riešenie tejto sústavy. Teraz už je podpriestor riešení iba jednorozmerný, čiže máme o čosi menšiu voľnosť. (Vektor $\vec{\gamma}_3$ je určený jednoznačne až na násobok.) Zoberme $\vec{\gamma}_3 = (1, -1, 3, 1)$.

Ak by sme chceli dostať ortonormálnu bázu, tak tieto vektory ešte treba vydeliť ich veľkosťou. Všimnime si, že takto dostaneme presne rovnaké vektory ako pri predošlom postupe. (Samozrejme, $\vec{\gamma}_2$ sme mohli voliť aj inak, čo by potom ovplyvnilo aj to, ako by vyzeralo $\vec{\gamma}_3$. Tu sme ich naschvál volili tak, aby vyšiel rovnaký výsledok. Môžete si vyskúšať nejaký iný výber, ktorým dostanete inú ortogonálnu resp. ortonormálnu bázu.)

Existenciu ortogonálnej bázy môžeme použiť na dôkaz niektorých ďalších faktov o ortogonálnom doplnku.

Veta 7.2.7. *Nech S je podpriestor konečnorozmerného euklidovského priestoru V. Potom ľubovoľný vektor $\vec{\gamma} \in V$ sa dá jednoznačne vyjadriť ako*

$$\vec{\gamma} = \vec{\alpha} + \vec{\beta},$$

kde $\vec{\alpha} \in S$ a $\vec{\beta} \in S^\perp$.

Dôkaz. Existencia: Vieme, že S má ortonormálnu bázu a tú môžeme rozšíriť na ortonormálnu bázu celého V. (Presnejšie povedané: Vieme ju podľa Steinitzovej vety rozšíriť na bázu celého V, ak z tejto bázy postupom použitým v dôkaze vety 7.2.4 vytvoríme ortonormálnu bázu, tak bázové vektory patriace do S sa nezmenia, pretože boli ortonormálne už pred ortonormalizáciou.)

Nech teda vektory $\vec{\gamma}_1, \ldots, \vec{\gamma}_k$ tvoria ortonormálnu bázu S a vektory $\vec{\gamma}_{k+1}, \ldots, \vec{\gamma}_n$ sú ostatné vektory ortonormálnej bázy V. Ľubovoľný vektor $\vec{\gamma}$ sa dá jednoznačne zapísať ako

$$\vec{\gamma} = c_1\vec{\gamma}_1 + \cdots + c_k\vec{\gamma}_k + c_{k+1}\vec{\gamma}_{k+1} + \cdots + c_n\vec{\gamma}_n.$$

Ak zvolíme $\vec{\alpha} = c_1\vec{\gamma}_1 + \cdots + c_k\vec{\gamma}_k$ a $\vec{\beta} = c_{k+1}\vec{\gamma}_{k+1} + \cdots + c_n\vec{\gamma}_n$, tak $\vec{\alpha} \in S$ a $\vec{\beta} \in S^\perp$.

Jednoznačnosť: Majme dva rozklady uvedeným spôsobom, t.j.

$$\vec{\gamma} = \vec{\alpha}_1 + \vec{\beta}_1 = \vec{\alpha}_2 + \vec{\beta}_2,$$

pričom $\vec{\alpha}_1, \vec{\alpha}_2 \in S$ a $\vec{\beta}_1, \vec{\beta}_2 \in S^\perp$.

Z rovnosti

$$\vec{\alpha}_1 - \vec{\alpha}_2 = \vec{\beta}_2 - \vec{\beta}_1$$

vidíme, že vektor $\vec{\alpha}_1 - \vec{\alpha}_2$ patrí do $S \cap S^\perp$. Z toho ale potom vyplýva

$$\langle \vec{\alpha}_1 - \vec{\alpha}_2, \vec{\alpha}_1 - \vec{\alpha}_2 \rangle = 0$$

a $\vec{\alpha}_1 - \vec{\alpha}_2 = 0$, čiže $\vec{\alpha}_1 = \vec{\alpha}_2$. Samozrejme, potom musí platiť aj $\vec{\beta}_1 = \vec{\beta}_2$. □

Definícia 7.2.8. V situácii z predošlej vety sa vektor $\vec{\alpha}$ nazýva *ortogonálna projekcia* vektora $\vec{\gamma}$ na podpriestor S.

Termín *ortogonálna projekcia* sa často používa aj pre zobrazenie $P\colon V \to V$, ktoré danému vektoru priradí jeho ortogonálnu projekciu. Ľahko sa overí, že toto zobrazenie je lineárne (úloha 7.2.12).

{skal:DOSORTON}

Dôsledok 7.2.9. *Nech S, T sú podpriestory konečnorozmerného priestoru V. Potom:*

{skal:DOSORTONit1}

(i) $V = S \oplus S^\perp$

{skal:DOSORTONit2}

(ii) $(S^\perp)^\perp = S$

{skal:DOSORTONit3}

(iii) $(S \cap T)^\perp = S^\perp + T^\perp$.

Dôkaz. (i) Vyplýva z predchádzajúcej vety a z vety 4.5.6.

(ii) Priamo z definície vidno, že $S \subseteq (S^\perp)^\perp$. (Každý vektor z S je kolmý na všetky vektory z $S^\perp$.) Súčasne máme

$$V = S \oplus S^\perp = (S^\perp)^\perp \oplus S^\perp,$$

z čoho pre dimenzie dostaneme

$$d(S) + d(S^\perp) = d((S^\perp)^\perp) + d(S^\perp),$$

a teda $d(S) = d((S^\perp)^\perp)$. Keďže S je podpriestor $(S^\perp)^\perp$ a majú rovnakú dimenziu, platí $S = S^\perp$ (tvrdenie 4.4.18).

(iii) Použitím tvrdenia 7.1.16 a časti (ii) dostaneme

$$(S^\perp + T^\perp)^\perp = (S^\perp)^\perp \cap (T^\perp)^\perp = S \cap T.$$

Ak ešte raz aplikujeme operátor ortogonálneho doplnku a použijeme (ii), dostávame rovnosť

$$S^\perp + T^\perp = (S \cap T)^\perp.$$

□

Nasledujúci príklad ukazuje, že v nekonečnorozmerných priestoroch tvrdenia dokázané v predchádzajúcom dôsledku neplatia vo všeobecnosti v prípade, že euklidovský vektorový priestor a podpriestory vystupujúce v dôsledku sú nekonečnorozmerné. (Tento príklad je o čosi komplikovanejší, ale aspoň pre tých z vás, ktorých zaujíma analýza, by mohol byť zaujímavý.)

Príklad* 7.2.10. Priestor, v ktorom budeme pracovať je priestor postupností

$$V = \ell_2 = \{(x_n) \in \mathbb{R}^{\mathbb{N}}; \sum_{n=1}^{\infty} x_n^2 < +\infty\}.$$

Skalárny súčin, s ktorým budeme pracovať, je

$$\langle x, y\rangle = \sum_{n=1}^{\infty} x_n y_n.$$

Tento priestor hrá dôležitú úlohu v matematickej analýze.

Fakt, že tento predpis naozaj určuje zobrazenie $\ell_2 \times \ell_2 \to \mathbb{R}$ (teda, že pre každé 2 postupnosti z ℓ_2 je súčet $\sum_{n=1}^{\infty} x_n y_n$ konečný) vyplýva z nerovnosti (7.1). Stačí v nej zobrať limitu pre $n \to \infty$ a dostaneme nerovnosť

$$\left|\sum_{k=1}^{\infty} x_k y_k\right| \leq \sqrt{\sum_{k=1}^{\infty} x_k^2 \sum_{k=1}^{\infty} y_k^2},$$

čo je presne nerovnosť, ktorá sa nám hodí na tomto mieste.

Overenie jednotlivých vlastností skalárneho súčinu je len o niečo zložitejšie ako v príklade 7.1.3.

Stále sme však ešte neoverili, že ide o euklidovský vektorový priestor – chýba nám overenie podmienky, s ktorou obvykle začíname, t.j. to, že V je vektorový priestor. Keďže ide o podmnožinu vektorového priestoru $\mathbb{R}^{\mathbb{N}}$ (a operácie sú definované rovnako), stačí overiť uzavretosť na súčet a skalárny násobok. Netriviálna je iba uzavretosť na súčet. Ak $(x_n), (y_n) \in \ell_2$, znamená to, že rady $\sum_{n=1}^{\infty} x_n^2$ i $\sum_{n=1}^{\infty} y_n^2$ konvergujú. Potom máme

{skal:EQL2SUCET}
$$\sum_{n=1}^{\infty}(x_n + y_n)^2 = \sum_{n=1}^{\infty}(x_n^2 + 2x_n y_n + y_n^2) \overset{(*)}{=} \sum_{n=1}^{\infty} x_n^2 + 2\sum_{n=1}^{\infty} x_n y_n + \sum_{n=1}^{\infty} y_n^2. \tag{7.3}$$

Dôležité je uvedomiť si, či skutočne platí rovnosť (*), t.j. či môžeme takýmto spôsobom zmeniť poradie sumácie. Z matematickej analýzy vieme, že sa to dá urobiť, ak rady, ktoré sčitujeme sú absolútne konvergentné.[1] Keďže rady $\sum_{n=1}^{\infty} x_n^2$ a $\sum_{n=1}^{\infty} y_n^2$ sú rady s kladnými členmi, pre ne je absolútna konvergencia zrejmá. V prípade radu $\sum_{n=1}^{\infty} x_n y_n$ si stačí všimnúť, že platí nerovnosť

$$\sum_{n=1}^{\infty}|x_n y_n| \leq \sqrt{\sum_{n=1}^{\infty} x_n^2 . \sum_{n=1}^{\infty} y_n^2}.$$

Túto nerovnosť môžeme dostať napríklad limitným prechodom z (7.1) (z (7.1) vieme, že uvedená nerovnosť platí pre všetky čiastočné súčty radov, ktoré v nej vystupujú).

Vidíme teda, že rady vystupujúce na pravej strane rovnosti (7.3) sú absolútne konvergentné, čím mám dokázanú platnosť tejto rovnosti.

Navyše, keď si ešte uvedomíme, že platí

$$\sum_{n=1}^{\infty} x_n y_n \leq \sum_{n=1}^{\infty}|x_n y_n|,$$

tak vidíme, že všetky rady na pravej strane (7.3) majú konečný súčet, teda platí $\sum_{n=1}^{\infty}(x_n + y_n)^2 < +\infty$, čo znamená, že $(x_n + y_n) \in \ell_2$.

Zvoľme si teraz podpriestor

$$S = \{(x_n) \in \ell_2; x_n = 0 \text{ pre všetky } n \text{ okrem konečného počtu}\}$$

[1] Pripomeňme, že rad $\sum_{n=1}^{\infty} a_n$ je absolútne konvergentný, ak konverguje rad $\sum_{n=1}^{\infty}|a_n|$.

pozostávajúci z tých postupností, ktoré majú iba konečne veľa nenulových členov. Platí

$$S^\perp = \{0\}.$$

Stačí si uvedomiť, že ak ako e_n označíme postupnosť, ktorá má všetky členy okrem n-tého miesta nuly a jej n-tý člen 1, t.j. $e_n = (0, \dots, 0, 1, 0 \dots)$, tak všetky takéto postupnosti patria do S. Teda pre každú postupnosť z $S^\perp$ dostaneme

$$\langle x, e_n \rangle = x_n = 0.$$

Z toho dostaneme

$$S \oplus S^\perp = S \neq V,$$
$$(S^\perp)^\perp = \{0\}^\perp = V.$$

Cvičenia

{skalcvic:ULOBAZABOT}

Úloha 7.2.1. Nájdite bázu a dimenziu $S^\perp$ pre daný podpriestor S priestoru $\mathbb{R}^4$:
a) $S = [(1,1,0,1),(2,1,0,1)]$
b) $S = [(1,5,4,3),(2,-1,2,-1)]$
c) $S = [(1,2,1,1),(2,1,-1,-1)]$
d) $S = [(1,2,3,4),(1,1,1,1),(4,3,2,1)]$
e) $S = [(2,1,2,3),(0,1,-2,1),(1,0,2,1)]$
f) $S = [(1,1,1,2),(1,0,1,1),(0,1,2,1)]$

Úloha 7.2.2. Zistite, či daný predpis určuje skalárny súčin na $\mathbb{R}^3$. Nech $\vec{\alpha} = (a_1, a_2, a_3)$ a $\vec{\beta} = (b_1, b_2, b_3)$.
a) $\langle \vec{\alpha}, \vec{\beta} \rangle = a_1b_1 - a_1b_2 + a_1b_3 + a_2b_1 + 3a_2b_2 - a_3b_3$
b) $\langle \vec{\alpha}, \vec{\beta} \rangle = a_1b_1 + 2a_1b_2 + 2a_2b_1$
c) $\langle \vec{\alpha}, \vec{\beta} \rangle = 3a_1b_2 + 2a_2b_2 + a_3b_3$
d) $\langle \vec{\alpha}, \vec{\beta} \rangle = a_1b_1 + a_2b_2 + a_3b_3$
e) $\langle \vec{\alpha}, \vec{\beta} \rangle = a_1b_1 + a_1b_2 + a_2b_1 + 3a_2b_2 + a_3b_3$
f) $\langle \vec{\alpha}, \vec{\beta} \rangle = a_1b_1 + a_1b_2 + a_2b_1 + a_2b_2 + 2a_3b_3$
g) $\langle \vec{\alpha}, \vec{\beta} \rangle = a_1b_1 + 2a_1b_2 + 2a_2b_1 + a_2b_2 + 2a_3b_3$
h) $\langle \vec{\alpha}, \vec{\beta} \rangle = a_1b_2 + a_2b_1$
i) $\langle \vec{\alpha}, \vec{\beta} \rangle = 3a_1b_1 + 2a_1b_2 + a_2b_1 + 3a_3b_3$

Úloha 7.2.3. Zistite, či $\sin \pi x$ a $\cos \pi x$ sú kolmé v priestore $C(0,1)$ so skalárnym súčinom z príkladu 7.1.6. Akú majú tieto vektory veľkosť?

Úloha 7.2.4. Overte či predpis
a) $\langle f, g \rangle = f(0)g(0) + f(1)g(1)$
b) $\langle f, g \rangle = f(-1)g(-1) + f(0)g(0) + f(1)g(1)$
určuje skalárny súčin na priestore P_2 všetkých polynómov stupňa najviac 2 nad poľom $\mathbb{R}$.

Úloha 7.2.5. Ukáže, že pre ľubovoľné dva vektory $\vec{\alpha}$, $\vec{\beta}$ v euklidovskom vektorovom priestore platí $|\vec{\alpha}| = |\vec{\beta}|$ práve vtedy, keď vektory $\vec{\alpha} - \vec{\beta}$ a $\vec{\alpha} + \vec{\beta}$ sú na seba kolmé.

Úloha 7.2.6. Dokážte, že v ľubovoľnom euklidovskom priestore platí:
a) $\langle \vec{\alpha}, \vec{\beta} \rangle = 0 \Rightarrow |\vec{\alpha} + \vec{\beta}|^2 = |\alpha|^2 + |\beta|^2$ (Pytagorova veta)
b) $|\vec{\alpha} + \vec{\beta}|^2 = |\alpha|^2 + |\beta|^2 + 2\langle \vec{\alpha}, \vec{\beta} \rangle$ (kosínová veta)
c) $|\vec{\alpha} + \vec{\beta}|^2 + |\vec{\alpha} - \vec{\beta}|^2 = 2(|\vec{\alpha}|^2 + |\vec{\beta}|^2)$ (rovnobežníkové pravidlo)

Úloha 7.2.7*. Ukážte, že ak $|\cdot|\colon V \to \mathbb{R}$ je funkcia definovaná na vektorovom priestore V nad $\mathbb{R}$, ktorá spĺňa podmienky (i), (ii), (iii) a (v) z tvrdenia 7.1.8 i rovnobežníkové pravidlo, tak existuje skalárny súčin na V taký, že $|\alpha| = \sqrt{\langle\vec{\alpha},\vec{\alpha}\rangle}$ pre všetky $\vec{\alpha} \in V$.

Úloha 7.2.8. Ukážte, že funkcia $|\cdot|\colon \mathbb{R}^n \to \mathbb{R}$, $|(x_1,\ldots,x_n)| = \max\{|x_i|; i = 1,\ldots,n\}$ spĺňa podmienky (i), (ii), (iii) a (v) z tvrdenia 7.1.8, ale neexistuje skalárny súčin na $\mathbb{R}^n$ taký, že $|\alpha| = \sqrt{\langle\vec{\alpha},\vec{\alpha}\rangle}$ (pre všetky $\vec{\alpha} \in \mathbb{R}^n$).

{skalcvic:ULOSTOPA}

Úloha 7.2.9. Pre štvorcovú maticu typu $n \times n$ definujeme stopu matice ako súčet jej diagonálnych prvkov, t.j. $\mathrm{Tr}(A) = \sum_{i=1}^{n} a_{ii}$. Overte, či na vektorovom priestore $M_{n,n}(\mathbb{R})$ určuje predpis

$$\langle A, B\rangle = \mathrm{Tr}(AB^T)$$

skalárny súčin. (Hint 1: Pokúste sa vyjadriť hodnotu $\langle A, B\rangle$ pomocou prvkov matíc A, B. Hint 2: Možno vám pri tom pomôžu rovnosti $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$ a $\mathrm{Tr}(A) = \mathrm{Tr}(A^T)$. Prvá z nich sa dá ľahko overiť pomocou definície súčinu, druhá je zrejmá.)

Úloha 7.2.10. Overte, že v priestore $C(0, 2\pi)$ všetkých spojitých funkcií z uzavretého intervalu $\langle 0, 2\pi\rangle$ do $\mathbb{R}$ so skalárnym súčinom $\langle f, g\rangle = \int_0^{2\pi} f(x)g(x)dx$ sú ľubovoľné dve rôzne funkcie z množiny $\{1, \sin nx, \cos nx; n \in \mathbb{N}\}$ na seba kolmé. (Po vynormovaní by sme dostali množinu funkcií, ktorá má v tomto priestore do istej miery podobné vlastnosti ako ortonormálna báza v konečnorozmerných priestorov. Tento systém funkcií je dôležitý v matematickej analýze v súvislosti s *Fourierovými radmi.*)

Úloha 7.2.11. Nájdite ortonormálnu bázu pre priestory z úlohy 7.2.1.

{skalcvic:ULOPROJJELIN}

Úloha 7.2.12. Nech S je podpriestor konečnorozmerného euklidovského vektorového priestoru V. Nech $P\colon V \to V$ je ortogonálna projekcia na tento podpriestor. Overte, že:
a) P je lineárne zobrazenie;
b) $\mathrm{Im}\, P = S$ a $\mathrm{Ker}\, P = S^\perp$;
c) $P \circ P = P$.

{skalcvic:ULOMATPROJ}

Úloha 7.2.13. Nájdite maticu ortogonálnej projekcie pri obvyklom skalárnom súčine pre:
a) priestory z úlohy 7.2.1;
b) pre ľubovoľný podpriestor $S = [\vec{\alpha}]$, pričom vektor $\vec{\alpha}$ je normovaný (má jednotkovú dĺžku);
c) pre podpriestor $S = [\vec{\alpha}_1,\ldots,\vec{\alpha}_k]$, pričom vektory $\vec{\alpha}_1,\ldots,\vec{\alpha}_k$ sú ortonormálne.
[Odpovede: b) $\vec{\alpha}^T\vec{\alpha}$; c) A^TA, kde A je matica, ktorej riadky tvoria vektory $\vec{\alpha}_1,\ldots,\vec{\alpha}_k$; toto sa inak dá zapísať aj ako $\vec{\alpha}_1^T\vec{\alpha}_1 + \vec{\alpha}_2^T\vec{\alpha}_2 + \cdots + \vec{\alpha}_k^T\vec{\alpha}_k$.]

Úloha 7.2.14. Ukážte, že pre ľubovoľný podpriestor S euklidovského vektorového priestoru V platí $S^{\perp\perp\perp} = S^\perp$. (Hint: Skúste si uvedomiť, ktorú z inklúzií medzi S a $S^\perp$ sme v dôkaze dôsledku 7.2.9 dokázali bez použitia predpokladu o konečnorozmernosti. Túto inklúziu použite raz pre S a raz pre $S^\perp$.)

Kapitola 8

Kvadratické formy

{kvadform:CHAPTER}

Táto kapitola je spracovaná prevažne na základe [KGGS, Kapitola 9].

8.1 Definícia a základné vlastnosti

{kanon:SECTION}

Definícia 8.1.1. Výraz (polynóm)

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j,$$

kde $a_{ij} \in \mathbb{R}$ a $x_1, \ldots, x_n$ sú (komutujúce) premenné budeme nazývať *kvadratická forma* v premenných $x_1, \ldots, x_n$.

V súvislosti s touto definíciou je užitočné ozrejmiť si zopár faktov.

Poznámka 8.1.2.

1. Podobne by sme mohli definovať kvadratickú formu nad ľubovoľným poľom. Budeme sa však zaoberať iba reálnym prípadom.
2. Slovo polynóm je v definícii uvedené v zátvorke preto, že s polynómami viac premenných sme doteraz nepracovali. Intuitívne by však mohlo byť zrejmé, ako sa takéto polynómy sčitujú a násobia. Pri násobení polynómov viac premenných si treba uvedomiť, že platí $x_ix_j = x_jx_i$ – presne to je myslené tým, že v definícii sa hovorí, že premenné komutujú. Znamená to teda, že dva polynómy uvedeného tvaru sa rovnajú ak pre všetky i platí $a_{ii} = b_{ii}$ a súčasne pre $i \neq j$ platí $a_{ij} + a_{ji} = b_{ij} + b_{ji}$.
3. Vieme, že v prípade polynómov jednej premennej nad poľom $\mathbb{R}$ bola rovnosť polynómov ekvivalentná s rovnosťou polynomických funkcií, ktoré tieto polynómy určujú. Podobne je to aj v tomto prípade – čiže kvadratické formy môžeme chápať ako funkcie n premenných špeciálneho tvaru. (Takéto polynómy viacerých premenných, ktoré majú všetky členy rovnakého stupňa, sa nazývajú *homogénne polynómy.*)

{kanon:PRMATICA}

Príklad 8.1.3. Príkladom kvadratickej formy je $x_1^2 + 2x_1x_2 + 2x_2^2 + 4x_1x_3 + 2x_2x_3 + x_3^2$.

Všimnime si, že ju môžeme zapísať aj pomocou maticového zápisu

$$(x_1, x_2, x_3)\begin{pmatrix}1 & 2 & 4\\ 0 & 2 & 2\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}$$

Všeobecne, ak označíme $\vec{\alpha} = (x_1, \ldots, x_n)$, tak pre ľubovoľnú maticu $A \in M_{n,n}(\mathbb{R})$ je $\vec{\alpha}A\vec{\alpha}^T$ kvadratická forma.

Tú istú kvadratickú formu by sme mohli zapísať aj pomocou iných matíc. Nám sa bude hodiť hlavne reprezentácia pomocou symetrickej matice

$$(x_1, x_2, x_3) \begin{pmatrix} 1 & 1 & 2 \\ 1 & 2 & 1 \\ 2 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

Výhoda reprezentácie pomocou symetrickej matice je v tom, že takáto reprezentácia je už jednoznačná.

{kanon:VTSYMMATICA}

Veta 8.1.4. *Každá kvadratická forma sa dá jednoznačne zapísať ako $\vec{\alpha}B\vec{\alpha}^T$, kde B je symetrická matica.*

Dôkaz. Uvažujme kvadratickú formu $\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j$. Podľa toho, čo sme si povedali o rovnosti kvadratických foriem, matica B vyjadruje tú istú kvadratickú formu práve vtedy, keď platí

$$a_{ij} + a_{ji} = b_{ij} + b_{ji}$$

pre ľubovoľné i a j. Vďaka tomu, že matica b je symetrická je táto rovnosť ekvivalentná s rovnosťami

$$2b_{ij} = a_{ij} + a_{ji}$$
$$b_{ij} = \frac{a_{ij} + a_{ji}}{2}$$

Ľahko vidíme, že matica určená takýmto predpisom je skutočne symetrická ($b_{ij} = b_{ji}$). Súčasne sme ukázali, že toto je jediná možnosť ako voliť koeficienty matice B. □

8.2 Kanonický tvar kvadratickej formy

Pokúsme sa upraviť kvadratickú formu z príkladu 8.1.3 na iný tvar, ktorý môže byť na niektoré účely vhodnejší. Upravíme ju pomocou doplnenia na štvorec.

{kanon:PRIKLKANON1}

Príklad 8.2.1.

$$x_1^2 + 2x_1x_2 + 2x_2^2 + 4x_1x_3 + 2x_2x_3 + x_3^2 = (x_1 + x_2 + 2x_3)^2 + x_2^2 - 2x_2x_3 - 3x_3^2 =$$
$$(x_1 + x_2 + 2x_3)^2 + (x_2 - x_3)^2 - 4x_3^2 = (x_1 + x_2 + 2x_3)^2 + (x_2 - x_3)^2 - (2x_3)^2$$

To znamená, že ak by sme zaviedli nové premenné

$$y_1 = x_1 + x_2 + 2x_3$$
$$y_2 = x_2 - x_3$$
$$y_3 = 2x_3$$

tak pomocou týchto premenných môžeme kvadratickú formu zapísať v jednoduchšom tvare $y_1^2 + y_2^2 - y_3^2$.

Skúsme si ešte rozmyslieť, ako tento fakt môžeme zapísať pomocou maticového zápisu. Ak označíme $\vec{\alpha} = (x_1, x_2, x_3)$ a $\vec{\beta} = (y_1, y_2, y_3)$, tak uvedenú transformáciu premenných môžeme zapísať ako

$$\vec{\beta} = \vec{\alpha}P,$$

kde

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 2 & -1 & 2 \end{pmatrix}.$$

Vieme, že túto kvadratickú formu môžeme zapísať pomocou symetrickej matice $A = \left(\begin{smallmatrix} 1 & 1 & 2 \\ 1 & 2 & 1 \\ 2 & 1 & 1 \end{smallmatrix}\right)$ ako $\vec{\alpha} A \vec{\alpha}^T$.

Ak vyjadríme vektor $\vec{\alpha}$ pomocou vektoru $\vec{\beta}$, t.j. $\vec{\alpha} = \vec{\beta} P^{-1}$, tak dostaneme

$$\vec{\alpha} A \vec{\alpha}^T = \vec{\beta} P^{-1} A (P^{-1})^T \vec{\beta}^T.$$

Zistili sme, že matica kvadratickej formy $y_1^2 + y_2^2 - y_3^2$ sa dá vyjadriť ako $B = P^{-1} A (P^{-1})^T$. Všimnime si, že táto matica je symetrická – platí totiž

$$B^T = (P^{-1} A (P^{-1})^T)^T = P^{-1} A^T (P^{-1})^T.$$

Pretože podľa vety 8.1.4 je symetrická matica prislúchajúca danej kvadratickej formy jednoznačne určená, zistili sme vlastne, že pre maticu $D = \left(\begin{smallmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{smallmatrix}\right)$ platí

$$D = P^{-1} A {P^{-1}}^T$$
$$A = P D P^T$$

Môžeme to (pre tento konkrétny príklad) overiť aj priamym výpočtom:

$$PDP^T = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 2 & -1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 2 \end{pmatrix} =$$
$$= \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 2 & -1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & -2 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 2 & 1 \\ 2 & 1 & 1 \end{pmatrix} = A$$

Definícia 8.2.2. Hovoríme, že matice A a B sú *kongruentné,* ak existuje regulárna matica P taká, že

$$A = P B P^T.$$

Ľahko sa dá overiť, že kongruencia matíc je relácia ekvivalencie (úloha 8.2.1).

Poznámka 8.2.3. Z postupu použitého v predchádzajúcom príklade by mohlo byť vidno, že dve matice sú kongruentné práve vtedy, keď predstavujú tú istú kvadratickú formu, len vyjadrenú v iných premenných. (Pričom vzťah medzi pôvodnými a novými premennými je lineárny.)

Teraz by sme chceli ukázať, že každú kvadratickú formu môžeme pomocou vhodnej transformácie premenných previesť na podobný tvar – taký, ktorý zodpovedá diagonálnej matici majúcej na diagonále iba prvky 0, ± 1. Inak povedané, chceme ukázať, že každá symetrická matica je kongruentná s diagonálnou maticou takéhoto tvaru. Dôkaz bude konštruktívny a bude sa podobať na postup z predchádzajúceho príkladu. Ešte skôr než sa pustíme do dôkazu, vyskúšame si na jednom príklade jediný krok tohoto dôkazu, kde používame iný postup než doplnenie na štvorce.

{kanon:PRIKLX

Príklad 8.2.4. Uvažujme kvadratickú formu x_1x_2. Všimnime si, že

$$x_1x_2 = \frac{(x_1+x_2)^2}{4} - \frac{(x_1-x_2)^2}{4} = \left(\frac{x_1}{2}+\frac{x_2}{2}\right)^2 - \left(\frac{x_1}{2}-\frac{x_2}{2}\right)^2.$$

To znamená, že túto kvadratickú formu vieme previesť na tvar $y_1^2 - y_2^2$ pomocou transformácie

$$y_1 = \frac{x_1+x_2}{2}$$
$$y_2 = \frac{x_1-x_2}{2}$$

Ekvivalentne to môžeme vyjadriť tak, že premenné x_1 a x_2 sme transformovali ako

$$x_1 = y_1 + y_2$$
$$x_2 = y_1 - y_2$$

Môžeme si všimnúť, že pre matice $A = \left(\begin{smallmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{smallmatrix}\right)$, $B = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$, $P = \frac{1}{2}\left(\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix}\right)$ opäť platí $A = PBP^T$. Alebo tiež obrátene, $B = QAQ^T$, kde $Q = \left(\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix}\right)$. (Tieto matice sme dostali z vyjadrenia transformácií premenných podobným spôsobom ako v predchádzajúcom príklade.)

{kanon:VTKANON}

Veta 8.2.5. *Pre ľubovoľnú kvadratickú formu $\sum_{i=1}^n \sum_{j=1}^n a_{ij}x_ix_j$ existuje regulárna transformácia premenných $(y_1,\dots,y_n) = (x_1,\dots,x_n)P$ taká, že táto kvadratická forma sa dá v premenných $y_1,\dots,y_n$ vyjadriť ako*

$$\sum_{i=1}^n \sum_{j=1}^n a_{ij}x_ix_j = \sum_{k=1}^n d_k y_k^2,$$

kde $d_k \in \{0, \pm 1\}$.

Zápis v tvare $\sum_{k=1}^n d_k y_k^2$ budeme nazývať kanonický tvar *kvadratickej formy $\sum_{i=1}^n \sum_{j=1}^n a_{ij}x_ix_j$.*

Pod pojmom *regulárna transformácia premenných* v predchádzajúcej vete rozumieme to, že matica P určujúca túto transformáciu je regulárna.

Dôkaz. Dôkaz je v podstate konštruktívny a budeme v ňom používať postupy, ktoré sme si ukázali v predchádzajúcich príkladoch.

Ukážeme len, že ľubovoľnú kvadratickú formu možno previesť na *diagonálny tvar*

$$c_1y_1^2 + c_2y_2^2 + \cdots + c_ny_n^2.$$

Z tohoto tvaru už kanonický tvar dostaneme ľahko – stačí zaviesť nové premenné $z_i = y_i$ pre tie i, pre ktoré $c_i = 0$ a $z_i = \sqrt{|c_i|}y_i$ pre ostatné i. Takáto transformácia premenných je očividne regulárna.

Budeme postupovať matematickou indukciou vzhľadom na počet premenných n.

1° Ak máme len 1 premennú, tak kvadratická forma $a_{11}x_1^2$ je už v diagonálnom tvare.

2° Predpokladajme, že uvedené tvrdenie platí pre ľubovoľnú kvadratickú formu $n-1$ premenných. Uvažujme kvadratickú formu $\sum_{i=1}^n \sum_{j=1}^n a_{ij}x_ix_j$. Bez ujmy na všeobecnosti môžeme predpokladať, že matica $A = \|a_{ij}\|$ je symetrická.

Predpokladajme najprv, že $a_{11} \neq 0$. Všimnime si všetky členy, ktoré obsahujú premennú x_1 a vhodne ich upravme.

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j = a_{11}x_1^2 + 2a_{12}x_1x_2 + \cdots + 2a_{1n}x_1x_n + \sum_{i=2}^{n}\sum_{j=2}^{n} a_{ij}x_ix_j =$$

$$a_{11}\left(x_1^2 + 2\frac{a_{12}}{a_{11}}x_1x_2 + \cdots + 2\frac{a_{1n}}{a_{11}}x_1x_n\right) + \sum_{i=2}^{n}\sum_{j=2}^{n} a_{ij}x_ix_j =$$

$$a_{11}\left(x_1 + \frac{a_{12}}{a_{11}}x_2 + \cdots + \frac{a_{1n}}{a_{11}}x_n\right)^2 - \sum_{i=2}^{n}\frac{a_{1i}^2}{a_{11}}x_i^2 - \sum_{i=2}^{n}\sum_{j=i+1}^{n}\frac{a_{1i}a_{1j}}{a_{11}}x_ix_j + \sum_{i=2}^{n}\sum_{j=2}^{n} a_{ij}x_ix_j$$

Ak označíme $y_1 = x_1 + \frac{a_{12}}{a_{11}}x_2 + \cdots + \frac{a_{1n}}{a_{11}}x_n$, podarilo sa nám upraviť pôvodnú kvadratickú formu na tvar

$$a_{11}y_1^2 + B(x_2, \ldots, x_n),$$

kde $B(x_2, \ldots, x_n)$ je kvadratická forma v $n-1$ premenných $x_2, \ldots, x_n$.

Podľa indukčného predpokladu sa dá táto kvadratická forma previesť regulárnou transformáciou premenných na diagonálny tvar $c_2y_2^2 + \cdots + c_nx_n^2$. Kombináciou týchto 2 transformácií prevedieme pôvodnú kvadratickú formu na

$$a_{11}y_1^2 + c_2y_2^2 + \cdots + c_nx_n^2.$$

Ak ako P' označíme maticu transformácie pre kvadratickú formu $B(x_2, \ldots, x_n)$, tak matica transformácie pôvodnej kvadratickej formy je

$$\begin{pmatrix} 1 & 0 & \ldots & 0 \\ \frac{a_{21}}{a_{11}} & & & \\ \vdots & & P' & \\ \frac{a_{n1}}{a_{11}} & & & \end{pmatrix}$$

Ak urobíme Laplaceov rozvoj podľa prvého riadku, dostávame $|P| = |P'|$, čiže matica P je tiež regulárna.

Zostáva nám vyriešiť prípad, že $a_{11} = 0$. V prípade, že $a_{ii} \neq 0$ pre nejaké i stačí vymeniť premenné x_1 a x_i (čo je regulárna transformácia) a ďalej postupovať ako v predchádzajúcom prípade.

Ak sú všetky diagonálne prvky matice $\|a_{ij}\|$ nulové, ale existujú nejaké i a j také, že $a_{ij} \neq 0$, použijeme postup z príkladu 8.2.4. Ak totiž dosadíme $x_i = y_i + y_j$ a $x_j = y_i - y_j$, tak dostaneme novú kvadratickú formu, ktorá určite bude obsahovať y_i^2 s nenulový koeficientom. Ďalej môžeme opäť postupovať ako v predchádzajúcom prípade. Použitá transformácia je opäť regulárna.

Zostáva jediný prípad – že všetky čísla a_{ij} sú nulové. Vtedy je už kvadratická forma v kanonickom tvare $0x_1^2 + \cdots + 0x_n^2$. □

Z toho, čo sme si ukázali v príklade 8.2.1 vyplýva, že sme súčasne dokázali nasledujúce tvrdenie o symetrických reálnych maticiach.

Dôsledok 8.2.6. *Každá reálna symetrická matica typu $n \times n$ je kongruentná s nejakou diagonálnou maticou* $\operatorname{diag}(d_1, \ldots, d_n)$ *takou, že $d_i \in \{0, \pm 1\}$ pre $i = 1 \ldots n$.*

Vysvetlíme si ešte (aspoň na konkrétnom príklade) iný postup ako vieme z danej symetrickej matice dostať jej zodpovedajúcu diagonálnu maticu i príslušnú maticu prechodu. Predtým však pripomeňme niečo o tom, ako súvisia riadkové a stĺpcové operácie s násobením matíc (pozri podkapitolu 5.6).

Vykonaním elementárnej riadkovej operácie na matici A dostaneme maticu EA, kde E je matica elementárnej riadkovej operácie – je to taká matica, ktorú dostaneme z jednotkovej matice použitím tejto riadkovej operácie. Napríklad pripočítanie c-násobku prvého riadku k druhému zodpovedá vynásobením maticou $\left(\begin{smallmatrix}1&0&0\\c&1&0\\0&0&1\end{smallmatrix}\right)$ zľava.

Podobne ako riadkové operácie zodpovedajú násobeniu vhodnou maticou zľava, pre stĺpcové operácie treba použiť násobenie sprava. Všimnime si tiež, že ak E je matica riadkovej operácie, pre tú istú stĺpcovú operáciu dostaneme maticu E^T. Vidno to z toho, že ak riadková operácia vytvorila z matice A maticu EA, stĺpcovú operáciu si môžeme predstaviť ako vykonanie riadkovej operácie na matici A^T (a potom opätovné transponovanie), takže dostaneme $(EA^T)^T = AE^T$. Napríklad pripočítanie c-násobku prvého stĺpca k druhému je to isté ako vynásobenie maticou $\left(\begin{smallmatrix}1&c&0\\0&1&0\\0&0&1\end{smallmatrix}\right)$ sprava.

Z toho vidno, že ak by sme začali s maticou A a robili na nej riadkové aj stĺpcové operácie (t.j. po použití riadkovej operácie by sme hneď urobili aj tú istú operáciu na stĺpcoch) dostali by sme maticu

$$E_n \dots E_1 A E_1^T \dots E_n^T = E_n \dots E_1 A (E_n \dots E_1)^T.$$

Ak by sa nám takto podarilo upraviť maticu A na diagonálnu maticu, dostali by sme rovnosť

$$PAP^T = D,$$

kde P označuje $E_n \dots E_1$. Teda použitím riadkových a stĺpcových operácií by sme mohli dostať diagonálnu maticu a aj maticu transformácie premenných.

Príklad 8.2.7. Postup, ktorý sme si práve vysvetlili, si ukážeme na matici kvadratickej formy z príkladu 8.2.1.

Budeme teda robiť striedavo elementárne riadkové a stĺpcové operácie. Súčasne budeme na matici I robiť tie isté riadkové operácie, aby sme dostali maticu, ktorá transformuje A na diagonálnu maticu

$A = \left(\begin{smallmatrix}1&1&2\\1&2&1\\2&1&1\end{smallmatrix}\right) \overset{(1)}{\sim} \left(\begin{smallmatrix}1&1&2\\0&1&-1\\2&1&1\end{smallmatrix}\right) \overset{(1')}{\sim} \left(\begin{smallmatrix}1&0&2\\0&1&-1\\2&-1&1\end{smallmatrix}\right) \overset{(2)}{\sim} \left(\begin{smallmatrix}1&0&2\\0&1&-1\\0&-1&-3\end{smallmatrix}\right) \overset{(2')}{\sim} \left(\begin{smallmatrix}1&0&0\\0&1&-1\\0&-1&-3\end{smallmatrix}\right) \overset{(3)}{\sim} \left(\begin{smallmatrix}1&0&0\\0&1&-1\\0&0&-4\end{smallmatrix}\right) \overset{(3')}{\sim}$
$\left(\begin{smallmatrix}1&0&0\\0&1&0\\0&0&-4\end{smallmatrix}\right) \overset{(4)}{\sim} \left(\begin{smallmatrix}1&0&0\\0&1&0\\0&0&-2\end{smallmatrix}\right) \overset{(4')}{\sim} \left(\begin{smallmatrix}1&0&0\\0&1&0\\0&0&-1\end{smallmatrix}\right) = D$

(1) `2r-=1r` (od druhého riadku odrátam prvý); (1') je rovnaká operácia pre stĺpce (všimnite si, že vždy po vykonaní riadkovej aj stĺpcovej operácie musím dostať symetrickú maticu)
(2) `3r-=2*1r`
(3) `3r+=2r`
(4) `3r*=1/2`

$I = \left(\begin{smallmatrix}1&0&0\\0&1&0\\0&0&1\end{smallmatrix}\right) \overset{(1)}{\sim} \left(\begin{smallmatrix}1&0&0\\-1&1&0\\0&0&1\end{smallmatrix}\right) \overset{(2)}{\sim} \left(\begin{smallmatrix}1&0&0\\-1&1&0\\-2&0&1\end{smallmatrix}\right) \overset{(3)}{\sim} \left(\begin{smallmatrix}1&0&0\\-1&1&0\\-3&1&1\end{smallmatrix}\right) \overset{(4)}{\sim} \left(\begin{smallmatrix}1&0&0\\-1&1&0\\-\frac{3}{2}&\frac{1}{2}&\frac{1}{2}\end{smallmatrix}\right) =: Q$

Priamym výpočtom môžeme overiť, že skutočne platí $D = QAQ^T$. Tiež si môžeme všimnúť, že pre maticu P, ktorú sme dostali v príklade 8.2.1 platí $P = Q^{-1}$.

Na tomto sme videli príklade, že pokiaľ sa v priebehu úprav v tom riadku, ktorý práve upravujeme, na diagonále nevyskytne 0, je postup veľmi podobný na úpravu matice na redukovaný trojuholníkový tvar. Jediný rozdiel bol v tom, že ak sme z prvku c na diagonále chceli dostať ± 1, nedelili sme riadok c-čkom ale iba $\sqrt{|c|}$. V prípade, že by sa vyskytla nula na diagonále, mohli by sme si pomôcť pripočítaním riadku (a stĺpca), ktorý obsahuje nenulový prvok mimo diagonály – ako v nasledujúcom príklade.

Príklad 8.2.8. Pokúsme sa upraviť na kanonický tvar nasledujúcu maticu. V jednotlivých krokoch je urobená vždy riadková aj stĺpcová úprava
$A = \begin{pmatrix} 1&1&0\\1&1&1\\0&1&1 \end{pmatrix} \sim \begin{pmatrix} 1&0&0\\0&0&1\\0&1&1 \end{pmatrix} \sim \begin{pmatrix} 1&0&0\\0&-1&0\\0&0&1 \end{pmatrix} =: D$

V poslednom kroku sme od druhého riadku/stĺpca odrátali tretí riadok/stĺpec. (Na to, aby sme na diagonále dostali nenulový prvok a mohli pokračovať ďalej, stačilo by nám pripočítať ľubovoľný nenulový násobok tretieho riadku/stĺpca. Zhodou okolností sa pri tejto voľbe vynulovali aj zvyšné nediagonálne prvky.)

Použitím rovnakých riadkových úprav na jednotkovú maticu dostaneme
$\begin{pmatrix} 1&0&0\\0&1&0\\0&0&1 \end{pmatrix} \sim \begin{pmatrix} 1&0&0\\-1&1&0\\0&0&1 \end{pmatrix} \sim \begin{pmatrix} 1&0&0\\-1&1&-1\\0&0&1 \end{pmatrix} := P$

Pre túto maticu platí $PAP^T = D$.

Cvičenia

{kanoncvic:ULOKONGRELEKV

Úloha 8.2.1. Overte, že relácia $A \sim B$, t.j. kongruencia symetrických matíc, je relácia ekvivalencie na množine reálnych symetrických matíc typu $n \times n$.

Úloha 8.2.2. Upravte na diagonálny (prípadne kanonický) tvar a nájdite príslušnú transformáciu premenných. Zapíšte aj maticové rovnosti, ktoré z nich vyplývajú:
a) $x_1^2 + 2x_1x_2 + 2x_2^2 + 4x_2x_3 + 5x_3^2$
b) $x_1^2 - 4x_1x_2 + 2x_1x_3 + 4x_2^2 + x_3^2$
c) $x_1x_2 + x_2x_3 + x_3x_1$
d) $x_1^2 - 2x_1x_2 + 2x_1x_3 - 2x_1x_4 + x_2^2 + 2x_2x_3 - 4x_2x_4 + x_3^2 - 2x_4^2$
e) $x_1^2 + x_1x_2 + x_3x_4$

{kanoncvic:ULOHNN}

Úloha 8.2.3*. Preveďte kvadratickú formu $\sum_{i=1}^{n} x_i^2 + \sum_{1\le i<k\le n} x_ix_k$ na diagonálny tvar.

[Výsledok: $y_1^2 + \frac{3}{4}y_2^2 + \frac{4}{6}y_3^2 + \cdots + \frac{n+1}{2n}y_n^2$; $P = \begin{pmatrix} 1&0&\dots&\dots&\dots&0\\ \frac{1}{2}&1&0&\dots&\dots&0\\ \frac{1}{2}&\frac{1}{3}&1&0&\dots&0\\ \dots&\dots&\dots&\dots&\dots&\dots\\ \frac{1}{2}&\frac{1}{3}&\frac{1}{4}&\dots&\frac{1}{n}&1 \end{pmatrix}$]

Úloha 8.2.4*. Preveďte kvadratickú formu $\sum_{1\le i<k\le n} x_ix_k$ na diagonálny tvar.

8.3 Zákon zotrvačnosti

{zotrv:SECTZOTV}

Keďže sme tvar kvadratickej formy z vety 8.2.5 nazvali kanonický, dá sa očakávať, že bude v nejakom zmysle jednoznačný. Cieľom tejto kapitoly je práve sformulovať a dokázať túto jednoznačnosť.

Veta 8.3.1. *Pre danú kvadratickú formu je počet výskytov $+1$ a počet výskytov -1 v jej kanonickom tvare jednoznačne určený (nezávisí od transformácie, ktorou sme túto kvadratickú formu previedli na kanonický tvar).*

Dôkaz. Uvažujme kvadratickú formu $\vec{\alpha}A\vec{\alpha}^T$. Predpokladajme, že sa dá regulárnou transformáciou previesť na tvar

$$y_1^2 + \cdots + y_k^2 - y_{k+1}^2 - \cdots - y_s^2 \tag{8.1}$$

{zotrv:EQ1}

a súčasne (inou regulárnou transformáciou) na tvar

$$z_1^2 + \cdots + z_r^2 - z_{r+1}^2 - \cdots - z_t^2. \tag{8.2}$$

{zotrv:EQ2}

Označme ako P_1 a P_2 regulárne matice, ktoré zodpovedajú tejto transformácii premenných, t.j. $(y_1,\dots,y_n)=(x_1,\dots,x_n)P_1$ a $(z_1,\dots,z_n)=(x_1,\dots,x_n)P_2$.

Chceme ukázať, že $s=t$ a $k=r$.

Všimnime si, že s je presne hodnosť diagonálnej matice zodpovedajúcej kvadratickej forme (8.1) a t je hodnosť matice pre kvadratickú formu (8.2). Tieto matice môžeme vyjadriť ako $D_1=P_1^{-1}A(P_1^{-1})^T$ a $D_2=P_2^{-1}A(P_2^{-1})^T$. Súčasne vieme, že násobenie regulárnou maticou zodpovedá lineárnemu izomorfizmu, takže nemení hodnosť matice. Teda s aj t sa musí rovnať hodnosti matice A.

Zostáva nám dokázať, že $k=r$. Máme rovnosť

$$z_1^2+\dots+z_r^2-z_{r+1}^2-\dots-z_t^2=y_1^2+\dots+y_k^2-y_{k+1}^2-\dots-y_s^2,$$

ktorú môžeme upraviť na tvar

{zotrv:EQ4}
$$z_1^2+\dots+z_r^2+y_{k+1}^2+\dots+y_s^2=y_1^2+\dots+y_k^2+z_{r+1}^2+\dots+z_t^2. \tag{8.3}$$

Predpokladajme najprv, že $r<k$. Ukážeme, že za tohoto predpokladu sa dá nájsť nenulový vektor $(x_1,\dots,x_n)$ tak, aby platilo

{zotrv:EQ3}
$$z_1=\dots=z_r=y_{k+1}=\dots=y_n=0. \tag{8.4}$$

Všimnime si, že $(z_1,\dots,z_r,y_{k+1}\dots y_n)=(x_1,\dots,x_n)P$, kde matica P pozostáva z prvých r stĺpcov matice P_2 a z $(k+1)$-vého až n-tého stĺpca matice P_1. Hľadanie vektora $(x_1,\dots,x_n)$, ktorý vyhovuje rovnici (8.4) teda zodpovedá riešeniu homogénnej sústavy

$$\vec{\alpha}P=\vec{0},$$
$$P^T\vec{\alpha}^T=\vec{0}^T.$$

Keďže táto sústava má menej rovníc než neznámych, existuje aspoň jedno nenulové riešenie.

Ak však nájdeme $x_1,\dots,x_n$ také, že platí (8.4), na základe (8.3) musia byť nulové aj všetky ostatné premenné $z_{r+1},\dots,z_t$. Lenže potom máme

$$(x_1,\dots,x_n)=(z_1,\dots,z_n)P_2^{-1}=\vec{0},$$

čo je spor s tým, že vektor $(x_1,\dots,x_n)$ je nenulový.

Podobne aj predpoklad $r>k$ by viedol k sporu. Musí teda platiť $k=r$. □

Predchádzajúca veta nám teda vlastne hovorí, že kanonický tvar kvadratickej formy je až na výmenu premenných jednoznačne určený.

Niekedy nás zaujímajú kvadratické formy, ktorých kanonický tvar má na diagonále iba jednotky.

Definícia 8.3.2. Nech A je symetrická reálna matica. Hovoríme, že A je
a) *kladne semidefinitná*, ak pre každý vektor $\vec{\alpha}$ platí $\vec{\alpha}A\vec{\alpha}^T\geq 0$;
b) *kladne definitná*, ak pre každý nenulový vektor $\vec{\alpha}\neq\vec{0}$ platí $\vec{\alpha}A\vec{\alpha}^T>0$;
c) *záporne semidefinitná*, ak pre každý vektor $\vec{\alpha}$ platí $\vec{\alpha}A\vec{\alpha}^T\leq 0$;
d) *záporne definitná*, ak pre každý nenulový vektor $\vec{\alpha}\neq\vec{0}$ platí $\vec{\alpha}A\vec{\alpha}^T<0$.

Je ľahké si všimnúť, že A je kladne (semi)definitná práve vtedy, keď $-A$ je záporne (semi)definitná.

Nasledujúca veta charakterizuje kladne definitné matice. Tým súčasne charakterizuje symetrické matice, ktoré určujú skalárne súčiny na $\mathbb{R}^n$ (pozri príklad 7.1.5).

Veta 8.3.3. *Symetrická matica A je kladne definitná práve vtedy, keď existuje regulárna matica P taká, že $A = PP^T$.*

Chceli by sme dokázať kritérium, pomocou ktorého sa dá pomerne jednoducho určiť, či je daná matica kladne definitná. V dôkaze tohoto kritéria bude užitočné nasledujúce pomocné tvrdenie:

{zotrv:TVRMINORY}

Tvrdenie 8.3.4. *Nech A je symetrická reálna matica typu $n \times n$ taká, že všetky rohové determinanty*

$$D_1 = |a_{11}|$$
$$D_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$
$$\vdots$$
$$D_n = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$$

sú nenulové.

Potom matica A je kongruentná s diagonálnou maticou $\operatorname{diag}(D_1, D_2/D_1, D_3/D_2, \dots, D_n/D_{n-1})$.

Determinanty podmatíc vystupujúce v predchádzajúcom tvrdení sa niekedy zvyknú nazývať aj *hlavné minory* matice A.

Dôkaz. Ukážeme, že danú maticu možno upraviť na diagonálnu maticu len použitím operácií typu „pripočítanie násobku riadka/stĺpca k inému“ (t.j. nepoužívame výmeny riadkov a ani násobenie riadkov konštantou) tak, že dostaneme práve diagonálnu maticu tvaru, ktorý je uvedený v tvrdení.

Tvrdenie dokážeme matematickou indukciou vzhľadom na rozmer matice n. Platnosť tvrdenia pre $n = 1$ je zrejmá.

Vieme, že $D_1 = a_{11}$. Vďaka tomu môžeme vynulovať všetky prvky v prvom riadku a prvom stĺpci. (Odčítaním vhodného násobku prvého stĺpca/riadku.) Dostaneme tak maticu A do tvaru

$$\begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & b_{22} & \dots & b_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & b_{n2} & \dots & b_{nn} \end{pmatrix}$$

Pritom vieme, že úpravy, ktoré sme použili, nemenia determinant matice A ani žiaden z jej hlavných minorov (veta 6.3.9).

Označme $D'_2, \dots, D'_n$ minory podmatice, ktorá vznikne vynechaním prvého riadku a prvého stĺpca. Z tvaru matice vidíme, že pre $k = 2, \dots, n$ platí $a_{11} D'_k = D_k$ a teda

$$D'_k = \frac{D_k}{D_1}.$$

Podľa indukčného predpokladu vieme túto podmaticu upraviť na diagonálny tvar, kde ako prvý člen na diagonále bude $D'_2 = \frac{D_2}{D_1}$ a ďalšie členy budú tvaru $\frac{D'_{k+1}}{D'_k} = \frac{D_{k+1}}{D_k}$. Z toho dostávame diagonálny tvar

$$\operatorname{diag}(D_1, D_2/D_1, D_3/D_2, \dots, D_n/D_{n-1})$$

pre pôvodnú maticu. □

Veta 8.3.5. *Nech A je symetrická matica typu $n \times n$. Matica A je kladne definitná práve vtedy, keď všetky jej hlavné minory $D_1, \dots, D_n$ sú kladné.*

Dôkaz. $\boxed{\Rightarrow}$ Ak je matica kladne definitná, tak je kongruentná s jednotkovou maticou. Z toho máme $A = PP^T$ a

$$|A| = |PP^T| = |P||P^T| = |P|^2 > 0.$$

Podobné tvrdenie pre minory vyplynie z toho, že ak za premenné $x_{k+1}, \dots, x_n$ dosadíme 0, dostaneme tak kvadratickú formu v premenných $x_1, \dots, x_k$, ktorá je opäť kladne definitná a ktorej matica je presne podmatica určená prvými k riadkami a stĺpcami.

$\boxed{\Leftarrow}$ Ak všetky hlavné minory matice A sú kladné, tak podľa tvrdenia 8.3.4 možno príslušnú kvadratickú formu upraviť na diagonálny tvar, v ktorom sú všetky členy kladné. Preto je táto matica kladne definitná. □

Dôsledok 8.3.6. *Nech A je symetrická matica typu $n \times n$. Matica A je záporne definitná práve vtedy, keď hlavný minor D_k má rovnaké znamienko ako $(-1)^k$ pre všetky $k = 1, \dots, n$. (Teda znamienka hlavných minorov sú striedavo $(-, +, -, +, \dots)$.)*

Poznámka 8.3.7. Aby sme dostali kritérium pre kladne semidefinitné matice, nestačí v predchádzajúcej vete zmeniť slovo „kladné“ na „nezáporné“. (Ako jednoduchý kontrapríklad môžeme zobrať maticu $\left(\begin{smallmatrix} 0 & 0 \\ 0 & -1 \end{smallmatrix}\right)$.) V podobnom kritériu pre kladne semidefinitné matice vystupujú všetky minory matice A (=všetky determinanty štvorcových podmatíc).

Možnosť overiť, či je nejaká matica kladne alebo záporne definitná, má význam v matematickej analýze pre hľadaní extrémov funkcií viacerých premenných. Nutná podmienka na to, aby v nejakom bode x_0 mala nejaká funkcia lokálny extrém je, aby všetky parciálne derivácie boli nulové. (Podobne ako v jednorozmere bola nutná podmienka $f'(x_0) = 0$.)

$$\frac{\partial f}{\partial x_1}(x_0) = \dots = \frac{\partial f}{\partial x_n}(x_0) = 0.$$

V prípade, že je v jednorozmere splnená táto podmienka, skúmame ďalej to, či je kladná alebo záporná v danom bode jej druhá derivácia. Vo viacrozmere funkciu druhej derivácie hrá matice

$$H = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(x_0) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(x_0) & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(x_0) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(x_0) & \frac{\partial^2 f}{\partial x_2^2}(x_0) & \dots & \frac{\partial^2 f}{\partial x_2 \partial x_n}(x_0) \\ \dots & \dots & \dots & \dots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(x_0) & \frac{\partial^2 f}{\partial x_n \partial x_2}(x_0) & \dots & \frac{\partial^2 f}{\partial x_n^2}(x_0) \end{pmatrix}$$

Táto matica je symetrická pre každú funkciu, ktorá je dvakrát spojito diferencovateľná.

Z viacrozmernej verzie Taylorovej vety totiž vyplýva, že hodnotu funkcie v bode x_0 môžeme aproximovať ako

$$f(x) - f(x_0) = \frac{1}{2!}(x - x_0)H(x - x_0)^T$$

(kde $x - x_0$ sú body z $\mathbb{R}^n$, teda ich chápeme ako vektory.)

To znamená, že hodnota $f(x) - f(x_0)$ je aproximovaná (v nejakom okolí bodu x_0) kvadratickou formou s maticou H. Z toho vyplýva, že v bode $f(x_0)$ je lokálne minimum práve vtedy, keď táto matica je kladne definitná ($f(x) - f(x_0)$ je v nejakom okolí kladné), lokálne maximum ak je záporne definitná. (V prípade, že je kladne definitná, vieme dokonca povedať, že vo vhodných súradniciach sa táto funkcia lokálne podobná na funkciu $x_1^2 + \dots + x_n^2$, ktorú si vieme aspoň v dvojrozmere celkom dobre geometricky predstaviť. Inou podobnou funkciou vieme zasa aproximovať tie funkcie, ktoré majú maticu H záporne definitnú.)

Viac o tejto problematike sa môžete dozvedieť napríklad v [GĎ, Kapitola 9.4], [Ap], [P] (a v podstate v každej učebnici, ktorá sa zaoberá analýzou viac premenných).

Príklad 8.3.8. V dôkaze tvrdenia 8.3.4 sme videli, ako sa (za predpokladu, že daná symetrická matica má nenulové hlavné minory) dajú nájsť koeficienty v diagonálnom tvare, do ktorého túto maticu vieme previesť iba pomocou operácie pripočítavania niektorého násobku riadku/stĺpca k inému. Kombinácia takýchto operácií znamená to, že matica transformácie bude mať na diagonále jednotky.

Vyskúšajme si to na kvadratickej forme z úlohy 8.2.3*. Zodpovedajúca symetrická matica je

$$\begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} & \dots & \frac{1}{2} \\ \frac{1}{2} & 1 & \frac{1}{2} & \dots & \frac{1}{2} \\ \dots & \dots & \dots & \dots & \dots \\ \frac{1}{2} & \frac{1}{2} & \dots & 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \dots & \frac{1}{2} & 1 \end{pmatrix}$$

Ak počítame jej hlavné minory dostávame $D_1 = 1$ a pre $k > 1$

$$D_k = \begin{vmatrix} 1 & \frac{1}{2} & \frac{1}{2} & \dots & \frac{1}{2} \\ \frac{1}{2} & 1 & \frac{1}{2} & \dots & \frac{1}{2} \\ \dots & \dots & \dots & \dots & \dots \\ \frac{1}{2} & \frac{1}{2} & \dots & 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \dots & \frac{1}{2} & 1 \end{vmatrix} = \begin{vmatrix} \frac{k+1}{2} & \frac{k+1}{2} & \frac{k+1}{2} & \dots & \frac{k+1}{2} \\ \frac{1}{2} & 1 & \frac{1}{2} & \dots & \frac{1}{2} \\ \dots & \dots & \dots & \dots & \dots \\ \frac{1}{2} & \frac{1}{2} & \dots & 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \dots & \frac{1}{2} & 1 \end{vmatrix} = (k+1) \begin{vmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \dots & \frac{1}{2} \\ \frac{1}{2} & 1 & \frac{1}{2} & \dots & \frac{1}{2} \\ \dots & \dots & \dots & \dots & \dots \\ \frac{1}{2} & \frac{1}{2} & \dots & 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \dots & \frac{1}{2} & 1 \end{vmatrix} = (k+1) \begin{vmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \dots & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \frac{1}{2} & 0 \\ 0 & 0 & \dots & 0 & \frac{1}{2} \end{vmatrix} = \frac{k+1}{2^k}$$

(V prvom kroku sme k prvému riadku pripočítali všetky ostatné, v poslednom kroku sme odrátali prvý riadok od všetkých ostatných. Takéto operácie nemenia hodnotu determinantu.)

Pre koeficienty na diagonále potom dostávame

$$c_k = \frac{D_k}{D_{k-1}} = \frac{k+1}{2^k} \frac{2^{k-1}}{k} = \frac{k+1}{2k},$$

čiže rovnaký výsledok ako nám vyšiel v úlohe 8.2.3*.

Cvičenia

Úloha 8.3.1. Pre danú kvadratickú formu určte tie hodnoty parametra $t \in \mathbb{R}$, pre ktoré je kladne definitná.
a) $5x_1^2 + 3x_2^2 + tx_3^2 + 4x_1x_2 - 3x_1x_3 - 2x_2x_3$
b) $2x_1^2 + x_2^2 + 3x_3^2 + 2tx_1x_2 + 2x_1x_3$
c) $\frac{1}{2}x_1^2 + 2x_2^2 - 3tx_3^2 + 2x_1x_2 + 2tx_2x_3 + 2x_1x_3$
d) $(x_1^2 + x_2^2 + x_3^2 + 2x_1x_3 - 2x_2x_3) + t(6x_1x_2 - 2x_1x_3 - 2x_1^2) + t^2(x_1^2 + x_2^2)$
(Poznámka: Niekedy sa výpočet determinantov $D_1, D_2, \dots$ môže zjednodušiť, ak zmeníte poradie premenných. Takáto zmena neovplyvní to, či je matica kladne definitná.)

Úloha 8.3.2. Z údajov ktoré sú zadané o symetrickej matici A zistite, ako vyzerá kanonický tvar príslušnej kvadratickej formy.[1] (Dali by sa tieto úvahy použiť na zistenie kanonického tvaru pre niektoré kvadratické formy z predošlých príkladov?)
a) Matica A je *kladne definitná* symetrická matica rozmerov $n \times n$.
b) Matica A je *záporne definitná* symetrická matica rozmerov $n \times n$.
c*) A je nenulová symetrická matica rozmerov 3×3, ktorá má nulovú stopu aj determinant, t.j. $\det(A) = \mathrm{Tr}(A) = 0$.

Úloha 8.3.3. Nech A je symetrická reálna matica taká, že $D_1 > 0, D_2 > 0, \dots, D_n > 0$. (Determinanty D_k majú rovnaký význam ako v tvrdení 8.3.4.) Dokážte, že potom $a_{nn} > 0$.

[1] V časti c) treba použiť nejaké veci, ktoré budeme preberať v ďalších kapitolách – sem som úlohu zaradil iba preto, že sa podobá na ostatné časti a že súvisí s kanonickým tvarom kvadratickej formy.

Úloha 8.3.4. Nech V je euklidovský vektorový priestor a $\vec{\alpha}_1, \ldots, \vec{\alpha}_n \in V$. Definujme maticu $A = \|a_{ij}\|$ tak, že $a_{ij} = \langle \vec{\alpha_i}, \vec{\alpha_j} \rangle$. (Táto matica sa zvykne volať *Gramova matica.*) Dokážte, že $|A| \geq 0$ a že tieto vektory sú lineárne nezávislé práve vtedy, keď $|A| > 0$.

Úloha 8.3.5*. Pre kvadratické formy $f = \sum_{i,j} a_{ij} x_i x_j$ a $g = \sum_{i,j} b_{ij} x_i x_j$ definujeme kvadratickú formu $(f, g) = \sum_{i,j} a_{ij} b_{ij} x_i x_j$. Ukážte, že ak f a g sú kladne definitné, tak aj (f, g) je kladne definitná.

Kapitola 9

Podobnosť matíc

{podob:CHAPTER}

Úvodná časť tejto kapitoly je spracovaná na základe [KGGS, 9.4,9.5] a .... Podkapitola 9.2 obsahuje poznámky spracované J. Guričanom k tejto téme.[1]

V kapitole 8 sme sa zaoberali kvadratickými formami a ukázali sme, že pri vhodnej zmene premenných vieme kvadratickú formu upraviť na veľmi jednoduchý a pekný tvar (diagonálny, prípadne kanonický). Súčasne nám vzťah medzi týmito kvadratickými formami povedal niečo o vzťahoch medzi ich maticami.

V tejto kapitole sa budeme zaoberať do istej miery podobným problémom. Tentokrát sa však budeme snažiť pomocou zmeny premenných nájsť čo najkrajší tvar matice lineárnej transformácie.

9.1 Matica prechodu, podobnosť matíc

Pripomeňme, že ak $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ je nejaká báza konečnorozmerného vektorového priestoru V, tak ľubovoľný vektor z V sa dá jednoznačne vyjadriť v tvare $\vec{\gamma} = c_1\vec{\alpha}_1 + \cdots + c_n\vec{\alpha}_n$. (Pozri vetu 4.4.16.) Tento fakt nám vlastne hovorí, že báza nám poskytuje akúsi súradnicovú sústavu v priestore V – každý vektor má jednoznačne určené súradnice $c_1, \ldots, c_n$.

Definícia 9.1.1. Ak $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ je báza vektorového priestoru V nad poľom F a $\vec{\gamma} \in V$, tak n-ticu $(c_1, \ldots, c_n) \in F^n$ takú, že platí

$$\vec{\gamma} = c_1\vec{\alpha}_1 + \cdots + c_n\vec{\alpha}_n$$

nazývame *súradnicami vektora $\vec{\gamma}$ v báze $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$*.

Jednou z otázok, ktorými sa budeme v tejto podkapitole zaoberať, je to, ako sa zmenia súradnice vektora pri zmene bázy daného vektorového priestoru. Pri tom bude užitočná matica uvedená v nasledujúcej definícii, ktorá popisuje istým spôsobom vzťah medzi týmito dvoma bázami.

Definícia 9.1.2. Nech $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ a $\vec{\alpha}'_1, \ldots, \vec{\alpha}'_n$ sú dve bázy vektorového priestoru V nad

[1] http://modx.gurican.sk/assets/files/algebra_i_3/podobnost.pdf

poľom F. Nech $p_{ij} \in F$ sú také, že platí

$$\begin{aligned}
\vec{\alpha}'_1 =& p_{11}\vec{\alpha}_1 + p_{12}\vec{\alpha}_2 + \cdots + p_{1n}\vec{\alpha}_n \\
\vec{\alpha}'_2 =& p_{21}\vec{\alpha}_1 + p_{22}\vec{\alpha}_2 + \cdots + p_{2n}\vec{\alpha}_n \\
&\vdots \\
\vec{\alpha}'_n =& p_{n1}\vec{\alpha}_1 + p_{n2}\vec{\alpha}_2 + \cdots + p_{nn}\vec{\alpha}_n
\end{aligned} \tag{9.1}$$ {prechod:EQPRI

Potom maticu $P = \|p_{ij}\|$ nazývame *matica prechodu* od bázy $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ k báze $\vec{\alpha}'_1, \ldots, \vec{\alpha}'_n$.

Inak povedané, matica prechodu je taká matica P, ktorej i-ty riadok je tvorený súradnicami vektoru $\vec{\alpha}'_i$ v báze $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$.

Poznámka 9.1.3. V literatúre nájdete často i presne opačnú definíciu, než sme uviedli my. Teda niektorí autori by túto maticu nazvali maticou prechodu od $\vec{\alpha}'_1, \ldots, \vec{\alpha}'_n$ k $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$.

Skúsme si rozmyslieť, čo vieme povedať o matici prechodu opačným smerom, t.j. od $\vec{\alpha}'_1, \ldots, \vec{\alpha}'_n$ k $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$. Označme túto maticu $P' = \|p_{ij}\|$. Platí:

$$\begin{aligned}
\vec{\alpha}_i =& p'_{i1}\vec{\alpha}'_1 + p'_{i2}\vec{\alpha}'_2 + \cdots + p'_{in}\vec{\alpha}'_n = \\
=& p'_{i1}(p_{11}\vec{\alpha}_1 + p_{12}\vec{\alpha}_2 + \cdots + p_{1n}\vec{\alpha}_n) + \\
&+ p'_{i2}(p_{21}\vec{\alpha}_1 + p_{22}\vec{\alpha}_2 + \cdots + p_{2n}\vec{\alpha}_n) + \\
&\vdots \\
&+ p'_{in}(p_{n1}\vec{\alpha}_1 + p_{n2}\vec{\alpha}_2 + \cdots + p_{nn}\vec{\alpha}_n)
\end{aligned}$$

(Vektory $\vec{\alpha}'_1, \ldots, \vec{\alpha}'_n$ sme upravili pomocou (9.1).) Túto rovnosť teraz upravíme tak, že dáme dokopy členy obsahujúci ten istý vektor $\vec{\alpha}_i$ – inak povedané tak, ako sme ju zapísali pred chvíľou to znamená, že sčítance teraz usporiadame po stĺpcoch.

$$\begin{aligned}
\vec{\alpha}_i =& (p'_{i1}p_{11} + p'_{i2}p_{21} + \cdots + p'_{in}p_{n1})\vec{\alpha}_1 + \\
&+ (p'_{i1}p_{12} + p'_{i2}p_{22} + \cdots + p'_{in}p_{n2})\vec{\alpha}_2 + \\
&\vdots \\
&+ (p'_{i1}p_{1n} + p'_{i2}p_{2n} + \cdots + p'_{in}p_{nn})\vec{\alpha}_n
\end{aligned}$$

Obe predchádzajúce úpravy sme mohli stručnejšie zapísať takto:[2]

$$\vec{\alpha}_i = \sum_{j=1}^{n} p'_{ij}\vec{\alpha}'_j = \sum_{j=1}^{n}\sum_{k=1}^{n} p'_{ij}p_{jk}\vec{\alpha}_k = \sum_{k=1}^{n}\left(\sum_{j=1}^{n} p'_{ij}p_{jk}\right)\vec{\alpha}_k.$$

Z jednoznačnosti vyjadrenia vektora v báze $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ vyplýva, že koeficienty na pravej strane rovnosti sa musia rovnať

$$\sum_{j=1}^{n} p'_{ij}p_{jk} = \delta_{ik} = \begin{cases} 1, & \text{ak } i = k \\ 0, & \text{inak.} \end{cases}$$

[2] Aj v ďalšom budeme používať tento stručnejší zápis pomocou súm, chcel som však aspoň pri prvom použití celú úpravu rozpísať trochu podrobnejšie tak, aby súčasne bolo vidno, že sa tam skutočne násobil i-ty riadok matice P' s jednotlivými stĺpcami a aby sme si uvedomili, že na výmenu poradia sumácie sa dá v takýchto prípadoch pozerať tak, že namiesto toho, aby sme rovnosť prečítali po riadkoch, si ju prečítame po stĺpcoch. V prípade, že by výmena poradia sčitovania v niektorom z ďalších dôkazov robila problémy, môže pomôcť prepísať si ju tak ako tu.

Zistili sme teda, že platí $P'P = I$, čo znamená, že P' je inverzná matica k P (pozri poznámku 5.5.8).

Ukážme si, ako sme celé predchádzajúce odvodenie mohli stručnejšie odvodiť pomocou maticového zápisu.

V prvom rade si uvedomme, že vzťah (9.1) sa dá ekvivalentne zapísať takto:

$$\begin{pmatrix} \vec{\alpha}'_1 \\ \vdots \\ \vec{\alpha}'_n \end{pmatrix} = P \begin{pmatrix} \vec{\alpha}_1 \\ \vdots \\ \vec{\alpha}_n \end{pmatrix}. \tag{9.2}$$ {prechod:EQMATICOVAPRECH

Keďže vektory $\vec{\alpha}'_1, \ldots, \vec{\alpha}'_n$ aj $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ tvoria bázu, matice $\begin{pmatrix} \vec{\alpha}'_1 \\ \vdots \\ \vec{\alpha}'_n \end{pmatrix}$ a $\begin{pmatrix} \vec{\alpha}_1 \\ \vdots \\ \vec{\alpha}_n \end{pmatrix}$ majú hodnosť n, čiže sú regulárne. Ak pomocou nich vyjadríme maticu P, zistíme, že aj táto matica je regulárna (súčin dvoch regulárnych), čiže k nej existuje inverzná.

Hneď vidíme, že ak platí rovnosť (9.2), tak platí i

$$P^{-1} \begin{pmatrix} \vec{\alpha}'_1 \\ \vdots \\ \vec{\alpha}'_n \end{pmatrix} = \begin{pmatrix} \vec{\alpha}_1 \\ \vdots \\ \vec{\alpha}_n \end{pmatrix}.$$

Posledná rovnosť znamená presne to, že P^{-1} je matica prechodu od $\vec{\alpha}'_1, \ldots, \vec{\alpha}'_n$ k $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$.

Mohli by sme použiť aj postup, ktorý by úplne presne kopíroval predchádzajúce odvodenie, pričom by sme dostali

$$\begin{pmatrix} \vec{\alpha}_1 \\ \vdots \\ \vec{\alpha}_n \end{pmatrix} = P'P \begin{pmatrix} \vec{\alpha}_1 \\ \vdots \\ \vec{\alpha}_n \end{pmatrix}$$

z čoho (na základe regularity matice $\begin{pmatrix} \vec{\alpha}_1 \\ \vdots \\ \vec{\alpha}_n \end{pmatrix}$) už vyplýva $P'P = I$.

Poznámka 9.1.4. Predchádzajúce odvodenie v skutočnosti nebolo úplne korektné. Z vektorov $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ môžeme vytvoriť maticu typu $n \times n$ len vtedy, ak ide o vektory vo vektorovom priestore $V = F^n$. Toto však môžeme pomerne ľahko opraviť – stačí si uvedomiť, že každý n-rozmerný priestor je izomorfný s F^n (veta 5.5.15). Ak si pevne zvolíme nejaký izomorfizmus medzi V a F^n, môžeme potom už všetky úvahy robiť v F^n. Dôležité je uvedomiť si, že izomorfizmus neovplyvní veci ako dimenzia, lineárna kombinácia, lineárna nezávislosť, súradnice vektora v danej báze a pod. (Napríklad ak vektor $\vec{\gamma} \in V$ má v báze $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ súradnice $(c_1, \ldots, c_n)$ a $f\colon V \to F^n$ je ľubovoľný izomorfizmus, tak aj súradnice vektora $f(\vec{\gamma})$ v báze $f(\vec{\alpha}_1), \ldots, f(\vec{\alpha}_n)$ sú $(c_1, \ldots, c_n)$. Z tejto skutočnosti ďalej vyplýva, že sa zachová aj matica prechodu medzi dvoma bázami.)

V ďalších úvahách budeme niekedy používať podobné argumenty – bez toho, že by sme zdôraznili prechod do F^n. Čitateľ si môže na príslušných rozmyslieť, že to skutočne funguje. Budeme však vždy uvádzať aj odvodenie, ktoré sa neopiera o maticový zápis, a teda pri ňom takýto prechod nie je potrebný.

Dokázali sme nasledujúcu vetu:

{prechod:VTREGULARNA}

Veta 9.1.5. *Ak P je matica prechodu od bázy $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ k báze $\vec{\alpha}'_1, \ldots, \vec{\alpha}'_n$, tak matica P je regulárna a matica P^{-1} je matica prechodu opačným smerom, teda od $\vec{\alpha}'_1, \ldots, \vec{\alpha}'_n$ k $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$.*

Ukážeme, že to funguje aj naopak – pre každú bázu a regulárnu maticu P použitím predpisu (9.1) dostaneme opäť bázu.

{prechod:TVRBU

Tvrdenie 9.1.6. *Nech $P = \|p_{ij}\|$ je regulárna matica typu $n \times n$ a $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ je báza vektorového priestoru V. Potom aj vektory $\vec{\alpha}'_1, \ldots, \vec{\alpha}'_n$ určené vzťahmi*

$$\begin{aligned}\vec{\alpha}'_1 &= p_{11}\vec{\alpha}_1 + p_{12}\vec{\alpha}_2 + \cdots + p_{1n}\vec{\alpha}_n\\ \vec{\alpha}'_2 &= p_{21}\vec{\alpha}_1 + p_{22}\vec{\alpha}_2 + \cdots + p_{2n}\vec{\alpha}_n\\ &\vdots\\ \vec{\alpha}'_n &= p_{n1}\vec{\alpha}_1 + p_{n2}\vec{\alpha}_2 + \cdots + p_{nn}\vec{\alpha}_n\end{aligned}$$

tvoria bázu priestoru V.

Dôkaz. Podľa vety 4.4.14 nám stačí ukázať, že tieto vektory sú lineárne nezávislé. Nech teda $c_1\vec{\alpha}_1 + \cdots + c_n\vec{\alpha}_n = \vec{0}$. Ak do tejto rovnosti dosadíme vyjadrenia vektorov $\vec{\alpha}'_1, \ldots, \vec{\alpha}'_n$ v báze $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ dostaneme

$$\vec{0} = \sum_{i=1}^{n} c_i \left(\sum_{j=1}^{n} p_{ij}\vec{\alpha}_j \right) = \sum_{j=1}^{n} \left(\sum_{i=1}^{n} c_i p_{ij} \right) \vec{\alpha}_j.$$

Všimnime si, že koeficienty pri vektoroch $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ na pravej strane predchádzajúcej rovnosti sú presne zložky vektora $(c_1, \ldots, c_n)P$. Pretože vektory $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ sú lineárne nezávislé, aby platila táto rovnosť, musia sa všetky tieto koeficienty rovnať nule. Dostali sme teda rovnosti

$$\begin{aligned}(c_1, \ldots, c_n)P &= \vec{0}\\ (c_1, \ldots, c_n) &= \vec{0}P^{-1} = \vec{0}\\ c_1 &= \cdots = c_n = 0\end{aligned}$$

Tým sme ukázali, že $\vec{\alpha}'_1, \ldots, \vec{\alpha}'_n$ sú lineárne nezávislé. □

Opäť môžeme to isté tvrdenie dokázať aj s využitím (9.2).

Dôkaz. Keďže $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ tvoria bázu, tak hodnosť matice $\begin{pmatrix}\vec{\alpha}_1\\ \vdots\\ \vec{\alpha}_n\end{pmatrix}$ je n. Pretože násobenie regulárnou maticou nemení hodnosť, tak aj hodnosť matice

$$\begin{pmatrix}\vec{\alpha}'_1\\ \vdots\\ \vec{\alpha}'_n\end{pmatrix} = P \begin{pmatrix}\vec{\alpha}_1\\ \vdots\\ \vec{\alpha}_n\end{pmatrix}$$

je n, čo znamená, že riadky tejto matice tvoria bázu vektorového priestoru V (keďže jeho dimenzia je n; aj tu využívame vetu 4.4.14). □

Zmena súradníc vektora pri zmene bázy

Ukážeme si ako pomocou matice prechodu môžeme dostať vzťah medzi vyjadrením súradníc daného vektora v dvoch rôznych bázach.

{prechod:VTSURVEKTOR}

Veta 9.1.7. *Nech $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ a $\vec{\alpha}'_1, \ldots, \vec{\alpha}'_n$ sú bázy vektorového priestoru V. Nech P je matica prechodu od bázy $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ k báze $\vec{\alpha}'_1, \ldots, \vec{\alpha}'_n$. Nech $\vec{\gamma} \in V$ a $(x_1, \ldots, x_n)$ sú súradnice vektora $\vec{\gamma}$ v báze $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$, $(x'_1, \ldots, x'_n)$ sú jeho súradnice v báze $\vec{\alpha}'_1, \ldots, \vec{\alpha}'_n$. Potom platí*

$$(x_1, \ldots, x_n) = (x'_1, \ldots, x'_n)P.$$

Postup, ktorý použijeme v dôkaze je v podstate rovnaký, ako úpravy použité v dôkaze predchádzajúceho tvrdenia.

Dôkaz. To, že vektor $\vec{\gamma}$ má v báze $\vec{\alpha}'_1, \ldots, \vec{\alpha}'_n$ súradnice $(x'_1, \ldots, x'_n)$ znamená, že platí rovnosť

$$\vec{\gamma} = x'_1\vec{\alpha}'_1 + \cdots + x'_n\vec{\alpha}'_n.$$

Ak do tejto rovnosti dosadíme za $\vec{\alpha}'_1, \ldots, \vec{\alpha}'_n$ z (9.1), dostaneme

$$\vec{\gamma} = \sum_{i=1}^{n} x'_i \sum_{j=1}^{n} p_{ij}\vec{\alpha}_j.$$

Aby sme dostali koeficienty pri jednotlivých vektoroch z $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$, zmeníme poradie sčitovania.

$$\vec{\gamma} = \sum_{j=1}^{n} \vec{\alpha}_j \sum_{i=1}^{n} x'_i p_{ij}$$

Výraz $\sum_{i=1}^{n} x'_i p_{ij}$, ktorý sme dostali pri vektore $\vec{\alpha}_j$, je presne j-ty prvok z n-tice $(x'_1, \ldots, x'_n)P$. Tým je tvrdenie vety dokázané. □

Opäť môžeme predchádzajúci dôkaz zapísať stručnejšie maticovým zápisom.

Dôkaz. Uvedomme si najprv, že $\vec{\gamma}$ má súradnice $(x'_1, \ldots, x'_n)$ v báze $\vec{\alpha}'_1, \ldots, \vec{\alpha}'_n$ práve vtedy, keď platí rovnosť

$$\vec{\gamma} = (x'_1, \ldots, x'_n) \begin{pmatrix} \vec{\alpha}'_1 \\ \vdots \\ \vec{\alpha}'_n \end{pmatrix}.$$

Spolu s rovnosťou (9.2) potom dostaneme

$$(x'_1, \ldots, x'_n) \begin{pmatrix} \vec{\alpha}'_1 \\ \vdots \\ \vec{\alpha}'_n \end{pmatrix} = (x'_1, \ldots, x'_n)P \begin{pmatrix} \vec{\alpha}_1 \\ \vdots \\ \vec{\alpha}_n \end{pmatrix},$$

teda súradnice v báze $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ sú $(x'_1, \ldots, x'_n)P$. □

Matica zobrazenia v danej báze

Definícia 9.1.8. Nech V je vektorový priestor a $f\colon V \to V$ je lineárne zobrazenie. Nech $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ je báza V. *Matica zobrazenia f vzhľadom na bázu* $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ je matica $A = \|a_{ij}\|$ taká, že platí

$$f(\vec{\alpha}_i) = a_{i1}\vec{\alpha}_1 + \cdots + a_{in}\vec{\alpha}_n.$$

Predchádzajúca definícia teda hovorí, že matica zobrazenia f pri báze V je taká matica, ktorej i-ty riadok tvoria súradnice obrazu i-teho bázového vektora v tejto báze.

Táto definícia do istej miery pripomína definíciu matice zobrazenia, ktorú poznáme z prvého ročníka (definícia 5.3.8). Tam sme používali štandardnú bázu. Na rozdiel od prípadu, ktorý sme uviedli tu, nepožadovali sme, aby zobrazenie išlo z daného vektorového priestoru do toho istého priestoru. Podobne aj tu by sme mohli definovať o čosi všeobecnejší pojem matice lineárneho zobrazenia $f\colon V \to W$ vzhľadom na nejakú dvojicu báz (jedna z nich je bázou priestoru V a druhá je bázou priestoru W), zatiaľ sa však uspokojíme s týmto jednoduchším prípadom.

Nasledujúce tvrdenie je do istej miery analogické s podobným výsledkom z prvého ročníka, ktorý hovoril o rovnosti medzi obrazom vektora a súčinom vektora s maticou zobrazenia (poznámka 5.4.10).

{prechod:TVRO

Tvrdenie 9.1.9. *Nech V je vektorový priestor, $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ je báza V a $f\colon V \to V$ je lineárne zobrazenie. Ak A je matica zobrazenia f pri báze $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ a vektor $\vec{\gamma}$ má v tejto báze súradnice $(x_1, \ldots, x_n)$, tak jeho obraz $f(\vec{\gamma})$ má v tej istej báze súradnice*

$$(x_1, \ldots, x_n)A.$$

Dôkaz. Podľa predpokladov platí $\vec{\gamma} = x_1\vec{\alpha}_1 + \cdots + x_n\vec{\alpha}_n$. Ak použijeme na obe strany rovnosti zobrazenie f, tak (s využitím linearity f) dostaneme

$$f(\vec{\gamma}) = f\left(\sum_{i=1}^{n} x_i\vec{\alpha}_i\right) = \sum_{i=1}^{n} x_i f(\vec{\alpha}_i) = \sum_{i=1}^{n} x_i \sum_{j=1}^{n} a_{ij}\vec{\alpha}_j = \sum_{j=1}^{n} \vec{\alpha}_j \sum_{i=1}^{n} x_i a_{ij}.$$

Vidíme, že j-ta súradnica vektora $f(\vec{\gamma})$ v báze $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ je $\sum_{i=1}^{n} x_i a_{ij}$, čo je skutočne j-ta súradnica vektora $(x_1, \ldots, x_n)A$. □

Opäť nás bude zaujímať to, ako sa zmení matica zobrazenia, ak zmeníme bázu vektorového priestoru.

{prechod:VTPODOB}

Veta 9.1.10. *Nech V je vektorové priestory, $f\colon V \to V$ je lineárne zobrazenie a $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$, $\vec{\alpha}'_1, \ldots, \vec{\alpha}'_n$ sú bázy priestoru V. Ak P je matica prechodu od $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ k $\vec{\alpha}'_1, \ldots, \vec{\alpha}'_n$, A je matica zobrazenia f pri báze $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ a B je matica tohoto zobrazenia pri báze $\vec{\alpha}'_1, \ldots, \vec{\alpha}'_n$, tak platí*

{prechod:EQPODOB}

$$B = PAP^{-1}. \tag{9.3}$$

Dôkaz. Uvažujme vektor $\vec{\gamma}$, ktorý má v báze $\vec{\alpha}'_1, \ldots, \vec{\alpha}'_n$ súradnice $(x'_1, \ldots, x'_n)$. Podľa vety 9.1.7 má tento vektor v báze $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ súradnice $(x'_1, \ldots, x'_n)P$. Ďalej z tvrdenia 9.1.9 vieme, že tento vektor sa v zobrazení f zobrazí na taký vektor, ktorý má súradnice $(x'_1, \ldots, x'_n)PA$ (ide opäť o súradnice v báze $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$.)

Schematicky môžeme situáciu v súradniciach $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ znázorniť takto

$$(x'_1, \ldots, x'_n)P \mapsto (x'_1, \ldots, x'_n)PA.$$

Čo dostaneme, ak sa na situáciu pozrieme v súradniciach $\vec{\alpha}'_1, \ldots, \vec{\alpha}'_n$? Súradnice vektora $f(\vec{\gamma})$ vieme použitím vety 9.1.7 previesť do tejto bázy použitím matice prechodu od $\vec{\alpha}'_1, \ldots, \vec{\alpha}'_n$ k $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$. Podľa vety 9.1.5 je to matica P^{-1}.

Vektor $f(\vec{\gamma})$ má teda v báze $\vec{\alpha}'_1, \ldots, \vec{\alpha}'_n$ súradnice $(x'_1, \ldots, x'_n)PAP^{-1}$. V súradniciach $\vec{\alpha}'_1, \ldots, \vec{\alpha}'_n$ to teda vyzerá takto:

$$(x'_1, \ldots, x'_n) \mapsto (x'_1, \ldots, x'_n)PAP^{-1}$$

Podľa tvrdenia 9.1.9 však súčasne musí platiť, že $f(\vec{\gamma})$ má v báze $\vec{\alpha}'_1, \ldots, \vec{\alpha}'_n$ súradnice $(x'_1, \ldots, x'_n)B$.

$$(x'_1, \ldots, x'_n) \mapsto (x'_1, \ldots, x'_n)B$$

Z toho dostávame rovnosť

$$(x'_1, \ldots, x'_n)B = (x'_1, \ldots, x'_n)PAP^{-1}.$$

Pretože táto rovnosť platí pre ľubovoľnú n-ticu $(x'_1, \ldots, x'_n) \in F^n$, musí platiť maticová rovnosť

$$B = PAP^{-1}.$$

□

Definícia 9.1.11. Nech A, B sú štvorcové matice nad poľom F. Ak existuje matica P taká, že $B = PAP^{-1}$, hovoríme, že matice A a B sú *podobné.*

Z vety 9.1.10 vyplýva, že 2 matice sú podobné práve vtedy, keď existujú lineárne zobrazenie a dvojica báz také, že toto zobrazenie má v jednej báze maticu A a v druhej maticu B.

Je pomerne ľahké overiť, že podobnosť matíc je relácia ekvivalencie.

Cvičenia

Úloha 9.1.1. Ukážte, že podobnosť matíc (chápaná ako relácia na $M_{n,n}(F)$) je relácia ekvivalencie.

Úloha 9.1.2. Pre $\vec{\alpha}_1 = (2,1)$, $\vec{\alpha}_2 = (1,2)$, $\vec{\beta}_1 = (-1,1)$, $\vec{\beta}_2 = (2,3)$, $\vec{\gamma}_1 = (1,1)$, $\vec{\gamma}_2 = (3,1)$. Nájdite:
a) Maticu P_1 prechodu od bázy $\vec{\alpha}_1, \vec{\alpha}_2$ k báze $\vec{\beta}_1$, $\vec{\beta}_2$.
b) Maticu P_2 prechodu od bázy $\vec{\beta}_1, \vec{\beta}_2$ k báze $\vec{\gamma}_1$, $\vec{\gamma}_2$.
c) Maticu P_3 prechodu od bázy $\vec{\alpha}_1, \vec{\alpha}_2$ k báze $\vec{\gamma}_1$, $\vec{\gamma}_2$.
d) Aký je vzťah medzi maticami P_1, P_2 a P_3?

Úloha 9.1.3. Nájdite všetky matice, ktoré sú podobné s nulovou maticou.

Úloha 9.1.4. Ak aspoň jedna zo štvorcových matíc A, B stupňa n je regulárna, tak AB a BA sú podobné. Platí to aj za predpokladu, že nie sú regulárne?

Úloha 9.1.5. Nech $A = cI$. Aké matice sú podobné s maticou A?

Úloha 9.1.6. Pre vektory $\vec{\gamma}_i \in \mathbb{R}^3$, $i = 1,2,3$, označme ako $\vec{x}_i$ súradnice vektora v báze $\vec{\alpha}_1, \vec{\alpha}_2, \vec{\alpha}_3$ a $\vec{x}'_i$ súradnice toho istého v báze $\vec{\alpha}'_1, \vec{\alpha}'_2, \vec{\alpha}'_3$. Nájdite matice prechodu od $\vec{\alpha}_1, \vec{\alpha}_2, \vec{\alpha}_3$ k $\vec{\alpha}'_1, \vec{\alpha}'_2, \vec{\alpha}'_3$ ak viete, že $\vec{x}_1 = (1,2,1)$, $\vec{x}'_1 = (-1,1,1)$, $\vec{x}_2 = (-1,0,3)$, $\vec{x}'_2 = (1,-1,1)$, $\vec{x}_3 = (3,1,2)$ a $\vec{x}'_3 = (2,1,-2)$. (Návod: Bude to matica istého lineárneho zobrazenia.)

Úloha 9.1.7. Ukážte, že ak matica A je podobná matici B, tak aj matice A^{-1} a B^{-1} sú podobné.

Úloha 9.1.8. Ukážte, že ak A a B sú podobné, tak majú rovnakú hodnosť, determinant a stopu. (Stopu matice sme definovali v úlohe 7.2.9.)

Úloha 9.1.9. Nájdite všetky matice A také, že jediná matica, ktorá je podobná s A, je práve matica A. (Inak povedané, trieda ekvivalencie matice A je jednoprvková.)

Nasledujúce úlohy sa týkajú „dosadzovania“ matíc do polynómov. Ak máme polynóm tvaru $p(x) = c_n x^n + \cdots + c_1 x + c_0$, kde $c_0, \ldots, c_n \in F$, tak symbolom $p(A)$ rozumieme maticu

$$p(A) = c_n A^n + \cdots + c_1 A + c_0 I.$$

Toto sa skutočne podobá na dosadenie matice do toho istého polynómu - s tým, že posledný člen sme interpretovali ako $c_0 A^0 = c_0 I$. (Nemohli sme tam napísať iba c_0 – dostali by sme tak súčet matice a čísla, čo nedáva zmysel.)

Môžeme sa pýtať na to, aké polynómy vynulujú danú maticu, t.j. platí

$$p(A) = c_n A^n + \cdots + c_1 Ax + c_0 I = 0.$$

Neskôr uvidíme výsledok, ktorým sa dá dostať, že maticu A vynuluje tzv. *charakteristický polynóm* – tento polynóm vieme vypočítať priamo z matice A (veta 9.2.14, nazývaná Cayley-Hamiltonova veta). Ale aj s vedomosťami, ktoré máme už teraz, vieme pomerne ľahko ukázať, že maticu A vynuluje nejaký polynóm stupňa nanajvýš n^2. A s trochu väčšou námahou sa nám to podarí aj pre polynóm stupňa najviac n.

{ULOMATPOLYN^

Úloha 9.1.10. Nech $A \in M_{n,n}(F)$. Ukážte, že existuje nenulový polynóm stupňa nanajvýš n^2 taký, že $p(A) = 0$.

{ULOMATPOLYVE

Úloha 9.1.11. Nech $A \in M_{n,n}(F)$ a $\vec{x} \in F^n$. Dokážte, že existuje nenulový polynóm stupňa nanajvýš n taký, že $\vec{x}p(A) = 0$.

Výsledok pre stupeň n je asi výhodnejšie dokazovať v podobe pre lineárne transformácie. Ak máme vektorové priestor V a lineárne zobrazenie $f\colon V \to V$ tak pre polynóm $p(x) = c_n x^n + \cdots + c_1 x + c_0$ môžeme zobrať

$$p(f) = c_n f^n + \cdots + c_1 f + c_0 id_V.$$

Ak naše zobrazenie je zobrazenie $f(\vec{x}) = \vec{x}A$, tak $p(f)$ je zobrazenie určené maticou $p(A)$. (Všeobecnejšie: Ak f má pri nejakej báze maticu A, tak $p(f)$ má pri tej istej báze maticu $p(A)$.)

Úloha 9.1.12. Nech V je vektorový priestor nad poľom F a $f\colon V \to V$ je lineárna transformácia. Ukážte, že pre ľubovoľný polynóm $p(x) \in F[x]$ je aj $p(f)\colon V \to V$ lineárne transformácia.

{ULOMATPOLYSKLAD}

Úloha 9.1.13. Nech V je vektorový priestor nad poľom F a $f\colon V \to V$ je lineárna transformácia. Ukážte, že ak $p_1, p_2 \in F[x]$, tak pre ich súčin $p(x) = p_1(x) \cdot p_2(x)$ platí

$$p(f) = p_1(f) \circ p_2(f).$$

{ULOMATPOLYN}

Úloha 9.1.14*. Nech V je konečnorozmerný vektorový priestor dimenzie n nad poľom F a $f\colon V \to V$ je lineárna transformácia. Dokážte, že existuje nenulový polynóm $p(x) \in F[x]$ taký, že $p(f) = 0$.

Z toho dostávame pre $A \in M_{n,n(F)}$ existenciu polynómu stupňa najviac n takého, že $p(f) = 0$.

(Hint 1: Ak si zvolíme nejaký nenulový vektor $\vec{x} \in V$, tak dostaneme nejaký polynóm $p_1(x)$ na základe úlohy 9.1.11. Čo vieme povedať o dimenzii podpriestoru $\mathrm{Ker}(p_1(f))$? Hint 2: Vedeli by ste nejako použiť podpriestor $\mathrm{Im}(p_2(f))$ a matematickú indukciu na dokončenie dôkazu tohoto tvrdenia?)

9.2 Podobnosť s diagonálnou maticou

{diag:SECT}

9.2.1 Nutné a postačujúce podmienky

Pre štvorcovú maticu A je zaujímavé zistiť, či je podobná s diagonálnou maticou, t.j. či existuje regulárna matica P taká, že PAP^{-1} je diagonálna (matica). Kvôli zjednodušeniu budeme diagonálnu maticu

$$D = \begin{pmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & d_n \end{pmatrix}$$

skrátene zapisovať ako $\mathrm{diag}(d_1, d_2, \dots, d_n)$.

Ak teda A je podobná diagonálnej, je $PAP^{-1} = \mathrm{diag}(d_1, d_2, \dots, d_n) = D$ pre vhodnú maticu P a vhodné čísla $d_1, \dots, d_n$, potom vieme ľahko vypočítať napr. A^{100} ako

$$P^{-1}\mathrm{diag}(d_1^{100}, d_2^{100}, \dots, d_n^{100})P,$$

alebo ak v danom poli existujú napr. $\sqrt{d}_1, \ldots, \sqrt{d}_n$, tak vieme vypočítať niečo ako $\sqrt{A}$ (t.j. maticu B takú, že $B^2 = A$) pomocou $P^{-1} \cdot \mathrm{diag}(\sqrt{d}_1, \sqrt{d}_2, \ldots, \sqrt{d}_n) \cdot P$ (matica B nie je vo všeobecnosti určená jednoznačne, toto je jedno možné riešenie). Alebo keby sme chceli vypočítať niečo typu e^A, mohli by sme použiť Taylorov rozvoj $e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots$ a potom počítať

$$e^A = P^{-1}(I + D + \frac{D^2}{2!} + \frac{D^3}{3!} + \ldots)P = P^{-1} \cdot \mathrm{diag}(e^{d_1}, \ldots, e^{d_n}) \cdot P$$

Tento výpočet je urobený formálne, bez toho, aby sme strážili konvergenciu potrebných radov, ale pri troche starostlivosti by sa dalo ukázať, že je to pomerne zmysluplný postup. Dá sa potom robiť pre napr. funkcie $f(x)$, ktoré majú Taylorov rozvoj, ktorý pri dosadení všetkých čísiel $d_1, \ldots, d_n$ konverguje - t.j. všetky sa nachádzajú vnútri polomeru konvergencie príslušného Taylorovho radu.

Dá sa potom definovať aj niečo ako $e^{At} = P^{-1} \cdot \mathrm{diag}(e^{d_1 t}, \ldots, e^{d_n t}) \cdot P$. (Ide o funkciu, ktorá každému reálnemu číslu t priradí maticu.) Pre funkciu $f(t) = e^{At}$, potom platí $f'(t) = Af(t)$, čo aspoň trochu naznačuje, že takáto funkcia by mohla súvisieť s riešením diferenciálnych rovníc. (Na overenie rovnosti $f'(t) = Af(t)$ sa stačí presvedčiť o tom, že $(e^{At})' = P(e^{Dt})'P^{-1} = PDe^{Dt}P^{-1} = PDP^{-1}Pe^{Dt}P^{-1} = Ae^{At}$.)

Jedna vec, ktorá v danom momente nie je zrejmá je, či čísla $d_1, \ldots, d_n$ závisia od P - matice prechodu - alebo nie. Z nasledujúceho postupu bude jasné, že nezávisia (okrem poradia).

Skúsme si uvedomiť, čo presne znamená, že matica A je podobná s diagonálnou maticou D. To nám hovorí, že zobrazenie, ktoré má pri štandardnej báze maticu A, má pri vhodnej báze $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ maticu $D = \mathrm{diag}(d_1, d_2, \ldots, d_n)$. Pre každý z vektorov bázy teda platí

$$\vec{\alpha}_i A = d_i \vec{\alpha}_i.$$

Skúsme sa na to ešte pozrieť trochu inak. To, že A a D sú podobné nám dáva rovnosti

$$PAP^{-1} = D$$
$$PA = DP$$

Označme teraz riadky matice P ako $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$. Potom z predchádzajúcej rovnosti dostaneme:

$$\begin{pmatrix} \vec{\alpha}_1 \\ \vdots \\ \vec{\alpha}_n \end{pmatrix} A = D \begin{pmatrix} \vec{\alpha}_1 \\ \vdots \\ \vec{\alpha}_n \end{pmatrix},$$
$$\begin{pmatrix} \vec{\alpha}_1 A \\ \vdots \\ \vec{\alpha}_n A \end{pmatrix} = \begin{pmatrix} d_1\vec{\alpha}_1 \\ \vdots \\ d_n\vec{\alpha}_n \end{pmatrix}. \qquad (9.4) \quad \texttt{\{diagon:EQ1\}}$$

Keď porovnáme jednotlivé riadky matíc v poslednej rovnosti, opäť dostávame presne rovnaký vzťah

$$\vec{\alpha}_i A = d_i \vec{\alpha}_i.$$

Zdá sa, že pri zisťovaní, či je daná matica podobná s diagonálnou, by mohli hrať zaujímavú úlohu dvojice $c \in F$ a $\vec{\alpha} \in F^n$ s vlastnosťou $\vec{\alpha}A = c\vec{\alpha}$. Budeme sa teda teraz chvíľu zaoberať tým, ako takéto dvojice nájsť.

Definícia 9.2.1. Nech A je štvorcová matica nad poľom F. Prvok $c \in F$ nazveme *vlastným číslom* matice A, ak existuje nenulový vektor $\vec{\alpha} \in F^n$ taký, že $\vec{\alpha}A = c\vec{\alpha}$.

Nenulový vektor $\vec{\alpha} \in F^n$ nazývame *vlastným vektorom* matice A, ak existuje $c \in F$ (c môže byť aj 0) také, že $\vec{\alpha}A = c\vec{\alpha}$.

Ak $\vec{\alpha}$ je nenulový vektor a pre $c \in F$ platí $\vec{\alpha}A = c\vec{\alpha}$, hovoríme, aj, že (vlastný) vektor $\vec{\alpha}$ prislúcha ku vlastnému číslu c, alebo že (vlastné) číslo c prislúcha ku vlastnému vektoru $\vec{\alpha}$.

Ako nájdeme vlastné čísla a vlastné vektory matice A?
Najprv vlastné čísla: Pozrime sa nasledujúce ekvivalentné tvrdenia:

1. c je vlastné číslo matice A
2. Existuje nenulový vektor $\vec{\alpha}$ taký, že $\vec{\alpha}A = c\vec{\alpha}$
3. Existuje nenulový vektor $\vec{\alpha}$ taký, že $\vec{\alpha}A = \vec{\alpha}(cI)$ (I je identická matica)
4. Existuje nenulový vektor $\vec{\alpha}$ taký, že $\vec{\alpha}(A - cI) = \vec{0}$
5. Jadro zobrazenia s maticou $A - cI$ je netriviálne (t.j. toto zobrazenie nie je injektívne, t.j. matica $A - cI$ je singulárna).
6. Determinant matice $A - cI$ je nulový.

Keďže v tomto momente hľadáme vhodné prvky c, môžeme sa na determinant matice $A - cI$ v poslednom tvrdení pozrieť ako na výraz v „neznámej“ c - skúsme radšej použiť premennú x. Je dobre si uvedomiť, že $|A - xI|$ je polynóm v premennej x, napríklad ak

$$A = \begin{pmatrix} 1 & 2 \\ 0 & 4 \end{pmatrix}, \text{ tak } \quad |A - xI| = \begin{vmatrix} 1-x & 2 \\ 0 & 4-x \end{vmatrix} = (1-x)(4-x) - 0 \cdot 2 = 4 - 5x + x^2$$

Posledne menovaný determinant budeme nazývať *charakteristický polynóm* matice A a označovať ako $ch_A(x)$, t.j. $ch_A(x) = |A - xI|$. Zistiť vlastné čísla matice A teda znamená nájsť korene jej charakteristického polynómu $ch_A(x)$.

Pre uvedený príklad teda dostávame, že vlastné čísla sú 1 a 4.

Nájsť vlastné vektory znamená teraz pre dané vlastné číslo c nájsť netriviálne riešenia rovnice $\vec{\alpha}(A - cI) = \vec{0}$. Ak si napíšeme $\vec{\alpha} = (x_1, \ldots, x_n)$, rovnicu môžeme prepísať do tvaru

$$(A - cI)^T \vec{\alpha}^T = (A - cI)^T \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

čo je vlastne homogénny systém rovníc s neznámymi $x_1, \ldots, x_n$ a maticou systému $(A - cI)^T$.

Pre uvedenú maticu

$$A = \begin{pmatrix} 1 & 2 \\ 0 & 4 \end{pmatrix}$$

teda vieme, že jej vlastné čísla sú 1 a 4. Nájdime vlastné vektory:

Pre vlastné číslo 1:

$$A - 1 \cdot I = \begin{pmatrix} 1-1 & 2 \\ 0 & 4-1 \end{pmatrix} = \begin{pmatrix} 0 & 2 \\ 0 & 3 \end{pmatrix}, \text{ teda } \quad (A - 1 \cdot I)^T = \begin{pmatrix} 0 & 0 \\ 2 & 3 \end{pmatrix}$$

Hľadáme riešenia homogénneho systému rovníc s poslednou maticou, t.j.

$$(A - 1 \cdot I)^T = \begin{pmatrix} 0 & 0 \\ 2 & 3 \end{pmatrix} \sim \begin{pmatrix} 2 & 3 \\ 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & \frac{3}{2} \\ 0 & 0 \end{pmatrix}$$

odkiaľ je vidieť, že vlastné vektory prislúchajúce vlastnému číslu 1 sú nenulové vektory z podpriestoru $[(-\frac{3}{2}, 1)] = [(-3, 2)]$. (Overte si, že napríklad $(-3, 2)A = (-3, 2)$.)

Pre vlastné číslo 4:

$$A - 4 \cdot I = \begin{pmatrix} 1-4 & 2 \\ 0 & 4-4 \end{pmatrix} = \begin{pmatrix} -3 & 2 \\ 0 & 0 \end{pmatrix}, \text{ teda } \quad (A - 4 \cdot I)^T = \begin{pmatrix} -3 & 0 \\ 2 & 0 \end{pmatrix}$$

Hľadáme riešenia homogénneho systému rovníc s poslednou maticou, t.j.

$$(A - 4 \cdot I)^T = \begin{pmatrix} -3 & 0 \\ 2 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 \\ 2 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

odkiaľ je vidieť, že vlastné vektory prislúchajúce vlastnému číslu 4 sú nenulové vektory z podpriestoru $[(0,1)]$. (Overte si, že $(0,1)A = 4(0,1)$.)

Teraz sa môžeme zamyslieť nad tým, akú maticu má lineárna transformácia určená v báze $(1,0),(0,1)$ maticou A v báze $(-3,2),(0,1)$, t.j. ak $\vec{\varepsilon}_1 = (1,0), \vec{\varepsilon}_2 = (0,1)$ a $\vec{\alpha}_1 = (-3,2), \vec{\alpha}_2 = (0,1)$, $A = A_\varphi^{\vec{\varepsilon}_1,\vec{\varepsilon}_2} (= A_\varphi^\varepsilon)$, čo bude $A_\varphi^{\vec{\alpha}_1,\vec{\alpha}_2} (= A_\varphi^\alpha)$. Podľa vzorca (9.3) je

$$A_\varphi^\alpha = P A_\varphi^\varepsilon P^{-1}, \text{ kde } \quad P = \begin{pmatrix} -3 & 2 \\ 0 & 1 \end{pmatrix},$$

t.j. riadky matice P sú vektory $\vec{\alpha}_1, \vec{\alpha}_2$, t.j. generátory podpriestorov $[(-3,2)]$ (ktorého nenulové vektory sú vlastné vektory prislúchajúce ku vlastnému číslu 1) a $[(0,1)]$ (ktorého nenulové vektory sú vlastné vektory prislúchajúce ku vlastnému číslu 4). P je samozrejme matica prechodu od epsilonovej ku alfovej báze. Ale A_φ^α v $i-$tom riadku obsahuje súradnice obrazu vektora $\vec{\alpha}_i$ vyjadrené v báze $\vec{\alpha}_1, \vec{\alpha}_2$, a keďže $\vec{\alpha}_1$ je vlastný vektor prislúchajúci ku vlastnému číslu 1, je

$$\vec{\alpha}_1 A = 1.\vec{\alpha}_1 + 0.\vec{\alpha}_2$$

a podobne, keďže $\vec{\alpha}_2$ je vlastný vektor prislúchajúci ku vlastnému číslu 4, je

$$\vec{\alpha}_2 A = 0.\vec{\alpha}_1 + 4.\vec{\alpha}_2$$

a preto je $A_\varphi^\alpha = \operatorname{diag}(1,4) = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix} = PAP^{-1}$

Iný spôsob, ako môžeme zdôvodniť túto rovnosť je použiť rovnaký postup ako pri odvodení (9.4).

V predchádzajúcom príklade tvorili vlastné vektory bázu priestoru $\mathbb{R}^2$, v ktorom sme pracovali. Vďaka tomu sme z nich dostali regulárnu maticu P. Nasledujúci príklad ukazuje, že to tak nemusí byť vždy.

Príklad 9.2.2. Vypočítajme vlastné čísla a vlastné vektory pre maticu

$$A = \begin{pmatrix} 1 & -3 & 3 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}$$

Tu je $ch_A(x) = \begin{vmatrix} 1-x & -3 & 3 \\ 0 & 1-x & -2 \\ 0 & 0 & 1-x \end{vmatrix} = (1-x)^3$, t.j. máme jediné vlastné číslo, 1 - je trojnásobný koreň charakteristického polynómu - (to samo ešte nemusí byť na závadu). Ale keď hľadáme vlastné vektory, zistíme, že sú to riešenia homogénneho systému rovníc s maticou $(A-I)^T$, t.j. $\begin{pmatrix} 0 & -3 & 3 \\ 0 & 0 & -2 \\ 0 & 0 & 0 \end{pmatrix}^T = \begin{pmatrix} 0 & 0 & 0 \\ -3 & 0 & 0 \\ 3 & -2 & 0 \end{pmatrix}$. Táto matica má očividne hodnosť

2 a preto riešenia tohoto systému tvoria jednorozmerný podpriestor ($[(0,0,1)]$), ktorého nenulové prvky sú jediné vlastné vektory tejto matice. Ako ukážeme, toto je problém (málo lineárne nezávislých vlastných vektorov), ktorý spôsobuje, že uvedená matica nie je podobná so žiadnou diagonálnou maticou.

{sec:SKU:veta

Veta 9.2.3. *Nech $A = ||a_{ij}||$ je štvorcová matica typu $n \times n$ nad poľom F. Potom A je podobná s diagonálnou maticou práve vtedy, keď spomedzi vlastných vektorov vieme vybrať bázu.*

Dôkaz. $\Rightarrow$: Nech A je podobná diagonálnej matici $\mathrm{diag}(d_1, d_2, \ldots, d_n)$, t.j. existuje regulárna matica P taká, že $PAP^{-1} = \mathrm{diag}(d_1, d_2, \ldots, d_n)$. Potom ak $\varphi\colon F^n \to F^n$ je zobrazenie s maticou A, t.j. $A = A_\varphi^\varepsilon = A_\varphi^{\vec{\varepsilon}_1,\ldots,\vec{\varepsilon}_n}$, tak maticu P môžeme považovať za maticu prechodu od bázy $\vec{\varepsilon}_1, \ldots, \vec{\varepsilon}_n$ ku báze $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$, kde $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ sú postupne riadky matice P. Podľa vzorca (9.3) platí

$$PA_\varphi^\epsilon P^{-1} = A_\varphi^{\vec{\alpha}_1,\ldots,\vec{\alpha}_n},$$

a teda $A_\varphi^{\vec{\alpha}_1,\ldots,\vec{\alpha}_n} = \mathrm{diag}(d_1, d_2, \ldots, d_n)$. Toto a význam matice $A_\varphi^{\vec{\alpha}_1,\ldots,\vec{\alpha}_n}$ hovorí, že $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ sú vektory s vlastnosťou $\vec{\alpha}_i A = 0 \cdot \vec{\alpha}_1 + \cdots + d_i\vec{\alpha}_i + \cdots + 0 \cdot \vec{\alpha}_n = d_i\vec{\alpha}_i$, t.j. $\vec{\alpha}_i, i = 1, \ldots, n$ sú vlastné vektory matice A prislúchajú po rade vlastným číslam $d_i, i = 1, \ldots, n$ (a tiež, že $d_i, i = 1, \ldots, n$ sú vlastné čísla). Ale keďže vektory $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ sú riadky regulárnej matice P, tvoria bázu F^n, takže vieme, že existujú vlastné vektory tvoriace bázu.

Opäť i v tomto prípade by sme ako alternatívne zdôvodnenie mohli použiť rovnaký postup ako pri odvodení (9.4).

$\Leftarrow$: Nech sa z vlastných vektorov matice A dá vybrať báza $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$. Potom ak tieto vektory uložíme ako riadky do matice P, tak rovnako ako v prvej časti dôkazu vidíme, že $PAP^{-1} = PA_\varphi^\epsilon P^{-1} = A_\varphi^{\vec{\alpha}_1,\ldots,\vec{\alpha}_n}$. Ale matica $A_\varphi^{\vec{\alpha}_1,\ldots,\vec{\alpha}_n}$ je diagonálna, lebo $c_1, \ldots, c_n$ sú vlastné čísla. □

Táto veta umožňuje zistiť, či je matica podobná s diagonálnou alebo nie, ale jej použitie je numericky pomerne náročné, a preto je vhodné (ak nás naozaj zaujíma len táto skutočnosť a nie aj matica prechodu P) nájsť iné, aspoň postačujúce podmienky, ktoré zabezpečia to, že matica A je podobná diagonálnej. Jedna z dvoch, ktoré si uvedieme, je založená na leme

islostVlastnychVektorov}

Lema 9.2.4. *Nech $A = ||a_{ij}||$ je štvorcová matica typu $n \times n$ nad poľom F a vlastné čísla $c_1, \ldots, c_k$ matice A sú navzájom rôzne prvky poľa F, $\vec{\alpha}_1, \ldots, \vec{\alpha}_k$ sú vlastné vektory po rade prislúchajúce $c_1, \ldots, c_k$. Potom sú vektory $\vec{\alpha}_1, \ldots, \vec{\alpha}_k$ lineárne nezávislé.*

(*Stručne: Rôznym vlastným číslam zodpovedajú lineárne nezávislé vlastné vektory.*)

Dôkaz. Sporom. Vektor $\vec{\alpha}_1$ je vlastný vektor a preto je nenulový. Preto je závislosť vektorov $\vec{\alpha}_1, \ldots, \vec{\alpha}_k$ ekvivalentná s tým, že jeden z nich (pre $i > 1$) lineárnou kombináciu predchádzajúcich, t.j. existujú $a_1, \ldots, a_{i-1}$ také, že

$$\vec{\alpha}_i = a_1\vec{\alpha}_1 + \cdots + a_{i-1}\vec{\alpha}_{i-1}$$

Predpokladajme, že sme vybrali najmenšie i s takouto vlastnosťou, t.j. že vektory $\vec{\alpha}_1, \ldots, \vec{\alpha}_{i-1}$ už sú lineárne nezávislé. Teraz využijeme, že $\vec{\alpha}_1, \ldots, \vec{\alpha}_k$ sú vlastné vektory, preto

$$\vec{\alpha}_i A = c_i\vec{\alpha}_i = c_i(a_1\vec{\alpha}_1 + \cdots + a_{i-1}\vec{\alpha}_{i-1}) = c_i a_1\vec{\alpha}_1 + \cdots + c_i a_{i-i}\vec{\alpha}_{i-1},$$

kde aspoň jedno z čísiel $a_1, \ldots, a_{i-1}$ je nenulové (inak by bol vektor $\vec{\alpha}_i$ nulový, čo sa vlastnému vektoru nemôže stať) ale aj

$$(a_1\vec{\alpha}_1 + \cdots + a_{i-1}\vec{\alpha}_{i-1})A = a_1\vec{\alpha}_1 A + \cdots + a_{i-i}\vec{\alpha}_{i-1}A = c_1 a_1\vec{\alpha}_1 + \cdots + c_{i-1}a_{i-i}\vec{\alpha}_{i-1}$$

odkiaľ porovnaním a prehodením všetkých členov na ľavú stranu získame rovnosť

$$(c_i - c_1)a_1\vec{\alpha}_1 + \cdots + (c_i - c_{i-1})a_i\vec{\alpha}_{i-1} = \vec{0}$$

Keďže všetky čísla $c_i - c_1, \ldots, c_i - c_{i-1}$ sú nenulové, tak aspoň jedno z čísiel $(c_i - c_1)a_1, \ldots, (c_i - c_{i-1})a_{i-1}$ je nenulové, čím dostávame spor s lineárnou nezávislosťou vektorov $\vec{\alpha}_1, \ldots, \vec{\alpha}_{i-1}$. □

Príslušná postačujúca podmienka (ešte stále numericky pomerne náročná) potom znie

{sec:SKU:veta_PodobnaSDi

Dôsledok 9.2.5. *Nech $A = ||a_{ij}||$ je štvorcová matica typu $n \times n$ nad poľom F a vlastné čísla $c_1, \ldots, c_n$ matice A sú navzájom rôzne prvky poľa F (t.j. A má n navzájom rôznych vlastných čísiel z poľa F). Potom A je podobná s diagonálnou maticou.*

Dôkaz. Treba spojiť výsledok lemy 9.2.4 a vety 9.2.3 a uvedomiť si, že n lineárne nezávislých vektorov v priestore F^n tvorí bázu. □

Ešte sformulujme niekoľko pomocných kritérií, ktoré pomôžu zistiť skutočnosť, že dve konkrétne matice A, B nie sú podobné (obe tieto kritériá naozaj fungujú len jedným smerom, t.j. žiadne z nich nevie potvrdiť, že matice A, B sú podobné, vedia len vylúčiť tento fakt - sú to nutné podmienky na podobnosť).

{diagon:LMROVNAKYCHAR}

Lema 9.2.6. *Nech A, B sú štvorcové matice typu $n \times n$ nad poľom F. Ak A a B sú podobné, tak $ch_A(x) = ch_B(x)$.*

Dôkaz. Nech sú A, B podobné, t.j. existuje taká regulárna matica P, že $PAP^{-1} = B$. Počítajme

$$\begin{aligned} ch_B(x) &= |B - xI| = |PAP^{-1} - xI| = |PAP^{-1} - xPIP^{-1}| = |P(A - xI)P^{-1}| = \\ &= |P||A - xI||P^{-1}| = |PP^{-1}||A - xI| = |I||A - xI| = |A - xI| = ch_A(x) \end{aligned}$$

□

a ešte jednoduchšie kritérium

{diagon:DOSSTOPA}

Dôsledok 9.2.7. *Pre maticu $A = ||a_{ij}||$ - štvorcová matica typu $n \times n$ nad poľom F položme $\mathrm{Tr}(A) = a_{11} + a_{22} + \cdots + a_{nn}$ - t.j. $\mathrm{Tr}(A)$ je súčet prvkov na diagonále matice A. Ak sú matice A, B podobné, tak $\mathrm{Tr}(A) = \mathrm{Tr}(B)$.*

Hodnota $\mathrm{Tr}(A)$ sa nazýva *stopa matice A*.

Dôkaz. Treba si uvedomiť, ako vyzerá charakteristický polynóm matice A, je to

$$A - xI = \begin{vmatrix} a_{11} - x & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} - x & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & \ldots & a_{n,n-1} & a_{nn} - x \end{vmatrix}$$

Podľa definície determinant $|B| = ||b_{ij}||$ je súčet súčinov $(-1)^{i(\varphi)} b_{1\varphi(1)} \ldots b_{n\varphi(n)}$, kde $\varphi \in S_n$. Špeciálne pre našu maticu $A - xI$ si treba všimnúť, že pre φ identickú permutáciu, t.j. výber diagonálnych prvkov tam máme člen $+(a_{11} - x)) \cdots (a_{nn} - x)$. Ak zoberieme akýkoľvek iný súčin, neobsahuje aspoň jeden diagonálny prvok, ale musí obsahovať z každého riadka a

každého stĺpca práve jeden prvok, tak musí existovať ešte jeden diagonálny prvok, ktorý neobsahuje a preto ako polynóm v premennej x má stupeň najviac $n-2$. Keďže

$$\begin{aligned}(a_{11}-x)\cdots(a_{nn}-x) &= (-1)^n x^n + (-1)^{n-1}(a_{11}+\cdots+a_{nn})x^{n-1}+\dots \\ &= (-1)^n x + (-1)^{n-1}\operatorname{Tr}(A)x^{n-1}+\dots\end{aligned}$$

a ostatné súčiny neovplyvnia koeficienty pri x^n a x^{n-1} v $ch_A(x)$, a podľa predošlej lemy $ch_A(x) = ch_B(x)$. Preto ich koeficienty pri x^{n-1} sú rovnaké, ale tieto koeficienty sú (až na znamienko) $\operatorname{Tr}(A)$, respektíve $\operatorname{Tr}(B)$, dostávame rovnosť $\operatorname{Tr}(A) = \operatorname{Tr}(B)$. □

Iný dôkaz. Ľahko sa dá overiť, že platí $\operatorname{Tr}(AB) = \operatorname{Tr}(BA)$ (úloha 5.4.4). Z toho máme $\operatorname{Tr}(PAP^{-1}) = \operatorname{Tr}(APP^{-1}) = \operatorname{Tr}(A)$. □

Podobným spôsobom sa dá dokázať, že podobné matice A, B majú rovnaké determinanty (až na znamienko sú to absolútne koeficienty - koeficienty pri x^0 - v charakteristických polynómoch). Opäť, ten istý fakt môžeme dokázať aj použitím rovnosti $|AB| = |A||B| = |BA|$.

{diagon:PRROVNAKYCHAR}

Príklad 9.2.8. Ak si vezmeme matice $A = \left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)$ a $B = \left(\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right)$, tak vidíme, že obe matice majú rovnaký charakteristický polynóm $ch_A(x) = ch_B(x) = (x-1)^2$. Ľahko sa overí, že matica B nie je podobná so žiadnou diagonálnou maticou

Tento príklad teda ukazuje, že implikácie v leme 9.2.6 platí iba jedným smerom. (Iný kontrapríklad môžete nájsť v úlohe 9.2.14.)

9.2.2 Symetrické matice – veta o hlavných osiach

Teraz prejdeme ku druhému sľubovanému kritériu, ktoré zabezpečí, že matica je podobná s diagonálnou. Je založené na úplne inom princípe ako kritérium z dôsledku 9.2.5. Nižšie je sformulované pod názvom „veta o hlavných osiach“ (veta 9.2.11). Numericky je veľmi jednoduché, ale má pomerne úzky „rozsah aplikovateľnosti“ – hovorí, že každá reálna symetrická matica je podobná s diagonálnou maticou a že dokonca v tomto prípade vieme podobnosť zabezpečiť pomocou ortogonálnej matice P, t.j. táto podobnosť je zároveň kongruencia matíc (lebo *ortogonálna matica* je definovaná ako matica P spĺňajúca podmienku $P^T = P^{-1}$, a teda $PAP^{-1} = PAP^T$). Takúto podobnosť budeme nazývať ortogonálna podobnosť.

Zastavme sa chvíľu pri definícii ortogonálnej matice. Podľa definície je to matica, ktorá spĺňa $PP^T = P^TP = I$. Rovnosť $PP^T = I$ vlastne znamená, že riadky matice P sú ortonormálne vektory. Rovnosť $P^TP = I$ hovorí to isté o stĺpcoch matice P.

Pred uvedením samotnej vety o hlavných osiach potrebujeme dve tvrdenia.

eta_HornaTrojuholnikova}

Veta 9.2.9 (Schurova veta)**.** *Nech $A = ||a_{ij}||$ je štvorcová matica typu $n \times n$ nad poľom $\mathbb{R}$. Nech všetky vlastné čísla matice A sú z poľa $\mathbb{R}$. Potom existuje horná trojuholníková matica T, ktorá je ortogonálne podobná s maticou A.*

Dôkaz. Nech $a_n \in \mathbb{R}$ je vlastné číslo matice A, nech $\vec{\alpha}_n \in \mathbb{R}^n$ je vlastný vektor matice A prislúchajúci ku a_n a ktorý má dĺžku 1. Nech $\vec{\alpha}_1, \dots, \vec{\alpha}_{n-1}$ sú vektory v $\mathbb{R}^n$, ktoré dopĺňajú $\vec{\alpha}_n$ do ortonormálnej bázy priestoru $\mathbb{R}^n$. Transformácia s maticou A (t.j. A je matica tejto transformácie pri štandardnej báze) má v báze $\vec{\alpha}_1, \dots, \vec{\alpha}_n$ maticu A', ktorej posledný riadok obsahuje len jeden zaujímavý prvok — posledný prvok je a_n a ostatné prvky v poslednom riadku sú nuly, t.j. ak sú riadky matice P po rade vektory $\vec{\alpha}_1, \dots, \vec{\alpha}_n$, tak

$$\left(\begin{array}{c|c} & c_1 \\ \mathrm{B} & \vdots \\ & c_{n-1} \\ \hline 0\dots0 & a_n \end{array}\right) = A' = PAP^T$$

Keďže riadky matice P sú ortonormálne vektory, platí $P^T = P^{-1}$ (matica P je ortogonálna).

Celý dôkaz teraz urobíme indukciou. Štart indukcie — matica A je 1×1, teda A je horná trojuholníková a nemáme čo dokazovať.

Nech to teraz platí pre všetky matice B typu $(n-1) \times (n-1)$. Urobme vyššie uvedenú úvahu. Keďže pre charakteristické polynómy matíc A a B platí vzťah $ch_A(x) = (x - a_n)ch_B(x)$ (lebo matice A, A' sú podobné), každá vlastná hodnota matice B je vlastná hodnota matice A a preto všetky vlastné hodnoty matice B ležia v poli $\mathbb{R}$. Matica B je podľa indukčného predpokladu ortogonálne podobná hornej trojuholníkovej matici, označme ju T'. Teda existuje ortogonálna matica Q typu $(n-1) \times (n-1)$ taká, že $QBQ^T = T'$. Potom

$$\left(\begin{array}{c|c} Q & \begin{matrix}0\\ \vdots\\ 0\end{matrix} \\ \hline 0 \dots 0 & 1 \end{array}\right) \cdot \left(\begin{array}{c|c} B & \begin{matrix}c_1\\ \vdots\\ c_{n-1}\end{matrix} \\ \hline 0 \dots 0 & a_n \end{array}\right) \cdot \left(\begin{array}{c|c} Q^T & \begin{matrix}0\\ \vdots\\ 0\end{matrix} \\ \hline 0 \dots 0 & 1 \end{array}\right) =$$

$$= \left(\begin{array}{c|c} QBQ^T & Q\gamma^T \\ \hline 0 \dots 0 & a_n \end{array}\right) = \left(\begin{array}{c|c} T' & Q\gamma^T \\ \hline 0 \dots 0 & a_n \end{array}\right),$$

kde $\gamma = (c_1, \dots, c_{n-1})$. Posledná matica je ale vďaka indukčnému predpokladu horná trojuholníková matica. Keďže je matica

$$Q' = \left(\begin{array}{c|c} Q & \begin{matrix}0\\ \vdots\\ 0\end{matrix} \\ \hline 0 \dots 0 & 1 \end{array}\right)$$

ortogonálna (overte!), je tým je ukončený dôkaz indukčného kroku. □

Ortogonálna podobnosť zachováva pojmy ako symetričnosť, kososymetričnosť a ortogonálnosť, t.j. ak je A symetrická (kososymetrická, ortogonálna) a P je ortogonálna matica, potom PAP^T je tiež symetrická (kososymetrická, ortogonálna). Ak je A symetrická, a T je horná trojuholníková ortogonálne podobná s A, tak T je tiež symetrická, t.j. diagonálna. Ak by sme vedeli, že reálna symetrická matica A je ortogonálne podobná hornej trojuholníkovej - t.j. vďaka Schurovej vete sa stačí presvedčiť, že reálna symetrická matica má vždy všetky vlastné čísla reálne - vedeli by sme, že je (dokonca ortogonálne) podobná diagonálnej matici.

{sec:SKU:veta_VlastneCis

Veta 9.2.10. *Nech $A = ||a_{ij}||$ je štvorcová symetrická matica typu $n \times n$ nad poľom $\mathbb{R}$, potom všetky vlastné čísla matice A sú z poľa $\mathbb{R}$.*

Dôkaz. Podľa predpokladu je polynóm $ch_A(x)$ polynóm s reálnymi koeficientami, t.j. jeho korene sú buď reálne alebo komplexné čísla. Budeme predpokladať, že sú to komplexné čísla $a + bi$ $(a, b \in \mathbb{R})$ a dokážeme, že pre symetrickú maticu A je vždy $b = 0$, t.j. je to reálne číslo.

Keďže $A = ||a_{ij}||$ je matica nad $\mathbb{R}$, určite je to aj matica nad poľom komplexných čísiel $\mathbb{C}$. Ak uvažujeme o jej (možno komplexnom) vlastnom čísle $z = a + bi$, musí k nemu prislúchať (možno komplexný) vlastný vektor $(z_1, \dots, z_n) = (a_1 + b_1 i, \dots, a_n + b_n i) = (a_1, \dots, a_n) + (b_1, \dots, b_n)i$ $(a_i, b_i \in R)$, t.j. platí $\vec{z}A = (a + bi)\vec{z}$. Ak položíme $\vec{\alpha} = (a_1, \dots, a_n)$ a $\vec{\beta} = (b_1, \dots, b_n)$, tak vďaka tomu, že A je reálna matica môžeme túto rovnosť napísať ako

$$(\vec{\alpha} + \vec{\beta}i)A = (a + bi)(\vec{\alpha} + \vec{\beta}i) = (a\vec{\alpha} - b\vec{\beta}) + (a\vec{\beta} + b\vec{\alpha})i$$

Ale platí aj $(\vec{\alpha}+\vec{\beta}i)A = (\vec{\alpha}A)+(\vec{\beta}A)i$, kde vektory $\vec{\alpha}A$, $\vec{\beta}A$ sú reálne vektory, aspoň jeden z nich nie je nulový. Porovnaním reálnych a imaginárnych častí dostaneme

$$\begin{aligned}\vec{\alpha}A &= a\vec{\alpha}-b\vec{\beta}\\ \vec{\beta}A &= a\vec{\beta}+b\vec{\alpha}\end{aligned}$$

Pozrime sa teraz na skalárny súčin $\langle\vec{\alpha}A,\vec{\beta}\rangle$. Jedna z možností, ako počítať tento skalárny súčin je použiť maticové násobenie, presnejšie pre štandardný skalárny súčin platí $\langle\vec{x},\vec{y}\rangle = \vec{x}\vec{y}^T$. Použitím tohoto vzorca dostaneme

$$\langle\vec{\alpha}A,\vec{\beta}\rangle = \vec{\alpha}A\vec{\beta}^T = \vec{\alpha}A^T\vec{\beta}^T = \vec{\alpha}(\vec{\beta}A)^T = \langle\vec{\alpha},\vec{\beta}A\rangle$$

Druhá rovnosť je dôsledok symetrie matice A. Ale použitím vzorcov pre $\vec{\alpha}A$ a $\vec{\beta}A$, ktoré sme získali vyššie dostaneme

$$\langle\vec{\alpha}A,\vec{\beta}\rangle = \langle a\vec{\alpha}-b\vec{\beta},\vec{\beta}\rangle = a\langle\vec{\alpha},\vec{\beta}\rangle - b\langle\vec{\beta},\vec{\beta}\rangle$$

a

$$\langle\vec{\alpha},\vec{\beta}A\rangle = \langle\vec{\alpha},a\vec{\beta}+b\vec{\alpha}\rangle = a\langle\vec{\alpha},\vec{\beta}\rangle + b\langle\vec{\alpha},\vec{\alpha}\rangle$$

čiže $a\langle\vec{\alpha},\vec{\beta}\rangle - b\langle\vec{\beta},\vec{\beta}\rangle = a\langle\vec{\alpha},\vec{\beta}\rangle + b\langle\vec{\alpha},\vec{\alpha}\rangle$, a teda

$$\begin{aligned}-b\langle\vec{\beta},\vec{\beta}\rangle &= b\langle\vec{\alpha},\vec{\alpha}\rangle,\\ b(\langle\vec{\beta},\vec{\beta}\rangle + \langle\vec{\alpha},\vec{\alpha}\rangle) &= 0\end{aligned}$$

Keďže aspoň jeden z vektorov $\vec{\alpha}$, $\vec{\beta}$ je nenulový, platí $\langle\vec{\beta},\vec{\beta}\rangle + \langle\vec{\alpha},\vec{\alpha}\rangle > 0$. Preto z poslednej rovnosti vyplýva $b=0$, a teda vlastné číslo $z = a+0i$ je reálne číslo. □

Na základe Schurovej vety teda dostávame tvrdenie, ktoré je známe ako „Veta o hlavných osiach“:

KU:veta_OHlavnychOsiach}

Veta 9.2.11 (o hlavných osiach)**.** *Nech $A = ||a_{ij}||$ je štvorcová symetrická matica typu $n\times n$ nad poľom $\mathbb{R}$, potom A je ortogonálne podobná s diagonálnou maticou.*

Pri hľadaní príslušnej ortogonálnej matice prechodu je užitočné vedieť nasledujúci fakt

eta_VlastneVektoryKolme}

Veta 9.2.12. *Nech $A = ||a_{ij}||$ je štvorcová symetrická matica typu $n\times n$ nad poľom $\mathbb{R}$, nech $a\neq b$ sú dve vlastné čísla matice A a nech $\vec{\alpha}$ je vlastný vektor prislúchajúci ku vlastnému číslu a, podobne $\vec{\beta}$ je vlastný vektor prislúchajúci ku vlastnému číslu b. Potom $\vec{\alpha}\perp\vec{\beta}$ (t.j. $\vec{\alpha}$ a $\vec{\beta}$ sú na seba kolmé v zmysle štandardného skalárneho súčinu).*

Dôkaz. Jedno z čísiel a,b je nenulové, nech je to a. Vypočítajme hodnotu $a\langle\vec{\alpha},\vec{\beta}\rangle$:

$$a\langle\vec{\alpha},\vec{\beta}\rangle = \langle a\vec{\alpha},\beta\rangle = \langle\vec{\alpha}A,\beta\rangle = \vec{\alpha}A\vec{\beta}^T = \vec{\alpha}A^T\vec{\beta}^T = \vec{\alpha}(\vec{\beta}A)^T = \langle\vec{\alpha},\vec{\beta}A\rangle = \langle\vec{\alpha},b\vec{\beta}\rangle = b\langle\vec{\alpha},\vec{\beta}\rangle$$

čiže $(a-b)\langle\vec{\alpha},\vec{\beta}\rangle = 0$ a pretože $a-b\neq 0$, je skalárny súčin $\langle\vec{\alpha},\vec{\beta}\rangle = 0$, t.j. $\vec{\alpha}\perp\vec{\beta}$. □

Ilustrujme si použitie tejto vety na konkrétnom príklade.

Príklad 9.2.13.

$$A = \begin{pmatrix} 1 & 4 & -2\\ 4 & 1 & -2\\ -2 & -2 & -2\end{pmatrix}$$

Ak existujú, nájdite ortogonálnu maticu P a diagonálnu maticu D tak, aby $D = PAP^T$.

Charakteristický polynóm je $-(x+3)^2(x-6)$, čiže vlastné hodnoty sú -3 a 6.

Riešením sústav s maticami $(A+3I)^T$, resp. $(A-6I)^T$ dostaneme vlastné vektory. Dôležité je vlastné vektory normalizovať, prípadne ak je niektorá vlastné hodnota viacnásobná, tak aj z nich urobiť ortonormálnu bázu. (Aby sme dostali ortogonálnu maticu.) Vďaka predchádzajúcej vete máme automaticky zabezpečené, že vlastné vektory pre rôzne vlastné hodnoty budú na seba kolmé.

Vlastné vektory prislúchajúce k -3 sú $(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0)$, $(\frac{1}{3\sqrt{2}}, \frac{1}{3\sqrt{2}}, \frac{4}{3\sqrt{2}})$. Vlastný vektor prislúchajúci k 6 je $(\frac{2}{3}, \frac{2}{3}, -\frac{1}{3})$. Keď tieto vektory dáme do stĺpcov dostaneme hľadanú maticu P.

$$\begin{pmatrix} 1 & 4 & -2 \\ 4 & 1 & -2 \\ -2 & -2 & -2 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{3\sqrt{2}} & \frac{2}{3} \\ -\frac{1}{\sqrt{2}} & \frac{1}{3\sqrt{2}} & \frac{2}{3} \\ 0 & \frac{4}{3\sqrt{2}} & -\frac{1}{3} \end{pmatrix} \begin{pmatrix} -3 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 6 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{3\sqrt{2}} & \frac{1}{3\sqrt{2}} & \frac{4}{3\sqrt{2}} \\ \frac{2}{3} & \frac{2}{3} & -\frac{1}{3} \end{pmatrix}$$

Veta o hlavných osiach nám poskytuje rozklad symetrickej matice na matice projekcie. Ortogonálna podobnosť matice A s diagonálnou maticou nám totiž hovorí, že existuje ortogonálna matica P taká, že

$$A = P^T D P,$$

pričom vieme, že na diagonále matice D sú vlastné čísla $d_1, \ldots, d_n$ matice A a riadky matice P sú vlastné vektory matice A. Predchádzajúcu rovnosť potom môžeme upraviť na tvar

$$A = \begin{pmatrix} \vec{\alpha}_1^T & \ldots & \vec{\alpha}_n^T \end{pmatrix} \begin{pmatrix} d_1 & 0 & \ldots & 0 \\ 0 & d_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & d_n \end{pmatrix} \begin{pmatrix} \vec{\alpha}_1 \\ \vdots \\ \vec{\alpha}_n \end{pmatrix} = \begin{pmatrix} \vec{\alpha}_1^T & \ldots & \vec{\alpha}_n^T \end{pmatrix} \begin{pmatrix} d_1\vec{\alpha}_1 \\ \vdots \\ d_n\vec{\alpha}_n \end{pmatrix}$$

$$A = d_1\vec{\alpha}_1^T\vec{\alpha}_1 + d_2\vec{\alpha}_2^T\vec{\alpha}_2 + \cdots + d_n\vec{\alpha}_n^T\vec{\alpha}_n \tag{9.5}$$ {diagon:EQSPEKROZKLAD}

Predchádzajúci zápis sa niekedy zvykne nazývať *spektrálny rozklad* matice A.

Všimnime si, že maticu A sme rozložili na súčet násobkov ortogonálnych projekcií do smerov vlastných vektorov. (Matica $\vec{\alpha}_i^T\vec{\alpha}_i$ je presne matica ortogonálnej projekcie na podpriestor $[\vec{\alpha}_i]$ – pozri úlohu 7.2.13).

Môžeme si všimnúť, že matice tvaru $A = \vec{\alpha}^T\vec{\alpha}$ majú niektoré pekné vlastnosti. Očividne platí $A^T = A$, čiže takáto matica je symetrická. Vďaka tomu, že matica P je ortogonálna, sú vlastné vektory $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ ortonormálne. Špeciálne vďaka tomu, že majú veľkosť 1, dostaneme $AA = \vec{\alpha}^T(\vec{\alpha}\vec{\alpha}^T)\vec{\alpha} = \vec{\alpha}^T.1.\vec{\alpha} = \vec{\alpha}^T\vec{\alpha} = A$. Matice (a im zodpovedajúce lineárne zobrazenia), pre ktoré platí $A^2 = A$ sa zvyknú nazývať *projekcie*.

9.2.3 Cayley-Hamiltonova veta

{diagon:SSECTCAYLEYHAM}

Ak by sme chceli teóriu načatú v tomto texte chceli študovať serióznejšie, potrebovali by sme tzv. Caley-Hamiltonovu vetu, my ju teraz uvedieme len ako malé doplnenie problematiky a možno ako zaujímavosť.

{diagon:VTCAYLEYHAM}

Veta 9.2.14 (Cayley-Hamilton)**.** *Nech A je štvorcová matica nad poľom F. Potom A je koreňom svojho charakteristického polynómu, presnejšie ak je $ch_A(x) = (-1)^n x^n + c_{n-1}x^{n-1} + \cdots + c_0$ tak $(-1)^n A^n + c_{n-1}A^{n-1} + \cdots + c_0 I = \|0\|_{n\times n}$.*

Dôkaz. Vzhľadom na to, že charakteristický polynóm je determinant, je vhodné pripomenúť jednu zaujímavú vlastnosť - v prvom semestri sme ju používali pre výpočet inverznej matice ku regulárnej matici. Teraz ju uvedieme v trochu všeobecnejšej formulácii, ktorú využijeme v dôkaze: majme štvorcovú maticu A typu $n \times n$. Znakom A_{ij} $(1 \leq i, j \leq n)$ označme determinant matice, ktorá z A vznikne vynechaním i-teho riadku a j-teho stĺpca vynásobený

číslom $(-1)^{i+j}$ (tzv. algebraický doplnok prvku a_{ij} matice A). Známy vzorec na výpočet inverznej matice ku regulárnej matici A (veta 6.5.1) hovorí, že

$$A^{-1} = \begin{pmatrix} \frac{A_{11}}{|A|} & \frac{A_{21}}{|A|} & \dots & \frac{A_{n1}}{|A|} \\ \frac{A_{12}}{|A|} & \frac{A_{22}}{|A|} & \dots & \frac{A_{n2}}{|A|} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{A_{1n}}{|A|} & \frac{A_{2n}}{|A|} & \dots & \frac{A_{nn}}{|A|} \end{pmatrix}$$

čo je špeciálny prípad nasledujúceho vzorca:

$$A \cdot \begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{pmatrix} = |A| \cdot \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} = |A| \cdot I,$$

ktorý platí aj v prípade, že je $|A| = 0$. (A dá sa dokázať podobne ako veta 6.5.1.)

Matica

$$\begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{pmatrix}$$

sa označuje znakom $\mathrm{adj}(A)$, t.j. platí $A \cdot \mathrm{adj}(A) = |A| \cdot I$.

Naviac je dôležité to, že pojem determinantu môžeme rovnako ako pre matice nad (nejakým) poľom F zaviesť pre matice nad ľubovoľným komutatívnym okruhom a v prípade, že je to okruh s 1, uvedený vzorec bude platiť (i keď možno napr. nemôžeme maticu upraviť na redukovaný trojuholníkový tvar a na základe známych viet o tom (tieto vety ostanú v platnosti aj pre matice nad komutatívnym okruhom), ako sa správa determinant pri elementárnych riadkových operáciách potom počítať determinant - to je vo všeobecnosti výsada matíc nad poliami).

Teraz môžeme pristúpiť ku samotnému dôkazu. Nech $B = A - xI$. Potom B_{ij} ako (až na znamienko) determinant „podmatice“ matice B, ktorá vznikla vynechaním jedného riadku a jedného stĺpca je polynóm stupňa menej ako n, t.j. nanajvýš stupňa $n-1$ v premennej x. T.j. matica $\mathrm{adj}(B)$ je matica s prvkami z $F[x]$, čo je komutatívny okruh s 1.

Vzhľadom na uvedené možné stupne polynómov existujú také matice $C_{n-1}, \dots C_1, C_0$ - všetko matice $n \times n$ nad poľom F (t.j. ich prvky už sú konštanty, nie polynómy), že

$$\mathrm{adj}(A - xI) = \mathrm{adj}(B) = C_{n-1}x^{n-1} + \dots + C_1 x + C_0$$

Označme si $ch_A(x) = |B| = (-1)^n x^n + c_{n-1}x^{n-1} + \dots + c_1 x + c_0$

$$|B| \cdot I = (-1)^n I x^n + c_{n-1} \cdot I x^{n-1} + \dots + c_1 \cdot I x + c_0 I = B \cdot \mathrm{adj}(B) = (A - xI) \cdot (C_{n-1}x^{n-1} + \dots + C_1 x + C_0),$$

odkiaľ porovnaním „koeficientov“ pri rovnakých mocninách x dostaneme

$$\begin{aligned} c_0 I &= AC_0 \\ c_1 I &= AC_1 - C_0 \\ c_2 I &= AC_2 - C_1 \\ &\dots \\ c_{n-1} I &= AC_{n-1} - C_{n-2} \\ \text{a nakoniec} & \\ (-1)^n I &= -C_{n-1} \end{aligned}$$

Vynásobme zľava teraz postupne prvú rovnicu maticou I, druhú rovnicu maticou A, tretiu rovnicu maticou $A^2, \dots$, a poslednú, t.j. rovnicu s poradovým číslom $n+1$ maticou A^n, dostaneme

$$\begin{aligned} c_0 I &= AC_0 \\ c_1 A &= A^2 C_1 - AC_0 \\ c_2 A^2 &= A^3 C_2 - A^2 C_1 \\ &\dots \\ c_{n-1} A^{n-1} &= A^n C_{n-1} - A^{n-1} C_{n-2} \\ \text{a nakoniec} & \\ (-1)^n A^n &= -A^n C_{n-1} \end{aligned}$$

Po sčítaní ľavých a pravých strán týchto rovníc dostaneme

$$\begin{aligned} (-1)^n A^n + c_{n-1} A^{n-1} + \dots + c_1 A + c_o I &= \\ -A^n C_{n-1} + (A^n C_{n-1} - A^{n-1} C_{n-2}) + \dots + (A^3 C_2 - A^2 C_1) + (A^2 C_1 - AC_0) + AC_0 &= \|0\|_{n\times n} \end{aligned}$$

t.j. $ch_A(A) = \|0\|$. □

Cvičenia

Úloha 9.2.1. Nájdite vlastné hodnoty a vlastné vektory daných matíc nad poľom $\mathbb{C}$:
a) $\left(\begin{smallmatrix} 2 & 4 \\ 5 & 3 \end{smallmatrix}\right)$
b) $\left(\begin{smallmatrix} 3 & 1 \\ 1 & 3 \end{smallmatrix}\right)$
c) $\left(\begin{smallmatrix} 1 & 2 \\ 2 & -2 \end{smallmatrix}\right)$
d) $\left(\begin{smallmatrix} -1 & 2i \\ -2i & 2 \end{smallmatrix}\right)$
e) $\left(\begin{smallmatrix} 2 & 1 \\ -1 & 0 \end{smallmatrix}\right)$
f) $\left(\begin{smallmatrix} 4 & 1 \\ 3 & 2 \end{smallmatrix}\right)$
Ak taká matica existuje, nájdite regulárnu maticu P s vlastnosťou, že PAP^{-1} je diagonálna.

Úloha 9.2.2. Ukážte, že vlastné vektory matice A typu $n \times n$ prislúchajúce k danej vlastnej hodnote c spolu s nulovým vektorom tvoria podpriestor priestoru F^n.

Úloha 9.2.3. Ako vyzerá matica A zodpovedajúca otočeniu v rovine okolo počiatku súradnicovej sústavy o nenulový uhol φ? Nájdite jej vlastné hodnoty a vlastné vektory v $\mathbb{C}$? Ako možno geometricky interpretovať fakt, že táto matica nemá reálne vlastné vektory?

Úloha 9.2.4. Ukážte, že pre $c \in \mathbb{R} \setminus \{0\}$ matica $\left(\begin{smallmatrix} 1 & c \\ 0 & 1 \end{smallmatrix}\right)$ nie je podobná s diagonálnou maticou. Aká je geometrická interpretácia tohoto výsledku?

Úloha 9.2.5. Ukážte, že ak k je smernica vlastného vektora matice A typu 2×2, tak k spĺňa kvadratickú rovnicu $a_{21}k^2 + (a_{11} - a_{22})k - a_{12} = 0$.

Úloha 9.2.6. Ak A, B sú regulárne matice, tak AB a BA majú rovnaké vlastné hodnoty.

Úloha 9.2.7. Dokážte: Štvorcová matica A je regulárna práve vtedy, keď 0 nie je vlastné číslo matice A.

Ak A je regulárna, tak c je vlastné číslo matice A práve vtedy, keď c^{-1} je vlastné číslo matice A^{-1}.

Úloha 9.2.8. Ak A je idempotentná matica, čiže $A^2 = A$, tak jej vlastné hodnoty môžu byť jedine 0 alebo 1.

Úloha 9.2.9. Nech A je štvorcová matica. Ukážte, že λ je vlastné číslo matice A práve vtedy, keď $\lambda + a$ je vlastné číslo matice $A + aI$.

Úloha 9.2.10. Nájdite (ak taká matica existuje) maticu P takú, že $PAP^{-1} = D$ je diagonálna matica.

a) $\begin{pmatrix} -2 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

b) $\begin{pmatrix} 1 & 2 & 1 \\ 6 & -1 & 0 \\ -1 & -2 & -1 \end{pmatrix}$

c) $\begin{pmatrix} -1 & -1 & 1 \\ 0 & -2 & 1 \\ 0 & 0 & -1 \end{pmatrix}$

d) $\begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix}$

Úloha 9.2.11. Nájdite (ak taká matica existuje) ortogonálnu maticu P takú, že $PAP^T = D$ je diagonálna matica.

a) $\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$

b) $\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 5 \\ 0 & 5 & 1 \end{pmatrix}$

c) $\begin{pmatrix} -1 & 2 & 0 \\ 2 & -1 & 0 \\ 0 & 0 & 4 \end{pmatrix}$

d) $\begin{pmatrix} -2 & 3 & 0 \\ 3 & -2 & 0 \\ 0 & 0 & 7 \end{pmatrix}$

e) $\begin{pmatrix} -1 & 2 & 0 & 0 \\ 2 & -1 & 0 & 0 \\ 0 & 0 & -1 & 4 \\ 0 & 0 & 4 & -1 \end{pmatrix}$

f) $\begin{pmatrix} -1 & 2 & 0 & 0 \\ 2 & -1 & 0 & 0 \\ 0 & 0 & 2 & 3 \\ 0 & 0 & 3 & 2 \end{pmatrix}$

g) $\begin{pmatrix} 3 & 2 & 2 \\ 2 & 4 & 1 \\ 2 & 1 & 4 \end{pmatrix}$

h) $\begin{pmatrix} 0 & 1 & 2 \\ 1 & 0 & 2 \\ 2 & 2 & -1 \end{pmatrix}$

i) $\begin{pmatrix} 2 & 1 & -2 \\ 1 & 2 & -2 \\ -2 & -2 & 5 \end{pmatrix}$

j) $\begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix}$

k) $\begin{pmatrix} 2 & 1 & 2 \\ 1 & 2 & 2 \\ 2 & 2 & 5 \end{pmatrix}$

Úloha 9.2.12. Sú matice $A = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 2 & \dots & 0 & 0 \\ 0 & 0 & \dots & n-1 & 0 \\ 0 & 0 & \dots & 0 & n \end{pmatrix}$ a $B = \begin{pmatrix} n & 0 & \dots & 0 & 0 \\ 0 & n-1 & \dots & 0 & 0 \\ 0 & 0 & \dots & 2 & 0 \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix}$ podobné? Ak áno, nájdite maticu P takú, že $B = PAP^{-1}$.

Úloha 9.2.13. Pre dané matice vyrátajte charakteristické polynómy $ch_A(x)$, $ch_B(x)$. Vyrátajte aj stopu a determinant týchto matíc a porovnajte ich s príslušnými koeficientami charakteristického polynómu. Sú tieto matice podobné?
$A = \begin{pmatrix} 1 & 4 & -2 \\ 4 & 1 & -2 \\ -2 & -2 & -2 \end{pmatrix}$; $B = \begin{pmatrix} 1 & 5 & -2 \\ 2 & 1 & -4 \\ -2 & -4 & -2 \end{pmatrix}$.

{diagoncvic:ULOROVNCHAR}

Úloha 9.2.14. Pre dané matice vyrátajte charakteristické polynómy $ch_A(x)$, $ch_B(x)$. Sú dané matice podobné?
$A = \begin{pmatrix} 1 & 4 & -2 \\ 4 & 1 & -2 \\ -2 & -2 & -2 \end{pmatrix}$; $B = \begin{pmatrix} 1 & 4 & -2 \\ 1 & 1 & -8 \\ -2 & -2 & -2 \end{pmatrix}$

Úloha 9.2.15*. Nájdite symetrickú a ortogonálnu maticu P takú, že PAP^{-1} je diagonálna matica ak

$$A = \begin{pmatrix} a^2 & ab & ab & b^2 \\ ab & a^2 & b^2 & ab \\ ab & b^2 & a^2 & ab \\ b^2 & ab & ab & a^2 \end{pmatrix}$$

Úloha 9.2.16*. Nech V je vektorový priestor všetkých matíc typu $n \times n$ nad $\mathbb{R}$, nech $A \in V$ nech $T\colon V \to V$ je definované ako $T(X) = AX$. Nájdite charakteristický polynóm matice zobrazenia T a ukážte, že ak matica A je podobná s diagonálnou maticou, tak aj T je podobná s diagonálnou maticou. (Poznámka: Matica zobrazenia T síce závisí od voľby bázy priestoru V, nie je však ťažké si uvedomiť, že charakteristický polynóm ani diagonalizovateľnosť matice sa nemenia prechodom k inej báze, čiže od voľby bázy nezávisia.)

9.3 Krivky druhého rádu

Ako aplikáciu vety o hlavných osiach si popíšeme ako vyzerajú množiny bodov v rovine popísané polynómom 2 premenných stupňa 2. Z toho, čo si o nich povieme, by snáď mohlo byť zrejmé, prečo sa táto veta nazýva „veta o hlavných osiach“.

9.3.1 Ortogonálne matice 2×2

Najprv sa na chvíľu zastavme pri ortogonálnych maticiach. Keďže chceme skúmať situáciu v $\mathbb{R}^2$, budú nás hlavne zaujímať reálne symetrické matice rozmeru 2×2.

Pripomeňme, že ortogonálna matica je štvorcová matica O, ktorá spĺňa podmienku $OO^T = I$. Táto podmienka znamená, že riadky tejto matice tvoria ortonormálnu bázu v F^n.

Ekvivalentne môžeme definovať ortogonálne matice podmienkou $O^TO = I$, čo znamená, že ortonormálnu bázu tvoria stĺpce. Iná ekvivalentná formulácia je, že transponovaná matica k O je súčasne k tejto matici inverzná, t.j. $O^T = O^{-1}$. Takisto priamo z definície vidno, že ortogonálna matica musí byť regulárna.

Ľahko sa dá ukázať, že ortogonálne matice daného rozmeru tvoria vzhľadom na násobenie matíc grupu (úloha 9.3.1). Všimnime si ešte jednu vlastnosť ortogonálnych matíc. Uvažujme štandardný skalárny súčin na $\mathbb{R}^n$. Pre ľubovoľné 2 vektory $\vec{\alpha}, \vec{\beta} \in \mathbb{R}^n$ a ortogonálnu maticu O dostaneme

$$\langle \vec{\alpha}O, \vec{\beta}O \rangle = \vec{\alpha}O(\vec{\beta}O)^T = \vec{\alpha}OO^T\vec{\beta}^T = \vec{\alpha}\vec{\beta}^T = \langle \vec{\alpha}, \vec{\beta} \rangle.$$

Zistili sme, že lineárne zobrazenie zodpovedajúce matici O zachováva skalárny súčin. Z toho špeciálne vyplýva, že zachováva veľkosť a uhly vektorov.

Teraz sa pozrime len na matice rozmeru 2×2. Ak reálna matica $O = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ je ortogonálna, jej prvky musia spĺňať

$$\begin{aligned} a^2 + b^2 &= 1 \\ c^2 + d^2 &= 1 \\ ac + bd &= 0 \end{aligned}$$

Ak $a = 0$, dostaneme z prvej rovnice $b = \pm 1$ a z tretej rovnice $d = 0$. Z toho potom vyplýva $c = \pm 1$.

Ak $b = 0$, dostaneme $a = \pm 1$, $c = 0$ a $d = \pm 1$.

Ak $ab \neq 0$ môžeme poslednú rovnicu vydeliť číslom ab a dostaneme $\frac{c}{b} + \frac{d}{a} = 0$, čiže $\frac{c}{b} = -\frac{d}{a}$. Ak označíme $\frac{c}{b} = -\frac{d}{a} =: k$, máme $c = bk$ a $d = -ak$. Po dosadení do druhej rovnice máme

$$(a^2 + b^2)k^2 = 1.$$

Spolu s prvou rovnicou to znamená, že $k^2 = 1$, a teda $k = \pm 1$.

Ľubovoľné riešenie prvej rovnice je tvaru $a = \cos\varphi$, $b = \sin\varphi$. V závislosti od voľby k dostaneme buď $c = \sin\varphi$ a $d = -\cos\varphi$ alebo $c = -\sin\varphi$ a $d = \cos\varphi$. Všimnime si, že tieto riešenia zahŕňajú aj prípad $a = 0$ a $b = 0$, ktoré sme riešili zvlášť (pre $\varphi = 0, \frac{\pi}{2}, \pi, \frac{3}{2}\pi$).

Zistili sme teda, že všetky ortogonálne matice 2×2 sú tvaru

$$\begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix} \qquad \begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix}$$

pre $\varphi \in \langle 0, 2\pi)$.

(Tieto riešenia možno ľahko nájsť aj na základe geometrického významu rovníc, ktoré sme používali. Hľadali sme vlastne vektory (a, b) a (c, d), ktoré sú navzájom kolmé a majú veľkosť 1. Skúste si nakresliť obrázok.)

Prvá z uvedených matíc je presne matica otočenia okolo počiatku súradnicovej sústavy o uhol φ proti smeru pohybu hodinových ručičiek (stačí si všimnúť, že pri tomto lineárnom zobrazení sa vektor $(1, 0)$ zobrazí na $(\cos\varphi, \sin\varphi)$ a vektor $(0, 1)$ na $(-\sin\varphi, \cos\varphi)$).

Z rovnosti

$$\begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}$$

vidíme, že ostatné ortogonálne matice zodpovedajú zobrazeniam, ktoré sú zložením osovej súmernosti podľa osi x a otočenia o uhol φ.

Záver: Lineárne zobrazenia zodpovedajúce ortogonálnym maticiam sú práve zobrazenia, ktoré vzniknú zložením osových súmerností (podľa osi prechádzajúcej počiatkom) a otočení (okolo počiatku).

Niečo podobné platí aj vo všeobecnosti – každá reálna ortogonálna matica sa dá napísať ako súčin matice nejakej rotácie okolo počiatku a matice, ktorá zodpovedá lineárnemu zobrazeniu takému, že jednotkové vektory sa pri ňom nejakým spôsobom povymieňajú a niektoré z nich sa zmenia na opačné.

9.3.2 Popis kriviek druhého rádu

Teraz sa už skúsme dostať k otázke, ktorou sme sa chceli zaoberať pôvodne – preskúmať krivky v rovine popísané rovnicami druhého stupňa. Presnejšie, ak

{krivky2radu:EQF} $$f(x_1, x_2) = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 + 2a_1x_1 + 2a_2x_2 + d, \tag{9.6}$$

kde aspoň jedno z čísel a_{11}, a_{12}, a_{22} je nenulové (t.j. člen druhého stupňa v tejto funkcii je nenulový), tak nás zaujíma ako vyzerá množina bodov

$$K = \{(x_1, x_2) \in \mathbb{R}^2; f(x_1, x_2) = 0\}.$$

Ako sa dá uhádnuť z označenia použitého v (9.6), tento problém bude nejako súvisieť s kvadratickými formami.

Ako sa dá uhádnuť z označenia použitého v (9.6), tento problém bude nejako súvisieť s kvadratickými formami. Ak si všímame len kvadratickú časť predpisu $f(x_1, x_2)$, vidíme, že ide o kvadratickú formu

$$g(x_1, x_2) = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2.$$

Matica tejto kvadratickej formy $A = \left(\begin{smallmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{smallmatrix}\right)$ je symetrická. Podľa vety o hlavných osiach 9.2.11 teda existuje ortogonálna matica P tak, že $PAP^T = \operatorname{diag}(\lambda_1, \lambda_2)$, kde $\lambda_{1,2}$ sú vlastné čísla matice A. Bez ujmy na všeobecnosti môžeme predpokladať, že P je maticou otočenia okolo počiatku súradnicovej sústavy. (Ak by to nebola matica otočenia, stačí výraz PAP^T zľava aj sprava vynásobiť maticou $\left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$.)

Matica P zodpovedá zmene premenných, ktorá je lineárna, teda v nových súradniciach bude rovnica našej krivky vyzerať

$$f(y_1, y_2) = \lambda_1 y_1^2 + \lambda_2 y_2^2 + 2b_1 y_1 + 2b_2 y_2 + d' = 0.$$

Uvažujme najprv prípad, že $\lambda_1\lambda_2 \neq 0$. Potom môžeme túto rovnicu ďalej upraviť doplnením na štvorec. Zavedieme nové premenné $z_1 = y_1 + \frac{b_1}{\lambda_1}$, $z_2 = y_2 + \frac{b_2}{\lambda_2}$. Dostaneme

$$f(z_1, z_2) = \lambda_1 z_1^2 + \lambda_2 z_2^2 + d'' = 0.$$

Geometricky zmena premenných, ktorú sme urobili, zodpovedá posunutiu o vektor $(-\frac{b_1}{\lambda_1}, -\frac{b_2}{\lambda_2})$.

V závislosti od znamienka koeficientov vystupujúcich v tejto rovnici už z nej vieme vyčítať tvar krivky. V prípade, že $\lambda_1, \lambda_2 > 0$ a $d'' < 0$, ako aj v prípade $\lambda_1, \lambda_2 < 0$ a $d'' > 0$ ide o *elipsu.*

Ak $\lambda_1, \lambda_2 > 0$ a $d'' > 0$ alebo $\lambda_1, \lambda_2 < 0$ a $d'' < 0$, tak uvedené rovnica nemá riešenie.

V prípade, že λ_1 a λ_2 majú rôzne znamienka a $d'' \neq 0$, je to *hyperbola.*

Ak $d'' = 0$ tak ide buď o *jednobodovú množinu* (ak λ_1 a λ_2 majú rovnaké znamienko) alebo o *dvojicu pretínajúcich sa priamok* (ak majú rôzne znamienka).

Zostáva nám vyriešiť prípad, keď niektoré vlastné číslo je nulové. Nech napríklad $\lambda_1 = 0$. V tomto prípade môžeme doplnenie na štvorec použiť len v druhej premennej a dostaneme

$$\lambda_2 z_2^2 + 2b_1 z_1 + d'' = 0.$$

Ak $b_1 \neq 0$, je to *parabola.* Ak $b_1 = 0$, tak v závislosti od znamienka d'' to môže byť prázdna množina (rovnaké znamienko ako λ_2), *priamka* (ak $d'' = 0$) alebo *dvojica rovnobežných priamok.*

Zistili sme teda, že – s výnimkou degenerovaných prípadov – každá krivka vyjadrená rovnicou druhého stupňa bude vhodne posunutá a otočená elipsa, hyperbola alebo parabola. Vlastné hodnoty matice A nám udávajú hlavnú a vedľajšiu poloos týchto kužeľosečiek.

9.3.3 Invarianty kriviek druhého rádu

V tejto časti si ukážeme, ako možno zistiť typ krivky druhého rádu bez toho, aby sme ju museli upravovať na kanonický tvar.

Definícia 9.3.1. *Invariantom* krivky druhého rádu

$$a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 + 2a_1x_1 + 2a_2x_2 + d = 0$$

rozumieme taký algebraický výraz závisiaci od a_{11}, a_{12}, a_{22}, a_1, a_2 a d, ktorý sa nezmení, ak túto krivku vyjadríme v iných súradniciach, ktoré dostaneme posunutím a otočením.

Tvrdenie 9.3.2. *Výrazy* $s = \mathrm{Tr}(A) = a_{11} + a_{22}$, $\delta = |A| = \begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix}$ *a* $\Delta = \begin{vmatrix} a_{11} & a_{12} & a_1 \\ a_{12} & a_{22} & a_2 \\ a_1 & a_2 & d \end{vmatrix}$

sú invariantmi krivky druhého rádu

$$a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 + 2a_1x_1 + 2a_2x_2 + d = 0$$

Dôkaz. Overme najprv, že tieto výrazy sa nezmenia pri posunutí. Položme $x_1 = y_1 + d_1$ a $x_2 = y_2 + d_2$. Dostaneme

$$\begin{aligned} &a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 + 2a_1x_1 + 2a_2x_2 + d = \\ &a_{11}y_1^2 + 2a_{12}y_1y_2 + a_{22}y_2^2 + 2(a_1 + a_{11}d_1 + a_{12}d_2)y_1 + 2(a_2 + a_{22}d_2 + a_{12}d_1)y_2 + a_{11}d_1^2 + \\ &\qquad 2a_{12}d_1d_2 + a_{22}d_2^2 + 2a_1d_1 + 2a_2d_2 + d. \end{aligned}$$

Pre vyjadrenie krivky v nových súradniciach máme $s = a_{11} + a_{22}$, $\delta = |A| = \begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix}$ a

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & a_1 + a_{11}d_1 + a_{12}d_2 \\ a_{12} & a_{22} & a_2 + a_{22}d_2 + a_{12}d_1 \\ a_1 + a_{11}d_1 + a_{12}d_2 & a_2 + a_{22}d_2 + a_{12}d_1 & a_{11}d_1^2 + 2a_{12}d_1d_2 + a_{22}d_2^2 + 2a_1d_1 + 2a_2d_2 + d \end{vmatrix}$$

Stačí si všimnúť, že maticu v determinante Δ môžeme dostať ako

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ d_1 & d_2 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_1 \\ a_{12} & a_{22} & a_2 \\ a_1 & a_2 & d \end{pmatrix} \begin{pmatrix} 1 & 0 & d_1 \\ 0 & 1 & d_2 \\ 0 & 0 & 1 \end{pmatrix}.$$

(Pri hľadaní matíc, ktorými musíme prenásobiť pôvodnú maticu, je môže pomôcť všimnúť si, aké riadkové a stĺpcové operácie sa dajú použiť.) Keďže sme pôvodnú maticu násobili maticami s determinantom 1, determinant sa tým nezmení.

Teraz skúsme to isté overiť pre otočenie o uhol φ, čiže $x_1 = y_1\cos\varphi + y_2\sin\varphi$ a $x_2 = -y_1\sin\varphi + y_2\cos\varphi$. Dostaneme

$$\begin{aligned} &a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 + 2a_1x_1 + 2a_2x_2 + d = \\ &(a_{11}\cos^2\varphi - 2a_{12}\cos\varphi\sin\varphi + a_{22}\sin^2\varphi)y_1^2 + 2(a_{11}\cos\varphi\sin\varphi + a_{12}(\cos^2\varphi - \sin^2\varphi) - a_{22}\cos\varphi\sin\varphi)y_1y_2 + \\ &(a_{11}\sin^2\varphi - 2a_{12}\sin^2\varphi + a_{22}\cos^2\varphi)y_2^2 + 2(a_1\cos\varphi - a_2\sin\varphi)y_1 + 2(a_1\sin\varphi + a_2\cos\varphi) + d \end{aligned}$$

Potom dostaneme
$s = \mathrm{Tr}(A) = (a_{11}\cos^2\varphi - 2a_{12}\cos\varphi\sin\varphi + a_{22}\sin^2\varphi) + (a_{11}\sin^2\varphi - 2a_{12}\sin^2\varphi + a_{22}\cos^2\varphi) =$
$a_{11}(\cos^2\varphi + \sin^2\varphi) + a_{22}(\cos^2\varphi + \sin^2\varphi) = a_{11} + a_{22}$.
$\delta = \begin{vmatrix} a_{11}\cos^2\varphi - 2a_{12}\cos\varphi\sin\varphi + a_{22}\sin^2\varphi & a_{11}\cos\varphi\sin\varphi + a_{12}(\cos^2\varphi - \sin^2\varphi) - a_{22}\cos\varphi\sin\varphi \\ a_{11}\cos\varphi\sin\varphi + a_{12}(\cos^2\varphi - \sin^2\varphi) - a_{22}\cos\varphi\sin\varphi & a_{11}\sin^2\varphi + 2a_{12}\cos\varphi\sin\varphi + a_{22}\cos^2\varphi \end{vmatrix}$
a súčasne

$$\begin{pmatrix} a_{11}\cos^2\varphi - 2a_{12}\cos\varphi\sin\varphi + a_{22}\sin^2\varphi & a_{11}\cos\varphi\sin\varphi + a_{12}(\cos^2\varphi - \sin^2\varphi) - a_{22}\cos\varphi\sin\varphi \\ a_{11}\cos\varphi\sin\varphi + a_{12}(\cos^2\varphi - \sin^2\varphi) - a_{22}\cos\varphi\sin\varphi & a_{11}\sin^2\varphi + 2a_{12}\cos\varphi\sin\varphi + a_{22}\cos^2\varphi \end{pmatrix} =$$
$$\begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix} \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}$$

Keďže sme pôvodnú maticu vynásobili ortogonálnou maticou (a tá má determinant rovný 1), determinant sa nezmení.

Podobne dostaneme

$$\begin{pmatrix} a'_{11} & a'_{12} & a_1\cos\varphi - a_2\sin\varphi \\ a'_{21} & a'_{22} & a_1\sin\varphi + a_2\cos\varphi \\ a_1\cos\varphi - a_2\sin\varphi & a_1\sin\varphi + a_2\cos\varphi & d \end{pmatrix} = \begin{pmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_1 \\ a_{12} & a_{22} & a_2 \\ a_1 & a_2 & d \end{pmatrix} \begin{pmatrix} \cos\varphi & \sin\varphi & 0 \\ -\sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Ani v tomto prípade sa determinant nezmení. □

V predchádzajúcej časti sme rozanalyzovali akú krivku dostaneme na základe znamienok $\lambda_{1,2}$ a d''. Na určenie týchto znamienok nám však budú stačiť aj spomínané invarianty.

Ukázali sme, že rovnicu krivky druhého rádu môžeme upraviť na tvar $\lambda_1 z_1^2 + \lambda_2 z_2^2 + d'' = 0$ (v prípade, že obe vlastné hodnoty sú nenulové). V tomto prípade platí $s = \lambda_1 + \lambda_2$, $\delta = \lambda_1\lambda_2$, $\Delta = \lambda_1\lambda_2 d''$.

V prípade, že $\lambda_1 = 0$ sa rovnica danej krivky dala upraviť na tvar $\lambda_2 z_2^2 + 2b_1 z_1 + d'' = 0$. Hodnoty invariantov sú $s = \lambda_2$, $\delta = 0$ a $\Delta = -b_1^2\lambda_2$.

Uvažujme najprv prípad $\delta \neq 0$, ktorý zodpovedá tomu, že obe vlastné hodnoty sú nenulové. Ak $\delta > 0$ znamená to, že vlastné hodnoty majú rovnaké znamienka, ak $\delta < 0$ tak majú opačné znamienka.

Ak majú vlastné hodnoty rovnaké znamienka, tak dostávame buď prázdnu množinu – ak aj d'' má rovnaké znamienko ako vlastné hodnoty – alebo elipsu, ak má opačné znamienko.

Ak $\Delta = 0$, znamená to, že $d'' = 0$, čiže ide o jednobodovú množinu.

Nech teraz $\delta < 0$, čiže vlastné hodnoty majú rôzne znamienka. Pre $d'' \neq 0$ (čiže $\Delta \neq 0$) dostaneme hyperbolu. Ak je d'' nulové, je to dvojica pretínajúcich sa priamok.

Zostáva nám prípad, že niektorá z vlastných hodnôt je nulová, podobne ako doteraz budeme predpokladať, že je to λ_1. Potom ak b_1 je nenulové, ide o parabolu. To zodpovedá tomu, že $\Delta \neq 0$. V opačnom prípade ide o dvojicu rovnobežných priamok, jedinú priamku alebo prázdnu množinu (v závislosti od toho, či λ_2 a d'' majú rovnaké znamienka).

$\delta > 0$	$\Delta \neq 0$	elipsa alebo prázdna množina
$\delta > 0$	$\Delta = 0$	jediný bod
$\delta < 0$	$\Delta \neq 0$	hyperbola
$\delta < 0$	$\Delta = 0$	pretínajúce sa priamky
$\delta = 0$	$\Delta \neq 0$	parabola
$\delta = 0$	$\Delta = 0$	rovnobežné priamky alebo $\emptyset$

9.3.4 Kužeľosečky

Krivky, ktoré takýmto spôsobom dostaneme sa zvyknú nazývať kužeľosečky. Vieme ich totiž dostať ako prienik kužeľa s vhodnou rovinou. (Dvojicu rovnobežných priamok vieme dostať ako prienik valca s vhodnou rovinou.) Na prvý pohľad vidno, že ak vezmeme kužeľ $z^2 = x^2 + y^2$ a rovinu $ax + by + cz = d$, a ak napríklad koeficient c je nenulový, tak môžeme z druhej rovnice vyjadriť $z = \frac{d - ax - by}{c}$ a dosadiť do prvej, očividne tak dostaneme rovnicu druhého stupňa.

Pokúsme sa vyjadriť kužeľosečku v súradniciach určených danou rovinou. Naša rovina nech je daná rovnicou

$$(x_1, x_2, x_3) = (b_1, b_2, b_3) + t(u_1, u_2, u_3) + s(v_1, v_2, v_3).$$

Vhodné bude predpokladať, že vektory $\vec{u} = (u_1, u_2, u_3)$ a $\vec{v} = (v_1, v_2, v_3)$ sú na seba kolmé. (Takto vyjadríme hľadanú krivku v pravouhlej súradnicovej sústave.) Vyrátajme jej priesečník s kužeľom $x_3^2 = x_1^2 + x_2^2$. Po dosadení za x_1, x_2 a x_3 do rovnice kužeľa dostaneme

$$(b_3 + tu_3 + sv_3)^2 - (b_1 + tu_1 + sv_1)^2 - (b_2 + tu_2 + sv_2)^2 = 0$$

a po úprave

$$(u_3^2 - u_1^2 - u_2^2)t^2 + 2(u_3v_3 - u_1v_1 - u_2v_2)st + (v_3^2 - v_1^2 - v_2^2)s^2 + \\ 2(b_3u_3 - b_1u_1 - b_2u_2)t + 2(b_3v_3 - b_1v_1 - b_2v_2)s + (b_3^2 - b_1^2 - b_1^2) = 0$$

Ak s a t chápeme ako súradnice, vidíme, že ide skutočne o krivku druhého stupňa. Pokúsme sa vyrátať aspoň δ – tento invariant určuje typ krivky.

$$\begin{vmatrix} u_3^2-u_1^2-u_2^2 & u_3v_3-u_1v_1-u_2v_2 \\ u_3v_3-u_1v_1-u_2v_2 & v_3^2-v_1^2-v_2^2 \end{vmatrix} = (u_3^2 - u_1^2 - u_2^2)(v_3^2 - v_1^2 - v_2^2) - (u_3v_3 - u_1v_1 - u_2v_2)^2 = \\ = u_1^2v_1^2 + u_2^2v_2^2 + u_3^2v_3^2 + u_1^2v_2^2 + u_2^2v_1^2 - u_1^2v_3^2 - u_3^2v_1^2 - u_2^2v_3^2 - u_3^2v_2^2 \\ - u_1^2v_1^2 - u_2^2v_2^2 - u_3^2v_3^2 - 2u_1u_2v_1v_2 + 2u_1u_3v_1v_3 + 2u_2u_3v_2v_3 = \\ = u_1^2v_2^2 + u_2^2v_1^2 - 2u_1u_2v_1v_2 - u_1^2v_3^2 - u_3^2v_1^2 + 2u_1u_3v_1v_3 - u_2^2v_3^2 - u_3^2v_2^2 + 2u_2u_3v_2v_3 = \\ = (u_1v_2 - u_2v_1)^2 - (u_1v_3 - u_3v_1)^2 - (u_2v_3 - u_3v_2)^2$$

Ak smerové vektory roviny sú (u_1, u_2, u_3) a (v_1, v_2, v_3), tak jej normálový vektor je

$$(n_1, n_2, n_3) = (u_1, u_2, u_3) \times (v_1, v_2, v_3) = (u_2v_3 - u_3v_2, u_3v_1 - u_1v_3, u_1v_2 - u_2v_1).$$

Dostali sme teda

$$\delta = n_3^2 - n_1^2 - n_2^2.$$

Elipsu dostaneme v prípade, že $\delta > 0$, čo zodpovedá tomu, že normálový vektor smeruje do vnútra uvažovaného kužeľa. Aby sme dostali parabolu, musí platiť $\delta = 0$, t.j. normálový vektor patrí uvažovanému kužeľu. Vzhľadom k tomu, že sme zobrali kužeľ s rovnicou $x_3^2 = x_1^2 + x_2^2$, je to ekvivalentné s tým, že rovina je rovnobežná s niektorou priamkou ležiacou na povrchu kužeľa. Zostávajúci prípad $\delta < 0$ zodpovedá tomu, že normálový vektor smeruje mimo daný kužeľ.

Pokúsme sa pozrieť na to, či by sme niečo podobné dostali aj keby sme rátali s ľubovoľným kužeľom. Kužeľ, ktorý má os orientovanú v smere vektora $(0, 0, 1)$ a vrchol v počiatku súradnicovej sústavy bude mať rovnicu $x_1^2 + x_2^2 - ax_3^2 = 0$, kde $a > 0$ je nejaká kladná konštanta. Túto rovnicu môžeme prepísať v tvare $\vec{x}K\vec{x}^T = 0$ pre

$$K = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -a \end{pmatrix}.$$

Táto matica je diagonálna, čo znamená, že je symetrická a tiež, že ľahko vieme vyrátať inverznú maticu $K^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -\frac{1}{a} \end{pmatrix}$.

Zapíšme rovnicu roviny ako $\vec{x} = \vec{x}_0 + \vec{u}s + \vec{v}t$, pričom predpokladáme, že $\vec{u}$, $\vec{v}$ majú jednotkovú veľkosť a sú na seba kolmé. Po dosadení do rovnice $\vec{x}K\vec{x}^T = 0$ dostaneme $\vec{u}K\vec{u}^T s^2 + (\vec{u}K\vec{v}^T + \vec{v}K\vec{u})^T st + \vec{v}K\vec{v}^T t^2 + \cdots = 0$. (Chceme vyrátať δ, čiže nás zaujímajú len členy, ktoré sú stupňa 2.) Máme teda

$$\delta = \begin{vmatrix} \vec{u}K\vec{u}^T & \vec{u}K\vec{v}^T \\ \vec{v}K\vec{u}^T & \vec{v}K\vec{v}^T \end{vmatrix}.$$

Všimnime si, že $\vec{u}K\vec{v}^T = u_1v_1 + u_2v_2 - au_3v_3 = \vec{v}K\vec{u}^T$ (alebo tiež $\vec{u}K\vec{v}^T = (\vec{u}K\vec{v}^T)^T = \vec{v}K^T\vec{u}^T = \vec{v}K\vec{u}^T$), teda uvedená matica je skutočne symetrická a môžeme použiť to, čo sme odvodili v predošlej časti. (Používame tu výsledky, ktoré platia pre symetrické matice.)

Na vyjadrenie tohoto determinantu nám pomôže, keď si všimneme

$$\begin{pmatrix} \vec{u} \\ \vec{v} \\ \vec{n}K^{-1} \end{pmatrix} K \begin{pmatrix} \vec{u}^T & \vec{v}^T & K^{-1}\vec{n}^T \end{pmatrix} = \begin{pmatrix} \vec{u}K\vec{u}^T & \vec{u}K\vec{v}^T & 0 \\ \vec{v}K\vec{u}^T & \vec{v}K\vec{v}^T & 0 \\ 0 & 0 & \vec{n}K^{-1}\vec{n}^T \end{pmatrix},$$

kde $\vec{n} = \vec{u} \times \vec{v}$ je vektorový súčin vektorov $\vec{u}$ a $\vec{v}$, čiže je to normálový vektor danej roviny.

Ak na obe strany rovnosti použijeme determinant, tak máme:

$$\left| \begin{matrix} \vec{u} \\ \vec{v} \\ \vec{n}K^{-1} \end{matrix} \right|^2 |K| = \delta \cdot \vec{n}K^{-1}\vec{n}^T$$

Súčasne si môžeme všimnúť, že $\vec{n}K^{-1}\vec{n}^T = n_1^2 + n_2^2 - \frac{1}{a}n_3^2$, $|K| = -a$ a

$$\left| \begin{matrix} \vec{u} \\ \vec{v} \\ \vec{n}K^{-1} \end{matrix} \right| = \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ n_1 & n_2 & -\frac{n_3}{a} \end{vmatrix} = n_1^2 + n_2^2 - \frac{1}{a}n_3^2,$$

z čoho už dostaneme

$$\delta = -a\left(n_1^2 + n_2^2 - \frac{1}{a}n_3^2\right),$$

čiže znamienko δ je rovnaké ako znamienko výrazu $n_1^2 + n_2^2 - \frac{1}{a}n_3^2$, ktorý určuje, akú polohu má daná rovina vzhľadom ku kužeľu.

9.3.5 Maximálna a minimálna vlastná hodnota

Ešte sa pozrime na geometrický význam, ktorý má najväčšia a najmenšia vlastná hodnota v prípade, že ide o elipsu. Uvažujme rovnicu už upravenú na diagonálny tvar

$$\lambda_1 x_1^2 + \lambda_2 x_2^2 = d.$$

Nech napríklad $\lambda_1 \geq \lambda_2$. Pre jednoduchosť predpokladajme, že obe vlastné hodnoty sú kladné. (V opačnom prípade by sme ich v predchádzajúcej rovnici nahradili ich absolútnymi hodnotami.)

Skúsme hľadať bod na elipse s najväčšou možnou vzdialenosťou od jej stredu. Máme

$$d = \lambda_1 x_1^2 + \lambda_2 x_2^2 \geq \lambda_2(x_1^2 + x_2^2)$$
$$x_1^2 + x_2^2 \leq \frac{d}{\lambda_2}$$

Vidíme teda, že najväčšia možná hodnota, akú môže výraz $x_1^2 + x_2^2$ nadobúdať, je $\frac{d}{\lambda_2}$. Táto hodnota sa skutočne aj nadobúda pre $x_1 = 0$. Keďže táto rovnica je už v nových súradniciach, znamená to, že bod z najväčšou vzdialenosťou od stredu leží v smere vlastného vektora prislúchajúceho k λ_2.

Všeobecne – vlastná hodnota s najmenšou absolútnou hodnotou a jej vlastný vektor nám určujú najvzdialenejší bod elipsy od stredu, podobne ak vezmeme v absolútnej hodnote najväčšiu vlastnú hodnotu a jej vlastný vektor, nádjeme tak najbližší bod. Vlastné vektory a vlastné čísla nám teda udávajú *hlavné osi* tejto elipsy.

Maximum z absolútnych hodnôt vlastných hodnôt matice A sa zvykne nazývať *spektrálny polomer* matice A. Je dôležitý napríklad z toho dôvodu, že – ako sme už spomínali – na zistenie či nejaký mocninový rad obsahujúci matice konverguje je potrebné zistiť, či všetky vlastné hodnoty sú menšie ako polomer konvergencie. Samozrejme, na to nám stačí skúmať najväčšiu vlastnú hodnotu.

Cvičenia

{krivky2rcvic

Úloha 9.3.1. Dokážte, že ortogonálne matice typu $n \times n$ tvoria s operáciou násobenia matíc grupu.

9.4 Jordanov normálny tvar

Keď sme sa zaoberali kvadratickými formami a kongruenciou matíc, podarilo sa nám ukázať, že každú maticu (kvadratickú formu) možno upraviť na diagonálny tvar. Tento diagonálny tvar bol teda spoločným reprezentantom celej triedy kongruentných matíc.

Podobne aj pri podobnosti matíc sa dá z každej triedy vybrať „pekný“ reprezentant. Ako sme však videli v predchádzajúcej kapitole, nie každá matica je podobná s diagonálnou. Preto v tomto prípade matice reprezentujúce jednotlivé triedy vyzerajú o čosi komplikovanejšie. Ukážeme si jednu z možností výberu takéhoto reprezentanta, ktorá sa nazýva Jordanov normálny tvar matice.

Vetu o Jordanovom normálnom tvare uvedieme bez dôkazu. Jeden možný dôkaz, ktorý využíva teóriu modulov (=isté zovšeobecnenie pojmu vektorového priestoru), sa môžete dozvedieť na predmete Vybrané kapitoly z algebry [G3]. Dôkaz toho tvrdenia môžete nájsť aj v [Z1, Kapitola 20]. Iný dôkaz založený na teórii matíc (využíva okrem iného aj komplexnú verziu Schurrovej vety) možno nájsť v [HJ, Theorem 3.1.11].

Definícia 9.4.1. *Jordanov blok* veľkosti k prislúchajúci číslu λ je matica typu $k \times k$ tvaru

$$J_k(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \dots & \dots & 0 \\ 0 & \lambda & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \lambda & 1 & 0 \\ 0 & \dots & \dots & 0 & \lambda & 1 \\ 0 & \dots & \dots & \dots & 0 & \lambda \end{pmatrix}$$

{jordan:VTJORDAN}

Veta 9.4.2 (Jordanov normálny tvar)**.** *Pre každú maticu A nad $\mathbb{C}$ existuje blokovo diagonálna matica J, ktorej diagonálne bloky sú Jordanove bloky taká, že A je podobná s maticou J.*

$$A \sim \begin{pmatrix} J_{k_1}(\lambda_1) & 0 & \dots & \dots & 0 \\ 0 & J_{k_2}(\lambda_2) & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & J_{k_j}(\lambda_{j-1}) & 0 \\ 0 & \dots & \dots & 0 & J_{k_j}(\lambda_j) \end{pmatrix}$$

Matica J sa nazýva Jordanov normálny tvar *matice A.*

Navyše platí, že matica J je jednoznačne určená až na poradie Jordanových blokov na diagonále.

Ďalej platí, že dve matice A a B sú podobné práve vtedy, keď majú rovnaký Jordanov tvar (až na poradie Jordanových blokov).

Hodnoty $\lambda_1, \dots, \lambda_k$ vystupujúce v predchádzajúcej vete sú vlastné hodnoty matice A. V prípade, že je matica A diagonalizovateľná, všetky Jordanove bloky v jej Jordanovom normálnom tvare majú veľkosť 1.

Vieme, že podobné matice majú rovnakú stopu i determinant. Keďže stopu a determinant matice v Jordanovom normálnom tvare možno vypočítať veľmi jednoducho, vidíme, že stopa

matice je presne súčet jej vlastných hodnôt (vrátane násobnosti[3]) a determinant matice dostaneme ako súčin jej vlastných hodnôt (vrátane násobnosti). To isté sme už vlastne odvodili v dôkaze dôsledku 9.2.7 a poznámke za ním.

Keď už poznáme kanonický tvar z vety 9.4.2, na spôsob, ako ho nájsť, môžeme prísť veľmi podobne ako pri hľadaní diagonálnej matice podobnej s danou maticou.

Pre jednoduchosť sa pozrime najprv na to, čo znamená, že daná matica je podobná s Jordanovým blokom, t.j. existuje taká regulárna matica P, že platí $PAP^{-1} = J_n(\lambda)$. Predchádzajúcu rovnosť môžeme prepísať ako

$$PA = J_n(\lambda)P$$

Označme riadky matice A ako $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ a prečítajme si tú istú maticovú rovnosť po riadkoch.

$$\begin{pmatrix} \vec{\alpha}_1 \\ \vec{\alpha}_2 \\ \vdots \\ \vec{\alpha}_{n-1} \\ \vec{\alpha}_n \end{pmatrix} A = \begin{pmatrix} \lambda & 1 & 0 & \ldots & 0 \\ 0 & \lambda & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & \lambda & 1 \\ 0 & \ldots & \ldots & 0 & \lambda \end{pmatrix} \begin{pmatrix} \vec{\alpha}_1 \\ \vec{\alpha}_2 \\ \vdots \\ \vec{\alpha}_{n-1} \\ \vec{\alpha}_n \end{pmatrix} \tag{9.7}$$ {jordan:EQRIADKY1}

$$\begin{pmatrix} \vec{\alpha}_1 A \\ \vec{\alpha}_2 A \\ \vdots \\ \vec{\alpha}_{n-1} A \\ \vec{\alpha}_n A \end{pmatrix} = \begin{pmatrix} \lambda\vec{\alpha}_1 + \alpha_2 \\ \lambda\vec{\alpha}_2 + \alpha_3 \\ \vdots \\ \lambda\vec{\alpha}_{n-1} + \alpha_n \\ \lambda\vec{\alpha}_n \end{pmatrix} \tag{9.8}$$ {jordan:EQRIADKY2}

Porovnaním týchto matíc vidíme, že musí platiť

$$\vec{\alpha}_n A = \lambda\vec{\alpha}_n.$$

To znamená, že λ je vlastná hodnota matice A a $\vec{\alpha}_n$ je vlastný vektor, ktorý k nej prislúcha. Vlastné hodnoty a vlastné vektory už vieme hľadať – nájdeme korene charakteristického polynómu $|A - xI|$ a pre ne potom riešime homogénnu sústavu určenú maticou $(A - \lambda I)^T$.

Predchádzajúci vektor má spĺňať rovnosť

$$\vec{\alpha}_{n-1} A = \lambda\vec{\alpha}_{n-1} + \alpha_n$$

alebo, ekvivalentne,

$$\vec{\alpha}_{n-1}(A - \lambda I) = \vec{\alpha}_n.$$

Vektor $\vec{\alpha}_{n-1}$ teda môžeme nájsť riešením nehomogénnej sústavy rovníc

$$(A - \lambda I)^T \vec{\alpha}_{n-1}^T = \vec{\alpha}_n^T.$$

Podobne postupujeme ďalej – v každom kroku nájdeme nový vektor riešením sústavy

$$(A - \lambda I)^T \vec{\alpha}_{j-1}^T = \vec{\alpha}_j^T.$$

Ak nájdeme vektory spĺňajúce uvedené rovnosti, tak matica P skutočne spĺňa rovnosti (9.8) a $PAP^{-1} = J_n(\lambda)$. Navyše, ukázali sme, že aby pre maticu P tieto rovnosti platili, jej riadky musia spĺňať všetky uvedené podmienky.

[3]Rozumieme tým násobnosť vlastnej hodnoty ako koreňa charakteristického polynómu.

Na ten istý problém sa môžeme pozrieť ešte aj iným spôsobom. Vieme, že podobnosť matíc znamená, že obe matice predstavujú pri rôznych bázach to isté zobrazenie. Čiže hľadáme bázu $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$, pri ktorej má zobrazenie určené (pri štandardnej báze) maticou A maticu $J_\lambda(n)$. To znamená špeciálne, že má platiť

$$\begin{aligned}\vec{\alpha}_n A &= \lambda\vec{\alpha}_n\\ \vec{\alpha}_{n-1}A &= \lambda\vec{\alpha}_{n-1} + \vec{\alpha}_n\\ &\vdots\\ \vec{\alpha}_1 &= \lambda\vec{\alpha}_1 + \vec{\alpha}_2\end{aligned}$$

Dostali sme teda presne tie isté rovnice pre $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$.

Vo všeobecnosti môžeme mať Jordanových blokov viac, nie iba jeden ako v predchádzajúcej úvahe. Ukážme si na konkrétnom príklade, ako môžeme potom nájsť Jordanov normálny tvar danej matice.

Príklad 9.4.3. Nájdime Jordanov normálny tvar pre maticu

$$A = \begin{pmatrix} 6 & 1 & -3 & 2 & 5\\ -1 & 2 & 1 & -3 & 0\\ 1 & 0 & 1 & 3 & 0\\ -1 & 0 & 1 & 3 & -2\\ -2 & 0 & 2 & 2 & -2 \end{pmatrix}$$

Najprv chceme nájsť vlastné hodnoty. Vypočítajme charakteristický polynóm.

$$ch_A(x) = \begin{vmatrix} 6-x & 1 & -3 & 2 & 5\\ -1 & 2-x & 1 & -3 & 0\\ 1 & 0 & 1-x & 3 & 0\\ -1 & 0 & 1 & 3-x & -2\\ -2 & 0 & 2 & 2 & -2-x \end{vmatrix} \overset{(1)}{=} \begin{vmatrix} 6-x & 1 & -3 & 2 & 5\\ 0 & 2-x & 2-x & 0 & 0\\ 1 & 0 & 1-x & 3 & 0\\ -1 & 0 & 1 & 3-x & -2\\ -2 & 0 & 2 & 2 & -2-x \end{vmatrix} = (2-x)\begin{vmatrix} 6-x & 1 & -3 & 2 & 5\\ 0 & 1 & 1 & 0 & 0\\ 1 & 0 & 1-x & 3 & 0\\ -1 & 0 & 1 & 3-x & -2\\ -2 & 0 & 2 & 2 & -2-x \end{vmatrix} =$$

$$= (2-x)\begin{vmatrix} 6-x & 0 & -4 & 2 & 5\\ 0 & 1 & 1 & 0 & 0\\ 1 & 0 & 1-x & 3 & 0\\ -1 & 0 & 1 & 3-x & -2\\ -2 & 0 & 2 & 2 & -2-x \end{vmatrix} \overset{(2)}{=} (2-x)\begin{vmatrix} 6-x & -4 & 2 & 5\\ 1 & 1-x & 3 & 0\\ -1 & 1 & 3-x & -2\\ -2 & 2 & 2 & -2-x \end{vmatrix} \overset{(3)}{=} (2-x)\begin{vmatrix} 6-x & -4 & 2 & 5\\ 1 & 1-x & 3 & 0\\ -1 & 1 & 3-x & -2\\ 0 & 0 & 2x-4 & 2-x \end{vmatrix} =$$

$$= (2-x)^2\begin{vmatrix} 6-x & -4 & 2 & 5\\ 1 & 1-x & 3 & 0\\ -1 & 1 & 3-x & -2\\ 0 & 0 & -2 & 1 \end{vmatrix} = (2-x)^2\begin{vmatrix} 6-x & -4 & 12 & 0\\ 1 & 1-x & 3 & 0\\ -1 & 1 & -1-x & 0\\ 0 & 0 & -2 & 1 \end{vmatrix} = (2-x)^2\begin{vmatrix} 6-x & -4 & 12\\ 1 & 1-x & 3\\ -1 & 1 & -1-x \end{vmatrix} \overset{4}{=} (2-x)^2\begin{vmatrix} 6-x & -4 & 12\\ 1 & 1-x & 3\\ 0 & 2-x & 2-x \end{vmatrix} =$$

$$= (2-x)^3\begin{vmatrix} 6-x & -4 & 12\\ 1 & 1-x & 3\\ 0 & 1 & 1 \end{vmatrix} = (2-x)^3\begin{vmatrix} 6-x & -16 & 0\\ 1 & -2-x & 0\\ 0 & 1 & 1 \end{vmatrix} = (2-x)^3\begin{vmatrix} 6-x & -16\\ 1 & -2-x \end{vmatrix} =$$

$$(2-x)^3[(x-6)(x+2)+16] = (2-x)^3(x^2-4x+4) = (2-x)^5$$

Použité úpravy:
(1) Pripočítanie tretieho riadku k druhému
(2) Laplaceov rozvoj podľa druhého stĺpca
(3) Odpočítanie dvojnásobku tretieho riadku od štvrtého.
(4) Pripočítanie druhého riadku k tretiemu

Našli sme charakteristický polynóm

$$ch_A(x) = -(x-2)^5.$$

Jediný koreň charakteristického polynómu (a jediná vlastná hodnota) je teda číslo 2.

Teraz nájdeme k tejto vlastnej hodnote vlastné vektory. Riešime sústavu danú maticou $(A-2I)^T$.

$$\begin{pmatrix} 4 & -1 & 1 & -1 & -2 \\ 1 & 0 & 0 & 0 & 0 \\ -3 & 1 & -1 & 1 & 2 \\ 2 & -3 & 3 & 1 & 2 \\ 5 & 0 & 0 & -2 & -4 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & -1 & -2 \\ 0 & 1 & -1 & 1 & 2 \\ 0 & -3 & 3 & 1 & 2 \\ 0 & 0 & 0 & -2 & -4 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & -1 & -2 \\ 0 & 1 & -1 & 1 & 2 \\ 0 & -3 & 3 & 1 & 2 \\ 0 & 0 & 0 & 1 & 2 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & -3 & 3 & 0 & 0 \\ 0 & 0 & 0 & 1 & 2 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Podpriestor riešení tejto sústavy je $[(0,1,1,0,0),(0,0,0,2,-1)]$. Ľahko overíme, že tieto vektory sú skutočne vlastné vektory prislúchajúce k vlastnej hodnote 2. Podobne je vlastným vektorom aj každá ich lineárna kombinácia $a(0,1,1,0,0)+b(0,0,0,2,-1)$.

Keďže množina vlastných vektorov je dvojrozmerná, pre maticu A sa dajú nájsť 2 lineárne nezávislé vlastné vektory. Znamená to, že jej Jordanov normálny tvar bude mať 2 Jordanove bloky.

Teraz budeme riešiť rovnicu zadanú maticou $(A-2I)^T$, v ktorej pravú stranu tvoria práve vypočítané vlastné vektory.

$$\left(\begin{array}{ccccc|c} 4 & -1 & 1 & -1 & -2 & 0 \\ 1 & 0 & 0 & 0 & 0 & a \\ -3 & 1 & -1 & 1 & 2 & a \\ 2 & -3 & 3 & 1 & 2 & 2b \\ 5 & 0 & 0 & -2 & -4 & -b \end{array}\right) \sim \left(\begin{array}{ccccc|c} 1 & 0 & 0 & 0 & 0 & a \\ 4 & -1 & 1 & -1 & -2 & 0 \\ -3 & 1 & -1 & 1 & 2 & a \\ 2 & -3 & 3 & 1 & 2 & 2b \\ 5 & 0 & 0 & -2 & -4 & -b \end{array}\right) \sim \left(\begin{array}{ccccc|c} 1 & 0 & 0 & 0 & 0 & a \\ 0 & -1 & 1 & -1 & -2 & -4a \\ 0 & 1 & -1 & 1 & 2 & 4a \\ 0 & -3 & 3 & 1 & 2 & -2a+2b \\ 0 & 0 & 0 & -2 & -4 & -5a-b \end{array}\right) \sim$$

$$\left(\begin{array}{ccccc|c} 1 & 0 & 0 & 0 & 0 & a \\ 0 & -1 & 1 & -1 & -2 & -4a \\ 0 & 1 & -1 & 1 & 2 & 4a \\ 0 & -3 & 3 & 1 & 2 & -2a+2b \\ 0 & 0 & 0 & 1 & 2 & \frac{5}{2}a+\frac{1}{2}b \end{array}\right) \sim \left(\begin{array}{ccccc|c} 1 & 0 & 0 & 0 & 0 & a \\ 0 & -1 & 1 & & & -\frac{3}{2}a+\frac{1}{2}b \\ 0 & 1 & -1 & 0 & 0 & \frac{3}{2}a-\frac{1}{2}b \\ 0 & -3 & 3 & 0 & 0 & -\frac{9}{2}a+\frac{3}{2}b \\ 0 & 0 & 0 & 1 & 2 & \frac{5}{2}a+\frac{1}{2}b \end{array}\right) \sim \left(\begin{array}{ccccc|c} 1 & 0 & 0 & 0 & 0 & a \\ 0 & 1 & -1 & 0 & 0 & \frac{3}{2}a-\frac{1}{2}b \\ 0 & 0 & 0 & 1 & 2 & \frac{5}{2}a+\frac{1}{2}b \end{array}\right)$$

Našli sme riešenie $a(1,\frac{3}{2},0,\frac{5}{2},0)+b(0,-\frac{1}{2},0,\frac{1}{2},0)$ pre vlastný vektor $a(0,1,1,0,0)+b(0,0,0,2,-1)$. Keďže sme našli riešenie pre každý vlastný vektor, oba Jordanove bloky musia mať veľkosť aspoň 2. Vzhľadom k tomu, že súčet ich veľkostí je 5, musia to byť bloky veľkosti 2 a 3. To znamená, že už vieme, ako vyzerá Jordanov normálny tvar našej matice, pokúsime sa však ešte dopočítať aj maticu prechodu.

Opäť riešime sústavu danú tou istou maticou, pričom za pravú stranu vezmeme ľubovoľný vektor tvaru $a(1,\frac{3}{2},0,\frac{5}{2},0)+b(0,-\frac{1}{2},0,\frac{1}{2},0)$.

$$\left(\begin{array}{ccccc|c} 4 & -1 & 1 & -1 & -2 & a \\ 1 & 0 & 0 & 0 & 0 & \frac{3}{2}a-\frac{1}{2}b \\ -3 & 1 & -1 & 1 & 2 & 0 \\ 2 & -3 & 3 & 1 & 2 & \frac{5}{2}a+\frac{1}{2}b \\ 5 & 0 & 0 & -2 & -4 & 0 \end{array}\right) \sim \left(\begin{array}{ccccc|c} 1 & 0 & 0 & 0 & 0 & \frac{3}{2}a-\frac{1}{2}b \\ 4 & -1 & 1 & -1 & -2 & a \\ -3 & 1 & -1 & 1 & 2 & 0 \\ 2 & -3 & 3 & 1 & 2 & \frac{5}{2}a+\frac{1}{2}b \\ 5 & 0 & 0 & -2 & -4 & 0 \end{array}\right) \sim \left(\begin{array}{ccccc|c} 1 & 0 & 0 & 0 & 0 & \frac{3}{2}a-\frac{1}{2}b \\ 0 & -1 & 1 & -1 & -2 & -5a+2b \\ 0 & 1 & -1 & 1 & 2 & \frac{9}{2}a-\frac{3}{2}b \\ 0 & -3 & 3 & 1 & 2 & -\frac{1}{2}a+\frac{3}{2}b \\ 0 & 0 & 0 & -2 & -4 & -\frac{15}{2}a+\frac{5}{2}b \end{array}\right) \sim$$

$$\left(\begin{array}{ccccc|c} 1 & 0 & 0 & 0 & 0 & \frac{3}{2}a-\frac{1}{2}b \\ 0 & -1 & 1 & -1 & -2 & -5a+2b \\ 0 & 1 & -1 & 1 & 2 & \frac{9}{2}a-\frac{3}{2}b \\ 0 & -3 & 3 & 1 & 2 & -\frac{1}{2}a+\frac{3}{2}b \\ 0 & 0 & 0 & 1 & 2 & \frac{15}{4}a-\frac{5}{4}b \end{array}\right) \sim \left(\begin{array}{ccccc|c} 1 & 0 & 0 & 0 & 0 & \frac{3}{2}a-\frac{1}{2}b \\ 0 & -1 & 1 & 0 & 0 & -\frac{5}{4}a+\frac{3}{4}b \\ 0 & 1 & -1 & 0 & 0 & \frac{3}{4}a-\frac{1}{4}b \\ 0 & -3 & 3 & 0 & 0 & -\frac{17}{4}a+\frac{11}{4}b \\ 0 & 0 & 0 & 1 & 2 & \frac{15}{4}a-\frac{5}{4}b \end{array}\right)$$

Aby predchádzajúca sústava mala riešenie, musí platiť $-\frac{5}{4}a+\frac{3}{4}b=-\frac{3}{4}a+\frac{1}{4}b$ (dostaneme to porovnaním druhého a tretieho riadku). Z toho dostaneme $a=b$. Tým prejde sústava na tvar

$$\left(\begin{array}{ccccc|c} 1 & 0 & 0 & 0 & 0 & a \\ 0 & -1 & 1 & 0 & 0 & -\frac{a}{2} \\ 0 & 1 & -1 & 0 & 0 & \frac{a}{2} \\ 0 & -3 & 3 & 0 & 0 & -\frac{3}{2}a \\ 0 & 0 & 0 & 1 & 2 & \frac{5}{2}a \end{array}\right) \sim \left(\begin{array}{ccccc|c} 1 & 0 & 0 & 0 & 0 & a \\ 0 & 1 & -1 & 0 & 0 & \frac{a}{2} \\ 0 & 0 & 0 & 1 & 2 & \frac{5}{2}a \end{array}\right)$$

Ako jedno z možných riešení dostaneme $a(1,\frac{1}{2},0,\frac{5}{2},0)$. Pri voľbe $a=1$ máme

$\vec{\alpha}_1=(1,\frac{1}{2},0,\frac{5}{2},0)$

$\vec{\alpha}_2=\vec{\alpha}_1(A-2I)=(1,1,0,3,0)$

$\vec{\alpha}_3=\vec{\alpha}_2(A-2I)=(0,1,1,2,-1)$

Dostali sme presne vlastný vektor zodpovedajúci hodnotám $a=b=1$. Tieto 3 vektory určujú jeden Jordanov blok. Aby sme dostali druhý, potrebujeme použiť nejaký vlastný vektor lineárne nezávislý od $(0,1,1,2,-1)$. Môžeme zvoliť napríklad $a=1$, $b=0$:

$\vec{\alpha}_4=(1,\frac{3}{2},0,\frac{5}{2},0)$

$\vec{\alpha}_5=\vec{\alpha}_4(A-2I)=(0,1,1,0,0)$

Dostali sme teda

$$P = \begin{pmatrix} 1 & \frac{1}{2} & 0 & \frac{5}{2} & 0 \\ 1 & 1 & 0 & 3 & 0 \\ 0 & 1 & 1 & 2 & -1 \\ 1 & \frac{3}{2} & 0 & \frac{5}{2} & 0 \\ 0 & 1 & 1 & 0 & 0 \end{pmatrix} \qquad J = \begin{pmatrix} 2 & 1 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}$$

Pre tieto matice skutočne platí

$$J = PAP^{-1}.$$

Vlastné vektory sú určené rovnosťou $\vec{\alpha}(A - 2I) = \vec{0}$ Všimnime si, že pre ostatné vektory, ktoré sme dostali v predošlom príklade platí $\vec{\alpha}_2(A-2I)^2 = \vec{\alpha}_4(A-2I)^2 = \vec{0}$ a $\vec{\alpha}_1(A-2I)^3 = \vec{0}$. Vektory, ktoré vyhovujú rovnici $\vec{\alpha}(A - \lambda I)^n$ pre nejaké $n \in \mathbb{N}$ a $\lambda \in \mathbb{C}$, sa nazývajú *zovšeobecnené vlastné vektory.*

Vďaka tomuto pozorovaniu môžeme hľadať vlastné vektory aj iným spôsobom. Konkrétne ide o to, že sa môžeme pozrieť na mocniny matice $A - \lambda I$, teda na matice tvaru $(A - \lambda I)^n$.

Ak je matica A podobná s blokovo-diagonálnou maticou J pozostávajúcich Jordanových blokov, tak $A - \lambda I$ je podobná s maticou $J - \lambda I$. Táto matica vyzerá tak, že v blokoch zodpovedajúcich vlastnej hodnote λ sa λ nahradila nulou a v blokoch zodpovedajúcich nejakej vlastnej hodnote $\mu \neq \lambda$ nahradíme na diagonále μ číslom $\mu - \lambda$.

Ak nás zaujíma hodnosť matice $(A - \lambda I)^n$, stačí sa pozrieť na maticu $(J - \lambda I)^n$, lebo podobné matice majú rovnakú hodnosť. Táto matica je výrazne jednoduchšia, takže ju budeme asi vedieť ľahšie umocniť. Skutočne, ak umocňujeme blokovo-diagonálnu maticu, tak výsledok je opäť blokovo-diagonálna matica, pričom jednotlivé bloky sú mocninami blokov vystupujúcich v pôvodnej matici. Takže sa nám stačí pozrieť na to, čo sa deje pri umocňovaní jednotlivých blokov.

Z blokov zodpovedajúcich vlastnej hodnote λ sme dostali bloky takéhoto tvaru:

$$\begin{pmatrix} 0 & 1 & 0 & \dots & \dots & 0 \\ 0 & 0 & 1 & 0 & & 0 \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & \vdots \\ 0 & & & 0 & 1 & 0 \\ 0 & & & & 0 & 1 \\ 0 & \dots & \dots & \dots & \dots & 0 \end{pmatrix}$$

Vidíme, že v každom taktomto bloku je jeden riadok nulový a ostatné sú lineárne nezávislé. Bloky zodpovedajúce ostatným vlastným hodnotám vyzerajú takto

$$\begin{pmatrix} \mu-\lambda & 1 & 0 & \dots & \dots & 0 \\ 0 & \mu-\lambda & 1 & 0 & & 0 \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & \vdots \\ 0 & & & \mu-\lambda & 1 & 0 \\ 0 & & & & \mu-\lambda & 1 \\ 0 & \dots & \dots & \dots & \dots & \mu-\lambda \end{pmatrix}$$

Fakt, že $\mu - \lambda \neq 0$ zabezpečí, že takáto bloková matica bude regulárna.

Keď sa teraz pozrieme na celú maticu $A - \lambda I$, tak táto matica obsahuje toľko nulových riadkov, koľko je v nej Jordanových blokov prislúchajúcich k vlastnej hodnote λ a ostatné riadky sú lineárne nezávislé. Teda počet Jordanových blokov pre túto vlastnú hodnotu je $n - h(A - \lambda I)$, ak pôvodná matica A má rozmery $n \times n$.

Ešte by sme si mali rozmyslieť, čo dostaneme umocňovaním takýchto matíc. Nie je ťažké uvedomiť si, že ak umocníme ľubovoľnú maticu tvaru

$$\begin{pmatrix} 0 & c_{1,2} & c_{1,3} & \dots & \dots & c_{1,n} \\ 0 & 0 & c_{2,3} & c_{2,4} & & c_{2,n} \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ 0 & & & 0 & c_{n-2,n-1} & c_{n-1,n} \\ 0 & & & & 0 & c_{n-1,n} \\ 0 & \dots & \dots & \dots & \dots & 0 \end{pmatrix}$$

t.j. ľubovoľnú maticu, ktorá má na hlavnej diagonále aj pod ňou samé nuly, tak v druhej mocnine nám pribudnú nuly tesne nad diagonálou. V tretej mocnine nám pribudne ďalšia vedľajšia diagonála pozostávajúca zo samých núl. (Do istej miery podobná úvaha ako robíme za rovnosťou (9.13).) Čiže keď porovnáme $(k-1)$-vú a k-tu mocninu, tak nám pre každý Jordanov blok veľkosti aspoň k pribudol jeden nulový riadok.

Bloky, ktoré mali na diagonále nenulové čísla $\mu-\lambda$ budú vo svojich mocninách na diagonále nenulové čísla $(\mu-\lambda)^n$ a zostanú regulárne. Podobná úvaha by fungovala aj pre ľubovoľnú maticu tvaru

$$\begin{pmatrix} d_1 & c_{1,2} & c_{1,3} & \dots & \dots & c_{1,n} \\ 0 & d_2 & c_{2,3} & c_{2,4} & & c_{2,n} \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ 0 & & & d_{n-2} & c_{n-2,n-1} & c_{n-1,n} \\ 0 & & & & d_{n-1} & c_{n-1,n} \\ 0 & \dots & \dots & \dots & \dots & d_n \end{pmatrix}.$$

Keď zhrnieme doterajšie úvahy, zistili sme, že v matici $(A-\lambda I)^k$ je počet nulových riadkov rovný $n_1+\dots+n_k$, kde n_k označuje počet Jordanových blokov veľkosti aspoň k prislúchajúcich vlastnej hodnote λ. Ostatné riadky sú lineárne nezávislé. Teda z hodností takýchto matíc vieme zistiť počty Jordanových blokov jednotlivých veľkostí.

Ukážme si tento postup na tej istej matici ako v predošlom príklade.

Príklad 9.4.4. Opäť teda budeme pracovať s maticou $A=\begin{pmatrix} 6 & 1 & -3 & 2 & 5 \\ -1 & 2 & 1 & -3 & 0 \\ 1 & 0 & 1 & 3 & 0 \\ -1 & 0 & 1 & 3 & -2 \\ -2 & 0 & 2 & 2 & -2 \end{pmatrix}$. Už sme vyrátali, že charakteristický polynóm je $ch_A(x)=-(x-2)^5$ a jediné vlastné číslo je 2.

Pozrime sa teraz na to, ako vyzerajú mocniny matice $(A-2I)$. Z Cayley-Hamiltonovej vety vieme, že $(A-2I)^5=0$, teda budeme musieť ísť nanajvýš po piatu mocninu. (To vyzerá na prvý pohľad veľmi prácne – ideme násobiť maticu 5×5 – keď si však všimneme, že druhý a tretí riadok sa líšia až na skalárny násobok; podobne je to pre štvrtý a piaty riadok; tak keď máme druhý a štvrtý riadok v matici $(A-2I)^2$, hneď poznáme aj tretí a piaty, ktoré dostaneme ako príslušné násobky.)

$$A-2I=\begin{pmatrix} 4 & 1 & -3 & 2 & 5 \\ -1 & 0 & 1 & -3 & 0 \\ 1 & 0 & -1 & 3 & 0 \\ -1 & 0 & 1 & 1 & -2 \\ -2 & 0 & 2 & 2 & -4 \end{pmatrix}$$

$$(A-2I)^2=\begin{pmatrix} 0 & 4 & 4 & 8 & -4 \\ 0 & -1 & -1 & -2 & 1 \\ 0 & 1 & 1 & 2 & -1 \\ 0 & -1 & -1 & -2 & 1 \\ 0 & -2 & -2 & -4 & 2 \end{pmatrix}$$

$$(A-2I)^3=\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Vidíme, že $h(A-2I)=3$, $h((A-2I)^2)=1$ a $h((A-2I)^3)=0$.

Z hodností, ktoré sme vyrátali, vidíme, že počet Jordanových blokov je $2=5-3$. Počet blokov, ktoré majú veľkosť aspoň dva je $2=3-1$ a počet blokov veľkosti aspoň tri je $1=1-0$. Teda máme jeden blok veľkosti 2 a jeden blok veľkosti 3. (Môžeme si všimnúť aj to, že nám stačilo rátať prvé dve mocniny.)

Teraz už teda vieme, ako vyzerá Jordanov normálny tvar:

$$J=\begin{pmatrix} 2 & 1 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}$$

Nevýhoda oproti predošlému postupu je tá, že sme nezostavili maticu P. (Čiže pri predošlom postupe sme mohli urobiť kontrolu vynásobením PAP^{-1}.)

Tak sa skúsme pozrieť na to, či by sme tu vedeli nájsť zovšeobecnené vlastné vektory. Vektor $\vec{\alpha}_3$ je tvaru $\vec{\alpha}_1(A-2I)^2$. Každý taký vektor je násobkom vektora $(0,1,1,2,-1)$. (Pozeráme sa na vektory, ktoré sú v podpriestore generovanom riadkami matice $(A-2I)^2$, v našom konkrétnom prípade je tento podpriestor jednorozmerný.) Môžeme teda zvoliť $\vec{\alpha}_3 = (0,1,1,2,-1)$.

Mali by sme teraz nájsť vektor $\vec{\alpha}_2$ taký, že $\vec{\alpha}_2(A-2I) = \vec{\alpha}_3$ a vektor $\vec{\alpha}_1$ vyhovujúci rovnosti $\vec{\alpha}_1(A-2I) = \vec{\alpha}_2$. To sa dá urobiť riešením sústavy; v podstate rovnakú sústavu sme riešili pri počítaní predošlým postupom, vieme teda, že možné riešenia sú $\vec{\alpha}_1 = (1, \frac{1}{2}, 0, \frac{5}{2}, 0)$, $\vec{\alpha}_2 = \vec{\alpha}_1(A-2I) = (1,1,0,3,0)$.

Ďalej by nás zaujímal vektor $\vec{\alpha}_5$, ktorý je vlastným vektorom a súčasne sa dá dostať ako $\vec{\alpha}_5 = \vec{\alpha}_4(A-2I)$ pre nejaký vektor $\vec{\alpha}_4$. (Týmto podmienkam vyhovuje aj vektor $\vec{\alpha}_3$, my chceme nejaký vektor, ktorý nie je jeho násobkom.)

Vektor $\vec{\alpha}_5$ teda patrí do podpriestoru $[(4,1,-3,2,5),(1,0,-1,3,0),(1,0,-1,-1,2)]$ generovaného riadkami matice $(A-2I)$. Súčasne by mal byť vlastným vektorom, teda by mal ležať v podpriestore $[(0,1,1,0,0),(0,0,0,2,-1)]$. (Tento podpriestor sme našli ako podpriestor riešení homogénneho systému s maticou $(A-2I)^T$ už v predošlom postupe.)

Prienik dvoch podpriestorov by sme vedeli vyrátať, v tomto konkrétnom prípade máme však situáciu výrazne zjednodušenú. Pretože celý podpriestor $[(0,1,1,0,0),(0,0,0,2,-1)]$ je dvojrozmerný a máme v ňom dva vektory lineárne nezávislé $\vec{\alpha}_3$ a $\vec{\alpha}_5$, tak vieme, že ten prienik musí byť dvojrozmerný, bude sa teda zhodovať s podpriestorom $[(0,1,1,0,0),(0,0,0,2,-1)]$. Preto je vhodným kandidátom pre $\vec{\alpha}_5$ ľubovoľný nenulový vektor z tohoto podpriestoru rôzny od $\vec{\alpha}_3$. Ak napríklad zvolíme $\vec{\alpha}_5 = (0,1,1,0,0)$, tak riešením sústavy dostaneme $\vec{\alpha}_4 = (1, \frac{3}{2}, 0, \frac{5}{2}, 0)$.

Poznámka 9.4.5. Na základe toho, že z hodností matíc $h((A-\lambda I)^k)$ sa dá vyčítať počet blokov jednotlivých veľkostí, by sme vedeli dokázať aspoň to, že Jordanov normálny tvar je jednoznačne určený až na poradie blokov.

Cvičenia

Úloha 9.4.1. Nájdite Jordanov normálny tvar J matice A a regulárnu maticu P takú, že platí $PAP^{-1} = J$.

a) $A = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 3 & -3 \\ 0 & 1 & 0 \end{pmatrix}$

b) $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 3 & 1 \\ 1 & -3 & 0 \end{pmatrix}$

c) $A = \begin{pmatrix} 1 & -1 & -1 & 1 \\ 0 & 0 & -1 & 1 \\ -2 & -2 & -1 & 4 \\ -1 & -2 & -2 & 4 \end{pmatrix}$

d) $A = \begin{pmatrix} 0 & 2 & 2 \\ -1 & 3 & 1 \\ -1 & 1 & 3 \end{pmatrix}$

e) $A = \begin{pmatrix} 2 & 0 & -1 \\ -1 & 2 & 1 \\ 2 & -1 & -1 \end{pmatrix}$

[a) $ch_A(x) = -(x-1)^3$, vlastný vektor $(1,1,-2)$, $J = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$; b) $ch_A(x) = -(x-1)^3$, jediný vlastný vektor $(1,1,1)$, $J = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$; c) $ch_A(x) = -(x-1)^4$, vlastné vektory $[(-1,1,0,0),(-2,0,-1,2)]$, $J = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$; d) $ch_A(x) = -(x-2)^3$, vlastné vektory $[(1,0,-2),(0,1,-1)]$, $J = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$; e) $ch_A(x) = -(x-1)^3$, vlastné vektory $[(1,-1,-1)]$, $J = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$]

Úloha 9.4.2. Nájdite Jordanov normálny tvar daných matíc.
$\begin{pmatrix} 0 & 1 & 0 \\ -4 & 4 & 0 \\ -2 & 1 & 2 \end{pmatrix}$ $\begin{pmatrix} 2 & 6 & -15 \\ 1 & 1 & -5 \\ 1 & 2 & -6 \end{pmatrix}$ $\begin{pmatrix} 9 & -6 & -2 \\ 18 & -12 & -3 \\ 18 & -9 & -6 \end{pmatrix}$ $\begin{pmatrix} 4 & -5 & 2 \\ 5 & -7 & 3 \\ 6 & -9 & 4 \end{pmatrix}$ $\begin{pmatrix} 1 & -3 & 3 \\ -2 & -6 & 13 \\ -1 & -4 & 8 \end{pmatrix}$ $\begin{pmatrix} 7 & -12 & 6 \\ 10 & -19 & 10 \\ 12 & -24 & 13 \end{pmatrix}$ $\begin{pmatrix} t & 0 & 0 \\ 0 & t & 0 \\ t & 0 & t \end{pmatrix}$ pre $t \neq 0$
Riešenia: $\begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$ $\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{pmatrix}$ $\begin{pmatrix} -3 & 0 & 0 \\ 0 & -3 & 1 \\ 0 & 0 & -3 \end{pmatrix}$ $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ $\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$ $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$ $\begin{pmatrix} 1 & 0 & 0 \\ 0 & t & 1 \\ 0 & 0 & t \end{pmatrix}$

Úloha 9.4.3. Nájdite Jordanov normálny tvar daných matíc.
$\begin{pmatrix} 1 & -3 & 0 & 3 \\ -2 & -6 & 0 & 13 \\ 0 & -3 & 1 & 3 \\ -1 & -4 & 0 & 8 \end{pmatrix}$ $\begin{pmatrix} 3 & -1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 3 & 0 & 5 & -3 \\ 4 & -1 & 3 & -1 \end{pmatrix}$ $\begin{pmatrix} 3 & -4 & 0 & 2 \\ 4 & -5 & -2 & 4 \\ 0 & 0 & 3 & -2 \\ 0 & 0 & 2 & -1 \end{pmatrix}$ $\begin{pmatrix} 3 & -1 & 1 & -7 \\ 9 & -3 & -7 & -1 \\ 0 & 0 & 4 & -8 \\ 0 & 0 & 2 & -4 \end{pmatrix}$
Riešenia $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ $\begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}$ $\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \end{pmatrix}$ $\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$

Úloha 9.4.4. Nájdite Jordanov tvar danej matice:
a) $\begin{pmatrix} -2 & -5 & 3 \\ 1 & 0 & -1 \\ 0 & -4 & 1 \end{pmatrix}$
b) $\begin{pmatrix} 0 & 1 & -1 \\ 1 & 0 & 1 \\ 2 & -2 & 3 \end{pmatrix}$
Výsledky: a) $\begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ b) $\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

Úloha 9.4.5. Nájdite charakteristický polynóm a Jordanov tvar matice Nájdite aj príslušnú maticu prechodu a zapíšte príslušnú maticovú rovnosť.
$\begin{pmatrix} -3 & 0 & 4 \\ 3 & -1 & -5 \\ -2 & 0 & 3 \end{pmatrix}$ $\begin{pmatrix} 1 & 4 & -3 \\ 0 & -3 & 1 \\ 0 & -4 & 1 \end{pmatrix}$ $\begin{pmatrix} -2 & -5 & 3 \\ 1 & 0 & -1 \\ 0 & -4 & 1 \end{pmatrix}$ $\begin{pmatrix} 2 & -1 & 0 \\ 1 & 0 & 0 \\ 2 & 2 & 1 \end{pmatrix}$ $\begin{pmatrix} 3 & -3 & 1 \\ 2 & -2 & 1 \\ 2 & -3 & 2 \end{pmatrix}$ $\begin{pmatrix} 3 & -3 & 1 \\ 2 & -2 & 1 \\ 2 & -3 & -2 \end{pmatrix}$
Výsledky: $\begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ $\begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ $\begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ $\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$ $\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ $\begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

9.5 Aplikácie podobnosti a Jordanovho normálneho tvaru

9.5.1 Lineárne rekurencie

Pomocou Jordanovho normálneho tvaru sa dá odvodiť riešenie lineárnych rekurentných rovníc. My sa pre jednoduchosť obmedzíme na rekurencie druhého rádu.

Pod *lineárnou rekurenciou druhého rádu* rozumieme predpis

$$A_{n+1} = aA_n + bA_{n-1}, \tag{9.9}$$ {aplik:REK2}

kde $a, b \in \mathbb{C}$. Je zrejmé, že ak nejaká postupnosť $(A_n)_{n=1}^{\infty}$ vyhovuje rovnici (9.9) pre všetky $n \in \mathbb{N}$ a ak poznáme jej počiatočné hodnoty A_0 a A_1, tak tým je už postupnosť $(A_n)_{n=1}^{\infty}$ jednoznačne určená.

Základom pre to, aby sme mohli aplikovať naše poznatky o podobnosti matíc na lineárne rekurencie je uvedomiť si, že s rovnosťou (9.9) je ekvivalentný nasledovný maticový zápis

$$\begin{pmatrix} A_{n+1} \\ A_n \end{pmatrix} = \begin{pmatrix} a & b \\ 1 & 0 \end{pmatrix} \begin{pmatrix} A_n \\ A_{n-1} \end{pmatrix} \tag{9.10}$$ {aplik:EQREKMAT}

Charakteristický polynóm tejto matice je

$$ch_A(x) = \begin{vmatrix} a-x & b \\ 1 & -x \end{vmatrix} = x(x-a) - b = x^2 - ax - b.$$

Vlastné hodnoty sú riešenia rovnice $ch_A(x) = 0$. Táto rovnica sa zvykne nazývať *charakteristická rovnica* rekurencie (9.9).

Pre vlastné hodnoty $\lambda_{1,2}$ matice $A = \begin{pmatrix} a & b \\ 1 & 0 \end{pmatrix}$ teda platí

$$\begin{aligned} \lambda_1 + \lambda_2 &= a \\ \lambda_1 \cdot \lambda_2 &= -b \end{aligned} \tag{9.11}$$ {aplik:EQVIET

Z rovnosti (9.10) dostaneme postupne

{aplik:EQMOCN}
$$\begin{pmatrix} A_{n+1} \\ A_n \end{pmatrix} = A \begin{pmatrix} A_n \\ A_{n-1} \end{pmatrix} = A^2 \begin{pmatrix} A_{n-1} \\ A_{n-2} \end{pmatrix} = \cdots = A^n \begin{pmatrix} A_1 \\ A_0 \end{pmatrix} \tag{9.12}$$

Ak poznáme Jordanov tvar matice A, t.j. ak platí $A = P^{-1}JP$, tak túto rovnosť vieme prepísať ako

$$\begin{pmatrix} A_{n+1} \\ A_n \end{pmatrix} = A^n \begin{pmatrix} A_1 \\ A_0 \end{pmatrix} = P^{-1} J^n P \begin{pmatrix} A_1 \\ A_0 \end{pmatrix}.$$

Na ďalšie výpočty potrebujeme vedieť ako vyzerajú mocniny matice v Jordanovom normálnom tvare. V prípade, že J je diagonálna matica, jednoducho umocníme prvky na diagonále. Pre matice 2×2 máme.

$$J = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \qquad \Rightarrow \qquad J^n = \begin{pmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{pmatrix}$$

Vo všeobecnosti nám stačí umocňovať Jordanove bloky. V prípade matice 2×2 je situácia pomerne jednoduchá:

{aplik:EQMOCNJORD}
$$J = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} \qquad J^2 = \begin{pmatrix} \lambda^2 & 2\lambda \\ 0 & \lambda^2 \end{pmatrix} \qquad \ldots \qquad J^n = \begin{pmatrix} \lambda^n & n\lambda^{n-1} \\ 0 & \lambda^n \end{pmatrix} \tag{9.13}$$

Keďže pracujeme iba s rekurenciami druhého stupňa, vystačíme s maticami 2×2. Môžete si ale skúsiť rozmyslieť, že podobne to bude fungovať aj pre Jordanove bloky väčších rozmerov, t.j. k-ta mocnina Jordanovho bloku obsahuje (počnúc od diagonály) prvky $\lambda^k, k\lambda^k, \binom{k}{2}\lambda^{k-1}, \ldots$

Pozrime sa teraz nato, čo z predchádzajúcich rovností dostaneme pre rekurentné postupnosti.

V prípade, že Jordanov tvar je diagonálny, máme

$$\begin{pmatrix} A_{n+1} \\ A_n \end{pmatrix} = \begin{pmatrix} p'_{11} & p'_{12} \\ p'_{21} & p'_{22} \end{pmatrix} \begin{pmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{pmatrix} \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix} \begin{pmatrix} A_1 \\ A_0 \end{pmatrix} = \begin{pmatrix} p'_{11} & p'_{12} \\ p'_{21} & p'_{22} \end{pmatrix} \begin{pmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{pmatrix} \begin{pmatrix} a'_1 \\ a'_0 \end{pmatrix} =$$
$$= \begin{pmatrix} p'_{11} & p'_{12} \\ p'_{21} & p'_{22} \end{pmatrix} \begin{pmatrix} a'_1\lambda_1^n \\ a'_0\lambda_2^n \end{pmatrix} = \begin{pmatrix} p'_{11}a'_1\lambda_1^n + p'_{12}a'_0\lambda_2^n \\ p'_{21}a'_1\lambda_1^n + p'_{22}a'_0\lambda_2^n \end{pmatrix}$$

Ak porovnáme druhý riadok matíc na ľavej a pravej strane v predošlej rovnosti, vyšlo nám, že

$$A_n = c_1\lambda_1^n + c_2\lambda_2^n.$$

Na vyrátanie konštánt $c_{1,2}$ môžeme použiť to, že poznáme iniciálne hodnoty $A_{0,1}$.

Pozrime sa na prípad, že matica A nie je diagonalizovateľná. Jej Jordanov tvar potom obsahuje jediný Jordanov blok veľkosti 2. Vlastné hodnoty sú rovnaké, ich spoločnú hodnotu

označme λ. Potom máme

$$\begin{pmatrix} A_{n+1} \\ A_n \end{pmatrix} = \begin{pmatrix} p'_{11} & p'_{12} \\ p'_{21} & p'_{22} \end{pmatrix} \begin{pmatrix} \lambda^n & n\lambda^{n-1} \\ 0 & \lambda^n \end{pmatrix} \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix} \begin{pmatrix} A_1 \\ A_0 \end{pmatrix} = \begin{pmatrix} p'_{11} & p'_{12} \\ p'_{21} & p'_{22} \end{pmatrix} \begin{pmatrix} \lambda^n & n\lambda^{n-1} \\ 0 & \lambda^n \end{pmatrix} \begin{pmatrix} a'_1 \\ a'_0 \end{pmatrix} =$$
$$= \begin{pmatrix} p'_{11} & p'_{12} \\ p'_{21} & p'_{22} \end{pmatrix} \begin{pmatrix} a'_1\lambda^n + a'_0 n\lambda^{n-1} \\ a'_0\lambda^n \end{pmatrix} = \begin{pmatrix} (p'_{11}a'_1 + p'_{12}a'_0)\lambda^n + p'_{11}n\lambda^{n-1} \\ (p'_{21}a'_1 + p'_{22}a'_0)\lambda^n + p'_{21}n\lambda^{n-1} \end{pmatrix}$$

Zistili sme, že

$$A_n = c_1\lambda^n + c_2 n\lambda^{n-1}.$$

Opäť, konštanty $c_{1,2}$ môžeme vyrátať z počiatočných podmienok.

Podobným spôsobom sa dá odvodiť aj všeobecný vzťah pre lineárne rekurencie k-tého stupňa

$$A_{n+k} = c_{k-1}A_{n+k-1} + c_{k-2}A_{n+k-2} + \cdots + c_1A_{n+1} + c_0A_n.$$

Opäť, riešenie môžeme nájsť ako lineárnu kombináciu geometrických postupností určených koreňmi charakteristickej rovnice, v prípade, že je niektorý koreň viacnásobný, treba brať do úvahy okrem geometrickej postupnosti λ^n aj postupnosti $n\lambda^{n-1}, n^2\lambda^{n-2}, \ldots$ (toľko z nich, koľko je násobnosť príslušného koreňa).

Viac o lineárnych rekurenciách ako aj dôkaz vety o tvare riešení založený práve na použití Jordanovho normálneho tvaru môžete nájsť napríklad v [CFR, Section 2.2]. Iný dôkaz môžete nájsť v [W].

{aplik:PRFIBON}

Príklad 9.5.1. Nájdime vyjadrenie n-tého člena Fibonacciho postupnosti určenej predpisom

$$F_{n+1} = F_n + F_{n-1} \tag{9.14}$$

{aplik:EQFIB}

a počiatočnými hodnotami

$$F_0 = 0, F_1 = 1.$$

Maticový zápis rovnice (9.14) je

$$\begin{pmatrix} F_{n+1} \\ F_n \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} F_n \\ F_{n-1} \end{pmatrix} \tag{9.15}$$

{aplik:EQFIBMAT}

Charakteristická rovnica je $x^2 - x - 1 = 0$ a jej korene sú

$$\lambda_{1,2} = \frac{1 \pm \sqrt{5}}{2}.$$

Na základe Viétových vzťahov pre ne platí

$$\lambda_1 + \lambda_2 = 1,$$
$$\lambda_1\lambda_2 = -1.$$

Vlastné vektory pre vlastnú hodnotu λ_1 nájdeme riešením sústavy s maticou

$$\begin{pmatrix} 1-\lambda_1 & 1 \\ 1 & -\lambda_1 \end{pmatrix} = \begin{pmatrix} \lambda_2 & 1 \\ 1 & -\lambda_1 \end{pmatrix}$$

Vidíme, že obom rovniciam vyhovuje napríklad vektor $(1, -\lambda_2)$.

Podobne vlastným vektorom pre vlastnú hodnotu λ_2 je $(1, -\lambda_1)$. Teda pre maticu

$$P = \begin{pmatrix} 1 & -\lambda_2 \\ 1 & -\lambda_1 \end{pmatrix}$$

platí $PAP^{-1} = \operatorname{diag}(\lambda_1, \lambda_2)$.

Inverzná matica k matici P je

$$P^{-1} = \frac{1}{\lambda_2 - \lambda_1} \begin{pmatrix} -\lambda_1 & \lambda_2 \\ -1 & 1 \end{pmatrix}$$

Dosadením do (9.12) potom dostaneme

$$\begin{pmatrix} F_{n+1} \\ F_n \end{pmatrix} = P^{-1} \begin{pmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{pmatrix} P \begin{pmatrix} F_1 \\ F_0 \end{pmatrix} = \frac{1}{\lambda_2 - \lambda_1} \begin{pmatrix} -\lambda_1^{n+1} & \lambda_2^{n+1} \\ -\lambda_1^n & \lambda_2^n \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{\lambda_2 - \lambda_1} \begin{pmatrix} \lambda_2^{n+1} - \lambda_1^{n+1} \\ \lambda_2^n - \lambda_1^n \end{pmatrix}$$

Na základe porovnania spodného riadku na ľavej a pravej strane predchádzajúcej rovnosti vidno, že

{aplikrekur:EQBINETAB}
$$F_n = \frac{\lambda_2^n - \lambda_1^n}{\lambda_2 - \lambda_1}, \tag{9.16}$$

Ak dosadíme $\lambda_2 = \frac{1+\sqrt{5}}{2}$ a $\lambda_1 = \frac{1-\sqrt{5}}{2}$, tak máme

{aplikrekur:EQBINET}
$$F_n = \frac{\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n}{\sqrt{5}}. \tag{9.17}$$

Vzorec (9.17) (resp. (9.16)) sa volá *Cauchy-Binetova formula.*

Maticové rovnosti (9.10) a (9.15) nám môžu pomôcť nielen odvodiť vzorec pre n-tý člen postupnosti ale aj na odvodenie niektorých vzťahov platných pre rekurentné postupnosti. Ako jednoduchý príklad si môžeme ukázať odvodenie vzorca pre súčet prvých n členov Fibonacciho postupnosti.

Viac o využití matíc pri odvodzovaní rôznych identít platných pre členy Fibonacciho postupnosti (prípadne aj všeobecnejšie pre lineárne rekurencie druhého stupňa) sa môžete dočítať napríklad v [J, Š].

Príklad 9.5.2. Využijeme rovnosti

$$\begin{pmatrix} F_{n+1} \\ F_n \end{pmatrix} = A^{n-1} \begin{pmatrix} F_2 \\ F_1 \end{pmatrix},$$
$$\begin{pmatrix} F_n \\ F_{n-1} \end{pmatrix} = A^{n-2} \begin{pmatrix} F_2 \\ F_1 \end{pmatrix},$$
$$\vdots$$
$$\begin{pmatrix} F_2 \\ F_1 \end{pmatrix} = I \begin{pmatrix} F_2 \\ F_1 \end{pmatrix}.$$

Ich sčítaním dostaneme

$$\begin{pmatrix} F_2 + \cdots + F_{n+1} \\ F_1 + \cdots + F_n \end{pmatrix} = (I + A + \cdots + A^{n-1}) \begin{pmatrix} F_2 \\ F_1 \end{pmatrix}.$$

Všimnime si, že platí

$$(A - I)(I + A + \cdots + A^{n-1}) = A^n - I,$$

a teda

$$I + A + \cdots + A^{n-1} = (A - I)^{-1}(A^n - I).$$

Ľahko vypočítame, že

$$(A-I)^{-1} = \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} = A.$$

(Tento fakt sme mohli spozorovať aj z charakteristickej rovnice. Podľa Cayley-Hamiltonovej vety 9.2.14 totiž musí platiť $ch_A(A) = A^2 - A - I = 0$, z čoho dostaneme $A(A-I) = I$ a $(A-I)^{-1} = A$.)

Máme teda

$$\begin{pmatrix} F_2 + \cdots + F_{n+1} \\ F_1 + \cdots + F_n \end{pmatrix} = (A-I)^{-1}(A^n - I)\begin{pmatrix} F_2 \\ F_1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} F_{n+2} - 1 \\ F_{n+1} - 2 \end{pmatrix} = \begin{pmatrix} F_{n+3} - 1 \\ F_{n+2} - 1 \end{pmatrix}.$$

Vypočítali sme teda, že

$$\sum_{k=1}^{n} F_k = F_{n+2} - 1.$$

Poznamenajme, že identitu odvodenú v predchádzajúcom príklade by sme mohli ľahko overiť indukciou. Za výhodu uvedeného prístupu možno považovať to, že takto sme boli schopní tento vzorec objaviť – pri dôkaze indukciou musíme najprv uhádnuť ako vzorec vyzerá. (Táto výhoda je možno zjavnejšia pri odvodzovaní niektorých komplikovanejších rovností; my sme sa uspokojili s týmto veľmi jednoduchým príkladom.)

Cvičenia

Úloha 9.5.1. Nájdite predpis pre n-tý člen danej rekurentnej postupnosti:
a) $a_n = 5a_{n-1} - 6a_{n-2}$, pričom $a_0 = 4$ a $a_1 = 7$;
b) $a_n = a_{n-1} + 2a_{n-2}$, pričom $a_0 = 4$ a $a_1 = 5$;
c) $a_n = 6a_{n-1} - 9a_{n-2}$, pričom $a_0 = 2$ a $a_1 = 3$;
d) $a_n = 2a_{n-1} - 2a_{n-2}$, pričom $a_0 = a_1 = 2$.
[Výsledky: a) $5.2^n - 3^n$; b) $3.2^n + (-1)^n$; c) $2.3^n - n.3^n$ d) $(1+i)^n + (1-i)^n$.]

Viaceré cvičenia v tejto časti sú zamerané na dôkaz niektorých identít týkajúcich sa Fibonacciho čísel pomocou matíc. Pre mnohé z nich sa dajú dokázať podobné výsledky aj pre ľubovoľné rekurentné postupnosti druhého rádu. Na porovnanie obtiažnosti si môžete niektoré z nich skúsiť dokázať aj matematickou indukciou, dosadením vzorca (9.17) pre n-tý člen alebo inými spôsobmi (generujúce funkcie, rôzne metódy na výpočet súm, ...).

{aplikrekurcvic:FIBDOUBL

Úloha 9.5.2. Ukážte, že pre maticu $A = \left(\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ a pre $n \in \mathbb{N}$ platí $A^n = \begin{pmatrix} F_{n+1} & F_n \\ F_n & F_{n-1} \end{pmatrix}$.

Na základe toho ukážte, že:
a) $F_{n+1}F_{n-1} - F_n^2 = (-1)^n$ (Cassiniho identita)
b) $F_{m+n} = F_mF_{n+1} + F_{m-1}F_n$ (konvolučná vlastnosť)
c) $F_{2n} = F_n(F_{n+1} + F_{n-1})$ a $F_{2n+1} = F_{n+1}^2 + F_n^2$.

Vedeli by ste pomocou výsledkov z časti c) navrhnúť efektívny algoritmus na výpočet n-tého Fibonacciho čísla.

Úloha 9.5.3. Dokážte, že $\sum_{k=0}^{n} F_{2k+1} = F_{2(n+1)}$ a $\sum_{k=0}^{n} F_{2k} = F_{2n+1} - 1$.

Úloha 9.5.4. Ukážte, že $\sum_{k=0}^{n} F_k^2 = F_nF_{n+1}$. (Hint k maticovému odvodeniu: Čomu sa rovná $(A^k)^2$? Iná možnosť: Použiť nejako úlohu 9.5.2c).)

Úloha 9.5.5. Ukážte, že $F_{2n} = \sum_{k=0}^{n} \binom{n}{k} F_k$. (Hint k maticovému odvodeniu: Skúste využiť rovnosť $A^2 = I + A$.)

Úloha 9.5.6. Nájdite vzorec pre F_{j+k+l} a pre F_{3n}.

V zostávajúcich úlohách budeme používať aj iné postupy ako použitie matíc. Cieľom je ukázať niektoré zaujímavé vlastnosti Fibonacciho postupnosti.

Úloha 9.5.7. a) Vyjadrite F_{n+3} a F_n pomocou F_{n+1} a F_{n+2}.
b) Ukážte, že $F_{n+2}^2 - F_{n+1}^2 = F_n F_{n+3}$ platí pre ľubovoľné $n \in \mathbb{N}$.

Úloha 9.5.8. Lucasova postupnosť je postupnosť určená predpisom $L_{n+1} = L_n + L_{n-1}$ a podmienkami $L_0 = 2$, $L_1 = 1$.
a) Nájdite vyjadrenie n-tého člena Lucasovej postupnosti.
b) Ukážte, že $L_n = F_{n-1} + F_{n+1}$.

Úloha 9.5.9. Ukážte, že pre Fibonacciho postupnosť platí:
a) $F_n \mid F_{kn}$,
b) $(F_n, F_{n+1}) = 1$,
c) $(F_{kn+r}, F_n) = (F_r, F_n)$,
d) $(F_m, F_n) = F_{(m,n)}$.

Koľko delení so zvyškom je potrebné vykonať, ak hľadáme (F_n, F_{n-1}) pomocou Euklidovho algoritmu?

9.5.2 Sústavy lineárnych homogénnych diferenciálnych rovníc

V tejto časti sa budeme zaoberať sústavami lineárnych homogénnych diferenciálnych rovníc. Pre jednoduchosť sa opäť obmedzíme na sústavy 2 rovníc. Sú to teda sústavy tvaru

$$\begin{aligned} x'(t) &= ax(t) + by(t) \\ y'(t) &= cx(t) + dy(t) \end{aligned}$$

pričom $a, b, c, d \in \mathbb{C}$ sú konštanty, $x(t), y(t)$ sú hľadané funkcie reálnej premennej a všetky derivácie vystupujúce v sústave chápeme ako derivácie podľa t.

Stručnejšie budeme predchádzajúcu sústavu zapisovať ako

{aplik:EQDIF}

$$\begin{aligned} x' &= ax + by \\ y' &= cx + dy \end{aligned} \tag{9.18}$$

Budeme využívať to, že vieme, že riešením diferenciálnej rovnice

$$u' = ku$$

pre dané $k \in \mathbb{R}$ je funkcia

$$u(t) = Ce^{kt},$$

pričom $C = u(0)$. (Ľahko overíme, že táto funkcia danej rovnici skutočne vyhovuje. Akonáhle je dané $u(0)$, je tým už funkcia u jednoznačne určená.)

Pozrime sa najprv na nasledujúci jednoduchý príklad, kde sa dá uhádnuť ako môžeme sústavu previesť na tvar, ktorý už vieme riešiť.

Príklad 9.5.3. Riešme sústavu

$$\begin{aligned} x' &= x + 2y \\ y' &= 2x + y \end{aligned}$$

Upravme túto sústavu tak, že jednotlivé rovnice sčítame a odčítame. Dostaneme tak rovnice

$$x' + y' = 3x + 3y$$
$$x' - y' = -x + y$$

čiže

$$(x + y)' = 3(x + y)$$
$$(x - y)' = -(x - y)$$

Z týchto 2 rovníc máme

$$x + y = c_1 e^{3t}$$
$$x - y = c_2 e^{-t}$$

a po vyjadrení x a y dostaneme riešenie

$$x = a_1 e^{3t} + a_2 e^{-t}$$
$$y = a_1 e^{3t} - a_2 e^{-t}$$

(Kvôli „krajšiemu“ zápisu sme zaviedli nové konštanty $a_{1,2}$ také, že $c_1 = 2a_1$, $c_2 = 2a_2$.)

Dosadením do pôvodnej rovnice sa ľahko presvedčíme, že je to skutočne riešenie pôvodnej sústavy.

Opäť, podobne ako v prípade rekurencií, vieme sústavu (9.18) zapísať maticovo ako

$$(x', y') = (x, y)A \tag{9.19}$$ {aplik:EQDIFMAT}

pričom

$$A = \begin{pmatrix} a & c \\ b & d \end{pmatrix}.$$

Všimnime si, čo sme vlastne v predchádzajúcom príklade urobili. Najprv sme zaviedli nové funkcie u a v (zmena súradníc), pre ktoré sme sústavu vedeli riešiť, a potom sme urobili opačnú zmenu súradníc, aby sme dostali riešenie pre pôvodné funkcie. To presne zodpovedá podobnosti matíc – pri nej tiež robíme zmenu súradníc P a zmenu súradníc opačným smerom P^{-1}. Na tento problém sa teda možno pozerať tak, že hľadáme maticu podobnú matici A, pričom chceme, aby táto matica (a tým pádom aj zodpovedajúca sústava) boli čo najjednoduchšie. Vhodným kandidátom by mohol byť Jordanov normálny tvar.

Čo dostaneme ak maticu P prevedieme na Jordanov normálny tvar? Máme potom

$$(x', y') = (x, y)P^{-1}JP$$

čiže

$$(x', y')P^{-1} = (x, y)P^{-1}J.$$

Uvažujme nové funkcie $u(t)$ a $v(t)$ určené ako

$$(u, v) = (x, y)P^{-1}.$$

Z nich vieme vyrátať hľadané funkcie x a y a získali sme pre ne o čosi jednoduchšiu sústavu

$$(u', v') = (u, v)J.$$

Rozmyslime si aspoň najjednoduchší prípad, matica J je diagonálna a obe vlastné hodnoty sú reálne čísla. Potom sústava, ktorej majú vyhovovať u a v je

$$u' = \lambda_1 u$$
$$v' = \lambda_2 v$$

a jej riešenie je

$$u = c_1 e^{\lambda_1 t}$$
$$v = c_2 e^{\lambda_2 t}$$

Riešenia pôvodnej sústavy dostaneme vynásobením sprava maticou P.

$$(x(t), y(t)) = (c_1 e^{\lambda_1 t}, c_2 e^{\lambda_2 t})P$$

Keď riadky matice P (čo sú súčasne vlastné vektory matice A) označíme ako $\vec{\alpha}_1$, $\vec{\alpha}_2$, tak predchádazjúca rovnosť znamená

$$(x(t), y(t)) = c_1 e^{\lambda_1 t}\vec{\alpha}_1 + c_2 e^{\lambda_2 t}\vec{\alpha}_2,$$

čiže riešenia pôvodnej sústavy sú lineárne kombinácie riešení $e^{\lambda_1 t}\vec{\alpha}_1$ a $e^{\lambda_2 t}\vec{\alpha}_2$.

Viac o riešení sústav lineárnych diferenciálnych rovníc a tiež spôsob, akým sa riešia prípady komplexných alebo viacnásobných vlastných hodnôt, možno nájsť napríklad v [GŠŠ].

9.6 PageRank algoritmus

V tejto časti chceme z hľadiska lineárnej algebry popísať PageRank algoritmus, ktorý sa dá použiť na ohodnotenie dôležitosti webových stránok. Je dôležitý pri vyhľadávaní na zoradenie výsledkov. Podrobnejšie o tejto téme sa môžete dozvedieť napríklad v článku [BL] alebo v knihe [LM] (táto podkapitola spracovaná z týchto dvoch zdrojov a z [Me]). Ďalším dostupným zdrojom je bakalárska práca [Mi], v ktorej nájdete dokázaný i všeobecný prípad niektorých tvrdení, ktoré tu dokážeme iba v zjednodušenom prípade, že ide o matice podobné s diagonálnou maticou.

Autormi tohoto algoritmu sú zakladatelia firmy Google Sergey Brin a Larry Page, začali na ňom pracovať v druhej polovici 90-tych rokov. Oproti dovtedy používaným spôsobom hodnotenia dôležitosti nájdených výsledkov je významným rozdielom to, že dôležitosť sa tu neurčuje na základe obsahu stránky, ale podľa hypertextových odkazov na danú stránku z iných stránok. Približne v tom istom čase navrhol Jon Kleinberg algoritmus HITS, ktorý bol do značnej miery podobný, nesnažil sa ho však komerčne využiť. V súčasnosti niektoré vyhľadávače používajú tento algoritmus. Hlavný rozdiel medzi oboma algoritmami je v tom, že HITS okrem liniek smerujúcich na danú stránku zohľadňuje aj linky, ktoré smerujú z nej.

V tejto súvislosti treba spomenúť, že PageRank nie je jediné kritérium, na základe ktorého Google vytvára poradie nájdených výsledkov.

Základná idea PageRank algoritmu sa dá popísať veľmi jednoducho. Ak dôležitosť posudzujeme podľa toho, koľko stránok sa na ňu odkazuje, môžeme sa na to pozrieť ako na „hlasovanie“. Každá stránka dostane 1 hlas a ak na nej je n liniek, tak s váhou $1/n$ hlasuje za dôležitosť stránok, na ktoré sa odkazuje. Teda každá stránka má rovnocenné „hlasovacie právo“, a hlas stránky sa rozdelí medzi tie, na ktoré odkazuje.

Znamená to, že dôležitosť stránky bude úmerná tomu, koľko liniek na ňu odkazuje.

Práve popísaný výpočet môžeme popísať veľmi jednoducho ako

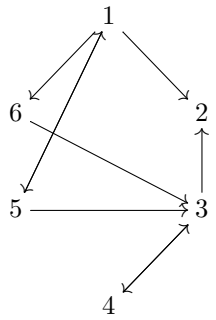
$$\vec{x} = \vec{e}A,$$

kde vektor $\vec{x} = (x_1, \ldots, x_n)$ obsahuje ohodnotenia stránok, $\vec{e} = (1, 1, \ldots, 1)$ a matica A je určená ako

$$a_{ij} = \begin{cases} \frac{1}{n_i} & \text{ak } i\text{-ta stránka obsahuje odkaz na } j\text{-tu,} \\ 0 & \text{inak.} \end{cases}$$

pričom n_i označuje počet liniek na i-tej stránke.

Napríklad pre web naznačený v nasledujúcom diagrame



vyzerá matica A ako

$$A = \begin{pmatrix} 0 & \frac{1}{3} & 0 & 0 & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}.$$

Poznamenajme, že aj v reálnych aplikáciach bude táto matica obsahovať veľa núl – ide o takzvanú *riedku maticu.* To prináša dve výhody – maticu možno uložiť (pomocou vhodnej dátovej štruktúry) tak, aby nezaberala zbytočne veľa priestoru. Takisto niektoré výpočty, napríklad násobenie takouto maticou, možno implementovať tak, že budú oveľa rýchlejšie ako pre matice, ktoré majú skoro všetky hodnoty nenulové.

Nedostatok hodnotenia, ktoré získame doteraz popísaným spôsobom, je v tom, že sme rovnakú vážnosť prikladali linkám z menej dôležitých stránok ako linkám z významných stránok. Ako však odlíšiť významné a menej významné stránky a zabezpečiť, aby hlas dôležitých stránok zavážil viac? Môžeme jednoducho zobrať práve vypočítaný odhad pre dôležitosť ako váhy stránok a znovu urobiť to isté. To znamená, že budeme postupne počítať

$$\begin{aligned} \vec{x}_0 &= \vec{e} \\ \vec{x}_1 &= \vec{x}_0 A \\ \vec{x}_2 &= \vec{x}_1 A = \vec{x}_0 A^2 \\ &\ldots \\ \vec{x}_n &= \vec{x}_{n-1} A = \vec{x}_0 A^n \end{aligned}$$

V prípade, že váhový vektor $\vec{x}_n$ konverguje k nejakej limite, je pomerne prirodzené túto limitu považovať za ohodnotenie významnosti jednotlivých stránok.

Spomeňme ešte iný pohľad, ako možno interpretovať tieto výpočty. Dá sa na to hľadieť tak, že popisujeme browsovanie náhodného surfera, ktorý sa správa tak, že z niektorej stránky si náhodne vyberie ľubovoľnú linku a na tú stránku ide ďalej. Ak stránka neobsahuje žiadne linky, tak si náhodne vyberie ľubovoľnú stránku a zase browsuje ďalej. Takto možno považovať vektor, ku ktorému konverguje $\vec{x}_n$, chápať ako hodnotu pravdepodobnosti, že sa v danom

okamihu surfer nachádza na danej stránke za predpokladu, že ho necháme surfovať neobmedzene dlho. (Aby sme boli presní, ak chceme použiť túto pravdepodobnostnú interpretáciu, použijeme $\vec{x}_0 = \frac{1}{n}\vec{e}$ alebo iný vektor, ktorý má kladné súradnice a ich súčet je 1 – pretože táto podmienka musí platiť pre pravdepodobnosť.) Práve tento pohľad dáva do súvisu PageRank algoritmus s markovovskými reťazcami. Tie sa študujú v teórii stochastických procesov, čo je matematická oblasť patriaca do teórie pravdepodobnosti.

Teraz sa budeme zaoberať tým, pre aké matice A si môžeme byť istý, že vektor $\vec{x}_n$ bude skutočne konvergovať k nejakej hodnote. Ako uvidíme, aby to fungovalo v praxi, bude potrebné maticu, ktorú sme práve popísali, ešte trochu upraviť.

Najprv si všimnime, že ak $\vec{x}_0 A^n$ konverguje k nejakému nenulovému vektoru, musí to byť vlastný vektor matice A prislúchajúci k vlastnej hodnote 1.

Ak totiž platí

$$\vec{x} = \lim_{n\to\infty} \vec{x}_0 A^n,$$

tak máme aj

$$\vec{x}A = \lim_{n\to\infty} \vec{x}_0 A^{n+1} = \vec{x}.$$

(Tu sme využili fakt, že v priestore $\mathbb{R}^n$ je každé lineárne zobrazenie spojité.[4] Lineárne zobrazenie je vlastne násobenie maticou A.)

Všimnime si, že súčet prvkov v každom nenulovom riadku matici A je rovný 1 (prvá iterácia bola demokratická – každá stránka mala rovnaké „hlasovacie právo“).

Definícia 9.6.1. Matica A sa nazýva *riadkovo stochastická,* ak súčet prvkov jej ľubovoľného riadku je 1.

{google:TVR1JEVC}

Tvrdenie 9.6.2. *Ak matica A je riadkovo stochastická, tak číslo 1 je jej vlastnou hodnotou.*

Dôkaz. Stačí si všimnúť, že súčet stĺpcov matice $A - I$ je $\vec{0}$, čo znamená, že jej stĺpce sú lineárne závislé a táto matica je teda singulárna. □

Všimnime si ešte jednu vlastnosť riadkovo stochastických matíc – lineárne zobrazenie prislúchajúce takejto matici nemení súčet jednotlivých zložiek vektora.

Tvrdenie 9.6.3. *Nech A je riadkovo stochastická matica, $\vec{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ a $\vec{y} = (y_1, \ldots, y_n) = \vec{x}A$. Potom platí $\sum_{i=1}^{n} y_i = \sum_{i=1}^{n} x_i$.*

Dôkaz. Máme

$$y_i = \sum_{j=1}^{n} a_{ji} x_j.$$

Sčítaním týchto rovníc dostaneme

$$\sum_{i=1}^{n} y_i = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ji}x_j = \sum_{j=1}^{n} x_j \sum_{i=1}^{n} a_{ji} = \sum_{j=1}^{n} x_j,$$

kde posledná rovnosť vyplýva z faktu, že súčet prvkov matice A v danom riadku je 1. □

[4] Stručné zdôvodnenie tohoto faktu: Ak označíme $\|A\|_{max} = \max_{i,j}|a_{ij}|$, tak očividne pre každý vektor $\vec{z}$ taký, že $\max|z_i| < \delta$ platí, že všetky súradnice vektora $\vec{z}A$ nepresahujú v absolútnej hodnote $\delta n\|A\|_{max}$. Takže ak máme dané $\varepsilon > 0$ a dvojicu vektorov takú, že pre všetky ich súradnice platí $|x_i - y_i| < \frac{\varepsilon}{n\|A\|_{max}}$, tak dostaneme $\vec{x}A - \vec{y}A = (\vec{x} - \vec{y})A < n\|A\|_{max}\frac{\varepsilon}{n\|A\|_{max}} = \varepsilon$.

Iný dôkaz. Fakt, že riadky majú súčet jedna, sa dá stručne jedným vzťahom vyjadriť ako

$$A\vec{e}^T = \vec{e}^T.$$

Všimnime si tiež, že súčet jednotlivých súradníc vektora je presne rovný skalárnemu súčinu $\vec{x}\vec{e}^T$. Potom dostaneme

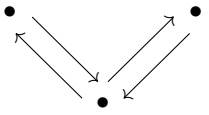
$$\vec{x}A\vec{e}^T = \vec{x}\vec{e}^T,$$

čo je presne dokazované tvrdenie. □

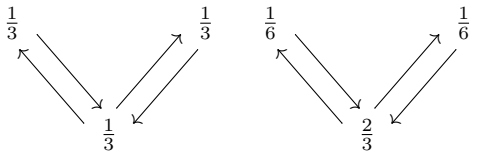
Z tvrdenia 9.6.2 teda vidíme, že keby sme v matici A nemali nulové riadky, mali by sme zaručenú existenciu aspoň jedného kadidáta na výsledné ohodnotenie. Aby sme dostali riadkovo stochastickú maticu, nahradíme každý nulový riadok vektorom $\frac{1}{n}\vec{e}$. (Na stránke bez liniek sa surfer rozhodne úplne náhodne pre ľubovoľnú stránku na internete.)

Zatiaľ však stále nemáme nijako zaručené, že vlastný vektor pre vlastné číslo 1 bude jediný a ani to, že k nemu budú naše vektory skutočne konvergovať.

Veľmi jednoduchý príklad ukazujúci, že tento proces nemusí konvergovať je takáto sieť



Pomerne ľahko vidíme, že ak na začiatku sú pravdepodobnosti výskytu v týchto troch vrcholoch $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$, tak v ďalšej iterácii dostaneme $(\frac{1}{6}, \frac{2}{3}, \frac{1}{6})$.



Zodpovedajúca matica je

$$\begin{pmatrix} 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \end{pmatrix}.$$

Ak skontrolujeme, že vlastné hodnoty sú 0 a ± 1, tak pre zodpovedajúcu diagonálnu maticu máme $D^n = \mathrm{diag}(1, 0, (-1)^n)$; teda aj tu vid9me striedanie dvoch hodnôt.

Ukážeme, že vlastný podpriestor prislúchajúci k vlastnej hodnote 1 bude jednorozmerný, v prípade, že matica A mala všetky prvky kladné.

Lema 9.6.4. *Nech matica A je riadkovo stochastická a $a_{ij} > 0$ pre $i, j = 1, \ldots, n$. Potom každý jej vlastný vektor prislúchajúci k vlastnej hodnote 1 má všetky súradnice rovnakého znamienka (všetky nezáporné alebo všetky nekladné).*

Dôkaz. V dôkaze použijeme fakt, že nerovnosť $|\sum_{i=1}^n y_i| \leq \sum_{i=1}^n |y_i|$ je ostrá pre každý vektor obsahujúci prvky so zmiešanými znamienkami.

Budeme postupovať sporom. Nech by $\vec{x}A = \vec{x}$ a vektor $\vec{x}$ má na niektorých súradniciach rôzne znamienka. Potom máme ostré nerovnosti

$$|x_i| = \left| \sum_{j=1}^n a_{ji} x_j \right| < \sum_{j=1}^n a_{ji} |x_j|.$$

Sčítaním týchto nerovností dostaneme

$$\sum_{i=1}^{n}|x_i| < \sum_{i=1}^{n}\sum_{j=1}^{n}a_{ji}|x_j| = \sum_{j=1}^{n}|x_j|\sum_{i=1}^{n}a_{ji} = \sum_{j=1}^{n}|x_j|.$$

(V poslednej rovnosti sme využili, že súčet prvkov j-teho riadku matice a je 1.)

Dostali sme nerovnosť

$$\sum_{i=1}^{n}|x_i| < \sum_{j=1}^{n}|x_j|,$$

čo je samozrejme spor. □

Veľmi podobným spôsobom ako v predchádzajúcej leme môžeme odvodiť nasledujúci fakt, ktorý bude neskôr užitočný pri dôkaze konvergencie použitej metódy.

Lema 9.6.5. *Nech matica A je riadkovo stochastická a $a_{ij} \geq 0$ pre $i, j = 1, \ldots, n$. Ak λ je jej vlastná hodnota, tak $|\lambda| \leq 1$.*

Dôkaz. Pre vlastnú hodnotu λ a príslušný vlastný vektor máme.

$$\lambda x_i = \sum_{j=1}^{n} a_{ji}x_j$$

$$|\lambda||x_i| = \left|\sum_{j=1}^{n} a_{ji}x_j\right| \leq \sum_{j=1}^{n} a_{ji}|x_j|$$

Sčítaním týchto nerovností dostaneme

$$|\lambda|\sum_{i=1}^{n}|x_i| \leq \sum_{i=1}^{n}\sum_{j=1}^{n}a_{ji}|x_j| = \sum_{j=1}^{n}|x_j|\sum_{i=1}^{n}a_{ji} = \sum_{j=1}^{n}|x_j|$$

$$|\lambda| \leq 1$$

(V poslednom kroku sme využili, že $\vec{x} \neq \vec{0}$, čiže aj $\sum_{j=1}^{n}|x_j| \neq 0$.) □

Lema 9.6.6. *Ak $\vec{\alpha}$, $\vec{\beta}$ sú lineárne nezávislé vektory, tak existujú koeficienty $c, d \in \mathbb{R}$ také, že $c\vec{\alpha} + d\vec{\beta}$ obsahuje prvky so zmiešanými znamienkami.*

Dôkaz. Ak pre vektor $\vec{\alpha} = (a_1, \ldots, a_n)$ platí $\sum_{i=1}^{n} a_i = 0$, tak tento vektor obsahuje zmiešané znamienka (keďže je nenulový) a stačí zvoliť $c = 1$ a $d = 0$. Podobne v prípade, že to platí pre vektor $\vec{\beta}$.

Ak žiadny z vektorov nedáva nulový súčet, stačí nám zvoliť c a d tak, aby $c\sum_{i=1}^{n} a_i + d\sum_{i=1}^{n} b_i = 0$. Lineárna nezávislosť zabezpečí to, že vektor $c\vec{\alpha} + d\vec{\beta}$ je nenulový, dostávame teda, že znamienka jeho súradníc nemôžu byť všetky rovnaké. □

Dôsledok 9.6.7. *Ak A je riadkovo stochastická matica, ktorej prvky sú kladné, tak podpriestor tvorený vlastnými vektormi k vlastnému číslu 1 je jednorozmerný.*

Aby matica A mala všetky členy kladné, zabezpečíme tak, že namiesto pôvodnej matice použijeme maticu

$$G = \alpha A + (1 - \alpha)\frac{1}{n}\vec{e}^T\vec{e},$$

kde $\alpha \in (0,1)$. Všimnime si, že aj matica G je riadkovo stochastická.

V predchádzajúcej rovnosti sme skombinovali maticu A s maticou $\frac{1}{n}\vec{e}^T\vec{e}$, ktorá na každom svojom mieste obsahuje hodnotu $\frac{1}{n}$. Zodpovedá to tomu, že náhodne sa pohybujúci surfer sa v nejakom percente prípadov (určenom koeficientom $1-\alpha$) rozhodne nepokračovať linkou zo stránky, na ktorej sa nachádza, ale vyberie si novú stránku úplne náhodne. (Z takejto interpretácie vidno aj to, že „náhodnému surferovi“ týmto zabránime zacykliť sa. Cykly, podobne ako stránky, z ktorých nevychádzajú linky, by spôsobili, že pravdepodobnosť výskytu stránky pri náhodnom surfovaní – a teda jej váha – sa nám pri iteráciách postupne naakumuluje vo vrcholoch cyklu.) Súčasne sme touto zmenou nezvýhodnili niektorú stránku oproti iným (všade sme pripočítali rovnakú hodnotu).

Neskôr ukážeme, že nastavenie parametra α ovplyvňuje rýchlosť konvergencie k hľadanému riešeniu a tým aj počet krokov potrebných na dosiahnutie dostatočnej presnosti. Okrem toho na tomto parametre závisí aj citlivosť metódy na zmenu matice A.

Vďaka tomu, že súčet použitých koeficientov je 1 a obe matice sú riadkovo stochastické, aj výsledná matica je riadkovo stochastická.

Na prvý pohľad by sa mohlo zdať, že sme touto zmenou matice stratili výhodu, ktorú nám prinášala riedkosť matice A. Keďže sme ale pripočítali maticu, ktorá má všetky prvky rovnaké, je jasné, že ju nemusíme mať v pamäti uloženú po jednotlivých prvkoch a aj to, že násobiť takouto maticou sa tiež dá dosť jednoducho.

Už teda vieme, že ak bude $\vec{x_0}A^n$ konvergovať, môže konvergovať jedine k hľadanej vlastnej hodnote. Pre jednoduchosť konvergenciu overíme len pre diagonalizovateľné matice. Na to najprv dokážeme vetu o spektrálnom rozklade diagonalizovateľnej matice.

Veta 9.6.8. *Ak je matica A typu $n \times n$ podobná s diagonálnou maticou* $\operatorname{diag}(\lambda_1, \ldots, \lambda_n)$*, tak existujú matice $G_1, \ldots, G_n$ také, že platí*

$$A = \lambda_1 G_1 + \lambda_2 G_2 + \cdots + \lambda_n G_n$$

a súčasne

$$G_1 + \cdots + G_n = I,$$

pre každé i platí $G_i^2 = G_i$ a pre $i \neq j$ platí $G_i G_j = 0$.

Všimnime si, že z podmienok uvedených v predchádzajúcej vete vyplýva

$$A^k = \lambda_1^k G_1 + \lambda_2^k G_2 + \cdots + \lambda_n^k G_n.$$

Dôkaz. Podľa predpokladu existuje regulárna matica P taká, že $PAP^{-1} = D$, čiže $P^{-1}DP = A$. Označme stĺpce matice P^{-1} ako $\vec{\alpha}_1, \ldots, \vec{\alpha}_n$ a riadky matice P ako $\vec{\beta}_1, \ldots, \vec{\beta}_n$.

$$P^{-1} = (\vec{\alpha}_1^T, \ldots, \vec{\alpha}_n^T) \qquad P = \begin{pmatrix} \vec{\beta}_1 \\ \vdots \\ \vec{\beta}_n \end{pmatrix}$$

Z rovnosti $P^{-1}DP = A$ potom dostaneme (podobným spôsobom ako v dôkaze (9.5))

$$A = (\vec{\alpha}_1^T, \ldots, \vec{\alpha}_n^T) \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \begin{pmatrix} \vec{\beta}_1 \\ \vdots \\ \vec{\beta}_n \end{pmatrix} = \lambda_1 \vec{\alpha}_1^T \vec{\beta}_1 + \cdots + \lambda_n \vec{\alpha}_n^T \vec{\beta}_n.$$

Označme $G_i = \vec{\alpha}_i^T \vec{\beta}_i$. Takto definované G_i sú matice $n \times n$ a platí pre ne $A = \lambda_1 G_1 + \lambda_2 G_2 + \cdots + \lambda_n G_n$.

Overme, že platia aj ostatné rovnosti uvedené vo vete. Rovnosť $I = P^{-1}P$ môžeme prepísať ako

$$I = (\vec{\alpha}_1^T, \ldots, \vec{\alpha}_n^T) \begin{pmatrix} \vec{\beta}_1 \\ \vdots \\ \vec{\beta}_n \end{pmatrix} = \vec{\alpha}_1^T \vec{\beta}_1 + \cdots + \vec{\alpha}_n^T \vec{\beta}_n = G_1 + \cdots + G_n.$$

Ak násobíme PP^{-1}, tak dostaneme

$$I = \begin{pmatrix} \vec{\beta}_1 \\ \vdots \\ \vec{\beta}_n \end{pmatrix} (\vec{\alpha}_1^T, \ldots, \vec{\alpha}_n^T) = \begin{pmatrix} \vec{\beta}_1\vec{\alpha}_1^T & \ldots & \vec{\beta}_1\vec{\alpha}_n^T \\ \ldots & \ldots & \ldots \\ \vec{\beta}_n\vec{\alpha}_1^T & \ldots & \vec{\beta}_n\vec{\alpha}_n^T \end{pmatrix}$$

Porovnaním oboch matíc dostaneme $\vec{\beta}_i\vec{\alpha}_i^T = 1$ a $\vec{\beta}_i\vec{\alpha}_i^T = 0$ pre $i \neq j$. To znamená, že

$$G_i^2 = \vec{\alpha}_i^T(\vec{\beta}_i\vec{\alpha}_i^T)\vec{\beta}_i = \vec{\alpha}_i^T\vec{\beta}_i = G_i$$

a

$$G_iG_j = \vec{\alpha}_i^T(\vec{\beta}_i\vec{\alpha}_j^T)\vec{\beta}_j = \vec{\alpha}_i^T.0.\vec{\beta}_j = 0.$$

□

Z vety o spektrálnom rozklade dostaneme, v prípade, že naša matica je diagonalizovateľná,

$$A^k = G_1 + \lambda_2^k G_2 + \cdots + \lambda_n^k G_n$$

a

$$\vec{x_0}A^k = \vec{x_0}G_1 + \lambda_2^k\vec{x_0}G_2 + \cdots + \lambda_n^k\vec{x_0}G_n$$

Z nasledujúceho tvrdenia vyplynie, že vlastné hodnoty $\lambda_2, \ldots, \lambda_n$ sú v absolútnej hodnote ostro menšie ako 1, čiže pre ne platí $\lambda_i^k \longrightarrow 0$. To znamená, že uvedený výraz skutočne konverguje.

Z predchádzajúcej rovnice vidíme tiež, že rýchlosť konvergencie závisí od vlastnej hodnoty, ktorá je druhá najväčšia v absolútnej hodnote. Nasledujúce tvrdenie súčasne ukazuje, že veľkosť tejto vlastnej hodnoty (a teda aj rýchlosť konvergencie) závisí od parametra α.

Tvrdenie 9.6.9. *Nech A je riadkovo stochastická matica a jej vlastné čísla sú $1, \lambda_2, \ldots, \lambda_n$. Potom matica*

$$G = \alpha A + (1-\alpha)\frac{1}{n}\vec{e}^T\vec{e}$$

má vlastné hodnoty $1, \alpha\lambda_2, \ldots, \alpha\lambda_n$.

Dôkaz. Skutočnosť, že súčty v riadkoch matice A sú rovné 1, je ekvivalentná s rovnosťou

$$A\vec{e}^T = \vec{e}^T.$$

Nech Q je ľubovoľná regulárna matica, ktorej prvý stĺpec je $\vec{e}^T$.

$$Q = (\vec{e}^T, X)$$

Ak prvý riadok inverznej matice Q^{-1} označíme $\vec{y}$, tak máme

{google:EQINVQQ}

$$Q^{-1}Q = \begin{pmatrix} \vec{y} \\ Y \end{pmatrix} (\vec{e}^T, X) = \begin{pmatrix} \vec{y}\vec{e}^T & \vec{y}X \\ Y\vec{e}^T & YX \end{pmatrix} = \begin{pmatrix} 1 & \vec{0} \\ \vec{0}^T & I \end{pmatrix}. \tag{9.20}$$

(Všimnime si, že X a Y nie sú štvorcové matice, takže rovnosť $YX = I$ vyplývajúca z predchádzajúcej rovnosti neznamená, že tieto matice sú navzájom inverzné.)

Ak vynásobíme týmito maticami maticu A, tak dostaneme

$$Q^{-1}AQ = \begin{pmatrix} \vec{y} \\ Y \end{pmatrix} A(\vec{e}^T, X) = \begin{pmatrix} \vec{y}A \\ YA \end{pmatrix} (\vec{e}^T, X) = \begin{pmatrix} \vec{y}A\vec{e}^T & \vec{y}AX \\ YA\vec{e}^T & YAX \end{pmatrix}$$

Súčasne použitím (9.20) dostaneme

$$\vec{y}A\vec{e}^T = \vec{y}\vec{e}^T = 1$$

a

$$YA\vec{e}^T = Y\vec{e}^T = \vec{0}^T.$$

Dostali sme teda

$$Q^{-1}AQ = \begin{pmatrix} 1 & \vec{y}AX \\ \vec{0}^T & YAX \end{pmatrix}.$$

Z poslednej rovnosti a z toho, čo vieme o vlastných hodnotách matice A, vyplýva, že matica YAX je podobná s hornou trojuholníkovou maticou, ktorá má na diagonále hodnoty $\lambda_2, \ldots, \lambda_n$. (Stačí si všimnúť, že pre hornú trojuholníkovú maticu je charakteristický polynóm jednoznačne určený hodnotami na diagonále. Jordanov normálny tvar ľubovoľnej matice je horná trojuholníková matica.)

Vypočítajme teraz to isté pre maticu $\vec{e}^T\vec{e}$. Dostaneme

$$Q^{-1}\vec{e}^T\vec{e}Q = \begin{pmatrix} \vec{y} \\ Y \end{pmatrix} \vec{e}^T\vec{e}(\vec{e}^T, X) = \begin{pmatrix} \vec{y}\vec{e}^T \\ Y\vec{e}^T \end{pmatrix} (\vec{e}\vec{e}^T, \vec{e}X) = \begin{pmatrix} 1 \\ \vec{0}^T \end{pmatrix} (n, \vec{e}X) = \begin{pmatrix} n & \vec{e}X \\ \vec{0}^T & 0 \end{pmatrix}$$

Pre maticu G potom dostaneme

$$Q^{-1}GQ = \alpha Q^{-1}AQ + (1-\alpha)\frac{1}{n}Q^{-1}\vec{e}^T\vec{e}Q =$$
$$\alpha \begin{pmatrix} 1 & \vec{y}AX \\ \vec{0}^T & YAX \end{pmatrix} + (1-\alpha) \begin{pmatrix} 1 & \frac{1}{n}\vec{e}X \\ \vec{0}^T & 0 \end{pmatrix} = \begin{pmatrix} 1 & \alpha\vec{y}AX + (1-\alpha)\frac{1}{n}\vec{e}X \\ \vec{0}^T & \alpha YAX \end{pmatrix}$$

Matica αYAX je podobná s hornou trojuholníkovou maticou, ktorá má na diagonále prvky $\alpha\lambda_1, \ldots, \alpha\lambda_n$, z čoho už vyplýva dokazované tvrdenie. □

Hodnota α, ktorú Google reálne používa, je $0,85$. Vďaka tomu veľkosť druhej najväčšej vlastnej hodnoty je nanajvýš $0,85$. Všimnime si, že čím menšiu hodnotu α zvolíme, tým rýchlejšie bude táto metóda konvergovať, spôsobíme tým však súčasne to, že nad maticou A popisujúcou vzhľad webu prevládne matica $\frac{1}{n}\vec{e}\vec{e}^T$, ktorú sme pridali umelo na zabezpečenie konvergencie.

Kapitola 10

Symetrické polynómy

10.1 Základná veta o symetrických polynómoch

Symetrické polynómy som robil presne podľa [KGGS]. Nebudem tu preto k ním písať poznámky. Jediné miesto, ktoré som robil inak, bol dôkaz jednoznačnosti v základnej vete o symetrických polynómoch (veta 5.8.3). Argument, ktorý som použil, je zhruba rovnaký ako v [T, Theorem 8.7] – snažil som sa ho tu rozpísať o čosi podrobnejšie.

Dôkaz jednoznačnosti. Máme ukázať, že ak pre nejaké polynómy p, q v n-premenných platí

$$p(A_1(x_1,\ldots,x_n),\ldots,A_n(x_1,\ldots,x_n)) = q(A_1(x_1,\ldots,x_n),\ldots,A_n(x_1,\ldots,x_n))$$

tak sa polynómy p a q rovnajú.

Namiesto predchádzajúceho zdĺhavého zápisu budeme písať stručnejšie $p(A_1,\ldots,A_n)$, vždy sa tým myslí, že ide o polynómy v premenných $x_1,\ldots,x_n$.

Najprv ukážeme, že tvrdenie platí v prípade, že jeden z polynómov je nulový, t.j. platí

{sympoly:IMPNUL}
$$t(A_1,\ldots,A_n) = 0 \Rightarrow t(x_1,\ldots,x_n) = 0. \tag{10.1}$$

(Obidve rovnosti v predchádzajúcom riadku chápeme ako rovnosti polynómov – čiže musia sa zhodovať príslušné koeficienty.)

Postupujme sporom. Predpokladajme, že v polynóme $t(x_1,\ldots,x_n)$ zapísanom v normálnom tvare máme nejaké nenulové členy. Pripomeňme, že z nenulového členu $ax_1^{k_1}\ldots x_n^{k_n}$ takto dostaneme polynóm $aA_1^{k_1}\ldots A_n^{k_n}$, ktorý má podľa lemy 5.8.3 vedúci člen $ax_1^{k_1+k_2\cdots+k_n}x_2^{k_2\cdots+k_n}\ldots x_n^{k_n}$.

Nie je ťažké overiť, že priradenie $(k_1,\ldots,k_n)\mapsto(k_1+k_2\cdots+k_n,k_2\cdots+k_n,\ldots,k_n)$ je injektívne. To znamená, že člen s rovnakým systémom exponentov nemôžeme dostať ako vedúci člen z niektorého z členov polynómu $t(x_1,\ldots,x_n)$.

Máme však dostať $t(A_1,\ldots,A_n)=0$, to znamená, že člen $ax_1^{k_1+k_2\cdots+k_n}x_2^{k_2\cdots+k_n}\ldots x_n^{k_n}$ sa musí niekde vykrátiť. Z predchádzajúceho vyplýva, že sa môže vykrátiť len s členom polynómu $bA_1^{l_1}\ldots B_n^{l_n}$, ktorého vedúci člen má (v lexikografickom usporiadaní) ostro vyšší systém exponentov. Vedúci člen tohto polynómu sa opäť musí nejako vykrátiť a opäť to môže byť jedine s nejakým polynómom, ktorý má vyšší vedúci člen. Takto by sme museli postupovať do nekonečna, čo sa však nedá, lebo polynóm má iba konečne veľa členov.

Akonáhle máme dokázanú implikáciu (10.1), stačí ju použiť pre rozdiel polynómov $p-q$ a dostaneme dokazované tvrdenie. □

ic:ULOSYMPOL}

Cvičenia

Úloha 10.1.1. Zapíšte daný symetrický polynóm pomocou základných symetrických polynómov:
a) $f(x_1,x_2,x_3)=\sum x_1^3x_2=x_1^3x_2+x_1x_2^3+x_1^3x_3+x_1x_3^3+x_2^3x_3+x_2x_3^3$
b) $f(x_1,x_2,x_3)=\sum x_1^2x_2=x_1^2x_2+x_1x_2^2+x_1^2x_3+x_1x_3^2+x_2^2x_3+x_2x_3^2$
c) $f(x_1,x_2,x_3)=(x_1+x_2)(x_1+x_3)(x_2+x_3)$
d) $f(x_1,x_2,x_3)=x_1^2+x_2^2+x_3^2$
e) $f(x_1,x_2,x_3)=x_1^3+x_2^3+x_3^3$
[Výsledky: a) $f=A_1^2A_2-2A_2^2-A_1A_3$; b) $f=A_1A_2-3A_3$; c) $A_1A_2-A_3$; d) $A_1^2-2A_2$; e) $A_1^3-3A_1A_2+3A_3$]

Úloha 10.1.2. Je zadaný polynóm symetrický? Vyjadrite ho pomocou základných symetrických polynómov:
a) $x_1^3+x_2^3+x_3^3-3x_1x_2x_3$;
b) $x_1^4+x_2^4+x_3^4-2x_1^2x_2^2-2x_1^2x_3^2-2x_2^2x_3^2$;
c) $x_1^5x_2^2+x_1^2x_2^5+x_1^5x_3^2+x_1^2x_3^5+x_2^5x_3^2+x_2^2x_3^5$;
d) $(x_1^2+x_2^2)(x_1^2+x_3^2)(x_2^2+x_3^2)$;
e) $(2x_1-x_2-x_3)(2x_2-x_1-x_3)(2x_3-x_1-x_2)$;
f) $(x_1-x_2)^2(x_1-x_3)^2(x_2-x_3)^2$

Úloha 10.1.3. Ukážte, že pre symetrické polynómy v n premenných, kde $n\geq 3$, platí $\sum x_1^2x_2=A_1A_2-3A_3$.

Úloha 10.1.4. Nech $x_{1,2,3}\in\mathbb{C}$ sú korene rovnice $x^3-5x+11=0$.
a) Aká je hodnota výrazu $(x_1-x_2)^2+(x_1-x_3)^2+(x_2-x_3)^2$?
b) Aká je hodnota výrazu $x_1^2+x_2^2+x_3^2$?

Úloha 10.1.5. Nájdite (v $\mathbb{C}$) riešenia sústavy rovníc

$$\begin{aligned}x+y+z&=4\\x^2+y^2+z^2&=4\\x^3+y^3+z^3&=4\end{aligned}$$

(Môžu sa vám hodiť nejaké veci, ktoré ste už vyrátali v úlohe 10.1.1 a Vietove vzťahy.)

Úloha 10.1.6. Nájdite riešenia všetky danej sústavy nad $\mathbb{C}$:

$$\begin{aligned}a+b+c&=4\\(a+b)\,(b+c)\,(c+a)&=18\\\frac{1}{a}+\frac{1}{b}+\frac{1}{c}&=\frac{5}{2}\end{aligned}$$

Úloha 10.1.7. Nájdite všetky riešenia danej sústavy v $\mathbb{C}$:

$$\begin{aligned}x+y+z&=1\\x^2+y^2+z^2&=35\\x^3+y^3+z^3&=97\end{aligned}$$

Kapitola 11

Grupy a podgrupy

{grupy2:CHAPTER}

11.1 Základné vlastnosti grúp

Definíciu a základné vlastnosti grúp sme sa naučili už v prvom semestri – pouívali sme ich pri definícii poľa aj pri definícii vektorového priestoru. Pre zopakovanie však uveďme aspoň stručný prehľad.

{grupy2:DEFINITIONGROUP}

Definícia 11.1.1. Dvojicu $(G, *)$, kde G je množina a $*$ je binárna operácia na G nazývame *grupa,* ak

(i) operácia $*$ je asociatívna

$$(\forall g, h, k \in G) g * (h * k) = (g * h) * k,$$

(ii) operácia $*$ má neutrálny prvok

$$(\exists e \in G)(\forall g \in G) e * g = g * e = g,$$

(iii) pre každý prvok $g \in G$ existuje inverzný prvok vzhľadom na operáciu $*$

$$(\forall g \in G)(\exists g^{-1} \in G) g * g^{-1} = g^{-1} * g = e.$$

Ak je binárna operácia $*$ komutatívna

$$(\forall g, h \in G) g * h = h * g,$$

hovoríme o *komutatívnej grupe.*

Ak platí len prvá z podmienok definície grupy, t.j. ak $*$ je asociatívna binárna operácia na množine G, tak dvojicu $(G, *)$ nazývame *pologrupa.* Ak navyše existuje neutrálny prvok pre operáciu $*$, tak $(G, *)$ voláme *pologrupa s jednotkou* alebo tiež *monoid.*

Často označenie pre grupovú operáciu vynechávame a píšeme ab namiesto $a * b$.

Asociatívnosť vlastne znamená, že môžeme vynechávať zátvorky – pri ľubovoľnom uzátvorkovaní dostaneme ten istý prvok. (V tvrdení 3.1.14 sme dokázali zovšeobecnený asociatívny zákon, ktorý hovorí, že zátvorky môžeme vynechávať aj pri väčšom počte prvkov.)

V grupe platia zákony o krátení

$$\begin{aligned} g * h_1 = g * h_2 \qquad &\Rightarrow h_1 = h_2 \\ h_1 * g = h_2 * g \qquad &\Rightarrow h_1 = h_2 \end{aligned}$$

Zo zákonov o krátení sa dá odvodiť jednoznačnosť neutrálneho prvku aj inverzného prvku. Pre inverzný prvok v grupe platí

$$(g^{-1})^{-1} = g$$
$$(g * h)^{-1} = h^{-1} * g^{-1}$$

Veľa príkladov grúp poznáme z minulého semestra.

Príklad 11.1.2. Príklady grúp:
$(V, +)$ kde V je ľubovoľný vektorový priestor,
$(F, +)$ a $(F \setminus \{0\}, \cdot)$ pre ľubovoľné pole $(F, +, \cdot)$,
$(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$, $(\mathbb{C}, +)$, $(\mathbb{Q} \setminus \{0\}, \cdot)$, $(\mathbb{R} \setminus \{0\}, \cdot)$, $(\mathbb{C} \setminus \{0\}, \cdot)$,
$(\mathbb{Z}_n, \oplus)$ pre $n \in \mathbb{N}$, $n \geq 2$,
$(\mathbb{Z}_p \setminus \{0\}, \odot)$ kde p je prvočíslo.
Príkladom nekomutatívnej grupy je grupa S_n všetkých permutácií n-prvkovej množiny s operáciou skladania zobrazení pre $n \geq 3$ (úloha 3.2.2, permutáciam sa budeme venovať aj tento semester v časti 11.5).

{grp2:PRCN}

Príklad 11.1.3. Množina $\mathbb{C}_n = \{x \in \mathbb{C}; x^n = 1\}$ s operáciou násobenia komplexných čísel tvorí grupu. Asociatívnosť je zrejmá (zdedí sa z komplexných čísel), $1 \in \mathbb{C}_n$ je neutrálny prvok. Takisto ak $x^n = 1$ tak aj $\left(\frac{1}{x}\right)^n = \frac{1}{x^n} = \frac{1}{1} = 1$, čiže $\frac{1}{x}$ patrí do $\mathbb{C}_n$. Číslo $\frac{1}{x}$ je inverzný prvok k prvku x.

{grp2:PRDIREKT}

Príklad 11.1.4. Ak $(G, *_G)$ a $(H, *_H)$ sú grupy, tak aj $G \times H$ s operáciou $(a, b) * (a', b') = (a *_G a', b *_H b')$ je grupa (úloha 11.1.2).

{grp2:DEFDIREKT}

Definícia 11.1.5. Grupu $G \times H$ z predchádzajúceho príkladu nazývame *priamy súčin grúp* G *a* H (alebo tiež *direktný súčin* grúp.)

Cvičenia

Úloha 11.1.1. Nech $(G, *)$ je grupa a e je jej neutrálny prvok. Dokážte:
a) $x * y = y * x \Leftrightarrow x * y * x^{-1} * y^{-1} = e$.
b) Ak $x * x = e$ pre všetky $x \in G$, tak G je komutatívna.

{grpcvic:DIREKT}

Úloha 11.1.2. Overte, že $G \times H$ spolu s operáciou $*$ definovanou v príklade 11.1.4 tvorí grupu pre ľubovoľné grupy $(G, *_G)$ a $(H, *_H)$.

Úloha 11.1.3. Nech G je grupa, e je jej neutrálny prvok a $a, b \in G$. Ukážte, že ak $(ab)^2 = e$, tak aj $(ba)^2 = e$.

Úloha 11.1.4. Nech G je grupa s vlastnosťou, že pre ľubovoľné $a, b, c, d, x \in G$ platí implikácia $axb = cxd \Rightarrow ab = cd$. (Táto vlastnosť by sa dala nazvať „krátenie v strede“.) Dokážte, že potom grupa G je komutatívna.

Platí aj obrátené tvrdenie? T.j. ak G je komutatívna grupa, platí pre ľubovoľné jej prvky uvedená implikácia?

Úloha 11.1.5*. Nech $*$ je asociatívna binárna operácia na množine $G \neq \emptyset$. Nech pre každé $a, b \in G$ majú rovnice $a * x = b$, $y * a = b$ riešenia v G. (Inými slovami, pre každé $a, b \in G$ existujú $x \in G$ a $y \in G$ také, že $a * x = b$, $y * a = b$.) Dokážte, že $(G, *)$ je grupa. (Hint: Skúste začať dôkazom existencie ľavého a pravého neutrálneho prvku.)

{grpcvic:NASK

Úloha 11.1.6. Nech $(G, \cdot)$ je grupa a $P(G)$ je systém všetkých podmnožín G. Dokážte, že operácia $\cdot$ na množine $P(G)$ daná predpisom

$$A \cdot B = \{a \cdot b; a, b \in G\}$$

je asociatívna. Tvorí $P(G) \setminus \{\emptyset\}$ s touto operáciou grupu?

Úloha 11.1.7*. Každá konečná grupa s párnym počtom prvkov obsahuje prvok x taký, že $x = x^{-1}$.

Úloha 11.1.8. Nech $G = (\mathbb{Q} \times \mathbb{Q} \setminus \{(0,0)\})$. Definujme na tejto množine binárnu operáciu $*$ predpisom $(a,b) * (c,d) = (ac + 2bd, ad + bc)$. Je to skutočne binárna operácia? Je $(G, *)$ grupa? Je to komutatívna grupa?

Úloha 11.1.9. Uvažujme funkcie $f_i \colon \mathbb{R} \setminus \{0,1\} \to \mathbb{R} \setminus \{0,1\}$ definované ako $f_1(x) = x$, $f_2(x) = 1/x$, $f_3(x) = 1-x$, $f_4(x) = 1/(1-x)$, $f_5(x) = (x-1)/x$, $f_6(x) = x/(x-1)$. Dokážte, že $G = \{f_1, f_2, f_3, f_4, f_5, f_6\}$ s operáciou skladania zobrazení tvorí grupu.

Úloha 11.1.10*. Nech G je grupa a $a, b \in G$. Nech pre tieto prvky platia rovnosti $aba = ba^2b$, $a^3 = e$ a pre nejaké $n \in \mathbb{N}$ platí $b^{2n-1} = e$. Dokážte, že $b = e$. (Hint vedeli by ste ukázať $ab^2 = b^2a$? Dá sa to ďalej použiť na dôkaz, že pre tieto prvky platí $ab = ba$?)

Úloha 11.1.11*. Nech $*$ je asociatívna operácia na *konečnej* neprázdnej množine M. Ukážte, že existuje prvok $x \in M$ taký, že $x * x = x$.

11.2 Podgrupy

Pojem podgrupy, ktorý teraz zadefinujeme, predstavuje podmnožinu nejakej grupy, ktorá s tou istou operáciou opäť tvorí grupu. Je to do istej miery analógia pojmu podpriestoru vektorového priestoru, s ktorým sme sa zoznámili v minulom semestri.

Definícia 11.2.1. Nech $(G, *)$ je grupa a $H \subseteq G$ je ľubovoľná podmnožina G. Hovoríme, že H je *podgrupa* grupy G, ak H s binárnou operáciou $*$ zúženou na podmnožinu H tvorí grupu.

Budeme používať označenie $H \leq G$, prípadne $(H, *) \leq (G, *)$.

Pod zúžením operácie na podmnožinu rozumieme operáciu danú predpisom

$$h_1 *_H h_2 = h_1 *_G h_2$$

pre ľubovoľné $h_1, h_2 \in H$. (Kvôli zrozumiteľnosti sme tu použili rozličné označenie pre operáciu na grupe G a jej podgrupe H, ďalej však budeme používať rovnaké označenie pre obe operácie.)

{grp2:POZNUZAV}

Poznámka 11.2.2. Dôležité je všimnúť si, že definícia podgrupy zahŕňa aj požiadavku, aby zúženie operácie $*$ na podmnožinu H bola binárna operácia na H. To znamená, že množina H je uzavretá vzhľadom na operáciu $*$, čiže

$$h_1, h_2 \in H \Rightarrow h_1 * h_2 \in H.$$

Pretože každá grupa musí obsahovať neutrálny prvok, priamo z definície vyplýva, že $H \neq \emptyset$.

Z toho, že podgrupa má rovnako definovanú grupovú operáciu vyplýva, že aj inverzné prvky a neutrálny prvok sú v podgrupe rovnaké ako v celej grupe.

rp2:LMPODGNP}

Lema 11.2.3. *Nech $(G, *)$ je grupa a H je jej podgrupa.*

:LMPODGNPit1}

(i) *Ak e_H je neutrálny prvok grupy H a e_G je neutrálny prvok grupy G, tak $e_H = e_G$. (Z toho špeciálne vyplýva $e_G \in H$, teda každá podgrupa musí obsahovať neutrálny prvok.)*

:LMPODGNPit2}

(ii) *Ak $a \in H$, b je inverzný prvok k a v G a c je inverzný prvok k a v H, tak $b = c$.*

Dôkaz. (i) Z toho, že e_G je neutrálny prvok grupy G dostaneme

$$e_G * e_H = e_H.$$

Súčasne platí

$$e_H * e_H = e_H$$

lebo e_H je neutrálny prvok grupy H. Dostávame teda rovnosť

$$e_G * e_H = e_H * e_H$$

a zo zákona o krátení (v grupe G) potom vyplýva $e_G = e_H$.

(ii) Opäť využijeme zákon o krátení. Z prvej časti vieme, že $e_G = e_H$, označme teda neutrálny prvok oboch grúp ako $e := e_G = e_H$. Ak b je inverzný prvok k a v grupe G, tak $a * b = e$. Podobne, z toho, že c je inverzný prvok k a v H máme $a * c = e$. Z rovnosti

$$a * b = a * c$$

vyplýva $b = c$. □

Príklad 11.2.4. $(\mathbb{Q}, +)$ je podgrupa grupy $(\mathbb{R}, +)$, lebo operácia $+$ na $\mathbb{Q}$ funguje rovnako ako sčitovanie reálnych čísel. Podobne $(\mathbb{Z}, +)$ je podgrupa $(\mathbb{Q}, +)$ a $(\mathbb{R}, +)$ je podgrupa grupy $(\mathbb{C}, +)$.

Podobne ako pri vektorových priestoroch, aj pri podgrupách máme pomerne jednoduché kritérium na zistenie, či nejaká podmnožina tvorí podgrupu danej grupy.

{grp2:VTKRIT}

Veta 11.2.5 (Kritérium podgrupy)**.** *Nech $(G, *)$ je grupa a $H \subseteq G$, $H \neq \emptyset$. Nasledujúce podmienky sú ekvivalentné*

{grp2:KRITit1}

(i) *H je podgrupa grupy G;*

{grp2:KRITit2}

(ii) *množina H je uzavretá vzhľadom na operáciu $*$ a na tvorenie inverzných prvkov v G, čiže platí (pre ľubovoľné $a, b \in H$)*

$$a, b \in H \Rightarrow a * b \in H$$
$$a \in H \Rightarrow a^{-1} \in H$$

{grp2:KRITit3}

(iii) *pre ľubovoľné $a, b \in H$ platí aj $a^{-1} * b \in H$*

Dôkaz. (i) ⇒ (ii): Uzavretosť množiny H vzhľadom na operáciu $*$ vyplýva z toho, že zúženie operácie $*$ na H je binárna operácia na podmnožine H (poznámka 11.2.2). Ďalej z lemy 11.2.3 vieme, že inverzné prvky v H sú rovnaké ako v G. Preto inverzný prvok k ľubovoľnému $a \in H$ (v grupe G) musí patriť do H (je to inverzný prvok k a v grupe H).

(ii) ⇒ (iii): Ak $a, b \in H$ tak aj $a^{-1} \in H$ (na základe druhej z dvoch podmienok uvedených v (ii)), na základe prvej podmienky (použitej pre a^{-1} a b) potom dostaneme $a^{-1} * b \in H$.

(iii) $\Rightarrow$ (ii): Pretože H je neprázdna množina, existuje aspoň jeden prvok $x \in H$. Použitím (iii) pre $b = a = x$ dostaneme $x^{-1} * x = e \in H$. Majme teraz ľubovoľné $a, b \in H$. Z (iii) pre prvky a a e dostaneme $a^{-1} * e = a^{-1} \in H$. Ak teraz použijeme tú istú podmienku pre a^{-1} a b, dostaneme $(a^{-1})^{-1} * b = a * b \in H$. Obe implikácie z (ii) sú teda splnené.

(ii) $\Rightarrow$ (i): Máme dokázať, že H so zúženou operáciou $*$ spĺňa podmienky z definície grupy. Uzavretosť na operáciu $*$ znamená, že zúženie operácie $*$ je binárna operácia na H (poznámka 11.2.2). Asociatívnosť sa automaticky zdedí z grupy G (pozri poznámku 4.2.6). Ďalej by sme mali ukázať, že $e \in H$. Použijeme podobný postup, ako v predchádzajúcej časti dôkazu. Keďže $H \neq \emptyset$, existuje nejaký prvok $a \in H$ a z (ii) máme $a^{-1} * a = e \in H$. Zostáva nám dokázať, že každý prvok má v H inverzný prvok. Z predchádzajúcej lemy však vieme, že inverzné prvky v G a v H sú rovnaké a druhá časť podmienky (ii) hovorí, že podmnožina H je uzavretá vzhľadom na inverzné prvky. □

Pri použití kritéria podgrupy potrebujeme overiť aj to, že podmnožina H je neprázdna. Pretože každá podgrupa musí obsahovať neutrálny prvok, je často najjednoduchšie začať overením, či ho daná podmnožina H naozaj obsahuje. (Ak zistíme, že $e \in H$, tak $H \neq \emptyset$ a môžeme použiť kritérium podgrupy. V opačnom prípade H nemôže byť podgrupa, takže ďalšie podmienky už nemusíme overovať.)

Poznámka 11.2.6. V skutočnosti ak je množina H konečná, je možné uvedené kritérium ešte o čosi zjednodušiť – stačí overovať podmienku $a, b \in H \Rightarrow a*b \in H$. Tento fakt dokážeme o niečo neskôr v dôsledku 11.4.5.

Poznámka 11.2.7. Kritérium podgrupy nám dáva inú možnosť ako dokázať, že nejaká množina so zadanou binárnou operáciou tvorí grupu. Pre konzistentnosť s uvedeným tvrdením, označme našu množinu H a binárnu operáciu $*$. Chceme overiť, že $(H, *)$ je grupa.

Kritérium podgrupy môžeme využiť v situácii, keď ide o zúženie nejakej binárnej operácie na väčšej množine G, o ktorej sme už ukázali, že tvorí grupu. Namiesto toho, aby sme overovali pre H podmienky z definície grupy, stačí nám overiť, že H spĺňa kritérium podgrupy.

Príklad 11.2.8. Ak $(G, *)$ je grupa, tak G je podgrupa grupy G. Ak e je neutrálny prvok grupy G, tak $\{e\}$ je podgrupa grupy G. Čiže každá grupa obsahuje dve podgrupy – celú grupu G a jednoprvkovú podgrupu $\{e\}$ obsahujúcu len neutrálny prvok.

Uvedieme teraz dva príklady podgrúp grupy $(\mathbb{C} \setminus \{0\}, \cdot)$.

{grp2:PRCIRCLE}

Príklad 11.2.9. Označme $S = \{x \in \mathbb{C}; |x| = 1\}$. Množina S tvorí podgrupu grupy $(\mathbb{C} \setminus \{0\}, \cdot)$. Skutočne:
Platí $1 \in S$, teda S je neprázdna.
Ak $x, y \in S$, znamená to $|x| = |y| = 1$. Potom $|xy| = |x| \cdot |y| = 1$, teda aj $xy \in S$.
Ak $x \in S$, čiže $|x| = 1$, tak $\left|\frac{1}{x}\right| = \frac{1}{|x|} = 1$, čiže $\frac{1}{x} \in S$.

Príklad 11.2.10. Podmnožiny $\mathbb{R} \setminus \{0\}$, $\mathbb{Q} \setminus \{0\}$ sú podgrupy grupy $(\mathbb{C} \setminus \{0\}, \cdot)$. Stačí si uvedomiť, že podiel dvoch nenulových reálnych (racionálnych) čísel je opäť reálne (racionálne) číslo.

Príklad 11.2.11. Podmnožina $\mathbb{R}^+$ je podgrupa grupy $(\mathbb{R} \setminus \{0\}, \cdot)$. Vyplýva to z toho, že podiel dvoch kladných čísel je opäť kladné číslo.

Kritérium podgrupy môžeme využiť na dôkaz nasledujúcej jednoduchej lemy ako aj dôležitého tvrdenia 11.2.13 o prieniku podgrúp.

{grp2:LMPODPODGRP}

Lema 11.2.12. *Nech $(G, *)$ je grupa. Potom*

(i) *Ak $K \subset H \subset G$ a K, H sú podgrupy G, tak K je podgrupa H.*

(ii) *Ak H je podgrupa G a K je podgrupa H, tak K je aj podgrupa G.*

Dôkaz. Cvičenie. □

{grp2:TVRPRIEN}

Tvrdenie 11.2.13. *Nech $(G, *)$ je grupa a H_i je podgrupa grupy G pre každé $i \in I$. Potom prienik týchto podgrúp*

$$H := \bigcap_{i \in I} H_i$$

je opäť podgrupa grupy G.

Dôkaz. Potrebujeme ukázať, že $H \neq \emptyset$ a H spĺňa podmienku (iii) z vety 11.2.5.

Pretože $e \in H_i$ pre všetky podgrupy H_i, neutrálny prvok e leží aj v ich prieniku, a teda $H \neq \emptyset$.

Nech teraz $a, b \in H$. To znamená, že $a, b \in H_i$ pre každé $i \in I$. Z kritéria podgrupy potom dostaneme $a^{-1} * b \in H_i$. Pretože tento prvok patrí do každej z množín H_i pre $i \in I$, patrí aj do ich prieniku, čiže $a^{-1} * b \in H = \bigcap_{i \in I} H_i$. □

Príklad 11.2.14. Už sme videli, že $S = \{x \in \mathbb{C}; |x| = 1\}$, $\mathbb{R} \setminus \{0\}$ aj $\mathbb{R}^+$ sú podgrupy grupy $(\mathbb{C} \setminus \{0\}, \cdot)$. Môžeme si všimnúť, že $S \cap (\mathbb{R} \setminus \{0\}) = \{\pm 1\}$ aj $S \cap \mathbb{R}^+ = \{1\}$ sú podgrupy $(\mathbb{C} \setminus \{0\}, \cdot)$.

Priamo z tvrdenia 11.2.13 dostaneme nasledujúci dôležitý dôsledok.

Dôsledok 11.2.15. *Ak $(G, *)$ je grupa a $A \subset G$ je ľubovoľná podmnožina G, tak prienik všetkých podgrúp obsahujúcich množinu A je tiež podgrupa G. Túto podgrupu nazývame* podgrupa generovaná podmnožinou A *a označujeme $[A]$. Ak $[A] = G$, hovoríme, že grupa G je generovaná podmnožinou A (alebo tiež, že A generuje G). V prípade, že $A = \{a\}$ je jednoprvková množina, tak namiesto $[\{a\}]$ budeme používať označenie $[a]$ a hovoríme o* podgrupe generovanej prvkom a.

Definíciu podgrupy generovanej množinou A môžeme stručne zapísať ako

$$[A] = \bigcap \{H \subseteq G; H \supseteq A \ \wedge \ H \text{ je podgrupa } G\}. \tag{11.1}$$

{grp2:EQGENPRIEN}

Poznámka 11.2.16. Podgrupa generovaná množinou A je najmenšia (vzhľadom na inklúziu) podgrupa grupy G, ktorá obsahuje A. Pod pojmom „najmenšia vzhľadom na inklúziu“ rozumieme to, že pre ľubovoľnú podgrupu H, ktorá obsahuje A platí $[A] \subseteq H$. Inak povedané, je to najmenší prvok množiny tých podgrúp, ktoré obsahujú A, vzhľadom na čiastočné usporiadanie $\subseteq$ na tejto množine.[1]

Z toho, čo sme doteraz uviedli o podgrupe generovanej nejakou množinou je zrejmá analógia s pojmom vektorového podpriestoru generovaného nejakou množinou vektorov.

S týmto pojmom sa ešte stretneme, podrobnejšie sa budeme venovať najmä podgrupám generovaným jediným prvkom. Zatiaľ uveďme aspoň jeden jednoduchý príklad.

[1] S pojmom čiastočné usporiadanie a najmenší prvok ste sa pravedpodobne už stretli alebo ešte stretnete na iných prednáškach, pozri [OŠ].

{grp2:GENERZ}

Príklad 11.2.17. Uvažujme grupu $(\mathbb{R}, +)$. Potom podgrupa generovaná prvkom 1 je $\mathbb{Z}$, čiže $\mathbb{Z} = [1]$.

Aby sme videli, že $\mathbb{Z} \subseteq [1]$, stačí si všimnúť, že keď nejaká podgrupa $(\mathbb{R}, +)$ obsahuje 1, musí obsahovať aj $2 = 1 + 1$, $3 = 2 + 1$ atď. Indukciou môžeme dokázať, že obsahuje všetky prirodzené čísla. Samozrejme, ako každá podgrupa, obsahuje aj neutrálny prvok 0 a z uzavretosti na inverzné prvky vyplýva, že musí obsahovať aj všetky záporné celé čísla. Takže podgrupa [1] určite obsahuje všetky prvky množiny $\mathbb{Z}$.

Na druhej strane, $\mathbb{Z}$ je podgrupa $(\mathbb{R}, +)$, ktorá obsahuje číslo 1. Takže je jednou z podgrúp vystupujúcich v prieniku (11.1), z čoho vyplýva $[1] \subseteq \mathbb{Z}$.

Takmer rovnakým spôsobom by sa dalo ukázať, že aj $[-1] = \mathbb{Z}$.

Pomerne ľahko sa môžete presvedčiť aj o tom, že $[\{2, 3\}] = \mathbb{Z}$ (úloha 11.2.12).

{grp2:GENERZ2Z3}

Príklad 11.2.18. Pozrime sa teraz na 6-prvkovú grupu $\mathbb{Z}_2 \times \mathbb{Z}_3$. (Usporiadané dvojice, kde na prvej súradnici sčitujeme modulo 2 a na druhej súradnici súradnici modulo 3 - pozri definíciu 11.1.5).

V tomto prípade (ako čitateľ ľahko overí) platí napríklad:
$[(1, 0)] = \mathbb{Z}_2 \times \{0\}$;
$[(0, 1)] = \{0\} \times \mathbb{Z}_3$;
$[(1, 1)] = \mathbb{Z}_2 \times \mathbb{Z}_3$, keďže pomocou prvku $(1, 1)$ postupne dostaneme prvky $(0, 2)$, $(1, 0)$, $(0, 1)$, $(1, 2)$ a $(0, 0)$, čiže všetky prvky grupy.

Príklad 11.2.19. Iným príkladom by bola podgrupa $H = [\{1, \sqrt{2}\}]$ grupy $(\mathbb{R}, +)$. T.j. H je najmenšia podmnožina množiny $\mathbb{R}$, ktorá obsahuje čísla 1 a $\sqrt{2}$ a je uzavretá na súčet aj na opačné prvky. Pomerne ľahko sa dá overiť, že

$$H = \{a + b\sqrt{2}; a, b \in \mathbb{Z}\}.$$

O trochu náročnejšie je ukázať, že táto grupa nie je cyklická; t.j. nedá sa v nej nájsť generátor (úloha 11.2.18*).

Cvičenia

Úloha 11.2.1. Dokážte lemu 11.2.12.

Úloha 11.2.2. Nájdite všetky podgrupy $(\mathbb{Z}_6, \oplus)$.

Úloha 11.2.3. Dokážte, že matice typu $n \times n$, ktorých determinant je rovný 1, s operáciou násobenia matíc tvoria grupu.

Úloha 11.2.4. Dokážte: Ak H je podgrupa grupy $(G, \cdot)$ tak $H^2 = H \cdot H = H$.

Úloha 11.2.5. Ak A, B, C sú podgrupy G a $C \subseteq A \cup B$, tak $C \subseteq A$ alebo $C \subseteq B$.

Úloha 11.2.6. Tvoria pri sčitovaní/násobení matíc grupu štvorcové matice $n \times n$, ktoré sú: symetrické, antisymetrické, diagonálne, regulárne, horné trojuholníkové...

Úloha 11.2.7. Nájdite príklad nekonečnej grupy, ktorá obsahuje netriviálnu konečnú podgrupu. (Pod netriviálnou podgrupou tu rozumieme podgrupu, ktorá má viac ako jeden prvok.)

Úloha 11.2.8. Matice typu $n \times n$, ktoré v každom riadku a každom stĺpci majú práve jednu jednotku a ostatné prvky sú nulové, s operáciou násobenia matíc tvoria grupu. (Hint: Súvisia tieto matice nejako s permutáciami? Akým lineárnym zobrazeniam zodpovedajú?)

Úloha 11.2.9. Ukážte, že $H = \{\frac{m}{n}; m, n$ sú nepárne$\}$ je podgrupa grupy $(\mathbb{Q} \setminus \{0\}, \cdot)$.

Úloha 11.2.10. Nájdite všetky podgrupy grupy $\mathbb{Z}_2 \times \mathbb{Z}_2$ a všetky podgrupy grupy $\mathbb{Z}_4$ (v oboch prípadoch operácia $\oplus$). Majú tieto grupy rovnaký počet dvojprvkových podgrúp? (Z toho, čo sa naučíme v ďalšej podkapitole sa na základe tejto úvahy bude dať zdôvodniť, že tieto dve grupy nie sú izomorfné.)

Úloha 11.2.11. Je množina $H = \{\ln a; a \in \mathbb{Q}, a > 0\}$ podgrupou grupy $(\mathbb{R}, +)$?

{podgrupcvic:ULOGENERZ}

Úloha 11.2.12. Budeme pracovať v grupe $(\mathbb{R}, +)$.
a) Dokážte, že $[\{2, 3\}] = \mathbb{Z}$;
b) Dokážte, že $[\{1, \sqrt{2}\}] = \{a + b\sqrt{2}; a, b \in \mathbb{Z}\}$.
c*) Je možné podgrupu $[\{1, \sqrt{2}\}]$ generovať jediným prvkom? (V terminológii, ktorú zavedieme v časti 11.4 sa dá táto otázka sformulovať takto: Je podgrupa $[\{1, \sqrt{2}\}]$ cyklická?)

Úloha 11.2.13. Dokážte, alebo vyvráťte: Ak H_1 je podgrupa G_1 a H_2 je podgrupa G_2, tak $H_1 \times H_2$ je podgrupa $G_1 \times G_2$.

Úloha 11.2.14. Nech H je podgrupa grupy G. Nech $g \in G$. Ukážte, že $gHg^{-1} = \{ghg^{-1}; h \in H\}$ je podgrupa grupy G.

Úloha 11.2.15. Nech V je vektorový priestor nad poľom $\mathbb{R}$. Je aj každá podgrupa grupy $(V, +)$ podpriestorom priestoru V? Ako je to s vektorovými priestormi nad poľom $\mathbb{Z}_p$?

Úloha 11.2.16. Nech H je vlastná podgrupa grupy G (t.j. $H \neq G$). Dokážte, že $[G-H] = G$.

Úloha 11.2.17. Nech A, B sú podgrupy grupy G. Dokážte, že AB je podgrupa G práve vtedy, keď $AB = BA$.

{podgrpcvic:GENER1SQRT2}

Úloha 11.2.18*. Ukážte, že podgrupa H grupy $(\mathbb{R}, +)$ vygenerovaná prvkami 1 a $\sqrt{2}$ je presne $H = \{a + b\sqrt{2}; a, b \in \mathbb{Z}\}$. Ukážte, že táto grupa nie je cyklická, t.j. neexistuje prvok $h \in H$, pre ktorý by platilo $H = [h] = \{kh; k \in \mathbb{Z}\}$.

11.3 Homomorfizmy grúp

Pri vektorových priestoroch boli dôležité lineárne zobrazenia – zobrazenia zachovávajúce operácie určujúce vektorový priestor. Podobne aj pri štúdiu grúp sú užitočné zobrazenia medzi grupami, ktoré zachovávajú grupové operácie. Takéto zobrazenia voláme homomorfizmy. [2]

Definícia 11.3.1. Nech $(G, \circ)$, $(H, *)$ sú grupy. Potom zobrazenie $f\colon G \to H$ je *homomorfizmus,* ak

$$f(g_1 \circ g_2) = f(g_1) * f(g_2)$$

platí pre ľubovoľné $g_1, g_2 \in G$.

Na označenie homomorfizmu budeme niekedy používať stručnejší zápis $f\colon (G, \circ) \to (H, *)$ (t.j. týmto zápisom súčasne popíšeme ako označujeme homomorfizmus a aj ako označujeme grupové operácie.)

Skôr než si tento pojem ilustrujeme na príkladoch, dokážeme si dve jednoduché vlastnosti, ktoré musí každý homomorfizmus spĺňať.

{grp2:VTHOMNEU}

Veta 11.3.2. *Nech $(G, \circ)$, $(H, *)$ sú grupy a $f\colon G \to H$ je homomorfizmus. Označme ďalej e_G neutrálny prvok grupy G a e_H neutrálny prvok grupy H. (Inverzné prvky budeme v oboch prípadoch označovať pomocou horného indexu -1 ako obvykle.) Potom platí:*

[2] Keďže pojem homomorfizmu sa definuje aj pre iné štruktúry než sú grupy, niekedy sa používa aj termín *grupový homomorfizmus.*

{grp2:itHN1}
{grp2:itHN2}

(i) $f(e_G) = e_H$ (*teda homomorfizmus musí zobraziť neutrálny prvok na neutrálny prvok*);

(ii) $f(a^{-1}) = (f(a))^{-1}$ (*teda homomorfizmy zachovávajú aj inverzné prvky*).

Dôkaz. (i) Pretože

$$f(e_G) = f(e_G \circ e_G) = f(e_G) * f(e_G),$$

dostávame rovnosť $f(e_G) * e_H = f(e_G) = f(e_G) * f(e_G)$. Zo zákona o krátení (použitého pre grupu H) potom dostaneme $f(e_G) = e_H$.

(ii) Z definície homomorfizmu tentokrát máme

$$f(a^{-1}) * f(a) = f(a^{-1} \circ a) = f(e_G) \overset{\text{(i)}}{=} e_H,$$

teda $f(a^{-1}) * f(a) = (f(a))^{-1} * f(a)$ a opäť stačí použiť zákon o krátení, aby sme dostali $f(a^{-1}) = (f(a))^{-1}$. □

{grp2:PRHOMTRIV}

Príklad 11.3.3. Ak $(G, *)$ a $(H, \circ)$ sú ľubovoľné grupy, tak $f\colon G \to H$ určené predpisom $f(g) = e_H$ pre všetky $g \in G$ je homomorfizmus.

Zobrazenie $id_G\colon G \to G$ je homomorfizmus pre každú grupu $(G, *)$.

{grp2:PRHOMEXP}

Príklad 11.3.4. Uvažujme zobrazenie

$$f\colon (\mathbb{R}, +) \to (\mathbb{R}^+, \cdot), \qquad f\colon x \mapsto e^x.$$

Priamo z vlastností exponenciálnej funkcie vyplýva, že f je homomorfizmus

$$f(x + y) = e^{x+y} = e^x \cdot e^y = f(x) \cdot f(y)$$

{grp2:PRHOMEXPKOMPL}

Príklad 11.3.5. Definujme

$$g\colon (\mathbb{R}, +) \to (S, \cdot), \qquad g\colon \varphi \mapsto e^{i\varphi} = \cos\varphi + i\sin\varphi,$$

kde $S = \{z \in \mathbb{C}; |z| = 1\}$ je grupa z príkladu 11.2.9. (Zápis $e^{i\varphi}$ budeme chápať jednoducho ako skratku zápisu $\cos\varphi + i\sin\varphi$. Ide o tzv. goniometrický a exponenciálny tvar komplexného čísla, pozri podkapitolu C.4.)

Fakt, že ide o homomorfizmus vyplýva z Moivrovej vety C.2.3, ktorá hovorí, že pri násobení komplexných čísel sa uhly sčitujú (a absolútne hodnoty sa násobia).

$$g(\varphi_1+\varphi_2) = (\cos(\varphi_1+\varphi_2)+i\sin(\varphi_1+\varphi_2)) = (\cos\varphi_1+i\sin\varphi_1)\cdot(\cos\varphi_2+i\sin\varphi_2) = g(\varphi_1)\cdot g(\varphi_2)$$

{grp2:PRHOMEXPMOD}

Príklad 11.3.6. Homomorfizmus z predchádzajúceho príkladu teraz trochu zmodifikujeme.

Budeme pracovať s grupou $(\langle 0, 2\pi), +)$, v ktorej je sčitovanie definované modulo 2π. Formálne môžeme operáciu $+$ definovať ako

$$\alpha + \beta = 2\pi \left\{ \frac{\alpha + \beta}{2\pi} \right\},$$

kde $\{x\}$ označuje desatinnú časť čísla x. (Význam znamienka $+$ na pravej strane predstavuje sčitovanie reálnych čísel, zatiaľčo na ľavej strane je operácia, ktorú definujeme. Na to ste si však už pravedpodobne zvykli, že niekedy označujeme rôzne veci rovnakým symbolom.)

Potom dostávame homomorfizmus

$$h\colon (\langle 0, 2\pi), +) \to (S, \cdot), \qquad h\colon \varphi \mapsto e^{i\varphi} = \cos\varphi + i\sin\varphi.$$

(Je to naozaj homomorfizmus, lebo zmena uhla φ o nejaký násobok 2π neovplyvní hodnotu čísla $\cos\varphi + i\sin\varphi$. Výsledok je teda rovnaký ako pri použití homomorfizmu z predchádzajúceho príkladu. Tento homomorfizmus je navyše bijektívny.)

{grp2:PRHOMZZN}

Príklad 11.3.7. Nech $n \in \mathbb{N}$, $n \geq 2$. Zobrazenie $f\colon (\mathbb{Z}, +) \to (\mathbb{Z}_n, \oplus)$ dané predpisom

$$f(k) = k \bmod n$$

je homomorfizmus. (Rovnosť $f(k+l) = f(k) \oplus f(l)$ vyplýva z toho, že dostaneme rovnaký výsledok keď dve čísla sčítame a potom urobíme zvyšok po delení n a keď to urobíme v obrátenom poradí.)

Pripomeňme, čo rozumieme pod obrazom a vzorom množiny v danom zobrazení.

Definícia 11.3.8. Nech $f\colon X \to Y$, $A \subseteq X$, $B \subseteq Y$.
Potom *obraz množiny* A v zobrazení f je

$$f[A] = \{f(a); a \in A\}$$

a *vzor množiny* B v zobrazení f je

$$f^{-1}(B) = \{a \in A; f(a) \in B\}.$$

V prípade, že $B = \{b\}$, t.j. že množina B je jednoprvková, používame niekedy stručnejšie označenie $f^{-1}(b)$ namiesto $f^{-1}(\{b\})$. (Hoci niekedy by sa mohlo toto označenie pliesť s označením pre obraz prvku b v inverznej funkcii, z kontextu snáď vždy bude jasné, čo máme na mysli.)

Všimnime si, že $f[A]$ je podmnožina množiny Y a pre $y \in Y$ platí

$$y \in f[A] \Leftrightarrow (\exists a \in A) y = f(a).$$

Množina $f^{-1}(B)$ je zasa podmnožinou množiny X a pre $x \in X$ máme

$$x \in f^{-1}(B) \Leftrightarrow f(x) \in B.$$

{grp2:TVROBRGRP}
{grp2:OBRGRPit1}
{grp2:OBRGRPit2}

Tvrdenie 11.3.9. *Nech $(G, \circ)$, $(H, *)$ sú grupy a $f\colon G \to H$ je homomorfizmus.*

(i) *Ak G' je podgrupa grupy G, tak aj jej obraz $f[G']$ je podgrupa grupy H.*

(ii) *Ak H' je podgrupa grupy H, tak aj jej vzor $f^{-1}(H')$ je podgrupa grupy G.*

Dôkaz. V oboch prípadoch použijeme kritérium podgrupy.
(i) Pretože $e_G \in G'$, z vety 11.3.2 máme $f(e_G) = e_H \in f[G']$, a teda $f[G'] \neq \emptyset$.
Nech $a, b \in f[G']$. To znamená, že existujú $a_1, b_1 \in G'$ také, že $f(a_1) = a$ a $f(b_1) = b$. Potom dostávame

$$f(a_1^{-1} b_1) = f(a_1)^{-1} f(b_1) = a^{-1} b.$$

Pretože G' je podgrupa, máme $a_1^{-1} b_1 \in G'$, a teda $a^{-1} b = f(a_1^{-1} b_1) \in f[G']$.
(ii) Opäť sa ľahko ukáže, že $e_G \in f^{-1}(H')$, teda táto podmnožina je neprázdna.
Ak $a, b \in f^{-1}(H')$, znamená to, že $f(a), f(b) \in H'$. Potom

$$f(a^{-1} b) = f(a^{-1}) f(b) = f(a)^{-1} f(b) \in H',$$

lebo H' je podgrupa. Zistili sme, že $a^{-1} b \in f^{-1}(H')$, teda $f^{-1}(H')$ vyhovuje podmienke z kritéria podgrupy. □

{grp2:DOSKER}

Dôsledok 11.3.10. *Nech $(G,\circ)$, $(H,*)$ sú grupy a $f\colon G\to H$ je homomorfizmus. Potom* jadro

$$\operatorname{Ker} f=\{g\in G; f(g)=e_H\}$$

je podgrupa grupy G a

$$\operatorname{Im} f=\{f(g); g\in G\}$$

je podgrupa grupy H.

Dôkaz. Vyplýva z toho, že $\{e_H\}$ je podgrupa grupy H a G je podgrupa grupy G. □

Definícia 11.3.11. Nech $(G,\circ)$, $(H,*)$ sú grupy. Ak $f\colon G\to H$ je bijektívny homomorfizmus, hovoríme, že f je *izomorfizmus* alebo tiež, že grupy G a H sú izomorfné (označujeme $G\cong H$).

Opäť, podobne ako v prípade vektorových priestorov, existencia izomorfizmu znamená, že grupy G a H sú v podstate rovnaké, len ich prvky sú inak pomenované. Bijektívne zobrazenie f je „slovníkom“, ktorý prekladá medzi týmito dvoma pomenovaniami.

{grp2:LMZLOZHOM}

Lema 11.3.12. *Nech $(G,*)$, $(H,\circ)$, $(K,\odot)$ sú grupy.*

{grp2:ZHit1}

(i) *Ak $f\colon G\to H$ je izomorfizmus, tak aj $f^{-1}\colon H\to G$ je izomorfizmus.*

{grp2:ZHit2}

(ii) *Ak $f\colon G\to H$ a $g\colon H\to K$ sú homomorfizmy, tak aj $g\circ f\colon G\to K$ je homomorfizmus.*

{grp2:ZHit3}

(iii) *Ak $f\colon G\to H$ a $g\colon H\to K$ sú izomorfizmy, tak aj $g\circ f\colon G\to K$ je izomorfizmus.*

Dôkaz. (i): Nech $a,b\in H$. Pretože f je surjekcia, existujú $a_1,b_1\in G$ také, že $f(a_1)=a$, $f(b_1)=b$. Z definície homomorfizmu potom máme

$$a\circ b=f(a_1)\circ f(b_1)=f(a_1*b_1).$$

Potom priamo z definície inverzného zobrazenia vyplýva

$$f^{-1}(a\circ b)=a_1*b_1=f^{-1}(a)*f^{-1}(b).$$

(ii): Ak $a,b\in G$, dvojnásobným použitím definície homomorfizmu dostaneme

$$g(f(a*b))=g(f(a)\circ f(b))=g(f(a))\odot g(f(b))$$

(iii): Podľa (ii) je zloženie homomorfizmov opäť homomorfizmus. Súčasne vieme (tvrdenie 2.2.13), že zloženie bijekcií je bijekcia. □

Z predchádzajúcej lemy vidíme, že:
a) ak grupa G je izomorfná s grupou H, tak aj H je izomorfná s G,
b) ak G je izomorfná s H a H je izomorfná s K, tak aj grupy G a K sú izomorfné.
Ak si navyše uvedomíme, že každá grupa je izomorfná sama so sebou (identické zobrazenie $id_G\colon G\to G$ je izomorfizmus), tak vidíme, že vzťah „byť izomorfný“ má podobné vlastnosti ako relácia ekvivalencie.

Príklad 11.3.13. Zobrazenie $f\colon(\mathbb{R},+)\to(\mathbb{R}^+,\cdot)$, $f(x)=e^x$, z príkladu 11.3.4 je izomorfizmus. Aby sme videli, že f je bijekcia, stačí si všimnúť, že zobrazenie $f^{-1}\colon\mathbb{R}^+\to\mathbb{R}$ dané predpisom $f^{-1}(x)=\ln x$ je inverzné k zobrazeniu f. (Podľa predchádzajúcej lemy je teda aj toto zobrazenie homomorfizmom.)

Príklad 11.3.14. Homomorfizmus $h(\varphi) = e^{i\varphi} = \cos\varphi + i\sin\varphi,\ h\colon (\langle 0, 2\pi), +) \to (S, \cdot)$, z príkladu 11.3.6 je tiež izomorfizmom.

Aby sme videli, že h je bijekcia, stačí si uvedomiť, že prvky grupy S predstavujú body jednotkovej kružnice a každý bod na jednotkovej kružnici je jednoznačne určený uhlom z intervalu $\langle 0, 2\pi)$, ktorý sa naň zobrazí zobrazením g.

Aj surjektívne a injektívne homomorfizmy majú niektoré zaujímavé vlastnosti, preto sa nám v budúcnosti bude hodiť nasledovná terminológia.

Definícia 11.3.15. Nech $f\colon (G, *) \to (H, \circ)$ je homomorfizmus. Ak f je injektívne zobrazenie, tak hovoríme, že f je *monomorfizmus.* Ak f je surjektívne zobrazenie, tak hovoríme, že f je *epimorfizmus.*

Hovoríme, že grupa $(H, \circ)$ je *homomorfný obraz* grupy $(G, *)$, ak existuje epimorfizmus $f\colon (G, *) \to (H, \circ)$.

Videli sme napríklad, že pre každé $n \in \mathbb{N}$ je $(\mathbb{Z}_n, \oplus)$ homomorfný obraz grupy $(\mathbb{Z}, +)$ (príklad 11.3.7), grupa $(S, \cdot)$ je homomorfný obraz grupy $(\mathbb{R}, +)$ (príklad 11.3.4).

Cvičenia

Úloha 11.3.1. Zistite, či sú grupy G a H izomorfné:
a) $G = (\mathbb{R} \setminus \{0\}, \cdot) \times (\mathbb{R} \setminus \{0\}, \cdot)$, $H = (\mathbb{C} \setminus \{0\}, \cdot)$
b) $G = (\mathbb{Z}_6, \oplus)$, $H = (\mathbb{Z}_2 \times \mathbb{Z}_3, \oplus)$
c) $G = (\mathbb{R} \setminus \{0\}, \cdot)$, $H = (\mathbb{Q} \setminus \{0\}, \cdot)$
d) $G = (\mathbb{Z} \times \mathbb{Z}_2, +)$, $H = (\mathbb{Z}, +)$ (Sčitovanie v $\mathbb{Z} \times \mathbb{Z}_2$ chápeme tak, že na prvej súradnici používame obvyklé sčitovanie a na druhej sčitujeme modulo 2, t.j. tak ako sme definovali priamy súčin grúp.)

Úloha 11.3.2. Zistite, či sú grupy G a H izomorfné a či je grupa H homomorfným obrazom grupy G. Svoju odpoveď zdôvodnite!
a) $G = (\mathbb{R}, +) \times (\mathbb{R}, +)$, $H = (\mathbb{C}, +)$
b) $G = (\mathbb{Q}, +)$, $H = (\mathbb{R}, +)$
c) $G = (\mathbb{Q}, +)$, $H = (\mathbb{Q}^+, \cdot)$
d) $G = (\mathbb{C} \setminus \{0\}, \cdot)$, $H = (\mathbb{R} \setminus \{0\}, \cdot)$
e) $G = (\mathbb{Q}, +)$, $H = (\mathbb{Q} \setminus \{0\}, \cdot)$

Úloha 11.3.3. Zistite, či sú grupy G a H izomorfné a či je niektorá z nich homomorfným obrazom druhej. Svoju odpoveď zdôvodnite!
a) $G = (\mathbb{R}, +)$, $H = (\mathbb{R}^+, \cdot)$
b) $G = (\mathbb{R} \setminus \{0\}, \cdot)$, $H = (\mathbb{R}^+, \cdot)$
c) $G = (\mathbb{Z}_4, \oplus)$, $H = (\mathbb{Z}_2 \times \mathbb{Z}_2, \oplus)$
d) $G = (\mathbb{Z}, +)$, $H = (\mathbb{Z}, +) \times (\mathbb{Z}, +)$

Úloha 11.3.4. Nech $(G, \circ)$ je grupa. Je zobrazenie $g \mapsto g^{-1}$ izomorfizmus z G na G? Ak nie, vedeli by ste definovať binárnu operáciu $*$ na G, tak, aby toto zobrazenie bol izomorfizmus grúp $(G, \circ)$ a $(G, *)$? Je uvedené zobrazenie izomorfizmom, ak G je komutatívna?

Úloha 11.3.5. Nech $(G, *)$ je ľubovoľná grupa. Dokážte, že zobrazenie $g \mapsto g * g$ je homomorfizmus z G do G práve vtedy, keď G je komutatívna.

Úloha 11.3.6. a) Dokážte, že $\mathbb{R} \times \mathbb{R} \setminus \{(0, 0)\}$ s operáciou $*$ definovanou ako $(a, b) * (c, d) = (ac - bd, ad + bc)$ tvorí grupu.

b) Dokážte, že všetky nenulové matice tvaru $\left(\begin{smallmatrix} a & b \\ -b & a \end{smallmatrix}\right)$ tvoria s násobením matíc grupu. (Hint k obom častiam úlohy: Možno vám pomôže nájsť jednoduchšie riešenie to, že táto úloha je v časti o homomorfizmoch.)

{homcvic:MONOJADRO}

Úloha 11.3.7. Nech $f\colon G \to H$ je homomorfizmus grúp. Dokážte:
a) Zobrazenie f je surjektívne práve vtedy, keď $\operatorname{Im} f = H$.
b) Zobrazenie g je injektívne práve vtedy, keď $\operatorname{Ker} f = \{e\}$.

{homcvic:HOMPROJ}

Úloha 11.3.8. Majme dané grupy G_1, G_2 a definujme zobrazenia $p_1\colon G_1 \times G_2 \to G_1$ a $p_2\colon G_1 \times G_2 \to G_2$ ako

$$p_1(x, y) = x,$$
$$p_2(x, y) = y.$$

Dokážte, že p_1 aj p_2 sú surjektívne homomorfizmy.

{homcvic:SUCINHOMNEW}

Úloha 11.3.9. Nech $f_1\colon G \to H_1$ a $f_2\colon G \to H_2$ sú homomorfizmy grúp. Potom aj zobrazenie $f\colon G \to H_1 \times H_2$ dané predpisom $f(x, y) = (g(x), h(y))$ je homomorfizmus.

{homcvic:SUCINHOM}

Úloha 11.3.10. Nech $g\colon G \to G'$ a $h\colon H \to H'$ sú homomorfizmy grúp. Potom aj zobrazenie $f\colon G \times H \to G' \times H'$ dané predpisom $f(x, y) = (g(x), h(y))$ je homomorfizmus. Ak g a h sú izomorfizmy (surjektívne homomorfizmy/injektívne homomorfizmy), tak f je izomorfizmus (surjektívny homomorfizmus/injektívny homomorfizmus).

{homcvic:ULODIAG}

Úloha 11.3.11. Nech G je grupa. Dokážte, že potom množina

$$\Delta = \{(x, x); x \in G\}$$

je podgrupa grupy $G \times G$. Ukážte, že Δ je izomorfná s G.

{homcvic:ULOHOMROV}

Úloha 11.3.12. Nech $f, g\colon G \to H$ sú homomorfizmy grúp. Je množina $\{a \in G; f(a) = g(a)\}$ podgrupa grupy G?

Úloha 11.3.13. Nech $f, g\colon (G, \circ) \to (H, *)$ sú homomorfizmy grúp. Definujme zobrazeníe $h\colon G \to H$ ako $h(x) = f(x) * g(x)$. Bude aj h homomorfizmus? Bude to platiť v prípade, že H je komutatívna?

Úloha 11.3.14. Nech $(H, \circ)$, $(G, *)$ sú grupy a $H \subseteq G$. Potom H je podgrupa G práve vtedy, keď zobrazenie $i\colon H \to G$ dané predpisom $i(h) = h$ je homomorfizmus.

Úloha 11.3.15. Nech $f\colon (G, *) \to (H, \circ)$ je homomorfizmus, ktorý je surjektívny ale nie injektívny.
a) Dokážte, že existuje aspoň jedno pravé inverzné zobrazenie k f, ktoré nie je homomorfizmom.
b) Ukážte na príklade, že môže nastať taká situácia, že žiadne pravé inverzné zobrazenie nie je homomorfizmom.
c) Ukážte na príklade, že môže nastať aj taká situácia, že aspoň jedno pravé inverzné zobrazenie je homomorfizmom.

Rozhodnite podobné otázky pre injektívny ale nesurjektívny homomorfizmus a ľavé inverzné zobrazenie.

Úloha 11.3.16. Dokážte, že ak grupa H je homomorfným obrazom komutatívnej grupy G, tak aj H je komutatívna.

Úloha 11.3.17. Nech $(G, \cdot)$ je grupa. Pre $a \in G$ označíme $C_G(a) = \{x \in G; xa = ax\}$. Dokážte, že $C_G(a)$ je podgrupa grupy G. Čomu sa rovná $C_G(e)$? Čomu sa rovná $C_G(a)$, ak G je komutatívna grupa?

Úloha 11.3.18. Nech $(G, *)$ je grupa. Ukážte, že $(G, *')$ s operáciou definovanou ako[3]

$$x *' y = y * x$$

je tiež grupa a že tieto dve grupy sú izomorfné, t.j. $(G, *) \cong (G, *')$.

11.4 Cyklické grupy

{cykl:SECTCYKL}

V tejto kapitole budeme často využívať označenie pre opakované použitie binárnej operácie, t.j.

$$x^n = \underbrace{x \circ \ldots \circ x}_{n\text{-krát}}.$$

Formálne toto označenie zavedieme pomocou definície matematickou indukciou.

Definícia 11.4.1. Nech $(G, \circ)$ je grupa a $x \in G$. Potom pre $n \in \mathbb{N}$ definujeme indukciou $x^1 = x$ a

$$x^{n+1} = x^n \circ x.$$

Ďalej definujeme $x^0 = e$, kde e je neutrálny prvok grupy G a $x^{-n} = (x^{-1})^n$ pre ľubovoľné $n \in \mathbb{N}$. (Tým je výraz x^k definovaný pre ľubovoľné $k \in \mathbb{Z}$.)

Ukážeme, že práve zadefinovaná mocnina v grupe sa správa podobne, ako celočíselné mocniny. (Tvrdenia v nasledujúcej leme sú na prvý pohľad jasné a ich formálny dôkaz je len cvičením na matematickú indukciu.)

{cykl:LMPOWER}

Lema 11.4.2. *Nech $(G, \circ)$ je grupa, $x, y \in G$, $m, n \in \mathbb{Z}$. Potom platí:*

(i) *Ak $x \circ y = y \circ x$, tak $x \circ y^n = y^n \circ x$.* {cykl:POWit5}

(ii) *Ak $x \circ y = y \circ x$, tak platí aj $(x \circ y)^n = x^n \circ y^n$.* {cykl:POWit2}

(iii) $x^{-n} = (x^n)^{-1}$ {cykl:POWit3}

(iv) $x^{m+n} = x^m \circ x^n = x^n \circ x^m$ {cykl:POWit1}

(v) $(x^m)^n = x^{mn}$ {cykl:POWit4}

Dôkaz. (i): Uvedené tvrdenie dokážeme najprv pre $n \in \mathbb{N}$. Budeme postupovať matematickou indukciou vzhľadom na n.

1° Pre $n = 0$ máme $x \circ e = e \circ x$, pre $n = 1$ máme $x \circ y = y \circ x$, čo je náš predpoklad.

2° Ak platí $x \circ y^n = y^n \circ x$ (indukčný predpoklad), tak dostaneme

$$x \circ y^{n+1} = x \circ (y^n \circ y) = (x \circ y^n) \circ y \overset{IP}{=} (y^n \circ x) \circ y = y^n \circ (x \circ y) = y^n \circ (y \circ x) = (y^n \circ y) \circ x = y^{n+1} \circ x.$$

Ak $n < 0$, označme $k = -n$. Potom z už dokázanej časti tvrdenia máme

$$x \circ y^n = x \circ y^{-k} = x \circ (y^{-1})^k = (y^{-1})^k \circ x = y^{-k} \circ x = y^n \circ x$$

[3] Wikipédia: Opposite group `https://en.wikipedia.org/wiki/Opposite_group`

(ii): Aj túto časť najprv overíme pre $n \in \mathbb{N}$.

Opäť pre $n = 0$ je platnosť tvrdenia zrejmá a pre $n = 1$ sa zhoduje priamo s predpokladom.

Predpokladajme, že (ii) platí pre n. Pre $n + 1$ dostávame postupne

$$(x \circ y)^{n+1} = (x \circ y)^n \circ (x \circ y) \overset{IP}{=} (x^n \circ y^n) \circ (x \circ y) = x^n \circ (y^n \circ x) \circ y \overset{\text{(i)}}{=} \\ x^n \circ (x \circ y^n) \circ y = (x^n \circ x) \circ (y^n \circ y) = x^{n+1} \circ y^{n+1}.$$

Rovnosť (ii) rozšírime na záporné čísla priamo použitím definície. Pre $n \in \mathbb{N}$ máme

$$(x \circ y)^{-n} = ((x \circ y)^{-1})^n = ((y \circ x)^{-1})^n = (x^{-1} \circ y^{-1})^n = (x^{-1})^n \circ (y^{-1})^n = x^{-n} \circ y^{-n}.$$

(iii): Najprv nech $n \in \mathbb{N}$. Z (ii) dostávame

$$x^{-n} \circ x^n = (x^{-1})^n \circ x^n = (x^{-1} \circ x)^n = e^n = e$$

(posledná rovnosť sa ľahko overí indukciou na n).

Rovnosť, ktorú sme odvodili, znamená, že $x^{-n} = (x^n)^{-1}$.

Ak $n < 0$, tak $n = -k$ pre nejaké $k \in \mathbb{N}$. Pre k už máme tvrdenie dokázané, čo znamená, že $x^{-k} = (x^k)^{-1}$ a

$$(x^n)^{-1} = ((x^k)^{-1})^{-1} = x^k = x^{-n}.$$

(iv) Overíme tvrdenie najprv pre $n \in \mathbb{N}$ a $m \geq -n$. Budeme postupovať indukciou vzhľadom na n. (T.j. indukciou na n dokazujeme výrok $(\forall m \geq -n)x^{m+n} = x^m \circ x^n = x^n \circ x^m$.)

1° Pre $n = 0$ máme $x^{m+0} = x^m \circ e = e \circ x^m$, čo evidentne platí.

2° Nech uvedená rovnosť platí pre n (a pre ľubovoľné $m \in \mathbb{N}$). Potom

$$x^{m+(n+1)} = x^{(m+n)+1} = x^{m+n} \circ x = (x^m \circ x^n) \circ x = x^m \circ (x^n \circ x) \overset{\text{IP}}{=} x^m \circ x^{n+1}$$

(Nerovnosť $m \geq -n$ sme potrebovali na to, aby $m + n \geq 0$, lebo iba v tomto prípade je $x^{(m+n)+1}$ definované uvedeným spôsobom. Takisto sme využili rovnosť $x^{n+1} = x^n \circ x$, ktorú zatiaľ máme len pre $n \geq 0$.)

Podobne dostaneme

$$x^{m+(n+1)} = x^{(m+n)+1} = x^{m+n} \circ x = (x^n \circ x^m) \circ x = x^n \circ (x^m \circ x) \overset{\text{(i)}}{=} \\ x^n \circ (x \circ x^m) = (x^n \circ x) \circ x^m = x^{n+1} \circ x^m.$$

Tým sme dokázali (iv) pre ľubovoľné $n \in \mathbb{N}$ a $m + n \geq 0$. Zo symetrie vyplýva, že platí aj pre $m \in \mathbb{N}$ a $m + n \geq 0$, čiže vlastne už máme túto rovnosť dokázanú pre ľubovoľné celé čísla m, n také, že $m + n \geq 0$.

Teraz, ak $m, n \in \mathbb{Z}$ sú také, že $m + n < 0$, tak

$$x^{m+n} = (x^{-1})^{(-m)+(-n)} = (x^{-1})^{-m} \circ (x^{-1})^{-n} = x^m \circ x^n,$$

podobne možno odvodiť druhú časť rovnosti (kde sú iba vymenené m a n.)

(v): Najprv túto rovnosť dokážeme pre $n \in \mathbb{N}$, $m \in \mathbb{Z}$ pomocou (iv) indukciou vzhľadom na n. (T.j. výrok $V(n)$, ktorý dokazujeme indukciou, je $(\forall m \in \mathbb{Z})(x^m)^n = x^{mn}$.)

1° Pre $n = 0$ máme rovnosť $e = e$, pre $n = 1$ máme $x^m = x^m$; v oboch prípadoch rovnosť platí.

2° Nech (v) platí pre n. Pre $n + 1$ potom dostaneme

$$(x^m)^{n+1} = (x^m)^n \circ x^m \overset{IP}{=} x^{mn} \circ x^m \overset{\text{(iv)}}{=} x^{mn+m} = x^{m(n+1)}.$$

Na záporné čísla túto rovnosť rozšírime nasledovne

$$(x^m)^{-n} = ((x^m)^{-1})^n \overset{\text{(iii)}}{=} ((x)^{-m})^n = (x)^{-mn} \overset{\text{(iii)}}{=} x^{mn}.$$

□

Ako sme už spomínali pre definícii 3.3.12, niekedy namiesto a^n používame zápis $n \times a$ (hlavne ak je grupová operácia označená ako $+$ alebo $\oplus$) – hovoríme o multiplikatívnom a aditívnom zápise grupovej operácie.

Definícia 11.4.3. Nech $(G, \circ)$ je grupa a $x \in G$. *Rád prvku* x v grupe G je najmenšie číslo $n \in \mathbb{N}$ také, že $n > 0$ a

$$x^n = e.$$

Ak také číslo neexistuje, rád prvku x sa definuje ako ∞.

Indukciou sa dá ľahko dokázať (úloha 11.4.1), že pre ľubovoľný homomorfizmus $f\colon (G, *) \to (H, \circ)$ platí

$$f(a^n) = f(a)^n. \tag{11.2}$$ {cykl:EQUMOCHOM}

Z toho môžeme odvodiť, že ak f je izomorfizmus, tak f zachováva rády prvkov. (T.j. rád prvku a v grupe G je rovnaký ako rád $f(a)$ v H. Pozri úlohu 11.4.1.)

Tento fakt môžeme využiť, ak chceme dokázať, že medzi niektorými dvoma grupami neexistuje izomorfizmus.

Príklad 11.4.4. Ukážeme, že grupy $\mathbb{Z}_2 \times \mathbb{Z}_2$ a $\mathbb{Z}_4$ nie sú izomorfné.

Stačí si uvedomiť, že v grupe $\mathbb{Z}_4$ je rád prvku 1 rovný 4. Ak by bola táto grupa izomorfná s grupou $\mathbb{Z}_2 \times \mathbb{Z}_2$, tak by aj v nej musel existovať prvok rádu 4. Prvky $(1,0)$, $(1,1)$, $(0,1)$ však majú rád 2 a neutrálny prvok $(0,0)$ má rád 1. Teda v $\mathbb{Z}_2 \times \mathbb{Z}_2$ nie je žiadny prvok rádu 4.

Postup z predchádzajúceho príkladu je dosť často používaný v prípade, že chceme dokázať neexistenciu izomorfizmu medzi dvoma grupami. Nájdeme nejakú vlastnosť, ktorú izomorfizmy zachovávajú (invariant) a ukážeme, že jedna z grúp túto vlastnosť nemá. Z vlastností s ktorými sme sa doteraz stretli sa dajú použiť napríklad rád prvku, veľkosť grupy alebo jej podgrúp (v zmysle počtu prvkov alebo kardinality[4]), existencia prvku spĺňajúceho nejakú identitu (ako napríklad $x * x = x$, $x^3 = a^2$ alebo $x * y \neq y * x$ – existencia prvkov, ktoré nekomutujú), atď. Niektoré z týchto vlastností sa dajú použiť aj na dôkaz, že neexistuje (surjektívny) homomorfizmus z jednej grupy do druhej.

Ľahko sa dá ukázať, že v konečnej grupe má každý prvok konečný rád (úloha 11.4.4). Tento fakt možno overiť veľmi podobným spôsobom, aký použijeme v nasledujúcom dôkaze.

{cykl:TVRKRITKON}

Tvrdenie 11.4.5. *Nech $(G, *)$ je grupa a $H \subseteq G$, $H \neq \emptyset$ je jej konečná podmnožina. Potom H je podgrupa G práve vtedy, keď platí*

$$a, b \in H \;\Rightarrow a * b \in H. \tag{11.3}$$ {cykl:EQKRIT}

Dôkaz. Implikácia $\boxed{\Rightarrow}$ je zrejmá.

$\boxed{\Leftarrow}$ Podľa vety 11.2.5(ii) nám stačí overiť, že pre každý prvok $a \in H$ aj inverzný prvok a^{-1} patrí do H.

Na to nám stačí ukázať, že prvok $a \in H$ má konečný rád – ak totiž vieme, že existuje prirodzené číslo n také, že $a^n = e$, tak $a^{n-1} * a = a * a^{n-1} = e$, čo znamená, že a^{n-1} je inverzný

[4] S pojmom kardinality (mohutnosti) množiny ste sa už pravdepodobne stretli na niektorej inej prednáške.

prvok k a. Z (11.3) však indukciou ľahko dostaneme, že $a^{n-1} \in H$. (Môžeme predpokladať, že $a \neq e$, a teda $n > 1$. Pre neutrálny prvok e dokazovaná implikácia očividne platí.)

Nech teda $a \in H$. Vieme, že pre ľubovoľné n aj $a^n \in H$. Keďže množina H je konečná, existujú $m \neq k$ také, že $a^m = a^k$. Bez ujmy na všeobecnosti, nech $m > k$. Potom z rovnosti $a^m = a^k$ dostaneme $a^{m-k} = e$. Ukázali sme teda existenciu prirodzeného čísla $n := m - k$ takého, že $a^n = e$. □

Definícia 11.4.6. *Cyklická grupa* je grupa G, ktorá je generovaná nejakým jej prvkom $a \in G$.

Prvok a, ktorý generuje grupu G, nazývame *generátor* grupy G.

Príklad 11.4.7. V príklade 11.2.17 sme videli, že $\mathbb{Z} = [1]$, teda $(\mathbb{Z}, +)$ je cyklická grupa. Súčasne platí $\mathbb{Z} = [-1]$, teda generátor cyklickej grupy nemusí byť jednoznačne určený.

{cykl:LMSUBAN}

Lema 11.4.8. *Ak $(G, *)$ je grupa a $a \in G$, tak $H = \{a^n; n \in \mathbb{Z}\}$ je podgrupa grupy G.*

Dôkaz. Pretože $e = a^0 \in H$, množina H je neprázdna. Overme, či pre H platí kritérium podgrupy.

Ak $a^n, a^m \in H$, tak aj $a^n * a^m = a^{n+m} \in H$.

Ak $a^n \in H$, tak aj $(a^n)^{-1} = a^{-n} \in H$. □

{cykl:VTTVARPRVKOV}

Veta 11.4.9. *Ak G je cyklická grupa a a je jej generátor, tak*

$$G = \{a^n; n \in \mathbb{Z}\},$$

t.j. G pozostáva práve z mocnín generátora a.

Dôkaz. Z predchádzajúcej lemy vieme, že $H = \{a^n; n \in \mathbb{Z}\}$ je podgrupa G obsahujúca prvok a. Preto $[a] \subseteq H$. Z predpokladu, že G je generovaná prvkom a potom máme $G \subseteq H$ a keďže H je podmnožina G, musí platiť $G = H$. □

V predchádzajúcom dôkaze sme vlastne súčasne ukázali, že

$$[a] = \{a^n; n \in \mathbb{Z}\}.$$

Lema 11.4.8 hovorí vlastne to, že pre každý prvok a obsahuje G cyklickú podgrupu generovanú prvkom a.

Pripomeňme, že pod označením $a \mid b$ (a delí b), kde $a, b \in \mathbb{Z}$ rozumieme to, že existuje nejaké celé číslo c také, že $b = c \cdot a$.

$$a \mid b \qquad \Leftrightarrow \qquad (\exists c \in \mathbb{Z})(b = c \cdot a)$$

{cykl:LMRADMOD}

Lema 11.4.10. *Ak $a \in G$, kde G je grupa, a rád prvku a je $n \in \mathbb{N}$, tak*

$$a^m = a^k \qquad \Leftrightarrow \qquad n \mid m - k.$$

Ak rád prvku a je ∞, tak

$$a^m = a^k \qquad \Leftrightarrow \qquad m = k.$$

Dôkaz. Uvažujme najprv prípad, že rád a je n.

$\boxed{\Rightarrow}$ Ak $a^m = a^k$, tak $a^{m-k} = e$. Nech $l = (m-k) \bmod n$ je zvyšok čísla $m-k$ po delení n. Potom $a^l = e$ a $0 \leq l < n$. Ak by platilo $l > 0$, dostali by sme spor s tým, že n je najmenšie prirodzené číslo s touto vlastnosťou. Preto musí platiť $l = 0$. Keďže $m - k$ má nulový zvyšok po delení číslom n, je to násobok čísla n.

$\boxed{\Leftarrow}$ Ak $m-k=cn$, t.j. $m=k+cn$, tak $a^m=a^{k+cn}=a^k(a^n)^c=a^ke^c=a^ke=a^k$.

Dôkaz prípadu, keď rád a je ∞, je veľmi podobný, snáď len s tým rozdielom, že implikácia $\boxed{\Leftarrow}$ je v tomto prípade zrejmá. Dokážme teda netriviálny smer.

$\boxed{\Rightarrow}$ Bez ujmy na všeobecnosti môžeme predpokladať $m \geq k$. Ak $a^m=a^k$, tak $a^{m-k}=e$. Pretože rád a je ∞, neexistuje číslo $n>0$ také, že $a^n=e$. To znamená, že $m-k=0$ a $m=k$. □

Ak sa pozrieme na to, čo nám táto lema hovorí v prípade $k=0$, tak dostaneme:

{cykl:DOSRADMOD}

Dôsledok 11.4.11. *Ak $a \in G$, kde G je grupa, a rád prvku a je $n \in \mathbb{N}$, tak*

$$a^m = e \qquad \Leftrightarrow \qquad n \mid m.$$

Ak rád prvku a je ∞, tak

$$a^m = e \qquad \Leftrightarrow \qquad m = 0.$$

Môžete sa pokúsiť dokázať si tento dôsledok priamo, bez použitia lemy 11.4.10 – úloha 11.4.11. Vedeli by ste dokázať lemu 11.4.10 pomocou dôsledku 11.4.11. (Teda dokázať tieto dve tvrdenia v opačnom poradí, než sme to urobili v tomto texte.)

{cykl:VTZN}

Veta 11.4.12. *Nech G je cyklická grupa a a je jej generátor. Ak rád prvku a je $n \in \mathbb{N}$, tak $G \cong (\mathbb{Z}_n, \oplus)$. Ak rád prvku a je ∞, tak $G \cong (\mathbb{Z}, +)$. (Teda každá cyklická grupa je izomorfná so $\mathbb{Z}$ alebo so $\mathbb{Z}_n$).*

Dôkaz. Najprv nech rád generátora a je rovný n. Definujme zobrazenie $f\colon k \mapsto a^k$, $f\colon \mathbb{Z}_n \to G$. Ukážeme, že f je izomorfizmus. Z lemy 11.4.2 máme

$$a^{k+l} = a^k a^l.$$

Súčasne, z lemy 11.4.10, máme

$$a^{k+l} = a^{k \oplus l}$$

pretože čísla $k+l$ a $k \oplus l$ sa líšia o nejaký násobok čísla n.

Ďalej ukážeme, že zobrazenie f je bijektívne.

Surjektívnosť: Keďže a je generátor grupy G, každý prvok tejto grupy má tvar a^s pre nejaké s (veta 11.4.9). Ak s' je zvyšok čísla s po delení číslom n, tak $s' \in \mathbb{Z}_n$ a navyše $n \mid s-s'$, takže (podľa predchádzajúcej lemy) $a^s = a^{s'} = f(s')$.

Injektívnosť: Ak $f(s)=f(t)$, čiže $a^s=a^t$, máme $n \mid s-t$. Ak $s,t \in \{0,1,\ldots,n-1\}$, tak $s-t \in \{0, \pm 1, \ldots, \pm(n-1)\}$. Jediné číslo z tejto množiny, ktoré je deliteľné n, je 0, a teda $s-t=0$ a $s=t$.

Zostáva nám ešte ukázať druhý prípad, t.j. rád a je ∞. (Dôkaz v tomto prípade je veľmi podobný.) Definujeme $f\colon \mathbb{Z} \to G$ ako $f(n)=a^n$.

Homomorfizmus:

$$f(k+l) = a^{k+l} = a^k a^l = f(k)f(l)$$

podľa lemy 11.4.2.

Surjektívnosť: Každý prvok z G je tvaru $a^n = f(n)$ pre nejaké $n \in \mathbb{Z}$ podľa vety 11.4.9.

Injektívnosť: Druhá časť lemy 11.4.10. □

Špeciálne z predchádzajúcej vety vidíme, že každá netriviálna grupa musí obsahovať podgrupu homomorfnú so $(\mathbb{Z},+)$ alebo s niektorým $(\mathbb{Z}_n, \oplus)$. Takisto si môžeme všimnúť, že rád prvku je presne počet prvkov podgrupy generovanej týmto prvkom.

Ešte sa skúsme zaoberať otázkou, či z cyklickej grupy dostaneme opäť cyklickú grupu pomocou niektorých základných operácií – podgrupa, homomorfný obraz, priamy súčin grúp (definícia 11.1.5).

Veta 11.4.13. *Každá podgrupa cyklickej grupy je cyklická.*

Dôkaz. Nech G je cyklická grupa s generátorom a a H je nejaká jej podgrupa. Nech d je najmenšie prirodzené číslo také, že $d > 0$ a $a^d \in H$. (Ak $H \neq \{e\}$, tak existuje aspoň jedno také číslo.) Dokážeme, že a^d generuje H (a teda H je cyklická).

Najprv si uvedomme, aké prvky patria do podgrupy $[a^d]$ generovanej prvkom a^d. Podľa vety 11.4.9 sú to presne prvky tvaru $(a^d)^k = a^{kd}$ pre $k \in \mathbb{Z}$, čiže tie mocniny generátora a, pre ktoré je exponent násobkom d.

Postupujme sporom. Nech by existoval prvok $a^s \in H$ taký, že. $a^s \notin [a^d]$. Číslo s môžeme prepísať ako $s = k \cdot d + s'$, kde $0 \leq s' < d$ (číslo s sme vydelili číslom d, zvyšok sme označili s'). Ak predpokladáme, že $a^s \notin [a^d]$, tak $s' \neq 0$. (V opačnom prípade by totiž platilo $a^s = a^{kd}$ a už sme ukázali, že prvky takéhoto tvaru patria do $[a^d]$.) Z rovnosti $a^s = a^{kd}a^{s'}$ dostaneme

$$a^{s'} = (a^{kd})^{-1}a^s,$$

a pretože $a^{kd}, a^s \in H$ a H je podgrupa, z tejto rovnosti vyplýva $a^{s'} \in H$. Pretože $0 < s' < d$, dostávame tak spor s predpokladom, že d je najmenší možný exponent taký, že $a^d \in H$. □

Veta 11.4.14. *Homomorfný obraz cyklickej grupy je cyklická grupa.*

Dôkaz. Ak $f\colon G \to H$ je epimorfizmus a a je generátor G, tak ľahko možno vidieť, že $f(a)$ je generátor grupy H.

Skutočne, každý prvok z H je tvaru $f(a^n)$ pre nejaké $n \in \mathbb{Z}$ (na základ surjektívnosti f) a podľa (11.2) platí $f(a^n) = f(a)^n$, čiže ho vieme dostať ako mocninu prvku $f(a)$. □

Na vyriešenie otázky, kedy je súčin cyklických grúp opäť cyklická grupa, budeme potrebovať pomocné tvrdenie týkajúce sa najväčšieho spoločného deliteľa. Nebudeme ho na tomto mieste dokazovať, keďže neskôr dokážeme všeobecnejšiu verziu tohoto výsledku – konkrétne v tvrdení 13.4.22. Ak by Vás zaujímal jeho dôkaz už teraz, môžete si ho pozrieť napríklad v [ŠHHK, Lema 3.1.3], [Č], [Sl3] (a v podstate v takmer každom úvodnom texte zaoberajúcom sa deliteľnosťou). Najprv však pripomeňme definíciu najväčšieho spoločného deliteľa.

{cykl:DEFGCD}

Definícia 11.4.15. Najväčší spoločný deliteľ čísel $a, b \in \mathbb{Z}$ je číslo $d \in \mathbb{N} \setminus \{0\}$ s vlastnosťami

(i) $d \mid a \wedge d \mid b$ (čiže d delí obe čísla);

(ii) ak $c \mid a$ a $c \mid b$, tak $c \leq d$ (čiže d je najväčšie číslo s uvedenou vlastnosťou).

Najväčší spoločný deliteľ budeme označovať ako $d = \gcd(a, b)$.

{cykl:TVRBEZOUT}

Tvrdenie 11.4.16. *Ak $d = \gcd(m, n)$ je najväčší spoločný deliteľ dvoch čísel $m, n \in \mathbb{N}$ tak existujú také $u, v \in \mathbb{Z}$, že platí $um + vn = d$.*

Napríklad pre čísla 3 a 7 máme $\gcd(3, 7) = 1$ a skutočne platí $(-2) \cdot 3 + 1 \cdot 7 = 1$. Ak skúsime o niečo väčšie čísla, tak napríklad $\gcd(14, 9) = 1$ a $2 \cdot 14 - 3 \cdot 9 = 1$.

Veta 11.4.17. *Grupa $\mathbb{Z}_m \times \mathbb{Z}_n$ je cyklická práve vtedy, keď m a n sú nesúdeliteľné, t.j. ich najväčší spoločný deliteľ $\gcd(m, n) = 1$. V takomto prípade je prvok $(1, 1)$ jej generátorom.*

Dôkaz. $\boxed{\Rightarrow}$ Najprv ukážeme, že ak $\mathbb{Z}_m \times \mathbb{Z}_n$ je cyklická grupa, tak dvojica $(1, 1)$ je jej generátorom. Nech (g_1, g_2) je ľubovoľný generátor. Zrejme potom g_1 je generátor $\mathbb{Z}_m$ a g_2 je generátor $\mathbb{Z}_n$. Potom priradenie $g_1 \mapsto 1$ jednoznačne určuje izomorfizmus $f_1\colon \mathbb{Z}_m \to \mathbb{Z}_m$ a priradenie $g_2 \mapsto 1$ nám dáva izomorfizmus $f_2\colon \mathbb{Z}_n \to \mathbb{Z}_n$. Potom $f\colon \mathbb{Z}_m \times \mathbb{Z}_n \to \mathbb{Z}_m \times \mathbb{Z}_n$ určené predpisom

$$f(x, y) = (f_1(x), f_2(y))$$

je tiež izomorfizmus (úloha 11.3.10). Platí $f(g_1, g_2) = (1, 1)$, čiže $(1, 1)$ je generátor.

Ak $d = \gcd(m, n) > 1$, tak podgrupa $[(1, 1)]$ generovaná dvojicou $(1, 1)$ neobsahuje dvojicu $(1, 0)$. Skutočne, ak by platilo pre nejaké $k \in \mathbb{Z}$ v tejto grupe $k \times (1, 1) = (1, 0)$, znamenalo by to, že v celých číslach platia rovnosti

$$k = am + 1$$
$$k = bn$$

z čoho $bn - am = 1$. Pritom však d delí obe čísla m aj n, teda $d \mid bn - am = 1$, čo je spor. (Jediné celočíselné delitele čísla 1 sú ± 1, my sme predpokladali $d > 1$.)

$\boxed{\Leftarrow}$ Očividne $[(1, 0), (0, 1)] = \mathbb{Z}_m \times \mathbb{Z}_n$, stačí teda ukázať, že pomocou $(1, 1)$ vieme vygenerovať dvojice $(1, 0)$ a $(0, 1)$. Ak $am + bn = 1$, tak v grupe $\mathbb{Z}_m \times \mathbb{Z}_n$ platia rovnosti (zapísané aditívne)

$$am \times (1, 1) = (0, 1)$$
$$bn \times (1, 1) = (1, 0)$$

□

V príklade 11.2.18 sme videli, že grupa $\mathbb{Z}_2 \times \mathbb{Z}_3$ je skutočne generovaná prvkom $(1, 1)$.

Cvičenia

{cykl:ULOIZORAD}

Úloha 11.4.1. Nech $f\colon (G, *) \to (H, \circ)$ je homomorfizmus. Dokážte, že potom platí:
a) $f(a^n) = f(a)^n$
b) $a^n = e_G \Rightarrow f(a)^n = e_H$
c) Ak f je navyše izomorfizmus, tak $a^n = e_G \Leftrightarrow f(a)^n = e_H$.
d) Izomorfizmus zachováva rád prvku, t.j. rád prvku a v grupe G je rovnaký ako rád prvku $f(a)$ v grupe H.
e) Viete niečo povedať o vzťahu medzi rádmi prvkov a a $f(a)$ aj ak f je homomorfizmus? (T.j. bez predpokladu o bijektívnosti.)

Úloha 11.4.2. Nech $(G, *)$ je cyklická grupa a a je jej generátor. Potom ľubovoľný homomorfizmus z G do nejakej grupy $(H, \circ)$ je jednoznačne určený obrazom prvku a.

Úloha 11.4.3. Nájdite všetky izomorfizmy medzi $(\mathbb{Z}_4, \oplus)$ a $(\mathbb{Z}_5 \setminus \{0\}, \odot)$.

{cykl:ULOKONRAD}

Úloha 11.4.4. Ukážte, že ak G je konečná grupa, tak každý jej prvok má konečný rád.

Úloha 11.4.5. a) Ukážte, že grupa $(\mathbb{Q}, +)$ nie je cyklická.
b) Ukážte, že grupa $(\mathbb{R}, +)$ nie je cyklická.

Úloha 11.4.6. Nájdite všetky homomorfizmy:
a) zo $\mathbb{Z}_4$ do $\mathbb{Z}_2 \times \mathbb{Z}_2$,
b) zo $\mathbb{Z}_2 \times \mathbb{Z}_2$ do $\mathbb{Z}_4$,

Úloha 11.4.7. Zistite, či sú grupy G a H izomorfné. Svoju odpoveď zdôvodnite!
a) $G = (\mathbb{Z}_7 \setminus \{0\}, \odot)$, $H = (\mathbb{Z}_6, \oplus)$
b) $G = (\mathbb{Z}, +)$, $H = (\mathbb{Q}, +)$
c) $G = (\mathbb{Z}_6, \oplus)$, $(\mathbb{Z}_2, \oplus) \times (\mathbb{Z}_3, \oplus)$

Úloha 11.4.8. V každej grupe majú nasledujúce prvky rovnaký rád: x a yxy^{-1}; ab a ba; abc, bca a cab. Naopak, prvky abc a cba môžu mať rôzny rád. (Hint: Jedna z možností ako dokázať, že dva prvky $g, h \in G$ majú rovnaký rád je dokázať ekvivalenciu $g^n = e \Leftrightarrow h^n = e$. Iná možnosť je nájsť izomorfizmus $f\colon G \to G$ taký, že $f(g) = h$, a použiť úlohu 11.4.1a).)

Úloha 11.4.9. Nech $a, b \in G$, kde G je grupa, $a, b \neq e$ také, že $ab = ba$ a $b^3 = 1$. Dokážte, že $\{a^n, ba^n, b^2a^n; n \in \mathbb{Z}\}$ je podgrupa grupy G.

Úloha 11.4.10. Nech $n \in \mathbb{N} \setminus \{0\}$. Pre každé $k \mid n$ existuje k-prvková podgrupa grupy $(\mathbb{Z}_n, \oplus)$.

{cyklcvic:ULORADMOD}

Úloha 11.4.11. Ak rád prvku a v grupe G je n a e je neutrálny prvok tejto grupy, tak pre prirodzené čísla $k \in \mathbb{N}$ platí $a^k = e$ práve vtedy, keď $n \mid k$. Ďalej pre každé $s \in \mathbb{N}$ existuje $m \in \mathbb{N}$ také, že $a^s = a^m$ a $0 \leq m \leq n - 1$.

Úloha 11.4.12. Nech x je prvok rádu n. Potom:
a) Rád prvku x^m delí n.
b*) Rád prvku x^m je $\frac{n}{(m,n)}$, kde (m, n) označuje najväčší spoločný deliteľ čísel m a n. (Hint: Tvrdenie 11.4.16.)

Úloha 11.4.13. Ak $f\colon G \to H$ je homomorfizmus grúp a $g \in G$ je prvok konečného rádu, tak rád prvku $f(g)$ delí rád prvku G.

Úloha 11.4.14. Nech G je grupa, $a \in G$ je prvok konečného rádu n. Nech H je podgrupa grupy G a $k = \min\{j \in \mathbb{N}; j > 0, a^j \in H\}$, t.j. k je najmenšie kladné prirodzené číslo také, že $a^k \in H$. Dokážte, že $k \mid n$; t.j. k delí rád prvku g.

Úloha 11.4.15. Nech G je grupa regulárnych matíc typu 2×2 nad $\mathbb{R}$ s násobením a $A = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$, $B = \left(\begin{smallmatrix} 0 & 1 \\ -1 & -1 \end{smallmatrix}\right)$. Dokážte, že A má rád 4, B má rád 3, ale AB má rád ∞.

Úloha 11.4.16. Nech G je grupa a $a, b \in G$ sú prvky také, že a, b aj ab majú rád 2. Dokážte, že $ab = ba$.

Úloha 11.4.17. Nech $x, y \in G$, kde G je grupa. Ukážte, že ak $x^3y = yx^3$ a x má rád 7, tak $xy = yx$.

Úloha 11.4.18*. Nech $a, b \in G$, kde G je grupa. Ukážte, že ak $ba = a^4b^3$, tak prvky a^4b a a^2b^3 majú rovnaký rád.

Úloha 11.4.19. Nech x a y sú prvky konečného rádu grupy G a nech $xy = yx$. Dokážte, že ak $[x] \cap [y] = \{e\}$, čiže prienik podgrúp generovaných týmito prvkami je triviálna podgrupa, tak rád prvku xy je najmenší spoločný násobok rádov prvkov x a y. Platia tieto tvrdenia aj ak x a y nekomutujú?

Úloha 11.4.20*. Nech x a y sú prvky konečného rádu grupy G a nech $xy = yx$. Dokážte, že ak ich rády k a l sú nesúdeliteľné, tak rád prvku xy je kl. Dokážte, že existujú exponenty m a n, také, že rád prvku $x^m y^n$ je rovný $[k, l]$ (najmenší spoločný násobok rádov). Platia tieto tvrdenia aj ak x a y nekomutujú? (Hint: V tomto príklade môže byť užitočné tvrdenie 11.4.16.)

11.5 Permutácie

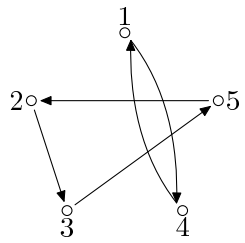
{SECTPERM}

S permutáciami sme sa už stretli pri definícii determinantu (kapitola 6).

Z minulého semestra už vieme, že pod *permutáciou* konečnej[5] množiny M rozumieme bijekciu z M od M. Tiež sme sa dohodli na označení permutácií množiny $\{1, \ldots, n\}$ pomocou zápisu

$$\left(\begin{smallmatrix} 1 & 2 & \ldots & n \\ \varphi(1) & \varphi(2) & \ldots & \varphi(n) \end{smallmatrix}\right),$$

[5]Názov permutácia sa niekedy vyskytuje aj ako pomenovanie bijekcie nekonečnej množiny na seba. My sa budeme držať terminológie, ktorú sme zaviedli už v minulom semestri – t.j. pri permutáciach automaticky predpokladáme, že ide o konečnú množinu.



Obr. 11.1: Príklad permutácie 5-prvkovej množiny. Vidíme, že z každého prvku vychádza aj do každého prvku vchádza práve jedna šípka. {perm:FIG5A}

Napríklad permutáciu znázornenú na obrázku 11.1 by sme zapísali ako

$$\left(\begin{smallmatrix}1&2&3&4&5\\4&3&5&1&2\end{smallmatrix}\right)$$

Tiež vieme, že zloženie permutácií je permutácia a inverzné zobrazenie k danej permutácii je permutácia. Pomocou týchto faktov vieme ukázať, že permutácie danej konečnej množiny tvoria grupu. Množinu všetkých permutácií n-prvkovej množiny $\{1, 2, \ldots, n\}$ budeme označovať S_n a nazývať *symetrická grupa.*

V tejto časti sa budeme zaoberať o niečo podrobnejšie grupou S_n všetkých permutácií n-prvkovej množiny a niektorými jej podgrupami. Začneme tým, že si povieme o rozklade permutácie na jednoduchšie permutácie (budeme ich volať cykly), ktorý nám (okrem iného) umožní vyrátať rád ľubovoľného prvku v grupe S_n.

11.5.1 Rozklad na súčin disjunktných cyklov

{perm:SSECTCYKLY}

Ak sa pozrieme na permutáciu na obrázku 11.1 vidíme, že na tomto obrázku sa po jednotlivých šípkach z prvku 1 dostaneme do prvku 4 a potom naspäť znovu do prvku 1. Podobne z prvku 2 sa dostaneme do 3 a 5 a potom opäť do 2.

Toto musí platiť pre ľubovoľnú permutáciu. Ak začneme s ľubovoľným prvkom, opakovaným aplikovaním tej istej permutácie po konečnom počte krokov dostaneme znovu ten istý prvok. Dôvodom je, že máme len konečný počet prvkov – teda po istom čase sa niektorý prvok zopakuje. Ak by to bol iný prvok, ako ten z ktorého sme začali, dané zobrazenie by nebola bijekcia. (Podrobnejšie tento argument vysvetlíme v dôkaze tvrdenia 11.5.7.) Vďaka tomu môžeme každú permutáciu rozdeliť na cykly.

Celá táto časť je venovaná spôsobu ako môžeme práve uvedený jednoduchý fakt formálne zapísať a dokázať. Tiež si povieme, prečo je tento poznatok užitočný.

Najprv potrebujeme zaviesť pojem cyklu.

Definícia 11.5.1. Permutáciu φ konečnej množiny M nazveme *cyklus,* alebo *cyklická permutácia* ak existujú prvky $a_1, a_2, \ldots, a_k$ také, že

$$\begin{cases}\varphi(a_i) = a_{i+1} \text{ pre } i = 1, 2, \ldots, k-1,\\ \varphi(a_k) = a_1,\\ \varphi(a) = a \text{ pre ostatné prvky } a \neq a_i.\end{cases}$$

Pre cyklus tohoto tvaru budeme používať zápis $(a_1 a_2 \ldots a_k)$.

V definícii cyklu pripúšťame aj nulový počet prvkov. *Prázdny cyklus,* ktorý označujeme $()$, sa rovná identickej permutácii.

Všimnime si, že zápis pomocou cyklov môžeme použiť pre ľubovoľnú množinu. Tento zápis však neidentifikuje množinu, ktorej permutáciu robíme – tá musí byť zadaná zvlášť.

Príklad 11.5.2. Uvažujme permutácie množiny $\{1, 2, 3, 4, 5\}$

Zápis $\varphi = (14)$ označuje permutáciu s vlastnosťou $\varphi(1) = 4$ a $\varphi(4) = 1$, ktorá prvky 2, 3 a 5 nemení (teda $\varphi(2) = 2$, $\varphi(3) = 3$, $\varphi(5) = 5$).

Zápis $\tau = (235)$ znamená permutáciu určenú predpisom $\tau(1) = 1$, $\tau(2) = 3$, $\tau(3) = 5$, $\tau(4) = 4$, $\tau(5) = 2$.

Všimnime si, že ten istý cyklus môže byť zapísaný viacerými spôsobmi: $\tau = (235) = (352) = (523)$.

Tiež si môžeme všimnúť, že identickú permutáciu môžeme zapísať nielen ako prázdny cyklus ale aj ako ľubovoľný jednoprvkový cyklus: $id = () = (1) = (2) = (3) = (4) = (5)$

Príklad 11.5.3. Vieme, že inverzné zobrazenie k permutácii je tiež permutácia. Ak máme cyklus $\varphi = (a_1 a_2 \dots a_n)$, tak inverzná permutácia je tiež cyklus $\varphi^{-1} = (a_n a_{n-1} \dots a_1)$. Zodpovedá to tomu, že poprechádzame po tých istých šípkach tvoriacich cyklus, ale v opačnom poradí.

Pre cykly z predchádzajúceho príkladu máme $\varphi^{-1} = (14)$ a $\tau^{-1} = (253)$.

Definícia 11.5.4. Dve permutácie φ a τ tej istej množiny M nazveme *disjunktné,* ak pre každý prvok $a \in M$ platí $\varphi(a) = a$ alebo $\tau(a) = a$. (Teda každý prvok zostáva nezmenený pri aspoň jednej z týchto dvoch permutácií.)

Špeciálne, dva cykly $(a_1 a_2 \dots a_k)$ a $(b_1 b_2 \dots b_l)$, kde $k, l \geq 2$, sú disjunktné, ak platí rovnosť $\{a_1, a_2, \dots, a_k\} \cap \{b_1, b_2, \dots, b_l\} = \emptyset$.

Nasledujúca lema hovorí, že poradie disjunktných permutácií môžeme pri skladaní vymeniť – disjunktné permutácie komutujú.

{perm:LMDISJKOM}

Lema 11.5.5. *Ak φ a τ sú disjunktné permutácie, tak*

$$\varphi \circ \tau = \tau \circ \varphi.$$

Dôkaz. Stačí priamym výpočtom overiť, že permutácia na ľavej aj pravej strane rovnosti nadobúda rovnaké hodnoty; čiže $\tau(\varphi(m)) = \varphi(\tau(m))$.

Dokážme túto rovnosť najprv pre prípad, že $\varphi(m) \neq m$. Potom z injektívnosti zobrazenia φ máme $\varphi(\varphi(m)) \neq \varphi(m)$. Pretože φ a τ sú disjunktné, permutácia τ nemení prvky m ani $\varphi(m)$, teda $\tau(m) = m$ a $\tau(\varphi(m)) = \varphi(m)$. Použitím týchto rovností dostaneme

$$\tau(\varphi(m)) = \varphi(m) = \varphi(\tau(m)).$$

Prípad $\tau(m) \neq m$ je symetrický.

Zostáva už len možnosť $\tau(m) = \varphi(m) = m$, vtedy však na oboch stranách rovnosti dostaneme m:

$$\tau(\varphi(m)) = \varphi(\tau(m)).$$

□

Príklad 11.5.6. Ľahko si môžeme všimnúť, že pre permutácie, ktoré nie sú disjunktné, predchádzajúce tvrdenie neplatí. Zoberme napríklad cykly $\varphi = (12)$ a $\tau = (135)$. Potom dostaneme

$$\varphi \circ \tau = (1352) \qquad \text{a} \qquad \tau \circ \varphi = (1235).$$

{perm:TVRCYKL}

Tvrdenie 11.5.7. *Každú permutáciu možno zapísať ako zloženie disjunktných cyklov. Tento zápis je jednoznačný až na poradie cyklov (a vynechanie prázdneho cyklu a jednoprvkových cyklov). Nazývame ho* rozklad permutácie na súčin disjunktných cyklov.

Dôkaz. Uvažujme ľubovoľnú permutáciu φ množiny M.

Existenciu rozkladu permutácie φ na súčin disjunktných dokážeme indukciou vzhľadom na počet prvkov množiny M. Pre jednoprvkovú množinu máme jedinú možnú permutáciu id_M, ktorá sa rovná prázdnemu cyklu.

Predpokladajme teraz, že M má n prvkov a pre ľubovoľnú množinu s menej ako n prvkami tvrdenie platí. Zoberme teraz ľubovoľný prvok $a \in M$. Položme $a_1 = a$, $a_{i+1} = \varphi(a_i)$ pre všetky $i = 1, 2, \ldots$. Nech k je najmenšie číslo také, že $k \geq 1$ a $a_{k+1} = a$. Ukážeme najprv, že také číslo musí existovať.

Pretože množina M je konečná, musia sa niektoré prvky v postupnosti (a_k) opakovať, t.j. existujú nejaké čísla s vlastnosťou $a_s = a_t$ a $s \neq t$. Tvrdíme, že ak s je najmenšie možné takéto číslo, tak $s = 1$. Skutočne, v opačnom prípade máme $\varphi(a_{s-1}) = \varphi(a_{t-1})$ a z injektívnosti zobrazenia φ dostaneme $a_{s-1} = a_{t-1}$.

Zistili sme, že existuje aspoň jedno číslo s horeuvedenými vlastnosťami, preto môžeme definovať k ako najmenšie číslo, ktoré má tieto vlastnosti. Pomocou neho definujeme cyklus

$$\tau = (a_1 a_2 \ldots a_k).$$

Položme ďalej $\psi = \tau^{-1}\varphi$. Ukážeme, že permutácie ψ a τ sú disjunktné a že $\psi(a) = a$.

Skutočne, pre ľubovoľný z prvkov $a_1, a_2, \ldots, a_{k-1}$ platí $\varphi(a_i) = a_{i+1}$ a $\tau^{-1}(a_{i+1}) = a_i$, teda $\psi(a_i) = \tau^{-1}(\varphi(a_i)) = a_i$. Podobne môžeme overiť, že $\psi(a_k)\tau^{-1}(\varphi(a_k)) = \tau^{-1}(a_1) = a_k$. Vidíme, že permutácia ψ nemení žiadny z prvkov, ktoré mení cyklus τ teda tieto permutácie sú disjunktné. Špeciálne, ψ nemení prvok $a_1 = a$.

Teraz si stačí uvedomiť, že ψ (resp. zúženie tohoto zobrazenia) môžeme chápať ako permutáciu množiny $M \setminus \{a_1, \ldots, a_k\}$, ktorá má menej prvkov ako množina M. Podľa indukčného predpokladu teda existuje rozklad $\psi = \tau_1\tau_2 \ldots \tau_l$ tejto permutácie na disjunktné cykly. Potom

$$\varphi = \tau(\tau^{-1}\varphi) = \tau\psi = \tau\tau_1 \ldots \tau_l$$

je rozklad permutácie φ na disjunktné cykly.

Jednoznačnosť. Predpokladajme, že máme 2 rozklady

$$\varphi_1 \ldots \varphi_s = \psi_1 \ldots \psi_t$$

tej istej permutácie na disjunktné cykly, pričom v rozkladoch sa nevyskytuje prázdny cyklus. Postupujme indukciou vzhľadom na s.

Ak $s = 0$, teda na ľavej strane rovnosti je identita, musí byť aj počet disjunktných cyklov na pravej strane nulový. V opačnom prípade by sme mali na pravej strane aspoň jeden neprázdny cyklus, teda permutácia na pravej strane by menila aspoň jeden prvok, čiže permutácia na pravej strane by nebola identita.

Predpokladajme teraz, že $s > 0$, čiže daná permutácia nie je identita. Nech a je nejaký prvok, ktorý táto permutácia mení. Tento prvok sa musí vyskytovať v aspoň jednom z cyklov, súčasne z disjunktnosti vyplýva, že sa nemôže vyskytovať vo viacerých cykloch. Teda prvok a sa vyskytuje práve v jednom z cyklov vystupujúcich na ľavej strane rovnosti, podobne v práve jednom z cyklov na pravej strane. Bez ujmy na všeobecnosti (cykly môžeme poprehadzovať) predpokladajme, že sú to cykly φ_1 a τ_1. Potom dostaneme $\varphi_1(a) = \varphi(a) = \tau_1(a)$, podobne $\varphi_1^2(a) = \varphi^2(a) = \tau_1^2(a)$ atď. Čiže všetky prvky vyskytujúce sa v cykle φ_1 sa zobrazia rovnako aj cyklom τ_1 a obrátene. Platí teda $\varphi_1 = \tau_1$. Z toho dostaneme rovnosť $\varphi_2 \ldots \varphi_s = \psi_2 \ldots \psi_t$. Pre tieto cykly už môžeme použiť indukčný predpoklad. □

Príklad 11.5.8. Pre permutáciu $\varphi = \left(\begin{smallmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 3 & 5 & 1 & 2 \end{smallmatrix}\right)$ (obrázok 11.1) dostaneme rozklad $\varphi = (14)(235)$.

Z toho vieme hneď zistiť aj $\varphi^{-1} = (14)(253)$.

Rozklad na súčin disjunktných cyklov môže byť užitočný pri výpočte rádu permutácie (v zmysle rádu prvku grupy všetkých permutácií danej množiny).

Definícia 11.5.9. Ak φ je permutácia konečnej množiny M, tak *rád permutácie* φ je najmenšie prirodzené číslo $n \geq 1$ také, že

$$\varphi^n = id_M.$$

Napríklad v príklade 2.3.2 sme zistili, že rád permutácie $\varphi = \left(\begin{smallmatrix}1&2&3&4\\1&3&4&2\end{smallmatrix}\right)$ je 3. (Všimnime si, že ide o cyklickú permutáciu.)

{perm:LMRADCYKLU}

Lema 11.5.10. *Rád cyklu je rovný jeho dĺžke, t.j. ak $\varphi = (a_1 \dots a_n)$, tak rád φ je rovný n.*

Dôkaz. Pre $k = 1, 2, \dots, n-1$ je $\varphi^k(a_1) = a_k \neq a_1$. Až pre $k = n$ prvýkrát dostaneme $\varphi^k(a_1) = a_1$. Úplne rovnaké zdôvodnenie prejde aj pre ostatné prvky cyklu, čiže máme $\varphi^n = id$ a navyše n je najmenšie číslo z množiny $\{1, 2, 3, \dots\}$ s touto vlastnosťou. □

Z predchádzajúcej lemy vidíme, že ak n je rád cyklu φ, tak $\varphi^k = id_M$ platí práve vtedy, keď k je násobkom n, t.j. $n \mid k$.

Veta 11.5.11. *Rád permutácie je najmenší spoločný násobok dĺžok disjunktných cyklov, ktoré vystupujú v jej rozklade.*

Dôkaz. Nech $\varphi = \tau_1 \dots \tau_m$ je rozklad permutácie na φ na disjunktné cykly. Pretože disjunktné cykly komutujú, mocniny permutácie φ môžeme vyjadriť ako

$$\varphi^n = \tau_1^n \dots \tau_m^n.$$

Pritom permutácie $\tau_1^n, \dots, \tau_m^n$ sú opäť disjunktné. Z toho vyplýva, že aby sme dostali identické zobrazenie, musí pre každé $i = 1, 2, \dots, m$ platiť $\tau_i^n = id_M$. To nastane práve vtedy, keď n je násobkom rádu τ_i (pre všetky i). Teda najmenšie možné n s takouto vlastnosťou je najmenší spoločný násobok rádov jednotlivých disjunktných cyklov. Podľa lemy 11.5.10 sú to presne dĺžky jednotlivých cyklov. □

Príklad 11.5.12. Vypočítajme rád permutácie $\varphi = \left(\begin{smallmatrix}1&2&3&4&5\\4&3&5&1&2\end{smallmatrix}\right) = (14)(235)$. Podľa predchádzajúcej vety by to mal byť najmenší spoločný násobok čísel 2 a 3, čiže 6.

Keď sa to pokúsime overiť na základe definície, dospejeme k tomu istému výsledku:
$\varphi^2 = (253)$,
$\varphi^3 = (14)$,
$\varphi^4 = (235)$,
$\varphi^5 = (14)(253)$,
$\varphi^6 = id$.

11.5.2 Parita permutácie

S pojmom inverzie sme sa už stretli pri definícii determinantu. Teraz ho použijeme na zadefinovanie pojmu párnej a nepárnej permutácie.

Definícia 11.5.13. Dvojica $(\varphi(k), \varphi(s))$ sa volá *inverzia* permutácie φ, ak $k < s$ ale $\varphi(k) > \varphi(s)$.

Ak má permutácia φ párny počet inverzií, hovoríme, že je to *párna permutácia,* v opačnom prípade hovoríme o *nepárnej permutácii.*

Príklad 11.5.14. Permutácia $\varphi = \left(\begin{smallmatrix}1&2&3&4&5\\4&3&5&1&2\end{smallmatrix}\right)$ má 7 inverzií: (4,3), (4,1), (4,2), (3,1), (3,2), (5,1) a (5,2). Teda táto permutácia je nepárna.

Na dôkaz niektorých vlastností parity permutácií sa nám budú hodiť cykly dĺžky dva.

Definícia 11.5.15. Permutáciu tvaru (a_1a_2) (t.j. dvojprvkový cyklus) budeme nazývať *transpozícia.*

{perm:PRIKLTRANSNEPAR}

Príklad 11.5.16. Ukážeme, že transpozícia je nepárna permutácia. Bez ujmy na všeobecnosti, nech $a_1 < a_2$. Nech $\varphi = (a_1a_2)$, t.j. $\varphi(a_1) = a_2$, $\varphi(a_2) = a_1$ a $\varphi(a) = a$ pre $a \neq a_1, a_2$. Potom všetky inverzie permutácie φ sú $(a_2 = \varphi(a_1), i)$ a $(i, a_1 = \varphi(a_2))$ pre $a_1 < i < a_2$ (týchto je párny počet) a inverzia (a_2, a_1).

{perm:TVRPERMTRANS}

Tvrdenie 11.5.17. *Každú permutáciu možno zapísať ako zloženie transpozícií. Navyše, pri každom takomto zápise je parita počtu transpozícií rovná parite permutácie.* (*Teda permutácia je párna práve vtedy, keď ju je možné získať zložením párneho počtu transpozícií. Permutácia je nepárna práve vtedy, keď sa dá dostať zložením nepárneho počtu transpozícií.*)

Poznámka 11.5.18. Prvá časť tohoto tvrdenia hovoríme o tom, že každá permutácia sa dá rozložiť na transpozície. Na rozdiel od rozkladu na disjunktné cykly, v tomto prípade už rozklad nemusí byť jednoznačný. Ľahko sa presvedčíte o tom, že napríklad pre permutácie množiny $\{1, 2, 3, 4\}$ máme

$$(123) = (13)(12) = (12)(23) = (12)(23)(13)(13) = (14)(34)(24)(14).$$

Je to príklad permutácie, ktorú môžeme rozložiť na transpozície rôznymi spôsobmi, dokonca aj počet použitých transpozícií môže byť rôzny. (Keďže sme v našom príklade zobrali párnu permutáciu, tak z tvrdenia 11.5.17 vieme, že pri akomkoľvek rozklade na transpozície bude ich počet párny.)

Dôkaz tvrdenia 11.5.17. Na dôkaz prvej časti stačí ukázať, že každý cyklus sa dá rozložiť na transpozície (pretože podľa tvrdenia 11.5.7 sa dá každá permutácia rozložiť na cykly). Jedna z možností, ako rozložiť cyklus dĺžky $n \geq 2$ na transpozície je

$$(a_1a_2 \dots a_n) = (a_1a_n)(a_1a_{n-1}) \dots (a_1a_2).$$

Na dôkaz druhej časti si stačí uvedomiť dve skutočnosti. Prvou z nich je, že transpozícia je nepárna permutácia – pozri príklad 11.5.16 Keď ďalej dokážeme, že pri zložení ľubovoľnej permutácie s nejakou transpozíciou sa zmení parita, dôkaz je hotový.

Majme teda ľubovoľnú permutáciu φ a uvažujme transpozíciu $\tau = (ij)$, pričom $i < j$. Potom

$$\psi = \varphi \circ \tau = \begin{pmatrix} 1 & \dots & i & \dots & j & \dots & n \\ \varphi(1) & \dots & \varphi(i) & \dots & \varphi(j) & \dots & \varphi(n) \end{pmatrix} (ij) = \begin{pmatrix} 1 & \dots & i & \dots & j & \dots & n \\ \varphi(1) & \dots & \varphi(j) & \dots & \varphi(i) & \dots & \varphi(n) \end{pmatrix}.$$

(Predchádzajúcim zápisom sa myslí to, že jediné miesta, na ktorých sa ψ a φ líšia, sú i-te a j-te miesto, t.j. líšia sa len obrazy prvkov i a j. Preto sme vyznačili iba tieto prvky a okrem nich prvý a posledný prvok a ich obrazy.)

Pokúsme sa zistiť, ako sa líši počet inverzií permutácií φ a ψ. Zrejme jediné inverzie, ktoré sme mohli ovplyvniť, sú tie, ktoré obsahovali prvky i a j.

Je zrejmé, že ak $(\varphi(i), \varphi(j))$ je inverzia, t.j. ak $i < j$, tak dvojica $(\psi(i), \psi(j))$ už inverziu netvorí. Takisto obrátene, ak tieto dva prvky netvoria inverziu permutácie φ, dostaneme z nich inverziu v ψ.

Skúsme nájsť ešte aj ostatné inverzie, ktoré mohli „vzniknúť“ alebo „zaniknúť“.

Inverzie, ktoré sa ešte mohli zmeniť môžu byť jedine také, že jeden prvok z dvojice, ktorá je „zanikajúcou“ alebo „vznikajúcou“ inverziou je buď $\varphi(i)$ alebo $\varphi(j)$ a druhý prvok je $\varphi(k)$ pre niektoré $i < k < j$.

Navyše, charakter každej takejto dvojice sa zmení na opačný. T.j. ak $(\varphi(i), \varphi(k))$ je inverzia φ, tak $(\varphi(k), \varphi(i)) = (\psi(k), \psi(j))$ už nie je inverzia. A takisto obrátene, ak táto dvojica pôvodne netvorila inverziu, v permutácii ψ ju už určite tvorí. Takých dvojíc je práve $j - i - 1$.

Podobná úvaha však samozrejme platí aj keď zoberieme j namiesto i.

Celkove sme teda zistili, že pri zložení s transpozíciou (ij) sa zmení počet inverzií o ± 1 pre $2(j - i - 1) + 1$ dvojíc, čo je nepárne číslo. Teda parita permutácie sa musí zmeniť. □

Všimnime si, že v dôkaze predchádzajúceho tvrdenia sme okrem iného ukázali aj to, že cyklus párnej dĺžky je nepárna permutácia a obrátene cyklus nepárnej dĺžky je párna permutácia. (Pretože cyklus dĺžky n sme rozložili na $n + 1$ transpozícií.)

Dôsledok 11.5.19. *Zložením dvoch permutácií rovnakej parity dostaneme párnu permutáciu. Zložením párnej a nepárnej permutácie dostaneme nepárnu permutáciu.*

Príklad 11.5.20. Uvažujeme permutáciu z predchádzajúceho príkladu a pozrime sa na jej rozklad na disjunktné cykly.

$$\varphi = \left(\begin{smallmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 3 & 5 & 1 & 2 \end{smallmatrix}\right) = (14)(235).$$

Je zložená z 2 cyklov, jeden z nich je párnej dĺžky, teda je to nepárna permutácia. Druhý z nich má nepárnu dĺžku, čiže ide o párnu permutáciu. Permutácia φ je teda zložením párnej a nepárnej permutácie, preto φ je nepárna permutácia.

{perm:DOSALTGRUPA}

Dôsledok 11.5.21. *Párne permutácie tvoria podgrupu grupy S_n. Grupu tvorenú párnymi permutáciami množiny $\{1, 2, \ldots, n\}$ budeme označovať A_n a nazývať* alternujúca grupa.

Dôkaz. Zrejme $id \in A_n$, preto $A_n \neq \emptyset$. Podľa predchádzajúceho dôsledku je táto množina uzavretá na skladanie.

Keďže A_n je konečná množina, na základe tvrdenia 11.4.5 už dostávame, že to je podgrupa. □

Hoci sme vďaka tvrdeniu 11.4.5 nepotrebovali v predošlom dôkaze kontrolovať uzavretosť na inverzné prvky, môžeme si uvedomiť, že ak je permutácia vyjadrená pomocou transpozícií, tak veľmi ľahko nájdeme vyjadrenie pre inverznú permutáciu. Najprv si všimnime, že pre ľubovoľnú transpozíciu τ platí $\tau^{-1} = \tau$. Z toho dostávame, že ak $\varphi = \tau_1 \tau_2 \ldots \tau_k$, kde $\tau_1, \ldots, \tau_k$ sú transpozície, tak

$$\varphi^{-1} = \tau_k^{-1} \tau_{k-1}^{-1} \ldots \tau_2^{-1} \tau_1^{-1} = \tau_k \tau_{k-1} \ldots \tau_2 \tau_1,$$

teda jednoducho stačí vymeniť poradie permutácií. (Túto úvahu by sme mohli využiť na dôkaz dôsledku 11.5.21 bez použitia tvrdenia 11.4.5.)

V oblasti zábavnej matematiky sa tiež môžeme stretnúť s podgrupami S_n. Ako grupy permutácií možno chápať povolené transformácie Rubikovej kocky (pozri [KGGS, Kapitola 3.2]) alebo tiež „hra 15“ (pozri [KGGS, Cvičenia 3.6.9*, 3.6.10*]).

Cvičenia

Úloha 11.5.1. V tomto cvičení budeme pracovať s permutáciami množiny $\{1, 2, 3, 4, 5, 6, 7, 8\}$ (čiže prvkami grupy S_8) a budeme zadané permutácie aj výsledky vždy zapisovať ako súčiny disjunktných cyklov: Označme
$\varphi = (14)(235)(78)$
$\psi = (234)(67)$
$\tau = (135)(24)(68)$
a) Vypočítajte $(\varphi \circ \psi) \circ \tau$ a $\varphi \circ (\psi \circ \tau)$

b) Ku každej z uvedených permutácií vypočítajte inverznú permutáciu.
c) Zistite rád a paritu permutácií φ, ψ, τ a aj permutácií, ktoré sme dostali ako výsledky v predchádzajúcich častiach tejto úlohy.

Úloha 11.5.2. Pre dané permutácie určte rád, paritu, a rozklad na disjunktné cykly:
$\varphi = \left(\begin{smallmatrix}1&2&3&4&5\\2&3&4&5&1\end{smallmatrix}\right)$, $\tau = \left(\begin{smallmatrix}1&2&3&4&5\\2&1&4&5&3\end{smallmatrix}\right)$, $\psi = \left(\begin{smallmatrix}1&2&3&4&5\\3&2&1&5&4\end{smallmatrix}\right)$.
Ďalej vypočítajte permutácie $\varphi\tau\psi$, φ^{-1}, τ^{-1}, ψ^{-1}.

Úloha 11.5.3. Vypočítajte $\varphi \circ \psi$ a $\psi \circ \varphi$ pre:
a) $\varphi = (14)(5678)$, $\psi = (23)(5678)$
b) $\varphi = (124)(5678)$, $\psi = (23)(5678)$.
Je niektorý z týchto prípadov príkladom nedisjunktných permutácií, ktoré komutujú?

Úloha 11.5.4. V grupe $(S_7, \circ)$ nájdite rád a paritu permutácie $\varphi = (12)(14)(35)(26)(21)(67)$. Vypočítajte φ^{127}. Nájdite cyklickú podgrupu S_7 generovanú touto permutáciou. S ktorou grupou $(\mathbb{Z}_n, \oplus)$ je táto podgrupa izomorfná?

Úloha 11.5.5. Zostavte tabuľku grupovej operácie pre grupu S_3.

Úloha 11.5.6. Nájdite všetky podgrupy grupy S_3.

Úloha 11.5.7. Ak počet inverzií permutácie $\left(\begin{smallmatrix}1&2&\cdots&n\\a_1&a_2&\cdots&a_n\end{smallmatrix}\right)$ je k, zistite počet inverzií permutácie $\left(\begin{smallmatrix}1&2&\cdots&n\\a_n&a_{n-1}&\cdots&a_1\end{smallmatrix}\right)$.

{permcvic:PERMTRANSP}

Úloha 11.5.8. Dokážte, že grupa S_n je generovaná:
a) Množinou všetkých transpozícií.
b) Množinou transpozícií $\{(12), (13), \ldots, (1n)\}$.
c) Množinou transpozícií $\{(12), (23), \ldots, (n-1, n)\}$.
d) Transpozíciou (12) a cyklom $(12\ldots n)$. (Hint: Skúste vyrátať $(12\ldots n)^{-k}(12)(12\ldots n)^k$.)
Dokážte, že S_4 nie je generovaná pomocou (13) a (1234). (Teda S_4 síce je generovaná cyklom dĺžky 4 a transpozíciou, túto transpozíciu však nemôžem vybrať ľubovoľne.)

{permcvic:ALTTROJCYK}

Úloha 11.5.9. Dokážte, že alternujúca grupa A_n je generovaná:
a) Množinou všetkých cyklov (ijk) dĺžky 3.
b) Množinou cyklov dĺžky 3 tvaru $(123), (124), \ldots, (12n)$.

Úloha 11.5.10. Popíšte permutácie z grupy S_n, ktoré majú rád 2. Aký je ich počet? (Stačí nájsť rekurzívny predpis pre počet takých prvkov.)

{NAJVRAD}

Úloha 11.5.11. Čomu sa rovná najväčší možný rád prvku grupy S_{12}.

11.6 Cayleyho veta*

Ľahko vieme overiť že pre danú množinu M všetky bijekcie z M do M tvoria s operáciou skladania zobrazení grupu. (Dôkaz je presne rovnaký ako pre permutácie v prípade, že M bola konečná.) Túto grupu budeme označovať $(S(M), \circ)$ alebo stručnejšie $S(M)$.

Definícia 11.6.1. Pod *grupou transformácií množiny* M budeme rozumieť ľubovoľnú podgrupu grupy $(S(M), \circ)$.

Ekvivalentne by sme mohli grupu transformácií definovať tak, že je to množina bijekcií z M do M uzavretá na skladanie a inverzné zobrazenie. (V [KGGS] sa grupa transformácií definuje takto, pretože tento pojem je tu zavedený skôr než pojem grupy. Aj historické poradie, v akom matematici definovali tieto pojmy, je rovnaké.)

V tejto časti ukážeme Cayleyho vetu, ktorá hovorí, že každá grupa je izomorfná s nejakou grupou transformácií. To znamená, že keby sme sa zaoberali len grupami, v ktorých je binárnou operáciou skladanie zobrazení, v podstate by sme nič nestratili.

Definícia 11.6.2. Nech $(G, *)$ je grupa a $a \in G$.
Zobrazenie $f_a \colon G \to G$ dané predpisom $f_a(x) = a * x$ voláme *ľavá translácia.*
Zobrazenie $g_a \colon G \to G$ dané predpisom $g_a(x) = x * a$ voláme *pravá translácia.*

Lema 11.6.3. *Každá ľavá (pravá) translácia je bijekcia.*

Dôkaz. Surjektívnosť f_a: Pre každé $x \in G$ máme $f(a^{-1} * x) = a * a^{-1} * x = x$.
Injektívnosť f_a: Ak $f_a(x) = f_a(y)$, znamená to, že $a * x = a * y$. Zo zákona o krátení potom vyplýva $x = y$.
Ukázali sme, že ľavé translácie sú bijekcie; dôkaz pre pravé translácie je úplne analogický. □

Dôsledok 11.6.4. *Pre všetky $a \in G$ platí $f_a, g_a \in S(G)$.*

Lema 11.6.5. *Pre ľavé translácie platí*

$$f_b \circ f_a = f_{b*a}$$

Dôkaz. Pre $x \in G$ máme

$$(f_b \circ f_a)(x) = f_b(f_a(x)) = b * (a * x) = (b * a) * x = f_{b*a}(x).$$

□

Dôsledok 11.6.6. *Zobrazenie $a \mapsto f_a$ je homomorfizmus z G do $S(G)$.*

Dôsledok 11.6.7. *Ľavé translácie tvoria podgrupu grupy $S(G)$.*

Dôkaz. Ľavé translácie sú presne obrazom G v zobrazení $a \mapsto f_a$ z G do $S(G)$. Pretože toto zobrazenie je homomorfizmus, podľa tvrdenia 11.3.9 tvoria ľavé translácie podgrupu. □

Aby sme dokázali Cayleyho vetu, stačí už len ukázať, že práve uvedený homomorfizmus je v skutočnosti izomorfizmus na svoj obraz, t.j. že je injektívny.

{cayl:VTCAYL}

Veta 11.6.8 (Cayley)**.** *Každá grupa $(G, *)$ je izomorfná s nejakou grupou transformácií. Presnejšie, $(G, *)$ je izomorfná s podgrupou grupy $S(G)$ tvorenou všetkými ľavými transláciami.*

Dôkaz. Z toho, čo sme dokázali doteraz už vieme, že ľavé translácie sú skutočne podgrupou $S(G)$ a zobrazenie $a \mapsto f_a$ je surjektívny homomorfizmus z G na túto podgrupu. (Surjektívnosť vyplýva z toho, že táto podgrupa je priamo obrazom grupy G v uvedenom zobrazení.)
Zostáva teda dokázať len injektívnosť. Ak platí $f_a = f_b$ (v zmysle rovnosti zobrazení), tak potom aj $a = a * e = f_a(e) = f_b(e) = b * e = b$. □

Dôsledok 11.6.9. *Ľubovoľná konečná grupa rádu n (t.j. taká, ktorá má n prvkov) je izomorfná s podgrupou grupy permutácií S_n.*

Príklad 11.6.10. Ilustrujme si Cayleyho vetu na príklade grupy $(\mathbb{Z}, +)$. (Keďže ide o komutatívnu grupu, v tomto prípade sú ľavé a pravé translácie totožné.)
V tomto prípade pre $a \in \mathbb{Z}$ máme zobrazenie $f_a \colon \mathbb{Z} \to \mathbb{Z}$

$$f_a(x) = a + x,$$

ktoré je očividne bijektívne, čiže $f_a \in S(\mathbb{Z})$. Takisto sa ľahko overí, že $f_{a+b} = f_a \circ f_b$, z čoho vidíme, že $a \mapsto f_a$ je homomorfizmus. Fakt, že tento homomorfizmus je injektívny, môžeme overiť podobne ako v dôkaze Cayleyho vety.

{cayl:PRSROT}

Príklad 11.6.11. Skúsme sa pozrieť na reprezentáciu grupy $(S, \cdot)$, kde $S = \{z \in \mathbb{C}; |z| = 1\}$ pomocou Cayleyho vety.

V tomto prípade máme $f_z(x) = zx$. Z Moivrovej vety vieme, že vynásobenie komplexných číslom z s jednotkovou veľkosťou presne zodpovedá otočeniu bodu v komplexnej rovine okolo počiatku o uhol φ taký, že $z = \cos\varphi + i\sin\varphi$. Čiže v tomto prípade tvoria grupu transformácií z Cayleyho vety všetky otočenia kružnice okolo nuly.

Poznámka 11.6.12. Ďalšou zaujímavou oblasťou, ktorá súvisí s grupami transformácií sú grupy symetrií rovinných útvarov alebo telies. V prípade, že by Vás táto téma zaujímala, môžete si o nej niečo prečítať napríklad v [KGGS, Kapitola 3.1].

Cvičenia

{caylcvic:ULOVAUT}

Úloha 11.6.1.

Izomorfizmus grupy G na samú seba voláme *automorfizmus.* Dokážte, že:
a) Množina $\operatorname{Aut} G$ všetkých automorfizmov grupy G je grupou transformácií.
b) Definujeme pre ľubovoľné $a \in G$ zobrazenie $f_a\colon G \to G$ ako $f_a(g) = aga^{-1}$. Potom platí $f_{ab} = f_a \circ f_b$ pre ľubovoľné $a, b \in G$.
c) Pre každé $a \in G$ je zobrazenie f_a automorfizmus grupy G (takýto automorfizmus nazveme vnútorný).
d) Množina $\operatorname{VAut} G$ všetkých vnútorných automorfizmov grupy G je grupou transformácií.
e) Grupa G je komutatívna práve vtedy, keď množina $\operatorname{VAut} G$ všetkých jej vnútorných automorfizmov je jednoprvková.
f) Zobrazenie $a \mapsto f_a$ je surjektívny homomorfizmus z G na $\operatorname{VAut} G$. Nájdite jadro tohoto automorfizmu.

Kapitola 12

Faktorizácia

12.1 Relácie ekvivalencie a rozklady

So základnými vlastnosťami ekvivalencií a rozkladov a s ich vzájomným súvisom ste sa už stretli na iných prednáškach (pozri [KGGS, časť 1.4], [OŠ]), napriek tomu tu však zopakujeme niektoré ich základné vlastnosti.

Definícia 12.1.1. *Relácia ekvivalencie* je relácia R na množine A, ktorá je reflexívna, symetrická a tranzitívna; t.j. pre všetky $a, b, c \in A$ platí:

$$aRa$$
$$aRb \Rightarrow bRa$$
$$aRb \wedge bRc \Rightarrow aRc$$

Množina $\{b \in A; aRb\}$ sa nazýva *triedou ekvivalencie s reprezentantom* a a označuje sa $[a]_R$, prípadne len $[a]$.

Definícia 12.1.2. *Rozklad množiny* A je taká množina $\mathcal{A} = \{A_i; i \in I\}$ neprázdnych podmnožín množiny A, že platí:

(i) Pre všetky $i, j \in I$ platí buď $A_i = A_j$ alebo $A_i \cap A_j = \emptyset$.

(ii) $\bigcup_{i \in I} A_i = A$.

Pred hlavnými výsledkami týkajúcimi sa rozkladov a ekvivalencií uvedieme si ešte jednu lemu:

{ekv:LMTRIEDY}

Lema 12.1.3. *Nech R je relácia ekvivalencie. Potom*

$$aRb \Leftrightarrow [a]_R = [b]_R.$$

Veta 12.1.4. *Ak R je relácia ekvivalencie na A, tak množina všetkých tried ekvivalencie tvorí rozklad množiny A.*

Veta 12.1.5. *Ak $\mathcal{A} = \{A_i; i \in I\}$ je rozklad množiny A, tak relácia R definovaná tak, že*

$$aRb \qquad \Leftrightarrow \qquad (\exists i \in I) a \in A_i \wedge b \in A_i$$

je relácia ekvivalencie. (*Definícia relácie R vlastne hovorí, že dva prvky sú v relácii R práve vtedy, keď ležia v tej istej množine rozkladu $\mathcal{A}$.*)

Dôkaz týchto viet môžete nájsť napríklad v [OŠ].

Videli sme, že relácii ekvivalencie na množine A môžeme priradiť rozklad množiny A a opačne. Chceli by sme ukázať, že táto korešpondencia medzi reláciami ekvivalencie a rozkladmi je jednojednoznačná; čiže relácie ekvivalencie a rozklady sú vlastne len 2 rôzne pohľady na tú istú vec.

Označme rozklad prislúchajúci relácii ekvivalencie R ako $\mathcal{A}_R$ a reláciu ekvivalencie danú rozkladom $\mathcal{A}$ ako $R_{\mathcal{A}}$. My vlastne chceme ukázať, že tieto 2 priradenia sú navzájom inverzné, čiže $R_{\mathcal{A}_R} = R$ a $\mathcal{A}_{R_{\mathcal{A}}} = \mathcal{A}$.

(Tu je tiež dôležité si uvedomiť, čo znamená že 2 relácie resp. 2 rozklady sú rovnaké. Relácie chápeme ako podmnožiny $A \times A$, 2 relácie sa R a R' sa rovnajú práve vtedy, keď platí $aRb \Leftrightarrow aR'b$ pre všetky $a, b \in A$. Rovnosť pre rozklady takisto chápeme ako rovnosť množín – to znamená, že rovnaké rozklady pozostávajú z tých istých podmnožín.)

Z lemy 12.1.3 vidíme, že ak priradíme relácii ekvivalencie rozklad, tak v rovnakých podmnožinách budú práve tie prvky, ktoré sú v relácii R, a teda skutočne platí $R_{\mathcal{A}_R} = R$. Platnosť rovnosti $\mathcal{A}_{R_{\mathcal{A}}} = \mathcal{A}$ pre ľubovoľný rozklad sa tiež ukáže pomerne jednoducho (úloha 12.1.2 – dá sa opäť použiť lema 12.1.3).

Cvičenia

Úloha 12.1.1. Dokážte lemu 12.1.3.

{ekv:ULOBIJEK}

Úloha 12.1.2. Dokážte, že pre ľubovoľný rozklad $\mathcal{A}$ platí $\mathcal{A}_{R_{\mathcal{A}}} = \mathcal{A}$.

{ekvcvic:ULOEKVZOBR}

Úloha 12.1.3. a) Nech $f\colon A \to B$ je surjektívne zobrazenie. Dokážte, že relácia R na množine A určená predpisom $aRa' \Leftrightarrow f(a) = f(a')$ je relácia ekvivalencie a triedy rozkladu sú množiny $f^{-1}(\{b\}) = f^{-1}(b)$ pre $b \in B$.
b) Nech R je relácia ekvivalencie na množine A a nech B je množina všetkých tried ekvivalencie. Dokážte, že zobrazenie $f\colon A \to B$, ktoré každému prvku priradí jeho triedu ekvivalencie (teda $f\colon a \mapsto [a]$) je surjektívne.
c) V predchádzajúcej časti sme každému surjektívnemu zobrazeniu priradili reláciu ekvivalencie a obrátene. Sú tieto dve priradenia sú navzájom inverzné?

Úloha 12.1.4. Dokážte, že relácia $R_1 = A \times A$ na A je relácia ekvivalencie. Dokážte, že $R_2 = \{(a, a); a \in A\}$ je relácia ekvivalencie na A.

Úloha 12.1.5. Ak R_1, R_2 sú relácie ekvivalencie na A, tak aj $R_1 \cap R_2$ je relácia ekvivalencie na A.

{ekvcvic:ULOPRIENIK}

Úloha 12.1.6. Nech pre každé $i \in I$ je R_i relácie na množine A a nech $R = \bigcap_{i \in I} R_i$. Dokážte, že:
a) Ak všetky relácie R_i sú reflexívne, tak aj ich prienik R je reflexívna relácia.
b) Ak všetky relácie R_i sú symetrické, tak aj ich prienik R je symetrická relácia.
c) Ak všetky relácie R_i sú tranzitívne, tak aj ich prienik R je tranzitívna relácia.
d) Ak všetky relácie R_i sú relácie ekvivalencie, tak aj ich prienik R je relácia ekvivalencie.

{ekvcvic:ULOOBAL}

Úloha 12.1.7. Nech S je relácia na množine A. Ukážte, že existuje relácia R na množine A, ktorá je najmenšia vzhľadom na inklúziu medzi reláciami s vlastnosťami:
a) $S \subseteq R$ a R je reflexívna.
b) $S \subseteq R$ a R je symetrická.
a) $S \subseteq R$ a R je tranzitívna.
d) $S \subseteq R$ a R je relácia ekvivalencie.

Úloha 12.1.8. Dokážte, že nasledujúce relácie sú relácie ekvivalencie:
a) Relácia R na množine S_n taká, že $\varphi R\tau$ práve vtedy, keď permutácie φ a τ majú rovnaký počet inverzií;
b) relácia R na množine $\mathbb{Z}$ definovaná ako $xRy \Leftrightarrow x$ a y majú rovnaký ciferný súčet;
c) relácia R na množine $\mathbb{Z}$ definovaná predpisom $xRy \Leftrightarrow x + y$ je párne;
d) pre ľubovoľnú (konečnú) množinu M relácia R na množine $\mathcal{P}(M)$ definovaná ako $ARB \Leftrightarrow |A| = |B|$. (Pod označením $|A|$ rozumieme buď počet prvkov – v prípade, že M je konečná, alebo mohutnosť množiny.)

{ekvcvic:KONGR}

Úloha 12.1.9. Dokážte, že nasledujúce relácie sú relácie ekvivalencie:
a) relácia R na množine $\mathbb{R}$ definovaná ako $(x, y, z)R(x', y', z') \Leftrightarrow x + y + z = x' + y' + z'$,
b) relácia R na množine S_n definovaná ako $\varphi R\tau \Leftrightarrow \varphi$ a τ majú rovnakú paritu,
c) relácia R na množine $\mathbb{Z}$ definovaná ako $xRy \Leftrightarrow 5 \mid x - y$
d) pre ľubovoľnú maticu typu $m \times n$ nad poľom F je relácia $\vec{\alpha}R\vec{\beta} \Leftrightarrow A\vec{\alpha}^T = A\vec{\beta}^T$ relácia ekvivalencie na množine F^n;
e) relácia R na množine $\mathbb{R}$ definovaná tak, že $xRy \Leftrightarrow x = y \vee x = -y$;
f) relácia R na množine $\mathbb{R}^{\mathbb{R}}$ všetkých zobrazení z $\mathbb{R}$ do $\mathbb{R}$ definovaná ako $fRg \Leftrightarrow f(0) = g(0)$.

Úloha 12.1.10. Nech R je relácia ekvivalencie na množine X a S je relácia ekvivalencie na množine Y. Potom relácia T určená ako $(x, y)T(x', y') \Leftrightarrow xRx' \wedge ySy'$ je relácia ekvivalencie na množine $X \times Y$.

Úloha 12.1.11. Nech $f\colon X \to Y$ je surjektívne zobrazenie.
a) Ak R je relácia ekvivalencie na X, tak relácia S na Y daná predpisom $ySy' \Leftrightarrow$ existujú $x, x' \in X$ také, že $f(x) = y$, $f(x') = y'$ a xRx' je tiež relácia ekvivalencie.
b) Ak S je relácia ekvivalencie na množine Y tak relácia R na X daná predpisom $xRx' \Leftrightarrow f(x)Sf(x')$ je tiež relácia ekvivalencie. (Ako špeciálny prípad dostaneme tvrdenie z úlohy 12.1.3b).)

12.2 Rozklad grupy podľa podgrupy

Najprv si zadefinujeme jeden pomocný pojem – násobenie podmnožín grupy.

Definícia 12.2.1. Nech G je grupa a $A, B \subseteq G$ sú jej ľubovoľné podmnožiny. Potom definujeme *súčin AB podmnožín A, B* ako

$$AB = \{ab; a \in A, b \in B\}.$$

V prípade, že jedna z množín je jednoprvková, budeme používať stručnejší zápis aB namiesto $\{a\}B$ a Ab namiesto $A\{b\}$.

Niektoré užitočné vlastnosti násobenia podmnožín zhrnieme v nasledujúcej leme:

{rozkl:LMKOMPL}

Lema 12.2.2. *Nech G je grupa.*

(i) *Násobenie podmnožín je asociatívne, t.j. $A(BC) = (AB)C$ pre ľubovoľné podmnožiny $A, B, C \subseteq G$.*

(ii) *Pre ľubovoľnú podmnožinu $A \subset G$ platí $eA = Ae = A$.*

(iii) *Ak $B \subseteq C$, tak $AB \subseteq AC$ a $BA \subseteq CA$.*

(iv) *Ak H je podgrupa grupy G a $h \in H$, tak $hH = H$.*

(v) *Ak H je podgrupa grupy G, tak $H^2 = H \cdot H = H$.*

ozkl:lmitem5}

(vi) *Pre ľubovoľnú podmnožinu $A \subseteq G$ platí $(A^{-1})^{-1} = A$, kde používame označenie $A^{-1} = \{a^{-1}; a \in A\}$.*

(vii) *Ak H je podgrupa grupy G, tak $H^{-1} = \{h^{-1}; h \in H\} = H$.*

{rozkl:lmitem6}

(viii) *Pre ľubovoľné podmnožiny $A, B \subseteq G$ platí $(AB)^{-1} = B^{-1} \cdot A^{-1}$.*

{rozkl:lmitem4}

(ix) *Ak K, H sú podgrupy grupy G, tak $(HK)^{-1} = K^{-1} \cdot H^{-1} = KH$.*

Označenie H^{-1} v predchádzajúcej leme neznamená, že by táto množina bola inverzným prvkom ku H v $\mathcal{P}(G) \setminus \{\emptyset\}$ s operáciou násobenia podmnožín – H^{-1} jednoducho len označuje množinu inverzných prvkov ku prvkom z H.

Nebudeme dokazovať všetky časti tejto lemy – väčšinu z nich ponecháme ako cvičenie (úloha 12.2.1). Na ukážku si dokážme (vi).

Dôkaz. (vi): Pretože $(a^{-1})^{-1} = a$, každý prvok z A patrí aj do $(A^{-1})^{-1}$, čiže $A \subseteq (A^{-1})^{-1}$.

Obrátene, ak $b \in (A^{-1})^{-1}$, tak $b = (a^{-1})^{-1}$ pre nejaké $a \in A$, ale $(a^{-1})^{-1} = a$, teda $b = a \in A$. Ukázali sme aj inklúziu $(A^{-1})^{-1} \subseteq A$. □

Násobenie podmnožín vo všeobecnosti nemusí byť komutatívne (ako kontrapríklad stačí zobrať v nekomutatívnej grupe 2 jednoprvkové množiny $\{a\}$ a $\{b\}$ pre prvky a a b, ktoré nekomutujú). Samozrejme, pre komutatívnu grupu je aj násobenie podmnožín komutatívne.

Predchádzajúce tvrdenie hovorí o tom, že takto definovaný súčin množín je asociatívny a má aj neutrálny prvok. Nie pre každú množinu však existuje inverzný prvok. Špeciálne nemusí platiť $AA^{-1} = \{e\}$ (úloha 12.2.2).

Definícia 12.2.3. Ak H je podgrupa grupy G, tak označíme pre $a \in G$

$$aH = \{ah; h \in H\},$$
$$Ha = \{ha; h \in H\}.$$

Množiny aH nazývame *ľavé triedy grupy G podľa H* (alebo ľavé triedy grupy G modulo H), množiny Ha sú *pravé triedy grupy G podľa H*.

Ako sme už spomenuli, násobenie podmnožín vo všeobecnosti nemusí byť komutatívne, takisto ani nemusí vo všeobecnosti platiť $aH = Ha$. V ďalšej časti uvidíme, že podgrupy, ktoré majú túto vlastnosť sú z istého hľadiska zaujímavé. Je zrejmé, že táto rovnosť platí ak G je komutatívna.

Začali by sme však s tým, že ukážeme, že ľavé triedy G podľa H tvoria rozklad G (a podobne to platí pre pravé triedy).

{rozkl:PRZ3}

Príklad 12.2.4. Triedy $\mathbb{Z}$ podľa $3\mathbb{Z}$ sú $\{3k; k \in \mathbb{Z}\}$, $\{3k+1; k \in \mathbb{Z}\}$ a $\{3k+2; k \in \mathbb{Z}\}$. Je zrejmé, že tvoria rozklad množiny $\mathbb{Z}$ – každé celé číslo je buď tvaru $3k$, $3k+1$ alebo $3k+2$. (V tomto prípade ide o komutatívnu grupu, preto sú ľavé a pravé triedy totožné.)

{rozkl:LMAINVB}

Lema 12.2.5. *Nech H je podgrupa G a $a, b \in G$. Potom $aH = bH$ práve vtedy, keď $b^{-1}a \in H$. (Ďalšou ekvivalentnou podmienkou je $a^{-1}b \in H$).*

Podobne platí $Ha = Hb \Leftrightarrow ab^{-1} \in H \Leftrightarrow ba^{-1} \in H$.

Dôkaz. $\boxed{\Rightarrow}$ Ak $aH = bH$, tak $a \in bH$, čiže $a = bh$ pre nejaké $h \in H$. Z toho $b^{-1}a = h \in H$.
$\boxed{\Leftarrow}$ Ak $b^{-1}a \in H$, tak $b^{-1}aH = H$, a teda $bH = b(b^{-1}aH) = (bb^{-1})aH = eaH = aH$.
Pretože $(b^{-1}a)^{-1} = a^{-1}(b^{-1})^{-1} = a^{-1}b$, aj tretia uvedená podmienka je ekvivalentná.
Dôkaz druhej časti tejto lemy je analogický. □

Tvrdenie 12.2.6. *Ľavé triedy grupy G podľa jej podgrupy H tvoria rozklad G.* (*Inak:* $\{aH; a \in G\}$ *je rozklad množiny G.*)
Pravé triedy grupy G podľa jej podgrupy H tvoria rozklad G.

Dôkaz. Každá trieda aH obsahuje prvok a, je teda neprázdna. Overme teda ešte ostatné dve podmienky z definície rozkladu.
Platí $\bigcup_{a\in G} aH \supseteq \bigcup_{a\in G}\{a\} = G$, teda zjednotenie všetkých ľavých tried je celé G.
Nech $a, b \in G$. Stačí ukázať, že ak $aH \cap bH \neq \emptyset$, tak $aH = bH$. Nech teda $x \in aH \cap bH$. To znamená, že $x = ah = bh'$ pre nejaké $h, h' \in H$. Z rovnosti $ah = bh'$ ľahko dostaneme $b^{-1}a = h'h^{-1}$, a teda $b^{-1}a \in H$. Podľa lemy 12.2.5 potom platí $aH = bH$
Dôkaz pre pravé triedy je skoro identický. □

Napriek tomu, že tu budeme tak trochu opakovať podobné úvahy ako v dôkaze lemy 12.2.5, možno sa oplatí urobiť aj iný dôkaz. Z neho navyše hneď vidno aj to, že ľavé triedy sú vlastne triedy rozkladu zodpovedajúcemu istej relácii ekvivalencie.

Dôkaz. Pomerne ľahko sa ukáže, že $a \sim b \Leftrightarrow a^{-1}b \in H$ je relácia ekvivalencie (úloha 12.2.3). Vieme, že relácia ekvivalencie nám dáva rozklad, teda nám stačí ukázať, že triedy rozkladu pre túto reláciu ekvivalencie sú presne ľavé triedy G podľa H.
Chceme teda zdôvodniť, že $[a] = aH$. Platí:

$$\begin{aligned}[a] &= \{b \in G; a \sim b\} \\ &= \{b \in G; a^{-1}b \in H\} \overset{(*)}{=} aH.\end{aligned}$$

Azda jediná rovnosť, ktorá potrebuje trochu detailnejšie zdôvodnenie je rovnosť označená $(*)$.
$\boxed{\subseteq}$ Ak $a^{-1}b \in H$, tak z rovnosti $b = a(a^{-1}b)$ máme $b \in aH$.
$\boxed{\supseteq}$ Ak $b \in aH$, tak $b = ah$ pre nejaké $h \in H$. Z rovnosti $b = ah$ dostaneme $h = a^{-1}b$, čiže vidíme, že $a^{-1}b \in H$. □

Definícia 12.2.7. Nech G je grupa a H je podgrupa. Rozklad $\{aH; a \in G\}$ sa nazýva *ľavý rozklad G podľa H* a rozklad $\{Ha; a \in G\}$ sa nazýva *pravý rozklad G podľa H.*

Všimnime si, že $eH = He = H$, teda ako jedna z ľavých (pravých) tried sa vždy vyskytne podgrupa H.

{rozkl:PRRxR}

Príklad 12.2.8. Uvažujme podgrupu $H = \{(x, x); x \in \mathbb{R}\}$ grupy $G = (\mathbb{R} \times \mathbb{R}, +)$. Ide o komutatívnu grupu, takže ľavý aj pravý rozklad sú rovnaké. Dva prvky $(a, b), (c, d) \in G$ ležia v tej istej triede rozkladu práve vtedy, keď $(a, b) - (c, d) = (a - c, b - d) \in H$, teda keď $a - c = b - d$, čo je ekvivalentné s

$$a - b = c - d.$$

To znamená, že každá trieda rozkladu je určená rozdielom $r = a-b$. Inak povedané, jednotlivé triedy rozkladu sú práve množiny

$$\{(x, y) \in \mathbb{R}^2; x - y = r\}$$

pre $r \in \mathbb{R}$. Všimnime si, že rôznym r zodpovedajú rôzne triedy ekvivalencie, takže každú triedu ekvivalencie dostaneme takýmto spôsobom iba raz. Z každej triedy rozkladu by sme mohli vybrať napríklad reprezentanta tvaru $\{(r,0); r \in \mathbb{R}\}$, t.j. triedy rozkladu môžeme zapísať ako

$$(r,0) + H$$

pre $r \in \mathbb{R}$.

Uveďme ešte aspoň jeden príklad, kde G je konečná grupa.

{rozkl:PRZ8}

Príklad 12.2.9. Nech $G = (\mathbb{Z}_8, \oplus)$ a $H = 4\mathbb{Z}_8 = \{0,4\}$. Opäť sú obe grupy komutatívne, takže ľavý a pravý rozklad sú totožné. Rozklad G podľa H obsahuje 4 triedy: $\{0,4\}$, $\{1,5\}$, $\{2,6\}$, $\{3,7\}$.

Teraz ukážeme nejaké tvrdenia hovoriace o počtoch (mohutnostiach) tried rozkladu G podľa H. (Môžete si ich podrobnejšie rozmyslieť na rozkladoch z príkladov 12.2.4, 12.2.8, 12.2.9.)

Lema 12.2.10. *Nech H je podgrupa grupy G a $a \in G$. Potom zobrazenie $\varphi\colon H \to aH$ definované ako*

$$\varphi\colon h \mapsto ah$$

je bijekcia.

Podobne zobrazenie $\psi\colon H \to Ha$, $\psi\colon h \mapsto ha$ je bijekcia.

Dôkaz. Zobrazenie φ je injekcia: $\varphi(h_1) = \varphi(h_2) \Rightarrow ah_1 = ah_2 \Rightarrow h_1 = h_2$ (podľa zákonov o krátení).

Priamo z definície množiny aH vyplýva, že φ je aj surjekcia.

Dôkaz pre zobrazenie ψ by bol analogický. □

Z predchádzajúcej lemy okamžite dostávame, že:

Veta 12.2.11. *Nech H je konečná podgrupa G. Potom počet prvkov každej ľavej triedy aH je rovnaký (a rovná sa počtu prvkov podgrupy H). Takisto sa rovná počtu prvkov ľubovoľnej pravej triedy Hb.*

Lema 12.2.12. *Nech H je podgrupa grupy G. Potom zobrazenie*

$$\varphi\colon aH \mapsto Ha^{-1}$$

je bijekcia medzi množinami tried $\{aH; a \in G\}$ a $\{Ha; a \in G\}$.

Dôkaz. Všimnime si, že platí $(aH)^{-1} = H^{-1}a^{-1} = Ha^{-1}$.

Z toho vyplýva, že zobrazenie φ je dobre definované. (Nezávisí od výberu reprezentanta triedy aH, keďže sme ukázali, že $\varphi(aH)$ je práve množina všetkých inverzných prvkov k prvkom z aH.)

Zobrazenie φ je zrejme surjekcia (vzorom pre triedu Hb je ľavá trieda $b^{-1}H$).

Na základe rovnosti $Ha^{-1} = (aH)^{-1}$ ďalej dostaneme $Ha^{-1} = Hb^{-1} \Leftrightarrow (aH)^{-1} = (bH)^{-1} \Leftrightarrow aH = bH$. (Využili sme lemu 12.2.2(vi,viii).) Vidíme teda, že φ je injekcia. □

Na základe predchádzajúcej vety, ktorá hovorí, že počet ľavých a pravých tried je rovnaký, má zmysel nasledujúca definícia.

Definícia 12.2.13. Nech H je podgrupa konečnej grupy. Potom $[G\colon H]$ je počet všetkých ľavých (pravých) tried rozkladu G podľa H. Toto číslo nazývame *indexom grupy G podľa H.*

{rozkl:VTLAG}

Veta 12.2.14 (Lagrangeova veta)**.** *Ak G je konečná grupa a H je jej podgrupa, tak platí*

$$|G| = |H|.[G\colon H].$$

Teda počet prvkov podgrupy H delí počet prvkov G.

Dôkaz. Máme rozklad množiny G na $[G\colon H]$ tried rovnakej veľkosti $|H|$. Potom $|G| = [G\colon H]|H|$.

Z toho je zrejmé aj to, že $|H| \mid |G|$ (počet prvkov H delí počet prvkov G). □

Na tomto mieste treba spomenúť, že neplatí obrátenie Lagrangeovej vety v tom zmysle, že pre každý deliteľ k čísla $|G|$ (počtu prvkov grupy G) by musela existovať k-prvková podgrupa. Pozri úlohu 12.3.3. (Pre cyklické grupy však toto tvrdenie platí, tam dokonca existuje jediná k-prvková podgrupa. Takéto tvrdenie – že by počtom prvkov bola podgrupa jednoznačne určená – takisto vo všeobecnosti neplatí.)

Nasledujúci výsledok by snáď mohol vysvetlovať, prečo namiesto počtu prvkov konečnej grupy niekedy používame aj termín *rád grupy.*

{rozkl:DOS1}

Dôsledok 12.2.15. *Ak G je konečná grupa, tak rád každého prvku delí rád grupy G (počet prvkov grupy G).*

Dôkaz. Stačí si uvedomiť, že rád prvku a je počet prvkov podgrupy $[a]$. □

{rozkl:DOSPRV}

Dôsledok 12.2.16. *Ak G je p-prvková grupa a p je prvočíslo, tak každý jej prvok okrem neutrálneho prvku je generátorom G (a teda G je cyklická).*

Dôkaz. Rád prvku $a \neq e$ nie je 1 a keďže je deliteľ prvočísla p, musí byť rovný p. Teda $[a]$ obsahuje p rôznych prvkov $e, a^1, a^2, \ldots, a^{p-1}$, čiže $[a] = G$. □

Dôsledok 12.2.17. *Každá 4-prvková grupa je izomorfná buď so $\mathbb{Z}_4$ alebo so $\mathbb{Z}_2 \times \mathbb{Z}_2$.*

Dôkaz. Nech G je 4-prvková grupa. Podľa dôsledku 12.2.15 rády jej prvkov môžu byť jedine 1, 2 alebo 4. Ak G obsahuje prvok rádu 4, tak tento prvok je jej generátor. V tomto prípade dostávame, že G je cyklická a $G \cong \mathbb{Z}_4$.

Druhá možnosť je, že všetky prvky s výnimkou neutrálneho majú rád 2, čiže pre každý prvok platí $a^2 = e$, kde e je neutrálny prvok G. Inak povedané, pre všetky $a \in G$ platí $a = a^{-1}$. Z toho dostávame aj to, že G je komutatívna: $xy = (xy)^{-1} = y^{-1}x^{-1} = yx$.

Označme prvky tejto grupy e, a, b, c. Zatiaľ o nich vieme toto:

	e	a	b	c
e	e	a	b	c
a	a	e		
b	b		e	
c	c			e

Podľa zákonov o krátení sa každý prvok vyskytne v ľubovoľnom riadku a v ľubovoľnom stĺpci tabuľky grupovej operácie práve raz. Tento fakt nám umožní jednoznačne doplniť prázdne miesta v tabuľke. Všimnime si napríklad, že prvok ab nemôže byť a, e ani b (inak by sme mali v niektorom riadku alebo stĺpci tento prvok dvakrát). Podobnú úvahu môžeme urobiť pre prvok ba. Dostávame:

	e	a	b	c
e	e	a	b	c
a	a	e	c	
b	b	c	e	
c	c			e

Teraz už v každom riadku a stĺpci máme jediné voľné miesto, teda zostávajúci prvok je jednoznačne určený

	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e

Pretože aj $\mathbb{Z}_2 \times \mathbb{Z}_2$ má tú vlastnosť, že všetky prvky okrem neutrálneho majú rád 2, a práve sme ukázali, že touto podmienkou je grupa jednoznačne určená (až na označenie prvkov – čiže až na izomorfizmus), máme $G \cong \mathbb{Z}_2 \times \mathbb{Z}_2$. □

Cvičenia

{rozklcvic:LMKOMPL}

Úloha 12.2.1. Dokážte zvyšné časti lemy 12.2.2.

{rozklcvic:KOMPLINV}

Úloha 12.2.2. Ukážte na konkrétnom príklad, že v grupe nemusí platiť $AA^{-1} = \{e\}$.
Platí vo všeobecnosti $AA^{-1} = A^{-1}A$?
Tvorí množina $\mathcal{P}(G)$ s operáciou násobenia množín grupu?

{rozklcvic:ULORELEKV}

Úloha 12.2.3. Ukážte, že ak H je podgrupa grupy G, tak relácia určená podmienkou

$$a \sim b \qquad \Leftrightarrow \qquad a^{-1}b \in H$$

je relácia ekvivalencie.

Úloha 12.2.4. Ak G je konečná grupa, H je podgrupa G a K je podgrupa H, tak $[G : K] = [G : H][H : K]$.

Úloha 12.2.5. Dokážte, že každá grupa, ktorá má menej ako 6 prvkov, je komutatívna.

Úloha 12.2.6. Nájdite všetky ľavé (pravé) triedy grupy G podľa podgrupy H, ak
a) $G = (\mathbb{R}, +)$, $H = \mathbb{Z}$;
b) $G = (\mathbb{R} \times \mathbb{R}, +)$, $H = \{(x, y) \in \mathbb{R} \times \mathbb{R}; y = 3x\}$;
c) $G = (\mathbb{Z}_6, \oplus)$, $H = 2\mathbb{Z}_3$;
d) $G = S_n$, $H = A_n$;
e) $G = S_3$, $H = [(12)]$;
f) $G = (\mathbb{Z}, +)$, $H = 3\mathbb{Z}$.
(Pod „nájdite všetky triedy“ sa rozumie to, že pre každú triedu vyberieme jedného reprezentanta, v prípade, že ide o konečné množiny ich môžeme aj vypísať.)

{rozklcvic:PNA8}

Úloha 12.2.7. Dokážte, že každá 8-prvková grupa obsahuje dvojprvkovú podgrupu.

Úloha 12.2.8. Ak G má p^2 prvkov, kde p je prvočíslo, tak každá vlastná podgrupa G je cyklická.

{rozklcvic:PNAK}

Úloha 12.2.9. Ak p je prvočíslo a $k \geq 1$ prirodzené číslo, tak každá p^k-prvková grupa má p-prvkovú podgrupu.

Úloha 12.2.10. Nech G je konečná grupa, počet jej prvkov označme n. Ak $A \subseteq G$ a A má viac než $n/2$ prvkov, tak A generuje G.

Úloha 12.2.11. Nech $\varphi\colon G \to H$ je homomorfizmus grúp. Nech $g \in G$. Označme $h := \varphi(g)$, $A := \varphi^{-1}(h) = \{x \in G; \varphi(x) = h\}$ a $K := \operatorname{Ker}\varphi$. (Z dôsledku 11.3.10 vieme, že K je podgrupa grupy H.) Dokážte, že $A = gK$.

Úloha 12.2.12. Nech S_i je podgrupa G a $t_i \in G$ pre každé $i \in I$. Označme $D = \bigcap_{i\in I} S_i$. Dokážte, že buď $\bigcap_{i\in I} S_i t_i = \emptyset$ alebo existuje $g \in G$ také, že $\bigcap_{i\in I} S_i t_i = Dg$. (T.j. prienik pravých tried $S_i t_i$ je buď prázdna množina alebo niektorá pravá trieda rozkladu podľa D.)

{rozklcvic:GRUPNEPAR}

Úloha 12.2.13. Dokážte, že v každej grupe s nepárnym počtom prvkov je ľubovoľný prvok štvorcom nejakého (a navyše jednoznačne určeného) prvku tejto grupy.

Úloha 12.2.14. Nech H je podgrupa G. Ukážte, že systém $\{HaH; a \in G\}$ je rozklad množiny G.

Úloha 12.2.15. Nech $(G, *)$ je grupa. Dokážte alebo vyvráťte: $H = \{g^2; g \in G\}$ je podgrupa grupy G.

12.3 Normálne podgrupy

Dostali sme teda rozklad G na ľavé triedy a rozklad G na pravé triedy. Tieto dva rozklady môžu byť vo všeobecnosti rôzne – my sa budeme snažiť nájsť podmienky, kedy sú rovnaké.

{normal:TVRAHHB}

Tvrdenie 12.3.1. *Nech H je podgrupa grupy G. Ak $aH = Hb$, tak $Ha = Hb$.* (*Takisto za týchto predpokladov platí $aH = bH$.*)

Dôkaz. Ak $aH = Hb$, tak $a \in Hb$, čiže $a = hb$ pre nejaké $h \in H$. Potom $h = ab^{-1} \in H$ a podľa lemy 12.2.5 máme $Ha = Hb$.

Dôkaz druhej časti tvrdenia je analogický. □

{normal:VTINV}

Veta 12.3.2. *Nech H je podgrupa G. Nasledujúce podmienky sú ekvivalentné:*

{normal:item:IT1} (i) $aH = Ha$ *pre všetky* $a \in G$,

{normal:item:IT2} (ii) $aH \subseteq Ha$ *pre všetky* $a \in G$,

{normal:item:IT3} (iii) $Ha \subseteq aH$ *pre všetky* $a \in G$,

{normal:item:DEFINV} (iv) $aHa^{-1} \subseteq H$ *pre všetky* $a \in G$,

{normal:item:IT5} (v) $H \subseteq aHa^{-1}$ *pre všetky* $a \in G$,

{normal:item:IT4} (vi) $aHa^{-1} = H$ *pre všetky* $a \in G$,

{normal:item:AHHB} (vii) $\{aH; a \in G\} = \{Hb; b \in G\}$.

Všimnime si, že podmienku (v) môžeme zapísať aj tak, že platí $aha^{-1} \in H$ pre všetky $h \in H$ a $a \in G$, čiže

{normal:EQDEFINV}
$$h \in H \qquad \Rightarrow \qquad aha^{-1} \in H. \tag{12.1}$$

V nasledujúcom dôkaze budeme využívať vlastnosti násobenia podmnožín sformulované v leme 12.2.2 ako aj pomerne ľahko dokázateľný fakt, že násobenie podmnožín zachováva inklúzie. (Vlastne nám budú stačiť implikácie $B \subseteq C \Rightarrow aB \subseteq aC$ a $B \subseteq C \Rightarrow Ba \subseteq Ca$.)

Dôkaz. Zrejmé: (i) $\Rightarrow$ (ii), (i) $\Rightarrow$ (iii)
(ii) $\Rightarrow$ (iv): Ak platí $aH \subseteq Ha$, tak platí aj

$$aHa^{-1} \subseteq (Ha)a^{-1} = H(aa^{-1}) = He = H.$$

Podobne sa ukáže (iii) $\Rightarrow$ (iv): Keď použijeme (iii) pre prvok a^{-1}, tak máme $Ha^{-1} \subseteq a^{-1}H$. Z tejto inklúzie dostaneme

$$aHa^{-1} \subseteq aa^{-1}H = H.$$

(iv) $\Rightarrow$ (v): Použitím (iv) pre prvok a^{-1} dostaneme $a^{-1}Ha \subseteq H$. Keď túto inklúziu vynásobíme zľava a a sprava a^{-1}, tak máme

$$H = a(a^{-1}Ha)a^{-1} \subseteq aHa^{-1}.$$

(v) $\Rightarrow$ (vi): Stačí v skutočnosti ukázať, že (v) $\Rightarrow$ (iv), pretože z (v) a (iv) okamžite vyplýva (vi). Dôkaz implikácie (v) $\Rightarrow$ (iv) je takmer rovnaký ako dôkaz predchádzajúcej implikácie. Ak použijeme (v) pre prvok a^{-1}, dostaneme $a^{-1}Ha \subseteq H$. Keď túto inklúziu vynásobíme zľava a a sprava a^{-1}, tak máme

$$H = a(a^{-1}Ha)a^{-1} \subseteq aHa^{-1}.$$

(vi) $\Rightarrow$ (i): Ak $H = aHa^{-1}$, tak $Ha = (aHa^{-1})a = aH(a^{-1}a) = a(He) = aH$.
Takisto implikácia (i) $\Rightarrow$ (vii) je zrejmá. Opačná implikácia (vii) $\Rightarrow$ (i) vyplýva z tvrdenia 12.3.1. Z rovnosti $\{aH; a \in G\} = \{Hb; b \in G\}$ totiž vyplýva, že každé aH sa musí rovnať nejakému Hb, ale potom podľa tvrdenia 12.3.1 platí aj $aH = Ha$.
Dokázali sme:
(i) $\Rightarrow$ (ii) $\Rightarrow$ (iv) $\Rightarrow$ (v) $\Rightarrow$ (vi) $\Rightarrow$ (i),
(i) $\Rightarrow$ (iii) $\Rightarrow$ (iv) $\Rightarrow$ (v) $\Rightarrow$ (vi) $\Rightarrow$ (i),
(i) $\Leftrightarrow$ (vii),
z čoho vyplýva, že všetky uvedené podmienky sú ekvivalentné. □

Definícia 12.3.3. Podgrupa H grupy G sa nazýva *normálna* (*invariantná*) *podgrupa,* ak spĺňa niektorú z ekvivalentných podmienok uvedených vo vete 12.3.2. Označujeme $H \triangleleft G$.

Ak G je komutatívna grupa, tak každá jej podgrupa je invariantná.
Z vety 12.3.2 vidíme, že pre invariantnú podgrupu ľavé a pravé triedy rozkladu sú totožné.

Príklad 12.3.4. Pre každú grupu G sú jej podgrupy G a $\{e\}$ normálnymi podgrupami.

Pretože v komutatívnej grupe je každá podgrupa normálna, úloha zistiť, či nejaká podgrupa je normálna, je zaujímavá len v nekomutatívnom prípade.

{normal:PRS3}

Príklad 12.3.5. Preskúmajme, ktoré podgrupy S_3 sú normálne. Zostavme najprv tabuľku grupovej operácie. (Do riadku φ a stĺpca τ zapisujeme $\varphi \circ \tau$.)

	id	(12)	(13)	(23)	(123)	(132)
id	id	(12)	(13)	(23)	(123)	(132)
(12)	(12)	id	(132)	(123)	(23)	(13)
(13)	(13)	(123)	id	(132)	(12)	(23)
(23)	(23)	(132)	(123)	id	(13)	(12)
(123)	(123)	(13)	(23)	(12)	(132)	id
(132)	(132)	(23)	(12)	(13)	id	(123)

Teraz skúsme nájsť všetky podgrupy grupy S_3. Z Lagrangeovej vety 12.2.14 vieme, že (okrem podgrúp $\{e\}$ a S_3) stačí hladať podgrupy rádu 2 a 3. Podľa dôsledku 12.2.16 ide o cyklické grupy, teda nám stačí nájsť všetky prvky rádu 2 resp. 3.

Prvky rádu 2 sú práve cykly dĺžky 2. Tie vygenerujú podgrupy $H_1 = \{id, (12)\}$, $H_2 = \{id, (13)\}$ a $H_3 = \{id, (23)\}$.

Napríklad pre podgrupu $H_1 = \{id, (12)\}$ máme $(13)H_1 = \{(13), (123)\}$ a $H_1(13) = \{(13), (132)\}$. Keďže sme dostali pre ten istý prvok inú ľavú a pravú triedu, podgrupa H_1 nespĺňa podmienku (i) z vety 12.3.2, a teda nie je normálna.

Podobným spôsobom môžeme overiť, že ani ostatné 2-prvkové podgrupy nie sú normálne.

Prvky rádu 3 sú trojcykly (123) a (132). Obe generujú tú istú 3-prvkovú podgrupu $A_3 = \{id, (123), (132)\}$ pozostávajúcu z párnych permutácií množiny $\{1, 2, 3\}$. Vidíme, že pravý i ľavý rozklad je rovnaký, jeho triedy sú množina A_3 (párne permutácie) a jej doplnok $S_3 \setminus A_3$ (nepárne permutácie). Teda A_3 spĺňa podmienku (vii) z vety 12.3.2, čiže je normálna. (Na zdôvodnenie toho, že A_4 je normálna sme mohli použiť aj všeobecnejší fakt, že každá podgrupa indexu 2 je normálna – úloha 12.3.2.)

Všimnime si, že H_1 je komutatívna grupa. Teda komutatívna podgrupa ešte nutne nemusí byť normálna. (Ako sme už spomínali, ak vieme, že *celá* je *komutatívna,* tak každá podgrupa je normálna.)

Cvičenia

rmalcvic:ULOPRIENIKNORM}

Úloha 12.3.1. Dokážte, že prienik (ľubovoľného systému) normálnych podgrúp danej grupy G je opäť normálna podgrupa G.

{normalcvic:INDEX2}

Úloha 12.3.2. Ak H je podgrupa G a $[G : H] = 2$, tak H je normálna podgrupa. Navyše, pre každý prvok $x \in G$ platí $x^2 \in H$. (Poznamenám, že riešenie druhej časti – hovoriacej o $x^2 \in H$ – sa dá urobiť pomocou faktorových grúp, tie sú až v nasledujúcej podkapitole. Dá sa však vymyslieť aj riešenie bez nich.)

{permcvic:A4LAGR}

Úloha 12.3.3*. Dokážte, že grupa A_4 párnych permutácií 4-prvkovej množiny nemá žiadnu 6-prvkovú podgrupu.

Úloha 12.3.4. Pre všetky $n \in \mathbb{N}$ je A_n normálna podgrupa grupy S_n.

{normalcvic:CYKLNORM}

Úloha 12.3.5. Dokážte, že ľubovoľná normálna podgrupa A_n pre $n \geq 5$, ktorá obsahuje aspoň jeden cyklus dĺžky 3 je celá grupa A_n.

Úloha 12.3.6. Nájdite všetky normálne podgrupy v grupe: a) $(\mathbb{Z}_5, \oplus)$, b) $(\mathbb{Z}_6, \oplus)$, c) $(\mathbb{Z}_4, \oplus)$, d) $(\mathbb{Z}_2 \times \mathbb{Z}_2, \oplus)$, e) $(\mathbb{Z}_{12}, \oplus)$, f) $(\mathbb{Z}_p, \oplus)$, kde p je prvočíslo, g*) $(A_4, \circ)$, h*) $(S_4, \circ)$.

Úloha 12.3.7. Dokážte, že každá jednoduchá podgrupa grupy S_n, ktorá má viac ako 2 prvky, je obsiahnutá v grupe A_n. (Grupa sa nazýva jednoduchá, ak nemá žiadne normálne podgrupy okrem seba samej a triviálnej podgrupy.)

{normalcvic:CENTRUM}

Úloha 12.3.8. Centrom grupy G nazývame množinu $Z(G) = \{g \in G; (\forall h \in G)gh = hg\}$ takých prvkov, ktoré komutujú so všetkými prvkami G. Ukážte, že $Z(G)$ je normálna podgrupa grupy G.

Úloha 12.3.9. Ak A a B sú normálne podgrupy G, $a \in A$ a $b \in B$, tak $aba^{-1}b^{-1} \in A \cap B$.

{normalcvic:PRIENNORM}

Úloha 12.3.10. Ak H a H' sú normálne podgrupy G také, že $H \cap H' = \{e\}$, tak $hh' = h'h$ pre ľubovoľné $h \in H$ a $h' \in H'$ (ľubovoľný prvok H komutuje s ľubovoľným prvkom H'.)

Úloha 12.3.11. Uvažujme grupu $G = \{A \in M_{2,2}(F); |A| = 1\}$ s operáciou násobenia. (Pričom F je ľubovoľné pole.) Definujme $H = \{\left(\begin{smallmatrix} 1 & 0 \\ a & 1 \end{smallmatrix}\right); a \in F\}$.
a) Overte, že H je podgrupa G.
b) Je táto podgrupa komutatívna?
c) Je táto podgrupa normálna?
Aké lineárne zobrazenia zodpovedajú maticiam patriacim do H?

Úloha 12.3.12. Nech $G = \{\left(\begin{smallmatrix} a & 0 \\ b & a \end{smallmatrix}\right); a, b \in \mathbb{R}, a \neq 0\}$ a $H = \{\left(\begin{smallmatrix} 1 & 0 \\ c & 1 \end{smallmatrix}\right); c \in \mathbb{R}\}$.
a) Overte, že G tvorí s operáciou násobenia matíc grupu a že H je podgrupa grupy G.
b) Je H normálna podgrupa grupy G?
(Všimnite si, že H je tá istá grupa, ako v predošlej úlohe, ale teraz sa na ňu pozeráme ako na podgrupu inej grupy.)

Úloha 12.3.13. Uvažujme grupu G všetkých zhodných izometrií[1] roviny nemeniacich orientáciu. Inak povedané, sú to všetky zobrazenia, ktoré môžeme dostať ako zloženie posunutia o nejaký vektor $\vec{u}$ a otočenia o nejaký uhol α, čiže zobrazenia dané predpisom $(x, y) \mapsto (c + x\cos\alpha + y\sin\alpha, d - x\sin\alpha + y\cos\alpha)$. Sú nasledujúce grupy normálnymi podgrupami grupy G?
a) H = všetky posunutia;
b) H = rotácie okolo počiatku súradnicovej sústavy;
c) H_x = všetky zobrazenia z G také, že $f(x) = x$, pričom $x \in \mathbb{R}^2$ je nejaký pevne zvolený bod roviny.

12.4 Faktorové grupy

{fakt:VTFAKT}

Veta 12.4.1. *Ak G je grupa a H je jej invariantná podgrupa, tak na množine všetkých tried G podľa H môžeme definovať operáciu $\cdot$ ako*

$$(aH) \cdot (bH) = (ab)H \tag{12.2}$$

{fakt:EQOP}

Táto operácia je dobre definovaná (nezávisí od výberu reprezentanta triedy) a množina všetkých tried G podľa H s touto operáciou tvorí grupu. Túto grupu označujeme G/H a nazývame faktorová grupa *grupy G podľa H.*

Je dôležité si uvedomiť, že faktorovú grupu môžeme definovať iba pre invariantnú podgrupu.

Dôkaz. Všetky tvrdenia vety vlastne vyplývajú z toho, že takto definované násobenie je to isté ako násobenie podmnožín grupy G. Platí totiž

$$(aH)(bH) = (aH)(Hb) = a(HH)b = aHb = a(Hb) = a(bH) = (ab)H.$$

Z toho vyplýva, že operácia, ktorú sme definovali je dobre definovaná a takisto, že je asociatívna.

Pretože $eH = H$ a $HH = H$, trieda eH je neutrálny prvok.

Inverzný prvok k aH je $a^{-1}H$, pretože $(aH)(a^{-1}H) = (aa^{-1})H = eH = H$. □

{fakt:PRZ3}

Príklad 12.4.2. Ak $G = \mathbb{Z}$ a $H = 3\mathbb{Z}$ tak G/H obsahuje 3 triedy H, $1 + H$ a $2 + H$ (príklad 12.2.4).

[1] izometria=zobrazenie zachovávajúce vzdialenosti; zhodné = zachovávajú aj orientáciu

	H	$1+H$	$2+H$
H	H	$1+H$	$2+H$
$1+H$	$1+H$	$2+H$	H
$2+H$	$2+H$	H	$1+H$

Z predchádzajúcej tabuľky vidíme, že $H \mapsto 0$, $1+H \mapsto 1$, $2+H \mapsto 2$ je izomorfizmus medzi G/H a $(\mathbb{Z}_3, \oplus)$, čiže v tomto prípade je faktorová grupa izomorfná s grupou $\mathbb{Z}_3$.

{fakt:PRRxR}

Príklad 12.4.3. Nech $G = (\mathbb{R} \times \mathbb{R}, +)$ a $H = \{(x, x); x \in \mathbb{R}\}$.

V príklade 12.2.8 sme videli, že každú triedu rozkladu môžeme reprezentovať ako $(r, 0) + H$ (a navyše takto dostaneme každú triedu práve raz). Z toho pomerne ľahko vidno, že zobrazenie

$$(r, 0) + H \mapsto r$$

je izomorfizmus medzi grupami G/H a $(\mathbb{R}, +)$, čiže $G/H \cong \mathbb{R}$.

Podobne môžeme vidieť, že pre $G = \mathbb{Z}_8$ a $H = 4\mathbb{Z}_8$ máme $G/H \cong \mathbb{Z}_4$. V ďalšej podkapitole dokážeme vetu, ktorá nám umožní takéto vlastnosti dokazovať pomerne jednoducho a elegantne.

Cvičenia

Úloha 12.4.1. Dokážte, že ak H je normálna podgrupa G a $[G : H] = n$, tak $x^n \in H$ pre ľubovoľné $x \in G$. Ukážte na príklade, že toto tvrdenie nemusí platiť, ak H nie je normálna.

12.5 Vety o izomorfizme

V úlohe 12.1.3 sme videli vzťah medzi surjektívnymi zobrazeniami a reláciami ekvivalencie. V prípade, že na danej množine máme navyše grupovú štruktúru, surjektívne homomorfizmy budú podobným spôsobom zodpovedať normálnym podgrupám (a navyše, ako uvidíme v cvičeniach za touto časťou, istým špeciálnym reláciam ekvivalencie, ktoré voláme kongruencie).

{izom:VTKANON}

Veta 12.5.1 (Kanonický homomorfizmus)**.** *Ak G je grupa a H je normálna podgrupa G, tak zobrazenie $f\colon G \to G/H$ dané predpisom*

$$f\colon a \mapsto aH$$

je surjektívny homomorfizmus. Tento homomorfizmus voláme kanonický homomorfizmus.

Navyše, jadro kanonického homomorfizmu je práve podgrupa H.

Dôkaz. Z vlastností násobenia podmnožín grupy (lema 12.2.2) a z toho, že H je normálna podgrupa dostaneme

$$f(a)f(b) = (aH)(bH) = a(Hb)H = a(bH)H = (ab)H^2 = (ab)H = f(ab).$$

Teda toto zobrazenie je skutočne homomorfizmus.

Surjektívnosť vyplýva priamo z definície.

Pretože neutrálny prvok faktorovej grupy G/H je $eH = H$, jadro zobrazenia f je množina tých $a \in G$, pre ktoré platí $aH = eH$, čo je presne podgrupa H (vyplýva to napríklad z lemy 12.2.5, ľahko to však môžeme overiť aj priamo.) □

Vidíme teda, že pre každú faktorovú grupu máme surjektívny homomorfizmus. Obrátené tvrdenie dáva nasledujúca veta:

{izom:VTIZOM}

Veta 12.5.2 (Veta o izomorfizme)**.** *Ak $f\colon G \to G'$ je homomorfizmus grúp, tak* $\operatorname{Ker} f$ *je normálna podgrupa grupy G a faktorová grupa $G/\operatorname{Ker} f$ je izomorfná s podgrupou* $\operatorname{Im} f$ *grupy G'.*

Dôkaz. Označme $H = \operatorname{Ker} f$ a neutrálny prvok grupy G' označme ako e'. Z dôsledku 11.3.10 vieme, že H je podgrupa G. Ukážeme, že táto podgrupa je normálna. Skutočne, ak $h \in \operatorname{Ker} f$, t.j. $f(h) = e'$, tak aj

$$f(aha^{-1}) = f(a)f(h)f(a)^{-1} = f(a)e'f(a)^{-1} = f(a)f(a)^{-1} = e'$$

a $aha^{-1} \in \operatorname{Ker} f = H$.

Definujme zobrazenie $\varphi\colon G/H \to \operatorname{Im} f$ ako

$$\varphi\colon aH \mapsto f(a).$$

Najprv ukážeme, že toto zobrazenie je dobre definované (nezávisí od výberu reprezentanta ľavej triedy aH). Skutočne, ak $aH = bH$, tak $b^{-1}a \in H = \operatorname{Ker} f$, čiže $f(b^{-1}a) = e'$. Potom

$$f(b) = f(b)e' = f(b)f(b^{-1}a) = f(bb^{-1}a) = f(a).$$

Zostáva dokázať, že takto definované zobrazenie je bijektívny homomorfizmus. Máme

$$\varphi(abH) = f(ab) = f(a)f(b) = \varphi(aH)\varphi(bH),$$

teda φ je homomorfizmus.

Surjektívnosť vyplýva z toho, že za obor hodnôt sme zobrali $\operatorname{Im} f$. Aby sme ukázali, že homomorfizmus φ je injektívny, stačí ukázať, že $\operatorname{Ker}\varphi$ obsahuje iba neutrálny prvok (úloha 11.3.7). Skutočne, ak $\varphi(aH) = e'$, znamená to, že $f(a) = e'$ a $a \in \operatorname{Ker} f = H$, teda $aH = H$. □

Dôsledok 12.5.3. *Ak $f\colon G \to H$ je surjektívny homomorfizmus grúp, tak grupa H je izomorfná s faktorovou grupou $G/\operatorname{Ker} f$.*

Vety 12.5.1 a 12.5.2 nám hovoria, že normálne podgrupy sú práve jadrá homomorfizmov. (Jadro každého homomorfizmu je normálna podgrupa a obrátene, pre každú normálnu podgrupu máme epimorfizmus na faktorovú grupu, ktorého jadrom je práve táto podgrupa.)

Môžeme si všimnúť, že veta o izomorfizme nám dáva ďalšiu možnosť ako ukázať, že nejaká podgrupa grupy G je normálna – ak sa nám podarí nájsť homomorfizmus z G do inej grupy, ktorého jadrom je daná podgrupa. Dokonca vieme jednoducho popísať aj triedy rozkladu – do jednej triedy patria tie prvky z G, ktoré majú v tomto homomorfizme ten istý obraz. (Úloha 12.5.12, viac-menej to vidno už aj z dôkazu vety o izomorfizme.)

Z vety o izomorfizme okamžite dostaneme nasledujúce jednoduché dôsledky.

Príklad 12.5.4. Použime vetu o izomorfizme pre homomorfizmy $f\colon G \to H$, $f(x) = e$ (kde G, H sú ľubovoľné grupy a e označuje neutrálny prvok grupy H) a $id_G\colon G \to G$. Platí $\operatorname{Ker} f = G$ a $\operatorname{Ker} id_G = \{e\}$, z čoho vyplýva na základe vety 12.5.2 $G/G \cong \{e\}$ a $G/\{e\} \cong G$.

V predchádzajúcej podkapitole sme uviedli niekoľko príkladov faktorových grúp pričom sme spomenuli, že sú izomorfné s niektorými známymi grupami. Predchádzajúca veta je jednoduchým prostriedkom ako ukázať, že ide skutočne o izomorfné grupy.

{izom:PRZ3}

Príklad 12.5.5. Z príkladu 12.2.4 vieme, že rozklad grupy $\mathbb{Z}$ podľa podgrupy $3\mathbb{Z}$ má 3 prvky. Z toho je jasné, že faktorová grupa $\mathbb{Z}/3\mathbb{Z}$ je izomorfná s grupou $(\mathbb{Z}_3, \oplus)$. (Z dôsledku 12.2.16 a vety 11.4.12 vyplýva, že $\mathbb{Z}_3$ je, až na izomorfizmus, jediná trojprvková grupa.)

Skúsme však tento fakt odvodiť na základe vety 12.5.2. Na to stačí nájsť surjektívny homomorfizmus so $\mathbb{Z}$ na $\mathbb{Z}_3$, ktorého jadro je práve $3\mathbb{Z}$. Takýmto homomorfizmom je zobrazenie $f\colon \mathbb{Z} \to \mathbb{Z}_3$ dané predpisom

$$f\colon n \mapsto n \bmod 3,$$

t.j. každému prvku priradí zvyšok po delení 3.

{izom:PRRxR}

Príklad 12.5.6. Uvažujme situáciu z príkladu 12.2.8, t.j. grupu $G = (\mathbb{R} \times \mathbb{R}, +)$ a jej podgrupu $H = \{(x, x); x \in \mathbb{R}\}$. Jednoducho možno overiť, že zobrazenie $f\colon (G, +) \to (\mathbb{R}, +)$

$$f\colon (x, y) \mapsto x - y$$

je epimorfizmus a $\operatorname{Ker} f = H$. Z toho vyplýva, že $G/H \cong (\mathbb{R}, +)$.

{izom:PRZ8}

Príklad 12.5.7. V príklade 12.2.9 sme mali $G = \mathbb{Z}_8$ a $H = 4\mathbb{Z}_8 = \{0, 4\}$. V tomto prípade máme surjektívny homomorfizmus $f\colon \mathbb{Z}_8 \to \mathbb{Z}_4$

$$f\colon n \mapsto n \bmod 4$$

a faktorová grupa G/H je izomorfná s grupou $(\mathbb{Z}_4, \oplus)$.

{izom:PRS3}

Príklad 12.5.8. V príklade 12.3.5 sme videli, že jedinou normálnou podgrupou grupy S_3 je podgrupa A_3 všetkých párnych permutácií. Zobrazenie $f\colon S_3 \to \mathbb{Z}_2$, ktoré priradí párnym permutáciam 0 a nepárnym 1 je surjektívny homomorfizmus taký, že $\operatorname{Ker} f = A_3$.

12.5.1 Druhá a tretia veta o izomorfizme*

Dôkaz nasledujúceho tvrdenia je veľmi podobný tej časti dôkazu vety 12.5.2, v ktorej sme dokazovali, že ide o homomorfizmus.

{izom:LMKERHOM}

Lema 12.5.9. *Nech $f\colon G \to G'$ je grupový homomorfizmus. Nech H je normálna podgrupa G taká, že $H \subseteq \operatorname{Ker} f$. Potom zobrazenie $\varphi\colon G/H \to G'$ dané predpisom*

$$\varphi(aH) = f(a)$$

je dobre definované a je to grupový homomorfizmus.

Navyše, ak f je epimorfizmus, tak aj φ je epimorfizmus.

Dôkaz. Najprv ukážeme, že φ je dobre definované. Ak máme 2 rôznych reprezentantov tej istej triedy, t.j. $aH = bH$, tak platí $b^{-1}a \in H \subseteq \operatorname{Ker} f$. To znamená, že $f(b^{-1}a) = e'$, a teda

$$f(b) = f(b)e' = f(b)f(b^{-1}a) = f(bb^{-1}a) = f(a).$$

Overíme teraz, že φ je homomorfizmus.

$$\varphi(abH) = f(ab) = f(a)f(b) = \varphi(aH)\varphi(bH).$$

Ak f je surjektívne zobrazenie, tak pre každé $b \in G'$ existuje $a \in G$ také, že $f(a) = b$. Potom platí $\varphi(aH) = b$, teda aH je vzor b v zobrazení φ. Z toho vyplýva, že aj zobrazenie φ je surjektívne. □

Ak túto lemu použijeme na kanonický homomorfizmus $\varphi\colon G \to G/K$, dostaneme:

om:DOSKERHOM}

Dôsledok 12.5.10. *Ak H, K sú normálne podgrupy grupy G a $H \subseteq K$, tak zobrazenie $f\colon G/H \to G/K$*

$$f\colon aH \mapsto aK$$

je surjektívny homomorfizmus.

Pomocou predchádzajúcej vety môžeme odvodiť výsledok, ktorý pripomína „krátenie“ pre faktorové grupy. Dôležité je uvedomiť si, že ak H, K sú normálne podgrupy G a $H \subseteq K$, tak H je normálna podgrupa K. Navyše K/H je podmnožina G/H tvorená triedami aH pre ktoré $a \in K$.

{izom:VTIZOM2}

Veta 12.5.11 (Tretia veta o izomorfizme)**.** *Ak H, K sú normálne podgrupy G, pričom $H \subseteq K \subseteq G$, tak K/H je normálna podgrupa G/H a platí*

$$G/K \cong (G/H)/(K/H).$$

Dôkaz. Pretože $H \subseteq K$, dostávame z dôsledku 12.5.10, že zobrazenie $\psi\colon G/H \to G/K$ určené predpisom

$$\psi\colon aH \mapsto aK$$

je surjektívny homomorfizmus. Potom podľa vety 12.5.2 je grupa G/K izomorfná s grupou $(G/H)/(\operatorname{Ker}\psi)$. Pokúsme sa teda určiť jadro homomorfizmu ψ.

Do $\operatorname{Ker}\psi$ patria tie ľavé triedy aH grupy G/H, ktoré sa zobrazia na neutrálny prvok grupy G/K, čiže na $eK = K$. Teda $aH \in \operatorname{Ker}\psi$ platí práve vtedy, keď $aK = K$, čiže $a \in K$. To znamená, že $\operatorname{Ker}\psi = K/H$ (keďže $\operatorname{Ker}\psi$ pozostáva práve z tých ľavých tried aH, pre ktoré $a \in K$). Vďaka tomu vidíme z vety o izomorfizme, že $K/H = \operatorname{Ker}\psi$ je normálna podgrupa a

$$(G/H)/(K/H) \cong G/K.$$

□

Dôkaz predchádzajúcej vety opäť ilustruje užitočnosť vety 12.5.2. Keby sme namiesto použitia tejto vety robili priamy dôkaz, museli by sme pracovať s prvkami grupy $(G/H)/(K/H)$, čo sú triedy rozkladu faktorovej grupy G/H podľa jej podgrupy K/H, ktorých reprezentantmi sú opäť triedy rozkladu, tentokrát rozkladu K podľa H. Zdá sa, že prístup využívajúci vetu o izomorfizme je prehľadnejší.

Ukážeme si ešte jeden výsledok o faktorových grupách.

Veta 12.5.12 (Druhá veta o izomorfizme)**.** *Nech G je grupa, N je normálna podgrupa G a S je podgrupa G. Potom množina SN tvorí podgrupu grupy G, N je normálna podgrupa SN, $S \cap N$ je normálna podgrupa S a platí*

$$S/(S \cap N) \cong SN/N.$$

Dôkaz. Najprv overme, že $SN = \{ab; a \in S, b \in N\}$ je podgrupa G. Z toho, že N je normálna, máme $gN = Ng$ pre každé $g \in G$, teda pre ľubovoľné $g \in G$ a $n \in N$ existuje $n' \in \mathbb{N}$ také, že $gn = n'g$. Pomocou tejto vlastnosti už ľahko overíme kritérium podgrupy.

Ak máme 2 prvky z SN tvaru a_1b_1, a_2b_2, kde $a_{1,2} \in S$ a $b_{1,2} \in S$, tak ich súčin môžeme vyjadriť ako

$$a_1b_1a_2b_2 = a_1a_2b_1'b_2$$

pre nejaké $b_1' \in N$. Keďže $a_1a_2 \in S$ a $b_1'b_2 \in N$, vidíme, že uvedený súčin patrí do SN.

Ak máme prvok z SN tvaru ab, pričom $a \in S$, $b \in N$, tak tento prvok môžeme prepísať ako $b'a$ pre vhodné $b \in N$. Potom inverzný prvok

$$(b'a)^{-1} = a^{-1}b'^{-1}$$

je opäť z SN.

Z toho, že $N \subseteq SN$ dostaneme, že N je podgrupa SN. Fakt, že $N \triangleleft SN$, t.j. ide o normálnu podgrupu, overíme použitím podmienky (vi) z vety 12.3.2. Skutočne, pre ľubovoľné $a \in S$, $b \in N$ máme

$$(ab)N(ab)^{-1} = (ab)N(b^{-1}a^{-1}) = a(bNb^{-1})a^{-1} = aNa^{-1} = N.$$

Na dôkaz ostatných častí tvrdenia použijeme kanonický homomorfizmus

$$f\colon a \mapsto aN,$$
$$f\colon G \to G/N.$$

Aj jeho zúženie $f|_S\colon S \to G/N$ na množinu S je homomorfizmus. Pokúsme sa zistiť, čomu sa rovná jeho jadro.

Máme

$$\operatorname{Ker} f|_S = \{a \in S; aN = N\} = \{a \in S; a \in N\} = S \cap N.$$

Podľa vety o izomorfizme je $S \cap N$ normálna podgrupa S a platí

$$S/(S \cap N) \cong \operatorname{Im} f|_S.$$

Stačí nám už teda len dokázať, že $\operatorname{Im} f|_S = SN/N$.

Množina $\operatorname{Im} f|_S$ pozostáva práve z tých tried aN, pre ktoré $a \in S$;

$$\operatorname{Im} f|_S = \{aN; a \in S\}.$$

Súčasne však máme

$$SN/N = \{abN; a \in S, b \in N\} = \{aN; a \in S\},$$

teda skutočne platí rovnosť $\operatorname{Im} f|_S = SN/N$. □

Cvičenia

Úloha 12.5.1. Overte, či H je normálna podgrupa grupy G a opíšte faktorovú grupu G/H (aké má triedy, vybrať z každej triedy práve jedného reprezentanta, zistiť, či je izomorfná s nejakou známou grupou).

a) $G = (\mathbb{R} \times \mathbb{R}, +)$, $H = \{(x, y); x + 2y = 0\}$
b) $G = (\mathbb{R} \times \mathbb{R}, +)$, $H = \{(x, 3x); x \in \mathbb{R}\}$
c) $G = (\mathbb{C}, +)$, $H = \mathbb{R}$
d) $G = (\mathbb{C} \setminus \{0\}, \cdot)$, $H = \mathbb{R} \setminus \{0\}$
e) $G = (\mathbb{C} \setminus \{0\}, \cdot)$, $H = \mathbb{R}^+ = \{x \in \mathbb{R}; x > 0\}$
f) $G = (\mathbb{Z}, +)$, $H = 4\mathbb{Z} = \{4z; z \in \mathbb{Z}\}$
g) $G = (\mathbb{Z}_4 \times \mathbb{Z}_6, +)$, $H = [(2, 2)]$
h) $G = (\mathbb{Z} \times \mathbb{Z} \times \mathbb{Z}, +)$, $H = \{(n, m, 0); n, m \in \mathbb{Z}\}$
i) $G = (\{c \in \mathbb{C}; c^{12} = 1\}, \cdot)$, $H = \{c \in \mathbb{C}; c^4 = 1\}$
j) $G = (S_n, \circ)$, $H = A_n$
k) $G = (\mathbb{Z}, +)$, $H = n\mathbb{Z}$
l) $G = (\mathbb{C} \setminus \{0\}, \cdot)$, $H = \{c \in \mathbb{C}; c^6 \in \mathbb{R} \setminus \{0\}\}$
m) $G = (\mathbb{C} \setminus \{0\}, \cdot)$, $H = \{c \in \mathbb{C}; c^n \in \mathbb{R} \setminus \{0\}\}$

Úloha 12.5.2. Zistite, či dané grupy sú izomorfné. V celom cvičení budeme ako S označovať grupu $(\{c \in \mathbb{C}; |c| = 1\}, \cdot)$ (prípadne množinu prvkov tejto grupy) a $C_n = (\{c \in \mathbb{C}; c^n = 1\}, \cdot)$
a) $(\mathbb{C} \setminus \{0\}, \cdot)/\{c \in \mathbb{C}; c^n \in \mathbb{R} \setminus \{0\}\}$, $(\mathbb{C} \setminus \{0\}, \cdot)/\mathbb{R}^+$ (pod $\mathbb{R}^+$ tu myslíme kladné reálne čísla, čiže $0 \notin \mathbb{R}^+$), S
b) $(\mathbb{R}, +)/\mathbb{Z}$, S/C_n, S
c) $(\mathbb{C} \setminus \{0\}, \cdot)$, $(\mathbb{C} \setminus \{0\}, \cdot)/C_n$
d) $(\{c \in \mathbb{C}; c^n \in \mathbb{R} \setminus \{0\}\}, \cdot)/\mathbb{R}^+$, C_n
e) $(\{c \in \mathbb{C}; c^n \in \mathbb{R} \setminus \{0\}\}, \cdot)/C_n$, $\mathbb{R}^+$
f) C_{12}/C_4, $\mathbb{Z}_3$
g) $(\mathbb{Z}_2 \times \mathbb{Z}_3, +)/(\mathbb{Z}_2 \times \{0\})$, $\mathbb{Z}_3$
h) $S_3/[(123)]$, $(\mathbb{Z}_2, +)$

Úloha 12.5.3. Nech G je grupa všetkých regulárnych matíc typu $n \times n$ (s operáciou násobenia matíc). Ako H označme tie z nich, ktoré majú determinant $|A| = 1$. Dokážte, že H je invariantná podgrupa G! Vedeli by ste nájsť grupu izomorfnú s G/H?

Úloha 12.5.4. Nech $S = \{z \in \mathbb{C}; |z| = 1\}$ (jednotková kružnica v komplexnej rovine). Ukážte, že zobrazenie $\varphi\colon \mathbb{R} \to S$ definované ako $\varphi(x) = e^{2\pi x i} = \cos 2\pi x + i \sin 2\pi x$ je surjektívny homomorfizmus. Nájdite $\operatorname{Ker}\varphi$. Aká faktorová grupa je potom izomorfná s kružnicou?

{izomcvic:NORMA4}

Úloha 12.5.5. Je podgrupa $\{id, (12)(34), (13)(24), (14)(23)\}$ normálna podgrupa A_4?

Úloha 12.5.6. Ukážte, že $\mathbb{Q}/\mathbb{Z}$ (obe grupy berieme so sčitovaním) je nekonečná grupa, v ktorej má každý prvok konečný rád.

Úloha 12.5.7. Nech $\varphi\colon G \to G/H$ je kanonický homomorfizmus a $X \subset G$. Dokážte: Ak $\varphi[X]$ generuje G/H, tak $H \cup X$ generuje G.

Úloha 12.5.8. Ukážte na príklade, že ak H je normálna podgrupa G, tak G nemusí obsahovať podgrupu izomorfnú s G/H.

Úloha 12.5.9. Nech H je podgrupa grupy G a $[G : H] = n$.
a) Ukážte, že ak H je normálna podgrupa, tak pre každé $x \in G$ platí $x^n \in H$.
b) Platí toto tvrdenie pre ľubovoľnú podgrupu (t.j. aj bez predpokladu, že H je normálna)?

Úloha 12.5.10. Nech G je množina všetkých matíc tvaru $\left(\begin{smallmatrix} a & 0 \\ b & 1 \end{smallmatrix}\right)$, kde $a, b \in \mathbb{R}$ a $a \neq 0$. Dokážte:
a) G s násobením matíc tvorí grupu.
b) Zobrazenia $f\colon G \to (\mathbb{R}, \cdot)$ a $g\colon G \to (\mathbb{R}, +)$ dané ako $f\colon \left(\begin{smallmatrix} a & 0 \\ b & 1 \end{smallmatrix}\right) \mapsto a$; $g\colon \left(\begin{smallmatrix} a & 0 \\ b & 1 \end{smallmatrix}\right) \mapsto b$ sú homomorfizmy.
c) Izomorfizmus akých grúp dostaneme podľa časti b) na základe vety o izomorfizme?
d) Popíšte lineárne zobrazenia zodpovedajúce maticiam z $\operatorname{Ker} f$, $\operatorname{Ker} g$ a G.

Úloha 12.5.11. Nech $G = \{\left(\begin{smallmatrix} a & b \\ -b & a \end{smallmatrix}\right), \left(\begin{smallmatrix} a & b \\ b & -a \end{smallmatrix}\right); a, b \in \mathbb{R}; a^2 + b^2 = 1\}$ (operácia je násobenie matíc.) Nech ďalej $H = \{\left(\begin{smallmatrix} a & b \\ -b & a \end{smallmatrix}\right); a, b \in \mathbb{R}; a^2 + b^2 = 1\}$.
a) Zistite, či G je grupa a či H je podgrupa G.
b) Je H normálna podgrupa grupy G? Ak áno, s akou grupou je izomorfná grupa G/H?
c) Ako sa dajú geometricky popísať lineárne zobrazenia zodpovedajúce maticiam z G a H?

{izomcvic:ULOTRIEDYSUVZO

Úloha 12.5.12. Dokážte, že ak $f\colon G \to H$ je surjektívny homomorfizmus grúp, tak ľavý (pravý) rozklad grupy G podľa normálnej podgrupy $\operatorname{Ker} f$ pozostáva presne z množín $f^{-1}(x) = \{g \in G; f(g) = x\}$ pre $x \in H$. (Teda triedy rozkladu sú presne vzory prvkov z H. Takto dostávame bijekciu medzi triedami rozkladu G podľa $\operatorname{Ker} f$ a prvkami H, táto bijekcia sa už vyskytla v dôkaze vety o izomorfizme.)

Úloha 12.5.13. V úlohe 12.3.8 sme videli, že centrum grupy G, t.j. množina $Z(G) = \{g \in G; (\forall h \in G)gh = hg\}$, je normálna podgrupa G. Dokážte, že faktorová grupa $G/Z(G)$ je izomorfná s grupou $\mathrm{VAut}\, G$ všetkých vnútorných automorfizmov grupy G (pozri úlohu 11.6.1).

Úloha 12.5.14. Vo vete 12.3.1 sme ukázali, že ak $H \triangleleft G$, tak predpis

$$(aH) \cdot (bH) = (ab)H$$

dobre definuje binárnu operáciu na množine ľavých tried rozkladu G podľa H. Ukážte, že platí aj opačná implikácia: Ak uvedený predpis dobre definuje binárnu operáciu, tak $H \triangleleft G$. (Teda invariantnosť podgrupy H je nielen postačujúca ale aj nutná podmienka na to, aby táto binárna operácia bola dobre definovaná.)

Úloha 12.5.15*. Dokážte, že v grupe $(\mathbb{Q}, +)$ neexistuje maximálna (vzhľadom na inklúziu) vlastná podgrupa. T.j. neexistuje podgrupa S grupy $(\mathbb{Q}, +)$ taká, že ak $S \subseteq T$ a T je podgrupa, tak $T = S$ alebo $T = \mathbb{Q}$. (Inak povedané: Jediné podgrupy obsahujúce S by boli S a $\mathbb{Q}$.)

12.6 Grupové kongruencie

{grpkong:SECTGRPKONG}

Už vieme, že normálne podgrupy sú presne jadrá homomorfizmov. Veta o izomorfizme nám tiež ukazuje súvis medzi surjektívnymi homomorfizmami a faktorovými grupami.

Táto podkapitola nám dá ešte jeden možný pohľad na faktorové grupy a normálne podgrupy. Uvedomiť si vzťah medzi týmito rôznymi prístupmi môže byť užitočné pre lepšie pochopenie týchto pojmov.

My sme najprv zaviedli normálne grupy a pomocou nich faktorové grupy. V princípe by sa dalo postupovať aj naopak – začať s pojmom kongruencie, pomocou neho definovať faktorovú grupu a takto sa dostať aj k pojmu faktorovej grupy. Je na vašom posúdení, ktorý postup sa vám zdá prirodzenejší a zrozumiteľnejší.

To, čo budeme robiť v tejto kapitole sa dá zhruba zhrnúť takto: Ak máme nejakú reláciu ekvivalencie na danej množine G, tak dostávame rozklad na triedy ekvivalencie. Predpokladajme, že máme k dispozícii aj binárnu operáciu na tejto množine – primárne nás bude zaujímať prípad, že máme grupu $(G, \cdot)$. Vieme potom nejako dostať binárnu operáciu alebo dokonca grupu na množine tried ekvivalencie? Aké podmienky by mala spĺňať táto relácia ekvivalencie, aby sa to dalo urobiť?

12.6.1 Pojem kongruencie pre celé čísla

Neskôr aspoň stručne spomenieme aj kongruencie pre okruhy (definícia 13.2.23). Na začiatok je ale možno jednoduchšie pozerať sa na prípad, keď máme iba jednu operáciu – takže hlavný cieľ tejto časti je pozrieť sa na kongruencie v grupách.

Napriek tomu však najprv pripomeniem situáciu, v ktorej ste sa už so slovom *kongruencia* stretli – a je to v skutočnosti špeciálny prípad okruhovej kongruencie.

Pre ľubovoľné dve celé čísla vieme povedať čo to znamená, že sú kongruentné modulo n.

Definícia 12.6.1. Nech $n \in \mathbb{N}$, $n > 0$. Pre $x, y \in \mathbb{Z}$ hovoríme, že x a y sú *kongruentné modulo n,* ak platí $n \mid x - y$. Fakt, že x a y sú kongruentné modulo n označujeme ako

$$x \equiv y \pmod{n}.$$

Dve celé čísla sú teda kongruentné ako ich rozdiel je násobok n – čo je vlastne to isté, ako povedať, že dávajú rovnaký zvyšok po delení číslom n.

Tiež si môžeme všimnúť, že čísla x a y sú kongruentné modulo n práve vtedy, keď $x-y \in n\mathbb{Z} = \{nk; k \in \mathbb{Z}\}$ a množina $n\mathbb{Z}$ je podgrupa grupy $(\mathbb{Z}, +)$. Pripomenúť si tento základný príklad môže byť užitočné neskôr, keď sa budeme snažiť pozrieť na vzťah medzi grupovými kongruenciami a normálnymi podgrupami.

Všimnime si, že tento vzťah sa pekne správa vzhľadom na sčitovanie aj násobenie. Konkrétne, ak platí

$$\begin{aligned} x_1 &\equiv y_1 \pmod{n} \\ x_2 &\equiv y_2 \pmod{n} \end{aligned}$$

tak platí aj

$$\begin{aligned} x_1 + y_1 &\equiv x_2 + y_2 \pmod{n} \\ x_1 \cdot y_1 &\equiv x_2 \cdot y_2 \pmod{n} \end{aligned}$$

(Ak to nie je jasné, skúste si rozmyslieť prečo.)

Tiež by malo byť pomerne ľahké skontrolovať, že vzťah „byť kongruentný modulo n“ je relácia ekvivalencie.

Cieľom tejto kapitoly je pozerať sa na veľmi podobný pojem – budeme mať nejakú reláciu ekvivalencie, ktorá rešpektuje našu grupovú operáciu. A budeme sa chcieť pozrieť na to, aké ďalšie veci z toho vieme dostať. (A tiež na to, ako to celé súvisí s faktorovými grupami a s normálnymi podgrupami.)

Na množine celých čísel máme dve operácie, obe sa slušne správajú vzhľadom na kongruenciu. Teda asi prirodzenejšie by bolo pozerať sa na tento príklad ako na príklad okruhovej kongruencie. Ak chceme teda celé čísla použiť ako ilustráciu pojmu, ktorému sa venujeme teraz, tak si všímajme iba sčitovanie a pracujme s grupou $(\mathbb{Z}, +)$. Potom tento vzťah je špeciálnym prípadom grupovej kongruencie, ktorú teraz ideme zadefinovať.

12.6.2 Relácia kongruencie

Pre danú grupu G sa budeme pozerať na vhodné relácie ekvivalencie na množine G.

{grpkong:DEFKONG}

Definícia 12.6.2. Nech $(G, \cdot)$ je grupa. Relácia ekvivalencie $\sim$ na množine G sa nazýva *kongruencia,* ak pre ľubovoľné $x_{1,2} \in G$, $y_{1,2} \in G$ platí

$$x_1 \sim y_1, x_2 \sim y_2 \qquad \Rightarrow \qquad x_1 \cdot x_2 \sim y_1 \cdot y_2. \tag{12.3}$$

{grpkong:EQDEF}

Ako sme už spomenuli, kongruencie sa dajú zaviesť aj pre iné typy štruktúr. Niekedy teda budeme používať aj názov *grupová kongruencia.*

Podmienka (12.3) je pomerne prirodzená. Vlastne hovorí, že nás zaujímajú iba také relácie ekvivalencie, ktoré sa správajú rozumne vzhľadom na našu operáciu $\cdot$.

Predtým, než budeme s kongruenciami pracovať ďalej, skúsme si uvedomiť, že kongruencie zachovávajú aj inverzné prvky.

Ako to často pri zavádzaní nového pojmu býva, aj tu vieme vymyslieť nejaké veľmi triviálne prípady

Príklad 12.6.3. Pre ľubovoľnú grupu $(G, \cdot)$ sú relácie

$$\begin{aligned} R_1 &= G \times G \\ R_2 &= \{(x, y) \in G \times G; x = y\} \end{aligned}$$

grupové kongruencie.

Príklad 12.6.4. Iný príklad grupovej kongruencie sme už spomenuli. Zoberme si napríklad $(\mathbb{Z}, +)$ a reláciu

$$x \sim y \Leftrightarrow 6 \mid x - y,$$

t.j. reláciu hovoriacu, že x a y sú kongruentné modulo 6. Dostaneme tak grupovú kongruenciu, pretože

$$\left.\begin{array}{ll} a_1 \equiv b_1 & (\mathrm{mod}\ 6) \\ a_2 \equiv b_2 & (\mathrm{mod}\ 6) \end{array}\right\} \quad \Longrightarrow \quad a_1 + a_2 \equiv b_1 + b_2 \pmod 6$$

A takisto to funguje ak namiesto čísla šesť vezmeme ľubovoľné $n \in \mathbb{N}$, $n > 0$.

{grpkong:LMKONGINV}

Lema 12.6.5. *Nech $(G, \cdot)$ je grupa a $\sim$ je grupová kongruencia. Potom pre ľubovoľné prvky $x, y \in G$ také, že $x \sim y$ platí aj $x^{-1} \sim y^{-1}$.*

$$(\forall x, y \in G) x \sim y \Rightarrow x^{-1} \sim y^{-1}$$

Dôkaz nie je príliš ťažký – oplatí sa nad ním skúsiť zamyslieť aj samostatne. Ale pre úplnosť som ho uviedol.

Dôkaz. Zo vzťahu $x \sim y$, z definície kongruencie a z vlastností grupy postupne dostaneme:

$$\begin{aligned} x &\sim y \\ xx^{-1} &\sim yx^{-1} \\ e &\sim yx^{-1} \\ y^{-1} &\sim y^{-1}yx^{-1} \\ y^{-1} &\sim x^{-1} \\ x^{-1} &\sim y^{-1} \end{aligned}$$

(V dôkaze sme vlastne využívali aj asociatívnosť – vďaka tomu sme mohli vynechať zátvorky.) □

12.6.3 Faktorová grupa pre danú reláciu

Po tejto drobnej príprave sa už môžeme pozrieť na to, že kongruencia je presne ten typ relácie, pre ktorú vieme vyrobiť binárnu operáciu (a grupu) aj na množine tried.

Pre reláciu ekvivalencie $\sim$ na množine G si označme množinu tried ekvivalencie ako

$$G/\sim = \{[a]; a \in G\}.$$

{grpkong:TVRKONGFAKTGRP}

Tvrdenie 12.6.6. *Nech $(G, \cdot)$ je grupa a $\sim$ je grupová kongruencia na G.*

{grpkong:itDOBREDEF}

(i) *Predpis*

{grpkong:EQOP}

$$[a] \cdot [b] = [a \cdot b] \tag{12.4}$$

určuje dobre definovanú *binárnu operáciu na množine G.*

{grpkong:itGRUPA}

(ii) *Množina G s uvedenou operáciou tvorí grupu. Ak G je komutatívna, tak aj grupa $G/\sim$ je komutatívna.*

Grupu $G/\sim$ budeme nazývať faktorová grupa.

Síce sme ten istý názov názov faktorová grupa použili v dvoch rôznych definíciach – vysvetlíme si však, že sú to vlastne iba dva rôzne pohľady na tú istú vec. (A aj teraz – kým sme ešte nepovedali nič o vzťahu medzi grupovými kongruenciami a normálnymi podgrupami – by malo byť z kontextu jasné, či faktorizujeme podľa podgrupy alebo podľa kongruencie.)

Dôkaz. (i): Vlastne chceme ukázať, že pre ľubovoľné $a_1, a_2, b_1, b_2 \in G$ platí

$$\left.\begin{aligned}[a_1] &= [b_1]\\ [a_2] &= [b_2]\end{aligned}\right\} \qquad \Longrightarrow \qquad [a_1 a_2] = [b_1 b_2]$$

Pre reláciu ekvivalencie sú však podmienky $[x] = [y]$ a $x \sim y$ ekvivalentné. Po prepísaní teda vidíme, že uvedená podmienka je ekvivalentná s takouto implikáciou.

$$\left.\begin{aligned}a_1 &\sim b_1\\ a_2 &\sim b_2\end{aligned}\right\} \qquad \Longrightarrow \qquad a_1 a_2 \sim b_1 b_2$$

To je ale presne podmienka (12.3) z definície kongruencie.[2]

(ii): Ak už vieme, že sme dostali binárnu operáciu, tak ostatné časti dôkazu sú už vcelku priamočiare.[3]

Na overenie asociatívnosti stačí skontrolovať, že platí:

$$\begin{aligned}[a] \cdot ([b] \cdot [c]) = [a] \cdot ([b \cdot c]) = [a \cdot (b \cdot c)] = \\ = [(a \cdot b) \cdot c] = [a \cdot b] \cdot [c] = ([a] \cdot [b]) \cdot [c]\end{aligned}$$

Neutrálnym prvkom je trieda $[e]$:

$$\begin{aligned}[a] \cdot [e] = [a \cdot e] = [a]\\ [e] \cdot [a] = [e \cdot a] = [a]\end{aligned}$$

A inverzným prvkom ku $[a]$ je $[a^{-1}]$:

$$\begin{aligned}[a] \cdot [a^{-1}] = [a \cdot a^{-1}] = [e]\\ [a^{-1}] \cdot [a] = [a^{-1} \cdot a] = [e]\end{aligned}$$

V dôkaze tejto časti sme v skutočnosti využili lemu 12.6.5. Naschvál som nezdôraznil, kde je táto lema potrebná – a ktorá časť uvedeného dôkazu nie je celkom v poriadku tak, ako som ju napísal tu – pretože chcem, aby ste sa nad tým skúsili zamyslieť samostatne: úloha 12.6.1.

Dôkaz poslednej časti týkajúcej sa komutatívnosti je tiež veľmi priamočiary – úloha 12.6.2. □

Ďalšia vec, ktorú vidno vcelku ľahko, je existencia pomerne prirodzeného homomorfizmu z G do $G/\sim$.

{grpkong:TVRKANHOM}

Tvrdenie 12.6.7. *Nech $\sim$ je grupová kongruencia na grupa $(G, \cdot)$. Potom zobrazenie*

$$\begin{aligned}\varphi\colon G &\to G/\sim\\ \varphi\colon x &\mapsto [x]\end{aligned}$$

je surjektívny grupový homomorfizmus. Tento homomorfizmus budeme nazývať kanonický homomorfizmus.

[2] Z tohto dôkazu vidíme aj to, že podmienka (12.3) je veľmi prirodzená. Inak povedané, grupovú kongruenciu sme definovali presne takým spôsobom, aby sme na množine tried dostali dobre definovanú binárnu operáciu.

[3] Aj tu – a aj pri všetkých ostatných dôkazoch v tejto časti – platí, že sa oplatí nad dôkazom jednotlivých tvrdení zamyslieť samostatne.

Dôkaz. Vyplýva z definície homomorfizmu a z definície operácie na $G/\sim$.

$$\varphi(x \cdot y) = [x \cdot y] = [x] \cdot [y] = \varphi(x) \cdot \varphi(y)$$

Tiež pomerne ľahko vidno, že zobrazenie φ je surjektívne – každý prvok z $G/\sim$ má tvar $[x] = \varphi(x)$ pre nejaké $x \in G$. □

Tvrdenie 12.6.7 môžeme v istom zmysle obrátiť – každý surjektívny homomorfizmus grúp vyzerá „v podstate“ takto.

{grpkong:TVRHOMREL}
{grpkong:itHOMRELKONG}
Tvrdenie 12.6.8. *Nech $f\colon G \to G'$ je surjektívny grupový homomorfizmus. Potom:*
(i) *Relácia $\sim$ na množine G zadaná ako*

$$a \sim b \qquad \Leftrightarrow \qquad f(a) = f(b)$$

{grpkong:itHOMRELIZOM}
je grupová kongruencia.
(ii) *Pre faktorovú grupu platí*

$$G/\sim \cong G'.$$

Dôkaz. (i) Vieme, že takýmto spôsobom dostaneme reláciu ekvivalencie – úloha 12.1.3. Treba ešte ukázať, že relácia $\sim$ sa správa rozumne vzhľadom na grupovú operáciu. Skutočne však pre $a_1 \sim b_1$ a $a_2 \sim b_2$ máme $f(a_1) = f(b_1)$ aj $f(a_2) = f(b_2)$. Z čoho vyplýva

$$f(a_1a_2) \overset{(*)}{=} f(a_1)f(a_2) = f(b_1)f(b_2) \overset{(*)}{=} f(b_1b_2),$$

a teda

$$a_1a_2 \cong b_1b_2.$$

Symbolom $(*)$ sme vyznačili miesta, na ktorých sme využili fakt, že f je homomorfizmus. (A v prvej časti tvrdenia sme dokonca ani nikde nepotrebovali surjektívnosť.)

(ii) Chceli by sme nájsť izomorfizmus medzi $G/\sim$ a G'. Pokúsme sa ho zadefinovať ako

$$\varphi\colon G/\sim \to G'$$
$$\varphi\colon [a] \mapsto f(a)$$

Ako prvú vec si musíme rozmyslieť, či sme skutočne dostali zobrazenie – t.j. či φ je *dobre definované.*

Ak platí $[a] = [b]$, znamená to, že $a \sim b$, a teda

$$f(a) = f(b).$$

Čiže uvedený predpis naozaj definuje zobrazenie z $G/\sim$ do G'.

Zo surjektívnosti zobrazenia f vidíme, že aj φ je *surjektívne.*

Navyše φ je aj *injektívne,* z rovnosti $\varphi([a]) = \varphi([b])$ totiž vyplýva $f(a) = f(b)$, čo znamená, že $a \sim b$ a $[a] = [b]$.

Zostáva ukázať, že φ je homomorfizmus – úloha 12.6.3. □

Vidíme teda, že relácie kongruencie súvisia so surjektívnymi homomorfizmami. Chceli by sme sa dostať aj k súvisu s normálnymi podgrupami a k tomu, aký je vzťah medzi faktorizáciou podľa relácie kongruencie a faktorizáciou podľa normálnej podgrupy.

Najprv sa pozrime na to, ako vyzerá jadro kanonického homomorfizmu. Pomerne ľahko si uvedomíme, že platí:

$$\begin{aligned}\operatorname{Ker}\varphi &= \{a \in G; \varphi(a) = [e]\}\\ &= \{a \in G; [a] = [e]\}\\ &= \{a \in G; a \sim e\}\\ &= [a].\end{aligned}$$

Je to teda presne trieda neutrálneho prvku.

Už sme videli, že normálne podgrupy sú presne tie podgrupy, ktoré sa dajú dostať ako jadrá homomorfizmov. Dá sa teda očakávať, že trieda neutrálneho prvku môže hrať významnú úlohu.

{grpkong:TVRKONGNORM}

Tvrdenie 12.6.9. *Nech $(G, \cdot)$ je grupa $a \sim$ je kongruencia na grupe G. Potom $H = [e]$ je normálna podgrupa grupy G.*

Navyše platí

$$a \sim b \qquad \Leftrightarrow \qquad ab^{-1} \in H.$$

Dôkaz. H je podgrupa. Evidentne platí $e \in [e]$, a teda množina $[e]$ je neprázdna.

Ak $x, y \in [e]$ tak z (12.3) máme

$$x \cdot y \sim e \cdot e = e,$$

a teda aj $xy \in [e]$.

Lema 12.6.5 zabezpečí uzavretosť množiny $[e]$ vzhľadom na inverzné prvky.

H je normálna. Pre ľubovoľný prvok $x \in H$ a ľubovoľné $g \in G$ dostaneme

$$gxg^{-1} \sim geg^{-1} = e,$$

teda aj $gxg^{-1} \in H$.

Vzťah medzi $\sim$ a H. Ak platí $a \sim b$, tak máme aj

$$ab^{-1} \sim bb^{-1} = e,$$

a teda $ab^{-1} \in H$.

Obrátene, ak $ab^{-1} \in H$, znamená to, že

$$\begin{aligned}ab^{-1} &\sim e\\ (ab^{-1}) \cdot b &\sim e \cdot b\\ a &\sim b\end{aligned}$$

□

Vidíme teda, ako môžeme z kongruencie dostať normálnu podgrupu. Ale aj obrátene z normálnej podgrupy vieme dostať kongruenciu.

{grpkong:TVRNORMKONG}

Tvrdenie 12.6.10. *Nech $(G, \cdot)$ je grupa a H je normálna podgrupa grupy G. Potom predpis*

$$x \sim y \qquad \Leftrightarrow \qquad xy^{-1} \in H$$

dáva kongruenciu na grupe G.

Navyše pre túto kongruenciu platí $H = [e]$, t.j. podgrupa H je presne trieda neutrálneho prvku.

Dôkaz. Úloha 12.6.4. □

Keď poskladáme doteraz spomínané pojmy dokopy, tak vidíme, že oba doteraz spomenuté pojmy, ktoré sme nazvali faktorovou grupou – t.j. $G/\sim$ a G/H – vlastne predstavujú to isté.

Ak máme na grupe G nejakú reláciu kongruencie $\sim$, tak:

- Trieda $H = [e]$ je normálna podgrupa.
- Pre túto podgrupu máme $x \sim y \Leftrightarrow xy^{-1} \in H$.
- Pre každý prvok $x \in G$ platí $xH = [x]$, t.j. triedy rozkladu G podľa relácie $\sim$ sú presne ľavé (pravé) triedy rozkladu G podľa podgrupy H.
- Teda množiny $G/\sim$ a G/H sú presne rovnaké. A operácia definovaná vzťahmi (12.2) a (12.4) je presne tá istá operácia.

Podobne, ak začneme s normálnou podgrupou H, tak z nej vieme dostať uvedeným spôsobom reláciu kongruencie $\sim$ a opäť platí, že $G/\sim$ a G/H je presne tá istá grupa.

Máme teda viacero prístupov k tomu, ako sa dá pozerať na faktorové grupy resp. na normálne podgrupy:

- rozklad grupy podľa normálnej podgupy;
- rozklad grupy podľa grupovej kongruencie;
- obraz grupy v surjektívnom homomorfizme.

faktorová grupa	normálna podgrupa
$G/\sim$	$[e]$
G/H	H
$f\colon G \to G'$	$\operatorname{Ker} f$

Súčasne máme tri pohľady na normálne podgrupy.

- Sú to podgrupy, pre ktoré je ľavý a pravý rozklad rovnaký.
- Sú to presne tie podgrupy, ktoré vieme dostať z relácie kongruencie ako triedu neutrálneho prvku.
- Sú to práve tie podgrupy, ktoré sú jadrami grupových homomorfizmov.

Príklad 12.6.11. Pozrime na najjednoduchší príklad, ktorý sme spomínali – grupa $(\mathbb{Z}, +)$, relácia „byť kongruentný modulo n“, podgrupa $H = n\mathbb{Z}$. Pomerne ľahko sa dá skontrolovať, že všetky doteraz spomínané veci fungujú:

- Trieda $[0]$ je presne podgrupa $n\mathbb{Z}$.
- Z podgrupy $H = n\mathbb{Z}$ dostaneme reláciu určenú podmienkou $x - y \in n\mathbb{Z}$, čo je presne kongruencia modulo n.
- Trieda celého čísla x je presne $x + H = \{x + kn; k \in \mathbb{Z}\}$.
- Ak vytvoríme faktorovú grupu, tak dostaneme presne $\mathbb{Z}/n\mathbb{Z}$.

12.6.4 Faktorizácia v ďalších štruktúrach

Faktorizácii sa budeme venovať neskôr pre okruhy – v časti 13.2. Uvidíme, že okruh sa dá faktorizovať podľa ideálu – teda ideály tu hrajú podobnú úlohu ako normálne podgrupy pri faktorových grupách. Aj tu by sa dalo začať s okruhovými kongruenciami a vybudovať pojem faktorového okruhu pomocou nich. Okruhové kongruencie sú spomenuté v definícii 13.2.23 a v niektorých cvičeniach za kapitolou o faktorových okruhoch.

Z podobnosti týchto konštrukcií sa dá prísť na to, že kongruencie by sme boli schopní zadefinovať aj v iných situáciách. Ak pracujeme so štruktúrou určenou niekoľkými binárnymi operáciami na množine, tak môžeme zadefinovať kongruencie – sú to relácie ekvivalencie rešpektujúce všetky uvedené operácie. A potom môžeme definovať faktorovú štruktúru. Pomerne všeobecne sa takéto niečo dá zadefinovať pre *univerzálne algebry.*

Ako jeden špeciálny prípad môžeme spomenúť *pologrupy* (definícia 11.1.1). Môžete si všimnúť, že pri niektorých častiach v tejto kapitole sme vlastne nepotrebovali štruktúru grupy – iba binárnu operáciu na množine G. Takže by sme boli schopní zadefinovať reláciu kongruencie a pomocou nej faktorovú pologrupu.

Je tu však jeden podstatný rozdiel. V každej grupe máme jeden význačný prvok – neutrálny prvok. Pomocou neho sme sa teda vedeli dostať k normálnym podgrupám. (Normálna podgrupa je trieda ekvivalencie neutrálneho prvku. Okrem toho sme normálne podgrupy vedeli dostať ako vzory neutráneho prvku v homomorfizme.) V definícii pologrupy sa nijaký špeciálny prvok nevyskytuje – nevieme tu teda dostať podpologrupy, ktoré by v nejakom zmysle zodpovedali normálny podgrupám.

Iný typ faktorovej štruktúry, s ktorou by ste sa možno mohli stretnúť, sú faktorové boolovské algebry.

Okrem toho v časti 12.8 stručne spomíname aj faktorové vektorové priestory. Tu je situácia trochu iná než to, čo sme spomenuli vyššie – násobenie skalárom totiž nie je binárna operácia na V. Konštrukcia faktorového vektorového priestoru je však výrazne jednoduchšia než to, čo sme robili pre grupy. (Faktorizovať sa dá podľa každého podpriestoru.)

Cvičenia

{grpkong:ULOINVDOBREDEF}

Úloha 12.6.1. Akým spôsobom sa v dôkaze existencie inverzného prvku faktorovej grupy v tvrdení 12.6.6 využíva fakt, že relácia kongruencie zachováva inverzné prvky (t.j. lema 12.6.5)?

{grpkong:ULOKOMUT}

Úloha 12.6.2. Ukážte, že ak G je komutatívna grupa, tak aj faktorová grupa $G/\sim$ je komutatívna.

{grpkong:ULOHOMREL}

Úloha 12.6.3. Dokážte, že zobrazenie φ z dôkazu druhej časti tvrdenia 12.6.8 je skutočne homomorfizmus.

{grpkong:ULONORMKONG}

Úloha 12.6.4. Ukážte, že pre ľubovoľnú normálnu podgrupu H grupy G predpis

$$x \sim y \qquad \Leftrightarrow \qquad xy^{-1} \in H$$

definuje reláciu kongruencie na grupe G a že pre túto reláciu platí $H = [e]$.

Úloha 12.6.5. Nech G je grupa a $R \subseteq G \times G$ je relácia na množine G. Ukážte, že R je (grupová) kongruencia práve vtedy, keď R je podgrupa grupy $G \times G$.

{grpkong:ULOPRIENIK}

Úloha 12.6.6. Nech pre každé $i \in I$ je R_i relácie grupová kongruencia na G. Dokážte, že aj relácia $R = \bigcap_{i \in I} R_i$ je kongruencia na grupe G.

12.7 Komutátor a komutant*

Skúsme sa pozrieť na problém, kedy je faktorová grupa komutatívna. Ľahko dostaneme nasledujúce kritérium.

Lema 12.7.1. *Ak G je grupa a H je jej normálna podgrupa, tak G/H je komutatívna práve vtedy, keď pre ľubovoľné $a, b \in G$ platí*

$$a^{-1}b^{-1}ab \in H.$$

Dôkaz. Grupa G/H je komutatívna práve vtedy, keď pre ľubovoľné $a, b \in G$ platí $(ab)H = (ba)H$. Podľa lemy 12.2.5 je to ekvivalentné s podmienkou

$$(ba)^{-1}ab = a^{-1}b^{-1}ab \in H.$$

□

Definícia 12.7.2. Nech G je grupa. Pre $a, b \in G$ nazývame prvok

$$[a, b] = a^{-1}b^{-1}ab$$

sa *komutátor* prvkov a a b.

Podgrupa generovaná všetkými komutátormi sa nazýva *komutant* grupy G a označuje sa ako $[G, G]$.

Poznamenajme, že množina všetkých komutátorov ešte nemusí tvoriť podgrupu, hoci nie je celkom ľahké nájsť kontrapríklad. (Najmenší možný kontrapríklad je 96-prvková grupa [Rot, Exercise 2.43]).

Ukážeme, že komutant je vždy normálna podgrupa. Na to sa nám bude hodiť explicitný popis podgrupy generovanej danou množinou.

Lema 12.7.3. *Ak G je grupa a $A \subseteq G$, tak podgrupa $[A]$ generovaná množinou A pozostáva práve z prvkov tvaru*

$$a_1^{\varepsilon_1} a_2^{\varepsilon_2} \dots a_n^{\varepsilon_n}$$

pre $n \in \mathbb{N}$, $a_i \in A$ a $\varepsilon_i \in \{\pm 1\}$. (V prípade, že $n = 0$, pod týmto súčinom chápeme neutrálny prvok.)

Dôkaz. □

Tvrdenie 12.7.4. *Nech G je grupa. Jej podgrupa $[G, G]$ je normálna.*

Dôkaz. □

12.8 Faktorové vektorové priestory *

{dopln:SECTFAKTVP}

Keď sme zvládli faktorové grupy, môžeme si uvedomiť, že podobná konštrukcia bude fungovať aj pre vektorové priestory.

V tomto prípade je situácia výrazne jednoduchšia – vektorový priestor sa dá faktorizovať podľa ľubovoľného podpriestoru. (Pri faktorových grupách sme nemohli použiť ľubovoľnú podgrupu – fungovalo to iba pre normálne podgrupy. Podobne pri faktorových okruhoch sa nedá použiť ľubovoľný podokruh – faktorizovať sa dá podľa ideálu.)

Definícia 12.8.1. Nech $(V, +, \cdot)$ je vektorový priestor nad poľom F a S je jeho podpriestor. Triedy rozkladu grupy $(V, +)$ podľa podgrupy S budeme v tomto prípade označovať ako $\vec{\alpha} + S$ pre $\vec{\alpha} \in V$. Potom $(V/S, +)$ s operáciou $(\vec{\alpha} + S) + (\vec{\beta} + S) = (\vec{\alpha} + \vec{\beta}) + S$ tvorí komutatívnu grupu. Ukážeme, že aj násobenie dané predpisom

$$c.(\vec{\alpha} + S) = (c.\vec{\alpha}) + S$$

pre $c \in F$ a $\vec{\alpha} \in V$ je dobre definované a V/S spolu s týmito operáciami tvorí vektorový priestor nad poľom F. Tento vektorový priestor nazývame *faktorový vektorový priestor V* podľa S.

Dôkazy faktov, ktoré sme spomínali v predchádzajúcej definícii sú pomerne jednoduché a mohli by sme ich ponechať ako cvičenie, pre úplnosť ich však aspoň naznačíme.

Dôkaz. Operácia $\cdot\colon F \times V/S \to V/S$ *je dobre definovaná.* Ak $\vec{\alpha} + S = \vec{\beta} + S$, tak $\vec{\alpha} - \vec{\beta} \in S$. Pretože S je podpriestor, platí potom aj $c(\vec{\alpha} - \vec{\beta}) = c\vec{\alpha} - c\vec{\beta} \in S$, čiže

$$c\vec{\alpha} + S = c\vec{\beta} + S.$$

V/S s uvedenými operáciami tvorí vektorový priestor nad F. Vieme, že V/S je komutatívna grupa.

Z ostatných podmienok vystupujúcich v definícii vektorového priestoru overme na ukážku jednu, všetky ostatné sa overia veľmi podobne.

Nech napríklad $\vec{\alpha}, \vec{\beta} \in V$ a $c \in F$. Potom

$$c[(\vec{\alpha} + S) + (\vec{\beta} + S)] = c[(\vec{\alpha} + \vec{\beta}) + S] = c(\vec{\alpha} + \vec{\beta}) + S = (c\vec{\alpha} + c\vec{\beta}) + S = (c\vec{\alpha} + S) + (c\vec{\beta} + S).$$

□

Príklad 12.4.3, t.j. faktorová grupa grupy $G = (\mathbb{R}\times\mathbb{R}, +)$ podľa podgrupy $H = \{(x, x); x \in \mathbb{R}\}$ je súčasne aj príkladom faktorového vektorového priestoru (keďže G je súčasne vektorový priestor a H je jeho podpriestor).

Kapitola 13

Okruhy a polia

13.1 Okruhy (a súvisiace pojmy)

Definícia 13.1.1. Trojicu $(R, +, \cdot)$ nazývame *okruh* ak $+$ a $\cdot$ sú binárne operácie na množine R také, že

(i) $(R, +)$ je komutatívna grupa,

(ii) operácia $\cdot$ je asociatívna[1],

$$(\forall a, b, c \in R) \qquad a \cdot (b \cdot c) = (a \cdot b) \cdot c$$

(iii) pre operácie $+$ a $\cdot$ platia *distributívne zákony*

$$(\forall a, b, c \in R) \qquad a \cdot (b + c) = a \cdot b + a \cdot c$$
$$(\forall a, b, c \in R) \qquad (b + c) \cdot a = b \cdot a + c \cdot a$$

Neutrálny prvok operácie $+$ budeme označovať 0. Podobne ako sme to robili pre polia, inverzný prvok k prvku a vzhľadom na operáciu $+$ budeme označovať $-a$. Označenie $b - a$ bude znamenať $b + (-a)$.

Ak je navyše operácia $\cdot$ komutatívna, t.j.

$$(\forall a, b \in R) \qquad a \cdot b = b \cdot a,$$

tak $(R, +, \cdot)$ voláme *komutatívny okruh.*

Ak existuje neutrálny prvok e operácie $\cdot$ a súčasne $e \neq 0$ (ako sme sa dohodli, 0 označuje neutrálny prvok operácie $+$), tak tento neutrálny prvok označujeme 1 a hovoríme že, že $(R, +, \cdot)$ je (*komutatívny*) *okruh s jednotkou.*[2]

Poznámka 13.1.2. Označenie pre operáciu $\cdot$ obvykle vynechávame, čiže namiesto $a \cdot b$ častejšie budeme používať označenie ab.

[1] t.j. $(R, \cdot)$ je pologrupa

[2] Prípad, že neutrálny prvok oboch operácií je ten istý, ktorý sme z tejto definície vylúčili, nastane iba pre jednoprvkový okruh $\{0\}$.
Z minulého semestra vieme, že jednotka v okruhu musí byť jednoznačne určená – tvrdenie 3.1.7.
V niektorých učebniciach sa v definícii okruhu s jednotkou nepožaduje podmienka $1 \neq 0$, potom sa však táto podmienka objaví ako jeden z predpokladov vo väčšine viet, ktoré o okruhoch s jednotkou dokazujeme, preto sme tu zvolili túto formu definície.

Takisto, keď budú uvažované binárne operácie jasné z kontextu, budeme písať stručne R namiesto $(R, +, \cdot)$.

Pri grupách sme spomínali aditívny a multiplikatívny zápis – v okruhu vždy pre operáciu $+$ používame aditívny a pre operáciu $\cdot$ multiplikatívny zápis. Teda použitie operácie viackrát na ten istý prvok označíme ako $n \times a$ pre operáciu $+$ a a^n pre operáciu $\cdot$ (kde $n \in \mathbb{N} \setminus \{0\}$).

Príklad 13.1.3. $(\mathbb{Z}, +, \cdot)$ – celé čísla s obvyklým sčitovaním a násobením tvoria komutatívny okruh s jednotkou.
$(\mathbb{Z}_n, \oplus, \odot)$ – množina $\mathbb{Z}_n = \{0, 1, \ldots, n-1\}$ so sčitovaním modulo n tvorí komutatívny okruh s jednotkou.

Príklad 13.1.4. Príklad komutatívneho okruhu, ktorý nemá jednotku: $(2\mathbb{Z}, +, \cdot)$.

Dôkaz nasledujúcej lemy ponechávame ako cvičenie, keďže je veľmi podobný dôkazom, ktoré sme robili pre polia.

Lema 13.1.5. *Nech $(R, +, \cdot)$ je okruh, $a, b \in R$. Potom platí*

$$\begin{aligned} 0a = a0 &= 0 \\ a(-b) = -ab &= (-a)b \\ (-a)(-b) &= ab \end{aligned}$$

{okr:PRZxZ}

Príklad 13.1.6. Na množine $\mathbb{Z} \times \mathbb{Z}$ definujeme operácie $+$ a $\cdot$ ako sčitovanie a násobenie po zložkách, t.j.

$$\begin{aligned} (a, b) + (a', b') &= (a + a', b + b'), \\ (a, b)(a', b') &= (aa', bb'). \end{aligned}$$

Potom $(\mathbb{Z} \times \mathbb{Z}, +, \cdot)$ je komutatívny okruh s jednotkou. (Jednotka je dvojica $(1, 1)$, nula je dvojica $(0, 0) = 0$.)

Všimnime si, že $(1, 0) \cdot (0, 1) = (0, 0)$, teda v okruhu môže byť súčin nenulových prvkov rovný nule.

Predchádzajúci príklad možno jednoducho zovšeobecniť:

{okr:PRSUCIN}

Príklad 13.1.7. Ak $(R_1, +, \cdot)$ a $(R_2, +, \cdot)$ sú okruhy, tak $R_1 \times R_2$ tvorí s operáciami definovanými po zložkách

$$\begin{aligned} (a_1, a_2) + (b_1, b_2) &= (a_1 + b_1, a_2 + b_2) \\ (a_1, a_2) \cdot (b_1, b_2) &= (a_1 \cdot b_1, a_2 \cdot b_2) \end{aligned}$$

tiež okruh.

Podobne, ak pre každé $i \in M$ je $(R_i, +, \cdot)$ okruh, tak aj množina[3] $\prod\limits_{i \in M} R_i = \{f \colon M \to \bigcup\limits_{i \in M} R_i \mid (\forall i \in M)(f(i) \in R_i)\}$ tvorí s operáciami definovanými po zložkách

$$\begin{aligned} (f + g)(i) &= f(i) + g(i) \\ (f \cdot g)(i) &= f(i) \cdot g(i) \end{aligned}$$

okruh.

V prípade, že všetky použité okruhy sú rovnaké, t.j. $R_i = R$ pre každé $i \in M$, budeme používať označenie R^M. Okruh R^M pozostáva zo všetkých zobrazení z M do R.

[3]Takto sa definuje karteziánsky súčin pre ľubovoľný (teda nie len konečný) počet množín. V prípade, že ste to nemali na žiadnom inom predmete, bude asi jednoduchšie, keď túto definíciu budete čítať tak, ako keby $R_i = R$ pre všetky $i \in M$ – pozri poznámku na konci tohoto príkladu.

Príklad 13.1.8. Dôležitý príklad okruhu tvoria matice $M_{n,n}(F)$ typu $n \times n$ nad poľom F spolu s násobením matíc. To, že sčitovanie a násobenie matíc spĺňajú podmienky z definície okruhu, sme ukázali v minulom semestri. Tento okruh má jednotku, je ňou jednotková matica I. Tento okruh nie je komutatívny.

Definícia 13.1.9. Nech $(R, +, \cdot)$ je okruh a $S \subseteq R$ je neprázdna podmnožina množiny R. Hovoríme, že S je *podokruh* okruhu R, ak pre ľubovoľné $a, b \in S$ platí $a - b \in S$, $ab \in S$.

$$a, b \in S \qquad \Rightarrow \qquad a - b \in S, ab \in S$$

Inými slovami, podokruh je podgrupa grupy $(R, +)$, ktorá je navyše uzavretá vzhľadom na násobenie.

Pomerne jednoducho sa dá overiť, že platí

Tvrdenie 13.1.10. *Nech $(R, +, \cdot)$ je okruh a $S \subseteq R$, $S \neq \emptyset$. Množina S je podokruh okruhu $(R, +, \cdot)$ práve vtedy, keď S s operáciami $+$ a $\cdot$ zúženými na množinu S tvorí okruh.*

Príklad 13.1.11. $2\mathbb{Z}$ je podokruh $(\mathbb{Z}, +, \cdot)$.

$\mathbb{N}$ nie je podokruh $(\mathbb{Z}, +, \cdot)$ (je uzavretý na násobenie a súčet, nie však na rozdiel).

{okr:PRC01}

Príklad 13.1.12. Uvažujme zobrazenia z uzavretého intervalu $\langle 0, 1\rangle$ do $\mathbb{R}$. Z matematickej analýzy vieme, že rozdiel a súčin spojitých funkcií je opäť spojitá funkcia. Vďaka tomu spojité funkcie $f\colon \langle 0, 1\rangle \to \mathbb{R}$ tvoria, so sčitovaním a násobením funkcií po bodoch, podokruh okruhu $\mathbb{R}^{\langle 0,1\rangle}$. Tento okruh označujeme $C(0, 1)$.

Definícia 13.1.13. Ak v okruhu $(R, +, \cdot)$ neexistujú prvky a, b také, že $a, b \neq 0$ a

$$ab = 0,$$

tak hovoríme, že R je *okruh bez deliteľov nuly* (alebo tiež, že R nemá delitele nuly).

Ak $(R, +, \cdot)$ je komutatívny okruh s jednotkou bez deliteľov nuly, hovoríme, že $(R, +, \cdot)$ je *obor integrity.*

Fakt, že R je okruh bez deliteľov nuly môžeme vyjadriť pomocou nasledovnej implikácie[4]

$$(\forall a, b \in R) \qquad ab = 0 \Rightarrow a = 0 \vee b = 0.$$

Príklad okruhu, ktorý nie je oborom integrity, je okruh $\mathbb{Z} \times \mathbb{Z}$ z príkladu 13.1.6. Dokonca ľubovoľný okruh tvaru $R_1 \times R_2$ (pozri príklad 13.1.7), kde ani jeden z okruhov R_1, R_2 nie je nulový, nám dáva takýto príklad.

Ľahko sa overí, že v okruhu bez deliteľov nuly môžeme krátiť nenulovými prvkami:

{okr:TVROIKRATENIE}

Tvrdenie 13.1.14. *Nech R je okruh bez deliteľov nuly a $a, b, c \in R$. Ak $a \neq 0$ a platí $ab = ac$, tak $b = c$.*

Dôkaz. Z rovnosti $ab = ac$ dostaneme pomocou distributívnosti $a(b - c) = 0$. Keďže $a \neq 0$, máme $b - c = 0$, a teda $b = c$. □

{okr:DEFTEL}

Definícia 13.1.15. Okruh R s jednotkou nazývame *telesom,* ak ku každému nenulovému prvku $a \in R \setminus \{0\}$ existuje inverzný prvok vzhľadom na násobenie, t.j.

$$(\forall a \in R \setminus \{0\})(\exists b \in R) \qquad ab = ba = 1$$

Komutatívne teleso voláme *pole.*

[4] Je to negácia výroku $(\exists a, b \in \mathbb{R})\ \ ab = 0 \wedge (a \neq 0 \wedge b \neq 0)$.

Tvrdenie 13.1.16. *Každé teleso je okruh bez deliteľov nuly.*
Každé pole je oborom integrity.

Dôkaz. Nech R je teleso a pre $a, b \in R$ platí $ab = 0$. Predpokladajme, že $a \neq 0$. Potom existuje $c \in R$ taký, že $ca = 1$. Z toho dostaneme

$$b = 1b = cab = c0 = 0,$$

čiže $b = 0$. Podobne, z predpokladu $b \neq 0$ by sme dostali $a = 0$.
Druhá časť tvrdenia ľahko vyplýva z prvej časti. □

Definícia 13.1.15 vlastne hovorí, že ak $(R, +, \cdot)$ je okruh a navyše $(R \setminus \{0\}, \cdot)$ je grupa, ide o teleso. Ak je to komutatívna grupa, ide o pole. Táto definícia poľa je teda ekvivalentná s definíciou 3.3.1, ktorú sme uviedli v minulom semestri. Z minulého semestra poznáme veľa príkladov polí – $\mathbb{C}$, $\mathbb{R}$, $\mathbb{Q}$ s obvyklým sčitovaním a násobením, $(\mathbb{Z}_p, \oplus, \odot)$ pre ľubovoľné prvočíslo p.

Príkladom telesa, ktoré nie je poľom (t.j. nekomutatívneho telesa) sú kvaternióny. Viac sa o nich môžete dozvedieť v [KGGS, Kapitola 4.7].

Cvičenia

Úloha 13.1.1. Zistite (a svoje tvrdenie zdôvodnite) ktoré z uvedených vlastností sa z okruhu R prenesú na uvedené konštrukcie: [5]

	$R \times R$	R/I	R^M	podokruh
pole				
obor integrity				
nemá delitele nuly				
má delitele nuly				
komutatívny okruh				
okruh s jednotkou				

Úloha 13.1.2. Je každý podokruh poľa okruh bez deliteľov nuly? Je každý podokruh poľa obsahujúci 1 oborom integrity?

{okrcvic:ULOSQRT2}

Úloha 13.1.3. Zistite, či nasledujúce množiny tvoria podokruhy poľa $(\mathbb{C}, +, \cdot)$. Zistite, ktoré z nich sú navyše poliami.
a) $\mathbb{Z}[\sqrt{2}] = \{a + b\sqrt{2}; a, b \in \mathbb{Z}\}$
b) $\mathbb{Q}[\sqrt{2}] = \{a + b\sqrt{2}; a, b \in \mathbb{Q}\}$

Úloha 13.1.4. Zistite, či nasledujúce množiny tvoria podokruhy poľa $(\mathbb{Q}, +, \cdot)$. Sú niektoré z nich polia?
a) Všetky zlomky také, že v základnom tvare je menovateľ nepárne číslo.
b) Všetky zlomky také, že v základnom tvare je menovateľ párne číslo.
c) Všetky zlomky také, že v základnom tvare je čitateľ nepárne číslo.
d) Všetky zlomky také, že v základnom tvare je čitateľ párne číslo.
e) Všetky druhé mocniny racionálnych čísel.

[5] Označenie R/I označuje faktorový okruh okruhu R podľa ideálu I. O faktorových okruhoch sa dozviete v nasledujúcej podkapitole, čiže tento stĺpec zatiaľ nechajte nevyplnený a vráťte sa k nemu neskôr.

Úloha 13.1.5. Dokážte: Ak R je obor integrity a $x^2 = 1$, tak $x = 1$ alebo $x = -1$.

Platí takéto tvrdenie aj za predpokladu, že R je komutatívny okruh s jednotkou? (T.j. ak vynecháme predpoklad, že R nemá delitele nuly.)

Úloha 13.1.6. Ak R je okruh bez deliteľov nuly a $ab = 1$, tak aj $ba = 1$.

Úloha 13.1.7. Nech $(R, +, \cdot)$ je okruh. Definujme binárnu operáciu $*$ ako $a * b = b \cdot a$. Dokážte, že aj $(R, +, *)$ je okruh.

Úloha 13.1.8. Dokážte, že $\{(r, r); r \in R\}$ je podokruh okruhu $R \times R$. Je tento podokruh izomorfný s okruhom R?

{okrcvic:ULOPODOKJEDN}

Úloha 13.1.9. Zistite, či S je podokruhom R, a tiež či v okruhoch S a R existuje jednotka (a ak áno, tak či je v oboch prípadoch rovnaká).
a) $R = \mathbb{Z}_6$, $S = 2\mathbb{Z}_3 = \{0, 2, 4\}$.
b) $R = A \times A$ a $S = A \times \{0\}$, kde A je ľubovoľný okruh s jednotkou.
c) $R = M_{2,2}(\mathbb{R})$, t.j. reálne matice rozmerov 2×2 a $S = \{\left(\begin{smallmatrix} a & 0 \\ 0 & 0 \end{smallmatrix}\right); a \in \mathbb{R}\}$.

Úloha 13.1.10. Nech R je okruh. Centrom okruhu R nazveme množinu $Z = \{z \in R; (\forall r \in R) zr = rz\}$. Ukážte, že:
a) Z je podokruh R.
b) Ak R má jednotku, tak $1 \in Z$.
c) Centrum telesa je pole.

Úloha 13.1.11*. Nech $(R, +, \cdot)$ je okruh s jednotkou. Ak existuje inverzný prvok vzhľadom na operáciu $\cdot$ k $1 - ab$, tak existuje aj inverzný prvok k $1 - ba$.

13.2 Homomorfizmy, ideály a faktorové okruhy

{ide:SECTFAKTOKR}

Definícia 13.2.1. Nech $(R, +, \cdot)$, $(S, +, \cdot)$ sú okruhy. Zobrazenie $f\colon R \to S$ nazývame *homomorfizmus,* ak platí

$$f(a + b) = f(a) + f(b),$$
$$f(ab) = f(a)f(b).$$

Surjektívny homomorfizmus nazývame *epimorfizmus,* injektívny homomorfizmus nazývame *monomorfizmus* a bijektívny homomorfizmus nazývame *izomorfizmus.* Ak existuje izomorfizmus medzi $(R, +, \cdot)$ a $(S, +, \cdot)$, hovoríme, že okruhy R a S sú izomorfné a píšeme $R \cong S$.

Pretože oba tieto pojmy používame aj pre grupy, občas sa vyskytne situácia, že budeme potrebovať rozlíšiť, či hovoríme o homomorfizme (izomorfizme) grúp alebo okruhov. V takomto prípadne použijeme termín *grupový homomorfizmus* (*izomorfizmus*) alebo *okruhový homomorfizmus* (*izomorfizmus*).

Dôkaz nasledujúceho tvrdenia vynechávame – je skoro identický s dôkazom analogického tvrdenia pre grupy.

Tvrdenie 13.2.2. *Zloženie homomorfizmov je homomorfizmus. Zloženie izomorfizmov je izomorfizmus.*

Príklad 13.2.3. Jednoduché príklady homomorfizmov:
$f\colon \mathbb{Z} \to \mathbb{Z}_n$, $f\colon k \mapsto k \bmod n$
$g\colon \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}$, $g\colon (a, b) \mapsto a$
$h\colon \mathbb{C} \to \mathbb{C}$, $h\colon a + bi \mapsto a - bi$

Opäť, podobne ako pri grupách, existencia izomorfizmu medzi dvoma okruhmi znamená, že tieto okruhy sú z hľadiska teórie okruhov rovnaké – nie sú rozlíšiteľné pomocou pojmov definovaných pre ľubovoľné okruhy. (Obe operácie pracujú rovnako, len prvky sú inak pomenované a izomorfizmus je bijektívne zobrazenie, ktoré poskytuje „slovník“ na preklad medzi týmito dvoma pomenovaniami.)

Túto myšlienku je možné použiť aj keď chceme ukázať, že nejaká množina s danými binárnymi operáciami tvorí okruh – nájdeme bijekciu medzi touto množinou a nejakým známym okruhom, ktorá zachováva operácie.

{ide:PRKOMPLMAT}

Príklad 13.2.4. Uvažujme podmnožinu S okruhu $M_{2,2}(\mathbb{R})$ tvorenú maticami tvaru

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix},$$

kde $a, b \in \mathbb{R}$.

Overme najprv, že ide o podokruh. Zrejme rozdiel 2 matíc takéhoto tvaru má opäť uvedený tvar. Pre súčin máme

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix} \begin{pmatrix} c & d \\ -d & c \end{pmatrix} = \begin{pmatrix} ac - bd & ad + bc \\ -(ad + bc) & ac - bd \end{pmatrix}, \tag{13.1}$$

{ide:EQKOMPLMAT}

čiže súčin matíc z S opäť patrí do S.

Definujme teraz zobrazenie $f\colon S \to \mathbb{C}$ predpisom

$$f\colon \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \mapsto a + bi.$$

Z rovnosti (13.1) vidíme, že pre $A = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}, B = \begin{pmatrix} c & d \\ -d & c \end{pmatrix} \in S$ máme

$$f(AB) = (ac - bd) + (ad + bc)i = (a + bi)(c + di) = f(A)f(B).$$

Overenie, že f zachováva súčty je jednoduché, takže ide o okruhový homomorfizmus.

Navyše, f je bijekcia, je to teda izomorfizmus medzi okruhom A a poľom komplexných čísel.

Toto zobrazenie by sme mohli použiť napríklad na dôkaz, že komplexné čísla tvoria okruh; alebo tiež na dôkaz, že A je podokruh (ak by sme už mali dokázané, že $\mathbb{C}$ je okruh; t.j. stačilo by nám overiť, že sa zachovávajú operácie). Vďaka tomu, že sme našli izomorfizmus medzi uvedenými dvoma okruhmi, hneď vieme, že A je pole – aj bez toho, že by sme to museli overovať priamym výpočtom.

Môžeme si položiť otázku, či sa na túto maticovú reprezentáciu komplexných čísel dá prísť aj nejakým priamočiarym spôsobom, bez toho, aby nám ho niekto povedal, alebo aby sme ho „uhádli“.

Skúsme sa, pre dané komplexné číslo $z = a + bi$ pozrieť na zobrazenie $f_z\colon \mathbb{C} \to \mathbb{C}$, $f_z\colon x \mapsto zx$ (toto je presne zobrazenie, ktoré sme priradili komplexnému číslu v Cayleyho vete 11.6.8, pozri tiež príklad 11.6.11). Ak komplexné číslo z vyjadríme v goniometrickom tvare ako $z = r(\cos\varphi + i\sin\varphi)$ tak z Moivrovej vety vieme, že násobenie číslom z znamená otočenie bodu (komplexné čísla chápeme ako body v rovine) okolo bodu 0 o uhol φ a potom jeho r-násobné zväčšenie.

Obidve tieto zobrazenia – otočenie aj natiahnutie – sú lineárne zobrazenia. Skúsme sa pozrieť na maticu takéhoto zobrazenia – nato stačí vedieť kam sa zobrazia vektory $(1, 0)$ a $(0, 1)$.

Vektor $(1,0)$ sa zobrazí otočením o uhol φ na $(\cos\varphi, \sin\varphi)$ a vektor $(1,0)$ na $(-\sin\varphi, \cos\varphi)$. To znamená, že otočeniu zodpovedá matica

$$\begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}$$

Ak ešte použijeme natiahnutie s koeficientom r, dostaneme maticu

$$\begin{pmatrix} r\cos\varphi & r\sin\varphi \\ -r\sin\varphi & r\cos\varphi \end{pmatrix} = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$$

čiže presne tú, ktorú sme použili v našom izomorfizme.

Príklad 13.2.5. Uvažujme lineárne zobrazenia z F^n do F^n. Označme množinu všetkých takýchto zobrazení A. Ukážeme, že $(A, +, \circ)$ je okruh s jednotkou ($+$ znamená obvyklé sčitovanie funkcií a $\circ$) je skladanie funkcií.

Namiesto toho, aby sme priamo overovali definíciu okruhu, uvažujme zobrazenie $f\colon A \to M_{n,n}(F)$, ktoré každému zobrazeniu priradí jeho maticu. Toto zobrazenie je bijektívne a navyše, keď ho berieme ako zobrazenie medzi $(A, +, \circ)$ a okruhom $(M_{n,n}(F), +, \cdot)$ rešpektuje binárne operácie (matica súčtu zobrazení je súčet matíc, matica súčinu zobrazení je súčin matíc). Keďže už vieme, že matice typu $n \times n$ tvoria okruh, vyplýva z toho, že aj $(A, +, \circ)$ je okruh. (Samozrejme, nie je ťažké overiť vlastnosti okruhu aj priamo, bez toho, aby sme si pomáhali maticami.)

Môžeme si ešte všimnúť, že podobné tvrdenie už neplatí, ak vezmeme ľubovoľné zobrazenia (t.j. nielen lineárne). Napríklad pre $F^n = \mathbb{R}^1$ a zobrazenia $f(x) = x^2$, $g(x) = x$, $h(x) = 1$ máme $f(g(x) + h(x)) = (x+1)^2 = x^2 + 2x + 1$, zatiaľčo $f(g(x)) + f(h(x)) = x^2 + 1$, čiže

$$f \circ (g + h) \neq f \circ g + f \circ h.$$

Ľahko sa dá ukázať, že jadro a obraz homomorfizmu musia byť podokruhy. Pri grupách sme videli, že nie každá podgrupa danej grupy môže byť jadrom nejakého homomorfizmu – túto vlastnosť mali len invariantné podgrupy. V prípade okruhov je zodpovedajúcim pojmom pojem ideálu.

Definícia 13.2.6. Nech R je okruh. Neprázdna podmnožina $I \subseteq R$ je *ideál* v okruhu R, ak platí

$$\begin{aligned} (\forall a, b \in I) \qquad & a - b \in I \\ (\forall a \in I)(\forall r \in R) \qquad & ar \in I, ra \in I \end{aligned}$$

t.j. ak je táto množina uzavretá vzhľadom na sčitovanie (prvkov z I) a násobenie ľubovoľným prvkom z R.

Inak povedané, ideál je taký podokruh, ktorý je uzavretý vzhľadom na násobenie všetkými prvkami z R.

Príklad 13.2.7. V každom okruhu máme ideály $\{0\}$ a R. Ideály I také, že $I \neq R$ voláme *vlastné.*

Pre ľubovoľné $k \in \mathbb{Z}$ je podmnožina $k\mathbb{Z} = \{kz; z \in \mathbb{Z}\}$ ideálom v okruhu $(\mathbb{Z}, +, \cdot)$.

V okruhu $R_1 \times R_2$ tvoria podmnožiny $R_1 \times \{0\}$ aj $\{0\} \times R_2$ ideály.

Príkladom podokruhu, ktorý nie je ideálom, je napríklad $\mathbb{Z}$ v $(\mathbb{R}, +, \cdot)$.

Často sa budú vyskytovať ideály určené jediným prvkom.

Definícia 13.2.8. Ak R je komutatívny okruh a $a \in R$, tak množina

$$(a) = \{ax; x \in R\}$$

je ideálom v R (úloha 13.2.4). Ideály takéhoto tvaru voláme *hlavné ideály.*

Nasledujúce pozorovanie je veľmi jednoduché, sformulujeme ho však do lemy, aby sme sa naň neskôr mohli odkazovať.

{ide:LMID1}

Lema 13.2.9. *Nech R je okruh s jednotkou a I je ideál v R. Potom $I = R$ práve vtedy, keď $1 \in I$.*

Dôkaz. Implikácia $\Rightarrow$ je úplne triviálna. Na dôkaz opačnej implikácie si stačí všimnúť, že pre ľubovoľné $c \in R$ máme

$$c = c.1$$

a ak $1 \in I$, tak aj $c.1$ patrí do ideálu I. □

Dôsledok 13.2.10. *Ak R je pole, tak jediné ideály v R sú $\{0\}$ a R.*

Dôkaz. Ak ideál I obsahuje prvok $a \neq 0$, tak k prvku a existuje inverzný prvok b, t.j. taký prvok, že $ba = 1$. Potom ale priamo z definície ideálu vyplýva, že aj $1 = ba \in I$, a teda $I = R$. □

Prvý krok na ceste k tomu, aby sme ukázali, že ideály majú pre okruhy podobnú úlohu ako normálne podgrupy pre grupy, je nasledujúca lema.

{ide:LMKER}

Lema 13.2.11. *Ak $\varphi\colon R \to S$ je homomorfizmus okruhov, tak jeho jadro* $\operatorname{Ker}\varphi$ *je ideál v R.*

Dôkaz. Ak $a \in \operatorname{Ker}\varphi$, znamená to, že $\varphi(a) = 0$. Potom pre ľubovoľné $x \in R$ máme

$$\varphi(ax) = \varphi(a)\varphi(x) = 0\varphi(x) = 0$$
$$\varphi(xa) = \varphi(x)\varphi(a) = \varphi(x)0 = 0$$

čiže aj $ax, xa \in \operatorname{Ker}\varphi$. □

Podobne ako pri grupách sme pre invariantné podgrupy boli schopní zadefinovať faktorovú grupu aj v tomto prípade vieme definovať faktorový okruh.

Ak na chvíľu zabudneme na operáciu $\cdot$, tak máme komutatívnu grupu $(R, +)$ a I je jej podgrupa. Pretože každá podgrupa komutatívnej podgrupy je normálna, máme potom faktorovú grupu $(R/I, +)$. Dôkaz nasledujúcej vety v podstate spočíva v overení, že sa dá pridať operácia $\cdot$ a že tak dostaneme okruh. Overenie týchto podmienok je vlastne jednoduchým mechanickým výpočtom, pričom používame iba definíciu ideálu a podmienku

$$a + I = b + I \qquad \Leftrightarrow \qquad a - b \in I,$$

ktorú poznáme z lemy 12.2.5. Takisto v ďalších výsledkoch o faktorových okruhoch sa budeme často odvolávať na to, čo už vieme o faktorových grupách; potom zostane overiť niektoré vlastnosti operácie $\cdot$.

Veta 13.2.12. *Nech $(R, +, \cdot)$ je ľubovoľný okruh a I je ideál v R. Ak na prvkoch faktorovej*[6] *grupy $(R, +)$ podľa podgrupy I*

$$R/I = \{a + I; a \in R\}$$

[6] Grupa $(R, +)$ je komutatívna, takže jej podgrupa I je invariantná. Má teda zmysel hovoriť o faktorovej grupe.

definujeme binárnu operáciu $\cdot$ ako

$$(a+I)\cdot(b+I)=(a\cdot b)+I,$$

tak je táto binárna operácia dobre definovaná a $(R/I,+,\cdot)$ je okruh. Tento okruh voláme faktorový okruh *R podľa I.*

Ak je okruh R komutatívny, tak aj R/I je komutatívny. Ak R je okruh s jednotkou a $I\neq R$, tak $1+I$ je jednotka faktorového okruhu R/I.

Dôkaz. Najprv ukážeme, že uvedená operácia je dobre definovaná. T.j. potrebujeme dokázať, že ak $a+I=a'+I$ a $b+I=b'+I$, tak aj $ab+I=a'b'+I$. Rovnosť $a+I=a'+I$ je však ekvivalentná s tým, že $a-a'\in I$ (lema 12.2.5), podobne druhú podmienku môžeme nahradiť podmienkou $b-b'\in I$.

Ak $a-a'\in I$, $b-b'\in I$, tak $ab-a'b'=a(b-b')+(a-a')b'\in I$. (Máme $a(b-b')\in I$, lebo $b-b'\in I$, podobne $(a-a')b'\in I$ lebo $a-a'\in I$, uvedený prvok je teda súčet dvoch prvkov z I.) Z $ab-a'b'\in I$ už vyplýva, že $ab+I=a'b'+I$.

Keď už vieme, že uvedený predpis definuje binárnu operáciu na R/I, zostáva overiť podmienky z definície okruhu. Vieme, že $(R/I,+)$ je grupa, navyše je aj komutatívna (lebo grupa R je komutatívna). Zostáva overiť asociatívnosť a distributívnosť. Máme

$$\begin{gathered}(a+I)((b+I)(c+I))=a(bc)+I=(ab)c+I=((a+I)(b+I))(c+I)\\(a+I)((b+I)+(c+I))=a(b+c)+I=(ab+ac)+I=(ab+I)+(ac+I)\\((b+c)+I)(a+I)=(b+c)a+I=(ba+ca)+I=(ba+I)+(ca+I)\end{gathered}$$

Úplne rovnako sa dokáže komutatívnosť R/I v prípade, že R je komutatívny a takisto, že $1+I$ je neutrálny prvok operácie $\cdot$. Podmienka $I\neq R$ zabezpečí, že $1=1-0\notin I$, t.j. $1+I\neq 0+I$ (v okruhu s jednotkou požadujeme aj aby $1\neq 0$.) □

Aj pre faktorové okruhy platí veta o izomorfizme.

{ide:VTIZOMOKR}

Veta 13.2.13 (Veta o izomorfizme)**.** *Ak $f\colon R\to R'$ je homomorfizmus okruhov, tak* Ker f *je ideál v okruhu R a faktorový okruh $R/\operatorname{Ker} f$ je izomorfný s podokruhom* Im f *okruhu R'.*

Dôkaz. Z lemy 13.2.11 vieme, že Ker f je ideál.

Označme $I=\operatorname{Ker} f$. Pretože $(R,+)$ je komutatívna grupa, jej podgrupa I je invariantná podgrupa. Potom (podľa vety o izomorfizme pre grupy) je zobrazenie $\varphi\colon R/I\to R'$ určené predpisom

$$\varphi\colon a+I\mapsto f(a)$$

dobre definované a je to injektívny grupový homomorfizmus. Zostáva teda len dokázať, že je to aj okruhový homomorfizmus, t.j. že zachováva aj operáciu $\cdot$. To však ľahko vyplýva z toho, že f je okruhový homomorfizmus:

$$\varphi(ab+I)=f(ab)=f(a)f(b)=\varphi(a+I)\varphi(b+I).$$

□

Postupom z predchádzajúceho dôkazu sa dá ukázať, že pre každý ideál I je zobrazenie $\varphi\colon R\to R/I$ určené predpisom

$$\varphi\colon a\mapsto a+I$$

okruhový homomorfizmus. Toto zobrazenie voláme *kanonický homomorfizmus.* Pre kanonický homomorfizmus platí $I=\operatorname{Ker}\varphi$.

Videli sme, že faktorový okruh komutatívneho okruhu je opäť komutatívny okruh a (s výnimkou prípadu $I = R$) dostaneme aj z okruhu s jednotkou znovu okruh s jednotkou. Otázka, či sa na faktorový okruh prenesie aj vlastnosť „byť oborom integrity“ alebo „byť poľom“ je o čosi komplikovanejšia.

Definícia 13.2.14. Ideál I v okruhu R sa nazýva prvoideál, ak pre ľubovoľné $a, b \in R$ také, že $a \cdot b \in I$ aspoň jeden z prvkov a, b patrí do I čiže ak platí

$$a \cdot b \in I \qquad \Rightarrow \qquad a \in I \vee b \in I.$$

Môžeme si všimnúť, že $\{0\}$ je prvoideál v R práve vtedy, keď R nemá delitele nuly.

Príklad 13.2.15. Pozrime sa na ideály v okruhu $(\mathbb{Z}, +, \cdot)$. Nie je príliš ťažké ukázať, že každý ideál v $\mathbb{Z}$ je hlavný – nebudeme sa tomu venovať na tomto mieste, lebo neskôr ukážeme všeobecnejší výsledok v tvrdení 13.4.16.

Otázka teda je, ktoré hlavné ideály v $\mathbb{Z}$ sú prvoideály.

Chceme sa pozrieť na ideál (p) pre nejaké celé číslo $p \neq 0$. Pýtame sa teda na to, kedy platí

$$ab \in (p) \Rightarrow a \in (p) \vee b \in (p).$$

Ideál p obsahuje presne násobky čísla p. Teda sa vlastne pýtame na platnosť podmienky

$$p \mid ab \Rightarrow p \mid a \vee p \mid b.$$

Vieme, že táto podmienka je splnená pre všetky prvočísla. Keď pracujeme v celých číslach, musíme pridať aj opačné čísla.

Naopak, pre zložené čísla táto podmienka neplatí. Ak $p = ab$ pre nejaké čísla $a, b \neq \pm 1$, tak máme $ab \in (p)$ ale a ani b do (p) nepatria.

Vidíme teda, že (p) je prvoideál v $\mathbb{Z}$ práve vtedy, keď $\pm p$ je prvočíslo (alebo 0).

{ide:VTOIPRV}

Veta 13.2.16. *Nech R je komutatívny okruh s jednotkou a I je ideál v R. Potom faktorový okruh R/I je oborom integrity práve vtedy, keď I je vlastný prvoideál.*

Dôkaz. $\boxed{\Rightarrow}$ Nech R/I je obor integrity. Z toho hneď vyplýva $1 + I \neq 0 + I$, a teda ideál I je vlastný. Využijeme fakt, že $I = \operatorname{Ker}\varphi$ pre kanonický homomorfizmus $\varphi\colon R \to R/I$; $\varphi(a) = a + I$. Z toho vyplýva, že ak $ab \in I$, tak

$$\varphi(ab) = \varphi(a)\varphi(b) = 0.$$

Pretože R/I je obor integrity, z predchádzajúcej rovnosti vyplýva, že $\varphi(a) = 0$ alebo $\varphi(b) = 0$, čiže $a \in I = \operatorname{Ker}\varphi$ alebo $b \in I = \operatorname{Ker}\varphi$.

$\boxed{\Leftarrow}$ Podobne ako v prvej časti využijeme surjektívny homomorfizmus $\varphi\colon R \to R/I$; $\varphi(a) = a + I$. Ak $x, y \in R/I$ sú také, že $xy = 0$ a $x = \varphi(a)$, $y = \varphi(b)$ (zo surjektívnosti vyplýva, že také $a, b \in R$ existujú) tak máme

$$\varphi(ab) = \varphi(a)\varphi(b) = xy = 0,$$

čiže $ab \in \operatorname{Ker}\varphi = I$. Pretože I je prvoideál, tak z toho vyplýva $a \in I = \operatorname{Ker}\varphi$ alebo $b \in I = \operatorname{Ker}\varphi$, čo však znamená, že

$$x = \varphi(a) = 0 \qquad \vee \qquad y = \varphi(b) = 0.$$

□

Príklad 13.2.17. Ako ilustráciu predchádzajúcej vety si môžeme všimnúť, že $\mathbb{Z}/(3) \cong \mathbb{Z}_3$ je pole, a teda aj obor integrity. Vcelku ľahko skontrolujeme, že (3) je maximálny ideál v $\mathbb{Z}$. (V skutočnosti môžeme číslo 3 nahradiť ľubovoľným prvočíslom.)

Ale napríklad $\mathbb{Z}/(4) \cong \mathbb{Z}_4$ nie je obor integrity (a (4) nie je prvoideál v $\mathbb{Z}$).

Definícia 13.2.18. Ideál I v okruhu R nazývame *maximálny,* ak $I \neq R$ a súčasne pre každý ideál J s vlastnosťnou $I \subseteq J \subseteq R$ platí $I = J$ alebo $J = R$.

Predchádzajúca definícia vlastne hovorí, že maximálne ideály sú práve maximálne prvky množiny vlastných ideálov okruhu R vzhľadom na usporiadanie $\subseteq$.

Poznámka 13.2.19. Bez dôkazu[7] spomeňme, že pre každý ideál I taký, že $I \neq R$ existuje maximálny ideál M obsahujúci I, t.j. $I \subseteq M$.

{ide:VTPOLEMAX}

Veta 13.2.20. *Nech R je komutatívny okruh s jednotkou a I je ideál v R. Potom faktorový okruh R/I je pole práve vtedy, keď I je maximálny ideál.*

Dôkaz. $\boxed{\Rightarrow}$ Predpokladajme, že R/I je pole. Potom musí platiť $0 + I \neq 1 + I$, čiže $1 \notin I$ a I je vlastný ideál.

Ďalej nech $I \subseteq J \subseteq R$. Predpokladajme, že $I \neq J$, teda existuje prvok $a \in J$ taký, že $a \notin I$. Potom $a + I \neq 0 + I$, čiže k $a + I$ existuje v poli R/I inverzný prvok. To znamená, že existuje $c \in R$ také, že

$$(ac) + I = 1 + I,$$

čiže $1 - ac \in I \subseteq J$. Potom z toho, že $ac \in J$ (lebo $a \in J$) a $1 - ac \in J$ vyplýva $1 \in J$ a $J = R$ (lema 13.2.9).

$\boxed{\Leftarrow}$ Nech I je maximálny ideál. Ak $a \notin I$ (čiže $a + I \neq 0 + I$), chceme ukázať, že k $a + I$ existuje v R/I inverzný prvok. Definujme

$$J = \{j + ca; j \in I, c \in R\}.$$

Overme najprv, že J je ideál. Skutočne, $(j + ca) - (j' + c'a) = (j - j') + (c - c')a$ a $j - j' \in I$, $c - c' \in R$ pre $j, j' \in J$, $c, c' \in R$. Ďalej $(j + ca)r = jr + car = jr + (cr)a$ a z $jr \in I$ dostávame, že $(j + ca)r \in I$ pre ľubovoľné $r \in R$.

Navyše, pre ideál J platí $I \subsetneqq J \subseteq R$. Pretože I je maximálny ideál, máme potom $J = R$, a teda $1 \in J$. To znamená, že existujú $c \in R$, $j \in I$ také, že $j + ca = 1$. Potom máme

$$ca - 1 \in I,$$
$$ca + I = 1 + I,$$

čiže $c + I$ je inverzný prvok vzhľadom na násobenie k $a + I$ v okruhu R/I. □

Pretože každé pole je oborom integrity, dokázali sme súčasne:

{ide:DOSMAXJEPRV}

Dôsledok 13.2.21. *V komutatívnom okruhu s jednotkou je každý maximálny ideál prvoideál.*

Príklad 13.2.22. Je prirodzené sa pýtať, či implikáciu v dôsledku 13.2.21 nie je možné obrátiť, teda či vieme nájsť príklad prvoideálu (v nejakom komutatívnom okruhu s jednotkou), ktorý nie je maximálny.

Z viet, ktoré sme práve ukázali a z toho, že $R/(0) \cong R$ ľahko vidno, že ak R je obor integrity, ktorý nie je poľom, tak (0) je prvoideál, ktorý nie je maximálny.

[7] Dôkaz vynechávame z toho dôvodu, že sa nedá urobiť pomocou vecí, ktoré ste sa zatiaľ učili. Obvyklý dôkaz je založený na Zornovej leme (resp. Axióme výberu). V prípade záujmu môžete dôkaz nájsť napríklad v [Sl1].

Teda napríklad ideál (0) v okruhu $(\mathbb{Z}, +, \cdot)$ je prvoideál, ktorý nie je maximálny. Ľahko to vidno aj priamo, bez odvolávania sa na predošlé vety. To, že ide o prvoideál, vidno z platnosti implikácie $ab = 0 \Rightarrow a = 0 \vee b = 0$ v okruhu $\mathbb{Z}$. Nie je maximálny, lebo napríklad $(0) \subsetneq (2) \subsetneq \mathbb{Z}$.

O trochu menej triviálny (aj keď veľmi podobný) príklad máte v úlohe 13.2.20.

V tvrdení 13.4.19 budeme vidieť, že okruhu $\mathbb{Z}$ by sme už nijaký iný kontrapríklad ako (0) nájsť nemohli.

Opäť, podobne ako v prípade grúp a normálnych podgrúp, zodpovedajú ideály kongruenciám na okruhu R (pozri úlohy 13.2.25, 13.2.26).

{ide:DEFKONG}

Definícia 13.2.23. Nech $(R, +, \cdot)$ je okruh. Relácia ekvivalencie E na R sa nazýva *kongruencia*, ak platí

$$aEa', bEb' \qquad \Rightarrow \qquad (a+b)E(a'+b'), (ab)E(a'b')$$

Aj tu by sa dali zopakovať všetky tvrdenia z časti 12.6 – v podstate bezo zmeny, len by sme všade nahradili výraz „normálna podgrupa“ výrazom ideál a výraz „faktorová grupa“ výrazom „faktorový okruh“. (A pracovali by sme s okruhovými kongruenciami namiesto grupových kongruencií.)

Aspoň niečo je sformulované v cvičeniach (úlohy 13.2.25 a 13.2.26). Dá sa však predpokladať, že čitateľ, ktorý dostatočne porozumel faktorizácii pri grupách bude schopní si zobrať všetky tvrdenia o grupových kongruenciách a samostatne si rozmyslieť, ako vyzerajú analogické tvrdenia pre okruhové kongruencie a ako by sa dali dokázať.

Cvičenia

{idecvic:ULOSYMDIF}

Úloha 13.2.1. Nech $X \neq \emptyset$ je ľubovoľná neprázdna množina. Dokážte, že potenčná množina $(P(X), \triangle, \cap)$ s operáciami $\triangle$ (symetrická diferencia množín) a $\cap$ (prienik množín) tvorí okruh. Nájdite izomorfizmus medzi týmto okruhom a okruhom $\mathbb{Z}_2^X$. (Poznámka: Bijekcia, ktorú nájdete v druhej časti, by sa dala použiť aj na dôkaz tvrdenia uvedeného v prvej časti.)

Úloha 13.2.2. Nech F je pole a $A \neq \emptyset$ je neprázdna množina. Dokážte, že v okruhu F^A (príklad 13.1.7) je každý ideál tvaru $M_p = \{f \in F^A; f(p) = 0\}$, kde p je nejaký prvok z A, maximálny. (Hint: Dá sa využiť veta 13.2.20. Ale dá sa postupovať aj priamo z definície.)

{idecvic:PRIENPODOKR}

Úloha 13.2.3. Prienik ľubovoľného systému podokruhov je podokruh. Prienik ľubovoľného systému ideálov je ideál.

{idecvic:ULOHI}

Úloha 13.2.4. Overte, že $(a) = \{ax; x \in R\}$ je ideál v komutatívnom okruhu R (teda hlavné ideály sú skutočne ideály.)

{idecvic:HOMSUCIN}

Úloha 13.2.5. Dokážte, že zobrazenie $f_1 \colon R_1 \times R_2 \to R_1$ určené predpisom $f_1(r_1, r_2) = r_1$ je homomorfizmus.

Dokážte, že pre každé $i \in I$ je zobrazenie $f_i \colon R^I \to R$ dané predpisom $f_i(g) = g(i)$ (pre ľubovoľné $g \colon I \to R$) je homomorfizmus.

Úloha 13.2.6. Nech $R \neq \{0\}$ je komutatívny okruh s jednotkou taký, že jediné ideály v R sú $\{0\}$ a R. Dokážte, že R je pole.

Úloha 13.2.7. Ak I_1 je ideál v okruhu R_1 a I_2 je ideál v okruhu R_2, tak podmnožina $I_1 \times I_2$ je ideál v okruhu $R_1 \times R_2$.

Úloha 13.2.8. Ak I_1, I_2 sú ideály v komutatívnom okruhu $(R, +, \cdot)$, tak aj
a) $I_1 + I_2 = \{a + b; a \in I_1, b \in I_2\}$ je ideál v R.
b) $I_1 \cdot I_2 = \{a_1b_1 + \cdots + a_nb_n; n \in \mathbb{N}, a_i \in I_1, b_i \in I_2\}$ je ideál v R.

Úloha 13.2.9. Nech $(G, *)$ je cyklická grupa, a je jej generátor, t.j. $G = [a]$. Ak definujeme operáciu $\cdot$ ako $a^k \cdot a^l = a^{k \cdot l}$ (pre ľubovoľné $k, l \in \mathbb{Z}$), tak $(G, *, \cdot)$ je okruh. Viete povedať (v závislosti od rádu generátora a) s akým okruhom je tento okruh izomorfný?

Úloha 13.2.10. Ak pre každé $n \in \mathbb{N}$ je I_n ideál v okruhu R a navyše platí $I_n \subseteq I_{n+1}$, tak aj zjednotenie $\bigcup_{i=1}^{\infty} I_i$ je ideál v R.

Úloha 13.2.11. Okruh R sa volá boolovský okruh, ak pre každé $a \in R$ platí $a^2 = a$. Dokážte, že každý boolovský okruh je komutatívny. (Boolovským okruhom je napríklad okruh z úlohy 13.2.1.)

Úloha 13.2.12. Dokážte, že okruhy $(2\mathbb{Z}, +, \cdot)$ a $(3\mathbb{Z}, +, \cdot)$ nie sú izomorfné.

Úloha 13.2.13. Nájdite všetky homomorfné obrazy okruhu $\mathbb{Z}$.

Úloha 13.2.14. Nájdite všetky homomorfizmy zo $\mathbb{Z}$ do $\mathbb{Z}_{30}$.

Úloha 13.2.15. Nájdite všetky homomorfizmy:
a) zo $\mathbb{Z}[\sqrt{2}]$ do $\mathbb{Z}[\sqrt{2}]$,
b) z $\mathbb{Q}[\sqrt{2}]$ d $\mathbb{Q}[\sqrt{2}]$.
(Tieto okruhy sú definované v úlohe 13.1.3.)

Úloha 13.2.16. Nájdite všetky homomorfizmy $\mathbb{Z} \times \mathbb{Z} \to \mathbb{Z} \times \mathbb{Z}$.

Úloha 13.2.17. Zistite, ktoré z nasledujúcich zobrazení sú homomorfizmy medzi okruhom A všetkých matíc typu 2×2 s celočíselnými koeficientami a okruhom $\mathbb{Z}$.
a) $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto a$
b) $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto a + d$ (*stopa* matice)
c) $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto ad - bc$ (determinant matice)

Úloha 13.2.18. Zistite, či tieto množiny tvoria ideály v okruhu $\mathbb{Z} \times \mathbb{Z}$:
a) $\{(a, a); a \in \mathbb{Z}\}$
b) $\{(2a, 2b); a, b \in \mathbb{Z}\}$
c) $\{(2a, 0); a \in \mathbb{Z}\}$
d) $\{(a, -a); a \in \mathbb{Z}\}$

{idecvic:ULOPRVNIEMAX}

Úloha 13.2.19. Zistite, s akými okruhmi sú izomorfné okruhy $\mathbb{Z}_{60}/(15)$, $\mathbb{Z}_{60}/(20)$, $\mathbb{Z}_{60}/(12)$.

Úloha 13.2.20. Ukážte, že $\mathbb{Z} \times \{0\}$ je ideál v okruhu $\mathbb{Z} \times \mathbb{Z}$ (s obvyklým sčitovaním a násobením). Ukážte, že je to prvoideál, ktorý nie je maximálny.

Úloha 13.2.21. Zistite, či dané ideály v okruhu $\mathbb{Z}[i] = \{a + bi; a, b \in \mathbb{Z}\}$ sú maximálne ideály/prvoideály.
a) $(1 + i) = \{(1 + i)z; z \in \mathbb{Z}[i]\}$
b) $(2) = \{2z; z \in \mathbb{Z}[i]\}$
c*) $(2 + i) = \{(2 + i)z; z \in \mathbb{Z}[i]\}$

decvic:BINOM}

Úloha 13.2.22. Nech R je komutatívny okruh s jednotkou. Dokážte, že v ňom platí binomická veta

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} \times a^k b^{n-k}.$$

Úloha 13.2.23*. a) Dokážte, že v okruhu $C(0,1)$ (príklad 13.1.12) je každý ideál tvaru $M_p = \{f \in C(0,1); f(p) = 0\}$ pre $p \in \langle 0,1\rangle$ maximálny.
b) Dokážte, že všetky maximálne ideály v $C(0,1)$ majú takýto tvar.

Úloha 13.2.24. Nájdite príklad takých okruhov $(R,+,\cdot)$, $(S,+,\cdot)$ a zobrazenie $f\colon R \to S$, že f je grupový homomorfizmus (medzi grupami $(R,+)$ a $(S,+)$), ale nie je to okruhový homomorfizmus.

{idecvic:KONGRNG}

Úloha 13.2.25$^+$. Nech R je okruh, I je ideál v R. Definujeme reláciu E na R ako $aEb \Leftrightarrow a - b \in I$. Dokážte, že E je (okruhová) kongruencia (definícia 13.2.23).

Obrátene ak E je ľubovoľná kongruencia na R, tak trieda ekvivalencie $[0]_E$ je ideál v R.

{idecvic:KONGHOM}

Úloha 13.2.26$^+$. Nech $(R,+,\cdot)$ je okruh.
a) Ak $f\colon R \to S$ je homomorfizmus, tak relácia E na množine R daná predpisom $xEy \Leftrightarrow f(x) = f(y)$ je kongruencia (pozri úlohu 12.1.3).
b) Ak E je kongruencia na R, tak na množine R/E tried ekvivalencie tejto relácie predpisy $[a]+[b] = [a+b]$, $[a]\cdot[b] = [ab]$ dobre definuje binárne operácie $+, \cdot$ a R/E s týmito binárnymi operáciami tvorí grupu. Navyše, zobrazenie $a \mapsto [a]$ je surjektívny homomorfizmus z R do R/E a jeho jadro je $[0]$.

13.3 Okruhy polynómov – definícia a delenie so zvyškom

Na strednej škole ste strávili veľa času s kvadratickými rovnicami $ax^2+bx+c = 0$. Venovali ste sa aj všeobecnejším rovniciam vyššieho stupňa. Tieto rovnice súvisia s funkciami $f\colon \mathbb{R} \to \mathbb{R}$ tvaru

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0.$$

Takéto funkcie budeme volať polynomické funkcie.

V tejto časti by sme chceli zaviesť podobný pojem pre ľubovoľný komutatívny okruh s jednotkou. V minulom semestri sme pracovali s polynomickými funkciami nad $\mathbb{R}$ ako s prvkami vektorového priestoru $\mathbb{R}^{\mathbb{R}}$ všetkých zobrazení z $\mathbb{R}$ do $\mathbb{R}$. Vtedy sme často používali fakt, že polynomická funkcia $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0$ sa rovná nulovej funkcii (t.j. v každom bode nadobúda hodnotu 0) práve vtedy, keď všetky koeficienty sú nulové, t.j.

$$a_n = a_{n-1} = \ldots = a_0 = 0.$$

(Ako uvidíme, táto vlastnosť neplatí pre všetky polia, v prípade poľa $\mathbb{R}$ však platí, ukážeme to v tvrdení 13.3.13.)

Práve toto je vlastnosť, ktorú budeme požadovať od pojmu polynómu, ktorý teraz ideme definovať.

13.3.1 Definícia okruhu polynómov

{polyn1:DEFRX}

Definícia 13.3.1. Nech R je komutatívny okruh s jednotkou. Potom formálne zápisy tvaru

$$p = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0,$$

kde n je prirodzené číslo a $a_i \in R$ pre $i = 0, \ldots, n$ nazývame *polynómy* v premennej x nad okruhom R.

Namiesto zápisu $a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0$ budeme často používať stručnejší zápis

$$p = \sum_{i=0}^{n} a_i x^i.$$

Prvky $a_n, a_{n-1}, \ldots, a_0 \in R$ voláme *koeficienty* polynómu p.

Ak navyše $a_n \neq 0$, tak n voláme *stupeň polynómu* p, označujeme $\operatorname{st} p = n$. V prípade nulového polynómu (všetky koeficienty sú nulové) definujeme $\operatorname{st} p = -\infty$. (Všimnite si, že s výnimkou nulového polynómu je možné také n zvoliť, t.j. stupeň je definovaný pre každý polynóm.) Polynómy stupňa menšieho ako 1 voláme *konštantné polynómy*.

Koeficient $a_n \neq 0$ pre $n = \operatorname{st} p$ voláme *vedúci koeficient* polynómu p.

Dva polynómy považujeme za rovnaké, ak majú rovnaké koeficienty, t.j. ak $p = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0$, $q = b_m x^m + b_{m-1}x^{m-1} + \cdots + b_0$ a $n \geq m$, tak $p = q$ práve vtedy, keď

$$a_i = b_i \qquad \text{pre } i = 0, 1, \ldots, m$$

a $a_i = 0$ pre $i = m+1, \ldots, n$.

V tejto definícii môže byť trochu nejasné, čo je x a prečo sa volá premenná. Ešte sa tohoto problému dotkneme na konci tejto časti, je to však možné jednoducho brať tak, že doplnenie symbolov x^i je len spôsob zápisu – polynóm je jednoznačne určený svojimi koeficientami.

Príklad 13.3.2. Napríklad $0x^3 + 1x^2 + 2x + 1 = 1x^2 + 2x + 1$ chápeme ako dva rôzne zápisy toho istého polynómu z $\mathbb{R}[x]$.

Vidíme teda, že pridanie alebo odobranie nulových koeficientov polynóm tento polynóm nemení.

Ďalej by sme radi rozumným spôsobom zadefinovali sčitovanie a násobenie polynómov. „Rozumný“ spôsob by mal spĺňať prinajmenšom to, že nejakým spôsobom bude rešpektovať násobenie v okruhu R a tiež by bolo vhodné, aby výsledný okruh bol komutatívny.

Pritom polynóm budeme chápať ako súčet výrazov $a_i x^i$. To vlastne jednoznačne určuje sčitovanie, napríklad pre $p = x^2 + 2x + 1$ a $q = 2x + 1$ máme

$$p + q = (x^2 + 2x + 1) + (2x + 1) = x^2 + 2x + 2x + 1 + 1 = x^2 + 4x + 2.$$

(Využili sme iba to, že polynóm vieme rozložiť na jednotlivé členy a distributívnosť.)

Tieto požiadavky (t.j. vlastnosti komutatívneho okruhu a to, že koeficienty sa násobia rovnako ako v R) už takmer určujú násobenie. Ak chceme napríklad vynásobiť polynómy $p = x^2 + 2x + 1$ a $q = 2x + 1$ v $\mathbb{Z}[x]$, tak z distributívnosti dostaneme

$$(x^2 + 2x + 1)(2x + 1) = x^2 \cdot 2x + 2x \cdot 2x + 1 \cdot 2x + x^2 \cdot 1 + 2x \cdot 1 + 1 \cdot 1.$$

Na základe komutatívnosti dostaneme

$$(x^2 + 2x + 1)(2x + 1) = 2x^2. \cdot x + 4x \cdot x + 2 \cdot x + 1x^2 + 2x + 1.$$

Predpokladajme, že násobenie výrazov obsahujúcich iba x^k funguje takým spôsobom, že $x^m \cdot x^n = x^{m+n}$. Potom predchádzajúci výraz môžeme upraviť na tvar

$$(x^2 + 2x + 1)(2x + 1) = 2x^3 + 4x^2 + 2x + 1x^2 + 2x + 1.$$

Opäť z distributívnosti dostaneme

$$(x^2+2x+1)(2x+1) = 2x^3+5x^2+4x+1.$$

Možno sa tento jednoduchý výpočet zdá rozpísaný zbytočne priveľmi podrobne, cieľom však bolo ukázať, aké vlastnosti potrebujeme, keď chceme niečo podobné definovať nad ľubovoľným komutatívnym okruhom s jednotkou. Zopakovaním rovnakej úvahy pre všeobecný prípad dostaneme:

Definícia 13.3.3. Nech R je komutatívny okruh s jednotkou. Nech $p = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0$ a $q = b_n x^n + b_{n-1}x^{n-1} + \cdots + b_0$ sú ľubovoľné polynómy nad R. (Tým, že u oboch polynómov predpokladáme rovnaký počet koeficientov sme sa nijako neobmedzili – v prípade potreby je možné niektorý polynóm doplniť nulami.)

Potom *súčet polynómov* p a q je

$$p+q = \sum_{i=0}^{n}(a_i+b_i)x^i.$$

Súčin polynómov p a q je polynóm $r = \sum_{i=0}^{2n} c_i x^i$, kde

$$c_k = \sum_{j=0}^{k} a_j b_{k-j}.$$

Teda obe operácie sme definovali rovnako ako v predchádzajúcom príklade – pri sčitovaní sa jednoducho sčítajú koeficienty a pri násobení sú koeficienty výsledného polynómu práve tie výrazy, ktoré by sme dostali roznásobením (koeficient c_k je súčet všetkých možných $a_s b_l$ pre $s+l=k$, čo sú presne všetky možnosti, ako môžeme dostať $x^k = x^s.x^l$).

Definíciu súčinu by sme mohli ekvivalentne prepísať ako

$$c_k = \sum_{m+n=k} a_m b_n.$$

Z tejto ekvivalentnej definície vidno, že pre násobenie polynómov platí asociatívnosť: pre $p = \sum_{i=0}^{n} a_i x^i$, $q = \sum_{i=0}^{n} b_i x^i$, $r = \sum_{i=0}^{n} c_i x^i$ dostaneme $(pq)r = \sum_{i=0}^{3n} d_i x^i$, kde koeficienty d_k majú hodnoty

$$d_k = \sum_{m+n=k} a_m \sum_{s+t=n} b_s c_t = \sum_{m+s+t=k} a_m b_s c_t.$$

Vďaka tomu dostaneme, že

{polyn1:TVRRXJEOKR}

Tvrdenie 13.3.4. *Nech R je komutatívny okruh s jednotkou. Množina všetkých polynómov nad R s násobením a sčitovaním definovaným v predchádzajúcej definícii tvorí komutatívny okruh s jednotkou. Tento okruh označujeme $R[x]$ a voláme ho* okruh polynómov nad R.

Sčitovanie a násobenie polynómov sme vlastne definovali tak, aby akákoľvek rovnosť, ktorá platí pre polynómy platila aj keď namiesto x napíšeme akýkoľvek prvok okruhu R (alebo nejakého nadokruhu, ktorý obsahuje R). To zdôvodňuje použitie názvu premenná – namiesto x môžeme napísať (dosadiť) hocijaký prvok, čiže sa môže meniť. (Aj dosadzovaniu do polynómov sa budeme ešte venovať.)

Dohoda. V ďalšom budeme polynómy zapisovať ako $p(x)$, $q(x)$ atď., čím označíme o polynóm v akej premennej ide. (Ak budeme niekde hovoriť súčasne o polynómoch aj o funkciách, tak opäť použijeme radšej jednopísmenkové označenie p, q; aby nemohlo dôjsť k omylu, že máme na mysli nejakú funkciu resp. jej funkčnú hodnotu.)

Poznámka 13.3.5. Všimnime si, že sčitovanie a násobenie konštantných polynómov funguje rovnako ako násobenie v okruhu R. To znamená, že keď prvky okruhu R stotožníme s im prislúchajúcimi konštantnými polynómami, môžeme R chápať ako podokruh okruhu $R[x]$. (Formálne by sme tento fakt sformulovali tak, že zobrazenie, ktoré prvku $a \in R$ priradí konštantný polynóm $a \in R[x]$ je okruhový homomorfizmus, ktorý je navyše injektívny.) V ďalšom budeme toto stotožnenie často používať (aj bez toho, že by sme na to výslovne upozornili.) To znamená, že R budeme chápať priamo ako podmnožinu $R[x]$.

Všimnime si ešte, že vlastnosť „byť oborom integrity“ sa prenesie z okruhu R na okruh $R[x]$ polynómov nad týmto okruhom.

{polyn1:TVRRXJEOI}

Tvrdenie 13.3.6. *Ak R je obor integrity, tak pre ľubovoľné nenulové polynómy $f, g \in R[x]$ platí*

$$\operatorname{st}(fg) = \operatorname{st}(f) + \operatorname{st}(g)$$

a okruh $R[x]$ polynómov nad okruhom R je obor integrity.

Dôkaz. Ak f a g sú nenulové polynómy, môžeme ich zapísať ako

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0,$$
$$g(x) = b_m x^m + b_{m-1}x^{m-1} + \cdots + b_0,$$

pričom $n = \operatorname{st} f$, $m = \operatorname{st} g$. Vyrátajme, aký bude koeficient c_{n+m} polynómu $f.g$ pri x^{n+m}. Priamo z definície máme, že

$$c_{n+m} = a_n b_m,$$

a pretože R je obor integrity, dostávame $c_{n+m} \neq 0$. To znamená, že polynóm $f.g$ je nenulový (teda $R[x]$ je obor integrity) a tiež, že

$$\operatorname{st}(fg) = m + n = \operatorname{st}(f) + \operatorname{st}(g).$$

□

13.3.2 Delenie so zvyškom

Pre nás bude dôležitý hlavne prípad keď okruh R je pole. Ako sme už ukázali, v tomto prípade platí

$$\operatorname{st}(pq) = \operatorname{st} p + \operatorname{st} q.$$

Neskôr bude pre nás dôležitá nasledujúca veta:

{polyn1:VTDEL}

Veta 13.3.7 (Veta o delení so zvyškom)**.** *Nech F je pole, $f(x), g(x) \in F[x]$ a $g(x) \neq 0$. Potom existujú $q(x), r(x) \in F[x]$ také, že*

$$f(x) = q(x) \cdot g(x) + r(x)$$

a $\operatorname{st} r(x) < \operatorname{st} g(x)$.

Navyše, $q(x)$ a $r(x)$ sú týmito podmienkami jednoznačne určené.

Definícia 13.3.8. Polynómy $q(x)$ a $r(x)$ jednoznačne určené podmienkami z vety 13.3.7 sa nazývajú *podiel* a *zvyšok po delení* polynómu $f(x)$ polynómom $g(x)$. Zvyšok po delení označujeme $f(x) \bmod g(x)$.

Dôkaz. Existencia. Matematickou indukciou vzhľadom na $n = \mathrm{st}(f)$.

1° Ak $\mathrm{st}\, f(x) < \mathrm{st}\, g(x)$, stačí položiť $q(x) = 0$ a $r(x) = f(x)$.

2° Nech $n = \mathrm{st}\, f(x) \geq \mathrm{st}\, g(x)$ a každý polynóm stupňa menej ako n sa dá vydeliť so zvyškom polynómom $g(x)$ (indukčný predpoklad).

Označme $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0$, $g(x) = b_m x^m + b_{m-1}x^{m-1} + \cdots + b_0$ pričom $a_n, b_m \neq 0$. Položme $h(x) = f(x) - a_n b_m^{-1} x^{n-m} g(x)$. Koeficient pri x^n v polynóme $h(x)$ je $a_n - a_n b_m^{-1} b_m = 0$. Teda $\mathrm{st}(h) < \mathrm{st}(f)$, čiže pre polynóm h (podľa indukčného predpokladu) existujú $s(x), r(x) \in F[x]$ také, že

$$h(x) = s(x)g(x) + r(x)$$

a $\mathrm{st}(r) < \mathrm{st}(g)$. Potom

$$f(x) = (s(x) + a_n b_m^{-1} x^{n-m})g(x) + r(x).$$

Jednoznačnosť. Nech platí

$$f(x) = q_1(x)g(x) + r_1(x) = q_2(x)g(x) + r_2(x),$$

pričom $\mathrm{st}(r_1) < \mathrm{st}(g)$, $\mathrm{st}(r_2) < \mathrm{st}(g)$. Potom máme

$$(q_1(x) - q_2(x))g(x) = r_2(x) - r_1(x).$$

Na pravej strane je polynóm stupňa menšieho ako $\mathrm{st}(g)$. Ak by platilo $q_1(x) - q_2(x) \neq 0$, tak na ľavej strane tejto rovnosti dostaneme polynóm stupňa aspoň $\mathrm{st}(g)$, čo je spor. Preto musí platiť $q_1(x) - q_2(x) = 0$ a $q_1(x) = q_2(x)$.

Z toho potom dostávame aj $r_1(x) - r_2(x) = 0$ a $r_1(x) = r_2(x)$. □

Všimnime si, že dôkaz predchádzajúcej vety nám súčasne dáva návod, ako rátať pre dané polynómy ich podiel a zvyšok.

Príklad 13.3.9. Vydeľme so zvyškom polynóm $f(x) = x^4+6x^3+12x^2+12x+10$ polynómom $g(x) = x^2 + x + 1$. Podľa návodu z dôkazu by sme sa mali pozrieť najprv na vedúce členy – vidíme, že $x^4 = x^2 \cdot x^2$. Vypočítame teda

$$f(x) - x^2 g(x) = (x^4 + 6x^3 + 12x^2 + 12x + 10) - x^2(x^2 + x + 1) = 5x^3 + 11x^2 + 12x + 10.$$

Výsledok by sme opäť mali deliť polynómom $g(x)$ a postup opakovať, až kým nedostaneme polynóm stupňa menšieho ako $g(x)$.

$$5x^3 + 11x^2 + 12x + 10 - 5x(x^2 + x + 1) = 6x^2 + 7x + 10$$
$$6x^2 + 7x + 10 - 6(x^2 + x + 1) = x + 4$$

Celkovo sme dostali, že $f(x) - (x^2 + 5x + 6)g(x) = x + 4$, čiže

$$f(x) = (x^2 + 5x + 6)g(x) + (x + 4),$$

teda podiel je $x^2 + 5x + 6$ a zvyšok po delení je $x + 4$.

V prípade, že je polynóm $g(x)$ (=stupňa 1) môžeme podiel vyrátať jednoduchším spôsobom, ktorý sa naučíme v časti 13.5.1.

Neskôr bude pre nás užitočný fakt, že analogická veta platí aj v okruhu $(\mathbb{Z}, +, \cdot)$. Dala by sa dokazovať podobným spôsobom ako predchádzajúca veta , tu si ukážeme o trochu iný dôkaz.

{VTDELvZ}

Veta 13.3.10. *Nech a, b sú celé čísla, $b > 0$. Potom existujú celé čísla q a r také, že*

$$a = q \cdot b + r \qquad a \qquad 0 \leq r < q.$$

Navyše, q a r sú týmito podmienkami jednoznačne určené.

Definícia 13.3.11. Číslo r z predchádzajúcej vety sa nazýva *zvyšok a po delení b* a označuje sa $a \mod b$.

Dôkaz. Existencia: Množina $\{k; kb \leq pa\}$ je zhora ohraničená. Preto existuje $q := \max\{k; kb \leq a\}$. Položme $r = a - qb$. Očividne platí $a = qb + r$ a $r \geq 0$.

Tvrdíme, že $r < b$. Nech by to tak nebolo. Z nerovnosti $r \geq b$ dostaneme $a \geq (q+1)b$, čo je spor s definíciou čísla q.

Jednoznačnosť: Predpokladajme, že $a = qb + r = q'b + r'$, kde $0 \leq r, r' < b$. Potom

$$(q - q')b = r' - r.$$

Predpokladajme, že by $|q - q'| > 0$. Potom $|r - r'| \geq b$, čo je spor s tým, že $0 \leq r, r' < b$.

Preto platí

$$(q - q')b = r - r' = 0,$$

a $q = q'$, $r = r'$. □

13.3.3 Polynómy a polynomické funkcie

V tomto článku budeme polynómy vždy označovať ako $p, q, \ldots$ (t.j. jedným písmenom).

{polyn1:DEFPOLFCIA}

Definícia 13.3.12. Nech R je komutatívny okruh s jednotkou. *Polynomickou funkciou* nad R budeme rozumieť ľubovoľnú funkciu $f\colon R \to R$ určenú predpisom

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0$$

pre nejaké $n \in \mathbb{N}$ a $a_1 \ldots a_n \in R$.

Množina všetkých polynomických funkcií s obvyklým násobením a sčitovaním funkcií opäť tvorí okruh (je to podokruh okruhu R^R – úloha 13.3.1), tento okruh budeme označovať $R\langle x\rangle$.

Pri zavedení polynómov sme spomínali polynomické funkcie nad poľom $\mathbb{R}$. Zaujíma nás, aký je vo všeobecnosti vzťah medzi okruhmi $F[x]$ a $F\langle x\rangle$, ak F je ľubovoľné pole.

Máme prirodzené priradenie medzi polynómami a polynomickými funkciami $\varphi\colon F[x] \to F\langle x\rangle$, ktoré polynómu $p = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0$ priradí funkciu danú predpisom $f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0$. Dá sa overiť, že toto zobrazenie je surjektívny homomorfizmus z okruhu $F[x]$ na okruh $F\langle x\rangle$.

V prípade, že homomorfizmus φ je injektívny, tak je to izomorfizmus. Čiže na to, aby sme zistili, či sú tieto dva okruhy izomorfné, stačí zistiť, ako vyzerá $\operatorname{Ker}\varphi$. Ukážeme, že pre nekonečné polia sú okruhy $F[x]$ a $F\langle x\rangle$ izomorfné, zatiaľčo pre konečné polia to platiť nemusí.

{polyn1:TVRNEKPOLE}

Tvrdenie 13.3.13. *Ak F je nekonečné pole tak polynomická funkcia $f\colon F \to F$*

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0$$

sa rovná nulovej funkcii práve vtedy, keď $a_0 = a_1 = \cdots = a_n = 0$, t.j. vtedy, keď sú všetky koeficienty nulové.

Náčrt dôkazu. Vyberme $n+1$ navzájom rôznych prvkov $x_0, \ldots, x_n$ poľa F. Potom koeficienty $a_0, \ldots, a_n$ spĺňajú sústavu $n+1$ lineárnych rovníc

$$\begin{aligned} a_n x_0^n + a_{n-1} x_0^{n-1} + \cdots + a_0 &= 0 \\ a_n x_1^n + a_{n-1} x_1^{n-1} + \cdots + a_0 &= 0 \\ &\cdots \\ a_n x_n^n + a_{n-1} x_n^{n-1} + \cdots + a_0 &= 0 \end{aligned}$$

Z úlohy 6.5.8 (pozri tiež napríklad [K, Príklad 6.2.17(2)], [KGGS, s.114/7]) vieme, že determinant matice tejto sústavy je

$$\begin{vmatrix} 1 & x_0 & x_0^2 & \ldots & x_0^n \\ 1 & x_1 & x_1^2 & \ldots & x_1^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \ldots & x_x^n \end{vmatrix} = \prod_{0 \le i < j \le n} (x_j - x_i)$$

čiže ak prvky x_i sú navzájom rôzne, je nenulový. To znamená, že táto matica je regulárna a uvedenej sústave rovníc vyhovuje iba nulové riešenie.

Teda $\operatorname{Ker} f$ v tomto prípade pozostáva iba z nulového polynómu (všetky koeficienty sú nuly). □

Príklad 13.3.14. Homomorfizmus $\varphi\colon \mathbb{Z}_2[x] \to \mathbb{Z}_2\langle x \rangle$, ktorý polynómu priraďuje zodpovedajúcu polynomickú funkciu, nie je injektívny.

Stačí si všimnúť, že pre každé $x \in \mathbb{Z}_2$ platí $x^2 + x = 0$, teda polynomická funkcia $x^2 + x$ je nulová a

$$x^2 + x \in \operatorname{Ker} \varphi.$$

Homomorfizmus $\varphi\colon R[x] \to R\langle x \rangle$ nám súčasne dáva možnosť „dosadzovať“ do polynómov. Ak totiž máme daný prvok $b \in R$ a nejaký polynóm $f = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0 \in R[x]$, tak mu vieme priradiť funkciu $\varphi(f)\colon R \to R$. Potom môžeme b dosadiť do tejto funkcie, čiže dostaneme

$$\varphi(f)(b) = a_n b^n + a_{n-1} b^{n-1} + \cdots + a_0.$$

Navyše, zobrazenie $f_b\colon R[x] \to R$ určené predpisom

$$f_b\colon f \mapsto a_n b^n + a_{n-1} b^{n-1} + \cdots + a_0$$

je okruhový homomorfizmus taký, že $f(x) = b$ (t.j. polynóm x sa zobrazí na prvok b.)

To, že f_b je skutočne homomorfizmus možno vidieť napríklad z toho, že $f_b = g_b \circ \varphi$, kde $g_b\colon R^R \to R$ je homomorfizmus daný predpisom $g_b(f) = f(b)$ (úloha 13.2.5).

{polyn1:DEFDOSHOM}

Definícia 13.3.15. Ak R je komutatívny okruh a $b \in R$, tak homomorfizmus $f_b\colon R[x] \to R$ daný predpisom

$$f_b\colon f \mapsto a_n b^n + a_{n-1} b^{n-1} + \cdots + a_0$$

voláme *dosadzovací homomorfizmus.*

13.3.4 Iné možnosti, ako definovať okruh polynómov

Keďže považujeme izomorfné okruhy za rovnaké (z toho dôvodu, že sú nerozlíšiteľné pomocou pojmov definovaných „v jazyku okruhov“, t.j. nemožno ich odlíšiť žiadnou vlastnosťou sformulovanou len s použitím sčitovania a násobenia v okruhu), je jasné, že akákoľvek iná

definícia okruhov, ktorá by ako výsledok poskytla okruh izomorfný s okruhom $R[x]$, by bola rovnako dobrá.

Pomerne jednoduchá definícia, s ktorou by sa nám dobre pracovalo a ktorej by sme intuitívne celkom dobre rozumeli, by bola definícia okruhu $R[x]$ ako okruhu všetkých polynomických funkcií. Ako sme už videli, takto okruh $R[x]$ nemôžeme definovať, pretože pre konečné polia by sme takto zadefinovali niečo úplne iné než chceme.

Iná možná definícia okruhu polynómov nad okruhom R by bola nasledovná (takto sa definujú okruhy polynómov v [KGGS]):

{polyn1:DEFRX2}

Definícia 13.3.16. Nech R je komutatívny okruh s jednotkou. Predpokladajme, že R je podokruh nejakého komutatívneho okruhu R' a existuje prvok $x \in R'$ taký, že rovnosť

$$a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0 = 0$$

pre $a_1, \ldots, a_n \in R$ platí práve vtedy, keď $a_1 = \cdots = a_n = 0$. Potom prvok x voláme *transcendentný prvok* nad R.

Podokruh

$$R[x] = \{a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0; n \in \mathbb{N}, a_1, \ldots, a_n \in R\}$$

okruhu R' potom voláme *okruhom polynómov* v premennej x nad R.

Overiť, že množina $R[x]$ zadefinovaná v predchádzajúcej definícii je skutočne podokruhom R' je jednoduché – dá sa dokonca ukázať, že je to najmenší podokruh obsahujúci $R \cup \{x\}$. Z toho vyplýva výhoda tejto definície – automaticky vidíme, že $R[x]$ je okruh, z toho, že ide o podokruh okruhu R'. (V našej definícii sme to museli dokazovať.)

Táto definícia si vyžaduje istú prácu navyše – aby sme mohli definovať $R[x]$ pre ľubovoľný komutatívny okruh R s jednotkou, treba dokázať, že pre každý takýto okruh R existuje vhodný nadokruh R', t.j. existuje nadokruh obsahujúci aspoň jeden transcendentný prvok. O chvíľu sa dozvieme, ako sa dá dokázať takéto niečo.

Ďalej pri použití takejto definície musíme ukázať aj to, že bez ohľadu na voľbu nadokruhu R' a transcendentného prvku $x \in R'$ dostaneme vždy (až na izomorfizmus) to isté.

Skúsme sa ešte na chvíľu pozrieť na našu definíciu 13.3.1. K nej by sme mohli mať jednu vážnu výhradu – kedysi v minulom semestri sme tvrdili, že pre nás bude pojem množiny základným pojmom, ktorý síce nedefinujeme (iba popíšeme niektoré jeho vlastnosti), pomocou množín a operácii s nimi už však budeme schopní vystavať celú potrebnú teóriu, teda všetky ďalšie pojmy budeme schopný preformulovať v jazyku množín.

V tejto definícii sme použili „symbol x“ a „formálne zápisy tvaru“ $a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0$ – čo rozhodne nesedí s našou koncepciou definovať všetko pomocou množín. (Čo je symbol? Čo znamená *formálny zápis*?)

Ukážeme si, ako to môžeme zachrániť – t.j. zadefinujeme okruh polynómov tak, aby naša definícia bola „množinová“. Súčasne nám táto definícia poskytne aj riešenie jedného z problémov s definíciou 13.3.16 – existenciu nadokruhu, ktorý obsahuje transcendentný prvok.

Napriek tomu sa však menej formálna definícia 13.3.1 zdá byť lepšia – pretože pri nej s prvkami okruhu $R[x]$ pracujeme rovnako ako s výrazmi obsahujúcimi nejaké prvky okruhu R. Násobenie, ako sme ho definovali v tejto definícii je teda veľmi prirodzené. V nasledujúcej definícii bude o niečo komplikovanejšie a striktné používanie tejto definície by viedlo k zložitejším zápisom polynómov.

{polyn1:DEFRX3}

Definícia 13.3.17. Nech R je ľubovoľný komutatívny okruh s jednotkou. Ako $R[x]$ označíme množinu všetkých postupností prvkov z R takých, že iba konečne veľa členov tejto postupnosti je nenulových. Ďalej zadefinujeme sčitovanie dvoch postupností ako

$$(a_n)_{n=1}^{\infty} + (b_n)_{n=1}^{\infty} = (a_n + b_n)_{n=1}^{\infty}$$

a súčin postupností $(a_n)_{n=1}^{\infty}$, $(b_n)_{n=1}^{\infty}$ definujeme ako postupnosť $(c_n)_{n=1}^{\infty}$, ktorej členy sú určené predpisom

$$c_k = \sum_{j=0}^{k} a_j b_{k-j} = \sum_{m+n=k} a_m b_n.$$

Táto množina postupností s uvedeným sčitovaním a násobením tvorí okruh, ktorý voláme *okruh polynómov* nad R.

Táto definícia v istom zmysle presne zodpovedá definícii 13.3.1 – postupnosti sú tiež jednoznačne určené svojimi členmi, takisto ako dva polynómy sme v definícii 13.3.1 prehlásili za rovnaké, ak mali rovnaké koeficienty.

Dôkaz, že takýmto spôsobom dostaneme okruh je takmer totožný s dôkazom tvrdenia 13.3.4.

Polynóm $3x^2 - 1x + 0$ v tejto definícii zodpovedá postupnosti $(0, -1, 3, 0, 0, \ldots)$. Prvky z R môžeme stotožniť s postupnosťami tvaru $(a, 0, 0, 0, \ldots)$, kde $a \in R$. Všimnime si, že polynóm x zodpovedá postupnosti $(0, 1, 0, 0, \ldots)$ a dá sa ukázať, že v okruhu $R[x]$ (chápanom ako postupnosti, čiže ako v poslednej uvedenej definícii) je tento prvok transcendentným prvkom nad R.

Cvičenia

{polyn1cvic:POLFCIE}

Úloha 13.3.1. Dokážte, že polynomické funkcie (definícia 13.3.12) tvoria podokruh okruhu F^F.

13.4 Deliteľnosť v okruhoch

V tejto časti sa budeme zaoberať deliteľnosťou v okruhoch. Najdôležitejšími príkladmi budú pre nás okruh $(\mathbb{Z}, +, \cdot)$ celých čísel a okruh $(F[x], +, \cdot)$ polynómov nad poľom F.

V celej podkapitole budeme predpokladať, že okruh, s ktorým pracujeme, je obor integrity. Nasledujúcu vlastnosť oborov integrity budeme často používať, preto ju sformulujeme ako samostatnú lemu. (Vlastne to je špeciálny prípad krátenia nenulovým prvkom v obore integrity – tvrdenie 13.1.14.)

{euklid:LMAB1}

Lema 13.4.1. *Nech R je obor integrity, $a, b \in R$. Ak platí $ab = a$ pre $a \neq 0$, tak $b = 1$.*

Dôkaz. Z rovnosti $ab = a = a1$ vyplýva

$$ab - a1 = a(b-1) = 0,$$

čiže v obore integrity pre $a \neq 0$ máme $b - 1 = 0$, čiže $b = 1$. □

{euklid:DEFDELI}

Definícia 13.4.2. Nech R je obor integrity. Hovoríme, že *a delí b*, označujeme $a \mid b$, ak existuje $c \in R$ také, že $b = ca$.

{euklid:LMDELI}

Lema 13.4.3. *Nech R je obor integrity. Potom pre ľubovoľné $a, b, c, d \in R$, $a_i, r_i \in R$ platí*

(i) $a \mid a$

(ii) $a \mid b \wedge b \mid c \Rightarrow a \mid c$

(iii) $a \mid b \wedge c \mid d \Rightarrow ac \mid bd$

(iv) $a \mid 0$, $1 \mid a$

(v) $0 \mid a \Leftrightarrow a = 0$

(vi) $ac \mid bc \wedge c \neq 0 \Rightarrow a \mid b$

(vii) $a \mid a_i$ *pre* $i = 1, \ldots n \Rightarrow a \mid a_1 r_1 + \cdots + a_n r_n$

Dôkaz. Jednoduchý – ponecháme ako cvičenie. □

Príklad 13.4.4. V prípade okruhu $\mathbb{Z}$ je relácia $\mid$ tá istá relácia deliteľnosti, ktorú poznáte zo strednej školy, t.j. napríklad $3 \mid 12$, lebo $12 = 3 \cdot 4$, zatiaľčo $3 \nmid 7$.

Všimnime si, že $a \mid b$ znamená to isté, ako že zvyšok čísla b po delení číslom a je 0.

Príklad 13.4.5. V okruhoch $\mathbb{Z}[x]$, $\mathbb{R}[x]$ platí $x - 1 \mid x^2 - 1$, pretože $x^2 - 1 = (x - 1)(x + 1)$.

Pritom si môžeme všimnúť, že v $\mathbb{R}[x]$ platí aj $2x - 2 \mid x^2 - 1$ (lebo $x^2 - 1 = (2x - 2)(\frac{1}{2}x + \frac{1}{2})$), ale v okruhu $\mathbb{Z}[x]$ už táto relácia neplatí. Deliteľnosť polynómov, ak ich chápeme ako polynómy nad $\mathbb{Z}$ a nad $\mathbb{R}$, sú rôzne pojmy, hoci $\mathbb{R}$ je nadpoľom $\mathbb{Z}$.

Všimnime si, že aj v okruhoch $F[x]$ platí $f(x) \mid g(x)$ práve vtedy, keď zvyšok polynómu $g(x)$ po delení $f(x)$ je 0. (Neskôr si to zdôvodníme podrobnejšie vo všeobecnejšom prípade)

Definícia 13.4.6. Ak $a, b \in R$, kde R je obor integrity, hovoríme, že prvky a a b sú *asociované,* označujeme $a \sim b$, ak $a \mid b$ a súčasne $b \mid a$

$$a \mid b \ \wedge \ b \mid a \Leftrightarrow a \sim b$$

{euklid:LMKONJ}

Lema 13.4.7. *Nech R je obor integrity. Pre ľubovoľné $a, b, c, d \in R$ platí*

(i) $a \sim b \wedge b \sim c \Rightarrow a \sim c$

(ii) $a \sim a$

(iii) $a \sim b \Rightarrow b \sim a$

(iv) $a \sim b \wedge c \sim d \Rightarrow ac \sim bd$

Dôkaz lemy 13.4.7 pre jednoduchosť vynechávame. Môžeme si všimnúť, že prvé tri vlastnosti nám hovoria, že relácia „byť asociovaný“ je relácia ekvivalencie. (Podobným spôsobom môžeme dostať z ľubovoľného čiastočného usporiadania reláciu ekvivalencie – úloha 13.4.2.) Posledná podmienka hovorí, že relácia $\sim$ sa správa rozumne vzhľadom na násobenie.

Definícia 13.4.8. Ak okruh R má jednotku a $ab = 1$, hovoríme, že a je *deliteľ jednotky.* Množinu všetkých deliteľov jednotky budeme označovať $U(R)$.

{euklid:TVRURSIM}

{euklid:URSIMit1}

{euklid:URSIMit2}

Tvrdenie 13.4.9. *Nech R je obor integrity. Potom*

(i) *Delitele jednotky s operáciou násobenia tvoria grupu, t.j. $(U(R), \cdot)$ je grupa.*

(ii) *$a \sim b$ práve vtedy, keď existuje deliteľ jednotky u taký, že $a = bu$.*

Dôkaz. (i) Uzavretosť na násobenie: Ak $a, b \in U(R)$, znamená to existenciu $c, d \in R$ takých, že $ac = 1$, $bd = 1$. Potom $acbd = (ab)(cd) = 1$, čiže aj ab je deliteľ jednotky.

Asociatívnosť máme priamo z definície okruhu, neutrálny prvok je 1.

Existencia inverzného prvku: Ak a je deliteľ jednotky, znamená to, že existuje $b \in R$ také, že $ab = 1$. To znamená, že $b \in U(R)$ a tento prvok je inverzný k a vzhľadom na násobenie.

(ii) Ľahko vidno, že $a \sim 0$ platí práve vtedy, keď $a = 0$ (z lemy 13.4.3 vieme, že $0 \mid a$ iba pre $a = 0$). Samozrejme, $u0 = 0$ pre ľubovoľné $u \in U(R)$.

Zostáva nám teda dokázať tvrdenie pre prípade $a \neq 0$.

Ak $a \mid b$ a $b \mid a$, tak existujú $c, d \in R$ také, že $ac = b$ a $bd = a$. Potom máme

$$a = bd = (ac)d = a(cd)$$

a z lemy 13.4.1 dostaneme $cd = 1$, čiže c aj d sú delitele jednotky. □

lid:PRDELJED}

Príklad 13.4.10. Ľahko sa dá overiť, že ± 1 sú delitele jednotky v $\mathbb{Z}$ a všetky nenulové konštantné polynómy sú delitele jednotky v $F[x]$. (Tento fakt vyplýva aj z lemy 13.4.14, ktorú o chvíľu dokážeme.)

Takisto nie je ťažké ukázať, že iné delitele jednotky tam už nie sú. Skutočne, ak $ab = 1$ v $\mathbb{Z}$, tak $a, b \neq 0$, z čoho máme $|a| \geq 1$, $|ab| = |a||b| \geq 1$. Aby v predchádzajúcej rovnosti nastala rovnosť, musí byť $|a| = 1$, čiže $a = \pm 1$.

Ak $f(x)$ je deliteľ jednotky v $F[x]$, tak máme $f(x)g(x) = 1$. Pritom $g(x) \neq 0$ (lebo potom by sme dostali $f(x)g(x) = 0$), preto $\operatorname{st} g \geq 0$. Potom (tvrdenie 13.3.6) $\operatorname{st}(fg) = \operatorname{st} f + \operatorname{st} g \geq \operatorname{st} f$. Súčasne vieme $\operatorname{st}(fg) = \operatorname{st} 1 = 0$, preto aj $\operatorname{st} f = 0$ a $f(x)$ je konštantný polynóm. (Nemôže platiť $f(x) = 0$; zdôvodniť to môžeme rovnako ako sme to spravili pre polynóm $g(x)$.)

13.4.1 Euklidovské okruhy

Veta 13.3.7 o delení so zvyškom je dôležitou vlastnosťou okruhu $F[x]$ polynómov nad poľom F. Veta 13.3.10 nám hovorí, že analogickú vlastnosť má aj okruh celých čísel $(\mathbb{Z}, +, \cdot)$.

Na základe tejto vety môžeme odvodiť mnohé vlastnosti, ktoré sú spoločné pre oba spomínané okruhy – najjednoduchšie bude odvodiť ich všeobecne pre oba spomínané okruhy.

{euklid:DEFOI}

Definícia 13.4.11. Obor integrity R sa nazýva *euklidovský okruh,* ak existuje funkcia $N\colon R \setminus \{0\} \to \mathbb{N}$ taká, že pre ľubovoľné $a, b \in R$, $b \neq 0$ existujú $q, r \in R$ také, že $a = qb + r$ a buď $r = 0$ alebo $N(r) < N(b)$.

Funkciu N budeme nazývať *norma.*

Okruh je euklidovský, ak existuje funkcia N s uvedenými vlastnosťami. Samozrejme, ako uvidíme aj v nasledujúcom príklade, pre nejaký euklidovský okruh môže existovať viacero noriem.

V prípade, že sa nám to hodí, môžeme uvažovať o norme aj ako o funkcii definovanej na celom R; z definície je jasné, že voľba hodnoty $N(0)$ nijako neovplyvní či norma spĺňa dané podmienky.

Poznámka 13.4.12. Niektorí autori v definícii euklidovského okruhu navyše požadujú, aby norma spĺňala podmienku $N(a) \leq N(ab)$. V skutočnosti sú tieto 2 definície ekvivalentné, t.j. ak na obore integrity existuje norma s vlastnosťami z definície 13.4.11, tak existuje aj taká norma, ktorá navyše spĺňa $N(a) \leq N(ab)$ (pozri napríklad [Rog]).

Príklad 13.4.13. Okruh $\mathbb{Z}$ je euklidovský okruh. Ako normu môžeme zvoliť absolútnu hodnotu čísla z, čiže $N(z) = |z|$. Takisto norma $N(z) = |z| + 1$ vyhovuje definícii euklidovského okruhu.

Okruh $F[x]$, kde F je ľubovoľné pole, je euklidovský okruh. Za normu môžeme zvoliť stupeň polynómu (ten je pre každý nenulový polynóm definovaný ako prirodzené číslo).

Ľahko si môžeme všimnúť, že

{euklid:LMNORMO}

Lema 13.4.14. *Ak R je euklidovský okruh, $u \neq 0$ a $N(u) = 0$, tak u je deliteľ jednotky.*

Dôkaz. Priamo z definície máme, že $1 = uc + d$, pričom $N(d) < 0$ alebo $d = 0$. Pretože prípad $N(d) < 0$ nemôže nastať, máme $d = 0$. □

13.4.2 Okruhy hlavných ideálov

Ďalším typom okruhov, ktorý bude pre nás užitočný sú okruhy hlavných ideálov.

Definícia 13.4.15. Ak R je obor integrity, hovoríme, že R je okruh hlavných ideálov, ak každý ideál v R je hlavný, t.j. ak je tvaru

$$I = (a) = \{ax; x \in R\}$$

pre nejaké $a \in R$.

{euklid:TVREOJEOHI}

Tvrdenie 13.4.16. *Každý euklidovský okruh je okruh hlavných ideálov.*

Dôkaz. Nech R je euklidovský okruh, $I \neq \emptyset$ je ideál v R.

Ak $I = \{0\}$, tak $I = (0)$. Môžeme teda predpokladať, že I obsahuje aspoň jeden nenulový prvok.

[8] Ak v $I \setminus \{0\}$ existuje prvok s nulovou normou, tak tento prvok je deliteľom jednotky (podľa lemy 13.4.14). To by ale znamenalo, že $I = R = (1)$. V ďalšej časti dôkazu teda môžeme predpokladať, že všetky prvky $I \setminus \{0\}$ majú nenulovú normu.

Nech b je nenulový prvok z I s najmenšou normou. (Taký prvok existuje, lebo $\{N(b); b \in I \setminus \{0\}\}$ je neprázdna podmnožina prirodzených čísel. Každá neprázdna podmnožina prirodzených čísel má najmenší prvok – princíp dobrého usporiadania.)

Tvrdíme, že $I = (b)$. Pre každý prvok $a \in I$ máme $a = bc + d$. Pritom $d = bc - a \in I$, čiže opäť nemôže nastať možnosť $N(d) < N(b)$. Teda $d = 0$ a $a = bc$. Tým sme ukázali, že $I \subseteq (b)$. Inklúzia $(b) \subseteq I$ je zrejmá. □

Obrátené tvrdenie neplatí, ale príklad, ktorý to ukazuje nie je úplne jednoduchý.

Príklad 13.4.17. Z predchádzajúceho tvrdenia špeciálne dostávame, že $\mathbb{Z}$ a $F[x]$ sú okruhy hlavných ideálov, teda v $\mathbb{Z}$ neexistujú iné ideály ako ideály tvaru $(k) = k\mathbb{Z}$ a takisto v $F[x]$ každý ideál pozostáva z násobkov nejakého polynómu $f(x)$.

{eukld:PRID2X}

Príklad 13.4.18. Okruh $\mathbb{Z}[x]$ je príklad oboru integrity, ktorý nie je okruhom hlavných ideálov. Ak uvažujeme ideál

$$(2, x) = \{a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0 \in \mathbb{Z}[x]; a_0 \in 2\mathbb{Z}, a_i \in \mathbb{Z} \text{ pre } i \geq 1\}$$

v okruhu $\mathbb{Z}[x]$ (t.j. ideál generovaný polynómom x a konštantným polynómom 2; pozri poznámku 13.4.20), tak tento ideál nie je hlavný.

Ak by bol totiž generovaný jediným polynómom, musel by to byť polynóm stupňa 0. (V ideále $(f(x))$ generovanom polynómom $f(x)$ majú všetky polynómy stupeň väčší alebo rovný st f – vyplýva to z tvrdenia 13.3.6.) Generátorom by teda musel byť nejaký konštantný polynóm c. Potom však c musí byť párne (lebo iné konštantné polynómy v ideále $(2, x)$ nie sú.) V hlavnom ideále (c) generovanom nejakou párnou konštantou však nevyhnutne musia mať všetky polynómy iba párne koeficienty, čiže nedostali by sme tak všetky polynómy patriace do $(2, x)$.

Špeciálne, keďže sme ukázali, že $\mathbb{Z}[x]$ nie je okruh hlavných ideálov, vyplýva z tvrdenia 13.4.16, že to nie ani euklidovský okruh.

[8] Toto bolo chybne:
Ak by všetky nenulové prvky v I mali nulovú normu, tak sú deliteľmi jednotky (podľa lemy 13.4.14). To by ale znamenalo, že $I = R = (1)$. V ďalšej časti dôkazu teda môžeme predpokladať, že v I existuje nenulový prvok s nenulovou normou.

V dôsledku 13.2.21 sme ukázali, že každý maximálny ideál je prvoideál. V okruhu hlavných ideálov platí aj obrátená implikácia:

VROHIPRVOMAX}

Tvrdenie 13.4.19. *Ak $I = (m)$, $I \neq \{0\}$, je prvoideál v OHI R, tak I je maximálny.*

Dôkaz. Nech $I \subseteq J \subseteq R$. Pretože R je OHI existuje prvok $a \in R$ taký, že $J = (a)$. Zrejme $a \neq 0$ (inak by platilo $I \subseteq (0)$, teda I by bol nulový ideál). Máme teda $(m) \subseteq (a)$, čiže $m = ac$ pre nejaké $c \in R$. Potom buď $a \in I$ a $I = (a)$ alebo $c \in I$, čiže $c = md$ a $m = ac = m(ad)$. Z toho máme $ad = 1$ (pretože R je OI), čiže a je deliteľ jednotky a $(a) = R$. □

Deliteľnosť v okruhoch hlavných ideálov

Všimnime si, že v OHI platí nasledovný vzťah medzi deliteľnosťou v okruhu a hlavnými ideálmi:

$$a \mid b \Leftrightarrow b \in (a) \Leftrightarrow (b) \subseteq (a). \tag{13.2}$$

{euklid:EQIDEDELI}

(Vyplýva to priamo z definície deliteľnosti a z definície hlavného ideálu.)

V súvislosti s hlavnými ideálmi si tiež môžeme všimnúť, že $(a) = R$ práve vtedy, keď a je deliteľ jednotky. (Pozri lemu 13.2.9.)

{euklid:POZNIDEGENER}

Poznámka 13.4.20. Podobne, ako (a) označuje ideál generovaný prvkom a, znakom $(a_1, \ldots, a_n)$ budeme označovať najmenší ideál obsahujúci všetky prvky $a_1, \ldots, a_n$. Ľahko sa dá overiť, že v komutatívnom okruhu s jednotkou

$$(a_1, \ldots, a_n) = \{\sum_{i=1}^{n} x_i a_i; x_i \in R\}$$

(Je zjavné, že táto množina obsahuje prvky $a_1, \ldots, a_n$. Stačí teda overiť, že to je ideál – to ide ľahko z definície ideálu.)

Špeciálne máme

$$(a, b) = \{ax + by; x, y \in R\}.$$

Podobne ako pre celé čísla, aj v oboroch integrity vieme definovať pojem najväčší spoločný deliteľ.

Definícia 13.4.21. *Najväčší spoločný deliteľ* prvkov $a, b \in R$ je taký prvok $c \in R$, že

(i) $c \mid a$, $c \mid b$,

(ii) pre ľubovoľný prvok $d \in R$ taký, že $d \mid a$ a $d \mid b$ platí aj $d \mid c$.

Označujeme ho $\gcd(a, b)$.

Inak povedané, $\gcd(a, b)$ je najväčší (vzhľadom na usporiadanie $|$) prvok z množiny čísel, ktoré súčasne delia a aj b (=spoločné delitele čísel a, b).

Priamo z definície vidno, že najväčší spoločný deliteľ (ak existuje) je určený jednoznačne až na asociovanosť.[9]

[9] Pretože najväčší spoločný deliteľ nie je jednoznačne určený, nemal by som používať označenie $c = \gcd(a, b)$; znamienko rovnosti typicky používame keď na oboch stranách máme jednoznačne určený objekt. Na druhej strane napríklad pri okruhoch $\mathbb{Z}$ a $F[x]$ sme z viacerých možností schopný vybrať „pekného“ reprezentanta. V $\mathbb{Z}$ vyberieme nezáporné číslo, v $F[x]$ vyberieme polynóm kde vedúci koeficient je jednotka. Čiže v týchto prípadoch ste asi zvyknutí na takýto výber najväčšieho spoločného deliteľa. Používam v tomto texte zápis s rovnosťou – ale upozornil som na to, že nie je úplne korektný.

{euklid:TVRBE

Tvrdenie 13.4.22. *[Bézoutova identita] Ak R je okruh hlavných ideálov, tak pre ľubovoľné $a, b \in R$ existuje v R najväčší spoločný deliteľ $c = \gcd(a, b)$.*

Navyše, existujú také $x, y \in R$, že

$$c = xa + yb.$$

Dôkaz. Vieme, že $(a, b) = \{ax + by; x, y \in R\}$ je ideál v R. Pretože R je okruh hlavných ideálov, existuje $c \in R$ také, že $(c) = (a, b)$. Z toho špeciálne máme $a, b \in (c)$, čiže $c \mid a$, $c \mid b$.

Navyše, pretože $c \in (a, b)$, máme zaručenú existenciu $x, y \in R$ s vlastnosťou $ax + by = c$.

Z toho potom dostávame, že pre ľubovoľné $d \in R$ také, že $d \mid a$, $d \mid b$, platí

$$d \mid ax + by = c.$$

□

{euklid:DOSEUKLEMA}

Z predchádzajúceho tvrdenia dostávame nasledujúci dôsledok, ktorý je často užitočný.

Dôsledok 13.4.23. *Nech R je okruh hlavných ideálov, $a, b, c \in R$, $a, b \neq 0$. Ak $\gcd(a, b) = 1$ a $a \mid bc$, tak $a \mid c$.*

$$\gcd(a, b) = 1 \quad \wedge \quad a \mid bc \qquad \Rightarrow \qquad a \mid c$$

Dôkaz. Z tvrdenia 13.4.22 máme existenciu $x, y \in R$ takých, že

$$ax + by = 1.$$

Potom

$$a \mid ac \cdot x + bc \cdot y = (ax + by)c = c.$$

□

Tvrdenie 13.4.22 hovorí o existencii najväčšieho spoločného deliteľa čísel a, b a o existencii $x, y \in R$ s vlastnosťou $\gcd(a, b) = xa + yb$, nehovorí však, ako by sme $\gcd(a, b)$, x a y vedeli vyrátať.

V prípade, že vieme v našom obore integrity (algoritmicky) deliť so zvyškom, dá sa to urobiť pomocou *Euklidovho algoritmu.*

{euklid:LMEUKLIDGCD}

Základom Euklidovho algoritmu je nasledujúca lema:

Lema 13.4.24. *Ak R je obor integrity a $a, b \in R$, tak*

$$\gcd(a, b) = \gcd(a + bx, b)$$

pre ľubovoľné $x \in R$.

Dôkaz. Keďže najväčší spoločný deliteľ je generátor ideálu (a, b), stačí dokazovať rovnosť ideálov $(a, b) = (a + bx, b)$.

Priamo z definície ideálu máme $bx \in (a, b)$, teda aj $a + bx \in (a, b)$ a $(a + bx, b) \subseteq (a, b)$.

Podobne sa ukáže $a = (a + bx) - bx \in (a + bx, b)$ a $(a, b) \subseteq (a + bx, b)$. □

Ak postupne počítame zvyšky po delení, vieme ich vyjadriť ako kombináciu čísel a, b.

$$\begin{aligned}
a &= q_1 \cdot b + r_1 && N(r_1) < N(b) && r_1 = a - q_1.b \\
b &= q_2 \cdot r_1 + r_2 && N(r_2) < N(r_1) && r_2 = b - q_2.r_1 = (1 + q_1q_2)b - q_2a \\
r_1 &= q_3 \cdot r_2 + r_3 && N(r_3) < N(r_2) && r_3 = r_2 - q_3.r_2 = \cdots = x_3a + y_3b \\
&\vdots && && \vdots \\
r_{l-2} &= q_l \cdot r_{l-1} + r_l && N(r_l) < N(r_{l-1}) && r_l = r_{l-2} - q_l.r_{l-1} = \cdots = x_la + y_lb \\
r_{l-1} &= q_{l+1} \cdot r_l && \text{zvyšok } 0
\end{aligned}$$

Pretože v každom kroku norma zvyšku klesá, po istom čase sa algoritmus musí zastaviť a dostaneme nulový zvyšok. Navyše, z predchádzajúcej lemy vidíme, že v každom kroku platí $\gcd(r_k, r_{k-1}) = \gcd(a, b)$, preto na konci platí $\gcd(a, b) = \gcd(r_{l-1}, r_l) = \gcd(q_{l+1}r_l, r_l) = r_l$. Ďalej každý zvyšok sme vedeli vyjadriť v tvare $r_k = x_k a + y_k b$, kde $x_k, y_k \in R$, čiže týmto algoritmom vieme získať takéto vyjadrenie pre $\gcd(a, b)$.

Ukážeme si tento postup na konkrétnych príkladoch – najprv v $\mathbb{Z}$. (Najväčší spoločný deliteľ v $\mathbb{Z}$ viete zo strednej školy rátať pomocou rozkladu na prvočísla – niečo podobné platí všeobecne, ako uvidíme v tvrdení 13.4.39. Takýto postup nám však neposkytuje najväčší spoločný deliteľ ako kombináciu daných čísel – v príklade 13.4.26 uvidíme, že takéto vyjadrenie pre n.s.d. môže byť užitočné. Navyše to predpokladá, že poznáme rozklad na ireducibilné prvky – čo zatiaľ v $F[x]$ nevieme robiť vôbec, v $\mathbb{Z}$ to vieme robiť pre malé čísla. Pre veľké čísla je výpočtovo efektívnejší Euklidov algoritmus.)

{euklid:PREUKLID}

Príklad 13.4.25. Chceme vyrátať $d = \gcd(89, 16)$ a vyjadriť ho v tvare $89u + 16v$.

Keď použijeme viackrát vetu o delení so zvyškom, tak dostaneme:

$$\begin{aligned} 89 &= 5 \cdot 16 + 9 & 9 &= 89 - 5 \cdot 16 \\ 16 &= 1 \cdot 9 + 7 & 7 &= 16 - 9 = 6 \cdot 16 - 89 \\ 9 &= 1 \cdot 7 + 2 & 2 &= 9 - 7 = 2 \cdot 89 - 11 \cdot 16 \\ 7 &= 3 \cdot 2 + 1 & 1 &= 7 - 3 \cdot 2 = 39 \cdot 16 - 7 \cdot 89 \\ 2 &= 2 \cdot 1 + 0 \end{aligned}$$

Z lemy 13.4.24 potom vidíme, že $\gcd(89, 16) = \gcd(16, 9) = \gcd(9, 7) = \gcd(7, 2) = \gcd(2, 1) = 1$. V pravom stĺpci sme dostali hľadané vyjadrenie

$$1 = 39 \cdot 16 - 7 \cdot 89.$$

Tento postup môžeme prehľadne zapísať aj do tabuľky.

89	1	0	
16	0	1	
9	1	-5	`1r-5*2r`
7	-1	6	`2r-3r`
2	2	-11	`3r-4r`
1	-7	39	`4r-3*5r`

Tento postup do istej miery pripomína riadkové úpravy na matici. V každom riadku máme koeficienty, pomocou ktorých vieme číslo z prvého stĺpca vyjadriť ako celočíselnú kombináciu čísel 89 a 16.

Posledný stĺpec tabuľky sme doplnili len na to, aby bolo vidno, aké úpravy sme robili. Môže to byť užitočné pri hľadaní prípadnej chyby – tento stĺpec ale v podstate nie je nutný. Postupovali sme presne podľa vety o delení so zvyškom. Ak rátate niečo takéto ručne, pri malých číslach si občas môžete všimnúť aj nejaké veci, ktoré vám trochu urýchlia výpočet. Napríklad ak si všimnete, že $2 = 2.9 - 16$, ušetríte jeden riadok. (Treba si ale dávať pozor, či zrýchlený postup je správny – v podstate si stačí pamätať lemu 13.4.24 a postupovať podľa nej.)

89	1	0	
16	0	1	
9	1	-5	`1r-5*2r`
2	2	-11	`2*4r-3r`
1	-7	39	`3r-4*4r`

V predošlom príklade sme našli jednu dvojicu (u, v) takú, že $89u + 16v = 1$. Je to jediná možnosť? Vedeli by ste nájsť všetky ostatné také dvojice?

{euklid:PRINVZP}

Príklad 13.4.26. Inverzné prvky v poli $\mathbb{Z}_p$ (kde p je prvočíslo) sme zatiaľ vedeli počítať iba takým spôsobom, že sme postupne skúšali všetky prvky poľa. Euklidov algoritmus, ktorý sme sa teraz naučili, môžeme využiť na ten istý účel.

Pokúsme sa vypočítať 5^{-1} v $\mathbb{Z}_{13}$. Pretože 13 je prvočíslo platí $\gcd(5, 13) = 1$, čiže vieme nájsť čísla $x, y \in \mathbb{Z}$ také, že $1 = 5x + 13y$.

Postupným delením dostaneme

$$\begin{aligned} 13 &= 2 \cdot 5 + 3 & 3 &= 1 \cdot 13 - 2 \cdot 5 \\ 5 &= 1 \cdot 3 + 2 & 2 &= 5 - 3 = 3 \cdot 5 - 1 \cdot 13 \\ 3 &= 1 \cdot 2 + 1 & 1 &= 3 - 2 = 2 \cdot 13 - 5 \cdot 5 \end{aligned}$$

Ak pre všetky čísla v rovnosti $1 = 2.13 - 5.5$ urobíme zvyšok po delení 13, dostaneme rovnosť

$$1 = -5 \odot 5 = 8 \odot 5,$$

ktorá platí v $\mathbb{Z}_{13}$. Teda v $\mathbb{Z}_{13}$ platí $5^{-1} = 8$.

Opäť ten istý postup by sme mohli zapísať tabuľkou:

13	1	0	
5	0	1	
3	1	-2	`1r-2*2r`
2	-1	3	`2r-3r`
1	2	-5	`3r-4r`

Vyskúšajme si aspoň jeden konkrétny príklad v $\mathbb{Q}[x]$. Vieme, že najväčší spoločný deliteľ je určený jednoznačne až na asociovanosť – čiže v tomto prípade až na vynásobenie konštantou. Dohodnime sa, že si vyberieme ten, ktorý má vedúci koeficient 1 (tzv. normovaný polynóm) – potom už je najvúčší spoločný deliteľ určený jednoznačne.

Príklad 13.4.27. Vypočítajte $d(x) = \gcd(f(x), g(x))$ a vyjadrite ho v tvare $d(x) = u(x)f(x) + v(x)g(x)$ pre polynómy $f(x) = 3x^5 + 5x^4 - 16x^3 - 6x^2 - 5x - 6$, $g(x) = 3x^4 - 4x^3 - x^2 - x - 2$.

Podobne ako v predchádzajúcom príklade, budeme polynómy postupne deliť so zvyškom a zvyšok si v každom kroku vyjadríme ako kombináciu $f(x)$ a $g(x)$.

Kvôli prehľadnosti som zapísal zvlášť delenie polynómov a zvlášť vyjadrenie zvyšku v tvare kombinácie $f(x)$ a $g(x)$.

$$3x^5 + 5x^4 - 16x^3 - 6x^2 - 5x - 6 = (x + 3)(3x^4 - 4x^3 - x^2 - x - 2) - 3x^3 - 2x^2$$
$$3x^4 - 4x^3 - x^2 - x - 2 = (-3x^3 - 2x^2)(-x + 2) + 3x^2 - x - 2$$
$$-3x^3 - 2x^2 = (3x^2 - x - 2)(-x - 1) + (-3x - 2)$$

Vieme, že posledný nenulový zvyšok $-3x - 2$ v Euklidovom algoritme je hľadaný najväčší spoločný deliteľ. Pretože chceme dostať normovaný polynóm, vydelíme ho ešte vedúcim koeficientom -3.

$$\gcd(f(x), g(x)) = x + \frac{2}{3}$$

Zvyšky v jednotlivých deleniach vyjadríme pomocou $f(x)$ a $g(x)$ takto

$$-3x^3 - 2x^2 = f(x) - g(x)(x + 3)$$

$$\begin{aligned}3x^2 - x - 2 &= g(x) - (-3x^3 - 2x^2)(-x + 2) = \\ &g(x) - (f(x) - g(x)(x + 3))(-x + 2) = \\ &f(x)(x - 2) + [1 - (x - 2)(x + 3)g(x)] = \\ &(x - 2)f(x) - (x^2 + x - 7)g(x)\end{aligned}$$

$$\begin{aligned}-3x - 2 = &- 3x^3 - 2x^2 - (3x^2 - x - 2)(-x - 1) = \\ &f(x) - g(x)(x + 3) + [(x - 2)f(x) - (x^2 + x - 7)g(x)](x + 1) = \\ &f(x)[1 + (x - 2)(x + 1)] - g(x)[(x + 3) + (x + 1)(x^2 + x - 7)] = \\ &f(x)(x^2 - x - 1) - g(x)(x^3 + 2x^2 - 5x - 4)\end{aligned}$$

Po vydelení poslednej rovnosti číslom -3 dostávame

$$\gcd(f(x), g(x)) = x + \frac{2}{3} = -f(x)\frac{x^2 - x - 1}{3} + g(x)\frac{x^3 + 2x^2 - 5x - 4}{3}.$$

Opäť, pokiaľ by Vám to lepšie vyhovovalo, celý postup si môžete zapísať do tabuľky.

$f(x) = 3x^5 + 5x^4 - 16x^3 - 6x^2 - 5x - 6$	1	0
$g(x) = 3x^4 - 4x^3 - x^2 - x - 2$	0	1
$h_1(x) = f(x) - (x + 3)g(x) = -3x^3 - 2x^2$	1	$-(x + 3)$
$h_2(x) = g(x) + (x - 2)h_1(x) = 3x^2 - x - 2$	$x - 2$	$-(x^2 + x - 7)$
$h_3(x) = h_1(x) + (x + 1)h_2(x) = -3x - 2$	$x^2 - x - 1$	$-(x^3 + 2x^2 - 5x - 4)$
$h_2(x) + (x - 1)h_3(x) = 0$		

V predposlednom riadku sa naposledy vyskytol nenulový zvyšok, čiže ide o $\gcd(f(x), g(x))$. Z tohoto riadku vieme aj jeho vyčítať vyjadrenie – také isté, ako sme dostali v predchádzajúcom postupe. (Presnejšie povedané, dostali sme rovnaké vyjadrenie až na prenásobenie konštantou – vynormovanie.)

Pri výpočtoch takého typu ako sme robili v predchádzajúcom príklade sa celkom ľahko dá pomýliť – preto je užitočné občas (povedzme po každom kroku) vyskúšať, či rovnosti, ktoré sme dostali pre polynómy skutočne platí aj po dosadení nejakých čísel. (Je rozumné skúšať malé, čísla, napríklad 0, ± 1 – aby sa nám ľahko počítali hodnoty polynómu v týchto číslach.) Pri takejto čiastočnej skúške správnosti máme veľkú šancu prípadnú chybu odhaliť. (Samozrejme, dá sa urobiť skúška aj tak, že kombináciu $f(x)$ a $g(x)$, ktorú sme dostali, skutočne poroznásobujeme a zistíme, či vyjde rovnaký polynóm ako na druhej strane rovnosti – čo je však o dosť prácnejšie.)

Podobne ako pri počítaní racionálnych koreňov, ak v priebehu výpočtu nám vyjde ako jeden zo zvyškov polynóm, v ktorom všetky koeficienty sú násobkom toho istého celého čísla, môžeme polynóm týmto číslom vydeliť – dostaneme opäť polynóm s celočíselnými koeficientami (teda sa nám s ním bude dobre počítať) a neovplyvníme hodnotu najväčšieho spoločného deliteľa (v okruhu $F[x]$ sme tento polynóm zmenili len o deliteľ jednotky). Je ale dôležité pri vyjadrovaní najväčšieho spoločného deliteľa pomocou $f(x)$ a $g(x)$ nezabudnúť zarátať aj toto vydelenie.

13.4.3 Gaussove okruhy

Pojem analogický k pojmu prvočísla je v okruhu pojem ireducibilného prvku.

Definícia 13.4.28. Prvok $a \neq 0$ okruhu R sa nazýva *ireducibilný,* ak a je nenulový, nie je to deliteľ jednotky a ak z rovnosti $a = bc$ vyplýva, že niektorý z prvkov b, c je deliteľ jednotky v R.

Inými slovami, ireducibilný prvok sa (až na asociovanosť a výmenu poradia) nedá zapísať ako súčin dvoch prvkov z R inak ako $1 \cdot a$.

Príklad 13.4.29. Vieme, že prvočísla boli definované tak, že ich rozklad na súčin $p = a \cdot b$ je možný iba vtedy, ak niektoré z čísel a, b je rovné 1. Z toho vidno, že ireducibilné prvky v $\mathbb{Z}$ sú práve čísla tvaru $\pm p$, kde p je prvočíslo.

Ireducibilnými prvkami v okruhu $F[x]$ (volajú sa ireducibilné polynómy) sa budeme zaoberať neskôr.

Naším najbližším cieľom je dokázať, že v okruhoch hlavných ideálov platí tvrdenie zodpovedajúce rozkladu prirodzených (celých) čísel na súčin prvočísel.

Definícia 13.4.30. *Okruh s jednoznačným rozkladom* (alebo tiež *Gaussov okruh*) je obor integrity, v ktorom pre každý prvok $x \in R$, ktorý je nenulový a nie je deliteľom jednotky, existuje rozklad

$$x = p_1 \dots p_k$$

na súčin ireducibilných prvkov a navyše je tento rozklad jednoznačný až na asociovanosť a poradie.

Tvrdenie 13.4.31. *Ak ideál (p) v obore integrity R je vlastný prvoideál a $p \neq 0$, tak p je ireducibilný v R.*

Dôkaz. Ak (p) je prvoideál a $ab = p$, tak jeden prvok z dvojice a, b musí byť násobkom p (pretože patrí do (p)). Bez ujmy na všeobecnosti, nech $a = kp$. Potom $p = ab = (kb)p$, z čoho $kb = 1$ (lema 13.4.1), čiže b je deliteľ jednotky.

Keďže ideál p je vlastný, p nie je deliteľ jednotky. □

V OHI platí aj obrátená implikácia.

{euklid:TVROHIIREDPRVO}

Tvrdenie 13.4.32. *Ak p je ireducibilný prvok v OHI R, tak (p) je prvoideál.*

Dôkaz. Nech p je ireducibilný. Ukážeme, že ideál p je maximálny (a teda je to prvoideál). Nech by $(p) \subseteq (m)$. Z toho vyplýva $p = mc$. Potom buď m je asociovaný s p a $(p) = (m)$, alebo m je invertibilný a $(m) = R$. □

Z toho dostávame (pomocou (13.2)) nasledujúci veľmi dôležitý vzťah.

{gcc:DOSPRVOC}

Dôsledok 13.4.33. *V OHI pre ľubovoľný ireducibilný prvok p platí implikácia*

$$p \mid ab \qquad \Rightarrow \qquad p \mid a \vee p \mid b.$$

Teraz už sme schopní vysloviť a dokázať tvrdenie o rozklade na súčin ireducibilných prvkov.

{gcc:TVROHIJEUFD}

Tvrdenie 13.4.34. *Každý okruh hlavných ideálov je okruhom s jednoznačným rozkladom.*

Dôkaz. Chceme dokázať existenciu a jednoznačnosť rozkladu na súčin ireducibilných prvkov. Jednoznačnosť vyplýva z dôsledku 13.4.33.

Existencia. Sporom. Nech by x bol taký prvok, ktorý sa nedá v R rozložiť na súčin ireducibilných prvkov (pričom $x \neq 0$, x nie je deliteľ jednotky). Pretože x nie je ireducibilný, vieme ho zapísať ako $x = r_1 \cdot q_1$. Keby obidva prvky r_1 aj q_1 boli ireducibilné, máme rozklad x. Teda jeden z nich nie je ireducibilný, bez ujmy na všeobecnosti nech je to q_1. Potom $q_1 = r_2 \cdot q_2$ pre nejaké $r_2, q_2 \in R$. Takýmto spôsobom indukciou zostrojíme nekonečnú postupnosť prvkov $r_n \in R$ takú, že nasledujúci vždy delí predchádzajúci, teda $r_{n+1} \mid r_n$. To je ekvivalentné s tým,

že $(r_n) \subseteq (r_{n+1})$ a takto dostávame nekonečnú postupnosť ideálov $I_1 \subsetneq I_2 \subsetneq \ldots \subsetneq I_n \subsetneq \ldots$, kde I_k označuje ideál (r_k). Ukážeme, že v OHI takáto postupnosť nemôže existovať, čím dostaneme požadovaný spor.

Skutočne, ak by sme mali takýto rastúci reťazec ideálov. Potom aj $I = \bigcup_{n=1}^{\infty} I_n$ je ideál. Pretože R je OHI, existuje $a \in R$ také, že $(a) = I$. Lenže z toho, že $a \in \bigcup_{n=1}^{\infty} I_n$ vyplýva existencia čísla n_0 s vlastnosťou $a \in I_{n_0}$. Potom pre všetky $n > n_0$ máme $(a) \subseteq I_{n_0} \subseteq I_n \subseteq I$, čiže od n_0 počnúc sa už všetky ideály I_n rovnajú. □

Poznamenajme, že okruhy, ktoré spĺňajú podmienku, že v nich neexistuje nekonečný rastúci reťazec ideálov, sa nazývajú *noetherovské.*

Z predchádzajúceho tvrdenia špeciálne dostávame, že každé prirodzené číslo vieme napísať ako súčin prvočísel jednoznačne až na poradie. (A po pridaní deliteľov jednotky ± 1 dostaneme všetky prvé čísla.)

Analogickému tvrdeniu pre okruh polynómov $F[x]$ sa budeme venovať v nasledujúcej kapitole.

Skúsme nájsť príklad oboru integrity, ktorý nie je okruh s jednoznačným rozkladom.

Príklad 13.4.35. Budeme pracovať v okruhu $\mathbb{Z}[2i] = \{a + 2bi; a, b \in \mathbb{Z}\}$. Zrejme ide o obor integrity (je to podokruh poľa $\mathbb{C}$). Jediné delitele jednotky v tomto okruhu sú ± 1. Pozrime sa na rozklad $4 = 2.2 = (2i)(-2i)$.

Prvky 2 aj $\pm 2i$ sú ireducibilné. Ak totiž máme $2 = ab$, tak platí aj $2 = |a|.|b|$, pričom $|a| = |b|$ sú celé čísla. Potom pre niektoré z čísel a, b musí platiť, že má veľkosť 1. Takéto prvky v $\mathbb{Z}[i]$ sú však iba ± 1, zistili sme teda, že niektoré z čísel a, b je deliteľ jednotky. Tým sme overili, že 2 je ireducibilný prvok, zdôvodnenie pre $\pm 2i$ je presne rovnaké, opäť využijeme, že $|\pm 2i| = 2$.

Súčasne 2 je asociovaný iba s prvkami ± 2. Našli sme teda dva rozklady čísla 4 na súčin ireducibilných prvkov, ktoré sa nelíšia iba asociovanosťou. Teda $\mathbb{Z}[2i]$ nie je okruh s jednoznačným rozkladom.

Príklad 13.4.36. Ďalším takýmto príkladom je $\mathbb{Z}[\sqrt{5}i] = \{a + b\sqrt{5}i; a, b \in \mathbb{Z}\}$. V tomto okruhu máme rozklady

$$6 = 2.3 = (1 + \sqrt{5}i)(1 - \sqrt{5}i).$$

Vidno, že 2 nedelí žiaden z činiteľov na pravej strane. Ak ukážeme, že 2 je ireducibilný prvok, tak z dôsledku 13.4.33 vyplýva, že to nie je okruh s jednoznačným rozkladom.

Nech $2 = x.y$, kde $x, y \in \mathbb{Z}[\sqrt{5}i]$. Potom $|x| \leq 2$ aj $|y| \leq 2$, lebo všetky prvky tohoto okruhu majú vlastnosť $|x| \geq 1$. Ak $x = a + \sqrt{5}i$, tak sme dostali

$$|x|^2 = a^2 + 5b^2 \leq 4,$$

čo je možné jedine v prípade $b = 0$. Rozklad $2 = x.y$ je teda v skutočnosti rozklad na súčin dvoch celých čísel. V takomto rozklade musí byť nevyhnutne niektorý z činiteľov rovný ± 1.

Poznamenajme, že podobný spôsobom sa dá ukázať, že aj 3 a $1 \pm \sqrt{5}i$ sú ireducibilné.

Príklad 13.4.37. Dá sa dokázať, že ak R je okruh s jednoznačným rozkladom, tak aj okruh polynómov $R[x]$ je okruh s jednoznačným rozkladom. (Pozri napríklad [KGGS, Lema 7.4.1], [DF, Corollary 9.6]). Ak sme ochotní uveriť tomuto tvrdeniu, tak máme $\mathbb{Z}[x]$ ako príklad Gaussovho okruhu, ktorý nie je okruh hlavných ideálov. (Pozri príklad 13.4.18.)

V prípade, že máme rozklad prvkov a, b Gaussovho okruhu R, môžeme z neho zistiť, či $a \mid b$ ako aj určiť rozklad ich najväčšieho spoločného deliteľa $\gcd(a, b)$.

Lema 13.4.38. *Nech R je Gaussov okruh a $a, b \in R$. Ak $a = p_1 \dots p_n$ a $b = q_1 \dots q_m$ sú rozklady týchto prvkov na súčin ireducibilných činiteľov, tak $a \mid b$ práve vtedy, keď existuje injekcia $f\colon \{1, 2, \dots, n\} \to \{1, 2, \dots, m\}$ s vlastnosťou $q_{f(m)} \sim p_m$.*

(Toto tvrdenie je len formálny zápis faktu, že všetky ireducibilné prvky z rozkladu a sa musia vyskytnúť aj v rozklade b, pričom ak sa tam vyskytuje viackrát prvok z tej istej triedy asociovanosti, tak sa toľkokrát musí vyskytnúť aj v rozklade b.)

Dôkaz. □

{euklid:TVRGCDIRED}

Tvrdenie 13.4.39. *Nech R je Gaussov okruh, $a, b \in R \setminus \{0\}$. Majme tieto prvky vyjadrené v tvare $a = up_1^{k_1} \dots p_n^{k_n}$ a $b = u'p_1^{l_1} \dots p_n^{l_n}$, kde $u, u' \in U(R)$ a $p_1, \dots, p_n$ sú po dvoch neasociované ireducibilné prvky v R. Potom*

$$d = p_1^{m_1} \dots p_n^{m_n},$$

kde $m_i = \min\{k_i, l_i\}$ pre $i = 1, \dots, n$ je ich najväčší spoločný deliteľ.

Dôkaz. □

Cvičenia

Úloha 13.4.1. Ak u je deliteľ jednotky v okruhu R, tak aj $-u$ je deliteľ jednotky.

{euklidcvic:CUMRE}

Úloha 13.4.2.

Úloha 13.4.3. Nech R je euklidovský okruh a S je jeho podokruh, ktorý obsahuje jednotku. Musí byť aj S euklidovský okruh?

Úloha 13.4.4. Dokážte, že okruhy polynómov $\mathbb{Z}[x]$ a $\mathbb{Q}[x]$ nie sú izomorfné.

13.5 Okruhy polynómov II

V tejto časti sa budeme zaoberať polynómami, pričom často budeme využívať niektoré fakty, ktoré sme dokázali v predchádzajúcej podkapitole pre euklidovské okruhy, resp. pre okruhy s jednoznačným rozkladom. (Vieme, že $R[x]$ je euklidovský okruh, ak R je pole. Bez dôkazu sme si spomenuli, že ak R je Gaussov okruh, tak aj $R[x]$ je Gaussov okruh.)

13.5.1 Korene polynómov

{polyn2:SSECTKORENE}

Do polynómu $f(x) \in F[x]$ môžeme dosadiť ľubovoľný prvok c poľa F a vypočítať hodnotu polynómu v tomto prvku. (Zobrazenie, ktoré polynómu priradilo jeho hodnotu v c sme nazvali dosadzovací homomorfizmus – definícia 13.3.15.)

Definícia 13.5.1. Nech F je pole a F' je jeho nadpole. Prvok $c \in F'$ nazývame *koreňom* polynómu $f(x) \in F[x] \subset F'[x]$, ak $f(c) = 0$ (t.j. po dosadení c do polynómu F dostaneme 0).

V predchádzajúcej definícii dosadzujeme do polynómu z $F[x]$ prvok z nadpoľa F'. To však nie je problém – keďže koeficienty polynómu $f(x)$ sú z $F \subseteq F'[x]$, tento polynóm súčasne patrí do $F'[x]$.

Príklad 13.5.2. Číslo i je koreňom polynómu $x^2 + 1$, lebo $i^2 + 1 = 0$.

Všimnime si, aký je vzťah medzi koreňmi polynómu a deliteľnosťou lineárnymi polynómami.

LMZVYSOKJEFC}

Lema 13.5.3. *Ak $f(x) \in F[x]$, kde F je pole, a $c \in F$, tak zvyšok polynómu $f(x)$ po delení polynómom $x - c$ je rovný $f(c)$, t.j. existuje polynóm $g(x) \in F[x]$ taký, že*

EQZVYSOKJEFC}

$$f(x) = (x - c)g(x) + f(c). \tag{13.3}$$

Dôkaz. Z vety o delení so zvyškom vieme

$$f(x) = g(x)(x - c) + r,$$

pričom zvyšok je polynóm stupňa menšieho ako 1, preto je to nejaká konštanta $r \in F$.

Ak do predošlej rovnosti dosadíme c za x, tak máme

$$f(c) = g(c)(c - c) + r = r,$$

čiže táto konštanta musí byť rovná práve $f(c)$, t.j. hodnote polynómu f v bode c. □

Z predchádzajúcej lemy už ľahko dostaneme

{polyn2:LMLINDELI}

Lema 13.5.4. *Nech F je pole a F' je jeho nadpole. Nech $f(x) \in F[x]$. Potom $c \in F'$ je koreňom $f(x)$ práve vtedy, keď $x - c \mid f(x)$ v $F'[x]$, t.j. existuje polynóm $g(x) \in F'[x]$ taký, že $f(x) = g(x)(x - c)$.*

Dôkaz. Podľa (13.5.3) máme $f(x) = (x - c)g(x) + f(c)$, čiže ak $f(c) = 0$, tak $f(x) = (x - c)g(x)$, čiže $x - c \mid f(x)$.

Obrátene, ak $x - c \mid f(x)$, tak zvyšok po delení polynómu $f(x)$ polynómom $x - c$ je 0, čiže (opäť z lemy 13.5.3) $f(c) = 0$ a c je koreň polynómu f. □

{polyn2:DEFNASKOR}

Definícia 13.5.5. Nech F' je nadpole poľa F, $f(x) \in F[x]$ a c je koreň $f(x)$. Hovoríme, že *násobnosť* koreňa c je k (alebo tiež, že c je k-násobný koreň $f(x)$), ak $(x - c)^k \mid f(x)$ (t.j. ak existuje polynóm $g(x) \in F'[x]$ taký, že $f(x) = g(x)(x - c)^k$) a súčasne $(x - c)^{k+1} \nmid f(x)$.

Pre $k = 1$ voláme k-násobný koreň *jednoduchý koreň* polynómu $f(x)$, ak $k > 1$ tak hovoríme o násobnom koreni.

Príklad 13.5.6. Čísla ± 1 sú dvojnásobné korene polynómu $x^4 - 2x^2 + 1$, lebo $x^4 - 2x^2 + 1 = (x^2 - 1)^2 = (x - 1)^2(x + 1)^2$

Jednoduchý spôsob ako ručne spočítať hodnotu polynómu v danom čísle (a tým zistiť, či toto číslo je koreňom polynómu) je použitie Hornerovej schémy.

Základná idea Hornerovej schémy je, že hodnotu polynómu môžeme vyjadriť ako

$$a_n c^n + a_{n-1} c^{n-1} + \cdots + a_0 = \left(a_n c^{n-1} + \cdots + a_1\right) c + a_0 = \\ ((\ldots(a_n c + a_{n-1})c + \ldots)c + a_1)c + a_0$$

Stačí nám teda postupne počítať čísla a_n, $a_n c + a_{n-1}$, $(a_n c + a_{n-1})c + a_{n-2}$ atď., t.j. predchádzajúci výsledok vždy vynásobíme číslom c a pripočítame k nemu nasledujúci koeficient.

Príklad 13.5.7. Vypočítajte hodnotu polynómu $f(x) = x^4 - 3x^3 + 2x - 1$ nad poľom $\mathbb{R}$ v bode $c = 2$.

Do tabuľky si zapíšeme koeficienty polynómu (dôležité je nezabudnúť na nulový koeficient pochádzajúci z člena $0x^2$) a postupujeme postupom, ktorý sme naznačili.

	1	-3	0	2	-1
2		2	-2	-4	-4
	1	-1	-2	-2	$\boxed{-5}$

Všimnime sme, že súčasne sme vypočítali, že

$$x^4 - 3x^3 + 2x - 1 = (x^3 - x^2 - 2x - 2)(x - 2) - 5.$$

(Stačí si uvedomiť, že pri Hornerovej schéme vlastne robíme to isté, čo pri algoritme na delenie polynómov.)

Aby sme si uvedomili, čo vlastne v Hornerovej schéme počítame, pokúsme sa ju zapísať o čosi všeobecnejšie (kvôli šírke rozdelené na 2 tabuľky)

	a_n	a_{n-1}	a_{n-2}	$\dots$
c		$a_n c$	$(a_n c + a_{n-1})c$	$\dots$
	a_n	$a_n c + a_{n-1}$	$a_n c^2 + a_{n-1}c + a_{n-2}$	$\dots$

$\dots$	a_1	a_0	
$\dots$	$\dots$	$(a_n c^{n-1} + a_{n-1}c^{n-2} + \dots + a_1)c$	
$\dots$	$a_n c^{n-1} + a_{n-1}c^{n-2} + \dots + a_1$	$a_n c^n + a_{n-1}c^{n-1} + \dots + a_1 c + a_0 = f(c)$	

Príklad 13.5.8. Overte, že 1 je koreňom polynómu $f(x) = x^4 - 3x^3 + 3x - 1 \in \mathbb{R}[x]$. Zistite násobnosť tohoto koreňa.

Budeme postupovať pomocou Hornerovej schémy – pri vypočítaní hodnoty $f(1)$ súčasne nájdeme polynóm $g(x)$ taký, že $f(x) = g(x)(x - 1) + f(1)$. Ak $f(1) = 0$, na zistenie, či ide násobnosť tohoto koreňa je aspoň 2, stačí overiť, či aj $g(1) = 0$. Analogicky postupujeme ďalej, až kým nedostaneme nenulový zvyšok.

	1	-3	0	3	-1
1		1	-2	-2	1
	1	-2	-2	1	$\boxed{0}$
1		1	-1	-3	
	1	-1	-3	$\boxed{-2}$	

Zistili sme, že 1 je jednoduchým (jednonásobným) koreňom polynómu $f(x)$ a že

$$f(x) = (x - 1)(x^3 - 2x^2 - 2x + 1),$$

pričom $x - 1 \nmid x^3 - 2x^2 - 2x + 1$.

Rátať korene polynómov je vo všeobecnosti ťažká úloha. Zo strednej školy poznáte vzorec na hľadanie koreňov polynómov druhého stupňa – kvadratických polynómov. (Podobné vzorce, aj keď zložitejšie, sa dajú nájsť aj pre rovnice tretieho a štvrtého stupňa. Vo všeobecnosti však také vzorce neexistujú.) Okrem nich vieme ešte v komplexných číslach riešiť binomické rovnice, t.j. rovnice tvaru $x^n = a$, kde $a \in \mathbb{C}$ (pozri C.3.2 alebo [KGGS, kapitola 6.1]).

Povieme si, ako pre polynóm s celočíselnými koeficientami vieme nájsť všetky korene, ktoré sú racionálnymi číslami (t.j. všetky korene daného polynómu ležiace v poli $\mathbb{Q}$).

13.5.2 Racionálne korene polynómu s celočíselnými koeficientami

olyn2:RACKOR}

Tvrdenie 13.5.9. *Ak $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0 \in \mathbb{Z}[x]$ je polynóm s celočíselnými koeficientami a racionálne číslo $c = \frac{p}{q}$ je koreň $f(x)$ (pričom $\gcd(p,q) = 1$, t.j. racionálne číslo c je zapísané v základnom tvare), tak*

$$p \mid a_0 \qquad a \qquad q \mid a_n.$$

Dôkaz. Ak $c = \frac{p}{q}$ je koreň $f(x)$, tak máme rovnosť

$$f(c) = a_n \frac{p^n}{q^n} + a_{n-1}\frac{p^{n-1}}{q^{n-1}} + \cdots + a_1 \frac{p}{q} + a_0 = 0.$$

Ak túto rovnosť vynásobíme q^n, dostaneme

$$a_n p^n + a_{n-1}p^{n-1}q + \cdots + a_1 p q^{n-1} + a_0 q^n = 0.$$

(Všimnime si, že v predchádzajúcej rovnosti vystupujú iba celé čísla.)

Túto rovnosť môžeme upraviť ako

$$-a_n p^n = (a_{n-1}p^{n-1} + \cdots + a_1 p q^{n-2} + a_0 q^{n-1})q,$$

čo znamená, že $q \mid a_n p^n$. Pretože $\gcd(p,q) = 1$ (p a q sú nesúdeliteľné), vyplýva z toho $q \mid a_n$ (dôsledok 13.4.23).

Pri dôkaze toho, že $p \mid a_0$ postupujeme takmer rovnako. Máme

$$-a_0 q^n = (a_n p^{n-1} + a_{n-1}p^{n-2}q + \cdots + a_1 q^{n-1})p,$$

čiže $p \mid a_0 q^n$, a teda (na základe nesúdeliteľnosti) $p \mid a_0$. □

Predchádzajúce tvrdenie môžeme použiť na nájdenie všetkých racionálnych koreňov daného polynómu zo $\mathbb{Z}[x]$. Predchádzajúce tvrdenie nám poskytuje obmedzenie na všetkých možných kandidátov na korene v množine racionálnych čísel. Postupným vyskúšaním nájdeme všetky korene, ktoré patria do $\mathbb{Q}$.

Ďalšie obmedzenie, ktoré nám môže pomôcť pri skúšaní jednotlivých možností, nám poskytne nasledujúce pozorovanie (ktorého špeciálnym prípadom je tvrdenie 13.5.9).

{polyn2:RACKOR2}

Tvrdenie 13.5.10. *Nech $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0 \in \mathbb{Z}[x]$ je polynóm s celočíselnými koeficientami a racionálne číslo $c = \frac{p}{q}$ je koreň $f(x)$ (pričom $\gcd(p,q) = 1$, t.j. racionálne číslo c je zapísané v základnom tvare). Nech $g(x) = b_{n-1}x^{n-1} + \cdots + b_0$ je polynóm z $\mathbb{Q}[x]$ taký, že*

$$f(x) = g(x)\left(x - \frac{p}{q}\right). \tag{13.4}$$

{polyn2:EQSUCIN}

Potom aj $g(x) \in \mathbb{Z}[x]$, t.j. koeficienty polynómu $g(x)$ sú celočíselné.

Budeme sa snažiť dokázať toto tvrdenie indukciou. Ale začnime tým, že sa pozrieme aspoň na prvé dva prípady - z toho azda budeme vedieť vymyslieť, čo vlastne chceme indukciou dokazovať (čo všetko potrebujeme, aby prešiel indukčný krok).

Dôkaz. V dôkaze budeme samozrejme využívať rovnosť (13.4), ktorá nám vlastne dáva vzťah medzi koeficientami polynómu $f(x)$ a koeficientami polynómu $g(x)$.

Okrem toho budeme často používať to, že vieme

$$a_n \frac{p^n}{q^n} + a_{n-1}\frac{p^{n-1}}{q^{n-1}} + \cdots + a_1 \frac{p}{q} + a_0 = 0. \tag{13.5}$$

{polyn2:EQDOSAD}

Pre najvyšší koeficient polynómu $g(x)$ platí $b_{n-1} = a_n$.

Všimnime si tiež, že z rovnosti (13.5) máme

$$b_{n-1}\frac{p^n}{q^n} = -a_{n-1}\frac{p^{n-1}}{q^{n-1}} - \cdots - a_1\frac{p}{q} - a_0.$$

Ak túto rovnosť prenásobíme q^n, tak máme

$$b_{n-1}p^n = -(a_{n-1}p^{n-1} + a_{n-2}p^{n-2}q \cdots + a_1pq^{n-2} + a_0q_{n-1})q.$$

Pretože výraz v zátvorke je celé číslo, máme $q \mid b_{n-1}p^n$. Z toho, že p a q sú nesúdeliteľné, dostávame $q \mid b_{n-1}$. (Vlastne sme zatiaľ iba zopakovali úvahu z dôkazu tvrdenia 13.5.9. Ale asi sa ju oplatí zopakovať, keďže podobnú úvahu budeme používať aj ďalej.)

Pozrime sa na ďalší koeficient. Tento koeficient môžeme vyjadriť ako

$$b_{n-2} = b_{n-1}\frac{p}{q} + a_{n-1} = a_n\frac{p}{q} + a_{n-1}.$$

(Prvú rovnosť dostaneme z Hornerovej schémy. Alebo tiež z porovnania koeficientov pri x^{n-1} v polynómoch $f(x)$ a $g(x)\left(x - \frac{p}{q}\right)$ vidíme, že $a_{n-1} = b_{n-2} - b_{n-1}\frac{p}{q}$.)

Pozrime sa na výraz

$$b_{n-2}\frac{p^{n-1}}{q^{n-1}} = a_n\frac{p^n}{q^n} + a_{n-1}\frac{p^{n-1}}{q^{n-1}},$$

ktorý môžeme použitím (13.5) prepísať do tvaru

$$b_{n-2}\frac{p^{n-1}}{q^{n-1}} = -a_{n-2}\frac{p^{n-2}}{q^{n-2}} - \cdots - a_1\frac{p}{q} - a_0.$$

Opäť stačí túto rovnosť vynásobiť q^{n-1} a dostaneme

$$b_{n-2}p^{n-1} = -(a_{n-2}p^{n-2} + a_{n-3}p^{n-3}q \cdots + a_1pq^{n-3} + a_0q_{n-2})q.$$

A podobne ako pri predošlom koeficiente, z toho, že p a q sú nesúdeliteľné máme $q \mid b_{n-2}$.

Teraz sa už dá uhádnuť, že sa zrejme budeme snažiť indukciou dokázať tieto dve veci pre $k = 1, \ldots, n-1$: Platí rovnosť

{polyn2:EQIND}
$$b_{n-k} = a_n\frac{p^{k-1}}{q^{k-1}} + a_{n-1}\frac{p^{k-2}}{q^{k-2}} + \cdots + a_{n-k+1} \tag{13.6}$$

a navyše platí, že b_{n-k} je celé číslo, ktoré je deliteľné číslom q.

Indukčný krok bude vyzerať takto: Predpokladáme, že uvedené tvrdenie platí pre k, pričom $k < n-1$. Chceme dokázať, že platí aj pre $k+1$. Máme rovnosť

$$b_{n-(k+1)} = b_{n-k}\frac{p}{q} + a_{n-k}.$$

Ak za b_{n-k} dosadíme výraz, ktorý máme z indukčného predpokladu, tak dostaneme

$$\begin{aligned} b_{n-(k+1)} &= \left(a_n\frac{p^{k-1}}{q^{k-1}} + a_{n-1}\frac{p^{k-2}}{q^{k-2}} + \cdots + a_{n-k+1}\right)\frac{p}{q} + a_{n-k} \\ &= a_n\frac{p^k}{q^k} + a_{n-1}\frac{p^{k-1}}{q^{k-1}} + \cdots + a_{n-k+1}\frac{p}{q} + a_{n-k} \end{aligned}$$

Teda aj pre $k+1$ platí rovnosť (13.6).

Po vynásobení predošlej rovnosti číslom $\frac{p^{n-k}}{q^{n-k}}$ máme

$$b_{n-(k+1)}\frac{p^{n-k}}{q^{n-k)}} = a_n\frac{p^n}{q^n} + a_{n-1}\frac{p^{n-1}}{q^{n-1}} + \cdots + a_{n-k+1}\frac{p^{n-k+1}}{q^{n-k+1}} + a_{n-k}\frac{p^{n-k}}{q^{n-k}}.$$

z čoho použitím (13.5) dostaneme

$$b_{n-(k+1)}\frac{p^{n-k}}{q^{n-k}} = -a_{n-k-1}\frac{p^{n-k-1}}{q^{n-k-1}} - a_{n-k-2}\frac{p^{n-k-2}}{q^{n-k-2}} - \cdots - a_1\frac{p}{q} - a_0.$$

Opäť stačí túto rovnosť vynásobiť číslom q^{n-k} a máme

$$b_{n-(k+1)}p^{n-k} = (-a_{n-k-1}p^{n-k-1} - a_{n-k-2}p^{n-k-2}q - \cdots - a_1pq^{n-k-2} - a_0q^{n-k-1})q.$$

A znovu si stačí všimnúť, že číslo v zátvorke na pravej strane rovnosti je celé, čiže pravá strana je násobok q. Na základe toho, že p a q sú nesúdeliteľné, dostaneme $q \mid b_{n-(k+1)}$. □

Z predchádzajúceho tvrdenia vyplýva, že ak overujeme, či nejaké racionálne číslo je koreňom polynómu s celočíselnými koeficientami, v okamihu, keď nám v priebehu výpočtu vyjde v spodnom riadku zlomok, už nemusíme rátať ďalej. (Vieme totiž, že čísla v spodnom riadku Hornerovej schémy sú presne koeficienty polynómu $g(x)$, teda ak je dané racionálne číslo koreňom, musia všetky tieto koeficienty podľa predchádzajúceho tvrdenia byť celé čísla.)

Ukážme si teda hľadanie racionálnych koreňov daného polynómu zo $\mathbb{Z}[x]$ na konkrétnom príklade.

{polyn2:PRRACKOR}

Príklad 13.5.11. Nájdite racionálne korene polynómu $f(x) = 24x^5+10x^4-x^3-19x^2-5x+6$ (aj s násobnosťami).

Podľa tvrdenia 13.5.9 má platiť $p \mid 6$, $q \mid 24$. Dostávame teda možnosti:
$p \in \{\pm1, \pm2, \pm3, \pm6\}$
$q \in \{1, 2, 3, 4, 6, 8, 12, 24\}$
$\frac{p}{q} \in \{\pm1, \pm\frac{1}{2}, \pm\frac{1}{3}, \dots\}$
(Pre q nám stačí skúšať kladné hodnoty, pretože voľba znamienok pre číslo p nám zabezpečí obidve možnosti – kladné aj záporné korene.)

Začnime najprv skúšať tých kandidátov na korene, kde čitateľ je ±1.

	24	10	-1	-19	-5	6
1		24	34	33	14	9
	24	34	33	14	9	15

	24	10	-1	-19	-5	6
-1		-24	14	-13	32	-27
	24	-14	13	-32	27	−21

	24	10	-1	-19	-5	6
$\frac{1}{2}$		12	11	5	-7	-6
	24	22	10	-14	-12	0
$\frac{1}{2}$		12	17	$\frac{27}{2}$		
	24	34	27	$-\frac{1}{2}$	$\neq 0$	

Zistili sme, že $\frac{1}{2}$ je jednoduchý koreň polynómu $f(x)$. (V poslednom výpočte sme nerátali do konca – zastavili sme sa pri zlomku $-\frac{1}{2}$.)

Mohli by sme pokračovať v skúšaní možností ďalej, trochu nám však zjednoduší prácu, ak si uvedomíme, že všetky ďalšie korene musia byť koreňmi polynómu $g(x) = 24x^4 + 22x^3 + 10x^2 - 14x - 12$. (Tento polynóm je podiel polynómu $f(x)$ a polynómu $x - \frac{1}{2}$, jeho koeficienty vieme vyčítať z predchádzajúcej Hornerovej schémy.)

Každý koeficient tohoto polynómu je párny – môžeme teda celý polynóm vydeliť číslom 2 a dostaneme polynóm $12x^4 + 11x^3 + 5x^2 - 7x - 6$, ktorý má tiež celočíselné koeficienty a má rovnaké korene ako $g(x)$. Keď hľadáme racionálne korene tohoto polynómu, dostávame pre čitateľ a menovateľ podmienky $p \mid 6$, $q \mid 12$, čiže
$p \in \{\pm 1, \pm 2, \pm 3, \pm 6\}$
$q \in \{1, 2, 3, 4, 6, 12\}$
$\frac{p}{q} \in \{\pm 1, \pm\frac{1}{2}, \pm\frac{1}{3} \pm \frac{1}{4}, \dots\}$
Pritom samozrejme čísla, ktoré sme už vyskúšali pre $f(x)$, pre polynóm $g(x)$ skúšať nemusíme. Získali sme teda dve zjednodušenia – budeme pracovať s polynómom nižšieho stupňa a máme menej možností, ktoré treba vyskúšať.

	12	11	5	-7	-6
$-\frac{1}{2}$		-6	$-\frac{5}{2}$		
	12	5	$\frac{5}{2}$		$\neq 0$

	12	11	5	-7	-6
$\frac{1}{3}$		4	5	$\frac{10}{3}$	
	12	15	10	$\frac{-11}{3}$	$\neq 0$

	12	11	5	-7	-6
$-\frac{1}{3}$		-4	$-\frac{7}{3}$		
	12	7	$-\frac{28}{3}$		$\neq 0$

	12	11	5	-7	-6
$\frac{1}{4}$		3	$\frac{14}{4}$		
	12	14			$\neq 0$

	12	11	5	-7	-6
$-\frac{1}{4}$		-2	$\frac{9}{4}$		
	12	9			$\neq 0$

	12	11	5	-7	-6
$\frac{1}{6}$		2	$\frac{13}{6}$		
	12	13			$\neq 0$

	12	11	5	-7	-6
$-\frac{1}{6}$		-2	$\frac{9}{6}$		
	12	9			$\neq 0$

	12	11	5	-7	-6
2		24	70	150	286
	12	35	75	143	$\boxed{280}$

	12	11	5	-7	-6
-2		-24	26	-62	138
	12	-13	31	-69	$\boxed{132}$

	12	11	5	-7	-6
$\frac{2}{3}$		8	$\frac{38}{3}$		
	12	19			$\neq 0$

	12	11	5	-7	-6
$-\frac{2}{3}$		-8	-2	-2	6
	12	3	3	-9	$\boxed{0}$
$-\frac{2}{3}$		-8	$\frac{10}{3}$		
	12	-5		$\neq 0$	

Dostali sme ďalší jednoduchý koreň $-\frac{2}{3}$. Nový polynóm, s ktorým budeme pracovať, je $h(x) = 12x^3+3x^2+3x-9$. Po vydelení koeficientov číslom 3 dostaneme jednoduchší polynóm $4x^3 + x^2 + x - 3$ a podmienky pre korene $p \mid 3$, $q \mid 4$, čiže
$p \in \{\pm 1, \pm 3\}$
$q \in \{1, 2, 4\}$
$\frac{p}{q} \in \{\pm 1, \pm\frac{1}{2}, \pm\frac{1}{4}, \pm 3, \pm\frac{3}{2} \pm \frac{3}{4}\}$

	4	1	1	-3
3		12	39	120
	4	13	40	$\boxed{117}$

	4	1	1	-3
-3		-12	33	-102
	4	-11	34	$\boxed{-105}$

	4	1	1	-3
$\frac{3}{2}$		6	$\frac{21}{2}$	
	4	7		$\neq 0$

	4	1	1	-3
$-\frac{3}{2}$		-6	$-\frac{15}{2}$	
	4	-5		$\neq 0$

	4	1	1	-3
$\frac{3}{4}$		3	3	3
	4	4	4	$\boxed{0}$
$\frac{3}{4}$		3	$\frac{21}{4}$	
	4	7	$\neq 0$	

Našli sme ďalší jednoduchý koreň $\frac{3}{4}$

Ďalej môžeme pracovať s polynómom $x^2 + x + 1$. Tu sú však jediní možní kandidáti na korene čísla ± 1 a tie sme už vyskúšali.

Záver: Daný polynóm má tieto 3 racionálne korene: $\frac{1}{2}$, $-\frac{2}{3}$, $\frac{3}{4}$; násobnosť každého z nich je 1.

Všimnime si, že sme vlastne súčasne dostali, že

$$f(x) = 24x^5 + 10x^4 - x^3 - 19x^2 - 5x + 6 = 24(x - \frac{1}{2})(x + \frac{2}{3})(x - \frac{3}{4})(x^2 + x + 1).$$

(Pri poslednom delení nám vyšiel podiel $4(x^2 + x + 1)$ a v priebehu výpočtu sme polynóm vydelili raz číslom 2 a raz číslom 3.) Predchádzajúcu rovnosť môžeme tiež prepísať ako

$$f(x) = (2x - 1)(3x + 2)(4x - 3)(x^2 + x + 1).$$

Pozrime sa na to, ako z tvrdenia 13.5.9 vyplýva, že $\sqrt{2}$ nie je racionálne číslo.

Príklad 13.5.12. Číslo $\sqrt{2}$ je očividne koreňom polynómu $p(x) = x^2 - 2$.

Ak by tento polynóm mal racionálny koreň, tak na základe tvrdenia 13.5.9 to môže byť jedine niektoré z čísel ± 1, ± 2. Ľahko skontrolujeme, že čísla ± 1 ani ± 2 nevyhovujú danej rovnici.

Teda polynóm $p(x)$ nemá racionálne korene a číslo $\sqrt{2}$ je iracionálne.

Môžete si všimnúť, že zdôvodnenie z tohoto príkladu prejde bez zmeny ak číslo 2 nahradíme ľubovoľným prvočíslom.

13.5.3 Algebraicky uzavreté polia

Definícia 13.5.13. Pole F sa nazýva *algebraicky uzavreté,* ak každý polynóm $f(x) \in F[x]$ stupňa aspoň jedna má v poli F aspoň jeden koreň.

V prípade, že $f(x)$ má koreň c, môžeme ho vydeliť koreňovým činiteľom $x - c$ a dostaneme jeho deliteľ nižšieho stupňa. Ten opäť musí mať nejaký koreň (ak nie je konštatný), preto takýmto spôsobom postupne dostaneme rozklad polynómu $f(x)$ na koreňové činitele. Dostávame:

Tvrdenie 13.5.14. *Ak F je algebraicky uzavreté pole, tak každý polynóm $f(x)$ je v $F[x]$ rozložiteľný na koreňové činitele.*

Z toho ďalej vidno, že ak F je algebraicky uzavreté pole, tak súčet násobností koreňov polynómu $f(x)$ je rovný jeho stupňu. (Toto tvrdenie sa zvyčajne formuluje tak, že polynóm stupňa n má práve n koreňov, ak zarátame aj ich násobnosti.)

Vieme, že pole komplexných čísel $\mathbb{C}$ má túto vlastnosť (aj keď dôkaz tejto vety nie je jednoduchý).

Veta 13.5.15 (Základná veta algebry)**.** *Pole komplexných čísel $\mathbb{C}$ je algebraicky uzavreté.*

Spomeňme (opäť bez dôkazu), že ku každému poľu sa dá zostrojiť nadpole, v ktorom už každý polynóm z $F[x]$ bude mať koreň. Dokonca platí:

{polyn2:VTSTEINITZ}

Veta 13.5.16 (Steinitz)**.** *Pre každé pole F existuje algebraicky uzavreté nadpole F'.*

Všimnime si ešte jednu užitočnú vlastnosť komplexných koreňov polynómov s reálnymi koeficientami.

{polyn2:TVRKOMPZDR}

Tvrdenie 13.5.17. *Ak $f(x) \in \mathbb{R}[x]$ je polynóm s reálnymi koeficientami a $z = a + bi \in \mathbb{C}$ je koreň polynómu $f(x)$, tak aj komplexne združené číslo $\overline{z} = a - bi$ je koreňom polynómu $f(x)$. Pritom násobnosť koreňa $\overline{z}$ je rovnaká ako násobnosť z.*

Dôkaz. Stačí si všimnúť, že zobrazenie $z \mapsto \overline{z}$ je homomorfizmus (súčet/súčin komplexne združených čísel je komplexne združené číslo k súčtu/súčinu) a že pre $z \in \mathbb{R}$ platí $\overline{z} = z$. Z toho potom dostávame rovnosť

$$\overline{f(z)} = \overline{a_n z^n + a_{n-1} z^{n-1} + \cdots + a_0} = a_n (\overline{z})^n + a_{n-1} (\overline{z})^{n-1} + \cdots + a_0 = f(\overline{z})$$

pre ľubovoľné $z \in \mathbb{C}$.

Z tejto rovnosti špeciálne vyplýva, že ak $f(z) = 0$, tak aj $f(\overline{z}) = 0$.

Druhá časť vyplýva z prvej použitej pre polynóm zapísaný v tvare $f(x) = g(x)(x - z)^k$, kde k je násobnosť koreňa z. □

Veľmi prirodzeným zovšeobecnením tohoto výsledku je tvrdenie sformulované v úlohe 13.5.1.

DOSREALNEPAR}

Dôsledok 13.5.18. *Každý polynóm $f(x) \in \mathbb{R}[x]$ nepárneho stupňa má aspoň 1 reálny koreň.*

Dôkaz. Ak by polynóm mal iba komplexné korene, tak môžeme popárovať dvojice komplexne združených koreňov. Komplexne združené korene majú podľa predchádzajúceho tvrdenia rovnakú násobnosť. Preto súčet násobností všetkých komplexných koreňov je párne číslo. Súčet násobností sa však rovná stupňu polynómu $f(x)$ (pretože $\mathbb{C}$ je algebraicky uzavreté pole). □

13.5.4 Ireducibilné polynómy

Definícia 13.5.19. Polynóm $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0$ sa nazýva *normovaný* (alebo tiež *monický,*) ak $a_n = 1$ (vedúci koeficient sa rovná 1).

Definícia 13.5.20. Ak R je obor integrity, tak ireducibilné prvky okruhu $R[x]$ nazývame *ireducibilné polynómy* v $R[x]$.

V prípade, že ide o pole, tak z predchádzajúcej kapitoly vieme, že $F[x]$ je euklidovský okruh (a teda je to aj okruh hlavných ideálov a okruh s jednoznačným rozkladom). Tento fakt nám umožní používať všetky výsledky z predchádzajúcej kapitoly aj pre polynómy nad nejakým poľom.

Veta 13.5.21 (Rozklad na ireducibilné polynómy)**.** *Ak F je pole, tak každý polynóm $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0$ možno vyjadriť v tvare*

$$f(x) = a_n p_1(x) \dots p_n(x),$$

kde $p_1, \dots, p_n$ sú ireducibilné normované polynómy. Navyše, tento rozklad je (až na poradie činiteľov) jednoznačne určený.

Dôkaz. Pretože $F[x]$ je okruh s jednoznačným rozkladom, vieme, že každý polynóm sa dá rozložiť na súčin ireducibilných polynómov a ten rozklad je jednoznačný až na asociovanosť. V okruhu $F[x]$ sú dva prvky asociované práve vtedy, keď sa líšia iba konštantným násobkom. Tým, že vo vete požadujeme normované polynómy, sú teda už jednoznačne určené (z ľubovoľného polynómu dostaneme normovaný, keď ho vynásobíme b_m^{-1}, kde b_m je jeho vedúci koeficient; súčin vedúcich koeficientov sme dali pred súčin normovaných činiteľov – tento súčin sa rovná a_n). □

Zatiaľ však o ireducibilných polynómoch vieme iba to, že existujú – nevieme, ako overiť, či je daný polynóm ireducibilný ani ako rozklad na súčin ireducibilných polynómov hľadať.

Je zrejmé, že každý polynóm stupňa 1 je ireducibilný – nedá sa rozložiť na súčin polynómov nižších stupňov. Teda ak c je k-násobný koreň, v rozklade polynómu $f(x)$ sa musí vyskytnúť $(x-c)^k$. V prípade, že súčet násobností koreňov je rovný stupňu polynómu vieme teda ten polynóm rozložiť ako

$$f(x) = a_n(x-c_1)^{k_1}(x-c_2)^{k_2} \dots (x-c_m)^{k_m},$$

kde $c_1, \dots, c_m$ sú všetky korene $f(x)$ a $k_1, \dots, k_m$ sú ich násobnosti. Takýto rozklad (ak existuje) voláme rozklad na súčin *koreňových činiteľov.*

V niektorých prípadoch vieme o ireducibilite rozhodnúť, ak poznáme korene polynómu.

{polyn2:TVRIREDSTUP23}

Tvrdenie 13.5.22. *Ak F je pole a $f(x) \in F[x]$ je polynóm stupňa 2 alebo 3, tak polynóm $f(x)$ je ireducibilný v F práve vtedy, keď $f(x)$ nemá koreň v F.*

Dôkaz. Stačí si všimnúť, že ak chceme polynóm stupňa 2 alebo 3 rozložiť ako súčin polynómov nižších stupňov, nevyhnutne sa tam musí vyskytnúť polynóm stupňa 1. Z lemy 13.5.4 vieme, ako súvisia lineárne delitele polynómu a jeho korene. □

Všimnime si, že ireducibilita polynómu závisí od toho, nad akým poľom ho uvažujeme (pretože polynóm nad poľom F môžeme súčasne chápať aj ako polynóm nad ľubovoľným nadpoľom $F' \supseteq F$).

Príklad 13.5.23. Uvažujme polynóm $f(x) = x^4 + 1$. Tento polynóm má celočíselné koeficienty, môžeme sa teda skúmať jeho ireducibilitu v okruhoch polynómov $\mathbb{Z}[x]$, $\mathbb{Q}[x]$, $\mathbb{R}[x]$ aj $\mathbb{C}[x]$.

V poli $\mathbb{C}$ má tento polynóm 4 korene $\frac{\pm\sqrt{2}\pm\sqrt{2}i}{2}$ (vieme ich nájsť riešením binomickej rovnice $x^4 = -1$). Teda v $\mathbb{C}$ máme rozklad

$$x^4 + 1 = \left(x - \frac{\sqrt{2}+\sqrt{2}i}{2}\right)\left(x - \frac{\sqrt{2}-\sqrt{2}i}{2}\right)\left(x + \frac{\sqrt{2}+\sqrt{2}i}{2}\right)\left(x + \frac{\sqrt{2}-\sqrt{2}i}{2}\right)$$

Nad poľom $\mathbb{R}$ máme rozklad

$$x^4 + 1 = (x^2 + \sqrt{2}x + 1)(x^2 - \sqrt{2}x + 1).$$

(Polynómy v rozklade môžeme získať napríklad ako súčin koreňových činiteľov pre komplexne združené korene. Alebo tento rozklad môžeme dostať tak, že si všimneme rovnosť $x^4 + 1 = (x^2+1)^2 - (\sqrt{2}x)^2$.) Pritom oba polynómy $x^2 \pm \sqrt{2}x + 1$ sú už nad $\mathbb{R}$ nerozložiteľné – pretože nemajú reálne korene.

Nad poľom $\mathbb{Q}$ je tento polynóm ireducibilný. Ak by sa totiž dal rozložiť na súčin nejakých polynómov, bol by súčasne aj súčinom týchto polynómov v $\mathbb{R}[x]$. Ako sme však videli, jediný (až na poradie a asociovanosť) rozklad na súčin polynómov nižšieho stupňa v $\mathbb{R}[x]$ je $x^4+1 = (x^2 + \sqrt{2}x + 1)(x^2 - \sqrt{2}x + 1)$ a polynómy, ktoré vystupujú v tomto rozklade, nepatria do $\mathbb{Q}[x]$.

13.5.5 Ireducibilné polynómy nad $\mathbb{Q}$ a $\mathbb{R}$

Z toho, čo doteraz vieme, sme schopní aspoň v niektorých konkrétnych prípadoch nájsť rozklad daného polynómu na súčin ireducibilných polynómov.

Postupovať môžeme tak, že hľadáme korene polynómu – pomocou hľadania racionálnych koreňov, riešením kvadratickej alebo binomickej rovnice (prípadne iných typov rovníc, ktoré vieme riešiť, ako sú reciproké rovnice, bikvadratické rovnice, kubické rovnice, rovnice štvrtého stupňa). Po nájdení koreňov môžeme polynóm vydeliť koreňovými činiteľmi (a znovu sa pokúsiť riešiť novú rovnicu nižšieho stupňa než bola pôvodná). V prípade, že by polynóm mal násobné korene, dá sa znížiť jeho stupeň použitím derivácie – o tom si ešte v tejto kapitole povieme.

V prípade, že po vydelení dostaneme polynóm dostatočne nízkeho stupňa, ktorý nemá korene, vieme už, že je ireducibilný.

Príklad 13.5.24. V príklade 13.5.11 sme zistili, že

$$f(x) = 24x^5 + 10x^4 - x^3 - 19x^2 - 5x + 6 = 24\left(x - \frac{1}{2}\right)\left(x + \frac{2}{3}\right)\left(x - \frac{3}{4}\right)(x^2 + x + 1).$$

Pretože polynóm $x^2 + x + 1$ nemá reálne korene (a je to polynóm druhého stupňa), je to rozklad na ireducibilné polynómy nad $\mathbb{R}$ (a tým pádom aj nad $\mathbb{Q}$). Rozklad nad $\mathbb{C}$ by sme získali, keby sme ešte $x^2 + x + 1$ rozložili na ireducibilné činitele.

Všimnime si, že sme vlastne dostali aj rozklad

$$f(x) = (2x-1)(3x+2)(4x-3)(x^2+x+1)$$

v $\mathbb{Z}[x]$.

Viac o rozklade polynómov na ireducibilné činitele (a o algoritmoch používaných na jeho výpočet) sa môžete dozvedieť na predmete počítačová algebra, pozri napríklad [G1, G2].

13.5.6 Derivácia a Taylorov rozvoj polynómov

Definícia 13.5.25. *Formálna derivácia* polynómu $f(x) = \sum_{k=0}^{n} a_k x^k$ je polynóm $Df(x) = \sum_{k=1}^{n} k \times a_k x^{k-1}$.

V prípade, že pracujeme nad ľubovoľným poľom, môže sa stať, že nenulový polynóm má nulovú deriváciu.

Príklad 13.5.26. Pre $f(x) = x^p$ v $\mathbb{Z}_p[x]$ dostávame $Df(x) = p \times x^{p-1} = 0$.

Priamo z definície sa dá overiť, že takto definovaná formálna derivácia má podobné vlastnosti, na aké sme zvyknutí z analýzy.

{polyn2:TVRLEIBNIZ}

Tvrdenie 13.5.27. *Nech F je pole. Pre ľubovoľné $c \in F$, $f(x), g(x) \in F[x]$ platí*

$$D(f(x)+g(x)) = Df(x) + Dg(x)$$
$$D(cf(x)) = cDf(x)$$
$$D(f(x)g(x)) = Df(x).g(x) + f(x).Dg(x)$$

Dôkaz. Overme iba tretiu rovnosť (prvé dve sú skutočne jednoduché). Koeficient pri x^n v polynóme na ľavej strane tejto rovnosti je $(n+1)$-násobok koeficientu polynómu $f(x).g(x)$ pri x^n.

Označme koeficienty polynómu $f(x)$ ako a_k, koeficienty polynómu $g(x)$ ako b_k. Pre koeficienty polynómu na ľavej strane rovnosti potom máme

$$l_n = (n+1) \times \sum_{k=0}^{n+1} a_k b_{n+1-k}$$

Na pravej strane rovnosti dostávame

$$p_n = \sum_{k=0}^{n} (k+1) \times a_{k+1} b_{n-k} + \sum_{k=0}^{n} (n+1-k) \times a_k b_{n+1-k}.$$

Zmenou sumačného indexu v prvej sume dostaneme vyjadrenie

$$p_n = \sum_{k=1}^{n+1} k \times a_k b_{n+1-k} + \sum_{k=0}^{n} (n+1-k) \times a_k b_{n+1-k} = \sum_{k=0}^{n+1} (n+1) \times a_k b_{n+1-k} = l_n$$

(aby sme uvedené členy mohli zlúčiť do jednej sumy, pridali sme dva nulové členy – v prvej sume pre $k=0$ člen $0 \times a_0 b_{n+1}$ a v druhej sume pre $k = n+1$ člen $0 \times a_{n+1} b_0$). □

Uvedieme si dve tvrdenia, ktoré ukazujú, prečo je tento pojem užitočný – prvé z nich je vyjadrenie Taylorovho polynómu v nejakom $c \in F$; druhé z nich hovorí o tom, či nejaký polynóm má násobné korene.

{polyn2:TVREXI

Tvrdenie 13.5.28. *Nech F je pole, $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0 \in F[x]$. Potom existujú jednoznačne určené $b_0, b_1, \ldots, b_n \in F$ také, že*

$$f(x) = b_n(x-c)^n + \cdots + b_1(x-c) + b_0. \tag{13.7}$$

{polyn2:EQEXRO

Dôkaz. Indukciou vzhľadom na n. Ak $n = 0$, tak stačí položiť $a_0 = b_0$ (a inú možnosť očividne nemáme).

Predpokladajme, že uvedené tvrdenie platí pre polynómy stupňa nanajvýš $n-1$. Podľa lemy 13.5.3

{polyn2:EQEXROZVOJ}

$$f(x) = g(x)(x-c) + f(c). \tag{13.8}$$

Polynóm $g(x)$ je stupňa najviac $n-1$- Podľa indukčného predpokladu existujú $b_1, \ldots, b_n \in F$ také, že $g(x) = b_n(x-c)^{n-1} + \cdots + b_2(x-c) + b_1$. Položme $b_0 = f(c)$. Potom pre $f(x)$ platí

$$f(x) = (b_n(x-c)^{n-1} + \cdots + b_2(x-c) + b_1)(x-c) + b_0 = b_n(x-c)^n + \cdots + b_1(x-c) + b_0.$$

Tým máme dokázanú existenciu.

Ak dosadíme do rovnosti (13.7) $x = c$, tak vidíme, že $b_0 = f(c)$. Ďalej polynóm $g(x)$ z (13.8) je podľa vety o delení so zvyškom jednoznačne určený. K tomuto polynómu sú podľa indukčného predpokladu jednoznačne určené $b_1, b_2, \ldots, b_n \in F$. □

Tvrdenie 13.5.29. *Ak F je pole charakteristiky ∞, tak koeficienty $b_0, \ldots, b_n$ z tvrdenia* 13.5.28 *možno vyjadriť ako*

$$b_n = \frac{D^{(n)}f(c)}{n!},$$

kde znak $D^{(n)}$ znamená, že polynóm $f(x)$ zderivujeme n-krát.

Dôkaz. Toto tvrdenie dostaneme priamo z rovnosti (13.7) viacnásobným zderivovaním (resp. ho môžeme ukázať pomocou indukcie). □

Rozvoj v tvare (13.7) môžeme dostať aj pomocou Hornerovej schémy – teda Hornerova schéma nám poskytuje možnosť vypočítať hodnoty $D^{(n)}f(c)$ pre daný polynóm $f(x)$ a $c \in F$.

Pred dôkazom nasledujúceho tvrdenia si všimnime jednu dôležitú vlastnosť najväčšieho spoločného deliteľa – konkrétne fakt, že zostane taký istý, aj keď prejdeme k nejakému nadpoľu.

{polyn2:POZNNSDNEZAV}

Poznámka 13.5.30. Už sme spomínali, že ak $f(x), g(x) \in F[x]$ a $F' \supseteq F$ je nadpole poľa F, tak polynómy $f(x)$ a $g(x)$ sú súčasne aj prvkami $F'[x]$. To znamená, že sa môžeme pýtať na najväčší spoločný deliteľ týchto 2 polynómov v okruhu $F[x]$ i v okruhu $F'[x]$. V obidvoch prípadoch je tento polynóm rovnaký.

Vyplýva to z toho, že podiel a zvyšok pri delení dvoch polynómov z $F[x]$ vyjde rovnako, bez ohľadu na to, či delíme so zvyškom v $F[x]$ alebo v $F'[x]$. (V $F[x]$ sa dajú vydeliť tak, aby podiel i zvyšok mali koeficienty z F, podiel v $F'[x]$ je rovnaký, pretože vo vete o delení so zvyškom máme jednoznačnosť.)

Z toho vyplýva aj to, že relácia „delí“ nezávisí od toho, či sa na polynómy $f(x)$, $g(x)$ pozeráme ako na prvky $F[x]$ alebo ako na prvky $F'[x]$.

{polyn2:TVRNASKORNSD}

Tvrdenie 13.5.31. *Nech F je pole, $F' \supseteq F$ je jeho nadpole. Nech $f(x) \in F[x]$ je polynóm nad poľom F. Ak v nadpoli F' existuje násobný koreň polynómu $f(x)$, tak polynómy $f(x)$ a $Df(x)$ sú súdeliteľné, t.j.*

$$\mathrm{st}(\gcd(f(x), D(f(x)))) \geq 1.$$

Dôkaz. Ak c je násobný koreň $f(x)$, tak podľa definície 13.5.5 $f(x) = g(x)(x-c)^k$, kde $k > 1$. Potom

$$Df(x) = Dg(x)(x-c)^k + k \times g(x)(x-c)^{k-1} = (x-c)^{k-1}(Dg(x)(x-c) + k \times g(x)),$$

teda $x - c \mid Df(x)$. Keďže súčasne $x - c \mid f(x)$, máme

$$x - c \mid \gcd(f(x), Df(x))$$

a $\mathrm{st}(\gcd(f(x), D(f(x)))) \geq 1$. (Predchádzajúcu nerovnosť sme dokázali pre najväčší spoločný deliteľ v $F'[x]$. Na základe poznámky 13.5.30 je však najväčší spoločný deliteľ v $F[x]$ rovnaký.) □

Predchádzajúce tvrdenie nám umožní nájsť polynóm, ktorý má rovnaké korene ako daný polynóm, ale každý koreň má násobnosť 1. Pred uvedením tohoto výsledku však potrebujeme zaviesť pojem charakteristiky poľa.

Definícia 13.5.32. *Charakteristika poľa* F je najmenšie prirodzené číslo $k > 0$ s vlastnosťou $k \times 1 = 0$. Označujeme ju $\mathrm{char}(F)$. Ak neexistuje k s uvedenou vlastnosťou, tak definujeme $\mathrm{char}(F) = \infty$.

Ak $\mathrm{char}(F) = k$, tak pre každé $c \in F$ platí $k \times c = c.(k \times 1) = c.0 = 0$.

{polyn2:TVRBEZNASKOR}

Tvrdenie 13.5.33. *Nech F je pole s nekonečnou charakteristikou. Nech $f(x) \in F[x]$ a $h(x)$ je najväčší spoločný deliteľ $f(x)$ a $Df(x)$. Potom existuje polynóm $g(x)$ s vlastnosťami*

(i) $f(x) = g(x).h(x)$,

(ii) *$g(x)$ má v každom nadpoli poľa F tie isté korene ako $f(x)$,*

(iii) *násobnosť každého koreňa $g(x)$ je 1.*

Dôkaz. Pretože $\mathrm{char}(F) = \infty$, máme $Df(x) \neq 0$. (Vedúci koeficient $Df(x)$ je $n \times a_n$, kde a_n je vedúci koeficient $f(x)$. V poli s nekonečnou charakteristikou z $a \neq 0$ vyplýva $n \times a \neq 0$.)

Potom aj $h(x)$ je nenulový polynóm. Navyše $h(x) \mid f(x)$, takže pri delení so zvyškom dostaneme

$$f(x) = g(x)h(x) + 0.$$

Ak c je násobný koreň $f(x)$ s násobnosťou k, tak platí $f(x) = (x-c)^k f_1(x)$, pričom c nie je koreňom $f_1(x)$. Z predchádzajúcej rovnosti dostaneme

$$Df(x) = Df_1(x)(x-c)^k + k \times f_1(x)(x-c)^{k-1} = (x-c)^{k-1}(Df_1(x)(x-c) + k \times f_1(x)).$$

Potom

$$h(x) = \gcd(f(x), Df(x)) = (x-c)^{k-1} \gcd((x-c)f_1(x), Df_1(x)(x-c) + k \times f_1(x)).$$

Pritom $x - c \nmid f_1(x)$, z čoho vyplýva $x - c \nmid Df_1(x)(x-c) + k \times f_1(x)$ a

$$x - c \nmid \gcd((x-c)f_1(x), Df_1(x)(x-c) + k \times f_1(x)).$$

Teda

$$(x-c)^k \nmid h(x)$$

(c je len $k-1$-násobným koreňom $h(x)$). T.j., ak vyjadríme $h(x)$ v tvare $h(x) = (x-c)^{k-1}h_1(x)$, tak $x-c \nmid h(x)$. Potom máme

$$(x-c)^k \mid g(x)h(x) = g(x)h_1(x)(x-c)^{k-1}$$
$$x-c \mid g(x)h_1(x)$$

Pretože $x-c$ je ireducibilný a $x-c \nmid h_1(x)$, vyplýva z toho už $x-c \mid g(x)$, čiže c je koreňom $g(x)$.

Navyše, c je iba jednoduchý koreň $g(x)$, v opačnom prípade by sme mali $(x-c)^2 \mid g(x)$, a teda

$$(x-c)^{k+1} \mid g(x)h_1(x-c)^{k-1} = g(x)h(x) = f(x).$$

To je spor s tým, že násobnosť koreňa c je k. □

Príklad 13.5.34. Majme polynóm $f(x) = x^4 + 2x^2 + 1 \in \mathbb{R}[x]$. Potom $Df(x) = 4x^3 + 4x$ a ich normovaný najväčší spoločný deliteľ je

$$h(x) = \gcd(f(x), Df(x)) = x^2 + 1 = x^4 + 2x^2 + 1 - \frac{x}{4}(4x^3 + 4x).$$

Po vydelení $f(x)$ polynómom $h(x)$ dostaneme $g(x) = x^2 + 1$.

Skutočne, polynómy $f(x) = (x^2+1)^2$ a $g(x) = x^2 + 1$ majú v $\mathbb{C}$ tie isté korene $\pm i$, v prípade polynómu $g(x)$ sú to jednoduché korene.

{polyn2cvic:KORINVHOM}

Cvičenia

Úloha 13.5.1. Nech F je pole, F' je jeho nadpole a $\varphi\colon F' \to F'$ je homomorfizmus taký, že $\varphi(x) = x$ pre každé $x \in F$ (nemení prvky poľa F). Potom pre každý koreň c polynómu $f(x)$ je aj $\varphi(c)$ koreňom $f(x)$.

Úloha 13.5.2. Vedeli by ste dokázať dôsledok 13.5.18 na základe poznatkov, ktoré máte z analýzy?

Úloha 13.5.3. Vydeľte dané polynómy so zvyškom v $\mathbb{C}[x]$.
a) $f(x) = x^4 + 3x^3 - 4x + 2$, $g(x) = x^2 + x - 2$
b) $f(x) = x^5 + 2x^3 + 3x + 4$, $g(x) = x^3 + x + 1$
c) $f(x) = x^3 + (2+2i)x^2 + 3ix + 1$, $g(x) = x^2 + (2+i)x + i$

Úloha 13.5.4. Použitím Hornerovej schémy zistite, či c je koreň polynómu $f(x) \in \mathbb{C}[x]$ a vyjadrite tento polynóm v tvare $f(x) = g(x)(x-c) + f(c)$.
a) $f(x) = x^4 + 3x^3 - 4x + 2$, $c = -2$
b) $f(x) = x^5 + 2x^3 + 3x + 4$, $c = -1$
c) $f(x) = x^3 + (2+2i)x^2 + 3ix + 1$, $c = -i$

Úloha 13.5.5. Pomocou Hornerovej schémy vyjadriť:
a) $f(x+3)$ pre $f(x) = x^4 - x^3 + 1$
b) $(x-2)^4 + 4(x-2)^3 + 6(x-2)^2 + 10(x-2) + 20$

Úloha 13.5.6. Nájdite všetky racionálne korene daných polynómov a zistite ich násobnosť. Aká je ich násobnosť?
a) $f(x) = 4x^4 - 7x^2 - 5x - 1$
b) $f(x) = 24x^5 + 10x^4 - x^3 - 19x^2 - 5x + 6$
c) $f(x) = 6x^4 + 19x^3 - 7x^2 - 26x + 12$
d) $f(x) = 6x^4 - 7x^3 + 8x^2 - 7x + 2$
e) $f(x) = 4x^4 + x^2 - 3x + 1$

Úloha 13.5.7. Ukážte, že reálny polynóm $f(x) = (x-4)(x-1)(x+1)(x+4) - 1$ nemá racionálne korene.

Úloha 13.5.8. Vypočítajte $d(x) = \gcd(f(x), g(x))$ a vyjadrite ho v tvare $d(x) = u(x)f(x) + v(x)g(x)$.
a) $f(x) = 3x^5 + 5x^4 - 16x^3 - 6x^2 - 5x - 6$, $g(x) = 3x^4 - 4x^3 - x^2 - x - 2$;
b) $f(x) = 4x^4 - 2x^3 - 16x^2 + 5x + 9$, $g(x) = 2x^3 - x^2 - 5x + 4$;
c) $f(x) = x^4 + 6x^3 + 9x^2 - 2x - 9$, $g(x) = x^3 + 4x^2 + 2x - 7$;
d) $f(x) = x^8 - 1$, $g(x) = x^5 - 1$;
e) $f(x) = x^{10} - 1$, $g(x) = x^4 - 1$.
(Výsledky: a) $u(x) = -\frac{1}{3}x^2 + \frac{1}{3}x + \frac{1}{3}$, $v(x) = \frac{1}{3}x^3 + \frac{2}{3}x^2 - \frac{5}{3}x - \frac{4}{3}$, $d(x) = x + \frac{2}{3}$
b) $u(x) = -\frac{x-1}{3}$, $v(x) = \frac{2x^2-2x-3}{3}$, $d(x) = x - 1$
c) $d(x) = 1$, $u(x) = 1/30(2x^2 + 5x - 1)$, $v(x) = -1/30(2x^3 + 9x^2 + 7x + 3)$)

Úloha 13.5.9. Vypočítajte $d(x) = \gcd(f(x), g(x))$ a vyjadrite ho v tvare $d(x) = u(x)f(x) + v(x)g(x)$.
a) $f(x) = x^4 - x^3 - 4x^2 - x + 5$, $g(x) = x^2 + x - 2$;
b) $f(x) = x^5 - 6x + 5$, $g(x) = x^3 + 4x^2 + x - 6$;
c) $f(x) = x^4 - 2x^2 - 3x - 2$, $g(x) = x^3 + 4x^2 + 4x + 1$

Úloha 13.5.10*. Nech $m, n > 1$ sú prirodzené čísla. Ukážte, že $\gcd(x^m - 1, x^n - 1) = x^d - 1$, kde $d = \gcd(m, n)$.

Úloha 13.5.11. Dokážte, že $x^2 + x + 1 \mid x^{3m} + x^{3n+1} + x^{3p+2}$ (v $F[x]$ pre ľubovoľné pole F).

Úloha 13.5.12. Dokážte, že $x^2 + x + 1 \mid x^{3m} + x^{3n+1} + x^{3p+2}$ v $\mathbb{C}[x]$. (Využite to, čo viete o koreňoch týchto polynómov.)

Úloha 13.5.13. Rozložiť na koreňové činitele (nad $\mathbb{C}$):
a) $x^3 - 6x^2 + 11x - 6$
b) $x^4 + 4$,
c) $x^4 + 4x^3 + 4x^2 - 1$,
d*) $x^4 + 4x^3 + 4x^2 + 1$,
e) $x^4 - 10x^2 + 1$,
f) $x^4 - 4x^3 + 4x - 1$.

Úloha 13.5.14. Rozložiť na súčin ireducibilných polynómov nad $\mathbb{R}$:
a) $x^4 + 4$
b) $x^6 + 27$
c) $x^4 + 4x^3 + 4x^2 - 1$
d) $x^5 + x^4 + x^3 + x^2 + x + 1$
e*) $x^{2n} - 2x^n + 2$
f*) $x^4 - ax^2 + 1$ pre $a \in (-2, 2)$
g*) $x^{2n} + x^n + 1$.

Úloha 13.5.15. Dokážte: Ak $a + bi$ je koreň polynómu $f(x) \in \mathbb{R}[x]$ a $b \neq 0$, tak $x^2 - 2ax + a^2 + b^2 \mid f(x)$.

Úloha 13.5.16. Nájdite všetky ireducibilné polynómy nad $\mathbb{Z}_2$ stupňov 2,3,4.

Úloha 13.5.17. Nájdite rozklad $f(x)$ na ireducibilné polynómy v $F[x]$.
a) $f(x) = 4x^4 + 3x^3 + 4x^2 + 4x + 6$, $F = \mathbb{Z}_7$
b) $f(x) = x^4 - 1$, $F = \mathbb{Z}_{11}$
c) $f(x) = x^4 - 1$, $F = \mathbb{Z}_{13}$

Úloha 13.5.18. Zistite, či dané ideály sú maximálne v $\mathbb{R}[x]$:
a) $I_1 = (x^2 - 1)$;
b) $I_2 = (x^2 + 1)$.

Úloha 13.5.19*. Nech $f(x) \in \mathbb{Z}[x]$ je polynóm s celočíselnými koeficientami. Dokážte, že ak $a + b\sqrt{3}$ je koreň $f(x)$, tak aj $a - b\sqrt{3}$ je koreň $f(x)$. Dokážte, že podobné tvrdenie platí, ak c nahradíme ľubovoľným prirodzeným číslom, ktoré nie je druhou mocninou prirodzeného čísla.

Úloha 13.5.20. Dokážte, že polynóm $f(x) = 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \ldots + \frac{x^n}{n!}$ nemá viacnásobný koreň (nad $\mathbb{R}$ resp. nad $\mathbb{C}$).

Kapitola 14

Polia

14.1 Podielové pole

Die ganzen Zahlen hat der liebe Gott gemacht, alles andere ist Menschenwerk.
(Celé čísla dal ľuďom dobrotivý Boh, všetko ostatné už je ľudským dielom.)
Leopold Kronecker

V tejto časti zodpovieme otázku, ktoré okruhy môžu byť podokruhmi polí. Je zrejmé, že podokruh poľa musí byť komutatívny okruh. Takisto nemôže obsahovať delitele nuly – boli by deliteľmi nuly aj v jeho nadpoli.

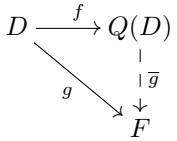
Táto otázka má teda zmysel hlavne pre obory integrity. V nasledujúcej vete dokážeme, že každý obor integrity je podokruhom nejakého poľa.

Definícia 14.1.1. Hovoríme, že okruh R je *vnorený* do okruhu R' ak existuje injektívny homomorfizmus $f\colon R \to R'$. Injektívny homomorfizmus $f\colon R \to R'$ nazývame *vnorenie*.

Vnorenie je vlastne izomorfizmus na podokruh $\operatorname{Im} f$, to znamená, že okruh R môžeme chápať priamo ako podokruh R'.

{podielove:VTQD}

Veta 14.1.2. *Pre každý obor integrity D existuje pole $Q(D)$ a vnorenie $f\colon D \to Q(D)$ s nasledujúcou vlastnosťou: Pre každé vnorenie $g\colon D \to F$ do poľa F existuje práve jedno vnorenie $\overline{g}\colon Q(D) \to F$ také, že $\overline{g} \circ f = g$.*



Význam podmienky v predchádzajúcej vete je o trochu jasnejší, keď si uvedomíme, že injektívne zobrazenie f nám hovorí, že obor integrity D môžeme chápať ako podokruh $Q(D)$. V prípade, že stotožníme prvky z D s ich obrazmi nám teda táto podmienka vlastne hovorí, že každé vloženie $g\colon D \to F$ do nejakého poľa F možno rozšíriť na vloženie celého $Q(D)$. (Teda $Q(D)$ je v istom zmysle najmenšie pole obsahujúce D.)

Dôkaz vety 14.1.2 urobíme vo viacerých krokoch – najprv zadefinujeme, ako vyzerá pre daný obor integrity pole $Q(D)$, postupne overíme, že spĺňa vlastnosti z definície poľa aj vlastnosť uvedenú vo vete.

Lema 14.1.3. *Nech D je obor integrity. Na množine $D \times (D \setminus \{0\})$ definujeme reláciu $\equiv$ predpisom*

$$(a,b) \equiv (c,d) \qquad \overset{def}{\Leftrightarrow} \qquad ad = bc.$$

Potom táto relácia je reláciou ekvivalencie a jej triedy $[(a,b)]$ nazývame zlomkami *nad oborom integrity D.*

Všimnite si, že relácia ekvivalencie je definovaná rovnako ako rovnosť zlomkov predstavujúcich racionálne čísla – ako uvidíme, celá konštrukcia podielového poľa $Q(D)$, ktorá bude nasledovať, pripomína spôsob, ktorým z okruhu celých čísel $\mathbb{Z}$ dostaneme pole racionálnych čísel $\mathbb{Q}$. (Väčšina dôkazov je skoro rovnaká, ako keby sme overovali, že $\mathbb{Q}$ je pole a spĺňa vlastnosť uvedenú vo vete 14.1.2.) V mnohých učebniciach, aby sa zdôraznila podobnosť s konštrukciou racionálnych čísel, sa pre zlomky nad D používa priamo označenie $\frac{a}{b}$ namiesto nášho označenia $[(a,b)]$.

Dôkaz. Dôkaz, že relácia $\equiv$ je reflexívna a symetrická je úplne priamočiary – cvičenie.

Tranzitívnosť: Nech $(a,b) \equiv (c,d)$ a $(c,d) \equiv (e,f)$. To znamená, že $ad = bc$ a $cf = de$.

Ak prvú rovnosť vynásobíme prvkom f a druhú prvkom b, dostaneme $adf = bcf = bde$, čiže $d(af - be) = 0$. Pretože D je obor integrity a $d \neq 0$, máme $af - be = 0$, teda

$$af = be$$

a $(a,b) \equiv (e,f)$. □

Lema 14.1.4. *Označme $Q(D)$ množinu všetkých tried ekvivalencie $\equiv$ na množine $D \times (D \setminus \{0\})$ (čiže množinu všetkých zlomkov nad D). Na tejto množine definujeme operácie $+$ a $\cdot$ predpismi*

$$[(a,b)] + [(c,d)] = [(ad + bc, bd)],$$
$$[(a,b)] \cdot [(c,d)] = [(ac, bd)].$$

Potom $+$ a $\cdot$ sú dobre definované a $(Q(D), +, \cdot)$ je pole.

Dôkaz. Najprv ukážeme, že obe operácie sú dobre definované, čiže nezávisia od výberu reprezentantov. Nech teda $(a,b) \equiv (a',b')$, čiže

$$ab' = a'b.$$

Potom máme

$$(ad + bc)b'd = ab'd^2 + bb'cd = a'bd^2 + bb'cd = (a'd + b'c)bd$$
$$acb'd = ab'cd = a'bcd = a'cbd$$

čo znamená

$$[(ad + bc, bd)] = [(a'd + b'c, b'd)]$$
$$[(ac, bd)] = [(a'c, b'd)]$$

čiže $[(a,b)] + [(c,d)] = [(a',b')] + [(c,d)]$ a $[(a,b)] \cdot [(c,d)] = [(a',b')] \cdot [(c,d)]$.

$(Q(D), +)$ *je komutatívna grupa.* Komutatívnosť je zrejmá. Asociatívnosť overíme priamym výpočtom.

$$([(a,b)] + [(c,d)]) + [(e,f)] = [(ad + bc, bd)] + [(e,f)] = [(adf + bcf + bde, bdf)]$$
$$[(a,b)] + ([(c,d)] + [(e,f)]) = [(a,b)] + [(cf + ed, df)] = [(adf + bcf + bed, bdf)]$$

Neutrálny prvok pre sčitovanie je $[(0,1)]$, opačný prvok k triede $[(a,b)]$ je $[(-a,b)]$.

$(Q(D) \setminus \{0\}, \cdot)$ *je komutatívna grupa.* Komutatívnosť je zrejmá, asociatívnosť sa ľahko overí priamym výpočtom. Neutrálny prvok vzhľadom na násobenie je $[(1,1)]$. Všimnime si, že $[(a,b)] \neq [(0,1)]$ práve vtedy, keď $a \neq 0$. Preto každý nenulový prvok $[(a,b)]$ má inverzný prvok $[(b,a)]$.

Distributívnosť. Keďže ide o komutatívny okruh, stačí overovať iba jednu z podmienok distributívnosti. Distributívnosť sa overí priamočiarym prepísaním z definície.

$$[(a,b)] \cdot ([(c,d)] + [(e,f)]) = [(a,b)] \cdot [(cf+de, df)] = [(a(cf+de), bdf)] =$$
$$= [(acf, bdf)] + [(ade, bdf)] = [(ac,bd)] + [(ae,bf)] = [(a,b)] \cdot [(c,d)] + [(a,b)] \cdot [(e,f)]$$

□

Lema 14.1.5. *Zobrazenie* $f\colon D \to Q(D)$

$$f\colon a \mapsto [(a,1)]$$

je injektívny homomorfizmus okruhov.

Ďalej pre ľubovoľný injektívny homomorfizmus $g\colon D \to F$*, kde* F *je pole existuje práve jeden injektívny homomorfizmus* $\overline{g}\colon Q(D) \to F$ *s vlastnosťou* $\overline{g} \circ f = g$*.*

Dôkaz. Zobrazenie f je homomorfizmus:

$$a + b \mapsto [(a+b,1)] = [(a,1)] + [(b,1)]$$
$$ab \mapsto [(ab,1)] = [(a,1)] \cdot [(b,1)]$$

Zobrazenie f je injektívne: rovnosť $[(a,1)] = [(b,1)]$ znamená, že $a.1 = b.1$, čiže $a = b$.

Ak $g\colon D \to F$ je injektívny homomorfizmus z D do nejakého poľa F, definujme $\overline{g}\colon Q(D) \to F$ predpisom

$$\overline{g}\colon [(a,b)] \mapsto g(a)g(b)^{-1}.$$

(Pretože g je injektívny homomorfizmus, máme $\operatorname{Ker} g = \{0\}$, čiže $g(b) \neq 0$ pre každé $b \in D \setminus \{0\}$. Uvedený predpis teda skutočne má zmysel. Navyše je jasné, že toto je jediná možnosť ako definovať $\overline{g}$, pretože toto zobrazenie musí spĺňať $\overline{g}([a,1]) = g(a)$ a $\overline{g}([1,b]) = g(b)^{-1}$.)

Zobrazenie $\overline{g}$ *je dobre definované.* Zobrazenie $\overline{g}$ sme definovali pomocou nejakého reprezentanta triedy $[(a,b)]$ – chceme ukázať, že výsledok zobrazenia nezávisí od výberu reprezentanta. Nech teda $(c,d) \equiv (a,b)$, čiže $ad = bc$. Potom dostávame (z toho, že g je homomorfizmus)

$$g(a)g(d) = g(b)g(c)$$

Po vynásobení tejto rovnosti $g(b)^{-1}g(d)^{-1}$ máme

$$g(a)g(b)^{-1} = g(c)g(d)^{-1}.$$

Čiže hodnota $\overline{g}$ skutočne nezávisí od výberu reprezentanta.

Zobrazenie $\overline{g}$ *je homomorfizmus.* Zachováva sčitovanie:

$$[(a,b)] + [(c,d)] = [(ad+bc, bd)] \mapsto g(ad+bc)g(bd)^{-1} = (g(ad) + g(bc))g(bd)^{-1} =$$
$$= g(a)g(d)g(b)^{-1}g(d)^{-1} + g(b)g(c)g(b)^{-1}g(d)^{-1} =$$
$$= g(a)g(b)^{-1} + g(c)g(d)^{-1} = \overline{g}([(a,b)]) + \overline{g}([(c,d)])$$

Zachováva násobenie:

$$[(a,b)]\cdot[(c,d)] = [(ac,bd)] \mapsto g(ac)g(bd)^{-1} =$$
$$= g(a)g(b)^{-1}g(c)g(d)^{-1} = \overline{g}([(a,b)])\cdot\overline{g}([(c,d)])$$

Homomorfizmus $\overline{g}$ je injektívny. Stačí overiť, že $\operatorname{Ker}\overline{g} = \{0\}$. Ak $g(a)g(b)^{-1} = 0$, znamená to, že $g(a) = 0$ (lebo prvok $g(b)^{-1} \in F \setminus \{0\}$ je nenulový). Potom (keďže homomorfizmus g je injektívny) máme $a = 0$, čiže $[(a,b)] = [(0,b)] = [(0,1)]$ je nulový prvok poľa $Q(D)$. □

Predchádzajúce tri lemy už spolu dokazujú vetu 14.1.2.

Poznámka 14.1.6. Existuje o niečo všeobecnejšia konštrukcia, ktorá sa nazýva *okruh zlomkov* alebo *lokalizácia* (napríklad [DF, Section 15.4], [AM, Chapter 3]). V tomto prípade sa pracuje s komutatívnym okruhom R s jednotkou, vyberie sa nejaká podmnožina $U \subseteq R$, ktorej prvky budú predstavovať menovatele zlomkov. (Inak povedané, U sú tie prvky, ktoré budú po urobení tejto konštrukcie mať inverzy vzhľadom na násobenie. Treba vyžadovať, aby množina U bola uzavretá vzhľadom na násobenie. V prípade konštrukcie podielového poľa sme mali $U = D \setminus \{0\}$.) Konštrukcia okruhu zlomkov je veľmi podobná konštrukcii podielového poľa, má aj podobné vlastnosti. Dôležitý rozdiel je, že v tomto prípade už zobrazenia spomínané vo vete 14.1.2 nemusia byť (vo všeobecnosti) injektívne. (Od zobrazenia g sa požaduje, aby zobrazovalo všetky prvky z U na delitele jednotky.) Táto konštrukcia je dôležitá napríklad v algebraickej geometrii a komutatívnej algebre.

Poznámka 14.1.7. Veľmi podobná konštrukcia ako vytvorenie podielového poľa z oboru integrity sa dá urobiť pre ľubovoľnú pologrupu s krátením (=pologrupa, v ktorej platia zákony o krátení). Dokážeme tak, že každú pologrupu s krátením možno vložiť do nejakej podgrupy (je podgrupou nejakej grupy.) Pozri úlohu 14.1.3.

Poznámka 14.1.8. Citát na začiatku tejto podkapitoly sa spája s konštrukciou reálnych čísel z celých čísel. My sme si ukázali, ako z celých čísel vytvoriť racionálne. Ďalším krokom by bolo pomocou racionálnych čísel nejakým spôsobom zaviesť reálne čísla. Existuje veľa ekvivalentných spôsobov ako to dosiahnuť (zúplnenie, Dedekinove rezy, reťazové zlomky, desatinné rozvoje...), viac sa o nich môžete dozvedieť napríklad v [Š]. Azda najčastejšie vyučovaným spôsobom je konštrukcia pomocou tried ekvivalencie cauchyovských postupností – zúplnenie racionálnych čísel – s ktorou by ste sa mohli stretnúť v niektorom pokročilejšom kurze analýzy. (Každý zo spomínaných spôsobov konštrukcie reálnych čísel nejakým spôsobom využíva pojem spojitosti.)

Cvičenia

Úloha 14.1.1. Dokážte, že každý komutatívny okruh možno vložiť do komutatívneho okruhu s jednotkou.

Úloha 14.1.2. Dokážte, že podielové pole je vlastnosťami uvedenými vo vete 14.1.2 určené jednoznačne až na izomorfizmus.

{podielovecvic:POLOGSKR}

Úloha 14.1.3*. Dokážte, že každú komutatívnu pologrupu s krátením možno vložiť do grupy. (Teda ak S je komutatívna pologrupa, v ktorej platia zákony o krátení, tak existuje grupa G a prostý homomorfizmus $f\colon S \to G$; pričom pod homomorfizmom pologrúp sa rozumie zobrazenie zachovávajúce operáciu, podobne ako pri grupách.)

14.2 Charakteristika poľa

S charakteristikou poľa sme sa už stretli v tvrdení 13.5.33. Pripomenieme si jej definíciu a dokážeme si niektoré fakty o charakteristike poľa, ktoré budú pre nás v nasledujúcich častiach prednášky užitočné.

{charpole:DEFCHAR}

Definícia 14.2.1. *Charakteristika poľa F* je najmenšie prirodzené číslo $k > 0$ s vlastnosťou $k \times 1 = 0$. Označujeme ju $\mathrm{char}(F)$. Ak neexistuje k s uvedenou vlastnosťou, tak definujeme $\mathrm{char}(F) = \infty$.

Predchádzajúcu definíciu môžeme preformulovať aj nasledovne

$$\mathrm{char}\, F = \min\{k \in \mathbb{N}, k > 0; k \times 1 = 0\},$$

pričom minimum z prázdnej množiny chápeme ako nekonečno.

Charakteristiku možno definovať aj všeobecnejšie – pre ľubovoľný okruh, pozri napríklad [KGGS, Kapitola 4.4]. V prípade okruhu s jednotkou je táto všeobecnejšia definícia ekvivalentná s definíciou, ktorú sme tu uviedli pre polia. (My budeme charakteristiku potrebovať iba pre polia.)

Príklad 14.2.2. $\mathrm{char}(\mathbb{Z}_p) = p$, pretože $p \times 1 = 0$ (počítame modulo p) pre $0 < k < p$ platí v $\mathbb{Z}_p$ $k \times 1 = k \neq 0$.

$\mathrm{char}(\mathbb{Q}) = \infty$, pretože žiadny násobok jednotky $k \times 1$, pre kladné celé čísla k, nie je 0.

Lema 14.2.3. *Každé konečné pole F má konečnú charakteristiku.*

Dôkaz. Vyplýva z Dirichletovho princípu.

Uvažujme množinu $\{k \times 1; k \in \mathbb{N}, k > 0\}$. Táto množina je konečná (je to podmnožina konečnej množiny F). Preto (na základe Dirichletovho princípu) existujú rôzne prirodzené čísla $m, n > 0$ také, že

$$m \times 1 = n \times 1.$$

Bez ujmy na všeobecnosti, nech $m > n$. Potom pre $k = m - n$ máme

$$k \times 1 = (m - n) \times 1 = m \times 1 - n \times 1 = 0,$$

čiže množina $\{k \in \mathbb{N}, k > 0; k \times 1 = 0\}$ je neprázdna, teda charakteristika (najmenší prvok tejto množiny) je konečná. □

Poznámka 14.2.4. Existujú aj nekonečné polia s konečnou charakteristikou.

{charpola:TVRPRVOCISLO}

Tvrdenie 14.2.5. *Charakteristika ľubovoľného poľa F je prvočíslo alebo ∞.*

Dôkaz. Stačí ukázať, že v prípade, že ak je charakteristika konečná, nemôže byť zložené číslo.

Sporom. Predpokladajme, že charakteristika poľa F je zložené číslo m, teda $m = nk$ pre nejaké prirodzené čísla $1 < n, k < m$. Potom platí

$$m \times 1 = (nk) \times 1 = (n \times 1)(k \times 1) = 0.$$

Každé pole je okruh bez deliteľov nuly, preto jeden z prvkov $n \times 1$, $k \times 1$ poľa F musí byť 0. Pritom $n, k < m$, čo je spor s tým, že m je (podľa definície charakteristiky) najmenšie kladné celé číslo s touto vlastnosťou. □

Nasledujúce tvrdenie budeme v ďalších kapitolách často využívať.

{charpola:TVR

Tvrdenie 14.2.6. *Nech F, F' sú polia a zobrazenie $\varphi\colon F \to F'$ je okruhový homomorfizmus. Potom buď $\varphi[F] = \{0\}$, alebo $\varphi[F]$ je podpole F', ktoré je izomorfné s F. (Inými slovami: zobrazenie φ je buď nulové alebo injektívne; čiže vnorenie – izomorfizmus na svoj obraz.)*

Dôkaz. Vieme, že $\operatorname{Ker}\varphi$ je ideál v F. Jediné ideály v poli sú však $\{0\}$ a F. V prvom prípade je homomorfizmus φ injektívny, v druhom prípade sa každý prvok zobrazí na nulu. □

Pomocou predchádzajúceho tvrdenia môžeme ukázať, že (v závislosti od charakteristiky) každé pole obsahuje podpole (izomorfné s) $\mathbb{Q}$ alebo $\mathbb{Z}_p$.

{charpola:TVROBSAHUJEZP}

Tvrdenie 14.2.7. *Ak* $\operatorname{char} F = \infty$*, tak existuje injektívny homomorfizmus z $\mathbb{Q}$ do F.*
Ak $\operatorname{char} F = p$ *pre nejaké prvočíslo p, tak existuje injektívny homomorfizmus zo $\mathbb{Z}_p$ do F.*

Dôkaz. Zobrazenie $\varphi\colon \mathbb{Z} \to F$

$$\varphi\colon z \mapsto z \times 1$$

je okruhový homomorfizmus, pričom $\operatorname{Ker}\varphi$ obsahuje práve tie celé čísla, ktoré sú násobky $\operatorname{char}(F)$ (a v prípade, že $\operatorname{char} F = \infty$ je $\operatorname{Ker}\varphi = \{0\}$).

Ak $\operatorname{char} F = p$, tak máme (na základe vety o faktorovom izomorfizme) izomorfizmus z $\mathbb{Z}/\operatorname{Ker}\varphi = \mathbb{Z}/(\operatorname{char}(F)) = \mathbb{Z}/(p) \cong \mathbb{Z}_p$ na $\operatorname{Im}\varphi$. Tým dostávame hľadaný injektívny homomorfizmus zo $\mathbb{Z}_p$ do F.

Ak $\operatorname{char} F = \infty$, tak $\operatorname{Ker} F = (0)$ a φ je injektívny homomorfizmus zo $\mathbb{Z}$ do F. Ten sa podľa vety 14.1.2 dá rozšíriť na injektívny homomorfizmus $\overline{\varphi}\colon Q(\mathbb{Z}) \to F$ z podielového poľa oboru integrity $\mathbb{Z}$ do F. Podielové pole $\mathbb{Z}$ je však práve pole racionálnych čísel. □

Nasledujúce tvrdenie má síce veľmi jednoduchý dôkaz, podarí sa nám však z neho odvodiť veľmi zaujímavé dôsledky.

{charpola:NADPOLEJEVP}

Tvrdenie 14.2.8. *Nech K, F sú polia a K je nadpole poľa F (t.j. $K \supseteq F$ a operácie na F sú zúženia operácií na K). Potom K je vektorový priestor nad poľom F (so sčitovaním a násobením skalárom rovnakým ako je sčitovanie a násobenie v K).*

Dôkaz. Jednoduchý – jednotlivé vlastnosti z definície vektorového priestoru po prepísaní na tento konkrétny príklad sú vlastne známe vlastnosti poľa ako distributívnosť, asociatívnosť násobenia atď. (Dôkaz je skoro identický s postupom použitým v úlohách 4.1.10 a 4.1.5) □

Ako sme už spomenuli, napriek jednoduchému dôkazu bude mať pohľad na nadpole ako na vektorový priestor nad daným poľom mnohé zaujímavé dôsledky. Ako prvý z nich si ukážeme, aký počet prvkov môže mať konečné pole.

{charpola:DOSPOCETPNAN}

Dôsledok 14.2.9. *Konečné pole charakteristiky p má p^n prvkov pre nejaké $n \in \mathbb{N}$.*

Dôkaz. Podľa tvrdenia 14.2.7 každé konečné pole F s $\operatorname{char}(F) = p$ obsahuje podpole (izomorfné so) $\mathbb{Z}_p$. Teda ho môžeme chápať ako vektorový priestor nad $\mathbb{Z}_p$. Keďže množina F je konečná, ide o konečnorozmerný vektorový priestor, teda F je (ako vektorový priestor) izomorfný s priestorom $(\mathbb{Z}_p)^n$ pre nejaké n (veta 5.5.15). □

Z toho vyplýva, že počet prvkov konečného poľa musí byť mocnina prvočísla. Napríklad dostávame, že nemôže existovať 6-prvkové pole. (Platí aj obrátené tvrdenie, pre každú mocninu prvočísla $k = p^n$ existuje k-prvkové pole.)

Môžeme si všimnúť ešte jeden užitočný fakt súvisiaci s charakteristikou poľa.

a:TVRFROBHOM}

Tvrdenie 14.2.10. *Nech* $\mathrm{char}(F) = p$ *(*p *je prvočíslo). Potom pre ľubovoľné* $a, b \in F$ *platí*

$$(a+b)^p = a^p + b^p$$
$$(ab)^p = a^p b^p$$

čiže zobrazenie $f\colon F \to F$, $f(x) = x^p$, *je homomorfizmus (endomorfizmus poľa* F*).*

Ďalej pre ľubovoľné $n \in \mathbb{N}$ *a* $q = p^n$ *máme*

$$(a+b)^q = a^q + b^q$$
$$(ab)^q = a^q b^q$$

Dôkaz. Jediná netriviálna časť je rovnosť $(a+b)^p = a^p + b^p$. Použitím binomickej vety (ktorá platí v každom komutatívnom okruhu s jednotkou, úloha 13.2.22) máme

$$(a+b)^p = \sum_{k=0}^{p} \binom{p}{k} \times a^k b^{p-k}.$$

Chceme ukázať, že všetky sčítance s výnimkou prvého a posledného (t.j. $k = 0$ a $k = p$) sú nulové.

Na to nám stačí ukázať, že $p \mid \binom{p}{k}$ v $\mathbb{Z}$,keďže p je charakteristika poľa, s ktorým pracujeme. (Vieme, že binomický koeficient je vždy celé číslo.) Ak však p je prvočíslo, tak p delí čitateľ zlomku

$$\binom{p}{k} = \frac{p!}{k!(p-k!)}$$

pričom v menovateli sa (pre $k \neq 0, p$) vyskytnú len čísla ostro menšie ako p, t.j. žiadne z nich nie je deliteľné p. Z toho už vyplýva, že p delí toto číslo.

Posledná časť tvrdenia, ktorá hovorí o $q = p^n$, sa ľahko odvodí z prvej časti indukciou vzhľadom na n. □

Cvičenia

{charpolacvic:ULOKxA}

Úloha 14.2.1. Nech F je pole, $a \in F$, $a \neq 0$. Dokážte, že ak $k \times a = 0$, tak platí $k = 0$ alebo F má konečnú charakteristiku a

$$\mathrm{char}(F) \mid k,$$

t.j. k je celočíselný násobok charakteristiky.

14.3 Rozšírenia polí

Prezentácia výsledkov v tejto kapitole i v nasledujúcich kapitolách je podobná ako v [KGGS, Kapitola 8], [DF, Chapter 13].

Na začiatok začnime s definíciou pojmu rozšírenia poľa.

Definícia 14.3.1. Ak K, F sú polia a súčasne F je podokruhom K, tak hovoríme, že K je *rozšírením* poľa F.

Vidíme, že rozšírenie poľa je vlastne len iné pomenovanie pre dvojicu pozostávajúcu z poľa F a jeho nadpoľa K (čiže vždy, keď hovoríme o rozšírení poľa, máme na mysli dve polia).

Ak K je rozšírenie poľa F, tak K môžeme chápať ako vektorový priestor nad F (tvrdenie 14.2.8). Pre nás bude zaujímavý hlavne ten prípad, keď je to konečnorozmerný vektorový priestor.

Definícia 14.3.2. Ak K je rozšírenie poľa F také, že K je konečnorozmerný vektorový priestor nad F, tak K nazývame *konečné rozšírenie* poľa F.

Dimenziu $d_F(K)$ poľa K ako vektorového priestoru nad F nazývame *stupeň rozšírenia* a označujeme $[K : F]$.

$$[K : F] = d_F(K)$$

Príklad 14.3.3. Pole $\mathbb{C}$ je rozšírením poľa $\mathbb{R}$. Všimnime si, že $\mathbb{C} = \{a + bi; a, b \in \mathbb{R}\}$, a teda 1, i tvoria bázu $\mathbb{C}$ ako vektorového priestoru nad $\mathbb{R}$. Preto $[\mathbb{C} : \mathbb{R}] = 2$.

Na $\mathbb{C}$ sa môžeme pozerať tak, že k poľu $\mathbb{R}$ sme pridali koreň polynómu $x^2 + 1$ (a aj všetky ďalšie prvky, ktoré si pridanie tohoto koreňa vynútilo, aby novovytvorená štruktúra bola opäť poľom). V poli $\mathbb{R}$ polynóm $x^2 + 1$ nemá koreň.

Príklad 14.3.4. Vieme, že $\mathbb{Q}[\sqrt{2}] = \{a+b\sqrt{2}; a, b \in \mathbb{Q}\}$ je pole (úloha 3.3.2). Je to rozšírenie poľa $\mathbb{Q}$. Ak sa na $\mathbb{Q}[\sqrt{2}]$ pozrieme ako na vektorový priestor nad $\mathbb{Q}$, tak jeho bázu tvoria $1, \sqrt{2}$. (Rozmyslite si, prečo sú lineárne nezávislé nad $\mathbb{Q}$.) Teda $[\mathbb{Q}[\sqrt{2}] : \mathbb{Q}] = 2$.

Aj v tomto prípade môžeme toto rozšírenie chápať tak, že k poľu $\mathbb{Q}$ sme pridali koreň polynómu $x^2 - 2$. (V poli $\mathbb{Q}$ tento polynóm nemá koreň.)

V predchádzajúcich dvoch príkladoch sme videli, že k ireducibilným polynómom $x^2 - 2 \in \mathbb{Q}[x]$, $x^2 + 1 \in \mathbb{R}[x]$ existujú konečné rozšírenia, v ktorých už tieto polynómy majú koreň. Ukážeme, že podobné tvrdenie platí pre ľubovoľný ireducibilný polynóm.

{rozs:VTIREDKOREN}

Veta 14.3.5. *Nech F je pole a $p(x)$ je ireducibilný polynóm v $F[x]$. Potom existuje rozšírenie poľa F, v ktorom $p(x)$ má koreň.*

Dôkaz. Keďže $p(x)$ je ireducibilný, $(p(x))$ je maximálny ideál v $F[x]$ (tvrdenia 13.4.19 a 13.4.32). Teda faktorový okruh $K = F[x]/(p(x))$ je pole. O tomto poli K ukážeme, že má požadované vlastnosti.

Máme kanonický homomorfizmus $\varphi\colon F[x] \to K$ taký, že $\operatorname{Ker}\varphi = (p(x))$. Súčasne F je podmnožinou $F[x]$, teda máme aj homomorfizmus $\varphi|_F\colon F \to K$ (zúženie homomorfizmu φ na podmnožinu F). Tento homomorfizmus je nenulový, keďže na nulu sa zobrazia iba prvky z $\operatorname{Ker}\varphi = (p(x))$, kam patria iba 0 a polynómy stupňa aspoň $\operatorname{st} p(x) \geq 1$ (čiže žiadny nenulový konštantný polynóm – žiadny nenulový prvok poľa F). Podľa tvrdenia 14.2.6 je to teda injektívny homomorfizmus (vnorenie) a F môžeme chápať ako podpole K. Ide teda skutočne o rozšírenie poľa F.

Treba ešte dokázať, že $p(x)$ má v tomto poli koreň. Ukážeme, že koreňom je prvok $\varphi(x) = \underline{x + (p(x))}$. Kvôli zjednodušeniu zápisu budeme používať označenie $\varphi(x) = \overline{x}$, resp. $\varphi(f(x)) = \overline{f(x)}$ pre ľubovoľné $f(x) \in F[x]$.

Máme rovnosť

$$p(\overline{x}) \overset{(*)}{=} \overline{p(x)} = p(x) + (p(x)) = 0 + (p(x)),$$

ktorá znamená, že $\overline{x}$ je skutočne koreňom polynómu $p(x)$. (V predchádzajúcom odvodení bola najdôležitejším krokom rovnosť označená $(*)$, ktorá je založená na tom, že φ je homomorfizmus medzi komutatívnymi okruhmi, teda zachováva súčet, súčin a teda aj všetky polynomické výrazy).[1] □

[1] O niečo podrobnejšie zdôvodnenie rovnosti $(*)$: Ak $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$, tak

$$\begin{aligned}
\overline{p(x)} &= \overline{a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0} \\
&= \overline{a_n x^n} + \overline{a_{n-1}x^{n-1}} + \cdots + \overline{a_1 x} + \overline{a_0} \\
&= \overline{a_n} \cdot \overline{x}^n + \overline{a_{n-1}} \cdot \overline{x}^{n-1} + \cdots + \overline{a_1} \cdot \overline{x} + \overline{a_0} \\
&\overset{(\triangle)}{=} a_n \overline{x}^n + a_{n-1}\overline{x}^{n-1} + \cdots + a_1 \overline{x} + a_0 \\
&= p(\overline{x})
\end{aligned}$$

Teraz ukážeme, že rozšírenie K poľa F zostrojené v predchádzajúcej vete je konečným rozšírením.

{rozs:VTSTUPPOL}

Veta 14.3.6. *Nech $p(x) \in F[x]$ je ireducibilný polynóm a $K = F[x]/(p(x))$. Nech $n = \operatorname{st} p$. Označme $u = x + (p(x)) = \varphi(x)$ (kde $\varphi\colon F[x] \to K$ označuje kanonický homomorfizmus). Potom $1, u, \ldots, u^{n-1}$ je báza K ako vektorového priestoru nad F, čiže*

$$K = \{a_{n-1}u^{n-1} + \cdots + a_1 u + a_0; a_0, \ldots, a_{n-1} \in F\}.$$

Dôkaz. Máme surjektívny homomorfizmus $\varphi\colon F[x] \to K$, ktorý polynóm $f(x)$ zobrazí na triedu $f(x) + (p(x)) = f(u)$. (Teda každý prvok K možno vyjadriť ako $f(u)$ pre nejaké $f \in F[x]$.)

Ak $f(x)$ je ľubovoľný polynóm z $F[x]$, tak podľa vety o delení so zvyškom existujú $q(x)$ a $r(x)$ také, že

$$f(x) = q(x)p(x) + r(x),$$

pričom $\operatorname{st} r \leq \operatorname{st} p = n$ Potom máme

$$f(u) = f(x) + (p(x)) = r(x) + (p(x)) = r(u) = a_{n-1}u^{n-1} + \cdots + a_1 u + a_0.$$

Čiže každý prvok z K sa skutočne vyjadriť ako lineárna kombinácia $1, u, \ldots, u^{n-1}$ (s koeficientami z F), t.j. vektory $1, u, \ldots, u^{n-1}$ generujú vektorový priestor K.

Ešte zostáva ukázať, že $1, u, \ldots, u^{n-1}$ sú lineárne nezávislé nad F. Predpokladajme, že pre nejaké $b_0, \ldots, b_{n-1}$ by platilo v $K = F[x]/(p(x))$

$$b_{n-1}u^{n-1} + \cdots + b_1 u + b_0 = 0.$$

Táto rovnosť vo faktorovom okruhu $F[x]/(p(x))$ znamená, že v okruhu $F[x]$ platí

$$b_{n-1}x^{n-1} + \cdots + b_1 x + b_0 \in (p(x)).$$

Jediný polynóm v $(p(x))$, ktorý má stupeň menej ako n, je však nulový polynóm, preto $b_0 = b_1 = \cdots = b_{n-1} = 0$, čiže $1, u, \ldots, u^{n-1}$ sú skutočne lineárne nezávislé. □

{rozs:DOSSTUPPOL}

Dôsledok 14.3.7. *Ak $p(x) \in F[x]$ je ireducibilný polynóm stupňa n, tak $K = F[x]/(p(x))$ je konečné rozšírenie F a stupeň rozšírenia $[K : F]$ je tiež rovný n.*

$$[K : F] = \operatorname{st} p(x)$$

Predchádzajúca veta nám hovorí, že každý prvok poľa $F[x]/(p(x))$ môžeme vyjadriť ako $a_{n-1}u^{n-1}+\cdots+a_1u+a_0$ pre nejaké $a_0, \ldots, a_{n-1} \in F$. V poli $F[x]/(p(x))$ vieme jednoduchým spôsobom sčitovať a násobiť – ide jednoducho o sčitovanie a násobenie modulo $p(x)$. Presnejšie ak máme 2 prvky vyjadrené ako $f(u)$ a $g(u)$ pre nejaké polynómy $f(x), g(x) \in F[x]$ stupňa menšieho ako n, tak ich súčet zodpovedá priamo súčtu polynómov $f(x) + g(x)$. Ich súčin dostaneme tak, že vypočítame súčin $f(x)g(x)$ a zistíme jeho zvyšok po delení $p(x)$. (Fakt, že sčitovanie a násobenie v poli $F[x]/(p(x))$ sa správa takýmto spôsobom, vyplýva priamo z definície faktorového okruhu.)

{rozs:PRGF4}

Príklad 14.3.8. Uvažujme polynóm $p(x) = x^2 + x + 1$ nad poľom $\mathbb{Z}_2$. Tento polynóm je ireducibilný, lebo ide o polynóm druhého stupňa, ktorý nemá v danom poli koreň (tvrdenie 13.5.22). Ak označíme ako u triedu polynómu x vo faktorovom okruhu $GF_4 = \mathbb{Z}_2[x]/(p(x))$,

Všetky rovnosti sú vlastne založené iba na definícii homomorfizmu, jedine zdôvodnenie rovnosti (△) je iné. Tu využívame, že sme stotožnili (cez kanonický homomorfizmus) prvok a z poľa F s jeho triedou $\overline{a}$ v $F[x]/(p(x))$.

tak prvky poľa GF_4 sú $\{0,1,u,u+1\}$. Na základe predchádzajúcich úvah vieme vyplniť tabuľku násobenia a sčitovania v tomto poli:

$$(au+b)+(cu+d)=(a+b)u+(b+d)$$
$$(au+b)(cu+d)=acu^2+(bc+ad)u+bd=ac(u+1)+(bc+ad)u+bd=(ac+bc+ad)u+(ac+bd)$$

$+$	0	1	u	$u+1$
0	0	1	u	$u+1$
1	1	0	$u+1$	u
u	u	$u+1$	0	1
$u+1$	$u+1$	u	1	0

$\cdot$	0	1	u	$u+1$
0	0	0	0	0
1	0	1	u	$u+1$
u	0	u	$u+1$	1
$u+1$	0	$u+1$	1	u

Samozrejme, keďže polynóm x^2+x+1 je polynóm druhého stupňa a má v poli GF_4 koreň, musí sa dať rozložiť na lineárne činitele. Skutočne v GF_4 platí $x^2+x+1=(x+u)(x+u+1)$.

Príklad 14.3.9. Polynóm $p(x)=x^2+1$ je ireducibilný nad $\mathbb{R}$. Uvažujme pole $\mathbb{R}[x]/(x^2+1)$. Pokúsme sa zistiť, čomu sa v tomu poli rovná súčin $(au+b)(cu+d)$. V $\mathbb{R}[x]$ máme rovnosť

$$(ax+b)(cx+d)=acx^2+(cb+ad)x+bc=ac(x^2+1)+(cb+ad)x+(bd-ac).$$

Z toho dostávame rovnosť v poli $\mathbb{R}[x]/(x^2+1)$

$$(ax+b)(cx+d)+(p(x))=(cb+ad)x+(bd-ac)+(p(x)),$$
$$(au+b)(cu+d)=(cb+ad)u+(bd-ac).$$

Vidíme, že predpis pre sčitovanie násobenie je rovnaký ako pre komplexné čísla, čiže sme takto (až na izomorfizmus) získali pole $\mathbb{C}$ t.j. $\mathbb{C}\cong\mathbb{R}[x]/(x^2+1)$

Podobne dostaneme $\mathbb{Q}[\sqrt{2}]\cong\mathbb{Q}[x]/(x^2-2)$.

Definícia 14.3.10. Ak K je rozšírenie F a $u_1,\dots,u_n\in K$, tak symbolom $F(u_1,\dots,u_n)$ označujeme podpole generované množinou $F\cup\{u_1,\dots,u_n\}$. (T.j. najmenšie podpole, ktoré obsahuje túto množinu, čiže prienik všetkých podpolí, ktoré ju obsahujú.)[2]

V prípade, že existuje $u\in K$ také, že $K=F(u)$ hovoríme o *jednoduchom rozšírení.*

Vo vete 14.3.5 sme ukázali, že pre ireducibilný polynóm $p(x)$ existuje rozšírenie, v ktorom tento polynóm má koreň. Teraz ukážeme, že toto pole je jednoznačne určené až na izomorfizmus.

{rozs:VTFUJEFXPX}

Veta 14.3.11. *Nech F je pole, $p(x)\in F[x]$ je ireducibilný polynóm nad F a K je rozšírenie F, ktoré obsahuje koreň u polynómu $p(x)$. Potom*

$$F(u)\cong F[x]/(p(x)).$$

Dôkaz. Máme dosadzovací homomorfizmus (definícia 13.3.15) $\varphi_u\colon F[x]\to F(u)$

$$\varphi_u\colon a(x)\mapsto a(u).$$

Keďže u je koreň $p(x)$, platí $p(x)\in\operatorname{Ker}\varphi_u$. Vďaka tomu je homomorfizmus $\overline{\varphi}_u\colon F[x]/(p(x))\to F(u)$ určený ako

$$\overline{\varphi_u}\colon a(x)+(p(x))\mapsto a(u)$$

[2]Takéto podpole vždy existuje, stačí ukázať, že prienik podpolí je opäť podpole; úloha 14.3.1. (Pripomeniem, že už vieme, že prienik podokruhov je podokruh; úloha 13.2.3.)

dobre definovaný. (Ak $a(x)$ a $b(x)$ patria do tej istej triedy, tak $b(x) - a(x) = g(x)p(x)$, čiže $b(u) - a(u) = g(u)p(u) = 0$ a $b(u) = a(u)$. Teda definícia zobrazenia $\overline{\varphi}_u$ nezávisí od výberu reprezentanta.)

Zobrazenie $\overline{\varphi}_u$ je homomorfizmus polí. Je to nenulový homomorfizmus, lebo $\overline{\varphi}_u(1) = 1$. Z toho vyplýva, že tento homomorfizmus je injektívny (tvrdenie 14.2.6).

Navyše v tomto zobrazení sa každý prvok F zobrazí sám na seba a x sa zobrazí na u. Keďže $\operatorname{Im}\overline{\varphi}_u$ je podpole K a obsahuje F aj u, musí obsahovať celé $F(u)$. Teda homomorfizmus $\overline{\varphi}_u$ je i surjektívny.

Ukázali sme, že $\overline{\varphi}_u$ je izomorfizmus a teda polia $F(u)$ a $F[x]/(p(x))$ sú skutočne izomorfné. □

Z predchádzajúcej vety vyplýva, že dva korene ireducibilného polynómu sú algebraicky nerozlíšiteľné v tom zmysle, že po ich pridaní k poľu F dostaneme izomorfné polia. Tento fakt o čosi zovšeobecníme v nasledujúcej vete, kde nebudeme vychádzať z toho istého poľa ale z dvoch izomorfných polí.

Najprv si všimnime ako sa izomorfizmus medzi poľami dá rozšíriť na izomorfizmus medzi ich okruhmi polynómov.

{rozs:POZNIZOFX}

Poznámka 14.3.12. Nech $\varphi\colon F \to F'$ je izomorfizmus, F aj F' sú polia. Potom môžeme definovať zobrazenie $\hat{\varphi}\colon F[x] \to F'[x]$, ktoré polynómu z $F[x]$ priradí polynóm rovnakého stupňa, ktorého koeficienty dostaneme ako obrazy koeficientov pôvodného polynómu v homomorfizme φ.

$$\hat{\varphi}\colon \sum_{i=0}^{n} a_i x^i \mapsto \sum_{i=0}^{n} \varphi(a_i) x^i$$

Pomerne jednoducho sa overí, že ide opäť o izomorfizmus. Keďže ide o izomorfizmus, toto zobrazenie musí zachovávať maximálne ideály, prvoideály, ireducibilné prvky (=ireducibilné polynómy) a mnohé ďalšie vlastnosti.

Všimnime si tiež, že tento izomorfizmus navyše zachováva aj stupne polynómov.

{rozs:VTFUIZOM}

Veta 14.3.13. *Nech $\varphi\colon F \to F'$ je izomorfizmus polí. Nech $p(x)$ je ireducibilný polynóm nad F a $p'(x) \in F'[x]$ je polynóm $\hat{\varphi}(p)$ (čiže polynóm, ktorý získame použitím izomorfizmu $\varphi\colon F \to F'$ na všetky koeficienty polynómu $f(x)$). Potom $p'(x)$ je tiež ireducibilný polynóm (nad F').*

Nech u je koreň $p(x)$ (v nejakom nadpoli F) a v je koreň $p'(x)$ (v nejakom nadpoli F'). Potom existuje izomorfizmus

$$\sigma\colon F(u) \to F'(v),$$

ktorý zobrazí u na v a rozširuje φ, t.j. $\sigma(u) = v$ a $\sigma|_F = \varphi$.

Dôkaz. Fakt, že $p'(x)$ je ireducibilný vyplýva priamo z existencie izomorfizmu $\hat{\varphi}\colon F[x] \to F'[x]$.

Podľa vety 14.3.11 je $F(u) \cong F[x]/(p(x))$ a $F'(v) \cong F'[x]/(p'(x))$ (pričom izomorfizmy medzi uvedenými poliami nemenia prvky z F, resp. prvky z F'). V skutočnosti nám teda stačí hľadať izomorfizmus (s požadovanými vlastnosťami) medzi $F[x]/(p(x))$ a $F'[x]/(p'(x))$.

Máme zobrazenie $\hat{\varphi}\colon F[x] \to F'[x]$. Definujme $\psi\colon F[x]/(p(x)) \to F'[x]/(p'(x))$ predpisom

$$\psi\colon f(x) + (p(x)) \mapsto \hat{\varphi}(f(x)) + (p'(x)).$$

Ukážeme, že toto zobrazenie je dobre definované a je to izomorfizmus s požadovanými vlastnosťami.

Ak $f(x) = g(x)p(x) + r(x)$, tak

$$\hat{\varphi}(f(x)) = \hat{\varphi}(g(x))p'(x) + \hat{\varphi}(r(x)).$$

(V ďalšom budeme namiesto $\hat{\varphi}(r(x))$ používať stručnejšie označenie $r'(x)$.) Keďže zobrazenie $\hat{\varphi}$ zachováva stupne polynómov, predchádzajúca rovnosť nám hovorí, že zachováva aj zvyšky po delení $p(x)$ a $p'(x)$. (T.j. zvyšok polynómu $f(x)$ po delení $p(x)$ sa zobrazí na zvyšok polynómu $\hat{\varphi}(f(x))$ po delení $p'(x)$.)

To špeciálne znamená, že ak dva polynómy $f_1(x), f_2(x) \in F[x]$ majú rovnaký zvyšok po delení $p(x)$ (=sú reprezentantmi tej istej triedy rozkladu v $F[x]/(p(x))$), tak aj ich obrazy budú mať rovnaký zvyšok po delení $p'(x)$. Z toho vidíme, že zobrazenie ψ je dobre definované.

Z predchádzajúcej úvahy vyplýva aj to, že ψ je homomorfizmus – zachováva operácie $+$ a $\cdot$. S operáciou $+$ nemáme žiadne problémy, pretože sčitovanie v $F[x]/(p(x))$ pracuje rovnako ako sčitovanie polynómov a vieme, že $\hat{\varphi}$ zachováva sčitovanie. Násobenie funguje ako násobenie v $F[x]$ (resp. v $F'[x]$) s tým rozdielom, že musíme ešte urobiť zvyšok po delení $p(x)$ (v druhom prípade $p'(x)$). Práve sme si ozrejmili, že $\hat{\varphi}$ zachováva zvyšky po delení.

Zobrazenie $\hat{\varphi}$ zobrazí polynóm x na polynóm x (lebo $\varphi(1) = 1$ pre ľubovoľný izomorfizmus polí). Z toho vyplýva, že koreň $x + (p(x))$ polynómu $p(x)$ sa zobrazí na koreň $x + (p'(x))$ polynómu $p'(x)$.

Zatiaľ teda vieme, že ψ je homomorfizmus polí a je nenulový (prvok $x + (p(x))$ sa zobrazí na $x + (p'(x))$, ktorý je nenulový). Podľa tvrdenia 14.2.6 je tento homomorfizmus injektívny. Navyše, pretože $\hat{\varphi}$ je surjektívne zobrazenie, priamo z definície ψ vyplýva, že aj zobrazenie ψ je surjektívne. Je to teda izomorfizmus.

Už sme videli, že $\hat{\varphi}$ zobrazí koreň $p(x)$ na koreň $p'(x)$. Z toho, že $\hat{\varphi}$ nemení prvky poľa F vyplýva, že rovnakú vlastnosť má aj ψ. □

Cvičenia

{rozscvic:ULOPRIENPODP}

Úloha 14.3.1. Nech K je pole a $\{F_i, i \in I\}$ je systém podpolí poľa K. Dokážte, že aj prienik $\bigcap_{i\in I} F_i$ je podpoľom poľa K.

14.4 Algebraické rozšírenia

Definícia 14.4.1. Nech K je rozšírenie poľa F. Nech $u \in K$. Hovoríme, že prvok u je *algebraický* nad F, ak existuje nenulový polynóm $f(x) \in F[x]$, ktorého koreňom je u.

Ak každý prvok rozšírenia K je algebraický, hovoríme, že K je *algebraické rozšírenie.*

Príklad 14.4.2. Napríklad $\sqrt{3}$ je algebraický prvok na $\mathbb{Q}$, lebo je to koreň polynómu $x^2 - 3 \in \mathbb{Q}[x]$.

Nájsť konkrétny príklad prvku, ktorý nie je algebraický nad $\mathbb{Q}$, je ťažšie. (Také prvky sa zvyknú nazývať aj *transcendentné.*) Platí napríklad, že čísla π, e sú transcendentné nad $\mathbb{Q}$. Dôkaz však nie je jednoduchý. Na základe kardinality sa však dá zdôvodniť, že takéto čísla existujú – úloha 14.4.8. (Je to však iba existenčný dôkaz; nenašli sme konkrétny príklad takého čísla.)

Pomerne ľahko sa dá ukázať, že každé komplexné číslo je algebraické nad $\mathbb{R}$. Konkrétne $z = a + bi$ je koreňom polynómu $(x - a - bi)(x - a + bi) = x^2 - 2ax + a^2 + b^2$.

Ak u je algebraický nad F, znamená to, že množina všetkých polynómov, ktorých koreňom je u, je neprázdna. Ľahko sa overí, že táto množina

$$\{f(x) \in F[x]; f(u) = 0\}$$

je ideál v $F[x]$. Keďže $F[x]$ je okruh hlavných ideálov, existuje polynóm, ktorý generuje tento ideál.

Definícia 14.4.3. Ak u je algebraický prvok nad F, tak *minimálny polynóm* prvku u je normovaný polynóm, ktorý generuje ideál $\{f(x) \in F[x]; f(u) = 0\}$. Označujeme ho $m_u(x)$.

Stupeň algebraického prvku definujeme ako stupeň jeho minimálneho polynómu. Označujeme ho $[u : F]$.

$$[u : F] = \operatorname{st} m_u(x)$$

Pretože v definícii máme požiadavku normovanosti, minimálny polynóm je určený jednoznačne. Je to nenulový normovaný polynóm najnižšieho možného stupňa, ktorý patrí do ideálu $\{f(x) \in F[x]; f(u) = 0\}$.

Algebraický prvok môže patriť do rôznych rozšírení poľa F (napríklad $\sqrt{3}$ je prvkom $\mathbb{R}$ i $\mathbb{C}$, obe sú rozšírenia $\mathbb{Q}$). Pretože jeho definícia používa len ideál v $F[x]$, minimálny polynóm nezávisí od toho, aké rozšírenie obsahujúce u uvažujeme.

Príklad 14.4.4. Polynóm $x^2 - 3$ je minimálny polynóm čísla $\sqrt{3}$ nad $\mathbb{Q}$. (Nad $\mathbb{R}$ by mal tento prvok minimálny polynóm $x - \sqrt{3}$.)

Polynóm x^2+1 je minimálny polynóm čísla i nad $\mathbb{R}$. (Nad $\mathbb{C}$ by mal tento prvok minimálny polynóm $x - i$. Všeobecne, ak $u \in F$, tak $m_u(x) = x - u$.)

{algroz:VTFUAMINPOLY}

Veta 14.4.5. *Ak u je algebraický prvok nad F a $m_u(x) \in F[x]$ je jeho minimálny polynóm. Potom $m_u(x)$ je ireducibilný polynóm nad F,*

$$F(u) \cong F[x]/(m_u(x))$$

a $[u : F] = [F(u) : F]$.

Dôkaz. Ak by bol polynóm $m_u(x)$ reducibilný, t.j. $m_u(x) = f(x)g(x)$ pre nejaké nekonštantné polynómy $f(x), g(x) \in F[x]$, tak z rovnosti $m_u(u) = f(u)g(u) = 0$ vyplýva $f(u) = 0$ alebo $g(u) = 0$. To znamená, že jeden z polynómov $f(x)$, $g(x)$ by patril do ideálu $(m_u(x))$ a súčasne by mal nižší stupeň ako $m_u(x)$, čo je spor.

Z vety 14.3.11 potom vyplýva $F(u) \cong F[x]/(m_u(x))$ a z dôsledku 14.3.7 máme $[F(u) : F] = \operatorname{st} m_u = [u : F]$. □

Veta 14.4.6. *Nech K je rozšírenie F a $u \in K$. Prvok u je algebraický nad F práve vtedy, keď $F(u)$ je konečné rozšírenie F.*

Dôkaz. Ak u je algebraický, tak $F(u) \cong F[x]/(m_u(x))$ podľa vety 14.4.5, čo je konečné rozšírenie podľa dôsledku 14.3.7.

Obrátene, nech $F(u)$ je konečné rozšírenie F. Označme jeho stupeň n. Nech ďalej $a \in F(u)$. Potom $1, u, \ldots, u^n$ sú lineárne závislé v $F(u)$ (chápanom ako vektorový priestor nad F). Vyplýva to z toho, že máme $(n + 1)$ vektorov vo vektorovom priestore dimenzie n.

Teda existujú $c_0, c_1, \ldots, c_n$ (nie všetky nulové) tak, že $c_n u^n + \cdots + c_1 u + c_0 = 0$. Čiže $c_n x^n + \cdots + c_1 x + c_0 \in F[x]$ je nenulový polynóm, ktorého koreňom je u. Teda u je algebraický prvok nad F. □

{algroz:DOSINVJEALG}

Dôsledok 14.4.7. *Nech K je rozšírenie F a $u \in K$. Ak $u \neq 0$ a u je algebraický prvok nad poľom F, tak aj u^{-1} je algebraický nad F.*

Dôkaz. Stačí si uvedomiť, že $F(u) = F(u^{-1})$. □

Dôkaz by sme vedeli urobiť aj priamo z definície – úloha 14.4.1.

Dôsledok 14.4.8. *Každé konečné rozšírenie je algebraické.*

Dôkaz. Ak $u \in K$, kde K je konečné rozšírenie F, tak $F(u)$ je vektorový podpriestor priestoru K. Teda $F(u)$ je tiež konečnorozmerný priestor (konečné rozšírenie) a u je, na základe predchádzajúcej vety, algebraický prvok nad F. □

Tvrdenie 14.4.9. *Nech K je konečné rozšírenie poľa L a prvky $x_1, \ldots, x_n$ tvoria bázu K ako vektorového priestoru nad L. Nech L je konečné rozšírenie poľa F a prvky $y_1, \ldots, y_s$ tvoria bázu L ako vektorového priestoru nad F. Potom množina $\{x_i y_j; i = 1, \ldots, n, j = 1, \ldots, s\}$ tvorí bázu K ako vektorového priestoru nad F.*

Dôkaz. Podľa predpokladov každý prvok $k \in K$ možno vyjadriť ako

$$k = \sum_{i=1}^{n} c_i x_i$$

pre vhodné $c_1, \ldots, c_n \in L$. Ďalej každé $c_i \in L$ sa dá vyjadriť v tvare

$$c_i = \sum_{j=1}^{s} d_{ij} y_j,$$

kde $d_{ij} \in F$. Z týchto dvoch rovností dostávame vyjadrenie prvku x

$$x = \sum_{i=1}^{n} \sum_{j=1}^{s} d_{ij} x_i y_j$$

ako lineárnej kombinácie prvkov $x_i y_j$ s koeficientami z F.

Tým sme ukázali, že množina $\{x_i y_j; i = 1, \ldots, n, j = 1, \ldots, s\}$ generuje K ako vektorový priestor nad F. Aby sme ukázali, že ide o bázu, stačí nám už len overiť jej lineárnu nezávislosť.

Predpokladajme teda, že

$$\sum_{i=1}^{n} \sum_{j=1}^{s} a_{ij} x_i y_j = 0$$

pre nejaké koeficienty $a_{ij} \in F$.

Túto rovnosť môžeme prepísať do tvaru

$$\sum_{i=1}^{n} \left(\sum_{j=1}^{s} a_{ij} y_j \right) x_i = 0.$$

Dostali sme, že lineárna kombinácia prvkov $x_1, \ldots, x_n$ (s koeficientami z L) je rovná 0, pretože $x_1, \ldots, x_n$ je báza, každý koeficient musí byť nulový, teda pre každé $i = 1, \ldots, n$ máme

$$\sum_{j=1}^{s} a_{ij} y_j = 0.$$

Použitím rovnakého argumentu, tentoraz pre bázu $y_1, \ldots, y_s$ priestoru L nad F, máme, že všetky koeficienty a_{ij} sú nulové. Teda uvedené vektory sú skutočne lineárne nezávislé (ako prvky vektorového priestoru K nad poľom F). □

{algroz:DOSSTUPSUCIN}

Dôsledok 14.4.10. *Ak K je konečné rozšírenie poľa L a L je konečné rozšírenie poľa F, tak aj K je konečné rozšírenie poľa F a pre stupne rozšírení platí*

$$[K : F] = [K : L] \cdot [L : F].$$

Iný dôkaz tejto rovnosti pre prípad, keď F *je konečné pole,* je naznačený v úlohe 14.4.2.
{DOSSTUPUDELI}

Dôsledok 14.4.11. *Ak $u \in L$, kde L je konečné rozšírenie poľa F, tak*

$$[u : F] \mid [L : F].$$

{algroz:DOSUSCETSUCIN}

Dôsledok 14.4.12. *Ak u, v sú algebraické prvky nad F, tak aj ich súčet $u + v$ a súčin $u \cdot v$ sú algebraické prvky.*

Dôkaz. Vieme, že v je algebraický prvok nad F, teda je to aj algebraický prvok nad $F(u)$. (Je koreňom polynómu s koeficientami z F, tieto koeficienty súčasne patria do $F(u)$.)

To znamená, že $F(v)$ je konečné rozšírenie $F(u)$. Súčasne $F(u)$ je konečné rozšírenie F, čiže z dôsledku 14.4.10 máme, že $F(u)(v)$ je konečné rozšírenie poľa F.

Preto každý prvok poľa $F(u)(v)$ je algebraický nad F, špeciálne to platí aj pre prvky $u + v$, $u \cdot v$. □

{algroz:POZNALGPRVKYSUPO

Poznámka 14.4.13. Ak sa v nadpoli K pozeráme na všetky algebraické prvky, tak táto množina je uzavretá na súčin a súčet. Už skôr sme ukázali, že je uzavretá aj na inverzné prvky (dôsledok 14.4.7). Teda množina všetkých algebraických prvkov nad F (ležiacich v rozšírení K) opäť tvorí pole.

Príklad 14.4.14. Uvažujme rozšírenie $\mathbb{Q}(\sqrt{2}, \sqrt{3})$ poľa $\mathbb{Q}$, t.j. najmenšie podpole $\mathbb{C}$ obsahujúce $\mathbb{Q} \cup \{\sqrt{2}, \sqrt{3}\}$.

Vieme, že $\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2}; a, b \in \mathbb{Q}\}$ a $\mathbb{Q}(\sqrt{3}) = \{a + b\sqrt{3}; a, b \in \mathbb{Q}\}$, čiže $[\mathbb{Q}(\sqrt{2}) : \mathbb{Q}] = 2$ aj $[\mathbb{Q}(\sqrt{3}) : \mathbb{Q}] = 2$.

Pole $\mathbb{Q}(\sqrt{2}, \sqrt{3})$ môžeme chápať ako rozšírenie poľa $\mathbb{Q}(\sqrt{2})$, konkrétne platí

$$\mathbb{Q}(\sqrt{2}, \sqrt{3}) = \mathbb{Q}(\sqrt{2})(\sqrt{3}).$$

(Na oboch stranách rovnosti je najmenšie pole obsahujúce $\mathbb{Q} \cup \{\sqrt{2}, \sqrt{3}\}$.)

Vypočítajme stupeň $[\sqrt{3} : \mathbb{Q}(\sqrt{2})]$. Pretože $\sqrt{3}$ je koreň polynómu $x^2 - 3$ (s koeficientami z $\mathbb{Q}(\sqrt{2})$), jeho stupeň je najviac 2. Stupeň 1 by tento prvok mal iba ak by patril do $\mathbb{Q}(\sqrt{2})$. Z predpokladu $\sqrt{3} = a + b\sqrt{2}$ pre nejaké $a, b \in \mathbb{Q}$ však dostaneme

$$3 = a^2 + 2b^2 + 2ab\sqrt{2}$$

a z tejto rovnosti:
a) pre $ab \neq 0$ vyplýva $\sqrt{2} \in \mathbb{Q}$, čo je spor;
b) pre $b = 0$ vyplýva $\sqrt{3} = a \in \mathbb{Q}$, spor;
c) pre $a = 0$ vyplýva $\sqrt{3} = b\sqrt{2}$, čiže $\sqrt{6} = 2b \in \mathbb{Q}$, spor.

Teda platí $[\mathbb{Q}(\sqrt{3}) : \mathbb{Q}(\sqrt{2})] = [\sqrt{3} : \mathbb{Q}(\sqrt{2})] = 2$, z čoho dostaneme na základe predchádzajúcej vety

$$[\mathbb{Q}(\sqrt{2}, \sqrt{3}) : \mathbb{Q}] = [\mathbb{Q}(\sqrt{3}) : \mathbb{Q}(\sqrt{2})] \cdot [\mathbb{Q}(\sqrt{2}) : \mathbb{Q}] = 4$$

a z báz $1, \sqrt{2}$ pre $\mathbb{Q}(\sqrt{2})$ nad $\mathbb{Q}$ a $1, \sqrt{3}$ pre $\mathbb{Q}(\sqrt{2}, \sqrt{3})$ nad $\mathbb{Q}(\sqrt{2})$ vieme vytvoriť bázu $1, \sqrt{2}, \sqrt{3}, \sqrt{6}$ pre $\mathbb{Q}(\sqrt{2}, \sqrt{3})$ nad $\mathbb{Q}$, teda platí

$$\mathbb{Q}(\sqrt{2}, \sqrt{3}) = \{a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6}; a, b, c, d \in \mathbb{Q}\}.$$

Všimnime si, že

$$\mathbb{Q}(\sqrt{2}, \sqrt{3}) = \mathbb{Q}(\sqrt{2} + \sqrt{3}),$$

lebo každé nadpole $\mathbb{Q}$, ktoré obsahuje $u = \sqrt{2} + \sqrt{3}$, musí obsahovať aj

$$\frac{1}{u} = \frac{1}{\sqrt{2} + \sqrt{3}} \cdot \frac{\sqrt{3} - \sqrt{2}}{\sqrt{3} - \sqrt{2}} = \sqrt{3} - \sqrt{2},$$

a teda obsahuje aj prvky

$$\sqrt{3} = \frac{1}{2}(\sqrt{3}+\sqrt{2}) + \frac{1}{2}(\sqrt{3}-\sqrt{2}),$$
$$\sqrt{2} = \frac{1}{2}(\sqrt{3}+\sqrt{2}) - \frac{1}{2}(\sqrt{3}-\sqrt{2}).$$

Dá sa dokázať, že niečo podobné platí všeobecne – každé konečné rozšírenie $\mathbb{Q}$ je jednoduché. (Podobne pre ľubovoľné pole nekonečnej charakteristiky.)

Všimnime si, že mocniny prvku $\sqrt{2}+\sqrt{3}$ vieme vyjadriť ako lineárne kombinácie 1, $\sqrt{2}$, $\sqrt{3}$, $\sqrt{6}$, konkrétne

$$(\sqrt{2}+\sqrt{3})^0 = 1$$
$$(\sqrt{2}+\sqrt{3})^1 = \sqrt{2}+\sqrt{3}$$
$$(\sqrt{2}+\sqrt{3})^2 = 5+2\sqrt{6}$$
$$(\sqrt{2}+\sqrt{3})^3 = 11\sqrt{2}+9\sqrt{3}$$
$$(\sqrt{2}+\sqrt{3})^4 = 49+20\sqrt{6}$$

Máme teda 5 vektorov $(1,0,0,0)$, $(0,1,1,0)$, $(5,0,0,2)$, $(0,11,9,0)$, $(49,0,0,20)$ v priestore dimenzie 4 – sú teda lineárne závislé a riešením sústavy lineárnych rovníc vieme nájsť nenulové koeficienty také, že príslušná lineárna kombinácia týchto vektorov je nulový vektor.

Dostaneme tak $1-10u^2+u^4=0$, čo znamená, že

$$x^4-10x^2+1$$

je minimálny polynóm prvku $u=\sqrt{2}+\sqrt{3}$. Teda

$$\mathbb{Q}(\sqrt{2},\sqrt{3}) \cong \mathbb{Q}[x]/(x^4-10x^2+1).$$

Môžeme si ukázať aj inú možnosť, ako sa dá nájsť $m_u(x)$. Pre prvok $u=\sqrt{2}+\sqrt{3}$ máme

$$u-\sqrt{2} = \sqrt{3}$$
$$(u-\sqrt{2})^2 = 3$$
$$u^2-2\sqrt{2}u+2 = 3$$
$$u^2-1 = 2\sqrt{2}u$$
$$(u^2-1)^2 = 8u^2$$
$$u^4-2u^2+1 = 8u^2$$
$$u^4-10u^2+1 = 0$$

Zistili sme, že u je koreňom polynómu x^4-10x^2+1, pričom tento polynóm je normovaný a má racionálne koeficienty. Keďže sme už predtým videli, že $[u:F]=4$, tak toto je presne mininmálny polynóm $m_u(x)$.

Cvičenia

{algrozcvic:ULOINVKOREN}

Úloha 14.4.1. Nech K je rozšírenie poľa F a $u \in K \setminus \{0\}$ je koreňom polynómu $f(x) = c_n x^n + \cdots + c_1 x + c_0 \in F[x]$. Nájdite taký polynóm z $F[x]$, ktorého koreňom je u^{-1}.

SUCINKONECNE}

Úloha 14.4.2. Nech F je konečné pole a L je jeho rozšírenie stupňa $d = [K : F]$. Ukážte, že potom pre počty prvkov platí $|K| = |F|^d$. Vedeli by ste s využitím tohto faktu urobiť iný dôkaz rovnosti $[K : F] = [K : L] \cdot [L : F]$ pre prípad, že F *je konečné* pole?

Úloha 14.4.3. Nech L je rozšírenie poľa F a $u \in L$. Dokážte, že ak $[F(u) : F] = 5$, tak $[F(u^2) : F] = 5$.

Úloha 14.4.4. Ak $[L : F]$ je prvočíslo, tak pre každé $u \in L$ platí $u \in F$ alebo $F(u) = L$.

Úloha 14.4.5. V poli $\mathbb{Q}(\sqrt[3]{2})$ nájdite inverzný prvok ku $1 - 2\sqrt[3]{2} + \sqrt[3]{4}$ (treba ho vyjadriť ako $a + b\sqrt[3]{2} + c\sqrt[3]{4}$ pre vhodné $a, b, c \in \mathbb{Q}$.)

Úloha 14.4.6. Nájdite minimálne polynómy týchto čísel nad $\mathbb{Q}$:
a) $\sqrt{2} + 1$; b) $2 - 3\sqrt{5}$; c) $\sqrt[3]{3} + \sqrt{3}$; d) $\sqrt{2} - \sqrt{3}$; e) $\sqrt[3]{2} + i$; f) $1 + \sqrt[3]{2} - \sqrt[3]{4}$; g) $\frac{1}{2-\sqrt[3]{2}} + \sqrt[3]{4}$; h) $\frac{3+\sqrt{7}}{1+2\sqrt{7}}$; i) $\sqrt[3]{2} + \sqrt{2}$

Úloha 14.4.7. Určite stupeň viacnásobného rozšírenia a nájdite bázu nad $\mathbb{Q}$:
a) $\mathbb{Q}(\sqrt{3}, \sqrt{5})$; b) $\mathbb{Q}(i, \sqrt{2})$; c) $\mathbb{Q}(\sqrt[3]{5}, \sqrt[3]{25})$; d) $\mathbb{Q}(1 + \sqrt{2}, 1 - \sqrt{8})$

{algrozcvic:ULOKARD}

Úloha 14.4.8*. Aká je kardinalita množiny $\mathbb{Q}[x]$? Aká je kardinalita množiny všetkých čísel, ktoré môžeme dostať ako korene polynómov s racionálnymi koeficientami? Viete na základe toho zdôvodniť, že existuje aspoň jedno reálne číslo, ktoré je transcendentné? (T.j. aspoň jedno číslo $a \in \mathbb{R}$, ktoré nie je koreňom žiadneho polynómu $f(x) \in \mathbb{Q}[x]$.)

14.5 Rozkladové polia

{rozklpol:DEF}

Definícia 14.5.1. Nech F je pole, $f(x) \in F[x]$ je nekonštantný polynóm. Rozšírenie K poľa F nazývame *rozkladovým poľom polynómu $f(x)$ nad F,* ak existujú $c \in F$, $u_1, \ldots, u_n \in K$ také, že $L = F(u_1, \ldots, u_n)$ a f sa dá nad L rozložiť ako

$$f(x) = c(x - u_1)(x - u_2) \ldots (x - u_n).$$

V príklade 14.3.8 sme vlastne zostrojili rozkladové pole polynómu $x^2 + x + 1$ nad $\mathbb{Z}_2$.

{rozklpol:VTEXIS}

Veta 14.5.2. *Nech F je pole, $f(x) \in F[x]$ a* st $f = n > 0$. *Potom existuje rozšírenie K poľa F, ktoré je rozkladovým poľom polynómu $f(x)$.*

Dôkaz. Indukciou vzhľadom na n. Ak $n = 1$, tak je rozkladovým poľom priamo F.

Nech teraz $n > 1$ a tvrdenie platí pre všetky polynómy stupňa menšieho ako n nad ľubovoľným poľom. Podľa vety 14.3.5 existuje rozšírenie poľa F, v ktorom má f aspoň jeden koreň u. (Stačí vo vete 14.3.5 za $p(x)$ zobrať ktorýkoľvek ireducibilný polynóm deliaci $f(x)$.) Uvažujme pole $F(u)$. Polynóm $f(x)$ možno nad $F(u)$ rozložiť ako $(x-u)g(x)$, pričom st $g < n$. Nech $F(u)(u_2, \ldots, u_n)$ je rozkladové pole $g(x)$ nad $F(u)$. Ľahko vidíme, že $F(u)(u_2, \ldots, u_n) = F(u, u_2, \ldots, u_n)$ je rozkladové pole polynómu $f(x)$. (Polynóm $f(x)$ v ňom možno rozložiť na súčin koreňových činiteľov a toto pole je generované n koreňmi polynómu $f(x)$.) □

Predchádzajúca veta hovorí, že pre daný polynóm stupňa n existuje pole v ktorom tento polynóm má n koreňov. Steinitzova veta 13.5.16, ktorú sme si uviedli bez dôkazu, je podstatne silnejší výsledok – tam máme jediné pole, ktoré spĺňa túto vlastnosť pre všetky polynómy z $F[x]$.

Ďalej ukážeme že rozkladové pole polynómu $f(x)$ nad poľom F je určené jednoznačne až na izomorfizmus. (V dôkaze sme svedkami situácie, s ktorou sme sa už viackrát stretli – pri dôkaze indukciou je niekedy výhodnejšie dokazovať o niečo silnejšie tvrdenie, pretože potom nám silnejšie predpoklady môžu zjednodušiť dôkaz indukčného kroku.)

{rozklpol:VTJ

Veta 14.5.3. *Nech $\varphi\colon F \to F'$ je izomorfizmus polí, $f(x) \in F[x]$ a $f'(x) \in F'[x]$ je polynóm, ktorý získame z $f(x)$ aplikovaním φ na všetky koeficienty polynómu $f(x)$. (V označení z poznámky 14.3.12 to znamená $f'(x) = \hat{\varphi}(f(x))$.) Ak K je rozkladové pole polynómu $f(x)$ a L je rozkladové pole polynómu $f'(x)$, tak existuje izomorfizmus $\sigma\colon K \to L$, ktorý navyše rozširuje φ, t.j. $\sigma|_F = \varphi$.*

Dôkaz. Dôkaz urobíme indukciou vzhľadom na stupeň n polynómu $f(x)$.

1° Ak stupeň $f(x)$ je 1, tak jeho rozkladovým poľom je priamo pole F. Teda v tomto prípade tvrdenie platí.

2° Predpokladajme, že tvrdenie platí pre ľubovoľnú dvojicu izomorfných polí a ľubovoľný polynóm stupňa menšieho ako n. Nech $K = F(u_1 \dots u_n)$ je rozkladové pole polynómu $f(x)$ nad poľom F a nech $L = F'(v_1 \dots v_n)$ je rozkladové pole $f'(x)$ nad F'. Potom tieto polynómy môžeme rozložiť ako $f(x) = c(x-u_1)\dots(x-u_n)$ a $f'(x) = c'(x-v_1)\dots(x-v_n)$.

Súčasne máme $K = F(u_1)(u_2 \dots u_n)$ (t.j. K je rozkladové pole polynómu $f(x)$ nad $F(u_1)$) a takisto $L = F'(v_1)(v_2 \dots v_n)$. Navyše môžeme predpokladať, že u_1 a v_1 sú koreňmi navzájom si zodpovedajúcich ireducibilných faktorov polynómov $f(x)$ a $f'(x)$. (To sa dá dosiahnuť prípadnou výmenou koreňov.) Potom sú podľa vety 14.3.13 možno izomorfizmus φ rozšíriť na izomorfizmus $\sigma'\colon F(u_1) \to F'(v_1)$ taký, že $\sigma'|_F = \varphi$. Na základe indukčného predpokladu môžeme potom tento izomorfizmus rozšíriť na izomorfizmus $\sigma\colon K \to L$. □

{rozklpol:DOSJEDN}

Dôsledok 14.5.4. *Ľubovoľné dve rozkladové polia polynómu $f(x)$ nad F sú izomorfné.*

Predchádzajúci výsledok nám umožňuje dokázať úplnú charakterizáciu konečných polí. Vieme už, že počet prvkov konečného poľa musí byť mocninou prvočísla. Dokážeme, že pre každé $q = p^n$ (p je prvočíslo) existuje q-prvkové pole a je určené jednoznačne až na izomorfizmus.

Veta 14.5.5. *Nech $q = p^n$, kde p je prvočíslo a $n > 0$ je prirodzené číslo. Potom existuje (až na izomorfizmus jediné) q-prvkové pole. Je to rozkladové pole polynómu $x^q - x$ nad $\mathbb{Z}_p$.*

Dôkaz. Keďže pole s uvedenými vlastnosťami má charakteristiku p, obsahuje ako svoje podpole $\mathbb{Z}_p$.

Najprv si všimnime, že ak q-prvkové pole existuje, musí to byť skutočne rozkladové pole polynómu $x^q - x$ nad $\mathbb{Z}_p$. Vyplýva to z toho, že pre každé $x \neq 0$ platí $x^{q-1} = 1$. (Dostaneme to z Lagrangeovej vety – pozri dôsledok 12.2.15; alebo tiež z úloh 3.3.14 či 3.3.15.) Teda skutočne každý prvok poľa F je koreňom polynómu $x^q - x$. Zistili sme teda, že $x^q - x$ má v tomto poli q navzájom rôznych koreňov a teda sa dá rozložiť na súčin lineárnych faktorov. (Konkrétne sa dá rozložiť ako $x^q - x = \prod_{u \in F} (x - u)$.)

Stačí nám teda overiť, že rozkladové pole polynómu $x^q - x$ má práve q prvkov. Všimnime si, že v poli charakteristiky p platí $(a+b)^p = a^p + b^p$ a dokonca aj

$$(a+b)^q = a^q + b^q$$

(pozri tvrdenie 14.2.10) To znamená, že korene polynómu $x^q - x$ sú uzavreté vzhľadom na sčitovanie. Ľahko potom vidno, že sú uzavreté aj na rozdiel a násobenie.

Vieme dostať aj uzavretosť vzhľadom na inverzné prvky - ak nejaké nenulové a spĺňa $a^q = a$, a teda aj $a^{q-1} = 1$, tak aj pre inverzný prvok platí $(a^{-1})^{q-1} = 1$ a aj $(a^{-1})^q = a^{-1}$. Je teda koreňom toho polynómu $x^q - x$.

Teda samotné korene už tvoria pole – bude to rozkladové pole polynómu $x^q - x$, ktoré má práve q prvkov – q rôznych koreňov tohoto polynómu. (Pomocou tvrdenia 13.5.31 ľahko ukážeme, že polynóm $x^q - x$ nemá násobné korene, keďže jeho derivácia je -1. Tento fakt môžeme zdôvodniť aj bez použitia formálnej derivácie – úloha 14.5.2.) □

Príklad 14.5.6. V príklade 14.3.8 sme videli príklad 4-prvkového poľa. Podľa predošlej vety by to malo byť, až na izomorfizmus, práve rozkladové pole polynómu $x^4 - x$.

Skutočne sa môžeme presvedčiť, že v tomto poli dostaneme

$$x^4 - x = x^4 + x = x(x^3 + 1) = x(x + 1)(x^2 + x + 1) = x(x + 1)(x + u)(x + u + 1).$$

Teda tento polynóm sa dá rozložiť na koreňové činitele, pričom korene sú práve všetky prvky poľa: 0, 1, u, $u + 1$.

Cvičenia

{rozkpolcvic:ULOXN-1}

Úloha 14.5.1. Nech F je pole a $c \in F$. Označme $f(x) = x^n - 1$.
a) Ukážte, že ak c je koreň polynómu $f(x)$, tak

$$f(x) = (x - c)(x^{n-1} + cx^{n-2} + c^2x^{n-3} + \cdots + c^{n-2}x + c^{n-1}).$$

b) Ukážte, že ak $f(x)$ má násobný koreň F, tak F má konečnú charakteristiku a $\mathrm{char}(F) \mid n$.

{rozkpolcvic:ULONEMANASO

Úloha 14.5.2. Ukážte s použitím úlohy 14.5.1 že polynóm

$$f(x) = x^{p^n} - x$$

v poli charakteristiky p nemá násobné korene.

Úloha 14.5.3. Nájdite izomorfizmus medzi poľami $\mathbb{Z}_3[x]/(x^2 + 1)$ a $\mathbb{Z}_3[x]/(x^2 + x + 2)$. (Poznámka: Z výsledkov v tejto kapitole vieme dostať, že obe sú 9-prvkové polia a teda nutne musia byť izomorfné. V tejto úlohe by sme chceli priamo nájsť konkrétny izomorfizmus.)

Úloha 14.5.4*. a) Ukážte, že $\mathbb{Z}_3[x]/(x^2 + x + 2)$ je pole. Koľko má prvkov?
b) Nájdite rozklad polynómu $x^2 + 1$ v tomto poli na ireducibilné polynómy.
c) Nájdite rozklad polynómu $x^4 + 1$ v tomto poli na ireducibilné polynómy.

Dodatok A

Základné fakty z teórie čísel

{APTC}

V tejto časti spomenieme niektoré základné fakty z teórie čísel. Mnohé z nich by ste mali poznať zo strednej školy a tiež z iných predmetov na vysokej škole.

V texte používame najmä fakty týkajúce sa deliteľnosti a prvočísel – ktoré by vám mali byť známe. Uvádzame tu nejaký prehľad, pričom sú uvedené bez dôkazu. Spomenuli sme aj základné veci o kongruenciách – hoci ich v tomto texte nevyhnutne nepotrebujeme. Zdalo sa to však prirodzené, pretože takýto typ kongruencie je špeciálny typ okruhovej kongruencie, ktorú zavádzame pri faktorových okruhoch. Takisto bolo vhodné na tomto mieste spomenúť najväčší spoločný deliteľ a Euklidov algoritmus – ide o pomerne známy algoritmus a úzko súvisí aj s euklidovskými okruhmi, ktorým sa v tomto texte zaoberáme.

A.1 Veta o delení so zvyškom

A.2 Deliteľnosť a prvočísla

A.3 Kongruencie

Veta A.3.1 (Malá Fermatova veta)**.** *Nech p je prvočíslo, $a \in \mathbb{Z}$. Potom platí*

$$a^p \equiv a \pmod{p}.$$

Ak $p \nmid a$, tak platí

$$a^{p-1} \equiv 1 \pmod{p}.$$

Nejaké možnosti ako sa dá dokázať táto veta sme videli v úlohách 3.3.14 a 3.3.15.

A.4 Najväčší spoločný deliteľ a Euklidov algoritmus

Dodatok B

Lineárne rekurencie

{APLINREK}

TODO lineárne homogénne rekurencie s konštantnými koeficientami

Budeme sa teda zaoberať postupnosťami, ktoré sú určené podmienkou tvaru

$$A_n = c_{k-1}A_{n-1} + c_{k-2}A_{n-2} + \cdots + c_1 A_{n-k+1} + c_0 A_{n-k} \tag{B.1}$$

{rekur:EQREK}

kde $c_0, \ldots, c_k$ sú nejaké zadané konštanty a $(A_n)_{n=0}^{\infty}$ je postupnosť určená uvedeným predpisom. Bude predpokladať, že všetky konštanty c_k a aj členy postupnosti A_k sú komplexné čísla.

Asi je pomerne ľahké vidieť to, že ak požadujeme, aby postupnosť $(A_n)_{n=0}^{\infty}$ vyhovovala podmienke (B.1) a súčasne máme zadané hodnoty pre $A_0, \ldots, A_{k-1}$, tak už je celá postupnosť určená jednoznačne. Konkrétne vidíme, že zo zadaných hodnôt je jednoznačne určené čomu sa rovná A_k. Z týchto hodnôt už potom máme jednoznačne určené aj A_{k+1}, atď

$$\begin{aligned}
A_k &= c_{k-1}A_{k-1} + c_{k-2}A_{k-2} + \cdots + c_1 A_1 + c_0 A_0 \\
A_{k+1} &= c_{k-1}A_k + c_{k-2}A_{k-1} + \cdots + c_1 A_2 + c_0 A_1 \\
A_{k+2} &= c_{k-1}A_{k+1} + c_{k-2}A_k + \cdots + c_1 A_3 + c_0 A_2 \\
&\vdots
\end{aligned}$$

{rekur:PRIKLFIB}

Príklad B.0.1.

Definícia B.0.2.

Dodatok C

Komplexné čísla

{APKOMPL}

Základné vlastnosti komplexných čísel môžete nájsť vo veľkom množstve vysokoškolských i stredoškolských učebníc, ako napríklad [I, Kapitola 13], [KMŠ, Kapitola II.10], [Sm], [Bl] a mnoho iných. Kniha [AA] sa venuje komplexným číslam od základných poznatkov až po ich použitie v úlohách olympiádneho charakteru.

Niektorí z vás preberali komplexné čísla na strednej škole, pre tých, ktorí ich nemali, sa tu pokúsime zhrnúť ich najdôležitejšie vlastnosti, ktoré budete potrebovať.

C.1 Definícia komplexných čísel, algebraický tvar komplexného čísla

Vieme, že v reálnych číslach nemá rovnica

$$x^2 = -1$$

riešenie. (Pre každé reálne číslo platí $x^2 \geq 0$.) Čo by sme potrebovali urobiť, keby sme chceli dostať číselný obor, v ktorom táto rovnica bude mať riešenie? Znamená to vlastne, že chceme k reálnym číslam pridať nejaké „nové“ čísla a zadefinovať na nich sčitovanie a násobenie tak, aby sa tieto operácie správali podobne ako pre reálne čísla. (Pod slovom „podobne“ rozumieme to, že novovytvorený číselný obor má byť pole.) Ideme sa teraz pokúsiť zadefinovať takéto pole, ktoré potom nazveme poľom komplexných čísel.

Určite musíme teda pridať aspoň jedno riešenie rovnice $x^2 = -1$. Označíme ho i a budeme ho nazývať *imaginárna jednotka*. Teda

{komplex:EQIMAG}

$$i^2 = -1. \tag{C.1}$$

Pretože chceme, aby komplexné čísla obsahovali všetky reálne čísla musíme potom pridať aj čísla tvaru $b.i$, pre ľubovoľné $b \in \mathbb{R}$ a aj čísla tvaru $a + bi$ pre ľubovoľné $a \in \mathbb{R}$. Ukážeme si, že tieto čísla už postačia na to, aby sme vytvorili pole.

Definícia C.1.1. *Komplexným číslom* budeme nazývať ľubovoľné číslo tvaru

$$a + bi,$$

kde $a, b \in \mathbb{R}$. Množinu všetkých komplexných čísel označujeme

$$\mathbb{C} = \{a + bi; a, b \in \mathbb{R}\}.$$

Zápis komplexného čísla v tvare $a+bi$ nazývame *algebraický zápis* komplexného čísla. Pritom a sa nazýva *reálna časť* komplexného čísla a bi sa nazýva *imaginárna časť* komplexného čísla. Pre komplexné číslo $z = a+bi$ označujeme jeho reálnu časť $\operatorname{Re} z = a$ a imaginárnu časť $\operatorname{Im} z = bi$. (Niekedy sa tiež používa označenie $\Re z$ a $\Im z$.) Číslo, ktoré má nulovú reálnu časť, sa nazýva *rýdzoimaginárne.*

Komplexné číslo je jednoznačne určené svojou reálnou a imaginárnou časťou, teda dve komplexné čísla $z_1 = a_1 + b_1 i$ a $z_2 = a_2 + b_2 i$ sa rovnajú práve vtedy, keď

$$a_1 = a_2 \quad \text{a} \quad b_1 = b_2.$$

Poznámka C.1.2. Definíciu komplexných čísel môžeme chápať takým spôsobom, že sme zaviedli nejaký nový symbol i a komplexné čísla sú formálne zápisy tvaru $a + bi$, pričom a, b sú ľubovoľné reálne čísla.

Iná možnosť by bola definovať komplexné čísla ako usporiadané dvojice reálnych čísel a dohodnúť sa, že namiesto (a, b) budeme používať zápis $a+bi$. (Všimnite si, že takýto prístup zodpovedá tomu, že 2 komplexné čísla považujeme za rovnaké práve vtedy, keď majú rovnaké obe zložky.)

Pri prvom prístupe (formálne zápisy tvaru $a+bi$) je jasné, že reálne čísla sú podmnožinou komplexných čísel. (Každé reálne číslo a sa dá vyjadriť ako $a + 0i$, teda je prvkom množiny $\mathbb{C}$.) Pri druhom prístupe reálne čísla stotožníme s množinou $\{(a, 0); a \in \mathbb{R}\}$. Sčitovanie a násobenie definujeme tak, aby korešpondovali so sčitovaním a násobením reálnych čísel.

Pomocou nového symbolu i sme zaviedli nejakú množinu, ktorej prvky sme nazvali komplexné čísla. Ďalej by sme na tejto množine chceli zaviesť operácie sčitovania a násobenia tak, aby množina $\mathbb{C}$ s týmito operáciami tvorila pole.

Súčet 2 komplexných čísel definujeme veľmi prirodzeným spôsobom – sčítame ich reálne časti aj imaginárne časti:

$$(a + bi) + (c + di) = (a + c) + (b + d)i \tag{C.2}$$ {komplex:EQSUCET}

Ak chceme, aby platila distributívnosť, tak pre súčin čísel $a + bi$ a $c + di$ musí platiť

$$(a + bi)(c + di) = ac + bci + adi + bdi^2$$

a z rovnosti (C.1) potom máme

$$(a + bi)(c + di) = (ac - bd) + (bc + ad)i \tag{C.3}$$ {komplex:EQSUCIN} {komplex:PRSUCIN}

Príklad C.1.3. $(2 + 3i) + (2 - i) = 4 + 2i$
$(2 + 3i)(2 - i) = 4 - 2i + 6i + 3 = 7 + 4i$
$(\sqrt{2} + \sqrt{2}i)(\sqrt{2} + \sqrt{2}i) = 2 + 2i + 2i - 2 = 4i$

Oplatí sa zapamätať si, že

$$\begin{aligned} i^1 = i \quad i^2 = -1 \qquad i^3 = -i \quad i^4 = 1 \\ i^{4k+1} = i \quad i^{4k+2} = -1 \quad i^{4k+3} = -i \quad i^{4k} = 1 \end{aligned} \tag{C.4}$$ {komplex:EQMOCi}

Poznámka C.1.4. Násobenie a sčitovanie by sme definovali analogicky, keby sme komplexné čísla chápali ako dvojice reálnych čísel, pozri úloha 3.3.11.

Dôležité je, že takto definované operácie + a · sa správajú „rozumne“. Inak povedané, radi by sme ukázali, že $(\mathbb{C}, +, \cdot)$ je pole.

Niektoré z vlastností poľa sú zrejmé takmer okamžite. Komutatívnosť a asociatívnosť operácie + sa overí ľahko. Neutrálny prvok tejto operácie je $0 = 0 + 0i$ a inverzný prvok k $a + bi$ je $-(a + bi) = (-a) + (-b)i$. Z toho špeciálne vyplýva, že komplexné čísla vieme aj odčitovať,

$$(a + bi) - (c + di) = (a - c) + (b - d)i.$$

(Môžete si všimnúť, že ak sa na komplexné čísla pozeráme ako na dvojice reálnych čísel, tak je to tá istá operácia, ktorú sme zaviedli v úlohe 3.2.1g), resp. v príklade 4.1.3, kde sme videli, že dvojice reálnych čísel môžeme chápať ako vektorový priestor.)

V prípade operácie · máme o trochu komplikovanejšiu situáciu. Jej komutatívnosť je jasná priamo z definície. Skúsme overiť asociatívnosť. Teda máme 3 komplexné čísla $z_1 = a + bi$, $z_2 = c + di$ a $z_3 = e + fi$ a chceme priamym výpočtom (teda na základe definície násobenie) overiť $z_1(z_2z_3) = (z_1z_2)z_3$.

$$\begin{aligned} z_1(z_2z_3) &= (a + bi)[(c + di)(e + fi)] = (a + bi)[(ce - df) + (cf + de)i] = \\ &= (ace - adf - bcf - bde) + (acf + ade + bce - bdf)i \\ (z_1z_2)z_3 &= [(a + bi)(c + di)](e + fi) = [(ac - bd) + (bc + ad)i](e + fi) = \\ &= (ace - bde - bcf - adf) + (acf - bdf + bce + ade)i \end{aligned}$$

Vidíme, že v oboch prípadoch sme dostali taký istý výsledok.

Ľahko sa overí, že neutrálny prvok pre násobenie je $1 = 1 + 0i$.

Potrebovali by sme ešte zistiť, či vieme komplexné čísla deliť (z toho dostaneme existenciu inverzného prvku). Deliť komplexným číslom $c + di$ môžeme iba vtedy, ak je toto číslo rôzne od nuly. Teda $c + di \neq 0 + 0i$, čo znamená, že buď $c \neq 0$ alebo $d \neq 0$. Spôsob, akým to urobíme, sa podobá na trik, ktorým obvykle odstraňujeme z menovateľa odmocninu.

$$\frac{a + bi}{c + di} = \frac{a + bi}{c + di} \cdot \frac{c - di}{c - di} = \frac{(a + bi)(c - di)}{(c + di)(c - di)} = \frac{(ac + bd) + (bc - ad)i}{c^2 + d^2} = \frac{ac + bd}{c^2 + d^2} + \frac{bc - ad}{c^2 + d^2}i \tag{C.5}$$

{komplex:EQDEL}

Všimnime si, že ak je aspoň jedno z reálnych čísel c, d nenulové, tak $c^2 + d^2 > 0$, čiže v predchádzajúcom výraze nevystupuje v menovateli nula.

{komplex:PRDEL}

Príklad C.1.5. $\frac{1+2i}{2-i} = \frac{(1+2i)(2+i)}{(2-i)(2+i)} = \frac{5i}{5} = i$

Jediná vlastnosť z definície poľa, ktorú sme zatiaľ neoverili, je distributívnosť, t.j.

$$z_1(z_2 + z_3) = z_1z_3 + z_2z_3.$$

Táto vlastnosť sa opäť dá overiť priamym výpočtom (úloha C.4.3).

Overením jednotlivých vlastností sme dokázali:

Veta C.1.6. *Komplexné čísla s operáciami + a · definovanými vzťahmi* (C.2) *a* (C.3) *tvoria pole.*

Akonáhle máme dokázanú túto vetu, vieme, že pre komplexné čísla môžeme používať všetky vlastnosti z tvrdenia 3.3.4 a takisto vlastnosti, ktoré sme dokázali v cvičeniach v časti 3.3. (Takisto, keďže sme sa naučili riešiť sústavy lineárnych rovníc, počítať determinanty, inverzné matice a mnohé ďalšie veci v ľubovoľnom poli, vieme to robiť aj v poli komplexných čísel.)

Číslo $a - bi$, ktorým sme rozšírili čitateľ aj menovateľ pri výpočte podielu dvoch komplexných čísel (pozri (C.5) a príklad C.1.5) sa vyskytuje v súvislosti s číslom $a + bi$ pomerne často.

Definícia C.1.7. *Komplexne združeným číslom* k číslu $z = a + bi$ nazývame číslo $\overline{z} = a - bi$.

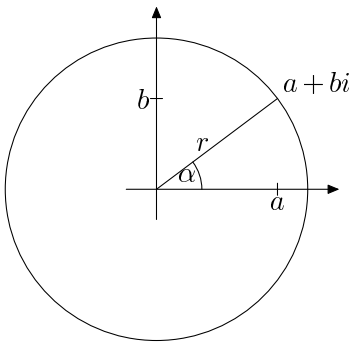
Úloha C.1.1. Overte, že platí (pre ľubovoľné $z, z_1, z_2 \in \mathbb{C}$)

$$\begin{aligned} \overline{z_1 + z_2} &= \overline{z_1} + \overline{z_2} \\ \overline{z_1 \cdot z_2} &= \overline{z_1} \cdot \overline{z_2} \\ z \cdot \overline{z} &= |z|^2 \\ z = \overline{z} \quad &\Leftrightarrow \quad z \text{ je reálne} \\ z = -\overline{z} \quad &\Leftrightarrow \quad z \text{ je rýdzoimaginárne} \end{aligned}$$

Symbol $|z|$ označuje absolútnu hodnotu komplexného čísla z. Ak $z = a + bi$, tak absolútna hodnota je definovaná ako $|z| = \sqrt{a^2 + b^2}$. (Budeme sa jej venovať o chvíľu.)

C.2 Geometrická interpretácia komplexných čísel, goniometrický tvar, Moivrova veta

Ako sme už spomenuli, komplexné čísla môžeme stotožniť s dvojicami reálnych čísel. Takisto vieme, že dvojiciam reálnych čísel vieme jednojednoznačne priradiť aj body v rovine. Čiže komplexné čísla môžeme chápať aj ako body v rovine. V tejto podkapitole uvidíme, že takáto interpretácia komplexných čísel poskytuje zaujímavú interpretáciu pre sčitovanie a násobenie komplexných čísel.



Obr. C.1: Znázornenie komplexného čísla v rovine

Keď stotožníme komplexné číslo s bodom v rovine, môžeme sa pozrieť na jeho vzdialenosť od počiatku súradnicovej sústavy a na uhol, ktorý zviera s osou x. Majme komplexné číslo $z = a + bi$, ktorému zodpovedá bod (a, b). Z Pytagorovej vety vieme vzdialenosť od počiatku určiť ako

$$r = \sqrt{a^2 + b^2}$$

a uhol medzi spojnicou bodov $(0, 0)$, (a, b) a osou x sa dá zistiť z rovností $\cos\varphi = \frac{a}{r}$ a $\sin\varphi = \frac{b}{r}$. Pre tieto hodnoty r a φ platí

$$a + bi = r(\cos\varphi + i\sin\varphi).$$

Definícia C.2.1. Zápis komplexného čísla v tvare

$$z = r(\cos\varphi + i\sin\varphi)$$

nazývame *goniometrický zápis* komplexného čísla. Číslo $r = \sqrt{a^2+b^2}$ nazývame *absolútna hodnota* alebo tiež *modul* komplexného čísla z a označujeme ho $|z|$. Číslo φ také, že $a = r\cos\varphi$ a $b = r\sin\varphi$ nazývame *argument* komplexného čísla z.

Príklad C.2.2. Pokúsme sa previesť do goniometrického tvaru číslo $z = 1-\sqrt{3}i$. Dostávame $r = |z| = \sqrt{1+3} = 2$. Z toho dostávame, že pre argument čísla z musí platiť

$$\cos\varphi = \frac{1}{2}$$
$$\sin\varphi = -\frac{\sqrt{3}}{2}$$

Riešeniami prvej rovnice sú práve uhly

$$\varphi = \pm\frac{\pi}{3} + 2k\pi$$

pre $k \in \mathbb{Z}$. Keď vezmeme do úvahy, že sínus má byť záporný, dostaneme

$$\varphi = -\frac{\pi}{3} + 2k\pi.$$

(Tým je uhol φ určený až na násobok 2π. Otočenie o uhol 2π okolo počiatku samozrejme bod v rovine nemení.)

Obrátene, ak máme daný goniometrický tvar komplexného čísla, ľahko ho prevedieme na algebraický tvar. Napríklad

$$\sqrt{2}\left(\cos\frac{\pi}{4} + i\sin\frac{\pi}{4}\right) = \sqrt{2}\left(\frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2}i\right) = 1+i.$$

Nasledujúca veta hovorí, že ak máme komplexné čísla zapísané v goniometrickom tvare, tak pri ich vynásobení sa vynásobia ich absolútne hodnoty a ich argumenty sa sčítajú.

{komplex:VTMOIVRE}

Veta C.2.3 (Moivrova veta)**.** *Nech* $z_1 = r_1(\cos\alpha + i\sin\alpha)$ *a* $z_2 = r_2(\cos\beta + i\sin\beta)$*. Potom pre ich súčin platí*

{komplex:EQMOIVRE}
$$z_1 z_2 = r_1 r_2(\cos(\alpha+\beta) + i\sin(\alpha+\beta)). \tag{C.6}$$

Špeciálne z toho vyplýva, že pre absolútne hodnoty platí

{komplex:EQABS}
$$|z_1 z_2| = |z_1|.|z_2|. \tag{C.7}$$

Dôkaz. Overme najprv rovnosť (C.7). Označme $z_1 = a+bi$ a $z_2 = c+di$, teda $z_1z_2 = (ac-bd)+(ad+bc)i$. Upravujme najprv $|z_1z_2|$. Lepšie sa nám bude pracovať bez druhej odmocniny, preto tento výraz umocnime na druhú.

$$|z_1z_2|^2 = (ac-bd)^2+(ad+bc)^2 = (a^2c^2-2abcd+b^2d^2)+(a^2d^2+2abcd+b^2c^2) = (ac)^2+(bd)^2+(ad)^2+(bc)^2$$

Teraz sa pokúsime upraviť $|z_1|^2.|z_2|^2$

$$|z_1|^2.|z_2|^2 = (a^2+b^2)(c^2+d^2) = (ac)^2+(ad)^2+(bc)^2+(bd)^2$$

Vidíme, že v oboch prípadoch sme dostali rovnaký výsledok. Teda $|z_1 z_2|^2 = (|z_1|.|z_2|)^2$. Pretože ide o nezáporné čísla, môžeme túto rovnosť odmocniť a máme

$$|z_1 z_2| = |z_1|.|z_2|.$$

Súčin čísel $z_1 = r_1(\cos\alpha + i\sin\alpha)$ a $z_2 = r_2(\cos\beta + i\sin\beta)$ môžeme upraviť ako

$$z_1 z_2 = r_1 r_2(\cos\alpha + i\sin\alpha)(\cos\beta + i\sin\beta) = r_1 r_2[(\cos\alpha\cos\beta - \sin\alpha\sin\beta) + i(\cos\alpha\sin\beta + \sin\alpha\cos\beta)].$$

Na základe goniometrických identít (známych zo strednej školy) vidíme, že

$$z_1 z_2 = r_1 r_2(\cos(\alpha+\beta) + i\sin(\alpha+\beta)).$$

Už vieme, že $|z_1 z_2| = r_1 r_2$. Teda číslo $z_1 z_2$ skutočne možno vyjadriť pomocou argumentu $\alpha + \beta$. □

Dôsledok C.2.4. *Ak $n \in \mathbb{N}$ a $z = r(\cos\alpha + i\sin\alpha)$, tak*

$$z^n = r^n\,(\cos(n\alpha) + i\sin(n\alpha))$$

Tento vzťah medzi násobením komplexných čísel a sčitovaním uhlom umožňuje elegantné odvodenie mnohých trigonometrických identít.

Príklad C.2.5. Umocnením čísla $\cos\alpha + i\sin\alpha$ na n-tú pre $n \in \mathbb{N}$ dostaneme

$$(\cos\alpha + i\sin\alpha)^n = \cos(n\alpha) + i\sin(n\alpha) = \sum_{j=0}^{n} \binom{n}{j} i^{n-j}\cos^j\alpha\sin^{n-j}\alpha$$

Keď teraz použijeme (C.4) a rozdelíme súčet na pravej strane na reálnu a imaginárnu časť, vidíme, že

$$\cos nx = \cos^n x - \binom{n}{2}\cos^{n-2} x\sin^2 x + \binom{n}{4}\cos^{n-4} x\sin^4 x - \dots$$
$$\sin nx = n\cos^{n-1} x\sin x - \binom{n}{3}\cos^{n-3} x\sin^3 x + \binom{n}{5}\cos^{n-5} x\sin^5 x - \dots$$

C.3 Riešenie rovníc v komplexných číslach

Dôležitá vlastnosť komplexných čísel je vyjadrená v nasledujúcej vete:

Veta C.3.1 (Základná veta algebry)**.** *Každý nenulový polynóm s komplexnými koeficientami má koreň v $\mathbb{C}$. T.j. ak*

$$f(x) = c_n x^n + \dots + c_1 x + c_0,$$

tak existuje $z \in \mathbb{C}$ také, že $f(z) = 0$.

Dokonca platí, že ak polynóm $f(x)$ je stupňa $n > 1$, tak počet koreňov vrátane násobnosti je práve n. (Ak $f(x) = (x-z)^k g(x)$ pre nejaký polynóm $g(x)$ a z nie je koreňom polynómu g, hovoríme, že *násobnosť* koreňa z je k. K polynómom a násobnosti koreňov sa ešte dostanete neskôr v rámci predmetu algebra.)

Nie všetky rovnice takéhoto tvaru však vieme jednoducho riešiť. Ukážeme si len dva typy rovníc, ktoré sa dajú riešiť vcelku ľahko. Najprv uvidíme, že pri kvadratických rovniciach s reálnymi koeficientmi môžeme postupovať podobne ako v reálnych číslach.

C.3.1 Kvadratické rovnice s reálnymi koeficientmi

Najprv si všimnime, že v komplexných číslach (na rozdiel od reálnych) vieme riešiť aj rovnicu $x^2 = r$, kde r je záporné reálne číslo.

Všimnime si, že každé záporné reálne číslo vieme zapísať v tvare $-a^2$ pre nejaké $a \in \mathbb{R}$. Teda vlastne riešime rovnicu

$$\begin{aligned} x^2 &= -a^2, \\ x^2 + a^2 &= 0, \\ (x - ia)(x + ia) &= 0, \end{aligned}$$

ktorej riešeniami sú práve $x = \pm ia$.

Predpokladajme teraz, že máme rovnicu tvaru

$$ax^2 + bx + c = 0,$$

kde $a, b, c \in \mathbb{R}$ a $a \neq 0$.

Zopakujeme presne ten istý postup, ktorým sa zvykne na strednej škole odvodzovať vzorec pre výpočet koreňov kvadratickej rovnice – použijeme doplnenie na štvorec.

$$\begin{aligned} ax^2 + bx + c &= 0 \\ a\left(x + \frac{b}{2a}\right)^2 + c - \frac{b^2}{4a} &= 0 \\ a\left(x + \frac{b}{2a}\right)^2 &= \frac{b^2 - 4ac}{4a} \\ \left(x + \frac{b}{2a}\right)^2 &= \frac{b^2 - 4ac}{4a^2} \end{aligned}$$

Označíme $D = b^2 - 4ac$. Ak $D > 0$, tak dostaneme

$$x_{1,2} = \frac{-b \pm \sqrt{D}}{2a}.$$

Ak $D = 0$, tak máme dvojnásobný koreň $x = -\frac{b}{2a}$. V prípade $D < 0$ máme rovnicu tvaru $z^2 = -(\sqrt{-D})^2$, pre $z = x + \frac{b}{2a}$, o ktorej už vieme, že jej riešeniami sú $z_{1,2} = \pm i\sqrt{-D}$. Z toho dostaneme

$$x_{1,2} = \frac{-b \pm i\sqrt{-D}}{2a}.$$

(Pretože $D < 0$, je $-D$ je kladné reálne číslo a výraz $\sqrt{-D}$ má zmysel.)

Príklad C.3.2. Riešme rovnicu $x^2 + 4x + 5 = 0$.

Dostaneme $D = 16 - 20 = -4$ a

$$x_{1,2} = \frac{-4 \pm 2i}{2} = -2 \pm i.$$

Mohli by sme rovnaký postup ako v odvodení vzorca pre korene kvadratickej rovnice použiť, keby boli koeficienty komplexné? V podstate áno – ale na mieste, kde sme použili odmocninu (inak povedané, riešili sme rovnicu $z^2 = \frac{D}{4a^2}$), zatiaľ nevieme, čo robiť v prípade, že D je komplexné číslo. Práve o niečom, čo sa dá nazvať „odmocninou“ z komplexného čísla, sa dozvieme o chvíľu.

x:SSECTBINOM}

C.3.2 Binomické rovnice

Rovnicu tvaru $x^2 = a$ vieme vyriešiť pre $a \in \mathbb{R}$ (či už kladné alebo záporné). Skúsme sa zamyslieť nad tým, čo by sa stalo, keby sme na pravej strane mali nejaké komplexné číslo.

Riešime teda rovnicu

$$x^n = z,$$

v obore komplexných čísel. Skúsme čísla vystupujúce v rovnici upraviť na goniometrický tvar. Nech teda $x = r(\cos\alpha + i\sin\alpha)$ a $z = |z|(\cos\varphi + i\sin\varphi)$. Potom dostaneme

$$r^n\left(\cos(n\alpha) + i\sin(n\alpha)\right) = |z|(\cos\varphi + i\sin\varphi)\,.$$

Predchádzajúca rovnosť je splnená práve vtedy, keď $r^n = |z|$, čiže

$$r = \sqrt[n]{|z|}$$

a $n\alpha = \varphi + 2k\pi$ čiže

$$\alpha = \frac{\varphi}{n} + \frac{2k}{n}\pi,$$

pre $k = 0, \ldots, n-1$. (Ten istý bod znamená rovnakú vzdialenosť od počiatku, uhol sa môže líšiť o 2π, lebo otočenie o násobok 2π je identické zobrazenie. Stačí použiť k od 0 po $n-1$, lebo potom sa už body začnú opakovať – uhly sa budú líšiť o násobok $\frac{2n}{n}\pi = 2\pi$.)

Príklad C.3.3. Riešme rovnicu $x^4 = 1+i$. Najprv prevedieme pravú stranu na goniometrický tvar: $x^4 = \sqrt{2}(\cos\frac{\pi}{4} + i\sin\frac{\pi}{4})$. Pre $x = r(\cos\varphi + i\sin\varphi)$ dostaneme

$$x^4 = r^4(\cos 4\varphi + i\sin 4\varphi),$$

čiže

$$r^4 = \sqrt{2} \qquad \Rightarrow \qquad r = \sqrt[8]{2}.$$

Pre uhol φ dostávame

$$4\varphi = \frac{\pi}{4} + 2k\pi \qquad \Rightarrow \qquad \varphi = \frac{\pi}{16} + k\frac{\pi}{2}.$$

Všetky riešenia danej rovnice sú teda $\sqrt[8]{2}\left(\cos(\frac{\pi}{16} + k\frac{\pi}{2}) + i\sin(\frac{\pi}{16} + k\frac{\pi}{2})\right)$ pre $k = 0, 1, 2, 3$.

Niekedy sa zvyknú všetky riešenia rovnice $x^n = z$ nazývať *n-tými odmocninami* komplexného čísla z. Môžeme si všimnúť, že vzorec

$$x = \frac{-b + \sqrt{D}}{2a}$$

bude platiť aj teraz ak symbol $\sqrt{D}$ chápeme v takom zmysle, že zaň možno dosadiť ktorúkoľvek z druhých odmocnín čísla D. (Teda ktorékoľvek komplexné čísla také, že $x^2 = D$.)

Príklad C.3.4. Vyriešme rovnicu $x^2 - (1+i)x - 2 - i = 0$.

Dostaneme $D = (1+i)^2 + 4(2+i) = 2i + 8 + 4i = 8 + 6i$. Diskriminant prevedieme na goniometrický tvar. Dostaneme

$$8 + 6i = 10(\cos\varphi + i\sin\varphi), \text{ kde } \cos\varphi = \frac{4}{5}, \sin\varphi = \frac{3}{5}.$$

Komplexné odmocniny z tohto čísla sú

$$u_{1,2} = \sqrt{10}\left(\cos\left(\frac{\varphi}{2} + k\pi\right) + i\sin\left(\frac{\varphi}{2} + k\pi\right)\right),$$

pre $k = 1, 2$. Pritom $|\cos\frac{\varphi}{2}| = \sqrt{\frac{1+\cos\varphi}{2}} = \sqrt{\frac{9}{10}} = \frac{3}{\sqrt{10}}$ a $|\sin\frac{\varphi}{2}| = \sqrt{\frac{1-\cos\varphi}{2}} = \sqrt{\frac{1}{10}} = \frac{1}{\sqrt{10}}$.

Z jednotkovej kružnice (na základe kvadrantu, v ktorom je φ) vieme určiť aj znamienko kosínu a sínu, čiže dostaneme

$$u_{1,2} = \pm(3 + i).$$

Z toho dostaneme riešenia kvadratickej rovnice ako

$$x_{1,2} = \frac{1 + i \pm (3 + i)}{2}$$
$$x_1 = 2 + i \qquad x_2 = -1$$

O správnosti riešenia sa môžeme presvedčiť dosadením alebo roznásobením $(x-2-i)(x+1) = x^2 - (1+i)x - 2 - i$.

Spôsoby riešenia niektorých ďalších typov rovníc (kubické, bikvadratické, reciproké) nájdete napríklad v [KGGS, Kapitola 6].

C.4 Zopár ďalších vecí súvisiacich s komplexnými číslami

{komplex:SECTDALSIE}

Spomenieme veľmi stručne niekoľko ďalších faktov o komplexných číslach.

Exponenciálny tvar komplexného čísla. Často sa stretnete so zápisom komplexného čísla v tvare

$$r(\cos\varphi + i\sin\varphi) = re^{i\varphi}$$

alebo

$$\cos\varphi + i\sin\varphi = i\varphi.$$

Bez toho, aby sme sa tým zaoberali hlbšie, tento zápis môžeme považovať jednoducho za skratku zápisu goniometrického zápisu. (Z Moivrovej vety vieme, že násobenie komplexných čísel s veľkosťou 1 funguje ako sčitovanie exponentov = sčitovanie uhlov.)

Kvaternióny. Podobným spôsobom ako komplexné čísla sa dajú vybudovať *kvaternióny.* V tomto prípade sa pridajú 3 nové prvky i, j, k, ktorých druhá mocnina je -1 a vhodne sa pre ne zadefinuje súčin. Kvaternióny tiež majú geometrický význam, ich násobenie súvisí s vektorovým súčinom. Na rozdiel od komplexných čísel však netvoria pole (násobenie nie je komutatívne.) Viac sa o nich môžete dočítať napríklad v [KGGS, Podkapitola 4.7].

Komplexné čísla sa nedajú usporiadať. Na reálnych číslach existuje relácia $\leq$, ktorá (okrem iných vlastností) spĺňa

(i) Lubovoľné dve reálne čísla sú porovnateľné, teda platí aspoň 1 z možností $x \leq y$ a $y \leq x$.

(ii) $x \leq y \wedge y \leq x \Rightarrow x = y$.

(iii) $x \leq y \Rightarrow x + z \leq y + z$.

(iv) $0 \leq z \wedge x \leq y \Rightarrow xz \leq yz$.

Na komplexných číslach sa nedá zadefinovať relácia $\leq$, ktorá by mala podobné vlastnosti. Z uvedených vlastností totiž pre každé x vieme odvodiť $x^2 \geq 0$. (Ak $x \geq 0$, tak túto rovnosť vynásobíme číslom x, ak $x \leq 0$, tak vynásobením číslom $-x$ dostaneme $-x^2 \leq 0$, z čoho vyplýva $0 \leq x^2$.)

Dostaneme teda, že $i^2 = -1 \geq 0$ a po pripočítaní 1 máme $0 \geq 1$. Súčasne však $(-1)^2 = 1 \geq 0$, teda $0 = 1$, čo je spor.

Vlastnosti, ktoré sme uviedli, sú niektoré z vlastností usporiadaných polí. O usporiadaných poliach sa viac môžete dočítať napríklad v [ŠHHK]. S pojmom relácia usporiadania ste sa pravdepodobne už stretli.

Cvičenia

Úloha C.4.1. Vypočítajte

a) $(3+2i)+(2-i)$	b) $(1+i)+(1-i)$	c) $(1+3i)+(\sqrt{3}+i)$
d) $(3+2i)(2-i)$	e) $(1+i)(1-i)$	f) $(1+\sqrt{3}i)(\sqrt{3}+i)$
e) $(3+2i)-(2-i)$	f) $(1+i)-(1-i)$	g) $(1+3i)-(\sqrt{3}+i)$
h) $(3+2i)/(2-i)$	i) $(1+i)/(1-i)$	j) $(1+\sqrt{3}i)/(\sqrt{3}+i)$

Úloha C.4.2. Overte výpočtom, že pri oboch uzátvorkovaniach výrazu $(1+2i)(1-i)(2-i)$ dostaneme ten istý výsledok.

{komplexcvic:DISTRIB}

Úloha C.4.3. Overte, že pre sčitovanie a násobenie komplexných čísel platí distributívnosť.

Úloha C.4.4. Overte, že pre komplexné čísla platí trojuholníková nerovnosť $|z_1 + z_2| \leq |z_1| + |z_2|$. Čo predstavuje táto nerovnosť geometricky?

Úloha C.4.5. Vieme, že na reálnej osi predstavujú riešenia nerovnice $|x - a| < r$ interval $(a - r, a + r)$ (pre $a, r \in \mathbb{R}$ a $r > 0$). Aký geometrický útvar v komplexnej rovine tvoria komplexné čísla vyhovujúce podmienke:
a) $|z - z_0| < r$,
a) $|z - z_0| = r$,
a) $|z - z_0| \leq r$,
kde z_0 je dané komplexné číslo a r je dané kladné reálne číslo?

Úloha C.4.6*. Ak $z_1, z_2 \in \mathbb{C}$ a $r \in \mathbb{R}$, $r > 0$, aký geometrický útvar tvoria body zodpovedajúce komplexným číslam s vlastnosťou $|z - z_1| + |z - z_2| = r$? Načrtnite ho pre $z_1 = 0$ a $z_2 = 3 + 2i$.

Úloha C.4.7. Nájdite goniometrický tvar daných komplexných čísel:
a) $1 - i$; b) $\sqrt{3} + i$; c) $-i$; d) $2 + i$; e) $(1 + i)(1 - i)$

Úloha C.4.8. Vyriešte rovnice:
a) $x^2-4x+13 = 0$ b) $4x^2+4x+2 = 0$ c) $x^2-6x+13 = 0$ d) $x^2+2x+50 = 0$ e) $x^2+x+1 = 0$

Úloha C.4.9. Vyriešte rovnice:
a) $z^2 = \frac{1-3i}{1+3i} - \frac{1}{5} + \frac{3}{5}i$; b) $z^6 = i$; c) $\frac{z^4}{8} + i\sqrt{3} = -1$; d) $z^4 = 1 + i$

Úloha C.4.10. Nájdite riešenia rovníc:
a) $x^2 + 2x + 3 = 0$; b) $x^4 + 5x^2 + 4 = 0$
a) $-1 \pm \sqrt{2}i$; b) $\pm i$, $\pm 2i$,

Úloha C.4.11. Vyriešte rovnice:
a) $x^2 - (1 + 2i)x - 3 + i = 0$ b)$x^2 - 2x + 1 - 2i = 0$ c) $x^2 - (4 + 3i)x + 1 + 5i = 0$ d) $x^2 - 3(1 + i)x + 5i = 0$ e) $x^2 + (1 + i)x - 4i = 0$

Úloha C.4.12. Riešte rovnice:
a) $z^3 - iz^2 + 4z - 4i = 0$ b) $x^4 + x^2 + 1 = 0$ c) $x^3 - (3 + 2i)x^2 + 2(1 + 3i)x - 4i = 0$ d) $x^3 - 2ix^2 - x + 2i = 0$

Úloha C.4.13. Riešte sústavy (môžete napr. použiť Gaussovu eliminačnú metódu, vyrátať inverznú maticu, použiť Cramerovo pravidlo):
$\left(\begin{smallmatrix} 1+i & 1-i & | & 1 \\ 1 & -1 & | & i \end{smallmatrix}\right)$ $\left(\begin{smallmatrix} 1+i & -i & | & 0 \\ i & -1 & | & 1 \end{smallmatrix}\right)$ $\left(\begin{smallmatrix} i & 1 & | & 0 \\ 1 & 1-i & | & 2i \end{smallmatrix}\right)$ $\left(\begin{smallmatrix} -1+i & 2-i & | & 1+i \\ -1+2i & 3-2i & | & 1-i \end{smallmatrix}\right)$

Úloha C.4.14. Nájdite všetky $x \in \mathbb{R}$, pre ktoré platí $\left(\frac{1+xi}{1-xi}\right)^6 = \frac{3+4i}{3-4i}$

Literatúra

[Ap] Tom M. Apostol. *Calculus II.* John Wiley and Sons, New York, 1969.

[Ax] Sheldon Axler. *Linear Algebra Done Right.* Springer-Verlag, New York, 2nd edition, 1997. Undergraduate Texts in Mathemtaics.

[AA] Titu Andreescu and Dorin Andrica. *Complex Numbers from A to ...Z.* Birkhäuser, Boston, 2006.

[AM] M. F. Atiyah and I. G. MacDonald. *Introduction to Commutative Algebra.* Adison-Wesley, Reading, 1969.

[Bl] Rudolf Blaško. Matematická analýza. `http://frcatel.fri.utc.sk/~beerb`.

[Bó] Miklós Bóna. *Combinatorics of Permutations.* CRC, Boca Raton, 2004.

[BL] Kurt Bryan and Tanya Leise. The $25,000,000,000 eigenvector – the linear algebra behind Google. *SIAM Review*, 48(3):569–581, 2006. `http://www.rose-hulman.edu/~bryan/googleFinalVersionFixed.pdf`.

[BŠ] Bohuslav Balcar and Petr Štěpánek. *Teorie množin.* Academia, Praha, 2001.

[Č] Juraj Činčura. Elementárna teória čísel. Poznámky k prednáške, `http://thales.doa.fmph.uniba.sk/sleziak/cvicenia/tc/`.

[CFR] Paul Cull, Mary Flahive, and Robby Robson. *Difference Equations - From Rabbits to Chaos.* Springer, New York, 2005. Undergraduate Texts in Mathematics.

[DF] David S. Dummit and Richard M. Foote. *Abstract Algebra.* John Willey and Sons, 3rd edition, 2004.

[G1] Jaroslav Guričan. Faktorizácia polynómov I. *Obzory matematiky, fyziky a informatiky.* `http://thales.doa.fmph.uniba.sk/katc/pages/member.php?clen=gurican`.

[G2] Jaroslav Guričan. Faktorizácia polynómov II. *Obzory matematiky, fyziky a informatiky.* `http://thales.doa.fmph.uniba.sk/katc/pages/member.php?clen=gurican`.

[G3] Jaroslav Guričan. Vybrané kapitoly z algebry. Poznámky k prednáške, `http://thales.doa.fmph.uniba.sk/katc/pages/member.php?clen=gurican`.

[GĎ] Milan Gera and Vladimír Ďurikovič. *Matematická analýza.* Alfa, Bratislava, 1990.

[GŠŠ] M. Greguš, M. Švec, and V. Šeda. *Obyčajné diferenciálne rovnice.*

[HJ] Roger A. Horn and Charles R. Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, 1985.

[HS] T. Hecht and Z. Sklenáriková. *Metódy riešenia matematických úloh*. SPN, Bratislava, 1992.

[HZK] Milan Hejný, Valent Zaťko, and Pavel Kršňák. *Geometria 1*. SPN, Bratislava, 1985.

[I] Ján Ivan. *Matematika 1*. Alfa, Bratislava, 1983.

[J] B. Johnson. Fibonacci numbers and matrices. `http://www.dur.ac.uk/bob.johnson/fibonacci/`.

[K] Július Korbaš. *Lineárna algebra a geometria I*. UK, Bratislava, 2003.

[KGGS] Tibor Katriňák, Martin Gavalec, Eva Gedeonová, and Jaroslav Smítal. *Algebra a teoretická aritmetika 1*. UK, Bratislava, 2002.

[KMŠ] Igor Kluvánek, Ladislav Mišík, and Marko Švec. *Matematika I*. Alfa, Bratislava, 4th edition, 1971.

[L] Loren C. Larson. *Metódy riešenia matematických problémov*. ALFA, Bratislava, 1990.

[LM] Amy N. Langville and Carl D. Meyer. *Google's PageRank and Beyond: The Science of Search Engine Rankings*. Princeton University Press, Princeton, 2006.

[Ma] Desmond MacHale. My favourite polynomial. *The Mathematical Gazette*, 75:157–165, 1991.

[Me] Carl D. Meyer. *Matrix Analysis and Applied Linear Algebra*. SIAM, 2000.

[Mi] Ondrej Mikuláš. PageRank algoritmus, 2010. Bakalárska práca, FMFI UK, Bratislava.

[NS] A. Naylor and G. Sell. *Teória lineárnych operátorov v technických a prírodných vedách (Linear Operator Theory in Engineering and Science)*. Alfa, Bratislava.

[OŠ] Daniel Olejár and Martin Škoviera. *Úvod do teórie diskrétnych matematických štruktúr*. Univerzita Komenského, Bratislava, 2007. `http://www.dcs.fmph.uniba.sk/texty/dsmain.pdf`.

[P] Murray H. Protter. *Basic Elements of Real Analysis*. Springer-Verlag, NY, 1998. Undergraduate Texts in Mathematics.

[Rog] Kenneth Rogers. The axioms for Euclidean domains. *Amer. Math. Monthly*, 78(10):1127–1128, 1971.

[Rot] Joseph J. Rotman. *An Introduction to the Theory of Groups*. Springer-Verlag, New York, 1995.

[Š] Tibor Šalát. *Reálne čísla*. Alfa, Bratislava, 1982.

[Si] John R. Silvester. Determinants of block matrices. *The Mathematical Gazette*, 84(501):460–467, 2000.

[Sl1] Martin Sleziak. 1-MAT-260 Algebra 2. Poznámky k prednáške, https://msleziak.com/vyuka/2011/alg2m/.

[Sl2] Martin Sleziak. 2-UMA-115 teória množín. Poznámky k prednáške, http://thales.doa.fmph.uniba.sk/sleziak/vyuka/.

[Sl3] Martin Sleziak. Teória čísel. Poznámky k prednáške, https://msleziak.com/vyuka.

[Sm] Jozef Smida. *Komplexné čísla, matematika pre 4. ročník gymnázií*. SPN, Bratislava, 1987.

[Š] Beáta Štupáková. Fibonacciho a Lucasove čísla, 2008. bakalárska práca, FMFI UK, Bratislava.

[ŠHHK] T. Šalát, A. Haviar, T. Hecht, and T. Katriňák. *Algebra a teoretická aritmetika 2*. Alfa, Bratislava, 1986.

[ŠS] Tibor Šalát and Jaroslav Smítal. *Teória množín*. UK, Bratislava, 1995.

[T] Jean-Pierre Tignol. *Galois' Theory of Algebraic Equations*. World Scientific, Singapore–New Jersey–London–Hong Kong, 2001.

[W] Michal Winczer. Diskrétna matematika. Poznámky k prednáške, http://edi.fmph.uniba.sk/~winczer/diskretna.html.

[Z1] Pavol Zlatoš. Lineárna algebra a geometria. http://thales.doa.fmph.uniba.sk/zlatos/.

[Z2] Pavol Zlatoš. *Ani matematika si nemôže byť istá sama sebou*. IRIS, Bratislava, 1995. http://thales.doa.fmph.uniba.sk/zlatos/animat/animat.pdf.

Register

Zoznam symbolov