Obsah

Úvod

Verzia: 15. apríla 2013

Predhovor

V rámci tohoto textu sa budeme občas odkazovať aj na veci z predchádzajúcich semestrov. Takéto odkazy budú označené napríklad ako veta I-3.2.6 alebo veta II-3.5.1. (Keďže ste absolvovali predmety Algebra 1,2 dá sa predpokladať, že budete vedieť o akú vetu ide. Toto je viac-menej pomôcka - ak si nebudete istí, môžete si tam znenie príslušenej vety, lemy, či definície skontrolovať.) Verzia poznámok z predchádzajúcich semestrov, na ktorú sa toto číslovanie vzťahuje, bude tiež na stránke predmetu.

Časti označené hviezdičkou sú nepovinné – doplnil som ich preto, že by Vás niektoré z nich mohli zaujímať. V cvičeniach hviezdička označuje náročnejšie cvičenia a + označuje nepovinné cvičenia (napríklad tie, ktoré sa týkajú nepovinných častí).

Sylaby a literatúra

Sylaby predmetu: 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.

Literatúra: Základnou literatúrou pre tento kurz je [KGGS]; objavia sa však aj témy, ktoré v tejto knihe spracované nie sú, k nim sem časom doplním vhodnú literatúru.

Cvičenia v texte som vyberal z kníh [KGGS, BM, FS, K, Pro1].

Označenia

Ako obvykle, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ označuje množinu celých, racionálnych, reálnych a komplexných čísel, $\mathbb{N} = \{1, 2, 3, \dots\}$ je množina prirodzených čísel a $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$.

Označenie A^B používame pre množinu zobrazení z B do A. Špeciálne v niektorých príkladoch sa vyskytuje množina $\mathbb{R}^\mathbb{R}$ a $\mathbb{R}^\mathbb{N}$, na oboch týchto množinách sa dajú prirodzeným spôsobom zaviesť sčitovanie a násobenie tak, že tieto množiny tvoria vektorové priestory.

Kapitola 1

Euklidovské vektorové priestory

Táto kapitola je spracovaná prevažne na základe [KGGS, 1.16,1.17].

1.1 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 1.1.1. 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

(i) $g(\vec{\alpha}, \vec{\beta}) = g(\vec{\beta}, \vec{\alpha})$,

(ii) $g(\vec{\alpha} + \vec{\beta}, \vec{\gamma}) = g(\vec{\alpha}, \vec{\gamma}) + g(\vec{\beta}, \vec{\gamma})$,

(iii) $g(c\vec{\alpha}, \vec{\beta}) = cg(\vec{\alpha}, \vec{\beta})$,

(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 2.3. 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:

(i) $\langle\vec{\alpha},\vec{\beta}\rangle = \langle\vec{\beta},\vec{\alpha}\rangle$,

(ii) $\langle\vec{\alpha}+\vec{\beta},\vec{\gamma}\rangle = \langle\vec{\alpha},\vec{\gamma}\rangle + \langle\vec{\beta},\vec{\gamma}\rangle$,

(iii) $\langle c\vec{\alpha},\vec{\beta}\rangle = c\langle\vec{\alpha},\vec{\beta}\rangle$,

(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 1.1.2. 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.

Príklad 1.1.3. Zoberme si vektorový priestor $\mathbb{R}^n$ s obvyklým sčitovaním a skalárnym násobkom (po zložkách). Potom pre vektory $\vec{\alpha}=(a_1,\dots,a_n)$ a $\vec{\beta}=(b_1,\dots,b_n)$ 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 1.1.4. Definujme na $\mathbb{R}^2$ skalárny súčin nasledovne:

$$\langle\vec{\alpha},\vec{\beta}\rangle = a_1b_1 + a_1b_2 + a_2b_1 + 2a_2b_2,$$

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

$$\langle\vec{\alpha},\vec{\alpha}\rangle = a_1^2 + 2a_1a_2 + 2a_2^2 = (a_1 + a_2)^2 + a_2^2,$$

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 $\vec{\alpha} = \vec{0}$.

Príklad 1.1.5. 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 matica typu $n \times n$. Takýto predpis priradí 2 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.

Príklad 1.1.6. 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 1.1.7. 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]).

Tvrdenie 1.1.8. *Nech V je euklidovský vektorový priestor. Pre ľubovoľné $\vec{\alpha}, \vec{\beta} \in V$ a $c \in \mathbb{R}$ platí:*

(i) $|\vec{\alpha}| \geq 0$

(ii) $|\vec{\alpha}| = 0 \Leftrightarrow \vec{\alpha} = \vec{0}$

(iii) $|c\vec{\alpha}| = |c||\vec{\alpha}|$

(iv) $|\langle \vec{\alpha}, \vec{\beta} \rangle| \leq |\vec{\alpha}||\vec{\beta}|$ *(Schwarzova nerovnosť)*

(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 $\vec{\beta}$.*

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

$$\begin{gathered} \langle \vec{\alpha}, \vec{\beta} \rangle^2 \leq |\vec{\alpha}|^2|\vec{\beta}|^2 \\ |\langle \vec{\alpha}, \vec{\beta} \rangle| \leq |\vec{\alpha}||\vec{\beta}| \end{gathered}$$

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 $\vec{\beta} \in \mathbb{R}$. Ľahko sa overí, že k rovnosti dôjde iba v prípade, že $k \geq 0$. (Stačí si uvedomiť, že $|\vec{\alpha} + k\vec{\alpha}| = |1+k|.|\vec{\alpha}|$ a $|\vec{\alpha}| + |k\vec{\alpha}| = (1+|k|).|\vec{\alpha}|$.) □

Schwarzova nerovnosť pre priestor $\mathbb{R}^n$ s obvyklým skalárnym súčinom sa často používa pri dôkaz rôznych nerovností.

$$\left|\sum_{k=1}^{n} x_k y_k\right| \leq \sqrt{\sum_{k=1}^{n} x_k^2 \sum_{k=1}^{n} y_k^2} \tag{1.1}$$

Definícia 1.1.9. 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 1.1.10. Vektory $\vec{\alpha}, \vec{\beta} \in V$ nazveme *kolmé* (*ortogonálne*), ak $\langle\vec{\alpha}, \vec{\beta}\rangle = 0$.

O k-tici vektorov $\vec{\alpha}_1, \dots, \vec{\alpha}_k$ 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 $i \neq j$.

Tvrdenie 1.1.11. *Nech V je euklidovský vektorový priestor. Ak nenulové vektory $\vec{\alpha}_1, \dots, \vec{\alpha}_k$ sú ortogonálne, tak sú lineárne nezávislé.*

Dôkaz. Nech $c_1, \dots, c_k \in \mathbb{R}$ sú také, že $c_1\vec{\alpha}_1 + \dots + c_k\vec{\alpha}_k = \vec{0}$. Zoberme ľubovoľné $i \in \{1, 2, \dots, k\}$. Potom dostaneme

$$0 = \langle\vec{\alpha}_i, \vec{0}\rangle = \langle\vec{\alpha}_i, c_1\vec{\alpha}_1 + \dots + c_k\vec{\alpha}_k\rangle = c_i|\vec{\alpha}_i|^2.$$

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 $i = 1, 2, \dots, k$ dostaneme $c_1 = c_2 = \ldots = c_k = 0$. Teda dané vektory sú lineárne nezávislé. □

Definícia 1.1.12. 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 1.1.13. *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 1.1.14. *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 1.1.15. *Nech V je euklidovský priestor a $\vec{\alpha}_1, \dots, \vec{\alpha}_k \in V$. Nech $S = [\vec{\alpha}_1, \dots, \vec{\alpha}_k]$ je podpriestor vygenerovaný týmito vektormi. Potom $S^{\perp} = \{\vec{\alpha}_1, \dots, \vec{\alpha}_k\}^{\perp}$.*

Dôkaz. Z toho, že $\{\vec{\alpha}_1, \dots, \vec{\alpha}_k\} \subseteq S$ vyplýva inklúzia $S^{\perp} \subseteq \{\vec{\alpha}_1, \dots, \vec{\alpha}_k\}^{\perp}$.

Naopak, nech $\vec{\beta} \in \{\vec{\alpha}_1, \dots, \vec{\alpha}_k\}^{\perp}$. To znamená, že $\langle \vec{\beta}, \vec{\alpha}_i \rangle = 0$ pre $i = 1, 2, \dots, k$. Potom $\langle \vec{\beta}, \vec{\alpha} \rangle = 0$ aj pre ľubovoľný vektor $\vec{\alpha} \in [\vec{\alpha}_1, \dots, \vec{\alpha}_k]$, pretože každý vektor z tohoto podpriestoru má tvar $\alpha = c_1\vec{\alpha}_1 + \dots + c_k\vec{\alpha}_k$ a

$$\langle \vec{\beta}, c_1\vec{\alpha}_1 + \dots + c_k\vec{\alpha}_k \rangle = c_1\langle \vec{\beta}, \vec{\alpha}_1 \rangle + \ldots + c_k\langle \vec{\beta}, \vec{\alpha}_k \rangle = 0.$$

□

Tvrdenie 1.1.16. *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

$$S^{\perp} \cap T^{\perp} \subseteq (S + T)^{\perp}.$$

□

1.2 Gram-Schmidtov ortogonalizačný proces

Definícia 1.2.1. Vektory $\vec{\alpha}_1, \dots, \vec{\alpha}_n$ 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 1.1.11 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 1.2.2. Ak vektory $\vec{\alpha}_1, \dots, \vec{\alpha}_n$ sú ortonormálne a tvoria bázu vektorového priestoru V, tak túto bázu nazývame *ortonormálna báza.*

Príklad 1.2.3. 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.

Veta 1.2.4. *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

$$\begin{aligned}
0 = \langle\vec{\gamma}_{k+1}, \vec{\gamma}_1\rangle =& \langle\vec{\alpha}_{k+1}, \vec{\gamma}_1\rangle + c_{k+1,1}\langle\vec{\gamma}_1, \vec{\gamma}_1\rangle \\
0 = \langle\vec{\gamma}_{k+1}, \vec{\gamma}_2\rangle =& \langle\vec{\alpha}_{k+1}, \vec{\gamma}_2\rangle + c_{k+1,2}\langle\vec{\gamma}_2, \vec{\gamma}_2\rangle \\
0 = \langle\vec{\gamma}_{k+1}, \vec{\gamma}_3\rangle =& \langle\vec{\alpha}_{k+1}, \vec{\gamma}_3\rangle + c_{k+1,3}\langle\vec{\gamma}_3, \vec{\gamma}_3\rangle \\
&\vdots \\
0 = \langle\vec{\gamma}_{k+1}, \vec{\gamma}_k\rangle =& \langle\vec{\alpha}_{k+1}, \vec{\gamma}_k\rangle + c_{k+1,k}\langle\vec{\gamma}_k, \vec{\gamma}_k\rangle
\end{aligned} \tag{1.2}$$

(V každej rovnici sme vynechali všetky členy obsahujúce $\langle\vec{\gamma}_i, \vec{\gamma}_j\rangle$ pre $i \neq j$, $i, j \in \{1, 2, \dots, k\}$, 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 $i = 1, 2, \dots, k$.

Z rovníc (1.2) vidno, že pre takéto hodnoty $c_{k+1,i}$ bude vektor $\vec{\gamma}_{k+1}$ skutočne kolmý na všetky predchádzajúce vektory.

Ďalej vieme, že $\vec{\alpha}_{k+1} \notin [\vec{\alpha}_1, \dots, \vec{\alpha}_k] = [\vec{\gamma}_1, \dots, \vec{\gamma}_k]$ (lebo vektory $\vec{\alpha}_1, \dots, \vec{\alpha}_n$ sú lineárne nezávislé). Teda aj $\vec{\alpha}_{k+1}, \vec{\gamma}_1, \dots, \vec{\gamma}_k$ sú lineárne nezávislé, čiže ich lineárnou kombináciou nemôžeme dostať $\vec{0}$. Pretože $\vec{\gamma}_{k+1} = \vec{\alpha}_{k+1} + c_{k+1,1}\vec{\gamma}_1 + c_{k+1,2}\vec{\gamma}_2 + \dots + c_{k+1,k}\vec{\gamma}_k$ 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 $\vec{\gamma}_{k+1} \in [\vec{\alpha}_{k+1}, \vec{\gamma}_1, \dots, \vec{\gamma}_k] = [\vec{\alpha}_1, \dots, \vec{\alpha}_{k+1}]$. Teda $[\vec{\gamma}_1, \dots, \vec{\gamma}_{k+1}] \subseteq [\vec{\alpha}_1, \dots, \vec{\alpha}_{k+1}]$.

Ďalej $\vec{\alpha}_{k+1} = \vec{\gamma}_{k+1} - (c_{k+1,1}\vec{\gamma}_1 + c_{k+1,2}\vec{\gamma}_2 + \dots + c_{k+1,k}\vec{\gamma}_k)$ je lineárna kombinácia vektorov $\vec{\gamma}_1, \dots, \vec{\gamma}_{k+1}$, čiže platí aj obrátená inklúzia $[\vec{\alpha}_1, \dots, \vec{\alpha}_{k+1}] \subseteq [\vec{\gamma}_1, \dots, \vec{\gamma}_{k+1}]$.

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 1.2.5. 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

$$\begin{aligned}
\langle\vec{\gamma}_3, \vec{\gamma}_1\rangle &= \langle\vec{\alpha}_3, \vec{\gamma}_1\rangle + d\langle\vec{\gamma}_1, \vec{\gamma}_1\rangle = 0 \\
\langle\vec{\gamma}_3, \vec{\gamma}_2\rangle &= \langle\vec{\alpha}_3, \vec{\gamma}_2\rangle + e\langle\vec{\gamma}_2, \vec{\gamma}_2\rangle = 0
\end{aligned}$$

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.

$$\vec{\beta}_1 = \frac{\vec{\gamma}_1}{|\vec{\gamma}_1|} = \frac{1}{\sqrt{2}}(1,0,0,-1)$$
$$\vec{\beta}_2 = \frac{\vec{\gamma}_2}{|\vec{\gamma}_2|} = \sqrt{\frac{2}{3}}\left(\frac{1}{2}, 1, 0, \frac{1}{2}\right)$$
$$\vec{\beta}_3 = \frac{\vec{\gamma}_3}{|\vec{\gamma}_3|} = \frac{\sqrt{3}}{2}\left(\frac{1}{3}, -\frac{1}{3}, 1, \frac{1}{3}\right)$$

Existenciu ortogonálnej bázy môžeme použiť na dôkaz niektorých ďalších faktov o ortogonálnom doplnku.

Veta 1.2.6. *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 1.2.4 vytvoríme ortonormálnu bázu, tak bázové vektory patriace do S sa nezmenia, pretože boli ortonormálne už pred ortonormalizáciou.)

Nech teda vektory $\vec{\gamma}_1, \dots, \vec{\gamma}_k$ tvoria ortonormálnu bázu S a vektory $\vec{\gamma}_{k+1}, \dots, \vec{\gamma}_n$ sú ostatné vektory ortonormálnej bázy V. Ľubovoľný vektor $\vec{\gamma}$ sa dá jednoznačne zapísať ako

$$\vec{\gamma} = c_1\vec{\alpha}_1 + \dots + c_k\vec{\alpha}_k + c_{k+1}\vec{\alpha}_{k+1} + \dots + c_n\vec{\alpha}_n.$$

Ak zvolíme $\vec{\alpha} = c_1\vec{\alpha}_1 + \dots + c_k\vec{\alpha}_k$ a $\vec{\beta} = c_{k+1}\vec{\alpha}_{k+1} + \dots + c_n\vec{\alpha}_n$, 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 1.2.7. 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 1.2.11).

Dôsledok 1.2.8. *Nech S, T sú podpriestory konečnorozmerného priestoru V. Potom:*

(i) $V = S \oplus S^\perp$

(ii) $(S^\perp)^\perp = S$

(iii) $(S \cap T)^\perp = S^\perp + T^\perp$.

Dôkaz. (i) Vyplýva z predchádzajúcej vety a z vety I-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 I-4.4.18).

(iii) Použitím tvrdenia 1.1.16 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* 1.2.9. 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 (1.1). (Stačí v nej zobrať limitu pre $n \to \infty$.)

Overenie jednotlivých vlastností skalárneho súčinu je len o niečo zložitejšie ako v príklade 1.1.3.

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

$$\sum_{n=1}^{\infty}(x_n+y_n)^2 = \sum_{n=1}^{\infty}(x_n^2+2x_ny_n+y_n^2) \overset{(*)}{=} \sum_{n=1}^{\infty} x_n^2 + 2\sum_{n=1}^{\infty} x_ny_n + \sum_{n=1}^{\infty} y_n^2. \tag{1.3}$$

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

$$\sum_{n=1}^{\infty}|x_ny_n| \le \sqrt{\sum_{n=1}^{\infty} x_n^2 . \sum_{n=1}^{\infty} y_n^2}.$$

Túto nerovnosť môžeme dostať napríklad limitným prechodom z (1.1) (z (1.1) vieme, že uvedená nerovnosť platí pre všetky čiastočné súčty radov, ktoré v nej vystupujú).

Vidíme teda, že rady vystupujúce na pravej strane rovnosti (1.3) sú absolútne konvergentné, čím mám dokázanú platnosť tejto rovnosti.

Navyše, keď si ešte uvedomíme, že platí

$$\sum_{n=1}^{\infty} x_ny_n \le \sum_{n=1}^{\infty}|x_ny_n|,$$

tak vidíme, že všetky rady na pravej strane (1.3) 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

Úloha 1.2.1. 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 1.2.2. 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 1.2.3. Zistite, či $\sin \pi x$ a $\cos \pi x$ sú kolmé v priestore $C(0,1)$ so skalárnym súčinom z príkladu 1.1.6. Akú majú tieto vektory veľkosť?

Úloha 1.2.4. 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 1.2.5. 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 1.2.6*. 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 1.1.8 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 1.2.7. Ukážte, že funkcia $|\cdot|\colon \mathbb{R}^n \to \mathbb{R}$, $|(x_1,\dots,x_n)| = \max\{|x_i|; i = 1,\dots,n\}$ spĺňa podmienky (i), (ii), (iii) a (v) z tvrdenia 1.1.8, ale neexistuje skalárny súčin na $\mathbb{R}^n$ taký, že $|\alpha| = \sqrt{\langle\vec{\alpha},\vec{\alpha}\rangle}$ (pre všetky $\vec{\alpha} \in \mathbb{R}^n$).

Úloha 1.2.8. 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 1.2.9. 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 1.2.10. Nájdite ortonormálnu bázu pre priestory z úlohy 1.2.1.

Úloha 1.2.11. 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 $\mathrm{Ker}\, P = S^\perp$.

Úloha 1.2.12. Nájdite maticu ortogonálnej projekcie pri obvyklom skalárnom súčine pre:
a) priestory z úlohy 1.2.1;
b) pre ľubovoľný podpriestor $S = [\vec{\alpha}]$, pričom vektor $\vec{\alpha}$ je normovaný (má jednotkovú dĺžku);
c) pre podpriestor $S = [\vec{\alpha}_1,\dots,\vec{\alpha}_k]$, pričom vektory $\vec{\alpha}_1,\dots,\vec{\alpha}_k$ sú ortonormálne.
[Odpovede: b) $\vec{\alpha}^T\vec{\alpha}$; c) A^TA, kde A je matica, ktorej riadky tvoria vektory $\vec{\alpha}_1,\dots,\vec{\alpha}_k$.]

Úloha 1.2.13. 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 1.2.8 dokázali bez použitia predpokladu o konečnorozmernosti. Túto inklúziu použite raz pre S a raz pre $S^\perp$.)

Kapitola 2

Kvadratické formy

Táto kapitola je spracovaná prevažne na základe [KGGS, Kapitola 9].

2.1 Definícia a základné vlastnosti

Definícia 2.1.1. Výraz (polynóm)

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_i x_j,$$

kde $a_{ij} \in \mathbb{R}$ a $x_1, \dots, x_n$ sú (komutujúce) premenné budeme nazývať *kvadratická forma* v premenných $x_1, \dots, x_n$.

V súvislosti s touto definíciou je užitočné ozrejmiť si zopár faktov.

Poznámka 2.1.2.

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 $x_i x_j = x_j x_i$ – 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 polynómy n premenných špeciálneho tvaru. (Takéto polynómy viacerých premenných, ktoré majú všetky členy rovnakého stupňa, sa nazývajú *homogénne polynómy.*)

Príklad 2.1.3. 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 $\vec{\alpha} = (x_1, \dots, x_n)$, 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.

Veta 2.1.4. *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. □

2.2 Kanonický tvar kvadratickej formy

Pokúsme sa upraviť kvadratickú formu z príkladu 2.1.3 na iný tvar, ktorý môže byť na niektoré účely vhodnejší. Upravíme ju pomocou doplnenia na štvorec.

Príklad 2.2.1.

$$x_1^2 + 2x_1x_2 + 2x_2^2 + 4x_1x_3 + 2x_2x_3 + x_3^2 = (x_1 + x_2 + 2x_3)^2 + x_2^2 - 2x_2x_3 - 3x_3^2 =$$
$$(x_1 + x_2 + 2x_3)^2 + (x_2 - x_3)^2 - 4x_3^2 = (x_1 + x_2 + 2x_3)^2 + (x_2 - x_3)^2 - (2x_3)^2$$

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 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

$$\vec{\alpha} A \vec{\alpha}^T = \vec{\beta} P^{-1} A (P^{-1})^T \vec{\beta}.$$

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 2.1.4 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

$$\begin{aligned} D &= P^{-1} A P^{-1^T} \\ A &= P D P^T \end{aligned}$$

Definícia 2.2.2. 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 2.2.1).

Poznámka 2.2.3. 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.

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.

Príklad 2.2.4. Uvažujme kvadratickú formu $x_1 x_2$. Všimnime si, že

$$x_1 x_2 = \frac{(x_1 + x_2)^2}{4} - \frac{(x_1 - x_2)^2}{4} = \left(\frac{x_1}{2} + \frac{x_2}{2}\right)^2 - \left(\frac{x_1}{2} - \frac{x_2}{2}\right)^2.$$

To znamená, že túto kvadratickú formu vieme previesť na tvar $y_1^2 - y_2^2$ pomocou transformácie

$$\begin{aligned} y_1 &= \frac{x_1 + x_2}{2} \\ y_2 &= \frac{x_1 - x_2}{2} \end{aligned}$$

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 $A = \begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, $P = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ opäť platí $A = PBP^T$. Alebo tiež obrátene, $B = QAQ^T$, kde $Q = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$. (Tieto matice sme dostali z vyjadrenia transformácií premenných podobným spôsobom ako v predchádzajúcom príklade.)

Veta 2.2.5. *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 $\sum\limits_{k=1}^n d_k y_k^2$ budeme nazývať kanonický tvar *kvadratickej formy $\sum\limits_{i=1}^n \sum\limits_{j=1}^n a_{ij}x_ix_j$.*

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*

$$c_1y_1^2 + c_2y_2^2 + \dots + c_ny_n^2.$$

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 $\sum\limits_{i=1}^n \sum\limits_{j=1}^n a_{ij}x_ix_j$. 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.

$$\sum_{i=1}^n \sum_{j=1}^n a_{ij}x_ix_j = a_{11}x_1^2 + 2a_{12}x_1x_2 + \dots + 2a_{1n}x_1x_n + \sum_{i=2}^n \sum_{j=2}^n a_{ij}x_ix_j =$$
$$a_{11}\left(x_1^2 + 2\frac{a_{12}}{a_{11}}x_1x_2 + \dots + 2\frac{a_{1n}}{a_{11}}x_1x_n\right) + \sum_{i=2}^n \sum_{j=2}^n a_{ij}x_ix_j =$$
$$a_{11}\left(x_1 + \frac{a_{12}}{a_{11}}x_2 + \dots + \frac{a_{1n}}{a_{11}}x_n\right)^2 - \sum_{i=2}^n \frac{a_{1i}^2}{a_{11}}x_i^2 - \sum_{i=2}^n \sum_{j=i+1}^n \frac{a_{1i}a_{1j}}{a_{11}}x_ix_j + \sum_{i=2}^n \sum_{j=2}^n a_{ij}x_ix_j$$

Ak označíme $y_1 = x_1 + \frac{a_{12}}{a_{11}}x_2 + \dots + \frac{a_{1n}}{a_{11}}x_n$, podarilo sa nám upraviť pôvodnú kvadratickú formu na tvar

$$a_{11}y_1^2 + B(x_2, \dots, x_n),$$

kde $B(x_2, \dots, x_n)$ je kvadratická forma v $n-1$ premenných $x_2, \dots, x_n$.

Podľa indukčného predpokladu sa dá táto kvadratická forma previesť regulárnou transformáciou premenných na diagonálny tvar $c_2y_2^2 + \dots + c_nx_n^2$. Kombináciou týchto 2 transformácií prevedieme pôvodnú kvadratickú formu na

$$a_{11}y_1^2 + c_2y_2^2 + \dots + c_nx_n^2.$$

Ak ako P' označíme maticu transformácie pre kvadratickú formu $B(x_2, \dots, x_n)$, tak matica transformácie pôvodnej kvadratickej formy je

$$\begin{pmatrix} 1 & 0 & \dots & 0 \\ \frac{a_{21}}{a_{11}} & & & \\ \vdots & & P' & \\ \frac{a_{n1}}{a_{11}} & & & \end{pmatrix}$$

Ak urobíme Laplaceov rozvoj podľa prvého riadku, dostávame $|P| = |P'|$, čiže matica P je tiež regulárna.

Zostáva nám vyriešiť prípad, že $a_{11} \neq 0$. 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 2.2.4. 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 $0x_1^2 + \dots + 0x_n^2$. $\square$

Z toho, čo sme si ukázali v príklade 2.2.1 vyplýva, že sme súčasne dokázali nasledujúce tvrdenie o symetrických reálnych maticiach.

Dôsledok 2.2.6. *Každá reálna symetrická matica typu $n \times n$ je kongruentná s nejakou diagonálnou maticou* $\operatorname{diag}(d_1, \dots, d_n)$ *takou, že* $d_i \in \{0, \pm 1\}$ *pre* $i = 1 \dots n$.

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 I-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 $\begin{pmatrix} 1 & c & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ 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 2.2.7. Postup, ktorý sme si práve vysvetlili, si ukážeme na matici kvadratickej formy z príkladu 2.2.1.

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

$A = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 2 & 1 \\ 2 & 1 & 1 \end{pmatrix} \overset{(1)}{\sim} \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & -1 \\ 2 & 1 & 1 \end{pmatrix} \overset{(1')}{\sim} \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & -1 \\ 2 & -1 & 1 \end{pmatrix} \overset{(2)}{\sim} \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & -1 \\ 0 & -1 & -3 \end{pmatrix} \overset{(2')}{\sim} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & -3 \end{pmatrix} \overset{(3)}{\sim} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & -4 \end{pmatrix} \overset{(3')}{\sim}$
$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -4 \end{pmatrix} \overset{(4)}{\sim} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \overset{(4')}{\sim} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} = D$

(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) `4r*=1/2`

$I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \overset{(1)}{\sim} \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \overset{(2)}{\sim} \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix} \overset{(3)}{\sim} \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -3 & 1 & 1 \end{pmatrix} \overset{(4)}{\sim} \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -\frac{3}{2} & \frac{1}{2} & \frac{1}{2} \end{pmatrix} =: Q$

Priamym výpočtom môžeme overiť, že skutočne platí $D = QAQ^T$. Tiež si môžeme všimnúť, že pre maticu P, ktorú sme dostali v príklade 2.2.1 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 2.2.8. 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

Úloha 2.2.1. Overte, že relácia $A \sim B$, t.j. kongruencia symetrických matíc, je relácia ekvivalencie na množine reálnych symetrických matíc typu $n \times n$.

Úloha 2.2.2. 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) $x_1^2 + x_1x_2 + x_3x_4$

Úloha 2.2.3*. Preveďte kvadratickú formu $\sum\limits_{i=1}^{n} x_i^2 + \sum\limits_{1\le i<k\le n} x_ix_k$ na diagonálny tvar.

[Výsledok: $y_1^2 + \frac{3}{4}y_2^2 + \frac{4}{6}y_3^2 + \dots + \frac{n+1}{2n}y_n^2$; $P = \begin{pmatrix} 1 & 0 & \dots & \dots & \dots & 0 \\ \frac12 & 1 & 0 & \dots & \dots & 0 \\ \frac12 & \frac13 & 1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \frac12 & \frac13 & \frac14 & \dots & \frac1n & 1 \end{pmatrix}$]

Úloha 2.2.4*. Preveďte kvadratickú formu $\sum\limits_{1\le i<k\le n} x_ix_k$ na diagonálny tvar.

2.3 Zákon zotrvačnosti

Keďže sme tvar kvadratickej formy z vety 2.2.5 nazvali kanonický, dá sa očakávať, že bude v nejakom zmysle jednoznačný. Cieľom tejto kapitoly je práve sformulovať a dokázať túto jednoznačnosť.

Veta 2.3.1. *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

$$y_1^2 + \dots + y_k^2 - y_{k+1}^2 - \dots - y_s^2 \tag{2.1}$$

a súčasne (inou regulárnou transformáciou) na tvar

$$z_1^2 + \dots + z_r^2 - z_{r+1}^2 - \dots - z_t^2. \tag{2.2}$$

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 (2.1) a t je hodnosť matice pre kvadratickú formu (2.2). 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

$$z_1^2 + \cdots + z_r^2 + y_{k+1}^2 + \cdots + y_s^2 = y_1^2 + \cdots + y_k^2 + z_{r+1}^2 + \cdots + z_t^2. \tag{2.3}$$

Predpokladajme najprv, že $r < k$. Ukážeme, že za tohoto predpokladu sa dá nájsť nenulový vektor $(x_1, \dots, x_n)$ tak, aby platilo

$$z_1 = \cdots = z_r = y_{k+1} = \cdots = y_n = 0. \tag{2.4}$$

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 rovnice (2.4) 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 (2.4), na základe (2.3) 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 2.3.2. 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 1.1.5).

Veta 2.3.3. *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:

Tvrdenie 2.3.4. *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 I-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 2.3.5. *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 2.3.4 možno príslušnú kvadratickú formu upraviť na diagonálny tvar, v ktorom sú všetky členy kladné. Preto je táto matica kladne definitná. □

Dôsledok 2.3.6. *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 2.3.7. 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], [A], [Pro2] (a v podstate v každej učebnici, ktorá sa zaoberá analýzou viac premenných).

Príklad 2.3.8. V dôkaze tvrdenia 2.3.4 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 2.2.3*. Zodpovedajúca symetrická matica je

$$\begin{pmatrix} 1 & \frac12 & \frac12 & \dots & \frac12 \\ \frac12 & 1 & \frac12 & \dots & \frac12 \\ \dots & \dots & \dots & \dots & \dots \\ \frac12 & \frac12 & \dots & 1 & \frac12 \\ \frac12 & \frac12 & \dots & \frac12 & 1 \end{pmatrix}$$

Ak počítame jej hlavné minory dostávame $D_1 = 1$ a pre $k > 1$

$$D_k = \begin{vmatrix} 1 & \frac12 & \frac12 & \dots & \frac12 \\ \frac12 & 1 & \frac12 & \dots & \frac12 \\ \dots & \dots & \dots & \dots & \dots \\ \frac12 & \frac12 & \dots & 1 & \frac12 \\ \frac12 & \frac12 & \dots & \frac12 & 1 \end{vmatrix} = \begin{vmatrix} \frac{k+1}{2} & \frac{k+1}{2} & \frac{k+1}{2} & \dots & \frac{k+1}{2} \\ \frac12 & 1 & \frac12 & \dots & \frac12 \\ \dots & \dots & \dots & \dots & \dots \\ \frac12 & \frac12 & \dots & 1 & \frac12 \\ \frac12 & \frac12 & \dots & \frac12 & 1 \end{vmatrix} = (k{+}1) \begin{vmatrix} \frac12 & \frac12 & \frac12 & \dots & \frac12 \\ \frac12 & 1 & \frac12 & \dots & \frac12 \\ \dots & \dots & \dots & \dots & \dots \\ \frac12 & \frac12 & \dots & 1 & \frac12 \\ \frac12 & \frac12 & \dots & \frac12 & 1 \end{vmatrix} = (k{+}1) \begin{vmatrix} \frac12 & \frac12 & \frac12 & \dots & \frac12 \\ 0 & \frac12 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \frac12 & 0 \\ 0 & 0 & \dots & 0 & \frac12 \end{vmatrix} = \frac{k+1}{2^k}$$

(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 2.2.3*.

Cvičenia

Úloha 2.3.1. Pre danú kvadratickú formu určte tie hodnoty parametra t, 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) $\frac12 x_1^2 + 2x_2^2 - 3tx_3^2 + 2x_1x_2 + 2tx_2x_3 + 2x_1x_3$
d) $(x_1^2 + x_2^2 + x_3^2 + 2x_1x_3 - 2x_2x_3) + t(6x_1x_2 - 2x_1x_3 - 2x_1^2) + t^2(x_1^2 + x_2^2)$

Úloha 2.3.2. 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 2.3.4.) Dokážte, že potom $a_{nn} > 0$.

Úloha 2.3.3. Nech V je euklidovský vektorový priestor a $\vec{\alpha}_1, \dots, \vec{\alpha}_n \in V$. 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 2.3.4*. Pre kvadratické formy $f = \sum\limits_{i,j} a_{ij}x_ix_j$ a $g = \sum\limits_{i,j} b_{ij}x_ix_j$ definujeme kvadratickú formu $(f, g) = \sum\limits_{i,j} a_{ij}b_{ij}x_ix_j$. Ukážte, že ak f a g sú kladne definitné, tak aj (f, g) je kladne definitná.

Kapitola 3

Podobnosť matíc

Úvodná časť tejto kapitoly je spracovaná na základe [KGGS, 9.4,9.5] a .... Podkapitola 3.2 obsahuje poznámky spracované J. Guričanom k tejto téme.[1]

V minulej kapitole 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.

3.1 Matica prechodu, podobnosť matíc

Pripomeňme, že ak $\vec{\alpha}_1, \dots, \vec{\alpha}_n$ je nejaká báza konečnorozmerného vektorového priestoru V, tak ľubovoľný vektor z V sa dá jednoznačne vyjadriť v tvare $\vec{\gamma} = c_1\vec{\alpha}_1 + \dots + c_n\vec{\alpha}_n$. (Pozri vetu I-4.4.16.) Tento fakt nám vlastne hovorí, že báza nám poskytuje akúsi súradnicovú sústavu v priestore V – každý vektor má jednoznačne určené súradnice $c_1, \dots, c_n$.

Definícia 3.1.1. Ak $\vec{\alpha}_1, \dots, \vec{\alpha}_n$ je báza vektorového priestoru V nad poľom F a $\vec{\gamma} \in V$, tak n-ticu $(c_1, \dots, c_n) \in F^n$ takú, že platí

$$\vec{\gamma} = c_1\vec{\alpha}_1 + \dots + c_n\vec{\alpha}_n$$

nazývame *súradnicami vektora* $\vec{\gamma}$ *v báze* $\vec{\alpha}_1, \dots, \vec{\alpha}_n$.

Jednou z otázok, ktorými sa budeme zaoberať 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 3.1.2. Nech $\vec{\alpha}_1, \dots, \vec{\alpha}_n$ a $\vec{\alpha}'_1, \dots, \vec{\alpha}'_n$ 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

$$\begin{aligned}\vec{\alpha}'_1 =& p_{11}\vec{\alpha}_1 + p_{12}\vec{\alpha}_2 + \cdots + p_{1n}\vec{\alpha}_n \\ \vec{\alpha}'_2 =& p_{21}\vec{\alpha}_1 + p_{22}\vec{\alpha}_2 + \cdots + p_{2n}\vec{\alpha}_n \\ &\vdots \\ \vec{\alpha}'_n =& p_{n1}\vec{\alpha}_1 + p_{n2}\vec{\alpha}_2 + \cdots + p_{nn}\vec{\alpha}_n\end{aligned} \tag{3.1}$$

Potom maticu $P = \|p_{ij}\|$ nazývame *matica prechodu* od bázy $\vec{\alpha}_1, \dots, \vec{\alpha}_n$ k báze $\vec{\alpha}'_1, \dots, \vec{\alpha}'_n$.

Inak povedané, matica prechodu je taká matica P, ktorej i-ty riadok je tvorený súradnicami vektoru $\vec{\alpha}'_i$ v báze $\vec{\alpha}_1, \dots, \vec{\alpha}_n$.

Poznámka 3.1.3. V literatúre nájdete často i presne opačnú definíciu, než sme uviedli mi. Teda niektorí autori by túto maticu nazvali maticou prechodu od $\vec{\alpha}'_1, \dots, \vec{\alpha}'_n$ k $\vec{\alpha}_1, \dots, \vec{\alpha}_n$.

Skúsme si rozmyslieť, čo vieme povedať o matici prechodu opačným smerom, t.j. od $\vec{\alpha}'_1, \dots, \vec{\alpha}'_n$ k $\vec{\alpha}_1, \dots, \vec{\alpha}_n$. 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 $\vec{\alpha}'_1, \dots, \vec{\alpha}'_n$ sme upravili pomocou (3.1).) 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 $\vec{\alpha}_1, \dots, \vec{\alpha}_n$ 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 I-5.5.8).

Ukážme si, ako sme celé predchádzajúce odvodenie mohli stručnejšie odvodiť pomocou maticového zápisu.

V prvom rade si uvedomme, že vzťah (3.1) sa dá ekvivalentne zapísať takto:

$$\begin{pmatrix} \vec{\alpha}'_1 \\ \vdots \\ \vec{\alpha}'_n \end{pmatrix} = P \begin{pmatrix} \vec{\alpha}_1 \\ \vdots \\ \vec{\alpha}_n \end{pmatrix}. \tag{3.2}$$

Keďže vektory $\vec{\alpha}'_1, \dots, \vec{\alpha}'_n$ aj $\vec{\alpha}_1, \dots, \vec{\alpha}_n$ 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 (3.2), 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 $\vec{\alpha}'_1, \dots, \vec{\alpha}'_n$ k $\vec{\alpha}_1, \dots, \vec{\alpha}_n$.

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 3.1.4. Predchádzajúce odvodenie v skutočnosti nebolo úplne korektné. Z vektorov $\vec{\alpha}_1, \dots, \vec{\alpha}_n$ 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 I-5.5.14). 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 $\vec{\alpha}_1, \dots, \vec{\alpha}_n$ súradnice $(c_1, \dots, c_n)$ a $f\colon V \to F^n$ je ľubovoľný izomorfizmus, tak aj súradnice vektora $f(\vec{\gamma})$ v báze $f(\vec{\alpha}_1), \dots, f(\vec{\alpha}_n)$ sú $(c_1, \dots, c_n)$. 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:

Veta 3.1.5. *Ak P je matica prechodu od bázy $\vec{\alpha}_1, \dots, \vec{\alpha}_n$ k báze $\vec{\alpha}'_1, \dots, \vec{\alpha}'_n$, tak matica P je regulárna a matica P^{-1} je matica prechodu opačným smerom, teda od $\vec{\alpha}'_1, \dots, \vec{\alpha}'_n$ k $\vec{\alpha}_1, \dots, \vec{\alpha}_n$.*

Ukážeme, že to funguje aj naopak – pre každú bázu a regulárnu maticu P použitím predpisu (3.1) dostaneme opäť bázu.

Tvrdenie 3.1.6. *Nech $P = \|p_{ij}\|$ je regulárna matica typu $n \times n$ a $\vec{\alpha}_1, \dots, \vec{\alpha}_n$ je báza vektorového priestoru V. Potom aj vektory $\vec{\alpha}'_1, \dots, \vec{\alpha}'_n$ 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 I-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 $\vec{\alpha}'_1, \dots, \vec{\alpha}'_n$ v báze $\vec{\alpha}_1, \dots, \vec{\alpha}_n$ 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 $\vec{\alpha}_1, \dots, \vec{\alpha}_n$ na pravej strane predchádzajúcej rovnosti sú presne zložky vektora $(c_1, \dots, c_n)P$. Pretože vektory $\vec{\alpha}_1, \dots, \vec{\alpha}_n$ sú lineárne nezávislé, aby platila táto rovnosť, musia sa všetky tieto koeficienty rovnať nule. Dostali sme teda rovnosti

$$\begin{gathered}
(c_1, \dots, c_n)P = \vec{0} \\
(c_1, \dots, c_n) = \vec{0}P^{-1} = \vec{0} \\
c_1 = \cdots = c_n = 0
\end{gathered}$$

Tým sme ukázali, že $\vec{\alpha}'_1, \dots, \vec{\alpha}'_n$ sú lineárne nezávislé. □

Opäť môžeme to isté tvrdenie dokázať aj s využitím (3.2).

Dôkaz. Keďže $\vec{\alpha}_1, \dots, \vec{\alpha}_n$ 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 I-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.

Veta 3.1.7. *Nech $\vec{\alpha}_1, \dots, \vec{\alpha}_n$ a $\vec{\alpha}'_1, \dots, \vec{\alpha}'_n$ sú bázy vektorového priestoru V. Nech P je matica prechodu od bázy $\vec{\alpha}_1, \dots, \vec{\alpha}_n$ k báze $\vec{\alpha}'_1, \dots, \vec{\alpha}'_n$. Nech $\vec{\gamma} \in V$ a $(x_1, \dots, x_n)$ sú súradnice vektora $\vec{\gamma}$ v báze $\vec{\alpha}_1, \dots, \vec{\alpha}_n$, $(x'_1, \dots, x'_n)$ sú jeho súradnice v báze $\vec{\alpha}'_1, \dots, \vec{\alpha}'_n$. Potom platí*

$$(x_1, \dots, x_n) = (x'_1, \dots, x'_n)P.$$

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 $\vec{\alpha}'_1, \dots, \vec{\alpha}'_n$ súradnice $(x'_1, \dots, x'_n)$ znamená, že platí rovnosť

$$\vec{\gamma} = x'_1\vec{\alpha}'_1 + \dots + x'_n\vec{\alpha}'_n.$$

Ak do tejto rovnosti dosadíme za $\vec{\alpha}'_1, \dots, \vec{\alpha}'_n$ z (3.1), 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 $\vec{\alpha}_1, \dots, \vec{\alpha}_n$, 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 $(x'_1, \dots, x'_n)P$. Tým je tvrdenie vety dokázané. □

Opäť môžeme predchádzajúci dôkaz zapísať stručnejšie maticovým zápisom.

Dôkaz. Uvedomme si najprv, že $\vec{\gamma}$ má súradnice $(x'_1, \dots, x'_n)$ v báze $\vec{\alpha}'_1, \dots, \vec{\alpha}'_n$ práve vtedy, keď platí rovnosť

$$\vec{\gamma} = (x'_1, \dots, x'_n)\begin{pmatrix} \vec{\alpha}'_1 \\ \vdots \\ \vec{\alpha}'_n \end{pmatrix}.$$

Spolu s rovnosťou (3.2) potom dostaneme

$$(x'_1, \dots, x'_n)\begin{pmatrix} \vec{\alpha}'_1 \\ \vdots \\ \vec{\alpha}'_n \end{pmatrix} = (x'_1, \dots, x'_n)P\begin{pmatrix} \vec{\alpha}_1 \\ \vdots \\ \vec{\alpha}_n \end{pmatrix},$$

teda súradnice v báze $\vec{\alpha}_1, \dots, \vec{\alpha}_n$ sú $(x'_1, \dots, x'_n)P$. □

Matica zobrazenia v danej báze

Definícia 3.1.8. Nech V je vektorový priestor a $f\colon V \to V$ je lineárne zobrazenie. Nech $\vec{\alpha}_1, \dots, \vec{\alpha}_n$ je báza V. *Matica zobrazenia f vzhľadom na bázu* $\vec{\alpha}_1, \dots, \vec{\alpha}_n$ je matica $A = \|a_{ij}\|$ taká, že platí

$$f(\vec{\alpha}_i) = a_{i1}\vec{\alpha}_1 + \dots + a_{in}\vec{\alpha}_n.$$

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 I-5.3.7). 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 I-5.4.9).

Tvrdenie 3.1.9. *Nech V je vektorový priestor, $\vec{\alpha}_1, \dots, \vec{\alpha}_n$ je báza V a $f\colon V \to V$ je lineárne zobrazenie. Ak A je matica zobrazenia f pri báze $\vec{\alpha}_1, \dots, \vec{\alpha}_n$ a vektor $\vec{\gamma}$ má v tejto báze súradnice $(x_1, \dots, x_n)$, tak jeho obraz $f(\vec{\gamma})$ má v tej istej báze súradnice*

$$(x_1, \dots, x_n)A.$$

Dôkaz. Podľa predpokladov platí $\vec{\gamma} = x_1\vec{\alpha}_1 + \dots + x_n\vec{\alpha}_n$. 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 $\vec{\alpha}_1, \dots, \vec{\alpha}_n$ je $\sum_{i=1}^{n} x_i a_{ij}$, čo je skutočne j-ta súradnica vektora $(x_1, \dots, x_n)A$. □

Opäť nás bude zaujímať to, ako sa zmení matica zobrazenia, ak zmeníme bázu vektorového priestoru.

Veta 3.1.10. *Nech V je vektorové priestory, $f\colon V \to V$ je lineárne zobrazenie a $\vec{\alpha}_1, \dots, \vec{\alpha}_n$, $\vec{\alpha}'_1, \dots, \vec{\alpha}'_n$ sú bázy priestoru V. Ak P je matica prechodu od $\vec{\alpha}_1, \dots, \vec{\alpha}_n$ k $\vec{\alpha}'_1, \dots, \vec{\alpha}'_n$, A je matica zobrazenia f pri báze $\vec{\alpha}_1, \dots, \vec{\alpha}_n$ a B je matica tohoto zobrazenia pri báze $\vec{\alpha}'_1, \dots, \vec{\alpha}'_n$, tak platí*

$$B = PAP^{-1}. \tag{3.3}$$

Proof. Uvažujme vektor $\vec{\gamma}$, ktorý má v báze $\vec{\alpha}'_1, \dots, \vec{\alpha}'_n$ súradnice $(x'_1, \dots, x'_n)$. Podľa vety 3.1.7 má tento vektor v báze $\vec{\alpha}_1, \dots, \vec{\alpha}_n$ súradnice $(x'_1, \dots, x'_n)P$. Ďalej z tvrdenia 3.1.9 vieme, že tento vektor sa v zobrazení f zobrazí na taký vektor, ktorý má súradnice $(x'_1, \dots, x'_n)PA$ (ide opäť o súradnice v báze $\vec{\alpha}_1, \dots, \vec{\alpha}_n$.)

Schematicky môžeme situáciu v súradniciach $\vec{\alpha}_1, \dots, \vec{\alpha}_n$ znázorniť takto

$$(x'_1, \dots, x'_n)P \mapsto (x'_1, \dots, x'_n)PA.$$

Čo dostaneme, ak sa na situáciu pozrieme v súradniciach $\vec{\alpha}'_1, \dots, \vec{\alpha}'_n$? Súradnice vektora $f(\vec{\gamma})$ vieme použitím vety 3.1.7 previesť do tejto bázy použitím matice prechodu od $\vec{\alpha}'_1, \dots, \vec{\alpha}'_n$ k $\vec{\alpha}_1, \dots, \vec{\alpha}_n$. Podľa vety 3.1.5 je to matica P^{-1}.

Vektor $f(\vec{\gamma})$ má teda v báze $\vec{\alpha}'_1, \dots, \vec{\alpha}'_n$ súradnice $(x'_1, \dots, x'_n)PAP^{-1}$. V súradniciach $\vec{\alpha}'_1, \dots, \vec{\alpha}'_n$ to teda vyzerá takto:

$$(x'_1, \dots, x'_n) \mapsto (x'_1, \dots, x'_n)PAP^{-1}$$

Podľa tvrdenia 3.1.9 však súčasne musí platiť, že $f(\vec{\gamma})$ má v báze $\vec{\alpha}'_1, \dots, \vec{\alpha}'_n$ súradnice $(x'_1, \dots, x'_n)B$.

$$(x'_1, \dots, x'_n) \mapsto (x'_1, \dots, x'_n)B$$

Z toho dostávame rovnosť

$$(x_1',\dots,x_n')B = (x_1',\dots,x_n')PAP^{-1}.$$

Pretože táto rovnosť platí pre ľubovoľnú n-ticu $(x_1',\dots,x_n') \in F^n$, musí platiť maticová rovnosť

$$B = PAP^{-1}.$$

□

Definícia 3.1.11. 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 3.1.10 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 3.1.1. Ukážte, že podobnosť matíc (chápaná ako relácia na $M_{n,n}(F)$) je relácia ekvivalencie.

Úloha 3.1.2. 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 3.1.3. 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 3.1.4. Nech $A = cI$. Aké matice sú podobné s maticou A?

Úloha 3.1.5. 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 3.1.6. Ukážte, že ak matica A je podobná matici B, tak aj matice A^{-1} a B^{-1} sú podobné.

Úloha 3.1.7. Ukážte, že ak A a B sú podobné, tak majú rovnakú hodnosť, determinant a stopu. (Stopu matice sme definovali v úlohe 1.2.8.)

3.2 Podobnosť s diagonálnou maticou

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 $\mathrm{diag}(d_1, d_2, \ldots, d_n)$.

Ak teda A je podobná diagonálnej, je $PAP^{-1} = \mathrm{diag}(d_1, d_2, \ldots, d_n) = D$ pre vhodnú maticu P a vhodné čísla $d_1, \ldots, d_n$, potom vieme ľahko vypočítať napr. A^{100} ako

$$P^{-1}\mathrm{diag}(d_1^{100}, d_2^{100}, \ldots, d_n^{100})P,$$

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:

$$\begin{pmatrix} \vec{\alpha}_1 \\ \vdots \\ \vec{\alpha}_n \end{pmatrix} A = D \begin{pmatrix} \vec{\alpha}_1 \\ \vdots \\ \vec{\alpha}_n \end{pmatrix},$$
$$\begin{pmatrix} \vec{\alpha}_1 A \\ \vdots \\ \vec{\alpha}_n A \end{pmatrix} = \begin{pmatrix} d_1\vec{\alpha}_1 \\ \vdots \\ d_n\vec{\alpha}_n \end{pmatrix}. \tag{3.4}$$

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 3.2.1. 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 (3.3) je

$$A_\varphi^\alpha = PA_\varphi^\varepsilon P^{-1}, \text{ kde } \quad P = \begin{pmatrix} -3 & 2 \\ 0 & 1 \end{pmatrix},$$

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 $A_\varphi^\alpha = \mathrm{diag}(1,4) = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix} = PAP^{-1}$

Iný spôsob, ako môžeme zdôvodniť túto rovnosť je použiť rovnaký postup ako pri odvodení (3.4).

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 3.2.2. 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.

Veta 3.2.3. *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 (3.3) 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 (3.4).

$\Leftarrow$: 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

Lema 3.2.4. *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. $\vec{\alpha}_1$ 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 $\vec{\alpha}_1, \dots, \vec{\alpha}_{i-1}$ už sú lineárne nezávislé. Teraz využijeme, že $\vec{\alpha}_1, \dots, \vec{\alpha}_k$ sú vlastné vektory, preto

$$\vec{\alpha}_i A = c_i \vec{\alpha}_i = c_i(a_1\vec{\alpha}_1 + \dots + a_{i-1}\vec{\alpha}_{i-1}) = c_i a_1 \vec{\alpha}_1 + \dots + c_i a_{i-i}\vec{\alpha}_{i-1},$$

kde aspoň jedno z čísiel $a_1, \dots, a_{i-1}$ je nenulové (inak by bol vektor $\vec{\alpha}_i$ nulový, čo sa vlastnému vektoru nemôže stať) ale aj

$$(a_1\vec{\alpha}_1 + \dots + a_{i-1}\vec{\alpha}_{i-1})A = a_1\vec{\alpha}_1 A + \dots + a_{i-i}\vec{\alpha}_{i-1}A = c_1 a_1\vec{\alpha}_1 + \dots + c_{i-1}a_{i-i}\vec{\alpha}_{i-1}$$

odkiaľ porovnaním a prehodením všetkých členov na ľavú stranu získame rovnosť

$$(c_i - c_1)a_1\vec{\alpha}_1 + \dots + (c_i - c_{i-1})a_i\vec{\alpha}_{i-1} = \vec{0}$$

Keďže všetky čísla $c_i - c_1, \dots, c_i - c_{i-1}$ sú nenulové, tak aspoň jedno z čísiel $(c_i - c_1)a_1, \dots, (c_i - c_{i-1})a_{i-1}$ je nenulové, čím dostávame spor s lineárnou nezávislosťou vektorov $\vec{\alpha}_1, \dots, \vec{\alpha}_{i-1}$. □

Príslušná postačujúca podmienka (ešte stále numericky pomerne náročná) potom znie

Dôsledok 3.2.5. *Nech $A = ||a_{ij}||$ je štvorcová matica typu $n \times n$ nad poľom F a vlastné čísla $c_1, \dots, c_n$ 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 3.2.4 a vety 3.2.3 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).

Lema 3.2.6. *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

Dôsledok 3.2.7. *Pre maticu $A = ||a_{ij}||$ - štvorcová matica typu $n \times n$ nad poľom F položme* $\mathrm{Tr}(A) = a_{11} + a_{22} + \dots + a_{nn}$ *- 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

$$A - xI = \begin{vmatrix} a_{11}-x & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22}-x & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{n-1,n} & a_{nn}-x \end{vmatrix}$$

Podľa definície determinant $|B| = ||b_{ij}||$ je súčet súčinov $(-1)^{i(\varphi)} b_{1\varphi(1)} \dots b_{n\varphi(n)}$, 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

$$+(a_{11}-x)) \cdots (a_{nn}-x) = (-1)^n x^n + (-1)^{n-1}(a_{11}+\dots+a_{nn})x^{n-1} + \dots = (-1)^n x + (-1)^{n-1} \operatorname{Tr}(A) x^{n-1} + \dots$$

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 I-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 $|AB| = |A||B| = |BA|$.

Príklad 3.2.8. 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 3.2.6 platí iba jedným smerom. (Iný kontrapríklad môžete nájsť v úlohe 3.2.13.)

Teraz prejdeme ku druhému sľubovanému kritériu, ktoré zabezpečí, že matica je podobná s diagonálnou. Je založené na úplne inom princípe ako kritérium z dôsledku 3.2.5. Nižšie je sformulované pod názvom „veta o hlavných osiach“ (veta 3.2.11). 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.

Veta 3.2.9 (Schurrova 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

$$\left(\begin{array}{c|c} & c_1 \\ \mathrm{B} & \vdots \\ & c_{n-1} \\ \hline 0 \dots 0 & a_n \end{array}\right) = A' = PAP^T$$

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, B 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

$$\left(\begin{array}{c|c} & 0 \\ Q & \vdots \\ & 0 \\ \hline 0 \dots 0 & 1 \end{array}\right) \cdot \left(\begin{array}{c|c} & c_1 \\ B & \vdots \\ & c_{n-1} \\ \hline 0 \dots 0 & a_n \end{array}\right) \cdot \left(\begin{array}{c|c} & 0 \\ Q^* & \vdots \\ & 0 \\ \hline 0 \dots 0 & 1 \end{array}\right) =$$

$$= \left(\begin{array}{c|c} QBQ^* & Q\gamma^T \\ \hline 0 \dots 0 & a_n \end{array}\right) = \left(\begin{array}{c|c} T' & Q\gamma^T \\ \hline 0 \dots 0 & a_n \end{array}\right),$$

kde $\gamma = (c_1, \dots, c_{n-1})$. Posledná matica je ale vďaka indukčnému predpokladu horná trojuholníková matica. Keďže je matica

$$Q' = \left(\begin{array}{c|c} & 0 \\ Q & \vdots \\ & 0 \\ \hline 0 \dots 0 & 1 \end{array}\right)$$

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 Schurrovej vete sa stačí presvedčiť, že reálna symetrická matica má vždy všetky vlastné čísla reálne - vedeli by sme, že je (dokonca ortogonálne) podobná diagonálnej matici.

Veta 3.2.10. *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 $(z_1,\dots,z_n) = (a_1+b_1i,\dots,a_n+b_ni) = (a_1,\dots,a_n) + (b_1,\dots,b_n)i$ $(a_i,b_i \in R)$, 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

$$\langle\vec{\alpha},\vec{\beta}A\rangle = \langle\alpha, a\vec{\beta}+b\vec{\alpha}\rangle = a\langle\vec{\alpha},\vec{\beta}\rangle + b\langle\vec{\alpha},\vec{\alpha}\rangle$$

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

$$-b\langle\vec{\beta},\vec{\beta}\rangle = b\langle\vec{\alpha},\vec{\alpha}\rangle,$$
$$b(\langle\vec{\beta},\vec{\beta}\rangle + \langle\vec{\alpha},\vec{\alpha}\rangle) = 0$$

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 Schurrovej vety teda dostávame tvrdenie, ktoré je známe ako „Veta o hlavných osiach“:

Veta 3.2.11 (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

Veta 3.2.12. *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. $\vec{\alpha} \perp \vec{\beta}$. $\square$

Ilustrujme si použitie tejto vety na konkrétnom príklade.

Príklad 3.2.13.

$$A = \begin{pmatrix} 1 & 4 & -2 \\ 4 & 1 & -2 \\ -2 & -2 & -2 \end{pmatrix}$$

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

$$A = P^T DP,$$

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

$$A = \begin{pmatrix} \vec{\alpha}_1^T & \ldots & \vec{\alpha}_n^T \end{pmatrix} \begin{pmatrix} d_1 & 0 & \ldots & 0 \\ 0 & d_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & d_n \end{pmatrix} \begin{pmatrix} \vec{\alpha}_1 \\ \vdots \\ \vec{\alpha}_n \end{pmatrix} = \begin{pmatrix} \vec{\alpha}_1^T & \ldots & \vec{\alpha}_n^T \end{pmatrix} \begin{pmatrix} d_1\vec{\alpha}_1 \\ \vdots \\ d_n\vec{\alpha}_n \end{pmatrix}$$

$$A = d_1\vec{\alpha}_1^T\vec{\alpha}_1 + d_2\vec{\alpha}_2^T\vec{\alpha}_2 + \cdots + d_n\vec{\alpha}_n^T\vec{\alpha}_n \tag{3.5}$$

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 1.2.12).

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*.

Ak by sme chceli teóriu načatú v tomto texte chceli študovať serióznejšie, potrebovali by sme tzv. Caley-Hamiltonova vetu, my ju teraz uvedieme len ako malé doplnenie problematiky a možno ako zaujímavosť.

Veta 3.2.14 (Cayley-Hamilton)**.** *Nech A je štvorcová matica nad poľom F. Potom A je koreňom svojho charakteristického polynómu, presnejšie ak je $ch_A(x) = (-1)^n x^n + c_{n-1}x^{n-1} + \dots + c_0$ tak $(-1)^n A^n + c_{n-1}A^{n-1} + \dots + c_0 I = \|0\|_{n\times n}$.*

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 I-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 I-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 $\operatorname{adj}(A)$, t.j. platí $A \cdot \operatorname{adj}(A) = |A| \cdot I$.

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 $\operatorname{adj}(B)$ 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

$$\operatorname{adj}(A - xI) = \operatorname{adj}(B) = C_{n-1}x^{n-1} + \dots + C_1 x + C_0$$

Označme si $ch_A(x) = |B| = (-1)^n x^n + c_{n-1}x^{n-1} + \dots + c_1 x + c_0$

$|B| \cdot I = (-1)^n I x^n + c_{n-1} \cdot I x^{n-1} + \dots + c_1 \cdot I x + c_0 I = B \cdot \operatorname{adj}(B) = (A - xI) \cdot (C_{n-1}x^{n-1} + \dots + C_1 x + C_0),$

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 $A^2, \ldots$, 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

$$\begin{aligned}
(-1)^n A^n + c_{n-1} A^{n-1} + \cdots + c_1 A + c_o I &= \\
-A^n C_{n-1} + (A^n C_{n-1} - A^{n-1} C_{n-2}) + \cdots + (A^3 C_2 - A^2 C_1) + (A^2 C_1 - AC_0) + AC_0 &= \|0\|_{n\times n}
\end{aligned}$$

t.j. $ch_A(A) = \|0\|$. □

Cvičenia

Úloha 3.2.1. Nájdite vlastné hodnoty a vlastné vektory daných matíc nad poľom $\mathbb{C}$:
a) $\begin{pmatrix} 2 & 4 \\ 5 & 3 \end{pmatrix}$
b) $\begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$
c) $\begin{pmatrix} 1 & 2 \\ 2 & -2 \end{pmatrix}$
d) $\begin{pmatrix} -1 & 2i \\ -2i & 2 \end{pmatrix}$
e) $\begin{pmatrix} 2 & 1 \\ -1 & 0 \end{pmatrix}$
f) $\begin{pmatrix} 4 & 1 \\ 3 & 2 \end{pmatrix}$
Ak taká matica existuje, nájdite regulárnu maticu P s vlastnosťou, že PAP^{-1} je diagonálna.

Úloha 3.2.2. 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 3.2.3. 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 3.2.4. Ukážte, že pre $c \in \mathbb{R} \setminus \{0\}$ matica $\begin{pmatrix} 1 & c \\ 0 & 1 \end{pmatrix}$ nie je podobná s diagonálnou maticou. Aká je geometrická interpretácia tohoto výsledku?

Úloha 3.2.5. 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 3.2.6. Ak A, B sú regulárne matice, tak AB a BA majú rovnaké vlastné hodnoty.

Úloha 3.2.7. 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 3.2.8. Ak A je idempotentná matica, čiže $A^2 = A$, tak jej vlastné hodnoty môžu byť jedine 0 alebo 1.

Úloha 3.2.9. 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) $\begin{pmatrix} -1 & -1 & 1 \\ 0 & -2 & 1 \\ 0 & 0 & -1 \end{pmatrix}$

Úloha 3.2.10. 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) $\begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix}$

Úloha 3.2.11. 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 3.2.12. 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?

$A = \begin{pmatrix} 1 & 4 & -2 \\ 4 & 1 & -2 \\ -2 & -2 & -2 \end{pmatrix}$; $B = \begin{pmatrix} 1 & 5 & -2 \\ 2 & 1 & -4 \\ -2 & -4 & -2 \end{pmatrix}$.

Úloha 3.2.13. 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 3.2.14*. 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 3.2.15*. 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.)

3.3 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“.

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 $O^T O = I$, 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 3.3.1). 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.

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

$$f(x_1,x_2)=a_{11}x_1^2+2a_{12}x_1x_2+a_{22}x_2^2+2a_1x_1+2a_2x_2+d, \tag{3.6}$$

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 (3.6), tento problém bude nejako súvisieť s kvadratickými formami.

Ako sa dá uhádnuť z označenia použitého v (3.6), 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 3.2.11 teda existuje ortogonálna matica P tak, že $PAP^T = \mathrm{diag}(\lambda_1, \lambda_2)$, 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.

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 3.3.1. *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 3.3.2. *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 detereminante 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 $\emptyset$

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

$$(b_3 + tu_3 + su_3)^2 - (b_1 + tu_1 + su_1)^2 - (b_2 + tu_2 + su_2)^2 = 0$$

a po úprave

$$\begin{aligned}(u_3^2 - u_1^2 - u_2^2)t^2 + 2(u_3v_3 - u_1v_1 - u_2v_2)st + (v_3^2 - v_1^2 - v_2^2)s^2 + \\ 2(b_3u_3 - b_1u_1 - b_2u_2)t + 2(b_3v_3 - b_1v_1 - b_2v_2)s + (b_3^2 - b_1^2 - b_1^2) = 0\end{aligned}$$

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.

$$\left|\begin{smallmatrix} u_3^2-u_1^2-u_2^2 & u_3v_3-u_1v_1-u_2v_2 \\ u_3v_3-u_1v_1-u_2v_2 & v_3^2-v_1^2-v_2^2 \end{smallmatrix}\right| = (u_3^2-u_1^2-u_2^2)(v_3^2-v_1^2-v_2^2)-(u_3v_3-u_1v_1-u_2v_2)^2 =$$
$$= u_1^2v_1^2+u_2^2v_2^2+u_3^2v_3^2+u_1^2v_2^2+u_2^2v_1^2-u_1^2v_3^2-u_3^2v_1^2-u_2^2v_3^2-u_3^2v_2^2$$
$$-u_1^2v_1^2-u_2^2v_2^2-u_3^2v_3^2-2u_1u_2v_1v_2+2u_1u_3v_1v_3+2u_2u_3v_2v_3 =$$
$$= u_1^2v_2^2+u_2^2v_1^2-2u_1u_2v_1v_2-u_1^2v_3^2-u_3^2v_1^2+2u_1u_3v_1v_3-u_2^2v_3^2-u_3^2v_2^2+2u_2u_3v_2v_3 =$$
$$= (u_1v_2-u_2v_1)^2-(u_1v_3-u_3v_1)^2-(u_2v_3-u_3v_2)^2$$

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

$$K=\left(\begin{smallmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -a \end{smallmatrix}\right).$$

Táto matica je diagonálna, čo znamená, že je symetrická a tiež, že ľahko vieme vyrátať inverznú maticu $K^{-1}=\left(\begin{smallmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -\frac{1}{a} \end{smallmatrix}\right)$.

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 $\vec{u}K\vec{u}^Ts^2+(\vec{u}K\vec{v}^T+\vec{v}K\vec{u})^Tst+\vec{v}K\vec{v}^Tt^2+\dots=0$. (Chceme vyrátať δ, čiže nás zaujímajú len členy, ktoré sú stupňa 2.) Máme teda

$$\delta=\left|\begin{smallmatrix} \vec{u}K\vec{u}^T & \vec{u}K\vec{v}^T \\ \vec{v}K\vec{u}^T & \vec{v}K\vec{v}^T \end{smallmatrix}\right|.$$

Na vyjadrenie tohoto determinantu nám pomôže, keď si všimneme

$$\left(\begin{smallmatrix} \vec{u} \\ \vec{v} \\ \vec{n}K^{-1} \end{smallmatrix}\right) K \left(\begin{smallmatrix} \vec{u}^T & \vec{v}^T & K^{-1}\vec{n}^T \end{smallmatrix}\right) = \left(\begin{smallmatrix} \vec{u}K\vec{u}^T & \vec{u}K\vec{v}^T & 0 \\ \vec{v}K\vec{u}^T & \vec{v}K\vec{v}^T & 0 \\ 0 & 0 & \vec{n}K^{-1}\vec{n}^T \end{smallmatrix}\right).$$

Ak na obe strany rovnosti použijeme determinant, tak máme:

$$\left|\begin{smallmatrix} \vec{u} \\ \vec{v} \\ \vec{n}K^{-1} \end{smallmatrix}\right|^2 |K| = \delta.\vec{n}K^{-1}\vec{n}^T$$

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

$$\left|\begin{smallmatrix} \vec{u} \\ \vec{v} \\ \vec{n}K^{-1} \end{smallmatrix}\right| = \left|\begin{smallmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ n_1 & n_2 & -n_3 \end{smallmatrix}\right| = n_1^2+n_2^2-\frac{1}{a}n_3^2,$$

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.

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

Úloha 3.3.1. Dokážte, že ortogonálne matice typu $n \times n$ tvoria s operáciou násobenia matíc grupu.

3.4 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 reprezentant, 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 [G]. Dôkaz toho tvrdenia môžete nájsť aj v [Z, 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 3.4.1. *Jordanov blok* veľkosti k prislúchajúci číslu λ je matica typu $k \times k$ tvaru

$$J_k(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \dots & \dots & 0 \\ 0 & \lambda & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \lambda & 1 & 0 \\ 0 & \dots & \dots & 0 & \lambda & 1 \\ 0 & \dots & \dots & \dots & 0 & \lambda \end{pmatrix}$$

Veta 3.4.2 (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 3.2.7 a poznámke za ním.

Keď už poznáme kanonický tvar z vety 3.4.2, 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 $\vec{\alpha}_1, \dots, \vec{\alpha}_n$ a prečítajme si tú istú maticovú rovnosť po riad-

[3]Rozumieme tým násobnosť vlastnej hodnoty ako koreňa charakteristického polynómu.

koch.

$$\begin{pmatrix} \vec{\alpha}_1 \\ \vec{\alpha}_2 \\ \vdots \\ \vec{\alpha}_{n-1} \\ \vec{\alpha}_n \end{pmatrix} A = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda & 1 \\ 0 & \dots & \dots & 0 & \lambda \end{pmatrix} \begin{pmatrix} \vec{\alpha}_1 \\ \vec{\alpha}_2 \\ \vdots \\ \vec{\alpha}_{n-1} \\ \vec{\alpha}_n \end{pmatrix} \tag{3.7}$$

$$\begin{pmatrix} \vec{\alpha}_1 A \\ \vec{\alpha}_2 A \\ \vdots \\ \vec{\alpha}_{n-1} A \\ \vec{\alpha}_n A \end{pmatrix} = \begin{pmatrix} \lambda\vec{\alpha}_1 + \alpha_2 \\ \lambda\vec{\alpha}_2 + \alpha_3 \\ \vdots \\ \lambda\vec{\alpha}_{n-1} + \alpha_n \\ \lambda\vec{\alpha}_n \end{pmatrix} \tag{3.8}$$

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 (3.8) a $PAP^{-1} = J_n(\lambda)$. Navyše, ukázali sme, že aby pre maticu P tieto rovnosti platili, jej riadky musia spĺňať všetky uvedené podmienky.

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 $\vec{\alpha}_1, \dots, \vec{\alpha}_n$, pri ktorej má zobrazenie určené (pri štandardnej báze) maticou A maticu $J_\lambda(n)$. To znamená špeciálne, že má platiť

$$\begin{gathered} \vec{\alpha}_n A = \lambda\vec{\alpha}_n \\ \vec{\alpha}_{n-1} A = \lambda\vec{\alpha}_{n-1} + \vec{\alpha}_n \\ \vdots \\ \vec{\alpha}_1 = \lambda\vec{\alpha}_1 + \vec{\alpha}_2 \end{gathered}$$

Dostali sme teda presne tie isté rovnice pre $\vec{\alpha}_1, \dots, \vec{\alpha}_n$.

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 3.4.3. 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.

$$ch_A(x) = \begin{vmatrix} 6-x & 1 & -3 & 2 & 5 \\ -1 & 2-x & 1 & -3 & 0 \\ 1 & 0 & 1-x & 3 & 0 \\ -1 & 0 & 1 & 3-x & -2 \\ -2 & 0 & 2 & 2 & -2-x \end{vmatrix} \overset{(1)}{=} (-1) \begin{vmatrix} -1 & 1 & -3 & 0 \\ 1 & 1-x & 3 & 0 \\ -1 & 1 & 3-x & -2 \\ -2 & 2 & 2 & -2-x \end{vmatrix} + (2-x) \begin{vmatrix} 6-x & -3 & 2 & 5 \\ 1 & 1-x & 3 & 0 \\ -1 & 1 & 3-x & -2 \\ -2 & 2 & 2 & -2-x \end{vmatrix}$$

$$\begin{vmatrix} -1 & 1 & -3 & 0 \\ 1 & 1-x & 3 & 0 \\ -1 & 1 & 3-x & -2 \\ -2 & 2 & 2 & -2-x \end{vmatrix} \overset{(2)}{=} \begin{vmatrix} -1 & 1 & -3 & 0 \\ 0 & 2-x & 0 & 0 \\ -1 & 1 & 3-x & -2 \\ -2 & 2 & 2 & -2-x \end{vmatrix} = (2-x) \begin{vmatrix} -1 & -3 & 0 \\ -1 & 3-x & -2 \\ -2 & 2 & -2-x \end{vmatrix} \overset{(3)}{=} (2-x) \begin{vmatrix} -1 & -3 & 0 \\ 0 & 6-x & -2 \\ 0 & 8 & -2-x \end{vmatrix} =$$
$$= (x-2) \begin{vmatrix} 6-x & -2 \\ 8 & -2-x \end{vmatrix} = (x-2)(x^2-4x-12+16) = (x-2)(x^2-4x+4) = (x-2)^3$$

$$\begin{vmatrix} 6-x & -3 & 2 & 5 \\ 1 & 1-x & 3 & 0 \\ -1 & 1 & 3-x & -2 \\ -2 & 2 & 2 & -2-x \end{vmatrix} \overset{(4)}{=} \begin{vmatrix} 6-x & -3 & 2 & 5 \\ 1 & 1-x & 3 & 0 \\ -1 & 1 & 3-x & -2 \\ 0 & 0 & 2x-4 & 2-x \end{vmatrix} = (x-2) \begin{vmatrix} 6-x & -3 & 2 & 5 \\ 1 & 1-x & 3 & 0 \\ -1 & 1 & 3-x & -2 \\ 0 & 0 & 2 & -1 \end{vmatrix} \overset{(5)}{=} (x-2) \begin{vmatrix} 6-x & -3 & 12 & 5 \\ 1 & 1-x & 3 & 0 \\ -1 & 1 & -1-x & -2 \\ 0 & 0 & 0 & -1 \end{vmatrix} =$$
$$= -(x-2) \begin{vmatrix} 6-x & -3 & 12 \\ 1 & 1-x & 3 \\ -1 & 1 & -1-x \end{vmatrix} \overset{(6)}{=} -(x-2) \begin{vmatrix} 6-x & -3 & 12 \\ 0 & 2-x & 2-x \\ -1 & 1 & -1-x \end{vmatrix} = (x-2)^2 \begin{vmatrix} 6-x & -3 & 12 \\ 0 & 1 & 1 \\ -1 & 1 & -1-x \end{vmatrix} =$$
$$\overset{(7)}{=} (x-2)^2 \begin{vmatrix} 6-x & 0 & 15 \\ 0 & 1 & 1 \\ -1 & 0 & -2-x \end{vmatrix} \overset{(8)}{=} (x-2)^2 \begin{vmatrix} 6-x & 15 \\ -1 & -2-x \end{vmatrix} = (x-2)^2(x^2-4x-12+15) = (x-2)^2(x^2-4x+3)$$

Použité úpravy:
(1) Rozvoj podľa druhého stĺpca
(2) Pripočítanie prvého riadku k druhému
(3) Od ostatných riadkov sme odčítali vhodný násobok prvého
(4) Odpočítanie dvojnásobku tretieho riadku od posledného
(5) Pripočítanie dvojnásobku posledného stĺpca k tretiemu
(6) Pripočítanie tretieho riadku k druhému
(7) Rozvoj podľa druhého stĺpca
Pomocou týchto čiastkových výpočtov už môžeme zrátať charakteristický polynóm.

$$ch_A(x) = -(x-2)^3 - (x-2)(x-2)^2(x^2-4x+3) = -(x-2)^3(x^2-4x+4) = -(x-2)^5$$

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:

$$\begin{pmatrix} 0 & \lambda & 0 & \dots & \dots & 0\\ 0 & 0 & \lambda & 0 & & 0\\ \vdots & & \ddots & \ddots & \ddots & \vdots\\ 0 & & & 0 & \lambda & 0\\ 0 & & & & 0 & \lambda\\ 0 & \dots & \dots & \dots & \dots & 0 \end{pmatrix}$$

Vidíme, že v každom taktomto bloku je jeden riadok nulový a ostatné sú lineárne nezávislé. Bloky zodpovedajúce ostatným vlastným hodnotám vyzerajú takto

$$\begin{pmatrix} \mu-\lambda & \mu & 0 & \dots & \dots & 0\\ 0 & \mu-\lambda & \mu & 0 & & 0\\ \vdots & & \ddots & \ddots & \ddots & \vdots\\ 0 & & & \mu-\lambda & \mu & 0\\ 0 & & & & \mu-\lambda & \mu\\ 0 & \dots & \dots & \dots & \dots & \mu-\lambda \end{pmatrix}$$

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 podobný úvaha ako robíme za rovnosťou (3.13).) 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 $n_1 + \cdots + n_k$, kde n_k označuje počet Jordanových blokov veľkosti aspoň k. 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 3.4.4. 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 3.4.5. 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 3.4.1. 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) $ch_A(x) = -(x-1)^3$, $J = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$; 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 3.4.2. 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 3.4.3. 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 3.4.4. Nájdite charakteristický polynóm a Jordanov tvar matice Nájdite aj príslušnú maticu prechodu a zapíšte príslušnú maticovú rovnosť.

$\begin{pmatrix} -3 & 0 & 4 \\ 3 & -1 & -5 \\ -2 & 0 & 3 \end{pmatrix}$ $\begin{pmatrix} 1 & 4 & -3 \\ 0 & -3 & 1 \\ 0 & -4 & 1 \end{pmatrix}$ $\begin{pmatrix} -2 & -5 & 3 \\ 1 & 0 & -1 \\ 0 & -4 & 1 \end{pmatrix}$ $\begin{pmatrix} 2 & -1 & 0 \\ 1 & 0 & 0 \\ 2 & 2 & 1 \end{pmatrix}$ $\begin{pmatrix} 3 & -3 & 1 \\ 2 & -2 & 1 \\ 2 & -3 & 2 \end{pmatrix}$ $\begin{pmatrix} 3 & -3 & 1 \\ 2 & -2 & 1 \\ 2 & -3 & -2 \end{pmatrix}$

3.5 Aplikácie podobnosti a Jordanovho normálneho tvaru

3.5.1 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

$$A_{n+1} = aA_n + bA_{n-1}, \tag{3.9}$$

kde $a, b \in \mathbb{C}$. Je zrejmé, že ak nejaká postupnosť $(A_n)_{n=1}^{\infty}$ vyhovuje rovnici (3.9) 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 (3.9) je ekvivalentný nasledovný maticový zápis

$$\begin{pmatrix} A_{n+1} \\ A_n \end{pmatrix} = \begin{pmatrix} a & b \\ 1 & 0 \end{pmatrix} \begin{pmatrix} A_n \\ A_{n-1} \end{pmatrix} \tag{3.10}$$

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 (3.9).

Pre vlastné hodnoty $\lambda_{1,2}$ matice $A = \begin{pmatrix} a & b \\ 1 & 0 \end{pmatrix}$ teda platí

$$\begin{aligned} \lambda_1 + \lambda_2 &= a \\ \lambda_1 \cdot \lambda_2 &= -b \end{aligned} \tag{3.11}$$

Z rovnosti (3.10) dostaneme postupne

$$\begin{pmatrix} A_{n+1} \\ A_n \end{pmatrix} = A \begin{pmatrix} A_n \\ A_{n-1} \end{pmatrix} = A^2 \begin{pmatrix} A_{n-1} \\ A_{n-2} \end{pmatrix} = \cdots = A^n \begin{pmatrix} A_1 \\ A_0 \end{pmatrix} \tag{3.12}$$

Ak poznáme Jordanov tvar matice A, t.j. ak platí $A = P^{-1}JP$, tak túto rovnosť vieme prepísať ako

$$\begin{pmatrix} A_{n+1} \\ A_n \end{pmatrix} = A^n \begin{pmatrix} A_1 \\ A_0 \end{pmatrix} = P^{-1}J^nP \begin{pmatrix} A_1 \\ A_0 \end{pmatrix}.$$

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:

$$J = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} \qquad J^2 = \begin{pmatrix} \lambda^2 & 2\lambda \\ 0 & \lambda^2 \end{pmatrix} \qquad \dots \qquad J^n = \begin{pmatrix} \lambda^n & n\lambda^{n-1} \\ 0 & \lambda^n \end{pmatrix} \tag{3.13}$$

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 $\lambda^k, k\lambda^k, \binom{k}{2}\lambda^{k-1}, \dots$

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, 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

$$A_{n+k-1} = c_{n+k-2}A_{n+k-2} + \dots + c_{n-1}A_{n-1}.$$

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 $n\lambda^{n-1}, n^2\lambda^{n-2}, \dots$ (toľko z nich, koľko je násobnosť príslušného koreňa).

Viac o riešeniach 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].

Príklad 3.5.1. Nájdime vyjadrenie n-tého člena Fibonacciho postupnosti určenej predpisom

$$F_{n+1} = F_n + F_{n-1} \tag{3.14}$$

a počiatočnými hodnotami

$$F_0 = 0, F_1 = 1.$$

Maticový zápis rovnice (3.14) je

$$\begin{pmatrix} F_{n+1} \\ F_n \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} F_n \\ F_{n-1} \end{pmatrix} \tag{3.15}$$

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 (3.12) 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

$$F_n = \frac{\lambda_2^n - \lambda_1^n}{\lambda_2 - \lambda_1}, \tag{3.16}$$

Ak dosadíme $\lambda_2 = \frac{1+\sqrt{5}}{2}$ a $\lambda_1 = \frac{1-\sqrt{5}}{2}$, tak máme

$$F_n = \frac{\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n}{\sqrt{5}}. \tag{3.17}$$

Vzorec (3.17) (resp. (3.16)) sa volá *Cauchy-Binetova formula.*

Maticové rovnosti (3.10) a (3.15) 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 3.5.2. Využijeme rovnosti

$$\begin{aligned}
\begin{pmatrix} F_{n+1} \\ F_n \end{pmatrix} &= A^{n-1} \begin{pmatrix} F_2 \\ F_1 \end{pmatrix}, \\
\begin{pmatrix} F_n \\ F_{n-1} \end{pmatrix} &= A^{n-2} \begin{pmatrix} F_2 \\ F_1 \end{pmatrix}, \\
&\vdots \\
\begin{pmatrix} F_2 \\ F_1 \end{pmatrix} &= I \begin{pmatrix} F_2 \\ F_1 \end{pmatrix}.
\end{aligned}$$

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 3.2.14 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

$$\begin{pmatrix} F_2 + \cdots + F_{n+1} \\ F_1 + \cdots + F_n \end{pmatrix} = (A-I)^{-1}(A^n - I) \begin{pmatrix} F_2 \\ F_1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} F_{n+2} - 1 \\ F_{n+1} - 1 \end{pmatrix} = \begin{pmatrix} F_{n+3} - 1 \\ F_{n+2} - 1 \end{pmatrix}.$$

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 schopný 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 3.5.1. 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 (3.17) pre n-tý člen alebo inými spôsobmi (generujúce funkcie, rôzne metódy na výpočet súm, ...).

Úloha 3.5.2. Ukážte, že pre maticu $A = \left(\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ a pre $n \in \mathbb{N}$ platí $A^n = \left(\begin{smallmatrix} F_{n+1} & F_n \\ F_n & F_{n-1} \end{smallmatrix}\right)$.

Na základe toho ukážte, že:
a) $F_{n+1}F_{n-1} - F_n^2 = (-1)^n$ (Cassiniho identita)
b) $F_{m+n} = F_m F_{n+1} + F_{m-1}F_n$ (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 3.5.3. 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 3.5.4. Ukážte, že $\sum_{k=0}^{n} F_k^2 = F_n F_{n+1}$. (Hint k maticovému odvodeniu: Čomu sa rovná $(A^k)^2$? Iná možnosť: Použiť nejako úlohu 3.5.2c).)

Úloha 3.5.5. 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 3.5.6. 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 3.5.7. 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 3.5.8. 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 3.5.9. 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?

3.5.2 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

$$\begin{aligned} x' &= ax + by \\ y' &= cx + dy \end{aligned} \tag{3.18}$$

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 3.5.3. 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

$$\begin{aligned} x' + y' &= 3x + 3y \\ x' - y' &= -x + y \end{aligned}$$

čiže

$$\begin{aligned} (x + y)' &= 3(x + y) \\ (x - y)' &= -(x - y) \end{aligned}$$

Z týchto 2 rovníc máme

$$\begin{aligned} x + y &= c_1 e^{3t} \\ x - y &= c_2 e^{-t} \end{aligned}$$

a po vyjadrení x a y dostaneme riešenie

$$\begin{aligned} x &= a_1 e^{3t} + a_2 e^{-t} \\ y &= a_1 e^{3t} - a_2 e^{-t} \end{aligned}$$

(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 (3.18) zapísať maticovo ako

$$(x', y') = (x, y)A \tag{3.19}$$

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

$$\begin{aligned} u' &= \lambda_1 u \\ v' &= \lambda_2 v \end{aligned}$$

a jej riešenie je

$$\begin{aligned} u &= c_1 e^{\lambda_1 t} \\ v &= c_2 e^{\lambda_2 t} \end{aligned}$$

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].

3.6 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 súčasnoti 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

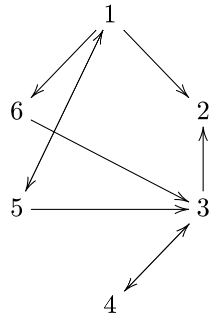
$$\vec{x} = \vec{e}A,$$

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

$$A = \begin{pmatrix} 0 & \frac{1}{3} & 0 & 0 & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}.$$

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

$$\begin{aligned} \vec{x}_0 &= \vec{e} \\ \vec{x}_1 &= \vec{x}_0 A \\ \vec{x}_2 &= \vec{x}_1 A = \vec{x}_0 A^2 \\ &\dots \\ \vec{x}_n &= \vec{x}_{n-1} A = \vec{x}_0 A^n \end{aligned}$$

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 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 3.6.1. Matica A sa nazýva *riadkovo stochastická,* ak súčet prvkov jej ľubovoľného riadku je 1.

Tvrdenie 3.6.2. *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 3.6.3. *Nech A je riadkovo stochastická matica, $\vec{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ a $\vec{y} = (y_1, \dots, y_n) = \vec{x}A$. Potom platí $\sum\limits_{i=1}^{n} y_i = \sum\limits_{i=1}^{n} x_i$.*

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. □

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 súradnic 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 3.6.2 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.

Ukážeme, že vlastný podpriestor prislúchajúci k vlastnej hodnote 1 bude jednorozmerný, v prípade, že matica A mala všetky prvky kladné.

[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 $\vec{x}A - \vec{y}A = (\vec{x} - \vec{y})A < n\|A\|_{max}\frac{\varepsilon}{n\|A\|_{max}} = \varepsilon$.

Lema 3.6.4. *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 3.6.5. *Nech matica A je riadkovo stochastická a $a_{ij} > 0$ 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

$$|\lambda| \sum_{i=1}^{n} |x_i| \leq \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ji} |x_j| = \sum_{j=1}^{n} |x_j| \sum_{i=1}^{n} a_{ji} = \sum_{j=1}^{n} |x_j|$$

$$|\lambda| \leq 1$$

□

Lema 3.6.6. *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 $\vec{\alpha} = (a_1, \dots, a_n)$ 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 3.6.7. *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 3.6.8. *Ak je matica A typu $n \times n$ podobná s diagonálnou maticou* $\mathrm{diag}(\lambda_1, \dots, \lambda_n)$, *tak existujú matice $G_1, \dots, G_n$ také, že platí*

$$A = \lambda_1 G_1 + \lambda_2 G_2 + \dots + \lambda_n G_n$$

a súčasne

$$G_1 + \dots + G_n = I,$$

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

$$A^k = \lambda_1^k G_1 + \lambda_2^k G_2 + \dots + \lambda_n^k G_n.$$

Dôkaz. Podľa predpokladu existuje regulárna matica P taká, že $PAP^{-1} = D$, čiže $P^{-1}DP = A$. Označme stĺpce matice P^{-1} ako $\vec{\alpha}_1, \dots, \vec{\alpha}_n$ a riadky matice P ako $\vec{\beta}_1, \dots, \vec{\beta}_n$.

$$P^{-1} = (\vec{\alpha}_1^T, \dots, \vec{\alpha}_n^T) \qquad P = \begin{pmatrix} \vec{\beta}_1 \\ \vdots \\ \vec{\beta}_n \end{pmatrix}$$

Z rovnosti $P^{-1}DP = A$ potom dostaneme (podobným spôsobom ako v dôkaze (3.5))

$$A = (\vec{\alpha}_1^T, \dots, \vec{\alpha}_n^T) \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \begin{pmatrix} \vec{\beta}_1 \\ \vdots \\ \vec{\beta}_n \end{pmatrix} = \lambda_1 \vec{\alpha}_1^T \vec{\beta}_1 + \dots + \lambda_n \vec{\alpha}_n^T \vec{\beta}_n.$$

Označme $G_i = \vec{\alpha}_i^T \vec{\beta}_i$. Takto definované G_i sú matice $n \times n$ a platí pre ne $A = \lambda_1 G_1 + \lambda_2 G_2 + \dots + \lambda_n G_n$.

Overme, že platia aj ostatné rovnosti uvedené vo vete. Rovnosť $I = P^{-1}P$ môžeme prepísať ako

$$I = (\vec{\alpha}_1^T, \dots, \vec{\alpha}_n^T) \begin{pmatrix} \vec{\beta}_1 \\ \vdots \\ \vec{\beta}_n \end{pmatrix} = \vec{\alpha}_1^T \vec{\beta}_1 + \dots + \vec{\alpha}_n^T \vec{\beta}_n = G_1 + \dots + G_n.$$

Ak násobíme PP^{-1}, tak dostaneme

$$I = \begin{pmatrix} \vec{\beta}_1 \\ \vdots \\ \vec{\beta}_n \end{pmatrix} (\vec{\alpha}_1^T, \dots, \vec{\alpha}_n^T) = \begin{pmatrix} \vec{\beta}_1 \vec{\alpha}_1^T & \dots & \vec{\beta}_1 \vec{\alpha}_n^T \\ \dots & \dots & \dots \\ \vec{\beta}_n \vec{\alpha}_1^T & \dots & \vec{\beta}_n \vec{\alpha}_n^T \end{pmatrix}$$

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

$$G_i G_j = \vec{\alpha}_i^T (\vec{\beta}_i \vec{\alpha}_j^T) \vec{\beta}_j = \vec{\alpha}_i^T .0. \vec{\beta}_j = 0.$$

□

Z vety o spektrálnom rozklade dostaneme, v prípade, že naša matica je diagonalizovateľná,

$$A^k = G_1 + \lambda_2^k G_2 + \dots + \lambda_n^k G_n$$

a

$$\vec{x_0} A^k = \vec{x_0} G_1 + \lambda_2^k \vec{x_0} G_2 + \dots + \lambda_n^k \vec{x_0} G_n$$

Z nasledujúceho tvrdenia vyplynie, že vlastné hodnoty $\lambda_2, \dots, \lambda_n$ 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 3.6.9. *Nech A je riadkovo stochastická matica a jej vlastné čísla sú $1, \lambda_2, \dots, \lambda_n$. Potom matica*

$$G = \alpha A + (1-\alpha)\frac{1}{n}\vec{e}^T\vec{e}$$

má vlastné hodnoty $1, \alpha\lambda_2, \dots, \alpha\lambda_n$.

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

$$Q^{-1}Q = \begin{pmatrix} \vec{y} \\ Y \end{pmatrix} (\vec{e}^T, X) = \begin{pmatrix} \vec{y}\vec{e}^T & \vec{y}X \\ Y\vec{e}^T & YX \end{pmatrix} = \begin{pmatrix} 1 & \vec{0} \\ \vec{0}^T & I \end{pmatrix}. \tag{3.20}$$

(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 (3.20) 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 $\lambda_2, \dots, \lambda_n$. (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

$$\begin{aligned} Q^{-1}GQ &= \alpha Q^{-1}AQ + (1-\alpha)\frac{1}{n}Q^{-1}\vec{e}^T\vec{e}Q = \\ &\alpha \begin{pmatrix} 1 & \vec{y}AX \\ \vec{0}^T & YAX \end{pmatrix} + (1-\alpha)\begin{pmatrix} 1 & \frac{1}{n}\vec{e}X \\ \vec{0}^T & 0 \end{pmatrix} = \begin{pmatrix} 1 & \alpha\vec{y}AX + (1-\alpha)\frac{1}{n}\vec{e}X \\ \vec{0}^T & \alpha YAX \end{pmatrix} \end{aligned}$$

Matica αYAX je podobná s hornou trojuholníkovou maticou, ktorá má na diagonále prvky $\alpha\lambda_1, \dots, \alpha\lambda_n$, 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 4

Symetrické polynómy

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

$$p(A_1(x_1,\dots,x_n),\dots,A_n(x_1,\dots,x_n)) = q(A_1(x_1,\dots,x_n),\dots,A_n(x_1,\dots,x_n))$$

tak sa polynómy p a q rovnajú.

Namiesto predchádzajúceho zdĺhavého zápisu budeme písať stručnejšie $p(A_1,\dots,A_n)$, vždy sa tým myslí, že ide o polynómy v premenných $x_1,\dots,x_n$.

Najprv ukážeme, že tvrdenie platí v prípade, že jeden z polynómov je nulový, t.j. platí

$$t(A_1,\dots,A_n) = 0 \Rightarrow t(x_1,\dots,x_n) = 0. \tag{4.1}$$

(Obidve rovnosti v predchádzajúcom riadku chápeme ako rovnosti polynómov – čiže musia sa zhodovať príslušné koeficienty.)

Postupujme sporom. Predpokladajme, že v polynóme $t(x_1,\dots,x_n)$ zapísanom v normálnom tvare máme nejaké nenulové členy. Pripomeňme, že z nenulového členu $ax_1^{k_1}\dots x_n^{k_n}$ takto dostaneme polynóm $aA_1^{k_1}\dots A_n^{k_n}$, ktorý má podľa lemy 5.8.3 vedúci člen $ax_1^{k_1+k_2\dots+k_n}x_2^{k_2\dots+k_n}\dots x_n^{k_n}$.

Nie je ťažké overiť, že priradenie $(k_1,\dots,k_n)\mapsto(k_1+k_2\dots+k_n,k_2\dots+k_n,k_n)$ je injektívne. To znamená, že člen s rovnakým systémom exponentov nemôžeme dostať ako vedúci člen z niektorého z členov polynómu $t(x_1,\dots,x_n)$.

Máme však dostať $t(A_1,\dots,A_n)=0$, to znamená, že člen $ax_1^{k_1+k_2\dots+k_n}x_2^{k_2\dots+k_n}\dots x_n^{k_n}$ sa musí niekde vykrátiť. Z predchádzajúceho vyplýva, že sa môže vykrátiť len s členom polynómu $bA_1^{l_1}\dots B_n^{l_n}$, 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 (4.1), stačí ju použiť pre rozdiel polynómov $p-q$ a dostaneme dokazované tvrdenie. □

Cvičenia

Úloha 4.0.1. Zapíšte daný symetrický polynóm pomocou základných symetrických polynómov:
a) $f(x_1, x_2, x_3) = \sum x_1^3 x_2 = x_1^3 x_2 + x_1 x_2^3 + x_1^3 x_3 + x_1 x_3^3 + x_2^3 x_3 + x_2 x_3^3$
b) $f(x_1, x_2, x_3) = \sum x_1^2 x_2 = x_1^2 x_2 + x_1 x_2^2 + x_1^2 x_3 + x_1 x_3^2 + x_2^2 x_3 + x_2 x_3^2$
[Výsledky: a) $f = A_1^2 A_2 - 2A_2^2 - A_1 A_3$; b) $f = A_1 A_2 - 3A_3$]

Úloha 4.0.2. Ukážte, že pre symetrické polynómy v n premenných, kde $n \geq 3$, platí $\sum x_1^2 x_2 = A_1 A_2 - 3A_3$.

Literatúra

[A] Tom M. Apostol. *Calculus II.* John Wiley and Sons, New York, 1969.

[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`.

[BM] Garrett Birkhoff and Saunders MacLane. *Prehľad modernej algebry.* Alfa, Bratislava, 1979.

[CFR] Paul Cull, Mary Flahive, and Robby Robson. *Difference Equations - From Rabbits to Chaos.* Springer, New York, 2005. Undergraduate Texts in Mathematics.

[FS] D. K. Faddeev and I. C. Sominskii. *Zadači po vyššej algebre.* Laň, St. Peterburg, 1999.

[G] 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.

[J] B. Johnson. Fibonacci numbers and matrices. `http://www.dur.ac.uk/bob.johnson/fibonacci/`.

[K] Thomas Koshy. *Discrete mathematics with applications.* Elsevier Academic Press, Burlington–San Diego–London, 2004.

[KGGS] Tibor Katriňák, Martin Gavalec, Eva Gedeonová, and Jaroslav Smítal. *Algebra a teoretická aritmetika 1.* UK, Bratislava, 2002.

[LM] Amy N. Langville and Carl D. Meyer. *Google's PageRank and Beyond: The Science of Search Engine Rankings.* Princeton University Press, Princeton, 2006.

[Me] Carl D. Meyer. *Matrix Analysis and Applied Linear Algebra.* SIAM, 2000.

[Mi] Ondrej Mikuláš. PageRank algoritmus, 2010. bakalárska práca, FMFI UK, Bratislava.

[Pro1] I. V. Proskurjakov. *Sbornik zadač po lineinoi algebre.* Moskva, 1966.

[Pro2] Murray H. Protter. *Basic Elements of Real Analysis.* Springer-Verlag, NY, 1998. Undergraduate Texts in Mathematics.

[Š] Beáta Štupáková. Fibonacciho a Lucasove čísla, 2008. bakalárska práca, FMFI UK, Bratislava.

[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`.

[Z] Pavol Zlatoš. Lineárna algebra a geometria. `http://thales.doa.fmph.uniba.sk/zlatos/`.

Register

Zoznam symbolov