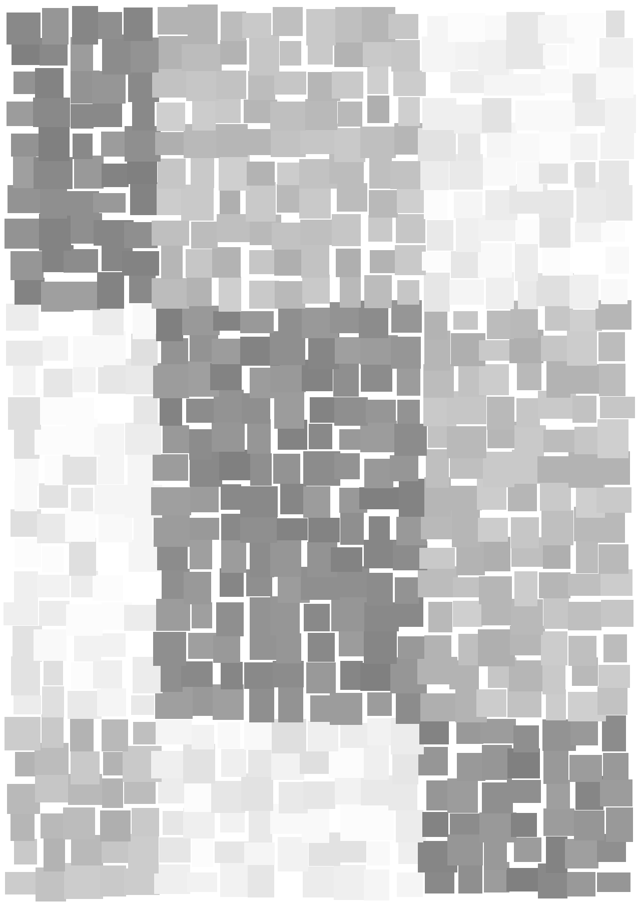
Maticová algebra pre štatistiku a dátovú vedu



Samuel Rosa

Radoslav Harman

Autori: Samuel Rosa, Radoslav Harman
Názov: Maticová algebra pre štatistiku a dátovú vedu
Vydavateľ: Knižničné a edičné centrum FMFI UK
Rok vydania: 2022
Miesto vydania: Bratislava
Vydanie: prvé
Počet strán: 131
ISBN 978-80-8147-109-4

Obsah

Prehľad značenia

$\mathbb{R}, \mathbb{N}, \mathbb{C}$	množina reálnych / prirodzených / komplexných čísel
$\mathbb{R}^n$	množina n-rozmerných stĺpcových vektorov
$\mathbb{R}^{m \times n}$	množina $m \times n$ matíc
$\mathcal{S}^n, \mathcal{S}^n_+, \mathcal{S}^n_{++}$	množina $n \times n$ symetrických / pozitívne semidefinitných / pozitívne definitných matíc
$a, \alpha, b, \beta, \ldots$	číslo
$\vec{x}, \vec{y}, \ldots$	prvok všeobecného vektorového priestoru
$\mathbf{x}, \mathbf{y}, \ldots$	stĺpcový vektor
$\mathbf{A}, \mathbf{B}, \ldots$	matica
x_i	i-ty prvok vektora $\mathbf{x}$
$(\mathbf{A})_{ij}$	(i, j)-ty prvok matice $\mathbf{A}$
$\mathbf{A}^T$	transpozícia matice $\mathbf{A}$
$\mathbf{A}^{-1}, \mathbf{A}^{-}, \mathbf{A}^{+}$	inverzia / zovšeobecnená inverzia / Mooreova-Penroseova pseudoinverzia matice $\mathbf{A}$
$\mathbf{P}_{\mathbf{A}}$	projekčná matica na $\mathcal{C}(\mathbf{A})$
$\mathcal{C}(\mathbf{A}), \mathcal{N}(\mathbf{A})$	stĺpcový / nulový priestor matice $\mathbf{A}$
$\mathrm{rank}(\mathbf{A})$	hodnosť matice $\mathbf{A}$
$\mathrm{tr}(\mathbf{A})$	stopa matice $\mathbf{A}$
$\det(\mathbf{A})$	determinant matice $\mathbf{A}$
$\kappa(\mathbf{A})$	číslo podmienenosti matice $\mathbf{A}$
$\mathbf{A} \otimes \mathbf{B}$, $\mathbf{A} \odot \mathbf{B}$	Kroneckerov / Hadamardov súčin matíc $\mathbf{A}$ a $\mathbf{B}$
$\dim(V)$	dimenzia vektorového (pod)priestoru V
$\langle\, , \rangle$	skalárny súčin
$\lVert . \rVert$	norma
$\lVert . \rVert_F$, $\lVert . \rVert_N$, $\lVert . \rVert_p$	Frobeniova / nukleárna / $\ell^{(p)}$-norma
$\mathrm{diag}(d_1, \ldots, d_n)$	diagonálna matica s prvkami $d_1, \ldots, d_n$ na diagonále
$\mathrm{diag}(\mathbf{A}_1, \ldots, \mathbf{A}_k)$	blokovo diagonálna matica s maticami $\mathbf{A}_1, \ldots, \mathbf{A}_k$ na diagonále
$\mathbf{1}_n, \mathbf{1}, \mathbf{0}_n, \mathbf{0}$	stĺpcový vektor jednotiek / núl (rozmeru $n \times 1$)
$\mathbf{e}_i$	stĺpcový vektor s jednotkou na i-tom mieste a nulami všade inde
$\mathbf{I}_n, \mathbf{I}$	matica identity (rozmeru $n \times n$)
$\mathbf{J}_{m \times n}, \mathbf{J}, \mathbf{0}_{m \times n}, \mathbf{0}$	matica jednotiek / núl (rozmeru $m \times n$)

Slovensko-anglický slovník

Pre lepšiu orientáciu v anglickej terminológii uvádzame slovník základných pojmov používaných v skriptách. Poznamenajme, že v anglickej odbornej literatúre sa môžeme stretnúť aj s odišnými prekladmi niektorých uvedených pojmov.

báza	basis
číslo podmienenosti	condition number
determinant	determinant
hodnosť	rank
– plná (riadková/stĺpcová) h.	– full (row/column) rank
kvadratická forma	quadratic form
lineárne nezávislý	linearly independent
lineárny obal	span
matica	matrix
– bloková	– block
– hustá	– dense
– idempotentná	– idempotent
– inverzná	– inverse
– ľavá/pravá inverzná	– left/right inverse
– ortogonálna	– orthogonal
– pozitívne definitná	– positive definite
– pozitívne semidefinitná	– non-negative definite, positive semidefinite
– pseudoinverzná	– pseudoinverse
– regulárna	– non-singular
– riedka	– sparse
– singulárna	– singular
– štvorcová	– square
– zovšeobecnená inverzná	– generalized inverse
maticový rozklad	matrix decomposition
násobnosť (vlastného čísla)	multiplicity (of an eigenvalue)
norma	norm
– indukovaná	– induced
– nukleárna n.	– nuclear norm, trace norm
nulový priestor	null space
problém najmenších štvorcov	least squares problem
Schurov doplnok	Schur complement

singulárne číslo	singular value
singulárny vektor	singular vector
singulárny rozklad	singular value decomposition (SVD)
– kompaktný s. r.	– compact SVD
skalárny súčin	dot product (pre $\mathbf{x}^T\mathbf{y}$), scalar product (pre všeobecný skal. s.)
spektrálny rozklad	spectral decomposition, eigendecomposition
stĺpcový priestor	column space
stopa	trace
vektor	vector
vlastné číslo	eigenvalue
vlastný vektor	eigenvector

Predslov

V matematike, na ktorej sú postavené metódy dátovej analýzy, hrajú matice kľúčovú rolu – od prípravy, uskladnenia, transformácie a vizualizácie dát, cez tvorbu štatistických modelov a riešenie optimalizačných úloh potrebných pre ich aplikovanie, až po predikciu. Pri svojej práci teda prichádza štatistik alebo dátový analytik neustále vedome či nevedome do kontaktu s maticami.

Môže sa zdať, že práve vďaka rôznym výpočtovým nástrojom, ktoré sú bežne k dispozícii (napr. R, Python, MATLAB, SAS), už nie je potrebné rozumieť princípom, na ktorých sú metódy štatistiky a dátovej analytiky postavené – a teda s maticami a matematikou všeobecne stačí spomínaný nevedomý kontakt. Nie je to však tak. Stačí, aby sme sa stretli s menej štandardnou, "neučebnicovou" úlohou a už musíme samostatne kreatívne pracovať s využitím hlbších princípov, ktoré sú často z oblasti maticovej algebry. Aj ak máme k dispozícii vhodný softvér, je nevyhnutné vedieť vybrať správnu metódu z obvykle veľkej ponuky, poznať limity zvolenej metódy a vedieť správne interpretovať výsledky. Pri všetkých týchto úlohách sa ťažko zaobídeme bez dostatočného porozumenia ako metódy fungujú, a teda bez matematiky a špeciálne maticovej algebry, na ktorej sú postavené.

Tieto skriptá sú určené pre študentov štatistiky a dátovej vedy, aby sa oboznámili s maticovou algebrou využívanou v moderných metódach spracovania dát, s dôrazom na pochopenie matematických princípov. Priebežne tiež načrtneme niektoré z týchto metód; priblížime napríklad pojmy a techniky ako kovariančná matica (základná charakteristika mnohorozmerných dát), metóda hlavných komponentov (redukcia dimenzie a vizualizácia dát), viacrozmerné normálne rozdelenie a lineárna regresia (štatistické modelovanie). Ukážeme tiež použitie stochastických matíc a vlastných čísel v Markovovských reťazcoch a cez číslo podmienenosti sa stručne dotkneme aj numerickej stability počítačových výpočtov.

Skriptá sú koncipované spôsobom veta-dôkaz-dôsledok, čiže sú vhodné skôr ako materiál k prednáškam, či ako referenčný text, nie ako učebnica určená na samoštúdium. Všetky tvrdenia sú očíslované a počínajúc kapitolou 2 aj dokázané (prípadne s odkazom na dôkaz v literatúre), skriptá teda môžu slúžiť ako referencia v slovenskom jazyku pre prácu s maticami. V zahraničnej odbornej literatúre môže čitateľ nájsť veľké množstvo textov o maticovej algebre, napríklad [2], [4], [5], [11], [13] a [15]. Numerickým algoritmom realizujúcim maticové výpočty sa v texte nevenujeme, tie sú obsahom numerickej lineárnej algebry – poznamenajme len, že sú bežnou súčasťou softvérov na prácu s dátami. Čitateľa zaujímajúceho sa o numerickú lineárnu algebru môžeme odkázať na monografiu [3].

Predpokladáme, že čitateľ je oboznámený so základmi lineárnej algebry a geometrie; tieto poznatky sú bez dôkazov zhrnuté v kapitole 1. V ďalších kapitolách (2 až 4) sa venujeme základom maticovej algebry, ktoré bývajú v rôznej miere pokryté v kurzoch lineárnej algebry a geometrie. Špeciálny dôraz kladieme na maticové rozklady a nulové a stĺpcové priestory, keďže tie umožňujú veľmi dobrý vhľad do práce s maticami. Jadro skrípt tvoria kapitoly 5 až 11, kde prejdeme od priestoru matíc a zovšeobecnených inverzných matíc cez projekčné matice až k vlastným číslam (spektrálny rozklad) a singulárnym číslam (singulárny rozklad).

V Bratislave, január 2022

Samuel Rosa; samuel.rosa@fmph.uniba.sk
Radoslav Harman; harman@fmph.uniba.sk
Fakulta matematiky, fyziky a informatiky
Univerzita Komenského v Bratislave

Kapitola 1

Základné pojmy lineárnej algebry

V tejto kapitole uvádzame prehľad základných poznatkov z lineárnej algebry, ktoré by mal čitateľ poznať pre pochopenie ďalších častí textu. Dôkladnejšie spracovanie základov lineárnej algebry a dôkazy tu spomínaných tvrdení môže čitateľ nájsť v učebnici [7]. V anglickom jazyku sú základy lineárnej algebry a geometrie spracované v množstve učebníc, napríklad v [8] a [14].

1.1 Vektorové priestory

Poznámka 1.1. V celých skriptách sa budeme zaoberať výlučne vektorovými priestormi definovanými nad poľom $\mathbb{R}$ reálnych čísel. Reálne čísla budeme obvykle značiť symbolmi a, b, c a podobne. Prvky všeobecného vektorového priestoru V budeme obvykle značiť $\vec{x}$, $\vec{y}$, $\vec{z}$ a podobne.

Definícia 1.2 (Vektorový priestor)**.** Množina V s operáciami sčítanie $+ : V \times V \to V$, $(\vec{x}, \vec{y}) \mapsto \vec{x} + \vec{y}$ a násobenie skalárom $\cdot : \mathbb{R} \times V \to V$, $(a, \vec{x}) \mapsto a \cdot \vec{x}$ sa nazýva *vektorovým priestorom*, ak pre každé $\vec{x}, \vec{y}, \vec{z} \in V$ a pre každé $a, b \in \mathbb{R}$ platí

1. $\vec{x} + (\vec{y} + \vec{z}) = (\vec{x} + \vec{y}) + \vec{z}$ (*asociatívnosť*),
2. $\vec{x} + \vec{y} = \vec{y} + \vec{x}$ (*komutatívnosť*),
3. existuje $\vec{0} \in V$, že $\vec{x} + \vec{0} = \vec{x}$ pre všetky $\vec{x} \in V$ (prvok $\vec{0}$ nazývame *nulový vektor*),
4. pre každé $\vec{x} \in V$ existuje $-\vec{x} \in V$, že $\vec{x} + (-\vec{x}) = \vec{0}$ (prvok $-\vec{x}$ nazývame *opačný vektor* k vektoru $\vec{x}$),
5. $a \cdot (\vec{x} + \vec{y}) = a \cdot \vec{x} + a \cdot \vec{y}$ (*distributívnosť* vzhľadom na sčítanie v priestore V),
6. $(a + b) \cdot \vec{x} = a \cdot \vec{x} + b \cdot \vec{x}$ (*distributívnosť* vzhľadom na sčítanie v poli $\mathbb{R}$),
7. $a \cdot (b \cdot \vec{x}) = (ab) \cdot \vec{x}$ (*kompatibilita násobenia* v $\mathbb{R}$ a násobenia $\cdot$) a
8. $1 \cdot \vec{x} = \vec{x}$.

Poznámka 1.3. Symbol $\cdot$ pre násobenie skalárom použitý v predchádzajúcej definícii sa kvôli jednoduchosti často vynecháva, čiže pre $a \in \mathbb{R}$ a $\vec{x} \in V$ píšeme namiesto $a \cdot \vec{x}$ len $a\vec{x}$.

Príklad 1.4. Jedným z najdôležitejších príkladov vektorového priestoru je $\mathbb{R}^n$ - množina n-rozmerných *stĺpcových* vektorov s prvkami z $\mathbb{R}$, spolu s pozložkovým sčítaním a štandardným násobením reálnym skalárom. Vektory v $\mathbb{R}^n$ značíme malými hrubými písmenami: $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$ a podobne. Pokúste sa pre tento vektorový priestor zdôvodniť platnosť všetkých vlastností z definície 1.2.

Definícia 1.5 (Vektorový podpriestor)**.** Nech V je vektorový priestor. Potom $\emptyset \neq W \subseteq V$ je *vektorový podpriestor* priestoru V, ak W je vektorový priestor (pričom sčítanie vo W a násobenie skalárom vo W sú zúžením sčítania a násobenia skalárom vo V).

Poznámka 1.6. Vektorový podpriestor priestoru V sa niekedy nazýva tiež *lineárny vektorový podpriestor* priestoru V alebo *lineárny podpriestor* priestoru V.

Veta 1.7 (Charakterizácia vektorového podpriestoru)**.** Nech V je vektorový priestor a $\emptyset \neq W \subseteq V$. Potom W je vektorový podpriestor V vtedy a len vtedy, keď W je uzavretý na sčítanie a násobenie skalárom, čiže ak pre každé $\vec{x}, \vec{y} \in W$ a $a \in \mathbb{R}$ platí $\vec{x}+\vec{y} \in W$ a $a\vec{x} \in W$.

Poznámka 1.8. Podmienky $\vec{x}+\vec{y} \in W$ a $a\vec{x} \in W$ pre každé $\vec{x}, \vec{y} \in W$ a $a \in \mathbb{R}$ sú ekvivalentné s podmienkou: $a_1\vec{x} + a_2\vec{y} \in W$ pre každé $\vec{x}, \vec{y} \in W$ a $a_1, a_2 \in \mathbb{R}$.

Príklad 1.9. Nech V je vektorový priestor. Množina pozostávajúca len z vektoru $\vec{0}$ a samotný priestor V sú vektorové podpriestory priestoru V, ktoré sa niekedy nazývajú triviálne podpriestory. Ak štandardným spôsobom stotožníme vektory vo vektorovom priestore $\mathbb{R}^2$ s bodmi v Euklidovskej rovine, tak netriviálnym vektorovým podpriestorom priestoru $\mathbb{R}^2$ budú zodpovedať všetky priamky prechádzajúce počiatkom Euklidovskej roviny. Pokúste sa formulovať analogické tvrdenie pre vektorový priestor $\mathbb{R}^3$.

Veta 1.10 (Prienik vektorových podpriestorov)**.** Nech $\{W_\alpha\}_{\alpha \in I}$ je neprázdny systém vektorových podpriestorov priestoru V, kde I je množina indexov. Potom $\bigcap_{\alpha \in I} W_\alpha$ je tiež vektorový podpriestor priestoru V.

Poznámka 1.11. Poznamenajme, že množina indexov I vo vete 1.10 nemusí byť konečná, dokonca ani spočítateľná.

Definícia 1.12 (Najmenší vektorový podpriestor)**.** Nech V je vektorový priestor a $\emptyset \neq M \subseteq V$. *Najmenší vektorový podpriestor* priestoru V obsahujúci M je vektorový podpriestor S obsahujúci M, ktorý spĺňa: ak W je vektorový podpriestor V a $M \subseteq W$, tak $S \subseteq W$.

Veta 1.13 (Najmenší vektorový podpriestor)**.** Ak $\emptyset \neq M \subseteq V$, tak existuje práve jeden najmenší vektorový podpriestor priestoru V obsahujúci M. Navyše, tento podpriestor je rovný prieniku všetkých podpriestorov priestoru V, ktoré obsahujú M.

Definícia 1.14 (Lineárna kombinácia vektorov, lineárny obal množiny vektorov)**.** Nech V je vektorový priestor, $\vec{x}_1, \ldots, \vec{x}_k \in V$ a nech $a_1, \ldots, a_k \in \mathbb{R}$. Potom:

1. *Lineárna kombinácia* vektorov $\vec{x}_1, \ldots, \vec{x}_k$ s koeficientmi $a_1, \ldots, a_k$ je vektor $a_1\vec{x}_1 + \ldots + a_k\vec{x}_k$.

2. Množinu $\{a_1\vec{x}_1 + \ldots + a_k\vec{x}_k \mid a_1, \ldots, a_k \in \mathbb{R}\}$ všetkých lineárnych kombinácií vektorov $\vec{x}_1, \ldots, \vec{x}_k$ nazývame *lineárny obal* vektorov $\vec{x}_1, \ldots, \vec{x}_k$ a značíme ju $\mathrm{span}(\vec{x}_1, \ldots, \vec{x}_k)$.

3. Ak vektorový podpriestor W spĺňa $W = \text{span}(\vec{x}_1, \dots, \vec{x}_k)$, potom hovoríme, že W je *generovaný* vektormi $\vec{x}_1, \dots, \vec{x}_k$.

Veta 1.15 (Najmenší vektorový podpriestor obsahujúci množinu vektorov)**.** Nech V je vektorový priestor a nech $\{\vec{x}_1, \dots, \vec{x}_k\} \subseteq V$. Potom najmenší vektorový priestor obsahujúci $\{\vec{x}_1, \dots, \vec{x}_k\}$ je $\text{span}(\vec{x}_1, \dots, \vec{x}_k)$.

Príklad 1.16. Nech $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^2$ a nech $W = \text{span}(\mathbf{x}_1, \mathbf{x}_2)$. Potom $W = \{\mathbf{0}\}$, alebo $W = \{c\mathbf{y} \mid c \in \mathbb{R}\}$ pre nejaké $\mathbf{y} \in \mathbb{R}^2$, alebo $W = \mathbb{R}^2$. Premyslite si, aké musia byť vektory $\mathbf{x}_1, \mathbf{x}_2$, aby sme dostali každú z týchto troch možností. Pokúste sa sformulovať analogické tvrdenie pre vektorový priestor $\mathbb{R}^3$.

1.2 Lineárna nezávislosť, báza, dimenzia

Definícia 1.17 (Lineárna závislosť a lineárna nezávislosť)**.** Nech V je vektorový priestor a nech $\vec{x}_1, \dots, \vec{x}_k \in V$. Potom $\vec{x}_1, \dots, \vec{x}_k$ sú *lineárne nezávislé*, ak rovnosť $a_1\vec{x}_1 + \ldots + a_k\vec{x}_k = \vec{0}$ pre $a_1, \dots, a_k \in \mathbb{R}$ implikuje $a_1 = \ldots = a_k = 0$. Hovoríme, že $\vec{x}_1, \dots, \vec{x}_k$ sú *lineárne závislé*, ak nie sú lineárne nezávislé.

Poznámka 1.18. Ak vektory $\vec{x}_1, \dots, \vec{x}_k \in V$ sú lineárne nezávislé, tak žiaden z nich nie je nulový, teda $\vec{x}_1 \neq \vec{0}, \dots, \vec{x}_k \neq \vec{0}$.

Veta 1.19 (Charakterizácia lineárne závislých vektorov)**.** Nech V je vektorový priestor a nech $\vec{x}_1, \dots, \vec{x}_k \in V$ sú nenulové vektory. Potom $\vec{x}_1, \dots, \vec{x}_k$ sú lineárne závislé práve vtedy, keď niektorý z nich je lineárnou kombináciou tých, čo sú napísané pred ním.

Definícia 1.20 (Konečne generovaný vektorový priestor)**.** Vektorový priestor V sa nazýva *konečne generovaný*, ak existuje konečná množina $\{\vec{x}_1, \dots, \vec{x}_k\} \in V$, že $V = \text{span}(\vec{x}_1, \dots, \vec{x}_k)$.

V ďalších kapitolách sa nebudeme venovať všeobecným priestorom, obmedzíme sa na konečne generované vektorové priestory.

Definícia 1.21 (Báza priestoru)**.** Nech $V \neq \{\vec{0}\}$ je konečne generovaný vektorový priestor. Potom postupnosť vektorov $\vec{b}_1, \dots, \vec{b}_k$ sa nazýva *báza* priestoru V, ak spĺňa:

1. $V = \text{span}(\vec{b}_1, \dots, \vec{b}_k)$,
2. $\vec{b}_1, \dots, \vec{b}_k$ sú nezávislé.

Definícia 1.22 (Súradnice)**.** Nech $V \neq \{\vec{0}\}$ je konečne generovaný vektorový priestor a $\vec{b}_1, \dots, \vec{b}_k$ je jeho báza. Potom *súradnice* vektora $\vec{x} \in V$ vzhľadom na bázu $\vec{b}_1, \dots, \vec{b}_k$ sú taká postupnosť čísel $c_1, \dots, c_k \in \mathbb{R}$, že $\vec{x} = c_1\vec{b}_1 + \ldots c_k\vec{b}_k$.

Príklad 1.23. Nech $V \neq \{\vec{0}\}$ je konečne generovaný vektorový priestor, $\vec{b}_1, \dots, \vec{b}_k$ je jeho báza a $\vec{x} \in V$. Ukážte, že súradnice vektora $\vec{x}$ vzhľadom na danú bázu sú jedinečné. Teda ukážte, že ak $c_1, \dots, c_k$ aj $d_1, \dots, d_k$ sú súradnice $\vec{x}$ vzhľadom na $\vec{b}_1, \dots, \vec{b}_k$, potom $c_1 = d_1, \dots, c_k = d_k$.

Veta 1.24 (O bázach)**.** Nech $V \neq \{\vec{0}\}$ je konečne generovaný vektorový priestor. Potom

(i) V má bázu.

(ii) Všetky bázy V majú rovnaký počet prvkov.

(iii) Každá lineárne nezávislá podmnožina vektorov z V sa dá doplniť na bázu V.

Definícia 1.25 (Dimenzia priestoru)**.** Počet prvkov ľubovoľnej bázy konečne generovaného priestoru $V \neq \{\vec{0}\}$ sa nazýva *dimenzia* priestoru V, značíme ju $\dim(V)$. Dimenzia nulového priestoru $\{\vec{0}\}$ je 0 a dimenzia priestoru, ktorý nie je konečne generovaný, je ∞.

Príklad 1.26. Báza priestoru $\mathbb{R}^n$ sú napríklad vektory $\mathbf{e}_1, \dots, \mathbf{e}_n$ vyjadrujúce štandardnú súradnicovú sústavu: $\mathbf{e}_i$ je vektor s jednotkou na i-tom mieste a nulami inde. Z toho vyplýva, že dimenzia $\mathbb{R}^n$ je n. Súradnice vektora $\mathbf{x} \in \mathbb{R}^n$ vzhľadom na štandardnú súradnicovú sústavu sú potom prvky vektora $\mathbf{x}$. Pre $n = 2$ a $n = 3$ nájdite nejakú inú bázu priestoru $\mathbb{R}^n$.

Veta 1.27 (O dimenziách)**.** Nech V je vektorový priestor a V_1, V_2 sú jeho vektorové podpriestory. Potom platí:

(i) Ak $V = \mathrm{span}(\vec{x}_1, \dots, \vec{x}_k)$, potom $\dim(V) \leq k$.

(ii) Ak $V_1 \subseteq V_2$, tak $\dim(V_1) \leq \dim(V_2)$.

(iii) Ak $V_1 \subseteq V_2$ a $\dim(V_1) = \dim(V_2)$, tak $V_1 = V_2$.

Veta 1.28 (Vlastnosti priestoru s dimenziou n)**.** Nech $\dim(V) = n$. Potom:

(i) Každých $n + 1$ vektorov z V je lineárne závislých.

(ii) $\vec{x}_1, \dots, \vec{x}_n \in V$ tvoria bázu V práve vtedy, keď $\vec{x}_1, \dots, \vec{x}_n$ sú lineárne nezávislé.

(iii) $\vec{x}_1, \dots, \vec{x}_n \in V$ tvoria bázu V práve vtedy, keď $V = \mathrm{span}(\vec{x}_1, \dots, \vec{x}_n)$.

1.3 Skalárny súčin, norma a uhol vektorov

V tejto podkapitole V vždy značí uvažovaný vektorový priestor.

Definícia 1.29 (Skalárny súčin)**.** *Skalárny súčin* na priestore V je zobrazenie $\langle\, , \rangle : V \times V \to \mathbb{R}$, ktoré spĺňa pre každé $\vec{x}, \vec{y}, \vec{z} \in V$ a $a \in \mathbb{R}$:

1. $\langle \vec{y}, \vec{x} \rangle = \langle \vec{x}, \vec{y} \rangle$ (*symetrickosť*),
2. $\langle \vec{x} + \vec{y}, \vec{z} \rangle = \langle \vec{x}, \vec{z} \rangle + \langle \vec{y}, \vec{z} \rangle$ (*distributívnosť* vzhľadom na súčet v prvej zložke),
3. $\langle a\vec{x}, \vec{y} \rangle = a\langle \vec{x}, \vec{y} \rangle$ (*homogénnosť* v prvej zložke),
4. $\langle \vec{x}, \vec{x} \rangle \geq 0$, pričom $\langle \vec{x}, \vec{x} \rangle = 0$ práve vtedy, keď $\vec{x} = \vec{0}$ (*kladná definitnosť*).

Poznámka 1.30. Vďaka prvej vlastnosti skalárneho súčinu (symetrickosti) vyplýva z vlastností 2 a 3, že skalárny súčin je distributívny a homogénny v každej zložke. Platí teda aj $\langle \vec{x}, \vec{y} + \vec{z} \rangle = \langle \vec{x}, \vec{y} \rangle + \langle \vec{x}, \vec{z} \rangle$ a $\langle \vec{x}, a\vec{y} \rangle = a\langle \vec{x}, \vec{y} \rangle$.

Príklad 1.31. Medzi skalárne súčiny patrí napríklad bežne používaný $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T\mathbf{y} = \sum_{i=1}^{n} x_i y_i$ na $\mathbb{R}^n$, kde x_i sú prvky vektora $\mathbf{x}$ a y_i sú prvky vektora $\mathbf{y}$. Overte, že tento skalárny súčin skutočne spĺňa vlastnosti 1-4.

Definícia 1.32 (Norma)**.** *Norma* na priestore V je zobrazenie $\|.\| : V \to \mathbb{R}$, ktoré spĺňa pre každé $\vec{x}$, $\vec{y}$ a $a \in \mathbb{R}$:

1. $\|a\vec{x}\| = |a| \cdot \|\vec{x}\|$ (*homogénnosť*),
2. $\|\vec{x}\| \geq 0$, pričom $\|\vec{x}\| = 0$ práve vtedy, keď $\vec{x} = \vec{0}$ (*kladná definitnosť*),
3. $\|\vec{x} + \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\|$ (*trojuholníková nerovnosť*).

Veta 1.33 (Norma prislúchajúca skalárnemu súčinu)**.** Nech $\langle\, , \rangle$ je skalárny súčin. Potom zobrazenie $\|.\|$ definované pre každé $\vec{x} \in V$ ako $\|\vec{x}\| = \langle \vec{x}, \vec{x} \rangle^{1/2}$ je norma.

Veta 1.34 (Cauchyho-Schwarzova nerovnosť)**.** Nech $\langle\, , \rangle$ je skalárny súčin a $\| \,.\, \|$ je ním definovaná norma. Potom pre každé $\vec{x}, \vec{y} \in V$ platí

$$|\langle \vec{x}, \vec{y} \rangle| \leq \|\vec{x}\| \cdot \|\vec{y}\|.$$

Príklad 1.35. Norma na $\mathbb{R}^n$ prislúchajúca skalárnemu súčinu $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T \mathbf{y}$ je euklidovská norma $\|\mathbf{x}\| = (\sum_{i=1}^n x_i^2)^{1/2}$. Overte, že spĺňa vlastnosti 1-3 z definície 1.32.

Definícia 1.36 (Uhol)**.** Nech $\langle\, , \rangle$ je skalárny súčin a nech $\vec{x}, \vec{y} \in V$. Ak $\vec{x}, \vec{y} \neq \vec{0}$, tak *uhol* medzi $\vec{x}$ a $\vec{y}$ je $\varphi \in [0, \pi]$, ktoré spĺňa

$$\cos(\varphi) = \frac{\langle \vec{x}, \vec{y} \rangle}{\|\vec{x}\| \|\vec{y}\|}. \tag{1.1}$$

Poznámka 1.37. Uhol medzi vektormi daný vzťahom (1.1) je dobre definovaný. Podľa Cauchyho-Schwarzovej nerovnosti totiž platí, že $\langle \vec{x}, \vec{y} \rangle / (\|\vec{x}\| \|\vec{y}\|) \in [-1, 1]$, pričom funkcia cos je na intervale $[0, \pi]$ klesajúcou spojitou funkciou s oborom hodnôt $[-1, 1]$. Pre každé $\vec{x}, \vec{y} \in V$ teda existuje jediné číslo $\varphi \in [0, \pi]$ spĺňajúce vzťah (1.1).

1.3.1 Ortogonálnosť

Definícia 1.38 (Ortogonálnosť a ortonormálnosť vektorov)**.** Hovoríme, že $\vec{x}$ a $\vec{y}$ sú *ortogonálne* (*kolmé*), ak $\langle \vec{x}, \vec{y} \rangle = 0$; značíme $\vec{x} \perp \vec{y}$. Množina $\{\vec{x}_1, \ldots, \vec{x}_k\} \subseteq V$ sa nazýva *ortogonálna*, ak každá dvojica $\vec{x}_i$, $\vec{x}_j$ pre $j \neq i$ je ortogonálna. Ak navyše platí $\|\vec{x}_i\| = 1$ pre $i = 1, \ldots, k$, tak táto množina sa nazýva *ortonormálna*.

Poznámka 1.39. Ak $\vec{x}, \vec{y} \neq \vec{0}$, tak $\vec{x} \perp \vec{y}$ znamená, že uhol medzi $\vec{x}$ a $\vec{y}$ je $\pi/2$.

Veta 1.40 (Ortogonálnosť a lineárna nezávislosť)**.** Nech nenulové vektory $\vec{x}_1, \ldots, \vec{x}_n$ sú navzájom ortogonálne. Potom $\vec{x}_1, \ldots, \vec{x}_n$ sú lineárne nezávislé.

Poznámka 1.41. Ak $\vec{b}_1, \ldots, \vec{b}_n$ je báza a zároveň $\{\vec{b}_1, \ldots, \vec{b}_n\}$ tvoria ortogonálnu množinu, hovoríme, že $\vec{b}_1, \ldots, \vec{b}_n$ je *ortogonálna báza*. Ak navyše $\|\vec{b}_i\| = 1$ pre $i = 1, \ldots, n$, tak $\vec{b}_1, \ldots, \vec{b}_n$ je *ortonormálna báza*.

Veta 1.42 (Gramova-Schmidtova ortogonalizácia)**.** Nech $\vec{x}_1, \dots, \vec{x}_k$ je množina lineárne nezávislých vektorov. Potom vektory $\vec{y}_1, \dots, \vec{y}_k$ definované nasledovne:

$$\begin{aligned}
\vec{y}_1 &= \vec{x}_1, \\
\vec{y}_2 &= \vec{x}_2 - \frac{\langle \vec{y}_1, \vec{x}_2 \rangle}{\langle \vec{y}_1, \vec{y}_1 \rangle} \vec{y}_1, \\
\vec{y}_3 &= \vec{x}_3 - \frac{\langle \vec{y}_1, \vec{x}_3 \rangle}{\langle \vec{y}_1, \vec{y}_1 \rangle} \vec{y}_1 - \frac{\langle \vec{y}_2, \vec{x}_3 \rangle}{\langle \vec{y}_2, \vec{y}_2 \rangle} \vec{y}_2, \\
&\vdots \\
\vec{y}_k &= \vec{x}_k - \sum_{j=1}^{k-1} \frac{\langle \vec{y}_j, \vec{x}_k \rangle}{\langle \vec{y}_j, \vec{y}_j \rangle} \vec{y}_j,
\end{aligned}$$

tvoria ortogonálnu bázu podpriestoru $\mathrm{span}(\vec{x}_1, \dots, \vec{x}_k)$. Navyše $\{\vec{z}_1, \dots, \vec{z}_k\}$, kde $\vec{z}_i = \vec{y}_i / \|\vec{y}_i\|$ pre $i = 1, \dots, k$, tvoria ortonormálnu bázu tohto podpriestoru.

Poznámka 1.43. Vektor $(\langle \vec{y}, \vec{x} \rangle / \langle \vec{y}, \vec{y} \rangle) \vec{y}$ vyjadruje kolmú (ortogonálnu) projekciu vektora $\vec{x}$ na priamku určenú vektorom $\vec{y}$. Ortogonálnym projekciám sa budeme bližšie venovať v kapitole 7. Pokúste sa pomocou tejto poznámky pochopiť Gramovu-Schmidtovu ortogonalizáciu v priestore $\mathbb{R}^3$.

Dôsledok 1.44 (Ortonormálne bázy)**.** Každý nenulový podpriestor konečne generovaného priestoru má ortonormálnu bázu.

Definícia 1.45 (Ortogonálnosť množín)**.** Nech $W_1, W_2 \subseteq V$. Potom hovoríme, že W_1 a W_2 sú navzájom *ortogonálne*, ak $\langle \vec{x}, \vec{y} \rangle = 0$ pre každé $\vec{x} \in W_1$, $\vec{y} \in W_2$; značíme $W_1 \perp W_2$.

Veta 1.46 (O ortogonálnosti podpriestorov)**.** Nech $\vec{x}_1, \dots, \vec{x}_s \in V$ a $\vec{y}_1, \dots, \vec{y}_t \in V$. Označme $W_1 = \mathrm{span}(\vec{x}_1, \dots, \vec{x}_s)$ a $W_2 = \mathrm{span}(\vec{y}_1, \dots, \vec{y}_t)$. Potom $W_1 \perp W_2$ práve vtedy, keď $\langle \vec{x}_i, \vec{y}_j \rangle = 0$ pre každé $i = 1, \dots, s$, $j = 1, \dots, t$.

Dôsledok 1.47 (Ortogonálnosť na podpriestor)**.** Nech $\vec{x}_1, \dots, \vec{x}_s \in V$ a nech $\vec{y} \in V$. Označme $W = \mathrm{span}(\vec{x}_1, \dots, \vec{x}_s)$. Potom $\vec{y} \perp W$ práve vtedy, keď $\langle \vec{x}_i, \vec{y} \rangle = 0$ pre každé $i = 1, \dots, s$.

Definícia 1.48 (Ortogonálny doplnok)**.** Nech M je množina vo vektorovom priestore V. Potom *ortogonálny doplnok* množiny M je

$$M^\perp := \{\vec{x} \in V \mid \vec{x} \perp \vec{y} \text{ pre každé } \vec{y} \in M\}.$$

Veta 1.49 (Vlastnosti ortogonálnych doplnkov)**.** Nech V je konečne generovaný vektorový priestor so skalárnym súčinom. Platí:

(i) Ak $M \subseteq V$, potom $M^\perp$ je vektorový podpriestor V.

(ii) Ak množiny M a N sú podmnožinami V a $M \subseteq N$, potom $N^\perp \subseteq M^\perp$.

(iii) Ak $W \subseteq V$ je vektorový podpriestor, tak $(W^\perp)^\perp = W$.

(iv) Ak $W \subseteq V$ je vektorový podpriestor, tak $\dim(W^\perp) = \dim(V) - \dim(W)$.

Príklad 1.50. Uvažujme nenulový vektor $\mathbf{x} \in \mathbb{R}^2$, ním určenú priamku $M_1 = \{a\mathbf{x} \mid a \in \mathbb{R}\}$, a jednoprvkovú množinu pozostávajúcu iba z tohto vektoru $M_2 = \{\mathbf{x}\}$. Nech $\mathbf{y} \in \mathbb{R}^2$ spĺňa $\mathbf{y} \perp \mathbf{x}$. Ukážte, že priamka $N = \{a\mathbf{y} \mid a \in \mathbb{R}\}$ je ortogonálnym doplnkom tak priamky M_1, ako aj množiny M_2, čiže $M_1^\perp = N$ a $M_2^\perp = N$. Čomu je rovné $N^\perp$? Poznamenajme, že každá z priamok M_1 a N tvorí podpriestor v $\mathbb{R}^2$, zatiaľ čo M_2 podpriestor netvorí.

1.4 Lineárne zobrazenia

Definícia 1.51 (Obraz a jadro)**.** Nech V a W sú vektorové priestory a nech f je zobrazenie z V do W. Potom:

1. *Obraz množiny* $M \subseteq V$ je $f(M) = \{\vec{y} \in W \mid \exists \vec{x} \in M : \vec{y} = f(\vec{x})\}$ a *vzor množiny* $N \subseteq W$ je $f^{-1}(N) = \{\vec{x} \in V \mid f(\vec{x}) \in N\}$.

2. *Obraz zobrazenia* f je $\mathrm{Im}(f) = f(V)$ a *jadro zobrazenia* je $\mathrm{Ker}(f) = f^{-1}(\{\vec{0}\})$, kde $\vec{0}$ je nulový vektor vo W.

Definícia 1.52 (Surjekcia, injekcia, bijekcia)**.** Zobrazenie $f : V \to W$ je

1. *surjektívne*, ak $f(V) = W$,

2. *injektívne*, ak z $f(\vec{x}) = f(\vec{y})$ vyplýva $\vec{x} = \vec{y}$, pre ľubovoľné $\vec{x}, \vec{y} \in V$,

3. *bijektívne*, ak je surjektívne aj injektívne.

Definícia 1.53 (Lineárne zobrazenie, transformácia)**.** Nech V a W sú vektorové priestory. Potom $f : V \to W$ je *lineárne zobrazenie* z V do W, ak pre každé $a, b \in \mathbb{R}$ a $\vec{x}, \vec{y} \in V$ platí

$$f(a\vec{x} + b\vec{y}) = af(\vec{x}) + bf(\vec{y}).$$

Ak f je lineárne zobrazenie z V do V, tak hovoríme, že f je *lineárna transformácia* priestoru V.

Príklad 1.54. Jedným z najjednoduchších príkladov lineárnej transformácie je lineárna funkcia $f : \mathbb{R} \to \mathbb{R}$ definovaná predpisom $f(x) = cx$ pre $c \in \mathbb{R}$. Overte, že táto funkcia spĺňa definíciu lineárnej transformácie. Overte tiež, že afinná funkcia $f(x) = cx + d$ pre $d \neq 0$ nie je lineárnou transformáciou.

Veta 1.55 (Obraz a vzor vektorového podpriestoru)**.** Nech $f : V \to W$ je lineárne zobrazenie. Potom:

(i) Ak S je vektorový podpriestor V, tak $f(S)$ je vektorový podpriestor W, a teda aj $\mathrm{Im}(f)$ je vektorový podpriestor W.

(ii) Ak T je vektorový podpriestor W, tak $f^{-1}(T)$ je vektorový podpriestor V, a teda aj $\mathrm{Ker}(f)$ je vektorový podpriestor V.

Veta 1.56 (Injektívne a surjektívne lineárne zobrazenie)**.** Lineárne zobrazenie $f : V \to W$ je

(i) injektívne práve vtedy, keď $\mathrm{Ker}(f) = \{\vec{0}\}$,

(ii) surjektívne práve vtedy, keď $\mathrm{Im}(f) = W$.

Definícia 1.57 (Lineárny izomorfizmus)**.** Zobrazenie $f : V \to W$ nazývame *lineárny izomorfizmus*, ak f je lineárne a bijektívne. Ak pre V, W existuje lineárny izomorfizmus z V do W, potom hovoríme, že V a W sú *lineárne izomorfné*.

Poznámka 1.58. Ak existuje lineárny izomorfizmus z V do W, tak existuje aj lineárny izomorfizmus z W do V.

Poznámka 1.59. Ak sú vektorové priestory V a W lineárne izomorfné, tak sú tieto priestory v istom zmysle také isté, napríklad majú rovnakú dimenziu. Existencia izomorfizmu f medzi V a W vyjadruje, že W vieme získať preznačením prvkov priestoru V (bijektívnosť f) tak, že zachováme štruktúru priestoru - násobenie skalárom a sčítanie (lineárnosť f).

Definícia 1.60 (Zložené zobrazenie)**.** Nech $f : V \to W$ a $g : W \to S$ sú zobrazenia. Potom znakom $g \circ f$ značíme *zložené zobrazenie* $g \circ f : V \to S$, také, že $(g \circ f)(\vec{x}) = g(f(\vec{x}))$.

Veta 1.61 (Zloženie lineárnych zobrazení)**.** Nech $f : V \to W$ a $g : W \to S$ sú lineárne zobrazenia. Potom aj $g \circ f : V \to S$ je lineárne zobrazenie. Navyše, ak f a g sú lineárne izomorfizmy, tak aj $g \circ f$ je lineárny izomorfizmus.

Veta 1.62 (Základná veta o lineárnych zobrazeniach)**.** Nech V a W sú konečne generované vektorové priestory, pričom sú dané báza $\vec{x}_1, \ldots, \vec{x}_n$ priestoru V a vektory $\vec{y}_1, \ldots, \vec{y}_n \in W$. Potom existuje jediné lineárne zobrazenie $f : V \to W$ také, že $f(\vec{x}_i) = \vec{y}_i$ pre každé $i = 1, \ldots, n$. Navyše, také f spĺňa $f(a_1\vec{x}_1 + \ldots + a_n\vec{x}_n) = a_1\vec{y}_1 + \ldots a_n\vec{y}_n$ pre ľubovoľné $a_1, \ldots, a_n \in \mathbb{R}$.

Kapitola 2

Základné pojmy maticovej algebry

2.1 Matice a maticové operácie

Definícia 2.1 (Matica a jej rozmery)**.** Nech $m, n \in \mathbb{N}$. *Matica* typu $m \times n$ je súbor reálnych čísel usporiadaných do obdĺžnikovej tabuľky s m riadkami a n stĺpcami. Matice budeme označovať veľkými hrubými písmenami $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ a podobne. Ak je $\mathbf{A}$ matica typu $m \times n$, tak čísla m a n sa nazývajú *rozmery* matice $\mathbf{A}$, niekedy tiež *dimenzie* matice $\mathbf{A}$. Množinu všetkých matíc typu $m \times n$ budeme označovať $\mathbb{R}^{m \times n}$.

Príklad 2.2. Uvažujte nasledovné matice

$$\mathbf{A} = \begin{pmatrix} 1 & -1 \\ -1 & -1 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 0{,}6 \\ 7 \\ 5{,}2 \end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix} -1 & 0 & 1 \end{pmatrix}, \quad \mathbf{D} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Potom $\mathbf{A}$ je typu 2×2 a $\mathbf{B}$ typu 3×1. Určte rozmery matíc $\mathbf{C}$ a $\mathbf{D}$.

Definícia 2.3 (Prvky matice)**.** Reálne čísla, z ktorých pozostáva matica, sa nazývajú *prvky* alebo *zložky* matice. Prvok matice $\mathbf{A}$ v i-tom riadku a j-tom stĺpci budeme označovať $(\mathbf{A})_{ij}$. Prvok $(\mathbf{A})_{ij}$ tiež označujeme ako prvok na (i, j)-tej pozícii. Ak $m, n \in \mathbb{N}$ a $a_{ij} \in \mathbb{R}$ pre $i = 1, \ldots, m$ a $j = 1, \ldots, n$, tak symbolom $\{a_{ij}\}_{i=1,\ldots,m; j=1,\ldots,n}$, alebo skrátene $\{a_{ij}\}$, označujeme maticu $\mathbf{A}$ typu $m \times n$, pre ktorú platí $(\mathbf{A})_{ij} = a_{ij}$.

Poznámka 2.4. Je možné sa stretnúť aj s maticami, ktorých prvky nemusia byť reálne čísla (napríklad matice s prvkami z $\mathbb{C}$, množiny komplexných čísel). V súlade s definíciou 2.1 však v týchto skriptách uvažujeme iba matice s reálnymi prvkami.

Poznámka 2.5. Matica je plne určená svojimi prvkami, takže pre $m \times n$ matice $\mathbf{A}$ a $\mathbf{B}$ platí $\mathbf{A} = \mathbf{B}$ vtedy a len vtedy keď $(\mathbf{A})_{ij} = (\mathbf{B})_{ij}$ pre všetky $i \in \{1, \ldots, m\}$ a $j \in \{1, \ldots, n\}$.

Poznámka 2.6. *Vektor* $\mathbf{x} \in \mathbb{R}^n$ je vlastne maticou typu $n \times 1$ a jeho prvky značíme

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Hovoríme, že $\mathbf{x}$ je vektor *rozmeru* n. *Riadkový vektor* $\mathbf{y}$ rozmeru n je maticou typu $1 \times n$, $\mathbf{y} = (y_1, \ldots, y_n)$.

Definícia 2.7 (Skalárny násobok matice, súčet matíc, rozdiel matíc, násobenie matíc)**.** Nech $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m\times n}$, $\mathbf{C} \in \mathbb{R}^{n\times k}$ a $c \in \mathbb{R}$. Potom

1. *násobok* matice $\mathbf{A}$ *skalárom* c je matica $c\mathbf{A} \in \mathbb{R}^{m\times n}$ spĺňajúca $(c\mathbf{A})_{ij} = c(\mathbf{A})_{ij}$ pre každé i, j,
2. *súčet* matíc $\mathbf{A}$ a $\mathbf{B}$ je matica $\mathbf{A}+\mathbf{B} \in \mathbb{R}^{m\times n}$ spĺňajúca $(\mathbf{A}+\mathbf{B})_{ij} = (\mathbf{A})_{ij} + (\mathbf{B})_{ij}$ pre každé i, j,
3. *rozdiel* matíc $\mathbf{A}$ a $\mathbf{B}$ je matica $\mathbf{A}-\mathbf{B} \in \mathbb{R}^{m\times n}$ spĺňajúca $\mathbf{A}-\mathbf{B} = \mathbf{A}+(-1)\mathbf{B}$, čiže $(\mathbf{A}-\mathbf{B})_{ij} = (\mathbf{A})_{ij} - (\mathbf{B})_{ij}$ pre každé i, j,
4. *súčin* matíc $\mathbf{A}$ a $\mathbf{C}$ je matica $\mathbf{AC} \in \mathbb{R}^{m\times k}$, ktorá pre každé i, j spĺňa
$$(\mathbf{AC})_{ij} = \sum_{u=1}^{n} (\mathbf{A})_{iu}(\mathbf{C})_{uj}.$$

Poznámka 2.8. Sčítať a odčítať môžeme iba matice rovnakých rozmerov (hovoríme, že také matice sú *konformné pre sčítanie*). Násobiť môžeme iba matice, kde počet stĺpcov prvej matice je rovnaký ako počet riadkov druhej matice (hovoríme, že také matice sú *konformné pre násobenie*).

Poznámka 2.9. Ak $c \in \mathbb{R}$, tak c je vlastne matica typu 1×1. Súčin matíc $\mathbf{AC}$ sa teda riadi podľa bodu 4 definície 2.7 iba ak ani jedna z matíc $\mathbf{A}$, $\mathbf{C}$ nie je typu 1×1. V opačnom prípade sa súčin matíc riadi bodom 1 tejto definície.

Veta 2.10 (Základná aritmetika matíc)**.** Pre základné operácie s maticami platí

(i) $\mathbf{A}+(\mathbf{B}+\mathbf{C}) = (\mathbf{A}+\mathbf{B})+\mathbf{C}$ (*asociatívnosť* súčtu),

(ii) $\mathbf{A}+\mathbf{B} = \mathbf{B}+\mathbf{A}$ (*komutatívnosť* súčtu)

(iii) $c(\mathbf{A}+\mathbf{B}) = c\mathbf{A}+c\mathbf{B}$ (*distributívnosť* násobenia skalárom vzhľadom na sčítanie),

(iv) $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$ (*asociatívnosť* súčinu),

(v) $\mathbf{A}(\mathbf{B}+\mathbf{C}) = \mathbf{AB}+\mathbf{AC}$ (*distributívnosť* násobenia zľava vzhľadom na sčítanie),

(vi) $(\mathbf{A}+\mathbf{B})\mathbf{C} = \mathbf{AC}+\mathbf{BC}$ (*distributívnosť* násobenia sprava vzhľadom na sčítanie),

pričom $\mathbf{A}, \mathbf{B}, \mathbf{C}$ sú v jednotlivých častiach matice vhodných rozmerov, aby boli konformné na sčítanie, resp. násobenie.

Dôkaz. Na dôkaz vlastnosti (v) predpokladajme, že $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{B}, \mathbf{C} \in \mathbb{R}^{n\times k}$. Potom

$$\begin{aligned}\Big(\mathbf{A}(\mathbf{B}+\mathbf{C})\Big)_{ij} &= \sum_{u=1}^{n} \mathbf{A}_{iu}(\mathbf{B}+\mathbf{C})_{uj} = \sum_{u=1}^{n} \Big((\mathbf{A})_{iu}(\mathbf{B})_{uj} + (\mathbf{A})_{iu}(\mathbf{C})_{uj}\Big) \\ &= \sum_{u=1}^{n}(\mathbf{A})_{iu}(\mathbf{B})_{uj} + \sum_{u=1}^{n}(\mathbf{A})_{iu}(\mathbf{C})_{uj} = (\mathbf{AB})_{ij} + (\mathbf{AC})_{ij}.\end{aligned}$$

Dôkaz ostatných vlastností nechávame ako cvičenie. $\square$

Poznámka 2.11. Vďaka asociatívnosti súčtu matíc môžeme súčet viacerých matíc značiť bez zátvoriek, napr. $\mathbf{A}+\mathbf{B}+\mathbf{C} = \mathbf{A}+(\mathbf{B}+\mathbf{C}) = (\mathbf{A}+\mathbf{B})+\mathbf{C}$. Podobne asociatívnosť maticového súčinu umožňuje značiť súčin matíc bez zátvoriek, napr. $\mathbf{ABC} = (\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$.

Poznámka 2.12 (Nekomutatívnosť súčinu matíc)**.** Súčin matíc nie je komutatívny, teda vo všeobecnosti neplatí $\mathbf{AB} = \mathbf{BA}$. Ak matice $\mathbf{A}$, $\mathbf{B}$ spĺňajú $\mathbf{AB} = \mathbf{BA}$ tak hovoríme, že *komutujú.*

Príklad 2.13. Nájdite dvojicu matíc typu 2×2, ktoré komutujú a dvojicu matíc typu 2×2, ktoré nekomutujú.

Definícia 2.14 (Transpozícia matice)**.** *Transpozícia* matice $\mathbf{A} \in \mathbb{R}^{m\times n}$ je matica $\mathbf{A}^T \in \mathbb{R}^{n\times m}$, ktorá spĺňa $(\mathbf{A}^T)_{ij} = (\mathbf{A})_{ji}$ pre každé $i = 1, \dots, n$, $j = 1, \dots, m$.

Veta 2.15 (Základné vlastnosti transpozície matíc)**.** Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{B} \in \mathbb{R}^{n\times k}$ a $\mathbf{C} \in \mathbb{R}^{m\times n}$. Potom $(\mathbf{A}^T)^T = \mathbf{A}$, $(\mathbf{A}+\mathbf{C})^T = \mathbf{A}^T + \mathbf{C}^T$, $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$.

Dôkaz. Dôkaz je triviálny. $\square$

Definícia 2.16 (Mocnina matice)**.** Nech $\mathbf{A} \in \mathbb{R}^{n\times n}$. Pre $k \in \mathbb{N}$ označujeme ako $\mathbf{A}^k$ maticu, ktorá je výsledkom k-násobného súčinu $\mathbf{A}\cdots\mathbf{A}$.

2.1.1 Špeciálne matice

Definícia 2.17 (Štvorcová matica, symetrická matica)**.** Každú maticu typu $n \times n$ nazývame *štvorcová*. Štvorcovú maticu $\mathbf{A}$ nazývame *symetrická*, ak $\mathbf{A} = \mathbf{A}^T$. Množinu všetkých symetrických matíc typu $n \times n$ značíme $\mathcal{S}^n$.

Definícia 2.18 (Diagonálne prvky, diagonálna matica)**.** Prvky $(\mathbf{A})_{11}, \dots, (\mathbf{A})_{nn}$ štvorcovej matice $\mathbf{A}$ typu $n \times n$ nazývame *diagonálne*, jej ostatné prvky nazývame *mimodiagonálne*. Maticu nazývame *diagonálna*, ak všetky mimodiagonálne prvky sú nulové. Diagonálnu maticu s prvkami $a_{11}, \dots, a_{nn}$ na diagonále označujeme $\mathrm{diag}(a_{11}, \dots, a_{nn})$.

Definícia 2.19 (Základné vektory a matice)**.** Nech $k, m, n \in \mathbb{N}$, $1 \leq k \leq n$. Vektor rozmeru n, ktorý má na k-tom mieste 1 a na ostatných miestach 0 nazývame *k-ty základný jednotkový vektor* rozmeru n a označujeme ho $\mathbf{e}_k$. Vektor rozmeru n, ktorý má všetky zložky rovné 1 (resp. rovné 0) voláme *n-rozmerný vektor jednotiek* (resp. *n-rozmerný vektor núl*) a označujeme ho $\mathbf{1}_n$ (resp. $\mathbf{0}_n$). Maticu $\mathrm{diag}(1, \dots, 1)$ typu $n \times n$ nazývame *matica identity* (niekedy tiež *jednotková matica*) a značíme ju $\mathbf{I}_n$. Maticu typu $m \times n$, ktorá má všetky zložky rovné 1 (rovné 0) voláme *matica jednotiek* (resp. *nulová matica*) a označujeme ju $\mathbf{J}_{m\times n}$ (resp. $\mathbf{0}_{m\times n}$).

Poznámka 2.20. Upozornime, že v zápise základného jednotkového vektoru $\mathbf{e}_k$ nie je špecifikovaná jeho dimenzia, tento zápis teda nie je formálne úplný. Vektor $\mathbf{e}_2$ môže napríklad vyjadrovať $(0,1)^T$ aj $(0,1,0)^T$. Zvyčajne je však dimenzia vektoru $\mathbf{e}_k$ zjavná z kontextu. V prípadoch, keď to tak nie je, budeme dimenziu priamo špecifikovať: napr. $\mathbf{e}_2 \in \mathbb{R}^3$ značí $(0,1,0)^T$.

Poznámka 2.21. Maticu $\mathbf{J}_{m\times n}$ vieme zapísať ako $\mathbf{J}_{m\times n} = \mathbf{1}_m\mathbf{1}_n^T$.

Poznámka 2.22. Niekedy kvôli prehľadnosti (obzvlášť v dôkazoch tvrdení) vynecháme dolné indexy pri matici identity a pri vektoroch a maticiach núl a jednotiek, t.j. používame aj značenie $\mathbf{I}$, $\mathbf{0}$, $\mathbf{1}$ a $\mathbf{J}$.

Definícia 2.23 (Horná a dolná trojuholníková matica)**.** Štvorcovú maticu $\mathbf{A}$ nazývame *horná trojuholníková*, ak $(\mathbf{A})_{ij} = 0$ pre všetky $i > j$ a *dolná trojuholníková*, ak $(\mathbf{A})_{ij} = 0$ pre všetky $i < j$.

Lema 2.24 (Súčin trojuholníkových matíc)**.** Ak $\mathbf{A}$ a $\mathbf{B}$ sú horné trojuholníkové matice rovnakého typu, tak aj $\mathbf{AB}$ je horná trojuholníková matica.

Dôkaz. Označme $\mathbf{A} = \{a_{ij}\}$, $\mathbf{B} = \{b_{ij}\}$ a $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n\times n}$. Prvok matice $\mathbf{AB}$ na (i,j)-tom mieste, kde $1 \leq j < i \leq n$, je

$$(\mathbf{AB})_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj} = \sum_{k=1}^{i-1} a_{ik}b_{kj} + \sum_{k=i}^{n} a_{ik}b_{kj} = \sum_{k=1}^{i-1} 0b_{kj} + \sum_{k=i}^{n} a_{ik}0 = 0.$$

$\square$

Príklad 2.25. Analogicky k leme 2.24 platí, že súčin dvoch dolných trojuholníkových matíc rovnakého typu je opäť dolná trojuholníková matica. Dokážte toto tvrdenie.

2.2 Blokové matice

Definícia 2.26 (Submatica, hlavná submatica, vedúca hlavná submatica)**.** *Submatica* matice $\mathbf{A}$ je matica, ktorú z nej získame odstránením niektorých riadkov a stĺpcov. *Hlavná submatica* štvorcovej matice $\mathbf{A}$ je taká štvorcová submatica, ktorá vznikla odstránením riadkov a stĺpcov s tými istými indexmi. *Vedúca hlavná submatica* štvorcovej matice $\mathbf{A}$ je hlavná submatica pozostávajúca z prvých k riadkov a stĺpcov matice $\mathbf{A}$ pre nejaké $k \in \mathbb{N}$.

Definícia 2.27 (Blokový zápis matice, bloková matica)**.** *Blokový zápis* matice $\mathbf{A} \in \mathbb{R}^{m\times n}$ je vyjadrenie $\mathbf{A}$ pomocou jej submatíc v tvare

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \cdots & \mathbf{A}_{1t} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \cdots & \mathbf{A}_{2t} \\ \vdots & \vdots & & \vdots \\ \mathbf{A}_{s1} & \mathbf{A}_{s2} & \cdots & \mathbf{A}_{st} \end{pmatrix} \tag{2.1}$$

pre nejaké $s \in \{1, \ldots, m\}$ a $t \in \{1, \ldots, n\}$, kde $\mathbf{A}_{ij}$ sú $m_i \times n_j$ matice, pričom $\sum_{i=1}^{s} m_i = m$ a $\sum_{j=1}^{t} n_j = n$. Matice $\mathbf{A}_{11}, \ldots, \mathbf{A}_{st}$ potom nazývame *bloky* matice $\mathbf{A}$. Ak maticu $\mathbf{A}$ vyjadríme v blokovom zápise, tak hovoríme, že $\mathbf{A}$ je *bloková matica*.

Poznámka 2.28. Ak blokový zápis matice má iba jeden riadok blokov (teda $s = 1$ v (2.1)), tak jednotlivé bloky niekedy oddeľujeme čiarkami. Napríklad sa často používa blokový zápis matice $\mathbf{A} \in \mathbb{R}^{m\times n}$ pomocou jej stĺpcov $\mathbf{a}_1, \ldots, \mathbf{a}_n \in \mathbb{R}^m$, čiže $\mathbf{A} = (\mathbf{a}_1, \ldots, \mathbf{a}_n)$.

Ak maticu vyjadríme v blokovom tvare, jej jednotlivé bloky niekedy oddeľujeme horizontálnymi a vertikálnymi čiarami, ako v nasledujúcom príklade.

Príklad 2.29. Nech

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 2 & 3 \\ 5 & 8 & 13 & 21 \\ 34 & 55 & 89 & 144 \\ 89 & 55 & 34 & 21 \end{pmatrix}.$$

Potom submaticou matice $\mathbf{A}$ sú napríklad matica $\mathbf{B}$, ktorú získame zmazaním tretieho riadku a druhého stĺpca, $\mathbf{C}$, ktorú získame zmazaním prvého a tretieho riadku aj stĺpca, a matica $\mathbf{D}$, ktorú získame odstránením tretieho a štvrtého riadku aj stĺpca:

$$\mathbf{B} = \begin{pmatrix} 1 & 2 & 3 \\ 5 & 13 & 21 \\ 89 & 34 & 21 \end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix} 8 & 21 \\ 55 & 21 \end{pmatrix}, \quad \mathbf{D} = \begin{pmatrix} 1 & 1 \\ 5 & 8 \end{pmatrix}.$$

Potom matice $\mathbf{C}$ a $\mathbf{D}$ sú hlavné submatice $\mathbf{A}$, pričom matica $\mathbf{D}$ je zároveň vedúca hlavná submatica. Maticu $\mathbf{A}$ môžeme vyjadriť v blokovom tvare napríklad ako

$$\mathbf{A} = \left(\begin{array}{cccc} 1 & 1 & 2 & 3 \\ \hline 5 & 8 & 13 & 21 \\ 34 & 55 & 89 & 144 \\ 89 & 55 & 34 & 21 \end{array}\right) \quad \text{alebo} \quad \mathbf{A} = \left(\begin{array}{cc|cc} 1 & 1 & 2 & 3 \\ 5 & 8 & 13 & 21 \\ \hline 34 & 55 & 89 & 144 \\ 89 & 55 & 34 & 21 \end{array}\right).$$

Nájdite ďalšie submatice, hlavné submatice a vedúce hlavné submatice matice $\mathbf{A}$.

Definícia 2.30 (Blokovo diagonálna matica)**.** Maticu $\mathbf{A}$ nazývame *blokovo diagonálna*, ak ju vieme zapísať v blokovom tvare (2.1), kde $s = t$ a $\mathbf{A}_{ij} = \mathbf{0}$ pre $i \neq j$, čiže

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{22} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{A}_{ss} \end{pmatrix}.$$

Takúto maticu značíme $\operatorname{diag}(\mathbf{A}_{11}, \dots, \mathbf{A}_{ss})$.

Definícia 2.31 (Konformnosť blokových zápisov pre sčítanie a násobenie)**.** Nech $\mathbf{A} \in \mathbb{R}^{m \times n}$ je v blokovom tvare (2.1), nech $\mathbf{B} \in \mathbb{R}^{p \times q}$ je v blokovom tvare

$$\mathbf{B} = \begin{pmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} & \cdots & \mathbf{B}_{1\ell} \\ \mathbf{B}_{21} & \mathbf{B}_{22} & \cdots & \mathbf{B}_{2\ell} \\ \vdots & \vdots & & \vdots \\ \mathbf{B}_{k1} & \mathbf{B}_{k2} & \cdots & \mathbf{B}_{k\ell} \end{pmatrix}, \tag{2.2}$$

pričom $\mathbf{B}_{ij}$ sú typu $p_i \times q_j$. Potom:

1. Ak $k = s$, $\ell = t$ a ak všetky dvojice $\mathbf{A}_{ij}$, $\mathbf{B}_{ij}$ sú konformné pre sčítanie (teda ak $m_i = p_i$ a $n_j = q_j$ pre všetky i, j), tak hovoríme, že blokové zápisy $\mathbf{A}$ a $\mathbf{B}$ sú *konformné pre sčítanie.*

2. Ak $k = t$ a ak všetky dvojice $\mathbf{A}_{iu}$, $\mathbf{B}_{uj}$ sú konformné pre násobenie (teda ak $n_u = p_u$ pre všetky u), tak hovoríme, že blokové zápisy $\mathbf{A}$ a $\mathbf{B}$ sú *konformné pre násobenie.*

Veta 2.32 (Základná aritmetika matíc vyjadrených v blokovom tvare)**.** Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$ je v blokovom tvare (2.1) a nech $\mathbf{B} \in \mathbb{R}^{p\times q}$ je v blokovom tvare (2.2). Potom:

(i)
$$c\mathbf{A} = \begin{pmatrix} c\mathbf{A}_{11} & c\mathbf{A}_{12} & \cdots & c\mathbf{A}_{1t} \\ c\mathbf{A}_{21} & c\mathbf{A}_{22} & \cdots & c\mathbf{A}_{2t} \\ \vdots & \vdots & & \vdots \\ c\mathbf{A}_{s1} & c\mathbf{A}_{s2} & \cdots & c\mathbf{A}_{st} \end{pmatrix} \quad \text{a} \quad \mathbf{A}^T = \begin{pmatrix} \mathbf{A}_{11}^T & \mathbf{A}_{21}^T & \cdots & \mathbf{A}_{s1}^T \\ \mathbf{A}_{12}^T & \mathbf{A}_{22}^T & \cdots & \mathbf{A}_{s2}^T \\ \vdots & \vdots & & \vdots \\ \mathbf{A}_{1t}^T & \mathbf{A}_{2t}^T & \cdots & \mathbf{A}_{st}^T \end{pmatrix}.$$

(ii) Ak blokové zápisy $\mathbf{A}$ a $\mathbf{B}$ sú konformné pre sčítanie, tak
$$\mathbf{A}+\mathbf{B} = \begin{pmatrix} \mathbf{A}_{11}+\mathbf{B}_{11} & \mathbf{A}_{12}+\mathbf{B}_{12} & \cdots & \mathbf{A}_{1t}+\mathbf{B}_{1t} \\ \mathbf{A}_{21}+\mathbf{B}_{21} & \mathbf{A}_{22}+\mathbf{B}_{22} & \cdots & \mathbf{A}_{2t}+\mathbf{B}_{2t} \\ \vdots & \vdots & & \vdots \\ \mathbf{A}_{s1}+\mathbf{B}_{s1} & \mathbf{A}_{s2}+\mathbf{B}_{s2} & \cdots & \mathbf{A}_{st}+\mathbf{B}_{st} \end{pmatrix}.$$

(iii) Ak blokové zápisy $\mathbf{A}$ a $\mathbf{B}$ sú konformné pre násobenie, tak
$$\mathbf{AB} = \begin{pmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} & \cdots & \mathbf{C}_{1\ell} \\ \mathbf{C}_{21} & \mathbf{C}_{22} & \cdots & \mathbf{C}_{2\ell} \\ \vdots & \vdots & & \vdots \\ \mathbf{C}_{s1} & \mathbf{C}_{s2} & \cdots & \mathbf{C}_{s\ell} \end{pmatrix},$$
kde $\mathbf{C}_{ij} = \sum_{u=1}^{t} \mathbf{A}_{iu}\mathbf{B}_{uj}$ $(i = 1,\ldots,s,\ j = 1,\ldots,\ell)$.

Dôkaz. Dôkazy častí (i) a (ii) sú jednoduché, nechávame ich na čitateľa. V časti (iii) počítajme prvky v bloku $\mathbf{C}_{ij}$ matice $\mathbf{C} = \mathbf{AB}$. Prvok na mieste (e, f) matice $\mathbf{C}_{ij}$ vznikol vynásobením príslušného (čiže $(m_1 + \ldots + m_{i-1} + e)$-teho) riadku matice $\mathbf{A}$ a príslušného (čiže $(q_1 + \ldots + q_{j-1} + f)$-tého) stĺpca matice $\mathbf{B}$. To je teda súčin e-teho riadku matice $(\mathbf{A}_{i1}, \ldots, \mathbf{A}_{it})$ a f-tého stĺpca matice
$$\begin{pmatrix} \mathbf{B}_{1j} \\ \vdots \\ \mathbf{B}_{tj} \end{pmatrix}.$$
Rozpísaním súčinu po zložkách vidíme, že hľadaný prvok je
$$(\mathbf{C}_{ij})_{ef} = \sum_{d=1}^{n_1} (\mathbf{A}_{i1})_{ed}(\mathbf{B}_{1j})_{df} + \ldots + \sum_{d=1}^{n_t} (\mathbf{A}_{it})_{ed}(\mathbf{B}_{tj})_{df}.$$
Zároveň prvok na pozícii (e, f) v matici $\mathbf{A}_{iu}\mathbf{B}_{uj}$ je z definície súčinu matíc:
$$(\mathbf{A}_{iu}\mathbf{B}_{uj})_{ef} = \sum_{d=1}^{n_u} (\mathbf{A}_{iu})_{ed}(\mathbf{B}_{uj})_{df},$$
a teda $(\mathbf{C}_{ij})_{ef} = \sum_{u=1}^{t} (\mathbf{A}_{iu}\mathbf{B}_{uj})_{ef}$, čiže $\mathbf{C}_{ij} = \mathbf{A}_{iu}\mathbf{B}_{uj}$. □

Dôsledok 2.33 (Násobenie matice a vektora)**.** Vyjadrime $\mathbf{A} \in \mathbb{R}^{m\times n}$ ako $\mathbf{A} = (\mathbf{a}_1, \ldots, \mathbf{a}_n)$, kde $\mathbf{a}_1, \ldots, \mathbf{a}_n$ sú stĺpce $\mathbf{A}$ a nech $\mathbf{x} \in \mathbb{R}^n$. Potom $\mathbf{Ax}$ vieme podľa vety 2.32 zapísať ako

$$\mathbf{Ax} = x_1\mathbf{a}_1 + \ldots + x_n\mathbf{a}_n.$$

Teda každý súčin $\mathbf{Ax}$ je lineárnou kombináciou stĺpcov matice $\mathbf{A}$, s koeficientami danými vektorom $\mathbf{x}$.

Poznámka 2.34. Blokové matice prirodzene vyvstávajú v štatistických modeloch, v ktorých sú dve alebo viaceré skupiny premenných. Také modely zodpovedajú napríklad situáciám, keď vplyvy niektorých premenných síce v modeli musíme uvážiť, ale nie je naším cieľom ich kvantifikovať (to sú tzv. rušivé premenné), alebo tiež štandardnému modelu lineárnej regresie, kde tzv. konštantný člen (angl. *intercept*) hrá prirodzene výnimočnú rolu. Zároveň sa so submaticami často pracuje v matematickom a štatistickom softvéri. Preto blokovým maticiam v týchto skriptách venujeme špeciálnu pozornosť.

2.3 Riedke matice

Voľne môžeme povedať, že matica je *riedka* (angl. *sparse*) ak obsahuje veľké množstvo núl, inak je *hustá* (angl. *dense*). Jedna z formálnych definícií riedkych matíc je nasledovná:

Definícia 2.35 (Riedka matica)**.** Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$. Ak viac ako polovica jej prvkov sú nuly, matica $\mathbf{A}$ sa nazýva *riedka*.

Definícia 2.36 (Dolné, horné pásmo matice)**.** Matica $\mathbf{A} \in \mathbb{R}^{m\times n}$ má *dolné pásmo šírky* (angl. *lower bandwith*) p, keď $(\mathbf{A})_{i,j} = 0$ pre každé i, j: $i > j + p$. $\mathbf{A} \in \mathbb{R}^{m\times n}$ má *horné pásmo šírky* (angl. *upper bandwith*) q, keď $(\mathbf{A})_{i,j} = 0$ pre každé i, j: $j > i + q$.

Riedka matica, ktorá má dolné alebo horné netriviálne pásmo sa nazýva *pásmová* (angl. *band matrix*) (čiže taká matica má nuly v ľavom dolnom rohu alebo v pravom hornom rohu).

Príklad 2.37. Matica

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & \mathbf{0} & \mathbf{0} \\ 3 & 4 & 5 & \mathbf{0} \\ 6 & 7 & 0 & 8 \\ \mathbf{0} & 7 & 6 & 5 \end{pmatrix}$$

má dolné pásmo šírky 2 a horné pásmo šírky 1 vďaka nulovým prvkom, ktoré sú vyznačené hrubým písmom. Skonštruujte 5×4 maticu, ktorá má horné pásom šírky 2 a dolné pásmo šírky 1.

Príklad 2.38. Všimnite si, že horná trojuholníková matica má dolné pásmo šírky 0. Určte šírky pásiem dolnej trojuholníkovej matice a diagonálnej matice.

S riedkymi maticami sa často stretávame pri numerickom derivovaní, resp. pri numerickom riešení diferenciálnych rovníc. Tiež môžu byť výstupom niektorých štatistických metód. Dáta majú tendenciu byť riedke špeciálne ak obsahujú kategorické premenné, ktoré sú kódované pomocou pomocných premenných (tzv. *dummy premenných*).

Z výpočtového hľadiska sa dá s riedkymi maticami jednoduchšie pracovať, keď algoritmy prispôsobíme ich veľkému množstvu núl a nemrháme pamäťou a výpočtovou kapacitou na zbytočné ukladanie núl a operácie s nimi. To už je však obsahom skôr výpočtových metód a numerickej matematiky; riedkymi maticami sa preto nebudeme ďalej zaoberať. Čitateľ sa viac môže dočítať napríklad v [3].

2.4 Ďalšie maticové operácie

Definícia 2.39 (Kroneckerov súčin)**.** Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$ a $\mathbf{B} \in \mathbb{R}^{p\times q}$. Potom *Kroneckerov súčin* matíc $\mathbf{A} = \{a_{ij}\}$ a $\mathbf{B}$ sa značí $\mathbf{A} \otimes \mathbf{B}$ a je v tvare blokovej matice definovaný ako

$$\mathbf{A} \otimes \mathbf{B} = \begin{pmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \dots & a_{1n}\mathbf{B} \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \dots & a_{2n}\mathbf{B} \\ \vdots & \vdots & & \vdots \\ a_{m1}\mathbf{B} & a_{m2}\mathbf{B} & \dots & a_{mn}\mathbf{B} \end{pmatrix}.$$

Veta 2.40 (Vlastnosti Kroneckerovho súčinu)**.** Nech $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\mathbf{D}$ sú matice vhodných rozmerov, aby príslušné súčty a súčiny boli dobre definované a nech $\mathbf{H} = \operatorname{diag}(h_1, \dots, h_k)$. Potom

(i) $(\mathbf{A} \otimes \mathbf{B})^T = \mathbf{A}^T \otimes \mathbf{B}^T$,

(ii) $(\mathbf{H} \otimes \mathbf{A}) = \operatorname{diag}(h_1\mathbf{A}, \dots, h_k\mathbf{A})$,

(iii) $\mathbf{A} \otimes (\mathbf{B} + \mathbf{C}) = \mathbf{A} \otimes \mathbf{B} + \mathbf{A} \otimes \mathbf{C}$,

(iv) $(\mathbf{A} + \mathbf{B}) \otimes \mathbf{C} = \mathbf{A} \otimes \mathbf{C} + \mathbf{B} \otimes \mathbf{C}$,

(v) $(\mathbf{A} \otimes \mathbf{B}) \otimes \mathbf{C} = \mathbf{A} \otimes (\mathbf{B} \otimes \mathbf{C})$,

(vi) $(\mathbf{A} \otimes \mathbf{B}) \cdot (\mathbf{C} \otimes \mathbf{D}) = (\mathbf{AC}) \otimes (\mathbf{BD})$.

Dôkaz. Časti (i)-(v) sú priamočiare, ponecháme ich na čitateľa. V dôkaze (vi) uvažujme $\mathbf{A} = \{a_{ij}\} \in \mathbb{R}^{m\times n}$, $\mathbf{B} \in \mathbb{R}^{p\times q}$, $\mathbf{C} = \{c_{ij}\} \in \mathbb{R}^{n\times s}$ a $\mathbf{D} \in \mathbb{R}^{q\times t}$. Ľavá strana rovnice v časti (vi) je

$$\begin{pmatrix} a_{11}\mathbf{B} & \dots & a_{1n}\mathbf{B} \\ a_{21}\mathbf{B} & \dots & a_{2n}\mathbf{B} \\ \vdots & & \vdots \\ a_{m1}\mathbf{B} & \dots & a_{mn}\mathbf{B} \end{pmatrix} \begin{pmatrix} c_{11}\mathbf{D} & \dots & c_{1s}\mathbf{D} \\ c_{21}\mathbf{D} & \dots & c_{2s}\mathbf{D} \\ \vdots & & \vdots \\ c_{n1}\mathbf{D} & \dots & c_{ns}\mathbf{D} \end{pmatrix}.$$

To je podľa vety 2.32 matica pozostávajúca z $m \times s$ blokov veľkosti $p \times t$, pričom blok na pozícii (i, j) $(i = 1, \dots, m,\ j = 1, \dots, s)$ je

$$\sum_{k=1}^{n} a_{ik}\mathbf{B} \cdot c_{kj}\mathbf{D} = \left(\sum_{k=1}^{n} a_{ik}c_{kj}\right)\mathbf{BD}.$$

Prvok na pozícii (i, j) matice $\mathbf{AC}$ je $\sum_{k=1}^{n} a_{ik}c_{kj}$, a teda blok na (i, j)-tej pozícii matice $(\mathbf{AC}) \otimes (\mathbf{BD})$ je $(\sum_{k=1}^{n} a_{ik}c_{kj})\,\mathbf{BD}$. Takže ľavá strana rovnice sa rovná pravej. □

Definícia 2.41 (Hadamardov súčin)**.** *Hadamardov súčin* (alebo *pozložkový súčin*) matíc $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m\times n}$ je matica typu $m \times n$, ktorú značíme $\mathbf{A} \odot \mathbf{B}$ a je určená predpisom

$$(\mathbf{A} \odot \mathbf{B})_{ij} = (\mathbf{A})_{ij}(\mathbf{B})_{ij}.$$

Poznámka 2.42. Hadamardov súčin sa v teoretickej maticovej algebre používa zriedka. Je ale užitočný pri používaní matematického alebo štatistického softvéru založeného na maticiach, ako sú R a MATLAB.

Definícia 2.43 (Operátor vec)**.** Zapíšme $\mathbf{A} = \{a_{ij}\} \in \mathbb{R}^{m\times n}$ pomocou jej stĺpcov ako $\mathbf{A} = (\mathbf{a}_1, \ldots, \mathbf{a}_n)$. *Operátor vec* je zobrazenie $\mathrm{vec} : \mathbb{R}^{m\times n} \to \mathbb{R}^{mn}$ definované nasledovne:

$$\mathrm{vec}(\mathbf{A}) = \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_n \end{pmatrix} = \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \\ a_{12} \\ \vdots \\ a_{m2} \\ \vdots \\ a_{mn} \end{pmatrix}.$$

Poznámka 2.44. Operátor vec (z anglického *vector*, vektorový zápis) umožňuje zapísať maticu typu $m \times n$ ako mn-rozmerný vektor a následne k nej teda pristupovať ako k vektoru.

Príklad 2.45. Nech

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 5 & -2 \\ 3 & -4 \\ 17 & 1 \end{pmatrix}.$$

Potom $\mathrm{vec}(\mathbf{A}) = (1, 4, 2, 5, 3, 6)^T$. Určte $\mathrm{vec}(\mathbf{B})$.

Veta 2.46 (Vlastnosti vec)**.** Operátor vec má nasledujúce vlastnosti.

(i) Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$, $c \in \mathbb{R}$, potom $\mathrm{vec}(c\mathbf{A}) = c \cdot \mathrm{vec}(\mathbf{A})$.

(ii) Nech $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m\times n}$, potom $\mathrm{vec}(\mathbf{A} + \mathbf{B}) = \mathrm{vec}(\mathbf{A}) + \mathrm{vec}(\mathbf{B})$.

(iii) Nech $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$, potom $\mathrm{vec}(\mathbf{b}\mathbf{a}^T) = \mathbf{a} \otimes \mathbf{b}$.

(iv) Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{B} = (\mathbf{b}_1, \ldots, \mathbf{b}_k) \in \mathbb{R}^{n\times k}$, potom

$$\mathrm{vec}(\mathbf{AB}) = \begin{pmatrix} \mathbf{Ab}_1 \\ \vdots \\ \mathbf{Ab}_k \end{pmatrix} = \mathrm{diag}(\mathbf{A}, \ldots, \mathbf{A})\mathrm{vec}(\mathbf{B}) = (\mathbf{I}_k \otimes \mathbf{A})\mathrm{vec}(\mathbf{B}).$$

(v) Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{B} \in \mathbb{R}^{n\times k}$, $\mathbf{C} \in \mathbb{R}^{k\times \ell}$, potom $\mathrm{vec}(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A})\mathrm{vec}(\mathbf{B})$.

Dôkaz. Tvrdenia (i) a (ii) sú triviálne.
(iii) Keďže $\mathbf{b}\mathbf{a}^T = (a_1\mathbf{b}, \ldots, a_n\mathbf{b})$, tak

$$\mathrm{vec}(\mathbf{b}\mathbf{a}^T) = \begin{pmatrix} a_1\mathbf{b} \\ \vdots \\ a_n\mathbf{b} \end{pmatrix} = \mathbf{a} \otimes \mathbf{b}.$$

(iv) Máme

$$\mathrm{vec}(\mathbf{AB}) = \mathrm{vec}\Big((\mathbf{Ab}_1, \ldots, \mathbf{Ab}_k)\Big) = \begin{pmatrix} \mathbf{Ab}_1 \\ \vdots \\ \mathbf{Ab}_k \end{pmatrix},$$

pričom zároveň

$$(\mathbf{I}_k \otimes \mathbf{A})\text{vec}(\mathbf{B}) = \text{diag}(\mathbf{A}, \dots, \mathbf{A})\text{vec}(\mathbf{B}) = \begin{pmatrix} \mathbf{Ab}_1 \\ \vdots \\ \mathbf{Ab}_k \end{pmatrix}.$$

(v) Zrejme

$$\mathbf{B} = (\mathbf{b}_1, \dots, \mathbf{b}_k) = \sum_{j=1}^{k} \mathbf{b}_j \mathbf{e}_j^T,$$

a teda

$$\text{vec}(\mathbf{ABC}) = \text{vec}\left(\mathbf{A}\left(\sum_{j=1}^{k} \mathbf{b}_j \mathbf{e}_j^T\right)\mathbf{C}\right) = \sum_{j=1}^{k} \text{vec}\left(\mathbf{Ab}_j \mathbf{e}_j^T \mathbf{C}\right)$$

podľa vlastnosti (ii). Keďže $\mathbf{Ab}_j$ a $\mathbf{C}^T\mathbf{e}_j$ sú vektory, tak z vlastnosti (iii) vyplýva

$$\text{vec}(\mathbf{ABC}) = \sum_{j=1}^{k} (\mathbf{C}^T\mathbf{e}_j) \otimes (\mathbf{Ab}_j) = \sum_{j=1}^{k} (\mathbf{C}^T \otimes \mathbf{A})(\mathbf{e}_j \otimes \mathbf{b}_j),$$

využijúc vetu 2.40(vi). Opäť môžeme použiť vlastnosti (iii) a (ii), čím dostaneme

$$\text{vec}(\mathbf{ABC}) = \sum_{j=1}^{k} (\mathbf{C}^T \otimes \mathbf{A})\text{vec}(\mathbf{b}_j \mathbf{e}_j^T) = (\mathbf{C}^T \otimes \mathbf{A})\text{vec}\left(\sum_{j=1}^{k} \mathbf{b}_j \mathbf{e}_j^T\right) = (\mathbf{C}^T \otimes \mathbf{A})\text{vec}(\mathbf{B}).$$

□

Definícia 2.47 (Operátor vech)**.** Zapíšme symetrickú maticu $\mathbf{A} = \{a_{ij}\} \in \mathcal{S}^n$ pomocou jej stĺpcov ako $\mathbf{A} = (\mathbf{a}_1, \dots, \mathbf{a}_n)$ a z každého stĺpca $\mathbf{a}_i$ vytvorme vektor $\mathbf{a}_i^*$ vymazaním jeho prvkov nad diagonálou, čiže $\mathbf{a}_i^* = (a_{ii}, a_{i+1,i}, \dots, a_{ni})^T$. *Operátor vech* je zobrazenie vech: $\mathcal{S}^n \to \mathbb{R}^{n(n+1)/2}$ definované nasledovne:

$$\text{vech}(\mathbf{A}) = \begin{pmatrix} \mathbf{a}_1^* \\ \mathbf{a}_2^* \\ \vdots \\ \mathbf{a}_n^* \end{pmatrix}.$$

Poznámka 2.48. V prípade symetrických matíc sa $n(n-1)/2$ mimodiagonálnych prvkov opakuje, takže symetrické matice stačí previesť do vektorového zápisu iba pomocou $n(n+1)/2$ prvkov na diagonále a pod diagonálou. Názov operátoru vech vychádza z anglického *vector-half*, vektorová polovica, keďže sa použije iba polovica mimodiagonálnych prvkov.

Príklad 2.49. Nech

$$\mathbf{A} = \begin{pmatrix} 2 & 4 & 6 \\ 4 & 3 & 5 \\ 6 & 5 & 9 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 7 & 8 \\ 8 & 2 \end{pmatrix}.$$

Potom $\text{vech}(\mathbf{A}) = (2, 4, 6, 3, 5, 9)^T$. Určte $\text{vech}(\mathbf{B})$.

Lema 2.50 (Vlastnosti operátora vech)**.** Nech $\mathbf{A}, \mathbf{B} \in \mathcal{S}^n$ a $c \in \mathbb{R}$. Potom

(i) $\text{vech}(c\mathbf{A}) = c \cdot \text{vech}(\mathbf{A})$,

(ii) $\text{vech}(\mathbf{A} + \mathbf{B}) = \text{vech}(\mathbf{A}) + \text{vech}(\mathbf{B})$.

Dôkaz. Tvrdenie vychádza z faktu, že $\text{vech}(\mathbf{A})$ je podvektorom $\text{vec}(\mathbf{A})$, ktorý tieto vlastnosti spĺňa. □

2.5 Matica ako lineárne zobrazenie

Veta 2.51 (Matica reprezentuje lineárne zobrazenie)**.** Nech $\mathbf{A} \in \mathbb{R}^{m \times n}$. Definujme zobrazenie $f : \mathbb{R}^n \to \mathbb{R}^m$, $f(\mathbf{x}) = \mathbf{A}\mathbf{x}$. Potom f je lineárne zobrazenie.

Dôkaz. Nech $a, b \in \mathbb{R}$ a $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Potom $f(a\mathbf{x} + b\mathbf{y}) = \mathbf{A}(a\mathbf{x} + b\mathbf{y}) = a\mathbf{A}\mathbf{x} + b\mathbf{A}y = af(\mathbf{x}) + bf(\mathbf{y})$, čiže f je podľa definície 1.53 lineárne. □

Veta 2.52 (Matica lineárneho zobrazenia)**.** Nech $f : \mathbb{R}^n \to \mathbb{R}^m$ je lineárne zobrazenie. Potom existuje matica $\mathbf{A}$ taká, že $f(\mathbf{x}) = \mathbf{A}\mathbf{x}$ pre každé $\mathbf{x} \in \mathbb{R}^n$. Táto matica má tvar

$$\mathbf{A} = \begin{pmatrix} f(\mathbf{e}_1), & \ldots, & f(\mathbf{e}_n) \end{pmatrix}$$

a nazýva sa *matica lineárneho zobrazenia* f.

Dôkaz. Nech $\mathbf{x} \in \mathbb{R}^n$. Potom $\mathbf{x} = x_1\mathbf{e}_1 + \ldots + x_n\mathbf{e}_n$. Keďže f je lineárne, tak $f(\mathbf{x}) = x_1 f(\mathbf{e}_1) + \ldots + x_n f(\mathbf{e}_n)$, čo je podľa dôsledku 2.33 totožné s $\mathbf{A}\mathbf{x}$. □

Príklad 2.53. Nech $\mathbf{A} = \text{diag}(3, 2)$. Potom $\mathbf{A}\mathbf{x}$ predstavuje zobrazenie, ktoré "natiahne" prvú zložku vektora $\mathbf{x}$ trojnásobne a druhú zložku dvojnásobne. Pre $\mathbf{x} = (0{,}8; 1{,}2)^T$ je toto zobrazenie zakreslené na obrázku 2.1a. Na obrázku 2.1b môžeme vidieť lineárne zobrazenie toho istého vektora $\mathbf{x}$ pre

$$\mathbf{A} = \begin{pmatrix} 3 & -3 \\ 1 & 3 \end{pmatrix}. \tag{2.3}$$

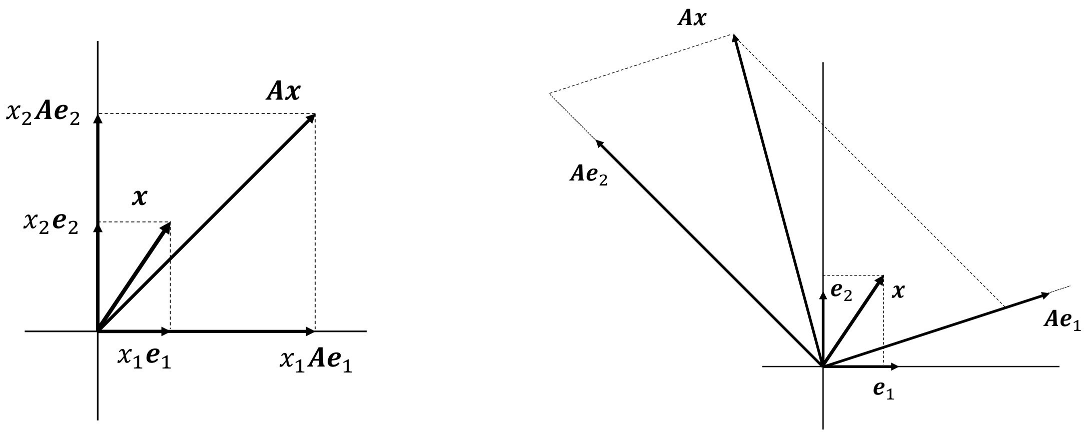
Transformáciu $\mathbf{A}\mathbf{x}$ pre $\mathbf{A} = (\mathbf{a}_1, \mathbf{a}_2) \in \mathbb{R}^{2\times 2}$ môžeme vo všeobecnosti znázorniť pomocou obrazov $\mathbf{e}_1$ a $\mathbf{e}_2$, čiže $\mathbf{A}\mathbf{e}_1 = \mathbf{a}_1$ a $\mathbf{A}\mathbf{e}_2 = \mathbf{a}_2$. Potom $\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2$ sa po transformácii zmení na $\mathbf{A}\mathbf{x} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2$. Dostávame teda nový vektor $\mathbf{y} = \mathbf{A}\mathbf{x}$, ktorý má tie isté súradnice x_1 a x_2, ale v novej báze $\mathbf{a}_1, \mathbf{a}_2$. Podobne zakreslite transformáciu vektora $\mathbf{x} = (1{,}5; -1)^T$ na $\mathbf{A}\mathbf{x}$ pomocou oboch spomínaných matíc $\mathbf{A}$.

Veta 2.54 (Zloženie lineárnych zobrazení a súčin matíc)**.** Nech $f : \mathbb{R}^n \to \mathbb{R}^m$ je lineárne zobrazenie dané maticou $\mathbf{A}$ a $g : \mathbb{R}^m \to \mathbb{R}^k$ je lineárne zobrazenie dané maticou $\mathbf{B}$. Potom matica lineárneho zobrazenia $g \circ f : \mathbb{R}^n \to \mathbb{R}^k$ je $\mathbf{BA}$.

Dôkaz. Označme i-ty stĺpec matice $\mathbf{A}$ ako $\mathbf{a}_i = \mathbf{A}\mathbf{e}_i$, $i = 1, \ldots, n$. Matica zobrazenia $g \circ f$ je daná obrazmi $g(f(\mathbf{e}_i)) = g(\mathbf{A}\mathbf{e}_i) = g(\mathbf{a}_i) = \mathbf{B}\mathbf{a}_i$ pre $i = 1, \ldots, n$, keďže $f(\mathbf{x}) = \mathbf{A}\mathbf{x}$ a $g(\mathbf{y}) = \mathbf{B}\mathbf{y}$ pre každé $\mathbf{x} \in \mathbb{R}^n$ a $\mathbf{y} \in \mathbb{R}^m$. Teda matica zobrazenia $g \circ f$ je

$$\left(\mathbf{B}\mathbf{a}_1, \ldots, \mathbf{B}\mathbf{a}_n\right) = \mathbf{B}\left(\mathbf{a}_1, \ldots, \mathbf{a}_n\right) = \mathbf{BA}.$$

□



(a) $\mathbf{A} = \mathrm{diag}(3, 2)$ (b) $\mathbf{A}$ je daná rovnicou (2.3)

Obr. 2.1: Lineárne zobrazenie $\mathbf{Ax}$ pre $\mathbf{x} \in \mathbb{R}^2$.

2.6 Dáta ako vektory a matice

Pri práci s dátami tieto zvyčajne predstavujú d-rozmerné vektory, kde d je počet meraných vlastností jednotlivých objektov. Napríklad poisťovňa môže mať o každom poistenom aute zaznamenané jeho nosnosť, hmotnosť a objem motora. Potom je každé auto reprezentované vektorom v $\mathbb{R}^3$. Údaje o n autách sú vektory $\mathbf{x}_1, \dots, \mathbf{x}_n \in \mathbb{R}^3$ (vo všeobecnosti v $\mathbb{R}^d$), spolu môžeme dáta zapísať v matici $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_n) \in \mathbb{R}^{d\times n}$. Vektory $\mathbf{x}_1, \dots, \mathbf{x}_n \in \mathbb{R}^d$ tiež nazývame dátové body.

Poznamenajme, že v praxi sa zvyčajne stretávame s maticou dát, kde jednotlivé objekty (poistenci) sú riadky a pozorované premenné sú stĺpce, teda $\mathbf{X}_2 = (\mathbf{x}_1, \dots, \mathbf{x}_n)^T \in \mathbb{R}^{n\times d}$. Táto forma zaznamenania dát je pohodlná, pretože objektov býva obvykle oveľa viac ako premenných, matematicky je však prirodzenejšie pracovať s maticou $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_n) \in \mathbb{R}^{d\times n}$, ktorú používame v týchto skriptách.

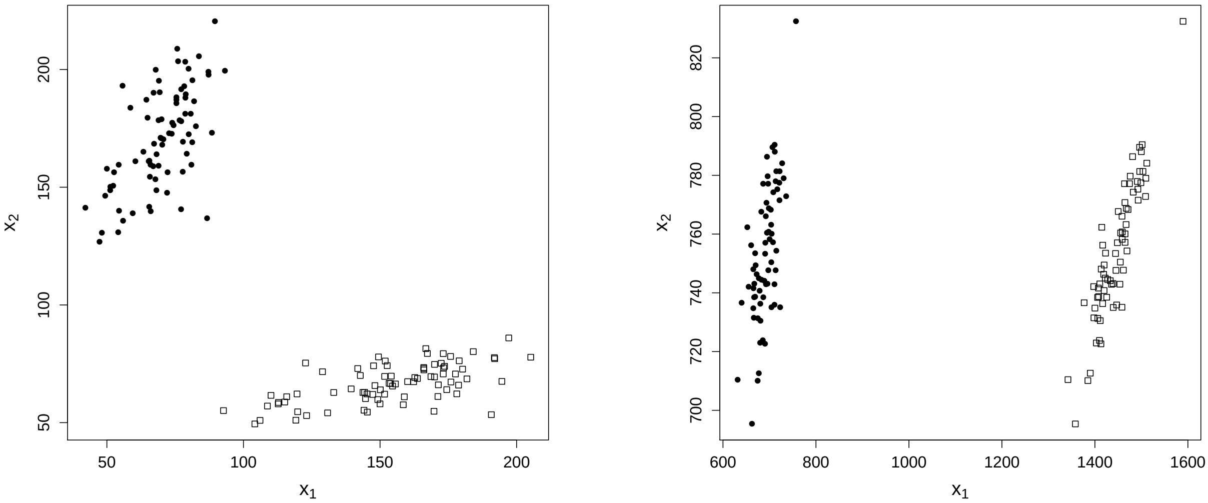
Ak chceme vykonať nejakú lineárnu transformáciu údajov, teda ich prenásobenie maticou $\mathbf{A} \in \mathbb{R}^{d\times d}$, tak to ľahko spravíme vynásobením matice $\mathbf{X}$. Upravené dáta získame ako $\tilde{\mathbf{X}} = (\tilde{\mathbf{x}}_1, \dots, \tilde{\mathbf{x}}_n) = (\mathbf{Ax}_1, \dots, \mathbf{Ax}_n) = \mathbf{AX}$. Elementárnym príkladom takej operácie je zmena mierky jednotlivých premenných.

Napríklad ak $\mathbf{x}_i \in \mathbb{R}^2$ predstavuje váhu $x_1^{(i)}$ i-teho človeka v kilogramoch a jeho výšku $x_2^{(i)}$ v centimetroch, $i = 1, \dots, n$, tak $2{,}2x_1^{(i)}$ bude jeho váha v librách a $0{,}39x_2^{(i)}$ bude jeho výška v palcoch. Teda pre potreby britského čitateľa môžeme dáta $\mathbf{x}_1, \dots, \mathbf{x}_n$ transformovať na $\mathbf{Ax}_1, \dots, \mathbf{Ax}_n$, kde $\mathbf{A} = \mathrm{diag}(2{,}2; 0{,}39)$. Na obrázku 2.2a sú zachytené (náhodne vygenerované) váhy a výšky 70 ľudí v pôvodných jednotkách a v britských jednotkách.

Uvažujme dáta o mesačných príjmoch 70 manželských párov. Tie môžeme reprezentovať vektormi $\mathbf{x}_1, \dots, \mathbf{x}_n \in \mathbb{R}^2$ $(n = 70)$, ktorých prvá zložka je príjem manželky a druhá zložka príjem manžela (obidve v eurách). Dáta môžeme reprezentovať aj pomocou vektorov $\tilde{\mathbf{x}}_1, \dots, \tilde{\mathbf{x}}_n$, ktorých prvá zložka je celkový príjem páru a druhá zložka ostáva príjmom muža. Teda $\tilde{x}_1^{(i)} = x_1^{(i)} + x_2^{(i)}$ a $\tilde{x}_2^{(i)} = x_2^{(i)}$ pre $i = 1, \dots, 70$. Takže transformáciu dát môžeme opäť vyjadriť pomocou matice ako $\tilde{\mathbf{x}}_i = \mathbf{Ax}_i$, kde

$$\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Vygenerované dáta aj ich transformácie pre 70 párov sú zakreslené na obrázku 2.2b.



(a) x_1 - váha, x_2 - výška. Pôvodné dáta (krúžky) sú v kilogramoch a centimetroch, transformované (štvorce) v librách a palcoch.

(b) Pôvodné dáta (krúžky): x_1 - príjem ženy, x_2 - príjem muža. Transformované (štvorce): x_1 - príjem páru, x_2 - príjem muža.

Obr. 2.2: Dáta z kapitoly 2.6 reprezentované vektormi v $\mathbb{R}^2$ a ich lineárne transformácie.

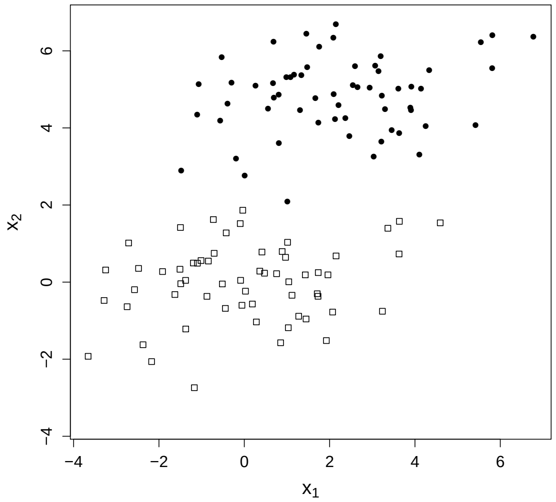
Príklad 2.55. Častým prvým krokom analýzy dát je ich centrovanie, čiže taká úprava, aby priemery dát po jednotlivých súradniciach boli nulové. Graficky to zodpovedá posunutiu dát tak, aby ich stred bol v počiatku súradnicovej sústavy. Ukážte, že keď $\mathbf{X} \in \mathbb{R}^{d\times n}$ je matica dát, potom $\tilde{\mathbf{X}} = \mathbf{X}(\mathbf{I}_n - n^{-1}\mathbf{J}_{n\times n})$ sú centrované dáta (angl. *centered data*), teda pre ne platí $n^{-1}\sum_{i=1}^n (\tilde{\mathbf{X}})_{ki} = 0$ pre každé $k = 1, \dots, d$. Príklad centrovaných dát je uvedený na obrázku 2.3; môžeme si všimnúť, že "stred" centrovaných dát je $\mathbf{0}_2$.

Príklad 2.56. Viaceré metódy analýzy dát, hlavne tie slúžiace na vizualizáciu alebo na zhlukovanie, pracujú so vzdialenosťami medzi jednotlivými dátovými bodmi. Ak $\mathbf{x}_i, \mathbf{x}_j \in \mathbb{R}^d$ sú dátové body (vektory dát), potom ich vzájomná vzdialenosť je $d_{ij} = \|\mathbf{x}_i - \mathbf{x}_j\| = \sqrt{(\mathbf{x}_i - \mathbf{x}_j)^T(\mathbf{x}_i - \mathbf{x}_j)}$. Pre maticu dát $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_n)$ definujme maticu $\mathbf{B} = \mathbf{X}^T\mathbf{X} \in \mathbb{R}^{n\times n}$. Všimnite si, že $\mathbf{B}$ je maticou skalárnych súčinov medzi vektormi $\mathbf{x}_1, \dots, \mathbf{x}_n$ (tzv. *Gramova matica*). Ukážte, že platí $d_{ij}^2 = (\mathbf{B})_{ii} + (\mathbf{B})_{jj} - 2(\mathbf{B})_{ij}$.

Maticu štvorcov vzájomných vzdialeností $\mathbf{D} = \{d_{ij}^2\}$ rozmeru $n \times n$ môžeme získať aj priamo maticovým zápisom. Označme $\mathbf{C} = \mathbf{y}\mathbf{1}_n^T$, kde $\mathbf{y} \in \mathbb{R}^n$ je vektor diagonálnych prvkov matice $\mathbf{B}$. Overte, že potom $\mathbf{D} = \mathbf{C} + \mathbf{C}^T - 2\mathbf{B}$. V softvéroch založených na práci s maticami umožňuje tento vzťah veľmi rýchly výpočet matice $\mathbf{D}$ štvorcov vzájomných vzdialeností.

2.6.1 Výberová kovariančná matica

V štatistike často používaným objektom opisujúcim skúmanú sadu dát je *výberová kovariančná matica* $\mathbf{S}$, ktorej prvky vyjadrujú rozptýlenosť jednotlivých premenných a závislosti medzi nimi. Celkovo tak výberová kovariančná matica reprezentuje tvar dát. Pre dáta



Obr. 2.3: Pôvodné dáta (vyplnené krúžky) a dáta vzniknuté ich centrovaním (štvorce).

$\mathbf{x}_1, \ldots, \mathbf{x}_n \in \mathbb{R}^d$ $(n \geq 2)$ je to $d \times d$ matica, ktorá má na diagonále tzv. *výberové disperzie* (*rozptyly*) $s_1^2, \ldots, s_d^2$ a mimo diagonály tzv. *výberové kovariancie* $c_{k\ell}$ $(k, \ell = 1, \ldots, d)$:

$$\mathbf{S} = \begin{pmatrix} s_1^2 & c_{12} & \ldots & c_{1d} \\ c_{12} & s_2^2 & \ldots & c_{2d} \\ \vdots & \vdots & & \vdots \\ c_{1d} & c_{2d} & \ldots & s_d^2 \end{pmatrix}.$$

Výberová disperzia s_k^2 vyjadruje mieru rozptýlenosti k-tej premennej (čiže k-tej zložky vektorov $\mathbf{x}_1, \ldots, \mathbf{x}_n$), $k = 1, \ldots, d$, zatiaľ čo výberová kovariancia $c_{k\ell}$ medzi k-tou a ℓ-tou premennou vyjadruje veľkosť (lineárnej) závislosti medzi týmito premennými. Poznamenajme, že $c_{k\ell} = c_{\ell k}$. Kladná vysoká (relatívne voči disperziám s_k^2 a s_ℓ^2) výberová kovariancia naznačuje, že pri vyšších hodnotách k-tej premennej má ℓ-tá premenná tiež tendenciu nadobúdať vyššie hodnoty; pri nižších nižšie. Naopak, výrazne záporná výberová kovariancia značí, že pri vyšších hodnotách jednej premennej nadobúda druhá premenná nižšie hodnoty. Nízka výberová kovariancia (t.j. blízka nule, relatívne voči disperziám – teda v absolútnej hodnote výrazne menšia než disperzie) vyjadruje, že medzi danými premennými nie je výrazná závislosť.

Formálne je matica $\mathbf{S}$ daná predpisom

$$\mathbf{S} = \frac{1}{n-1} \sum_{i=1}^{n} (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^T, \quad \text{kde } \bar{\mathbf{x}} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{x}_i.$$

Vektor $\bar{\mathbf{x}} \in \mathbb{R}^d$ je *výberový priemer*, ktorého k-ty prvok je rovný priemernej hodnote k-tej premennej (t.j. priemernej hodnote k-tych prvkov vektorov $\mathbf{x}_1, \ldots, \mathbf{x}_n$), $k = 1, \ldots, d$.

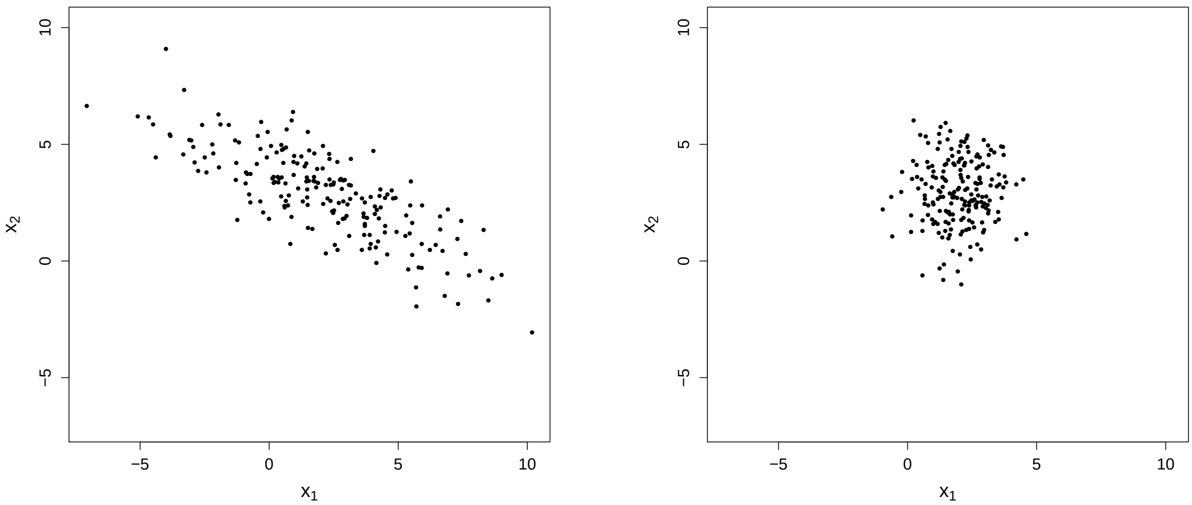
Príklad 2.57. Overte, že ak dáta vyjadríme v matici $\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_n)$, tak výberový priemer vieme vypočítať ako $\bar{\mathbf{x}} = n^{-1}\mathbf{X}\mathbf{1}_n$ a výberovú kovariančnú maticu ako

$$\begin{aligned} \mathbf{S} &= \frac{1}{n-1}(\mathbf{X} - \bar{\mathbf{x}}\mathbf{1}_n^T)(\mathbf{X} - \bar{\mathbf{x}}\mathbf{1}_n^T)^T = \frac{1}{n-1}(\mathbf{X} - n^{-1}\mathbf{X}\mathbf{J}_{n\times n})(\mathbf{X} - n^{-1}\mathbf{X}\mathbf{J}_{n\times n})^T \\ &= \frac{1}{n-1}\mathbf{X}(\mathbf{I}_n - n^{-1}\mathbf{J}_{n\times n})\mathbf{X}^T. \end{aligned}$$

Príklad 2.58. Výberová kovariančná matica vypočítaná z dát na obrázku 2.4a je

$$\mathbf{S} = \begin{pmatrix} 9{,}4 & -4{,}6 \\ -4{,}6 & 3{,}7 \end{pmatrix}.$$

Vidíme, že pozorovania sú skutočne viac rozptýlené v prvej súradnici než v druhej ($s_1^2 = 9{,}4 > 3{,}7 = s_2^2$) a tiež si môžeme všimnúť výraznú zápornú závislosť medzi premennými - keď jedna z nich je vyššia, tak druhá z nich nadobúda vo všeobecnosti nižšie hodnoty, čo potvrdzuje relatívne vysoká záporná hodnota $c_{12} = -4{,}6$. Skúste odhadnúť, ako asi bude vyzerať výberová kovariančná matica pre dáta na obrázku 2.4b (využite, že na obrázkoch 2.4a a 2.4b sú rovnaké rozsahy na súradnicových osiach). Následne váš odhad porovnajte so skutočnosťou[1].



(a) Prvé dáta. (b) Druhé dáta.

Obr. 2.4: Dáta z príkladu 2.58.

Príklad 2.59. Interpretačne jednoduchší než kovariancia je tzv. *výberový korelačný koeficient* medzi k-tou a ℓ-tou premennou definovaný ako $r_{k\ell} = c_{k\ell}/\sqrt{s_k^2 s_\ell^2}$. Korelačný koeficient nadobúda hodnoty iba v intervale $[-1, 1]$, pričom hodnoty blízko -1 a 1 hovoria o silnej zápornej (kladnej) lineárnej závislosti a hodnoty blízko 0 hovoria o veľmi slabej lineárnej závislosti príslušných premenných. *Výberová korelačná matica* je

$$\mathbf{R} = \begin{pmatrix} 1 & r_{12} & \dots & r_{1d} \\ r_{12} & 1 & \dots & r_{2d} \\ \vdots & \vdots & & \vdots \\ r_{1d} & r_{2d} & \dots & 1 \end{pmatrix}.$$

Ukážte, že $\mathbf{R}$ sa dá vyjadriť maticovým výrazom závislým od $\mathbf{S}$. Vypočítajte tiež výberovú korelačnú maticu pre príklad 2.58.

[1]Skutočné hodnoty sú $s_1^2 = 1{,}02$, $c_{12} = 0{,}03$, $s_2^2 = 1{,}99$.

2.7 Úlohy na precvičenie

Úloha 2.1. Nech
$$\mathbf{A} = \begin{pmatrix} 4 & 8 & 20 \\ 12 & 16 & 24 \end{pmatrix}.$$

Určte, s ktorými z nasledujúcich matíc sa dá matica $\mathbf{A}$ násobiť a) zľava, b) sprava. Určte tiež rozmery výsledných matíc.

$$\mathbf{B} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 2 \end{pmatrix}, \quad \mathbf{D} = \begin{pmatrix} 1 & 4 & 5 & 0 \\ 8 & 7 & 6 & 1 \\ -2 & 7 & 2 & 1 \end{pmatrix}, \quad \mathbf{E} = \begin{pmatrix} 1 & -1 \\ 0 & 0 \\ 3 & 5 \\ 0 & 0 \end{pmatrix}$$

$$\mathbf{F} = 5, \quad \mathbf{G} = \begin{pmatrix} 1 & 2 & 3 \end{pmatrix}, \quad \mathbf{H} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad \mathbf{K} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \quad \mathbf{L} = \begin{pmatrix} -1 & 1 \end{pmatrix}$$

Úloha 2.2. Ukážte, že pre všetky $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n\times n}$ platí: a) $(\mathbf{A}+\mathbf{B})(\mathbf{A}-\mathbf{B}) = \mathbf{A}^2 - \mathbf{B}^2$ vtedy a len vtedy, keď matice $\mathbf{A}, \mathbf{B}$ komutujú; b) ak sú matice $\mathbf{A}, \mathbf{B}, \mathbf{AB} - \mathbf{BA}$ symetrické, tak $\mathbf{A}$ a $\mathbf{B}$ komutujú.

Úloha 2.3. Overte, že súčet dvoch symetrických matíc rovnakých rozmerov je symetrická matica a že súčin dvoch symetrických matíc, ktoré komutujú, je symetrická matica. Nájdite príklad dvojice symetrických matíc rovnakých rozmerov, ktorých súčin nie je symetrická matica. Overte tiež, že ak $\mathbf{A} \in \mathbb{R}^{m\times n}$, tak $\mathbf{AA}^T$ a $\mathbf{A}^T\mathbf{A}$ sú symetrické matice.

Úloha 2.4. Maticu $\mathbf{A}$ typu $n \times n$ nazývame úplne symetrickou (angl. *completely symmetric*), ak $\mathbf{A} = a\mathbf{I}_n + b\mathbf{J}_{n\times n}$ pre nejaké $a, b \in \mathbb{R}$. Nahliadnite, že úplne symetrická matica je symetrická a že súčet úplne symetrických matíc je úplne symetrická matica. Dokážte, že súčin úplne symetrických matíc je úplne symetrická matica. Presvedčte sa, že úplne symetrické matice komutujú.

Úloha 2.5. Nech $\mathbf{A} \in \mathbb{R}^{n\times n}$ a nech $k \in \mathbb{N}$. Ukážte, že platí
$$\begin{pmatrix} \mathbf{I}_n & \mathbf{0}_{n\times n} \\ \mathbf{A} & \mathbf{I}_n \end{pmatrix}^k = \begin{pmatrix} \mathbf{I}_n & \mathbf{0}_{n\times n} \\ k\mathbf{A} & \mathbf{I}_n \end{pmatrix}.$$

Úloha 2.6. Nech $\mathbf{A}_1, \dots, \mathbf{A}_m$ sú štvorcové matice, nech $k \in \mathbb{N}$ a $\mathbf{A} = \mathrm{diag}(\mathbf{A}_1, \dots, \mathbf{A}_m)$. Vypočítajte $\mathbf{A}^k$.

Úloha 2.7. Nech $\mathbf{x} \in \mathbb{R}^n$, ktorého žiadna zložka nie je nulová. Definujme $\mathbf{y} \in \mathbb{R}^n$, kde $y_i = 1/x_i$, $i = 1, \dots, n$ a tiež
$$\mathbf{A} = \begin{pmatrix} \mathrm{diag}(\mathbf{x}) & \mathbf{x} \\ \mathbf{x}^T & 1 \end{pmatrix}, \quad \mathbf{G} = \begin{pmatrix} \mathrm{diag}(\mathbf{y}) & \mathbf{0}_n \\ \mathbf{0}_n^T & 0 \end{pmatrix}.$$

Vypočítajte $\mathbf{AGA}$.

Úloha 2.8. Nech $\mathbf{x} \in \mathbb{R}^n$ spĺňa $\mathbf{1}_n^T\mathbf{x} = 0$, $\mathbf{x}^T\mathbf{x} = 1$ a definujme

$$\mathbf{A} = \begin{pmatrix} \mathbf{1}_n & \mathbf{I}_n \\ \mathbf{1}_n & \mathbf{x}\mathbf{x}^T \end{pmatrix}.$$

Vypočítajte $\mathbf{A}^T\mathbf{A}$ a $\mathbf{A}\mathbf{A}^T$.

Úloha 2.9. Matica $\mathbf{A} \in \mathbb{R}^{n\times n}$ sa nazýva invertovateľná, ak existuje matica $\mathbf{A}^{-1} \in \mathbb{R}^{n\times n}$, nazývaná inverzia matice $\mathbf{A}$, pre ktorú platí $\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}_n$. H-maticou nazveme každú maticu tvaru

$$\begin{pmatrix} 1 & \mathbf{a}^T & c \\ \mathbf{0}_n & \mathbf{I}_n & \mathbf{b} \\ 0 & \mathbf{0}_n^T & 1 \end{pmatrix},$$

kde $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$, $c \in \mathbb{R}$. Dokážte, že H-matice tvoria grupu, čiže že súčin dvoch H-matíc je H-matica a ku každej H-matici $\mathbf{H}$ existuje H-matica $\mathbf{H}^{-1}$ taká, že $\mathbf{H}\mathbf{H}^{-1} = \mathbf{H}^{-1}\mathbf{H} = \mathbf{I}_{n+2}$.

Úloha 2.10. a) Nech

$$\mathbf{A} = \begin{pmatrix} 0 & 1 \\ 2 & 3 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 3 & -1 & 5 \end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}.$$

Vypočítajte $\mathbf{A} \otimes \mathbf{B}$, $\mathbf{B} \otimes \mathbf{A}$ a $\mathbf{C} \otimes \mathbf{C}$. b) Nech $\mathbf{D} = \operatorname{diag}(d_1, \ldots, d_k)$ a $\mathbf{A} \in \mathbb{R}^{m\times n}$. Ukážte, že $\mathbf{D} \otimes \mathbf{A} = \operatorname{diag}(d_1\mathbf{A}, \ldots, d_k\mathbf{A})$.

Úloha 2.11. Nech

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix} \quad \text{a} \quad \mathbf{B} = \begin{pmatrix} 10 & 2 & 4 & 6 \\ 5 & 1 & 2 & 3 \\ 15 & 3 & 6 & 9 \end{pmatrix}.$$

Vyjadrite $\operatorname{vec}(\mathbf{A})$ pomocou Kroneckerovho súčinu a ukážte, že $\operatorname{vec}(\mathbf{B}) = (5, 1, 2, 3)^T \otimes (2, 1, 3)^T$.

Úloha 2.12. Nech $\mathbf{x} = (-1, 5, 3, 8, 6, 4, 0, 0)^T$. Nájdite $\mathbf{A} \in \mathbb{R}^{2\times 4}$ a $\mathbf{B} \in \mathbb{R}^{4\times 2}$ také, že $\mathbf{x} = \operatorname{vec}(\mathbf{A}) = \operatorname{vec}(\mathbf{B})$.

Úloha* 2.13. Nech $\mathbf{A} \in \mathbb{R}^{n\times n}$ aj $\mathbf{B} \in \mathbb{R}^{n\times n}$. Koľko násobení a sčítaní reálnych čísiel je potrebných, aby sme vypočítali súčin $\mathbf{AB}$? Najprv sa nad touto otázkou zamyslite samostatne a potom si vyhľadajte odpoveď pomocou kľúčových slov “Matrix multiplication algorithm”. Zamyslite sa tiež nad tým, ako je možné efektívne vypočítať $\mathbf{A}^k$, kde k je veľké číslo.

Úloha 2.14. Nech $\mathbf{u} = (1, 1)^T$, $\mathbf{v} = (1, 0)^T$, $\mathbf{w} = (-1, 2)^T$. Opíšte alebo nakreslite nasledovné podpriestory a určte ich dimenzie:

a) $\operatorname{span}(\mathbf{u})$

b) $\operatorname{span}(\mathbf{v})$

c) $\operatorname{span}(\mathbf{u}, \mathbf{v})$

d) $\operatorname{span}(\mathbf{u}, \mathbf{w})$

e) $\operatorname{span}(\mathbf{u}, \mathbf{v}, \mathbf{w})$

Úloha 2.15. Nech $\mathbf{x} = (1, 0, 0)^T$, $\mathbf{y} = (0, 1, 0)^T$, $\mathbf{z} = (1, 1, 0)^T$, $\mathbf{w} = (0, 0, 1)^T$ a $\mathbf{q} = (1, 1, 1)^T$. Opíšte alebo nakreslite nasledovné podpriestory a určte ich dimenzie:

a) $\mathrm{span}(\mathbf{x})$

b) $\mathrm{span}(\mathbf{x}, \mathbf{y})$

c) $\mathrm{span}(\mathbf{x}, \mathbf{z})$

d) $\mathrm{span}(\mathbf{x}, \mathbf{y}, \mathbf{z})$

e) $\mathrm{span}(\mathbf{x}, \mathbf{y}, \mathbf{w})$

f) $\mathrm{span}(\mathbf{x}, \mathbf{y}, \mathbf{q})$

Kapitola 3

Stĺpcový a nulový priestor matice

3.1 Stĺpcový priestor matice

Definícia 3.1 (Stĺpcový priestor matice)**.** Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$. *Stĺpcový priestor* matice $\mathbf{A}$ je vektorový podpriestor $\mathcal{C}(\mathbf{A})$ generovaný stĺpcami tejto matice:

$$\mathcal{C}(\mathbf{A}) = \{\mathbf{y} \in \mathbb{R}^m \mid \text{existuje } \mathbf{x} \in \mathbb{R}^n \text{ také, že } \mathbf{y} = \mathbf{A}\mathbf{x}\}.$$

Poznámka 3.2. Stĺpcový priestor sa značí $\mathcal{C}(\mathbf{A})$ podľa anglického pomenovania *column space*, niekedy sa značí aj $\mathcal{M}(\mathbf{A})$. Je skutočne generovaný stĺpcami matice $\mathbf{A}$, keďže pozostáva zo všetkých vektorov $\mathbf{y} = \mathbf{A}\mathbf{x}$ pre $\mathbf{x} \in \mathbb{R}^n$, čiže zo všetkých lineárnych kombinácií $x_1\mathbf{a}_1 + \ldots + x_n\mathbf{a}_n$, kde $\mathbf{a}_1, \ldots, \mathbf{a}_n$ sú stĺpce matice a $x_1, \ldots, x_n \in \mathbb{R}$. Z toho tiež vyplýva, že $\mathcal{C}(\mathbf{A})$ je naozaj vektorový podpriestor.

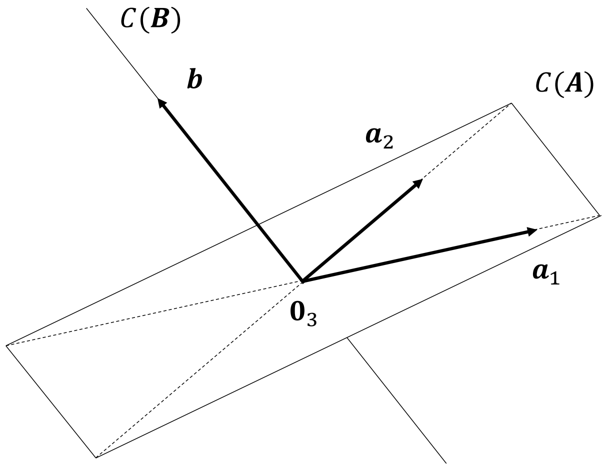
Príklad 3.3. Stĺpcové priestory matíc z $\mathbb{R}^{3\times n}$ sú vektorové podpriestory v $\mathbb{R}^3$. Sú to teda: nulový vektor $\mathbf{0}_3$ (formálne presne: $\{\mathbf{0}_3\}$); priamka alebo rovina prechádzajúca vektorom $\mathbf{0}_3$; alebo celý priestor $\mathbb{R}^3$. Na obrázku 3.1 sú znázornené priestory $\mathcal{C}(\mathbf{A})$ a $\mathcal{C}(\mathbf{B})$ pre maticu $\mathbf{A}$ s dvoma lineárne nezávislými stĺpcami $\mathbf{a}_1, \mathbf{a}_2 \in \mathbb{R}^3$ a pre maticu $\mathbf{B}$ s iba jedným stĺpcom $\mathbf{b}$ (ktorý je tiež lineárne nezávislý s $\mathbf{a}_1, \mathbf{a}_2$). Stĺpcový priestor matice $(\mathbf{A}, \mathbf{B}) = (\mathbf{a}_1, \mathbf{a}_2, \mathbf{b})$ je potom celé $\mathbb{R}^3$.

Určte, čo musia spĺňať stĺpce matice $\mathbf{A} \in \mathbb{R}^{3\times n}$, aby $\mathcal{C}(\mathbf{A})$ bol postupne $\{\mathbf{0}_3\}$, priamkou, rovinou, či $\mathbb{R}^3$. Opíšte tiež všetky možné stĺpcové priestory matíc z $\mathbb{R}^{2\times n}$ a určte, čo musia spĺňať matice z $\mathbb{R}^{2\times n}$, aby ich stĺpcové priestory boli jednotlivých opísaných typov.

Veta 3.4 (Základná veta o inklúzii stĺpcových priestorov)**.** Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$ a $\mathbf{B} \in \mathbb{R}^{m\times p}$. Potom $\mathcal{C}(\mathbf{B}) \subseteq \mathcal{C}(\mathbf{A})$ vtedy a len vtedy, keď $\mathbf{B} = \mathbf{A}\mathbf{X}$ pre nejakú maticu $\mathbf{X} \in \mathbb{R}^{n\times p}$.

Dôkaz. Nech $\mathcal{C}(\mathbf{B}) \subseteq \mathcal{C}(\mathbf{A})$. Nech $\mathbf{b}_1, \ldots, \mathbf{b}_p \in \mathbb{R}^m$ sú stĺpce matice $\mathbf{B}$. Keďže $\mathbf{b}_1, \ldots, \mathbf{b}_p \in \mathcal{C}(\mathbf{B})$, čiže $\mathbf{b}_1, \ldots, \mathbf{b}_p \in \mathcal{C}(\mathbf{A})$, tak existujú $\mathbf{x}_1, \ldots, \mathbf{x}_p \in \mathbb{R}^n$ také, že $\mathbf{b}_1 = \mathbf{A}\mathbf{x}_1, \ldots, \mathbf{b}_p = \mathbf{A}\mathbf{x}_p$, čo sa dá napísať v tvare $\mathbf{B} = \mathbf{A}\mathbf{X}$, kde $\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_p)$.

Nech $\mathbf{B} = \mathbf{A}\mathbf{X}$ pre nejakú maticu $\mathbf{X} \in \mathbb{R}^{n\times p}$ a nech $\mathbf{y} \in \mathcal{C}(\mathbf{B})$. Potom $\mathbf{y} = \mathbf{B}\mathbf{z}$ pre nejaké $\mathbf{z} \in \mathbb{R}^p$, z čoho vyplýva $\mathbf{y} = (\mathbf{A}\mathbf{X})\mathbf{z} = \mathbf{A}(\mathbf{X}\mathbf{z})$, takže $\mathbf{y} \in \mathcal{C}(\mathbf{A})$. Keďže $\mathbf{y}$ bol ľubovoľný prvok stĺpcového priestoru matice $\mathbf{B}$, dostali sme $\mathcal{C}(\mathbf{B}) \subseteq \mathcal{C}(\mathbf{A})$. □



Obr. 3.1: Stĺpcové priestory matíc $\mathbf{A}$ a $\mathbf{B}$ z príkladu 3.3.

3.2 Hodnosť matice

Definícia 3.5 (Hodnosť matice)**.** *Hodnosť matice* $\mathbf{A} \in \mathbb{R}^{m\times n}$ je dimenzia priestoru $\mathcal{C}(\mathbf{A})$, značíme ju $\operatorname{rank}(\mathbf{A})$.

Poznámka 3.6. V slovenských textoch sa na označenie hodnosti matice $\mathbf{A}$ niekedy tiež používa symbol $h(\mathbf{A})$.

Lema 3.7 (Rozklad matice pomocou hodnosti)**.** Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$ spĺňa $\mathbf{A} \neq \mathbf{0}_{m\times n}$ a označme $\dim(\mathcal{C}(\mathbf{A})) = c$, $\dim(\mathcal{C}(\mathbf{A}^T)) = r$. Potom

(i) existuje $\mathbf{B} \in \mathbb{R}^{m\times c}$, $\mathbf{L} \in \mathbb{R}^{c\times n}$ také, že $\mathbf{A} = \mathbf{BL}$;

(ii) existuje $\mathbf{K} \in \mathbb{R}^{m\times r}$, $\mathbf{T} \in \mathbb{R}^{r\times n}$ také, že $\mathbf{A} = \mathbf{KT}$.

Dôkaz. (i) Nech $\mathbf{b}_1, \dots, \mathbf{b}_c$ je báza $\mathcal{C}(\mathbf{A})$ a definujme $\mathbf{B} = (\mathbf{b}_1, \dots, \mathbf{b}_c)$. Potom $\mathcal{C}(\mathbf{A}) = \mathcal{C}(\mathbf{B})$ a podľa vety 3.4 existuje $\mathbf{L} \in \mathbb{R}^{c\times n}$ také, že $\mathbf{A} = \mathbf{BL}$.

(ii) Analogicky, nech $\mathbf{t}_1, \dots, \mathbf{t}_r$ je báza $\mathcal{C}(\mathbf{A}^T)$ a definujme $\mathbf{T} = (\mathbf{t}_1, \dots, \mathbf{t}_r)^T$. Potom $\mathcal{C}(\mathbf{A}^T) = \mathcal{C}(\mathbf{T}^T)$ a podľa vety 3.4 existuje $\mathbf{X} \in \mathbb{R}^{r\times m}$ také, že $\mathbf{A}^T = \mathbf{T}^T\mathbf{X}$, z čoho vyplýva $\mathbf{A} = \mathbf{X}^T\mathbf{T}$. Dôkaz ukončíme položením $\mathbf{K} = \mathbf{X}^T$. □

Poznámka 3.8 (Riadkový priestor matice)**.** Priestor $\mathcal{C}(\mathbf{A}^T)$ je generovaný riadkami matice $\mathbf{A}$, a preto sa nazýva tiež *riadkový priestor* (angl. *row space*) matice $\mathbf{A}$.

Veta 3.9 (O hodnosti matíc $\mathbf{A}$ a $\mathbf{A}^T$)**.** Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$. Potom $\operatorname{rank}(\mathbf{A}) = \operatorname{rank}(\mathbf{A}^T)$.

Dôkaz. Označme $\operatorname{rank}(\mathbf{A}) = \dim(\mathcal{C}(\mathbf{A})) = c$ a $\operatorname{rank}(\mathbf{A}^T) = \dim(\mathcal{C}(\mathbf{A}^T)) = r$. Podľa lemy 3.7 existujú matice $\mathbf{B}$, $\mathbf{L}$, $\mathbf{K}$, $\mathbf{T}$ spĺňajúce (i), (ii). Potom podľa vety 3.4 platí $\mathcal{C}(\mathbf{A}) \subseteq \mathcal{C}(\mathbf{K})$, a teda $c \leq \dim(\mathcal{C}(\mathbf{K}))$. Zároveň $\mathbf{A}^T = \mathbf{L}^T\mathbf{B}^T$, čiže $\mathcal{C}(\mathbf{A}^T) \subseteq \mathcal{C}(\mathbf{L}^T)$ a $r \leq \dim(\mathcal{C}(\mathbf{L}^T))$. Keďže $\mathbf{K}$ má r stĺpcov, tak $c \leq \dim(\mathcal{C}(\mathbf{K})) \leq r$ a $\mathbf{L}^T$ má c stĺpcov, čiže $r \leq \dim(\mathcal{C}(\mathbf{L}^T)) \leq c$. Dostávame $c = r$. □

Dôsledok 3.10. Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$. Potom $\operatorname{rank}(\mathbf{A}) = \operatorname{rank}(\mathbf{A}^T) \leq \min\{m, n\}$.

Veta 3.11 (**BT** rozklad matice)**.** Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$ spĺňa $\mathbf{A} \neq \mathbf{0}_{m\times n}$, $\text{rank}(\mathbf{A}) = c$. Potom existujú $\mathbf{B} \in \mathbb{R}^{m\times c}$ a $\mathbf{T} \in \mathbb{R}^{c\times n}$ také, že $\mathbf{A} = \mathbf{BT}$. Navyše pre akékoľvek matice $\mathbf{B} \in \mathbb{R}^{m\times c}$ a $\mathbf{T} \in \mathbb{R}^{c\times n}$ spĺňajúce $\mathbf{A} = \mathbf{BT}$, platí $\text{rank}(\mathbf{B}) = \text{rank}(\mathbf{T}) = c$.

Dôkaz. Existencia rozkladu vyplýva z lemy 3.7. Predpokladajme, že $\mathbf{A} = \mathbf{BT}$ pre $\mathbf{B} \in \mathbb{R}^{m\times c}$ a $\mathbf{T} \in \mathbb{R}^{c\times n}$. Potom $\text{rank}(\mathbf{B}) \leq c$, lebo $\mathbf{B}$ má c stĺpcov. Navyše, podľa vety 3.4 platí $\mathcal{C}(\mathbf{A}) \subseteq \mathcal{C}(\mathbf{B})$, a teda

$$c = \text{rank}(\mathbf{A}) = \dim(\mathcal{C}(\mathbf{A})) \leq \dim(\mathcal{C}(\mathbf{B})) = \text{rank}(\mathbf{B}) \leq c,$$

čiže $\text{rank}(\mathbf{B}) = c$. Dôkaz pre hodnosť $\mathbf{T}$ je analogický. □

Dôsledok 3.12. Ak $\mathbf{A} \in \mathbb{R}^{m\times n}$ a $\text{rank}(\mathbf{A}) = 1$, tak $\mathbf{A}$ vieme zapísať ako $\mathbf{A} = \mathbf{u}\mathbf{v}^T$, kde $\mathbf{u} \in \mathbb{R}^m$ a $\mathbf{v} \in \mathbb{R}^n$ sú nenulové vektory.

Lema 3.13 (Hodnosť nulovej matice)**.** Matica $\mathbf{A} \in \mathbb{R}^{m\times n}$ má nulovú hodnosť vtedy a len vtedy, keď $\mathbf{A} = \mathbf{0}_{m\times n}$.

Dôkaz. Ak $\mathbf{A} = \mathbf{0}_{m\times n}$, tak $\mathcal{C}(\mathbf{A}) = \mathbf{0}_m$, a teda $\text{rank}(\mathbf{A}) = 0$. Naopak, nech $\text{rank}(\mathbf{A}) = 0$, čiže $\dim(\mathcal{C}(\mathbf{A})) = 0$, čo nastáva podľa definície 1.25 iba ak $\mathcal{C}(\mathbf{A}) = \mathbf{0}_m$. Stĺpcový priestor pozostávajúci iba z vektora $\mathbf{0}_m$ má zjavne iba nulová matica. □

Veta 3.14 (O hodnosti súčinu matíc)**.** Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$ a $\mathbf{B} \in \mathbb{R}^{n\times k}$. Potom $\text{rank}(\mathbf{AB}) \leq \min\{\text{rank}(\mathbf{A}), \text{rank}(\mathbf{B})\}$.

Dôkaz. Podľa vety 3.4 vieme, že $\mathcal{C}(\mathbf{AB}) \subseteq \mathcal{C}(\mathbf{A})$, a teda $\text{rank}(\mathbf{AB}) \leq \text{rank}(\mathbf{A})$. Analogicky $\mathcal{C}(\mathbf{B}^T\mathbf{A}^T) \subseteq \mathcal{C}(\mathbf{B}^T)$, a teda $\text{rank}(\mathbf{AB}) = \text{rank}(\mathbf{B}^T\mathbf{A}^T) \leq \text{rank}(\mathbf{B}^T) = \text{rank}(\mathbf{B})$. □

Príklad 3.15. Nájdite matice $\mathbf{A}$, $\mathbf{B}$ také, že $\text{rank}(\mathbf{AB}) < \min\{\text{rank}(\mathbf{A}), \text{rank}(\mathbf{B})\}$. Nájdite tiež také matice $\mathbf{A}$, $\mathbf{B}$, že $\text{rank}(\mathbf{AB}) = \min\{\text{rank}(\mathbf{A}), \text{rank}(\mathbf{B})\}$.

Definícia 3.16 (Plná hodnosť matice, regulárna matica)**.** Ak $\mathbf{A} \in \mathbb{R}^{m\times n}$ a $\text{rank}(\mathbf{A}) = m$, tak hovoríme, že $\mathbf{A}$ má *plnú riadkovú hodnosť*. Ak $\text{rank}(\mathbf{A}) = n$, tak hovoríme, že $\mathbf{A}$ má *plnú stĺpcovú hodnosť*. Ak $m = n$ a $\mathbf{A}$ má plnú stĺpcovú hodnosť, tak hovoríme, že $\mathbf{A}$ je *regulárna*. Ak $m = n$ a $\mathbf{A}$ nemá plnú stĺpcovú hodnosť, tak $\mathbf{A}$ je *singulárna*.

Poznámka 3.17. Ak štvorcová matica $\mathbf{A}$ má plnú stĺpcovú hodnosť (a teda ak je regulárna), hovoríme skrátene, že $\mathbf{A}$ má *plnú hodnosť*.

Poznámka 3.18. Ak $\mathbf{A}$ je regulárna, tak má nielen plnú stĺpcovú, ale aj plnú riadkovú hodnosť.

3.3 Stĺpcový priestor a hodnosť blokovej matice

Poznámka 3.19. Stĺpcový priestor matice $\mathbf{C} = (\mathbf{A}, \mathbf{B})$ značíme skrátene $\mathcal{C}(\mathbf{C}) = \mathcal{C}(\mathbf{A}, \mathbf{B})$. Podobne $\text{rank}(\mathbf{C}) = \text{rank}(\mathbf{A}, \mathbf{B})$.

Veta 3.20 (Stĺpcový priestor a hodnosť blokovej matice)**.** Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{B} \in \mathbb{R}^{m\times k}$. Potom

(i) $\mathcal{C}(\mathbf{A}) \subseteq \mathcal{C}(\mathbf{A}, \mathbf{B})$, $\mathcal{C}(\mathbf{B}) \subseteq \mathcal{C}(\mathbf{A}, \mathbf{B})$,

(ii) $\text{rank}(\mathbf{A}) \leq \text{rank}(\mathbf{A}, \mathbf{B})$, $\text{rank}(\mathbf{B}) \leq \text{rank}(\mathbf{A}, \mathbf{B})$,

(iii) $\mathcal{C}(\mathbf{B}, \mathbf{A}) = \mathcal{C}(\mathbf{A}, \mathbf{B})$, $\operatorname{rank}(\mathbf{B}, \mathbf{A}) = \operatorname{rank}(\mathbf{A}, \mathbf{B})$,

(iv) $\operatorname{rank}(\mathbf{A}, \mathbf{B}) \leq \operatorname{rank}(\mathbf{A}) + \operatorname{rank}(\mathbf{B})$.

Nech $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m\times n}$. Potom

(v) $\mathcal{C}(\mathbf{A} + \mathbf{B}) \subseteq \mathcal{C}(\mathbf{A}, \mathbf{B})$, $\operatorname{rank}(\mathbf{A} + \mathbf{B}) \leq \operatorname{rank}(\mathbf{A}, \mathbf{B})$,

(vi) $\operatorname{rank}(\mathbf{A} + \mathbf{B}) \leq \operatorname{rank}(\mathbf{A}) + \operatorname{rank}(\mathbf{B})$.

Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{B} \in \mathbb{R}^{p\times q}$. Potom

(vii)
$$\operatorname{rank}\begin{pmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{B} \end{pmatrix} = \operatorname{rank}(\mathbf{A}) + \operatorname{rank}(\mathbf{B}).$$

Dôkaz. Dôkazy tvrdení (i), (iii) sú triviálne. Tvrdenie (ii) je dôsledkom (i).

Na dôkaz tvrdenia (iv) definujme $s = \operatorname{rank}(\mathbf{A})$, $r = \operatorname{rank}(\mathbf{B})$, $\mathbf{A}^* = (\mathbf{a}_1^*, \dots, \mathbf{a}_s^*)$ a $\mathbf{B}^* = (\mathbf{b}_1^*, \dots, \mathbf{b}_r^*)$, kde $\mathbf{a}_1^*, \dots, \mathbf{a}_s^*$ je báza $\mathcal{C}(\mathbf{A})$ a $\mathbf{b}_1^*, \dots, \mathbf{b}_r^*$ je báza $\mathcal{C}(\mathbf{B})$. Potom zrejme $\mathcal{C}(\mathbf{A}, \mathbf{B}) = \mathcal{C}(\mathbf{A}^*, \mathbf{B}^*)$ a $\operatorname{rank}(\mathbf{A}, \mathbf{B}) = \operatorname{rank}(\mathbf{A}^*, \mathbf{B}^*) \leq r+s = \operatorname{rank}(\mathbf{A})+\operatorname{rank}(\mathbf{B})$, keďže matica $(\mathbf{A}^*, \mathbf{B}^*)$ má $r + s$ stĺpcov.

Tvrdenie (v) vyplýva z vety 3.4, keďže $\mathbf{A} + \mathbf{B} = (\mathbf{A}, \mathbf{B})(\mathbf{I}_n, \mathbf{I}_n)^T$. Tvrdenie (vi) vyplýva z (v) a (iv).

Pre dôkaz (vii) definujme $\mathbf{A}^*$ a $\mathbf{B}^*$ ako v (iv) a tiež

$$\mathbf{C}^* = \begin{pmatrix} \mathbf{A}^* & \mathbf{0} \\ \mathbf{0} & \mathbf{B}^* \end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{B} \end{pmatrix}$$

Potom zrejme $\mathcal{C}(\mathbf{C}^*) = \mathcal{C}(\mathbf{C})$, a teda tieto matice majú aj rovnakú hodnosť. Zároveň

$$\begin{pmatrix} \mathbf{A}^* & \mathbf{0} \\ \mathbf{0} & \mathbf{B}^* \end{pmatrix} \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{0}_m \\ \mathbf{0}_p \end{pmatrix}$$

nastáva vtedy a len vtedy, keď $\mathbf{A}^*\mathbf{x}_1 = \mathbf{0}_m$ a $\mathbf{B}^*\mathbf{x}_2 = \mathbf{0}_p$, čo je práve vtedy, keď $\mathbf{x}_1 = \mathbf{0}_s$ a $\mathbf{x}_2 = \mathbf{0}_r$, keďže stĺpce $\mathbf{A}^*$ (a takisto $\mathbf{B}^*$) sú lineárne nezávislé. Z toho vyplýva, že stĺpce matice $\mathbf{C}^*$ sú lineárne nezávislé, čiže jej hodnosť je rovná počtu jej stĺpcov, $\operatorname{rank}(\mathbf{C}^*) = s + r$. Teda $\operatorname{rank}(\mathbf{C}) = s + r = \operatorname{rank}(\mathbf{A}) + \operatorname{rank}(\mathbf{B})$. □

Lema 3.21 (Submatica plnej hodnosti)**.** Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$ má hodnosť $0 < r < \min\{m, n\}$. Nech $1 \leq i_1 \leq \ldots \leq i_r \leq m$ sú také indexy, že riadky $i_1, \dots, i_r$ matice $\mathbf{A}$ sú lineárne nezávislé a nech $1 \leq j_1 \leq \ldots \leq j_r \leq n$ sú také indexy, že stĺpce $j_1, \dots, j_r$ matice $\mathbf{A}$ sú lineárne nezávislé. Potom matica $\mathbf{B} \in \mathbb{R}^{r\times r}$, ktorú z $\mathbf{A}$ získame odstránením všetkých riadkov okrem riadkov $i_1, \dots, i_r$ a všetkých stĺpcov okrem stĺpcov $j_1, \dots, j_r$, je regulárna.

Dôkaz. Pre jednoduchosť značenia predpokladajme, že máme $\{i_1, \dots, i_r\} = \{j_1, \dots, j_r\} = \{1, \dots, r\}$, teda že $\mathbf{B}$ sa dá vyjadriť ako blok $\mathbf{A}_{11}$ v blokovom zápise

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix},$$

kde $\mathbf{A}_{11} \in \mathbb{R}^{r\times r}$. Keďže prvých r riadkov $\mathbf{A}$ je lineárne nezávislých a ich počet je rovný hodnosti $\mathcal{C}(\mathbf{A}^T)$, tak tvoria bázu $\mathcal{C}(\mathbf{A}^T)$. Čiže každý riadok matice $\mathbf{A}$ sa dá zapísať ako nejaká lineárna kombinácia jej prvých r riadkov; existuje teda matica $\mathbf{C} \in \mathbb{R}^{m\times r}$, že $\mathbf{A} = \mathbf{C}(\mathbf{A}_{11}, \mathbf{A}_{12})$. To môžeme zapísať ako

$$\begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix} = \begin{pmatrix} \mathbf{C}\mathbf{A}_{11}, & \mathbf{C}\mathbf{A}_{12} \end{pmatrix}.$$

Keď označíme $\mathbf{S} \in \mathbb{R}^{m\times r}$ ako maticu tvorenú prvými r stĺpcami matice $\mathbf{A}$, tak dostávame $\mathbf{S} = \mathbf{C}\mathbf{A}_{11}$. Keďže stĺpce $\mathbf{S}$ sú nezávislé, tak $\operatorname{rank}(\mathbf{S}) = r$ a z vety 3.14 potom plynie $\operatorname{rank}(\mathbf{A}_{11}) \geq \operatorname{rank}(\mathbf{S}) = r$. Zároveň platí $\operatorname{rank}(\mathbf{A}_{11}) \leq r$, lebo $\mathbf{A}_{11} \in \mathbb{R}^{r\times r}$. Teda $\operatorname{rank}(\mathbf{A}_{11}) = r$. □

3.4 Nulový priestor matice

Definícia 3.22 (Nulový priestor matice)**.** *Nulovým priestorom* matice $\mathbf{A} \in \mathbb{R}^{m\times n}$ nazývame množinu $\mathcal{N}(\mathbf{A}) = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} = \mathbf{0}_m\}$.

Poznámka 3.23. Označenie $\mathcal{N}(\mathbf{A})$ vychádza z anglického názvu *null space*. Nulový priestor matice $\mathbf{A}$ sa niekedy označuje $\operatorname{Ker}(\mathbf{A})$ a nazýva sa jadro zobrazenia definovaného maticou $\mathbf{A}$, resp. kernel $\mathbf{A}$.

Lema 3.24 (Nulový priestor je lineárny podpriestor)**.** Pre každú maticu $\mathbf{A} \in \mathbb{R}^{m\times n}$ je $\mathcal{N}(\mathbf{A})$ lineárny podpriestor priestoru $\mathbb{R}^n$.

Dôkaz. Overme vlastnosti lineárneho podpriestoru podľa vety 1.7. Nech $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{N}(\mathbf{A})$ a $c_1, c_2 \in \mathbb{R}$. Potom $\mathbf{A}(c_1\mathbf{x}_1 + c_2\mathbf{x}_2) = c_1\mathbf{A}\mathbf{x}_1 + c_2\mathbf{A}\mathbf{x}_2 = \mathbf{0}_m$, čiže $c_1\mathbf{x}_1 + c_2\mathbf{x}_2 \in \mathcal{N}(\mathbf{A})$. □

Lema 3.25 (Nulový priestor nulovej matice)**.** Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$. Potom $\mathcal{N}(\mathbf{A}) = \mathbb{R}^n$ práve vtedy, keď $\mathbf{A} = \mathbf{0}_{m\times n}$.

Dôkaz. Označme $\mathbf{A} = (\mathbf{a}_1, \ldots, \mathbf{a}_n)$, $\mathbf{a}_i \in \mathbb{R}^m$, $i = 1, \ldots, n$. Nech $\mathcal{N}(\mathbf{A}) = \mathbb{R}^n$. Potom $\mathbf{A}\mathbf{x} = \mathbf{0}_m$ pre každé $\mathbf{x} \in \mathbb{R}^n$. Špeciálne pre $\mathbf{e}_i$ dostávame $\mathbf{0}_m = \mathbf{A}\mathbf{e}_i = \mathbf{a}_i$, $i = 1, \ldots, n$. To spolu dáva $\mathbf{A} = \mathbf{0}_{m\times n}$. □

Lema 3.26 (O ortogonálnosti nulového vektora)**.** Nech $\mathbf{z} \in \mathbb{R}^n$. Potom $\mathbf{z} = \mathbf{0}_n$ vtedy a len vtedy, keď $\mathbf{y}^T\mathbf{z} = 0$ pre každé $\mathbf{y} \in \mathbb{R}^n$.

Dôkaz. Nech $\mathbf{y}^T\mathbf{z} = 0$ pre každé $\mathbf{y} \in \mathbb{R}^n$. Potom aj $\mathbf{z}^T\mathbf{z} = 0$ a z kladnej definitnosti skalárneho súčinu (definícia 1.29, príklad 1.31) vyplýva, že $\mathbf{z} = \mathbf{0}_n$. Opačná implikácia je zjavná. □

Lema 3.27 (Nulovosť symetrickej matice)**.** Nech $\mathbf{A} \in \mathcal{S}^n$. Potom $\mathbf{A} = \mathbf{0}_{n\times n}$ práve vtedy, keď $\mathbf{x}^T\mathbf{A}\mathbf{x} = 0$ pre každé $\mathbf{x} \in \mathbb{R}^n$.

Dôkaz. Ak $\mathbf{x}^T\mathbf{A}\mathbf{x} = 0$ pre každé $\mathbf{x} \in \mathbb{R}^n$, tak to platí aj pre $\mathbf{e}_i$. Potom $0 = \mathbf{e}_i^T\mathbf{A}\mathbf{e}_i = (\mathbf{A})_{ii}$ pre každé $i = 1, \ldots, n$. Ak následne za $\mathbf{x}$ zvolíme $\mathbf{e}_i + \mathbf{e}_j$ ($j \neq i$), dostávame $(\mathbf{A})_{ii} + (\mathbf{A})_{jj} + (\mathbf{A})_{ij} + (\mathbf{A})_{ji} = 0$. Keďže už sme ukázali, že $(\mathbf{A})_{ii} = (\mathbf{A})_{jj} = 0$, a $\mathbf{A}$ je symetrická, tak máme $2(\mathbf{A})_{ij} = 0$ pre každé $j \neq i$. Spolu teda platí $(\mathbf{A})_{ij} = 0$ pre každé $i, j = 1, \ldots, n$, čiže $\mathbf{A} = \mathbf{0}$. Dôkaz opačnej implikácie je triviálny. □

Príklad 3.28. Nájdite takú nenulovú maticu $\mathbf{A} \in \mathbb{R}^{n\times n}$, že $\mathbf{x}^T\mathbf{A}\mathbf{x} = 0$ pre každé $\mathbf{x} \in \mathbb{R}^n$ pre a) $n = 2$, b) všeobecné n. Viete nájsť aj celú triedu matíc spĺňajúcich $\mathbf{x}^T\mathbf{A}\mathbf{x} = 0$ pre každé $\mathbf{x} \in \mathbb{R}^n$?

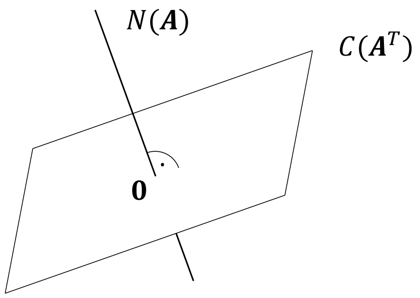
3.4.1 Vzťahy medzi nulovými a stĺpcovými priestormi

Pripomeňme, že $W^\perp$ značí ortogonálny doplnok podpriestoru W podľa definície 1.48.

Veta 3.29 (O vzťahu medzi $\mathcal{N}(\mathbf{A})$ a $\mathcal{C}(\mathbf{A}^T)$)**.** Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$. Potom $\mathcal{N}(\mathbf{A}) = \mathcal{C}(\mathbf{A}^T)^\perp$.

Dôkaz. Nech $\mathbf{x} \in \mathbb{R}^n$. Potom $\mathbf{x} \in \mathcal{N}(\mathbf{A})$ práve vtedy, keď $\mathbf{A}\mathbf{x} = \mathbf{0}_m$, čo podľa lemy 3.26 nastáva práve vtedy, keď $\mathbf{y}^T\mathbf{A}\mathbf{x} = 0$ pre každé $\mathbf{y} \in \mathbb{R}^m$, teda práve vtedy, keď $(\mathbf{A}^T\mathbf{y})^T\mathbf{x} = 0$ pre každé $\mathbf{y} \in \mathbb{R}^m$. Keďže množina vektorov $\mathbf{A}^T\mathbf{y} \in \mathbb{R}^n$ pre všetky $\mathbf{y} \in \mathbb{R}^m$ tvorí priestor $\mathcal{C}(\mathbf{A}^T)$, tak posledná podmienka je ekvivalentná s: $\langle \mathbf{w}, \mathbf{x}\rangle = 0$ pre všetky $\mathbf{w} \in \mathcal{C}(\mathbf{A}^T)$, čo nastáva práve vtedy, keď $\mathbf{x} \in \mathcal{C}(\mathbf{A}^T)^\perp$. $\square$

Zakreslenie vzťahu medzi $\mathcal{N}(\mathbf{A})$ a $\mathcal{C}(\mathbf{A}^T)$ môžeme vidieť na obrázku 3.2.



Obr. 3.2: Schematické znázornenie ortogonálnosti $\mathcal{N}(\mathbf{A})$ a $\mathcal{C}(\mathbf{A}^T)$.

Dôsledok 3.30. Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$. Potom $\mathcal{N}(\mathbf{A}^T) = \mathcal{C}(\mathbf{A})^\perp$, $\mathcal{N}(\mathbf{A})^\perp = \mathcal{C}(\mathbf{A}^T)$ a $\mathcal{N}(\mathbf{A}^T)^\perp = \mathcal{C}(\mathbf{A})$.

Dôkaz. Druhé tvrdenie platí, keďže $\mathcal{N}(\mathbf{A})^\perp = (\mathcal{C}(\mathbf{A}^T)^\perp)^\perp = \mathcal{C}(\mathbf{A}^T)$ (viď veta 1.49(iii)). Prvé a tretie tvrdenie vychádzajú zo zámeny $\mathbf{A}$ za $\mathbf{A}^T$. $\square$

Dôsledok 3.31. Nech $\mathbf{A} \in \mathbb{R}^{n\times n}$. Potom $\mathbf{A}$ je regulárna práve vtedy, keď $\mathcal{N}(\mathbf{A}) = \{\mathbf{0}_n\}$.

Príklad 3.32. Nech

$$\mathbf{A} = \begin{pmatrix} 3 & -6 \\ 5 & -10 \\ 7 & -14 \end{pmatrix}.$$

Nájdite $\mathcal{N}(\mathbf{A})$ a $\mathcal{C}(\mathbf{A}^T)$, a overte, že tieto dva priestory sú na seba kolmé.

Lema 3.33 (O inklúzii nulových priestorov)**.** Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$ a $\mathbf{B} \in \mathbb{R}^{m\times k}$. Potom $\mathcal{C}(\mathbf{A}) \subseteq \mathcal{C}(\mathbf{B})$ práve vtedy, keď $\mathcal{N}(\mathbf{B}^T) \subseteq \mathcal{N}(\mathbf{A}^T)$.

Dôkaz. Nech $\mathcal{C}(\mathbf{A}) \subseteq \mathcal{C}(\mathbf{B})$. Potom podľa vety 1.49 platí $\mathcal{C}(\mathbf{B})^\perp \subseteq \mathcal{C}(\mathbf{A})^\perp$, a teda $\mathcal{N}(\mathbf{B}^T) \subseteq \mathcal{N}(\mathbf{A}^T)$. Opačná implikácia sa dokazuje analogicky. $\square$

Veta 3.34 (O vzťahu $\mathcal{C}(\mathbf{A})$ a $\mathcal{C}(\mathbf{A}\mathbf{A}^T)$)**.** Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$. Potom $\mathcal{C}(\mathbf{A}) = \mathcal{C}(\mathbf{A}\mathbf{A}^T)$.

Dôkaz. Inklúzia $\mathcal{C}(\mathbf{A}\mathbf{A}^T) \subseteq \mathcal{C}(\mathbf{A})$ platí podľa vety 3.4. Na dôkaz opačnej inklúzie uvažujme $\mathbf{x} \in \mathcal{N}(\mathbf{A}\mathbf{A}^T)$, teda $\mathbf{A}\mathbf{A}^T\mathbf{x} = \mathbf{0}_m$. Potom $\mathbf{x}^T\mathbf{A}\mathbf{A}^T\mathbf{x} = 0$, čiže $\|\mathbf{A}^T\mathbf{x}\| = 0$. Z kladnej definitnosti normy vyplýva $\mathbf{A}^T\mathbf{x} = \mathbf{0}_n$, takže $\mathbf{x} \in \mathcal{N}(\mathbf{A}^T)$. Dostávame $\mathcal{N}(\mathbf{A}\mathbf{A}^T) \subseteq \mathcal{N}(\mathbf{A}^T)$, z čoho vyplýva $\mathcal{C}(\mathbf{A}) \subseteq \mathcal{C}(\mathbf{A}\mathbf{A}^T)$ podľa lemy 3.33. $\square$

Dôsledok 3.35. Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$. Potom

(i) $\mathcal{C}(\mathbf{A}^T) = \mathcal{C}(\mathbf{A}^T\mathbf{A})$.

(ii) Matice $\mathbf{A}$, $\mathbf{A}^T$, $\mathbf{A}\mathbf{A}^T$ a $\mathbf{A}^T\mathbf{A}$ majú rovnakú hodnosť.

Lema 3.36 (Vzťah hodnosti matice a nulového priestoru)**.** Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$. Potom platí $\dim(\mathcal{N}(\mathbf{A})) = n - \operatorname{rank}(\mathbf{A})$.

Dôkaz. Keďže $\mathcal{N}(\mathbf{A}) = \mathcal{C}(\mathbf{A}^T)^\perp$ a $\mathcal{N}(\mathbf{A}) \subseteq \mathbb{R}^n$, tak podľa vety 1.49(iv) platí $\dim(\mathcal{N}(\mathbf{A})) = n - \dim(\mathcal{C}(\mathbf{A}^T)) = n - \operatorname{rank}(\mathbf{A})$ podľa dôsledku 3.35. □

3.4.2 Hodnosť kovariančnej matice

Pri mnohých štatistických metódach je dôležité, aby mala výberová kovariančná matica $\mathbf{S}$ plnú hodnosť. Na analýzu toho, kedy je $\mathbf{S}$ regulárna, je užitočný jej maticový tvar z príkladu 2.57:

$$\mathbf{S} = \frac{1}{n-1}(\mathbf{X} - n^{-1}\mathbf{X}\mathbf{J}_{n\times n})(\mathbf{X} - n^{-1}\mathbf{X}\mathbf{J}_{n\times n})^T.$$

Konštanta $1/(n-1)$ nemá na hodnosť vplyv, čiže tú môžeme zanedbať a pracovať s maticou $\tilde{\mathbf{S}} = (n-1)\mathbf{S}$. Môžeme si všimnúť, že $\tilde{\mathbf{S}} = \mathbf{A}\mathbf{A}^T$, kde $\mathbf{A} = \mathbf{X} - n^{-1}\mathbf{X}\mathbf{J}_{n\times n}$, pričom $\mathbf{A}$ je $d \times n$ matica a $\tilde{\mathbf{S}}$ je rozmerov $d \times d$. Podľa dôsledku 3.35 platí $\operatorname{rank}(\tilde{\mathbf{S}}) = \operatorname{rank}(\mathbf{A})$. Takže výberová kovariančná matica má plnú hodnosť, keď $\operatorname{rank}(\mathbf{A}) = d$, a teda stačí skúmať maticu $\mathbf{A} = \mathbf{X} - n^{-1}\mathbf{X}\mathbf{J}_{n\times n} = \mathbf{X}(\mathbf{I}_n - n^{-1}\mathbf{J}_{n\times n})$.

Keďže $\mathbf{A}$ je rozmerov $d\times n$, tak $\operatorname{rank}(\mathbf{A})$ je nanajvýš $\min\{d, n\}$. Zároveň však platí $\mathbf{A}\mathbf{1}_n = \mathbf{X}(\mathbf{I}_n - n^{-1}\mathbf{J}_{n\times n})\mathbf{1}_n = \mathbf{X}(\mathbf{1}_n - \mathbf{1}_n) = \mathbf{0}_{d\times n}$, a teda $\mathbf{1}_n \in \mathcal{N}(\mathbf{A})$. Teda $\dim(\mathcal{N}(\mathbf{A})) \geq 1$, čiže $\operatorname{rank}(\mathbf{A}) \leq n-1$. Prichádzame tak k záveru, že aby bola výberová kovariančná matica plnej hodnosti, musí platiť $n - 1 \geq d$, resp. $n \geq d + 1$, teda počet pozorovaní musí byť aspoň o 1 vyšší než počet premenných.

Príklad 3.37. Zvoľte ľubovoľnú $d \times 2$ maticu pozorovaní $\mathbf{X}$ a overte, že potom $\operatorname{rank}(\mathbf{S}) \leq 1$. Viete nájsť aj takú nenulovú $d \times 2$ maticu $\mathbf{X}$, že $\operatorname{rank}(\mathbf{S}) = 0$?

Poznámka 3.38. Treba si tiež uvedomiť, že podmienka $n \geq d + 1$ je iba nutná pre plnú hodnosť kovariančnej matice, nie postačujúca. Pre zvolené d, n nájdite takú maticu pozorovaní $\mathbf{X}$, pre ktorú platí $n \geq d+1$, ale $\operatorname{rank}(\mathbf{S}) < d$. Aby $\mathbf{S}$ skutočne nadobudla plnú hodnosť, musí platiť $d = \operatorname{rank}(\mathbf{S}) = \operatorname{rank}(\mathbf{A})$, teda matica $\mathbf{X}(\mathbf{I}_n - n^{-1}\mathbf{J}_{n\times n}) \in \mathbb{R}^{d\times n}$ musí mať plnú riadkovú hodnosť. Takže $\mathbf{S}$ je regulárna práve vtedy, keď $\operatorname{rank}(\mathbf{X}(\mathbf{I}_n - n^{-1}\mathbf{J}_{n\times n})) = d$.

Príklad 3.39. Ukážte, že $\operatorname{rank}(\mathbf{X}(\mathbf{I}_n - n^{-1}\mathbf{J}_{n\times n})) = d$ nastáva práve vtedy, keď vektory $\mathbf{x}_1 - \bar{\mathbf{x}}, \ldots, \mathbf{x}_n - \bar{\mathbf{x}}$ sú lineárne nezávislé. Výberová kovariančná matica je teda regulárna práve vtedy, keď centrované dáta sú lineárne nezávislé.

3.5 Úlohy na precvičenie

Úloha 3.1. Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{B} \in \mathbb{R}^{m\times p}$, $\mathbf{C} \in \mathbb{R}^{r\times m}$. Ukážte, že a) ak $\mathcal{C}(\mathbf{A}) \subseteq \mathcal{C}(\mathbf{B})$ tak $\mathcal{C}(\mathbf{CA}) \subseteq \mathcal{C}(\mathbf{CB})$; b) ak $\mathcal{C}(\mathbf{A}) = \mathcal{C}(\mathbf{B})$ tak $\mathcal{C}(\mathbf{CA}) = \mathcal{C}(\mathbf{CB})$.

Úloha 3.2. Nech $\mathbf{A}$ je štvorcová matica. a) Dokážte, že $\mathcal{C}(\mathbf{A}^k) \subseteq \mathcal{C}(\mathbf{A})$ pre každé prirodzené číslo k. b) Dokážte, že ak $\mathcal{C}(\mathbf{A}^2) = \mathcal{C}(\mathbf{A})$, potom $\mathcal{C}(\mathbf{A}^k) = \mathcal{C}(\mathbf{A})$ pre každé prirodzené číslo k.

Úloha 3.3. Nech

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 & -3 & 2 \\ 0 & -2 & 0 & 6 & 2 \\ 0 & 2 & 2 & 5 & 2 \\ 0 & -4 & -2 & 1 & 0 \end{pmatrix}.$$

Nájdite (akúkoľvek) bázu priestoru $\mathcal{C}(\mathbf{A})$. Akú má matica $\mathbf{A}$ hodnosť?

Úloha 3.4. Zodpovedajte otázky o typoch stĺpcových priestorov z príkladu 3.3 pomocou hodnosti príslušných matíc. Vedeli by ste tvrdenie zovšeobecniť pre matice z $\mathbb{R}^{m\times n}$? Ako vo všeobecnosti vyzerá stĺpcový priestor matice s hodnosťou r?

Úloha 3.5. Presvedčte sa, že $\mathrm{rank}(\mathbf{A}_1 \cdots \mathbf{A}_k) \leq \min\{\mathrm{rank}(\mathbf{A}_1), \ldots, \mathrm{rank}(\mathbf{A}_k)\}$. (Predpokladáme, že $\mathbf{A}_1, \ldots, \mathbf{A}_k$ sú matice vhodných rozmerov.)

Úloha 3.6. Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$ má hodnosť r. Ukážte, že $\mathbf{A}$ sa dá napísať ako súčet r matíc, ktoré majú všetky hodnosť 1. Návod: Využite vetu 3.11.

Úloha 3.7. Pre všetky tvrdenia z tejto kapitoly ohľadom stĺpcového priestoru $\mathcal{C}(\mathbf{A})$ matice $\mathbf{A}$ sformulujte analogické tvrdenie ohľadom riadkového priestoru $\mathcal{C}(\mathbf{A}^T)$.

Úloha 3.8. Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$. Ukážte, že ak je $\mathbf{B}$ regulárna matica typu $n \times n$ (t.j. existuje matica $\mathbf{B}^{-1}$ taká, že $\mathbf{B}\mathbf{B}^{-1} = \mathbf{B}^{-1}\mathbf{B} = \mathbf{I}_n$) a $\mathbf{C}$ je akákoľvek matica typu $p \times m$, tak a) $\mathcal{C}(\mathbf{A}\mathbf{B}) = \mathcal{C}(\mathbf{A})$; b) $\mathcal{C}(\mathbf{C}\mathbf{A}\mathbf{B}\mathbf{B}^T\mathbf{A}^T) = \mathcal{C}(\mathbf{C}\mathbf{A})$.

Úloha 3.9. Ukážte, že $\mathrm{rank}(\mathbf{A}\mathbf{C} + \mathbf{B}\mathbf{D}) \leq \mathrm{rank}(\mathbf{A}, \mathbf{B})$ a $\mathrm{rank}(\mathbf{A}\mathbf{C} + \mathbf{B}\mathbf{D}) \leq \mathrm{rank}(\mathbf{C}^T, \mathbf{D}^T)$ pre ľubovoľné $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{B} \in \mathbb{R}^{m\times k}$, $\mathbf{C} \in \mathbb{R}^{n\times p}$, $\mathbf{D} \in \mathbb{R}^{k\times p}$.

Úloha 3.10. Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{B} \in \mathbb{R}^{m\times n}$. Ukážte, že

$$\mathrm{rank}\begin{pmatrix} \mathbf{A}\mathbf{A}^T & \mathbf{A}\mathbf{B}^T \\ \mathbf{B}\mathbf{A}^T & \mathbf{B}\mathbf{B}^T \end{pmatrix} = \mathrm{rank}(\mathbf{A}^T\mathbf{A} + \mathbf{B}^T\mathbf{B}).$$

Návod: vyjadrite ľavú aj pravú stranu ako súčin blokových matíc.

Úloha 3.11. Nech $\mathbf{X} \in \mathbb{R}^{p\times m}$ a $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m\times n}$. Ukážte, že:
a) $\mathcal{N}(\mathbf{A}) \subseteq \mathcal{N}(\mathbf{X}\mathbf{A})$; b) $\mathcal{N}((\mathbf{A}^T, \mathbf{B}^T)^T) \subseteq \mathcal{N}(\mathbf{A})$; c) $\mathcal{N}((\mathbf{A}^T, \mathbf{B}^T)^T) \subseteq \mathcal{N}(\mathbf{A} + \mathbf{B})$.

Úloha 3.12. Nájdite aspoň jeden **BT**-rozklad matice

$$\mathbf{A} = \begin{pmatrix} 2 & 4 & 1 & 0 \\ 1 & 2 & 7 & -13 \\ 3 & 6 & 3 & -3 \end{pmatrix}.$$

Úloha 3.13. Pre každú 2×2 regulárnu maticu $\mathbf{A} = \{a_{ij}\}$ definujme funkciu

$$T_{\mathbf{A}}(z) = \frac{a_{11}z + a_{12}}{a_{21}z + a_{22}}$$

pre každé $z \in \mathbb{R}$ také, že $a_{21}z + a_{22} \neq 0$. Ukážte, že potom pre ľubovoľné regulárne $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{2\times 2}$ platí $T_{\mathbf{B}}(T_{\mathbf{A}}(z)) = T_{\mathbf{B}\mathbf{A}}(z)$ pre každé $z \in \mathbb{R}$ také, že $a_{21}z + a_{22} \neq 0$ a $b_{21}T_{\mathbf{A}}(z) + b_{22} \neq 0$. Zamyslite sa, prečo $T_{\mathbf{A}}$ definujeme pre regulárne matice $\mathbf{A}$. Poznamenajme, že funkcia $T_{\mathbf{A}}(z)$ sa pre $z \in \mathbb{C}$ nazýva Möbiusova transformácia.

Kapitola 4

Inverzné matice

4.1 Systémy lineárnych rovníc

Definícia 4.1 (Systém lineárnych rovníc a jeho riešenie)**.** *Systém lineárnych rovníc* je sústava rovníc v maticovom zápise $\mathbf{AX} = \mathbf{B}$, kde $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{B} \in \mathbb{R}^{m \times k}$ sú matice známych koeficientov a $\mathbf{X} \in \mathbb{R}^{n \times k}$ je matica neznámych premenných. *Riešenie* systému rovníc $\mathbf{AX} = \mathbf{B}$ je ľubovoľná matica $\mathbf{X}^*$ typu $n \times k$, ktorá spĺňa $\mathbf{AX}^* = \mathbf{B}$.

Poznámka 4.2. Uvažujme systém rovníc

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ &\vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m, \end{aligned}$$

kde koeficienty a_{ij}, b_i ($i = 1, \ldots, m$, $j = 1, \ldots, n$) sú pevne dané reálne čísla a $x_1, \ldots, x_n \in \mathbb{R}$ sú neznáme. Takýto systém môžeme zapísať maticovo ako $\mathbf{Ax} = \mathbf{b}$, kde $\mathbf{A} = \{a_{ij}\}$, $\mathbf{x} \in \mathbb{R}^n$ a $\mathbf{b} \in \mathbb{R}^m$.

Systém rovníc $\mathbf{AX} = \mathbf{B}$ s maticami koeficientov $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{B} \in \mathbb{R}^{m \times k}$ a maticou neznámych $\mathbf{X} \in \mathbb{R}^{n \times k}$ môžeme vnímať ako k systémov rovníc s vektormi neznámych hodnôt:

$$\mathbf{Ax}_1 = \mathbf{b}_1,\ \ldots,\ \mathbf{Ax}_k = \mathbf{b}_k \quad \Leftrightarrow \quad \mathbf{AX} = \mathbf{B},$$

kde $\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_k)$ a $\mathbf{B} = (\mathbf{b}_1, \ldots, \mathbf{b}_k)$.

Definícia 4.3 (Homogénny a nehomogénny systém lineárnych rovníc)**.** Systém lineárnych rovníc $\mathbf{AX} = \mathbf{B}$ sa nazýva *homogénny*, ak $\mathbf{B} = \mathbf{0}$. Inak sa nazýva *nehomogénny*.

Definícia 4.4 (Konzistentný systém lineárnych rovníc)**.** Nech $\mathbf{A} \in \mathbb{R}^{m \times n}$ a $\mathbf{B} \in \mathbb{R}^{m \times k}$. Systém lineárnych rovníc $\mathbf{AX} = \mathbf{B}$ sa nazýva *konzistentný*, ak má aspoň jedno riešenie $\mathbf{X} \in \mathbb{R}^{n \times k}$.

Príklad 4.5. Každý homogénny systém lineárnych rovníc má riešenie $\mathbf{X} = \mathbf{0}$. Naproti tomu nehomogénne systémy rovníc nemusia mať riešenie. Overte, že systém $\mathbf{AX} = \mathbf{B}$, kde

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ -2 & -4 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 3 & 1 \\ 3 & 5 \end{pmatrix},$$

nemá riešenie.

Veta 4.6 (Charakterizácia konzistencie systému)**.** Každá z nasledujúcich podmienok je nutná a postačujúca na to, aby bol systém $\mathbf{AX} = \mathbf{B}$ konzistentný.

(i) $\mathcal{C}(\mathbf{B}) \subseteq \mathcal{C}(\mathbf{A})$,

(ii) $\mathcal{N}(\mathbf{A}^T) \subseteq \mathcal{N}(\mathbf{B}^T)$,

(iii) $\mathcal{C}(\mathbf{A}, \mathbf{B}) = \mathcal{C}(\mathbf{A})$,

(iv) $\text{rank}(\mathbf{A}, \mathbf{B}) = \text{rank}(\mathbf{A})$.

Dôkaz. Skutočnosť, že konzistentnosť systému $\mathbf{AX} = \mathbf{B}$ je ekvivalentná podmienke (i), plynie z vety 3.4 o inklúzii stĺpcových priestorov. Ekvivalencia (i) s (ii) vyplýva zo vzťahu $\mathcal{N}(\mathbf{A}^T) = \mathcal{C}(\mathbf{A})^\perp$.

Dokážeme ekvivalenciu medzi (iii) a (i). Predpokladajme $\mathcal{C}(\mathbf{A}, \mathbf{B}) = \mathcal{C}(\mathbf{A})$. Z vety 3.20 vyplýva, že vždy platí $\mathcal{C}(\mathbf{A}) \subseteq \mathcal{C}(\mathbf{A}, \mathbf{B})$. Aby nastala rovnosť, zjavne musí byť každý stĺpec $\mathbf{B}$ lineárnou kombináciou stĺpcov $\mathbf{A}$, a teda $\mathcal{C}(\mathbf{B}) \subseteq \mathcal{C}(\mathbf{A})$. Naopak, nech $\mathcal{C}(\mathbf{B}) \subseteq \mathcal{C}(\mathbf{A})$. Z toho vyplýva $\mathcal{C}(\mathbf{A}, \mathbf{B}) = \mathcal{C}(\mathbf{A})$, lebo každý stĺpec $\mathbf{B}$ je lineárnou kombináciou stĺpcov $\mathbf{A}$.

Ekvivalenciu medzi (iii) a (iv) dokážme tiež overením oboch implikácií. Zjavne platí, že ak $\mathcal{C}(\mathbf{A}, \mathbf{B}) = \mathcal{C}(\mathbf{A})$, tak aj $\text{rank}(\mathbf{A}, \mathbf{B}) = \text{rank}(\mathbf{A})$. Naopak, keďže $\mathcal{C}(\mathbf{A}) \subseteq \mathcal{C}(\mathbf{A}, \mathbf{B})$, potom podľa vety 1.27(iii) platí, že ak $\text{rank}(\mathbf{A}) = \text{rank}(\mathbf{A}, \mathbf{B})$, tak $\mathcal{C}(\mathbf{A}) = \mathcal{C}(\mathbf{A}, \mathbf{B})$. $\square$

4.2 Inverzné matice

Definícia 4.7 (Ľavá a pravá inverzná matica)**.** *Ľavou inverznou maticou* (*ľavou inverziou*) matice $\mathbf{A} \in \mathbb{R}^{m\times n}$ nazývame akúkoľvek maticu $\mathbf{L} \in \mathbb{R}^{n\times m}$ spĺňajúcu $\mathbf{LA} = \mathbf{I}_n$. *Pravá inverzná matica* (*pravá inverzia*) matice $\mathbf{A}$ je ľubovoľná matica $\mathbf{R} \in \mathbb{R}^{n\times m}$ spĺňajúca $\mathbf{AR} = \mathbf{I}_m$.

Lema 4.8 (O existencii ľavej a pravej inverzie)**.** Matica $\mathbf{A} \in \mathbb{R}^{m\times n}$ má ľavú inverziu práve vtedy, keď $\text{rank}(\mathbf{A}) = n$, a má pravú inverziu práve vtedy, keď $\text{rank}(\mathbf{A}) = m$.

Dôkaz. Nech $\mathbf{A}$ má pravú inverziu $\mathbf{R}$. Potom z vety 3.14 plynie $\text{rank}(\mathbf{A}) \geq \text{rank}(\mathbf{AR})$, a zároveň $\mathbf{AR} = \mathbf{I}_m$, čiže $\text{rank}(\mathbf{AR}) = m$. Dostávame $\text{rank}(\mathbf{A}) \geq m$. Keďže $\text{rank}(\mathbf{A}) \leq m$, lebo $\mathbf{A}$ má m riadkov, tak $\text{rank}(\mathbf{A}) = m$. Naopak uvažujme, že $\text{rank}(\mathbf{A}) = m$, čo nastáva práve vtedy, keď $\mathcal{C}(\mathbf{A}) = \mathbb{R}^m = \mathcal{C}(\mathbf{I}_m)$. Podľa vety 3.4 existuje $\mathbf{R}$, že $\mathbf{AR} = \mathbf{I}_m$. Dôkaz pre ľavú inverziu je analogický. $\square$

Dôsledok 4.9. Matica $\mathbf{A} \in \mathbb{R}^{m\times n}$ má ľavú aj pravú inverziu práve vtedy, keď $m = n$ a $\mathbf{A}$ je regulárna.

Veta 4.10 (O škrtaní sprava a zľava v maticovej rovnosti)**.** Nech $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m\times n}$, $\mathbf{C} \in \mathbb{R}^{p\times m}$, $\mathbf{D} \in \mathbb{R}^{n\times q}$ spĺňajú $\text{rank}(\mathbf{C}) = m$, $\text{rank}(\mathbf{D}) = n$ a $\mathbf{CAD} = \mathbf{CBD}$. Potom $\mathbf{A} = \mathbf{B}$.

Dôkaz. Keďže $\text{rank}(\mathbf{C}) = m$ a $\text{rank}(\mathbf{D}) = n$, tak podľa lemy 4.8 má $\mathbf{C}$ ľavú inverziu $\mathbf{L}$ a $\mathbf{D}$ má pravú inverziu $\mathbf{R}$. Preto

$$\mathbf{A} = \mathbf{I}_m\mathbf{A}\mathbf{I}_n = \mathbf{LCADR} = \mathbf{LCBDR} = \mathbf{I}_m\mathbf{B}\mathbf{I}_n = \mathbf{B}.$$

$\square$

Definícia 4.11 (Invertovateľnosť matice, inverzná matica)**.** Nech $\mathbf{A} \in \mathbb{R}^{n\times n}$. Ak existuje matica $\mathbf{B}$, ktorá je ľavou aj pravou inverziou matice $\mathbf{A}$, tak hovoríme, že $\mathbf{A}$ je *invertovateľná* a $\mathbf{B}$ nazývame *inverznou maticou* (*inverziou*) matice $\mathbf{A}$.

Lema 4.12 (Ľavá a pravá inverzia štvorcovej matice)**.** Ak $\mathbf{A} \in \mathbb{R}^{n\times n}$ má pravú inverziu $\mathbf{R}$, tak $\mathbf{R}$ je zároveň ľavou inverziou $\mathbf{A}$. Naopak, ak $\mathbf{L}$ je ľavá inverzia $\mathbf{A}$, tak $\mathbf{L}$ je zároveň pravou inverziou matice $\mathbf{A}$.

Dôkaz. Nech $\mathbf{A}$ má pravú inverziu $\mathbf{R}$, potom $\operatorname{rank}(\mathbf{A}) = n$, a teda existuje jej ľavá inverzia $\mathbf{L}$. Potom $\mathbf{L} = \mathbf{L}\mathbf{I}_n = \mathbf{L}\mathbf{A}\mathbf{R} = \mathbf{I}_m\mathbf{R} = \mathbf{R}$. □

Veta 4.13 (Charakterizácia invertovateľnosti matice)**.** Matica $\mathbf{A}$ je invertovateľná práve vtedy, keď $\mathbf{A}$ je štvorcová regulárna matica. Navyše akákoľvek regulárna matica má práve jednu inverziu.

Dôkaz. Prvá časť tvrdenia vyplýva z dôsledku 4.9. Nech $\mathbf{B}$ a $\mathbf{C}$ sú dve inverzie matice $\mathbf{A} \in \mathbb{R}^{n\times n}$. Potom $\mathbf{C} = \mathbf{C}\mathbf{I}_n = \mathbf{C}\mathbf{A}\mathbf{B} = \mathbf{I}_n\mathbf{B} = \mathbf{B}$. □

Poznámka 4.14. Inverznú maticu regulárnej matice $\mathbf{A}$ značíme $\mathbf{A}^{-1}$.

Lema 4.15 (Základné vlastnosti inverzie)**.** Nech $\mathbf{A}$ je invertovateľná. Potom

(i) $\mathbf{A}^{-1}$ je invertovateľná a $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$,

(ii) $\mathbf{A}^T$ je invertovateľná a $(\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T$.

Dôkaz. Dôkaz prvého tvrdenia je triviálny. Matica $\mathbf{A}^T$ je invertovateľná, keďže $\operatorname{rank}(\mathbf{A}^T) = \operatorname{rank}(\mathbf{A})$. Navyše $\mathbf{A}^T(\mathbf{A}^{-1})^T = (\mathbf{A}^{-1}\mathbf{A})^T = \mathbf{I}$. □

Dôsledok 4.16. Inverzná matica regulárnej symetrickej matice je tiež symetrická.

Poznámka 4.17. Keďže podľa lemy 4.15 pre regulárnu maticu $\mathbf{A}$ platí $(\mathbf{A}^{-1})^T = (\mathbf{A}^T)^{-1}$, niekedy tieto matice značíme jednotne ako $\mathbf{A}^{-T}$, čiže $\mathbf{A}^{-T} = (\mathbf{A}^{-1})^T = (\mathbf{A}^T)^{-1}$.

Veta 4.18 (Násobenie regulárnou maticou nemení hodnosť)**.** Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{B} \in \mathbb{R}^{n\times n}$, $\mathbf{C} \in \mathbb{R}^{m\times m}$ a nech $\mathbf{B}$ a $\mathbf{C}$ sú regulárne. Potom

(i) $\mathcal{C}(\mathbf{A}\mathbf{B}) = \mathcal{C}(\mathbf{A})$,

(ii) $\mathcal{C}((\mathbf{C}\mathbf{A})^T) = \mathcal{C}(\mathbf{A}^T)$,

(iii) $\operatorname{rank}(\mathbf{A}\mathbf{B}) = \operatorname{rank}(\mathbf{A}) = \operatorname{rank}(\mathbf{C}\mathbf{A})$.

Dôkaz. Podľa vety 3.4 platí $\mathcal{C}(\mathbf{A}\mathbf{B}) \subseteq \mathcal{C}(\mathbf{A})$. Keďže $\mathbf{B}$ je regulárna, tak existuje $\mathbf{B}^{-1}$ a potom $\mathcal{C}(\mathbf{A}) = \mathcal{C}(\mathbf{A}\mathbf{B}\mathbf{B}^{-1}) \subseteq \mathcal{C}(\mathbf{A}\mathbf{B})$, opäť podľa vety 3.4. Ukázali sme, že $\mathcal{C}(\mathbf{A}\mathbf{B}) = \mathcal{C}(\mathbf{A})$. Časť (ii) ľahko dokážeme pomocou (i): $\mathcal{C}((\mathbf{C}\mathbf{A})^T) = \mathcal{C}(\mathbf{A}^T\mathbf{C}^T) = \mathcal{C}(\mathbf{A}^T)$, keďže $\mathbf{C}^T$ je regulárna. Tvrdenie (iii) je dôsledkom predchádzajúcich dvoch tvrdení, pričom využijeme vetu 3.9. □

Dôsledok 4.19. Nech $\mathbf{A} \in \mathbb{R}^{n\times n}$. Ak $\mathbf{A} = \mathbf{X}\mathbf{D}\mathbf{Y}$, pričom $\mathbf{D} \in \mathbb{R}^{n\times n}$ je diagonálna a $\mathbf{X}, \mathbf{Y} \in \mathbb{R}^{n\times n}$ sú regulárne, tak hodnosť $\mathbf{A}$ je rovná počtu nenulových diagonálnych prvkov matice $\mathbf{D}$.

Dôkaz. Podľa vety 4.18 platí $\operatorname{rank}(\mathbf{A}) = \operatorname{rank}(\mathbf{X}\mathbf{D}\mathbf{Y}) = \operatorname{rank}(\mathbf{X}\mathbf{D}) = \operatorname{rank}(\mathbf{D})$, pričom hodnosť diagonálnej matice je zjavne počet nenulových prvkov na jej diagonále. □

Príklad 4.20. Nech

$$\mathbf{A} = \begin{pmatrix} 2 & 4 \\ 5 & 15 \\ 1 & 2 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

Určte $\mathcal{C}(\mathbf{A}), \mathcal{C}(\mathbf{A}\mathbf{B})$ a overte, že $\mathcal{C}(\mathbf{A}) = \mathcal{C}(\mathbf{A}\mathbf{B})$.

4.3 Hodnosť a inverzia blokovej matice

Lema 4.21 (Regularita základnej blokovo trojuholníkovej matice)**.** Nech $\mathbf{X} \in \mathbb{R}^{m\times n}$. Potom matice

$$\mathbf{A} = \begin{pmatrix} \mathbf{I}_m & \mathbf{X} \\ \mathbf{0}_{n\times m} & \mathbf{I}_n \end{pmatrix} \quad \text{a} \quad \mathbf{B} = \begin{pmatrix} \mathbf{I}_n & \mathbf{0}_{n\times m} \\ \mathbf{X} & \mathbf{I}_m \end{pmatrix}$$

sú regulárne.

Dôkaz. Matica $\mathbf{A}$ nie je iba blokovo horná trojuholníková, ale aj horná trojuholníková (v klasickom zmysle), pričom na diagonále má samé jednotky. Táto matica má teda lineárne nezávislé stĺpce, čiže je regulárna. Dôkaz pre maticu $\mathbf{B}$ je analogický. □

Lema 4.22 (Hodnosť blokovo hornej trojuholníkovej matice)**.** Nech $\mathbf{T} \in \mathbb{R}^{n\times n}$, $\mathbf{U} \in \mathbb{R}^{n\times q}$, $\mathbf{Q} \in \mathbb{R}^{m\times q}$ a nech $\mathbf{T}$ je regulárna. Potom

$$\operatorname{rank}\begin{pmatrix} \mathbf{T} & \mathbf{U} \\ \mathbf{0}_{m\times n} & \mathbf{Q} \end{pmatrix} = \operatorname{rank}(\mathbf{T}) + \operatorname{rank}(\mathbf{Q}).$$

Dôkaz. Keďže $\mathbf{T}$ je regulárna, tak existuje $\mathbf{T}^{-1}$. Ľahko sa presvedčíme, že

$$\begin{pmatrix} \mathbf{T} & \mathbf{U} \\ \mathbf{0} & \mathbf{Q} \end{pmatrix}\begin{pmatrix} \mathbf{I}_n & -\mathbf{T}^{-1}\mathbf{U} \\ \mathbf{0} & \mathbf{I}_q \end{pmatrix} = \begin{pmatrix} \mathbf{T} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q} \end{pmatrix},$$

pričom matica

$$\begin{pmatrix} \mathbf{I}_n & -\mathbf{T}^{-1}\mathbf{U} \\ \mathbf{0} & \mathbf{I}_q \end{pmatrix}$$

je podľa lemy 4.21 regulárna. Podľa viet 4.18 a 3.20 platí

$$\operatorname{rank}\begin{pmatrix} \mathbf{T} & \mathbf{U} \\ \mathbf{0} & \mathbf{Q} \end{pmatrix} = \operatorname{rank}\begin{pmatrix} \mathbf{T} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q} \end{pmatrix} = \operatorname{rank}(\mathbf{T}) + \operatorname{rank}(\mathbf{Q}).$$

□

Veta 4.23 (Hodnosť blokovej matice)**.** Nech $\mathbf{T} \in \mathbb{R}^{n\times n}$ je regulárna, nech $\mathbf{U} \in \mathbb{R}^{n\times q}$, $\mathbf{V} \in \mathbb{R}^{m\times n}$ a $\mathbf{W} \in \mathbb{R}^{m\times q}$. Potom

$$\operatorname{rank}\begin{pmatrix} \mathbf{T} & \mathbf{U} \\ \mathbf{V} & \mathbf{W} \end{pmatrix} = n + \operatorname{rank}(\mathbf{W} - \mathbf{V}\mathbf{T}^{-1}\mathbf{U}).$$

Dôkaz. Všimnime si, že

$$\begin{pmatrix} \mathbf{I}_n & \mathbf{0}_{n\times m} \\ -\mathbf{V}\mathbf{T}^{-1} & \mathbf{I}_m \end{pmatrix}\begin{pmatrix} \mathbf{T} & \mathbf{U} \\ \mathbf{V} & \mathbf{W} \end{pmatrix} = \begin{pmatrix} \mathbf{T} & \mathbf{U} \\ \mathbf{0} & \mathbf{W} - \mathbf{V}\mathbf{T}^{-1}\mathbf{U} \end{pmatrix},$$

pričom podľa lemy 4.21 je matica

$$\operatorname{rank}\begin{pmatrix} \mathbf{I}_n & \mathbf{0}_{n\times m} \\ -\mathbf{V}\mathbf{T}^{-1} & \mathbf{I}_m \end{pmatrix}$$

regulárna. Potom môžeme použiť vetu 4.18 a ešte raz lemu 4.22, čím dostávame

$$\operatorname{rank}\begin{pmatrix} \mathbf{T} & \mathbf{U} \\ \mathbf{V} & \mathbf{W} \end{pmatrix} = \operatorname{rank}\begin{pmatrix} \mathbf{T} & \mathbf{U} \\ \mathbf{0} & \mathbf{W} - \mathbf{V}\mathbf{T}^{-1}\mathbf{U} \end{pmatrix} = \operatorname{rank}(\mathbf{T}) + \operatorname{rank}(\mathbf{W} - \mathbf{V}\mathbf{T}^{-1}\mathbf{U}).$$

□

Definícia 4.24 (Schurov doplnok)**.** Za platnosti predpokladov vety 4.23 nazývame maticu $\mathbf{W} - \mathbf{V}\mathbf{T}^{-1}\mathbf{U}$ *Schurov doplnok* matice $\mathbf{T}$ v matici

$$\begin{pmatrix} \mathbf{T} & \mathbf{U} \\ \mathbf{V} & \mathbf{W} \end{pmatrix}.$$

Lema 4.25 (O nulovom Schurovom doplnku)**.** Nech $\mathbf{A} \in \mathbb{R}^{m \times n}$ má hodnosť r. Nech navyše

$$\mathbf{A} = \begin{pmatrix} \mathbf{T} & \mathbf{U} \\ \mathbf{V} & \mathbf{W} \end{pmatrix},$$

kde $\mathbf{T} \in \mathbb{R}^{r \times r}$. Potom ak $\mathrm{rank}(\mathbf{T}) = r$, tak $\mathbf{W} = \mathbf{V}\mathbf{T}^{-1}\mathbf{U}$.

Dôkaz. Podľa vety 4.23 platí $r = r + \mathrm{rank}(\mathbf{W} - \mathbf{V}\mathbf{T}^{-1}\mathbf{U})$, z čoho vyplýva $\mathrm{rank}(\mathbf{W} - \mathbf{V}\mathbf{T}^{-1}\mathbf{U}) = 0$. Využijúc lemu 3.13 dostávame $\mathbf{W} - \mathbf{V}\mathbf{T}^{-1}\mathbf{U} = \mathbf{0}$, čo značí $\mathbf{W} = \mathbf{V}\mathbf{T}^{-1}\mathbf{U}$. □

Veta 4.26 (Inverzia blokovej matice)**.** Nech $\mathbf{T} \in \mathbb{R}^{n \times n}$ je regulárna, nech $\mathbf{U} \in \mathbb{R}^{n \times m}$, $\mathbf{V} \in \mathbb{R}^{m \times n}$ a $\mathbf{W} \in \mathbb{R}^{m \times m}$. Potom matica

$$\begin{pmatrix} \mathbf{T} & \mathbf{U} \\ \mathbf{V} & \mathbf{W} \end{pmatrix}$$

je regulárna práve vtedy, keď $\mathbf{Q} = \mathbf{W} - \mathbf{V}\mathbf{T}^{-1}\mathbf{U}$ je regulárna. V takom prípade navyše platí

$$\begin{pmatrix} \mathbf{T} & \mathbf{U} \\ \mathbf{V} & \mathbf{W} \end{pmatrix}^{-1} = \begin{pmatrix} \mathbf{T}^{-1} + \mathbf{T}^{-1}\mathbf{U}\mathbf{Q}^{-1}\mathbf{V}\mathbf{T}^{-1} & -\mathbf{T}^{-1}\mathbf{U}\mathbf{Q}^{-1} \\ -\mathbf{Q}^{-1}\mathbf{V}\mathbf{T}^{-1} & \mathbf{Q}^{-1} \end{pmatrix}.$$

Dôkaz. Matica

$$\begin{pmatrix} \mathbf{T} & \mathbf{U} \\ \mathbf{V} & \mathbf{W} \end{pmatrix}$$

je regulárna práve vtedy, keď jej hodnosť je $m + n$. To nastáva podľa vety 4.23 práve vtedy, keď $n + \mathrm{rank}(\mathbf{Q}) = n + m$, čo je práve vtedy, keď $\mathbf{Q}$ je regulárna. Tvar inverznej matice overme priamym výpočtom:

$$\begin{aligned}
\begin{pmatrix} \mathbf{T} & \mathbf{U} \\ \mathbf{V} & \mathbf{W} \end{pmatrix} \begin{pmatrix} \mathbf{T}^{-1} + \mathbf{T}^{-1}\mathbf{U}\mathbf{Q}^{-1}\mathbf{V}\mathbf{T}^{-1} & -\mathbf{T}^{-1}\mathbf{U}\mathbf{Q}^{-1} \\ -\mathbf{Q}^{-1}\mathbf{V}\mathbf{T}^{-1} & \mathbf{Q}^{-1} \end{pmatrix} &= \\
= \begin{pmatrix} \mathbf{I} + \mathbf{U}\mathbf{Q}^{-1}\mathbf{V}\mathbf{T}^{-1} - \mathbf{U}\mathbf{Q}^{-1}\mathbf{V}\mathbf{T}^{-1} & -\mathbf{U}\mathbf{Q}^{-1} + \mathbf{U}\mathbf{Q}^{-1} \\ \mathbf{V}\mathbf{T}^{-1} + \mathbf{V}\mathbf{T}^{-1}\mathbf{U}\mathbf{Q}^{-1}\mathbf{V}\mathbf{T}^{-1} - \mathbf{W}\mathbf{Q}^{-1}\mathbf{V}\mathbf{T}^{-1} & -\mathbf{V}\mathbf{T}^{-1}\mathbf{U}\mathbf{Q}^{-1} + \mathbf{W}\mathbf{Q}^{-1} \end{pmatrix}& \\
= \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{pmatrix},&
\end{aligned}$$

keďže

$$\begin{aligned}
\mathbf{V}\mathbf{T}^{-1} + \mathbf{V}\mathbf{T}^{-1}\mathbf{U}\mathbf{Q}^{-1}\mathbf{V}\mathbf{T}^{-1} - \mathbf{W}\mathbf{Q}^{-1}\mathbf{V}\mathbf{T}^{-1} &= \mathbf{V}\mathbf{T}^{-1} + (\mathbf{V}\mathbf{T}^{-1}\mathbf{U} - \mathbf{W})\mathbf{Q}^{-1}\mathbf{V}\mathbf{T}^{-1} \\
&= \mathbf{V}\mathbf{T}^{-1} - \mathbf{Q}\mathbf{Q}^{-1}\mathbf{V}\mathbf{T}^{-1} = \mathbf{0}
\end{aligned}$$

a

$$-\mathbf{V}\mathbf{T}^{-1}\mathbf{U}\mathbf{Q}^{-1} + \mathbf{W}\mathbf{Q}^{-1} = (-\mathbf{V}\mathbf{T}^{-1}\mathbf{U} + \mathbf{W})\mathbf{Q}^{-1} = \mathbf{Q}\mathbf{Q}^{-1} = \mathbf{I}.$$

□

4.4 Ortogonálne matice

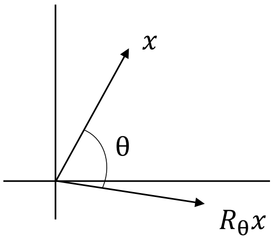
Definícia 4.27 (Ortogonálna matica)**.** Štvorcová matica $\mathbf{Q} \in \mathbb{R}^{n\times n}$ je *ortogonálna*, ak jej stĺpce sú ortonormálne, teda ak $\mathbf{Q}^T\mathbf{Q} = \mathbf{I}_n$.

Poznámka 4.28. Ekvivalentne môžeme definovať ako ortogonálnu maticu takú maticu $\mathbf{Q} \in \mathbb{R}^{n\times n}$, ktorá je regulárna a spĺňa $\mathbf{Q}^{-1} = \mathbf{Q}^T$. Z toho vyplýva, že aj $\mathbf{Q}\mathbf{Q}^{-1} = \mathbf{I}_n$, a teda $\mathbf{Q}\mathbf{Q}^T = \mathbf{I}_n$, čiže aj riadky ortogonálnej matice tvoria ortonormálny systém vektorov.

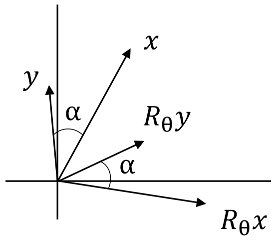
Príklad 4.29. Príkladmi ortogonálnych matíc sú napríklad $\mathbf{I}_n$,

$$\mathbf{Q} = \begin{pmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{pmatrix}, \quad \mathbf{R}_\theta = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$

pre ľubovoľný uhol $\theta \in [0, 2\pi)$. Matica $\mathbf{R}_\theta$ je tzv. maticou rotácie o uhol θ (v smere hodinových ručičiek), ako vidíme na obrázku 4.1a. Vynásobením vektoru $\mathbf{x} \in \mathbb{R}^2$ touto maticou ho otočíme v smere hodinových ručičiek o uhol θ. Overte túto vlastnosť pre $\theta = \pi/2$. Čo predstavuje matica $\mathbf{R}_\theta^2$? Nájdite tiež maticu rotácie o uhol θ proti smeru hodinových ručičiek a matice preklopenia okolo jednotlivých súradnicových osí.



(a) Otočenie o uhol θ.



(b) Zachovanie uhlov.

Obr. 4.1: Vynásobenie rotačnou maticou $\mathbf{R}_\theta$.

Lema 4.30 (Vlastnosti ortogonálnych matíc)**.** Medzi základné vlastnosti ortogonálnych matíc patrí:

(i) Ak $\mathbf{Q}$ je ortogonálna, tak aj $\mathbf{Q}^T$ je ortogonálna.

(ii) Ak $\mathbf{Q}_1, \dots, \mathbf{Q}_k$ sú ortogonálne, tak aj $\mathbf{Q}_1 \cdots \mathbf{Q}_k$ je ortogonálna.

Dôkaz. Dôkaz nechávame na čitateľa. □

Lema 4.31 (Ortogonálne matice zachovávajú dĺžku a uhol)**.** Nech $\mathbf{Q} \in \mathbb{R}^{n\times n}$ je ortogonálna a nech $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Potom $\|\mathbf{Qx}\| = \|\mathbf{x}\|$ a $\langle \mathbf{Qx}, \mathbf{Qy}\rangle = \langle \mathbf{x}, \mathbf{y}\rangle$.

Dôkaz. Tvrdenia ľahko overíme priamymi výpočtami: $\|\mathbf{Qx}\|^2 = \mathbf{x}^T\mathbf{Q}^T\mathbf{Q}\mathbf{x} = \mathbf{x}^T\mathbf{x} = \|\mathbf{x}\|^2$, a teda aj $\|\mathbf{Qx}\| = \|\mathbf{x}\|$. Zároveň $\langle \mathbf{Qx}, \mathbf{Qy}\rangle = \mathbf{x}^T\mathbf{Q}^T\mathbf{Q}\mathbf{y} = \mathbf{x}^T\mathbf{y} = \langle \mathbf{x}, \mathbf{y}\rangle$. □

Poznámka 4.32. Na obrázku 4.1b môžeme vidieť, že vynásobenie rotačnou maticou $\mathbb{R}_\theta$ zachováva uhly medzi vektormi.

Veta 4.33 (**QR** rozklad matice)**.** Nech $\mathbf{A} \in \mathbb{R}^{m \times n}$, kde $m \geq n$, má hodnosť n. Potom existuje ortogonálna matica $\mathbf{Q} \in \mathbb{R}^{m \times m}$ a horná trojuholníková matica $\mathbf{R} \in \mathbb{R}^{m \times n}$, že platí $\mathbf{A} = \mathbf{QR}$.

Dôkaz. Keďže $\mathbf{A}$ má plnú stĺpcovú hodnosť, jej stĺpce sú lineárne nezávislé, a teda pre ne môžeme použiť Gramovu-Schmidtovu ortogonalizáciu, pozri vetu 1.42. Čiže existujú ortogonálne a nenulové $\mathbf{b}_1, \ldots, \mathbf{b}_n \in \mathbb{R}^m$, že $\mathbf{b}_j = \mathbf{a}_j - \sum_{i=1}^{j-1} x_{i,j}\mathbf{b}_i$ pre $j = 1, \ldots, n$, čo sa dá zapísať ako $\mathbf{a}_j = \mathbf{b}_j + \sum_{i=1}^{j-1} x_{i,j}\mathbf{b}_i$, alebo v blokovom zápise

$$\mathbf{a}_j = \begin{pmatrix} \mathbf{b}_1 & \ldots & \mathbf{b}_n \end{pmatrix} \begin{pmatrix} x_{1,j} \\ \vdots \\ x_{j-1,j} \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \mathbf{B}\mathbf{x}_j,$$

kde $\mathbf{x}_j \in \mathbb{R}^n$ je príslušný vektor koeficientov. Spolu môžeme pre $j = 1, \ldots, n$ tieto rovnice zapísať ako $\mathbf{A} = \mathbf{BX}$, kde

$$(\mathbf{X})_{ij} = \begin{cases} x_{i,j}, & i < j, \\ 1, & i = j, \\ 0, & i > j. \end{cases}$$

Matica $\mathbf{X}$ je zjavne horná trojuholníková. Stĺpce matice $\mathbf{B}$ sú ortogonálne; ak chceme dosiahnuť ortonormálne stĺpce, zapíšme

$$\mathbf{A} = \mathbf{B}\,\mathrm{diag}(1/\|\mathbf{b}_1\|, \ldots, 1/\|\mathbf{b}_n\|) \cdot \mathrm{diag}(\|\mathbf{b}_1\|, \ldots, \|\mathbf{b}_n\|)\mathbf{X}.$$

Potom $\mathbf{Q} = \mathbf{B}\,\mathrm{diag}(1/\|\mathbf{b}_1\|, \ldots, 1/\|\mathbf{b}_n\|)$ je ortogonálna a $\mathbf{R} = \mathrm{diag}(\|\mathbf{b}_1\|, \ldots, \|\mathbf{b}_n\|)\mathbf{X}$ je horná trojuholníková. □

Dôsledok 4.34 (Úzky **QR** rozklad matice)**.** Nech $\mathbf{A} \in \mathbb{R}^{m \times n}$, kde $m \geq n$, má hodnosť n. Potom existuje matica $\mathbf{Q} \in \mathbb{R}^{m \times n}$ s ortonormálnymi stĺpcami a regulárna horná trojuholníková matica $\mathbf{R} \in \mathbb{R}^{n \times n}$, že platí $\mathbf{A} = \mathbf{QR}$.

Dôkaz. Zapíšme si matice z vety 4.33 ako $\mathbf{Q} = (\mathbf{Q}_1, \mathbf{Q}_2)$ a

$$\mathbf{R} = \begin{pmatrix} \mathbf{R}_1 \\ \mathbf{R}_2 \end{pmatrix},$$

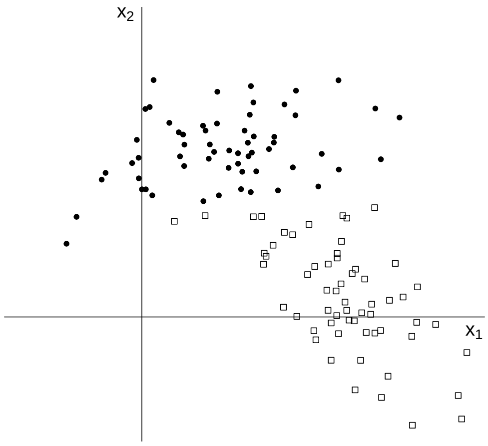
kde $\mathbf{Q}_1 \in \mathbb{R}^{m \times n}$ a $\mathbf{R}_1 \in \mathbb{R}^{n \times n}$. Keďže $\mathbf{Q}$ je ortogonálna, tak $\mathbf{Q}_1$ aj $\mathbf{Q}_2$ majú ortonormálne stĺpce. Keďže $\mathbf{R}$ je horná trojuholníková, tak $\mathbf{R}_1$ je horná trojuholníková a $\mathbf{R}_2 = \mathbf{0}$. Potom

$$\mathbf{A} = \mathbf{QR} = (\mathbf{Q}_1, \mathbf{Q}_2) \begin{pmatrix} \mathbf{R}_1 \\ \mathbf{0} \end{pmatrix} = \mathbf{Q}_1\mathbf{R}_1$$

a $\mathbf{Q}_1, \mathbf{R}_1$ sú hľadané matice. Matica $\mathbf{R}_1$ je regulárna, lebo na jej diagonále sú podľa dôkazu vety 4.33 nenulové čísla. □

Poznámka 4.35. Keďže ortogonálne matice reprezentujú rotácie a reflexie (preklopenia okolo nadrovín), sú kľúčovým nástrojom v mnohých metódach analýzy dát, napr. v metóde hlavných komponentov, faktorovej analýze, kanonických koreláciách a spektrálnom zhlukovaní. Na obrázku 4.2 je znázornená rotácia pozorovaní v $\mathbb{R}^2$ o $\pi/3$ (čiže o $60°$) v smere hodinových ručičiek.

S ortogonálnymi maticami sa veľmi dobre pracuje (viď poznámka 4.28 a lema 4.31), preto sú používané aj pri numerických výpočtoch, ktoré softvér vykonáva pri aplikácii rôznych metód; najmä prostredníctvom **QR** rozkladu. Napríklad v lineárnej regresii sa rieši tzv. problém najmenších štvorcov (s ktorým sa stretneme v kapitole 7), pričom pri numerickom riešení tohto problému sa bežne používajú ortogonálne matice.



Obr. 4.2: Rotácia dát o 60°okolo $\mathbf{0}_2$ v smere hodinových ručičiek. Pôvodné dáta sú znázornené vyplnenými krúžkami, transformované dáta sú štvorce.

4.4.1 Helmertova matica

Definícia 4.36 (Helmertova matica)**.** *Helmertovou maticou* nazývame maticu $\mathbf{H} \in \mathbb{R}^{n\times n}$, ktorá je v tvare

$$\mathbf{H} = \begin{pmatrix} c_1(x_1, \dots, x_n) \\ c_2(x_1, -\frac{x_1^2}{x_2}, 0, \dots, 0) \\ c_3(x_1, x_2, -\frac{x_1^2+x_2^2}{x_3}, 0, \dots, 0) \\ c_4(x_1, x_2, x_3, -\frac{x_1^2+x_2^2+x_3^2}{x_4}, 0, \dots, 0) \\ \vdots \\ c_n(x_1, x_2, \dots, x_{n-1}, -\frac{x_1^2+x_2^2+\dots+x_{n-1}^2}{x_n}) \end{pmatrix}$$

pre nejaké $\mathbf{x} \in \mathbb{R}^n$, ktorého každá zložka je nenulová, pričom $c_1, \dots, c_n$ sú konštanty zvolené tak, aby riadky matice $\mathbf{H}$ mali normu 1:

$$c_1 = \frac{1}{\|\mathbf{x}\|}, \quad c_2 = \frac{1}{\sqrt{x_1^2 + \frac{x_1^4}{x_2^2}}}, \quad c_3 = \frac{1}{\sqrt{x_1^2 + x_2^2 + \frac{(x_1^2+x_2^2)^2}{x_3^2}}}, \quad \dots$$

Lema 4.37 (Ortogonálnosť Helmertovej matice)**.** Každá Helmertova matica je ortogonálna.

Dôkaz. Ortogonálnosť riadkov Helmertovej matice sa dá nahliadnuť z jej konštrukcie. Ich ortonormálnosť je zaručená normalizačnými konštantami c_i. □

Príklad 4.38. Nájdime ortogonálnu maticu, ktorej prvý riadok je násobkom vektora $(1, 2, 3)$. Môžeme skonštruovať príslušnú Helmertovu maticu $\mathbf{H}$, kde $\mathbf{x} = (1, 2, 3)^T$:

$$\mathbf{H} = \begin{pmatrix} c_1(1, 2, 3) \\ c_2(1, -\frac{1}{2}, 0) \\ c_3(1, 2, -\frac{1+4}{3}) \end{pmatrix},$$

pričom

$$c_1 = \frac{1}{\sqrt{1+4+9}} = \frac{1}{\sqrt{14}}, \quad c_2 = \frac{1}{\sqrt{1+1/4}} = \frac{2}{\sqrt{5}}, \quad c_3 = \frac{1}{\sqrt{1+4+25/9}} = \frac{3}{\sqrt{70}}.$$

Teda

$$\mathbf{H} = \begin{pmatrix} \frac{1}{\sqrt{14}} & \frac{2}{\sqrt{14}} & \frac{3}{\sqrt{14}} \\ \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} & 0 \\ \frac{3}{\sqrt{70}} & \frac{6}{\sqrt{70}} & -\frac{5}{\sqrt{70}} \end{pmatrix}$$

je ortogonálna matica spĺňajúca zadanú podmienku. Skonštruujte Helmertovu maticu typu 4×4 pre $\mathbf{x} = (1, 2, 3, 4)^T$.

Poznámka 4.39. Helmertova matica nám umožňuje nájsť ortonormálnu bázu vektorového podpriestoru kolmého na zvolený vektor $\mathbf{x}$. Špeciálne pre $\mathbf{x} = \mathbf{1}_n$ sa vektorové podpriestory kolmé na $\mathbf{x}$ často používajú v niektorých štatistických metódach, keďže vektor samých jednotiek v nich hrá prirodzene výnimočnú rolu.

4.4.2 Permutačné matice

Definícia 4.40 (Permutácia)**.** *Permutácia* σ množiny $\{1, \dots, n\}$ je bijekcia medzi $\{1, \dots, n\}$ a $\{1, \dots, n\}$.

Definícia 4.41 (Permutačná matica)**.** Matica $\mathbf{P} \in \mathbb{R}^{n \times n}$ sa nazýva permutačná, ak vznikne permutovaním riadkov (alebo stĺpcov) matice identity $\mathbf{I}_n$. Formálne, $\sigma(i)$-ty riadok matice $\mathbf{P}$ je rovný vektoru $\mathbf{e}_i$ (čo je i-ty riadok matice $\mathbf{I}_n$), pre nejakú permutáciu σ množiny $\{1, ..., n\}$.

Príklad 4.42. Permutačnou maticou je napríklad $\mathbf{I}_n$, ale aj

$$\mathbf{P}_1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \quad \mathbf{P}_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

Je matica

$$\mathbf{P}_3 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

permutačnou maticou?

Poznámka 4.43. Ľahko sa dá nahliadnuť, že pre permutačné matice platí:

(i) Pre $n \in \mathbb{N}$ existuje $n!$ rôznych permutačných matíc typu $n \times n$.

(ii) Každá permutačná matica je ortogonálna.

(iii) $\mathbf{P}$ je permutačná matica práve vtedy, keď má v každom riadku aj v každom stĺpci práve jednu jednotku a jej ostatné prvky sú nuly.

(iv) Ak $\mathbf{A} \in \mathbb{R}^{m \times n}$ a $\mathbf{P} \in \mathbb{R}^{n \times n}$ je permutačná matica, tak $\mathbf{AP}$ je matica $\mathbf{A}$ so spermutovanými stĺpcami.

(v) Ak $\mathbf{A} \in \mathbb{R}^{n \times m}$ a $\mathbf{P} \in \mathbb{R}^{n \times n}$ je permutačná matica, tak $\mathbf{PA}$ je matica $\mathbf{A}$ so spermutovanými riadkami.

(vi) Súčin permutačných matíc je permutačná matica.

4.5 Úlohy na precvičenie

Úloha 4.1. Ukážte, že množina všetkých riešení konzistentného systému je afinná. Zdôvodnite teda, že ak $\mathbf{AX} = \mathbf{B}$ je konzistentný systém, $\mathbf{X}_1, \ldots, \mathbf{X}_k$ sú riešenia tohto systému a $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$ spĺňajú $\sum_{i=1}^k \alpha_i = 1$, tak aj $\sum_{i=1}^k \alpha_i \mathbf{X}_i$ je riešením systému $\mathbf{AX} = \mathbf{B}$.

Úloha 4.2. Nech $\mathbf{A} = a\mathbf{I}_n + b\mathbf{J}_{n \times n}$, $a, b \in \mathbb{R}$, je úplne symetrická matica (pozri príklad 2.4). Ukážte, že ak $a \neq 0$, $b \neq -a/n$, potom inverzia matice $\mathbf{A}$ je tiež úplne symetrická matica.

Úloha 4.3. Dokážte tvrdenie: Nech $\mathbf{A} \in \mathbb{R}^{m \times m}$, $\mathbf{D} \in \mathbb{R}^{n \times n}$ sú regulárne matice a nech $\mathbf{B} \in \mathbb{R}^{m \times n}$ a $\mathbf{C} \in \mathbb{R}^{n \times m}$ sú také matice, že $\mathbf{D}^{-1} + \mathbf{CA}^{-1}\mathbf{B}$ je tiež regulárna. Potom je aj matica $\mathbf{A} + \mathbf{BDC}$ regulárna a platí

$$(\mathbf{A} + \mathbf{BDC})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{D}^{-1} + \mathbf{CA}^{-1}\mathbf{B})^{-1}\mathbf{CA}^{-1}. \tag{4.1}$$

Rovnosť (4.1) sa nazýva Woodburyho formula.

Úloha 4.4. Presvedčte sa, že horná trojuholníková matica je regulárna práve vtedy, keď sú všetky jej diagonálne prvky nenulové. Na zvolenej hornej trojuholníkovej matici typu 4×4 so všetkými diagonálnymi a nad-diagonálnymi prvkami nenulovými demonštrujte, že inverzia regulárnej hornej trojuholníkovej matice je horná trojuholníková matica.

Úloha 4.5. Nech $\mathbf{A} \in \mathbb{R}^{n \times n}$ je regulárna matica, nech $\mathbf{G} \in \mathbb{R}^{n \times n}$ a nech p je prirodzené číslo. Ukážte, že nasledovné dva výroky sú ekvivalentné: a) Pre každú maticu $\mathbf{B} \in \mathbb{R}^{n \times p}$ má rovnica $\mathbf{AX} = \mathbf{B}$ riešenie $\mathbf{X} = \mathbf{GB}$; b) $\mathbf{G} = \mathbf{A}^{-1}$.

Úloha 4.6. Nech

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

Pomocou vety 4.26 o inverzii blokovej matice alebo pomocou blokovej eliminácie ukážte, že $\mathbf{A}$ je regulárna matica a nájdite $\mathbf{A}^{-1}$.

Úloha 4.7. Čo hovoria tvrdenia z podkapitoly 4.3 ak $\mathbf{T}, \mathbf{U}, \mathbf{V}, \mathbf{W}, \mathbf{Q} \in \mathbb{R}$, teda ak jednotlivé bloky sú iba reálne čísla? Interpretujte pomocou známeho vzorca pre determinant 2×2 matice.

Úloha 4.8. Pre každé $\theta \in \mathbb{R}$ definujme maticu

$$\mathbf{U}(\theta) = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}.$$

Zdôvodnite algebraicky aj geometricky tvrdenie: $\mathbf{U}(\alpha)\mathbf{U}(\beta) = \mathbf{U}(\alpha + \beta)$ pre každé $\alpha, \beta \in \mathbb{R}$.

Úloha 4.9. Opíšte množinu všetkých ortogonálnych horných trojuholníkových matíc typu $m \times m$.

Úloha 4.10. Nájdite nejakú ortogonálnu maticu typu 3×3, ktorej prvý riadok je $(\frac{1}{2}, \frac{1}{2}, \frac{1}{\sqrt{2}})$.

Úloha 4.11. Nájdite ortogonálnu maticu typu 4×4, ktorej prvý riadok je $(\frac{1}{\sqrt{3}}, 0, \frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}})$.

Úloha 4.12. Skonštruujte Helmertovu maticu typu $n \times n$, ktorej prvý riadok je $\mathbf{1}_n^T$.

Úloha 4.13. Nech $\mathbf{U} \in \mathbb{R}^{n \times n}$. Dokážte, že ak $\|\mathbf{U}\mathbf{x}\| = \|\mathbf{x}\|$ pre každé $\mathbf{x} \in \mathbb{R}^n$, tak $\mathbf{U}$ je ortogonálna. *Pomôcka:* môžete použiť lemu 3.27.

Úloha 4.14. Nech $\mathbf{P}_1, \mathbf{P}_2 \in \mathbb{R}^{m \times m}$ sú permutačné matice a nech $\mathbf{A} \in \mathbb{R}^{m \times m}$ je regulárna matica. Ukážte, že potom $\mathbf{P}_1\mathbf{A}\mathbf{P}_2$ je regulárna matica a platí $(\mathbf{P}_1\mathbf{A}\mathbf{P}_2)^{-1} = \mathbf{P}_2^T\mathbf{A}^{-1}\mathbf{P}_1^T$.

Úloha 4.15. Koľko obsahuje priestor $\mathbb{R}^{m \times m}$ rôznych permutačných matíc? Ak je m prvočíslo, koľko je permutačných matíc $\mathbf{P}$ spĺňajúcich $\mathbf{P}^m = \mathbf{I}_m$?

Úloha* 4.16. Koľko obsahuje priestor $\mathbb{R}^{m \times m}$ permutačných matíc $\mathbf{P}$ spĺňajúcich $\mathbf{P}^2 = \mathbf{I}_m$? (Ak úlohu neviete vyriešiť pre všeobecné m, vyriešte ju aspoň pre $m = 2, 3, 4$.)

Kapitola 5

Priestor matíc a jeho geometria

5.1 Priestor matíc

Veta 5.1 (O priestore $\mathbb{R}^{m\times n}$)**.** Množina $\mathbb{R}^{m\times n}$ všetkých matíc typu $m \times n$ spolu so sčítaním a násobením skalárom tvorí vektorový priestor.

Dôkaz. Požadované vlastnosti vyplývajú z vety 2.10; nulová matica (nulový prvok vektorového priestoru) je $\mathbf{0}_{m\times n}$ a opačná matica (opačný prvok) ku matici $\mathbf{A}$ je $-\mathbf{A} = (-1) \cdot \mathbf{A}$. □

Poznámka 5.2 (Základné pojmy pre vektorový priestor matíc)**.** Keďže priestor všetkých matíc typu $m \times n$ je vektorový priestor, môžeme používať terminológiu zavedenú v kapitole 1. Menovite platí:

1. Podmnožina $V \subseteq \mathbb{R}^{m\times n}$ je lineárny podpriestor, ak samotné V je vektorový priestor. Navyše, V je lineárny podpriestor priestoru $\mathbb{R}^{m\times n}$ vtedy a len vtedy, keď pre každé $\mathbf{A}, \mathbf{B} \in V$ a $c_1, c_2 \in \mathbb{R}$ platí $c_1\mathbf{A} + c_2\mathbf{B} \in V$.

2. *Vektorový obal matíc* $\mathbf{A}_1, \ldots, \mathbf{A}_k \in \mathbb{R}^{m\times n}$ je množina

$$\operatorname{span}(\mathbf{A}_1, \ldots, \mathbf{A}_k) = \{\mathbf{A} \in \mathbb{R}^{m\times n} \mid \mathbf{A} = c_1\mathbf{A}_1 + \ldots + c_k\mathbf{A}_k \text{ pre nejaké } c_1, \ldots, c_k \in \mathbb{R}\}$$

   a je to zároveň najmenší lineárny podpriestor generovaný týmito maticami.

3. Matice $\mathbf{A}_1, \ldots, \mathbf{A}_k \in \mathbb{R}^{m\times n}$ nazývame *lineárne nezávislé*, ak rovnosť $c_1\mathbf{A}_1+\ldots+c_k\mathbf{A}_k = \mathbf{0}_{m\times n}$ pre $c_1, \ldots, c_k \in \mathbb{R}$ implikuje $c_1 = \ldots = c_k = 0$.

4. Matice $\mathbf{A}_1, \ldots, \mathbf{A}_k \in \mathbb{R}^{m\times n}$ tvoria *bázu* podpriestoru $V \subseteq \mathbb{R}^{m\times n}$, ak sú nezávislé a $V = \operatorname{span}(\mathbf{A}_1, \ldots, \mathbf{A}_k)$. Počet prvkov bázy nazývame *dimenziou* V.

Príklad 5.3. Ukážte, že množina $\mathcal{S}^n$ všetkých symetrických matíc typu $n \times n$ tvorí lineárny podpriestor v $\mathbb{R}^{n\times n}$.

Príklad 5.4. Bázou priestoru $\mathbb{R}^{2\times 2}$ sú napríklad matice

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Bázou podpriestoru $\mathcal{S}^2$ všetkých symetrických matíc typu 2×2 sú napríklad matice

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Takže $\dim(\mathbb{R}^{2\times 2}) = 4$ a $\dim(\mathcal{S}^2) = 3$. Akú má vo všeobecnosti dimenziu priestor $\mathbb{R}^{m\times n}$ a akú dimenziu má $\mathcal{S}^n$?

Veta 5.5 (Operátor vec ako lineárny izomorfizmus)**.** Operátor vec je lineárny izomorfizmus medzi $\mathbb{R}^{m\times n}$ a $\mathbb{R}^{mn}$, teda priestor matíc typu $m \times n$ je lineárne izomorfný s priestorom mn-rozmerných vektorov.

Dôkaz. Vďaka vlastnostiam (i) a (ii) z vety 2.46 je vec lineárne zobrazenie (pozri definíciu 1.53). Na dôkaz surjektívnosti uvážme $\mathbf{x} \in \mathbb{R}^{mn}$ a zapíšme ho v blokovom zápise $\mathbf{x} = (\mathbf{x}_1^T, \ldots, \mathbf{x}_n^T)^T$, kde $\mathbf{x}_i \in \mathbb{R}^m$ pre $i = 1, \ldots, n$. Potom $\mathrm{vec}((\mathbf{x}_1, \ldots, \mathbf{x}_n)) = \mathbf{x}$. Toto zobrazenie je injektívne, lebo ak $\mathrm{vec}(\mathbf{A}) = \mathrm{vec}(\mathbf{B})$, tak zjavne $(\mathbf{A})_{ij} = (\mathbf{B})_{ij}$ pre každé i, j. $\square$

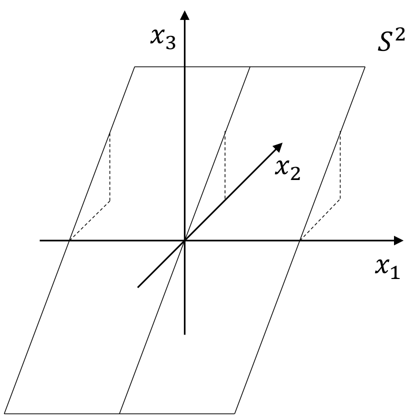
Príklad 5.6. Zobrazme priestor matíc $\mathbb{R}^{2\times 2}$ pomocou operátoru vec ako priestor vektorov $\mathbb{R}^4$, teda maticu

$$\begin{pmatrix} x_1 & x_3 \\ x_2 & x_4 \end{pmatrix}$$

vyjadríme ako vektor $\mathbf{x} = (x_1, x_2, x_3, x_4)^T$. Množinu symetrických matíc $\mathcal{S}^2$ potom vieme vyjadriť ako podpriestor priestoru $\mathbb{R}^4$, určený lineárnou rovnosťou $x_3 = x_2$ (čiže ako nadrovinu).

Ak uvážime iba matice s fixným prvkom $x_4 = 0$, tak priestor $\mathbb{R}^{2\times 2}$ vieme zakresliť ako $\mathbb{R}^3$, a $\mathcal{S}^2$ môžeme zobraziť ako rovinu v tomto priestore určenú rovnicou $x_3 = x_2$, ako môžeme vidieť na obrázku 5.1. Priestor $\mathcal{S}^2$ teda vieme vyjadriť ako nadrovinu v $\mathbb{R}^4$, ktorá v priestore určenom prvými troma súradnicami má tvar ako na obrázku 5.1, pričom okrem ľubovoľnej hodnoty x_1 môžu jej body nadobúdať aj ľubovoľnú hodnotu x_4. Kde by sa na obrázku 5.1 vyskytovala matica

$$\begin{pmatrix} 2 & 3 \\ 3 & 0 \end{pmatrix}?$$



Obr. 5.1: Znázornenie matíc z $\mathbb{R}^{2\times 2}$ s nulovou hodnotou na pozícii $(2, 2)$ ako $\mathbb{R}^3$ a zakreslenie symetrických matíc $\mathcal{S}^2$ (s tou istou fixovanou hodnotou) ako roviny v tomto priestore.

Veta 5.7 (Operátor vech ako lineárny izomorfizmus)**.** Operátor vech je lineárny izomorfizmus medzi $\mathcal{S}^n$ a $\mathbb{R}^{n(n+1)/2}$, teda priestor symetrických matíc typu $n \times n$ je lineárne izomorfný s priestorom $n(n+1)/2$-rozmerných vektorov.

Dôkaz. Z lemy 2.50 vyplýva, že vech je lineárne zobrazenie. Zároveň je injektívne, lebo ak $\mathbf{A}$, $\mathbf{B}$ sú symetrické matice a $\text{vech}(\mathbf{A}) = \text{vech}(\mathbf{B})$, tak $(\mathbf{A})_{ij} = (\mathbf{B})_{ij}$ pre $i \geq j$ a zo symetrickosti potom vyplýva aj $(\mathbf{A})_{ij} = (\mathbf{B})_{ij}$ pre $i < j$. Na dôkaz surjektívnosti uvážme $\mathbf{x} \in \mathbb{R}^{n(n+1)/2}$ a zapíšme ho v blokovom zápise $\mathbf{x} = (\mathbf{x}_1^T, \dots, \mathbf{x}_n^T)^T$, kde $\mathbf{x}_1 \in \mathbb{R}^n, \mathbf{x}_2 \in \mathbb{R}^{n-1}, \dots, \mathbf{x}_n \in \mathbb{R}^1$. Potom, ak $\mathbf{A} \in \mathcal{S}^n$ má v dolnom trojuholníku prvky $\mathbf{x}_1, \dots, \mathbf{x}_n$, tak $\mathbf{x} = \text{vech}(\mathbf{A})$. □

5.2 Stopa matice

Definícia 5.8 (Stopa matice)**.** *Stopa* štvorcovej matice $\mathbf{A} \in \mathbb{R}^{n\times n}$ je súčet jej diagonálnych prvkov, značíme ju $\text{tr}(\mathbf{A}) = \sum_{i=1}^n (\mathbf{A})_{ii}$.

Poznámka 5.9. Značenie $\text{tr}(\mathbf{A})$ vychádza z anglického *trace*. V slovenskej literatúre sa môžeme stretnúť aj so značením $\text{st}(\mathbf{A})$ od slova “stopa”.

Veta 5.10 (Základné vlastnosti stopy)**.** Nech $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n\times n}$ a $c \in \mathbb{R}$. Potom $\text{tr}(c\mathbf{A}) = c \cdot \text{tr}(\mathbf{A})$ a $\text{tr}(\mathbf{A}+\mathbf{B}) = \text{tr}(\mathbf{A}) + \text{tr}(\mathbf{B})$, čiže zobrazenie $\text{tr} : \mathbb{R}^{n\times n} \to \mathbb{R}$, ktoré priraďuje štvorcovej matici jej stopu, je lineárne.

Dôkaz. Dôkaz nechávame na čitateľa. □

Poznámka 5.11. Pre každú $\mathbf{A} \in \mathbb{R}^{n\times n}$ platí $\text{tr}(\mathbf{A}) = \text{tr}(\mathbf{A}^T)$.

Veta 5.12 (O výmene poradia násobenia v stope)**.** Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{B} \in \mathbb{R}^{n\times m}$. Potom $\text{tr}(\mathbf{AB}) = \text{tr}(\mathbf{BA})$.

Dôkaz. Počítajme

$$\text{tr}(\mathbf{AB}) = \sum_{i=1}^m (\mathbf{AB})_{ii} = \sum_{i=1}^m \sum_{j=1}^n (\mathbf{A})_{ij}(\mathbf{B})_{ji} = \sum_{j=1}^n \sum_{i=1}^m (\mathbf{B})_{ji}(\mathbf{A})_{ij} = \text{tr}(\mathbf{BA}).$$

□

Poznámka 5.13. Špeciálne pre $\mathbf{x} \in \mathbb{R}^{n\times 1}$ dostávame $\text{tr}(\mathbf{x}\mathbf{x}^T) = \text{tr}(\mathbf{x}^T\mathbf{x}) = \mathbf{x}^T\mathbf{x} = \|\mathbf{x}\|^2$.

Poznámka 5.14. Pre súčin viacerých matíc platí, že poradie násobenia môžeme meniť “cyklicky” (ak sedia rozmery), ale nie ľubovoľne. Teda napríklad pre $\mathbf{A}, \mathbf{B}, \mathbf{C} \in \mathbb{R}^{n\times n}$ platí $\text{tr}(\mathbf{ABC}) = \text{tr}(\mathbf{BCA}) = \text{tr}(\mathbf{CAB})$, ale vo všeobecnosti neplatí $\text{tr}(\mathbf{ABC}) = \text{tr}(\mathbf{CBA})$.

Dôsledok 5.15. Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{B} \in \mathbb{R}^{n\times m}$. Potom $\text{tr}(\mathbf{AB}) = \text{tr}(\mathbf{BA}) = \text{tr}(\mathbf{B}^T\mathbf{A}^T) = \text{tr}(\mathbf{A}^T\mathbf{B}^T)$.

Lema 5.16 (Kladná definitnosť stopy)**.** Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$. Potom $\text{tr}(\mathbf{A}^T\mathbf{A}) \geq 0$. Navyše $\mathbf{A} = \mathbf{0}_{m\times n}$ práve vtedy, keď $\text{tr}(\mathbf{A}^T\mathbf{A}) = 0$.

Dôkaz. Keďže $\text{tr}(\mathbf{A}^T\mathbf{A}) = \sum_{j=1}^n \sum_{i=1}^m (\mathbf{A})_{ij}^2$, tak $\text{tr}(\mathbf{A}^T\mathbf{A}) \geq 0$. Nech

$$0 = \text{tr}(\mathbf{A}^T\mathbf{A}) = \sum_{j=1}^n \sum_{i=1}^m (\mathbf{A})_{ij}^2.$$

Z toho vyplýva, že $(\mathbf{A})_{ij} = 0$ pre každé i a j, čiže $\mathbf{A} = \mathbf{0}$. Dôkaz opačnej implikácie je triviálny. □

5.3 Geometria priestoru matíc

5.3.1 Skalárny súčin a norma

Veta 5.17 (O skalárnom súčine v priestore matíc)**.** Zobrazenie $\langle\, ,\, \rangle : \mathbb{R}^{m\times n} \times \mathbb{R}^{m\times n} \to \mathbb{R}$ definované ako $\langle \mathbf{A}, \mathbf{B}\rangle = \mathrm{tr}(\mathbf{A}^T\mathbf{B})$ pre každé $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m\times n}$ predstavuje skalárny súčin na $\mathbb{R}^{m\times n}$.

Dôkaz. Overme vlastnosti skalárneho súčinu podľa definície 1.29. Symetrickosť vyplýva z dôsledku 5.15. Druhá a tretia vlastnosť vyplývajú z vety 5.10 a posledná vlastnosť platí podľa lemy 5.16. □

Dôsledok 5.18. Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$. Potom $\|\mathbf{A}\| = (\mathrm{tr}(\mathbf{A}^T\mathbf{A}))^{1/2}$ je norma na $\mathbb{R}^{m\times n}$ a uhol medzi nenulovými maticami $\mathbf{A}$ a $\mathbf{B}$ je $\alpha \in [0, \pi]$, kde

$$\cos(\alpha) = \frac{\mathrm{tr}(\mathbf{A}^T\mathbf{B})}{\mathrm{tr}^{1/2}(\mathbf{A}^T\mathbf{A}) \cdot \mathrm{tr}^{1/2}(\mathbf{B}^T\mathbf{B})} = \frac{\langle \mathbf{A}, \mathbf{B}\rangle}{\|\mathbf{A}\| \cdot \|\mathbf{B}\|}.$$

Poznámka 5.19. Skalárny súčin $\langle \mathbf{A}, \mathbf{B}\rangle = \mathrm{tr}(\mathbf{A}^T\mathbf{B})$ na $\mathbb{R}^{m\times n}$ sa vzhľadom na poznámku 5.11 a vetu 5.12 dá vyjadriť aj ako $\langle \mathbf{A}, \mathbf{B}\rangle = \mathrm{tr}(\mathbf{B}^T\mathbf{A}) = \mathrm{tr}(\mathbf{B}\mathbf{A}^T) = \mathrm{tr}(\mathbf{A}\mathbf{B}^T)$.

Príklad 5.20. Overte, že $\mathrm{tr}(\mathbf{A}^T\mathbf{B}) = \sum_{i=1}^m \sum_{j=1}^n (\mathbf{A})_{i,j}(\mathbf{B})_{i,j}$ pre každé $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m\times n}$. Teda $\mathrm{tr}(\mathbf{A}^T\mathbf{B})$ je súčet všetkých zložiek matice $\mathbf{A} \odot \mathbf{B}$, kde $\odot$ značí Hadamardov súčin.

Poznámka 5.21. Ďalej budeme v priestore $\mathbb{R}^{m\times n}$ vždy uvažovať skalárny súčin $\langle \mathbf{A}, \mathbf{B}\rangle = \mathrm{tr}(\mathbf{A}^T\mathbf{B})$ a príslušnú normu $\|\mathbf{A}\| = (\mathrm{tr}(\mathbf{A}^T\mathbf{A}))^{1/2} = (\sum_{j=1}^m \sum_{i=1}^n (\mathbf{A})_{ij}^2)^{1/2}$. Táto maticová norma sa nazýva tiež *Frobeniova norma* a niekedy sa preto značí $\|\,.\,\|_F$.

Poznámka 5.22. Podľa príkladu 5.20 skalárny súčin $\langle \mathbf{A}, \mathbf{B}\rangle = \mathrm{tr}(\mathbf{A}^T\mathbf{B})$ zodpovedá bežnému skalárnemu súčinu $\langle \mathbf{x}, \mathbf{y}\rangle = \mathbf{x}^T\mathbf{y}$ na priestore $\mathbb{R}^{mn}$. Presnejšie, $\mathrm{tr}(\mathbf{A}^T\mathbf{B}) = \mathrm{vec}(\mathbf{A})^T\mathrm{vec}(\mathbf{B})$, čiže $\langle \mathbf{A}, \mathbf{B}\rangle = \langle \mathrm{vec}(\mathbf{A}), \mathrm{vec}(\mathbf{B})\rangle$. Podobne teda maticová norma zodpovedá bežnej euklidovskej norme na $\mathbb{R}^{mn}$: $\|\mathbf{A}\| = \|\mathrm{vec}(\mathbf{A})\|$, a uhol medzi dvoma maticami $\mathbf{A}$ a $\mathbf{B}$ je rovný uhlu medzi $\mathrm{vec}(\mathbf{A})$ a $\mathrm{vec}(\mathbf{B})$.

Veta 5.23 (Cauchyho-Schwarzova nerovnosť pre matice)**.** Nech $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m\times n}$. Potom platí $\mathrm{tr}^2(\mathbf{A}^T\mathbf{B}) \leq \mathrm{tr}(\mathbf{A}^T\mathbf{A})\mathrm{tr}(\mathbf{B}^T\mathbf{B})$. Navyše rovnosť nastáva práve vtedy, keď $\mathbf{B} = \mathbf{0}_{m\times n}$ alebo $\mathbf{A} = c\mathbf{B}$ pre nejaké $c \in \mathbb{R}$.

Dôkaz. Veta platí podľa všeobecnejšej vety 1.34, dokážme ju ale priamo. Ak $\mathbf{B} = \mathbf{0}$, tak nerovnosť zjavne platí v tvare $0 \leq 0$. Môžeme predpokladať $\mathbf{B} \neq \mathbf{0}$, a teda $\|\mathbf{B}\| > 0$. Pre každé $x, y \in \mathbb{R}$ platí

$$\begin{aligned} 0 \leq \mathrm{tr}((x\mathbf{A} - y\mathbf{B})^T(x\mathbf{A} - y\mathbf{B})) &= \mathrm{tr}(x^2\mathbf{A}^T\mathbf{A} - xy\mathbf{A}^T\mathbf{B} - xy\mathbf{B}^T\mathbf{A} + y^2\mathbf{B}^T\mathbf{B}) \\ &= x^2\mathrm{tr}(\mathbf{A}^T\mathbf{A}) - 2xy\mathrm{tr}(\mathbf{A}^T\mathbf{B}) + y^2\mathrm{tr}(\mathbf{B}^T\mathbf{B}). \end{aligned} \tag{5.1}$$

Špeciálne môžeme položiť $x = \mathrm{tr}^{1/2}(\mathbf{B}^T\mathbf{B})$ a keďže predpokladáme $\|\mathbf{B}\| > 0$, môžeme tiež položiť $y = \mathrm{tr}(\mathbf{A}^T\mathbf{B})/\mathrm{tr}^{1/2}(\mathbf{B}^T\mathbf{B})$. Dostávame tak

$$0 \leq \mathrm{tr}(\mathbf{A}^T\mathbf{A})\mathrm{tr}(\mathbf{B}^T\mathbf{B}) - 2\mathrm{tr}^2(\mathbf{A}^T\mathbf{B}) + \mathrm{tr}^2(\mathbf{A}^T\mathbf{B}),$$

a teda $\mathrm{tr}^2(\mathbf{A}^T\mathbf{B}) \leq \mathrm{tr}(\mathbf{A}^T\mathbf{A})\mathrm{tr}(\mathbf{B}^T\mathbf{B})$.

Ak $\mathbf{A} = c\mathbf{B}$, tak zjavne nastáva rovnosť. Naopak, aby nastala rovnosť v Cauchyho-Schwarzovej nerovnosti, musí nastať aj v nerovnosti (5.1) pre x, y zvolené vyššie. Potom z lemy 5.16 vyplýva, že $x\mathbf{A} - y\mathbf{B} = \mathbf{0}$. Keďže $\|\mathbf{B}\| > 0$, tak $x > 0$. Preto rovnosť $x\mathbf{A} - y\mathbf{B} = \mathbf{0}$ môžeme zapísať ako $\mathbf{A} = (y/x)\mathbf{B}$. □

5.3.2 Ortogonálnosť

Poznámka 5.24. Keďže $\langle \mathbf{A}, \mathbf{B} \rangle = \mathrm{tr}(\mathbf{A}^T\mathbf{B})$ tvorí skalárny súčin na vektorovom priestore $\mathbb{R}^{m\times n}$, tak môžeme používať terminológiu týkajúcu sa vektorových priestorov so skalárnym súčinom. Špeciálne, množina matíc $\{\mathbf{A}_1, \ldots, \mathbf{A}_k\} \subseteq \mathbb{R}^{m\times n}$ sa nazýva *ortogonálna*, ak $\langle \mathbf{A}_i, \mathbf{A}_j \rangle = 0$ pre každé $i \neq j$. Ak navyše $\|\mathbf{A}_i\| = 1$ pre $i = 1, \ldots, k$, tak táto množina sa nazýva *ortonormálna*.

Lema 5.25 (Ortogonálnosť a nezávislosť matíc)**.** Nech množina matíc $\{\mathbf{A}_1, \ldots, \mathbf{A}_k\} \subseteq \mathbb{R}^{m\times n}$ je ortogonálna a nech $\mathbf{A}_i \neq \mathbf{0}_{m\times n}$ pre každé $i = 1, \ldots, k$. Potom $\mathbf{A}_1, \ldots, \mathbf{A}_k$ sú lineárne nezávislé.

Dôkaz. Nech $x_1\mathbf{A}_1 + \ldots + x_k\mathbf{A}_k = \mathbf{0}$ pre $x_1, \ldots, x_k \in \mathbb{R}$. Potom $\langle x_1\mathbf{A}_1 + \ldots + x_k\mathbf{A}_k, \mathbf{A}_i \rangle = 0$ pre $i = 1, \ldots, k$, čiže $x_1\langle \mathbf{A}_1, \mathbf{A}_i \rangle + \ldots + x_k\langle \mathbf{A}_k, \mathbf{A}_i \rangle = 0$, čiže špeciálne $x_i\langle \mathbf{A}_i, \mathbf{A}_i \rangle = 0$. Keďže $\mathbf{A}_i \neq \mathbf{0}$, tak $\langle \mathbf{A}_i, \mathbf{A}_i \rangle \neq 0$, a teda $x_i = 0$. Dostávame $x_1 = \ldots = x_k = 0$. □

Poznámka 5.26. Lema 5.25 je špeciálnym prípadom vety 1.40.

Veta 5.27 (Gramova-Schmidtova ortogonalizácia matíc)**.** Nech $\mathbf{A}_1, \ldots, \mathbf{A}_k \in \mathbb{R}^{m\times n}$ je množina lineárne nezávislých matíc. Definujme $\mathbf{B}_1, \ldots, \mathbf{B}_k$ nasledovne:

$$\begin{aligned}
\mathbf{B}_1 &= \mathbf{A}_1, \\
\mathbf{B}_2 &= \mathbf{A}_2 - x_{1,2}\mathbf{B}_1, \\
\mathbf{B}_3 &= \mathbf{A}_3 - x_{2,3}\mathbf{B}_2 - x_{1,3}\mathbf{B}_1, \\
&\ \vdots \\
\mathbf{B}_k &= \mathbf{A}_k - \sum_{i=1}^{k-1} x_{i,k}\mathbf{B}_i,
\end{aligned}$$

kde $x_{i,j} = \langle \mathbf{A}_j, \mathbf{B}_i \rangle / \langle \mathbf{B}_i, \mathbf{B}_i \rangle$, pre každé i, j. Potom $\mathbf{B}_1, \ldots, \mathbf{B}_k$ tvoria ortogonálnu bázu podpriestoru $\mathrm{span}(\mathbf{A}_1, \ldots, \mathbf{A}_k)$. Navyše $\mathbf{C}_1, \ldots, \mathbf{C}_k$, kde $\mathbf{C}_i = \mathbf{B}_i/\|\mathbf{B}_i\|$ pre $i = 1, \ldots, k$, tvoria ortonormálnu bázu tohto podpriestoru.

Dôkaz. Veta je dôsledkom vety 1.42 □

Veta 5.28 (O ortonormálnej báze)**.** Každý lineárny podpriestor priestoru $\mathbb{R}^{m\times n}$ má ortonormálnu bázu.

Dôkaz. Každý lineárny podpriestor má bázu $\mathbf{A}_1, \ldots, \mathbf{A}_k$ a tú môžeme pomocou Gramovej-Schmidtovej ortogonalizácie transformovať na ortonormálnu bázu $\mathbf{C}_1, \ldots, \mathbf{C}_k$. Veta je tiež špeciálnym prípadom dôsledku 1.44. □

Veta 5.29 (O súradniciach vzhľadom na ortonormálnu bázu)**.** Nech $\mathbf{A}_1, \ldots, \mathbf{A}_k$ tvoria ortonormálnu bázu lineárneho podpriestoru $V \subseteq \mathbb{R}^{m\times n}$. Potom pre každú maticu $\mathbf{A} \in V$ platí

$$\mathbf{A} = \langle \mathbf{A}, \mathbf{A}_1 \rangle \mathbf{A}_1 + \ldots + \langle \mathbf{A}, \mathbf{A}_k \rangle \mathbf{A}_k.$$

Dôkaz. Keďže $\mathbf{A}_1, \ldots, \mathbf{A}_k$ tvoria bázu, tak existujú $x_1, \ldots, x_k \in \mathbb{R}$, že $\mathbf{A} = x_1\mathbf{A}_1 + \ldots + x_k\mathbf{A}_k$. Potom $\langle \mathbf{A}, \mathbf{A}_i \rangle = x_1\langle \mathbf{A}_1, \mathbf{A}_i \rangle + \ldots + x_k\langle \mathbf{A}_k, \mathbf{A}_i \rangle = x_i\langle \mathbf{A}_i, \mathbf{A}_i \rangle = x_i$, keďže $\langle \mathbf{A}_j, \mathbf{A}_i \rangle = 0$ pre $j \neq i$ a $\langle \mathbf{A}_i, \mathbf{A}_i \rangle = 1$. Z toho vyplýva, že $x_i = \langle \mathbf{A}, \mathbf{A}_i \rangle$, $i = 1, \ldots, k$. □

5.4 Úlohy na precvičenie

Úloha 5.1. Nech $\mathbf{A}_1, \ldots, \mathbf{A}_k \in \mathbb{R}^{m\times n}$. Ukážte, že existujú indexy $1 \leq i_1 < \ldots < i_r \leq k$ také, že $\mathbf{A}_{i_1}, \ldots, \mathbf{A}_{i_r}$ tvoria bázu priestoru $\mathrm{span}(\mathbf{A}_1, \ldots, \mathbf{A}_k)$.

Úloha 5.2. Presvedčte sa, že množina $\mathcal{U}_n$ všetkých horných trojuholníkových matíc typu $n\times n$ a množina $\mathcal{L}_n$ všetkých dolných trojuholníkových matíc typu $n \times n$ sú lineárne podpriestory priestoru $\mathbb{R}^{n\times n}$. Akú majú tieto priestory dimenziu? Charakterizujte lineárny priestor $\mathcal{D}_n = \mathcal{U}_n \cap \mathcal{L}_n$. Akú dimenziu má priestor $\mathcal{D}_n$?

Úloha 5.3. Antisymetrickou nazývame každú maticu $\mathbf{A} \in \mathbb{R}^{n\times n}$, ktorá spĺňa $\mathbf{A}^T = -\mathbf{A}$. Zdôvodnite, prečo množina všetkých antisymetrických matíc typu $n\times n$ tvorí lineárny podpriestor priestoru $\mathbb{R}^{n\times n}$. Akú dimenziu má tento lineárny podpriestor?

Úloha 5.4. Nech $\mathcal{K}(\mathbf{A})$ je množina všetkých $n \times n$ matíc, ktoré komutujú s danou maticou $\mathbf{A} \in \mathbb{R}^{n\times n}$. Ukážte, že pre každdú $\mathbf{A} \in \mathbb{R}^{n\times n}$ platí: $\mathcal{K}(\mathbf{A}) \neq \emptyset$ a $\mathcal{K}(\mathbf{A})$ je lineárny podpriestor $\mathbb{R}^{n\times n}$.

Úloha 5.5. Dokážte, že ak sú matice $\mathbf{A}, \mathbf{B}, \mathbf{C} \in \mathbb{R}^{n\times n}$ symetrické, tak $\mathrm{tr}(\mathbf{ABC}) = \mathrm{tr}(\mathbf{BAC})$. Ukážte, že rovnosť $\mathrm{tr}(\mathbf{ABC}) = \mathrm{tr}(\mathbf{BAC})$ neplatí pre matice $\mathbf{A}, \mathbf{B}, \mathbf{C} \in \mathbb{R}^{n\times n}$, ak matice $\mathbf{A}, \mathbf{B}$ nekomutujú a $\mathbf{C} = (\mathbf{AB} - \mathbf{BA})^T$.

Úloha 5.6. a) Dokážte Pytagorovu vetu pre matice, čiže dokážte, že ak $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m\times n}$ sú navzájom ortogonálne nenulové matice (t.j. skalárny súčin matíc $\mathbf{A}$ a $\mathbf{B}$ je 0), tak $\|\mathbf{A}\|^2 + \|\mathbf{B}\|^2 = \|\mathbf{A} + \mathbf{B}\|^2$. b) Dokážte rovnobežníkovú vetu pre matice, čiže dokážte, že ak $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m\times n}$, tak $2\|\mathbf{A}\|^2 + 2\|\mathbf{B}\|^2 = \|\mathbf{A} + \mathbf{B}\|^2 + \|\mathbf{A} - \mathbf{B}\|^2$. c) Dokážte polarizačnú rovnosť pre matice, čiže dokážte, že ak $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m\times n}$, tak $\langle \mathbf{A}, \mathbf{B}\rangle = \frac{1}{4}(\|\mathbf{A} + \mathbf{B}\|^2 - \|\mathbf{A} - \mathbf{B}\|^2)$.

Úloha 5.7. Pomocou Cauchyho-Schwarzovej nerovnosti dokážte trojuholníkovú nerovnosť: $\|\mathbf{A} + \mathbf{B}\| \leq \|\mathbf{A}\| + \|\mathbf{B}\|$ pre všetky $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m\times n}$, pričom rovnosť platí len v prípade $\mathbf{B} = \mathbf{0}$ alebo ak $\mathbf{A} = c\mathbf{B}$ pre nejaké $c \in \mathbb{R}$.

Úloha 5.8. Dokážte nasledovné tvrdenie: Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$ a $\mathbf{b} \in \mathbb{R}^n$. Potom $\|\mathbf{Ab}\| \leq \|\mathbf{A}\| \cdot \|\mathbf{b}\|$. *Pomôcka:* $\|(\mathbf{a}_1, \ldots, \mathbf{a}_m)^T\mathbf{b}\|^2 = \sum_{i=1}^m (\mathbf{a}_i^T\mathbf{b})^2$ pre akékoľvek vektory $\mathbf{a}_1, \ldots, \mathbf{a}_m, \mathbf{b} \in \mathbb{R}^n$.

Úloha 5.9. Dokážte kosínusovú vetu pre matice, čiže dokážte, že ak $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m\times n}$ sú nenulové matice, tak $\|\mathbf{A} - \mathbf{B}\|^2 = \|\mathbf{A}\|^2 + \|\mathbf{B}\|^2 - 2\|\mathbf{A}\|\|\mathbf{B}\| \cos(\gamma)$, kde γ je uhol medzi $\mathbf{A}$ a $\mathbf{B}$.

Úloha 5.10. Nech $\mathbf{a}, \mathbf{b}$ sú nenulové n-rozmerné vektory. Nájdite uhol medzi maticami $\mathbf{I}_n, \mathbf{aa}^T$ a uhol medzi maticami $\mathbf{aa}^T, \mathbf{bb}^T$.

Úloha 5.11. a) Nájdite nejakú ortonormálnu bázu priestoru symetrických matíc typu 2×2, ktorá obsahuje maticu $(1, 0)^T(1, 0)$. b) Nájdite nejakú ortonormálnu bázu priestoru symetrických matíc typu 2×2, ktorá obsahuje maticu $(1/\sqrt{2})\mathbf{I}_2$.

Kapitola 6

Zovšeobecnené inverzné matice

V rôznych metódach na analýzu dát sa bežne stretávame s potrebou riešiť systémy lineárnych rovníc $\mathbf{AX} = \mathbf{B}$ alebo $\mathbf{Ax} = \mathbf{b}$. Ak je matica $\mathbf{A}$ regulárna, riešenie ľahko vyjadríme inverznou maticou. Ak však $\mathbf{A}$ regulárna nie je (prípadne ak nie je ani štvorcová), inverznú maticu nemôžeme použiť. Stále však riešenie konzistentného systému vieme vyjadriť maticovo – na to slúžia zovšeobecnené inverzné matice. Aj keď systém nie je konzistentný, pomocou zovšeobecnených inverzií vieme vyjadriť jeho približné riešenie, ako uvidíme v kapitole 7.

6.1 Základné vlastnosti

Definícia 6.1 (Zovšeobecnená inverzia)**.** *Zovšeobecnenou inverznou maticou* (*zovšeobecnenou inverziou*) matice $\mathbf{A} \in \mathbb{R}^{m\times n}$ nazývame ľubovoľnú maticu $\mathbf{G} \in \mathbb{R}^{n\times m}$, pre ktorú platí $\mathbf{AGA} = \mathbf{A}$. Ak $\mathbf{G}$ je zovšeobecnená inverzia matice $\mathbf{A}$, značíme ju $\mathbf{A}^-$ a množinu všetkých zovšeobecnených inverzií matice $\mathbf{A}$ značíme $\mathcal{G}^-(\mathbf{A})$.

Poznámka 6.2. Matica $\mathbf{A}^-$ teda reprezentuje ľubovoľnú zovšeobecnenú inverziu matice $\mathbf{A}$.

Príklad 6.3. Nech

$$\mathbf{A} = \begin{pmatrix} 1 & 3 & 2 \\ 2 & 6 & 4 \end{pmatrix}, \quad \mathbf{G} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Ukážte, že $\mathbf{G}$ je zovšeobecnenou inverziou matice $\mathbf{A}$. Skúste nájsť aspoň jednu zovšeobecnenú inverziu matice $\mathbf{J}_2$ a nájdite množinu $\mathcal{G}^-(\mathbf{0}_{m\times n})$.

Veta 6.4 (Základné vlastnosti zovšeobecnených inverzií)**.** Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$. Potom:

(i) Ak $\mathbf{G} \in \mathcal{G}^-(\mathbf{A})$, tak $c^{-1}\mathbf{G} \in \mathcal{G}^-(c\mathbf{A})$ pre $c \neq 0$.

(ii) Ak $\mathbf{G} \in \mathcal{G}^-(\mathbf{A})$, tak $\mathbf{G}^T \in \mathcal{G}^-(\mathbf{A}^T)$.

(iii) Ak $m = n$ a $\mathbf{A}$ je regulárna, tak $\mathcal{G}^-(\mathbf{A}) = \{\mathbf{A}^{-1}\}$.

(iv) Ak $\mathrm{rank}(\mathbf{A}) = m$, tak $\mathbf{G} \in \mathcal{G}^-(\mathbf{A})$ práve vtedy, keď $\mathbf{G}$ je pravou inverziou $\mathbf{A}$.

(v) Ak $\mathrm{rank}(\mathbf{A}) = n$, tak $\mathbf{G} \in \mathcal{G}^-(\mathbf{A})$ práve vtedy, keď $\mathbf{G}$ je ľavou inverziou $\mathbf{A}$.

Dôkaz. Dokážme tvrdenie (iv). Nech $\mathbf{G} \in \mathcal{G}^-(\mathbf{A})$, potom $\mathbf{AG} = \mathbf{AGI} = \mathbf{AGAR} = \mathbf{AR} = \mathbf{I}$, kde $\mathbf{R}$ je ľubovoľná pravá inverzia matice $\mathbf{A}$, ktorá existuje podľa lemy 4.8. Rovnosť $\mathbf{AG} = \mathbf{I}$ značí, že $\mathbf{G}$ je pravou inverziou $\mathbf{A}$. Naopak, nech $\mathbf{G}$ je pravá inverzia matice $\mathbf{A}$. Potom $\mathbf{AGA} = \mathbf{AI} = \mathbf{A}$, a teda $\mathbf{G} \in \mathcal{G}^-(\mathbf{A})$. Aj dôkazy ostatných tvrdení sú priamočiare, nechávame ich na čitateľa. □

Poznámka 6.5. Neformálne sa dajú tvrdenia (i) a (ii) vety 6.4 formulovať ako $(c\mathbf{A})^- = c^{-1}\mathbf{A}^-$, $(\mathbf{A}^-)^T = (\mathbf{A}^T)^-$. Takýto zápis ale nie je korektný, keďže zovšeobecnená inverzia vo všeobecnosti nie je jediná. Tvrdenie (iii) vieme neformálne zapísať ako $\mathbf{A}^- = \mathbf{A}^{-1}$ v prípade, že $\mathbf{A}$ je regulárna.

Veta 6.6 (O vyjadrení zovšeobecnenej inverzie pomocou $\mathbf{BT}$ rozkladu)**.** Nech $\mathbf{0} \neq \mathbf{A} \in \mathbb{R}^{m\times n}$ má hodnosť $\mathrm{rank}(\mathbf{A}) = r$ a nech $\mathbf{A} = \mathbf{BT}$, kde $\mathbf{B} \in \mathbb{R}^{m\times r}$, $\mathbf{T} \in \mathbb{R}^{r\times n}$ a $\mathrm{rank}(\mathbf{B}) = \mathrm{rank}(\mathbf{T}) = r$. Nech $\mathbf{L} \in \mathbb{R}^{r\times m}$ je ľavá inverzia matice $\mathbf{B}$ a $\mathbf{R} \in \mathbb{R}^{n\times r}$ je pravá inverzia matice $\mathbf{T}$. Potom $\mathbf{RL} \in \mathcal{G}^-(\mathbf{A})$.

Dôkaz. Overme definíciu zovšeobecnenej inverzie: $\mathbf{A}(\mathbf{RL})\mathbf{A} = \mathbf{BTRLBT} = \mathbf{BIIT} = \mathbf{A}$. □

Poznámka 6.7. Rozklad $\mathbf{A} = \mathbf{BT}$ vo vete 6.6 vždy existuje podľa vety 3.11.

Dôsledok 6.8 (Existencia zovšeobecnenej inverzie)**.** Pre každú $\mathbf{A} \in \mathbb{R}^{m\times n}$ existuje aspoň jedna zovšeobecnená inverzia $\mathbf{A}^-$.

Veta 6.9 (Množina zovšeobecnených inverzií je afinná)**.** Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$. Potom $\mathcal{G}^-(\mathbf{A})$ je afinná, čiže ak $\alpha_1, \ldots, \alpha_k$ sú reálne čísla, ktorých súčet je 1, a $\mathbf{G}_1, \ldots, \mathbf{G}_k$ sú zovšeobecnené inverzie matice $\mathbf{A}$, tak aj matica $\sum_{i=1}^k \alpha_i \mathbf{G}_i$ je zovšeobecnenou inverziou matice $\mathbf{A}$.

Dôkaz. Dôkaz je priamočiary, stačí overiť definíciu zovšeobecnenej inverzie pre $\sum_{i=1}^k \alpha_i \mathbf{G}_i$, detaily nechávame na čitateľa. □

Veta 6.10 (Symetrická zovšeobecnená inverzia symetrickej matice)**.** Nech $\mathbf{A} \in \mathbb{R}^{n\times n}$ je symetrická. Potom existuje symetrická $\mathbf{G}^* \in \mathcal{G}^-(\mathbf{A})$.

Dôkaz. Nech $\mathbf{G}$ je zovšeobecnená inverzia $\mathbf{A}$, teda $\mathbf{AGA} = \mathbf{A}$. Keďže $\mathbf{A}$ je symetrická, tak transponovaním tejto rovnice dostaneme $\mathbf{AG}^T\mathbf{A} = \mathbf{A}$. Položme $\mathbf{G}^* = (\mathbf{G} + \mathbf{G}^T)/2$. Potom $\mathbf{AG}^*\mathbf{A} = 2^{-1}(\mathbf{AGA} + \mathbf{AG}^T\mathbf{A}) = \mathbf{A}$, čiže $\mathbf{G}^*$ je zovšeobecnenou inverziou $\mathbf{A}$. Zároveň je $\mathbf{G}^*$ symetrická, keďže $(\mathbf{G}^*)_{ji} = (\mathbf{G}_{ji} + \mathbf{G}_{ij})/2 = (\mathbf{G}^*)_{ij}$ pre $i = 1, \ldots, n$, $j = 1, \ldots, n$. □

6.2 Vzťah zovšeobecnených inverzií so systémom lineárnych rovníc

Veta 6.11 (O vyjadrení jedného riešenia systému lineárnych rovníc pomocou zovšeobecnenej inverzie)**.** Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{G} \in \mathbb{R}^{n\times m}$ a $k \in \mathbb{N}$. Potom nasledujúce dve tvrdenia sú ekvivalentné:

(i) $\mathbf{GB}$ je riešením systému $\mathbf{AX} = \mathbf{B}$ (v premennej $\mathbf{X}$) pre každú maticu $\mathbf{B} \in \mathbb{R}^{m\times k}$, pre ktorú je daný systém konzistentný.

(ii) $\mathbf{G} \in \mathcal{G}^-(\mathbf{A})$.

Dôkaz. Nech platí (i). Zapíšme $\mathbf{A} = (\mathbf{a}_1, \ldots, \mathbf{a}_n)$. Nech $i \in \{1, \ldots, n\}$ a uvažujme systém $\mathbf{AX} = \mathbf{B}_i$, kde $\mathbf{B}_i = (\mathbf{a}_i, \mathbf{0}_m, \ldots, \mathbf{0}_m) \in \mathbb{R}^{m \times k}$. Ľahko sa možno presvedčiť, že $\mathbf{X} = (\mathbf{e}_i, \mathbf{0}_n, \ldots, \mathbf{0}_n) \in \mathbb{R}^{n \times k}$ je riešením tohto systému, a teda podľa (i) je $\mathbf{GB}_i$ tiež riešením. Takže $\mathbf{AG}(\mathbf{a}_i, \mathbf{0}_m, \ldots, \mathbf{0}_m) = (\mathbf{a}_i, \mathbf{0}_m, \ldots, \mathbf{0}_m)$, z čoho vyplýva $\mathbf{AGa}_i = \mathbf{a}_i$. Potom $\mathbf{AGa}_1 = \mathbf{a}_1, \ldots, \mathbf{AGa}_n = \mathbf{a}_n$, čo môžeme zapísať ako $\mathbf{AGA} = \mathbf{A}$, čiže $\mathbf{G}$ je zovšeobecnená inverzia matice $\mathbf{A}$.

Nech platí (ii) a nech $\mathbf{AX} = \mathbf{B}$ má nejaké riešenie $\mathbf{X}_0$. Potom $\mathbf{AGB} = \mathbf{AGAX}_0 = \mathbf{AX}_0 = \mathbf{B}$, a teda $\mathbf{GB}$ je riešením $\mathbf{AX} = \mathbf{B}$. □

Veta 6.12 (Vyjadrenie inklúzie stĺpcových priestorov pomocou zovšeobecnenej inverzie)**.** Nech $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{B} \in \mathbb{R}^{m \times k}$. Potom

(i) ak $\mathcal{C}(\mathbf{B}) \subseteq \mathcal{C}(\mathbf{A})$, tak $\mathbf{B} = \mathbf{AA}^-\mathbf{B}$ pre akúkoľvek $\mathbf{A}^- \in \mathcal{G}^-(\mathbf{A})$,

(ii) ak existuje $\mathbf{A}^- \in \mathcal{G}^-(\mathbf{A})$, ktorá spĺňa $\mathbf{B} = \mathbf{AA}^-\mathbf{B}$, potom $\mathcal{C}(\mathbf{B}) \subseteq \mathcal{C}(\mathbf{A})$.

Dôkaz. Nech $\mathcal{C}(\mathbf{B}) \subseteq \mathcal{C}(\mathbf{A})$. Potom podľa vety 3.4 existuje $\mathbf{X} \in \mathbb{R}^{n \times p}$, že $\mathbf{B} = \mathbf{AX}$, a teda $\mathbf{B} = \mathbf{AA}^-\mathbf{AX} = \mathbf{AA}^-\mathbf{B}$ pre akúkoľvek $\mathbf{A}^-$. Na dôkaz časti (ii) predpokladajme, že $\mathbf{B} = \mathbf{AA}^-\mathbf{B}$ pre nejakú $\mathbf{A}^-$. Potom $\mathbf{B} = \mathbf{AX}$ pre $\mathbf{X} = \mathbf{A}^-\mathbf{B}$, a teda veta 3.4 implikuje $\mathcal{C}(\mathbf{B}) \subseteq \mathcal{C}(\mathbf{A})$. □

Dôsledok 6.13. Nech $\mathbf{A} \in \mathbb{R}^{m \times n}$. Potom pre každé $\mathbf{x} \in \mathbb{R}^n$ platí

(i) ak $\mathbf{x} \in \mathcal{C}(\mathbf{A})$, tak $\mathbf{x} = \mathbf{AA}^-\mathbf{x}$ pre akúkoľvek $\mathbf{A}^- \in \mathcal{G}^-(\mathbf{A})$,

(ii) ak existuje $\mathbf{A}^- \in \mathcal{G}^-(\mathbf{A})$, ktorá spĺňa $\mathbf{x} = \mathbf{AA}^-\mathbf{x}$, potom $\mathbf{x} \in \mathcal{C}(\mathbf{A})$.

Dôkaz. Keď vo vete 6.12 zvolíme $\mathbf{B} = \mathbf{x}$, dostaneme požadované tvrdenie. □

Dôsledok 6.14. Nech $\mathbf{A} \in \mathbb{R}^{m \times n}$. Potom $\mathcal{C}(\mathbf{A}) = \mathcal{N}(\mathbf{I}_m - \mathbf{AA}^-)$ pre ľubovoľnú $\mathbf{A}^- \in \mathcal{G}^-(\mathbf{A})$.

Dôkaz. Pre ľubovoľnú $\mathbf{A}^-$ z dôsledku 6.13 plynie, že $\mathbf{x} \in \mathcal{C}(\mathbf{A})$ práve vtedy, keď $(\mathbf{I}_m - \mathbf{AA}^-)\mathbf{x} = \mathbf{0}_m$, čo dokazuje rovnosť $\mathcal{C}(\mathbf{A})$ a $\mathcal{N}(\mathbf{I}_m - \mathbf{AA}^-)$. □

Dôsledok 6.15 (Vyjadrenie konzistencie lineárneho systému pomocou zovšeobecnenej inverzie)**.** Nech $\mathbf{A} \in \mathbb{R}^{m \times n}$ a $\mathbf{B} \in \mathbb{R}^{m \times k}$. Potom systém $\mathbf{AX} = \mathbf{B}$ (v premennej $\mathbf{X}$) je konzistentný práve vtedy, keď $\mathbf{B} = \mathbf{AA}^-\mathbf{B}$ pre akúkoľvek $\mathbf{A}^- \in \mathcal{G}^-(\mathbf{A})$.

Veta 6.16 (O vzťahu medzi $\mathcal{C}(\mathbf{A})$ a $\mathcal{C}(\mathbf{AA}^-)$)**.** Nech $\mathbf{A} \in \mathbb{R}^{m \times n}$. Potom $\mathcal{C}(\mathbf{A}) = \mathcal{C}(\mathbf{AA}^-)$ pre ľubovoľnú $\mathbf{A}^- \in \mathcal{G}^-(\mathbf{A})$.

Dôkaz. Z vety 3.4 vyplýva $\mathcal{C}(\mathbf{AA}^-) \subseteq \mathcal{C}(\mathbf{A})$. Keďže $\mathbf{AA}^-\mathbf{A} = \mathbf{A}$, tak z tej istej vety vyplýva aj $\mathcal{C}(\mathbf{A}) \subseteq \mathcal{C}(\mathbf{AA}^-)$. □

Dôsledok 6.17. Nech $\mathbf{A} \in \mathbb{R}^{m \times n}$. Potom $\mathrm{rank}(\mathbf{A}) = \mathrm{rank}(\mathbf{AA}^-)$ pre ľubovoľnú $\mathbf{A}^- \in \mathcal{G}^-(\mathbf{A})$.

Veta 6.18 (Charakterizácia všetkých riešení konzistentného systému lineárnych rovníc)**.** Nech $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{B} \in \mathbb{R}^{m \times k}$. Matica $\mathbf{X}^* \in \mathbb{R}^{n \times k}$ je riešením konzistentného systému $\mathbf{AX} = \mathbf{B}$ vtedy a len vtedy, keď existujú $\mathbf{A}^- \in \mathcal{G}^-(\mathbf{A})$ a $\mathbf{Y} \in \mathbb{R}^{n \times k}$ také, že $\mathbf{X}^* = \mathbf{A}^-\mathbf{B} + (\mathbf{I}_n - \mathbf{A}^-\mathbf{A})\mathbf{Y}$.

Dôkaz. Nech $\mathbf{AX}^* = \mathbf{B}$. Potom $\mathbf{X}^*$ môžeme zapísať ako $\mathbf{X}^* = \mathbf{X}^* - \mathbf{A}^-(\mathbf{AX}^* - \mathbf{B}) = \mathbf{A}^-\mathbf{B} + (\mathbf{I} - \mathbf{A}^-\mathbf{A})\mathbf{X}^*$ pre ľubovoľnú $\mathbf{A}^-$. Naopak, nech existujú $\mathbf{A}^-$ a $\mathbf{Y} \in \mathbb{R}^{n\times k}$, že $\mathbf{X}^* = \mathbf{A}^-\mathbf{B} + (\mathbf{I} - \mathbf{A}^-\mathbf{A})\mathbf{Y}$. Potom $\mathbf{AX}^* = \mathbf{AA}^-\mathbf{B} + \mathbf{A}(\mathbf{I} - \mathbf{A}^-\mathbf{A})\mathbf{Y} = \mathbf{A}^-\mathbf{AB} + (\mathbf{A} - \mathbf{A})\mathbf{Y} = \mathbf{AA}^-\mathbf{B}$. Keďže systém je konzistentný, tak existuje $\mathbf{X}_0 \in \mathbb{R}^{n\times k}$, že $\mathbf{AX}_0 = \mathbf{B}$, čiže $\mathbf{AA}^-\mathbf{B} = \mathbf{AA}^-\mathbf{AX}_0 = \mathbf{AX}_0 = \mathbf{B}$, a teda $\mathbf{AX}^* = \mathbf{B}$. □

Dôsledok 6.19. Ak $\mathbf{A} \in \mathbb{R}^{m\times n}$, tak $\mathbf{X}^*$ je riešením $\mathbf{AX} = \mathbf{0}_{m\times k}$ práve vtedy, keď existujú $\mathbf{A}^- \in \mathcal{G}^-(\mathbf{A})$ a $\mathbf{Y} \in \mathbb{R}^{n\times k}$ také, že $\mathbf{X}^* = (\mathbf{I}_n - \mathbf{A}^-\mathbf{A})\mathbf{Y}$.

Dôsledok 6.20. Ak $\mathbf{A}$ je regulárna, tak $\mathbf{X}^*$ je riešením $\mathbf{AX} = \mathbf{B}$ vtedy a len vtedy, keď $\mathbf{X}^* = \mathbf{A}^{-1}\mathbf{B}$.

Príklad 6.21. Skúste sformulovať a dokázať tvrdenie podobné dôsledku 6.20 pre maticu $\mathbf{A}$ s plnou stĺpcovou hodnosťou.

6.3 Vyjadrenie zovšeobecnených inverzií

Veta 6.22 (Explicitné vyjadrenie všetkých zovšeobecnených inverzií matice $\mathbf{A}$)**.** Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$ má hodnosť $0 < r < \min\{m, n\}$. Nech $1 \le i_1 \le \ldots \le i_r \le m$ sú také indexy, že riadky $i_1, \ldots, i_r$ matice $\mathbf{A}$ sú lineárne nezávislé a nech $1 \le j_1 \le \ldots \le j_r \le n$ sú také indexy, že stĺpce $j_1, \ldots, j_r$ matice $\mathbf{A}$ sú lineárne nezávislé. Nech $\mathbf{P} \in \mathbb{R}^{m\times m}$ je akákoľvek permutačná matica, ktorej prvých r riadkov sú $1 \times m$ vektory $\mathbf{e}_{i_1}^T, \ldots, \mathbf{e}_{i_r}^T$. Nech $\mathbf{Q} \in \mathbb{R}^{n\times n}$ je akákoľvek permutačná matica, ktorej prvých r stĺpcov sú $\mathbf{e}_{j_1}, \ldots, \mathbf{e}_{j_r} \in \mathbb{R}^n$. Nech

$$\mathbf{B} = \mathbf{PAQ} = \begin{pmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{pmatrix},$$

kde $\mathbf{B}_{11} \in \mathbb{R}^{r\times r}$. Potom $\mathbf{B}_{11}$ je regulárna. Navyše platí $\mathbf{G} \in \mathcal{G}^-(\mathbf{A})$ vtedy a len vtedy, keď

$$\mathbf{G} = \mathbf{Q} \begin{pmatrix} \mathbf{B}_{11}^{-1} - \mathbf{XB}_{21}\mathbf{B}_{11}^{-1} - \mathbf{B}_{11}^{-1}\mathbf{B}_{12}\mathbf{Y} - \mathbf{B}_{11}^{-1}\mathbf{B}_{12}\mathbf{ZB}_{21}\mathbf{B}_{11}^{-1} & \mathbf{X} \\ \mathbf{Y} & \mathbf{Z} \end{pmatrix} \mathbf{P} \tag{6.1}$$

pre nejaké matice $\mathbf{X} \in \mathbb{R}^{r\times(m-r)}$, $\mathbf{Y} \in \mathbb{R}^{(n-r)\times r}$, $\mathbf{Z} \in \mathbb{R}^{(n-r)\times(m-r)}$.

Dôkaz. Blok $\mathbf{B}_{11}$ matice $\mathbf{B} = \mathbf{PAQ}$ je regulárna matica podľa lemy 3.21. Potom podľa lemy 4.25 platí $\mathbf{B}_{22} = \mathbf{B}_{21}\mathbf{B}_{11}^{-1}\mathbf{B}_{12}$, a preto môžeme $\mathbf{B}$ zapísať ako

$$\mathbf{B} = \begin{pmatrix} \mathbf{B}_{11} \\ \mathbf{B}_{21} \end{pmatrix} \mathbf{B}_{11}^{-1} \begin{pmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \end{pmatrix}.$$

Matica $\mathbf{H}$ je zovšeobecnenou inverziou $\mathbf{B}$ práve vtedy, keď $\mathbf{BHB} = \mathbf{B}$, čiže keď

$$\begin{pmatrix} \mathbf{B}_{11} \\ \mathbf{B}_{21} \end{pmatrix} \mathbf{B}_{11}^{-1} \begin{pmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \end{pmatrix} \mathbf{H} \begin{pmatrix} \mathbf{B}_{11} \\ \mathbf{B}_{21} \end{pmatrix} \mathbf{B}_{11}^{-1} \begin{pmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \end{pmatrix} = \begin{pmatrix} \mathbf{B}_{11} \\ \mathbf{B}_{21} \end{pmatrix} \mathbf{B}_{11}^{-1} \begin{pmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \end{pmatrix}.$$

Matice

$$\begin{pmatrix} \mathbf{B}_{11} \\ \mathbf{B}_{21} \end{pmatrix} \quad \text{a} \quad \begin{pmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \end{pmatrix}$$

majú podľa vety 3.20 hodnosť r, a teda rovnosť $\mathbf{BHB} = \mathbf{B}$ nastáva podľa vety 4.10 práve vtedy, keď

$$\mathbf{B}_{11}^{-1}\begin{pmatrix}\mathbf{B}_{11} & \mathbf{B}_{12}\end{pmatrix}\mathbf{H}\begin{pmatrix}\mathbf{B}_{11}\\ \mathbf{B}_{21}\end{pmatrix}\mathbf{B}_{11}^{-1} = \mathbf{B}_{11}^{-1}.$$

Zapíšme

$$\mathbf{H} = \begin{pmatrix}\mathbf{H}_{11} & \mathbf{H}_{12}\\ \mathbf{H}_{21} & \mathbf{H}_{22}\end{pmatrix},$$

kde $\mathbf{H}_{11} \in \mathbb{R}^{r\times r}$. Potom

$$\mathbf{B}_{11}^{-1}\begin{pmatrix}\mathbf{B}_{11} & \mathbf{B}_{12}\end{pmatrix}\mathbf{H}\begin{pmatrix}\mathbf{B}_{11}\\ \mathbf{B}_{21}\end{pmatrix}\mathbf{B}_{11}^{-1} = \mathbf{H}_{11} + \mathbf{B}_{11}^{-1}\mathbf{B}_{12}\mathbf{H}_{21} + \mathbf{H}_{12}\mathbf{B}_{21}\mathbf{B}_{11}^{-1} + \mathbf{B}_{11}^{-1}\mathbf{B}_{12}\mathbf{H}_{22}\mathbf{B}_{21}\mathbf{B}_{11}^{-1}.$$

Teda $\mathbf{BHB} = \mathbf{B}$ nastáva vtedy a len vtedy, keď

$$\mathbf{H}_{11} = \mathbf{B}_{11}^{-1} - \mathbf{B}_{11}^{-1}\mathbf{B}_{12}\mathbf{H}_{21} - \mathbf{H}_{12}\mathbf{B}_{21}\mathbf{B}_{11}^{-1} - \mathbf{B}_{11}^{-1}\mathbf{B}_{12}\mathbf{H}_{22}\mathbf{B}_{21}\mathbf{B}_{11}^{-1}. \tag{6.2}$$

Keďže $\mathbf{B} = \mathbf{PAQ}$ a každá permutačná matica je ortogonálna (čiže $\mathbf{P}^{-1} = \mathbf{P}^T$ a $\mathbf{Q}^{-1} = \mathbf{Q}^T$), tak $\mathbf{A} = \mathbf{P}^T\mathbf{B}\mathbf{Q}^T$. Potom $\mathbf{G}$ je zovšeobecnenou inverziou $\mathbf{A}$ práve vtedy, keď

$$\mathbf{P}^T\mathbf{B}\mathbf{Q}^T\mathbf{G}\mathbf{P}^T\mathbf{B}\mathbf{Q}^T = \mathbf{P}^T\mathbf{B}\mathbf{Q}^T,$$

čo je ekvivalentné s $\mathbf{B}\mathbf{Q}^T\mathbf{G}\mathbf{P}^T\mathbf{B} = \mathbf{B}$. Čiže $\mathbf{G} \in \mathcal{G}^-(\mathbf{A})$ práve vtedy, keď $\mathbf{Q}^T\mathbf{G}\mathbf{P}^T \in \mathcal{G}^-(\mathbf{B})$, čo z definície nastáva práve vtedy, keď $\mathbf{Q}^T\mathbf{G}\mathbf{P}^T = \mathbf{H}$ pre nejakú zovšeobecnenú inverziu $\mathbf{H}$ matice $\mathbf{B}$. To je ekvivalentné s $\mathbf{G} = \mathbf{QHP}$. Označme si matice $\mathbf{H}_{12} = \mathbf{X}$, $\mathbf{H}_{21} = \mathbf{Y}$, $\mathbf{H}_{22} = \mathbf{Z}$ a dosaďme tvar (6.2) do matice $\mathbf{H}$. Potom $\mathbf{G} \in \mathcal{G}^-(\mathbf{A})$ práve vtedy, keď je v tvare

$$\mathbf{G} = \mathbf{Q}\begin{pmatrix}\mathbf{B}_{11}^{-1} - \mathbf{X}\mathbf{B}_{21}\mathbf{B}_{11}^{-1} - \mathbf{B}_{11}^{-1}\mathbf{B}_{12}\mathbf{Y} - \mathbf{B}_{11}^{-1}\mathbf{B}_{12}\mathbf{Z}\mathbf{B}_{21}\mathbf{B}_{11}^{-1} & \mathbf{X}\\ \mathbf{Y} & \mathbf{Z}\end{pmatrix}\mathbf{P},$$

čo sme chceli dokázať. □

Príklad 6.23. Veta 6.22 neformálne platí aj keď $r = m$ alebo $r = n$, ak matice, ktorých niektorý rozmer je nulový, pokladáme za prázdne (neexistujúce). Sformulujte ako by táto veta znela korektne pre $r = m(\leq n)$ a pre $r = n(\leq m)$.

Poznámka 6.24. Špeciálne pre $\mathbf{X} = \mathbf{0}$, $\mathbf{Y} = \mathbf{0}$ a $\mathbf{Z} = \mathbf{0}$ dostávame vo vete 6.22, že

$$\mathbf{G} = \mathbf{Q}\begin{pmatrix}\mathbf{B}_{11}^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}\mathbf{P}$$

je zovšeobecnenou inverziou matice $\mathbf{A}$.

Príklad 6.25. Nech

$$\mathbf{A} = \begin{pmatrix}-6 & 2 & -3 & -2\\ 3 & -1 & 2 & 5\\ -3 & 1 & -1 & 3\end{pmatrix}.$$

Potom $\operatorname{rank}(\mathbf{A}) = 2$ a napríklad stĺpce 1 a 4 sú nezávislé a riadky 1 a 3 sú nezávislé. Maticu $\mathbf{B}$ môžeme získať pomocou

$$\mathbf{P} = \begin{pmatrix}1 & 0 & 0\\ 0 & 0 & 1\\ 0 & 1 & 0\end{pmatrix}, \quad \mathbf{Q} = \begin{pmatrix}1 & 0 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\\ 0 & 1 & 0 & 0\end{pmatrix},$$

čiže

$$\mathbf{B} = \mathbf{PAQ} = \begin{pmatrix} -6 & 2 & -3 & -2 \\ -3 & 1 & -1 & 3 \\ 3 & -1 & 2 & 5 \end{pmatrix} \mathbf{Q} = \begin{pmatrix} -6 & -2 & 2 & -3 \\ -3 & 3 & 1 & -1 \\ 3 & 5 & -1 & 2 \end{pmatrix}.$$

Potom

$$\mathbf{B}_{11} = \begin{pmatrix} -6 & -2 \\ -3 & 3 \end{pmatrix}$$

má plnú hodnosť a jej inverzia je

$$\mathbf{B}_{11}^{-1} = \begin{pmatrix} -3/24 & -2/24 \\ -3/24 & 6/24 \end{pmatrix}.$$

Zovšeobecnenú inverziu $\mathbf{G} \in \mathcal{G}^-(\mathbf{A})$ teda vieme získať ako

$$\mathbf{G} = \mathbf{Q} \begin{pmatrix} -3/24 & -2/24 & 0 \\ -3/24 & 6/24 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \mathbf{P} = \begin{pmatrix} -3/24 & 0 & -2/24 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ -3/24 & 0 & 6/24 \end{pmatrix}.$$

Nájdite inú zovšeobecnenú inverziu matice $\mathbf{A}$.

Veta 6.26 (Množina $\mathcal{G}^-(\mathbf{A})$ ako posunutý vektorový podpriestor)**.** Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$ má hodnosť r. Nech $\mathbf{P}, \mathbf{Q}$ sú ľubovoľné permutačné matice z vety 6.22 a $\mathbf{B} = \mathbf{PAQ}$ je v blokovom tvare ako vo vete 6.22. Potom $\mathcal{G}^-(\mathbf{A}) = \mathbf{G}^* + \mathcal{L}$, kde

$$\mathbf{G}^* = \mathbf{Q} \begin{pmatrix} \mathbf{B}_{11}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{P}$$

a

$$\mathcal{L} = \Big\{ \mathbf{Q} \begin{pmatrix} -\mathbf{X}\mathbf{B}_{21}\mathbf{B}_{11}^{-1} - \mathbf{B}_{11}^{-1}\mathbf{B}_{12}\mathbf{Y} - \mathbf{B}_{11}^{-1}\mathbf{B}_{12}\mathbf{Z}\mathbf{B}_{21}\mathbf{B}_{11}^{-1} & \mathbf{X} \\ \mathbf{Y} & \mathbf{Z} \end{pmatrix} \mathbf{P} \in \mathbb{R}^{n\times m} \,\Big|\, \mathbf{X} \in \mathbb{R}^{r\times(m-r)}, \mathbf{Y} \in \mathbb{R}^{(n-r)\times r}, \mathbf{Z} \in \mathbb{R}^{(n-r)\times(m-r)} \Big\}.$$

Navyše platí, že $\mathbf{G}^* \in \mathcal{G}^-(\mathbf{A})$ a $\mathcal{L}$ je vektorový podpriestor $\mathbb{R}^{n\times m}$ dimenzie $nm - r^2$.

Dôkaz. Podľa vety 6.22 platí $\mathbf{G} \in \mathcal{G}^-(\mathbf{A})$ práve vtedy, keď $\mathbf{G} = \mathbf{G}^* + \mathbf{C}$, pre nejakú $\mathbf{C} \in \mathcal{L}$, čím je prvá časť vety dokázaná. Matica $\mathbf{G}^*$ je zovšeobecnenou inverziou matice $\mathbf{A}$ podľa poznámky 6.24. Dôkaz toho, že $\mathcal{L}$ je vektorový podpriestor, necháme na čitateľa (viď úloha 6.6). Dimenzia $\mathcal{L}$ je rovná počtu voľných parametrov pri zápise $\mathcal{L}$, teda celkovému počtu prvkov matíc $\mathbf{X}$ $\mathbf{Y}$ a $\mathbf{Z}$. Teda

$$\dim(\mathcal{L}) = r(m-r) + (n-r)r + (n-r)(m-r) = nm - r^2.$$

□

Poznámka 6.27. Pod zápisom $\mathcal{G}^-(\mathbf{A}) = \mathbf{G}^* + \mathcal{L}$ vo vete 6.26 rozumieme, že

$$\mathcal{G}^-(\mathbf{A}) = \{\mathbf{G} \in \mathbb{R}^{n\times m} \,|\, \text{existuje } \mathbf{C} \in \mathcal{L} \text{ také, že } \mathbf{G} = \mathbf{G}^* + \mathbf{C}\}.$$

Spomínaná veta nám hovorí, že $\mathcal{G}^-(\mathbf{A})$ (čo je podľa vety 6.9 afinná množina) vieme zapísať ako vektorový podpriestor dimenzie $nm - r^2$ v nm-rozmernom priestore $\mathbb{R}^{n\times m}$ posunutý pomocou $\mathbf{G}^*$ tak, že neprechádza maticou $\mathbf{0}_{n\times m}$. Množina $\mathcal{G}^-(\mathbf{A})$ teda predstavuje $(nm - r^2)$-rozmernú posunutú "rovinu" v $\mathbb{R}^{n\times m}$. Preto $\mathcal{G}^-(\mathbf{A})$ môžeme schematicky kresliť ako rovinu posunutú z počiatku súradnicovej sústavy (ako na obrázku 6.1b).

Dôsledok 6.28 (Početnosť zovšeobecnených inverzií)**.** Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$. Potom ak $\mathbf{A}$ je štvorcová a regulárna, tak má práve jednu zovšeobecnenú inverziu. Inak má $\mathbf{A}$ nekonečne veľa zovšeobecnených inverzií.

Veta 6.29 (Vyjadrenie všetkých zovšeobecnených inverzií pomocou jednej zovšeobecnenej inverzie)**.** Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$ a nech $\mathbf{G}^* \in \mathcal{G}^-(\mathbf{A})$. Potom $\mathbf{G} \in \mathcal{G}^-(\mathbf{A})$ práve vtedy, keď existuje $\mathbf{Z} \in \mathbb{R}^{n\times m}$, že $\mathbf{G} = \mathbf{G}^* + \mathbf{Z} - \mathbf{G}^*\mathbf{AZAG}^*$.

Dôkaz. Nech $\mathbf{G} \in \mathcal{G}^-(\mathbf{A})$. Potom sa ľahko overí, že platí $\mathbf{G} = \mathbf{G}^* + (\mathbf{G} - \mathbf{G}^*) - \mathbf{G}^*\mathbf{A}(\mathbf{G} - \mathbf{G}^*)\mathbf{AG}^*$. Stačí teda položiť $\mathbf{Z} = \mathbf{G} - \mathbf{G}^*$. Naopak predpokladajme, že $\mathbf{G} = \mathbf{G}^* + \mathbf{Z} - \mathbf{G}^*\mathbf{AZAG}^*$ pre nejakú maticu $\mathbf{Z} \in \mathbb{R}^{n\times m}$. Overme definíciu zovšeobecnenej inverzie:

$$\mathbf{AGA} = \mathbf{AG}^*\mathbf{A} + \mathbf{AZA} - \mathbf{AG}^*\mathbf{AZAG}^*\mathbf{A} = \mathbf{A} + \mathbf{AZA} - \mathbf{AZA} = \mathbf{A}.$$

□

Príklad 6.30. Nájdime všetky zovšeobecnené inverzie matice

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 5 \\ -4 & -7 \end{pmatrix}.$$

Pomocou poznámky 6.24 sa dá získať zovšeobecnená inverzia

$$\mathbf{G}^* = \begin{pmatrix} -5 & 2 & 0 \\ 3 & -1 & 0 \end{pmatrix}.$$

Overte, že $\mathbf{AG}^*\mathbf{A} = \mathbf{A}$. Podľa vety 6.29 sú všetky zovšeobecnené inverzie $\mathbf{G} \in \mathbf{G}^-(\mathbf{A})$ dané ako $\mathbf{G} = \mathbf{G}^* + \mathbf{Z} - \mathbf{G}^*\mathbf{AZAG}^*$. Označme

$$\mathbf{Z} = \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}.$$

Potom

$$\mathbf{G}^*\mathbf{AZAG}^* = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & -1 & 0 \end{pmatrix} = \begin{pmatrix} a-c & b-c & 0 \\ d-f & e-f & 0 \end{pmatrix},$$

čiže

$$\mathbf{G} = \mathbf{G}^* + \mathbf{Z} - \mathbf{G}^*\mathbf{AZAG}^* = \begin{pmatrix} -5 & 2 & 0 \\ 3 & -1 & 0 \end{pmatrix} + \begin{pmatrix} c & c & c \\ f & f & f \end{pmatrix} = \begin{pmatrix} -5+c & 2+c & c \\ 3+f & -1+f & f \end{pmatrix}.$$

Teda

$$\mathcal{G}^-(\mathbf{A}) = \left\{ \mathbf{G} \in \mathbb{R}^{2\times 3} \;\middle|\; \mathbf{G} = \begin{pmatrix} -5+c & 2+c & c \\ 3+f & -1+f & f \end{pmatrix}, c, f \in \mathbb{R} \right\}.$$

Poznamenajme, že ten istý predpis $\mathcal{G}^-(\mathbf{A})$ sme mohli získať aplikovaním vety 6.22. Všimnite si, že $\mathcal{G}^-(\mathbf{A})$ je afinná množina, ako vieme z vety 6.9. Nájdite množinu $\mathcal{G}^-(\mathbf{A})$ aj pomocou vety 6.22.

6.4 Mooreova-Penroseova pseudoinverzia

Definícia 6.31 (Mooreova-Penroseova pseudoinverzia)**.** Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$. Potom maticu $\mathbf{A}^+ \in \mathbb{R}^{n\times m}$, ktorá spĺňa vlastnosti

1. $\mathbf{A}\mathbf{A}^+\mathbf{A} = \mathbf{A}$ ($\mathbf{A}^+$ je zovšeobecnenou inverziou $\mathbf{A}$),
2. $\mathbf{A}^+\mathbf{A}\mathbf{A}^+ = \mathbf{A}^+$ ($\mathbf{A}$ je zovšeobecnenou inverziou $\mathbf{A}^+$),
3. $(\mathbf{A}\mathbf{A}^+)^T = \mathbf{A}\mathbf{A}^+$ ($\mathbf{A}\mathbf{A}^+$ je symetrická),
4. $(\mathbf{A}^+\mathbf{A})^T = \mathbf{A}^+\mathbf{A}$ ($\mathbf{A}^+\mathbf{A}$ je symetrická),

nazývame *Mooreova-Penroseova pseudoinverzia* matice $\mathbf{A}$.

Poznámka 6.32. Niekedy sa Mooreova-Penroseova pseudoinverzia nazýva skrátene iba ako *pseudoinverzia.*

Veta 6.33 (Existencia a jedinečnosť pseudoinverzie)**.** Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$. Potom existuje jediná Mooreova-Penroseova pseudoinverzia $\mathbf{A}^+$. Ak $\mathbf{A} = \mathbf{0}_{m\times n}$, tak $\mathbf{A}^+ = \mathbf{0}_{n\times m}$. Ak $r = \text{rank}(\mathbf{A}) > 0$, nech $\mathbf{A} = \mathbf{B}\mathbf{T}$, kde $\mathbf{B} \in \mathbb{R}^{m\times r}$, $\mathbf{T} \in \mathbb{R}^{r\times n}$ a $\text{rank}(\mathbf{B}) = \text{rank}(\mathbf{T}) = r$. Potom $\mathbf{B}^T\mathbf{A}\mathbf{T}^T$ je regulárna a $\mathbf{A}^+ = \mathbf{T}^T(\mathbf{B}^T\mathbf{A}\mathbf{T}^T)^{-1}\mathbf{B}^T$.

Dôkaz. Pre $\mathbf{A} = \mathbf{0}_{m\times n}$ sú vlastnosti 1-4 z definície 6.31 zjavne splnené pre $\mathbf{A}^+ = \mathbf{0}_{n\times m}$. Navyše, 2. vlastnosť je splnená iba pre $\mathbf{A}^+ = \mathbf{0}$, lebo ľavá strana tejto rovnosti je pre $\mathbf{A} = \mathbf{0}$ vždy nulová. Nech $\mathbf{A} \neq \mathbf{0}$, potom existuje rozklad $\mathbf{A} = \mathbf{B}\mathbf{T}$ podľa vety 3.11. Všimnime si, že $\mathbf{B}^T\mathbf{B}$ a $\mathbf{T}\mathbf{T}^T$ sú $r \times r$ matice hodnosti r (na základe dôsledku 3.35), čiže sú regulárne. Potom hodnosť $r \times r$ matice $\mathbf{B}^T\mathbf{A}\mathbf{T}^T = \mathbf{B}^T\mathbf{B}\mathbf{T}\mathbf{T}^T$ je podľa vety 4.18 rovná hodnosti matice $\mathbf{B}^T\mathbf{B}$, čo je r. Teda $\mathbf{B}^T\mathbf{A}\mathbf{T}^T$ je regulárna. Položme $\mathbf{G}^* = \mathbf{T}^T(\mathbf{B}^T\mathbf{A}\mathbf{T}^T)^{-1}\mathbf{B}^T = \mathbf{T}^T(\mathbf{B}^T\mathbf{B}\mathbf{T}\mathbf{T}^T)^{-1}\mathbf{B}^T$. Keďže $\mathbf{T}\mathbf{T}^T$ a $\mathbf{B}^T\mathbf{B}$ sú regulárne, tak $\mathbf{G}^*$ môžeme zapísať ako $\mathbf{G}^* = \mathbf{T}^T(\mathbf{T}\mathbf{T}^T)^{-1}(\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T$. Potom

$$\mathbf{A}\mathbf{G}^*\mathbf{A} = \mathbf{B}\mathbf{T}\mathbf{T}^T(\mathbf{T}\mathbf{T}^T)^{-1}(\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T\mathbf{B}\mathbf{T} = \mathbf{B}\mathbf{T} = \mathbf{A}$$

a

$$\begin{aligned}\mathbf{G}^*\mathbf{A}\mathbf{G}^* &= \mathbf{T}^T(\mathbf{T}\mathbf{T}^T)^{-1}(\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T\mathbf{B}\mathbf{T}\mathbf{T}^T(\mathbf{T}\mathbf{T}^T)^{-1}(\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T \\ &= \mathbf{T}^T(\mathbf{T}\mathbf{T}^T)^{-1}(\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T = \mathbf{G}^*.\end{aligned}$$

Navyše

$$\mathbf{A}\mathbf{G}^* = \mathbf{B}\mathbf{T}\mathbf{T}^T(\mathbf{T}\mathbf{T}^T)^{-1}(\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T = \mathbf{B}(\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T$$

a

$$\mathbf{G}^*\mathbf{A} = \mathbf{T}^T(\mathbf{T}\mathbf{T}^T)^{-1}(\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T\mathbf{B}\mathbf{T} = \mathbf{T}^T(\mathbf{T}\mathbf{T}^T)^{-1}\mathbf{T}.$$

Podľa dôsledku 4.16 sú matice $(\mathbf{B}^T\mathbf{B})^{-1}$ a $(\mathbf{T}\mathbf{T}^T)^{-1}$ symetrické, čiže aj $\mathbf{A}\mathbf{G}^*$ a $\mathbf{G}^*\mathbf{A}$ sú symetrické. Teda $\mathbf{G}^*$ spĺňa vlastnosti 1-4, čím sme dokázali existenciu $\mathbf{A}^+$.

Nech $\mathbf{G}$ je ľubovoľná matica spĺňajúca vlastnosti 1-4. Potom $\mathbf{G} = \mathbf{G}\mathbf{A}\mathbf{G}$. Pred ďalšími úpravami najprv preskúmajme výrazy $\mathbf{G}\mathbf{A}$ a $\mathbf{A}\mathbf{G}$:

$$\mathbf{G}\mathbf{A} = (\mathbf{G}\mathbf{A})^T = (\mathbf{G}\mathbf{A}\mathbf{G}^*\mathbf{A})^T = (\mathbf{G}^*\mathbf{A})^T(\mathbf{G}\mathbf{A})^T = \mathbf{G}^*\mathbf{A}\mathbf{G}\mathbf{A} = \mathbf{G}^*\mathbf{A}$$

a podobne

$$\mathbf{A}\mathbf{G} = (\mathbf{A}\mathbf{G})^T = (\mathbf{A}\mathbf{G}^*\mathbf{A}\mathbf{G})^T = (\mathbf{A}\mathbf{G})^T(\mathbf{A}\mathbf{G}^*)^T = \mathbf{A}\mathbf{G}\mathbf{A}\mathbf{G}^* = \mathbf{A}\mathbf{G}^*.$$

Teda
$$\mathbf{G} = \mathbf{GAG} = \mathbf{G}^*\mathbf{AG} = \mathbf{G}^*\mathbf{AG}^* = \mathbf{G}^*,$$
čím je dokázaná jedinečnosť matice $\mathbf{A}^+$. □

Veta 6.34 (Základné vlastnosti pseudoinverzie)**.** Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$. Potom:

(i) Nech $c \neq 0$. Potom $(c\mathbf{A})^+ = c^{-1}\mathbf{A}^+$.

(ii) $(\mathbf{A}^T)^+ = (\mathbf{A}^+)^T$.

(iii) $(\mathbf{A}^+)^+ = \mathbf{A}$.

(iv) Ak $m = n$ a $\mathbf{A}$ je regulárna, potom $\mathbf{A}^+ = \mathbf{A}^{-1}$.

(v) Ak $m = n$ a $\mathbf{A} = \mathrm{diag}(d_1, \ldots, d_n)$, tak $\mathbf{A}^+ = \mathrm{diag}(d_1^+, \ldots, d_n^+)$, kde $d_i^+ = 1/d_i$ pre $d_i \neq 0$ a $d_i^+ = 0$ pre $d_i = 0$, $i = 1, \ldots, n$.

(vi) Ak $\mathrm{rank}(\mathbf{A}) = m$, tak $\mathbf{A}^+ = \mathbf{A}^T(\mathbf{AA}^T)^{-1}$.

(vii) Ak $\mathrm{rank}(\mathbf{A}) = n$, tak $\mathbf{A}^+ = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T$.

Dôkaz. Dôkazy (i)-(v) sú priamočiare, stačí overiť vlastnosti pseudoinverzie.

Pre dôkaz (vi), nech $\mathrm{rank}(\mathbf{A}) = m$. Potom matica $\mathbf{B}$ z $\mathbf{BT}$ rozkladu $\mathbf{A}$ je regulárna, a teda podľa vety 6.33 máme $\mathbf{A}^+ = \mathbf{T}^T(\mathbf{B}^T\mathbf{AT}^T)^{-1}\mathbf{B}^T = \mathbf{T}^T(\mathbf{AT}^T)^{-1}(\mathbf{B}^T)^{-1}\mathbf{B}^T = \mathbf{T}^T(\mathbf{AT}^T)^{-1}$. Zároveň $\mathbf{A}^T(\mathbf{AA}^T)^{-1} = \mathbf{T}^T\mathbf{B}^T(\mathbf{AT}^T\mathbf{B}^T)^{-1} = \mathbf{T}^T\mathbf{B}^T(\mathbf{B}^T)^{-1}(\mathbf{AT}^T)^{-1} = \mathbf{T}^T(\mathbf{AT}^T)^{-1}$, čím sme ukázali, že $\mathbf{A}^+ = \mathbf{A}^T(\mathbf{AA}^T)^{-1}$. Dôkaz (vii) je analogický. □

Lema 6.35 (Pseudoinverzia po ortogonálnej transformácii)**.** Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$ a nech stĺpce matice $\mathbf{U} \in \mathbb{R}^{p\times m}$ sú ortonormálne a stĺpce matice $\mathbf{V} \in \mathbb{R}^{q\times n}$ sú ortonormálne. Teda $\mathbf{U}^T\mathbf{U} = \mathbf{I}_m$ a $\mathbf{V}^T\mathbf{V} = \mathbf{I}_n$. Potom $(\mathbf{UAV}^T)^+ = \mathbf{VA}^+\mathbf{U}^T$.

Dôkaz. Overme definičné vlastnosti pseudoinverzie:
1. $\mathbf{UAV}^T(\mathbf{VA}^+\mathbf{U}^T)\mathbf{UAV}^T = \mathbf{UAA}^+\mathbf{AV}^T = \mathbf{UAV}^T$.
3. $(\mathbf{UAV}^T(\mathbf{VA}^+\mathbf{U}^T))^T = (\mathbf{UAA}^+\mathbf{U}^T)^T = \mathbf{U}(\mathbf{AA}^+)\mathbf{U}^T = \mathbf{UAV}^T\mathbf{VA}^+\mathbf{U}^T$.
Pri oboch odvodeniach sme využili príslušné vlastnosti (1., 3.) pseudoinverzie. Zvyšné dve vlastnosti majú analogické dôkazy. □

Lema 6.36 (Pseudoinverzia $\mathbf{A}$ pomocou $\mathbf{A}^T\mathbf{A}$)**.** Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$. Potom matica $(\mathbf{A}^T\mathbf{A})^+\mathbf{A}^T$ je pseudoinverziou matice $\mathbf{A}$.

Dôkaz. Označme $\mathbf{G} = (\mathbf{A}^T\mathbf{A})^+\mathbf{A}^T$. Overíme, že $\mathbf{G}$ spĺňa definičné vlastnosti pseudoinverzie.

1. $\mathbf{AGA} = \mathbf{A}(\mathbf{A}^T\mathbf{A})^+\mathbf{A}^T\mathbf{A}$. Keďže $\mathcal{C}(\mathbf{A}^T) = \mathcal{C}(\mathbf{A}^T\mathbf{A})$ (dôsledok 3.35), tak podľa základnej vety o inklúzii stĺpcových priestorov existuje matica $\mathbf{Z}$, že $\mathbf{A}^T = \mathbf{A}^T\mathbf{AZ}$. Transponovaním tejto rovnice získavame $\mathbf{A} = \mathbf{Z}^T\mathbf{A}^T\mathbf{A}$ a po dosadení tak dostávame
$$\mathbf{AGA} = \mathbf{Z}^T\mathbf{A}^T\mathbf{A}(\mathbf{A}^T\mathbf{A})^+\mathbf{A}^T\mathbf{A} = \mathbf{Z}^T\mathbf{A}^T\mathbf{A} = \mathbf{A},$$
pričom sme využili vlastnosť 1 pseudoinverzie $(\mathbf{A}^T\mathbf{A})^+$.

2. $\mathbf{GAG} = (\mathbf{A}^T\mathbf{A})^+\mathbf{A}^T\mathbf{A}(\mathbf{A}^T\mathbf{A})^+\mathbf{A}^T = (\mathbf{A}^T\mathbf{A})^+\mathbf{A}^T = \mathbf{G}$, využijúc vlastnosť 2 matice $(\mathbf{A}^T\mathbf{A})^+$.

3. $(\mathbf{AG})^T = (\mathbf{A}(\mathbf{A}^T\mathbf{A})^+\mathbf{A}^T)^T = \mathbf{A}(\mathbf{A}^T\mathbf{A})^+\mathbf{A}^T$, pričom sme použili vlastnosť, že symetrická matica $\mathbf{A}^T\mathbf{A}$ má symetrickú pseudoinverziu (dôsledok vety 6.34(ii)).

4. $(\mathbf{GA})^T = ((\mathbf{A}^T\mathbf{A})^+\mathbf{A}^T\mathbf{A})^T = (\mathbf{A}^T\mathbf{A})^+\mathbf{A}^T\mathbf{A}$ priamo z vlastnosti 4 matice $(\mathbf{A}^T\mathbf{A})^+$. □

Poznámka 6.37. Upozornime, že výraz $(\mathbf{A}^T\mathbf{A})^-\mathbf{A}^T$ závisí na voľbe $(\mathbf{A}^T\mathbf{A})^-$, na rozdiel od výrazu $\mathbf{A}(\mathbf{A}^T\mathbf{A})^-\mathbf{A}^T$, ktorý na voľbe zovšeobecnenej inverzie nezávisí. Matica $(\mathbf{A}^T\mathbf{A})^-\mathbf{A}^T$ pre ľubovoľnú $(\mathbf{A}^T\mathbf{A})^-$ teda nemusí byť pseudoinverziou matice $\mathbf{A}$.

Veta 6.38 (Riešenie systému rovníc s najmenšou normou pomocou pseudoinverzie)**.** Nech $\mathbf{AX} = \mathbf{B}$ je konzistentný systém rovníc v premennej $\mathbf{X} \in \mathbb{R}^{n\times k}$, kde $\mathbf{A} \in \mathbb{R}^{m\times n}$ a $\mathbf{B} \in \mathbb{R}^{m\times k}$. Potom $\mathbf{X}_+ = \mathbf{A}^+\mathbf{B}$ je riešenie tohto systému, a pre každé riešenie $\mathbf{X}^*$ systému $\mathbf{AX} = \mathbf{B}$ platí $\|\mathbf{X}^*\| \geq \|\mathbf{X}_+\|$. Teda $\mathbf{A}^+\mathbf{B}$ minimalizuje Frobeniovu normu $\|\mathbf{X}\|$ na množine všetkých riešení systému $\mathbf{AX} = \mathbf{B}$.

Dôkaz. Nech $\mathbf{X} \in \mathbb{R}^{n\times k}$ je riešením systému $\mathbf{AX} = \mathbf{B}$. Keďže $\mathbf{A}^+$ je zovšeobecnenou inverziou $\mathbf{A}$ (vlastnosť 1.), tak podľa vety 6.11 je $\mathbf{A}^+\mathbf{B}$ riešením $\mathbf{AX} = \mathbf{B}$. Z vety 6.18 plynie, že akékoľvek riešenie systému $\mathbf{AX} = \mathbf{B}$ možno vyjadriť v tvare $\mathbf{X} = \mathbf{A}^+\mathbf{B} + (\mathbf{I} - \mathbf{A}^+\mathbf{A})\mathbf{Y}$ pre nejaké $\mathbf{Y} \in \mathbb{R}^{n\times k}$. Z toho vyplýva, že

$$\begin{aligned}\|\mathbf{X}\|^2 &= \mathrm{tr}\big((\mathbf{Y}^T(\mathbf{I} - (\mathbf{A}^+\mathbf{A})^T) + \mathbf{B}^T(\mathbf{A}^+)^T)(\mathbf{A}^+\mathbf{B} + (\mathbf{I} - \mathbf{A}^+\mathbf{A})\mathbf{Y})\big)\\ &= \|\mathbf{A}^+\mathbf{B}\|^2 + \|(\mathbf{I} - \mathbf{A}^+\mathbf{A})\mathbf{Y}\|^2 + 2\mathrm{tr}(\mathbf{B}^T(\mathbf{A}^+)^T(\mathbf{I} - \mathbf{A}^+\mathbf{A})\mathbf{Y}).\end{aligned}$$

Navyše, $(\mathbf{A}^+\mathbf{A})^2 = \mathbf{A}^+\mathbf{A}\mathbf{A}^+\mathbf{A} = \mathbf{A}^+\mathbf{A}$ (porovnajte s vetou 7.5 z nasledujúcej kapitoly). Využijúc $\mathbf{B} = \mathbf{AX}$ a $(\mathbf{A}^+\mathbf{A})^T = \mathbf{A}^+\mathbf{A}$, získavame

$$\begin{aligned}\mathbf{B}^T(\mathbf{A}^+)^T(\mathbf{I} - \mathbf{A}^+\mathbf{A})\mathbf{Y} &= \mathbf{X}^T\mathbf{A}^T(\mathbf{A}^+)^T(\mathbf{I} - \mathbf{A}^+\mathbf{A})\mathbf{Y} = \mathbf{X}^T(\mathbf{A}^+\mathbf{A})^T(\mathbf{I} - \mathbf{A}^+\mathbf{A})\mathbf{Y}\\ &= \mathbf{X}^T\mathbf{A}^+\mathbf{A}(\mathbf{I} - \mathbf{A}^+\mathbf{A})\mathbf{Y} = \mathbf{X}^T(\mathbf{A}^+\mathbf{A} - (\mathbf{A}^+\mathbf{A})^2)\mathbf{Y}\\ &= \mathbf{X}^T(\mathbf{A}^+\mathbf{A} - \mathbf{A}^+\mathbf{A})\mathbf{Y} = \mathbf{0}_{k\times k}.\end{aligned}$$

Teda $\|\mathbf{X}\|^2 = \|\mathbf{A}^+\mathbf{B}\|^2 + \|(\mathbf{I} - \mathbf{A}^+\mathbf{A})\mathbf{Y}\|^2 \geq \|\mathbf{A}^+\mathbf{B}\|^2$. □

Dôsledok 6.39. Pre $k = 1$ vo vete 6.38 dostávame, že $\mathbf{x}^* = \mathbf{A}^+\mathbf{b}$ je riešenie s najmenšou euklidovskou normou konzistentného systému $\mathbf{Ax} = \mathbf{b}$.

Veta 6.40 (Zovšeobecnená inverzia s najmenšou normou pomocou pseudoinverzie)**.** Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$. Potom $\|\mathbf{A}^-\| \geq \|\mathbf{A}^+\|$ pre každú zovšeobecnenú inverziu $\mathbf{A}^-$ matice $\mathbf{A}$. Pseudoinverzia $\mathbf{A}^+$ je teda pre $\mathbf{A}$ zovšeobecnenou inverziou s najmenšou Frobeniovou normou.

Dôkaz. Nech $\mathbf{G}$ je zovšeobecnenou inverziou matice $\mathbf{A}$. Podľa vety 6.29 vieme $\mathbf{G}$ vyjadriť ako $\mathbf{G} = \mathbf{A}^+ + \mathbf{Z} - \mathbf{A}^+\mathbf{AZAA}^+$. Potom

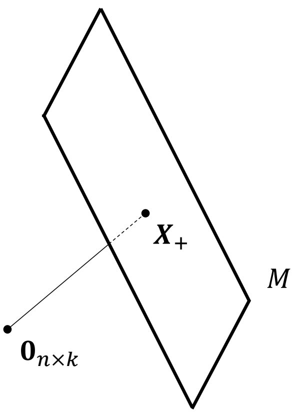
$$\begin{aligned}\|\mathbf{G}\|^2 &= \mathrm{tr}(\mathbf{G}^T\mathbf{G}) = \mathrm{tr}\Big(((\mathbf{A}^+)^T + (\mathbf{Z} - \mathbf{A}^+\mathbf{AZAA}^+)^T)(\mathbf{A}^+ + \mathbf{Z} - \mathbf{A}^+\mathbf{AZAA}^+)\Big)\\ &= \|\mathbf{A}^+\|^2 + \|\mathbf{Z} - \mathbf{A}^+\mathbf{AZAA}^+\|^2 + 2\mathrm{tr}((\mathbf{Z} - \mathbf{A}^+\mathbf{AZAA}^+)^T\mathbf{A}^+).\end{aligned}$$

Zároveň platí

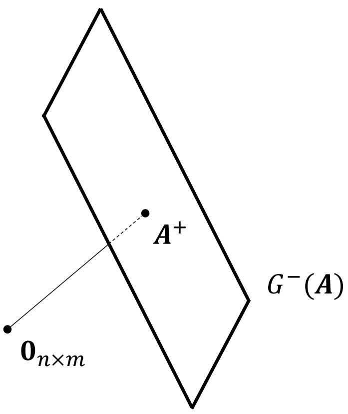
$$\begin{aligned}\mathrm{tr}((\mathbf{Z} - \mathbf{A}^+\mathbf{AZAA}^+)^T\mathbf{A}^+) &= \mathrm{tr}(\mathbf{Z}^T\mathbf{A}^+) - \mathrm{tr}((\mathbf{AA}^+)^T\mathbf{Z}^T(\mathbf{A}^+\mathbf{A})^T\mathbf{A}^+)\\ &= \mathrm{tr}(\mathbf{Z}^T\mathbf{A}^+) - \mathrm{tr}(\mathbf{AA}^+\mathbf{Z}^T\mathbf{A}^+\mathbf{AA}^+)\\ &= \mathrm{tr}(\mathbf{Z}^T\mathbf{A}^+) - \mathrm{tr}(\mathbf{AA}^+\mathbf{Z}^T\mathbf{A}^+)\\ &= \mathrm{tr}(\mathbf{Z}^T\mathbf{A}^+) - \mathrm{tr}(\mathbf{Z}^T\mathbf{A}^+\mathbf{AA}^+)\\ &= \mathrm{tr}(\mathbf{Z}^T\mathbf{A}^+) - \mathrm{tr}(\mathbf{Z}^T\mathbf{A}^+) = 0,\end{aligned}$$

pričom sme využili vlastnosti 2, 3 a 4 Mooreovej-Penroseovej pseudoinverzie a vetu 5.12. Z toho vyplýva $\|\mathbf{G}\|^2 = \|\mathbf{A}^+\|^2 + \|\mathbf{Z} - \mathbf{A}^+\mathbf{AZAA}^+\|^2 \geq \|\mathbf{A}^+\|^2$. □

Na obrázku 6.1a je symbolicky znázornené, že $\mathbf{X}_+ = \mathbf{A}^+\mathbf{B}$ má najmenšiu normu spomedzi všetkých riešení konzistentného systému $\mathbf{AX} = \mathbf{B}$. Podobne obrázok 6.1b reprezentuje to, že $\mathbf{A}^+$ má najmenšiu normu spomedzi všetkých zovšeobecnených inverzií $\mathbf{A}^-$.



(a) $\mathbf{X}_+ = \mathbf{A}^+\mathbf{B}$ má najmenšiu normu v afinnej množine $M = \{\mathbf{X} \mid \mathbf{AX} = \mathbf{B}\}$.



(b) $\mathbf{A}^+$ má najmenšiu normu v afinnej množine $\mathcal{G}^-(\mathbf{A})$.

Obr. 6.1: Mooreova-Penroseova pseudoinverzia minimalizujúca normu.

6.5 Úlohy na precvičenie

Úloha 6.1. Nech $\mathbf{A} \in \mathbb{R}^{n\times n}$ je antisymetrická matica (t.j. spĺňa $\mathbf{A}^T = -\mathbf{A}$). Ukážte, že potom ak $\mathbf{G} \in \mathcal{G}^-(\mathbf{A})$, tak aj $-\mathbf{G}^T \in \mathcal{G}^-(\mathbf{A})$. Na základe tohto tvrdenia dokážte, že pre každú antisymetrickú maticu existuje jej antisymetrická zovšeobecnená inverzia.

Úloha 6.2. Nech $\mathbf{0} \neq \mathbf{A} \in \mathbb{R}^{m\times n}$, $\operatorname{rank}(\mathbf{A}) = r$ a nech $\mathbf{B} \in \mathbb{R}^{m\times r}$, $\mathbf{T} \in \mathbb{R}^{r\times n}$ sú také matice, že $\mathbf{A} = \mathbf{BT}$. Nech $\mathbf{R} \in \mathbb{R}^{n\times r}$ je pravá inverzia matice $\mathbf{T}$. Zdôvodnite, prečo je $\mathbf{B}^T\mathbf{B}$ regulárna matica. Ukážte, že $\mathbf{R}(\mathbf{B}^T\mathbf{B})^{-1}\mathbf{R}^T$ je zovšeobecnená inverzia matice $\mathbf{A}^T\mathbf{A}$.

Úloha 6.3. Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{B} \in \mathbb{R}^{m\times p}$ a nech existuje matica $\mathbf{H} \in \mathbb{R}^{n\times m}$ taká, že $\mathbf{AHB} = \mathbf{B}$. Dokážte, že potom $\mathbf{AGB} = \mathbf{B}$ pre akúkoľvek zovšeobecnenú inverziu $\mathbf{G}$ matice $\mathbf{A}$.

Úloha 6.4. Nájdite aspoň dve rôzne zovšeobecnené inverzie matice

$$\mathbf{A} = \begin{pmatrix} 3 & 2 & 6 \\ -1 & 5 & -2 \\ 4 & 3 & 8 \\ 0 & 1 & 0 \\ -2 & -1 & -4 \end{pmatrix}$$

Úloha 6.5. Opíšte množinu všetkých zovšeobecnených inverzií $\mathcal{G}^-(\mathbf{A})$ matice

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 3 \\ -1 & 1 & 1 \end{pmatrix}$$

Úloha 6.6. Dokážte, že množina $\mathcal{L}$ z vety 6.26 je vektorový podpriestor.

Úloha 6.7. Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{B} \in \mathbb{R}^{p\times n}$, $\mathbf{C} \in \mathbb{R}^{m\times q}$, $\mathcal{C}(\mathbf{C}) \subseteq \mathcal{C}(\mathbf{A})$ a $\mathcal{C}(\mathbf{B}^T) \subseteq \mathcal{C}(\mathbf{A}^T)$. Potom matica $\mathbf{B}\mathbf{A}^-\mathbf{C}$ nezávisí na voľbe zovšeobecnenej inverzie $\mathbf{A}^-$. Dokážte.

Úloha 6.8. Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{X} \in \mathbb{R}^{m\times m}$, $\mathbf{Y} \in \mathbb{R}^{n\times n}$. Nech navyše $\mathbf{X}$ a $\mathbf{Y}$ sú regulárne a označme $\mathbf{B} = \mathbf{X}\mathbf{A}\mathbf{Y}$. Ukážte, že $\mathbf{G}$ je zovšeobecnenou inverziou $\mathbf{A}$ práve vtedy, keď $\mathbf{Y}^{-1}\mathbf{G}\mathbf{X}^{-1}$ je zovšeobecnenou inverziou $\mathbf{B}$.

Úloha 6.9. Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$ a nech $\mathbf{A}^-$ je zovšeobecnená inverzia matice $\mathbf{A}$. Ukážte, že potom $\mathcal{N}(\mathbf{A}) = \mathcal{C}(\mathbf{I}_n - \mathbf{A}^-\mathbf{A})$. Návod: Všimnite si, že ak $\mathbf{A}\mathbf{x} = \mathbf{0}_m$ pre $\mathbf{x} \in \mathbb{R}^n$, potom $\mathbf{x} = \mathbf{x} - \mathbf{A}^-\mathbf{A}\mathbf{x}$.

Úloha 6.10. Nech

$$\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}.$$

Ukážte, že

$$\mathbf{A}^+ = \begin{pmatrix} 0,1 & 0,2 \\ 0,1 & 0,2 \end{pmatrix}.$$

Úloha 6.11. Dokážte vlastnosti (i), (ii) z vety 6.34.

Úloha 6.12. Ukážte, že ak $\mathbf{A} = \operatorname{diag}(d_1, d_2)$, potom $\mathbf{A}^+ = \operatorname{diag}(d_1^+, d_2^+)$, kde $d_i^+ = 1/d_i$ ak $d_i \neq 0$, a $d_i^+ = 0$ ak $d_i = 0$ $(i = 1, 2)$. Následne overte vlastnosť (v) z vety 6.34.

Úloha 6.13. Ukážte, že ak $\mathbf{A} = \operatorname{diag}(\mathbf{B}, \mathbf{C})$, potom $\mathbf{A}^+ = \operatorname{diag}(\mathbf{B}^+, \mathbf{C}^+)$.

Kapitola 7

Idempotentné matice a projektory

7.1 Idempotentné matice

Definícia 7.1 (Idempotentná matica)**.** Štvorcovú maticu $\mathbf{A} \in \mathbb{R}^{n\times n}$ nazývame *idempotentná*, ak $\mathbf{A}^2 = \mathbf{A}$.

Príklad 7.2. Matice $\mathbf{0}_{n\times n}$, $\mathbf{I}_n$ sú idempotentné. Overte, že matica $\mathbf{J}_n$ idempotentná nie je a nájdite konštantu $c \neq 0$ takú, že $c\mathbf{J}_n$ je idempotentná. Ukážte tiež, že matice

$$\begin{pmatrix} 2 & 1 \\ -2 & -1 \end{pmatrix} \quad \text{a} \quad \begin{pmatrix} 3/4 & \sqrt{3}/4 \\ \sqrt{3}/4 & 1/4 \end{pmatrix}$$

sú idempotentné.

Lema 7.3 (Vlastnosti idempotentných matíc)**.** Pre idempotentné matice platí:

(i) Jediná regulárna idempotentná matica typu $n \times n$ je $\mathbf{I}_n$.

(ii) Ak $\mathbf{A}$ je idempotentná, tak aj $\mathbf{A}^T$ a $\mathbf{I} - \mathbf{A}$ sú idempotentné.

(iii) Ak $\mathbf{A}$ je idempotentná, tak $\mathbf{A}^k = \mathbf{A}$ pre každé $k = 1, 2, 3, \ldots$

(iv) Každá idempotentná matica je zovšeobecnenou inverziou seba samej.

Dôkaz. Pre dôkaz časti (i) predpokladajme, že $\mathbf{A}$ je regulárna a idempotentná. Keďže je regulárna, tak existuje $\mathbf{A}^{-1}$. Potom prenásobením rovnice $\mathbf{A}^2 = \mathbf{A}$ maticou $\mathbf{A}^{-1}$ dostávame $\mathbf{A} = \mathbf{I}$. Časti (ii)-(iv) sa ľahko overia priamym výpočtom. □

Veta 7.4 (O vzťahu medzi hodnosťou a stopou idempotentnej matice)**.** Nech $\mathbf{A}$ je idempotentná. Potom $\operatorname{rank}(\mathbf{A}) = \operatorname{tr}(\mathbf{A})$.

Dôkaz. Nech $\mathbf{A} \in \mathbb{R}^{n\times n}$ je idempotentná. Označme $r = \operatorname{rank}(\mathbf{A})$. Maticu $\mathbf{A}$ môžeme na základe vety 3.11 vyjadriť ako $\mathbf{A} = \mathbf{BT}$, kde $\mathbf{B} \in \mathbb{R}^{n\times r}$ a $\mathbf{T} \in \mathbb{R}^{r\times n}$ a $\operatorname{rank}(\mathbf{B}) = \operatorname{rank}(\mathbf{T}) = r$. Potom $\mathbf{BTBT} = \mathbf{A}^2 = \mathbf{A} = \mathbf{BT} = \mathbf{BI}_r\mathbf{T}$ a podľa vety 4.10 máme $\mathbf{TB} = \mathbf{I}_r$. Teda $\operatorname{tr}(\mathbf{A}) = \operatorname{tr}(\mathbf{BT}) = \operatorname{tr}(\mathbf{TB}) = \operatorname{tr}(\mathbf{I}_r) = r$. □

Lema 7.5 (O idempotentnosti $\mathbf{AA}^-$ a $\mathbf{A}^-\mathbf{A}$)**.** Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$. Potom $\mathbf{AA}^-$ aj $\mathbf{A}^-\mathbf{A}$ sú idempotentné matice pre ľubovoľnú $\mathbf{A}^- \in \mathcal{G}^-(\mathbf{A})$.

Dôkaz. Overme definíciu idempotentnej matice: $(\mathbf{A}\mathbf{A}^-)(\mathbf{A}\mathbf{A}^-) = (\mathbf{A}\mathbf{A}^-\mathbf{A})\mathbf{A}^- = \mathbf{A}\mathbf{A}^-$. Dôkaz pre $\mathbf{A}^-\mathbf{A}$ je analogický. $\square$

Veta 7.6 (Charakterizácia zovšeobecnenej inverzie pomocou idempotentnej matice)**.** Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$ a $\mathbf{B} \in \mathbb{R}^{n\times m}$. Potom nasledujúce dva výroky sú ekvivalentné:

(i) $\mathbf{AB}$ je idempotentná a $\operatorname{rank}(\mathbf{AB}) = \operatorname{rank}(\mathbf{A})$,

(ii) $\mathbf{B} \in \mathcal{G}^-(\mathbf{A})$.

Dôkaz. Nech $\mathbf{AB}$ je idempotentná a $\operatorname{rank}(\mathbf{AB}) = \operatorname{rank}(\mathbf{A})$. Z rovnosti hodností dostávame $\dim(\mathcal{C}(\mathbf{AB})) = \dim(\mathcal{C}(\mathbf{A}))$. Podľa vety 3.4 vieme, že $\mathcal{C}(\mathbf{AB}) \subseteq \mathcal{C}(\mathbf{A})$, a preto použitím vety 1.27(iii) získame $\mathcal{C}(\mathbf{AB}) = \mathcal{C}(\mathbf{A})$. Z toho opäť podľa vety 3.4 vyplýva $\mathbf{A} = \mathbf{ABX}$. Potom $\mathbf{ABA} = \mathbf{ABABX} = \mathbf{ABX} = \mathbf{A}$, keďže $\mathbf{AB}$ je idempotentná, a teda $\mathbf{B} \in \mathcal{G}^-(\mathbf{A})$.

Naopak, nech $\mathbf{B} \in \mathcal{G}^-(\mathbf{A})$. Podľa lemy 7.5 je $\mathbf{AB}$ idempotentná a podľa dôsledku 6.17 platí $\operatorname{rank}(\mathbf{AB}) = \operatorname{rank}(\mathbf{A})$. $\square$

7.2 Projekčné matice

7.2.1 Konštrukcia projekčnej matice

Poznámka 7.7. Nech V a W sú lineárne podpriestory $\mathbb{R}^n$ a nech $\mathbf{y} \in \mathbb{R}^n$. Pripomeňme, že podľa definície 1.45 platí $\mathbf{y} \perp V$ ak $\langle \mathbf{y}, \mathbf{x}\rangle = 0$ pre každé $\mathbf{x} \in V$. Navyše $W \perp V$ ak $\langle \mathbf{z}, \mathbf{x}\rangle = 0$ pre každé $\mathbf{z} \in W$ a $\mathbf{x} \in V$.

Lema 7.8 (O ortogonálnosti stĺpcových priestorov)**.** Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$ a $\mathbf{B} \in \mathbb{R}^{m\times k}$. Potom $\mathcal{C}(\mathbf{A}) \perp \mathcal{C}(\mathbf{B})$ práve vtedy, keď $\mathbf{A}^T\mathbf{B} = \mathbf{0}_{n\times k}$.

Dôkaz. Zapíšme $\mathbf{A}$ a $\mathbf{B}$ pomocou ich stĺpcov $\mathbf{A} = (\mathbf{a}_1, \ldots, \mathbf{a}_n)$, $\mathbf{B} = (\mathbf{b}_1, \ldots, \mathbf{b}_k)$, čím dostaneme $\mathcal{C}(\mathbf{A}) = \operatorname{span}(\mathbf{a}_1, \ldots, \mathbf{a}_n)$ a $\mathcal{C}(\mathbf{B}) = \operatorname{span}(\mathbf{b}_1, \ldots, \mathbf{b}_k)$. Potom podľa vety 1.46 platí $\mathcal{C}(\mathbf{A}) \perp \mathcal{C}(\mathbf{B})$ práve vtedy, keď $\mathbf{a}_i^T\mathbf{b}_j = 0$ pre každé $i = 1, \ldots, n$, $j = 1, \ldots, k$, čo sa dá zapísať ako $\mathbf{A}^T\mathbf{B} = \mathbf{0}$. $\square$

Veta 7.9 (Základná veta o ortogonálnej projekcii)**.** Nech $V \subseteq \mathbb{R}^n$ je lineárny podpriestor s $\dim(V) = r$ a nech $\mathbf{y} \in \mathbb{R}^n$. Potom existuje jednoznačne určený vektor $\mathbf{z} \in V$ taký, že $(\mathbf{y}-\mathbf{z}) \perp V$. Ak $r = 0$, tak $\mathbf{z} = \mathbf{0}_n$ a ak $r > 0$, tak $\mathbf{z}$ je možné vyjadriť v tvare $\mathbf{z} = c_1\mathbf{x}_1 + \ldots + c_r\mathbf{x}_r$, kde $\mathbf{x}_1, \ldots, \mathbf{x}_r$ je akákoľvek ortonormálna báza priestoru V a $c_i = \langle \mathbf{y}, \mathbf{x}_i\rangle$, $i = 1, \ldots, r$. Navyše platí $\mathbf{y} = \mathbf{z}$ práve vtedy, keď $\mathbf{y} \in V$.

Dôkaz. Dôkaz pre $r = 0$ je priamočiary, keďže vtedy $V = \{\mathbf{0}_n\}$. Nech $r > 0$ a nech $\mathbf{x}_1, \ldots, \mathbf{x}_r$ je ortonormálna báza V. Položme $c_i = \langle \mathbf{y}, \mathbf{x}_i\rangle$, $i = 1, \ldots, r$, a definujme $\mathbf{z} = c_1\mathbf{x}_1 + \ldots + c_r\mathbf{x}_r$. Zrejme $\mathbf{z} \in V$ a platí

$$\langle \mathbf{y}-\mathbf{z}, \mathbf{x}_i\rangle = \langle \mathbf{y}, \mathbf{x}_i\rangle - \langle \mathbf{z}, \mathbf{x}_i\rangle = \langle \mathbf{y}, \mathbf{x}_i\rangle - \langle \sum_{j=1}^r c_j\mathbf{x}_j, \mathbf{x}_i\rangle = c_i - \sum_{j=1}^r c_j\langle \mathbf{x}_j, \mathbf{x}_i\rangle = c_i - c_i = 0$$

vďaka tomu, že $\mathbf{x}_1, \ldots, \mathbf{x}_r$ tvoria ortonormálnu bázu. Keďže $V = \operatorname{span}(\mathbf{x}_1, \ldots, \mathbf{x}_r)$, tak z dôsledku 1.47 vyplýva $(\mathbf{y}-\mathbf{z}) \perp V$.

Nech $\mathbf{x} \in V$ je taký vektor, že $(\mathbf{y} - \mathbf{x}) \perp V$. Počítajme

$$\begin{aligned}\|\mathbf{x} - \mathbf{z}\|^2 &= \langle \mathbf{x} - \mathbf{z}, \mathbf{x} - \mathbf{z} \rangle = \langle \mathbf{x} - \mathbf{y} + \mathbf{y} - \mathbf{z}, \mathbf{x} - \sum_{i=1}^{r} c_i \mathbf{x}_i \rangle \\ &= \langle \mathbf{x} - \mathbf{y}, \mathbf{x} - \sum_{i=1}^{r} c_i \mathbf{x}_i \rangle + \langle \mathbf{y} - \mathbf{z}, \mathbf{x} - \sum_{i=1}^{r} c_i \mathbf{x}_i \rangle = 0 + 0 = 0,\end{aligned}$$

keďže $(\mathbf{x} - \mathbf{y}) \perp V$, $(\mathbf{y} - \mathbf{z}) \perp V$ a $(\mathbf{x} - \sum_{i=1}^{r} c_i \mathbf{x}_i) \in V$. Z vlastností normy (definícia 1.32) vyplýva $\mathbf{x} - \mathbf{z} = \mathbf{0}$, čiže $\mathbf{x} = \mathbf{z}$, čím je dokázaná jednoznačnosť vektora $\mathbf{z}$.

Na dôkaz poslednej časti tvrdenia uvážme, že zrejme ak $\mathbf{y} = \mathbf{z}$, tak $\mathbf{y} \in V$. Naopak, ak $\mathbf{y} \in V$, tak $\mathbf{y} = \sum_i b_i \mathbf{x}_i$, kde $b_i = \langle \mathbf{y}, \mathbf{x}_i \rangle$, $i = 1, \ldots, r$ podľa vety 5.29. Teda $b_i = c_i$ pre $i = 1, \ldots, r$, čiže $\mathbf{y} = \mathbf{z}$. □

Definícia 7.10 (Ortogonálna projekcia)**.** Nech $\mathbf{y} \in \mathbb{R}^n$ a nech V je lineárny podpriestor $\mathbb{R}^n$. Vektor $\mathbf{z} \in \mathbb{R}^n$ z vety 7.9 sa nazýva *ortogonálna projekcia* vektora $\mathbf{y}$ na podpriestor V.

Definícia 7.11 (Systém normálnych rovníc)**.** Systém

$$\mathbf{X}^T\mathbf{X}\mathbf{b} = \mathbf{X}^T\mathbf{y}$$

pre $\mathbf{X} \in \mathbb{R}^{n\times k}$ a $\mathbf{y} \in \mathbb{R}^n$ a neznámu premennú $\mathbf{b} \in \mathbb{R}^k$ nazývame *systém normálnych rovníc*. Všeobecnejšie, systém

$$\mathbf{X}^T\mathbf{X}\mathbf{B} = \mathbf{X}^T\mathbf{Y}$$

pre $\mathbf{X} \in \mathbb{R}^{n\times k}$ a $\mathbf{Y} \in \mathbb{R}^{n\times p}$ a neznámu premennú $\mathbf{B} \in \mathbb{R}^{k\times p}$ nazývame *maticový systém normálnych rovníc*.

Veta 7.12 (O konzistentnosti maticového systému normálnych rovníc)**.** Pre každé $\mathbf{X} \in \mathbb{R}^{n\times k}$ a $\mathbf{Y} \in \mathbb{R}^{n\times p}$ je systém $\mathbf{X}^T\mathbf{X}\mathbf{B} = \mathbf{X}^T\mathbf{Y}$ v premennej $\mathbf{B} \in \mathbb{R}^{k\times p}$ konzistentný.

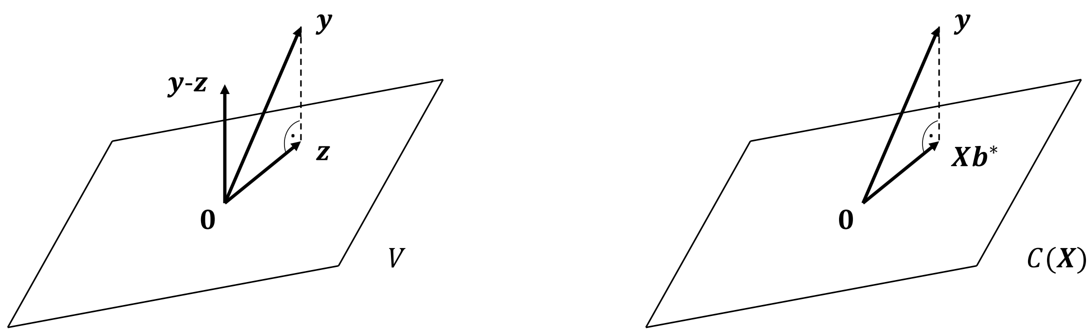
Dôkaz. Podľa dôsledku 3.35 platí $\mathcal{C}(\mathbf{X}^T\mathbf{X}) = \mathcal{C}(\mathbf{X}^T)$, pričom podľa vety 3.4 platí $\mathcal{C}(\mathbf{X}^T\mathbf{Y}) \subseteq \mathcal{C}(\mathbf{X}^T)$. Potom veta 4.6(i) zaručuje konzistenciu uvažovaného systému. □

Veta 7.13 (Ortogonálna projekcia pomocou riešení normálnych rovníc)**.** Nech $\mathbf{z}$ je ortogonálna projekcia vektora $\mathbf{y} \in \mathbb{R}^n$ na podpriestor $V \subseteq \mathbb{R}^n$ a nech $\mathbf{X} \in \mathbb{R}^{n\times k}$ spĺňa $\mathcal{C}(\mathbf{X}) = V$. Potom $\mathbf{z} = \mathbf{X}\mathbf{b}^*$ pre akékoľvek riešenie $\mathbf{b}^*$ systému $\mathbf{X}^T\mathbf{X}\mathbf{b} = \mathbf{X}^T\mathbf{y}$ v premennej $\mathbf{b}$.

Dôkaz. Podľa vety 7.12 je normálny systém rovníc konzistentný. Nech $\mathbf{b}^*$ je riešením $\mathbf{X}^T\mathbf{X}\mathbf{b} = \mathbf{X}^T\mathbf{y}$. Potom $\mathbf{X}^T(\mathbf{y} - \mathbf{X}\mathbf{b}^*) = \mathbf{0}$, čo znamená, že $(\mathbf{y} - \mathbf{X}\mathbf{b}^*) \perp \mathcal{C}(\mathbf{X}) = V$. Navyše $\mathbf{X}\mathbf{b}^* \in \mathcal{C}(\mathbf{X}) = V$, a teda z jednoznačnosti ortogonálnej projekcie dostávame $\mathbf{z} = \mathbf{X}\mathbf{b}^*$. □

Poznámka 7.14. Veta 7.13 vysvetľuje, prečo sa systém $\mathbf{X}^T\mathbf{X}\mathbf{b} = \mathbf{X}^T\mathbf{y}$ nazýva normálny. Ak hľadáme $\mathbf{b}^*$, ktoré je 'čo najbližšie' k riešeniu systému $\mathbf{y} = \mathbf{X}\mathbf{b}$ pomocou ortogonálnej projekcie vektora $\mathbf{y}$ do $\mathcal{C}(\mathbf{X})$, tak ho získame ako riešenie normálnych rovníc. Tieto rovnice teda vyjadrujú, že $\mathbf{y} - \mathbf{X}\mathbf{b}^*$ je kolmé na (teda tvorí *normálu* ku) $\mathcal{C}(\mathbf{X})$.

Ortogonálna projekcia daná vetou 3.4 je znázornená na obrázku 7.1a. Ortogonálna projekcia na stĺpcový priestor $\mathcal{C}(\mathbf{X})$ daná riešením normálnych rovníc je zakreslená na obrázku 7.1b.



(a) Projekcia $\mathbf{z}$ vektora $\mathbf{y}$ na podpriestor V. (b) Projekcia $\mathbf{Xb}^*$ vektora $\mathbf{y}$ na $\mathcal{C}(\mathbf{X})$.

Obr. 7.1: Ortogonálna projekcia.

Veta 7.15 (O vyjadrení projekcie maticou)**.** Nech $\mathbf{X} \in \mathbb{R}^{n\times k}$ a $\mathbf{y} \in \mathbb{R}^n$. Potom vektor $\mathbf{z} = \mathbf{X}(\mathbf{X}^T\mathbf{X})^-\mathbf{X}^T\mathbf{y}$ je ortogonálna projekcia $\mathbf{y}$ na $\mathcal{C}(\mathbf{X})$.

Dôkaz. Podľa vety 6.11 je $\mathbf{b}^* = (\mathbf{X}^T\mathbf{X})^-\mathbf{X}^T\mathbf{y}$ riešením normálnych rovníc. Potom $\mathbf{z} = \mathbf{Xb}^* = \mathbf{X}(\mathbf{X}^T\mathbf{X})^-\mathbf{X}^T\mathbf{y}$ je hľadanou projekciou. □

Definícia 7.16 (Projekčná matica)**.** Nech $\mathbf{X} \in \mathbb{R}^{n\times k}$. Maticu $\mathbf{P_X} = \mathbf{X}(\mathbf{X}^T\mathbf{X})^-\mathbf{X}^T$ nazývame *projekčná matica* (*projektor*) na priestor $\mathcal{C}(\mathbf{X})$.

Lema 7.17 (Nezávislosť projektoru na voľbe zovšeobecnenej inverzie)**.** Projekčná matica $\mathbf{P_X}$ nezávisí na voľbe zovšeobecnenej inverzie $(\mathbf{X}^T\mathbf{X})^-$.

Dôkaz. Podľa vety 3.34 platí $\mathcal{C}(\mathbf{X}^T) = \mathcal{C}(\mathbf{X}^T\mathbf{X})$, a teda použitím vety 3.4 dostávame $\mathbf{X}^T = \mathbf{X}^T\mathbf{XZ}$ pre nejakú maticu $\mathbf{Z}$. Potom $\mathbf{X} = \mathbf{Z}^T\mathbf{X}^T\mathbf{X}$, a teda $\mathbf{P_X} = \mathbf{Z}^T\mathbf{X}^T\mathbf{X}(\mathbf{X}^T\mathbf{X})^-\mathbf{X}^T\mathbf{XZ} = \mathbf{Z}^T\mathbf{X}^T\mathbf{XZ}$, čo je tvar, ktorý nezávisí na voľbe $(\mathbf{X}^T\mathbf{X})^-$. □

Poznámka 7.18. Jednoznačnosť projekčnej matice vyplýva aj priamo z toho, že projekcia je jednoznačne určené zobrazenie podľa vety 7.9.

Poznámka 7.19. Keďže projekciu ľubovoľného vektora $\mathbf{y}$ na $\mathcal{C}(\mathbf{X})$ vieme zapísať pomocou matice $\mathbf{P_X}$, tak zobrazenie ortogonálnej projekcie je lineárne.

Lema 7.20 (Ekvivalentná definícia projekčnej matice)**.** Nech $\mathbf{X} \in \mathbb{R}^{n\times k}$. Matica $\mathbf{P} \in \mathbb{R}^{n\times n}$ je projekčnou maticou na $\mathcal{C}(\mathbf{X})$ práve vtedy, keď spĺňa $\mathbf{Py} \in \mathcal{C}(\mathbf{X})$ a $(\mathbf{y} - \mathbf{Py}) \perp \mathcal{C}(\mathbf{X})$ pre každé $\mathbf{y} \in \mathbb{R}^n$.

Dôkaz. Zrejme ak $\mathbf{P}$ je projektor na $\mathcal{C}(\mathbf{X})$, tak spĺňa $\mathbf{Py} \in \mathcal{C}(\mathbf{X})$ a $(\mathbf{y} - \mathbf{Py}) \perp \mathcal{C}(\mathbf{X})$ pre každé $\mathbf{y} \in \mathbb{R}^n$. Naopak, ak $\mathbf{P}$ spĺňa $\mathbf{Py} \in \mathcal{C}(\mathbf{X})$ a $(\mathbf{y} - \mathbf{Py}) \perp \mathcal{C}(\mathbf{X})$ pre každé $\mathbf{y} \in \mathbb{R}^n$, tak $\mathbf{Py}$ je podľa vety 7.9 jednoznačne určenou projekciou $\mathbf{y}$ na $\mathcal{C}(\mathbf{X})$ pre každé $\mathbf{y} \in \mathbb{R}^n$. Podľa poznámky 7.18 je matica projekcie určená jednoznačne, čiže $\mathbf{P}$ je projektorom na $\mathcal{C}(\mathbf{X})$. □

Príklad 7.21. Nech $\mathbf{y} = (4, 8)^T$ a $\mathbf{X} = (3, 1)^T$. Nájdeme ortogonálnu projekciu $\mathbf{z}$ vektora $\mathbf{y}$ na $\mathcal{C}(\mathbf{X})$:

$$\begin{aligned}\mathbf{z} = \mathbf{X}(\mathbf{X}^T\mathbf{X})^-\mathbf{X}^T\mathbf{y} &= \begin{pmatrix}3\\1\end{pmatrix}\left[\begin{pmatrix}3 & 1\end{pmatrix}\begin{pmatrix}3\\1\end{pmatrix}\right]^- \begin{pmatrix}3 & 1\end{pmatrix}\begin{pmatrix}4\\8\end{pmatrix}\\ &= \begin{pmatrix}3\\1\end{pmatrix}\frac{1}{10}\begin{pmatrix}3 & 1\end{pmatrix}\begin{pmatrix}4\\8\end{pmatrix} = \frac{1}{10}\begin{pmatrix}9 & 3\\3 & 1\end{pmatrix}\begin{pmatrix}4\\8\end{pmatrix} = \begin{pmatrix}6\\2\end{pmatrix},\end{aligned}$$

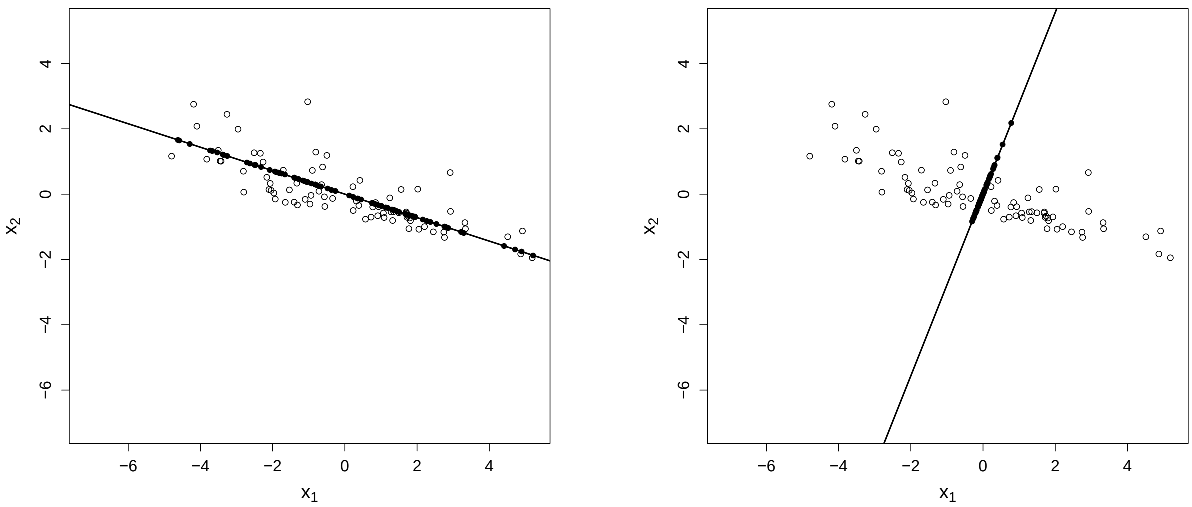
pričom matica

$$\frac{1}{10}\begin{pmatrix} 9 & 3 \\ 3 & 1 \end{pmatrix}$$

je projekčnou maticou na $\mathcal{C}(\mathbf{X})$. Overte, že $\mathbf{y} - \mathbf{z} \perp \mathcal{C}(\mathbf{X})$ a že $\mathbf{z} \in \mathcal{C}(\mathbf{X})$.

Poznámka 7.22. Niekedy je výhodné alebo potrebné znížiť dimenziu d-rozmerných dát. To je možné dosiahnuť projekciou dát do niektorého lineárneho (afinného) podpriestoru $\mathbb{R}^d$. Samozrejme, cenou za prácu s takýmito menejrozmernými dátami je strata istej časti informácie. Triviálnym prípadom je ignorovanie niektorých zložiek dát (napr. ignorovanie tretej zložky 3-rozmerných dát), čím vlastne projektujeme dáta do podpriestoru tvoreného ostatnými, neignorovanými, zložkami.

Je však možné projektovať dáta aj do iného menejrozmerného podpriestoru. Ak sa nám podarí nájsť podpriestor zvolenej dimenzie, ktorý čo najviac vystihuje rozptýlenosť dát, môže dôjsť iba k malej strate informácie. Poznamenajme, že zrejme najznámejšia metóda na hľadanie takých podpriestorov je tzv. metóda hlavných komponentov, ktorej sa bližšie budeme venovať v časti 10.2.1. Na obrázkoch 7.2a a 7.2b sú zachytené projekcie dvojrozmerných dát na zvolené priamky. Vidíme, že pôvodné dáta vystihujú omnoho lepšie projekcie na prvom obrázku, na rozdiel od projekcií na druhom obrázku, ktoré sa zdajú byť na vyjadrenie dát nevhodné.



(a) Priamka vystihujúca pôvodné dáta.

(b) Priamka nevystihujúca pôvodné dáta.

Obr. 7.2: Projekcie (plné krúžky) pôvodných dát (prázdne krúžky) na dve rôzne priamky.

7.2.2 Vlastnosti projekčnej matice

Veta 7.23 (Základné vlastnosti projekčných matíc)**.** Nech $\mathbf{X} \in \mathbb{R}^{n \times k}$. Potom

(i) $\mathbf{P_X X} = \mathbf{X}$,

(ii) $\mathbf{P_X} = \mathbf{X}\mathbf{B}^*$, kde $\mathbf{B}^*$ je akékoľvek riešenie systému $\mathbf{X}^T\mathbf{X}\mathbf{B} = \mathbf{X}^T$,

(iii) $\mathbf{P}_\mathbf{X}^T = \mathbf{P}_\mathbf{X}$ (projektor je symetrická matica),

(iv) $\mathbf{P}_\mathbf{X}^2 = \mathbf{P}_\mathbf{X}$ (projektor je idempotentná matica),

(v) $\mathcal{C}(\mathbf{P}_\mathbf{X}) = \mathcal{C}(\mathbf{X})$,

(vi) $\operatorname{rank}(\mathbf{P}_\mathbf{X}) = \operatorname{rank}(\mathbf{X})$,

(vii) $\mathbf{I} - \mathbf{P}_\mathbf{X}$ je symetrická a idempotentná,

(viii) $\operatorname{rank}(\mathbf{I} - \mathbf{P}_\mathbf{X}) = n - \operatorname{rank}(\mathbf{X})$.

Dôkaz. (i) Keďže každý stĺpec $\mathbf{x}_i$ ($i = 1, \dots, k$) matice $\mathbf{X}$ leží v $\mathcal{C}(\mathbf{X})$, tak $\mathbf{P}_\mathbf{X}\mathbf{x}_i = \mathbf{x}_i$, čiže $\mathbf{P}_\mathbf{X}\mathbf{X} = \mathbf{X}$. Túto vlastnosť použijeme pri dôkazoch nasledujúcich častí.

(ii) Ak $\mathbf{B}^*$ je riešením $\mathbf{X}^T\mathbf{X}\mathbf{B} = \mathbf{X}^T$, tak z vety 6.18 ju vieme vyjadriť ako $\mathbf{B}^* = (\mathbf{X}^T\mathbf{X})^-\mathbf{X}^T + (\mathbf{I} - (\mathbf{X}^T\mathbf{X})^-\mathbf{X}^T\mathbf{X})\mathbf{Y}$ pre nejakú zovšeobecnenú inverziu $(\mathbf{X}^T\mathbf{X})^-$ matice $\mathbf{X}^T\mathbf{X}$ a nejakú $\mathbf{Y} \in \mathbb{R}^{k\times n}$, a teda

$$\begin{aligned}\mathbf{X}\mathbf{B}^* &= \mathbf{X}(\mathbf{X}^T\mathbf{X})^-\mathbf{X}^T + (\mathbf{X} - \mathbf{X}(\mathbf{X}^T\mathbf{X})^-\mathbf{X}^T\mathbf{X})\mathbf{Y} = \mathbf{P}_\mathbf{X} + (\mathbf{X} - \mathbf{P}_\mathbf{X}\mathbf{X})\mathbf{Y}\\ &= \mathbf{P}_\mathbf{X} + (\mathbf{X} - \mathbf{X})\mathbf{Y} = \mathbf{P}_\mathbf{X},\end{aligned}$$

keďže $\mathbf{P}_\mathbf{X}\mathbf{X} = \mathbf{X}$.

(iii) Keďže $\mathbf{X}^T\mathbf{X}$ je symetrická, tak podľa vety 6.10 existuje jej zovšeobecnená inverzia $(\mathbf{X}^T\mathbf{X})^-$, ktorá je tiež symetrická. Preto aj matica $\mathbf{X}(\mathbf{X}^T\mathbf{X})^-\mathbf{X}^T$ je symetrická.

(iv) Keďže $\mathbf{P}_\mathbf{X}\mathbf{X} = \mathbf{X}$, tak $\mathbf{P}_\mathbf{X}^2 = \mathbf{P}_\mathbf{X}\mathbf{X}(\mathbf{X}^T\mathbf{X})^-\mathbf{X}^T = \mathbf{X}(\mathbf{X}^T\mathbf{X})^-\mathbf{X}^T = \mathbf{P}_\mathbf{X}$.

(v) Podľa vety 3.4 platí $\mathcal{C}(\mathbf{X}) \subseteq \mathcal{C}(\mathbf{P}_\mathbf{X})$, keďže $\mathbf{P}_\mathbf{X}\mathbf{X} = \mathbf{X}$. Zároveň $\mathbf{P}_\mathbf{X} = \mathbf{X}(\mathbf{X}^T\mathbf{X})^-\mathbf{X}^T$, a teda podľa tej istej vety dostávame $\mathcal{C}(\mathbf{P}_\mathbf{X}) \subseteq \mathcal{C}(\mathbf{X})$.

(vi) Tvrdenie platí podľa (v).

(vii) Dôkaz tohto tvrdenia je priamočiary.

(viii) Podľa vety 7.4 platí $\operatorname{rank}(\mathbf{I} - \mathbf{P}_\mathbf{X}) = \operatorname{tr}(\mathbf{I} - \mathbf{P}_\mathbf{X})$. To môžeme ďalej upraviť na $\operatorname{tr}(\mathbf{I}) - \operatorname{tr}(\mathbf{P}_\mathbf{X}) = n - \operatorname{rank}(\mathbf{P}_\mathbf{X}) = n - \operatorname{rank}(\mathbf{X})$, opäť využijúc vetu 7.4. □

Veta 7.24 (Projektor ako symetrická idempotentná matica)**.** Nech $\mathbf{P} \in \mathbb{R}^{n\times n}$. Potom $\mathbf{P}$ je projektor na $\mathcal{C}(\mathbf{P})$ práve vtedy, keď $\mathbf{P}$ je symetrická idempotentná matica.

Dôkaz. Ak $\mathbf{P}$ je projektor, tak podľa vety 7.23 je $\mathbf{P}$ symetrická a idempotentná. Naopak, nech $\mathbf{P}$ je symetrická a idempotentná matica. Potom projektor na $\mathcal{C}(\mathbf{P})$ je $\mathbf{P}(\mathbf{P}^T\mathbf{P})^-\mathbf{P}^T = \mathbf{P}(\mathbf{P}^2)^-\mathbf{P} = \mathbf{P}\mathbf{P}^-\mathbf{P} = \mathbf{P}$ vďaka symetrickosti a idempotentnosti $\mathbf{P}$. Teda $\mathbf{P}$ je projektor na $\mathcal{C}(\mathbf{P})$. □

Veta 7.25 (Vyjadrenie projektora na $\mathcal{C}(\mathbf{X})^\perp$ pomocou projektora na $\mathcal{C}(\mathbf{X})$)**.** Nech $\mathbf{X} \in \mathbb{R}^{n\times k}$. Potom $\mathbf{I}_n - \mathbf{P}_\mathbf{X}$ je projektor na $\mathcal{C}(\mathbf{X})^\perp$.

Dôkaz. Podľa vety 7.23 je $\mathbf{I} - \mathbf{P}_\mathbf{X}$ symetrická a idempotentná matica, teda podľa vety 7.24 je $\mathbf{I} - \mathbf{P}_\mathbf{X}$ projektor na $\mathcal{C}(\mathbf{I} - \mathbf{P}_\mathbf{X})$. Potrebujeme dokázať, že $\mathcal{C}(\mathbf{I} - \mathbf{P}_\mathbf{X}) = \mathcal{C}(\mathbf{X})^\perp$. Z dôsledku 6.14 plynie, že $\mathcal{C}(\mathbf{I} - \mathbf{P}_\mathbf{X}) = \mathcal{N}(\mathbf{I} - (\mathbf{I} - \mathbf{P}_\mathbf{X})(\mathbf{I} - \mathbf{P}_\mathbf{X})^-)$ pre akúkoľvek zovšeobecnenú inverziu $(\mathbf{I} - \mathbf{P}_\mathbf{X})^-$. Lenže $\mathbf{I} - \mathbf{P}_\mathbf{X}$ je idempotentná a je teda zovšeobecnenou inverziou sama sebe (lema 7.3), takže dostávame

$$\mathcal{C}(\mathbf{I} - \mathbf{P}_\mathbf{X}) = \mathcal{N}(\mathbf{I} - (\mathbf{I} - \mathbf{P}_\mathbf{X})(\mathbf{I} - \mathbf{P}_\mathbf{X})) = \mathcal{N}(\mathbf{I} - (\mathbf{I} - \mathbf{P}_\mathbf{X})) = \mathcal{N}(\mathbf{P}_\mathbf{X}),$$

pričom pri druhej rovnosti sme opäť použili idempotentnosť $(\mathbf{I}-\mathbf{P_X})$. Navyše ak použijeme vlastnosti (iii) a (v) z vety 7.23, dostávame $\mathcal{N}(\mathbf{P_X}) = \mathcal{C}(\mathbf{P_X^T})^\perp = \mathcal{C}(\mathbf{P_X})^\perp = \mathcal{C}(\mathbf{X})^\perp$, čím je dôkaz uzavretý. □

Veta 7.26 (Projektor pomocou pseudoinverzie)**.** Nech $\mathbf{X} \in \mathbb{R}^{n\times k}$. Potom $\mathbf{XX}^+$ je projekčnou maticou na $\mathcal{C}(\mathbf{X})$.

Dôkaz. Pomocou vlastností 1-3 pseudoinverzie ľahko overíme, že matica $\mathbf{XX}^+$ je symetrická a idempotentná. Teda podľa vety 7.24 je $\mathbf{XX}^+$ projektorom na $\mathcal{C}(\mathbf{XX}^+)$. Lenže podľa vety 6.16 platí $\mathcal{C}(\mathbf{XX}^+) = \mathcal{C}(\mathbf{X})$, čím je tvrdenie dokázané. □

Definícia 7.27 (Problém najmenších štvorcov)**.** *Problém najmenších štvorcov* pre zadané $\mathbf{X} \in \mathbb{R}^{n\times k}$ a $\mathbf{y} \in \mathbb{R}^n$ je nájsť také $\mathbf{b}^* \in \mathbb{R}^k$, ktoré minimalizuje $\|\mathbf{y}-\mathbf{Xb}\|^2 = \sum_{i=1}^n (y_i - (\mathbf{Xb})_i)^2$ cez všetky $\mathbf{b} \in \mathbb{R}^k$. Teda $\mathbf{b}^*$ spĺňa $\|\mathbf{y}-\mathbf{Xb}\| \geq \|\mathbf{y}-\mathbf{Xb}^*\|$ pre každé $\mathbf{b} \in \mathbb{R}^k$.

Poznámka 7.28. Ak $\mathbf{b}^*$ je riešením problému najmenších štvorcov pre $\mathbf{X} \in \mathbb{R}^{n\times k}$ a $\mathbf{y} \in \mathbb{R}^n$, tak hovoríme, že $\mathbf{b}^*$ *minimalizuje súčet štvorcov* pre dané $\mathbf{X}$, $\mathbf{y}$.

Problém najmenších štvorcov zodpovedá hľadaniu takého $\mathbf{b}$, aby $\mathbf{Xb}$ bolo čo najbližšie k $\mathbf{y}$, teda aby sme sa dostali čo najbližšie k riešeniu rovnice $\mathbf{Xb} = \mathbf{y}$. Výraz $\|\mathbf{y}-\mathbf{Xb}\|^2$ vyjadruje vzdialenosť danej aproximácie (vektora $\mathbf{Xb}$) od $\mathbf{y}$, takže sa dá vnímať ako chyba tejto aproximácie. Minimalizácia tejto chyby je aj intuitívne rozumný spôsob, ako sa dostať čo najbližšie k vektoru $\mathbf{y}$. Štatistickému pohľadu na problém najmenších štvorcov sa budeme venovať v podkapitole 7.2.3.

Veta 7.29 (O vzdialenosti projekcie)**.** Nech $\mathbf{y} \in \mathbb{R}^n$ a $\mathbf{X} \in \mathbb{R}^{n\times k}$. Potom pre každé $\mathbf{u} \in \mathcal{C}(\mathbf{X})$ platí $\|\mathbf{y}-\mathbf{u}\| \geq \|\mathbf{y}-\mathbf{P_X y}\|$. Navyše, rovnosť nastáva práve vtedy, keď $\mathbf{u} = \mathbf{P_X y}$.

Dôkaz. Položme $\mathbf{z} = \mathbf{P_X y}$. Potom pre každé $\mathbf{u} \in \mathcal{C}(\mathbf{X})$ platí

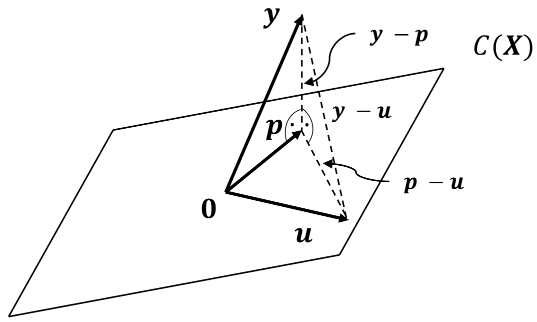
$$\begin{aligned}
\|\mathbf{y}-\mathbf{u}\|^2 &= \|(\mathbf{y}-\mathbf{z}) + (\mathbf{z}-\mathbf{u})\|^2 = \langle (\mathbf{y}-\mathbf{z}) + (\mathbf{z}-\mathbf{u}), (\mathbf{y}-\mathbf{z}) + (\mathbf{z}-\mathbf{u}) \rangle \\
&= \langle \mathbf{y}-\mathbf{z}, \mathbf{y}-\mathbf{z} \rangle + 2\langle \mathbf{y}-\mathbf{z}, \mathbf{z}-\mathbf{u} \rangle + \langle \mathbf{z}-\mathbf{u}, \mathbf{z}-\mathbf{u} \rangle \\
&= \|\mathbf{y}-\mathbf{P_X y}\|^2 + 2\langle \mathbf{y}-\mathbf{P_X y}, \mathbf{P_X y}-\mathbf{u} \rangle + \|\mathbf{P_X y}-\mathbf{u}\|^2 \\
&= \|\mathbf{y}-\mathbf{P_X y}\|^2 + \|\mathbf{P_X y}-\mathbf{u}\|^2,
\end{aligned}$$

keďže $(\mathbf{P_X y}-\mathbf{u}) \in \mathcal{C}(\mathbf{X})$, a teda $(\mathbf{y}-\mathbf{P_X y}) \perp (\mathbf{P_X y}-\mathbf{u})$. Potom vďaka $\|\mathbf{P_X y}-\mathbf{u}\|^2 \geq 0$ dostávame $\|\mathbf{y}-\mathbf{u}\|^2 \geq \|\mathbf{y}-\mathbf{P_X y}\|^2$.

Aby nastala rovnosť, musí platiť $\|\mathbf{P_X y}-\mathbf{u}\|^2 = 0$, z čoho podľa kladnej definitnosti normy vyplýva $\mathbf{u} = \mathbf{P_X y}$. □

Poznámka 7.30. Rovnica $\|\mathbf{y}-\mathbf{u}\|^2 = \|\mathbf{y}-\mathbf{P_X y}\|^2 + \|\mathbf{P_X y}-\mathbf{u}\|^2$ odvodená v dôkaze vety 7.29 má priamočiaru geometrickú interpretáciu. Ide vlastne o Pytagorovu vetu, kde vzhľadom na vlastnosti projekcie je medzi $\mathbf{y}-\mathbf{P_X y}$ a $\mathbf{P_X y}-\mathbf{u}$ pravý uhol, a $\mathbf{y}-\mathbf{u}$ predstavuje preponu vzniknutého trojuholníka (viď obrázok 7.3). Z Pytagorovej vety zároveň hneď dostávame, že odvesna $\mathbf{y}-\mathbf{P_X y}$ je nanajvýš taká dlhá ako prepona $\mathbf{y}-\mathbf{u}$, teda tvrdenie vety 7.29.

Dôsledok 7.31 (Minimalizácia súčtu štvorcov pomocou projekcie)**.** Nech $\mathbf{y} \in \mathbb{R}^n$ a $\mathbf{X} \in \mathbb{R}^{n\times k}$. Potom $\mathbf{b}^* \in \mathbb{R}^k$ je riešením normálnych rovníc $\mathbf{X}^T\mathbf{Xb} = \mathbf{X}^T\mathbf{y}$ práve vtedy, keď je riešením problému najmenších štvorcov. Teda $\|\mathbf{y}-\mathbf{Xb}\|^2 \geq \|\mathbf{y}-\mathbf{Xb}^*\|^2$ pre každé $\mathbf{b} \in \mathbb{R}^k$ práve vtedy, keď $\mathbf{X}^T\mathbf{Xb}^* = \mathbf{X}^T\mathbf{y}$.



Obr. 7.3: Geometrické znázornenie vety 7.29 a poznámky 7.30. Vektor $\mathbf{p}$ vyjadruje projekciu $\mathbf{P_X y}$, prerušované čiary znázorňujú vektory $\mathbf{y}-\mathbf{u}$, $\mathbf{y}-\mathbf{P_X y}$ a $\mathbf{P_X y}-\mathbf{u}$.

Dôkaz. Nech $\mathbf{X}^T\mathbf{X}\mathbf{b}^* = \mathbf{X}^T\mathbf{y}$. Potom podľa vety 7.13 je $\mathbf{X}\mathbf{b}^*$ projekciou $\mathbf{y}$ na $\mathcal{C}(\mathbf{X})$, čiže $\mathbf{X}\mathbf{b}^* = \mathbf{P_X y}$. Nech $\mathbf{b} \in \mathbb{R}^k$. Potom platí $\mathbf{X}\mathbf{b} \in \mathcal{C}(\mathbf{X})$, takže z vety 7.29 plynie, že $\|\mathbf{y}-\mathbf{X}\mathbf{b}\|^2 \geq \|\mathbf{y}-\mathbf{X}\mathbf{b}^*\|^2$.

Naopak, nech $\mathbf{b}^*$ je riešením problému najmenších štvorcov, teda spĺňa $\|\mathbf{y}-\mathbf{X}\mathbf{b}\|^2 \geq \|\mathbf{y}-\mathbf{X}\mathbf{b}^*\|^2$ pre ľubovoľné $\mathbf{b} \in \mathbb{R}^k$. Potom podľa vety 7.29 platí $\mathbf{X}\mathbf{b}^* = \mathbf{P_X y}$, teda $\mathbf{X}\mathbf{b}^*$ je projekcia vektoru $\mathbf{y}$ na $\mathcal{C}(\mathbf{X})$. Vzhľadom na vetu 7.13 potom $\mathbf{X}\mathbf{b}^*$ musí spĺňať $\mathbf{X}\mathbf{b}^* = \mathbf{X}\tilde{\mathbf{b}}$ pre nejaké $\tilde{\mathbf{b}}$ spĺňajúce $\mathbf{X}^T\mathbf{X}\tilde{\mathbf{b}} = \mathbf{X}^T\mathbf{y}$. Z toho vyplýva, že aj $\mathbf{X}^T\mathbf{X}\mathbf{b}^* = \mathbf{X}^T\mathbf{y}$. □

Poznámka 7.32. Keďže $\mathbf{b}^*$ v dôsledku 7.31 je riešením normálnych rovníc, tak $\mathbf{X}\mathbf{b}^*$ je projekciou $\mathbf{y}$ na $\mathcal{C}(\mathbf{X})$. Teda projekcia vektora $\mathbf{y}$ na $\mathcal{C}(\mathbf{X})$ zakreslená na obrázku 7.1b podľa tohto dôsledku zároveň reprezentuje riešenie problému najmenších štvorcov (resp. je určená riešením $\mathbf{b}^*$ problému najmenších štvorcov).

Dôsledok 7.33 (Minimalizácia súčtu štvorcov pomocou pseudoinverzie)**.** Nech $\mathbf{X} \in \mathbb{R}^{n\times k}$ a $\mathbf{y} \in \mathbb{R}^n$. Označme $\mathbf{b}^* = \mathbf{X}^+\mathbf{y}$. Potom $\|\mathbf{y}-\mathbf{X}\mathbf{b}\| \geq \|\mathbf{y}-\mathbf{X}\mathbf{b}^*\|$ pre každé $\mathbf{b} \in \mathbb{R}^k$.

Dôkaz. Vidíme, že $\mathbf{X}\mathbf{b}^* = \mathbf{X}\mathbf{X}^+\mathbf{y}$, pričom však $\mathbf{X}\mathbf{X}^+$ je podľa vety 7.26 projektor na $\mathcal{C}(\mathbf{X})$. Teda $\mathbf{X}\mathbf{b}^* = \mathbf{P_X y}$ a podľa vety 7.29 potom platí $\|\mathbf{y}-\mathbf{X}\mathbf{b}\| \geq \|\mathbf{y}-\mathbf{X}\mathbf{b}^*\|$ pre ľubovoľné $\mathbf{b} \in \mathbb{R}^k$. □

Príklad 7.34. Nech

$$\mathbf{X} = \begin{pmatrix} 1 & -1 \\ 1 & 5 \\ 1 & -3 \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

Vyriešte problém najmenších štvorcov pre dané $\mathbf{X}$ a $\mathbf{y}$, teda nájdite také $\mathbf{b}^* \in \mathbb{R}^2$, že $\|\mathbf{y}-\mathbf{X}\mathbf{b}\|^2 \geq \|\mathbf{y}-\mathbf{X}\mathbf{b}^*\|^2$ pre každé $\mathbf{b} \in \mathbb{R}^2$.

Poznámka 7.35. Pokiaľ matica $\mathbf{X} \in \mathbb{R}^{n\times k}$ nemá plnú stĺpcovú hodnosť $\text{rank}(\mathbf{X}) = k$, tak vzhľadom na vetu 6.18 existuje nekonečne veľa riešení systému normálnych rovníc $\mathbf{X}^T\mathbf{X}\mathbf{b} = \mathbf{X}^T\mathbf{y}$, a teda aj nekonečne veľa vektorov $\mathbf{b}^*$, ktoré riešia problém najmenších štvorcov.

Veta 7.36 (Riešenie problému najmenších štvorcov s najmenšou normou)**.** Nech $\mathbf{X} \in \mathbb{R}^{n\times k}$, $\mathbf{y} \in \mathbb{R}^n$ a nech $\tilde{\mathbf{b}}$ je riešenie príslušného problému najmenších štvorcov. Potom riešenie $\mathbf{b}^* = \mathbf{X}^+\mathbf{y}$ problému najmenších štvorcov spĺňa $\|\mathbf{b}^*\| \leq \|\tilde{\mathbf{b}}\|$.

Dôkaz. Keďže $\tilde{\mathbf{b}}$ je riešením problému najmenších štvorcov, tak podľa dôsledku 7.31 je zároveň riešením normálnych rovníc $\mathbf{X}^T\mathbf{X}\mathbf{b} = \mathbf{X}^T\mathbf{y}$. Podľa vety 6.38 potom platí $\|\tilde{\mathbf{b}}\| \geq \|\mathbf{b}_+\|$, kde $\mathbf{b}_+ = (\mathbf{X}^T\mathbf{X})^+\mathbf{X}^T\mathbf{y}$. Aplikovaním lemy 6.36 dostávame $(\mathbf{X}^T\mathbf{X})^+\mathbf{X}^T = \mathbf{X}^+$, čiže $\mathbf{b}_+ = \mathbf{b}^*$, čím je veta dokázaná. □

Veta 7.37 (Minimalizácia súčtu štvorcov pomocou **QR** rozkladu)**.** Nech $\mathbf{X} \in \mathbb{R}^{n\times k}$, kde $n \geq k$, má hodnosť k a nech $\mathbf{X} = \mathbf{QR}$ je jej úzky **QR** rozklad. Potom $\mathbf{b}^* = \mathbf{R}^{-1}\mathbf{Q}^T\mathbf{y}$ je riešením problému najmenších štvorcov $\min_{\mathbf{b}}\|\mathbf{y} - \mathbf{X}\mathbf{b}\|^2$ pre ľubovoľné $\mathbf{y} \in \mathbb{R}^n$.

Dôkaz. Podľa dôsledku 4.34 má $\mathbf{Q} \in \mathbb{R}^{m\times n}$ ortonormálne stĺpce, teda $\mathbf{Q}^T\mathbf{Q} = \mathbf{I}_n$. Ďalej $\mathbf{R} \in \mathbb{R}^{k\times k}$ je regulárna. Keďže $\mathbf{X} \in \mathbb{R}^{n\times k}$ má hodnosť k, tak $\operatorname{rank}(\mathbf{X}^T\mathbf{X}) = \operatorname{rank}(\mathbf{X}) = k$, a teda $\mathbf{X}^T\mathbf{X}$ je regulárna. Potom podľa dôsledku 7.31 je riešenie problému najmenších štvorcov:

$$\begin{aligned}\mathbf{b}^* &= (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y} = (\mathbf{R}^T\mathbf{Q}^T\mathbf{Q}\mathbf{R})^{-1}\mathbf{R}^T\mathbf{Q}^T\mathbf{y} = (\mathbf{R}^T\mathbf{R})^{-1}\mathbf{R}^T\mathbf{Q}^T\mathbf{y}\\ &= \mathbf{R}^{-1}(\mathbf{R}^T)^{-1}\mathbf{R}^T\mathbf{Q}^T\mathbf{y} = \mathbf{R}^{-1}\mathbf{Q}^T\mathbf{y}.\end{aligned}$$

□

Poznámka 7.38. Riešenie problému najmenších štvorcov pomocou vety 7.37 má oproti “klasickému” priamemu riešeniu pomocou normálnych rovníc výhodu v tom, že je numericky stabilnejšie. Preto sa aj v praxi často používa práve tento spôsob riešenia.

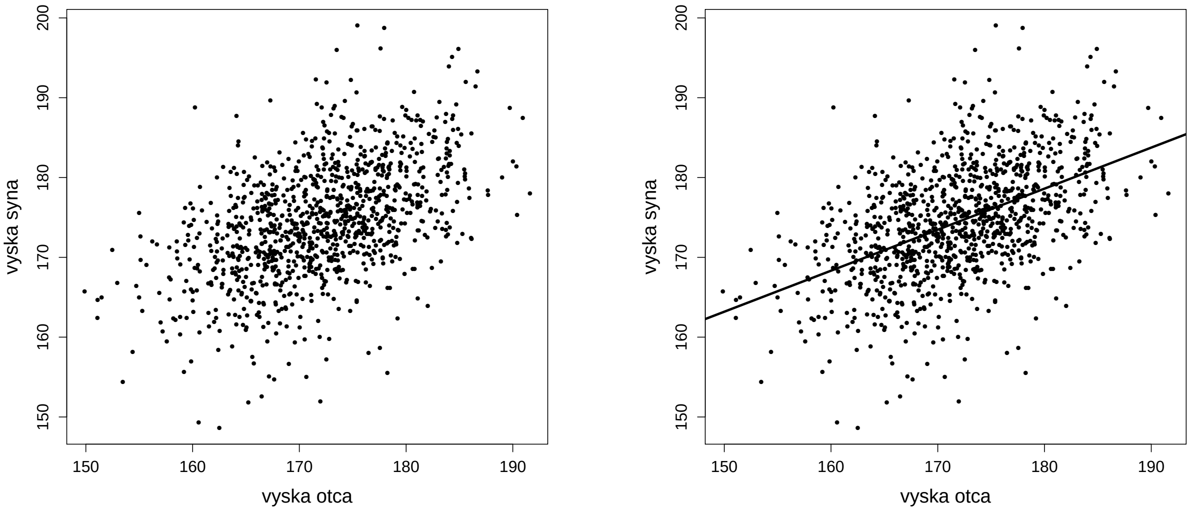
7.2.3 Minimalizácia súčtu štvorcov v štatistike

Problém najmenších štvorcov $\|\mathbf{y}-\mathbf{X}\mathbf{b}\|^2 \to \min_{\mathbf{b}}$ riešime nielen, keď hľadáme vektor $\mathbf{b}$, ktorý je “čo najbližšie” ku riešeniu systému $\mathbf{y} = \mathbf{X}\mathbf{b}$; tento systém hrá významnú úlohu aj v štatistike. Metóda najmenších štvorcov postavená na riešení problému najmenších štvorcov je jedna z najčastejšie používaných metód odhadovania neznámych parametrov v takzvanom lineárnom regresnom modeli. Pre potreby týchto skrípt metódu najmenších štvorcov iba stručne načrtneme, podrobnosti sa dajú nájsť napríklad v [6] a [9].

Uvažujme nasledujúci problém: máme namerané výšky 1078 mužov a ich otcov a snažíme sa zistiť, aký je medzi nimi vzťah. Tieto výšky sú zobrazené na obrázku 7.4a. Dáta pochádzajú z balíka *UsingR* štatistického softvéru R [10], pôvodne sú to dáta od K. Pearsona namerané v Anglicku približne v roku 1900. Pokúsime sa závislosť medzi výškami synov a ich otcov opísať ako afinnú, čiže dátami na obrázku preložíme priamku. Formálne, nech $y_1, \dots, y_n$ sú výšky synov a $x_1, \dots, x_n$ sú príslušné výšky otcov, kde $n = 1078$. Potom priamka prechádzajúca všetkými dátami by bola určená predpisom

$$y_i = b_1 + b_2 x_i, \quad i = 1, \dots, n. \tag{7.1}$$

Keďže dáta zjavne neležia na jednej priamke, tak systém (7.1) nemá riešenie: neexistujú b_1, b_2, ktoré by spĺňali všetkých n rovníc. Preto namiesto toho nájdeme priamku, ktorá je “najbližšie” k našim dátam. Nakoľko je priamka $y(x) = b_1 + b_2 x$ vzdialená od nameraných hodnôt $(x_1, y_1)^T, \dots, (x_n, y_n)^T$ budeme merať súčtom druhých mocnín (štvorcov) vzdialeností medzi skutočnými výškami synov y_i a priamkou odhadnutými výškami synov $y(x_i) = b_1 + b_2 x_i$. Minimalizujeme teda $\sum_{i=1}^n (y_i - (b_1 + b_2 x_i))^2$, získame tak odhady $\hat{b}_1, \hat{b}_2$. Táto metóda odhadu koeficientov priamky b_1 a b_2 čo najviac vystihujúcej dané dáta sa nazýva *metóda najmenších*



(a) Dáta (b) Dáta aj s regresnou priamkou

Obr. 7.4: Výšky otcov a synov z podkapitoly 7.2.3. Každý bod na obrázku predstavuje výšku syna na y-ovej súradnici a jeho otca na x-ovej súradnici, obidve v centimetroch. Priamka prechádzajúca dátami je daná ako $y = \hat{b}_1 + \hat{b}_2 x$, kde $\hat{b}_1, \hat{b}_2$ sú odhady neznámych parametrov získané metódou najmenších štvorcov.

štvorcov. Priamka získaná metódou najmenších štvorcov pre naše dáta je zakreslená na obrázku 7.4b, konkrétne hodnoty odhadov sú približne $\hat{b}_1 = 86{,}1$ a $\hat{b}_2 = 0{,}5$, čiže model odhaduje, že ak otec bol o jeden centimeter vyšší, tak syn je v priemere vyšší o pol centimetra.

Systém (7.1) môžeme zapísať vektorovo ako $\mathbf{y} = \mathbf{X}\mathbf{b}$, kde $\mathbf{y} = (y_1, \ldots, y_n)^T$, $\mathbf{b} = (b_1, b_2)^T$ a

$$\mathbf{X} = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix},$$

a teda metódu najmenších štvorcov vieme zapísať ako minimalizáciu $\|\mathbf{y} - \mathbf{X}\mathbf{b}\|^2$, čo je presne problém najmenších štvorcov z definície 7.27.

Vo všeobecnosti môžeme riešením problému najmenších štvorcov získať odhady neznámych koeficientov v lineárnom regresnom modeli, kde *vysvetľovanú premennú* y_i (napr. výšky synov) modelujeme pomocou viacerých *vysvetľujúcich premenných* $x_{i1}, \ldots, x_{ik}$ (napr. výšky otcov a výšky matiek). Dostaneme tak opäť (nekonzistentný) systém rovníc $\mathbf{y} = \mathbf{X}\mathbf{b}$, kde $\mathbf{y} \in \mathbb{R}^n$, $\mathbf{b} \in \mathbb{R}^k$ a $\mathbf{X} \in \mathbb{R}^{n \times k}$. Odhad koeficientov $\mathbf{b}$ metódou najmenších štvorcov bude opäť taký vektor $\hat{\mathbf{b}}$, ktorý je čo najbližšie riešeniu systému $\mathbf{y} = \mathbf{X}\mathbf{b}$, v zmysle minimalizácie súčtu štvorcov odchýlok $\|\mathbf{y} - \mathbf{X}\mathbf{b}\|^2$.

Príklad 7.39. Príkladom zložitejšieho lineárneho regresného modelu je situácia, kde pre každého syna zmeriame výšku jeho otca aj matky. Potom máme model $y_i = b_1 + b_2 m_i + b_3 z_i$, $i = 1, \ldots, n$, kde y_i je výška i-teho syna, m_i je výška i-teho otca a z_i je výška i-tej matky. Ako vyzerá matica $\mathbf{X}$ vo vektorovom zápise $\mathbf{y} = \mathbf{X}\mathbf{b}$ tohto modelu?

7.3 Úlohy na precvičenie

Úloha 7.1. Dokážte nasledovné tvrdenie. Nech $\mathbf{A} \in \mathbb{R}^{n\times n}$. Potom $\mathcal{N}(\mathbf{A}) = \mathcal{C}(\mathbf{I}_n - \mathbf{A})$ vtedy a len vtedy, keď $\mathbf{A}$ je idempotentná. Návod: Použite tvrdenie úlohy 6.9.

Úloha 7.2. Nech $\mathbf{A} \in \mathbb{R}^{n\times n}$ je idempotentná symetrická matica. Potom $\mathbf{I}_n - 2\mathbf{A}$ je ortogonálna matica. Dokážte! Nájdite geometrickú interpretáciu ortogonálnej matice $\mathbf{I}_n - 2\mathbf{A}$ pre prípad, že $\mathbf{A}$ je projektor na lineárny podpriestor $U \subseteq \mathbb{R}^n$ (aspoň pre $n = 2$).

Úloha 7.3. Najprv ukážte, že ak $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{B}, \mathbf{C} \in \mathbb{R}^{n\times p}$ a $\mathbf{A}^T\mathbf{A}\mathbf{B} = \mathbf{A}^T\mathbf{A}\mathbf{C}$, tak $\mathbf{A}\mathbf{B} = \mathbf{A}\mathbf{C}$. Návod: Využite rovnosť $(\mathbf{AB} - \mathbf{AC})^T(\mathbf{AB} - \mathbf{AC}) = (\mathbf{B}^T - \mathbf{C}^T)(\mathbf{A}^T\mathbf{AB} - \mathbf{A}^T\mathbf{AC})$. Pomocou tohto tvrdenia dokážte, že ak $\mathbf{A} \in \mathbb{R}^{n\times n}$ je symetrická matica a $\mathbf{A}^3 = \mathbf{A}^2$, tak $\mathbf{A}$ je idempotentná.

Úloha 7.4. Nech $\mathbf{A}, \mathbf{B}$ sú symetrické idempotentné matice, pre ktoré platí $\mathcal{C}(\mathbf{A}) = \mathcal{C}(\mathbf{B})$. Potom $\mathbf{A} = \mathbf{B}$. Dokážte!

Úloha 7.5. Nájdite nejakú nesymetrickú idempotentnú maticu typu 2×2.

Úloha 7.6. Nájdite ortogonálny projektor na a) priamku $\{c\mathbf{1}_n \in \mathbb{R}^n \mid c \in \mathbb{R}\}$; b) rovinu $\{(c, d, d)^T \in \mathbb{R}^3 \mid c, d \in \mathbb{R}\}$; c) priestor $\mathcal{C}(\mathbf{X})$, kde

$$\mathbf{X} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}.$$

Úloha 7.7. Nech vektory $\mathbf{u}_1, \ldots, \mathbf{u}_n \in \mathbb{R}^n$ tvoria ortonormálnu bázu priestoru $\mathbb{R}^n$, čiže $\mathbf{u}_1, \ldots, \mathbf{u}_n$ je systém navzájom ortogonálnych vektorov normy 1. Nech $k \in \{1, 2, \ldots, n\}$. Dokážte, že $\mathbf{P} = \sum_{i=1}^k \mathbf{u}_i\mathbf{u}_i^T$ je projektor na priestor $\mathcal{C}(\mathbf{u}_1, \ldots, \mathbf{u}_k)$ generovaný vektormi $\mathbf{u}_1, \ldots, \mathbf{u}_k$.

Úloha 7.8. Dokážte tvrdenie: Nech $\mathbf{X} \in \mathbb{R}^{n\times p}$ a $\mathbf{Y} \in \mathbb{R}^{n\times q}$, pričom $\mathcal{C}(\mathbf{Y}) \subseteq \mathcal{C}(\mathbf{X})$. Potom $\mathbf{P}_\mathbf{Y}\mathbf{P}_\mathbf{X} = \mathbf{P}_\mathbf{Y}$.

Úloha 7.9. Nech $\mathbf{x} \in \mathbb{R}^n$. a) Nájdite predpis pre projektor $\mathbf{P}_\mathbf{x}$ zobrazujúci na priestor $\mathcal{C}(\mathbf{x})$ generovaný vektorom $\mathbf{x}$. b) Ukážte, že $\{\mathbf{y} \in \mathbb{R}^2 \mid \mathbf{y} = \mathbf{P}_\mathbf{x}(1, 0)^T$ pre nejaké $\mathbf{x} \in \mathbb{R}^2\}$ je kružnica so stredom v bode $(1/2, 0)^T$ a polomerom $1/2$.

Úloha 7.10. Nech $\mathbf{X} \in \mathbb{R}^{m\times n}$. Nájdite tvar projekčnej matice na $\mathcal{N}(\mathbf{X})$.

Úloha 7.11. Nech $\mathbf{X}_1, \mathbf{X}_2 \in \mathbb{R}^{n\times k}$, pričom $\mathbf{X}_1^T\mathbf{X}_2 = \mathbf{0}$. Ukážte, že potom platí $\mathbf{P}_{\mathbf{X}_1} + \mathbf{P}_{\mathbf{X}_2} + (\mathbf{I}_n - \mathbf{P}_{\mathbf{X}_1})(\mathbf{I}_n - \mathbf{P}_{\mathbf{X}_2}) = \mathbf{I}_n$. Interpretujte toto tvrdenie geometricky pre $n = 3$.

Úloha 7.12. Nech $\mathbf{X} \in \mathbb{R}^{n\times k}$. Ukážte, že matica $\mathbf{X}\mathbf{X}^+$ je symetrická a idempotentná.

Úloha 7.13. Nech $\mathbf{X} \in \mathbb{R}^{n\times k}$ a nech $\mathbf{X} = \mathbf{QR}$ je jej úzky $\mathbf{QR}$ rozklad. Vyjadrite $\mathbf{P}_\mathbf{X}$ pomocou matíc $\mathbf{Q}$ a $\mathbf{R}$.

výška otca	169,7	174,2	185,9	175,4	160,0
výška syna	174,2	170,1	177,4	173,1	164,6

Tabuľka 7.1: Dáta pre úlohu 7.14.

Úloha 7.14. Zmerali sme výšky niekoľkých synov a ich otcov, výsledky sú v tabuľke 7.1.

Výšky synov sme uložili do vektora $\mathbf{y}$ a vytvorili sme maticu

$$\mathbf{X} = \begin{pmatrix} 1 & otec_1 \\ \vdots & \vdots \\ 1 & otec_5 \end{pmatrix} \in \mathbb{R}^{5\times 2},$$

kde $otec_i$ je výška i-teho otca. Nájdite riešenie problému najmenších štvorcov $\|\mathbf{y} - \mathbf{Xb}\|^2 \to \min$. (Táto úloha zodpovedá hľadaniu priamky $syn = b_1 + b_2 \cdot otec$ najlepšie vystihujúcej dané dáta.)

Kapitola 8

Determinant matice

8.1 Definícia a základné vlastnosti determinantu

Poznámka 8.1. Pripomeňme, že permutácia množiny $\{1, \ldots, n\}$ je podľa definície 4.40 bijekciou medzi $\{1, \ldots, n\}$ a $\{1, \ldots, n\}$.

Definícia 8.2 (Inverzie v permutácii a počet inverzií)**.** Nech σ je permutáciou množiny $\{1, \ldots, n\}$. *Inverzia* v permutácii σ je ľubovoľná usporiadaná dvojica $(\sigma(i), \sigma(j))$, $i, j \in \{1, \ldots, n\}$ taká, že $i < j$ a $\sigma(i) > \sigma(j)$. Počet všetkých inverzií v permutácii σ značíme $\Phi(\sigma)$.

Poznámka 8.3. Permutácia σ množiny $\{1, \ldots, n\}$ je vlastne preusporiadanie postupnosti $1, \ldots, n$: 1 je po preusporiadaní na mieste $\sigma(1)$, 2 na mieste $\sigma(2)$ atď., a n je na mieste $\sigma(n)$.

Príklad 8.4. Uvažujme permutáciu $3, 1, 2, 4$ množiny $\{1, 2, 3, 4\}$, teda formálne $\sigma(1) = 3$, $\sigma(2) = 1$, $\sigma(3) = 2$ a $\sigma(4) = 4$. Inverziami v permutácii σ sú dvojice $(3, 1)$ a $(3, 2)$. Nájdite všetky inverzie v permutácii $5, 3, 1, 2, 4$ množiny $\{1, 2, 3, 4, 5\}$.

Lema 8.5 (O počte inverzií v permutácii)**.** Nech σ je permutáciou množiny $\{1, \ldots, n\}$. Definujme p_k ako počet čísel spomedzi $\sigma(k+1), \ldots, \sigma(n)$, ktoré sú menšie ako $\sigma(k)$, pre $k = 1, \ldots, n-1$. Potom $\Phi(\sigma) = p_1 + \ldots + p_{n-1}$.

Dôkaz. Tvrdenie vyplýva z toho, že všetky inverzie rozdelíme na triedy: (i) tie, ktoré obsahujú číslo $\sigma(1)$ a s ním čísla vyššie od $\sigma(1)$, (ii) tie, ktoré obsahujú $\sigma(2)$ a s ním čísla vyššie od $\sigma(2)$, ..., nakoniec tie, ktoré obsahujú $\sigma(n-1)$ a s ním čísla vyššie od $\sigma(n-1)$. $\square$

Definícia 8.6 (Determinant matice)**.** Nech $\mathbf{A} \in \mathbb{R}^{n \times n}$. Potom *determinant* matice $\mathbf{A}$ je

$$\det(\mathbf{A}) = \sum_{\sigma-\text{perm.}} (-1)^{\Phi(\sigma)} (\mathbf{A})_{1,\sigma(1)} \cdots (\mathbf{A})_{n,\sigma(n)}, \tag{8.1}$$

kde suma ide cez všetky permutácie σ množiny $\{1, \ldots, n\}$.

Poznámka 8.7. Determinant matice $\mathbf{A}$ sa tiež niekedy značí $|\mathbf{A}|$. Vzorec (8.1) nazývame *Leibnizova formula* pre determinant.

Poznámka 8.8. Výraz $(-1)^{\Phi(\sigma)}$ nadobúda hodnotu 1 ak σ má párny počet inverzií, a nadobúda hodnotu -1 ak σ má nepárny počet inverzií.

Príklad 8.9. Ukážte, že pre $n = 2$ dostávame v (8.1) známy vzorec

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc.$$

Poznámka 8.10. Pre $n = 3$ platí pre determinant formula

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ - (a_{13}a_{22}a_{31} + a_{11}a_{23}a_{32} + a_{12}a_{21}a_{33}).$$

Túto formulu je možné si zapamätať pomocou *Sarrusovho pravidla*, ktoré hovorí, že keď k 3×3 matici doplníme jej prvé dva stĺpce:

$$\left(\begin{array}{ccc:cc} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{array}\right),$$

tak pripočítavame súčiny trojíc prvkov na "diagonálach" idúcich zľava hore doprava dole, a následne odpočítavame súčiny trojíc prvkov na opačných "diagonálach" idúcich sprava hore doľava dole.

Veta 8.11 (Základné vlastnosti determinantu)**.** Nech $\mathbf{A} \in \mathbb{R}^{n \times n}$. Vyjadrime ju pomocou jej stĺpcov $\mathbf{A} = (\mathbf{a}_1, \ldots, \mathbf{a}_n)$, kde $\mathbf{a}_1, \ldots, \mathbf{a}_n \in \mathbb{R}^n$. Potom:

(i) $\det(\mathbf{A}^T) = \det(\mathbf{A})$.

(ii) $\det(\mathbf{a}_1, \ldots, \mathbf{a}_{k-1}, c\mathbf{a}_k, \mathbf{a}_{k+1}, \ldots, \mathbf{a}_n) = c \det(\mathbf{A})$ pre ľubovoľné $k = 1, \ldots, n$ a $c \in \mathbb{R}$.

(iii) $\det(c\mathbf{A}) = c^n \det(\mathbf{A})$ pre každé $c \in \mathbb{R}$.

(iv) $\mathbf{A}$ je singulárna práve vtedy, keď $\det(\mathbf{A}) = 0$.

(v) Ak matica $\mathbf{B}$ vznikne z $\mathbf{A}$ prehodením dvojice riadkov alebo stĺpcov, tak $\det(\mathbf{B}) = -\det(\mathbf{A})$.

(vi) Ak matica $\mathbf{B}$ vznikne z $\mathbf{A}$ tak, že k niektorému riadku (stĺpcu) pripočítame skalárny násobok iného riadku (stĺpca), tak $\det(\mathbf{B}) = \det(\mathbf{A})$.

(vii) Ak $\mathbf{B} \in \mathbb{R}^{n \times n}$, tak $\det(\mathbf{AB}) = \det(\mathbf{A}) \det(\mathbf{B})$.

Dôkaz. Túto vetu uvádzame bez dôkazu; dôkazy jednotlivých tvrdení čitateľ nájde v [7]. □

Dôsledok 8.12. Nech $\mathbf{A} \in \mathbb{R}^{n \times n}$. Potom podľa vety 8.11(vii) platí $\det(\mathbf{A}^k) = \det(\mathbf{A})^k$ pre každé $k \in \mathbb{N}$.

Veta 8.13 (Determinant dolnej a hornej trojuholníkovej matice)**.** Nech $\mathbf{A} \in \mathbb{R}^{n \times n}$ je dolná alebo horná trojuholníková matica. Potom $\det(\mathbf{A}) = (\mathbf{A})_{11} \cdots (\mathbf{A})_{nn}$.

Dôkaz. Dôkaz vyplýva priamo z Leibnizovej formule – aby pre hornú trojuholníkovú maticu bolo $(\mathbf{A})_{1,\sigma(1)}, \ldots, (\mathbf{A})_{n,\sigma(n)}$ nenulové, musí byť $\sigma(2) \geq 2, \ldots, \sigma(n) \geq n$. To sa dá dosiahnuť iba pre $\sigma(1) = 1, \ldots, \sigma(n) = n$. Pre dolnú trojuholníkovú maticu to platí analogicky. □

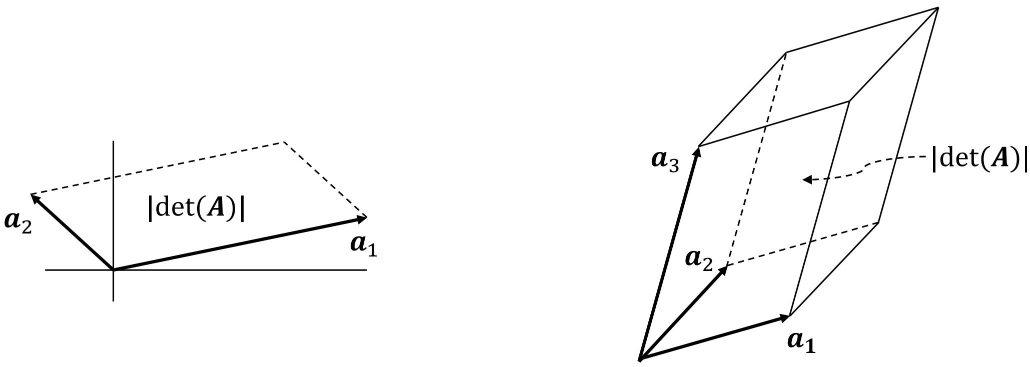
Dôsledok 8.14. Determinant každej diagonálnej matice je súčin jej diagonálnych prvkov. Špeciálne $\det(\mathbf{I}_n) = 1$.

Lema 8.15 (Determinant inverznej matice)**.** Nech $\mathbf{A} \in \mathbb{R}^{n\times n}$ je regulárna. Potom $\det(\mathbf{A}^{-1}) = 1/\det(\mathbf{A})$.

Dôkaz. Keďže $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}_n$, tak podľa vety 8.11 platí $\det(\mathbf{A})\det(\mathbf{A}^{-1}) = \det(\mathbf{I}_n)$. Výsledný tvar získame vďaka tomu, že $\det(\mathbf{I}_n) = 1$ podľa dôsledku 8.14. □

Dôsledok 8.16. Ak $\mathbf{A} \in \mathbb{R}^{n\times n}$ je regulárna, tak tvrdenie $\det(\mathbf{A}^k) = \det(\mathbf{A})^k$ z dôsledku 8.12 môžeme rozšíriť na všetky $k \in \mathbb{Z}$. Pričom pod $\mathbf{A}^{-m}$ pre $m \in \mathbb{N}$ rozumieme $(\mathbf{A}^{-1})^m$.

Poznámka 8.17. Nech $\mathbf{a}_1, \ldots, \mathbf{a}_n \in \mathbb{R}^n$ a označme $\mathbf{A} = (\mathbf{a}_1, \ldots, \mathbf{a}_n) \in \mathbb{R}^{n\times n}$. Potom $|\det(\mathbf{A})|$ je rovný n-rozmernému objemu rovnobežnostena, ktorého jeden vrchol je v počiatku súradnicovej sústavy, a vektory $\mathbf{a}_1, \ldots, \mathbf{a}_n$ určujú hrany vychádzajúce z tohto vrcholu, pozri [14]. Špeciálne pre $n = 2$ je $|\det(\mathbf{A})|$ rovný obsahu rovnobežníka so stranami určenými vektormi $\mathbf{a}_1$ a $\mathbf{a}_2$, ako je znázornené na obrázku 8.1a. Obrázok 8.1b zachytáva vzťah medzi determinantom a objemom rovnobežnostena pre $n = 3$.



(a) $n = 2$ (b) $n = 3$

Obr. 8.1: Obsah rovnobežníka, resp. objem rovnobežnostena, daného zakreslenými stĺpcami matice $\mathbf{A}$ je rovný $|\det(\mathbf{A})|$.

Príklad 8.18. Geometrická interpretácia determinantu z poznámky 8.17 dobre zodpovedá tvrdeniam (ii), (iii), (iv) a (vi) vo vete 8.11. Napríklad ak sledujúc (ii) vynásobíme jeden stĺpec matice $\mathbf{A}$ konštantou c, tak sa objem príslušného rovnobežnostena zjavne zmení na c-násobok, čo je v zhode s príslušným tvrdením. Skúste podobne interpretovať aj ostatné spomínané vlastnosti determinantu.

8.2 Laplaceov rozvoj determinantu a adjungovaná matica

Definícia 8.19 (Minor, kofaktor, adjungovaná matica)**.** Nech $\mathbf{A} \in \mathbb{R}^{n\times n}$ a $i, j \in \{1, \ldots, n\}$. Potom maticu, ktorá vznikne škrtnutím i-teho riadku a j-teho stĺpca matice $\mathbf{A}$, značíme

$\mathbf{A}_{ij}$. *Minor* prvku $(\mathbf{A})_{ij}$ v matici $\mathbf{A}$ je determinant matice $\mathbf{A}_{ij}$. *Kofaktor* prvku $(\mathbf{A})_{ij}$ je $\alpha_{ij} = (-1)^{i+j} \det(\mathbf{A}_{ij})$ a *adjungovaná matica* matice $\mathbf{A}$ je

$$\text{adj}(\mathbf{A}) = \begin{pmatrix} \alpha_{11} & \dots & \alpha_{n1} \\ \vdots & & \vdots \\ \alpha_{1n} & \dots & \alpha_{nn} \end{pmatrix},$$

čiže $\text{adj}(\mathbf{A}) = \{\alpha_{ji}\}_{i,j=1,\dots,n}$.

Poznámka 8.20. Všimnite si, že matica $\text{adj}(\mathbf{A})$ pozostáva z kofaktorov α_{ij} v "prehodenom poradí". Teda $(\text{adj}(\mathbf{A}))_{ij} = \alpha_{ji}$.

Veta 8.21 (Laplaceov rozvoj determinantu)**.** Nech $\mathbf{A} \in \mathbb{R}^{n\times n}$ a $i \in \{1, \dots, n\}$. Potom

$$\det(\mathbf{A}) = \sum_{j=1}^{n} (\mathbf{A})_{ij}\alpha_{ij}.$$

Dôkaz. Tvrdenie dokážeme najprv pre $i = 1$. Označme $a^{(j)}_{k,\ell}$ prvok na pozícii (k, ℓ) v $\mathbf{A}_{1j}$ pre $k, \ell = 1, \dots, n-1$, $j = 1, \dots, n$. Determinant $\mathbf{A}$ vieme vyjadriť ako

$$\begin{aligned}
\det(\mathbf{A}) &= \sum_{\sigma-\text{perm.}} (-1)^{\Phi(\sigma)} (\mathbf{A})_{1,\sigma(1)} \cdots (\mathbf{A})_{n,\sigma(n)} \\
&= (\mathbf{A})_{1,1} \sum_{\sigma(1)=1} (-1)^{\Phi(\sigma)} (\mathbf{A})_{2,\sigma(2)} \cdots (\mathbf{A})_{n,\sigma(n)} \\
&\qquad + (\mathbf{A})_{1,2} \sum_{\sigma(1)=2} (-1)^{\Phi(\sigma)} (\mathbf{A})_{2,\sigma(2)} \cdots (\mathbf{A})_{n,\sigma(n)} + \dots \\
&\qquad + (\mathbf{A})_{1,n} \sum_{\sigma(1)=n} (-1)^{\Phi(\sigma)} (\mathbf{A})_{2,\sigma(2)} \cdots (\mathbf{A})_{n,\sigma(n)}.
\end{aligned}$$

Pre $\sigma(1) = j$ $(j = 1, \dots, n)$ vieme každú $(n-1)$-ticu $(\mathbf{A})_{2,\sigma(2)}, \dots, (\mathbf{A})_{n,\sigma(n)}$ vyjadriť pomocou matice $\mathbf{A}_{1j} \in \mathbb{R}^{(n-1)\times(n-1)}$ a príslušnej permutácie prvkov $1, \dots, n-1$. Označme túto permutáciu π_j, a máme $(\mathbf{A})_{2,\sigma(2)} = a^{(j)}_{1,\pi_j(1)}, \dots, (\mathbf{A})_{n,\sigma(n)} = a^{(j)}_{n-1,\pi_j(n-1)}$. Keďže $\sigma(1) = j$, tak číslo 1 prispelo ku $j-1$ inverziám v σ, čiže $\Phi(\sigma) = \Phi(\pi_j) + j - 1$. Potom môžeme zapísať

$$\begin{aligned}
\det(\mathbf{A}) &= (\mathbf{A})_{1,1} \sum_{\pi_1-\text{perm.}} (-1)^{\Phi(\pi_1)} a^{(1)}_{1,\pi_1(1)} \cdots a^{(1)}_{n,\pi_1(n-1)} \\
&\qquad + (\mathbf{A})_{1,2} \sum_{\pi_2-\text{perm.}} (-1)^{\Phi(\pi_2)+1} a^{(2)}_{1,\pi_2(1)} \cdots a^{(2)}_{n,\pi_2(n-1)} + \dots \\
&\qquad + (\mathbf{A})_{1,n} \sum_{\pi_n-\text{perm.}} (-1)^{\Phi(\pi_n)+n-1} a^{(n)}_{1,\pi_n(1)} \cdots a^{(n)}_{n,\pi_n(n-1)} \\
&= (\mathbf{A})_{1,1} \det(\mathbf{A}_{11}) + (\mathbf{A})_{1,2}(-1)^1 \det(\mathbf{A}_{12}) + \dots + (\mathbf{A})_{1,n}(-1)^{n-1} \det(\mathbf{A}_{1n}) \\
&= \sum_{j=1}^{n} (\mathbf{A})_{1,j}(-1)^{j-1} \det(\mathbf{A}_{1j}) = \sum_{j=1}^{n} (\mathbf{A})_{1,j}(-1)^{j+1} \det(\mathbf{A}_{1j}) = \sum_{j=1}^{n} (\mathbf{A})_{1,j}\alpha_{1j}.
\end{aligned}$$

Pre $i > 1$ zapíšme $\mathbf{A}$ pomocou jej riadkov ako $\mathbf{A} = (\mathbf{c}_1, \ldots, \mathbf{c}_n)^T$ a definujme

$$\mathbf{B} = \begin{pmatrix} \mathbf{c}_i^T \\ \mathbf{c}_1^T \\ \vdots \\ \mathbf{c}_{i-1}^T \\ \mathbf{c}_{i+1}^T \\ \vdots \\ \mathbf{c}_n^T \end{pmatrix}.$$

Potom $(\mathbf{B})_{1,j} = (\mathbf{A})_{i,j}$ a $\mathbf{B}_{1j} = \mathbf{A}_{ij}$. Navyše $\mathbf{A}$ vieme z $\mathbf{B}$ získať postupnými prehodeniami dvojíc riadkov $(1,2), (2,3), \ldots, (i-1, i)$, a teda podľa vety 8.11 platí $\det(\mathbf{A}) = (-1)^{i-1} \det(\mathbf{B})$. Podľa dokázaného tvrdenia pre $i = 1$ dostávame

$$\det(\mathbf{A}) = (-1)^{i-1} \det(\mathbf{B}) = (-1)^{i-1} \sum_{j=1}^{n} (\mathbf{B})_{1,j} (-1)^{1+j} \det(\mathbf{B}_{1j}) = \sum_{j=1}^{n} (\mathbf{A})_{i,j} (-1)^{i+j} \det(\mathbf{A}_{ij}),$$

čo sme chceli dokázať. □

Veta 8.22 (O vzťahu medzi $\mathbf{A}$ a $\mathrm{adj}(\mathbf{A})$)**.** Nech $\mathbf{A} \in \mathbb{R}^{n \times n}$. Potom $\mathbf{A}\,\mathrm{adj}(\mathbf{A}) = \det(\mathbf{A}) \cdot \mathbf{I}_n$.

Dôkaz. Nech $i, k \in \{1, \ldots, n\}$. Prvok matice $\mathbf{A}\,\mathrm{adj}(\mathbf{A})$ na pozícii (i, k) je $\sum_{j=1}^{n} (\mathbf{A})_{ij} \alpha_{kj}$. Z vety 8.21 plynie, že pre $i = k$ je tento výraz rovný $\det(\mathbf{A})$. Pre $i \neq k$ definujme maticu $\mathbf{B} \in \mathbb{R}^{n \times n}$, ktorá je totožná s maticou $\mathbf{A}$, akurát k-ty riadok $\mathbf{B}$ je rovný i-temu riadku $\mathbf{A}$. Keďže $\mathbf{B}$ má dva rovnaké riadky, tak je singulárna a podľa vety 8.11 je $\det(\mathbf{B}) = 0$. Označme kofaktor prvku $(\mathbf{B})_{st}$ ako β_{st}, $s, t = 1, \ldots, n$. Potom $\beta_{kj} = \alpha_{kj}$ a zároveň $(\mathbf{B})_{kj} = (\mathbf{A})_{ij}$ pre $j = 1, \ldots, n$, čím získavame

$$\sum_{j=1}^{n} (\mathbf{A})_{ij} \alpha_{kj} = \sum_{i=1}^{n} (\mathbf{B})_{kj} \beta_{kj} = \det(\mathbf{B}) = 0,$$

podľa vety 8.21. Zistili sme, že diagonálne prvky matice $\mathbf{A}\,\mathrm{adj}(\mathbf{A})$ sú rovné $\det(\mathbf{A})$ a mimodiagonálne prvky sú rovné 0, čo môžeme zapísať ako $\mathbf{A}\,\mathrm{adj}(\mathbf{A}) = \det(\mathbf{A}) \cdot \mathbf{I}_n$. □

Dôsledok 8.23. Ak $\mathbf{A} \in \mathbb{R}^{n \times n}$ je regulárna, tak $\mathbf{A}^{-1} = \det(\mathbf{A})^{-1} \mathrm{adj}(\mathbf{A})$.

Lema 8.24 (Inverzia trojuholníkovej matice s jednotkovou diagonálou)**.** Nech $\mathbf{A} \in \mathbb{R}^{n \times n}$ je dolná (horná) trojuholníková matica s jednotkami na diagonále. Potom aj $\mathbf{A}^{-1}$ je dolná (horná) trojuholníková matica s jednotkami na diagonále.

Dôkaz. Tvrdenie dokážeme pre dolnú trojuholníkovú maticu. Podľa vety 8.13 je $\det(\mathbf{A})$ rovný súčinu jej diagonálnych prvkov, teda $\det(\mathbf{A}) = 1$, čiže $\mathbf{A}$ je regulárna. Inverzná matica má potom podľa vety 8.22 tvar $\mathbf{A}^{-1} = (\det(\mathbf{A}))^{-1} \mathrm{adj}(\mathbf{A}) = \mathrm{adj}(\mathbf{A})$. Matica $\mathbf{A}_{ii}$, ktorá vznikne odstránením i-teho riadku aj stĺpca, je opäť dolná trojuholníková s jednotkovou diagonálou, čiže $\alpha_{ii} = \det(\mathbf{A}_{ii}) = 1$. Nie je ťažké sa presvedčiť, že po odstránení j-teho riadku a i-teho stĺpca pre $i < j$ vzniknutá matica $\mathbf{A}_{ji}$ je tiež dolná trojuholníková, ale na diagonále má aspoň jednu nulu. Preto pre $i < j$ je $\det(\mathbf{A}_{ji}) = 0$, čiže $(\mathrm{adj}(\mathbf{A}))_{ij} = \alpha_{ji} = \det(\mathbf{A}_{ji}) = 0$, čiže $\mathrm{adj}(\mathbf{A})$ je dolná trojuholníková. Spolu s $\alpha_{ii} = 1$ dostávame požadované tvrdenie. □

8.3 Determinanty blokových matíc

Lema 8.25 (Determinanty špeciálnych blokových matíc)**.** Nech $\mathbf{A} \in \mathbb{R}^{m\times m}$, $\mathbf{B} \in \mathbb{R}^{n\times m}$ a $\mathbf{C} \in \mathbb{R}^{m\times n}$. Potom

$$\det\begin{pmatrix}\mathbf{A} & \mathbf{0}_{m\times n}\\ \mathbf{B} & \mathbf{I}_n\end{pmatrix} = \det\begin{pmatrix}\mathbf{I}_n & \mathbf{0}_{n\times m}\\ \mathbf{C} & \mathbf{A}\end{pmatrix} = \det(\mathbf{A}).$$

Dôkaz. Nech

$$\mathbf{X} = \begin{pmatrix}\mathbf{A} & \mathbf{0}_{m\times n}\\ \mathbf{B} & \mathbf{I}_n\end{pmatrix}.$$

Počítajme $\det(\mathbf{X})$ z definície. Zrejme, aby boli všetky členy $(\mathbf{X})_{1,\sigma(1)}, \ldots, (\mathbf{X})_{m+n,\sigma(m+n)}$ nenulové, musí platiť $\sigma(m+1) = m+1, \ldots, \sigma(m+n) = m+n$. Potom $\sigma(1), \ldots, \sigma(m)$ tvoria permutáciu $\{1, \ldots, m\}$, označme ju π, pričom $\Phi(\sigma) = \Phi(\pi)$, lebo prvky $\sigma(m+1) = m+1, \ldots, \sigma(m+n) = m+n$ neprispievajú k počtu inverzií. Teda

$$\begin{aligned}\det(\mathbf{X}) &= \sum_{\sigma-\text{perm.}} (-1)^{\Phi(\sigma)}(\mathbf{X})_{1,\sigma(1)}\cdots(\mathbf{X})_{m+n,\sigma(m+n)}\\ &= \sum_{\pi-\text{perm.}} (-1)^{\Phi(\pi)}(\mathbf{X})_{1,\pi(1)}\cdots(\mathbf{X})_{m,\pi(m)}\cdot(\mathbf{X})_{m+1,m+1}\cdots(\mathbf{X})_{m+n,m+n}\\ &= \sum_{\pi-\text{perm.}} (-1)^{\Phi(\pi)}(\mathbf{A})_{1,\pi(1)}\cdots(\mathbf{A})_{m,\pi(m)}\cdot 1\cdots 1 = \det(\mathbf{A}),\end{aligned}$$

pričom prvá suma ide cez všetky permutácie σ množiny $\{1, \ldots, m+n\}$ a ostatné sumy sú cez všetky permutácie π množiny $\{1, \ldots, m\}$. Dôkaz rovnosti

$$\det\begin{pmatrix}\mathbf{I}_n & \mathbf{0}_{n\times m}\\ \mathbf{C} & \mathbf{A}\end{pmatrix} = \det(\mathbf{A})$$

je analogický, nechávame ho na čitateľa. □

Lema 8.26 (Determinant blokovo trojuholníkovej matice)**.** Nech $\mathbf{T} \in \mathbb{R}^{m\times m}$, $\mathbf{U} \in \mathbb{R}^{m\times n}$, $\mathbf{V} \in \mathbb{R}^{n\times m}$, $\mathbf{W} \in \mathbb{R}^{n\times n}$. Potom

$$\det\begin{pmatrix}\mathbf{T} & \mathbf{U}\\ \mathbf{0}_{n\times m} & \mathbf{W}\end{pmatrix} = \det(\mathbf{T})\det(\mathbf{W}) \quad \text{a} \quad \det\begin{pmatrix}\mathbf{T} & \mathbf{0}_{m\times n}\\ \mathbf{V} & \mathbf{W}\end{pmatrix} = \det(\mathbf{T})\det(\mathbf{W}).$$

Dôkaz. Maticu

$$\mathbf{A} = \begin{pmatrix}\mathbf{T} & \mathbf{U}\\ \mathbf{0}_{n\times m} & \mathbf{W}\end{pmatrix}$$

vieme zapísať ako

$$\mathbf{A} = \begin{pmatrix}\mathbf{I}_m & \mathbf{0}_{m\times n}\\ \mathbf{0}_{n\times m} & \mathbf{W}\end{pmatrix}\begin{pmatrix}\mathbf{T} & \mathbf{U}\\ \mathbf{0}_{n\times m} & \mathbf{I}_n\end{pmatrix}.$$

Podľa vety 8.11 a lemy 8.25 dostávame

$$\det(\mathbf{A}) = \det\begin{pmatrix}\mathbf{I}_m & \mathbf{0}_{m\times n}\\ \mathbf{0}_{n\times m} & \mathbf{W}\end{pmatrix}\det\begin{pmatrix}\mathbf{T} & \mathbf{U}\\ \mathbf{0}_{n\times m} & \mathbf{I}_n\end{pmatrix} = \det(\mathbf{W})\det(\mathbf{T}).$$

Dôkaz pre blokovo dolnú trojuholníkovú maticu je analogický, nechávame ho na čitateľa. Priamo ho tiež vieme skonštruovať pomocou práve dokázaného tvrdenia a uvedomenia si faktu, že transponovanie nemení determinant. □

Veta 8.27 (Determinant blokovej matice)**.** Nech $\mathbf{T} \in \mathbb{R}^{m\times m}, \mathbf{U} \in \mathbb{R}^{m\times n}, \mathbf{V} \in \mathbb{R}^{n\times m}, \mathbf{W} \in \mathbb{R}^{n\times n}$, pričom $\mathbf{T}$ je regulárna. Potom

$$\det\begin{pmatrix}\mathbf{T} & \mathbf{U}\\ \mathbf{V} & \mathbf{W}\end{pmatrix} = \det\begin{pmatrix}\mathbf{W} & \mathbf{V}\\ \mathbf{U} & \mathbf{T}\end{pmatrix} = \det(\mathbf{T})\det(\mathbf{W}-\mathbf{V}\mathbf{T}^{-1}\mathbf{U}).$$

Dôkaz. Maticu

$$\mathbf{A} = \begin{pmatrix}\mathbf{T} & \mathbf{U}\\ \mathbf{V} & \mathbf{W}\end{pmatrix}$$

vieme vyjadriť ako

$$\begin{pmatrix}\mathbf{T} & \mathbf{U}\\ \mathbf{V} & \mathbf{W}\end{pmatrix} = \begin{pmatrix}\mathbf{I}_m & \mathbf{0}_{m\times n}\\ \mathbf{V}\mathbf{T}^{-1} & \mathbf{W}-\mathbf{V}\mathbf{T}^{-1}\mathbf{U}\end{pmatrix}\begin{pmatrix}\mathbf{T} & \mathbf{U}\\ \mathbf{0}_{n\times m} & \mathbf{I}_n\end{pmatrix},$$

z čoho vyplýva

$$\det\begin{pmatrix}\mathbf{T} & \mathbf{U}\\ \mathbf{V} & \mathbf{W}\end{pmatrix} = \det\begin{pmatrix}\mathbf{I}_m & \mathbf{0}_{m\times n}\\ \mathbf{V}\mathbf{T}^{-1} & \mathbf{W}-\mathbf{V}\mathbf{T}^{-1}\mathbf{U}\end{pmatrix}\det\begin{pmatrix}\mathbf{T} & \mathbf{U}\\ \mathbf{0}_{n\times m} & \mathbf{I}_n\end{pmatrix}.$$

Podľa lemy 8.25 máme

$$\det\begin{pmatrix}\mathbf{I}_m & \mathbf{0}_{m\times n}\\ \mathbf{V}\mathbf{T}^{-1} & \mathbf{W}-\mathbf{V}\mathbf{T}^{-1}\mathbf{U}\end{pmatrix} = \det(\mathbf{W}-\mathbf{V}\mathbf{T}^{-1}\mathbf{U}) \quad \text{a} \quad \det\begin{pmatrix}\mathbf{T} & \mathbf{U}\\ \mathbf{0}_{n\times m} & \mathbf{I}_n\end{pmatrix} = \det(\mathbf{T}),$$

čiže $\det(\mathbf{A}) = \det(\mathbf{T})\det(\mathbf{W}-\mathbf{V}\mathbf{T}^{-1}\mathbf{U})$. Druhú rovnosť získame pomocou rozkladu

$$\begin{pmatrix}\mathbf{W} & \mathbf{V}\\ \mathbf{U} & \mathbf{T}\end{pmatrix} = \begin{pmatrix}\mathbf{W}-\mathbf{V}\mathbf{T}^{-1}\mathbf{U} & \mathbf{V}\mathbf{T}^{-1}\\ \mathbf{0}_{m\times n} & \mathbf{I}_m\end{pmatrix}\begin{pmatrix}\mathbf{I}_n & \mathbf{0}_{n\times m}\\ \mathbf{U} & \mathbf{T}\end{pmatrix}.$$

□

8.4 Vandermondove matice

Definícia 8.28 (Vandermondova matica)**.** Nech $x_1, \ldots, x_n \in \mathbb{R}$. *Vandermondova matica* pre $x_1, \ldots, x_n$ je matica $\mathbf{V} \in \mathbb{R}^{n\times n}$ tvaru

$$\mathbf{V} = \begin{pmatrix}1 & x_1 & x_1^2 & \ldots & x_1^{n-1}\\ \vdots & \vdots & \vdots & & \vdots\\ 1 & x_n & x_n^2 & \ldots & x_n^{n-1}\end{pmatrix}.$$

Poznámka 8.29. Vandermondove matice sa používajú pri polynomiálnej interpolácii, t. j. pri hľadaní polynómu, ktorý presne prechádza danými dátami (bodmi $(x_i, y_i) \in \mathbb{R}^2$, $i = 1, \ldots, n$). Všimnite si, že ak by sme vyriešili sústavu rovníc $\mathbf{y} = \mathbf{V}\mathbf{b}$ v premennej $\mathbf{b} \in \mathbb{R}^n$, kde $\mathbf{y} \in \mathbb{R}^n$ a $\mathbf{V}$ je Vandermondova matica pre $x_1, \ldots, x_n$, tak polynóm $b_1 + b_2 x + \ldots + b_n x^{n-1}$ nadobúda v bodoch $x_1, \ldots, x_n$ hodnoty $y_1, \ldots, y_n$. Tieto matice sa tiež používajú v tzv. Reedovych-Solomonovych samoopravných kódoch (angl. *error correction codes*) pre dátové médiá.

Veta 8.30 (Determinant Vandermondovej matice)**.** Nech $x_1, \ldots, x_n \in \mathbb{R}$ a nech $\mathbf{V}_n$ je *Vandermondova matica* pre $x_1, \ldots, x_n$. Potom

$$\det(\mathbf{V}_n) = \prod_{i,j\in\{1,\ldots,n\}, i<j} (x_j - x_i).$$

Dôkaz. Tvrdenie dokážeme matematickou indukciou. Pre $\mathbf{V}_1$ aj $\mathbf{V}_2$ zjavne platí. Nech $n \in \mathbb{N}$. Predpokladajme, že $\mathbf{V}_{n-1} = \prod_{i<j}(x_j - x_i)$ pre $i, j = 1, \ldots, n-1$. Všimnime si, že pre

$$\mathbf{T}_n = \begin{pmatrix} 1 & -x_n & 0 & \ldots & 0 \\ 0 & 1 & -x_n & \ldots & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & \ldots & \ldots & \ldots & 1 \end{pmatrix} \in \mathbb{R}^{n\times n}$$

a $\mathbf{D}_{n-1} = \operatorname{diag}(x_1 - x_n, \ldots, x_{n-1} - x_n)$ platí $\mathbf{V}_n\mathbf{T}_n = \mathbf{A}_n$, kde

$$\mathbf{A}_n = \begin{pmatrix} \mathbf{1}_{n-1} & \mathbf{D}_{n-1}\mathbf{V}_{n-1} \\ 1 & \mathbf{0}_{n-1}^T \end{pmatrix}.$$

Postupným prehodením dvojíc riadkov $(n, n-1), (n-1, n-2), \ldots, (2, 1)$ dostaneme z matice $\mathbf{A}_n$ maticu

$$\begin{pmatrix} 1 & \mathbf{0}_{n-1}^T \\ \mathbf{1}_{n-1} & \mathbf{D}_{n-1}\mathbf{V}_{n-1} \end{pmatrix},$$

ktorej determinant je podľa vety 8.27 rovný $\det(\mathbf{D}_{n-1}\mathbf{V}_{n-1})$. Z vety 8.11 plynie $\det(\mathbf{A}_n) = (-1)^{n-1}\det(\mathbf{D}_{n-1}\mathbf{V}_{n-1})$, a teda $\det(\mathbf{V}_n)\det(\mathbf{T}_n) = \det(\mathbf{A}_n) = (-1)^{n-1}\det(\mathbf{D}_{n-1})\det(\mathbf{V}_{n-1})$. Podľa vety 8.13 máme $\det(\mathbf{T}_n) = 1$, podľa dôsledku 8.14 je $\det(\mathbf{D}_n) = \prod_{i=1}^{n-1}(x_i - x_n)$. S využitím indukčného predpokladu dostávame

$$\begin{aligned}\det(\mathbf{V}_n) &= (-1)^{n-1}\left(\prod_{i=1}^{n-1}(x_i - x_n)\right)\det(\mathbf{V}_{n-1}) = \left(\prod_{i=1}^{n-1}(x_n - x_i)\right)\prod_{i,j\in\{1,\ldots,n-1\}, i<j}(x_j - x_i) \\ &= \prod_{i,j\in\{1,\ldots,n\}, i<j}(x_j - x_i).\end{aligned}$$

□

Poznámka 8.31. Z vety 8.30 vidíme napríklad, že ak sú všetky $x_1, \ldots, x_n$ rôzne, tak $\det(\mathbf{V})$ je nenulový, teda $\mathbf{V}$ je regulárna. Z toho plynie, že pre ľubovoľné $\mathbf{y} \in \mathbb{R}^n$ existuje jednoznačné riešenie $\mathbf{b} = \mathbf{V}^{-1}\mathbf{y}$ systému $\mathbf{y} = \mathbf{V}\mathbf{b}$. Vzhľadom na poznámku 8.29 teda dostávame, že ak $x_1, \ldots, x_n$ sú rôzne, tak pre každé $y_1, \ldots, y_n$ existuje jednoznačne určený polynóm stupňa $n-1$, ktorý prechádza bodmi $(x_1, y_1), \ldots, (x_n, y_n)$.

8.5 Úlohy na precvičenie

Úloha 8.1. Ukážte, že determinant ortogonálnej matice môže byť len 1 alebo -1. Nech $\mathbf{Q} = (\mathbf{q}_1, \mathbf{q}_2) \in \mathbb{R}^{2\times 2}$ je ortogonálna. Nájdite geometrickú charakterizáciu vzťahu vektorov $\mathbf{q}_1, \mathbf{q}_2$, na základe ktorej je možné rozhodnúť, či $\det(\mathbf{Q}) = 1$ alebo $\det(\mathbf{Q}) = -1$.

Úloha 8.2. Ukážte, že determinant idempotentnej matice môže byť len 0 alebo 1. Ukážte, že ak má idempotentná matica determinant 1, tak to nutne musí byť matica identity.

Úloha 8.3. Z definície determinantu ukážte, že

$$\det\begin{pmatrix} a_{11} & a_{12} & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix} = \det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\det\begin{pmatrix} a_{33} & a_{34} \\ a_{43} & a_{44} \end{pmatrix}.$$

Úloha 8.4. Nech $\mathbf{T} \in \mathbb{R}^{m\times m}$ a $\mathbf{W} \in \mathbb{R}^{n\times n}$ sú regulárne matice a nech $\mathbf{U} \in \mathbb{R}^{m\times n}$, $\mathbf{V} \in \mathbb{R}^{n\times m}$. Pomocou vety o determinante blokovej matice ukážte, že $\det(\mathbf{T})\det(\mathbf{W} - \mathbf{V}\mathbf{T}^{-1}\mathbf{U}) = \det(\mathbf{W})\det(\mathbf{T} - \mathbf{U}\mathbf{W}^{-1}\mathbf{V})$.

Úloha 8.5. Použite úlohu 8.4 na dôkaz tvrdení: a) $\det(\mathbf{A} + \mathbf{x}\mathbf{x}^T) = \det(\mathbf{A})(1 + \mathbf{x}^T\mathbf{A}^{-1}\mathbf{x})$ pre akúkoľvek regulárnu maticu $\mathbf{A} \in \mathbb{R}^{n\times n}$ a akýkoľvek vektor $\mathbf{x} \in \mathbb{R}^n$; b) $\det(\mathbf{I}_m + \mathbf{C}\mathbf{D}) = \det(\mathbf{I}_n + \mathbf{D}\mathbf{C})$ pre akékoľvek matice $\mathbf{C} \in \mathbb{R}^{n\times m}, \mathbf{D} \in \mathbb{R}^{m\times n}$.

Úloha 8.6. Nech $a, b \in \mathbb{R}$. Dokážte, že determinant úplne symetrickej matice $\mathbf{A} = a\mathbf{I}_n + b\mathbf{1}_n\mathbf{1}_n^T$ je $\det(\mathbf{A}) = a^{n-1}(a+nb)$. Návod: K prvému riadku matice $\mathbf{A}$ pripočítajte všetky zvyšné riadky a následne odpočítajte prvý stĺpec od všetkých zvyšných stĺpcov.

Úloha 8.7. Nech $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n\times n}$ sú regulárne matice. Dokážte, že $\mathrm{adj}(\mathbf{A}\mathbf{B}) = \mathrm{adj}(\mathbf{B})\mathrm{adj}(\mathbf{A})$. (Toto tvrdenie platí aj bez predpokladu regularity, ale dôkaz je už komplikovanejší.)

Úloha 8.8. Nech $\mathbf{A}, \mathbf{B}, \mathbf{C} \in \mathbb{R}^{m\times n}$ sú nezávislé matice. Pre ktoré reálne čísla k sú matice $k\mathbf{A} + \mathbf{B}, \mathbf{A} + k\mathbf{B} + \mathbf{C}, \mathbf{B} + k\mathbf{C}$ nezávislé? (Môžete použiť vzorec pre determinant matice typu 3×3 a vzťah medzi nenulovým riešením homogénneho systému a determinantom.)

Úloha 8.9. Nech $\mathbf{A} \in \mathbb{R}^{2\times 2}$ je regulárna. Ukážte, že potom platí $\mathbf{A} + \det(\mathbf{A})\mathbf{A}^{-1} = \mathrm{tr}(\mathbf{A})\mathbf{I}_2$.

Úloha* 8.10. Dokážte poznámku 8.17 pre $n = 2$.

Kapitola 9

Pozitívne semidefinitné a pozitívne definitné matice

9.1 Kvadratické formy

Definícia 9.1 (Kvadratická forma)**.** Nech $\mathbf{A} \in \mathcal{S}^n$. Potom funkcia $f : \mathbb{R}^n \to \mathbb{R}$, $f(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x}$ pre každé $\mathbf{x} \in \mathbb{R}^n$, sa nazýva *kvadratická forma.*

Poznámka 9.2. Kvadratická forma je kvadratickou funkciou (t.j. polynómom stupňa 2), keďže $\mathbf{x}^T\mathbf{A}\mathbf{x} = \sum_i (\mathbf{A})_{ii}x_i^2 + 2\sum_i \sum_{j>i}(\mathbf{A})_{ij}x_i x_j$.

Lema 9.3 (Jednoznačnosť kvadratickej formy)**.** Nech $\mathbf{A}, \mathbf{B} \in \mathcal{S}^n$. Potom $\mathbf{x}^T\mathbf{A}\mathbf{x} = \mathbf{x}^T\mathbf{B}\mathbf{x}$ pre každé $\mathbf{x} \in \mathbb{R}^n$ práve vtedy, keď $\mathbf{A} = \mathbf{B}$.

Dôkaz. Zjavne ak $\mathbf{A} = \mathbf{B}$, tak aj $\mathbf{x}^T\mathbf{A}\mathbf{x} = \mathbf{x}^T\mathbf{B}\mathbf{x}$ pre ľubovoľné $\mathbf{x}$. Pre opačnú implikáciu si označme $\mathbf{C} = \mathbf{A} - \mathbf{B}$. Potom $\mathbf{C} \in \mathcal{S}^n$ a $\mathbf{x}^T\mathbf{C}\mathbf{x} = 0$ pre každé $\mathbf{x} \in \mathbb{R}^n$, čiže $\mathbf{C} = \mathbf{0}_{n\times n}$ podľa lemy 3.27. Teda máme $\mathbf{A} - \mathbf{B} = \mathbf{0}_{n\times n}$, z čoho vyplýva $\mathbf{A} = \mathbf{B}$. □

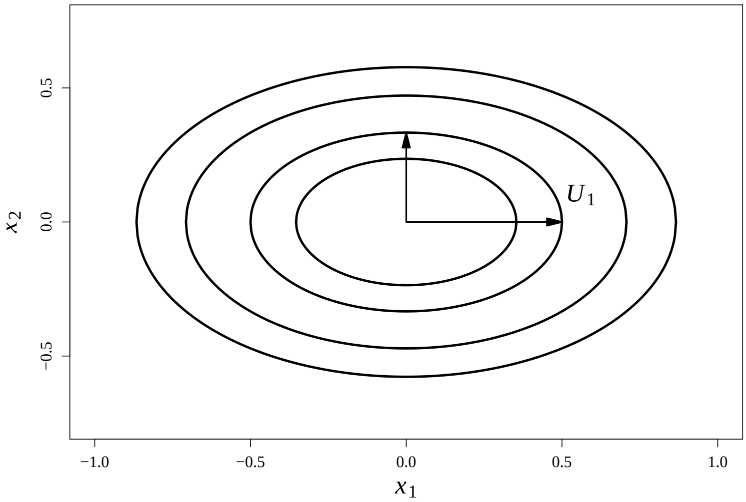
Poznámka 9.4. Z lemy 9.3 plynie, že kvadratická forma $\mathbf{x}^T\mathbf{A}\mathbf{x}$ je jednoznačne charakterizovaná symetrickou maticou $\mathbf{A}$. Maticu $\mathbf{A}$ nazývame *maticou kvadratickej formy* $\mathbf{x}^T\mathbf{A}\mathbf{x}$.

Definícia 9.5 (Pozitívne semidefinitná a pozitívne definitná kvadratická forma)**.** Nech $\mathbf{A} \in \mathcal{S}^n$. Potom kvadratická forma $\mathbf{x}^T\mathbf{A}\mathbf{x}$ sa nazýva *pozitívne semidefinitná* ak $\mathbf{x}^T\mathbf{A}\mathbf{x} \geq 0$ pre každé $\mathbf{x} \in \mathbb{R}^n$. Kvadratická forma $\mathbf{x}^T\mathbf{A}\mathbf{x}$ sa nazýva *pozitívne definitná* ak $\mathbf{x}^T\mathbf{A}\mathbf{x} > 0$ pre každé $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{x} \neq \mathbf{0}_n$.

Poznámka 9.6. Všimnite si, že každá pozitívne definitná kvadratická forma je zároveň pozitívne semidefinitnou.

Príklad 9.7. Overte, že $\mathbf{x}^T\mathbf{I}_n\mathbf{x}$ je pozitívne semidefinitná kvadratická forma a následne overte aj, že je pozitívne definitná. Nech $\mathbf{D} = \mathrm{diag}(d_1, \dots, d_n)$, kde $d_i \in \mathbb{R}$ pre $i = 1, \dots, n$. Určte, pre aké prvky $d_1, \dots, d_n$ je $\mathbf{x}^T\mathbf{D}\mathbf{x}$ pozitívne semidefinitná a pre aké prvky je $\mathbf{x}^T\mathbf{D}\mathbf{x}$ pozitívne definitná kvadratická forma.

Poznámka 9.8. Nech $\mathbf{A} = \mathrm{diag}(4, 9)$. Potom graf kvadratickej formy $\mathbf{x}^T\mathbf{A}\mathbf{x}$ je paraboloid s minimom v bode $\mathbf{x} = \mathbf{0}_2$. Môžeme ju zakresliť pomocou jej "vrstevníc" (úrovňových množín) v $\mathbb{R}^2$, čiže kriviek $U_t = \{\mathbf{x} \in \mathbb{R}^2 \mid \mathbf{x}^T\mathbf{A}\mathbf{x} = t\}$, $t > 0$. Tieto vrstevnice sú elipsy so stredom v počiatku súradnicovej sústavy a s polosami v smere súradnicových osí, ako je zakreslené na obrázku 9.1. Na obrázku môžeme napríklad vidieť, že U_1 je elipsa so stredom v bode $(0,0)^T$ a prechádzajúca bodmi $(0,1/3)^T$, $(1/2,0)^T$ na polosiach. Takéto tvrdenie platí aj vo všeobecnosti: grafy pozitívne definitných kvadratických foriem v $\mathbb{R}^2$ sú paraboloidy a ich vrstevnice sú elipsy, ako ukážeme v poznámke 10.46. Platí dokonca silnejšie tvrdenie: existuje bijekcia medzi pozitívne definitívnymi kvadratickými formami a štandardizovanými elipsoidmi (so stredom v bode $\mathbf{0}$ a "polomerom" 1, teda elipsoidmi typu U_1).



Obr. 9.1: Znázornenie úrovňových množín $U_{1/2}$, U_1, U_2 a U_3 kvadratickej formy zadanej v poznámke 9.8. Šípky znázorňujú polosi elipsy U_1.

9.2 Pozitívne semidefinitné a pozitívne definitné matice

Definícia 9.9 (Pozitívne semidefinitná a pozitívne definitná matica)**.** Symetrická matica $\mathbf{A} \in \mathcal{S}^n$ je

1. *pozitívne semidefinitná* ak $\mathbf{x}^T\mathbf{A}\mathbf{x} \geq 0$ pre každé $\mathbf{x} \in \mathbb{R}^n$,

2. *pozitívne definitná* ak $\mathbf{x}^T\mathbf{A}\mathbf{x} > 0$ pre každé $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{x} \neq \mathbf{0}_n$.

Množinu všetkých pozitívne semidefinitných matíc typu $n \times n$ značíme $\mathcal{S}^n_+$. Množinu všetkých pozitívne definitných matíc typu $n \times n$ značíme $\mathcal{S}^n_{++}$.

Poznámka 9.10. Symetrická matica $\mathbf{A}$ je pozitívne (semi-)definitná, ak príslušná kvadratická forma $\mathbf{x}^T\mathbf{A}\mathbf{x}$ je pozitívne (semi-)definitná.

Poznámka 9.11. Pozitívne semidefinitné (pozitívne definitné) matice sa nazývajú aj *kladne semidefinitné* (*kladne definitné*).

Príklad 9.12. Nech $\mathbf{A} = \mathbf{J}_{2\times 2}$ a nech $\mathbf{x} \in \mathbb{R}^2$. Potom $\mathbf{x}^T\mathbf{A}\mathbf{x} = x_1^2+2x_1x_2+x_2^2 = (x_1+x_2)^2 \geq 0$, a teda $\mathbf{A}$ je pozitívne semidefinitná. Zároveň pre $\mathbf{x} = (1, -1)^T$ platí $\mathbf{x}^T\mathbf{A}\mathbf{x} = (1-1)^2 = 0$, a teda $\mathbf{A}$ nie je pozitívne definitná.

Definujme

$$\mathbf{B} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

Potom $\mathbf{x}^T\mathbf{B}\mathbf{x} = 2x_1^2 + 2x_2^2 + 2x_1x_2 = (x_1 + x_2)^2 + x_1^2 + x_2^2 \geq 0$, čiže aj $\mathbf{B}$ je pozitívne semidefinitná. Zároveň ak $\mathbf{x} \neq \mathbf{0}$, tak aspoň jedno z x_1^2 a x_2^2 je väčšie ako 0, a teda $\mathbf{x}^T\mathbf{B}\mathbf{x} > 0$ pre $\mathbf{x} \neq \mathbf{0}$. Takže $\mathbf{B}$ je pozitívne definitná.

Určte, či je matica

$$\mathbf{C} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$$

pozitívne semidefinitná a či je pozitívne definitná.

Lema 9.13 (O vzťahu $\mathcal{S}^n_{++}$ a $\mathcal{S}^n_+$)**.** Nech $n \in \mathbb{N}$. Potom $\mathcal{S}^n_{++} \subseteq \mathcal{S}^n_+$.

Dôkaz. Nech $\mathbf{A} \in \mathcal{S}^n_{++}$. Potom pre $\mathbf{x} \neq \mathbf{0}_n$ platí $\mathbf{x}^T\mathbf{A}\mathbf{x} > 0$, pričom pre $\mathbf{x} = \mathbf{0}_n$ platí $\mathbf{x}^T\mathbf{A}\mathbf{x} = 0$ pre ľubovoľnú maticu typu $n \times n$. Teda $\mathbf{x}^T\mathbf{A}\mathbf{x} \geq 0$ pre každé $\mathbf{x} \in \mathbb{R}^n$, čiže $\mathbf{A} \in \mathcal{S}^n_+$. □

Lema 9.14 (Pozitívna semidefinitnosť projektorov)**.** Nech $\mathbf{A} \in \mathbb{R}^{n\times n}$ je symetrická a idempotentná. Potom $\mathbf{A} \in \mathcal{S}^n_+$.

Dôkaz. Nech $\mathbf{x} \in \mathbb{R}^n$. Počítajme $\mathbf{x}^T\mathbf{A}\mathbf{x} = \mathbf{x}^T\mathbf{A}\mathbf{A}\mathbf{x} = \mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} = (\mathbf{A}\mathbf{x})^T\mathbf{A}\mathbf{x} = \|\mathbf{A}\mathbf{x}\|^2 \geq 0$. □

Poznámka 9.15. Pripomeňme, že matica je podľa vety 7.24 projekčná práve vtedy, keď je symetrická a idempotentná

Definícia 9.16 ($\mathbf{LDL}^T$ rozklad)**.** Hovoríme, že $\mathbf{A} = \mathbf{LDL}^T$ je $\mathbf{LDL}^T$ *rozklad* matice $\mathbf{A} \in \mathcal{S}^n$, ak $\mathbf{L} \in \mathbb{R}^{n\times n}$ je dolná trojuholníková matica s jednotkami na diagonále a $\mathbf{D} \in \mathbb{R}^{n\times n}$ je diagonálna matica.

Poznámka 9.17. Podobný $\mathbf{LDL}^T$ rozkladu, ale pre všeobecnú $\mathbf{A} \in \mathbb{R}^{n\times n}$, je $\mathbf{LDU}$ rozklad $\mathbf{A} = \mathbf{LDU}$, kde $\mathbf{L} \in \mathbb{R}^{n\times n}$ je dolná trojuholníková s jednotkami na diagonále, $\mathbf{U} \in \mathbb{R}^{n\times n}$ je horná trojuholníková s jednotkami na diagonále a $\mathbf{D} \in \mathbb{R}^{n\times n}$ je diagonálna matica. $\mathbf{LDU}$ rozkladu sa však z priestorových dôvodov v týchto skriptách nebudeme ďalej venovať.

Lema 9.18 (Nula na diagonále pozitívne semidefinitnej matice)**.** Nech $\mathbf{A} \in \mathcal{S}^n_+$ a nech $(\mathbf{A})_{ii} = 0$ pre nejaké $i \in \{1, \ldots, n\}$. Potom $(\mathbf{A})_{ij} = (\mathbf{A})_{ji} = 0$ pre každé $j = 1, \ldots, n$, teda taká matica má nulový celý príslušný riadok aj stĺpec.

Dôkaz. Nech $(\mathbf{A})_{ii} = 0$ a nech $j \neq i$. Vezmime $\mathbf{x} \in \mathbb{R}^n$, ktoré spĺňa $x_i < -(\mathbf{A})_{jj}$, $x_j = 2(\mathbf{A})_{ij}$ a $x_k = 0$ pre $k \neq i, j$. Potom

$$\begin{aligned}\mathbf{x}^T\mathbf{A}\mathbf{x} &= (\mathbf{A})_{ii}x_i^2 + (\mathbf{A})_{jj}x_j^2 + 2(\mathbf{A})_{ij}x_ix_j = 0 + (\mathbf{A})_{jj}4(\mathbf{A})_{ij}^2 + 2(\mathbf{A})_{ij}x_i2(\mathbf{A})_{ij} \\ &= 4(\mathbf{A})_{ij}^2((\mathbf{A})_{jj} + x_i) \leq 0,\end{aligned}$$

lebo $x_i < -(\mathbf{A})_{jj}$. Zároveň však $\mathbf{x}^T\mathbf{A}\mathbf{x} \geq 0$, lebo $\mathbf{A}$ je pozitívne semidefinitná. Preto $\mathbf{x}^T\mathbf{A}\mathbf{x} = 0$, z čoho vyplýva $(\mathbf{A})_{ij}^2 = 0$, čiže $(\mathbf{A})_{ij} = 0$, keďže $(\mathbf{A})_{jj} + x_i < 0$. □

Veta 9.19 (**LDL**^{T} rozklad pozitívne semidefinitnej matice)**.** Pre každú $\mathbf{A} \in \mathcal{S}_+^n$ existuje **LDL**^{T} rozklad $\mathbf{A} = \mathbf{LDL}^T$, pričom diagonálne prvky matice $\mathbf{D}$ sú v každom takom rozklade nezáporné. Navyše, matica $\mathbf{D}$ je jediná a stĺpce $\mathbf{L}$ prislúchajúce nenulovým diagonálnym prvkom matice $\mathbf{D}$ sú tiež určené jednoznačne. Čiže ak $\mathbf{A} = \tilde{\mathbf{L}}\tilde{\mathbf{D}}\tilde{\mathbf{L}}^T$ je ľubovoľný **LDL**^{T} rozklad, tak $\tilde{\mathbf{D}} = \mathbf{D}$ a stĺpce $\tilde{\mathbf{L}}$ prislúchajúce nenulovým prvkom $\mathbf{D}$ sú totožné s príslušnými stĺpcami $\mathbf{L}$.

Dôkaz. Ukážeme najprv, že existuje dolná trojuholníková matica $\mathbf{B} \in \mathbb{R}^{n\times n}$ s jednotkami na diagonále taká, že $\mathbf{BAB}^T$ je diagonálna matica. To budeme dokazovať matematickou indukciou. Tvrdenie zjavne platí pre 1×1 matice. Predpokladajme, že platí pre každú $(n-1)\times(n-1)$ symetrickú maticu, a vezmime $\mathbf{A} \in \mathcal{S}_+^n$. Budeme uvažovať dva prípady:

(i) Ak $(\mathbf{A})_{11} = 0$, potom podľa lemy 9.18 je prvý riadok aj stĺpec matice $\mathbf{A}$ nulový, teda $\mathbf{A}$ má tvar $\mathbf{A} = \operatorname{diag}(0, \mathbf{A}_2)$ pre nejakú $\mathbf{A}_2$ typu $(n-1) \times (n-1)$. Keďže $\mathbf{A}$ je symetrická a $\mathbf{A}^T = \operatorname{diag}(0, \mathbf{A}_2^T)$, tak aj $\mathbf{A}_2$ je symetrická. Zároveň, ak by $\mathbf{x}^T\mathbf{A}_2\mathbf{x} < 0$ pre nejaké $\mathbf{x} \in \mathbb{R}^{n-1}$, tak $(0, \mathbf{x}^T)\mathbf{A}(0, \mathbf{x}^T)^T = \mathbf{x}^T\mathbf{A}_2\mathbf{x} < 0$, čiže $\mathbf{A}$ by nemohla byť pozitívne semidefinitná. Preto $\mathbf{x}^T\mathbf{A}_2\mathbf{x} \geq 0$ pre každé $\mathbf{x} \in \mathbb{R}^{n-1}$, čiže $\mathbf{A}_2$ je pozitívne semidefinitná. Potom podľa indukčného predpokladu existujú dolná trojuholníková $\mathbf{B}_2$ s jednotkami na diagonále a diagonálna $\mathbf{D}_2$, že $\mathbf{B}_2\mathbf{A}_2\mathbf{B}_2^T = \mathbf{D}_2$. Potom pre $\mathbf{B} = \operatorname{diag}(1, \mathbf{B}_2)$ platí

$$\mathbf{BAB}^T = \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{B}_2 \end{pmatrix} \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{A}_2 \end{pmatrix} \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{B}_2^T \end{pmatrix} = \begin{pmatrix} 0 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{B}_2\mathbf{A}_2\mathbf{B}_2^T \end{pmatrix} = \begin{pmatrix} 0 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{D}_2 \end{pmatrix}.$$

Čiže $\mathbf{BAB}^T$ je diagonálna matica, pričom $\mathbf{B}$ je dolná trojuholníková matica s jednotkami na diagonále.

(ii) Ak $(\mathbf{A})_{11} \neq 0$, zapíšme $\mathbf{A}$ v tvare

$$\mathbf{A} = \begin{pmatrix} (\mathbf{A})_{11} & \mathbf{a}^T \\ \mathbf{a} & \mathbf{A}_2 \end{pmatrix}.$$

Potom pre

$$\mathbf{B}_1 = \begin{pmatrix} 1 & \mathbf{0}^T \\ -(\mathbf{A})_{11}^{-1}\mathbf{a} & \mathbf{I}_{n-1} \end{pmatrix}$$

sa ľahko dá vypočítať, že

$$\mathbf{B}_1\mathbf{AB}_1^T = \begin{pmatrix} (\mathbf{A})_{11} & \mathbf{0}^T \\ \mathbf{0} & \mathbf{A}_2 - (\mathbf{A})_{11}^{-1}\mathbf{aa}^T \end{pmatrix},$$

čo môžeme zapísať ako $\mathbf{B}_1\mathbf{AB}_1^T = \operatorname{diag}((\mathbf{A})_{11}, \mathbf{C}_2)$, kde $\mathbf{C}_2 \in \mathbb{R}^{(n-1)\times(n-1)}$. Keďže $\mathbf{A}$ je symetrická, tak $(\mathbf{B}_1\mathbf{AB}_1^T)^T = \mathbf{B}_1\mathbf{AB}_1^T$. Potom podobným argumentom ako v (i) dostávame, že $\mathbf{C}_2 \in \mathcal{S}^{n-1}$. Keďže $\mathbf{A} \in \mathcal{S}_+^n$, tak aj $\mathbf{B}_1\mathbf{AB}_1^T \in \mathcal{S}_+^n$ (dôkaz necháme na čitateľa, pozri úlohu 9.3). Potom analogicky ako pre $\mathbf{A}_2$ v časti (i) platí, že $\mathbf{C}_2$ je pozitívne semidefinitná. Teda pre $\mathbf{C}_2$ vzhľadom na indukčný predpoklad existuje dolná trojuholníková $\mathbf{B}_2 \in \mathbb{R}^{(n-1)\times(n-1)}$ s jednotkami na diagonále a diagonálna $\mathbf{D}_2$, že $\mathbf{B}_2\mathbf{C}_2\mathbf{B}_2^T = \mathbf{D}_2$. Potom pre $\mathbf{B} = \operatorname{diag}(1, \mathbf{B}_2)\mathbf{B}_1$ dostávame

$$\begin{aligned}\mathbf{BAB}^T &= \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{B}_2 \end{pmatrix} \mathbf{B}_1\mathbf{AB}_1^T \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{B}_2^T \end{pmatrix} = \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{B}_2 \end{pmatrix} \begin{pmatrix} (\mathbf{A})_{11} & \mathbf{0}^T \\ \mathbf{0} & \mathbf{C}_2 \end{pmatrix} \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{B}_2^T \end{pmatrix} \\ &= \begin{pmatrix} (\mathbf{A})_{11} & \mathbf{0}^T \\ \mathbf{0} & \mathbf{D}_2 \end{pmatrix},\end{aligned}$$

čo je diagonálna matica, pričom v poslednom kroku sme použili $\mathbf{B}_2\mathbf{C}_2\mathbf{B}_2^T = \mathbf{D}_2$. Zároveň platí, že $\mathbf{B}$ je dolná trojuholníková matica (lebo je súčinom dvoch dolných trojuholníkových matíc, pozri lemu 2.24) a má jednotky na diagonále (lebo $\mathrm{diag}(1, \mathbf{B}_2)$ aj $\mathbf{B}_1$ majú jednotky na diagonále).

Keďže v každom prípade sme našli diagonálnu maticu $\mathbf{D}$ a dolnú trojuholníkovú $\mathbf{B}$ s jednotkovou diagonálou také, že $\mathbf{BAB}^T = \mathbf{D}$, tak $\mathbf{A} = \mathbf{B}^{-1}\mathbf{D}(\mathbf{B}^{-1})^T$. Zároveň, matica $\mathbf{B}^{-1}$ je podľa lemy 8.24 dolná trojuholníková a má na diagonále tiež samé jednotky, čiže je to hľadaná matica $\mathbf{L}$. Navyše, keďže $\mathbf{A} \in \mathcal{S}_+^n$, tak aj $\mathbf{D} = \mathbf{BAB}^T \in \mathcal{S}_+^n$ (opäť pozri úlohu 9.3). Aby diagonálna matica $\mathbf{D}$ bola pozitívne semidefinitná, musí mať nezápornú diagonálu, čím je dôkaz prvej časti hotový.

Uvedenú jedinečnosť $\mathbf{LDL}^T$ rozkladu nebudeme dokazovať; dôkaz môže čitateľ nájsť napríklad v [4] (veta 14.5.5). □

Poznámka 9.20. Všimnite si, že matica $\mathbf{C}_2$ v dôkaze vety 9.19 je vlastne Schurov doplnok prvku $(\mathbf{A})_{11}$ v matici $\mathbf{B}_1\mathbf{AB}_1^T$.

Veta 9.21 (Choleského rozklad)**.** Nech $\mathbf{A} \in \mathcal{S}_+^n$ má hodnosť r. Potom existuje jediná horná trojuholníková matica $\mathbf{T}$ s r kladnými zložkami na diagonále a s $(n-r)$ nulovými riadkami taká, že $\mathbf{A} = \mathbf{T}^T\mathbf{T}$. Navyše, ak $\mathbf{A} = \mathbf{LDL}^T$ je $\mathbf{LDL}^T$ rozklad, kde $\mathbf{D} = \mathrm{diag}(d_1, \dots, d_n)$, potom $\mathbf{T} = \mathbf{D}^{1/2}\mathbf{L}^T$, kde $\mathbf{D}^{1/2} = \mathrm{diag}(\sqrt{d_1}, \dots, \sqrt{d_n})$.

Dôkaz. Podľa vety 9.19 existuje $\mathbf{LDL}^T$ rozklad $\mathbf{A} = \mathbf{LDL}^T$, pričom $\mathbf{D}$ má nezáporné prvky a je jediná. Podľa dôsledku 4.19 je r prvkov $\mathbf{D}$ nenulových, čiže kladných. Matica $\mathbf{T} = \mathbf{D}^{1/2}\mathbf{L}^T$ je horná trojuholníková so zložkami $\sqrt{d_1}, \dots, \sqrt{d_n}$ na diagonále, z ktorých je r kladných. Zároveň má $(n-r)$ nulových riadkov, ktoré prislúchajú nulovým zložkám diagonály matice $\mathbf{D}^{1/2}$. Maticu $\mathbf{A}$ môžeme vyjadriť ako $\mathbf{A} = \mathbf{LD}^{1/2}\mathbf{D}^{1/2}\mathbf{L}^T = \mathbf{T}^T\mathbf{T}$, takže hľadaný rozklad existuje.

Na dôkaz jedinečnosti uvážme $\mathbf{LDL}^T$ rozklad $\mathbf{A} = \mathbf{LDL}^T$ a predpokladajme, že $\mathbf{A} = \mathbf{T}^T\mathbf{T}$, kde $\mathbf{T}$ je horná trojuholníková matica s r kladnými zložkami na diagonále a s $(n-r)$ nulovými riadkami. Pre jednoduchosť predpokladajme, že prvých r diagonálnych zložiek $\mathbf{T}$ je kladných, a označme riadky $\mathbf{T}$ ako $\mathbf{t}_1^T, \dots, \mathbf{t}_n^T$. Definujme $\tilde{\mathbf{D}} = \mathrm{diag}((\mathbf{T})_{11}^2, \dots, (\mathbf{T})_{nn}^2) = \mathrm{diag}((\mathbf{T})_{11}^2, \dots, (\mathbf{T})_{rr}^2, 0, \dots, 0)$, a $\tilde{\mathbf{L}}$ nech je ľubovoľná dolná trojuholníková matica s prvými r stĺpcami $(1/(\mathbf{T})_{11})\mathbf{t}_1, \dots, (1/(\mathbf{T})_{rr})\mathbf{t}_r$. Potom $\tilde{\mathbf{D}}^{1/2}\tilde{\mathbf{L}}^T = \mathbf{T}$, teda $\mathbf{A} = \mathbf{T}^T\mathbf{T} = \tilde{\mathbf{L}}\tilde{\mathbf{D}}\tilde{\mathbf{L}}^T$. Čiže $\tilde{\mathbf{L}}\tilde{\mathbf{D}}\tilde{\mathbf{L}}^T$ je tiež $\mathbf{LDL}^T$ rozklad matice $\mathbf{A}$. Podľa vety 9.19 je $\mathbf{D}$ jediná a prvých r stĺpcov $\mathbf{L}$ je určených jednoznačne, čiže $\tilde{\mathbf{D}} = \mathbf{D}$ a prvých r stĺpcov $\tilde{\mathbf{L}}$ je zhodných s príslušnými stĺpcami $\mathbf{L}$. Potom $\mathbf{T} = \tilde{\mathbf{D}}^{1/2}\tilde{\mathbf{L}}^T = \mathbf{D}^{1/2}\mathbf{L}^T$, z čoho vyplýva, že jediná horná trojuholníková matica $\mathbf{T}$ s r kladnými zložkami na diagonále a s $(n-r)$ nulovými riadkami, spĺňajúca $\mathbf{A} = \mathbf{T}^T\mathbf{T}$, je $\mathbf{T} = \mathbf{D}^{1/2}\mathbf{L}^T$. □

Veta 9.22 (Základné vlastnosti $\mathcal{S}_+^n$ a $\mathcal{S}_{++}^n$)**.** Nech $n \in \mathbb{N}$. Potom:

(i) Ak $\mathbf{A}, \mathbf{B} \in \mathcal{S}_+^n$, $\alpha \geq 0$, $\beta \geq 0$, potom $\alpha\mathbf{A} + \beta\mathbf{B} \in \mathcal{S}_+^n$.

(ii) Ak $\mathbf{A} \in \mathcal{S}_{++}^n$, $\mathbf{B} \in \mathcal{S}_+^n$, $\alpha > 0$, $\beta \geq 0$, potom $\alpha\mathbf{A} + \beta\mathbf{B} \in \mathcal{S}_{++}^n$.

(iii) $\mathbf{A} \in \mathcal{S}_{++}^n$ práve vtedy, keď $\mathbf{A} \in \mathcal{S}_+^n$ a $\mathbf{A}$ je regulárna.

(iv) $\mathbf{A} \in \mathcal{S}_{++}^n$ práve vtedy, keď $\mathbf{A}^{-1} \in \mathcal{S}_{++}^n$.

Dôkaz. V (i) platí $\mathbf{x}^T(\alpha\mathbf{A}+\beta\mathbf{B})\mathbf{x} = \alpha\mathbf{x}^T\mathbf{A}\mathbf{x}+\beta\mathbf{x}^T\mathbf{B}\mathbf{x} \geq 0$ pre každé $\mathbf{x} \in \mathbb{R}^n$. Podobne pre (ii) nech $\mathbf{x} \neq \mathbf{0}_n$. Potom $\mathbf{x}^T(\alpha\mathbf{A}+\beta\mathbf{B})\mathbf{x} = \alpha\mathbf{x}^T\mathbf{A}\mathbf{x}+\beta\mathbf{x}^T\mathbf{B}\mathbf{x} > 0$, lebo $\alpha\mathbf{x}^T\mathbf{A}\mathbf{x} > 0$ a $\beta\mathbf{x}^T\mathbf{B}\mathbf{x} \geq 0$.

Pre dôkaz (iii) najprv predpokladajme, že $\mathbf{A} \in \mathcal{S}^n_{++}$. Potom $\mathbf{A} \in \mathcal{S}^n_+$ podľa lemy 9.13 a teda stačí dokázať už len, že je regulárna. Pre dôkaz sporom uvažujme, že $\mathbf{A}$ je singulárna, čiže existuje $\mathbf{x} \neq \mathbf{0}_n$, že $\mathbf{A}\mathbf{x} = \mathbf{0}_n$. Potom $\mathbf{x}^T\mathbf{A}\mathbf{x} = 0$, čo je spor s pozitívnou definitnosťou. Naopak, nech $\mathbf{A} \in \mathcal{S}^n_+$ je regulárna. Potom podľa vety 9.21 máme $\mathbf{A} = \mathbf{T}^T\mathbf{T}$, pričom $\operatorname{rank}(\mathbf{T}) = \operatorname{rank}(\mathbf{A}) = n$ podľa dôsledku 3.35, čiže $\mathbf{T}$ je regulárna. Teda pre $\mathbf{x} \neq \mathbf{0}_n$ platí $\mathbf{x}^T\mathbf{A}\mathbf{x} = \mathbf{x}^T\mathbf{T}^T\mathbf{T}\mathbf{x} = \|\mathbf{T}\mathbf{x}\|^2 \geq 0$. Keďže $\mathbf{x} \neq \mathbf{0}_n$ a $\mathbf{T}$ je regulárna, tak $\mathbf{T}\mathbf{x} \neq \mathbf{0}_n$, čiže $\mathbf{x}^T\mathbf{A}\mathbf{x} = \|\mathbf{T}\mathbf{x}\|^2 > 0$.

Pre (iv) nech $\mathbf{A} \in \mathcal{S}^n_{++}$. Potom $\mathbf{A} = \mathbf{T}^T\mathbf{T}$, kde $\mathbf{T}$ je regulárna, analogicky k (iii). Z toho vyplýva $\mathbf{A}^{-1} = \mathbf{T}^{-1}\mathbf{T}^{-T}$. Pre $\mathbf{x} \neq \mathbf{0}_n$ teda platí $\mathbf{x}^T\mathbf{A}^{-1}\mathbf{x} = \mathbf{x}^T\mathbf{T}^{-1}\mathbf{T}^{-T}\mathbf{x} = (\mathbf{T}^{-T}\mathbf{x})^T\mathbf{T}^{-T}\mathbf{x} = \|\mathbf{T}^{-T}\mathbf{x}\|^2 \geq 0$. Keďže $\mathbf{T}^{-T}$ je regulárna a $\mathbf{x} \neq \mathbf{0}_n$, tak $\mathbf{T}^{-T}\mathbf{x} \neq \mathbf{0}_n$, čiže $\|\mathbf{T}^{-T}\mathbf{x}\|^2 > 0$. Naopak, nech $\mathbf{A}^{-1} \in \mathcal{S}^n_{++}$. Potom $\mathbf{A} = (\mathbf{A}^{-1})^{-1}$ a podľa predchádzajúcej implikácie platí $\mathbf{A} = (\mathbf{A}^{-1})^{-1} \in \mathcal{S}^n_{++}$. □

Definícia 9.23 (Loewnerovo usporiadanie)**.** *Loewnerovo usporiadanie* pozitívne semidefinitných matíc je definované nasledovne: ak $\mathbf{A}, \mathbf{B} \in \mathcal{S}^n_+$, potom $\mathbf{A} \succeq \mathbf{B}$ ak $\mathbf{A} - \mathbf{B} \in \mathcal{S}^n_+$.

Poznámka 9.24. Loewnerovo usporiadanie je iba čiastočné, teda nevieme porovnať všetky $\mathbf{A}, \mathbf{B} \in \mathcal{S}^n_+$, ako ukazuje príklad 9.25.

Príklad 9.25. Nech

$$\mathbf{A} = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}.$$

Potom $\mathbf{A} - \mathbf{B} = \mathbf{I}_2 \in \mathcal{S}^2_+$, čiže $\mathbf{A} \succeq \mathbf{B}$.

Nech

$$\mathbf{C} = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}, \quad \mathbf{D} = \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}.$$

Potom $\mathbf{C} - \mathbf{D} = \operatorname{diag}(-1, 1) \notin \mathcal{S}^2_+$ a $\mathbf{D} - \mathbf{C} = \operatorname{diag}(1, -1) \notin \mathcal{S}^2_+$, čiže neplatí ani $\mathbf{A} \succeq \mathbf{B}$, ani $\mathbf{B} \succeq \mathbf{A}$.

Nájdite príklady nediagonálnych matíc $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$, že $\mathbf{A} \succeq \mathbf{B}$ a matice $\mathbf{C}$ a $\mathbf{D}$ sa nedajú pomocou Loewnerovho usporiadania porovnať.

9.2.1 Pozitívna definitnosť výberovej kovariančnej matice

Skúmajme pozitívnu (semi)definitnosť výberovej kovariančnej matice $\mathbf{S} \in \mathbb{R}^{d\times d}$ z podkapitoly 2.6.1. Pripomeňme, že $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_n) \in \mathbb{R}^{d\times n}$ je maticou dát. Podľa príkladu 2.57 vieme $\mathbf{S}$ zapísať ako $\frac{1}{n-1}(\mathbf{X}-n^{-1}\mathbf{X}\mathbf{J}_{n\times n})(\mathbf{X}-n^{-1}\mathbf{X}\mathbf{J}_{n\times n})^T$, z čoho podľa úlohy 9.5 vidíme, že výberová kovariančná matica je vždy pozitívne semidefinitná. To je jeden z hlavných dôvodov, prečo sa v štatistike veľmi často pracuje s pozitívne semidefinitnými a pozitívne definitnými maticami.

Výberová kovariančná matica je potom vzhľadom na vetu 9.22 pozitívne definitná práve vtedy, keď je regulárna. To nastáva podľa poznámky 3.38 práve vtedy, keď má matica $\mathbf{X}(\mathbf{I}_n - n^{-1}\mathbf{J}_{n\times n}) \in \mathbb{R}^{d\times n}$ hodnosť d.

Príklad 9.26. Overte, že matica $n^{-1}\mathbf{J}_{n\times n}$ je maticou ortogonálnej projekcie na $\operatorname{span}(\mathbf{1}_n)$.

Podľa príkladu 9.26 môžeme výberovú kovariančnú maticu zapísať ako

$$\mathbf{S} = \frac{1}{n-1}(\mathbf{X} - \mathbf{X}\mathbf{P_1})(\mathbf{X} - \mathbf{X}\mathbf{P_1})^T = \frac{1}{n-1}\mathbf{X}(\mathbf{I}_n - \mathbf{P_1})[\mathbf{X}(\mathbf{I}_n - \mathbf{P_1})]^T,$$

kde $\mathbf{P_1}$ je maticou projekcie na $\mathrm{span}(\mathbf{1}_n)$. Z vety 7.25 potom vyplýva, že

$$\mathbf{S} = \frac{1}{n-1}(\mathbf{X}\mathbf{P}_2)(\mathbf{X}\mathbf{P}_2)^T,$$

kde $\mathbf{P}_2$ je projekčnou maticou na $\mathbf{1}_n^{\perp}$. Teda kovariančná matica je regulárna práve vtedy, keď $\mathrm{rank}(\mathbf{X}\mathbf{P}_2) = d$.

Všimnime si, že

$$\mathbf{X}\mathbf{P}_2 = \begin{pmatrix} \mathbf{v}_1^T \\ \vdots \\ \mathbf{v}_d^T \end{pmatrix} \mathbf{P}_2 = \begin{pmatrix} \mathbf{v}_1^T\mathbf{P}_2 \\ \vdots \\ \mathbf{v}_d^T\mathbf{P}_2 \end{pmatrix},$$

kde $\mathbf{v}_1^T, \ldots, \mathbf{v}_n^T$ sú riadky $\mathbf{X}$. Čiže $\mathbf{v}_i^T$ je riadkový vektor dĺžky n zachytávajúci hodnoty i-tej premennej pre jednotlivé pozorovania. Potom $\mathbf{v}_i^T\mathbf{P}_2$ je riadkový vektor hodnôt i-tej premennej sprojektovaný do podpriestoru kolmého na $\mathbf{1}_n$. Takže výberová kovariančná matica je pozitívne definitná práve vtedy, keď matica $\mathbf{X}\mathbf{P}_2 \in \mathbb{R}^{d\times n}$ pozorovaní, ktorých zložky sú sprojektované do podpriestoru kolmého na $\mathbf{1}_n$, má hodnosť d.

Z pohľadu dát majú riadky $\mathbf{X}\mathbf{P}_2$ priamočiaru interpretáciu. Matica $\mathbf{P}_2$ sa dá totiž spätne zapísať ako $\mathbf{P}_2 = \mathbf{I}_n - n^{-1}\mathbf{J}_{n\times n} = \mathbf{I}_n - n^{-1}\mathbf{1}_n\mathbf{1}_n^T$, a teda $\mathbf{X}\mathbf{P}_2 = \mathbf{X} - n^{-1}\mathbf{X}\mathbf{1}_n\mathbf{1}_n^T = \mathbf{X} - \bar{\mathbf{x}}\mathbf{1}_n^T$. Teda i-ty riadok tejto matice je $\mathbf{v}_i^T - \bar{\mathbf{x}}_i\mathbf{1}_n = (x_{i1} - \bar{\mathbf{x}}_i, \ldots, x_{in} - \bar{\mathbf{x}}_i)$, kde $x_{i1}, \ldots, x_{in}$ sú hodnoty i-tej premennej v jednotlivých pozorovaniach a $\bar{\mathbf{x}}_i$ je ich priemer. Takže i-ty riadok $\mathbf{X}\mathbf{P}_2$ je vektorom hodnôt i-tej premennej očistených o ich priemer, tzv. centrovaných hodnôt (pozri príklad 2.55), lebo priemer týchto očistených dát je nula. Výberová kovariančná matica je teda pozitívne definitná, ak matica pozorovaní, ktoré sú v jednotlivých premenných očistené o priemery, (matica centrovaných pozorovaní) $\mathbf{X}\mathbf{P}_2 \in \mathbb{R}^{d\times n}$ má plnú riadkovú hodnosť d.

Poznamenajme, že geometrický sa dá toto tvrdenie sformulovať nasledovne: výberová kovariančná matica je pozitívne definitná práve vtedy, keď dáta neležia v spoločnej afinnej množine (t.j. posunutom vektorovom podpriestore), ktorá má dimenziu nižšiu ako d.

9.3 Úlohy na precvičenie

Úloha 9.1. Presvedčte sa, že matica

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

je pozitívne definitná (z definície pozitívnej definitnosti).

Úloha 9.2. Nech

$$\mathbf{A} = \begin{pmatrix} \mathbf{T} & \mathbf{U} \\ \mathbf{U}^T & \mathbf{W} \end{pmatrix},$$

kde $\mathbf{T} \in \mathbb{R}^{m\times m}$, $\mathbf{U} \in \mathbb{R}^{m\times n}$ a $\mathbf{W} \in \mathbb{R}^{n\times n}$. Ukážte, že ak $\mathbf{A}$ je pozitívne semidefinitná, tak aj $\mathbf{T}$ a $\mathbf{W}$ sú pozitívne semidefinitné. Sformulujte a dokážte podobné tvrdenie pre pozitívnu definitnosť.

Úloha 9.3. Nech $\mathbf{A} \in \mathcal{S}^n$ a $\mathbf{B} \in \mathbb{R}^{n\times m}$. Ukážte, že potom

(i) ak $\mathbf{A} \in \mathcal{S}^n_+$, tak $\mathbf{B}^T\mathbf{A}\mathbf{B} \in \mathcal{S}^m_+$,

(ii) ak $\mathbf{A} \in \mathcal{S}^n_{++}$ a $\operatorname{rank}(\mathbf{B}) = m$, tak $\mathbf{B}^T\mathbf{A}\mathbf{B} \in \mathcal{S}^m_{++}$.

Návod: v časti (ii) môžete použiť lemu 3.36.

Úloha 9.4. Nech $\mathbf{T}$ v matici $\mathbf{A}$ z úlohy 9.2 je regulárna, a nech táto matica $\mathbf{A}$ je pozitívne semidefinitná. Ukážte, že potom Schurov doplnok matice $\mathbf{T}$ v matici $\mathbf{A}$, t.j. matica $\mathbf{S} = \mathbf{W} - \mathbf{U}^T\mathbf{T}^{-1}\mathbf{U}$, je pozitívne semidefinitná matica. Návod: využijúc úlohu 9.3 skúmajte maticu $\mathbf{B}^T\mathbf{A}\mathbf{B}$, kde

$$\mathbf{B} = \begin{pmatrix} \mathbf{I}_m & -\mathbf{T}^{-1}\mathbf{U} \\ \mathbf{0}_{n\times m} & \mathbf{I}_n \end{pmatrix}.$$

Úloha 9.5. Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$. Ukážte, že $\mathbf{A}^T\mathbf{A} \in \mathcal{S}^n_+$. Ukážte tiež, že $\mathbf{A}^T\mathbf{A} \in \mathcal{S}^n_{++}$ práve vtedy, keď $\operatorname{rank}(\mathbf{A}) = n$. Sformulujte podobné tvrdenie pre $\mathbf{A}\mathbf{A}^T$.

Úloha 9.6. Nech $\mathbf{x}_1, \ldots, \mathbf{x}_k \in \mathbb{R}^n$ a definujme $\mathbf{A} = \sum_{i=1}^k \mathbf{x}_i\mathbf{x}_i^T$. Ukážte, že $\mathbf{A} \in \mathcal{S}^n_+$ a dokážte, že hodnosť matice $\mathbf{A}$ je rovná počtu najväčšej množiny lineárne nezávislých vektorov spomedzi $\mathbf{x}_1, \ldots, \mathbf{x}_k$, teda $\operatorname{rank}(\mathbf{A}) = \dim(\operatorname{span}(\mathbf{x}_1, \ldots, \mathbf{x}_k))$.

Úloha 9.7. Ukážte, že ak $\mathbf{A} = \mathbf{L}\mathbf{D}\mathbf{L}^T$ je $\mathbf{L}\mathbf{D}\mathbf{L}^T$ rozklad symetrickej matice $\mathbf{A} \in \mathbb{R}^{n\times n}$ a ak $\mathbf{D}$ je diagonálna matica s kladnými (nezápornými) prvkami na diagonále, tak $\mathbf{A}$ je pozitívne definitná (pozitívne semidefinitná).

Úloha 9.8. Nech $\mathbf{A}, \mathbf{B} \in \mathcal{S}^n_+$ spĺňajú $\mathbf{A} \succeq \mathbf{B}$. Ukážte, že potom $\det(\mathbf{A}) \geq \det(\mathbf{B})$ a $\mathcal{C}(\mathbf{B}) \subseteq \mathcal{C}(\mathbf{A})$.

Kapitola 10

Vlastné čísla a vlastné vektory

10.1 Definície a základné vlastnosti

Definícia 10.1 (Vlastné číslo a vlastný vektor)**.** Nech $\mathbf{A} \in \mathbb{R}^{n\times n}$. Ak $\lambda \in \mathbb{R}$, $\mathbf{x} \in \mathbb{R}^n$ spĺňajú $\mathbf{x} \neq \mathbf{0}_n$ a $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$, tak λ nazývame *vlastné číslo* matice $\mathbf{A}$ a $\mathbf{x}$ nazývame *vlastný vektor* matice $\mathbf{A}$. Množinu všetkých vlastných čísel matice $\mathbf{A}$ nazývame *spektrum* matice $\mathbf{A}$.

Poznámka 10.2. Namiesto pojmu vlastné číslo sa niekedy používa pojem *vlastná hodnota.*

Poznámka 10.3. Ak $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$ pre $\mathbf{x} \neq \mathbf{0}$, potom hovoríme, že vlastné číslo λ prislúcha vlastnému vektoru $\mathbf{x}$, resp. že vlastný vektor $\mathbf{x}$ prislúcha vlastnému číslu λ.

Poznámka 10.4. Uvažujeme iba reálne vlastné čísla, takže niektoré známe tvrdenia o komplexných vlastných číslach sa na nami uvažované vlastné čísla nevzťahujú. Napríklad platí, že niektoré matice nemajú žiadne reálne vlastné číslo, ako ukážeme v príklade 10.14, hoci každá matica má aspoň jedno komplexné vlastné číslo.

Príklad 10.5. Rovnica $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$ vyjadruje, že v smere vlastného vektora $\mathbf{x}$ predstavuje lineárne zobrazenie $\mathbf{v} \mapsto \mathbf{A}\mathbf{v}$ jednoduché natiahnutie, či skrátenie vektora na jeho λ-násobok.

Overte, že matica

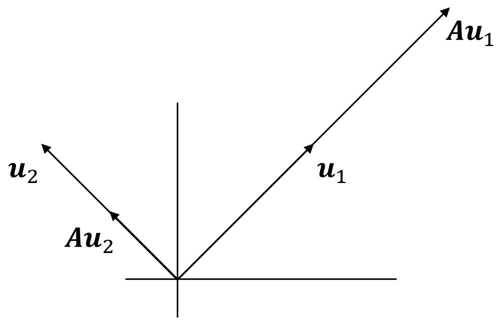
$$\mathbf{A} = \begin{pmatrix} 5/4 & 3/4 \\ 3/4 & 5/4 \end{pmatrix}$$

má vlastné čísla 2 a $1/2$ s prislúchajúcimi vlastnými vektormi $\mathbf{u}_1 = (1/\sqrt{2}, 1/\sqrt{2})^T$ a $\mathbf{u}_2 = (-1/\sqrt{2}, 1/\sqrt{2})^T$. Na obrázku 10.1 je znázornené natiahnutie/skrátenie príslušných vlastných vektorov po vynásobení maticou $\mathbf{A}$: $\mathbf{A}\mathbf{u}_1 = 2\mathbf{u}_1$ a $\mathbf{A}\mathbf{u}_2 = (1/2)\mathbf{u}_2$.

Lema 10.6 (O prislúchajúcom vlastnom čísle)**.** Nech $\mathbf{x} \neq \mathbf{0}_n$ je vlastný vektor matice $\mathbf{A} \in \mathbb{R}^{n\times n}$. Potom vektoru $\mathbf{x}$ prislúcha jediné vlastné číslo λ matice $\mathbf{A}$, a to $\lambda = \mathbf{x}^T\mathbf{A}\mathbf{x}/\|\mathbf{x}\|^2$.

Dôkaz. Nech λ je vlastné číslo prislúchajúce $\mathbf{x}$. Potom z $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$ vyplýva $\mathbf{x}^T\mathbf{A}\mathbf{x} = \lambda\mathbf{x}^T\mathbf{x}$. Keďže $\mathbf{x} \neq \mathbf{0}$, tak $\mathbf{x}^T\mathbf{x} = \|\mathbf{x}\|^2 > 0$, a teda jediné vlastné číslo prislúchajúce $\mathbf{x}$ je $\lambda = \mathbf{x}^T\mathbf{A}\mathbf{x}/\|\mathbf{x}\|^2$. □

Lema 10.7 (O skalárnom násobku vlastného vektora)**.** Nech $\mathbf{x}$ je vlastný vektor matice $\mathbf{A} \in \mathbb{R}^{n\times n}$ prislúchajúci vlastnému číslu λ. Potom pre ľubovoľné $c \in \mathbb{R}$, $c \neq 0$, je $c\mathbf{x}$ tiež vlastný vektor matice $\mathbf{A}$ prislúchajúci vlastnému číslu λ.



Obr. 10.1: Transformácia daná maticou $\mathbf{A}$ z príkladu 10.5 aplikovaná na jej vlastné vektory $\mathbf{u}_1$ a $\mathbf{u}_2$.

Dôkaz. Ľahko overíme, že $c\mathbf{x}$ spĺňa definíciu vlastného vektora: $\mathbf{A}(c\mathbf{x}) = c\mathbf{A}\mathbf{x} = c\lambda\mathbf{x} = \lambda(c\mathbf{x})$. □

Lema 10.8 (Ekvivalentné charakterizácie vlastného čísla)**.** Pre $\mathbf{A} \in \mathbb{R}^{n\times n}$ sú nasledujúce tvrdenia ekvivalentné:

(i) λ je vlastné číslo $\mathbf{A}$,

(ii) $\mathbf{A} - \lambda\mathbf{I}_n$ je singulárna,

(iii) $\det(\mathbf{A} - \lambda\mathbf{I}_n) = 0$.

Dôkaz. Zjavne λ je vlastné číslo práve vtedy, keď $\mathbf{A}\mathbf{x} - \lambda\mathbf{x} = \mathbf{0}$, čiže keď $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$, pre nejaké $\mathbf{x} \neq \mathbf{0}$. To je ekvivalentné s tým, že $\mathcal{N}(\mathbf{A} - \lambda\mathbf{I}) \neq \{\mathbf{0}\}$, teda podľa dôsledku 3.31 s tým, že $\mathbf{A} - \lambda\mathbf{I}$ je singulárna. Ekvivalencia (ii) s $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$ vyplýva z vety 8.11. □

Lema 10.9 (Ekvivalentné charakterizácie vlastného vektora)**.** Pre $\mathbf{A} \in \mathbb{R}^{n\times n}$ sú nasledujúce tvrdenia ekvivalentné:

(i) $\mathbf{u} \in \mathbb{R}^n$ je vlastný vektor matice $\mathbf{A}$ prislúchajúci vlastnému číslu λ,

(ii) $\mathbf{u}$ je riešením homogénneho systému $(\mathbf{A} - \lambda\mathbf{I}_n)\mathbf{x} = \mathbf{0}_n$ v premennej $\mathbf{x}$,

(iii) $\mathbf{u} \in \mathcal{N}(\mathbf{A} - \lambda\mathbf{I}_n)$.

Dôkaz. Dôkaz je analogický k dôkazu lemy 10.8, nechávame ho na čitateľa. □

Lema 10.10 (Stupeň polynómu $\det(\mathbf{A} - \lambda\mathbf{I}_n)$)**.** Pre $\mathbf{A} \in \mathbb{R}^{n\times n}$ je výraz $\det(\mathbf{A} - \lambda\mathbf{I}_n)$ polynómom stupňa n v premennej λ.

Dôkaz. Označme $\mathbf{B} = \mathbf{A} - \lambda\mathbf{I}$. Z definície determinantu máme

$$\det(\mathbf{B}) = \sum_{\sigma-\text{perm.}} (-1)^{\Phi(\sigma)} (\mathbf{B})_{1,\sigma(1)} \cdots (\mathbf{B})_{n,\sigma(n)},$$

kde každý výraz $(\mathbf{B})_{i,\sigma(i)}$ je buď tvaru $(\mathbf{A})_{ii} - \lambda$, alebo nezávisí od λ. Teda $\det(\mathbf{B})$ je polynómom stupňa nanajvýš n v premennej λ, keďže každý zo sčítancov $(-1)^{\Phi(\sigma)}(\mathbf{B})_{1,\sigma(1)} \cdots (\mathbf{B})_{n,\sigma(n)}$ je vzhľadom na λ polynómom stupňa nanajvýš n. Overme, že $\det(\mathbf{B})$ je polynómom presne stupňa n, teda že koeficient pri λ^n nie je 0. Jediný výraz, v ktorom vystupuje λ^n, zodpovedá tomu, keď každé $(\mathbf{B})_{i,\sigma(i)}$ je tvaru $(\mathbf{A})_{ii} - \lambda$, čo je vtedy, keď $\sigma(i) = i$ pre všetky $i = 1, \ldots, n$. Vtedy dostávame $(-1)^{\Phi(\sigma)}(\mathbf{B})_{1,\sigma(1)} \cdots (\mathbf{B})_{n,\sigma(n)} = \prod_{i=1}^{n}((\mathbf{A})_{ii} - \lambda)$, čo je polynóm s koeficientom $(-1)^n$ pri λ^n. Takže $\det(\mathbf{B})$ je skutočne polynómom stupňa n. □

Definícia 10.11 (Charakteristický polynóm)**.** Polynóm $\det(\mathbf{A}-\lambda\mathbf{I})$ v premennej λ nazývame *charakteristický polynóm* matice $\mathbf{A}$.

Definícia 10.12 (Vlastný priestor, násobnosť vlastného čísla)**.** Priestor $\mathcal{N}(\mathbf{A} - \lambda\mathbf{I})$ nazývame *vlastný priestor* vlastného čísla λ, jeho dimenziu nazývame (geometrická) *násobnosť* vlastného čísla λ.

Poznámka 10.13. Okrem geometrickej násobnosti môžeme definovať aj *algebraickú násobnosť* vlastného čísla: algebraická násobnosť vlastného čísla λ^* je násobnosť koreňa $\lambda = \lambda^*$ v charakteristickom polynóme $\det(\mathbf{A} - \lambda\mathbf{I})$. Poznamenajme, že geometrická a algebraická násobnosť vlastného čísla sa môžu líšiť. V tomto texte však budeme pracovať iba s geometrickou násobnosťou, čiže pod pojmom násobnosť vždy myslíme geometrickú násobnosť.

Príklad 10.14. Nech

$$\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 0 & -1 \end{pmatrix} \quad \text{a} \quad \mathbf{B} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Potom $\det(\mathbf{A} - \lambda\mathbf{I}) = (2 - \lambda)(-1 - \lambda)$. Vlastné čísla matice $\mathbf{A}$ sú riešenia $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$, čiže $\mathbf{A}$ má dve vlastné čísla $\lambda_1 = 2$ a $\lambda_2 = -1$, každé s násobnosťou 1.

Charakteristický polynóm matice $\mathbf{B}$ je $\lambda^2 + 1$, a ten má korene iba $\lambda_{1,2} = \pm i$, kde i je imaginárna jednotka. Takže $\mathbf{B}$ nemá žiadne (reálne) vlastné čísla. Nájdite 2×2 maticu, ktorá má iba jedno vlastné číslo, s násobnosťou 1.

Lema 10.15 (Vzťah nulového vlastného čísla a nulového priestoru)**.** Nech $\mathbf{A} \in \mathbb{R}^{n\times n}$. Násobnosť vlastného čísla $\lambda = 0$ je rovná dimenzii $\mathcal{N}(\mathbf{A})$. Pod násobnosťou 0 "vlastného čísla" λ rozumieme, že λ nie je vlastné číslo matice $\mathbf{A}$.

Dôkaz. Násobnosť $\lambda = 0$ je $\dim(\mathcal{N}(\mathbf{A} - 0\mathbf{I})) = \dim(\mathcal{N}(\mathbf{A}))$. □

Dôsledok 10.16 (Vzťah singulárnosti matice a nulového vlastného čísla)**.** Matica $\mathbf{A} \in \mathbb{R}^{n\times n}$ je singulárna práve vtedy, keď $\lambda = 0$ je jej vlastné číslo. Ekvivalentne, $\mathbf{A}$ je regulárna práve vtedy, keď $\lambda = 0$ nie je jej vlastné číslo.

Poznámka 10.17. Niekedy je výhodné pracovať s vlastnými číslami zopakovanými podľa násobnosti. Napríklad vlastné čísla matice $\mathbf{A} = \mathrm{diag}(3, 1, 5, 1, 3)$ sú 3 (s násobnosťou 2), 1 (s násobnosťou 2) a 5 (s násobnosťou 1), pričom vlastné čísla $\mathbf{A}$ zopakované podľa násobnosti sú $3, 3, 1, 1, 5$.

Lema 10.18 (Vlastné čísla a vektory transformácií matice $\mathbf{A}$)**.** Nech λ je vlastné číslo matice $\mathbf{A} \in \mathbb{R}^{n\times n}$ a $\mathbf{x}$ je k nemu prislúchajúci vlastný vektor. Potom:

(i) Pre každé $c \in \mathbb{R}$ je $c\lambda$ vlastné číslo matice $c\mathbf{A}$ prislúchajúce vlastnému vektoru $\mathbf{x}$.

(ii) Pre každé $k \in \mathbb{N}$ je λ^k vlastné číslo matice $\mathbf{A}^k$ prislúchajúce vlastnému vektoru $\mathbf{x}$.

(iii) Ak $\mathbf{A}$ je regulárna, tak λ^{-1} je vlastné číslo matice $\mathbf{A}^{-1}$ prislúchajúce vlastnému vektoru $\mathbf{x}$.

Dôkaz. Dôkaz tvrdenia (i) je triviálny. Overme časť (ii). Rovnicu $\mathbf{Ax} = \lambda\mathbf{x}$ prenásobíme maticou $\mathbf{A}$ zľava a dostaneme $\mathbf{A}^2\mathbf{x} = \lambda\mathbf{A}\mathbf{x}$. Pravá strana tejto rovnice je rovná $\lambda(\lambda\mathbf{x})$, keďže $\mathbf{Ax} = \lambda\mathbf{x}$. Dostávame $\mathbf{A}^2\mathbf{x} = \lambda^2\mathbf{x}$, čím je dokázané tvrdenie pre $k = 2$. Pre vyššie hodnoty k sa dá tvrdenie dokázať matematickou indukciou, resp. uvedomením si, že každým postupným vynásobením rovnice $\mathbf{Ax} = \lambda\mathbf{x}$ zvýšime na pravej strane mocninu vlastného čísla λ.

Podobne v (iii) môžeme rovnosť $\mathbf{Ax} = \lambda\mathbf{x}$ prenásobiť maticou $\mathbf{A}^{-1}$, čím dostaneme $\mathbf{x} = \lambda\mathbf{A}^{-1}\mathbf{x}$, čo sa dá prepísať na $\mathbf{A}^{-1}\mathbf{x} = \lambda^{-1}\mathbf{x}$. Číslom λ môžeme deliť, lebo podľa dôsledku 10.16 platí $\lambda \neq 0$. □

Lema 10.19 (Vlastné čísla súčinu matíc)**.** Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$ a $\mathbf{B} \in \mathbb{R}^{n\times m}$. Potom matica $\mathbf{AB}$ má rovnaké nenulové vlastné čísla ako matica $\mathbf{BA}$, vrátane násobností.

Dôkaz. Počítajme charakteristické polynómy:

$$p_{\mathbf{AB}}(\lambda) = \det(\mathbf{AB} - \lambda\mathbf{I}_m) = (-\lambda)^m \det(-\lambda^{-1}\mathbf{AB} + \mathbf{I}_m),$$

$$p_{\mathbf{BA}}(\lambda) = \det(\mathbf{BA} - \lambda\mathbf{I}_n) = (-\lambda)^n \det(-\lambda^{-1}\mathbf{BA} + \mathbf{I}_n).$$

Podľa úlohy 8.5 potom dostávame $p_{\mathbf{BA}}(\lambda) = (-\lambda)^n \det(-\lambda^{-1}\mathbf{AB} + \mathbf{I}_m)$. Teda pre každé $\lambda \neq 0$ platí $(-\lambda)^n p_{\mathbf{AB}}(\lambda) = (-\lambda)^m p_{\mathbf{BA}}(\lambda)$. Špeciálne potom majú oba tieto polynómy rovnaké nenulové korene, čiže $\mathbf{A}$ a $\mathbf{B}$ majú rovnaké nenulové vlastné čísla.

Na dokázanie rovností násobností nech $\lambda \neq 0$ je vlastné číslo matice $\mathbf{AB}$ a nech $\mathbf{x}_1, \dots, \mathbf{x}_k$ sú lineárne nezávislé vlastné vektory $\mathbf{AB}$ zodpovedajúce vlastnému číslu λ. Keď označíme $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_k) \in \mathbb{R}^{m\times k}$, potom $\mathbf{ABX} = \lambda\mathbf{X}$, keďže $\mathbf{x}_i$ sú príslušné vlastné vektory. Potom $\mathbf{BA}(\mathbf{BX}) = \mathbf{B}(\mathbf{ABX}) = \lambda\mathbf{BX}$, čiže stĺpce matice $\mathbf{BX}$ sú vlastné vektory matice $\mathbf{BA}$ prislúchajúce vlastnému číslu λ.

Z rovnosti $\mathbf{ABX} = \lambda\mathbf{X}$, dostávame $\operatorname{rank}(\lambda\mathbf{X}) \leq \operatorname{rank}(\mathbf{BX})$ podľa vety 3.14. Podľa tej istej vety tiež platí $\operatorname{rank}(\mathbf{BX}) \leq \operatorname{rank}(\mathbf{X})$. Keďže $\lambda \neq 0$, tak $\operatorname{rank}(\lambda\mathbf{X}) = \operatorname{rank}(\mathbf{X})$ a teda $\operatorname{rank}(\mathbf{BX}) = \operatorname{rank}(\mathbf{X})$. Keďže stĺpce $\mathbf{X}$ sú lineárne nezávislé, tak $\operatorname{rank}(\mathbf{BX}) = \operatorname{rank}(\mathbf{X}) = k$. Zároveň platí $\mathbf{BX} \in \mathbb{R}^{n\times k}$, čiže aj stĺpce $\mathbf{BX}$ sú lineárne nezávislé. Teda ak $\mathbf{AB}$ má k lineárne nezávislých vlastných vektorov prislúchajúcich λ, tak sme našli k lineárne nezávislých vlastných vektorov matice $\mathbf{BA}$ prislúchajúcich λ. Preto $\dim(\mathbf{BA} - \lambda\mathbf{I}) \geq \dim(\mathbf{AB} - \lambda\mathbf{I})$. Podobným postupom dostaneme $\dim(\mathbf{AB} - \lambda\mathbf{I}) \geq \dim(\mathbf{BA} - \lambda\mathbf{I})$ (pomocou pozorovania, že ak $\mathbf{Y}$ sú vlastné vektory $\mathbf{BA}$ prislúchajúce λ, tak $\mathbf{AY}$ sú vlastné vektory $\mathbf{AB}$ prislúchajúce λ). Spolu teda dostávame požadovanú rovnosť násobností vlastného čísla λ. □

Lema 10.20 (O nezávislosti vlastných vektorov)**.** Nech $\mathbf{A} \in \mathbb{R}^{n\times n}$ a λ_1, λ_2 sú jej dve rôzne vlastné čísla. Potom ak $\mathbf{x}_1, \dots, \mathbf{x}_k$ sú lineárne nezávislé vlastné vektory $\mathbf{A}$ prislúchajúce λ_1 a $\mathbf{y}_1, \dots, \mathbf{y}_\ell$ sú lineárne nezávislé vlastné vektory $\mathbf{A}$ prislúchajúce λ_2, tak $\mathbf{x}_1, \dots, \mathbf{x}_k, \mathbf{y}_1, \dots, \mathbf{y}_\ell$ sú lineárne nezávislé.

Dôkaz. Označme $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_k)$ a $\mathbf{Y} = (\mathbf{y}_1, \dots, \mathbf{y}_\ell)$. Všimnime si, že

$$\mathbf{AX} = (\mathbf{Ax}_1, \dots, \mathbf{Ax}_k) = \lambda_1(\mathbf{x}_1, \dots, \mathbf{x}_k) = \lambda_1\mathbf{X},$$

a podobne $\mathbf{AY} = \lambda_2\mathbf{Y}$. Keďže λ_1, λ_2 sú rôzne, aspoň jedno z nich je nenulové. Bez ujmy na všeobecnosti predpokladajme, že $\lambda_1 \neq 0$.

Tvrdenie dokážeme sporom. Predpokladajme, že $\mathbf{x}_1, \dots, \mathbf{x}_k, \mathbf{y}_1, \dots, \mathbf{y}_\ell$ sú lineárne závislé. Existuje teda $\mathbf{c} = (\mathbf{c}_1^T, \mathbf{c}_2^T)^T \neq \mathbf{0}$, kde $\mathbf{c}_1 \in \mathbb{R}^k$ a $\mathbf{c}_2 \in \mathbb{R}^\ell$, že $(\mathbf{X}, \mathbf{Y})\mathbf{c} = \mathbf{0}$. Čiže $\mathbf{Xc}_1 + \mathbf{Yc}_2 = \mathbf{0}$.

Po vynásobení tejto rovnice maticou $\mathbf{A}$ zľava dostávame $\lambda_1\mathbf{X}\mathbf{c}_1 + \lambda_2\mathbf{Y}\mathbf{c}_2 = \mathbf{0}$, teda $\mathbf{X}\mathbf{c}_1 = -\lambda_1^{-1}\lambda_2\mathbf{Y}\mathbf{c}_2$. Po dosadení do $\mathbf{X}\mathbf{c}_1 + \mathbf{Y}\mathbf{c}_2 = \mathbf{0}$ dostávame $\mathbf{Y}\mathbf{c}_2 = \lambda_1^{-1}\lambda_2\mathbf{Y}\mathbf{c}_2$, z čoho vyplýva $\mathbf{Y}\mathbf{c}_2 = \mathbf{0}$, keďže $\lambda_1^{-1}\lambda_2 \neq 1$. Keďže stĺpce $\mathbf{Y}$ sú lineárne nezávislé, dostávame $\mathbf{c}_2 = \mathbf{0}$. Z toho vyplýva $\mathbf{X}\mathbf{c}_1 = \mathbf{0}$, a teda podobne aj $\mathbf{c}_1 = \mathbf{0}$. To je však spor s $\mathbf{c} \neq \mathbf{0}$. □

Veta 10.21 (Diagonalizácia)**.** Nech $\mathbf{A} \in \mathbb{R}^{n\times n}$ má n vlastných čísel $\lambda_1, \dots, \lambda_n$ (vrátane násobností). Potom $\mathbf{A}$ vieme zapísať ako $\mathbf{A} = \mathbf{S}\mathbf{\Lambda}\mathbf{S}^{-1}$, kde $\mathbf{\Lambda} = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$, $\mathbf{S} = (\mathbf{x}_1, \dots, \mathbf{x}_n)$ je regulárna matica a $\mathbf{x}_1, \dots, \mathbf{x}_n$ sú vlastné vektory prislúchajúce vlastným číslam $\lambda_1, \dots, \lambda_n$.

Dôkaz. Vlastné vektory $\mathbf{x}_1, \dots, \mathbf{x}_n$ vyberieme nasledovne: každé vlastné číslo má toľko lineárne nezávislých vlastných vektorov, aká je jeho násobnosť; máme teda spolu n vlastných vektorov. Potom podľa lemy 10.20 sú tieto vlastné vektory lineárne nezávislé.

Pre každé $i = 1, \dots, n$ platí: $\mathbf{A}\mathbf{x}_i = \lambda_i\mathbf{x}_i$, čo môžeme spolu zapísať ako

$$(\mathbf{A}\mathbf{x}_1, \dots, \mathbf{A}\mathbf{x}_n) = (\lambda_1\mathbf{x}_1, \dots, \lambda_n\mathbf{x}_n),$$

teda $\mathbf{A}\mathbf{S} = \mathbf{S}\mathbf{\Lambda}$. Keďže $\mathbf{x}_1, \dots, \mathbf{x}_n$ sú lineárne nezávislé, tak $\mathbf{S} \in \mathbb{R}^{n\times n}$ je regulárna a dostávame $\mathbf{A} = \mathbf{S}\mathbf{\Lambda}\mathbf{S}^{-1}$. □

Definícia 10.22 (Podobné matice)**.** Nech $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n\times n}$. Potom hovoríme, že $\mathbf{A}$ a $\mathbf{B}$ sú *podobné* ak existuje regulárna matica $\mathbf{X} \in \mathbb{R}^{n\times n}$, že $\mathbf{B} = \mathbf{X}^{-1}\mathbf{A}\mathbf{X}$.

Veta 10.23 (Vlastné čísla podobných matíc)**.** Nech $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n\times n}$ sú podobné. Potom $\mathbf{A}$ a $\mathbf{B}$ majú rovnaké vlastné čísla, vrátane násobností.

Dôkaz. Keďže $\mathbf{A}, \mathbf{B}$ sú podobné, tak existuje regulárna matica $\mathbf{X} \in \mathbb{R}^{n\times n}$, že $\mathbf{B} = \mathbf{X}^{-1}\mathbf{A}\mathbf{X}$. Pozrime sa na charakteristický polynóm matice $\mathbf{B}$:

$$\begin{aligned}\det(\mathbf{B} - \lambda\mathbf{I}) &= \det(\mathbf{X}^{-1}\mathbf{A}\mathbf{X} - \lambda\mathbf{I}) = \det(\mathbf{X}^{-1}\mathbf{A}\mathbf{X} - \lambda\mathbf{X}^{-1}\mathbf{X}) = \det(\mathbf{X}^{-1}(\mathbf{A} - \lambda\mathbf{I})\mathbf{X}) \\ &= \det(\mathbf{X}^{-1})\det(\mathbf{A} - \lambda\mathbf{I})\det(\mathbf{X}) = \det(\mathbf{A} - \lambda\mathbf{I}),\end{aligned}$$

kde poslednú rovnosť sme získali vďaka $\det(\mathbf{X}^{-1}) = 1/\det(\mathbf{X})$. Teda $\mathbf{A}$ a $\mathbf{B}$ majú rovnaké vlastné čísla. Ešte potrebujeme ukázať, že majú rovnaké násobnosti.

Nech λ je vlastné číslo matice $\mathbf{A}$ (a teda aj $\mathbf{B}$). Potom násobnosť λ pre $\mathbf{A}$ je $n - \operatorname{rank}(\mathbf{A} - \lambda\mathbf{I})$. Násobnosť λ pre $\mathbf{B}$ je

$$n - \operatorname{rank}(\mathbf{B} - \lambda\mathbf{I}) = n - \operatorname{rank}(\mathbf{X}^{-1}\mathbf{A}\mathbf{X} - \lambda\mathbf{I}) = n - \operatorname{rank}(\mathbf{X}^{-1}(\mathbf{A} - \lambda\mathbf{I})\mathbf{X}).$$

Lenže násobenie regulárnou maticou nemení hodnosť (veta 4.18), čiže násobnosť λ pre $\mathbf{B}$ je $n - \operatorname{rank}(\mathbf{A} - \lambda\mathbf{I})$, čo je rovné násobnosti λ pre $\mathbf{A}$. □

10.2 Vlastné čísla a vlastné vektory symetrických matíc

Poznámka 10.24. V tejto podkapitole najprv ukážeme, že každá symetrická matica má aspoň jedno vlastné číslo. Toto tvrdenie sa zvyčajne dokazuje pomocou základnej vety algebry, ktorá však pracuje s komplexnými číslami, a ktorej dôkaz vychádza z výrazne netriviálnych tvrdení matematickej analýzy. Namiesto toho budeme sledovať argumentáciu v [4], ktorá umožňuje dokázať existenciu vlastného čísla symetrickej matice aj bez základnej vety algebry. Stále však bez dôkazu použijeme niektoré základné tvrdenia z matematickej analýzy.

Lema 10.25 (Vlastné vektory symetrickej matice pomocou kvadratickej funkcie)**.** Nech $\mathbf{A} \in \mathcal{S}^n$. Potom:

(i) Existujú nenulové vektory $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^n$ také, že

$$\frac{\mathbf{x}_1^T \mathbf{A} \mathbf{x}_1}{\mathbf{x}_1^T \mathbf{x}_1} \leq \frac{\mathbf{x}^T \mathbf{A} \mathbf{x}}{\mathbf{x}^T \mathbf{x}} \leq \frac{\mathbf{x}_2^T \mathbf{A} \mathbf{x}_2}{\mathbf{x}_2^T \mathbf{x}_2} \tag{10.1}$$

pre každé $\mathbf{x} \in \mathbb{R}^n$.

(ii) Vektory $\mathbf{x}_1$ a $\mathbf{x}_2$ sú vlastné vektory matice $\mathbf{A}$.

Dôkaz. Poznamenajme, že bez dôkazu budeme používať niektoré poznatky z oblasti minimalizácie/maximalizácie spojitých a diferencovateľných funkcií. Uvažujme množinu $M = \{\mathbf{x} \in \mathbb{R}^n \,|\, \mathbf{x}^T\mathbf{x} = 1\}$ a funkciu $f(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x}$. Keďže množina M je ohraničená a uzavretá (a teda kompaktná) a funkcia $f(\mathbf{x})$ je spojitá, tak existuje maximum aj minimum funkcie $f(\mathbf{x})$ na množine M. Označme príslušný bod minima ako $\mathbf{x}_1$ a bod maxima ako $\mathbf{x}_2$. Platí teda $\mathbf{x}_1^T\mathbf{A}\mathbf{x}_1 \leq \mathbf{x}^T\mathbf{A}\mathbf{x} \leq \mathbf{x}_2^T\mathbf{A}\mathbf{x}_2$ pre každé $\mathbf{x} \in M$. Keďže $\mathbf{x}_1, \mathbf{x}_2 \in M$, čiže majú jednotkovú normu, tak $\mathbf{x}_1, \mathbf{x}_2 \neq \mathbf{0}_n$.

Nech $\mathbf{x} \in \mathbb{R}^n$. Potom $\tilde{\mathbf{x}} = \mathbf{x}/\sqrt{\mathbf{x}^T\mathbf{x}}$ patrí do množiny M, čiže platí $\mathbf{x}_1^T\mathbf{A}\mathbf{x}_1 \leq \tilde{\mathbf{x}}^T\mathbf{A}\tilde{\mathbf{x}} \leq \mathbf{x}_2^T\mathbf{A}\mathbf{x}_2$. Po dosadení za $\tilde{\mathbf{x}}$ dostávame

$$\mathbf{x}_1^T \mathbf{A} \mathbf{x}_1 \leq \frac{\mathbf{x}^T \mathbf{A} \mathbf{x}}{\mathbf{x}^T \mathbf{x}} \leq \mathbf{x}_2^T \mathbf{A} \mathbf{x}_2. \tag{10.2}$$

Keďže $\mathbf{x}_1, \mathbf{x}_2 \in M$, tak $\mathbf{x}_i^T\mathbf{x}_i = 1$ a teda $\mathbf{x}_i^T\mathbf{A}\mathbf{x}_i = \mathbf{x}_i^T\mathbf{A}\mathbf{x}_i/(\mathbf{x}_i^T\mathbf{x}_i)$, $i = 1, 2$. Čiže (10.2) môžeme zapísať ako

$$\frac{\mathbf{x}_1^T \mathbf{A} \mathbf{x}_1}{\mathbf{x}_1^T \mathbf{x}_1} \leq \frac{\mathbf{x}^T \mathbf{A} \mathbf{x}}{\mathbf{x}^T \mathbf{x}} \leq \frac{\mathbf{x}_2^T \mathbf{A} \mathbf{x}_2}{\mathbf{x}_2^T \mathbf{x}_2}, \tag{10.3}$$

čím sme dokázali časť (i).

V (ii) bez ujmy na všeobecnosti tvrdenie dokážeme pre $\mathbf{x}_1$; pre $\mathbf{x}_2$ stačí vymeniť minimalizáciu za maximalizáciu. Keďže funkcia $g(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x}/(\mathbf{x}^T\mathbf{x})$ je diferencovateľná na $\mathbb{R}^n \setminus \{\mathbf{0}_n\}$ a keďže v bode $\mathbf{x}_1 \neq \mathbf{0}_n$ nadobúda svoje minimum, tak vektor parciálnych derivácií v bode $\mathbf{x}_1$ je rovný $\mathbf{0}_n$. Formálne:

$$\left.\frac{\partial g(\mathbf{x})}{\partial \mathbf{x}}\right|_{\mathbf{x}=\mathbf{x}_1} = \mathbf{0}_n.$$

Využijeme tiež nasledovné poznatky z derivovania funkcií viac premenných:

$$\frac{\partial}{\partial \mathbf{x}}\left(\frac{u(\mathbf{x})}{v(\mathbf{x})}\right) = \frac{v(\mathbf{x})\frac{\partial u(\mathbf{x})}{\partial \mathbf{x}} - u(\mathbf{x})\frac{\partial v(\mathbf{x})}{\partial \mathbf{x}}}{v(\mathbf{x})^2}, \qquad \frac{\partial \mathbf{x}^T\mathbf{A}\mathbf{x}}{\partial \mathbf{x}} = 2\mathbf{A}\mathbf{x} \text{ pre } \mathbf{A} \in \mathcal{S}^n, \tag{10.4}$$

ktoré môže čitateľ nájsť v kapitole 15 v [4]. Ak v (10.4) dosadíme $\mathbf{A} = \mathbf{I}$, dostávame $\partial(\mathbf{x}^T\mathbf{x})/\partial\mathbf{x} = 2\mathbf{x}$. Spolu s (10.3) dostávame

$$\frac{\mathbf{x}_1^T\mathbf{x}_1(2\mathbf{A}\mathbf{x}_1) - \mathbf{x}_1^T\mathbf{A}\mathbf{x}_1(2\mathbf{x}_1)}{(\mathbf{x}_1^T\mathbf{x}_1)^2} = \mathbf{0}_n,$$

a teda $(\mathbf{x}_1^T\mathbf{x}_1)\mathbf{A}\mathbf{x}_1 = (\mathbf{x}_1^T\mathbf{A}\mathbf{x}_1)\mathbf{x}_1$, z čoho vyplýva

$$\mathbf{A}\mathbf{x}_1 = \frac{\mathbf{x}_1^T\mathbf{A}\mathbf{x}_1}{\mathbf{x}_1^T\mathbf{x}_1}\mathbf{x}_1.$$

Dostali sme $\mathbf{A}\mathbf{x}_1 = \lambda_1 \mathbf{x}_1$, kde $\lambda_1 = (\mathbf{x}_1^T \mathbf{A}\mathbf{x}_1)/(\mathbf{x}_1^T \mathbf{x}_1)$, pričom $\mathbf{x}_1 \neq \mathbf{0}_n$. Vektor $\mathbf{x}_1$ je teda vlastný vektor matice $\mathbf{A}$. □

Dôsledok 10.26. Každá symetrická matica má aspoň jedno (reálne) vlastné číslo.

Veta 10.27 (Spektrálny rozklad symetrickej matice)**.** Nech $\mathbf{A} \in \mathcal{S}^n$. Potom existujú ortogonálna matica $\mathbf{U} \in \mathbb{R}^{n\times n}$ a diagonálna matica $\mathbf{\Lambda} \in \mathbb{R}^{n\times n}$ také, že $\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T$. Navyše platí, že stĺpce matice $\mathbf{U}$ sú (ortonormálne) vlastné vektory matice $\mathbf{A}$ a diagonálne prvky matice $\mathbf{\Lambda}$ sú príslušné vlastné čísla.

Dôkaz. Prvú časť dokážeme matematickou indukciou. Pre 1×1 matice tvrdenie zjavne platí. Predpokladajme teraz, že veta platí pre každú maticu v $\mathcal{S}^{n-1}$, a vezmime $\mathbf{A} \in \mathcal{S}^n$. Podľa dôsledku 10.26 existuje vlastné číslo λ a príslušný vlastný vektor $\mathbf{u} \neq \mathbf{0}_n$ matice $\mathbf{A}$. Keďže násobok vlastného vektora je tiež vlastný vektor (prislúchajúci tomu istému vlastnému číslu), tak môžeme predpokladať, že $\|\mathbf{u}\| = 1$. Potom môžeme $\mathbf{u}$ doplniť pomocou $n-1$ vektorov na ortonormálnu bázu $\mathbb{R}^n$. Uložme tieto vektory do matice $\mathbf{V} \in \mathbb{R}^{n\times(n-1)}$. Platí teda, že matica $(\mathbf{u}, \mathbf{V})$ je ortogonálna, čiže

$$\mathbf{I}_n = \begin{pmatrix} \mathbf{u}^T \\ \mathbf{V}^T \end{pmatrix} \begin{pmatrix} \mathbf{u} & \mathbf{V} \end{pmatrix} = \begin{pmatrix} \mathbf{u}^T\mathbf{u} & \mathbf{u}^T\mathbf{V} \\ \mathbf{V}^T\mathbf{u} & \mathbf{V}^T\mathbf{V} \end{pmatrix},$$

z čoho vyplýva $\mathbf{u}^T\mathbf{V} = \mathbf{0}_{n-1}^T$ a $\mathbf{V}^T\mathbf{u} = \mathbf{0}_{n-1}$. Počítajme

$$\begin{pmatrix} \mathbf{u}^T \\ \mathbf{V}^T \end{pmatrix} \mathbf{A} \begin{pmatrix} \mathbf{u} & \mathbf{V} \end{pmatrix} = \begin{pmatrix} \mathbf{u}^T\mathbf{A}\mathbf{u} & \mathbf{u}^T\mathbf{A}\mathbf{V} \\ \mathbf{V}^T\mathbf{A}\mathbf{u} & \mathbf{V}^T\mathbf{A}\mathbf{V} \end{pmatrix} = \begin{pmatrix} \lambda & \mathbf{0}_{n-1}^T \\ \mathbf{0}_n & \mathbf{V}^T\mathbf{A}\mathbf{V} \end{pmatrix},$$

kde poslednú maticu sme dostali, lebo $\mathbf{u}^T\mathbf{A}\mathbf{u} = \mathbf{u}^T\lambda\mathbf{u} = \lambda\mathbf{u}^T\mathbf{u} = 1$, $\mathbf{V}^T\mathbf{A}\mathbf{u} = \lambda\mathbf{V}^T\mathbf{u} = \mathbf{0}_{n-1}$ a podobne $\mathbf{u}^T\mathbf{A}\mathbf{V} = \mathbf{0}_{n-1}^T$. Matica $\mathbf{V}^T\mathbf{A}\mathbf{V}$ je symetrická a typu $(n-1)\times(n-1)$, čiže podľa indukčného predpokladu existujú ortogonálna $\mathbf{U}_1$ a diagonálna $\mathbf{\Lambda}_1$, že $\mathbf{V}^T\mathbf{A}\mathbf{V} = \mathbf{U}_1\mathbf{\Lambda}_1\mathbf{U}_1^T$. Z toho vyplýva, že $\mathbf{U}_1^T\mathbf{V}^T\mathbf{A}\mathbf{V}\mathbf{U}_1 = \mathbf{\Lambda}_1$. Definujme

$$\mathbf{U} = \begin{pmatrix} \mathbf{u} & \mathbf{V} \end{pmatrix} \begin{pmatrix} 1 & \mathbf{0}_{n-1}^T \\ \mathbf{0}_{n-1} & \mathbf{U}_1^T \end{pmatrix}.$$

Potom

$$\begin{aligned}\mathbf{U}^T\mathbf{A}\mathbf{U} &= \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{U}_1 \end{pmatrix} \begin{pmatrix} \mathbf{u}^T \\ \mathbf{V}^T \end{pmatrix} \mathbf{A} \begin{pmatrix} \mathbf{u} & \mathbf{V} \end{pmatrix} \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{U}_1 \end{pmatrix} = \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{U}_1^T \end{pmatrix} \begin{pmatrix} \lambda & \mathbf{0}^T \\ \mathbf{0} & \mathbf{V}^T\mathbf{A}\mathbf{V} \end{pmatrix} \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{U}_1 \end{pmatrix} \\ &= \begin{pmatrix} \lambda & \mathbf{0}^T \\ \mathbf{0} & \mathbf{U}_1^T\mathbf{V}^T\mathbf{A}\mathbf{V}\mathbf{U}_1 \end{pmatrix} = \begin{pmatrix} \lambda & \mathbf{0}^T \\ \mathbf{0} & \mathbf{\Lambda}_1 \end{pmatrix},\end{aligned}$$

čo je diagonálna matica, môžeme ju označiť $\mathbf{\Lambda}$. Navyše matica $\mathbf{U}$ je ortogonálna, lebo sme ju skonštruovali ako súčin dvoch ortogonálnych matíc (pozri lemu 4.30) . Potom z $\mathbf{U}^T\mathbf{A}\mathbf{U} = \mathbf{\Lambda}$ vyplýva $\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T$.

Na dôkaz druhej časti si stačí zapísať rovnicu $\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T$ ako $\mathbf{A}\mathbf{U} = \mathbf{U}\mathbf{\Lambda}$. Keď potom označíme $\mathbf{U} = (\mathbf{u}_1, \ldots, \mathbf{u}_n)$ a $\mathbf{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, tak dostávame

$$(\mathbf{A}\mathbf{u}_1, \ldots, \mathbf{A}\mathbf{u}_n) = (\lambda_1\mathbf{u}_1, \ldots, \lambda_n\mathbf{u}_n),$$

teda $\mathbf{A}\mathbf{u}_i = \lambda_i\mathbf{u}_i$ $(i = 1, \ldots, n)$, čiže $\mathbf{u}_i$ sú vlastné vektory a λ_i sú príslušné vlastné čísla. □

Dôsledok 10.28. Nech $\mathbf{A} \in \mathcal{S}^n$. Potom $\mathbf{A}$ má n vlastných čísel $\lambda_1, \ldots, \lambda_n$ (po zarátaní násobností) a existuje ortonormálny systém vlastných vektorov $\mathbf{u}_1, \ldots, \mathbf{u}_n$ matice $\mathbf{A}$ taký, že $\mathbf{u}_i$ prislúcha k λ_i, $i = 1, \ldots, n$.

Dôsledok 10.29 (Spektrálny rozklad vo vektorovej forme)**.** Každá $\mathbf{A} \in \mathcal{S}^n$ sa dá zapísať v tvare $\mathbf{A} = \sum_{i=1}^n \lambda_i \mathbf{u}_i \mathbf{u}_i^T$, kde $\lambda_1, \ldots, \lambda_n$ sú vlastné čísla matice $\mathbf{A}$ a $\mathbf{u}_1, \ldots, \mathbf{u}_n$ je ortonormálny systém príslušných vlastných vektorov.

Dôkaz. Stačí rozpísať $\mathbf{A} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^T$ pomocou vlastných čísel a vlastných vektorov. □

Poznámka 10.30. Podľa vety 10.27 tvoria vlastné vektory $\mathbf{u}_1, \ldots, \mathbf{u}_n$ symetrickej matice $\mathbf{A} \in \mathcal{S}^n$ ortonormálnu bázu $\mathbb{R}^n$. Ľubovoľný vektor $\mathbf{x} \in \mathbb{R}^n$ môžeme teda zapísať ako $\mathbf{x} = \sum_{i=1}^n c_i \mathbf{u}_i$ pre nejaké $c_1, \ldots, c_n \in \mathbb{R}$. Potom $\mathbf{A}\mathbf{x} = \mathbf{A} \sum_{i=1}^n c_i \mathbf{u}_i = \sum_{i=1}^n c_i \mathbf{A}\mathbf{u}_i = \sum_{i=1}^n \lambda_i c_i \mathbf{u}_i$. Ak teda vyjadríme ľubovoľný vektor $\mathbf{x}$ v ortonormálnej báze danej vlastnými vektormi $\mathbf{A}$, tak násobenie maticou $\mathbf{A}$ je jednoduché natiahnutie, či skrátenie jednotlivých súradníc c_i na $\lambda_i c_i$.

Uvažujme maticu $\mathbf{A}$ z príkladu 10.5, podľa ktorej túto maticu vieme zapísať ako $\mathbf{A} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^T$, kde

$$\boldsymbol{\Lambda} = \begin{pmatrix} 2 & 0 \\ 0 & 1/2 \end{pmatrix} \quad \text{a} \quad \mathbf{U} = \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}.$$

Zvoľme vektor $\mathbf{x} = (-2/3, 1)^T/\sqrt{2}$. Potom ak $\mathbf{x}$ vyjadríme v ortogonálnej súradnicovej sústave danej vektormi $\mathbf{u}_1$ a $\mathbf{u}_2$ ako $\mathbf{x} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2$, tak $\mathbf{A}\mathbf{x}$ je vektor, ktorý ľahko získame natiahnutím prvej súradnice c_1 na dvojnásobok ($\lambda_1 = 2$) a skrátením druhej súradnice c_2 na polovicu ($\lambda_2 = 1/2$), ako je znázornené na obrázku 10.2a.

Poznámka 10.31. Pre iný geometrický pohľad na spektrálny rozklad $\mathbf{A} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^T$ si môžeme uvedomiť, že ortogonálne matice $\mathbf{U}$ a $\mathbf{U}^T$ zodpovedajú rotáciám (prípadne preklopeniam) a diagonálna matica $\boldsymbol{\Lambda}$ pozložkovému natiahnutiu/stiahnutiu. Navyše, "rotácie" $\mathbf{U}$ a $\mathbf{U}^T$ sú opačné, lebo príslušné matice sú inverzné. Takže vynásobenie vektora $\mathbf{x} \in \mathbb{R}^n$ symetrickou maticou $\mathbf{A} \in \mathcal{S}^n$ vieme rozložiť do troch krokov: 1. "rotácia" maticou $\mathbf{U}^T$, 2. natiahnutie/stiahnutie pomocou matice $\boldsymbol{\Lambda}$, 3. spätná "rotácia" maticou $\mathbf{U}$.

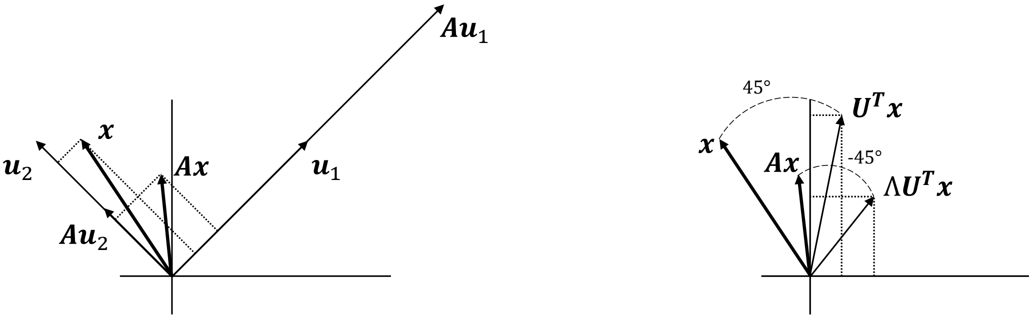
Pre maticu $\mathbf{A}$ z príkladu 10.5 zodpovedá $\mathbf{U}^T$ rotácii v smere hodinových ručičiek o $\pi/4$, teda o $45°$ (pozri príklad 4.29); $\boldsymbol{\Lambda}$ natiahnutiu prvej zložky na dvojnásobok a skráteniu druhej zložky o polovicu; a $\mathbf{U}$ rotácii o $45°$ proti smeru hodinových ručičiek. Pre vektor $\mathbf{x} = (-2/3, 1)^T/\sqrt{2}$ je toto rozloženie transformácie $\mathbf{A}\mathbf{x}$ zakreslené na obrázku 10.2b.

Príklad 10.32. Nech $\mathbf{A} \in \mathcal{S}^n$ a $\mathbf{u}_1, \ldots, \mathbf{u}_n$ sú jej ortonormálne vlastné vektory. Vyjadrite ľubovoľný vektor $\mathbf{x} \in \mathbb{R}^n$ v báze $\mathbf{u}_1, \ldots, \mathbf{u}_n$. Vyjadrite teda predpis pre $c_1, \ldots, c_n$ také, že $\mathbf{x} = c_1\mathbf{u}_1 + \ldots + c_n\mathbf{u}_n$. Viete tiež vektor súradníc $\mathbf{c}$ vyjadriť pomocou matíc $\boldsymbol{\Lambda}$ a $\mathbf{U}$ zo spektrálneho rozkladu $\mathbf{A} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^T$?

Nakoniec vypočítajte súradnice vektora $\mathbf{x}$ z príkladu 10.5 vzhľadom na bázu $\mathbf{u}_1, \mathbf{u}_2$ tvorenú vlastnými vektormi matice $\mathbf{A}$.

Veta 10.33 (Najväčšie a najmenšie vlastné číslo pomocou optimalizácie)**.** Nech $\mathbf{A} \in \mathcal{S}^n$. Potom

$$\max_{\mathbf{x} \in \mathbb{R}^n} \frac{\mathbf{x}^T\mathbf{A}\mathbf{x}}{\mathbf{x}^T\mathbf{x}} = \lambda_{\max}, \quad \min_{\mathbf{x} \in \mathbb{R}^n} \frac{\mathbf{x}^T\mathbf{A}\mathbf{x}}{\mathbf{x}^T\mathbf{x}} = \lambda_{\min},$$



(a) cez vyjadrenie $\mathbf{x}$ v báze $\mathbf{u}_1$, $\mathbf{u}_2$ (b) cez rozklad na rotáciu-natiahnutie-rotáciu

Obr. 10.2: Vyjadrenie lineárnej transformácie $\mathbf{Ax}$ pre maticu $\mathbf{A}$ z príkladu 10.5 a poznámok 10.30 a 10.31 pomocou spektrálneho rozkladu. (a) Súradnice vektora $\mathbf{x}$ v báze $\mathbf{u}_1, \mathbf{u}_2$ sú znázornené pomocou prerušovaných čiar. Následne sú tieto súradnice vynásobené príslušnými vlastnými číslami, čím opäť pomocou prerušovaných čiar získavame vektor $\mathbf{Ax}$. (b) Najprv $\mathbf{x}$ zrotujeme na $\mathbf{U}^T\mathbf{x}$, potom súradnice výsledného vektora natiahneme/stiahneme pomocou $\mathbf{\Lambda}$ na $\mathbf{\Lambda U}^T\mathbf{x}$, a nakoniec zrotujeme naspäť pomocou $\mathbf{U}$, čím dostávame $\mathbf{Ax} = \mathbf{U\Lambda U}^T\mathbf{x}$.

kde $\lambda_{\max}$ je najväčšie a $\lambda_{\min}$ je najmenšie vlastné číslo matice $\mathbf{A}$. Navyše, optimálnym riešením uvedenej maximalizačnej úlohy je ľubovoľný vlastný vektor matice $\mathbf{A}$ prislúchajúci $\lambda_{\max}$ a optimálnym riešením uvedenej minimalizačnej úlohy je ľubovoľný vlastný vektor matice $\mathbf{A}$ prislúchajúci $\lambda_{\min}$.

Dôkaz. Dôkaz spravíme pre $\lambda_{\max}$, pre $\lambda_{\min}$ je analogický. Keďže $\mathbf{A}$ je symetrická, tak ju môžeme zapísať pomocou spektrálneho rozkladu ako $\mathbf{A} = \mathbf{U\Lambda U}^T$. Potom pre ľubovoľné $\mathbf{x} \in \mathbb{R}^n$ platí

$$\frac{\mathbf{x}^T\mathbf{Ax}}{\mathbf{x}^T\mathbf{x}} = \frac{\mathbf{x}^T\mathbf{U\Lambda U}^T\mathbf{x}}{\mathbf{x}^T\mathbf{x}} = \frac{\mathbf{y}^T\mathbf{\Lambda y}}{\mathbf{x}^T\mathbf{x}} = \frac{\sum_i \lambda_i y_i^2}{\mathbf{x}^T\mathbf{x}} = \frac{\sum_i \lambda_i(\mathbf{u}_i^T\mathbf{x})^2}{\mathbf{x}^T\mathbf{x}},$$

kde $\mathbf{y} = \mathbf{U}^T\mathbf{x}$, $\mathbf{U} = (\mathbf{u}_1, \dots, \mathbf{u}_n)$ a $\mathbf{\Lambda} = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$. Keďže výrazy $(\mathbf{u}_i^T\mathbf{x})^2$ sú nezáporné, tak $\lambda_i(\mathbf{u}_i^T\mathbf{x})^2 \leq \lambda_{\max}(\mathbf{u}_i^T\mathbf{x})^2$ a dostávame

$$\frac{\mathbf{x}^T\mathbf{Ax}}{\mathbf{x}^T\mathbf{x}} \leq \frac{\sum_i \lambda_{\max}(\mathbf{u}_i^T\mathbf{x})^2}{\mathbf{x}^T\mathbf{x}} = \frac{\lambda_{\max}\sum_i(\mathbf{u}_i^T\mathbf{x})^2}{\mathbf{x}^T\mathbf{x}} = \frac{\lambda_{\max}\mathbf{x}^T\mathbf{UU}^T\mathbf{x}}{\mathbf{x}^T\mathbf{x}} = \frac{\lambda_{\max}\mathbf{x}^T\mathbf{x}}{\mathbf{x}^T\mathbf{x}} = \lambda_{\max}.$$

Pri tretej úprave si stačilo uvedomiť, že $\sum_i(\mathbf{u}_i^T\mathbf{x})^2$ je skalárny súčin vektoru $\mathbf{U}^T\mathbf{x}$ samého so sebou, v štvrtej úprave sme využili, že $\mathbf{U}$ je ortogonálna matica.

Navyše $\mathbf{x}_*^T\mathbf{Ax}_* = \lambda_{\max}\mathbf{x}_*^T\mathbf{x}_*$ pre ľubovoľný vlastný vektor $\mathbf{x}_*$ prislúchajúci $\lambda_{\max}$. Preto

$$\frac{\mathbf{x}_*^T\mathbf{Ax}_*}{\mathbf{x}_*^T\mathbf{x}_*} = \lambda_{\max},$$

čiže $\mathbf{x}_*$ je riešením maximalizačnej úlohy. □

Poznámka 10.34. Výraz $\mathbf{x}^T\mathbf{Ax}/(\mathbf{x}^T\mathbf{x})$, kde $\mathbf{A}$ je symetrická matica, sa nazýva *Rayleighov kvocient* (angl. *Rayleigh quotient*). Platí preň dokonca silnejšie tvrdenie než veta 10.33, ktoré formulujeme nižšie.

Veta 10.35 (Vlastné čísla a vektory pomocou Rayleighovho kvocientu)**.** Nech $\mathbf{A} \in \mathcal{S}^n$ a definujme funkciu $f : \mathbb{R}^n \setminus \{\mathbf{0}_n\} \to \mathbb{R}$,

$$f(\mathbf{x}) = \frac{\mathbf{x}^T\mathbf{A}\mathbf{x}}{\mathbf{x}^T\mathbf{x}}.$$

Potom vlastné vektory matice $\mathbf{A}$ sú stacionárne body funkcie f a jej funkčné hodnoty v týchto bodoch sú prislúchajúce vlastné čísla.

Dôkaz. Stacionárne body sú body, v ktorých je nulová derivácia, vypočítajme teda deriváciu f pomocou (10.4):

$$\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} = \frac{\mathbf{x}^T\mathbf{x}(2\mathbf{A}\mathbf{x}) - \mathbf{x}^T\mathbf{A}\mathbf{x}(2\mathbf{x})}{(\mathbf{x}^T\mathbf{x})^2}.$$

Táto derivácia je rovná nule ak $(\mathbf{x}^T\mathbf{x})\mathbf{A}\mathbf{x} = (\mathbf{x}^T\mathbf{A}\mathbf{x})\mathbf{x}$, teda ak

$$\mathbf{A}\mathbf{x} = \frac{\mathbf{x}^T\mathbf{A}\mathbf{x}}{\mathbf{x}^T\mathbf{x}}\mathbf{x},$$

čo je ekvivalentné tomu, že $\mathbf{x}$ je vlastný vektor matice $\mathbf{A}$ a že príslušné vlastné číslo je $\mathbf{x}^T\mathbf{A}\mathbf{x}/(\mathbf{x}^T\mathbf{x})$ (viď lema 10.6). □

Poznámka 10.36. Vlastné vektory prislúchajúce najmenšiemu a najväčšiemu vlastnému číslu sú (lokálnym aj globálnym) maximom resp. minimom Rayleighovho kvocientu, ako sme videli vo vete 10.33. Ostatné vlastné vektory sú sedlovými bodmi Rayleighovho kvocientu, teda v niektorých smeroch sa v ich okolí zvyšuje funkčná hodnota a v niektorých sa znižuje.

Lema 10.37 (Vlastné čísla pozitívne (semi-)definitných matíc)**.** Ak λ je vlastné číslo matice $\mathbf{A} \in \mathcal{S}^n_+$, tak $\lambda \geq 0$. Ak λ je vlastné číslo matice $\mathbf{A} \in \mathcal{S}^n_{++}$, tak $\lambda > 0$.

Dôkaz. Nech λ je vlastné číslo matice $\mathbf{A}$ a nech mu prislúcha vlastný vektor $\mathbf{x} \neq \mathbf{0}_n$. Podľa lemy 10.6 platí $\lambda = (\mathbf{x}^T\mathbf{A}\mathbf{x})/(\|\mathbf{x}\|^2)$, pričom zjavne $\|\mathbf{x}\|^2 > 0$. Pre $\mathbf{A} \in \mathcal{S}^n_+$ je $\mathbf{x}^T\mathbf{A}\mathbf{x} \geq 0$, a teda aj $\lambda \geq 0$. Pre $\mathbf{A} \in \mathcal{S}^n_{++}$ vyplýva $\lambda > 0$ z $\mathbf{x}^T\mathbf{A}\mathbf{x} > 0$. □

Dôsledok 10.38. Nech $\mathbf{A} \in \mathbb{R}^{n\times n}$ je symetrická a nech $\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T$ je jej spektrálny rozklad. Potom ak $\mathbf{A} \in \mathcal{S}^n_+$, tak na diagonále matice $\mathbf{\Lambda}$ sú nezáporné čísla. Ak $\mathbf{A} \in \mathcal{S}^n_{++}$, tak na diagonále matice $\mathbf{\Lambda}$ sú kladné čísla.

Príklad 10.39. Nech $\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T$, kde $\mathbf{U} \in \mathbb{R}^{n\times n}$ je ortogonálna, $\mathbf{\Lambda} = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$, $\lambda_i \geq 0$ pre každé $i = 1, \dots, n$. Ukážte, že potom $\mathbf{A} \in \mathcal{S}^n_+$. Ukážte tiež, že ak navyše $\lambda_i > 0$ pre každé $i = 1, \dots, n$, tak $\mathbf{A} \in \mathcal{S}^n_{++}$.

Podľa lemy 10.37 a príkladu 10.39 je symetrická matica $\mathbf{A}$ pozitívne semidefinitná (pozitívne definitná) vtedy a len vtedy, keď všetky jej vlastné čísla sú nezáporné (kladné). Iná charakterizácia pozitívnej definitnosti sa dá sformulovať pomocou determinantov ľavých horných submatíc matice $\mathbf{A}$, nazýva sa *Sylvestrovo kritérium.*

Poznámka 10.40. Nech $\mathbf{A} \in \mathbb{R}^{n\times n}$. Potom pre $k = 1, \dots, n$ pod *ľavou hornou submaticou* typu $k \times k$ matice $\mathbf{A}$ myslíme submaticu, ktorá vznikla z matice $\mathbf{A}$ zmazaním stĺpcov $k+1, \dots, n$ a riadkov $k+1, \dots, n$. Ľavou hornou submaticou typu $k \times k$ matice $\mathbf{A}$ pre $k = n$ je pôvodná matica $\mathbf{A}$.

Veta 10.41 (Sylvestrovo kritérium)**.** Nech $\mathbf{A} \in \mathbb{R}^{n \times n}$ je symetrická matica. Potom $\mathbf{A} \in \mathcal{S}^n_{++}$ práve vtedy, keď determinanty všetkých jej ľavých horných submatíc sú kladné.

Dôkaz. Túto vetu nebudeme dokazovať, jej dôkaz je možné nájsť v [14]. □

Poznámka 10.42. Ako prirodzené rozšírenie vety 10.41 sa ponúka zameniť $\mathcal{S}^n_{++}$ za $\mathcal{S}^n_+$ a kladné determinanty za nezáporné. Také tvrdenie však neplatí. Aj keď totiž determinanty všetkých ľavých horných submatíc matice $\mathbf{A}$ sú nezáporné, tak matica $\mathbf{A}$ ešte nemusí byť pozitívne semidefinitná. Ako protipríklad môžeme uvážiť napríklad maticu

$$\mathbf{A} = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}.$$

Príklad 10.43. Pomocou Sylvestrovho kritéria ukážte, že matica

$$\mathbf{A} = \begin{pmatrix} 6 & -1 & 3 \\ -1 & 5 & -1 \\ 3 & -1 & 9 \end{pmatrix}$$

je pozitívne definitná.

Lema 10.44 (Vyjadrenie charakteristík symetrickej matice pomocou vlastných čísel)**.** Nech $\mathbf{A} \in \mathcal{S}^n$ má vlastné čísla $\lambda_1, \dots, \lambda_n$ zopakované podľa násobnosti. Potom

(i) $\operatorname{tr}(\mathbf{A}) = \sum_{i=1}^n \lambda_i$,

(ii) $\|\mathbf{A}\|^2 = \sum_{i=1}^n \lambda_i^2$,

(iii) $\det(\mathbf{A}) = \prod_{i=1}^n \lambda_i$,

(iv) $\operatorname{rank}(\mathbf{A}) = |\{i \in \{1, \dots, n\} \mid \lambda_i \neq 0\}|$.

Dôkaz. Využijeme zápis $\mathbf{A} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^T$, kde $\mathbf{U}$ je ortogonálna a $\boldsymbol{\Lambda} = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$, podľa vety 10.27. Potom $\operatorname{tr}(\mathbf{A}) = \operatorname{tr}(\mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^T) = \operatorname{tr}(\boldsymbol{\Lambda}\mathbf{U}^T\mathbf{U}) = \operatorname{tr}(\boldsymbol{\Lambda}) = \sum_i \lambda_i$, vďaka vete 5.12 o výmene poradia v stope. Podobne získame $\|\mathbf{A}\|^2 = \operatorname{tr}(\mathbf{A}^T\mathbf{A}) = \operatorname{tr}(\mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^T\mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^T) = \operatorname{tr}(\mathbf{U}\boldsymbol{\Lambda}^2\mathbf{U}^T) = \operatorname{tr}(\boldsymbol{\Lambda}^2\mathbf{U}^T\mathbf{U}) = \operatorname{tr}(\boldsymbol{\Lambda}^2) = \sum_i \lambda_i^2$. Z vety 8.11(vii) plynie $\det(\mathbf{A}) = \det(\mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^T) = \det(\mathbf{U})\det(\boldsymbol{\Lambda})\det(\mathbf{U}^T) = \det(\boldsymbol{\Lambda})\det(\mathbf{U}\mathbf{U}^T) = \det(\boldsymbol{\Lambda})\det(\mathbf{I}_n) = \det(\boldsymbol{\Lambda}) = \prod_i \lambda_i$. Podľa lemy 3.36 a lemy 10.15 platí $\operatorname{rank}(\mathbf{A}) = n - \dim(\mathcal{N}(\mathbf{A})) = n - q$, kde q je násobnosť nulového vlastného čísla. Zároveň počet vlastných čísel (započítaných podľa násobnosti) je podľa dôsledku 10.28 rovný n, čiže $n - q$ je počet nenulových vlastných čísel. □

Dôsledok 10.45. Ak $\mathbf{A} \in \mathcal{S}^n_+$, tak $\det(\mathbf{A}) \geq 0$. Ak $\mathbf{A} \in \mathcal{S}^n_{++}$, tak $\det(\mathbf{A}) > 0$.

Poznámka 10.46. Úrovňové množiny (vrstevnice) $U_t = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x}^T\mathbf{A}\mathbf{x} = t\}$, $t > 0$, kvadratickej funkcie $\mathbf{x}^T\mathbf{A}\mathbf{x}$ pre $\mathbf{A} \in \mathcal{S}^n_{++}$ sú elipsoidy s polosami v smere vlastných vektorov matice $\mathbf{A}$, pričom dĺžky týchto polosí sú $\sqrt{t}/\sqrt{\lambda_1}, \dots, \sqrt{t}/\sqrt{\lambda_n}$, kde $\lambda_1, \dots, \lambda_n$ sú vlastné čísla $\mathbf{A}$ zopakované podľa násobnosti.

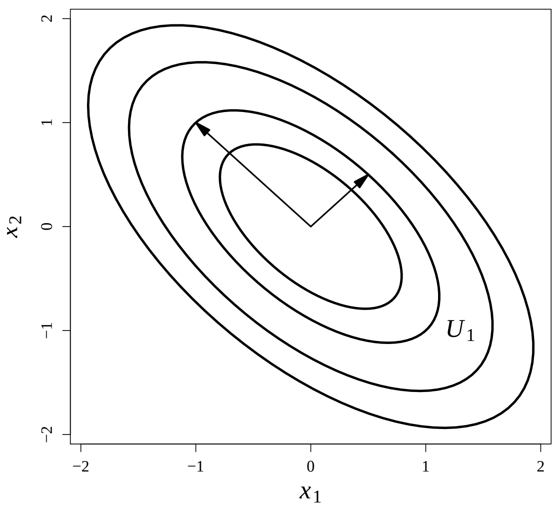
Toto tvrdenie nebudeme formálne dokazovať, iba naznačíme, prečo platí. Uvažujme pre jednoduchosť U_1, čiže rovnicu $\mathbf{x}^T\mathbf{A}\mathbf{x} = 1$. Využime spektrálny rozklad $\mathbf{A} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^T$. Potom $\mathbf{x}^T\mathbf{A}\mathbf{x} = \mathbf{x}^T\mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^T\mathbf{x} = \mathbf{y}^T\boldsymbol{\Lambda}\mathbf{y} = \lambda_1 y_1^2 + \ldots + \lambda_n y_n^2$, kde $\mathbf{y} = (y_1, \dots, y_n)^T = \mathbf{U}^T\mathbf{x}$. Rovnica $\lambda_1 y_1^2 + \ldots + \lambda_n y_n^2 = 1$ je zjavne rovnicou elipsoidu so stredom v počiatku súradnicovej sústavy a s polosami v smere súradnicových osí. Dĺžku i-tej polosi určíme tak, že položíme $y_j = 0$ pre $j \neq i$, z čoho dostávame $y_i = \pm 1/\sqrt{\lambda_i}$, čiže dĺžka i-tej polosi pre U_1 je $1/\sqrt{\lambda_i}$.

Takto sme ale vyjadrili vrstevnice vzhľadom na vektor $\mathbf{y} = \mathbf{U}^T\mathbf{x}$, určený vlastnými vektormi. Keďže ortogonálne matice sú maticami otočenia a preklopenia, zachovávajúce dĺžku vektora, tak polosi v pôvodných súradniciach majú rovnakú dĺžku, ale sú otočené tak, že sú v smeroch vlastných vektorov matice $\mathbf{A}$. To môžeme nahliadnuť tak, že uvážime, kedy nastáva $y_i = 1/\sqrt{\lambda_i}$ a $y_j = 0$ pre $j \neq i$, teda kedy sa nachádzame na i-tej polosi. Keďže $\mathbf{y} = \mathbf{U}^T\mathbf{x}$, tak riešime $\mathbf{U}^T\mathbf{x} = \mathbf{e}_i/\sqrt{\lambda_i}$, čo má riešenie $\mathbf{x} = \mathbf{U}\mathbf{e}_i/\sqrt{\lambda_i} = \mathbf{u}_i/\sqrt{\lambda_i}$, kde $\mathbf{u}_i$ je i-ty vlastný vektor $\mathbf{A}$. Teda i-ta polos je v smere vlastného vektora $\mathbf{u}_i$ a jej dĺžka je $\|\mathbf{u}_i/\sqrt{\lambda_i}\| = 1/\sqrt{\lambda_i}$, ako sme už ukázali vyššie.

Príklad 10.47. Pre maticu $\mathbf{A}$ z poznámky 10.5 už poznáme jej spektrálny rozklad. Potom

$$\mathbf{x}^T\mathbf{A}\mathbf{x} = \frac{1}{2}(x_1 + x_2, x_2 - x_1)\begin{pmatrix} 2 & 0 \\ 0 & 1/2 \end{pmatrix}\begin{pmatrix} x_1 + x_2 \\ x_2 - x_1 \end{pmatrix} = (x_1 + x_2)^2 + \frac{1}{4}(x_2 - x_1)^2$$

a úrovňové množiny tejto kvadratickej funkcie sú zakreslené na obrázku 10.3. Polosi jednotlivých elíps sú v smeroch vlastných vektorov, čiže v smeroch $(1/\sqrt{2}, 1/\sqrt{2})^T$ a $(-1/\sqrt{2}, 1/\sqrt{2})^T$. Dĺžka dlhšej polosi elipsy U_1 je $1/\sqrt{\lambda_2} = \sqrt{2}$, a teda táto polos je úsečka spájajúca $\mathbf{0}_2$ a $(-1, 1)^T$. Kratšia polos spája $\mathbf{0}_2$ s $\mathbf{u}_1/\sqrt{\lambda_1} = (1/2, 1/2)^T$. Overte, že body $\mathbf{u}_1/\sqrt{\lambda_1}$ a $\mathbf{u}_2/\sqrt{\lambda_2}$ ozaj ležia na elipse U_1.



Obr. 10.3: Znázornenie úrovňových množín $U_{1/2}$, U_1, U_2 a U_3 kvadratickej formy zadanej v príklade 10.47. Šípky znázorňujú polosi elipsy U_1, čiže $\mathbf{u}_1/\sqrt{\lambda_1}$ a $\mathbf{u}_2/\sqrt{\lambda_2}$.

10.2.1 Metóda hlavných komponentov

Spektrálny rozklad výberovej kovariančnej matice $\mathbf{S}$ (definovanej v podkapitole 2.6.1) je možné využiť na opísanie dát a redukciu dimenzie dát. Výberová kovariančná matica $\mathbf{S} = \frac{1}{n-1}(\mathbf{X} - n^{-1}\mathbf{X}\mathbf{J}_{n\times n})(\mathbf{X} - n^{-1}\mathbf{X}\mathbf{J}_{n\times n})^T \in \mathbb{R}^{d\times d}$ je zjavne symetrická a podľa úlohy 9.5 je aj pozitívne semidefinitná. Teda pre ňu platí veta 10.27. Usporiadajme vlastné čísla (zopakované podľa násobnosti) matice $\mathbf{S}$ zostupne: $\lambda_1 \geq \ldots \geq \lambda_d$ a pripomeňme, že príslušné vlastné vektory značíme $\mathbf{u}_1, \ldots, \mathbf{u}_d$.

Najprv dáta vycentrujme, t.j. od každej zložky odpočítajme jej výberový priemer: získavame $\mathbf{x}_1 - \bar{\mathbf{x}}, \ldots, \mathbf{x}_n - \bar{\mathbf{x}}$ (pripomeňme, že $\mathbf{x}_1, \ldots, \mathbf{x}_n$ sú stĺpce matice $\mathbf{X}$, a že $\bar{\mathbf{x}}$ je vektor priemerov $\bar{\mathbf{x}} = n^{-1}\mathbf{X}\mathbf{J}_{n\times n}$). Potom teoretické výsledky o spektrálnom rozklade kovariančnej matice náhodného vektora sa dajú neformálne sformulovať tak, že dáta sú najviac rozptýlené v smere prvého vlastného vektora. Následne druhý vlastný vektor určuje smer najväčšej rozptýlenosti v ortogonálnom doplnku k prvému vlastnému vektoru. Podobne tretí vlastný vektor hovorí o najväčšej rozptýlenosti v ortogonálnom doplnku k podpriestoru určenému prvými dvoma vlastnými vektormi atď. Poznamenajme, že tieto vlastnosti úzko súvisia s charakterizáciou vlastných vektorov a čísel pomocou extremálnych a sedlových bodov Rayleighovho kvocientu (viď vety 10.33 a 10.35).

Toto pozorovanie je ilustrované na obrázkoch 7.2a a 7.2b z kapitoly 7 - priamka na prvom obrázku je určená prvým vlastným vektorom výberovej kovariančnej matice pre príslušné dáta, zatiaľ čo priamka na druhom obrázku je určená druhým (a teda posledným) vlastným vektorom tejto matice. Skutočne, môžeme pozorovať, že dáta sú najviac rozptýlené v smere prvého vlastného vektora. Prvý vlastný vektor výberovej kovariančnej matice sa pri takomto použití nazýva prvý *hlavný komponent*, podobne i-ty vlastný vektor je i-ty hlavný komponent ($i = 1, \ldots, n$) a metóda, ktorá používa vlastné komponenty na opis a analýzu dát (najmä na redukciu dimenzie) sa nazýva *metóda hlavných komponentov.* Ak $\mathbf{x}_1, \ldots, \mathbf{x}_n$ sú mnohorozmerné dáta a snažíme sa ich zachytiť v menšom počte rozmerov (napríklad v dvoch rozmeroch, aby sa dali zakresliť do roviny), je metóda hlavných komponentov prirodzeným nástrojom na zníženie dimenzionality. Na opis, ďalšiu prácu, analýzu, či zakreslenie dát v takom prípade použijeme dáta vyjadrené v priestore prvých k hlavných komponentov pre zvolené $k < d$. Metódu hlavných komponentov sme tu iba stručne načrtli, podrobnejšie sa jej venujú pokročilé prednášky z matematickej štatistiky.

Poznámka 10.48. Vyjadrenie vektora $\mathbf{x}_i$ ($i = 1, \ldots, n$) v súradnicovej sústave danej prvými k hlavnými komponentmi $\mathbf{u}_1, \ldots, \mathbf{u}_k$ znamená nájdenie koeficientov $c_1, \ldots, c_k \in \mathbb{R}$, že $\mathbf{x}_i = c_1\mathbf{u}_1 + \ldots + c_k\mathbf{u}_k$. To sa dá prepísať ako $\mathbf{x}_i = \mathbf{U}_k\mathbf{c}$, kde $\mathbf{c} = (c_1, \ldots, c_k)^T$ a $\mathbf{U}_k = (\mathbf{u}_1, \ldots, \mathbf{u}_k)$. Keďže $\mathbf{u}_1, \ldots, \mathbf{u}_k$ sú ortonormálne vektory, tak $\mathbf{U}_k^T\mathbf{U}_k = \mathbf{I}_k$, čiže koeficienty vieme získať zo vzorca $\mathbf{c} = \mathbf{U}_k^T\mathbf{x}_i$.

10.3 Mocnina pozitívne semidefinitnej matice

Poznámka 10.49. Nech $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ a nech $k \in \mathbb{N}$. Označme $\mathbf{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. Potom zjavne $\mathbf{\Lambda}^k = \mathrm{diag}(\lambda_1^k, \ldots, \lambda_n^k)$. Navyše ak všetky $\lambda_i > 0$, tak $\mathbf{\Lambda}^{-1} = \mathrm{diag}(\lambda_1^{-1}, \ldots, \lambda_n^{-1})$. V nasledujúcej definícii rozšírime tento predpis na všetky $k > 0$ pre $\lambda_1, \ldots, \lambda_n \geq 0$ a na všetky $k \in \mathbb{R}$ pre $\lambda_1, \ldots, \lambda_n > 0$.

Definícia 10.50 (Mocnina pozitívne semidefinitnej diagonálnej matice)**.** Nech $\lambda_1, \ldots, \lambda_n \geq 0$. Potom pre ľubovoľné $k > 0$ je *k-tou mocninou* diagonálnej matice $\mathbf{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ matica $\mathbf{\Lambda}^k = \mathrm{diag}(\lambda_1^k, \ldots, \lambda_n^k)$. Ak $\lambda_1, \ldots, \lambda_n > 0$, tak pre ľubovoľné $k \in \mathbb{R}$ je *k-tou mocninou* diagonálnej matice $\mathbf{\Lambda}$ matica $\mathbf{\Lambda}^k = \mathrm{diag}(\lambda_1^k, \ldots, \lambda_n^k)$.

Príklad 10.51. Ukážte, že ak $\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T$ je spektrálny rozklad matice $\mathbf{A} \in \mathcal{S}_+^n$, tak $\mathbf{A}^k = \mathbf{U}\mathbf{\Lambda}^k\mathbf{U}^T$ pre akékoľvek $k \in \mathbb{N}$.

Lema 10.52 (Inverzia pozitívne definitnej matice pomocou spektrálneho rozkladu)**.** Nech $\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T$ je spektrálny rozklad matice $\mathbf{A} \in \mathcal{S}_{++}^n$. Potom $\mathbf{A}^{-1} = \mathbf{U}\mathbf{\Lambda}^{-1}\mathbf{U}^T$.

Dôkaz. Máme $\mathbf{A}^{-1} = (\mathbf{U}\mathbf{\Lambda}\mathbf{U}^T)^{-1} = \mathbf{U}\mathbf{\Lambda}^{-1}\mathbf{U}^T$, keďže $\mathbf{U}$ aj $\mathbf{\Lambda}$ sú regulárne, pričom $\mathbf{U}^{-1} = \mathbf{U}^T$. □

Definícia 10.53 (Odmocninová matica pozitívne semidefinitnej matice)**.** Nech $\mathbf{A} \in \mathcal{S}_+^n$ a nech $\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T$ je jej spektrálny rozklad. Potom *odmocninová matica* $\mathbf{A}^{1/2}$ matice $\mathbf{A}$ je $\mathbf{A}^{1/2} = \mathbf{U}\mathbf{\Lambda}^{1/2}\mathbf{U}^T$.

Lema 10.54 (Vlastnosti odmocninovej matice)**.** Nech $\mathbf{A} \in \mathcal{S}_+^n$ a nech $\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T$ je jej spektrálny rozklad, kde $\mathbf{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ a $\mathbf{U} = (\mathbf{u}_1, \ldots, \mathbf{u}_n)$. Potom

(i) $\mathbf{A}^{1/2} \in \mathcal{S}_+^n$, a ak $\mathbf{A} \in \mathcal{S}_{++}^n$, tak aj $\mathbf{A}^{1/2} \in \mathcal{S}_{++}^n$,

(ii) vlastné čísla matice $\mathbf{A}^{1/2}$ zopakované podľa násobnosti sú $\lambda_1^{1/2}, \ldots, \lambda_n^{1/2}$ a vlastné vektory k nim prislúchajúce sú v poradí $\mathbf{u}_1, \ldots, \mathbf{u}_n$,

(iii) $\mathbf{A}^{1/2}\mathbf{A}^{1/2} = \mathbf{A}$.

Dôkaz. Tvrdenie (i) platí podľa príkladu 10.39. Pre (ii) zvoľme ľubovoľné $i = 1, \ldots, n$ a počítajme $\mathbf{A}^{1/2}\mathbf{u}_i = \mathbf{U}\mathbf{\Lambda}^{1/2}\mathbf{U}^T\mathbf{u}_i = \mathbf{U}\mathbf{\Lambda}^{1/2}\mathbf{e}_i$, keďže $\mathbf{u}_1, \ldots, \mathbf{u}_n$ sú ortonormálne. To vieme ďalej upraviť na $\mathbf{U}\lambda_i^{1/2}\mathbf{e}_i = \lambda_i^{1/2}\mathbf{u}_i$, čím je tvrdenie dokázané. Dôkaz (iii) sa ľahko spraví priamym výpočtom, nechávame ho na čitateľa. □

Príklad 10.55. Nech

$$\mathbf{A} = \begin{pmatrix} 5 & 4 \\ 4 & 5 \end{pmatrix}.$$

Potom jej vlastné čísla sú $\lambda_1 = 9$, $\lambda_2 = 1$ a príslušné vlastné vektory sú $\mathbf{u}_1 = (1/\sqrt{2}, 1/\sqrt{2})^T$, $\mathbf{u}_2 = (-1/\sqrt{2}, 1/\sqrt{2})^T$. Teda

$$\mathbf{A}^{1/2} = \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

Presvedčte sa, že vlastné čísla $\mathbf{A}^{1/2}$ sú 3 a 1 s príslušnými vlastnými vektormi $\mathbf{u}_1$ a $\mathbf{u}_2$. Overte tiež, že $\mathbf{A}^{1/2}\mathbf{A}^{1/2} = \mathbf{A}$.

Definícia 10.56 (Mocnina pozitívne semidefinitnej matice)**.** Nech $\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T$ je spektrálny rozklad matice $\mathbf{A} \in \mathcal{S}_+^n$. Potom pre ľubovoľné $k > 0$ je *k-ta mocnina* matice $\mathbf{A}$ matica $\mathbf{A}^k = \mathbf{U}\mathbf{\Lambda}^k\mathbf{U}^T$. Ak $\mathbf{A} \in \mathcal{S}_{++}^n$, tak pre ľubovoľné $k \in \mathbb{R}$ je *k-ta mocnina* matice $\mathbf{A}$ matica $\mathbf{A}^k = \mathbf{U}\mathbf{\Lambda}^k\mathbf{U}^T$.

Poznámka 10.57. Definícia 10.56 rozširuje vzťah $\mathbf{A}^k = \mathbf{U}\mathbf{\Lambda}^k\mathbf{U}^T$ z príkladu 10.51 a lemy 10.52 na všetky $k > 0$, resp. $k \in \mathbb{R}$.

Veta 10.58 (Vlastnosti mocninovej matice)**.** Nech $\mathbf{A} \in \mathcal{S}_+^n$ a nech $\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T$ je jej spektrálny rozklad, kde $\mathbf{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ a $\mathbf{U} = (\mathbf{u}_1, \ldots, \mathbf{u}_n)$. Potom pre $k > 0$ platí

(i) $\mathbf{A}^k \in \mathcal{S}_+^n$,

(ii) vlastné čísla matice $\mathbf{A}^k$ zopakované podľa násobnosti sú $\lambda_1^k, \dots, \lambda_n^k$ a vlastné vektory k nim prislúchajúce sú v poradí $\mathbf{u}_1, \dots, \mathbf{u}_n$.

Ak $\mathbf{A} \in \mathcal{S}_{++}^n$, tak pre $k \in \mathbb{R}$ platí

(iii) $\mathbf{A}^k \in \mathcal{S}_{++}^n$,

(iv) vlastné čísla matice $\mathbf{A}^k$ zopakované podľa násobnosti sú $\lambda_1^k, \dots, \lambda_n^k$ a vlastné vektory k nim prislúchajúce sú v poradí $\mathbf{u}_1, \dots, \mathbf{u}_n$,

(v) ak $z > 0$, tak $\mathbf{A}^{-z} = (\mathbf{A}^{-1})^z = (\mathbf{A}^z)^{-1}$.

Dôkaz. Dôkazy častí (i)-(iv) sú analogické k dôkazu lemy 10.54, nechávame ich na čitateľa. Pre dôkaz (v) uvážme, že platí $\mathbf{A}^{-z} = \mathbf{U}\boldsymbol{\Lambda}^{-z}\mathbf{U}^T$. Zároveň $(\mathbf{A}^{-1})^z = \mathbf{U}(\boldsymbol{\Lambda}^{-1})^z\mathbf{U}^T = \mathbf{U}\boldsymbol{\Lambda}^{-z}\mathbf{U}^T$ podľa definície 10.56, keďže $\mathbf{U}\boldsymbol{\Lambda}^{-1}\mathbf{U}^T$ je spektrálny rozklad matice $\mathbf{A}^{-1}$ (lema 10.52). Teda $\mathbf{A}^{-z} = (\mathbf{A}^{-1})^z$. Navyše $\mathbf{A}^z = \mathbf{U}\boldsymbol{\Lambda}^z\mathbf{U}^T$, a teda $(\mathbf{A}^z)^{-1} = \mathbf{U}(\boldsymbol{\Lambda}^z)^{-1}\mathbf{U}^T$. Zjavne $(\boldsymbol{\Lambda}^z)^{-1} = \boldsymbol{\Lambda}^{-z}$, čiže $(\mathbf{A}^z)^{-1} = \mathbf{U}\boldsymbol{\Lambda}^{-z}\mathbf{U}^T = \mathbf{A}^{-z}$. □

Poznámka 10.59. Vo vete 10.58 máme špeciálne $\mathbf{A}^{-1/2} = \mathbf{U}\boldsymbol{\Lambda}^{-1/2}\mathbf{U}^T$ pre $\mathbf{A} \in \mathcal{S}_{++}^n$.

Lema 10.60 (O súčasnej diagonalizácii)**.** Nech $\mathbf{A} \in \mathcal{S}_{++}^n$ a $\mathbf{B} \in \mathcal{S}^n$. Potom existujú regulárna matica $\mathbf{V} \in \mathbb{R}^{n\times n}$ a diagonálna matica $\mathbf{D} \in \mathbb{R}^{n\times n}$, ktoré spĺňajú $\mathbf{V}^T\mathbf{A}\mathbf{V} = \mathbf{I}_n$ a $\mathbf{V}^T\mathbf{B}\mathbf{V} = \mathbf{D}$.

Dôkaz. Keďže $\mathbf{A}$ je pozitívne definitná, tak existuje $\mathbf{A}^{-1/2}$. Potom matica $\mathbf{A}^{-1/2}\mathbf{B}\mathbf{A}^{-1/2}$ je symetrická; označme teda jej spektrálny rozklad ako $\mathbf{A}^{-1/2}\mathbf{B}\mathbf{A}^{-1/2} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^T$. Ak vezmeme $\mathbf{V} = \mathbf{A}^{-1/2}\mathbf{U}$, tak dostávame $\mathbf{V}^T\mathbf{A}\mathbf{V} = \mathbf{U}^T\mathbf{A}^{-1/2}\mathbf{A}\mathbf{A}^{-1/2}\mathbf{U} = \mathbf{U}^T\mathbf{A}^{-1/2}\mathbf{A}^{1/2}\mathbf{A}^{1/2}\mathbf{A}^{-1/2}\mathbf{U} = \mathbf{U}^T\mathbf{U} = \mathbf{I}_n$, keďže matica $\mathbf{U}$ zo spektrálneho rozkladu je ortogonálna. Podobne platí $\mathbf{V}^T\mathbf{B}\mathbf{V} = \mathbf{U}^T\mathbf{A}^{-1/2}\mathbf{B}\mathbf{A}^{-1/2}\mathbf{U} = \mathbf{U}^T\mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^T\mathbf{U} = \boldsymbol{\Lambda}$, kde matica $\boldsymbol{\Lambda}$ je diagonálna. Matica $\mathbf{V}$ je regulárna, lebo je súčinom dvoch regulárnych matíc. Teda matice $\mathbf{V}$ a $\boldsymbol{\Lambda}$ spĺňajú požadované rovnosti. □

10.4 Viacrozmerné normálne rozdelenie

Pochopenie tejto časti vyžaduje od čitateľa znalosť niektorých základných pojmov z pravdepodobnosti a štatistiky, ako sú náhodná premenná a náhodný vektor, hustota (náhodného vektora), stredná hodnota, disperzia, kovariancia a normálne rozdelenie.

Medzi najčastejšie používané rozdelenia v štatistike patria jednorozmerné a viacrozmerné normálne rozdelenia. *Jednorozmerné normálne rozdelenie* so strednou hodnotou μ a disperziou $\sigma^2 > 0$ (značíme $N(\mu, \sigma^2)$) je definované svojou hustotou

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad x \in \mathbb{R}.$$

Viacrozmerné normálne rozdelenie je zovšeobecnením jednorozmerného, je to rozdelenie vektora náhodných premenných $\mathbf{X} = (X_1, \dots, X_d)^T$ (tzv. *náhodného vektora*). Značíme $\mathbf{X} \sim N_d(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, kde $\boldsymbol{\mu} = (\mu_1, \dots, \mu_d)$ je vektor stredných hodnôt náhodných premenných $X_1, \dots, X_d$ a $\boldsymbol{\Sigma}$ je tzv. *kovariančná matica* vektora $\mathbf{X}$. Pre prvky kovariančnej matice $\boldsymbol{\Sigma} \in \mathbb{R}^{d\times d}$ platí $(\boldsymbol{\Sigma})_{ij} = \mathrm{Cov}(X_i, X_j)$ pre $i \neq j$ a $(\boldsymbol{\Sigma})_{ii} = D(X_i)$. Kovariančná matica teda vyjadruje rozptýlenosti jednotlivých zložiek náhodného vektora a miery lineárnych závislostí medzi zložkami.

Nebudeme uvádzať formálnu definíciu viacrozmerného normálneho rozdelenia, platí však, že každá zložka viacrozmerného normálneho vektora X_i má jednorozmerné normálne rozdelenie $X_i \sim N(E(X_i), D(X_i))$. Zároveň, ak $\mathbf{X} \sim N_d(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ a $\boldsymbol{\Sigma}$ je pozitívne definitná, tak $\mathbf{X}$ má d-rozmernú hustotu

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{d/2}(\det(\boldsymbol{\Sigma}))^{1/2}} e^{-(\mathbf{x}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})/2}, \quad \mathbf{x} \in \mathbb{R}^d.$$

Príklad 10.61. Overte, že pre $\boldsymbol{\Sigma} \in \mathcal{S}^d_{++}$ je hustota viacrozmerného normálneho rozdelenia dobre definovaná funkcia. Zdôvodnite teda, že v predpise nedochádza k odmocneniu záporného čísla, deleniu nulou ani k invertovaniu singulárnej matice. Zdôvodnite tiež, že $f(\mathbf{x}) \geq 0$ pre každé $\mathbf{x} \in \mathbb{R}^d$, a že $f(\mathbf{x})$ má maximum v bode $\mathbf{x} = \boldsymbol{\mu}$.

Všimnite si, že vrstevnice $f(\mathbf{x}) = c$ ($c > 0$) hustoty viacrozmerného normálneho rozdelenia sú určené ako $(\mathbf{x}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}) = k$ pre nejaké $k > 0$, sú to teda elipsoidy dané pozitívne definitnou kvadratickou funkciou $\mathbf{y}^T\boldsymbol{\Sigma}^{-1}\mathbf{y}$ so stredom v bode $\boldsymbol{\mu}$. Tvar vrstevníc je podľa poznámky 10.46 určený vlastnými číslami a vlastnými vektormi matice $\boldsymbol{\Sigma}^{-1}$. Podobne ako hustota jednorozmerného normálneho rozdelenia tvorí “kopec” so stredom v bode $\mu \in \mathbb{R}$ (viď obrázok 10.4a), tak hustota viacrozmerného normálneho rozdelenia tvorí d-rozmerný “kopec” so stredom v bode $\boldsymbol{\mu} \in \mathbb{R}^d$. Tvar kopca je potom určený kvadratickou funkciou danou inverziou kovariančnej matice $\boldsymbol{\Sigma}$. Na obrázku 10.4b sú zobrazené vrstevnice hustoty dvojrozmerného normálneho rozdelenia $N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, kde

$$\boldsymbol{\mu} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \boldsymbol{\Sigma} = \begin{pmatrix} 3{,}7 & 0{,}7 \\ 0{,}7 & 2{,}3 \end{pmatrix}. \tag{10.5}$$

Poznamenajme, že vlastné čísla matice $\boldsymbol{\Sigma}$ sú približne $\lambda_1 = 4$, $\lambda_2 = 2$ a k nim prislúchajúce vlastné vektory sú približne $\mathbf{u}_1 = (0{,}92; 0{,}38)^T$ a $\mathbf{u}_2 = (0{,}38; -0{,}92)^T$. Všimnite si, že na získanie tvaru vrstevníc hustoty rozdelenia $N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ nepotrebujete vďaka leme 10.18 vypočítať $\boldsymbol{\Sigma}^{-1}$.

Príklad 10.62. Načrtnite úrovňové množiny hustoty rozdelenia $N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, kde $\boldsymbol{\mu} = \mathbf{0}_2$ a

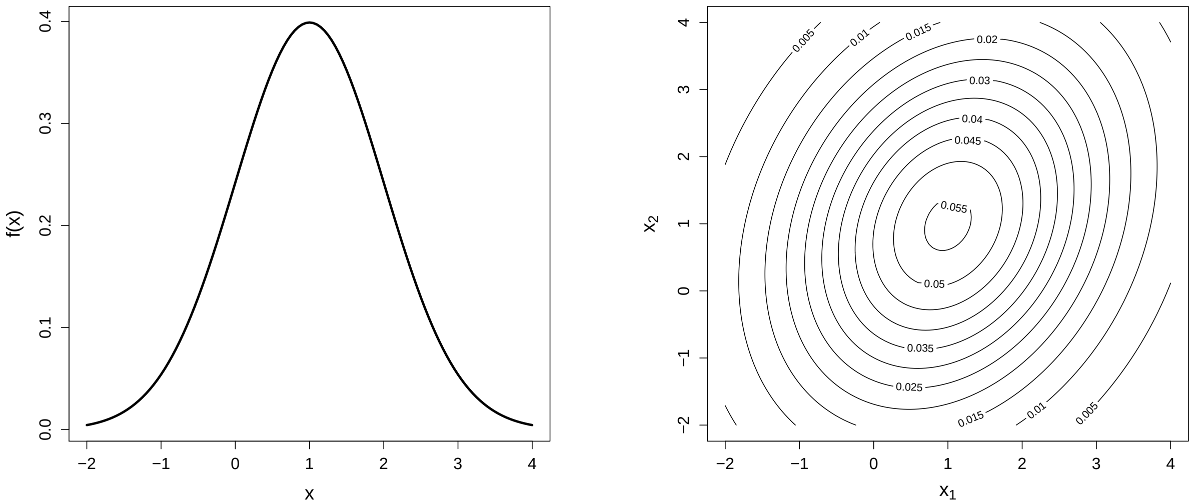
$$\boldsymbol{\Sigma} = \frac{1}{5}\begin{pmatrix} 9 & 2 \\ 2 & 6 \end{pmatrix}.$$

Ako budú tieto vrstevnice vyzerať ak stredná hodnota sa zmení na $\boldsymbol{\mu} = (-1, -1)^T$ a kovariančná matica $\boldsymbol{\Sigma}$ sa zachová?

10.5 Vlastné čísla stochastických matíc

Definícia 10.63 (Ľavá a pravá stochastická matica)**.** Maticu $\mathbf{A} \in \mathbb{R}^{n\times n}$, ktorej všetky prvky sú nezáporné, nazývame

1. *ľavá stochastická* ak súčet prvkov v každom z jej stĺpcov je rovný 1,
2. *pravá stochastická* ak súčet prvkov v každom z jej riadkov je rovný 1.



(a) Hustota premennej $X \sim N(1,1)$.

(b) Vrstevnice hustoty $\mathbf{X} \sim N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

Obr. 10.4: Zakreslenie hustoty jednorozmerného a dvojrozmerného normálneho rozdelenia. Parametre rozdelenia $N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ v časti (b) sú dané v (10.5), čísla na krivkách reprezentujú hodnoty, ktoré hustota na daných úrovňových množinách nadobúda.

Poznámka 10.64. Ak $\mathbf{A} \in \mathbb{R}^{n\times n}$ je pravá stochastická matica, tak spĺňa $\mathbf{A}\mathbf{1}_n = \mathbf{1}_n$. Vektor $\mathbf{1}_n$ je teda vlastným vektorom každej pravej stochastickej matice prislúchajúci vlastnému číslu 1. Analogicky, ak $\mathbf{A} \in \mathbb{R}^{n\times n}$ je ľavá stochastická matica, tak $\mathbf{1}_n$ je vlastným vektorom matice $\mathbf{A}^T$ prislúchajúci vlastnému číslu 1.

Poznámka 10.65. Vlastný vektor (vlastné číslo) matice $\mathbf{A}^T$ sa niekedy nazýva *ľavý vlastný vektor* (*ľavé vlastné číslo*) matice $\mathbf{A}$.

Veta 10.66 (Perronova-Frobeniova veta)**.** Nech $\mathbf{A} \in \mathbb{R}^{n\times n}$ je matica, ktorej všetky prvky sú kladné. Potom existuje $\lambda^* > 0$, ktoré spĺňa:

(i) λ^* je vlastné číslo matice $\mathbf{A}$ aj matice $\mathbf{A}^T$ a v oboch prípadoch má násobnosť 1.

(ii) Ak λ je vlastné číslo matice $\mathbf{A}$ alebo matice $\mathbf{A}^T$, tak $|\lambda| < \lambda^*$.

(iii) Existuje vlastný vektor $\mathbf{u} \in \mathbb{R}^n$ matice $\mathbf{A}$ prislúchajúci ku λ^*, ktorý má všetky zložky kladné. Existuje vlastný vektor $\mathbf{v} \in \mathbb{R}^n$ matice $\mathbf{A}^T$ prislúchajúci ku λ^*, ktorý má všetky zložky kladné.

(iv) Okrem vektora $\mathbf{u}$ (vektora $\mathbf{v}$) a jeho násobkov neexistujú iné vlastné vektory matice $\mathbf{A}$ (matice $\mathbf{A}^T$), ktorých všetky prvky sú nezáporné.

(v) Vezmime také násobky vektorov $\mathbf{u}$ a $\mathbf{v}$ z predchádzajúcich bodov, ktoré spĺňajú $\mathbf{u}^T\mathbf{v} = 1$. Potom $\lim_{k\to\infty} \mathbf{A}^k/(\lambda^*)^k = \mathbf{u}\mathbf{v}^T$.

Dôkaz. Túto vetu nebudeme dokazovať, dôkaz môže čitateľ nájsť v [12]. □

Poznámka 10.67. Perronova-Frobeniova veta v skutočnosti poskytuje analogické tvrdenia aj pre matice, ktorých všetky prvky sú nezáporné (nie nutne kladné) za istých technických predpokladov (tieto matice musia byť neredukovateľné, angl. *irreducible*). Pre jednoduchosť sa však v tejto podkapitole venujeme iba maticiam s kladnými zložkami.

10.5.1 Markovovské reťazce

Markovovský reťazec opisuje vývoj nejakého systému v čase. Uvažujeme diskrétny čas, t.j. časy $t = 0, 1, 2, \ldots$ a v každom čase môže skúmaný systém nadobúdať niektorú hodnotu z konečnej množiny stavov S. Formálne je v každom čase $t = 0, 1, 2, \ldots$ systém opísaný náhodnou premennou X_t, ktorá nadobúda hodnoty z množiny S. Systém je teda v každom čase t opísaný pravdepodobnosťami nadobudnutia jednotlivých stavov $P(X_t = i)$, $i \in S$.

Aby sme takýto proces mohli nazvať markovovským reťazcom, musí spĺňať tzv. *zabúdaciu (markovovskú) vlastnosť*, ktorá hovorí, že stav systému v čase t závisí iba od predchádzajúceho stavu v čase $t-1$, bez ohľadu na to, ako sa do stavu v čase $t-1$ systém dostal. Formálne $P(X_t = i_t \mid X_{t-1} = i_{t-1}, \ldots, X_1 = i_1) = P(X_t = i_t \mid X_{t-1} = i_{t-1})$ pre každé $t = 1, 2, \ldots$ a $i_1, \ldots, i_t \in S$.

Pravdepodobnosť prechodu zo stavu i do stavu j značíme p_{ij} $(i, j \in S)$; budeme predpokladať, že v každom čase sú pravdepodobnosti prechodu rovnaké, teda že p_{ij} nezávisí na čase t, v ktorom sa nachádzame. Pre jednoduchosť budeme uvažovať množinu stavov $S = \{1, \ldots, n\}$. Potom všetky pravdepodobnosti prechodu p_{ij} $(1 \le i, j \le n)$ vieme zapísať *maticou prechodu* $\mathbf{P} = \{p_{ij}\}_{i,j=1,\ldots,n}$. Táto matica je zjavne typu $n \times n$ a všetky jej prvky sú nezáporné. Navyše, keďže z ľubovoľného stavu i v čase $t-1$ sa systém musí dostať do nejakého stavu j v čase t, tak $\sum_{j=1}^{n} p_{ij} = 1$, čiže matica prechodu je pravá stochastická.

Príklad 10.68. Uvažujme jednoduchý príklad systému bonusov pri povinnom poistení áut. Po vstupe do systému platí vlastník auta štandardné poistné, ak však celý rok nemá žiadnu nehodu, za nasledujúci rok sa dostane do bonusovej triedy a bude platiť iba znížené poistné. Ak už sa v bonusovej triede nachádza, ani beznehodový priebeh roku mu už žiadne ďalšie výhody neprinesie. Naopak, ak sa vlastník dostal do bonusovej triedy a v priebehu roka mal nehodu, ďalší rok bude musieť platiť základné poistné. Ak má poistenec nehodu a nachádza sa v základnej triede, ďalší rok v nej ostane.

Stav vlastníka auta v tomto systéme v jednotlivých rokoch môžeme opísať pomocou markovovského reťazca. Máme teda dva stavy, nech 1 opisuje základnú triedu a 2 bonusovú triedu. Predpokladajme, že vodič v základnej triede má v každom roku aspoň jednu nehodu s pravdepodobnosťou 0,2 (a teda beznehodový rok s pravdepodobnosťou 0,8). Takže s pravdepodobnosťou $p_{11} = 0{,}2$ ostane v základnej triede a s pravdepodobnosťou $p_{12} = 0{,}8$ sa dostane do bonusovej triedy. Predpokladajme tiež, že ak sa vodič nachádza v bonusovej triede, tak tým, že nechce začať platiť vyššie poistné je motivovaný jazdiť bezpečnejšie a nehodu má iba s pravdepodobnosťou 0,15. Potom $p_{21} = 0{,}15$, $p_{22} = 0{,}85$ a celá matica prechodu je

$$\mathbf{P} = \begin{pmatrix} 0{,}2 & 0{,}8 \\ 0{,}15 & 0{,}85 \end{pmatrix}.$$

Stavy tohto markovovského reťazca a pravdepodobnosti prechodu medzi nimi sú zachytené na obrázku 10.5a.

Pri vstupe do systému, teda v čase $t = 0$, vieme, kde sa poistenec nachádza: v základnej triede. S akými pravdepodobnosťami sa bude nachádzať v jednotlivých triedach po roku (v čase $t = 1$)? Vedeli by ste tieto pravdepodobnosti určiť aj pre druhý rok?

Pravdepodobnosti, s akými sa systém v čase t nachádza v jednotlivých stavoch, značíme $\pi_1^{(t)}, \ldots, \pi_n^{(t)}$. Spoločne ich zapisujeme v *riadkovom* vektore $\pi^{(t)} = (\pi_1^{(t)}, \ldots, \pi_n^{(t)})$. Keďže sú to pravdepodobnosti, musia spĺňať $\pi^{(t)} \geq 0$ (teda $\pi_i^{(t)} \geq 0$ pre každé i) a $\sum_{i=1}^n \pi_i^{(t)} = 1$, čiže $\pi^{(t)}\mathbf{1}_n = 1$.

Systém sa v čase t môže dostať do stavu j zo stavu 1 s pravdepodobnosťou p_{1j}, zo stavu 2 s pravdepodobnosťou p_{2j} atď., až nakoniec zo stavu n s pravdepodobnosťou p_{nj}. Ak teda chceme určiť celkovú pravdepodobnosť, že sa systém v čase t nachádza v stave j na základe predchádzajúceho času, stačí spočítať pravdepodobnosti, že sa nachádzal v čase $t-1$ v jednotlivých stavoch $1, \ldots, n$ a následne prešiel do stavu j. Čiže

$$\pi_j^{(t)} = \pi_1^{(t-1)} p_{1j} + \ldots + \pi_n^{(t-1)} p_{nj} = \sum_{i=1}^n \pi_i^{(t-1)} p_{ij}.$$

To môžeme zapísať vektorovo ako $\pi^{(t)} = \pi^{(t-1)}\mathbf{P}$. Ak poznáme počiatočné rozdelenie pravdepodobnosti $\pi^{(0)}$ skúmaného systému, tak rozdelenie pravdepodobnosti v čase 1 je $\pi^{(1)} = \pi^{(0)}\mathbf{P}$, v čase 2 je to $\pi^{(2)} = \pi^{(1)}\mathbf{P} = \pi^{(0)}\mathbf{P}^2$, a podobne v čase $t > 0$ dostávame

$$\pi^{(t)} = \pi^{(0)}\mathbf{P}^t.$$

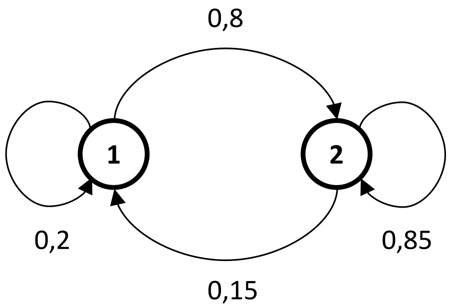
Vývoj markovovského reťazca je teda plne určený počiatočným stavom a týmto jednoduchým maticovým vzťahom.

V príklade 10.68 je $\pi^{(0)} = (1, 0)$ a $\pi^{(t)} = (1, 0)\mathbf{P}^t$. Vypočítajte $\pi^{(1)}$ a $\pi^{(2)}$. Vývoj pravdepodobnosti, že vlastník auta je v základnej triede (resp. v bonusovej triede), je zachytený na obrázku 10.5b. Všimnite si, že rozdelenie pravdepodobnosti sa pomerne rýchlo ustáli na hodnotách $\pi_1^{(t)} \approx 0{,}16$ a $\pi_2^{(t)} \approx 0{,}84$.

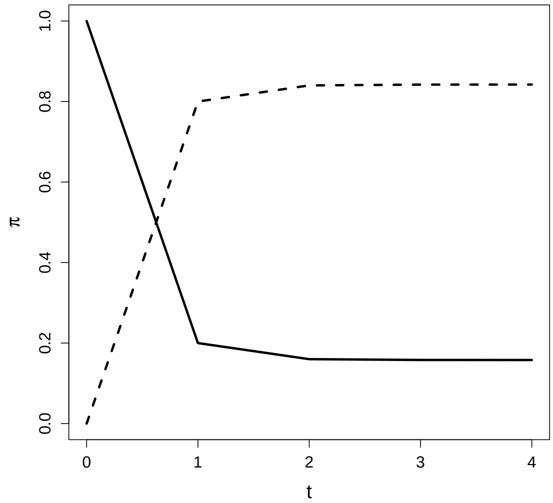
Ak pravdepodobnostné rozdelenie π spĺňa $\pi = \pi\mathbf{P}$, teda ak sa toto rozdelenie v markovovskom reťazci nemení v čase, tak hovoríme, že je stacionárne. Všimnite si, že takéto rozdelenie pravdepodobnosti je ľavým vlastným vektorom matice $\mathbf{P}$ prislúchajúcim vlastnému číslu $\lambda = 1$.

Veta 10.69 (Stacionárne rozdelenie markovovského reťazca)**.** Nech $\mathbf{P}$ je matica prechodu markovovského reťazca, ktorej všetky prvky sú kladné. Potom existuje jediné stacionárne rozdelenie π^* tohto markovovského procesu. Rozdelenie π^* vieme získať ako vlastný vektor zodpovedajúci najväčšiemu vlastnému číslu matice $\mathbf{P}^T$. Toto rozdelenie navyše spĺňa $\lim_{t\to\infty} \pi^{(t)} = \pi^*$ pre ľubovoľné $\pi^{(0)}$.

Dôkaz. Pre maticu $\mathbf{P}$ platí veta 10.66. Podľa tejto vety má $\mathbf{P}$ jediný kladný vlastný vektor (až na prenásobenie konštantou) a prislúcha vlastnému číslu λ^*, ktoré je najväčšie v absolútnej hodnote. Keďže $\mathbf{P}$ je pravá stochastická, tak $\mathbf{1}_n$, ktorý má všetky prvky kladné, je spomínaný vlastný vektor, prislúchajúci vlastnému číslu $\lambda^* = 1$. Potom podľa vety 10.66 existuje jediný (až na prenásobenie konštantou) vlastný vektor π^* matice $\mathbf{P}^T$ (teda ľavý vlastný vektor matice $\mathbf{P}$) spĺňajúci $\pi^* > 0$, ktorý navyše prislúcha vlastnému číslu $\lambda^* = 1$. Vektor π^* uvažujeme ako riadkový vektor, teda $\pi^* = \pi^*\mathbf{P}$. Vhodným prenásobením konštantou dosiahneme $\pi^*\mathbf{1}_n = 1$, čiže π^* je hľadané stacionárne rozdelenie pravdepodobnosti.



(a) Krúžky 1, 2 predstavujú stavy, šípky prechody medzi stavmi a čísla pri šípkach pravdepodobnosti prechodu.



(b) Vývoj $\pi_1^{(t)}$ (plná čiara) a $\pi_2^{(t)}$ (prerušovaná čiara) v čase.

Obr. 10.5: Markovovský reťazec z príkladu 10.68.

Nech $\pi^{(0)}$ je počiatočné rozdelenie pravdepodobností. Z vety 10.66 vyplýva $\lim_{t\to\infty} \mathbf{P}^t = \mathbf{1}_n\pi^*$, a teda

$$\lim_{t\to\infty} \pi^{(t)} = \pi^{(0)} \lim_{t\to\infty} \mathbf{P}^t = \pi^{(0)}\mathbf{1}_n\pi^* = \pi^*,$$

keďže $\pi^{(0)}\mathbf{1}_n = \sum_{i=1}^n \pi_i^{(0)} = 1$. □

Poznamenajme, že podobné tvrdenie platí aj pre markovovské reťazce, ktorých všetky pravdepodobnosti prechodu nemusia byť kladné, stačí ak markovovský reťazec je tzv. nerozložiteľný.

Vypočítajte $\lim_{t\to\infty} \pi^{(t)}$ pre príklad 10.68 pomocou vety 10.69 a overte, že je v zhode s obrázkom 10.5b.

10.6 Úlohy na precvičenie

Úloha 10.1. Nájdite vlastné čísla a vlastné vektory matice $\mathbf{A}$ z úlohy 9.1 a na základe nájdeného spektra sa presvedčte, že $\mathbf{A}$ je pozitívne definitná.

Úloha 10.2. Nájdite spektrum matíc

$$\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \quad \text{a} \quad \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Úloha 10.3. Ukážte, že každý projektor $\mathbf{P} \in \mathbb{R}^{n\times n}$ je pozitívne semidefinitná matica. Ukážte, že spektrum projektora neobsahuje iné čísla ako 0 a 1. Charakterizujte vlastný priestor projektora $\mathbf{P}$ zodpovedajúci vlastnému číslu 0 (ak $\mathbf{P} \neq \mathbf{I}_n$) a vlastný priestor projektora $\mathbf{P}$ zodpovedajúci vlastnému číslu 1 (ak $\mathbf{P} \neq \mathbf{0}_{n\times n}$).

Úloha 10.4. Nech $\mathbf{A}$ je symetrická matica typu $n \times n$, nech $\lambda_1, \ldots, \lambda_n$ sú jej vlastné čísla (so zopakovaním podľa násobnosti) a nech $\mathbf{u}_1, \ldots, \mathbf{u}_n$ je ortonormálny systém prislúchajúcich vlastných vektorov. Ukáže, že potom matica $\mathbf{A}^+ = \sum_{i;\lambda_i \neq 0} \lambda_i^{-1} \mathbf{u}_i \mathbf{u}_i^T$ je zovšeobecnená inverzia matice $\mathbf{A}$. Presvedčte sa, že $\mathbf{A}^+$ spĺňa vlastnosti Mooreovej-Penroseovej pseudoinverzie z definície 6.31.

Úloha 10.5. Nech $a, b \in \mathbb{R}$ a nech $\mathbf{A} = a\mathbf{I}_n + b\mathbf{1}_n\mathbf{1}_n^T$ je úplne symetrická matica. Ukážte, že potom je vlastným vektorom matice $\mathbf{A}$ vektor $\mathbf{1}_n$ a tiež akýkoľvek vektor kolmý na $\mathbf{1}_n$. Pre maticu $\mathbf{A}$ nájdite vlastné čísla a ich násobnosti.

Úloha 10.6. Na základe úlohy 10.5 určte, pre aké a a b je $\mathbf{A} = a\mathbf{I}_n + b\mathbf{J}_{n\times n}$ pozitívne semidefinitná. Pre každú takú dvojicu a, b nájdite maticu $\mathbf{A}^{1/2}$.

Úloha 10.7. Nech $\mathbf{A} \in \mathbb{R}^{n\times n}$ (nie nutne symetrická) má n vlastných hodnôt $\lambda_1, \ldots, \lambda_n$ (vrátane násobnosti). Vyjadrite pomocou λ_i: (i) $\det(\mathbf{A})$, (ii) $\mathrm{tr}(\mathbf{A})$, (iii) $\|\mathbf{A}\|^2$.

Úloha 10.8. Nech

$$\mathbf{A} = \begin{pmatrix} 25 & 10 \\ 10 & 40 \end{pmatrix}.$$

Načrtnite úrovňové množiny (vrstevnice) kvadratickej formy $\mathbf{x}^T\mathbf{A}\mathbf{x}$. Viete si predstaviť, ako vyzerá kvadratická forma $\mathbf{x}^T\mathbf{B}\mathbf{x}$, kde

$$\mathbf{B} = \begin{pmatrix} 25 & 10 & 0 \\ 10 & 40 & 0 \\ 0 & 0 & 1 \end{pmatrix}?$$

Úloha 10.9. Nech

$$\mathbf{A} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 2 \\ 0 & 2 & 2 \end{pmatrix}.$$

Nájdite $\mathbf{A}^{1/2}$.

Úloha 10.10. Nech $\mathbf{A}, \mathbf{B} \in \mathcal{S}_+^n$ spĺňajú $\mathbf{A} \succeq \mathbf{B}$, kde $\succeq$ značí Loewnerovo usporiadanie. Ukážte, že potom $\lambda_{\max}(\mathbf{A}) \geq \lambda_{\max}(\mathbf{B})$, kde $\lambda_{\max}$ značí najväčšie vlastné číslo. Návod: použite vetu 10.33.

Kapitola 11

Singulárny rozklad a maticové normy

11.1 Singulárny rozklad

Definícia 11.1 (Diagonálna obdĺžniková matica)**.** Pod *diagonálnou obdĺžnikovou maticou* $\mathbf{A} \in \mathbb{R}^{m\times n}$ rozumieme maticu, ktorá spĺňa $(\mathbf{A})_{ij} = 0$ pre všetky $i \neq j$.

Poznámka 11.2. V ďalšom texte budeme aj diagonálnu obdĺžnikovú maticu nazývať jednoducho diagonálna matica, keď bude z kontextu zrejmé, že matica je obdĺžniková.

Veta 11.3 (Singulárny rozklad matice)**.** Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$ má hodnosť $r > 0$. Potom $\mathbf{A}$ sa dá zapísať ako

$$\mathbf{A} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T,$$

kde $\mathbf{U} \in \mathbb{R}^{m\times m}$ a $\mathbf{V} \in \mathbb{R}^{n\times n}$ sú ortogonálne matice a diagonálna (obdĺžniková) matica $\boldsymbol{\Sigma} \in \mathbb{R}^{m\times n}$ spĺňa

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_1 & \mathbf{0}_{r\times(n-r)} \\ \mathbf{0}_{(m-r)\times r} & \mathbf{0}_{(m-r)\times(n-r)} \end{pmatrix},$$

kde $\boldsymbol{\Sigma}_1 \in \mathbb{R}^{r\times r}$ je diagonálna matica s kladnou diagonálou.

Dôkaz. Keďže matica $\mathbf{A}^T\mathbf{A} \in \mathbb{R}^{n\times n}$ je kladne semidefinitná, existuje jej spektrálny rozklad $\mathbf{A}^T\mathbf{A} = \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^T$ (veta 10.27), pričom jej vlastné čísla sú nezáporné. Označme

$$\boldsymbol{\Lambda} = \begin{pmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix},$$

kde $\mathbf{D}$ je diagonálnou maticou kladných vlastných čísel matice $\mathbf{A}^T\mathbf{A}$. Matica $\mathbf{D}$ je typu $r \times r$, lebo $\operatorname{rank}(\mathbf{A}^T\mathbf{A}) = \operatorname{rank}(\mathbf{A}) = r$, a teda $\mathbf{A}^T\mathbf{A}$ má r kladných vlastných čísel. Podobne zapíšme aj maticu $\mathbf{V}$ ako $\mathbf{V} = (\mathbf{V}_1, \mathbf{V}_2)$, kde $\mathbf{V}_1 \in \mathbb{R}^{n\times r}$. Potom

$$\begin{pmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} = \boldsymbol{\Lambda} = \mathbf{V}^T\mathbf{A}^T\mathbf{A}\mathbf{V} = \begin{pmatrix} \mathbf{V}_1^T\mathbf{A}^T\mathbf{A}\mathbf{V}_1 & \mathbf{V}_1^T\mathbf{A}^T\mathbf{A}\mathbf{V}_2 \\ \mathbf{V}_2^T\mathbf{A}^T\mathbf{A}\mathbf{V}_1 & \mathbf{V}_2^T\mathbf{A}^T\mathbf{A}\mathbf{V}_2 \end{pmatrix}.$$

Získavame tak $\mathbf{D} = \mathbf{V}_1^T\mathbf{A}^T\mathbf{A}\mathbf{V}_1$ a $\mathbf{0} = \mathbf{V}_2^T\mathbf{A}^T\mathbf{A}\mathbf{V}_2 = (\mathbf{A}\mathbf{V}_2)^T\mathbf{A}\mathbf{V}_2$. Z druhej rovnosti vyplýva $\|\mathbf{A}\mathbf{V}_2\|^2 = \operatorname{tr}((\mathbf{A}\mathbf{V}_2)^T\mathbf{A}\mathbf{V}_2) = 0$, čiže $\mathbf{A}\mathbf{V}_2 = \mathbf{0}$.

Označme $\boldsymbol{\Sigma}_1 = \mathbf{D}^{1/2}$ a $\mathbf{U}_1 = \mathbf{A}\mathbf{V}_1\boldsymbol{\Sigma}_1^{-1} \in \mathbb{R}^{m\times r}$. Všimnime si, že

$$\mathbf{U}_1^T\mathbf{U}_1 = \boldsymbol{\Sigma}_1^{-1}\mathbf{V}_1^T\mathbf{A}^T\mathbf{A}\mathbf{V}_1\boldsymbol{\Sigma}_1^{-1} = \boldsymbol{\Sigma}_1^{-1}\mathbf{D}\boldsymbol{\Sigma}_1^{-1} = \boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_1^2\boldsymbol{\Sigma}_1^{-1} = \mathbf{I}_r.$$

Preto $\text{rank}(\mathbf{U}_1) = \text{rank}(\mathbf{U}_1^T\mathbf{U}_1) = r$, a teda $\dim(\mathcal{N}(\mathbf{U}_1^T)) = m - r$. To znamená, že existuje ortonormálna báza $\mathcal{N}(\mathbf{U}_1^T)$ veľkosti $m-r$. Uložme túto bázu ako stĺpce matice $\mathbf{U}_2 \in \mathbb{R}^{m\times(m-r)}$; potom $\mathbf{U}_1^T\mathbf{U}_2 = \mathbf{0}$. Vytvorme $\mathbf{U} = (\mathbf{U}_1, \mathbf{U}_2)$ a overme

$$\mathbf{U}^T\mathbf{U} = \begin{pmatrix} \mathbf{U}_1^T\mathbf{U}_1 & \mathbf{U}_1^T\mathbf{U}_2 \\ \mathbf{U}_2^T\mathbf{U}_1 & \mathbf{U}_2^T\mathbf{U}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{m-r} \end{pmatrix} = \mathbf{I}_m.$$

Čiže $\mathbf{U}$ je ortogonálna. Počítajme:

$$\mathbf{U}^T\mathbf{A}\mathbf{V} = \begin{pmatrix} \mathbf{U}_1^T\mathbf{A}\mathbf{V}_1 & \mathbf{U}_1^T\mathbf{A}\mathbf{V}_2 \\ \mathbf{U}_2^T\mathbf{A}\mathbf{V}_1 & \mathbf{U}_2^T\mathbf{A}\mathbf{V}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{\Sigma}_1^{-1}\mathbf{V}_1^T\mathbf{A}^T\mathbf{A}\mathbf{V}_1 & \mathbf{0} \\ \mathbf{U}_2^T\mathbf{U}_1\mathbf{\Sigma}_1 & \mathbf{0} \end{pmatrix},$$

kde sme okrem $\mathbf{U}_1 = \mathbf{A}\mathbf{V}_1\mathbf{\Sigma}_1^{-1}$ a $\mathbf{A}\mathbf{V}_2 = \mathbf{0}$ využili ešte $\mathbf{U}_1\mathbf{\Sigma}_1 = \mathbf{A}\mathbf{V}_1\mathbf{\Sigma}_1^{-1}\mathbf{\Sigma}_1 = \mathbf{A}\mathbf{V}_1$. Maticu ďalej môžeme pomocou $\mathbf{D} = \mathbf{V}_1^T\mathbf{A}^T\mathbf{A}\mathbf{V}_1$ upraviť na

$$\mathbf{U}^T\mathbf{A}\mathbf{V} = \begin{pmatrix} \mathbf{\Sigma}_1^{-1}\mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} = \begin{pmatrix} \mathbf{\Sigma}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix},$$

keďže $\mathbf{\Sigma}_1^{-1}\mathbf{D} = \mathbf{\Sigma}_1^{-1}\mathbf{\Sigma}_1^2 = \mathbf{\Sigma}_1$. Získavame tak žiadaný rozklad, akurát stačí označiť

$$\mathbf{\Sigma} = \begin{pmatrix} \mathbf{\Sigma}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}.$$

□

Definícia 11.4 (Singulárny rozklad)**.** Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$ má hodnosť $r > 0$. Potom rozklad $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$ z vety 11.3 nazývame *singulárny rozklad* matice $\mathbf{A}$.

Dôsledok 11.5 (Kompaktný singulárny rozklad)**.** Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$ má hodnosť $r > 0$. Potom $\mathbf{A}$ sa dá zapísať ako

$$\mathbf{A} = \mathbf{U}_1\mathbf{\Sigma}_1\mathbf{V}_1^T,$$

kde $\mathbf{U}_1 \in \mathbb{R}^{m\times r}$ a $\mathbf{V}_1 \in \mathbb{R}^{n\times r}$ majú ortonormálne stĺpce a $\mathbf{\Sigma}_1 \in \mathbb{R}^{r\times r}$ je diagonálna matica s kladnou diagonálou. Tento rozklad nazývame *kompaktný (redukovaný, úzky) singulárny rozklad* matice $\mathbf{A}$.

Poznámka 11.6. V angličtine je singulárny rozklad známy pod skratkou *SVD* (*Singular Value Decomposition*).

Poznámka 11.7. Upozornime, že poradie vlastných čísel v matici $\mathbf{\Lambda}$ môžeme v spektrálnom rozklade $\mathbf{U}\mathbf{\Lambda}\mathbf{U}^T$ ľubovoľne meniť, ak príslušným spôsobom prehodíme aj stĺpce matice $\mathbf{U}$. Podobne môžeme meniť poradie diagonálnych prvkov v matici $\mathbf{\Sigma}_1$ v singulárnom rozklade $\mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$, ak rovnako prehodíme aj stĺpce matíc $\mathbf{U}$ a $\mathbf{V}^T$. Zvyčajne bývajú diagonálne prvky matice $\mathbf{\Sigma}_1$ usporiadané od najväčšieho po najmenší, ako uvidíme v dôsledku 11.12.

Poznámka 11.8. V dôkaze vety 11.3 sme vychádzali zo spektrálneho rozkladu matice $\mathbf{A}^T\mathbf{A}$. Pripomeňme, že podľa lemy 10.19 majú matice $\mathbf{A}^T\mathbf{A} \in \mathcal{S}_+^n$ a $\mathbf{A}\mathbf{A}^T \in \mathcal{S}_+^m$ rovnaké kladné vlastné čísla. Teda ekvivalentne sme mohli vychádzať zo spektrálneho rozkladu matice $\mathbf{A}\mathbf{A}^T$.

Veta 11.9 (Vlastnosti singulárneho rozkladu)**.** Nech matica $\mathbf{A} \in \mathbb{R}^{m\times n}$ sa dá zapísať ako $\mathbf{A} = \mathbf{U\Sigma V}^T$, kde $\mathbf{U} \in \mathbb{R}^{m\times m}$, $\mathbf{V} \in \mathbb{R}^{n\times n}$ sú ortogonálne matice,

$$\mathbf{\Sigma} = \begin{pmatrix} \mathbf{\Sigma}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$$

a $\mathbf{\Sigma}_1 \in \mathbb{R}^{r\times r}$ je diagonálna matica, ktorej všetky diagonálne prvky sú nenulové. Potom:

(i) $\operatorname{rank}(\mathbf{A}) = r$,

(ii)
$$\mathbf{A}^T\mathbf{A} = \mathbf{V}\begin{pmatrix} \mathbf{\Sigma}_1^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}\mathbf{V}^T \quad \text{a} \quad \mathbf{A}\mathbf{A}^T = \mathbf{U}\begin{pmatrix} \mathbf{\Sigma}_1^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}\mathbf{U}^T,$$

(iii) $\mathbf{U}_1 = \mathbf{AV}_1\mathbf{\Sigma}_1^{-1}$ a $\mathbf{V}_1 = \mathbf{A}^T\mathbf{U}_1\mathbf{\Sigma}_1^{-1}$, kde $\mathbf{U} = (\mathbf{U}_1, \mathbf{U}_2)$, $\mathbf{V} = (\mathbf{V}_1, \mathbf{V}_2)$, $\mathbf{U}_1 \in \mathbb{R}^{m\times r}$ a $\mathbf{V}_1 \in \mathbb{R}^{n\times r}$.

Dôkaz. V časti (i) dostávame $\operatorname{rank}(\mathbf{A}) = \operatorname{rank}(\mathbf{U\Sigma V}^T) = \operatorname{rank}(\mathbf{\Sigma}) = r$, keďže násobenie regulárnou maticou nemení hodnosť (veta 4.18). Časť (ii) sa ľahko dokáže priamym výpočtom.

Keďže $\mathbf{V}$ je ortogonálna, tak $\mathbf{V}_1^T\mathbf{V}_1 = \mathbf{I}$. Potom

$$\mathbf{AV}_1\mathbf{\Sigma}_1^{-1} = \mathbf{U}_1\mathbf{\Sigma}_1\mathbf{V}_1^T\mathbf{V}_1\mathbf{\Sigma}_1^{-1} = \mathbf{U}_1\mathbf{\Sigma}_1\mathbf{\Sigma}_1^{-1} = \mathbf{U}_1,$$

čím je dokázaná prvá rovnosť časti (iii). Dôkaz druhej rovnosti je analogický. □

Poznámka 11.10. Veta 11.9 opisuje previazanosť singulárneho a spektrálneho rozkladu. Označme $\mathbf{\Sigma}_1 = \operatorname{diag}(\sigma_1, \dots, \sigma_r)$, pričom budeme dodržiavať konvenciu, že σ_i sú usporiadané zostupne, teda $\sigma_1 \geq \ldots \geq \sigma_r$. Potom časť (ii) hovorí, že vlastné čísla matice $\mathbf{A}^T\mathbf{A}$ sú $\sigma_1^2, \dots, \sigma_r^2, 0, \dots, 0$ s príslušnými vlastnými vektormi uloženými v stĺpcoch matice $\mathbf{V}$. Podobne vlastné čísla matice $\mathbf{AA}^T$ sú $\sigma_1^2, \dots, \sigma_r^2, 0, \dots, 0$ s príslušnými vlastnými vektormi uloženými v stĺpcoch matice $\mathbf{U}$.

Definícia 11.11 (Singulárne čísla, singulárne vektory)**.** Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$ má hodnosť r. Potom čísla $\sigma_1 \geq \ldots \geq \sigma_r$ z poznámky 11.10 nazývame *singulárne čísla* (*singulárne hodnoty*) matice $\mathbf{A}$. Teda singulárne čísla sú odmocniny z kladných vlastných čísel matice $\mathbf{AA}^T$ (resp. $\mathbf{A}^T\mathbf{A}$). Stĺpce matíc $\mathbf{U}$ a $\mathbf{V}$ nazývame *singulárne vektory*, konkrétne stĺpce matice $\mathbf{U}$ sú *ľavé singulárne vektory* a stĺpce matice $\mathbf{V}$ sú *pravé singulárne vektory*.

Dôsledok 11.12 (Singulárny rozklad pomocou spektrálneho rozkladu)**.** Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$ má hodnosť r. Nech $\sigma_1^2 \geq \ldots \geq \sigma_r^2 > 0 = \ldots = 0$ sú vlastné čísla matice $\mathbf{A}^T\mathbf{A}$. Potom $\mathbf{A}$ sa dá zapísať ako

$$\mathbf{A} = \mathbf{U\Sigma V}^T,$$

kde

(i)
$$\mathbf{\Sigma} = \begin{pmatrix} \mathbf{\Sigma}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix},$$

pričom $\mathbf{\Sigma}_1 = \operatorname{diag}(\sigma_1, \dots, \sigma_r)$,

(ii) $\mathbf{U} \in \mathbb{R}^{m\times m}$ je ortogonálna matica vlastných vektorov matice $\mathbf{AA}^T$, ktoré v poradí zodpovedajú vlastným číslam $\sigma_1^2, \ldots, \sigma_r^2, 0, \ldots, 0$ tejto matice. $\mathbf{V} \in \mathbb{R}^{n\times n}$ je ortogonálna matica vlastných vektorov matice $\mathbf{A}^T\mathbf{A}$, ktoré v poradí zodpovedajú vlastným číslam $\sigma_1^2, \ldots, \sigma_r^2, 0, \ldots, 0$ tejto matice,

(iii) $\mathbf{U}_1 = \mathbf{AV}_1\mathbf{\Sigma}_1^{-1}$ a $\mathbf{V}_1 = \mathbf{A}^T\mathbf{U}_1\mathbf{\Sigma}_1^{-1}$, kde $\mathbf{U} = (\mathbf{U}_1, \mathbf{U}_2)$ a $\mathbf{V} = (\mathbf{V}_1, \mathbf{V}_2)$, $\mathbf{U}_1 \in \mathbb{R}^{m\times r}$ a $\mathbf{V}_1 \in \mathbb{R}^{n\times r}$.

Poznámka 11.13. Časť (ii) v dôsledku 11.12 sa dá maticovo zapísať nasledovne:

$$\mathbf{AA}^T = \mathbf{U}\begin{pmatrix}\mathbf{\Sigma}_1^2 & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}\mathbf{U}^T \quad \text{a} \quad \mathbf{A}^T\mathbf{A} = \mathbf{V}\begin{pmatrix}\mathbf{\Sigma}_1^2 & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}\mathbf{V}^T$$

sú spektrálne rozklady matíc $\mathbf{AA}^T$ a $\mathbf{A}^T\mathbf{A}$.

Poznámka 11.14. Veta 11.3 a dôsledok 11.12 sú formálne sformulované iba pre nehraničné prípady, čiže pre prípady keď $m \neq n$ a r je menšie od oboch týchto čísel. Ak pre niektorú dvojicu parametrov nastáva rovnosť, matica $\mathbf{\Sigma}$ sa náležite zjednoduší (zmiznú nulové riadky alebo stĺpce), pre $r = m$ zmiznú nulové vlastné čísla matice $\mathbf{AA}^T$ a pre $r = n$ zmiznú nulové vlastné čísla matice $\mathbf{A}^T\mathbf{A}$.

Singulárny rozklad matice $\mathbf{A}$ môžeme manuálne získať pomocou spektrálnych rozkladov matíc $\mathbf{A}^T\mathbf{A}$ a $\mathbf{AA}^T$ na základe dôsledku 11.12.

Príklad 11.15. Nájdime singulárny rozklad matice

$$\mathbf{A} = \begin{pmatrix}1 & 1 & 1\\ 1 & -1 & -1\end{pmatrix}.$$

Budeme vychádzať z dôsledku 11.12, preto vypočítajme najprv spektrálny rozklad matice

$$\mathbf{AA}^T = \begin{pmatrix}3 & -1\\ -1 & 3\end{pmatrix}.$$

Overte, že vlastné čísla $\mathbf{AA}^T$ sú $\lambda_1 = 4$, $\lambda_2 = 2$, a že vlastné vektory sú $\mathbf{u}_1 = (1, -1)^T/\sqrt{2}$, $\mathbf{u}_2 = (1, 1)^T/\sqrt{2}$. Vidíme, že $\text{rank}(\mathbf{AA}^T) = 2$, čiže $r = 2$, a teda $\mathbf{\Sigma}_1 \in \mathbb{R}^{2\times 2}$. Keďže

$$\mathbf{AA}^T = \mathbf{U}\begin{pmatrix}\mathbf{\Sigma}_1^2 & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}\mathbf{U}^T,$$

dostávame $\mathbf{U} = \mathbf{U}_1 = (\mathbf{u}_1, \mathbf{u}_2)$ a $\mathbf{\Sigma}_1 = \text{diag}(2, \sqrt{2})$.

Na výpočet príslušného spektrálneho rozkladu

$$\mathbf{A}^T\mathbf{A} = \begin{pmatrix}2 & 0 & 0\\ 0 & 2 & 2\\ 0 & 2 & 2\end{pmatrix} = \mathbf{V}\begin{pmatrix}\mathbf{\Sigma}_1^2 & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}\mathbf{V}^T$$

využijeme najprv vzťah

$$\mathbf{V}_1 = \mathbf{A}^T\mathbf{U}_1\mathbf{\Sigma}_1^{-1} = \begin{pmatrix}0 & 1\\ 1/\sqrt{2} & 0\\ 1/\sqrt{2} & 0\end{pmatrix}.$$

Keďže $\mathbf{V} \in \mathbb{R}^{3\times 3}$, ostáva určiť posledný vlastný vektor, ktorý prislúcha vlastnému číslu $\lambda = 0$. Tým je $\mathbf{v}_3 = (0, 1/\sqrt{2}, -1/\sqrt{2})^T$. Skontrolujte, že $\mathbf{v}_3$ ozaj prislúcha vlastnému číslu $\lambda = 0$. Získavame teda singulárny rozklad

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & \sqrt{2} & 0 \end{pmatrix} \begin{pmatrix} 0 & 1/\sqrt{2} & 1/\sqrt{2} \\ 1 & 0 & 0 \\ 0 & 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix}.$$

Lema 11.16 (Vlastnosti singulárnych čísel)**.** Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$ má singulárne čísla $\sigma_1, \dots, \sigma_r > 0$, kde $r = \operatorname{rank}(\mathbf{A})$. Potom

(i) $c\mathbf{A}$ má singulárne čísla $|c|\sigma_1, \dots, |c|\sigma_r$ pre ľubovoľné $c \neq 0$.

(ii) ak $\mathbf{A} \in \mathbb{R}^{n\times n}$ je regulárna, potom $\mathbf{A}^{-1}$ má singulárne čísla $\sigma_1^{-1}, \dots, \sigma_n^{-1}$.

Dôkaz. Pre $c\mathbf{A}$ platí $(c\mathbf{A})^T(c\mathbf{A}) = c^2(\mathbf{A}^T\mathbf{A})$, a teda jej singulárne čísla sú odmocniny z $c^2\lambda_1, \dots, c^2\lambda_r$, kde λ_i sú vlastné čísla matice $\mathbf{A}^T\mathbf{A}$. Takže jej singulárne čísla sú $|c|\sigma_1, \dots, |c|\sigma_r$.

Na dôkaz (ii) si podobne stačí uvedomiť, že $(\mathbf{A}^{-1})^T\mathbf{A}^{-1} = (\mathbf{A}\mathbf{A}^T)^{-1}$, pričom táto matica má vlastné čísla $\lambda_1^{-1}, \dots, \lambda_n^{-1}$, kde λ_i sú vlastné čísla matice $\mathbf{A}\mathbf{A}^T$ (resp. $\mathbf{A}^T\mathbf{A}$). Teda singulárne čísla matice $\mathbf{A}^{-1}$ (odmocniny z λ_i^{-1}) sú inverzné k singulárnym číslam matice $\mathbf{A}$ (odmocninám z λ_i). □

11.1.1 Použitie singulárneho rozkladu

Lema 11.17 (Singulárny rozklad pseudoinverzie)**.** Nech $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$ je singulárny rozklad matice $\mathbf{A}$ a $\mathbf{A} = \mathbf{U}_1\mathbf{\Sigma}_1\mathbf{V}_1^T$ je jej kompaktný singulárny rozklad. Potom

(i) singulárny rozklad matice $\mathbf{A}^+$ je $\mathbf{A}^+ = \mathbf{V}\mathbf{\Sigma}^+\mathbf{U}^T$, kde pseudoinverzia $\mathbf{\Sigma}^+$ má tvar

$$\mathbf{\Sigma}^+ = \begin{pmatrix} \mathbf{\Sigma}_1^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}. \tag{11.1}$$

(ii) kompaktný singulárny rozklad matice $\mathbf{A}^+$ je $\mathbf{A}^+ = \mathbf{V}_1\mathbf{\Sigma}_1^{-1}\mathbf{U}_1^T$.

Dôkaz. (i): Lema 6.35 hovorí, že $\mathbf{A}^+ = \mathbf{V}\mathbf{\Sigma}^+\mathbf{U}^T$, kde $\mathbf{\Sigma}^+$ má podľa úlohy 6.13 tvar (11.1). Tento rozklad zjavne spĺňa vlastnosti singulárneho rozkladu. Časť (ii) je dôsledkom časti (i). □

Poznámka 11.18. Lema 11.17 poskytuje metódu na konštrukciu pseudoinverznej matice: najprv vypočítame singulárny rozklad matice $\mathbf{A}$, a potom z neho skonštruujeme maticu $\mathbf{A}^+$.

Príklad 11.19. Nájdime pseudoinverziu matice

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & -1 \end{pmatrix}.$$

Podľa príkladu 11.15 vieme, že kompaktný singulárny rozklad matice $\mathbf{A}$ je

$$\mathbf{A} = \mathbf{U}_1\mathbf{\Sigma}_1\mathbf{U}_1^T = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & \sqrt{2} \end{pmatrix} \begin{pmatrix} 0 & 1/\sqrt{2} & 1/\sqrt{2} \\ 1 & 0 & 0 \end{pmatrix}.$$

Takže matica $\mathbf{A}^+$ je podľa lemy 11.17

$$\mathbf{A}^+ = \begin{pmatrix} 0 & 1 \\ 1/\sqrt{2} & 0 \\ 1/\sqrt{2} & 0 \end{pmatrix} \begin{pmatrix} 1/2 & 0 \\ 0 & 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} = \begin{pmatrix} 1/2 & 1/2 \\ 1/4 & -1/4 \\ 1/4 & -1/4 \end{pmatrix}.$$

Overte, že $\mathbf{A}^+$ spĺňa definičné vlastnosti pseudoinverznej matice.

Veta 11.20 (Eckartova-Youngova veta)**.** Nech $\mathbf{A} = \mathbf{U\Sigma V}^T$ je singulárny rozklad matice $\mathbf{A} \in \mathbb{R}^{m\times n}$ s hodnosťou $r > 0$ a nech $\sigma_1 \geq \ldots \geq \sigma_r$ sú jej singulárne čísla. Označme

$$\mathbf{A}^* = \mathbf{U} \begin{pmatrix} \mathbf{\Sigma}_1^* & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{V}^T,$$

kde $\mathbf{\Sigma}_1^* = \mathrm{diag}(\sigma_1, \ldots, \sigma_k)$ pre zvolené $k < r$. Potom

$$\|\mathbf{A}^* - \mathbf{A}\| \leq \|\mathbf{B} - \mathbf{A}\|$$

pre ľubovoľnú maticu $\mathbf{B} \in \mathbb{R}^{m\times n}$ hodnosti nanajvýš k.

Dôkaz. Vetu uvádzame bez dôkazu. Dôkaz čitateľ môže nájsť v [4] (veta 21.12.4). □

Maticu $\mathbf{A}$ teda podľa vety 11.20 spomedzi všetkých matíc s danou hodnosťou najlepšie aproximuje matica $\mathbf{A}^*$ získaná singulárnym rozkladom (z pohľadu Frobeniovej normy).

Veta 11.20 tak ukazuje, ako sa dá singulárny rozklad použiť na zníženie dimenzie dát. Ak v matici $\mathbf{A}$ máme uložené dáta (teda ak $\mathbf{A} = \mathbf{X}$ vo zvyčajnom značení), tak dostávame $\mathbf{A}^* = \mathbf{U}_*\mathbf{\Sigma}_1^*\mathbf{V}_*^T$, kde $\mathbf{U}_*$ pozostáva z prvých k stĺpcov $\mathbf{U}$ a $\mathbf{V}_*$ z prvých k stĺpcov $\mathbf{V}$. Pôvodnú $m \times n$ maticu $\mathbf{A}$ teda aproximujeme pomocou dvoch matíc $\mathbf{U}_*$ a $\mathbf{V}_*$ veľkostí $m \times k$ a $n \times k$ a pomocou k čísel $\sigma_1, \ldots, \sigma_k$. Obzvlášť ak m aj n sú vysoké čísla, môžeme tak dospieť k výraznej redukcii veľkosti dát.

Napríklad ak $m = 500\,000$ a $n = 10\,000$, tak $\mathbf{A}$ má 5 miliárd prvkov. Ak maticu aproximujeme hodnosťou 40, tak dostávame matice $\mathbf{U}_*$ a $\mathbf{V}_*$ veľkostí $40 \times 500\,000$ a $40 \times 10\,000$, čo je spolu so 40 singulárnymi číslami 20 400 040 údajov, teda približne 250-krát menej.

Príkladom dát s vysokým počtom m aj n sú užívateľské hodnotenia (napríklad filmov, či hudby). V takých prípadoch máme v matici $\mathbf{A}$ uložené hodnotenia m užívateľov o n filmoch a je bežné stretnúť sa s tisíckami filmov a stovkami tisícok užívateľov. Internetové distribučné spoločnosti potom riešia problém, ako na základe hodnotení niektorých filmov odporúčať používateľom ďalšie filmy. Jednou z techník, ktoré sa pri tvorbe systémov automatizovaného odporúčania používajú, je práve singulárny rozklad.

11.2 Vektorové a maticové normy

11.2.1 Vektorové normy

Doteraz sme pracovali s klasickými normami $\|\mathbf{x}\| = \sqrt{\mathbf{x}^T\mathbf{x}}$ pre $\mathbf{x} \in \mathbb{R}^n$ a $\|\mathbf{A}\| = (\mathrm{tr}(\mathbf{A}^T\mathbf{A}))^{1/2}$ pre $\mathbf{A} \in \mathbb{R}^{m\times n}$ (pozri kapitolu 5). Poznáme však aj mnohé iné normy.

Definícia 11.21 (ℓ^p-norma)**.** Nech $\mathbf{x} \in \mathbb{R}^n$. Potom ℓ^p*-norma* vektora $\mathbf{x}$ je

$$\|\mathbf{x}\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p} \qquad \text{pre } 1 \leq p < \infty,$$

$$\|\mathbf{x}\|_p = \max_{i=1,\dots,n} (|x_i|) \quad \text{pre } p = \infty.$$

Poznámka 11.22. Najpoužívanejšie špeciálne prípady ℓ^p-normy sú

(i) ℓ^1-norma $\|\mathbf{x}\|_1 = \sum_{i=1}^n |x_i|$,

(ii) ℓ^2-norma $\|\mathbf{x}\|_2 = (\sum_{i=1}^n x_i^2)^{1/2}$ (klasická euklidovská norma),

(iii) ℓ^∞-norma $\|\mathbf{x}\|_\infty = \max_i |x_i|$ (nazývaná tiež maximová norma),

Poznámka 11.23. Norma $\|\mathbf{x}\|_1$ sa nazýva tiež manhattanská norma alebo taxikárska norma. Maximová norma $\|\mathbf{x}\|_\infty$ je limitou $\|\mathbf{x}\|_p$ pre $p \to \infty$.

Napriek tomu, že ℓ^p-norma pre $p \neq 2$ nezodpovedá žiadnemu skalárnemu súčinu, stále spĺňa vlastnosti normy.

Veta 11.24 (O ℓ^p-normách)**.** Nech $1 \leq p \leq \infty$. Potom ℓ^p-norma $\|.\|_p$ spĺňa vlastnosti normy, teda definíciu 1.32.

Dôkaz. Ide o klasický výsledok z matematickej analýzy, tu ho ponecháme bez dôkazu. $\square$

Definícia 11.25 (Skalárny súčin a norma daná maticou)**.** Nech $\mathbf{A} \in \mathcal{S}^n_{++}$ a $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Potom skalárny súčin vektorov $\mathbf{x}$ a $\mathbf{y}$ daný maticou $\mathbf{A}$ je

$$\langle \mathbf{x}, \mathbf{y} \rangle_{\mathbf{A}} = \mathbf{x}^T \mathbf{A} \mathbf{y}$$

a norma $\mathbf{x}$ daná maticou $\mathbf{A}$ je

$$\|\mathbf{x}\|_{\mathbf{A}} = \sqrt{\mathbf{x}^T \mathbf{A} \mathbf{x}}.$$

Veta 11.26 (O skalárnom súčine danom maticou)**.** Nech $\mathbf{A} \in \mathcal{S}^n_{++}$ a $\langle \mathbf{x}, \mathbf{y} \rangle_{\mathbf{A}} = \mathbf{x}^T \mathbf{A} \mathbf{y}$ pre každé $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Potom $\langle , \rangle_{\mathbf{A}}$ je skalárny súčin na $\mathbb{R}^n$.

Dôkaz. Vlastnosti 1-3 sú priamočiare. Na dôkaz štvrtej vlastnosti potrebujeme $\mathbf{x}^T \mathbf{A} \mathbf{x} \geq 0$ pre každé $\mathbf{x} \in \mathbb{R}^n$, čo platí vďaka pozitívnej definitnosti $\mathbf{A}$. Zároveň z pozitívnej definitnosti $\mathbf{A}$ vyplýva $\mathbf{x}^T \mathbf{A} \mathbf{x} = 0$ iba pre $\mathbf{x} = \mathbf{0}$, čím je dokázaná aj druhá časť 4. vlastnosti. $\square$

Dôsledok 11.27. Nech $\mathbf{A} \in \mathcal{S}^n_{++}$. Potom podľa vety 1.33 je $\|.\|_{\mathbf{A}}$ norma na $\mathbb{R}^n$.

Príklad 11.28. Nech

$$\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

Nájdite a zakreslite ortogonálny doplnok vzhľadom na $\langle , \rangle_{\mathbf{A}}$ k priamke danej vektorom $\mathbf{x}$. Nájdite teda všetky vektory $\mathbf{y} \in \mathbb{R}^2$ také, že $\langle \mathbf{x}, \mathbf{y} \rangle_{\mathbf{A}} = 0$.

11.2.2 Projekcia vzhľadom na skalárny súčin daný maticou

V niektorých metódach analýzy dát sa pracuje aj s projekciami vzhľadom na skalárny súčin $\langle , \rangle_{\mathbf{A}}$ daný maticou $\mathbf{A} \in \mathcal{S}^n_{++}$ (napríklad v istých regresných modeloch, kde matica $\mathbf{A}$ charakterizuje "geometriu dát"). Projekcia vektora $\mathbf{y} \in \mathbb{R}^n$ na priestor $\mathcal{C}(\mathbf{X})$ pre $\mathbf{X} \in \mathbb{R}^{n\times k}$ vzhľadom na skalárny súčin $\langle , \rangle_{\mathbf{A}}$ je definovaná analogicky ako v podkapitole 7.2.

Definícia 11.29 (Projekcia vzhľadom na $\langle , \rangle_{\mathbf{A}}$)**.** Nech $\mathbf{A} \in \mathcal{S}^n_{++}$, $\mathbf{X} \in \mathbb{R}^{n\times k}$ a $\mathbf{y} \in \mathbb{R}^n$. Potom *projekcia vektora* $\mathbf{y}$ *na* $\mathcal{C}(\mathbf{X})$ *vzhľadom na* $\langle , \rangle_{\mathbf{A}}$ je taký vektor $\mathbf{z} \in \mathbb{R}^n$, ktorý spĺňa

1. $\mathbf{z} \in \mathcal{C}(\mathbf{X})$,
2. $\mathbf{y} - \mathbf{z} \perp_{\mathbf{A}} \mathcal{C}(\mathbf{X})$, čiže $\mathbf{x}^T\mathbf{A}(\mathbf{y}-\mathbf{z}) = 0$ pre každé $\mathbf{x} \in \mathcal{C}(\mathbf{X})$.

Lema 11.30 (Konzistentnosť normálnych rovníc pre skalárny súčin daný maticou)**.** Nech $\mathbf{A} \in \mathcal{S}^n_{++}$, $\mathbf{X} \in \mathbb{R}^{n\times k}$ a $\mathbf{y} \in \mathbb{R}^n$. Potom systém $\mathbf{X}^T\mathbf{A}\mathbf{X}\mathbf{b} = \mathbf{X}^T\mathbf{A}\mathbf{y}$ v premennej $\mathbf{b} \in \mathbb{R}^k$ je konzistentný.

Dôkaz. Vyjadrime $\mathbf{A} = \mathbf{A}^{1/2}\mathbf{A}^{1/2}$ a pripomeňme, že matica $\mathbf{A}^{1/2}$ je tiež symetrická. Potom $\mathbf{X}^T\mathbf{A}\mathbf{X}\mathbf{b} = \mathbf{X}^T\mathbf{A}\mathbf{y}$ vieme zapísať ako $\mathbf{X}^T\mathbf{A}^{1/2}\mathbf{A}^{1/2}\mathbf{X}\mathbf{b} = \mathbf{X}^T\mathbf{A}^{1/2}\mathbf{A}^{1/2}\mathbf{y}$ alebo tiež ako $\tilde{\mathbf{X}}^T\tilde{\mathbf{X}}\mathbf{b} = \tilde{\mathbf{X}}^T\tilde{\mathbf{y}}$, kde $\tilde{\mathbf{X}} = \mathbf{A}^{1/2}\mathbf{X}$ a $\tilde{\mathbf{y}} = \mathbf{A}^{1/2}\mathbf{y}$. Lenže systém $\tilde{\mathbf{X}}^T\tilde{\mathbf{X}}\mathbf{b} = \tilde{\mathbf{X}}^T\tilde{\mathbf{y}}$ je systémom normálnych rovníc pre dané $\tilde{\mathbf{X}}$ a $\tilde{\mathbf{y}}$, čiže je konzistentný podľa vety 7.12. Keďže tento systém sme získali len ako preznačenie pôvodného systému, tak aj pôvodný systém je konzistentný. □

Veta 11.31 (Projekcia vzhľadom na $\langle , \rangle_{\mathbf{A}}$)**.** Nech $\mathbf{A} \in \mathcal{S}^n_{++}$, $\mathbf{X} \in \mathbb{R}^{n\times k}$ a $\mathbf{y} \in \mathbb{R}^n$. Potom projekcia $\mathbf{z} \in \mathbb{R}^n$ vektora $\mathbf{y}$ na $\mathcal{C}(\mathbf{X})$ vzhľadom na skalárny súčin $\langle , \rangle_{\mathbf{A}}$ spĺňa

(i) $\mathbf{z} = \mathbf{X}\mathbf{b}$, kde $\mathbf{b}$ rieši systém rovníc $\mathbf{X}^T\mathbf{A}\mathbf{X}\mathbf{b} = \mathbf{X}^T\mathbf{A}\mathbf{y}$;

(ii) $\mathbf{z} = \mathbf{X}(\mathbf{X}^T\mathbf{A}\mathbf{X})^-\mathbf{X}^T\mathbf{A}\mathbf{y}$.

Dôkaz. Z vlastnosti 1 dostávame $\mathbf{z} = \mathbf{X}\mathbf{b}$ pre nejaký vektor $\mathbf{b} \in \mathbb{R}^k$. Z vlastnosti 2 dostávame, že $\mathbf{A}(\mathbf{y}-\mathbf{z})$ je vzhľadom na klasický skalárny súčin kolmé na $\mathcal{C}(\mathbf{X})$, teda že $\mathbf{X}^T\mathbf{A}(\mathbf{y}-\mathbf{z}) = \mathbf{0}_k$ (pozri lemu 7.8). Teda $\mathbf{X}^T\mathbf{A}\mathbf{y} = \mathbf{X}^T\mathbf{A}\mathbf{z}$, a po dosadení $\mathbf{z} = \mathbf{X}\mathbf{b}$ získavame $\mathbf{X}^T\mathbf{A}\mathbf{y} = \mathbf{X}^T\mathbf{A}\mathbf{X}\mathbf{b}$.

Jedno riešenie systému $\mathbf{X}^T\mathbf{A}\mathbf{X}\mathbf{b} = \mathbf{X}^T\mathbf{A}\mathbf{y}$ získame ako $\mathbf{b} = (\mathbf{X}^T\mathbf{A}\mathbf{X})^-\mathbf{X}^T\mathbf{A}\mathbf{y}$. Teda $\mathbf{z} = \mathbf{X}\mathbf{b} = \mathbf{X}(\mathbf{X}^T\mathbf{A}\mathbf{X})^-\mathbf{X}^T\mathbf{A}\mathbf{y}$, čím je dokázané (ii). □

Lema 11.32 (Existencia a jednoznačnosť projekcie vzhľadom na $\langle , \rangle_{\mathbf{A}}$)**.** Nech $\mathbf{A} \in \mathcal{S}^n_{++}$, $\mathbf{X} \in \mathbb{R}^{n\times k}$ a $\mathbf{y} \in \mathbb{R}^n$. Potom projekcia vektora $\mathbf{y}$ na $\mathcal{C}(\mathbf{X})$ vzhľadom na skalárny súčin $\langle , \rangle_{\mathbf{A}}$ existuje a je jediná.

Dôkaz. Hľadaná projekcia existuje podľa vety 11.31. Nech $\mathbf{z}_1, \mathbf{z}_2 \in \mathbb{R}^n$ spĺňajú vlastnosti 1,2 z definície 11.29. Potom

$$\|\mathbf{z}_1-\mathbf{z}_2\|^2_{\mathbf{A}} = \langle \mathbf{z}_1-\mathbf{z}_2, \mathbf{z}_1-\mathbf{z}_2\rangle_{\mathbf{A}} = \langle \mathbf{z}_1-\mathbf{y}+\mathbf{y}-\mathbf{z}_2, \mathbf{z}_1-\mathbf{z}_2\rangle_{\mathbf{A}} = \langle \mathbf{z}_1-\mathbf{y}, \mathbf{z}_1-\mathbf{z}_2\rangle_{\mathbf{A}}+\langle \mathbf{y}-\mathbf{z}_2, \mathbf{z}_1-\mathbf{z}_2\rangle_{\mathbf{A}}.$$

Zároveň platí, že $\mathbf{z}_1, \mathbf{z}_2 \in \mathcal{C}(\mathbf{X})$ podľa vlastnosti 1, a teda $\mathbf{z}_1 - \mathbf{z}_2 \in \mathcal{C}(\mathbf{X})$. Navyše, $\mathbf{z}_1 - \mathbf{y} \perp_{\mathbf{A}} \mathcal{C}(\mathbf{X})$ a $\mathbf{y}-\mathbf{z}_2 \perp_{\mathbf{A}} \mathcal{C}(\mathbf{X})$ podľa vlastnosti 2. Potom $\langle \mathbf{z}_1-\mathbf{y}, \mathbf{z}_1-\mathbf{z}_2\rangle_{\mathbf{A}} = 0$ a $\langle \mathbf{y}-\mathbf{z}_2, \mathbf{z}_1-\mathbf{z}_2\rangle_{\mathbf{A}} = 0$, čiže $\|\mathbf{z}_1 - \mathbf{z}_2\|^2_{\mathbf{A}} = 0$. Z kladnej definitnosti normy dostávame $\mathbf{z}_1 - \mathbf{z}_2 = \mathbf{0}$, teda $\mathbf{z}_1 = \mathbf{z}_2$, čím je dokázaná aj jednoznačnosť projekcie. □

Poznámka 11.33. Projekciu z vety 11.31 vieme vyjadriť aj ako $\mathbf{z} = \mathbf{P}_{\mathbf{A},\mathbf{X}}\mathbf{y}$, kde $\mathbf{P}_{\mathbf{A},\mathbf{X}} = \mathbf{X}(\mathbf{X}^T\mathbf{A}\mathbf{X})^{-}\mathbf{X}^T\mathbf{A}$.

Definícia 11.34 (Projektor vzhľadom na $\langle,\rangle_{\mathbf{A}}$)**.** Nech $\mathbf{A} \in \mathcal{S}^n_{++}$ a $\mathbf{X} \in \mathbb{R}^{n\times k}$. Potom *projekčná matica (projektor)* na $\mathcal{C}(\mathbf{X})$ vzhľadom na skalárny súčin $\langle,\rangle_{\mathbf{A}}$ je $\mathbf{P}_{\mathbf{A},\mathbf{X}} = \mathbf{X}(\mathbf{X}^T\mathbf{A}\mathbf{X})^{-}\mathbf{X}^T\mathbf{A}$.

Príklad 11.35. Dokážte, že matica $\mathbf{P}_{\mathbf{A},\mathbf{X}}$ je idempotentná. Projekčné matice vzhľadom na skalárne súčiny dané maticami však nemusia byť symetrické. Nájdite príklad takej matice $\mathbf{P}_{\mathbf{A},\mathbf{X}}$, ktorá nie je symetrická.

Príklad 11.36. Pre $\mathbf{A}$ a $\mathbf{x}$ z príkladu 11.28 vypočítajte maticu $\mathbf{P}_{\mathbf{A},\mathbf{x}}$.

11.2.3 Maticové normy

Lema 11.37 (Vzťah Frobeniovej normy a singulárnych čísel)**.** Nech matica $\mathbf{A}$ hodnosti r má singulárne čísla $\sigma_1, \ldots, \sigma_r$. Potom $\|\mathbf{A}\|_F^2 = \sigma_1^2 + \ldots + \sigma_r^2$.

Dôkaz. V značení poznámky 11.13 dostávame

$$\|\mathbf{A}\|_F^2 = \mathrm{tr}(\mathbf{A}^T\mathbf{A}) = \mathrm{tr}(\mathbf{U}\begin{pmatrix}\mathbf{\Sigma}_1^2 & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}\mathbf{U}^T) = \mathrm{tr}\begin{pmatrix}\mathbf{\Sigma}_1^2 & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix} = \mathrm{tr}(\mathbf{\Sigma}_1^2) = \sum_{i=1}^r \sigma_i^2.$$

□

Definícia 11.38 (Indukovaná norma)**.** Nech $\|.\|$ je vektorová norma na $\mathbb{R}^n$. Potom *maticová norma indukovaná vektorovou normou* $\|.\|$ je

$$\|\mathbf{A}\| = \max_{\mathbf{x}\neq\mathbf{0}_n} \frac{\|\mathbf{A}\mathbf{x}\|}{\|\mathbf{x}\|}$$

pre každé $\mathbf{A} \in \mathbb{R}^{m\times n}$.

Poznámka 11.39. Indukovanú maticovú normu tiež môžeme zapísať ako

$$\|\mathbf{A}\| = \max_{\mathbf{x}:\|\mathbf{x}\|=1} \|\mathbf{A}\mathbf{x}\|.$$

Lema 11.40 (O indukovanej norme)**.** Nech $\|.\|$ je vektorová norma. Potom príslušná indukovaná maticová norma spĺňa vlastnosti normy.

Dôkaz. Označme $\|.\|_M$ indukovanú maticovú normu. Dôkaz vlastnosti 1 z definície 1.32 je priamočiary. Na dôkaz vlastnosti 2 si najprv môžeme uvedomiť, že $\|\mathbf{A}\|_M \geq 0$, lebo $\|\mathbf{x}\| \geq 0$ pre ľubovoľné $\mathbf{x} \in \mathbb{R}^n$. Ostáva teda dokázať druhú časť vlastnosti. Ak $\mathbf{A} \neq \mathbf{0}_{m\times n}$, potom existuje $\mathbf{x} \in \mathbb{R}^n$, že $\mathbf{A}\mathbf{x} \neq \mathbf{0}_n$, čiže $\|\mathbf{A}\mathbf{x}\| > 0$. Potom aj

$$\|\mathbf{A}\|_M = \max_{\mathbf{x}\neq\mathbf{0}_n} \frac{\|\mathbf{A}\mathbf{x}\|}{\|\mathbf{x}\|} > 0.$$

Teda $\|\mathbf{A}\|_M = 0$ iba pre $\mathbf{A} = \mathbf{0}_{m\times n}$.

Na dôkaz vlastnosti 3 použime tvar z poznámky 11.39:

$$\begin{aligned}\|\mathbf{A}+\mathbf{B}\|_M &= \max_{\mathbf{x}:\|\mathbf{x}\|=1} \|(\mathbf{A}+\mathbf{B})\mathbf{x}\| = \max_{\mathbf{x}:\|\mathbf{x}\|=1} \|\mathbf{A}\mathbf{x}+\mathbf{B}\mathbf{x}\| \\ &\leq \max_{\mathbf{x}:\|\mathbf{x}\|=1} (\|\mathbf{A}\mathbf{x}\|+\|\mathbf{B}\mathbf{x}\|) \leq \max_{\mathbf{x}:\|\mathbf{x}\|=1} \|\mathbf{A}\mathbf{x}\| + \max_{\mathbf{x}:\|\mathbf{x}\|=1} \|\mathbf{B}\mathbf{x}\| \\ &= \|\mathbf{A}\|_M + \|\mathbf{B}\|_M,\end{aligned}$$

kde prvá nerovnosť vyplýva z trojuholníkovej nerovnosti pre vektorovú normu. □

Poznámka 11.41. Frobeniova norma $\|\mathbf{A}\|_F$ nie je indukovaná žiadnou vektorovou normou.

Veta 11.42 (Maticové normy indukované ℓ^p-normami)**.** Nech $\|.\|_p$ je ℓ^p-norma. Potom ňou indukovaná maticová norma $\|\mathbf{A}\|_p$ pre $\mathbf{A} \in \mathbb{R}^{m\times n}$ je:

(i) $\|\mathbf{A}\|_1 = \max_i \|\mathbf{a}_i\|_1$,

(ii) $\|\mathbf{A}\|_2 = \max_i \sigma_i$,

(iii) $\|\mathbf{A}\|_\infty = \max_j \|\mathbf{b}_j\|_1$,

kde $\sigma_1, \dots \sigma_r$ sú singulárne čísla matice $\mathbf{A}$, $\mathbf{a}_1, \dots \mathbf{a}_n$ sú jej stĺpce a $\mathbf{b}_1^T, \dots, \mathbf{b}_m^T$ sú jej riadky.

Dôkaz. Pre dôkaz časti (i) použime tvar z poznámky 11.39:

$$\|\mathbf{A}\|_1 = \max_{\sum_i |x_i|=1} \|\mathbf{A}\mathbf{x}\|_1. \tag{11.2}$$

Keďže $\mathbf{A}\mathbf{x} = x_1\mathbf{a}_1 + \dots x_n\mathbf{a}_n$, tak (11.2) je rovné

$$\max_{\sum_i |x_i|=1} \|x_1\mathbf{a}_1 + \dots + x_n\mathbf{a}_n\|_1,$$

čo je vzhľadom na trojuholníkovú nerovnosť pre vektorovú ℓ^1-normu menšie, nanajvýš rovné hodnote

$$\max_{\sum_i |x_i|=1} (\|x_1\mathbf{a}_1\|_1 + \dots + \|x_n\mathbf{a}_n\|_1) = \max_{\sum_i |x_i|=1} (|x_1|\|\mathbf{a}_1\|_1 + \dots + |x_n|\|\mathbf{a}_n\|_1).$$

Keď označíme $\mathbf{a}_{\max}$ ako stĺpec s najväčšou ℓ^1-normou, dostávame

$$\|\mathbf{A}\|_1 \leq \max_{\sum_i |x_i|=1} (|x_1|\|\mathbf{a}_{\max}\|_1 + \dots + |x_n|\|\mathbf{a}_{\max}\|_1) = \max_{\sum_i |x_i|=1} \|\mathbf{a}_{\max}\|_1 \sum_{i=1}^n |x_i| = \|\mathbf{a}_{\max}\|_1.$$

Túto nerovnosť dosiahneme ako rovnosť voľbou $\mathbf{x} = \mathbf{e}_{\max}$, kde $\mathbf{e}_{\max}$ má jednotku na pozícii zodpovedajúcej stĺpcu $\mathbf{a}_{\max}$ a nuly všade inde. Čiže $\|\mathbf{a}_{\max}\|_1$ je skutočne maximum, a teda $\|\mathbf{A}\|_1$ je rovné tejto hodnote.

Dôkaz časti (ii) je náročnejší, dá sa nájsť napríklad v [3] (časť 2.3.3).

Časť (iii) môžeme tiež dokázať v tvare poznámky 11.39:

$$\|\mathbf{A}\|_\infty = \max_{\max_i |x_i|=1} \max_{j=1,\dots,m} (|(\mathbf{A}\mathbf{x})_j|) = \max_{-1\leq x_i\leq 1} \max_{j=1,\dots,m} (|\mathbf{b}_j^T\mathbf{x}|).$$

Všimnime si, že hodnota $|\mathbf{b}_j^T\mathbf{x}|$ je pre $x_i \in [-1,1]$ rovná nanajvýš $|b_{j,1}| + \ldots + |b_{j,n}|$, kde $b_{j,k}$ je k-ty prvok $\mathbf{b}_j$. Potom

$$\|\mathbf{A}\|_\infty \leq \max_{-1\leq x_i \leq 1} \max_{j=1,\ldots,m} (|b_{j,1}| + \ldots + |b_{j,n}|) = \max_{-1\leq x_i \leq 1} \max_{j=1,\ldots,m} \|\mathbf{b}_j\|_1 = \max_{j=1,\ldots,m} \|\mathbf{b}_j\|_1,$$

kde poslednú rovnosť sme dosiahli uvedomením si, že $\max_{-1\leq x_i\leq 1}$ je už vo výraze vlastne zbytočné. Voľbou vektora $\mathbf{x}$ takou, že $x_i = 1$ ak $b_{\max,i} \geq 0$ a $x_i = -1$ ak $b_{\max,i} < 0$ ($b_{\max,i}$ vyjadruje i-ty prvok riadku s najväčšou ℓ^1-normou) dosiahneme nerovnosť ako rovnosť, čiže $\max_{j=1,\ldots,m} \|\mathbf{b}_j\|_1$ je skutočne hľadané maximum. □

Dôsledok 11.43. Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$. Potom $\|\mathbf{A}\|_1 = \|\mathbf{A}^T\|_\infty$.

Definícia 11.44 (Nukleárna norma)**.** Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$ má hodnosť r a $\sigma_1, \ldots, \sigma_r$ sú jej singulárne čísla. Potom *nukleárna norma* (angl. *nuclear norm* alebo tiež *trace norm*) matice $\mathbf{A}$ je

$$\|\mathbf{A}\|_N = \sigma_1 + \ldots + \sigma_r.$$

Poznámka 11.45. Ak $r = \mathrm{rank}(\mathbf{A}) = 0$, teda ak $\mathbf{A} = \mathbf{0}_{m\times n}$, tak pod $\|\mathbf{A}\|_N = \sigma_1 + \ldots + \sigma_r$ rozumieme $\|\mathbf{A}\|_N = 0$, lebo taká matica nemá žiadne (nenulové) singulárne čísla.

Lema 11.46 (O nukleárnej norme)**.** Nukleárna norma $\|.\|_N$ spĺňa vlastnosti normy.

Dôkaz. Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$, $c \in \mathbb{R}$ a $\mathbf{A} = \mathbf{U\Sigma V}^T$ je singulárny rozklad matice $\mathbf{A}$. Potom podľa lemy 11.16 singulárne čísla matice $c\mathbf{A}$ sú $|c|\sigma_1, \ldots, |c|\sigma_r$. Z toho vyplýva 1. vlastnosť normy z definície 1.32.

Hodnota $\|\mathbf{A}\|_N$ je zjavne nezáporná. Navyše, iba nulová matica má všetky singulárne čísla nulové, čiže $\|\mathbf{A}\|_N = 0$ iba pre $\mathbf{A} = \mathbf{0}_{m\times n}$, čím je dokázaná vlastnosť 2.

Dôkaz vlastnosti 3 je komplikovaný, dá sa získať napríklad pomocou vety IV.2.1 v [1]. □

Lema 11.47 (Alternatívne vyjadrenie nukleárnej normy)**.** Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$. Potom

$$\|\mathbf{A}\|_N = \mathrm{tr}((\mathbf{A}^T\mathbf{A})^{1/2})$$

Dôkaz. Označme $\sigma_1, \ldots, \sigma_r$ singulárne čísla matice $\mathbf{A}$. Potom matica $\mathbf{A}^T\mathbf{A}$ má vlastné čísla $\sigma_1^2, \ldots, \sigma_r^2, 0, \ldots, 0$ a matica $(\mathbf{A}^T\mathbf{A})^{1/2}$ má podľa lemy 10.54 vlastné čísla $\sigma_1, \ldots, \sigma_r, 0, \ldots, 0$ (keďže $\sigma_i > 0$, a teda $|\sigma_i| = \sigma_i$). Dôkaz uzavrieme pozorovaním, že stopa symetrickej matice je rovná súčtu jej vlastných čísel. □

Príklad 11.48. Všimnite si, že predpis pre Frobeniovu normu $\|\mathbf{A}\|_F = (\mathrm{tr}(\mathbf{A}^T\mathbf{A}))^{1/2}$ a vyjadrenie nukleárnej normy $\|\mathbf{A}\|_N = \mathrm{tr}((\mathbf{A}^T\mathbf{A})^{1/2})$ z lemy 11.47 sú vo všeobecnosti dva rôzne výrazy. Nájdite maticu $\mathbf{A}$, pre ktorú platí $\|\mathbf{A}\|_F \neq \|\mathbf{A}\|_N$, a takú maticu $\mathbf{A}$, pre ktorú $\|\mathbf{A}\|_F = \|\mathbf{A}\|_N$.

11.2.4 Číslo podmienenosti

Definícia 11.49 (Číslo podmienenosti)**.** *Číslo podmienenosti* regulárnej matice $\mathbf{A} \in \mathbb{R}^{n\times n}$ je $\kappa(\mathbf{A}) = \|\mathbf{A}\| \cdot \|\mathbf{A}^{-1}\|$.

Poznámka 11.50. Vidíme, že číslo podmienenosti závisí od zvolenej maticovej normy. Najčastejšie je používané číslo podmienenosti dané normou $\|\mathbf{A}\|_2 = \max_i \sigma_i$. Odteraz budeme uvažovať práve toto číslo podmienenosti.

Veta 11.51 (Číslo podmienenosti dané ℓ^2-normou)**.** Nech $\mathbf{A} \in \mathbb{R}^{n\times n}$ je regulárna a $\sigma_1 \geq \ldots \geq \sigma_n > 0$ sú jej singulárne čísla. Potom

$$\kappa(\mathbf{A}) = \frac{\sigma_1}{\sigma_n}.$$

Dôkaz. Z definície vidíme, že $\|\mathbf{A}\|_2 = \sigma_1$. Podľa lemy 11.16 vieme, že singulárne čísla $\mathbf{A}^{-1}$ sú $\sigma_1^{-1}, \ldots, \sigma_n^{-1}$. Najväčšie spomedzi týchto čísel je σ_n^{-1}, a teda $\|\mathbf{A}\|_2 = \sigma_n^{-1}$, čím je dôkaz ukončený. □

Veta 11.52 (Číslo podmienenosti a stabilita lineárneho systému)**.** Nech $\mathbf{A} \in \mathbb{R}^{n\times n}$ je regulárna a $\mathbf{b}, \mathbf{e} \in \mathbb{R}^m$. Potom

$$\max_{\mathbf{b},\mathbf{e}\neq\mathbf{0}_m} \left(\frac{\|\mathbf{A}^{-1}\mathbf{e}\|}{\|\mathbf{e}\|} \cdot \frac{\|\mathbf{b}\|}{\|\mathbf{A}^{-1}\mathbf{b}\|} \right) = \kappa(\mathbf{A}),$$

kde $\|.\|$ je euklidovská norma.

Dôkaz. Nech $\sigma_1 \geq \ldots \geq \sigma_n$ sú singulárne čísla matice $\mathbf{A}$. Máme

$$y := \max_{\mathbf{b},\mathbf{e}\neq\mathbf{0}_m} \left(\frac{\|\mathbf{A}^{-1}\mathbf{e}\|}{\|\mathbf{e}\|} \cdot \frac{\|\mathbf{b}\|}{\|\mathbf{A}^{-1}\mathbf{b}\|} \right) = \max_{\mathbf{e}\neq\mathbf{0}_m} \frac{\|\mathbf{A}^{-1}\mathbf{e}\|}{\|\mathbf{e}\|} \max_{\mathbf{b}\neq\mathbf{0}_m} \frac{\|\mathbf{b}\|}{\|\mathbf{A}^{-1}\mathbf{b}\|}.$$

Prvé maximum je z definície rovné $\|\mathbf{A}^{-1}\|_2$, čo je rovné σ_n^{-1}, keďže to je najväčšie singulárne číslo matice $\mathbf{A}^{-1}$ (lema 11.16). Označme $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$. Potom $\mathbf{b} = \mathbf{A}\mathbf{x}$, a teda

$$y = \sigma_n^{-1} \max_{\mathbf{x}\neq\mathbf{0}_n} \frac{\|\mathbf{A}\mathbf{x}\|}{\|\mathbf{x}\|} = \sigma_n^{-1}\sigma_1 = \kappa(\mathbf{A}).$$

z definície $\|\mathbf{A}\|_2$. □

Poznámka 11.53. Číslo podmienenosti vyjadruje, ako veľmi sa môže zmeniť riešenie rovnice $\mathbf{A}\mathbf{x} = \mathbf{b}$ pre regulárnu $\mathbf{A}$, ak sa mierne zmení pravá strana. Označme drobnú perturbáciu pravej strany ako $\mathbf{e} \in \mathbb{R}^m$. Nech $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ je riešenie $\mathbf{A}\mathbf{x} = \mathbf{b}$ a $\mathbf{x}_e = \mathbf{A}^{-1}(\mathbf{b} + \mathbf{e})$ je riešenie perturbovaného systému $\mathbf{A}\mathbf{x} = \mathbf{b} + \mathbf{e}$. Teda $\mathbf{x}_e = \mathbf{A}^{-1}\mathbf{b} + \mathbf{A}^{-1}\mathbf{e} = \mathbf{x} + \mathbf{A}^{-1}\mathbf{e}$. Zmena v riešení je teda $\delta\mathbf{x} = \mathbf{A}^{-1}\mathbf{e}$ a zmena v pravej strane je $\delta\mathbf{b} = \mathbf{e}$.

Potom pomer relatívnej zmeny v riešení oproti relatívnej zmene v pravej strane je

$$\frac{\|\delta\mathbf{x}\|}{\|\mathbf{x}\|} \Big/ \frac{\|\delta\mathbf{b}\|}{\|\mathbf{b}\|} = \frac{\|\mathbf{A}^{-1}\mathbf{e}\|}{\|\mathbf{A}^{-1}\mathbf{b}\|} \Big/ \frac{\|\mathbf{e}\|}{\|\mathbf{b}\|} = \frac{\|\mathbf{A}^{-1}\mathbf{e}\|}{\|\mathbf{e}\|} \cdot \frac{\|\mathbf{b}\|}{\|\mathbf{A}^{-1}\mathbf{b}\|}.$$

Číslo podmienenosti je maximum tohto výrazu – teda vyjadruje, ako veľmi sa môže riešenie zmeniť vplyvom zmeny v pravej strane. Práca s maticou, ktorá má vysoké číslo podmienenosti, môže byť teda numericky nestabilná.

Príklad 11.54. Nech

$$\mathbf{A} = \begin{pmatrix} 1 & 0.9 \\ 0.9 & 1 \end{pmatrix} \quad \text{a} \quad \mathbf{b} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Číslo podmienenosti matice $\mathbf{A}$ je $\kappa(\mathbf{A}) = 19$, teda zmena v pravej strane sa môže prejaviť až v 19-násobne väčšej zmene riešenia.

Riešenie rovnice $\mathbf{Ax} = \mathbf{b}$ je približne $\mathbf{x} = (0{,}5263; 0{,}5263)^T$. Ak na pravej strane uvážime drobnú chybu $\mathbf{e} = (-0{,}02; 0{,}01)^T$, tak riešenie $\mathbf{Ax} = \mathbf{b} + \mathbf{e}$ je $\mathbf{x}_e = (0{,}3737; 0{,}6737)^T$. Pomer relatívnych chýb je

$$\frac{\|\delta\mathbf{x}\|}{\|\mathbf{x}\|} \Big/ \frac{\|\delta\mathbf{b}\|}{\|\mathbf{b}\|} = \frac{\sqrt{(-0{,}1526)^2 + 0{,}1474^2}}{\sqrt{2} \cdot 0{,}5263} \Big/ \frac{\sqrt{(-0{,}02)^2 + (0{,}01)^2}}{\sqrt{2}} \approx 18{,}03.$$

Teda pre takúto chybu pravej strany sa riešenie relatívne zmení až 18-krát viac.

Keby sme namiesto matice $\mathbf{A}$ mali maticu

$$\mathbf{B} = \begin{pmatrix} 1 & 0{,}09 \\ 0{,}09 & 1 \end{pmatrix}$$

s číslom podmienenosti rovným približne 1,20, tak overte, že pri tom istom $\mathbf{b}$ a $\mathbf{e}$ by sme dostali

$$\frac{\|\delta\mathbf{x}\|}{\|\mathbf{x}\|} \Big/ \frac{\|\delta\mathbf{b}\|}{\|\mathbf{b}\|} = 1{,}18.$$

Nájdite tiež takú maticu $\mathbf{C}$ v tvare

$$\mathbf{C} = \begin{pmatrix} 1 & c \\ c & 1 \end{pmatrix}$$

pre nejaké $c \in \mathbb{R}$, že $\mathbf{C}$ má väčšie číslo podmienenosti ako $\mathbf{A}$. Následne vypočítajte, aká je relatívna zmena riešenia oproti relatívnej zmene pravej strany pre zvolenú $\mathbf{C}$ a pre $\mathbf{b}, \mathbf{e}$ určené vyššie.

11.3 Úlohy na precvičenie

Úloha 11.1. Nájdite singulárny rozklad matice

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ -1 & 0 & 1 & 0 \end{pmatrix}.$$

Úloha 11.2. Ukážte, že singulárny rozklad kladne semidefinitnej matice $\mathbf{A}$ je vlastne spektrálnym rozkladom tejto matice. *Pomôcka:* skúste vychádzať z dôsledku 11.12.

Úloha 11.3. Aký je vzťah medzi singulárnym rozkladom symetrickej matice $\mathbf{A}$ a jej spektrálnym rozkladom? Špeciálne určte vzťah medzi (i) singulárnymi a vlastnými číslami matice $\mathbf{A}$, (ii) singulárnymi a vlastnými vektormi matice $\mathbf{A}$.

Úloha 11.4. Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$ a $\|.\|$ je euklidovská norma. Ukážte, že $\max\|\mathbf{Ax}\| = \sigma_1$, kde σ_1 je najväčšie singulárne číslo matice $\mathbf{A}$ a maximum je cez všetky vektory $\mathbf{x} \in \mathbb{R}^n$ také, že $\|\mathbf{x}\| = 1$.

Úloha 11.5. Nech $\mathbf{A} \in \mathbb{R}^{m\times n}$ má hodnosť r a nech $\mathbf{u}_1, \ldots, \mathbf{u}_m$ a $\mathbf{v}_1, \ldots, \mathbf{v}_n$ sú jej singulárne vektory. Definujme

$$\mathbf{B} = \begin{pmatrix} \mathbf{0} & \mathbf{A} \\ \mathbf{A}^T & \mathbf{0} \end{pmatrix}.$$

a) Ukážte, že vektory $(\mathbf{u}_i^T, \mathbf{v}_i^T)^T$ a $(-\mathbf{u}_i^T, \mathbf{v}_i^T)^T$ $(i = 1, \ldots, \min\{m, n\})$ sú vlastné vektory matice $\mathbf{B}$. Určte k akým vlastným číslam prislúchajú. b) Demonštrujte časť (a) na matici $\mathbf{A}$ z úlohy 11.1.

Úloha 11.6. Pomocou singulárneho rozkladu nájdite pseudoinverziu matice

$$\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}.$$

Úloha 11.7. Nech $\mathbf{A} \in \mathbb{R}^{m \times n}$ má hodnosť r. Pomocou singulárneho rozkladu vyjadrite maticu $\mathbf{A}$ ako súčet r matíc hodnosti 1.

Úloha 11.8. Nech $\mathbf{Q} \in \mathbb{R}^{n \times n}$ je ortogonálna. Nájdite všetky jej singulárne čísla a jej singulárny rozklad. Pre $n = 2$ alebo $n = 3$ si zvoľte ortogonálnu $\mathbf{Q}$ a ukážte, že má aspoň dva rôzne singulárne rozklady (líšiace sa nie iba v znamienkach singulárnych vektorov).

Úloha 11.9. Nech

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & -1 \end{pmatrix}.$$

Nájdite $\|\mathbf{A}\|_F$, $\|\mathbf{A}\|_1$, $\|\mathbf{A}\|_2$, $\|\mathbf{A}\|_\infty$ a $\|\mathbf{A}\|_N$.

Úloha 11.10. Nech $f(\mathbf{x}) = \sqrt{2x_1^2 + x_2^2 + 2x_1x_2 + 3x_3^2}$ pre $\mathbf{x} \in \mathbb{R}^3$. Je funkcia $f(\mathbf{x})$ normou na $\mathbb{R}^3$? Návod: skúste vyjadriť $f(\mathbf{x})$ ako $\sqrt{\mathbf{x}^T\mathbf{A}\mathbf{x}}$.

Úloha 11.11. Nech $\mathbf{A} \in \mathcal{S}_+^n$ je singulárna a definujme $f(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T\mathbf{A}\mathbf{y}$ pre $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Odvoďte, ktoré vlastnosti skalárneho súčinu $f(\mathbf{x}, \mathbf{y})$ spĺňa a ktoré nespĺňa.

Úloha 11.12. Nech $\|.\|$ je Frobeniova norma a nech $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times n}$. Ukážte, že a) $\|\mathbf{A}\|^2 = \sum_{i=1}^n \|\mathbf{a}_i\|^2$, kde $\mathbf{a}_1, \dots, \mathbf{a}_n$ sú riadky matice $\mathbf{A}$; b) $\|\mathbf{AB}\| \leq \|\mathbf{A}\| \cdot \|\mathbf{B}\|$. Návod: použite (a) a Cauchyho-Schwarzovu nerovnosť pre $\mathbf{a}_i$ (i-ty riadok matice $\mathbf{A}$) a $\mathbf{b}_j$ (j-ty stĺpec matice $\mathbf{B}$).

Literatúra

[1] Bhatia, R.: *Matrix analysis*, Springer, New York, 1997

[2] Gentle, J. E.: *Matrix algebra: theory, computations and applications in statistics*, 2nd edition, Springer, New York, 2017

[3] Golub, G. H., Van Loan, C. F.: *Matrix computations*, 4th edition, The John Hopkins University Press, Baltimore, 2013

[4] Harville, D. A.: *Matrix algebra from a statistician's perspective*, Springer, New York, 1997.

[5] Horn, R. A., Johnson C. R.: *Matrix analysis*, 2nd edition, Cambridge University Press, Cambridge, 2012

[6] Janková, K., Pázman, A.: *Pravdepodobnosť a štatistika*, Univerzita Komenského, Bratislava, 2011.

[7] Korbaš, J.: *Prednášky z lineárnej algebry a geometrie*, Univerzita Komenského, Bratislava, 2013.

[8] Lay, D. C., Lay, S. R., McDonald, J. J.: *Linear algebra and its applications*, 5th edition, Pearson, Boston, 2015.

[9] Pázman, A., Lacko, V.: *Prednášky z regresných modelov*, Univerzita Komenského, Bratislava, 2012.

[10] R Core Team: *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, 2021

[11] Seber, G. A. F.: *A matrix handbook for statisticians*, Wiley, New Jersey, 2008.

[12] Seneta, E.: *Non-negative matrices and Markov chains*, revised printing, Springer, New York, 2006.

[13] Searle, S. R., Khuri, A. I.: *Matrix algebra useful for statistics*, 2nd edition, Wiley, New Jersey, 2017

[14] Strang, G.: *Linear algebra and its applications*, 4th edition, Thomson Brooks/Cole, Belmont, 2006.

[15] Strang, G.: *Linear algebra and learning from data*, Wellesley - Cambridge Press, Wellesley, 2019