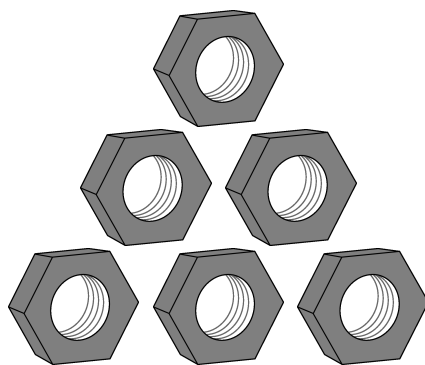


Milan Hladík

# Matice

(text k přednášce Lineární algebra III)

10. listopadu 2021





Tento text zatím slouží jako doprovodný text přednášky Lineární algebra III na studiu informatiky Matematicko-fyzikální fakulty Univerzity Karlovy v Praze, a navazuje na základní kurzy Lineární algebra I a II.

Autor čerpal především z knih Horn and Johnson [1985]; Horn and Johnson [1991]; Meyer [2000]; Prasolov [1994]; Rohn [2004]; Strang [1988].

Případné připomínky a chyby zasílejte prosím na adresu [hladik@kam.mff.cuni.cz](mailto:hladik@kam.mff.cuni.cz).

*Exempla movent, theoriae trahunt.*



# Obsah

<b>Obsah</b>	<b>5</b>
<b>1 Připomenutí</b>	<b>7</b>
<b>2 Vlastní čísla</b>	<b>9</b>
2.1 Vlastní čísla obecných matic . . . . .	9
2.2 Vlastní čísla symetrických matic . . . . .	10
<b>3 SVD rozklad</b>	<b>13</b>
3.1 Konstrukce a souvislost s vlastními čísly . . . . .	13
3.2 Aplikace SVD rozkladu . . . . .	15
<b>4 Pseudoinverzní matice</b>	<b>21</b>
4.1 Mooreova–Penroseova pseudoinverze . . . . .	21
4.2 Drazinova pseudoinverze . . . . .	23
<b>5 Maticová norma</b>	<b>25</b>
5.1 Definice a příklady, indukovaná norma . . . . .	25
5.2 Spektrální poloměr versus norma matice . . . . .	28
5.3 Mocninné posloupnosti . . . . .	30
5.4 Ortogonálně invariantní normy . . . . .	30
5.5 Další aplikace maticové normy . . . . .	31
<b>6 Číslo podmíněnosti</b>	<b>35</b>
6.1 Číslo podmíněnosti za spektrální normy . . . . .	36
6.2 Číslo podmíněnosti při řešení soustav lineárních rovnic . . . . .	37
6.3 Číslo podmíněnosti při počítání vlastních čísel . . . . .	38
<b>7 Perturbace</b>	<b>39</b>
7.1 Perturbace vlastních čísel . . . . .	39
7.2 Spojitost vlastních čísel . . . . .	41
7.3 Maticová derivace . . . . .	42
<b>8 Nezáporné a kladné matice</b>	<b>47</b>
8.1 Základní výsledky pro nezáporné matice . . . . .	47
8.2 Specifické výsledky pro kladné matice . . . . .	48
8.3 Specifické výsledky pro nezáporné matice . . . . .	49
8.4 Aplikace – verifikace lineárních soustav . . . . .	49
<b>9 Maticové funkce a mocninné řady</b>	<b>53</b>
9.1 Maticové funkce . . . . .	53
9.2 Maticová exponenciála . . . . .	54

<b>10 Nestandardní maticové součiny</b>	<b>59</b>
10.1 Kroneckerův součin . . . . .	59
10.1.1 Maticové soustavy rovnic . . . . .	61
10.2 Hadamardův součin . . . . .	62
<b>11 Kdo nemá dosti pozitivní semidefinitnosti</b>	<b>65</b>
<b>12 Speciální matice</b>	<b>69</b>
12.1 M-matice . . . . .	69
<b>13 Další témata</b>	<b>73</b>
<b>Značení</b>	<b>75</b>
<b>Literatura</b>	<b>77</b>

# Kapitola 1

## Připomenutí

Zde rychle připomeneme některé pojmy z výsledky ze základního kurzu lineární algebry [Hladík, 2017].

**Skalární součin a norma.** Pokud nezmíníme jinak, používáme standardní skalární součin v  $\mathbb{R}^n$ , definovaný jako  $\langle x, y \rangle = x^T y$ , a eukleidovskou normu  $\|x\| = \sqrt{x^T x}$ .

**Maticové prostory.** Pro matici  $A \in \mathbb{R}^{m \times n}$  platí

- $\mathcal{R}(A)^\perp = \text{Ker}(A)$ ,
- $\text{Ker}(A^T A) = \text{Ker}(A)$ ,
- $\mathcal{R}(A^T A) = \mathcal{R}(A)$ ,
- $\mathcal{S}(AA^T) = \mathcal{S}(A)$ ,
- $\text{rank}(A^T A) = \text{rank}(A)$ .

**Ortonormální systém.** Vektory  $x_1, \dots, x_m \in \mathbb{C}^n$  tvoří ortonormální systém, pokud mají eukleidovskou normu 1 a jsou navzájem kolmé, to jest  $\|x_i\| = 1$  pro všechna  $i$  a  $\langle x_i, x_j \rangle = 0$  pro  $i \neq j$ .

**Projekce.** (Ortogonalní) projekcí vektoru  $x \in V$  do podprostoru  $U$  je takový vektor  $x_U \in U$ , který splňuje

$$\|x - x_U\| = \min_{y \in U} \|x - y\|.$$

Tedy ze všech vektorů z  $U$  je to ten, který je k vektoru  $x$  nejbližší. Ekvivalentně projekci můžeme charakterizovat podmínkou  $x - x_U \in U^\perp$ , tudíž rozdíl vektoru a jeho projekce je kolmý na všechny vektory podprostoru  $U$  a žádný jiný vektor tuto vlastnost nemá.

Zobrazení projekce  $x \mapsto x_U$  je lineárním zobrazením, a proto jde reprezentovat maticově. Pokud projektujeme do sloupcového prostoru  $\mathcal{S}(A)$  matice  $A \in \mathbb{R}^{m \times n}$  hodnosti  $n$ , pak  $x_U = A(A^T A)^{-1} A^T x$ . Tudíž matice projekce je  $A(A^T A)^{-1} A^T$ .

**Ortogonalní a unitární matice.** Matice  $Q \in \mathbb{R}^{n \times n}$  je *ortogonalní*, pokud  $Q^T Q = I_n$ . Matice  $Q \in \mathbb{C}^{n \times n}$  je *unitární*, pokud  $\overline{Q}^T Q = I_n$ .

Ekvivalentní charakterizace, kdy je matice  $A \in \mathbb{R}^{n \times n}$  ortogonalní, jsou

- $Q$  je regulární a  $Q^{-1} = Q^T$ ,
- sloupce  $Q$  tvoří ortonormální bázi  $\mathbb{R}^n$ .

Ortogonalní matice jsou uzavřené na součin a transpozici. Lineární zobrazení  $x \mapsto Qx$  s ortogonalní maticí  $Q$  zachovává úhly a délky. Nijak tedy geometrické objekty nedeformuje, pouze je otočí resp. převrátí.

**Vlastní čísla.** Buď  $A \in \mathbb{C}^{n \times n}$ . Pak  $\lambda \in \mathbb{C}$  je *vlastní číslo* matice  $A$  a  $x \in \mathbb{C}^n$  jemu příslušný *vlastní vektor*, pokud  $Ax = \lambda x$ ,  $x \neq 0$ . Levý vlastní vektor matice  $A$  se pak definuje jako vlastní vektor matice  $A^T$ .

Vlastních čísel matice  $A \in \mathbb{C}^{n \times n}$  je  $n$  včetně násobností. Největší absolutní hodnota z vlastních čísel se nazývá spektrální poloměr a značí  $\rho(A)$ .

Každá matice  $A \in \mathbb{C}^{n \times n}$  je podobná matici v *Jordanově normálním tvaru*, to jest existuje regulární  $S$  taková, že  $A = SJS^{-1}$ , kde  $J$  je Jordanova matice. Jordanova matice je blokově diagonální matice a na blokové diagonále jsou Jordanovy buňky. Jordanova buňka příslušná vlastnímu číslu  $\lambda$  má tvar

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

. Pokud matice  $J$  je diagonální, pak říkáme, že  $A$  je diagonalizovatelná a píše se  $A = SAS^{-1}$ . Diagonální matice  $\Lambda$  má na diagonále vlastní čísla matice  $A$ , matice  $S$  má ve sloupcích příslušné vlastní vektory a matice  $S^{-1}$  má v řádcích příslušné levé vlastní vektory.

Speciální vlastnosti má reálná symetrická matice  $A \in \mathbb{R}^{n \times n}$ . Ta má vždy reálná vlastní čísla a je diagonalizovatelná. Můžeme ji též rozložit tzv. spektrálním rozkladem jako

$$A = Q\Lambda Q^T,$$

kde  $Q \in \mathbb{R}^{n \times n}$  je ortogonální a  $\Lambda \in \mathbb{R}^{n \times n}$  diagonální. Navíc matice  $\Lambda$  má na diagonále vlastní čísla matice  $A$  a matice  $Q$  má ve sloupcích příslušné vlastní vektory.

**Positivní semidefinitnost.** Symetrická matice  $A \in \mathbb{R}^{n \times n}$  je *pozitivně semidefinitní*, pokud je splněna jedna z následujících ekvivalentních podmínek:

- $x^T Ax \geq 0$  pro všechna  $x \in \mathbb{R}^n$ ,
- vlastní čísla  $A$  jsou nezáporná,
- existuje matice  $U \in \mathbb{R}^{m \times n}$  taková, že  $A = U^T U$ ,
- determinanty všech hlavních podmatic jsou nezáporné.  
(hlavní podmatice je matice, která vznikne z  $A$  odstraněním určitého, i nulového, počtu řádků a sloupců s týmiž indexy)

**Positivní definitnost.** Symetrická matice  $A \in \mathbb{R}^{n \times n}$  je *pozitivně definitní*, pokud je splněna jedna z následujících ekvivalentních podmínek:

- $x^T Ax > 0$  pro všechna  $x \in \mathbb{R}^n \setminus \{0\}$ ,
- vlastní čísla  $A$  jsou kladná,
- existuje matice  $U \in \mathbb{R}^{m \times n}$  hodnosti  $n$  taková, že  $A = U^T U$ ,
- determinanty všech hlavních vedoucích podmatic jsou kladné.  
(hlavní vedoucí podmatice je levá horní podmatice  $A$  velikosti  $i$ , tj. vznikne z  $A$  odstraněním posledních  $n - i$  řádků a sloupců)



# Kapitola 2

## Vlastní čísla

### 2.1 Vlastní čísla obecných matic

Schurův rozklad je důležitým výsledkem v teorii vlastních čísel. Stejně jako Jordanova normální forma existuje pro každou čtvercovou matici, ale využívá unitárních matic a v tom spočívá jeho síla.

**Věta 2.1** (Schurův rozklad). *Pro každou matici  $A \in \mathbb{C}^{n \times n}$  existuje unitární  $U \in \mathbb{C}^{n \times n}$  a horní trojúhelníková  $T \in \mathbb{C}^{n \times n}$  takové, že  $A = UTU^*$ .*

*Důkaz.* Matematickou indukcí podle  $n$  (analogicky jako důkaz spektrálního rozkladu symetrické matice). Pro  $n = 1$  je to zřejmé, tak uvažujme indukční krok.

Buď  $\lambda \in \mathbb{C}$  vlastní číslo  $A$  a  $x \in \mathbb{C}^n$ ,  $\|x\|_2 = 1$ , odpovídající vlastní vektor. Doplníme  $x$  na unitární matici  $\tilde{V} = (x \mid V)$ . Pak

$$\tilde{V}^* A \tilde{V} = \begin{pmatrix} x^* \\ V^* \end{pmatrix} (Ax \mid AV) = \begin{pmatrix} x^* \\ V^* \end{pmatrix} (\lambda x \mid AV) = \begin{pmatrix} \lambda & x^* AV \\ 0 & V^* AV \end{pmatrix}.$$

Podle indukčního předpokladu existuje unitární  $W \in \mathbb{C}^{(n-1) \times (n-1)}$  taková, že  $\tilde{T} := W^* V^* AVW$  je horní trojúhelníková. Nyní

$$\begin{pmatrix} 1 & 0^T \\ 0 & W^* \end{pmatrix} \tilde{V}^* A \tilde{V} \begin{pmatrix} 1 & 0^T \\ 0 & W \end{pmatrix} = \begin{pmatrix} 1 & 0^T \\ 0 & W^* \end{pmatrix} \begin{pmatrix} \lambda & x^* AV \\ 0 & V^* AV \end{pmatrix} \begin{pmatrix} 1 & 0^T \\ 0 & W \end{pmatrix} = \begin{pmatrix} \lambda & x^* AVW \\ 0 & W^* V^* AVW \end{pmatrix} = \begin{pmatrix} \lambda & x^* AVW \\ 0 & \tilde{T} \end{pmatrix}.$$

Tedy  $A = UTU^*$ , kde

$$U = \tilde{V} \begin{pmatrix} 1 & 0^T \\ 0 & W \end{pmatrix}, \quad T = \begin{pmatrix} \lambda & x^* AVW \\ 0 & \tilde{T} \end{pmatrix}. \quad \square$$

Protože matice  $T$  je podobná matici  $A$ , vlastní čísla matice  $A$  jsou na diagonále  $T$ . Pro reálnou matici  $A$  bohužel stále mohou být  $U, T$  komplexní, ale existují následující varianty.

**Věta 2.2** (Schurův rozklad, reálná varianta I.). *Pro každou matici  $A \in \mathbb{R}^{n \times n}$  existuje ortogonální  $Q \in \mathbb{R}^{n \times n}$  a blokově horní trojúhelníková  $T \in \mathbb{R}^{n \times n}$  takové, že  $A = QTQ^T$ . Bloky matice  $T$  jsou velikosti 1 nebo 2, přičemž ty velikosti 2 obsahují komplexně sdružená vlastní čísla.*

**Věta 2.3** (Schurův rozklad, reálná varianta II.). *Pro každou matici  $A \in \mathbb{R}^{n \times n}$  existuje regulární  $S \in \mathbb{R}^{n \times n}$  a blokově horní trojúhelníková  $T \in \mathbb{R}^{n \times n}$  takové, že  $A = STS^{-1}$ . Bloky matice  $T$  jsou velikosti 1 nebo 2, přičemž ty velikosti 2 mají tvar  $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$  a obsahují komplexně sdružená vlastní čísla  $a \pm ib$ .*

Jedním z důsledků Schurova rozkladu je pozorování, že libovolně blízko každé matice je diagonalizovatelná matice (viz věta 5.27).

Matice, jejichž Schurův rozklad má matici  $T$  diagonální, se nazývají *normální*. Typickým příkladem normální matice je symetrická reálná matice, protože její spektrální rozklad  $Q\Lambda Q^T$  ( $Q$  ortogonální a  $\Lambda$  diagonální) je přesně Schurův rozklad. Jiné než symetrické matice rozklad  $Q\Lambda Q^T$  nemají, protože výsledkem je symetrická matice. Nicméně díky komplexním maticím  $U, T$  v Schurově rozkladu mohou být normální i jiné reálné matice než jen symetrické – například antisymetrické ( $A = -A^T$ ) nebo ortogonální. Normální matice mají podobně pěkné vlastnosti jako symetrické matice, a některé z výsledků následující sekce 2.2 platí analogicky i pro ně.

## 2.2 Vlastní čísla symetrických matic

Rayleighova–Ritzova formule je elegantní vyjádření největšího a nejmenšího vlastního čísla symetrické matice. Říká, že jsou to největší resp. nejmenší hodnoty kvadratické funkce  $f(x) = x^T Ax$  na jednotkové eukleidovské sféře.

**Věta 2.4** (Rayleigh–Ritz). *Nechť  $\lambda_1 \geq \dots \geq \lambda_n$  jsou vlastní čísla symetrické matice  $A \in \mathbb{R}^{n \times n}$ . Pak*

$$\lambda_1 = \max_{x: \|x\|_2=1} x^T Ax, \quad \lambda_n = \min_{x: \|x\|_2=1} x^T Ax.$$

*Důkaz.* Pouze pro  $\lambda_1$ , druhá část je analogická.

Nerovnost „ $\leq$ “: Buď  $x_1$  vlastní vektor příslušný k  $\lambda_1$  normovaný  $\|x_1\|_2 = 1$ . Pak  $Ax_1 = \lambda_1 x_1$ . Přenásobením  $x_1^T$  zleva dostaneme

$$\lambda_1 = \lambda_1 x_1^T x_1 = x_1^T Ax_1 \leq \max_{x: \|x\|_2=1} x^T Ax.$$

Nerovnost „ $\geq$ “: Buď  $x \in \mathbb{R}^n$  libovolný vektor takový, že  $\|x\|_2 = 1$ . Označme  $y := Q^T x$ , pak  $\|y\|_2 = 1$ . S využitím spektrálního rozkladu  $A = Q\Lambda Q^T$  dostaneme:

$$x^T Ax = x^T Q\Lambda Q^T x = y^T \Lambda y = \sum_{i=1}^n \lambda_i y_i^2 \leq \sum_{i=1}^n \lambda_1 y_i^2 = \lambda_1 \|y\|_2^2 = \lambda_1. \quad \square$$

**Poznámka 2.5.** Symetrické matice nejen, že mají specifické vlastnosti, ale obecně se chovají „rozumněji“ než ty obecné. Uvažme například zobrazení  $x \mapsto Ax$ , kde

$$A = \begin{pmatrix} 1 & K \\ 0 & 1 \end{pmatrix},$$

kde  $K$  je libovolně velké. Matice  $A$  má jen jedno vlastní číslo, a to 1. Nicméně, vektor  $e_2 = (0, 1)^T$  se může zobrazit na libovolně velký vektor, neboť  $Ae_2 = (K, 1)^T$ . Tedy vlastní čísla neomezují velikost obrazů.

U symetrických matic je to jiné. Buď  $A \in \mathbb{R}^{n \times n}$  symetrická a  $x \in \mathbb{R}^n$ ,  $\|x\|_2 = 1$ . Pak podle Rayleighovy–Ritzovy formule

$$\|Ax\|_2 = \sqrt{(Ax)^T (Ax)} = \sqrt{x^T A^2 x} \leq \sqrt{\lambda_1(A^2)} = \sqrt{\max\{|\lambda_1(A)|, |\lambda_n(A)|\}^2} \leq \rho(A).$$

Tudíž velikost obrazů je omezená, maximálně se zvětší  $\rho(A)$ -krát.

Věta je přímým důsledkem následující věty (nahlédněte!). Ta dává formuli pro libovolné vlastní číslo, tedy nejen to nejmenší a největší.

**Věta 2.6** (Courant–Fischer). *Nechť  $\lambda_1 \geq \dots \geq \lambda_n$  jsou vlastní čísla symetrické matice  $A \in \mathbb{R}^{n \times n}$ . Pak*

$$\lambda_k = \max_{V \in \mathbb{R}^n: \dim V = k} \min_{x \in V: \|x\|_2=1} x^T Ax \quad (2.1)$$

$$= \min_{V \in \mathbb{R}^n: \dim V = n-k+1} \max_{x \in V: \|x\|_2=1} x^T Ax. \quad (2.2)$$

*Důkaz.* 1) Rovnost (2.1) Buďte  $q_1, \dots, q_n$  ortonormální vlastní vektory odpovídající vlastním číslům  $\lambda_1, \dots, \lambda_n$ . Označme  $Q := (q_1 \mid \dots \mid q_n)$

Nerovnost „ $\leq$ “: Definujme  $U := \text{span}\{q_1, \dots, q_k\}$ . Pak každé  $x \in U$  se dá vyjádřit jako

$$x = \sum_{i=1}^k \alpha_i q_i + \sum_{i=k+1}^n 0 q_i = Q(\alpha_1, \dots, \alpha_n)^T.$$

Je-li  $\|x\|_2 = 1$ , pak z ortogonality  $Q$  je také  $1 = \|Q^T x\|_2 = \sum_{i=1}^k \alpha_i^2$ . Nyní

$$\begin{aligned} x^T Ax &= (\sum_{i=1}^k \alpha_i q_i)^T A (\sum_{j=1}^k \alpha_j q_j) = (\sum_{i=1}^k \alpha_i q_i)^T (\sum_{j=1}^k \alpha_j \lambda_j q_j) \\ &= \sum_{i,j=1}^k \alpha_i \alpha_j q_i^T q_j \lambda_j = \sum_{i=1}^k \alpha_i^2 \lambda_i \geq \sum_{i=1}^k \alpha_i^2 \lambda_k = \lambda_k. \end{aligned}$$

Tudíž

$$\lambda_k \leq \min_{x \in U: \|x\|_2=1} x^T A x \leq \max_{V \in \mathbb{R}^n: \dim V=k} \min_{x \in V: \|x\|_2=1} x^T A x.$$

Nerovnost „ $\geq$ “: Buď  $V$  libovolný podprostor  $\mathbb{R}^n$  dimenze  $k$ . Pak  $V \cap \text{span}\{q_k, \dots, q_n\}$  má dimenzi aspoň 1, tudíž v něm existuje vektor  $v \equiv x_V$  jednotkové délky. Ten lze vyjádřit jako  $x = \sum_{i=k}^n \alpha_i q_i$ . Opět  $\|x\|_2 = 1$  implikuje  $\sum_{i=k}^n \alpha_i^2 = 1$ . Nyní

$$x^T A x = \left(\sum_{i=k}^n \alpha_i q_i\right)^T A \left(\sum_{j=k}^n \alpha_j q_j\right) = \sum_{i,j=k}^n \alpha_i \alpha_j q_i^T q_j \lambda_j = \sum_{i=1}^k \alpha_i^2 \lambda_i \leq \sum_{i=k}^n \alpha_i^2 \lambda_k = \lambda_k.$$

Tudíž

$$\max_{V \in \mathbb{R}^n: \dim V=k} \min_{x \in V: \|x\|_2=1} x^T A x \leq \max_{V \in \mathbb{R}^n: \dim V=k} x_V^T A x_V \leq \lambda_k.$$

2) Rovnost (2.2). Upravme

$$\begin{aligned} \lambda_k(A) &= -\lambda_{n-k+1}(-A) = - \max_{V \in \mathbb{R}^n: \dim V=n-k+1} \min_{x \in V: \|x\|_2=1} -x^T A x \\ &= \min_{V \in \mathbb{R}^n: \dim V=n-k+1} \max_{x \in V: \|x\|_2=1} x^T A x. \end{aligned} \quad \square$$

Jedním z důsledků je následující Cauchyho proplétací vlastnost (Cauchy interlacing theorem), spojující vlastní čísla matice a matice po odstranění  $i$ -tého řádku a sloupce.

**Věta 2.7** (Cauchyho proplétací vlastnost). *Buď  $A \in \mathbb{R}^{n \times n}$  symetrická a nechť matice  $B$  vznikne z  $A$  vyškrtnutím  $i$ -tého řádku a sloupce. Označme  $\lambda_1 \geq \dots \geq \lambda_n$  vlastní čísla  $A$  a  $\mu_1 \geq \dots \geq \mu_{n-1}$  vlastní čísla  $B$ . Pak*

$$\lambda_1 \geq \mu_1 \geq \lambda_2 \geq \dots \geq \lambda_{n-1} \geq \mu_{n-1} \geq \lambda_n.$$

*Důkaz.* Bez újmy na obecnost předpokládejme, že  $i = n$ . Označme  $W := \text{span}\{e_1, \dots, e_{n-1}\}$ .

Nerovnost „ $\lambda_k \geq \mu_k$ “:

$$\lambda_k = \max_{V \in \mathbb{R}^n: \dim V=k} \min_{x \in V: \|x\|_2=1} x^T A x \geq \max_{V \in W: \dim V=k} \min_{x \in V: \|x\|_2=1} x^T A x = \mu_k.$$

Nerovnost „ $\mu_k \geq \lambda_{k+1}$ “: Buď  $V_{k+1}$  prostor, pro který se nabyde maximum z (2.1). Pak  $V_{k+1} \cap W$  má dimenzi aspoň  $k$  a tudíž

$$\lambda_{k+1} = \min_{x \in V_{k+1}: \|x\|_2=1} x^T A x \leq \max_{V \in W: \dim V=k} \min_{x \in V_{k+1} \cap W: \|x\|_2=1} x^T A x \leq \mu_k. \quad \square$$

## Cvičení

2.1. Buď  $A \in \mathbb{R}^{n \times n}$  symetrická matice a  $D \in \mathbb{R}^{n \times n}$  diagonální s kladnou diagonálou. Dokažte, že  $\lambda_i(A) < \lambda_i(A + D)$  pro každé  $i = 1, \dots, n$ .



# Kapitola 3

## SVD rozklad

### 3.1 Konstrukce a souvislost s vlastními čísly

SVD rozklad jednou z nejdůležitějších numerických technik. Zkratka je za *Singular value decomposition*, rozklad na singulární čísla. Byl objeven r. 1873 nezávisle řadou autorů jako byli např. Ital Eugenio Beltrami, Francouz Marie E. Camille Jordan, Angličan James Sylvester, Němec Erhard Schmidt nebo Švýcar Hermann Weyl. Nicméně k praktické použitelnosti jej dovedla až implementace od Gene Goluba a Williama Kahana.

**Věta 3.1** (SVD rozklad). *Bud'  $A \in \mathbb{R}^{m \times n}$ ,  $q := \min\{m, n\}$ . Pak existuje diagonální matice  $\Sigma \in \mathbb{R}^{m \times n}$  s prvky  $\sigma_{11} \geq \dots \geq \sigma_{qq} \geq 0$  a ortogonální matice  $U \in \mathbb{R}^{m \times m}$ ,  $V \in \mathbb{R}^{n \times n}$  tak, že  $A = U\Sigma V^T$ .*

Ještě než řekneme ideu důkazu, zavedeme několik pojmů a ukážeme několik souvislostí.

Číslům na diagonále  $\sigma_{11}, \dots, \sigma_{qq}$  matice  $\Sigma$  říkáme *singulární čísla* matice  $A$  a značíme je obvykle  $\sigma_1, \dots, \sigma_q$ . Zjevně počet kladných singulárních čísel je roven hodnotě matice, tedy  $r = \text{rank}(A)$ , kde  $\sigma_r > 0$  a  $\sigma_{r+1} = 0$ .

*Redukovaný* tvar SVD rozkladu vypadá následovně: Rozložme  $U = (U_1 \mid U_2)$ ,  $V = (V_1 \mid V_2)$  na prvních  $r$  sloupců a zbytek, a dále  $S := \text{diag}(\sigma_1, \dots, \sigma_r)$ . Pak

$$A = U\Sigma V^T = (U_1 \mid U_2) \begin{pmatrix} S & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} V_1^T \\ V_2^T \end{pmatrix} = U_1 S V_1^T.$$

Redukovaný SVD používá jen část informace z SVD rozkladu, ale tu podstatnou, ze které můžeme plný SVD rozklad zrekonstruovat (doplněním  $U$ ,  $V$  na ortogonální matice).

**Věta 3.2** (Vztah singulárních a vlastních čísel I.). *Bud'  $A \in \mathbb{R}^{m \times n}$ ,  $r = \text{rank}(A)$ , a necht'  $A^T A$  má vlastní čísla  $\lambda_1 \geq \dots \geq \lambda_n$ . Pak kladná singulární čísla matice  $A$  jsou  $\sigma_i = \sqrt{\lambda_i}$ ,  $i = 1, \dots, r$ .*

*Důkaz.* Necht'  $A = U\Sigma V^T$  je SVD rozklad  $A$ . Pak

$$A^T A = V\Sigma^T U^T U\Sigma V^T = V\Sigma^T \Sigma V^T = V \text{diag}(\sigma_1^2, \dots, \sigma_q^2, 0, \dots, 0) V^T,$$

což je spektrální rozklad pozitivně definitní matice  $A^T A$ . Tudíž  $\lambda_i = \sigma_i^2$ . □

**Příklad 3.3.** Bud'  $Q \in \mathbb{R}^{n \times n}$  ortogonální. Pak  $Q^T Q = I_n$  má vlastní čísla samé jedničky. Tedy ortogonální matice  $Q$  má singulární čísla také samé jedničky.

Toto pozorování platí i opačným směrem: pokud má matice  $A \in \mathbb{R}^{n \times n}$  všechna singulární čísla rovna jedné, pak je ortogonální. To se nahlédne snadno, neboť  $\Sigma = I_n$  a tedy  $A = U\Sigma V^T = U I_n V^T = UV^T$  díky tomu, že součin ortogonálních matic je opět ortogonální matice. □

**Příklad 3.4.** Bud'  $A \in \mathbb{R}^{n \times n}$  symetrická. Pak  $A^T A = A^2$  má vlastní čísla druhé mocniny vlastních čísel  $\lambda_1, \dots, \lambda_n$  matice  $A$ . Tudíž singulární čísla  $A$  jsou  $|\lambda_1|, \dots, |\lambda_n|$  po sestupném setřídění. □

*Důkaz věty 3.1.* Buď  $A^T A = V \Lambda V^T$  spektrální rozklad matice  $A^T A$ , a označme  $v_1, \dots, v_m$  sloupce matice  $V$ . Buď  $\text{rank}(A) = r$ . Pak

$$v_j^T A^T A v_i = v_j^T \lambda_i v_i = \begin{cases} \lambda_i & \text{pokud } i = j, \\ 0 & \text{pokud } i \neq j. \end{cases} \quad (3.1)$$

Definujme  $u_i := \frac{1}{\sqrt{\lambda_i}} A v_i$  pro  $i = 1, \dots, r$ . Pak díky (3.1) máme  $v_i \perp v_j$  a navíc

$$\|u_i\|_2^2 = \langle u_i, u_i \rangle = \left\langle \frac{1}{\sqrt{\lambda_i}} A v_i, \frac{1}{\sqrt{\lambda_i}} A v_i \right\rangle = \frac{1}{\lambda_i} v_i^T A^T A v_i = 1.$$

Tudíž  $u_1, \dots, u_r$  tvoří ortonormální systém. Nyní (3.1) lze přepsat na tvar

$$\begin{pmatrix} u_1^T \\ \vdots \\ u_r^T \end{pmatrix} A V = \begin{pmatrix} \sqrt{\lambda_1} & & 0 \\ & \ddots & \vdots \\ & & \sqrt{\lambda_r} & 0 \end{pmatrix}. \quad (3.2)$$

Rozšíříme systém na  $u_1, \dots, u_r, \dots, u_m$  ortonormální bázi  $\mathbb{R}^m$ . Protože  $u_i \in \mathcal{S}(A)$  a  $\text{rank}(A) = r$ , tak doplněné vektory patří do  $\mathcal{S}(A)^\perp$  a proto  $u_i^T A = 0$  pro  $i = r + 1, \dots, m$ . Tudíž rovnost (3.2) můžeme rozšířit na

$$\begin{pmatrix} u_1^T \\ \vdots \\ u_m^T \end{pmatrix} A V = \begin{pmatrix} \sqrt{\lambda_1} & & 0 \\ & \ddots & \vdots \\ & & \sqrt{\lambda_r} & 0 \\ 0 & \dots & 0 & 0 \end{pmatrix}.$$

Označíme-li tuto rovnost jako  $U^T A V = \Sigma$ , dostaneme úpravou  $A = U \Sigma V^T$ .  $\square$

**Poznámka 3.5.** Důkaz věty prozradil navíc, že matice  $V$  je ortogonální maticí ze spektrálního rozkladu matice  $A^T A$ . Podobně, matice  $U$  je ortogonální maticí ze spektrálního rozkladu  $AA^T$ :

$$AA^T = U \Sigma V^T V \Sigma^T U^T = U \Sigma \Sigma^T U^T = U \text{diag}(\sigma_1^2, \dots, \sigma_q^2, 0, \dots, 0) U^T.$$

Bohužel, spektrální rozklady matic  $A^T A$  a  $AA^T$  nemůžeme použít ke konstrukci SVD rozkladu, protože nejsou jednoznačné. Použít můžeme jen jeden a druhý dopočítat trochu jinak.

Důkaz věty také naznačil, jak se dá SVD rozklad spočítat.

**Příklad 3.6.** Buď  $A = (1 \ 1 \ 1 \ 1)$ . Spektrální rozklad matice  $A^T A$  je

$$A^T A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1/2 & \dots \\ 1/2 & \dots \\ 1/2 & \dots \\ 1/2 & \dots \end{pmatrix} \begin{pmatrix} 4 & & \\ & 0 & \\ & & 0 \end{pmatrix} \begin{pmatrix} 1/2 & 1/2 & 1/2 & 1/2 \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

Nyní  $u_1 = \frac{1}{\sqrt{\lambda_1}} A v_1 = 1$ . Tudíž plný a redukovaný SVD rozklad je

$$A = (1)(2 \ 0 \ 0 \ 0) \begin{pmatrix} 1/2 & 1/2 & 1/2 & 1/2 \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix} = (1)(2)(1/2 \ 1/2 \ 1/2 \ 1/2). \quad \square$$

Kromě věty 3.2 existuje ještě jiný vztah mezi singulárními a vlastními čísly.

**Věta 3.7** (Vztah singulárních a vlastních čísel II.). *Matice  $A \in \mathbb{R}^{m \times n}$  má singulární čísla  $\sigma_1, \dots, \sigma_r$  právě tehdy, když nenulová vlastní čísla matice*

$$\begin{pmatrix} 0 & A \\ A^T & 0 \end{pmatrix} \quad (3.3)$$

*jsou  $\sigma_1, \dots, \sigma_r, -\sigma_r, \dots, -\sigma_1$ .*

*Důkaz.* Implikace „ $\Leftarrow$ “. Buď  $\lambda$  vlastní číslo a  $(y^T, x^T)^T$  odpovídající vlastní vektor  $\begin{pmatrix} 0 & A \\ A^T & 0 \end{pmatrix}$ , tj.

$$\begin{pmatrix} 0 & A \\ A^T & 0 \end{pmatrix} \begin{pmatrix} y \\ x \end{pmatrix} = \lambda \begin{pmatrix} y \\ x \end{pmatrix}.$$

To se rozpadá na dvě rovnosti  $A^T y = \lambda x$ ,  $Ax = \lambda y$ . Tudíž  $A^T A x = A^T (Ax) = A^T \lambda y = \lambda^2 x$ . Podle věty 3.2 je  $\sigma := \sqrt{\lambda^2} = |\lambda|$  singulární číslo  $A$ .

Implikace „ $\Rightarrow$ “. Buď  $\sigma$  singulární číslo  $A$ . Pak podle věty 3.2 je  $A^T A x = \sigma^2 x$  pro jisté  $x \neq 0$ . Označme  $y := \frac{1}{\sigma} A x$ . Pak  $Ax = \sigma y$ ,  $A^T y = \sigma x$ , neboli

$$\begin{pmatrix} 0 & A \\ A^T & 0 \end{pmatrix} \begin{pmatrix} y \\ x \end{pmatrix} = \sigma \begin{pmatrix} y \\ x \end{pmatrix}.$$

Proto je  $\sigma$  vlastním číslem  $\begin{pmatrix} 0 & A \\ A^T & 0 \end{pmatrix}$ . Analogicky označme  $y' := -\frac{1}{\sigma} A x$ . Pak  $Ax = -\sigma y'$ ,  $A^T y' = -\sigma x$ , neboli

$$\begin{pmatrix} 0 & A \\ A^T & 0 \end{pmatrix} \begin{pmatrix} y' \\ x \end{pmatrix} = -\sigma \begin{pmatrix} y' \\ x \end{pmatrix}.$$

Proto je i  $-\sigma$  vlastním číslem  $\begin{pmatrix} 0 & A \\ A^T & 0 \end{pmatrix}$ . □

Jako důsledek dostaneme proplétací vlastnost pro singulární čísla.

**Důsledek 3.8** (Proplétací vlastnost pro singulární čísla). *Buď  $A \in \mathbb{R}^{m \times n}$  hodnosti  $r$  a necht' matice  $B$  vznikne z  $A$  odstraněním jednoho řádku či sloupce. Pak*

$$\sigma_1(A) \geq \sigma_1(B) \geq \sigma_2(A) \geq \dots \geq \sigma_{r-1}(A) \geq \sigma_{r-1}(B) \geq \lambda_r(A).$$

*Důkaz.* Plyne přímočaře použitím Cauchyho proplétací vlastnosti z věty 2.7 na matici  $\begin{pmatrix} 0 & A \\ A^T & 0 \end{pmatrix}$ . Pokud  $B$  vznikne z  $A$  odstraněním  $i$ -tého řádku, tak z matice  $\begin{pmatrix} 0 & A \\ A^T & 0 \end{pmatrix}$  odstraníme  $i$ -tý řádek a sloupec. □

Jiným důsledkem věty 3.7 je min-maxová reprezentace singulárních čísel. Pro jednoduchost uvádíme jen adaptaci Rayleighovy–Ritzovy formule, namísto obecnější Courantovy–Fischerovy.

**Důsledek 3.9** (Min-maxová reprezentace singulárních čísel). *Buď  $A \in \mathbb{R}^{m \times n}$ . Pak*

$$\sigma_1(A) = \max_{\|x\|_2 = \|y\|_2 = 1} x^T A y. \quad (3.4)$$

*Důkaz.* Aplikací Rayleighovy–Ritzovy věty 2.4 na matici  $\begin{pmatrix} 0 & A \\ A^T & 0 \end{pmatrix}$  dostaneme

$$\begin{aligned} \sigma_1(A) &= \max_{\|(x,y)\|_2=1} (x^T \ y^T) \begin{pmatrix} 0 & A \\ A^T & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 2 \max_{\|(x,y)\|_2=1} x^T A y, \\ &\geq 2 \max_{\|(x,y)\|_2=1, \|x\|_2=\|y\|_2} x^T A y = \max_{\|x\|_2=\|y\|_2=1} x^T A y. \end{aligned}$$

Na druhou stranu, necht'  $A = U \Sigma V^T$  je SVD rozklad a necht'  $u$  je první sloupec matice  $U$  a  $v$  je první sloupec matice  $V$ . Pak  $u^T A v = u^T U \Sigma V^T v = e_1^T \Sigma e_1 = \sigma_1(A)$ . Tedy rovnost v (3.4) se nabyde pro  $x := u$  a  $y := v$ . □

## 3.2 Aplikace SVD rozkladu

**Poznámka 3.10.** Několik praktických aplikací (v angličtině):

- A pattern analysis of the second Rehnquist U.S. Supreme Court:  
<http://www.pnas.org/content/100/13/7432.abstract>
- Blockbuster Algorithm:  
<https://www2.bc.edu/~baglivo/MT210/SVDBlockbuster.pdf>

## SVD a ortogonalizace

SVD rozklad lze použít k nalezení ortonormální báze (nejen) sloupcového prostoru  $\mathcal{S}(A)$ . Na rozdíl od dosavadních přístupů nemusíme předpokládat lineární nezávislost sloupců matice  $A$ .

**Věta 3.11.** *Nechť  $A = U\Sigma V^T = U_1 S V_1^T$  je SVD rozklad matice  $A \in \mathbb{R}^{m \times n}$ . Pak*

- (1) Sloupce  $U_1$  tvoří ortonormální bázi prostoru  $\mathcal{S}(A)$ .
- (2) Sloupce  $U_2$  tvoří ortonormální bázi prostoru  $\text{Ker}(A^T)$ .
- (3) Sloupce  $V_1$  tvoří ortonormální bázi prostoru  $\mathcal{R}(A)$ .
- (4) Sloupce  $V_2$  tvoří ortonormální bázi prostoru  $\text{Ker}(A)$ .

*Důkaz.*

- (1) Redukovaný SVD rozklad je  $A = U_1 S V_1^T$ . Přenásobením  $V_1$  zprava dostaneme  $AV_1 = U_1 S$ . Nyní,  $\mathcal{S}(A) \ni \mathcal{S}(AV_1) = \mathcal{S}(U_1 S) = \mathcal{S}(U_1)$  díky regularitě matice  $S$ . Protože  $\text{rank}(A) = \text{rank}(U_1)$ , máme rovnost  $\mathcal{S}(A) = \mathcal{S}(U_1)$ .
- (2) Analogicky jako (4).
- (3) Plyne z předchozího díky  $\mathcal{R}(A) = \mathcal{S}(A^T)$ .
- (4) Z transpozice  $A^T = V_1 S U_1^T$  dostáváme redukovaný SVD rozklad matice  $A^T$ . Tedy sloupce  $V_1$  tvoří ortonormální bázi prostoru  $\mathcal{S}(A^T) = \mathcal{R}(A) = \text{Ker}(A)^\perp$ . Proto sloupce  $V_2$ , které doplňují sloupce  $V_1$  na ortonormální bázi  $\mathbb{R}^n$ , představují ortonormální bázi  $\text{Ker}(A)$ .  $\square$

## SVD a projekce do podprostoru

Pomocí SVD rozkladu můžeme snadno vyjádřit matici projekce do sloupcového (a řádkového) prostoru dané matice. Dokonce k tomu nepotřebujeme předpoklad na lineární nezávislost sloupců matice.

**Věta 3.12.** *Nechť  $A = U\Sigma V^T = U_1 S V_1^T$  je SVD rozklad matice  $A \in \mathbb{R}^{m \times n}$ . Pak matice projekce do*

- (1) sloupcového prostoru  $\mathcal{S}(A)$  je  $U_1 U_1^T$ ,
- (2) řádkového prostoru  $\mathcal{R}(A)$  je  $V_1 V_1^T$ ,

*Důkaz.*

- (1) Z věty 3.11 je  $\mathcal{S}(A) = \mathcal{S}(U_1)$ . Sloupce  $U_1$  jsou lineárně nezávislé, a proto matice projekce má dle známé věty tvar  $U_1(U_1^T U_1)^{-1} U_1^T = U_1(I_r)^{-1} U_1^T = U_1 U_1^T$ .
- (2) Plyne z předchozího díky  $\mathcal{R}(A) = \mathcal{S}(A^T)$ .  $\square$

Podobným způsobem lze odvodit i vzoreček pro přibližné řešení soustavy  $Ax = b$  metodou nejmenších čtverců. Nicméně, později ve větě 4.7 ukážeme silnější výsledek.

## SVD a geometrie lineárního zobrazení

Bud'  $A \in \mathbb{R}^{n \times n}$  regulární matice a studujme obraz jednotkové koule při zobrazení  $x \mapsto Ax$ . Z SVD rozkladu  $A = U\Sigma V^T$  plyne, že lineární zobrazení lze rozložit na složení tří základních zobrazení: ortogonální zobrazení s maticí  $V^T$ , škálování podle  $\Sigma$  a ortogonální zobrazení s maticí  $U$ . Konkrétně, zobrazení s maticí  $V^T$  zobrazí kouli na sebe sama,  $\Sigma$  ji zdeformuje na elipsoid a  $U$  ji otočí/převrátí. Tedy výsledkem bude elipsoid se středem v počátku, poloosy jsou ve směrech sloupců  $U$  a délky mají velikost  $\sigma_1, \dots, \sigma_n$ .

Hodnota  $\frac{\sigma_1}{\sigma_n} \geq 1$  se nazývá *míra deformace* a kvantitativně udává, jak moc zobrazení deformuje geometrické útvary. Je-li hodnota rovna 1, elipsoid bude mít tvar koule, a naopak čím větší bude hodnota, tím protáhlejší bude elipsoid. Význam této hodnoty je ale nejenom geometrický. V numerické matematice se podíl  $\frac{\sigma_1}{\sigma_n}$  nazývá číslo podmíněnosti a čím je větší, tím hůře podmíněná je matice  $A$  ve smyslu, že vykazuje špatné numerické vlastnosti – zaokrouhlování v počítačové aritmetice s pohyblivou řádkovou čárkou způsobuje chyby. Více budeme probírat v sekci 6.1.



## SVD a numerický rank

Hodnota matice  $A$  je rovna počtu (kladných) singulárních čísel. Nicméně, pro výpočetní účely se hodně malé kladné číslo považuje za praktickou nulu. Buď  $\varepsilon > 0$ , pak *numerický rank* matice  $A$  je  $\max\{s; \sigma_s > \varepsilon\}$ , tedy počet singulárních čísel větších než  $\varepsilon$ , ostatní se berou za nulová. Např. **Matlab / Octave** bere  $\varepsilon := \max\{m, n\} \cdot \sigma_1 \cdot \text{eps}$ , kde  $\text{eps} \approx 2 \cdot 10^{-16}$  je přesnost počítačové aritmetiky.

## SVD a low-rank aproximace

Buď  $A \in \mathbb{R}^{m \times n}$  a  $A = U\Sigma V^T$  její SVD rozklad. Jestliže ponecháme  $k$  největších singulárních čísel a ostatní vynulujeme  $\sigma_{k+1} := 0, \dots, \sigma_r := 0$ , tak dostaneme matici

$$A' = U \text{diag}(\sigma_1, \dots, \sigma_k, 0, \dots, 0) V^T$$

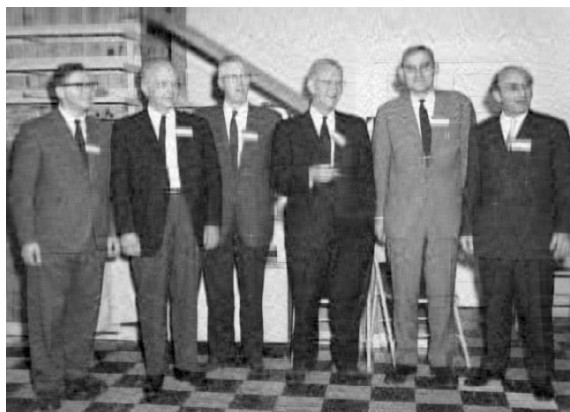
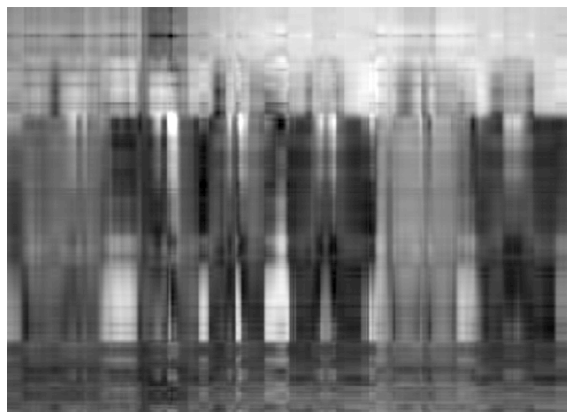
hodnosti  $k$ , která dobře aproximuje  $A$ . Navíc tato aproximace je v jistém smyslu nejlepší možná. To jest, v určité normě (viz věta 5.26) je ze všech matic hodnosti  $k$  právě  $A'$  nejblíže matici  $A$ . Low-rank aproximaci využijeme v následujícím:

## SVD a komprese dat

Předpokládejme, že matice  $A \in \mathbb{R}^{m \times n}$  reprezentuje data, které chceme zkomprimovat. Pokud  $\text{rank}(A) = r$ , tak pro redukovaný SVD rozklad  $A = U_1 S V_1^T$  si potřebujeme zapamatovat  $mr + r + nr = (m + n + 1)r$  hodnot. Při low-rank aproximaci  $A \approx U \text{diag}(\sigma_1, \dots, \sigma_k, 0, \dots, 0) V^T$  si stačí pamatovat jen  $(m + n + 1)k$  hodnot. Tedy kompresní poměr je  $k : r$ . Čím menší  $k$ , tím menší objem dat si stačí pamatovat. Ale na druhou stranu, menší  $k$  značí horší aproximaci.

**Příklad 3.13.** Zmíněný postup ilustrujeme na kompresi obrázku. Předpokládáme, že matice  $A \in \mathbb{R}^{m \times n}$  reprezentuje obrázek, ve kterém pixel na pozici  $(i, j)$  má barvu s číslem  $a_{ij}$ . Následující obrázky ilustrují kompresi pro různou volbu  $k$ . Pro  $k = 480$  máme originální obrázek, pro 150 asi třetinovou kompresi bez znatelné újmy na kvalitě obrázku, při  $k = 50$  už dochází k zrnění a při  $k = 5$  je obrázek značně rozmazán (ale na to, že máme zhruba 1% původního objemu dat, je výsledek stále slušný).

originál ( $k = 480$ ) $k = 150$

 $k = 50$  $k = 5$ 

Obrázek představuje foto z konference o numerické algebře v Gatlinburgu z r. 1964, a zaznamenává největší numerické matematiky své doby, zleva: James H. Wilkinson, Wallace Givens, George Forsythe, Alston Householder, Peter Henrici, a Fritz Bauer. Obrázek se skládá z  $480 \times 640$  pixelů, SVD rozklad trval cca 5 sec (11.5.2010). Zdrojový kód pro Matlab / Octave:

```
load gatlin,
[X,S,Y] = svd(X);
figure(2), clf,
k = 150;
Xk = X(:,1:k)*S(1:k,1:k)*Y(:,1:k)';
image(Xk),
colormap(map),
axis equal, axis off,
```

□

## SVD a míra regularity

Hodnota determinantu se jako míra regularity matice moc nehodí. Zato singulární čísla jsou pro to jako stvořená. Buď  $A \in \mathbb{R}^{n \times n}$ . Pak  $\sigma_n$  udává vzdálenost (v jisté normě, viz věta 5.26) k nejbližší singulární matici, podrobněji viz [Rohn, 2004]. Takže je to v souladu s tím, co bychom si pod takovou mírou představovali. Ortogonální matice mají míru 1, naproti tomu Hilbertovy matice mají malou míru regularity, tj. jsou téměř singulární:

$n$	$\sigma_n(H_n)$
3	$\approx 0.0027$
5	$\approx 10^{-6}$
10	$\approx 10^{-13}$
15	$\approx 10^{-18}$

## Cvičení

3.1. Určete singulární čísla matic:

- (a)  $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ ,
- (b)  $\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$ ,
- (c)  $a = (a_1, \dots, a_n)^T$ ,
- (d)  $a = (a_1, \dots, a_n)$ ,
- (e) pozitivně definitní matice s vlastními čísly  $\lambda_1, \dots, \lambda_n$ .

$$(f) A = \begin{pmatrix} 0 & a_1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \ddots & \ddots & \ddots & a_{n-1} \\ a_n & 0 & \dots & 0 & 0 \end{pmatrix}.$$

3.2. Určete SVD rozklad matic

- (a)  $A = 0_{m,n}$ ,
- (b)  $A = \text{diag}(a_1, \dots, a_n)$ ,  $a_1 \geq \dots \geq a_n \geq 0$ .
- (c)  $A = \text{diag}(a_1, \dots, a_n)$ .

3.3. Buď  $A = xy^T$ , kde  $x, y \in \mathbb{R}^n$ . Určete  $\sigma_1(A)$ ,  $\rho(A)$  a porovnejte je. Kdy se rovnají?

3.4. Ukažte, že matice  $A \in \mathbb{R}^{n \times n}$  má nulové singulární číslo právě tehdy, když má nulové vlastní číslo.

3.5. Nechť všechna singulární čísla matice  $A \in \mathbb{R}^{n \times n}$  jsou stejná. Ukažte, že  $A$  je násobek ortogonální matice.

3.6. Nechť regulární matice  $A \in \mathbb{R}^{n \times n}$  má singulární čísla  $\sigma_1 \geq \dots \geq \sigma_n$ . Najděte singulární čísla matic  $A^T$ ,  $A^{-1}$  a  $\text{adj}(A)$ .

3.7. Víme, že matice  $AB$  a  $BA$  mají stejná (nenulová) vlastní čísla. Co singulární čísla?

3.8. *Polární rozklad* matice  $A \in \mathbb{R}^{n \times n}$  je rozklad  $A = PQ$ , kde  $P$  je pozitivně semidefinitní a  $Q$  ortogonální.

- (a) Ukažte, že každá matice  $A \in \mathbb{R}^{n \times n}$  má polární rozklad.
- (b) Ukažte, že matice  $P$  je jednoznačně určená tak, že vyjádříte  $P = \sqrt{AA^T}$ .
- (c) Zobecněte tvrzení na obdélníkové matice: Každá matice  $A \in \mathbb{R}^{m \times n}$ ,  $m \geq n$ , má rozklad  $A = PQ$ , kde  $P$  je pozitivně semidefinitní a  $Q$  má ortogonální sloupce.
- (d) Najděte polární rozklad matice  $a \in \mathbb{R}^n$ .
- (e) Ukažte, že polární rozklad a SVD rozklad jsou ekvivalentní v tom smyslu, že jeden lze snadno odvodit z druhého.

\*3.9. Dokažte druhou část důsledku 3.9.



# Kapitola 4

## Pseudoinverzní matice

Přirozená snaha zobecnit pojem inverzní matice tak, aby tato zobecněná inverze (=pseudoinverze) existovala i pro singulární či dokonce obdélníkové matice, vedla k několika konceptům. Nejznámější je Mooreova–Penroseova pseudoinverze, která existuje pro každou matici a vyskytuje se u ortogonálních problémů jako je například metoda nejmenších čtverců. Drazinova pseudoinverze, která existuje jen pro čtvercové matice, se spíše vyskytuje u neortogonálních problémů.

Každopádně je dobré mít na paměti, že pseudoinverze se málokdy počítá explicitně. Spíš se používá pro stanovení a explicitní vyjádření určitých vlastností. Je to podobné jako pro klasickou inverzi: soustavu  $Ax = b$  neřešíme podle vzorce  $x = A^{-1}b$ , nicméně toto explicitní vyjádření řešení je velmi užitečné.

Podrobně se pseudoinverzemi zabývá například kniha Ben-Israel and Greville [2003].

### 4.1 Mooreova–Penroseova pseudoinverze

Nejčastější pseudoinverzí je tzv. Mooreova–Penroseova pseudoinverze<sup>1)</sup>, která spočívá na SVD rozkladu.

**Definice 4.1** (Mooreova–Penroseova pseudoinverze). Buď  $A \in \mathbb{R}^{m \times n}$  matice s redukovaným SVD rozkladem  $A = U_1 S V_1^T$ . Je-li  $A \neq 0$ , pak její *pseudoinverze* je  $A^\dagger = V_1 S^{-1} U_1^T \in \mathbb{R}^{n \times m}$ . Pro  $A = 0$  definujeme pseudoinverzi předpisem  $A^\dagger = A^T$ .

**Příklad 4.2.** Pseudoinverze nenulového vektoru  $a \in \mathbb{R}^n$  je  $a^\dagger = \frac{1}{a^T a} a^T$ , speciálně např.  $(1, 1, 1, 1)^\dagger = \frac{1}{4}(1, 1, 1, 1)^T$ .  $\square$

**Věta 4.3** (Vlastnosti pseudoinverze). Buď  $A \in \mathbb{R}^{m \times n}$ , pak

- (1) Je-li  $A$  regulární, tak  $A^{-1} = A^\dagger$ ,
- (2)  $(A^\dagger)^\dagger = A$ ,
- (3)  $(A^T)^\dagger = (A^\dagger)^T$ ,
- (4)  $A = AA^\dagger A$ ,
- (5)  $A^\dagger = A^\dagger AA^\dagger$ ,
- (6)  $AA^\dagger$  je symetrická,
- (7)  $A^\dagger A$  je symetrická,
- (8) má-li  $A$  lineárně nezávislé sloupce, pak  $A^\dagger = (A^T A)^{-1} A^T$ ,
- (9) má-li  $A$  lineárně nezávislé řádky, pak  $A^\dagger = A^T (AA^T)^{-1}$ .

*Důkaz.* Vlastnosti se dokážou jednoduše z definice. Pro ilustraci ukážeme jen dvě vlastnosti, zbytek necháváme čtenáři.

$$(4) \text{ Z definice } AA^\dagger A = U_1 S V_1^T V_1 S^{-1} U_1^T U_1 S V_1^T = U_1 S S^{-1} S V_1^T = U_1 S V_1^T = A.$$

<sup>1)</sup>Nezávisle ji objevili americký matematik Eliakim Hastings Moore r. 1920 v řeči ortogonálních projekcí a anglický fyzik, slavný popularizátor, Roger Penrose r. 1955 jako matici splňující vlastnosti (4)–(7) z věty 4.3.

(8) Z předpokladu je  $V_1$  čtvercová, tedy ortogonální. Pak

$$(A^T A)^{-1} = (V_1 S U_1^T U_1 S V_1^T)^{-1} = (V_1 S^2 V_1^T)^{-1} = V_1 S^{-2} V_1^T,$$

$$\text{z čehož } (A^T A)^{-1} A^T = V_1 S^{-2} V_1^T V_1 S U_1^T = V_1 S^{-1} U_1^T = A^\dagger. \quad \square$$

První vlastnost říká, že se skutečně jedná o zobecnění klasické inverze. Vlastnosti (4)–(7) jsou zajímavé v tom, že dávají alternativní definici pseudoinverze; ta se totiž ekvivalentně dá definovat jako matice, která splňuje podmínky (4)–(7), a taková matice kupodivu existuje vždy právě jedna.

Poznamenejme, že některé vlastnosti, u kterých bychom očekávali že platí, tak obecně platit nemusí. Např. obecně  $AA^\dagger \neq A^\dagger A$  a  $(AB)^\dagger \neq B^\dagger A^\dagger$ .

Pomocí pseudoinverze elegantně vyjádříme matice projekce do maticových prostorů.

**Věta 4.4.** *Bud'  $A \in \mathbb{R}^{m \times n}$ . Pak matice projekce do*

- (1) *sloupcového prostoru  $\mathcal{S}(A)$  je  $AA^\dagger$ ,*
- (2) *řádkového prostoru  $\mathcal{R}(A)$  je  $A^\dagger A$ ,*
- (3) *jádra  $\text{Ker}(A)$  je  $I_n - A^\dagger A$ .*

*Důkaz.*

- (1) S použitím redukovaného SVD rozkladu  $A = U_1 S V_1^T$  upravme

$$AA^\dagger = U_1 S V_1^T V_1 S^{-1} U_1^T = U_1 U_1^T.$$

Podle věty 3.12 je to hledaná matice projekce  $U_1 U_1^T$ .

- (2) Analogicky jako v předchozím je  $A^\dagger A = V_1 V_1^T$ , což je matice projekce do  $\mathcal{R}(A)$ .
- (3) Plyne z vlastnosti  $\text{Ker}(A) = \mathcal{R}(A)^\perp$ . □

Zajímavá je interpretace pseudoinverze z hlediska lineárních zobrazení.

**Věta 4.5** (Pseudoinverzní matice a lineární zobrazení). *Uvažujme lineární zobrazení  $f(x) = Ax$ , kde  $A \in \mathbb{R}^{m \times n}$ .*

- (1) *Pokud definiční obor  $f(x)$  omezíme pouze na prostor  $\mathcal{R}(A)$ , tak dostaneme isomorfismus mezi  $\mathcal{R}(A)$  a  $f(\mathbb{R}^n)$ .*
- (2) *Inverzní zobrazení k tomuto isomorfismu má tvar  $y \mapsto A^\dagger y$ .*

*Důkaz.*

- (1) Zobrazení s omezeným definičním oborem je „na“, protože

$$f(\mathbb{R}^n) = \mathcal{S}(A) = \mathcal{R}(A^T) = \mathcal{R}(AA^T) = \mathcal{S}(AA^T) = \{Ay; y \in \mathcal{R}(A)\} = f(\mathcal{R}(A)).$$

Jelikož prostory  $f(\mathbb{R}^n)$  a  $\mathcal{R}(A)$  mají stejnou dimenzi, musí být zobrazení isomorfismem.

- (2) Podle věty 4.4(2) se každý vektor  $x \in \mathcal{R}(A)$  při zobrazení  $x \mapsto A^\dagger Ax$  zobrazí na  $A^\dagger Ax = x$ . Tudíž  $A^\dagger$  je matice inverzního zobrazení k zobrazení  $x \mapsto Ax$ . □

Nejvýznačnější vlastnost pseudoinverze spočívá v popisu množiny řešení řešitelných soustav a množiny přibližných řešení metodou nejmenších čtverců neřešitelných soustav. V obou případech je  $A^\dagger b$  v jistém smyslu význačné řešení.

**Věta 4.6** (Pseudoinverzní matice a řešení soustav rovnic). *Bud'  $A \in \mathbb{R}^{m \times n}$ ,  $b \in \mathbb{R}^m$  a  $X$  množina řešení soustavy  $Ax = b$ . Je-li  $X \neq \emptyset$ , pak*

$$X = A^\dagger b + \text{Ker}(A).$$

*kde*

$$\text{Ker}(A) = \mathcal{S}(I_n - A^\dagger A).$$

*Navíc, ze všech vektorů z množiny  $X$  má  $A^\dagger b$  nejmenší eukleidovskou normu, a je to jediné řešení s touto vlastností.*

*Důkaz.* „=“ Bud'  $x \in X$ , tj.  $Ax = b$ . Potom  $AA^\dagger b = AA^\dagger Ax = Ax = b$ , tedy  $A^\dagger b \in X$ . Víme, že  $X = x_0 + \text{Ker}(A)$ , kde  $x_0$  je libovolné řešení. Podle věty 4.4(3) je  $\text{Ker}(A) = \mathcal{S}(I_n - A^\dagger A)$  a za  $x_0$  můžeme volit  $A^\dagger b$ .

„Norma.“ Podle věty 4.5(2) je  $A^\dagger b \in \mathcal{R}(A)$ , a dále platí  $\mathcal{R}(A)^\perp = \text{Ker}(A)$ . Nyní podle Pythagorovy věty pro každé  $y \in \mathcal{S}(I_n - A^\dagger A)$  platí

$$\|A^\dagger b + y\|_2^2 = \|A^\dagger b\|_2^2 + \|y\|_2^2 \geq \|A^\dagger b\|_2^2.$$

Tedy  $A^\dagger b$  nejmenší eukleidovskou normu. Každý jiný vektor z  $X$  má normu větší, protože  $y \neq 0$  implikuje  $\|y\|_2 > 0$ .  $\square$

**Věta 4.7** (Pseudoinverzní matice a metoda nejmenších čtverců). *Bud'  $A \in \mathbb{R}^{m \times n}$ ,  $b \in \mathbb{R}^m$  a  $X$  množina přibližných řešení soustavy  $Ax = b$  metodou nejmenších čtverců. Pak*

$$X = A^\dagger b + \text{Ker}(A).$$

*Navíc, ze všech vektorů z množiny  $X$  má  $A^\dagger b$  nejmenší eukleidovskou normu, a je to jediné řešení s touto vlastností.*

*Důkaz.* Množina přibližných řešení soustavy  $Ax = b$  metodou nejmenších čtverců je popsána soustavou  $A^T Ax = A^T b$ . Protože  $X \neq \emptyset$ , podle věty 4.6 máme

$$X = (A^T A)^\dagger (A^T b) + \text{Ker}(A^T A).$$

Jelikož  $(A^T A)^\dagger A^T = A^\dagger$  a  $\text{Ker}(A^T A) = \text{Ker}(A)$ , má množina  $X$  požadovaný popis a  $A^\dagger b$  požadovanou vlastnost.  $\square$

Předchozí dvě věty tedy mj. říkají, že  $A^\dagger b$  je význačný vektor. V případě, že soustava  $Ax = b$  má řešení, pak je jejím řešením s minimální normou. A v případě, že soustava  $Ax = b$  nemá řešení, pak je jejím přibližným řešením (opět s minimální normou) metodou nejmenších čtverců. Navíc není zapotřebí předpokladu na lineární nezávislost sloupců matice  $A$ .

## 4.2 Drazinova pseudoinverze

Druhý typ pseudoinverze, který zmíníme, je *Drazinova pseudoinverze*<sup>2)</sup>. Existuje pouze pro čtvercové matice.

**Definice 4.8** (Drazinova pseudoinverze). Bud'  $A \in \mathbb{R}^{n \times n}$  a necht'  $A = SJS^{-1}$ , kde  $J$  je Jordanova normální forma matice  $A$ . Necht'  $J$  je tvaru  $J = \begin{pmatrix} C & 0 \\ 0 & N \end{pmatrix}$ , kde  $C$  je regulární a  $N$  obsahuje Jordanovy buňky pro nulová vlastní čísla. Pak Drazinova pseudoinverze je  $A^D = S \begin{pmatrix} C^{-1} & 0 \\ 0 & 0 \end{pmatrix} S^{-1}$ .

**Příklad 4.9.** Bud'  $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ . Pak  $A^\dagger = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ , ale  $A^D = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ . Tudíž Mooreova–Penroseova a Drazinova pseudoinverze obecně nejsou shodné.  $\square$

**Věta 4.10** (Vlastnosti Drazinovy pseudoinverze). *Bud'  $A \in \mathbb{R}^{m \times n}$ , pak*

- (1) *Je-li  $A$  regulární, tak  $A^{-1} = A^D$ ,*
- (2)  *$A^D A A^D = A^D$ ,*
- (3)  *$A A^D = A^D A$ ,*
- (4)  *$A^{k+1} A^D = A^k$ , kde  $k$  je nejmenší takové, že  $\text{rank}(A^{k+1}) = \text{rank}(A^k)$  (tzv. index).*

*Důkaz.* Vlastnosti se dokážou jednoduše z definice. Pro ilustraci ukážeme jen dvě vlastnosti, zbytek necháváme čtenáři.

(1) Je-li  $A$  regulární, tak  $A = SCS^{-1}$ . Proto  $A^{-1} = SC^{-1}S^{-1} = A^D$ .

(2)  $A^D A A^D = S \begin{pmatrix} C^{-1} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} C & 0 \\ 0 & N \end{pmatrix} \begin{pmatrix} C^{-1} & 0 \\ 0 & 0 \end{pmatrix} S^{-1} = S \begin{pmatrix} C^{-1} C C^{-1} & 0 \\ 0 & 0 \end{pmatrix} S^{-1} = S \begin{pmatrix} C^{-1} & 0 \\ 0 & 0 \end{pmatrix} S^{-1} = A^D$ .  $\square$

<sup>2)</sup>Objevil ji americký matematik Michael P. Drazin, viz Drazin [1958].

Drazin původně definoval pseudoinverzi jako jednoznačnou matici, která splňuje vlastnosti (2)–(4). Tedy podobně jako Mooreova–Penroseova pseudoinverze se dá Drazinova pseudoinverze zavést výčtem vlastností, které má splňovat.

S Drazinovou pseudoinverzí se můžeme setkat např. při řešení lineárních diferenciálních rovnic, ale objevuje se i jako tzv. grupová inverze.

**Příklad 4.11** (Maticová grupa). Je dobře známo, že regulární matice z  $\mathbb{R}^{n \times n}$  s násobením tvoří grupu a inverzní prvek odpovídá klasické maticové inverzi. V prostoru  $\mathbb{R}^{n \times n}$  můžeme najít ale i jiné grupy s operací násobení.

Bud'  $S \in \mathbb{R}^{n \times n}$  pevná a regulární a buď  $k < n$ . Definujme množinu  $\mathcal{G}$  všech matic z  $\mathbb{R}^{n \times n}$  tvaru  $S \begin{pmatrix} C & 0 \\ 0 & 0 \end{pmatrix} S^{-1}$ , kde  $C \in \mathbb{R}^{k \times k}$  je regulární. Tedy všechny matice v  $\mathcal{G}$  jsou singulární. Není těžké nahlédnout, že  $\mathcal{G}$  tvoří grupu s násobením. Neutrální prvek je  $S \begin{pmatrix} I_k & 0 \\ 0 & 0 \end{pmatrix} S^{-1}$  a inverzní prvek k  $A = S \begin{pmatrix} C & 0 \\ 0 & 0 \end{pmatrix} S^{-1}$  je  $S \begin{pmatrix} C^{-1} & 0 \\ 0 & 0 \end{pmatrix} S^{-1} = A^D$ . Tedy inverzní prvky přesně odpovídají Drazinově pseudoinverzi.  $\square$

## Cvičení

4.1. Určete psudoinverzi matic

(a)  $A = \text{diag}(a_1, \dots, a_n)$ ,  $a_1 \geq \dots \geq a_n \geq 0$ .

(b)  $A = \text{diag}(a_1, \dots, a_n)$ .

4.2. Dokažte:

(a)  $AA^\dagger$  je symetrická,

(b)  $A = (A^\dagger)^T A^T A$ ,

4.3. Bud'  $A \in \mathbb{R}^{m \times n}$ ,  $b \in \mathbb{R}^m$  a pro soustavu  $Ax = b$  ukažte, že

(a) má aspoň jedno řešení právě tehdy, když  $AA^\dagger b = b$ ,

(b) má nanejvýš jedno řešení právě tehdy, když  $AA^\dagger = I_n$ ,

(c) má právě jedno řešení právě tehdy, když  $AA^\dagger b = b$  a  $AA^\dagger = I_n$ ,

(d) má-li aspoň jedno řešení, pak všechna řešení jsou vyjádřit jako  $A^\dagger b + y - A^\dagger A y$ , kde  $y \in \mathbb{R}^n$  je libovolné.

4.4. Bud'  $A \in \mathbb{R}^{m \times n}$ . Dokažte, že  $X = A^\dagger$  je optimálním řešením úlohy

$$\min_{X \in \mathbb{R}^{n \times m}} \|AX - I_m\|_F$$

a diskutujte jednoznačnost optima (zde je potřeba znalost z kapitoly 5). Pseudoinverze  $A^\dagger$  je tedy v jisté normě nejlepší jednostrannou inverzí k  $A$ .

\*4.5. Ukažte, že maticová soustava rovnic  $AXB = C$  s libovolnými maticemi vhodných rozměrů má řešení právě tehdy, když  $AA^\dagger CB^\dagger B = C$ . Navíc, množina řešení je tvaru  $X = A^\dagger CB^\dagger + Y - A^\dagger A Y B B^\dagger$ , kde  $Y$  je libovolná matice příslušných rozměrů.

4.6. Dokažte zbylé vlastnosti věty 4.10 o Drazinově pseudoinverzi.

4.7. Dokažte  $((A^D)^D)^D = A^D$ ,  $(A^D)^T = (A^T)^D$ .

4.8. Najděte příklad, kdy  $(A^D)^D \neq A$ .



# Kapitola 5

## Maticová norma

Norma matice je důležitá, protože umožňuje říci jak je matice „velká“, jaká je vzdálenost dvou matic, co znamená, že posloupnost matic konverguje k matici, nebo zavést nekonečný součet matic.

Připomeňme, že *vektorová norma* na reálném či komplexním vektorovém prostoru  $V$  je zobrazení  $\|\cdot\|: V \rightarrow \mathbb{R}$ , splňující:

- (1)  $\|x\| \geq 0$  pro všechna  $x \in V$ , a rovnost nastane pouze pro  $x = 0$ ,
- (2)  $\|\alpha x\| = |\alpha| \cdot \|x\|$  pro všechna  $x \in V$ , a pro všechna  $\alpha \in \mathbb{R}$  resp.  $\alpha \in \mathbb{C}$ ,
- (3)  $\|x + y\| \leq \|x\| + \|y\|$ .

Důležitým příkladem norm v  $\mathbb{R}^n$  jsou  $\ell_p$ -normy definované

$$\|x\|_{\ell_p} = \left( \sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}},$$

kde  $p \geq 1$ . Speciálně pro  $p \in \{1, 2\}$  a limitním přechodem pro  $p = \infty$  dostáváme součtovou, eukleidovskou a maximovou normu

$$\begin{aligned} \|x\|_{\ell_1} &= \sum_{i=1}^n |x_i|, \\ \|x\|_{\ell_2} &= \sqrt{\sum_{i=1}^n x_i^2}, \\ \|x\|_{\ell_\infty} &= \max_{i=1, \dots, n} |x_i|. \end{aligned}$$

Protože prostor matice  $\mathbb{R}^{n \times n}$  tvoří vektorový prostor, můžeme vektorové normy aplikovat i na matice. Tím dostáváme následující normy

$$\begin{aligned} \|A\|_{\ell_1} &= \sum_{i,j=1}^n |a_{ij}|, \\ \|A\|_{\ell_2} &= \sqrt{\sum_{i,j=1}^n a_{ij}^2}, \\ \|A\|_{\ell_\infty} &= \max_{i,j=1, \dots, n} |a_{ij}|. \end{aligned}$$

Speciálně, norma  $\|A\|_{\ell_2}$  se často nazývá *Frobeniova norma* a značí se  $\|A\|_F$ . Můžeme ji ekvivalentně vyjádřit jako

$$\|A\|_F = \sqrt{\text{trace}(A^T A)}.$$

### 5.1 Definice a příklady, indukovaná norma

Po maticové normě se oproti vektorové vyžaduje ještě jedna důležitá vlastnost navíc, tzv. *konsistence* či *sub-multiplikativita*,  $\|AB\| \leq \|A\| \cdot \|B\|$ . Proto maticovou normu definujeme následujícím způsobem. Omezíme se na reálné matice a normy, komplexní se definují analogicky.

**Definice 5.1** (Maticová norma). Zobrazení  $\|\cdot\|: \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$  je reálná *maticová norma*, pokud splňuje pro všechny  $A \in \mathbb{R}^{n \times n}$  a  $\alpha \in \mathbb{R}$ :

- (1)  $\|A\| \geq 0$ , a rovnost nastane pouze pro  $A = 0$ ,
- (2)  $\|\alpha A\| = |\alpha| \cdot \|A\|$ ,
- (3)  $\|A + B\| \leq \|A\| + \|B\|$ .
- (4)  $\|AB\| \leq \|A\| \cdot \|B\|$ .

**Tvrzení 5.2.** Pro každou maticovou normu a  $k \in \mathbb{N}$  platí  $\|A^k\| \leq \|A\|^k$ .

*Důkaz.* Matematickou indukcí: Aplikujeme iterativně vlastnost (4) z definice maticové normy.  $\square$

Přirozená otázka je, jestli  $\ell_p$ -normy splňují vlastnost (4), tedy jestli jsou skutečnými maticovými normami.

**Věta 5.3.**

- (1)  $\|A\|_{\ell_1}$  je maticová norma,
- (2)  $\|A\|_{\ell_2} = \|A\|_F$  je maticová norma,
- (3)  $\|A\|_{\ell_\infty}$  není maticovou normou, ale je jí

$$\|A\|_{\ell_\infty} = n \cdot \max_{i,j=1,\dots,n} |a_{ij}|.$$

*Důkaz.* Stačí ověřit jen vlastnost (4):

- (1)  $\|AB\|_{\ell_1} = \sum_{i,j} |\sum_k a_{ik} b_{kj}| \leq \sum_{i,j,k} |a_{ik} b_{kj}| \leq \sum_{i,j,k,\ell} |a_{ik} b_{\ell j}| = (\sum_{i,k} |a_{ik}|)(\sum_{\ell,j} |b_{\ell j}|) = \|A\|_{\ell_1} \|B\|_{\ell_1}$
- (2) S využitím Cauchyho–Schwarzovy nerovnosti  $\|AB\|_F^2 = \sum_{i,j} (\sum_k a_{ik} b_{kj})^2 \leq \sum_{i,j} (\sum_k a_{ik}^2)(\sum_\ell b_{\ell j}^2) = (\sum_{i,k} a_{ik}^2)(\sum_{\ell,j} b_{\ell j}^2) = \|A\|_F^2 \|B\|_F^2$ ,
- (3)  $\|A\|_{\ell_\infty}$  není maticovou normou, neboť například pro  $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$  je  $2 = \|A^2\| \not\leq \|A\|^2 = 1$ .  
Niméně, pro variantu  $\|A\|_{\ell_\infty}$  máme  $\|AB\|_{\ell_\infty} = n \cdot \max_{i,j} |\sum_k a_{ik} b_{kj}| \leq n \cdot \max_{i,j} \sum_k |a_{ik}| |b_{kj}| \leq n \cdot \max_{i,j} \sum_k \|A\|_{\ell_\infty} \|B\|_{\ell_\infty} = n \cdot \|A\|_{\ell_\infty} \cdot n \cdot \|B\|_{\ell_\infty} = \|A\|_{\ell_\infty} \|B\|_{\ell_\infty}$ .  $\square$

Je-li  $\|v\|$  vektorovou normou, pak  $\alpha\|v\|$  je vektorovou normou pro všechna  $\alpha > 0$ . Tato vlastnost pro maticovou normu už neplatí, jak ukázal protipříklad pro  $\|A\|_{\ell_\infty}$ . Niméně, platí následující slabší verze.

**Tvrzení 5.4.** Je-li  $\|A\|$  maticovou normou, pak  $\alpha\|A\|$  je maticovou normou pro všechna  $\alpha \geq 1$ .

*Důkaz.* Stačí opět ověřit jen vlastnost (4). Protože  $\|AB\| \leq \|A\| \cdot \|B\|$ , tak platí  $\alpha\|AB\| \leq \alpha\|A\| \cdot \|B\| \leq (\alpha\|A\|) \cdot (\alpha\|B\|)$ .  $\square$

Přestože již máme několik příkladů maticových norem, význačnější roli hrají indukované maticové normy.

**Definice 5.5** (Indukovaná maticová norma). Buď  $\|\cdot\|: \mathbb{R}^n \rightarrow \mathbb{R}$  vektorová norma. Pak indukovaná maticová norma na  $\mathbb{R}^{n \times n}$  je definovaná

$$\|A\| = \max_{x: \|x\|=1} \|Ax\|.$$

Protože jednotková sféra je kompaktní a norma spojitá, tak maximum z definice se nabyde a indukovaná norma je tak dobře definovaná. Alternativně můžeme indukovanou normu též vyjádřit jako

$$\|A\| = \max_{x: \|x\|=1} \|Ax\| = \max_{x: \|x\| \leq 1} \|Ax\| = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|}.$$

Druhá rovnost nastává díky tomu, že norma je konvexní funkcí a maximum konvexní funkce se nabyde na okraji množiny přípustných řešení.

**Poznámka 5.6** (Geometrický význam indukované normy). Uvažujme lineární zobrazení  $x \mapsto Ax$  s regulární maticí  $A$ . Jednotková kružnice  $\{x \in \mathbb{R}^n; \|x\| = 1\}$  se zobrazí na množinu  $\{Ax \in \mathbb{R}^n; \|x\| = 1\}$ . Největší velikost vektorů z této množiny je právě  $\|A\|$ , čili indukovaná norma odpovídá maximálnímu roztáhnutí jednotkové kružnice. Jak uvidíme ve tvrzení 5.10, nejmenší zúžení je naopak rovno  $\|A^{-1}\|^{-1}$ .

**Věta 5.7.** *Indukovaná norma je maticovou normou a navíc platí  $\|Ax\| \leq \|A\| \cdot \|x\|$  (tj., maticová norma je konsistentní s vektorovou normou, která ji indukuje).*

*Důkaz.* Ověříme axiomy z definice maticové normy

- (1) Zřejmě  $\|A\| \geq 0$ . Rovnost nastane pokud  $\|Ax\| = 0$  pro všechna  $x$ , což je jenom když  $A = 0$ .
- (2) Vlastnost  $\|\alpha A\| = |\alpha| \cdot \|A\|$  je zřejmá.
- (3)  $\|A+B\| = \max_{x:\|x\|=1} \|(A+B)x\| \leq \max_{x:\|x\|=1} (\|Ax\| + \|Bx\|) \leq \max_{x:\|x\|=1} \|Ax\| + \max_{x:\|x\|=1} \|Bx\| = \|A\| + \|B\|$ .
- (4) Odvodíme (za předpokladu  $Bx \neq 0$ , kterýžto lze snadno učinit)

$$\begin{aligned} \|AB\| &= \max_{x \neq 0} \frac{\|ABx\|}{\|x\|} = \max_{x \neq 0} \frac{\|ABx\|}{\|Bx\|} \frac{\|Bx\|}{\|x\|} \leq \max_{x \neq 0} \frac{\|ABx\|}{\|Bx\|} \max_{x \neq 0} \frac{\|Bx\|}{\|x\|} \\ &\leq \max_{y \neq 0} \frac{\|Ay\|}{\|y\|} \max_{x \neq 0} \frac{\|Bx\|}{\|x\|} = \|A\| \cdot \|B\|. \end{aligned}$$

- (5) Nakonec, z (ekvivalentního vyjádření) definice indukované normy máme pro každé  $x \neq 0$ :  $\|A\| \geq \frac{\|Ax\|}{\|x\|}$ . Tedy  $\|Ax\| \leq \|A\| \cdot \|x\|$ .  $\square$

Podíváme se, jaké maticové normy indukují vektorové  $\ell_p$ -normy. Pro ten účel označíme tzv. maticové  $p$ -normy

$$\|A\|_p = \max_{x:\|x\|_{\ell_p}=1} \|Ax\|_{\ell_p}.$$

Speciálně pro  $p = 2$  se maticová 2-norma nazývá *spektrální norma* a je jednou z nejdůležitějších maticových norem. (Schválně, která norma je defaultní v Matlabu či Octave?)

**Lemma 5.8.** *Maticová norma  $\|A\|_2$  je ortogonálně invariantní, tj.  $\|UAV\|_2 = \|A\|_2$  pro libovolné ortogonální matice  $U, V$ .*

*Důkaz.* S využitím toho, že eukleidovská norma je ortogonálně invariantní,

$$\begin{aligned} \|UAV\|_2 &= \max_{x:\|x\|_{\ell_2}=1} \|UAVx\|_{\ell_2} = \max_{x:\|x\|_{\ell_2}=1} \|AVx\|_{\ell_2} \\ &= \max_{x:\|Vx\|_{\ell_2}=1} \|AVx\|_{\ell_2} = \max_{y:\|y\|_{\ell_2}=1} \|Ay\|_{\ell_2} = \|A\|_2. \end{aligned} \quad \square$$

Pro  $p \in \{1, 2, \infty\}$  lze maticovou  $p$ -normu jednoduše vyjádřit. Pro jiné hodnoty  $p$  již žádný jednoduchý vzoreček není znám [Higham, 1996]. Pro hodnoty  $p \notin \{1, 2, \infty\}$  je dokonce výpočet  $\|A\|_p$  nebo její aproximace NP-těžký problém [Hendrickx and Olshevsky, 2010].

**Věta 5.9.** *Platí.<sup>1)</sup>*

- (1)  $\|A\|_1 = \max_j \sum_i |a_{ij}|$ ,
- (2)  $\|A\|_2 = \sigma_1(A)$  (největší singulární číslo  $A$ ),
- (3)  $\|A\|_\infty = \max_i \sum_j |a_{ij}| = \|A^T\|_1$ ,

*Důkaz.* Stačí ověřit jen vlastnost (4):

- (1) Označme  $h := \max_j \sum_i |a_{ij}| = \max_j \|A_{*j}\|_{\ell_1}$ .

„ $\geq$ “ Zvolme  $x := e_k$ , kde  $k$  je index, pro nějž se nabyde maximum v definici  $h$ . Pak  $\|A\|_1 \geq \|Ax\|_{\ell_1} = \|A_{*k}\|_{\ell_1} = h$ .

„ $\leq$ “ Odvodme  $\|A\|_1 = \max_{x:\|x\|_{\ell_1}=1} \|Ax\|_{\ell_1} = \max_{x:\|x\|_{\ell_1}=1} \|\sum_j x_j A_{*j}\|_{\ell_1} \leq \max_{x:\|x\|_{\ell_1}=1} \sum_j |x_j| \cdot \|A_{*j}\|_{\ell_1} \leq \max_{x:\|x\|_{\ell_1}=1} \sum_j |x_j| h = h$ .

<sup>1)</sup>Mnemotechnická pomůcka na zapamatování: symbol „1“ je svislý, proto 1-norma vybírá maximální sloupcový součet, zatímco „ $\infty$ “ je nalezato, proto  $\infty$ -norma prochází řádky.

(2) Nechť  $A = U\Sigma V^T$  je SVD rozklad. Dle lemmatu 5.8 je  $\|A\|_2 = \|\Sigma\|_2 = \max_{x:\|x\|_{\ell_2}=1} \|\Sigma x\|_{\ell_2} \leq \max_{x:\|x\|_{\ell_2}=1} \|\sigma_1 I_n x\|_{\ell_2} = \max_{x:\|x\|_{\ell_2}=1} \sigma_1 \|x\|_{\ell_2} = \sigma_1$ , a rovnost nastane pro  $x = e_1$ .

(3) Označme  $h := \max_i \sum_j |a_{ij}|$ .

„ $\geq$ “ Buď  $k$  index, pro nějž se nabyde maximum v definici  $h$ . Zvolme  $x := \text{sgn}(A_{k*})$  znaménkový vektor  $k$ -tého řádku  $A$ . Pak  $\|A\|_{\infty} \geq \|Ax\|_{\ell_{\infty}} = h$ .

„ $\leq$ “ Odvoďme

$$\begin{aligned} \|A\|_{\infty} &= \max_{x:\|x\|_{\ell_{\infty}}=1} \|Ax\|_{\ell_{\infty}} = \max_{x:\|x\|_{\ell_{\infty}}=1} \max_i \left| \sum_j a_{ij} x_j \right| \\ &\leq \max_{x:\|x\|_{\ell_{\infty}}=1} \max_i \sum_j |a_{ij}| |x_j| \leq \max_{x:\|x\|_{\ell_{\infty}}=1} \max_i \sum_j |a_{ij}| \|x\|_{\ell_{\infty}} = h. \quad \square \end{aligned}$$

**Tvrzení 5.10.** Pro indukovanou maticovou normu a regulární matici  $A$  platí  $\min_{x:\|x\|=1} \|Ax\| = \|A^{-1}\|^{-1}$ .

*Důkaz.*  $\min_{x:\|x\|=1} \|Ax\| = \min_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \max_{x \neq 0}^{-1} \frac{\|x\|}{\|Ax\|} = \max_{y \neq 0}^{-1} \frac{\|A^{-1}y\|}{\|y\|} = \|A^{-1}\|^{-1}$ .  $\square$

**Poznámka 5.11** (Ekvivalence maticových norem). Všechny maticové normy na  $\mathbb{R}^{n \times n}$  jsou ekvivalentní v následujícím smyslu: Buďte  $\|\cdot\|_{\alpha}, \|\cdot\|_{\beta}$  dvě normy. Pak existují čísla  $\gamma, \delta > 0$  takové, že pro každou matici  $A \in \mathbb{R}^{n \times n}$  platí:

$$\gamma \|A\|_{\alpha} \leq \|A\|_{\beta} \leq \delta \|A\|_{\alpha}.$$

Pro konkrétní normy pak můžeme určit i odpovídající koeficienty, například platí [Horn and Johnson, 1985, str. 365]:

$$\begin{aligned} \frac{1}{\sqrt{n}} \|A\|_1 &\leq \|A\|_2 \leq \sqrt{n} \|A\|_1, \\ \frac{1}{\sqrt{n}} \|A\|_{\infty} &\leq \|A\|_2 \leq \sqrt{n} \|A\|_{\infty}, \\ \|A\|_2 &\leq \|A\|_F \leq \sqrt{n} \|A\|_2. \end{aligned}$$

## 5.2 Spektrální poloměr versus norma matice

Jedna (a zdaleka ne jediná) z podstatných vlastností maticové normy je vztah ke spektrálnímu poloměru. Než k němu přikročíme, položme si otázku, jestli je spektrální poloměr maticovou normou.

**Příklad 5.12.** Spektrální poloměr není maticovou normou, protože

1.  $\rho(A) = 0$  i pro  $A \neq 0$ , například pro  $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ .
2. Neplatí  $\rho(A+B) \leq \rho(A) + \rho(B)$ , například pro  $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ ,  $B = A^T$ .
3. Neplatí  $\rho(AB) \leq \rho(A)\rho(B)$ , například pro  $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ ,  $B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ , neboť  $\rho(A) = \rho(B) = 1$ ,  $\rho(AB) = (1 + \sqrt{5})/2$ .

Nikoli náhodou matice z protipříkladů jsou nediagonalizovatelné. Bude spektrální poloměr maticovou normou, když se omezíme třeba jen na symetrické matice?  $\square$

Přestože tedy spektrální poloměr není maticovou normou, můžeme ho pomocí maticové normy více či méně těsně aproximovat.

**Věta 5.13** (Odhad spektrálního poloměru pomocí normy). Buď  $A \in \mathbb{R}^{n \times n}$ . Pak pro každou maticovou normu  $\rho(A) \leq \|A\|$ .

*Důkaz.* Buď  $\lambda \in \mathbb{C}$  libovolné vlastní číslo a  $x$  odpovídající vlastní vektor matice  $A$ , tedy  $Ax = \lambda x$ . Definujme matici  $X := (x \mid 0 \mid \dots \mid 0)$ . Protože platí  $AX = \lambda X$ , můžeme odvodit

$$|\lambda| \cdot \|X\| = \|\lambda X\| = \|AX\| \leq \|A\| \cdot \|X\|.$$

Vydělením  $\|X\| \neq 0$  dostáváme  $|\lambda| \leq \|A\|$ .  $\square$

Nerovnost z věty může být ostrá, například pro  $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$  je  $0 = \rho(A) < \|A\|$ , čili nerovnost je ostrá dokonce pro každou maticovou normu. Navíc rozdíl  $\|A\| - \rho(A)$  může být libovolně velký, stačí uvažovat matici  $A = \begin{pmatrix} 0 & \alpha \\ 0 & 0 \end{pmatrix}$  a  $\alpha > 0$  dost velké.

Na druhou stranu, jak říká věta dole, ke spektrálnímu poloměru se můžeme přiblížit libovolně blízko za použití vhodné normy.

**Lemma 5.14.** *Bud'  $S \in \mathbb{R}^{n \times n}$  regulární a  $\|\cdot\|$  maticová norma na  $\mathbb{R}^{n \times n}$ . Pak  $\|A\|_S := \|SAS^{-1}\|$  je také maticovou normou.*

*Důkaz.* Axiomy z definice se snadno ověří. Například konsistence se odvodí takto:

$$\|AB\|_S = \|SABS^{-1}\| = \|SAS^{-1}SBS^{-1}\| \leq \|SAS^{-1}\| \|SBS^{-1}\| = \|A\|_S \|B\|_S. \quad \square$$

**Věta 5.15.** *Bud'  $A \in \mathbb{R}^{n \times n}$  a  $\varepsilon > 0$ . Pak existuje maticová norma taková, že  $\rho(A) \leq \|A\| \leq \rho(A) + \varepsilon$ .*

*Důkaz.* Každá čtvercová matice je podobná matici v Jordanově normální formě. Tedy existuje regulární  $R \in \mathbb{C}^{n \times n}$  taková, že  $RAR^{-1} = J$  a matice  $J$  je v Jordanově normální formě. Definujme jako  $D$  diagonální matici s prvky  $t, t^2, \dots, t^n$  na diagonále, kde  $t > 0$  je určitý parametr. Nyní

$$DJD^{-1} = \begin{pmatrix} \lambda_1 & t^{-1}\{0, 1\} & 0 & \dots & 0 \\ 0 & \lambda_2 & t^{-1}\{0, 1\} & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \lambda_{n-1} & t^{-1}\{0, 1\} \\ 0 & \dots & \dots & 0 & \lambda_n \end{pmatrix},$$

kde  $\{0, 1\}$  symbolicky značí možnou hodnotu 0 nebo 1 (přesný tvar Jordanových buněk zde není důležitý). Tedy pro dost velké  $t > 0$  jsou mimodiagonální prvky libovolně malé, a součet prvků  $i$ -tého sloupce je libovolně blízko vlastnímu číslu  $\lambda_i$ . Podle lemmatu 5.14 definujme normu  $\|A\|_S := \|SAS^{-1}\|_1$  pro  $S = DR$ . Pak

$$\|A\|_S = \|SAS^{-1}\|_1 = \|DRAR^{-1}D^{-1}\|_1 = \|DJD^{-1}\|_1 \leq \rho(A) + \varepsilon. \quad \square$$

Jak jsme zmínili na začátku této sekce, jedna z vlastností maticové normy je, že nám dává horní odhad na spektrální poloměr. To se může hodit i v oblastech, kde se matice přímo nevyskytují.

**Příklad 5.16** (Odhady pro kořeny polynomů). Uvažujme polynom  $p(x) = x^n + a_{n-1}x^{n-1} + \dots + a_1x + a_0$ . Připomeňme, že matice společnice polynomu  $p(x)$  je čtvercová matice řádu  $n$  definovaná

$$C(p) := \begin{pmatrix} 0 & \dots & \dots & 0 & -a_0 \\ 1 & \ddots & & \vdots & -a_1 \\ 0 & \ddots & \ddots & \vdots & -a_2 \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \dots & 0 & 1 & -a_{n-1} \end{pmatrix}$$

a platí, že kořeny polynomu  $p(x)$  se shodují s vlastními čísly matice  $C(p)$ .

Bud'  $x^*$  kořen polynomu  $p(x)$ . Pak pro libovolnou maticovou normu platí  $|x^*| \leq \rho(C(p)) \leq \|C(p)\|$ . Různou volbou normy dostaneme různé odhady na umístění kořenů. Například použitím normy  $\|\cdot\|_\infty$  získáme

$$|x^*| \leq \|C(p)\|_\infty = \max\{|a_0|, 1 + |a_1|, \dots, 1 + |a_{n-1}|\} \leq 1 + \max_{i=0, \dots, n-1} |a_i|,$$

což je tzv. Cauchyho mez. □

Dalších vztahů mezi maticovou normou a spektrálním poloměrem je celá řada. V příští sekci zmíníme ještě jeden pozoruhodný výsledek (věta 5.19), ale k němu potřebujeme znát něco o mocninných posloupnostech.

### 5.3 Mocninné posloupnosti

Nyní se dostáváme k mocninné posloupnosti  $A, A^2, A^3, \dots$  a její konvergenci. Nejprve ale rozšíříme pojem konvergence posloupnosti na matice. Buď  $B_1, B_2, \dots \in \mathbb{R}^{n \times n}$  posloupnost matic a  $A \in \mathbb{R}^{n \times n}$ . Pak výrazem  $\lim_{k \rightarrow \infty} B_k = A$  myslíme  $\|B_k - A\| \rightarrow_{k \rightarrow \infty} 0$ . Protože podle poznámky 5.11 jsou všechny normy ekvivalentní, je jedno jakou použijeme. Použitím normy  $\|A\|_{\ell_\infty}$  z věty 5.3 vidíme, že maticová konvergence je ekvivalentní konvergenci po prvcích.

**Věta 5.17** (Postačující podmínka pro konvergenci). *Buď  $A \in \mathbb{R}^{n \times n}$ . Je-li  $\|A\| < 1$  pro nějakou maticovou normu, pak  $\lim_{k \rightarrow \infty} A^k = 0$ .*

*Důkaz.* Podle tvrzení 5.2 platí  $\|A^k\| \leq \|A\|^k \rightarrow_{k \rightarrow \infty} 0$ . Tudíž  $A^k \rightarrow_{k \rightarrow \infty} 0$ . □

Obrácená implikace obecně neplatí pro každou normu, stačí vzít matici  $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$  a určité normy. Norma matice  $A$  může být dokonce libovolně velká; stačí opět uvažovat  $A = \begin{pmatrix} 0 & \alpha \\ 0 & 0 \end{pmatrix}$  a  $\alpha > 0$  dost velké.

Na druhou stranu, ekvivalenci dosáhneme přechodem ke spektrálnímu poloměru. Navíc do hry kromě mocninných posloupností zapojíme ještě mocninné řady a ukážeme slavný výsledek o Neumannových řadách. Je to přímočaré zobecnění známé poučky o tom, že geometrická řada  $\sum_{k=0}^{\infty} q^k = \frac{1}{1-q}$  právě tehdy, když  $|q| < 1$ .

**Věta 5.18** (Neumannovy řady). *Buď  $A \in \mathbb{R}^{n \times n}$ . Následující jsou ekvivalentní:*

- (1)  $\rho(A) < 1$ ,
- (2)  $\lim_{k \rightarrow \infty} A^k = 0$ ,
- (3)  $\sum_{k=0}^{\infty} A^k$  konverguje.

*Pokud platí jedna z těchto vlastností, pak  $(I_n - A)^{-1} = \sum_{k=0}^{\infty} A^k$ .*

*Důkaz.*

„(1)  $\Rightarrow$  (2)“ Protože  $\rho(A) < 1$ , tak podle věty 5.15 existuje maticová norma taková, že  $\|A\| < 1$ . Nyní stačí použít větu 5.17.

„(1)  $\Leftarrow$  (2)“ Buď  $Ax = \lambda x$ , kde  $\lambda$  je vlastní číslo a  $x$  odpovídající vlastní vektor. Pak  $A^k x = \lambda^k x$ . Protože  $A^k \rightarrow_{k \rightarrow \infty} 0$ , musí i pravá strana konvergovat k nule. Tudíž  $|\lambda| < 1$ , z čehož  $\rho(A) < 1$ .

„(3)  $\Rightarrow$  (2)“ Jasně.

„(2)  $\Rightarrow$  (3)“ Matice  $(I_n - A)(\sum_{k=0}^m A^k) = I_n - A^{m+1}$  konverguje k  $I_n$  pro  $m \rightarrow \infty$ . Pro dost velké  $m$  je  $I_n - A^{m+1}$  regulární, a proto je  $I_n - A$  rovněž regulární. Tudíž můžeme psát  $\sum_{k=0}^m A^k = (I_n - A)^{-1}(I_n - A^{m+1}) \rightarrow_{m \rightarrow \infty} (I_n - A)^{-1}$ . □

Zde je dobré si uvědomit, že  $(I_n - A)^{-1}$  může existovat, aniž by platila jedna ze tří podmínek věty.

Neumannovy řady se hodí leckde. Jedna okamžitá aplikace je pro aproximaci inverzní matice  $(I_n - A)^{-1}$  pomocí několika prvních členů řady  $\sum_{k=0}^{\infty} A^k$ . Je-li  $\rho(A)$  dost malé, pak i aproximace  $(I_n - A)^{-1} \approx I + A$  může být použitelná.

**Věta 5.19** (Gelfandova formule). *Pro každou maticovou normu a  $A \in \mathbb{R}^{n \times n}$  platí  $\rho(A) = \lim_{k \rightarrow \infty} \|A^k\|^{1/k}$ .*

*Důkaz.* „ $\leq$ “ Platí  $\rho(A)^k = \rho(A^k) \leq \|A^k\|$ , z čehož  $\rho(A) \leq \|A^k\|^{1/k}$ . Tudíž nerovnost platí i v limitě.

„ $\geq$ “ Buď  $\varepsilon > 0$  libovolné a definujme  $B := \frac{1}{\rho(A) + \varepsilon} A$ . Pak  $\rho(B) = \frac{1}{\rho(A) + \varepsilon} \rho(A) < 1$  a proto  $B^k \rightarrow_{k \rightarrow \infty} 0$ . Buď  $k_0$  tak velké, že pro všechna  $k \geq k_0$  platí  $\|B^k\| < 1$ . To nám dává  $\|A^k\| < (\rho(A) + \varepsilon)^k$ , z čehož  $\|A^k\|^{1/k} < \rho(A) + \varepsilon$ . Tudíž pro všechna dost velká  $k$  je  $\|A^k\|^{1/k}$  je libovolně blízko  $\rho(A)$ . □

### 5.4 Ortogonálně invariantní normy

**Definice 5.20.** Maticová norma  $\|A\|$  je ortogonálně invariantní, pokud  $\|UAV\| = \|A\|$  pro libovolné ortogonální matice  $U, V$ .

Buď  $A \in \mathbb{R}^{m \times n}$  hodnosti  $r$  a buď  $A = U\Sigma V^T$  její SVD rozklad. Pro každou ortogonálně invariantní normu platí  $\|A\| = \|U\Sigma V^T\| = \|\Sigma\|$ , tudíž norma závisí pouze na singulárních číslech matice  $A$ .

Z lemmatu 5.8 víme, že spektrální norma  $\|A\|_2$  je ortogonálně invariantní. Nicméně, není to jediný případ. Následující věta říká, že i Frobeniova norma je ortogonálně invariantní, protože závisí pouze na singulárních číslech.

**Věta 5.21.** *Buď  $A \in \mathbb{R}^{m \times n}$  se singulárními čísly  $\sigma_1, \dots, \sigma_r$ . Pak  $\|A\|_F = \sqrt{\sum_{i=1}^r \sigma_i^2}$ .*

*Důkaz.* Podle věty 3.2 máme  $\|A\|_F = \sqrt{\sum_{i=1}^m \sum_{j=1}^n a_{ij}^2} = \sqrt{\text{trace}(A^T A)} = \sqrt{\sum_{i=1}^r \sigma_i^2}$ .  $\square$

Dalšími příklady ortogonálně invariantních norm je například *Ky Fanova* norma, definovaná jako součet  $k$  největších singulárních čísel, kde  $k$  je pevné. Tedy

$$\|A\|_{K_k} := \sum_{i=1}^k \sigma_i.$$

Speciálně, pro  $k = 1$  dostaneme spektrální normu a pro  $k = r$  dostaneme tzv. *trace norm* či *nukleární normu*. Nukleární norma se určitým způsobem dá využít pro aproximaci hodnosti matice – např. v robustní PCA namísto minimalizace hodnosti matice minimalizujeme její nukleární normu, protože je jejím nejlepším dolním konvexním odhadem uvnitř jednotkové koule. Nukleární norma se též značí  $\|A\|_*$  a je zajímavý její vztah s Frobeniovou normou:

$$\|A\|_F = \sqrt{\text{trace}(A^T A)}, \quad \|A\|_* = \text{trace}(\sqrt{A^T A}).$$

To *Schattenova* norma je při pevném  $p$  definovaná jako  $\ell_p$ -norma vektoru singulárních čísel  $(\sigma_1, \dots, \sigma_r)$ , tedy

$$\|A\|_{S_p} := (\sum_{i=1}^r \sigma_i^p)^{1/p}.$$

Ukázat ortogonální invariantnost dvou výše zmíněných norm je triviální. Stejně tak první dvě vlastnosti z definice maticové normy. Větší práci dá ukázat trojúhelníkovou nerovnost a konsistenci, to však dokazovat nebudeme.

Výše uvedené normy se používají například v kvantové teorii informace. Měří mj. vzdálenost dvou kvantových stavů a tím i podobnost výstupů (ideálního a skutečného) daného protokolu. To umožňuje určovat jak dobře protokol funguje [Wilde, 2017].

## 5.5 Další aplikace maticové normy

**Prokrustův problém.** Mějme  $A, B \in \mathbb{R}^{m \times n}$  a otázka zní, jestli  $A$  vznikne z  $B$  vhodnou rotací, tj. jestli existuje ortogonální  $Q \in \mathbb{R}^{m \times m}$  taková, že  $A = QB$ . Problémy tohoto typu vznikají přirozeně například když porovnáváme dvě chemické struktury a chceme vědět, jak moc si jsou podobné, případně jestli jsou shodné až na otočení resp. zrcadlení. V tom případě  $Q$  reprezentuje matici otočení a řádky  $A$  resp.  $B$  udávají např. pozici jednotlivých atomů v prostoru první resp. druhé struktury.

Řešme obecnější optimalizační problém

$$\min\{\|A - QB\|; Q \in \mathbb{R}^{m \times m} \text{ ortogonální}\}. \quad (5.1)$$

Zřejmě, pokud je optimální hodnota 0, pak  $A = QB$  pro vhodnou ortogonální  $Q$  a řešení původního problému existuje. Pokud je optimální hodnota kladná, pak řešení původního problému existuje. Volba vhodné maticové normy je na nás. Je přirozené volit ortogonálně invariantní normu, protože pracujeme s ortogonálními maticemi. Speciálně výhodné je vybrat Frobeniovu normu.

**Věta 5.22.** *Nechť  $AB^T = U\Sigma V^T$  je SVD rozklad matice  $AB^T$ . Pak optimum úlohy (5.1) je  $Q = UV^T$ . Dále, existuje ortogonální  $Q$  taková, že  $A = QB$  právě tehdy, když*

$$\|A\|_F = \|B\|_F = \sqrt{\sum_{i=1}^r \sigma_i}.$$

*Důkaz.* Připomínáme, že Frobeniova norma jako vektorová norma je indukovaná skalárním součinem  $\langle A, B \rangle = \text{trace}(A^T B)$ . Nyní upravme účelovou funkci minimalizační úlohy

$$\begin{aligned} \|A - QB\|_F^2 &= \langle A - QB, A - QB \rangle = \|A\|_F^2 + \|QB\|_F^2 - \langle A, QB \rangle - \langle QB, A \rangle \\ &= \|A\|_F^2 + \|B\|_F^2 - 2\langle QB, A \rangle. \end{aligned}$$

Nechť  $AB^T = U\Sigma V^T$  je SVD rozklad matice  $AB^T$ . Jelikož  $\|A\|_F^2 + \|B\|_F^2$  je konstantní, maximalizujeme vlastně na množině ortogonálních matic  $Q$  funkci

$$\langle QB, A \rangle = \text{trace}(B^T Q^T A) = \text{trace}(AB^T Q^T) = \text{trace}(U\Sigma V^T Q^T) = \text{trace}(\Sigma V^T Q^T U).$$

Označíme-li ortogonální matici  $H := V^T Q^T U$ , má výraz nahoře hodnotu

$$\langle QB, A \rangle = \text{trace}(\Sigma H) = \sum_{i=1}^r \sigma_i h_{ii}.$$

Jelikož  $h_{ii} \in [-1, 1]$  pro každé  $i$ , má výraz největší hodnotu  $\sum_{i=1}^r \sigma_i$ , která se nabyde pokud  $h_{ii} = 1$ , tedy  $H = I_m$ . Z rovnosti  $I_m = H = V^T Q^T U$  pak odvodíme  $Q = UV^T$ .

V druhé části tvrzení je zřejmé, že musí platit  $\|A\|_F = \|B\|_F$  jako nutná podmínka díky ortogonální invariantnosti Frobeniovy normy. Z první části máme, že existuje ortogonální  $Q$  taková, že  $A = QB$  právě tehdy, když

$$0 = \|A - QB\|_F^2 = 2\|A\|_F^2 - 2\langle QB, A \rangle,$$

neboli když

$$\|A\|_F^2 = \langle QB, A \rangle = \sum_{i=1}^r \sigma_i. \quad \square$$

**Poznámka 5.23.** Speciálně pro  $B = I$  pak dostáváme, že nejlepší aproximace matice  $A$  pomocí ortogonální matice  $Q$  je  $Q = UV^T$ , kde  $A = U\Sigma V^T$  je SVD rozklad matice  $A$ .

**Interpretace Mooreovy–Penroseovy pseudoinverze.** Podobnou úlohou je následující interpretace Mooreovy–Penroseovy pseudoinverze, která říká, že ve Frobeniově normě je matice  $AA^\dagger$  nejlepší aproximací jednotkové matice. Nejprve dokážeme pomocné tvrzení, které je maticovou variantou Pythagorovy věty.

**Lemma 5.24.** *Pokud pro matice  $X, Y \in \mathbb{R}^{m \times n}$  platí  $X^T Y = 0$  (čili  $\mathcal{S}(X) \perp \mathcal{S}(Y)$ ), pak*

$$\|X + Y\|_F^2 = \|X\|_F^2 + \|Y\|_F^2.$$

*Důkaz.* Pišme

$$\begin{aligned} \|X + Y\|_F^2 &= \text{trace}((X + Y)^T (X + Y)) = \text{trace}(X^T X + X^T Y + Y^T X + Y^T Y) \\ &= \text{trace}(X^T X + Y^T Y) = \|X\|_F^2 + \|Y\|_F^2. \end{aligned} \quad \square$$

**Věta 5.25** (Interpretace Mooreovy–Penroseovy pseudoinverze). *Bud'  $A \in \mathbb{R}^{m \times n}$ . Pak matice  $X = A^\dagger$  je optimálním řešením úlohy*

$$\min\{\|I_m - AX\|_F; X \in \mathbb{R}^{n \times m}\}.$$

*Důkaz.* Lemma 5.24 aplikujeme na matice  $X := I_m - AA^\dagger$  a  $Y := A(A^\dagger - X)$ . Protože  $\mathcal{S}(Y) \subseteq \mathcal{S}(A) = \text{Ker}(A^T)^\perp$  a  $\mathcal{S}(I_m - AA^\dagger) \subseteq \text{Ker}(A^T)$  (neboť podle věty 4.4 je  $I_m - AA^\dagger$  maticí projekce do  $\text{Ker}(A^T)$ ), máme předpoklady lemmatu. Nyní

$$\begin{aligned} \|I_m - AX\|_F^2 &= \|I_m - AA^\dagger + AA^\dagger - AX\|_F^2 = \|I_m - AA^\dagger\|_F^2 + \|AA^\dagger - AX\|_F^2 \\ &\geq \|I_m - AA^\dagger\|_F^2. \end{aligned}$$

Tedy matice  $X = A^\dagger$  minimalizuje hodnotu  $\|I_m - AX\|_F$ . □



**Interpretace singulárních čísel.** Další velmi zajímavá vlastnost singulárních čísel je, že  $\sigma_i$  udává v 2-normě vzdálenost matice k nejbližší matici hodnosti nanejvýš  $i - 1$ . V důkazu následující věty je schované i to, jak tuto matici sestavit – ne náhodou je to matice z low-rank aproximace.

**Věta 5.26** (Interpretace singulárních čísel). *Buď  $A \in \mathbb{R}^{m \times n}$  se singulárními čísly  $\sigma_1, \dots, \sigma_r$ . Pak  $\sigma_i = \min \{\|A - B\|_2; B \in \mathbb{R}^{m \times n}, \text{rank}(B) \leq i - 1\}$  pro každé  $i = 1, \dots, r$ .*

*Důkaz.* Nerovnost „ $\geq$ “. Nechť  $A = U\Sigma V^T$  je SVD rozklad matice  $A$ . Definujme matici hodnosti  $i - 1$  předpisem  $B := U \text{diag}(\sigma_1, \dots, \sigma_{i-1}, 0, \dots, 0) V^T$ . Pak

$$\|A - B\|_2 = \|U \text{diag}(0, \dots, 0, \sigma_i, \dots, \sigma_n) V^T\|_2 = \|\text{diag}(0, \dots, 0, \sigma_i, \dots, \sigma_n)\|_2 = \sigma_i.$$

Nerovnost „ $\leq$ “. Buď  $B \in \mathbb{R}^{n \times n}$  libovolná matice hodnosti nanejvýš  $i - 1$  a ukážeme, že  $\|A - B\|_2 \geq \sigma_i$ . Nechť  $V_1$  sestává z prvních  $i$  sloupců matice  $V$ . Buď  $0 \neq z \in \text{Ker}(B) \cap \mathcal{S}(V_1)$ , to jest  $Bz = 0$ , a navíc normujeme  $z$  tak, aby  $\|z\|_2 = 1$ . Takový vektor existuje, protože  $\dim \text{Ker}(B) \geq n - i + 1$  a  $\dim \mathcal{S}(V_1) = i$ . Pak

$$\|A - B\|_2^2 = \max_{x: \|x\|_2=1} \|(A - B)x\|_2^2 \geq \|(A - B)z\|_2^2 = \|Az\|_2^2 = \|U\Sigma V^T z\|_2^2.$$

Protože  $z \in \mathcal{S}(V_1)$ , lze psát  $z = Vy$  pro nějaký vektor  $y = (y_1, \dots, y_i, 0, \dots, 0)^T$ , přičemž  $\|y\|_2 = \|V^T z\|_2 = \|z\|_2 = 1$ . Nyní

$$\|U\Sigma V^T z\|_2^2 = \|U\Sigma V^T Vy\|_2^2 = \|\Sigma y\|_2^2 = \sum_{j=1}^i \sigma_j^2 y_j^2 \geq \sum_{j=1}^i \sigma_i^2 y_j^2 = \sigma_i^2 \|y\|_2^2 = \sigma_i^2. \quad \square$$

Speciálně, nejmenší singulární číslo  $\sigma_n$  matice  $A \in \mathbb{R}^{n \times n}$  udává vzdálenost k nejbližší singulární matici. To znamená, že matice  $A + C$  je regulární pro všechny matice  $C \in \mathbb{R}^{n \times n}$  splňující  $\|C\|_2 < \sigma_n$ .

Věta platí analogicky i pro jiné ortogonálně invariantní normy, například Frobeniovu. Tudiž vzdálenost matice  $A$  k nejbližší singulární matici ve Frobeniově normě je též  $\sigma_n$ .

Další zajímavé souvislosti viz např. Prasolov [1994].

**Další.** Pomocí maticové normy můžeme formálně vyjádřit tvrzení typu, že libovolně blízko dané matice je matice s nějakou vlastností. Tohoto typu je i následující věta, říkájící, že množina diagonalizovatelných matic je hustá v prostoru všech matic, tj. ke každé matici najdeme libovolně blízko matici diagonalizovatelnou. Díky ekvivalenci norem je jedno, kterou použijeme.

**Věta 5.27.** *Pro každou matici  $A \in \mathbb{C}^{n \times n}$  a každé  $\varepsilon > 0$  existuje diagonalizovatelná matice  $A' \in \mathbb{C}^{n \times n}$  taková, že  $\|A - A'\| < \varepsilon$ .*

*Důkaz.* Podle Schurova rozkladu (věta 2.1) je  $A = UTU^*$ , kde  $U$  je unitární a  $T$  horní trojúhelníková. Zvolme  $A' = UT'U^*$ , kde  $T' = T + \text{diag}(\delta, \dots, \delta^n)$  a  $\delta > 0$  je dost malé. Pokud  $\delta < |t_{ii} - t_{jj}|/2$  pro všechna  $i \neq j$ , tak  $T'$  má na diagonále různá čísla, tedy  $A'$  má navzájem různá vlastní čísla a je diagonalizovatelná. Navíc

$$\|A - A'\|_2 = \|U \text{diag}(\delta, \dots, \delta^n) U^*\|_2 = \|\text{diag}(\delta, \dots, \delta^n)\|_2 = \delta,$$

čili pro spektrální normu stačí brát libovolné  $\delta < \min_{i \neq j} \{\varepsilon, \frac{1}{2}|t_{ii} - t_{jj}|\}$ . □

## Cvičení

- 5.1. Ukažte, že maticová norma je konvexní funkcí na prostoru matic  $\mathbb{R}^{m \times n}$ .
- 5.2. Buďte  $\|\cdot\|_a, \|\cdot\|_b$  dvě maticové normy na prostoru  $\mathbb{R}^{n \times n}$ . Rozhodněte, zda  $\|A\| := \|A\|_a + \|A\|_b$  je také maticovou normou.
- 5.3. Dokažte, že pro  $P$  matici projekce do netriviálního vlastního podprostoru platí  $\|P\|_2 = \|I - P\|_2$ .
- 5.4. Buď  $A \in \mathbb{R}^{m \times n}$  a  $B := (A \mid I_m)$ . Dokažte, že  $\|B\|_2^2 = 1 + \|A\|_2^2$ .
- 5.5. Buď  $A \in \mathbb{R}^{n \times n}$ . Dokažte  $\|A\|_2 \leq \|A\|_F \leq \sqrt{n}\|A\|_2$  a ukažte, že meze nelze zlepšit (tj., jsou těsné pro určité matice).

- 5.6. Buď  $A \in \mathbb{R}^{n \times n}$  regulární a  $\lambda$  její vlastní číslo. Dokažte  $\|A^{-1}\|_2^{-1} \leq |\lambda| \leq \|A\|_2$ .
- 5.7. Ukažte, že pokud  $A \in \mathbb{R}^{n \times n}$  je singulární, pak  $\|I_n - A\| \geq 1$  pro každou maticovou normu.
- 5.8. Rozhodněte, pro které ze zmiňovaných norem platí  $\|A\| = \|A^T\|$ .
- 5.9. Rozhodněte, pro které ze zmiňovaných norem platí  $\left\| \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \right\| = \max\{\|A\|, \|B\|\}$ .
- 5.10. Ukažte, že  $\|A^T A\|_2 = \|A\|_2^2$ , ale obecně  $\|A^2\|_2 \neq \|A\|_2^2$ .
- 5.11. Dokažte, že  $\|I_n\| = 1$  pro každou indukovanou maticovou normu. Dále spočítejte  $\|I_n\|_F$ ,  $\|I_n\|_{\ell_1}$  a  $\|I_n\|_{\ell_\infty}$ , což prokáže, že tyto normy nejsou indukované žádnou vektorovou normou.
- 5.12. Buď  $A \in \mathbb{R}^{n \times n}$ . Dokažte, že ze všech symetrických matic je matice  $\frac{1}{2}(A + A^T)$  nejblíže k matici  $A$  ve Frobeniově a maticové 2-normě.
- 5.13. Buď matice  $A$  tvaru  $A = \alpha I_n - N$ , kde  $N \geq 0$  a  $\rho(N) < \alpha$ .
- Dokažte  $A^{-1} \geq 0$  pro případ  $\alpha = 1$ .
  - Dokažte  $A^{-1} \geq 0$  pro případ  $\alpha > 0$ .
  - Dokažte, že existuje  $x > 0$  takové, že  $Ax > 0$ .
  - Dokažte, že reálné části vlastních čísel matice  $A$  jsou kladné.
  - Dokažte  $\det(A) > 0$ .

# Kapitola 6

## Číslo podmíněnosti

Číslo podmíněnosti je charakteristika čtvercové matice a zhruba řečeno udává, jak nepřesně se s touto maticí numericky pracuje, to jest, jak velké numerické chyby způsobuje. Nejmenší číslo podmíněnosti je 1 a mají ho dobře podmíněné matice (typicky ortogonální matice). Čím větší je číslo podmíněnosti, tím hůře podmíněná matice je (typicky je blízko k singularní matici).

Číslo podmíněnosti bylo zavedeno Turingem roku 1948 pro Frobeniovu normu, ale podobný koncept už používali von Neumann a Goldstine roku 1947.

**Definice 6.1.** Číslo podmíněnosti regulární matice  $A \in \mathbb{R}^{n \times n}$  je hodnota  $k(A) = \|A\| \cdot \|A^{-1}\|$ , kde  $\|\cdot\|$  je indukovaná maticová norma. Speciálně pro  $p$ -normu pak značíme  $k_p(A) = \|A\|_p \cdot \|A^{-1}\|_p$ .

Přestože má definice smysl i pro neindukovanou maticovou normu, indukovaná norma se používá hlavně kvůli vlastnosti z tvrzení 6.3.

**Lemma 6.2.** Pro indukovanou maticovou normu platí:  $\|I_n\| = 1$ .

*Důkaz.* Podle definice  $\|I_n\| = \max_{x: \|x\|=1} \|I_n x\| = \max_{x: \|x\|=1} \|x\| = 1$ . □

Výše zmíněná vlastnost neplatí pro všechny maticové normy, například  $\|I_n\|_{\ell_1} = n$ .

**Tvrzení 6.3.** Platí:  $k(A) \geq 1$ .

*Důkaz.* Podle definice a lematu 6.2 je  $k(A) = \|A\| \cdot \|A^{-1}\| \geq \|AA^{-1}\| = \|I_n\| = 1$ . □

Shrneme některé základní vlastnosti čísla podmíněnosti.

**Tvrzení 6.4.** Platí:  $k(A) \geq \frac{|\lambda_{\max}(A)|}{|\lambda_{\min}(A)|}$ . Rovnost nastane mj. je-li  $A$  symetrická a  $k(A) = k_2(A)$ .

*Důkaz.* Podle věty 5.13 máme  $k(A) = \|A\| \cdot \|A^{-1}\| \geq \rho(A)\rho(A^{-1}) = |\lambda_{\max}(A)|/|\lambda_{\min}(A)|$ .

Je-li  $A$  symetrická a použijeme-li spektrální normu, pak  $k_2(A) = \|A\|_2 \cdot \|A^{-1}\|_2 = \rho(A)\rho(A^{-1}) = |\lambda_{\max}(A)|/|\lambda_{\min}(A)|$ . □

**Tvrzení 6.5.** Platí:

- (1)  $k(AB) \leq k(A)k(B)$ ,
- (2)  $k(\alpha A) = k(A)$ ,  $\alpha \neq 0$ .

*Důkaz.*

(1)  $k(AB) = \|AB\| \cdot \|B^{-1}A^{-1}\| \leq \|A\| \cdot \|B\| \cdot \|B^{-1}\| \cdot \|A^{-1}\| = k(A)k(B)$ .

(2)  $k(\alpha A) = \|\alpha A\| \cdot \|\alpha^{-1}A^{-1}\| = |\alpha| \cdot |\alpha^{-1}| \cdot \|A\| \cdot \|A^{-1}\| = k(A)$ . □

Podle tvrzení tedy číslo podmíněnosti nezávisí na škálování. Na druhou stranu, při násobení matic se zvětšuje, speciálně při umocňování matice se číslo podmíněnosti také umocňuje.

**Poznámka 6.6.** Číslo podmíněnosti slouží spíše k odvození teoretických vlastností. Málokdy se počítá, protože při výpočtu  $A^{-1}$  dochází k numerickým chybám, ale spíše se odhaduje.

Převrácená hodnota čísla podmíněnosti udává relativní vzdálenost k singulární matici v příslušné maticové normě (Gastinel & Kahan 1966). Proto matice s velkým číslem podmíněnosti se chovají numericky hůře.

**Poznámka 6.7.** Matice, které se běžně vyskytují v různých situacích, nejsou náhodné, a proto těžko lze určit distribuci jejich čísel podmíněnosti. Na druhou stranu je zajímavý výsledek [Demmel, 1988], který ukazuje, že náhodně zvolená matice (při rovnoměrném rozdělení) řádu  $n$  má číslo podmíněnosti (přesně řečeno, určitý jeho typ) hodnotu aspoň  $h$  s pravděpodobností řádově nanejvýš  $n(n^2 - 1)h^{-2}$ . Tedy pravděpodobnost velkého čísla podmíněnosti klesá kvadraticky s jeho převrácenou hodnotou.

## 6.1 Číslo podmíněnosti za spektrální normy

Spektrální norma je nejčastěji používanou normou pro číslo podmíněnosti, a to kvůli mnohým pěkným vlastnostem. Díky ekvivalenci maticových norem se tyto vlastnosti nepřímo přenáší i na  $k(A)$  s jinými normami.

**Tvrzení 6.8.** Platí:

- (1)  $k_2(A) = \sigma_1/\sigma_n$ ,
- (2)  $k_2(A) = k_2(A^T)$ ,
- (3)  $k_2(A^T A) = k_2(A)^2$ ,
- (4)  $k_2(A) = 1$  právě tehdy, když  $A$  je nenulový násobek ortogonální matice.

*Důkaz.* Nechť  $A = U\Sigma V^T$  je SVD rozklad matice  $A$ .

- (1) Platí  $k_2(A) = \|A\|_2 \cdot \|A^{-1}\|_2 = \sigma_1/\sigma_n$ .
- (2) Plyne z předchozího, neboť obě matice  $A$ ,  $A^T$  mají shodná singulární čísla.
- (3) Protože  $A^T A = (U\Sigma V^T)^T(U\Sigma V^T) = V\Sigma^2 V^T$ , tak singulární čísla matice  $A^T A$  jsou  $\sigma_1^2, \dots, \sigma_n^2$ .
- (4) „ $\Leftarrow$ “ Plyne přímo z tvrzení 6.5 a faktu, že ortogonální matice má všechny singulární čísla rovné 1.  
„ $\Rightarrow$ “ Protože  $k_2(A) = 1$ , musí být všechny singulární čísla matice  $A$  shodná, tedy  $\sigma_1 = \sigma_n = c > 0$ .  
Tudíž  $A = U\Sigma V^T = U(cI_n)V^T = c(UV^T)$ , kde  $UV^T$  je ortogonální.  $\square$

**Poznámka 6.9** (Geometrická interpretace I.). Geometrická interpretace čísla podmíněnosti je, že  $k_2(A) = \cot(\varphi/2)$ , kde  $\varphi$  je minimální úhel mezi  $Ax$  a  $Ay$ , kde  $x, y \in \mathbb{R}^n \setminus \{0\}$  probíhá všechny vektory takové, že  $x \perp y$ . Tudíž, pokud je  $A$  ortogonální, minimální úhel je  $90^\circ$  a  $k_2(A) = 1$ . Naopak, čím menší úhel jest, tím větší hodnotu dostaneme.

**Poznámka 6.10** (Geometrická interpretace II.). Buď  $A \in \mathbb{R}^{n \times n}$  regulární a studujme obraz jednotkové koule při zobrazení  $x \mapsto Ax$ . SVD rozklad  $A = U\Sigma V^T$  ukazuje, že lineární zobrazení lze rozložit na složení tří základních zobrazení: ortogonální zobrazení s maticí  $V^T$ , škálování podle  $\Sigma$  a ortogonální zobrazení s  $U$ . Konkrétně, zobrazení s maticí  $V^T$  zobrazí kouli na sebe sama,  $\Sigma$  ji zdeformuje na elipsoid a  $U$  ji otočí/převrátí. Tedy výsledkem bude elipsoid se středem v počátku, poloosy jsou ve směrech sloupců  $U$  a délky mají  $\sigma_1, \dots, \sigma_n$ .

Hodnota  $k_2(A) = \frac{\sigma_1}{\sigma_n} \geq 1$  je pak jakousi mírou deformace koule a kvantitativně udává, jak moc zobrazení deformuje geometrické útvary. Je-li hodnota rovna 1, elipsoid bude mít tvar koule, a naopak čím větší bude hodnota, tím protáhlejší bude elipsoid.

**Poznámka 6.11** (Numerická interpretace). Empirické pravidlo říká, že je-li číslo podmíněnosti řádově  $10^k$ , pak při výpočtech s maticí (inverze, řešení soustav, atp.) ztrácíme přesnost o  $k$  desetinných míst. Ortogonální matice mají číslo podmíněnosti rovné 1, a proto se v numerické matematice často používají.

Naproti tomu jsou matice velmi špatně podmíněné, například Hilbertovy matice. Hilbertova matice  $H_n$  řádu  $n$  je definovaná  $(H_n)_{ij} = \frac{1}{i+j-1}$ ,  $i, j = 1, \dots, n$ . Přestože  $H_n$  je regulární pro všechna  $n$ , číslo podmíněnosti má velmi vysoké:

$n$	číslo podmíněnosti $H_n$
3	$\approx 500$
5	$\approx 10^5$
10	$\approx 10^{13}$
15	$\approx 10^{17}$

## 6.2 Číslo podmíněnosti při řešení soustav lineárních rovnic

Uvažujme soustavu lineárních rovnic

$$Ax = b.$$

Nechť  $\tilde{x}$  je numericky spočítané řešení a zajímá nás, jak moc přesný výsledek jsme dostali. První nápad by mohl být spočítat tzv. *residuum*

$$r := b - A\tilde{x}.$$

Dalo by se čekat, že pro malé  $r$  budeme mít přesnější výsledek a pro velké  $r$  méně přesný. Takto jednoduché to však není, závisí právě ještě na čísle podmíněnosti matice  $A$ .

**Věta 6.12.** Označme  $x := A^{-1}b$  a necht' platí  $b \neq 0$ . Pak:

$$k(A)^{-1} \frac{\|r\|}{\|b\|} \leq \frac{\|x - \tilde{x}\|}{\|x\|} \leq k(A) \frac{\|r\|}{\|b\|}.$$

(Používáme libovolnou vektorovou normu a pro  $k(A)$  příslušnou indukovanou normu.)

*Důkaz.* (1) Použijeme odhady

$$x - \tilde{x} = A^{-1}b - \tilde{x} = A^{-1}(b - A\tilde{x}) = A^{-1}r,$$

a

$$\|b\| = \|Ax\| \leq \|A\| \cdot \|x\|.$$

Nyní máme

$$\|x - \tilde{x}\| = \|A^{-1}r\| \leq \|A^{-1}\| \cdot \|r\| \leq \frac{\|A\| \cdot \|x\|}{\|b\|} \|A^{-1}\| \cdot \|r\| = k(A) \cdot \|x\| \frac{\|r\|}{\|b\|},$$

z čehož plyne horní mez.

(2) Pišme

$$\begin{aligned} \|r\| &= \|b - A\tilde{x}\| = \|A(x - \tilde{x})\| \leq \|A\| \cdot \|x - \tilde{x}\| \frac{\|x\|}{\|x\|} = \|A\| \cdot \|x - \tilde{x}\| \frac{\|A^{-1}b\|}{\|x\|} \\ &\leq \|A\| \cdot \|x - \tilde{x}\| \frac{\|A^{-1}\| \cdot \|b\|}{\|x\|} = k(A) \cdot \|x - \tilde{x}\| \frac{\|b\|}{\|x\|} \end{aligned}$$

z čehož plyne dolní mez. □

Důležitější je horní mez, která říká, že  $\tilde{x}$  je dobrou aproximací pokud je residuum  $r$  a rovněž číslo podmíněnosti  $k(A)$  malé. Pokud je matice špatně podmíněná, tak i při malém residuu může být  $\tilde{x}$  daleko od skutečného řešení  $x$ .

Pokud  $k(A) = 1$ , pak podle věty máme  $\frac{\|x - \tilde{x}\|}{\|x\|} = \frac{\|r\|}{\|b\|}$ , čili znormované residuum udává přesnou relativní vzdálenost  $\tilde{x}$  od skutečného řešení  $\|x\|$ .

### 6.3 Číslo podmíněnosti při počítání vlastních čísel

Bud'  $\tilde{\lambda}$  odhad vlastního čísla a  $\tilde{x}$ ,  $\|\tilde{x}\| = 1$  odhad vlastního vektoru matice  $A \in \mathbb{R}^{n \times n}$ . Jak moc dobré odhady máme? Opět zkusme zkoumat kvalitu numericky spočítaných odhadů pomocí residua

$$r := A\tilde{x} - \tilde{\lambda}\tilde{x}.$$

I zde bychom očekávali, že pro malé  $r$  bude aproximace dobrá a při velkém  $r$  horší. A analogicky jako v předchozí sekci závisí na čísle podmíněnosti matice  $A$ .

V následující větě uvažujeme vektorovou  $\ell_p$ -normu a pro  $k_p(\cdot)$  jí indukovanou maticovou normu. Potřebujeme ji mj. proto, aby byla splněna absolutní monotonie maticové normy, tj. vlastnost

$$|A| \leq B \Rightarrow \|A\| \leq \|B\|.$$

Tato vlastnost není splněna pro každou maticovou normu, neplatí např. pro nukleární normu.

**Věta 6.13** (Bauer–Fike, 1960). *Bud'  $A \in \mathbb{R}^{n \times n}$  diagonalizovatelná, a bud'  $A = S\Lambda S^{-1}$  spektrální rozklad. Pak existuje vlastní číslo  $\lambda$  matice  $A$  takové, že*

$$|\lambda - \tilde{\lambda}| \leq k_p(S) \cdot \|r\|_{\ell_p}.$$

(Používáme libovolnou vektorovou normu a pro  $k(A)$  příslušnou indukovanou normu.)

*Důkaz.* Protože

$$r = A\tilde{x} - \tilde{\lambda}\tilde{x} = (S\Lambda S^{-1} - \tilde{\lambda}I_n)\tilde{x} = S(\Lambda - \tilde{\lambda}I_n)S^{-1}\tilde{x},$$

tak odvodíme  $\tilde{x} = S(\Lambda - \tilde{\lambda}I_n)^{-1}S^{-1}r$ . Bud'  $\lambda$  vlastní číslo  $A$ , které je nejbližší k  $\tilde{\lambda}$ . Nyní

$$\begin{aligned} 1 = \|\tilde{x}\| &= \|S(\Lambda - \tilde{\lambda}I_n)^{-1}S^{-1}r\| \leq \|S\| \cdot \|(\Lambda - \tilde{\lambda}I_n)^{-1}\| \cdot \|S^{-1}\| \cdot \|r\| \\ &= k_p(S) \cdot \|r\| \cdot \|(\Lambda - \tilde{\lambda}I_n)^{-1}\| \leq k_p(S) \cdot \|r\| \cdot \|(|\lambda - \tilde{\lambda}| \cdot I_n)^{-1}\| = k_p(S) \cdot \|r\| \cdot |\lambda - \tilde{\lambda}|^{-1}, \end{aligned}$$

z čehož už snadno odvodíme odhad pro  $|\lambda - \tilde{\lambda}|$ . □

Pro symetrickou matici  $A$  je matice  $S$  ortogonální, a proto  $k_2(S) = 1$ . Při použití eukleidovské normy má tudíž věta 6.13 jednodušší tvar.

**Důsledek 6.14.** *Je-li  $A \in \mathbb{R}^{n \times n}$  symetrická, pak*

$$|\lambda - \tilde{\lambda}| \leq \|r\|_2.$$

#### Cvičení

6.1. Ukažte

- (a)  $k(A) = k(A^{-1})$ ,
- (b)  $k_2(A) = k_2(UAV)$  pro  $U, V$  ortogonální.

6.2. Spočítejte  $k_2(A)$  pro  $A = I_n + 123xx^T$ , kde  $x \in \mathbb{R}^n$  je libovolný vektor normy 1.

6.3. Spočítejte číslo podmíněnosti ortogonální matice ve Frobeniově normě.

6.4. Spočítejte  $k_p(D)$  pro diagonální matici  $D$ .

6.5. Určete, jaký je vztah mezi  $k_1(A)$  a  $k_2(A)$ . Pro které matice nastává rovnost?

6.6. Odhadněte chybu při výpočtu  $A^{-1}$  analogicky jako ve větě 6.12.

# Kapitola 7

## Perturbace

V této kapitole si klademe otázky, jak se změní řešení daného problému (soustav lineárních rovnic, vlastní čísla matice apod.) když vstupní hodnoty trochu změníme.

U čísla podmíněnosti nás zajímaly numerické vlastnosti matice. Nyní se více zaměříme jak moc se řešení může změnit, a to jak pro diskrétní změnu, tak pro infinitesimální změny.

### 7.1 Perturbace vlastních čísel

Pro perturbaci vlastních čísel [Horn and Johnson, 1985] obecné matice můžeme použít následující verzi Bauerovy–Fikeovy věty.

**Věta 7.1** (Bauer–Fike, 1960). *Bud'  $A \in \mathbb{R}^{n \times n}$  diagonalizovatelná, a bud'  $A = SAS^{-1}$  spektrální rozklad. Nechť  $B \in \mathbb{R}^{n \times n}$  je perturbační matice. Pak pro každé vlastní číslo  $\mu$  matice  $A + B$  existuje vlastní číslo  $\lambda$  matice  $A$  takové, že*

$$|\lambda - \mu| \leq k_p(S) \cdot \|B\|_p.$$

*Důkaz.* Použijeme větu 6.13 pro  $\tilde{\lambda} = \mu$  a  $\tilde{x}$  odpovídající vlastní vektor matice  $A + B$  normy 1. Dostaneme

$$\begin{aligned} |\lambda - \mu| &\leq k_p(S) \cdot \|r\|_p = k_p(S) \cdot \|A\tilde{x} - \mu\tilde{x}\|_p \\ &= k_p(S) \cdot \|A\tilde{x} - (A + B)\tilde{x}\|_p = k_p(S) \cdot \|B\tilde{x}\|_p \leq k_p(S) \cdot \|B\|_p, \end{aligned}$$

kde poslední nerovnost vyplývá z definice indukované maticové  $p$ -normy. □

Je-li  $A$  symetrická, pak  $S$  z věty 7.1 je ortogonální, a tedy  $k_2(S) = 1$ . Čili máme odhad ve tvaru

$$|\lambda - \mu| \leq \|B\|_2. \tag{7.1}$$

Pro symetrické matice nicméně odvodíme silnější výsledky. Umíme totiž pro ně porovnat jednotlivá vlastní čísla, tedy  $i$ -té vlastní číslo matice  $A$  a matice  $A + B$ . Tuto interpretaci Bauerova–Fikeova věta nedává!

Pro symetrickou matici značíme vlastní čísla  $\lambda_1 \geq \dots \geq \lambda_n$ .

**Věta 7.2** (Weyl, 1912, speciální verze). *Bud'  $A, B \in \mathbb{R}^{n \times n}$  symetrické. Pak pro každé  $i = 1, \dots, n$*

$$\lambda_i(A) + \lambda_n(B) \leq \lambda_i(A + B) \leq \lambda_i(A) + \lambda_1(B). \tag{7.2}$$

*Důkaz.* Bud'  $i \in \{1, \dots, n\}$  pevné a označme  $C := A + B$ . Bud'  $a_1, \dots, a_n$ , resp.  $c_1, \dots, c_n$  ortonormální báze vlastních vektorů matice  $A$  resp.  $C$ , odpovídající vlastním číslům seříděným sestupně. Definujme prostory

$$U = \text{span}\{a_1, \dots, a_i\}, \quad V = \text{span}\{c_i, \dots, c_n\}.$$

Podle věty o dimenzi spojení a průniku prostorů máme

$$\dim U \cap V = \dim U + \dim V - \dim(U + V) \geq i + (n - i + 1) - n = 1.$$

Tedy existuje  $x \in U \cap V$ ,  $\|x\|_2 = 1$ . S využitím  $x \in V$  máme<sup>1)</sup>

$$\begin{aligned} x^T C x &= (\sum_{k=i}^n \gamma_k c_k)^T C (\sum_{\ell=i}^n \gamma_\ell c_\ell) = \sum_{k,\ell=i}^n \gamma_k \gamma_\ell c_k^T C c_\ell \\ &= \sum_{k,\ell=i}^n \gamma_k \gamma_\ell c_k^T \lambda_\ell(C) c_\ell = \sum_{k=i}^n \gamma_k^2 \lambda_k(C) \leq \sum_{k=i}^n \gamma_k^2 \lambda_i(C) = \lambda_i(C), \end{aligned}$$

neboť  $1 = \|x\|_2^2 = \sum_{k,\ell=i}^n \gamma_k \gamma_\ell c_k^T c_\ell = \sum_{k=i}^n \gamma_k^2$ . Analogicky díky  $x \in U$  máme

$$x^T C x = x^T (A + B) x = x^T A x + x^T B x \geq \lambda_i(A) + \lambda_n(B).$$

Dohromady to dává dolní odhad v (7.2). Druhá nerovnost se dokáže analogicky.  $\square$

**Věta 7.3** (Weyl, 1912, obecná verze). *Bud'  $A, B \in \mathbb{R}^{n \times n}$  symetrické. Pak pro každé  $i = 1, \dots, n$*

$$\lambda_i(A + B) \leq \lambda_j(A) + \lambda_{i-j+1}(B), \quad \forall j \leq i, \quad (7.3)$$

$$\lambda_i(A + B) \geq \lambda_j(A) + \lambda_{i-j+n}(B), \quad \forall j \geq i. \quad (7.4)$$

*Důkaz.* Bud'  $i \in \{1, \dots, n\}$  pevné. Bud'  $a_1, \dots, a_n$ , resp.  $b_1, \dots, b_n$ , resp.  $c_1, \dots, c_n$  ortonormální báze vlastních vektorů matice  $A$ , resp.  $B$ , resp.  $C := A + B$ . Definujme prostory

$$V_1 = \text{span}\{a_1, \dots, a_j\}, \quad V_2 = \text{span}\{b_1, \dots, b_{i-j+n}\}, \quad V_3 = \text{span}\{c_i, \dots, c_n\}.$$

Podle věty o dimenzi spojení a průniku prostorů máme

$$\begin{aligned} \dim V_1 \cap (V_2 \cap V_3) &= \dim V_1 + \dim V_2 \cap V_3 - \dim(V_1 + V_2 \cap V_3) \\ &= \dim V_1 + \dim V_2 + \dim V_3 - \dim(V_2 + V_3) - \dim(V_1 + V_2 \cap V_3) \\ &\geq j + (i - j + n) + (n - i + 1) - n - n = 1. \end{aligned}$$

Tedy existuje  $x \in V_1 \cap V_2 \cap V_3$ ,  $\|x\|_2 = 1$ . S využitím  $x \in V_3$  máme

$$\begin{aligned} x^T C x &= (\sum_{k=i}^n \gamma_k c_k)^T C (\sum_{\ell=i}^n \gamma_\ell c_\ell) = \sum_{k,\ell=i}^n \gamma_k \gamma_\ell c_k^T C c_\ell \\ &= \sum_{k,\ell=i}^n \gamma_k \gamma_\ell c_k^T \lambda_\ell(C) c_\ell = \sum_{k=i}^n \gamma_k^2 \lambda_k(C) \leq \sum_{k=i}^n \gamma_k^2 \lambda_i(C) = \lambda_i(C), \end{aligned}$$

neboť  $1 = \|x\|_2^2 = \sum_{k,\ell=i}^n \gamma_k \gamma_\ell c_k^T c_\ell = \sum_{k=i}^n \gamma_k^2$ . Analogicky díky  $x \in V_1 \cap V_2$  máme

$$x^T C x = x^T (A + B) x = x^T A x + x^T B x \geq \lambda_j(A) + \lambda_{i-j+n}(B).$$

Dohromady to dává nerovnost (7.4). Druhá nerovnost se dokáže analogicky.  $\square$

Věta tedy dává poměrně těsné perturbační odhady na všechna vlastní čísla. Jako důsledek máme i celkový odhad na změnu spektrálního poloměru.

**Důsledek 7.4.** *Bud'  $A, B \in \mathbb{R}^{n \times n}$  symetrické. Pak  $\rho(A + B) \leq \rho(A) + \rho(B)$ .*

*Důkaz.* Podle Weylové věty

$$\begin{aligned} \lambda_1(A + B) &\leq \lambda_1(A) + \lambda_1(B) \leq \rho(A) + \rho(B), \\ \lambda_n(A + B) &\geq \lambda_n(A) + \lambda_n(B) \leq -\rho(A) - \rho(B), \end{aligned}$$

Tedy  $\rho(A + B) = \max\{\lambda_1(A + B), -\lambda_n(A + B)\} \leq \rho(A) + \rho(B)$ .  $\square$

Takovýto odhad pro obecné matice neplatí. Například matice

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

mají vlastní čísla nulová, a tedy  $\rho(A) = \rho(B) = 0$ . Nicméně  $A + B$  má vlastní čísla  $\pm 1$ , a tedy  $\rho(A + B) = 1$ .

Jako další důsledek Weylové věty máme sice trochu hrubší, ale stejnoměrný odhad na velikost změny vlastních čísel.

<sup>1)</sup>Všimněte si analogie s Rayleighovou–Ritzovou větou 2.4.



**Důsledek 7.5.** *Budte  $A, B \in \mathbb{R}^{n \times n}$  symetrické. Pak pro každé  $i = 1, \dots, n$*

$$|\lambda_i(A + B) - \lambda_i(A)| \leq \rho(B).$$

*Důkaz.* Podle Weylovy věty máme z (7.2):

$$\begin{aligned} \lambda_i(A + B) &\leq \lambda_i(A) + \lambda_1(B) \leq \lambda_i(A) + \rho(B), \\ \lambda_i(A + B) &\geq \lambda_i(A) + \lambda_n(B) \leq \lambda_i(A) - \rho(B). \end{aligned}$$

□

Navíc podle věty 5.13 můžeme k odhadu použít i maticové normy:

$$|\lambda_i(A + B) - \lambda_i(A)| \leq \|B\|.$$

Speciálně pro spektrální normu dostaneme nerovnost (7.1), ale opět tento výsledek je silnější, protože narozdíl od (7.1) srovnává  $i$ -tá vlastní čísla obou matic.

## 7.2 Spojitost vlastních čísel

Nyní přejdeme od diskretní perturbace reprezentované aditivní maticí  $B$  k infinitesimálně malým perturbacím.

**Věta 7.6.** *Vlastní čísla matice  $A \in \mathbb{C}^{n \times n}$  jsou spojité funkce vzhledem ke složkám matice  $A$ .*

*Důkaz.* Vlastní čísla jsou kořeny charakteristického polynomu, a ty jsou spojitými funkcemi vzhledem ke koeficientům polynomu, viz Horn and Johnson [1985]; Meyer [2000] nebo [Mayer, 2017, Thm.1.4.1]. □

Vlastní vektory obecně spojité nejsou, například matice

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

má pro  $\alpha = 0$  vlastní číslo 0 a vlastní vektor jen  $(1, 0)^T$ . Ale pro  $\alpha > 0$  má navíc vlastní číslo  $\alpha$  a odpovídající vlastní vektor  $(1, \alpha)^T$ , čili vlastní vektory jsou dva.

Za celkem obecných předpokladů ale spjitost vlastních vektorů platí. Musíme ale ještě obejít jeden problém, a to, že vlastní vektory nejsou jednoznačné.

**Lemma 7.7.** *Bud'  $A \in \mathbb{C}^{n \times n}$  regulární a bud'  $b \in \mathbb{C}^n$ . Pak existuje okolí  $A$  a  $b$ , na kterém je řešení soustavy  $Ax = b$  spojitou funkcí.*

*Důkaz.* Je-li  $A$  regulární, pak je i na nějakém okolí, protože  $\det(A) \neq 0$  zůstane v platnosti i pro dost malou perturbaci. Nyní můžeme vyjádřit všechny složky řešení soustavy  $Ax = b$  podle Cramerova pravidla jako podíl dvou determinantů. To jsou spojité funkce, tedy i výsledek je spojitý. □

**Věta 7.8.** *Nechť matice  $A \in \mathbb{C}^{n \times n}$  má navzájem různá vlastní čísla. Pak lze vybrat vlastní vektory tak, že na okolí matice  $A$  jsou spojitými funkcemi vzhledem ke složkám matice  $A$ .*

*Důkaz.* Bud'  $\lambda$  vlastní číslo matice  $A$  a  $x$  odpovídající vlastní vektor. Pak  $x \in \text{Ker}(A - \lambda I_n)$  a matice  $A - \lambda I_n$  má hodnost  $n - 1$ . Protože  $x \neq 0$ , můžeme předpokládat bez újmy na obecnost, že má tvar  $x = (y^T, 1)^T$ . Pak rovnice  $(A - \lambda I_n)x = 0$  se dá vyjádřit jako  $By = b$ . To je soustava  $n$  rovnic o  $n - 1$  neznámých, čili redundantní rovnici odstraníme. Výsledná rovnice  $\tilde{B}y = \tilde{b}$  obsahuje regulární matici  $\tilde{B}$ . Podle lemmatu 7.7 pro dost malou perturbaci bude mít soustava spojitě řešení, a to odpovídá vlastnímu vektoru matice  $A$ . □

I když jsou vlastní vektory za výše uvedených předpokladů spojitě, pořád se můžou chovat velmi chaoticky. Uvažujme například

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} \alpha & \beta \\ 0 & 0 \end{pmatrix}.$$

Pak vlastní čísla  $A+B$  jsou  $1+\alpha$  a  $1$ . Prvnímu odpovídá vlastní vektor  $(1,0)^T$ , druhému  $\frac{1}{\sqrt{\alpha^2+\beta^2}}(-\beta, \alpha)^T$ . Tudíž i pro libovolně malou perturbaci  $B$  mohou mít vlastní vektory libovolný směr. Navíc, pokud  $\beta \rightarrow 0$  nekonečně mnohokrát změni znaménko a  $\alpha \rightarrow 0$ , pak vlastní vektor nekonečně mnohokrát změni diametrálně směr.

Spojitosť vlastních čísel je užitečná pro řadu věcí. Ukážeme si použití ve větě o Gerschgorinových discích. Ta říká, že každé vlastní číslo  $\lambda$  matice  $A \in \mathbb{C}^{n \times n}$  leží v kruhu o středu  $a_{ii}$  a poloměru  $\sum_{j \neq i} |a_{ij}|$  pro nějaké  $i \in \{1, \dots, n\}$ . Navíc k tomu lze ještě říci toto:

**Důsledek 7.9.** *V každé topologické komponentě souvislosti Gerschgorinových disků je tolik vlastních čísel, z kolika disků daná komponenta vznikla.*

*Důkaz.* Rozdělme  $A = D + N$  na diagonální a mimodiagonální část, tj.

$$D = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & a_{nn} \end{pmatrix}, \quad N = \begin{pmatrix} 0 & a_{12} & \dots & a_{1n} \\ a_{21} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_{n-1,n} \\ a_{n1} & \dots & a_{n,n-1} & 0 \end{pmatrix}.$$

Definujme matici  $A(t) = D + tN$ , kde  $t \in [0, 1]$  je proměnná. Matice  $A(0) = D$  má vlastní čísla  $a_{11}, \dots, a_{nn}$  a matice  $A(1) = A$  má vlastní čísla jako matice  $A$ . Protože se vlastní čísla mění spojitě, tak při průběhu  $t$  od 0 do 1 nemůže žádné vlastní číslo přeskočit do jiného, disjunktního, disku.  $\square$

### 7.3 Maticová derivace

Uvažujme matici  $A \in \mathbb{R}^{n \times n}$  závislou na parametru  $t \in \mathbb{R}$ . Každý prvek matice tak je reálnou funkcí v  $t$ , to jest  $a_{ij}(t): \mathbb{R} \rightarrow \mathbb{R}$ . Pak derivací matice  $A$  rozumíme matici slouženou z derivací jednotlivých složek matice  $A$ , tedy

$$A' = A(t)' = \begin{pmatrix} a_{11}(t)' & \dots & a_{1n}(t)' \\ \vdots & & \vdots \\ a_{n1}(t)' & \dots & a_{nn}(t)' \end{pmatrix}.$$

Připomeňme dále, že jakobián vektorové funkce  $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$  je definovaný

$$\frac{\partial f}{\partial x} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_n}{\partial x_1} \\ \vdots & & \vdots \\ \frac{\partial f_1}{\partial x_n} & \dots & \frac{\partial f_n}{\partial x_n} \end{pmatrix},$$

tedy ve sloupcích má gradienty jednotlivých funkcí  $f_1, \dots, f_n$ .

Ukážeme si několik základních pravidel maticové derivace.

**Tvrzení 7.10.** *Pro dvě parametrické matice platí  $(AB)' = A'B + AB'$ .*

*Důkaz.*  $(AB)'_{ij} = (\sum_k a_{ik}b_{kj})' = \sum_k (a_{ik}b_{kj})' = \sum_k a'_{ik}b_{kj} + \sum_k a_{ik}b'_{kj} = (A'B)_{ij} + (AB')_{ij}$ .  $\square$

**Tvrzení 7.11.** *Pro parametrickou regulární matici platí  $(A^{-1})' = -A^{-1}A'A^{-1}$ .*

*Důkaz.* Podle tvrzení 7.10 zderivujeme

$$0 = I'_n = (AA^{-1})' = A'A^{-1} + A(A^{-1})'$$

z čehož  $A(A^{-1})' = -A'A^{-1}$  a vynásobením zleva  $A^{-1}$  máme tvrzení.  $\square$

## Derivace řešení soustav lineárních rovnic

Mějme soustavu  $Ax = b$  s regulární maticí, tedy řešení je  $x = A^{-1}b$ . Uvažujme tři případy:

1. „ $A$  je pevná a  $b$  parametrická“. Pak zderivováním rovnice  $x = A^{-1}b$  dostaneme

$$x' = (A^{-1}b)' = A^{-1}b'.$$

Speciálně, pokud měníme pouze prvek  $b_k$ , to jest  $b_k(t) = b_k + t$ , tak

$$x' = A^{-1}b' = A^{-1}e_k = A_{*k}^{-1}.$$

Souhrnně pro všechna  $k$  to můžeme vyjádřit jako jakobián  $\frac{\partial x}{\partial b} = A^{-T}$ .

2. „ $A$  je parametrická a  $b$  pevná“. Pak zderivováním rovnice  $x = A^{-1}b$  dostaneme

$$x' = (A^{-1}b)' = (A^{-1})'b = -A^{-1}A'A^{-1}b = -A^{-1}A'x.$$

Speciálně, pokud měníme pouze prvek  $a_{ij}$ , to jest  $a_{ij}(t) = a_{ij} + t$ , tak

$$x' = -A^{-1}A'x = -A^{-1}e_i e_j^T x = -A_{*i}^{-1}x_j.$$

3. „ $A$  i  $b$  je parametrická“. Pak zderivováním dostaneme

$$x' = (A^{-1}b)' = (A^{-1})'b + A^{-1}b' = -A^{-1}A'A^{-1}b + A^{-1}b' = -A^{-1}A'x + A^{-1}b'.$$

Vidíme, že ve vzorcích pro derivaci se vždy vyskytuje matice  $A^{-1}$ . To samozřejmě není náhoda a souvisí to i s tím, že inverzní matice se vyskytuje v definici 6.1 čísla podmíněnosti.

**Příklad 7.12.** Uvažujme parametrickou soustavu

$$\left( \begin{array}{cc|c} 1+t & 2+t & 1 \\ 3+t & 4+t & 5 \end{array} \right)$$

Pro  $t = 0$  má soustava řešení  $x = (3, -1)^T$  a inverze matice soustavy je  $A^{-1} = \frac{1}{2} \begin{pmatrix} -4 & 2 \\ 3 & -1 \end{pmatrix}$ . Z toho spočítáme  $x' = (2, -2)^T$ . Tudíž, zvětšíme-li  $t$ , tak první složka řešení roste a druhá klesá stejnou měrou.  $\square$

## Derivace vlastních čísel

Připomeňme, že levý vlastní vektor matice  $A$  je definován jako (pravý) vlastní vektor matice  $A^T$ . Následující věta [Lax, 2007] dává elegantní formulku pro derivaci vlastního čísla. Pokud matice není symetrická, pak vlastní čísla jsou obecně komplexní, a tedy i derivace znamená derivaci komplexní proměnné.

**Věta 7.13.** *Nechť matice  $A \in \mathbb{C}^{n \times n}$  má navzájem různá vlastní čísla. Bud'  $\lambda \in \mathbb{C}$  její vlastní číslo,  $x, y \in \mathbb{C}^n$  její pravý a levý vlastní vektor znormované tak, že  $x^T y = 1$ . Pak*

$$\lambda' = y^T A' x.$$

Speciálně, pokud měníme pouze prvek  $a_{ij}$ , to jest  $a_{ij}(t) = a_{ij} + t$ , tak

$$\lambda' = y_i x_j.$$

Je-li ještě navíc matice  $A$  reálná symetrická, pak

$$\lambda' = x_i x_j.$$

*Důkaz.* Matice  $A$  je diagonalizovatelná, tudíž se dá vyjádřit jako  $A = X\Lambda X^{-1}$ , kde  $\Lambda$  je diagonální s vlastními čísly  $A$  na diagonále,  $X$  obsahuje ve sloupcích odpovídající vlastní vektory a  $X^{-1}$  obsahuje v řádcích odpovídající levé vlastní vektory. Díky vztahu  $X^{-1}X = I_n$  pak máme zaručeno normování  $x^T y = 1$  z předpokladu věty.

Zderivováním rovnice  $AX = X\Lambda$  dostaneme

$$A'X + AX' = X'\Lambda + X\Lambda',$$

z čehož

$$A'X - X\Lambda' = X'\Lambda - AX'.$$

Vynásobením  $X^{-1}$  zleva máme

$$\begin{aligned} X^{-1}A'X - \Lambda' &= X^{-1}X'\Lambda - X^{-1}AX' \\ &= X^{-1}X'\Lambda - X^{-1}X\Lambda X^{-1}X' \\ &= X^{-1}X'\Lambda - \Lambda X^{-1}X'. \end{aligned}$$

Diagonální prvky matice napravo jsou nulové, tudíž je nulová diagonála matice vlevo, z čehož máme pro každé  $k$

$$(X^{-1}A'X - \Lambda')_{kk} = 0.$$

Speciálně pro dané vlastní číslo dostaneme  $y^T A'x = \lambda'$ .

Pokud měníme pouze prvek  $a_{ij}$ , tak  $A' = e_i e_j^T$  a proto  $\lambda' = y^T A'x = y^T e_i e_j^T x = y_i x_j$ . Pokud je navíc matice  $A$  reálná symetrická, tak levé a pravé vlastní vektory se shodují.  $\square$

Pokud je matice  $A$  symetrická a měníme pouze diagonální prvek  $a_{ii}$ , tak  $\lambda' = x_i^2 \geq 0$ . To znamená, že při zvětšení jednoho či více diagonálních prvků se žádné vlastní číslo nezmenší – každé se buďto zvětší, nebo zůstane stejné.

**Příklad 7.14.** Uvažujme matici  $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ . Ta má vlastní čísla 3 a 1 a příslušné vlastní vektory jsou  $\frac{\sqrt{2}}{2}(1, 1)^T$  a  $\frac{\sqrt{2}}{2}(1, -1)^T$ . Tudíž podle věty platí

$$\frac{\partial \lambda_1}{\partial a_{11}} = \frac{1}{2}, \quad \frac{\partial \lambda_1}{\partial a_{12}} = \frac{1}{2}, \quad \frac{\partial \lambda_2}{\partial a_{11}} = \frac{1}{2}, \quad \frac{\partial \lambda_2}{\partial a_{12}} = -\frac{1}{2}.$$

Výsledek lze použít i k odhadu vlastních čísel po malé perturbaci. Tak například matice  $\begin{pmatrix} 2.1 & 1 \\ 1 & 2 \end{pmatrix}$  bude mít vlastní čísla  $\approx 3.05$  a  $\approx 1.05$  a matice  $\begin{pmatrix} 2 & 1.1 \\ 1 & 2 \end{pmatrix}$  bude mít vlastní čísla  $\approx 3.1$  a  $\approx 0.9$ .  $\square$

**Příklad 7.15** (Magnus [1985]). Uvažujme symetrickou matici

$$A = \begin{pmatrix} 1 + \varepsilon & \delta \\ \delta & 1 + \varepsilon \end{pmatrix}$$

závislou na dvou proměnných. Její vlastní čísla jsou  $\lambda_{1,2} = 1 \pm \sqrt{\varepsilon^2 + \delta^2}$  a v prostoru  $(\varepsilon, \delta, \lambda)$  představují dvojité kuželové pláště. Jediná možnost, jak vybrat dvě vlastní čísla  $\lambda_{1,2}$ , aby představovaly spojité funkce proměnných  $(\varepsilon, \delta)$  je ta, že jedno vlastní číslo odpovídá horní polovině kužele a druhé dolní polovině. Pak jsou ale vlastní čísla nehladká v počátku.

Poznamenejme, že zde byly zapotřebí obě proměnné  $(\varepsilon, \delta)$ , protože při jedné proměnné má množina vlastních čísel tvar podobný  $\lambda$ , a tudíž lze vybrat vlastní čísla jako hladké funkce tak, že jedno odpovídá tvaru  $\sqrt{\varepsilon^2 + \delta^2}$  a druhé  $-\sqrt{\varepsilon^2 + \delta^2}$ .  $\square$

## Další použití

Maticové derivace se vyskytují v řadě jiných oblastí, např. v optimalizaci. Ukážeme použití na známém příkladu metody nejmenších čtverců.

**Příklad 7.16** (Metoda nejmenších čtverců). Uvažujme soustavu  $Ax = b$ , kde  $A \in \mathbb{R}^{m \times n}$ ,  $b \in \mathbb{R}^m$  a matice  $A$  má hodnost  $n$ . Typicky  $m$  je mnohem větší než  $n$ . Protože soustava nemusí mít řešení, hledáme přibližné řešení optimalizační úlohou

$$\min_{x \in \mathbb{R}^n} \|Ax - b\|_2.$$

Zde je snaha najít takový vektor  $x$ , aby chyba mezi levou a pravou stranou byla v eukleidovské normě co nejmenší. Protože je druhá mocnina rostoucí funkce, optimum se nabyde ve stejném bodě jako pro úlohu s účelovou funkcí

$$\|Ax - b\|_2^2 = (Ax - b)^T(Ax - b) = x^T A^T A x - 2b^T A x + b^T b.$$

Účelová funkce je konvexní, tedy minimum se nabyde v bodě s nulovou derivací. Derivace (tj. gradient) účelové funkce je roven  $2A^T A x - 2A^T b$ . Pokud má být nulový, dostáváme podmínku  $A^T A x = A^T b$ . To je známá soustava normálních rovnic. Pokud jsou sloupce matice  $A$  lineárně nezávislé, je matice  $A^T A$  regulární a dostáváme jednoznačné řešení  $x = (A^T A)^{-1} A^T b$ .  $\square$

## Cvičení

- 7.1. Buď  $A \in \mathbb{R}^{n \times n}$  symetrická a taková, že  $A^k \rightarrow_{k \rightarrow \infty} 0$ . Určete co největší  $\delta > 0$  takové, že matice bude splňovat podmínku  $A^k \rightarrow_{k \rightarrow \infty} 0$  i když libovolné prvky matice nezávisle zperturbujeme až o hodnotu  $\delta$  (se zachováním symetrie matice).
- 7.2. Buď  $A \in \mathbb{R}^{n \times n}$  regulární. Je zobrazení  $A \mapsto A^{-1}$  spojitě na nějakém okolí matice  $A$ ?
- 7.3. Derivace determinantu.
  - (a) Buď  $A \in \mathbb{R}^{n \times n}$ . Určete matici  $\det(A)'$ , jejíž  $(i, j)$ -prvek je  $\frac{\partial \det(A)}{\partial a_{ij}}$ . (*Hint*: Laplaceův rozvoj determinantu.)
  - (b) Buď  $A \in \mathbb{R}^{n \times n}$  pozitivně definitní. Určete  $(\log \det(A))'$ . (*Poznámka*: Funkce  $-\log \det(A)$  je konvexní a vyskytuje se v geometrických úlohách minimalizace objemu, například nalezení nejmenšího elipsoidu, který obsahuje dané body.)
  - (c) Buďte  $A, B \in \mathbb{R}^{n \times n}$ . Určete  $\det(A + tB)'$ , tj. derivaci podle proměnné  $t$ . (*Hint*: Derivace složené funkce.)
- 7.4. Spočítejte:
  - (a)  $\frac{\partial \text{trace}(A)}{\partial A}$ ,
  - (b)  $\frac{\partial x^T A x}{\partial x}$ ,
  - (c)  $\frac{\partial A x}{\partial x}$ ,
  - (d)  $\frac{\partial \|A\|_F^2}{\partial A}$ .



# Kapitola 8

## Nezáporné a kladné matice

Nezáporné a kladné matice mají z hlediska vlastních čísel zajímavé vlastnosti, proto se vyplatí si o nich říci více. Matice je nezáporná resp. kladná pokud všechny její složky jsou nezáporné resp. kladné. Značíme  $A \geq 0$  resp.  $A > 0$ . S podobným významem „aplikace po složkách“ používáme i absolutní hodnotu.

Teorie nezáporných a kladných matic se nazývá často Perronova teorie podle německého matematika Oskara Perrona. Shrnuto, tato teorie říká:

**Věta 8.1** (Perron, 1907).

- (1) *Bud'  $A \in \mathbb{R}^{n \times n}$  nezáporná matice. Pak největší (v absolutní hodnotě) vlastní číslo je reálné nezáporné a příslušný vlastní vektor je nezáporný (ve všech složkách).*
- (2) *Bud'  $A \in \mathbb{R}^{n \times n}$  kladná matice. Pak největší (v absolutní hodnotě) vlastní číslo je reálné kladné, je jediné, násobnosti 1, a příslušný vlastní vektor je kladný (ve všech složkách). Navíc žádnému jinému vlastnímu číslu neodpovídá nezáporný vlastní vektor.*

V následujících odstavcích postupně vybudujeme Perronovu teorii, viz Horn and Johnson [1985]; Meyer [2000]. I dílčí mezikroky dávají silné a užitečné výsledky – např. věta 8.6 pro verifikaci ze Sekce 8.4.

Pracujeme se čtvercovými maticemi  $(A, B, \dots)$  řádu  $n$ .

### 8.1 Základní výsledky pro nezáporné matice

**Věta 8.2.** *Je-li  $|A| \leq B$ , pak  $\rho(A) \leq \rho(B)$ .*

*Důkaz.* Ze vztahu  $|A^k| \leq |A|^k \leq B^k$  odvodíme například pro Frobeniovu normu  $\|A^k\|_F \leq \| |A|^k \|_F \leq \| |A|^k \|_F \leq \|B^k\|_F$ . Tudíž  $\|A^k\|_F^{1/k} \leq \|B^k\|_F^{1/k}$  a podle Gelfandovy formule (věta 5.19) dostaneme limitním přechodem  $k \rightarrow \infty$  výsledný vztah  $\rho(A) \leq \rho(B)$ .  $\square$

**Důsledek 8.3.** *Je-li  $A \geq 0$  a  $\tilde{A}$  hlavní podmatice  $A$  (tj., vznikne z  $A$  odstraněním řádků a sloupců se stejnými indexy), pak  $\rho(\tilde{A}) \leq \rho(A)$ . Speciálně  $a_{ii} \leq \rho(A)$  pro všechna  $i$ .*

*Důkaz.* Podle věty 8.2 je  $\rho(\tilde{A}) = \rho \begin{pmatrix} \tilde{A} & 0 \\ 0 & 0 \end{pmatrix} \leq \rho(A)$ .  $\square$

**Lemma 8.4.** *Bud'  $A \geq 0$  s konstantními řádkovými součty. Pak  $\rho(A) = \|A\|_\infty$ .*

*Důkaz.*

„ $\leq$ “ Platí pro každou maticovou normu (věta 5.13).

„ $\geq$ “ Protože  $Ae = \|A\|_\infty e$ , tak  $\|A\|_\infty$  je vlastním číslem matice  $A$ .  $\square$

**Věta 8.5.** *Je-li  $A \geq 0$ , pak  $\min_i \sum_j a_{ij} \leq \rho(A) \leq \max_i \sum_j a_{ij}$ .*

*Důkaz.* Definujme  $\alpha := \min_i \sum_j a_{ij}$  a matici  $B$  s prvky  $b_{ij} := a_{ij} \frac{\alpha}{\sum_j a_{ij}}$  (pro  $\alpha = 0$  platí tvrzení triviálně, tudíž předpokládáme  $\alpha > 0$ ). Pak platí  $0 \leq B \leq A$  a řádkové součty matice  $B$  jsou rovny  $\alpha$ . Dle lemmatu 8.4 a věty 8.2 pak máme  $\alpha = \rho(B) \leq \rho(A)$ .

Druhá nerovnost je přímo rovna vztahu  $\rho(A) \leq \|A\|_\infty$ .  $\square$

**Věta 8.6.** *Bud'  $A \geq 0$  a  $x > 0$ . Pak platí implikace:*

$$\begin{aligned} \alpha x \leq Ax &\Rightarrow \alpha \leq \rho(A), \\ \alpha x < Ax &\Rightarrow \alpha < \rho(A), \\ Ax \leq \beta x &\Rightarrow \rho(A) \leq \beta, \\ Ax < \beta x &\Rightarrow \rho(A) < \beta. \end{aligned}$$

*Důkaz.* Dokážeme první implikaci, zbytek se dokáže analogicky. Je-li  $\alpha x \leq Ax$ , pak pro každé  $i$  je  $\alpha x_i \leq \sum_j a_{ij} x_j$ , a tedy  $\alpha \leq \min_i \sum_j a_{ij} \frac{x_j}{x_i}$ . Podle první nerovnosti z věty 8.5 pak máme

$$\alpha \leq \rho(\operatorname{diag}(x_1^{-1}, \dots, x_n^{-1})A \operatorname{diag}(x_1, \dots, x_n)) = \rho(A). \quad \square$$

Zvolíme-li nejlepší možné  $\alpha, \beta$  pak dostaneme následující meze pro  $\rho(A)$ . Kvalita odhadu závisí na volbě vektoru  $x > 0$ . Pokud se nabyde  $\rho(A)$  jako vlastní číslo a přísluší mu kladný vlastní vektor (což vždy platí, jak uvidíme později), tak nerovnosti v Collatzově větě se nabydou jako rovnosti. Tudíž, čím lepší je  $x$  odhadem vlastního vektoru, tím těsnější meze máme.

**Důsledek 8.7** (Collatz, 1942). *Je-li  $A \geq 0$  a  $x > 0$ , pak*

$$\min_i \frac{(Ax)_i}{x_i} \leq \rho(A) \leq \max_i \frac{(Ax)_i}{x_i}.$$

Další důsledek nám už dá dílčí tvrzení z Perronovy věty.

**Důsledek 8.8.** *Má-li  $A \geq 0$  vlastní vektor  $x > 0$ , pak příslušné vlastní číslo je  $\rho(A)$ .*

*Důkaz.* Bud'  $\lambda$  vlastní číslo příslušné  $x$ . Z rovnice  $Ax = \lambda x$  a  $x > 0$  plyne  $\lambda$  je reálné nezáporné. Přepíšeme-li rovnici jako  $\lambda x \leq Ax \leq \lambda x$ , tak z věty 8.6 dostáváme  $\lambda \leq \rho(A) \leq \lambda$ .  $\square$

## 8.2 Specifické výsledky pro kladné matice

Nyní dokážeme podstatnou část Perronovy věty – hodnota  $\rho(A) > 0$  se nabyde jako vlastní číslo a přísluší mu kladný vlastní vektor.

**Věta 8.9.** *Bud'  $A > 0$ ,  $Ax = \lambda x$ ,  $x \neq 0$ ,  $|\lambda| = \rho(A)$ . Pak  $A|x| = \rho(A)|x|$  a  $|x| > 0$ .*

*Důkaz.* Počítejme

$$\rho(A)|x| = |\lambda| \cdot |x| = |\lambda x| = |Ax| \leq |A| \cdot |x| = A|x|.$$

Pokud nastane rovnost  $A|x| = \rho(A)|x|$ , jsme hotovi a  $|x| > 0$  díky  $A|x| > 0$ . V opačném případě definujme  $z := A|x| > 0$ . Pak

$$0 < A(A|x| - \rho(A)|x|) = Az - \rho(A)z.$$

Tudíž  $\rho(A)z < Az$  a podle věty 8.6 je  $\rho(A) < \rho(A)$ , spor.  $\square$

Další část Perronovy věty je na řadě – největší vlastní číslo je jediné, tedy nejsou dvě různá vlastní čísla  $\lambda_1, \lambda_2$  s vlastností  $|\lambda_1| = |\lambda_2| = \rho(A)$ .

**Věta 8.10.** *Bud'  $A > 0$  a  $\lambda \neq \rho(A)$  její vlastní číslo. Pak  $|\lambda| < \rho(A)$ .*

*Důkaz.* Pro spor necht'  $|\lambda| = \rho(A)$ . Bud'  $x$  vlastní vektor odpovídající  $\lambda$ . Z důkazu věty 8.9 plyne  $|Ax| = A|x|$ , neboli pro každé  $k$  platí  $|\sum_j a_{kj} x_j| = \sum_j a_{kj} |x_j|$ . Trojúhelníková nerovnost pro komplexní čísla se nabyde jako rovnost právě tehdy když čísla leží na stejné polopřímce směřující od počátku. Tudíž existuje  $\gamma \in \mathbb{C}$  takové, že  $\gamma a_{kj} x_j > 0$  pro všechna  $j$  ( $\gamma$  reprezentuje otočení polopřímky tak, aby se shodovala s kladnou reálnou osou). Pro všechna  $j$  je  $\gamma x_j$  reálné kladné a tudíž z rovnice  $A(\gamma x) = \lambda(\gamma x)$  dostáváme, že vlastní vektor  $\gamma x > 0$ , tedy i  $\lambda$  je reálné kladné. Podle důsledku 8.8 pak  $\lambda = \rho(A)$ , spor.  $\square$

Žádnému jinému vlastnímu číslu než  $\rho(A)$  neodpovídá nezáporný vlastní vektor.



**Věta 8.11.** *Bud'  $A > 0$  a bud'  $x \geq 0$ ,  $x \neq 0$ , její vlastní vektor. Pak  $Ax = \rho(A)x$ .*

*Důkaz.* Pro spor necht'  $Ax = \lambda x$  a necht'  $|\lambda| < \rho(A)$ . Uvažujme vektor  $x^* := x + \varepsilon e$ , kde  $\varepsilon > 0$  je libovolně malé. Podle Collatzovy věty (důsledek 8.7) aplikované na vektor  $x^*$  je pak hodnota  $\rho(A)$  libovolně blízko k  $\lambda$ ,

$$\min_i \frac{\lambda x_i + \varepsilon(Ae)_i}{x_i + \varepsilon} = \min_i \frac{(Ax^*)_i}{x_i^*} \leq \rho(A) \leq \max_i \frac{(Ax^*)_i}{x_i^*} = \max_i \frac{\lambda x_i + \varepsilon(Ae)_i}{x_i + \varepsilon},$$

což je spor. □

Pro úplnost Perronovy věty pro kladné matice by ještě zbývalo ukázat, že  $\rho(A)$  má násobnost 1.

### 8.3 Specifické výsledky pro nezáporné matice

Část Perronovy věty pro nezáporné matice se ukáže celkem jednoduše limitním přechodem od kladných matic, s využitím spojitosti vlastních čísel a vlastních vektorů (sekce 7.2). Pro matici  $A \geq 0$  stačí uvažovat matici  $A + \varepsilon e e^T > 0$ , kde  $\varepsilon > 0$  je libovolně malé; limitním přechodem  $\varepsilon \rightarrow 0$  pak získáme požadované vlastnosti (Toto je klíčová myšlenka, ve skutečnosti je potřeba ještě doladit několik technických detailů).

Tím se ukáže, že největší vlastní číslo je reálné nezáporné a příslušný vlastní vektor je nezáporný. Další vlastnosti kladných matic obecně ztratíme:

- Matice  $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$  má vícenásobné největší vlastní číslo.
- Matice  $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$  má největší vlastní číslo 1 s vlastním vektorem  $(1, 0)^T$ , který není kladný. Navíc jiné vlastní číslo 0 má také nezáporný vlastní vektor  $(0, 1)^T$ .
- Matice  $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$  má největší vlastní číslo 0, i když je nezáporná nenulová.

Pro určité nezáporné matice ale můžeme zesílit vlastnosti na ty, co mají kladné matice. To vede na teorii ireducibilních matic, vybudovanou Frobeniem. Matice  $A$  je *ireducibilní* pokud neexistuje permutační matice  $P$  taková, aby  $P^T A P$  byla blokově horní trojúhelníková matice. Pro nezápornou ireducibilní matici pak platí, že  $\rho(A) > 0$  je vlastní číslo násobnosti 1 a odpovídá mu kladný vlastní vektor. Tedy platí skoro všechny vlastnosti kladných matic, jen se může stát, že  $\rho(A)$  se nabyde pro více vlastních čísel, např. pro  $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ , která má vlastní čísla  $\pm 1$ .

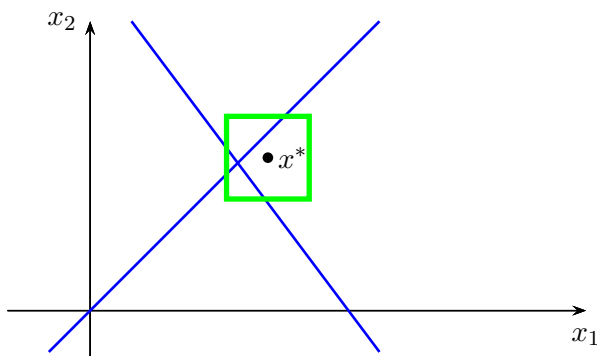
### 8.4 Aplikace – verifikace lineárních soustav

Mnohokrát jsme v tomto textu zkoumali numerické problémy při řešení soustav lineárních rovnic (jakožto jednoho vybraného problému). Označili jsme matice náchylné na numerické chyby skrze číslo podmíněnosti a zkoumali jsme změnu řešení při změnách prvků matice.

Cílem *verifikace* [Rump, 2010] je určit numericky rigorózní (horní) odhad vzdálenosti numericky spočítaného řešení od skutečného. To se může zdát podivné: Skutečné řešení přece neznáme! Navíc každý výpočet je náchylný na chyby, takže výpočet odhadu musí být chybový! Přesto to za určitých předpokladů umíme.

Uvažujme soustavu  $Ax = b$ , kde  $A \in \mathbb{R}^{n \times n}$  a  $b \in \mathbb{R}^n$ . Necht'  $x^* \in \mathbb{R}^n$  je numericky spočítané řešení. Náš cíl je najít odhad chyby v jakési vážené maximové vzdálenosti, tedy určit vektor  $y^\Delta > 0$  takový, že skutečné řešení  $x = A^{-1}b$  splňuje  $|x - x^*| \leq y^\Delta$ , neboli  $x^* - y^\Delta \leq x \leq x^* + y^\Delta$ . Zkráceně píšeme  $x \in x^* + [-y^\Delta, y^\Delta]$ . Geometrický význam je najít kolem bodu  $x^*$  hyperkvádr, ve kterém se nalézá  $x$  (obr. 8.1).

Soustavy rovnic se často předpokládají. Je-li  $C \in \mathbb{R}^{n \times n}$  regulární, pak  $C Ax = C b$  je ekvivalentní soustava. Pro naše účely je nejvhodnější volit  $C := A^{-1}$  nebo její aproximaci. V zásadě má ale věta dole obecnou platnost.



Obrázek 8.1: Ilustrace verifikace přibližného řešení  $x^*$  soustavy  $Ax = b$ .

Ve větě se dále používá intervalová aritmetika. Je jednoduše definovaná jako obraz intervalu při dané operaci. Tedy pro dva intervaly  $[\underline{a}, \bar{a}]$ ,  $[\underline{b}, \bar{b}]$  definujeme

$$\begin{aligned} [\underline{a}, \bar{a}] + [\underline{b}, \bar{b}] &= [\underline{a} + \underline{b}, \bar{a} + \bar{b}], \\ [\underline{a}, \bar{a}] - [\underline{b}, \bar{b}] &= [\underline{a} - \bar{b}, \bar{a} - \underline{b}], \\ [\underline{a}, \bar{a}] \cdot [\underline{b}, \bar{b}] &= [\min(\underline{a}\underline{b}, \underline{a}\bar{b}, \bar{a}\underline{b}, \bar{a}\bar{b}), \max(\underline{a}\underline{b}, \underline{a}\bar{b}, \bar{a}\underline{b}, \bar{a}\bar{b})]. \end{aligned}$$

Při implementaci intervalové aritmetiky pak volíme vhodný zaokrouhlovací mód tak, aby výsledný interval skutečně obsahoval všechny možné evaluace intervalových hodnot. Tím pádem zohledňujeme i numerické chyby.

**Věta 8.12** (Krawczyk, 1969, Moore, 1977, Rump, 1980). *Platí-li*

$$C(b - Ax^*) + (I - CA)[-y^\Delta, y^\Delta] \subseteq \text{int}[-y^\Delta, y^\Delta], \quad (8.1)$$

*pak  $A, C$  jsou regulární a  $A^{-1}b \in x^* + [-y^\Delta, y^\Delta]$ .*

*Důkaz.* Uvažujme zobrazení  $y \mapsto C(b - Ax^*) + (I - CA)y$ . Ze vztahu (8.1) plyne, že nadkvádr  $[-y^\Delta, y^\Delta]$  se zobrazí do sebe. Podle Brouwerovy věty o pevném bodě (o spojitém zobrazení konvexní kompaktní množiny do sebe) má zobrazení pevný bod  $y$ . Pro něj platí  $y = C(b - Ax^*) + (I - CA)y$ , neboli  $CA(x^* + y) = Cb$ . To znamená, že  $x^* + y$  je řešením soustavy.

Inkluze (8.1) dále znamená, že šířka intervalu vlevo je menší než šířka napravo, neboli  $|I - CA|y^\Delta < y^\Delta$ . Podle vět 8.2 a 8.6 pak  $\rho(I - CA) \leq \rho(|I - CA|) < 1$ . Vlastní čísla matice  $CA$  jsou nenulová. Tudíž  $CA$  je regulární a i jednotlivé matice  $C, A$  jakbysmet.  $\square$

Věta tedy dává garantovanou vzdálenost od skutečného řešení a verifikuje regularitu matice  $A$ . Pokud navíc celý výraz (8.1) vyhodnotíme intervalovou aritmetikou (tedy i operace s reálnými čísly nahradíme intervalovou aritmetikou se správným zaokrouhlováním), tak meze jsou numericky rigorózní.

Věta se typicky implementuje následujícím způsobem:  $C$  je numerická aproximace  $A^{-1}$ ,  $x^*$  je přibližné řešení soustavy a  $[-y^\Delta, y^\Delta]$  je malý počáteční nadkvádr. Pokud podmínka (8.1) uspěje, jsme hotovi. Jinak nadkvádr  $[-y^\Delta, y^\Delta]$  zvětšíme a postup opakujeme.

## Cvičení

8.1. Nespektrální vlastnosti:

- Buď  $A > 0$ . Může být  $A^{-1} > 0$ ?
- Buď  $A \geq 0$ . Ukažte, že  $A^{-1} \geq 0$  jen pokud kladné prvky matice  $A$  odpovídají permutaci.

8.2. Určete spektrální poloměr latinského čtverce velikosti  $n \times n$  (v každém řádku i sloupci má všechna čísla  $1, \dots, n$  právě jednou).

- 8.3. Buď  $A \geq 0$ . Ukažte, že  $\rho(A) \geq \min_i \{a_{ii}\}$ .
- 8.4. Buď  $A \geq 0$ . Ukažte, že  $\rho(A + \alpha I_n) = \rho(A) + \alpha$  pro každé  $\alpha \geq 0$ .
- 8.5. Buď  $0 \leq A < B$ . Ukažte, že  $\rho(A) < \rho(B)$ .
- 8.6. Buď  $A < 0$ . Rozhodněte zda platí:
- (a)  $\rho(A)$  se nabyde jako vlastní číslo,
  - (b) vlastní vektor, odpovídající dominantnímu vlastnímu číslu, je kladný,
- 8.7. Buď  $A \geq 0$  a  $A^k > 0$  pro nějaké přirozené  $k$ . Dokažte, že pro matici  $A$  platí spektrální vlastnosti kladných matic (věta 8.1(2)).
- 8.8. Ukažte, že pokud  $0 \leq A \leq B$ , pak pro spektrální normu platí  $\|A\|_2 \leq \|B\|_2$ .
- 8.9. Buď  $A \in \mathbb{R}^{n \times n}$  s kladnými mimodiagonálními prvky, tj.  $a_{ij} > 0$  pro  $i \neq j$ . Dokažte, že vlastní číslo s největší reálnou částí je reálné.
- 8.10. Buď  $A < 0$  a označme  $\alpha := \rho(A)$ . Ukažte, že limitní matice  $\lim_{k \rightarrow \infty} (\frac{1}{\alpha} A)^k$  existuje a má hodnotu 1.



# Kapitola 9

## Maticové funkce a mocninné řady

S mocninnými řadami matic jsme se letmo setkali ve větě 5.18 o Neumannových řadách. Nyní prostudujeme řady matic obecněji, a to v rámci tematiky maticových funkcí<sup>1)</sup>.

### 9.1 Maticové funkce

Položme si otázku: Jak zavést maticovou funkci jako např.  $\cos(A)$ ,  $e^A$ , atp.? Pro reálnou funkci  $f: \mathbb{R} \rightarrow \mathbb{R}$  a matici  $A \in \mathbb{R}^{n \times n}$  zavést  $f(A)$  tak, že aplikujeme funkci na každou složku matice zvlášť,

$$f(A) = \begin{pmatrix} f(a_{11}) & \dots & f(a_{1n}) \\ \vdots & & \vdots \\ f(a_{n1}) & \dots & f(a_{nn}) \end{pmatrix}, \quad (9.1)$$

je sice možné, ale moc pěkných vlastností to mít nebude. Zkusme jiný přístup. Předpokládejme, že funkce  $f: \mathbb{R} \rightarrow \mathbb{R}$  je dostatečně hladká a dá se vyjádřit nekonečným rozvojem  $f(x) = \sum_{k=0}^{\infty} a_k x^k$ ; reálné analytické funkce jako např.  $\sin(x)$ ,  $e^x$  aj. tento předpoklad splňují. Pak je tedy přirozené zavést  $f(A) = \sum_{k=0}^{\infty} a_k A^k$ . Tuto mocninnou řadu vyhodnotíme pomocí Jordanovy normální formy.

Buď  $A = SJS^{-1}$ , kde  $J$  je Jordanova normální forma matice  $A$ . Protože  $A^k = SJ^kS^{-1}$ , dostáváme

$$f(A) = \sum_{k=0}^{\infty} a_k SJ^kS^{-1} = Sf(J)S^{-1}.$$

Redukovali jsme tedy problém na vyhodnocení funkce  $f$  pro matice v Jordanově normální formě. Začneme s jednodušším případem, kdy  $J$  je diagonální (tj., matice  $A$  je diagonalizovatelná) a na diagonále má čísla  $\lambda_1, \dots, \lambda_n$ , což jsou vlastní čísla matice  $A$ . Pak

$$J^k = \text{diag}(\lambda_1, \dots, \lambda_n)^k = \text{diag}(\lambda_1^k, \dots, \lambda_n^k),$$

a tudíž

$$f(J) = \sum_{k=0}^{\infty} a_k \begin{pmatrix} \lambda_1^k & & \\ & \ddots & \\ & & \lambda_n^k \end{pmatrix} = \begin{pmatrix} \sum_{k=0}^{\infty} a_k \lambda_1^k & & \\ & \ddots & \\ & & \sum_{k=0}^{\infty} a_k \lambda_n^k \end{pmatrix} = \begin{pmatrix} f(\lambda_1) & & \\ & \ddots & \\ & & f(\lambda_n) \end{pmatrix}.$$

Nyní přikročíme k obecnému případu, kdy  $J$  je blokově diagonální a jednotlivé bloky jsou Jordanovy buňky  $J_{k_1}(\lambda_1), \dots, J_{k_m}(\lambda_m)$ . Snadno nahlédneme, že

$$f(J) := \begin{pmatrix} f(J_{k_1}(\lambda_1)) & & \\ & \ddots & \\ & & f(J_{k_m}(\lambda_m)) \end{pmatrix}.$$

<sup>1)</sup>Jako první o maticových funkcích pojednával J.J. Sylvester roku 1883.

Chybí tedy ještě zavést obraz Jordanových buněk  $J_{k_i}(\lambda_i)$ . Pro  $k_i = 1$  je to triviální, je to opět  $f(\lambda_i)$ . Pro  $k_i > 1$  je předpis složitější [Meyer, 2000]:

$$f(J_{k_i}(\lambda_i)) := \begin{pmatrix} f(\lambda_i) & f'(\lambda_i) & \dots & \frac{f^{(k_i-1)}(\lambda_i)}{(k_i-1)!} \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & f'(\lambda_i) \\ 0 & \dots & 0 & f(\lambda_i) \end{pmatrix}. \quad (9.2)$$

*Důkaz pro  $k_i = 2$ .* Nebudeme dokazovat (9.2) v celé obecnosti, ale ukážeme, proč vztah platí pro případ Jordanovy buňky velikosti 2. Základní myšlenka je založená na tom chtít, aby  $f(A)$  byla spojitá funkce. Protože  $A$  lze aproximovat libovolně přesně pomocí diagonalizovatelných matic (věta 5.27), můžeme hodnotu  $f(A)$  definovat jako limitu obrazů těchto matic.

Speciálně, uvažujme Jordanovu buňku velikosti  $k = 2$ , tedy  $A = J_2(\lambda) = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$ . Definujme matici  $A_\varepsilon := \begin{pmatrix} \lambda & 1 \\ 0 & \lambda + \varepsilon \end{pmatrix}$ , kde  $\varepsilon \neq 0$ . Tato matice už je diagonalizovatelná a

$$A_\varepsilon = S_\varepsilon \Lambda_\varepsilon S_\varepsilon^{-1}, \quad S_\varepsilon = \begin{pmatrix} 1 & 1 \\ 0 & \varepsilon \end{pmatrix}, \quad \Lambda_\varepsilon = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda + \varepsilon \end{pmatrix}, \quad S_\varepsilon^{-1} = \begin{pmatrix} 1 & -1/\varepsilon \\ 0 & 1/\varepsilon \end{pmatrix}.$$

Tudíž

$$f(A_\varepsilon) = S_\varepsilon f(\Lambda_\varepsilon) S_\varepsilon^{-1} = S_\varepsilon \begin{pmatrix} f(\lambda) & 0 \\ 0 & f(\lambda + \varepsilon) \end{pmatrix} S_\varepsilon^{-1} = \begin{pmatrix} f(\lambda) & (f(\lambda + \varepsilon) - f(\lambda))/\varepsilon \\ 0 & f(\lambda + \varepsilon) \end{pmatrix}.$$

Limitním přechodem pak dostáváme  $f(A) = \begin{pmatrix} f(\lambda) & f'(\lambda) \\ 0 & f(\lambda) \end{pmatrix}$ . □

**Příklad 9.1.** Uvažujme následující příklady:

1. Funkce  $f(x) = x^2$  má maticové rozšíření  $f(A) = A^2$ , jedná se tedy o klasické maticové mocnění. Na druhou stranu, předpis (9.1) by naznačoval mocnit jednotlivé prvky matice zvlášť, což není to, co bychom chtěli.
2. Pro funkci  $f(x) = \sqrt{x}$  a pozitivně semidefinitní matici dostaneme jako maticovou funkci  $f(A) = \sqrt{A}$  standardní odmocninu z matice [Hladík, 2017].
3. Pro funkci  $f(x) = \frac{1}{1-x}$  a matici  $A \in \mathbb{R}^{n \times n}$  splňující  $\rho(A) < 1$  dostaneme vlastnost z věty 5.18 o Neumannových řadách:  $(I_n - A)^{-1} = \sum_{k=0}^{\infty} A^k$ . □

Výše zmíněný postup umožňuje zavést maticovou funkci pro mnoho reálných funkcí, zejména analytické funkce. Tudíž můžeme uvažovat  $\sin(A)$ ,  $\cos(A)$ ,  $e^A$ , ... a studovat jejich vlastnosti. Některé známé identity se zobecní na maticové funkce [Horn and Johnson, 1991], jako například

$$\begin{aligned} \sin(2A) &= 2 \sin(A) \cos(A), \\ \sin^2(A) + \cos^2(A) &= I_n. \end{aligned}$$

Na druhou stranu, jiné se přímočaře zobecnit nedají (viz věta 9.4).

## 9.2 Maticová exponenciála

Exponenciála je důležitou reálnou funkcí a z pohledu maticových funkcí jednou z nejzajímavějších. Buď  $A \in \mathbb{R}^{n \times n}$ . Podle postupu z předchozí sekce je maticová exponenciála definovaná řadou

$$e^A = \sum_{k=0}^{\infty} \frac{1}{k!} A^k.$$

Vzorec pro maticovou exponenciálu můžeme odvodit přímo.

**Poznámka 9.2** (Odvození  $e^A$ ). Víme již, že maticovou funkci stačí vyjádřit pro jednotlivé Jordanovy buňky. Bez újmy na obecnost buď  $A$  Jordanova buňka, tedy  $A = \lambda I_n + N$ , kde

$$N = J_n(0) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ \vdots & \ddots & \ddots & 0 \\ \vdots & & \ddots & 1 \\ 0 & \dots & \dots & 0 \end{pmatrix}.$$

Pak

$$A^k = (\lambda I_n + N)^k = \sum_{m=0}^k \binom{k}{m} \lambda^{k-m} I_n^{k-m} N^m = \sum_{m=0}^k \binom{k}{m} \lambda^{k-m} N^m,$$

přičemž  $(N^m)_{i,i+m} = 1$  pro  $i = 1, \dots, n-m$  a jinak nuly. Speciálně,  $N^m = 0$  pro  $m \geq n$ . Nyní tedy

$$\begin{aligned} e^A &= \sum_{k=0}^{\infty} \frac{1}{k!} A^k = \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{m=0}^k \binom{k}{m} \lambda^{k-m} N^m \\ &= \sum_{k=0}^{\infty} \sum_{m=0}^k \frac{1}{k!} \binom{k}{m} \lambda^{k-m} N^m = \sum_{m=0}^{\infty} \sum_{k=m}^{\infty} \frac{1}{(k-m)!} \frac{1}{m!} \lambda^{k-m} N^m \\ &= \sum_{m=0}^{\infty} \frac{1}{m!} N^m \sum_{k=m}^{\infty} \frac{1}{(k-m)!} \lambda^{k-m} = \sum_{m=0}^{\infty} \frac{1}{m!} N^m e^{\lambda} \\ &= e^{\lambda} \sum_{m=0}^{n-1} \frac{1}{m!} N^m = e^{\lambda} \begin{pmatrix} 1 & \frac{1}{1!} & \dots & \frac{1}{(n-1)!} \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \frac{1}{1!} \\ 0 & \dots & 0 & 1 \end{pmatrix}. \end{aligned}$$

Dále zmíníme některé vlastnosti maticové exponenciály a příklady použití.

**Věta 9.3.** Platí  $\det(e^A) = e^{\text{trace}(A)}$ .

*Důkaz.* Buď  $A = SJS^{-1}$ , kde  $J$  je v Jordanově normální formě. Pak  $\det(e^A) = \det(Se^J S^{-1}) = \det(e^J)$ , přičemž

$$e^J = \sum_{k=0}^{\infty} \frac{1}{k!} J^k = \sum_{k=0}^{\infty} \frac{1}{k!} \begin{pmatrix} \lambda_1^k & & ? \\ & \ddots & \\ 0 & & \lambda_n^k \end{pmatrix} = \begin{pmatrix} e^{\lambda_1} & & ? \\ & \ddots & \\ 0 & & e^{\lambda_n} \end{pmatrix}.$$

Tudíž  $\det(e^J) = e^{\lambda_1} \dots e^{\lambda_n} = e^{\lambda_1 + \dots + \lambda_n} = e^{\text{trace}(A)}$ . □

**Věta 9.4.** Pokud  $AB = BA$ , pak  $e^{A+B} = e^A e^B$ .

*Důkaz.* Odvodíme

$$\begin{aligned} e^{A+B} &= \sum_{k=0}^{\infty} \frac{1}{k!} (A+B)^k = \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{m=0}^k \binom{k}{m} A^m B^{k-m} \\ &= \sum_{k=0}^{\infty} \sum_{m=0}^k \frac{1}{k!} \binom{k}{m} A^m B^{k-m} = \sum_{m=0}^{\infty} \sum_{k=m}^{\infty} \frac{1}{m!} A^m \frac{1}{(k-m)!} B^{k-m} \\ &= \sum_{m=0}^{\infty} \frac{1}{m!} A^m \sum_{k=m}^{\infty} \frac{1}{(k-m)!} B^{k-m} = \sum_{m=0}^{\infty} \frac{1}{m!} A^m e^B = e^A e^B. \end{aligned} \quad \square$$

Protipříklad pro výše zmíněnou vlastnost pro nekomutující matice je  $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ ,  $B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ , protože

$$e^{A+B} = \begin{pmatrix} e & e-1 \\ 0 & 1 \end{pmatrix} \neq \begin{pmatrix} e & e \\ 0 & 1 \end{pmatrix} = e^A e^B.$$

Čtenář si to může ověřit numericky v prostředí Matlab / Octave zavoláním funkce `expm()` na výpočet maticové exponenciály.

## Několik aplikací

**Diferenciální rovnice.** Uvažujme lineární diferenciální rovnici

$$y' = ay,$$

kde  $y = y(t): \mathbb{R} \rightarrow \mathbb{R}$  je hledaná funkce a  $a \in \mathbb{R}$  konstanta. Řešením je funkce  $y = c \cdot e^{at}$ , kde  $c \in \mathbb{R}$  je libovolné číslo (jeho hodnota vyplyne z tzv. počátečních podmínek, což je znalost hledané funkce v jenom bodě, např.  $y(0) = c$ ).

Nyní uvažujme soustavu lineárních diferenciálních rovnic řádu  $n$

$$y' = Ay,$$

kde  $y = y(t): \mathbb{R}^n \rightarrow \mathbb{R}^n$  je hledaný vektor funkcí a  $A \in \mathbb{R}^{n \times n}$  pevná matice. Pak řešením je  $y(t) = e^{At}c$ , kde  $c \in \mathbb{R}^n$  je libovolný vektor. Důkaz:

$$y(t)' = (e^{At}c)' = \left(\sum_{k=0}^{\infty} \frac{1}{k!} A^k t^k c\right)' = \sum_{k=1}^{\infty} \frac{1}{k!} A^k k t^{k-1} c = A \sum_{k=1}^{\infty} \frac{1}{(k-1)!} A^{k-1} t^{k-1} c = A e^{At} c = Ay.$$

**Centrum grafu [Estrada and Rodríguez-Velázquez, 2005].** Toto je jedna z možných definic centra grafu. Buď  $A$  matice sousednosti neorientovaného grafu, tj.  $A_{ij} = 1$  pokud  $i, j$  jsou spojeny hranou a  $A_{ij} = 0$  pokud ne. Pak  $(A^k)_{ij}$  udává počet sledů délky  $k$  mezi vrcholy  $i$  a  $j$ . Potom jakási průměrná či normalizovaný počet sledů můžeme definovat jako

$$e^A = \sum_{k=0}^{\infty} \frac{1}{k!} A^k.$$

Největší hodnota na diagonále matice  $e^A$  se nazývá *centrum grafu* a odpovídá vrcholu  $i$  s maximálním počtem sledů od  $i$  do  $i$ . Centrum grafu udává v jistém smyslu nejdůležitější vrchol. Existuje řada jiných definic centra grafu, ale výhoda tohoto přístupu je ta, že nezkoumá lokální chování (stupeň vrcholu atp.), ale zvažuje globální strukturu grafu.

**Matice rotace.** Maticová exponenciála se dá využít k vyjádření rotací v prostoru  $\mathbb{R}^3$ . Matice  $e^R$ , kde

$$R = \alpha \begin{pmatrix} 0 & -z & y \\ z & 0 & -x \\ -y & x & 0 \end{pmatrix} \quad (9.3)$$

totiž popisuje matici rotace kolem osy se směrnici  $(x, y, z)^T$  o úhel  $\alpha$  podle pravidla pravé ruky (viz cvičení 9.7).

**Další čtení.** Ještě jiné zajímavé aplikace maticové exponenciály najdeme v Higham [2008], například v definici průměrného oka v soustavě optických zařízení, nebo jak spočítat přechodovou matici šíření AIDS při limitovaných údajích.

## Cvičení

9.1. Pro matici  $A \in \mathbb{R}^{n \times n}$  a analytickou funkci  $f: \mathbb{R} \rightarrow \mathbb{R}$  ukažte

- $f(A^T) = f(A)^T$ ,
- $f(A)$  komutuje s  $A$ ,
- $f(SAS^{-1}) = Sf(A)S^{-1}$  pro každou regulární  $S \in \mathbb{R}^{n \times n}$ .

9.2. Vyjádřete  $f(A)$  pro matici  $A \in \mathbb{R}^{n \times n}$  a danou funkci:

- $f(t) = c$ , kde  $c$  je reálná konstanta,
- $f(t) = t$ ,



(c)  $f(t) = g(t) + h(t)$ .

9.3. Určete  $e^A$  pro matici  $A \in \mathbb{R}^{n \times n}$  pokud:

- (a)  $A = 0$ ,
- (b)  $A^2 = 0$ ,
- (c)  $A^2 = A$ .

9.4. Pro matici  $A \in \mathbb{R}^{n \times n}$  ukažte:

- (a)  $e^A$  je regulární a najděte její inverzi,
- (b) určete vlastní čísla a vlastní vektory matice  $e^A$ ,
- (c) je-li  $S \in \mathbb{R}^{n \times n}$  regulární, pak  $e^{SAS^{-1}} = Se^AS^{-1}$ ,
- (d) je-li  $A$  symetrická, pak  $e^A$  je pozitivně definitní,
- (e) je-li  $A = -A^T$ , pak  $e^A$  je ortogonální.

9.5. Najděte explicitní vyjádření pro  $e^A$ , kde  $A = \begin{pmatrix} 0 & -t \\ t & 0 \end{pmatrix}$ .

9.6. Odvoďte vzorec pro  $\frac{\partial}{\partial t} e^{At}$ , kde  $t$  je skalární proměnná.

9.7. Ukažte, že matice  $e^R$ , kde matice  $R$  byla definována v (9.3), skutečně popisuje rotaci o úhel  $\alpha$  kolem osy  $(x, y, z)^T$ . Stačí, když důkaz předvedete pouze pro rotace okolo souřadných os.

9.8. Jak vypadá  $e^A$  pokud  $A \geq 0$ ? A jak vypadá  $e^A$  pokud pouze mimodiagonální prvky matice  $A$  jsou nezáporné?

9.9. Jak byste definovali  $X^Y$  pro dvě matice  $X, Y \in \mathbb{R}^{n \times n}$ ?



# Kapitola 10

## Nestandardní maticové součiny

Kromě standardního maticového součinu existují i jiné. Jejich význam spočívá zejména ve zjednodušení složitých vyjádření a vyskytují se v určitých situacích. Přestože nemají tak důležitý význam jako ten standardní součin, je dobré o nich něco málo vědět.

### 10.1 Kroneckerův součin

Kroneckerův (nebo též tenzorový) součin umožňuje vynásobit matice libovolných rozměrů [Horn and Johnson, 1991].

**Definice 10.1.** Buď  $A \in \mathbb{R}^{m \times n}$ ,  $B \in \mathbb{R}^{p \times q}$ . Pak *Kroneckerův součin* matic  $A, B$  je matice  $A \otimes B \in \mathbb{R}^{mp \times nq}$  definována blokově jako

$$A \otimes B = \begin{pmatrix} a_{11}B & \dots & a_{1n}B \\ \vdots & & \vdots \\ a_{m1}B & \dots & a_{mn}B \end{pmatrix}.$$

**Poznámka 10.2** (Geometrický význam Kroneckerova součinu). Buďte  $f: U \rightarrow V$ ,  $\tilde{f}: \tilde{U} \rightarrow \tilde{V}$  dvě lineární zobrazení mezi vektorovými prostory. Buď dále

$$\begin{aligned} u_1, \dots, u_n \text{ báze } U, & \quad v_1, \dots, v_m \text{ báze } V, \\ \tilde{u}_1, \dots, \tilde{u}_q \text{ báze } \tilde{U}, & \quad \tilde{v}_1, \dots, \tilde{v}_p \text{ báze } \tilde{V}. \end{aligned}$$

Uvažujme vektorové prostory  $U \times \tilde{U}$  a  $V \times \tilde{V}$ . Pro vektory  $u \in U$  a  $\tilde{u} \in \tilde{U}$  však nebudeme uvažovat standardní množinu uspořádaných dvojic  $(u, \tilde{u})$  s operacemi

$$\begin{aligned} \alpha(u, \tilde{u}) &:= (\alpha u, \alpha \tilde{u}), \\ (u_1, \tilde{u}_1) + (u_2, \tilde{u}_2) &:= (u_1 + u_2, \tilde{u}_1 + \tilde{u}_2). \end{aligned}$$

Namísto toho budeme chtít, aby platily vlastnosti podobné bilineárním formám

$$\begin{aligned} \alpha(u, \tilde{u}) &= (\alpha u, \tilde{u}) = (u, \alpha \tilde{u}), \\ (u_1, \tilde{u}) + (u_2, \tilde{u}) &= (u_1 + u_2, \tilde{u}), \\ (u, \tilde{u}_1) + (u, \tilde{u}_2) &= (u, \tilde{u}_1 + \tilde{u}_2). \end{aligned}$$

Tyto vlastnosti pak pro vektory  $u = \sum_i \alpha_i u_i \in U$  a  $\tilde{u} = \sum_j \beta_j \tilde{u}_j \in \tilde{U}$  implikují

$$(u, \tilde{u}) = \left( \sum_i \alpha_i u_i, \sum_j \beta_j \tilde{u}_j \right) = \sum_{i,j} \alpha_i \beta_j (u_i, \tilde{u}_j).$$

Prostor  $U \times \tilde{U}$  je pak lepší si představit jako množinu výrazů  $\alpha_i \beta_j (u_i, \tilde{u}_j)$ , kde  $\alpha_i \beta_j \in \mathbb{R}$ , než jako množinu uspořádaných dvojic  $(u, \tilde{u})$ . Báze prostorů lze volit například jako

$$(u_1, \tilde{u}_1), \dots, (u_1, \tilde{u}_q), \dots, (u_n, \tilde{u}_1), \dots, (u_n, \tilde{u}_q) \text{ pro prostor } U \times \tilde{U},$$

$(v_1, \tilde{v}_1), \dots, (v_1, \tilde{v}_p), \dots, (v_m, \tilde{v}_1), \dots, (v_m, \tilde{v}_p)$  pro prostor  $V \times \tilde{V}$ .

Pak  $f \otimes \tilde{f}: U \times \tilde{U} \rightarrow V \times \tilde{V}$  je lineární zobrazení definované  $(u_j, \tilde{u}_\ell) \mapsto (f(u_j), \tilde{f}(\tilde{u}_\ell))$ . Buďte  $A, B$  matice zobrazení  $f, \tilde{f}$  vzhledem k daným bázím, tedy

$$f(u_j) = \sum_i a_{ij} v_i, \quad \tilde{f}(\tilde{u}_\ell) = \sum_k b_{k\ell} \tilde{v}_k.$$

Jaká bude matice zobrazení  $f \otimes \tilde{f}$ ?

Z definice se  $(u_j, \tilde{u}_\ell)$  zobrazí na

$$(f(u_j), \tilde{f}(\tilde{u}_\ell)) = (\sum_i a_{ij} v_i, \sum_k b_{k\ell} \tilde{v}_k) = \sum_{i,k} a_{ij} b_{k\ell} (v_i, \tilde{v}_k).$$

Tudíž hledaná matice je  $A \otimes B$ .

**Věta 10.3** (Vlastnosti Kroneckerova součinu). *Platí pro matice vhodných rozměrů:*

- (1)  $(A \otimes B)(C \otimes D) = AC \otimes BD$ , má-li  $AC, BD$  smysl,
- (2)  $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$ , jsou-li  $A, B$  regulární,
- (3)  $(A \otimes B)^T = A^T \otimes B^T$ ,
- (4)  $(A \otimes B) \otimes C = A \otimes (B \otimes C)$ ,
- (5) obecně  $A \otimes B \neq B \otimes A$ .

*Důkaz.*

- (1) Porovnáme bloky na pozici  $i, j$ :

$$\begin{pmatrix} a_{i1}B & \dots & a_{in}B \end{pmatrix} \begin{pmatrix} c_{1j}D \\ \vdots \\ c_{nj}D \end{pmatrix} = (\sum_{k=1}^n a_{ik}c_{kj})BD = (AC)_{ij}BD.$$

- (2) Podle předchozího  $(A \otimes B)(A^{-1} \otimes B^{-1}) = AA^{-1} \otimes BB^{-1} = I_{mn}$ .

Ostatní ponecháváme za cvičení 10.1. □

**Věta 10.4.** *Nechť matice  $A \in \mathbb{R}^{m \times m}$  má vlastní čísla  $\lambda_1, \dots, \lambda_m$  a matice  $B \in \mathbb{R}^{n \times n}$  má vlastní čísla  $\mu_1, \dots, \mu_n$ . Pak matice  $A \otimes B$  má vlastní čísla  $\lambda_i \mu_j \forall i, j$ .*

*Důkaz.* Víme, že pro lineární zobrazení  $x \mapsto Ax$  změna báze znamená změnu matice na  $S^{-1}AS$  pro určitou regulární  $A$ . Tedy pro lineární zobrazení (podle poznámky 10.2) změna báze prvního prostoru povede na tvar  $S^{-1}AS \otimes B$ , a podobně pro  $B$ . Pokud matice  $A, B$  převedeme na Jordanovy normální formy  $J_A, J_B$ , tak odpovídající matice bude  $J_A \otimes J_B$ . To je horní trojúhelníková matice, a proto čísla na diagonále  $\lambda_1 \mu_1, \dots, \lambda_1 \mu_n, \dots, \lambda_m \mu_1, \dots, \lambda_m \mu_n$  jsou vlastní čísla  $A \otimes B$ . □

**Důsledek 10.5.** *Platí pro matice  $A \in \mathbb{R}^{m \times m}, B \in \mathbb{R}^{n \times n}$  s vlastními čísly  $\lambda_1, \dots, \lambda_m$  a  $\mu_1, \dots, \mu_n$ :*

- (1) Jsou-li  $A, B$  pozitivně (semi-)definitní, pak  $A \otimes B$  je pozitivně (semi-)definitní,
- (2)  $\det(A \otimes B) = \det(A)^n \det(B)^m$ ,
- (3) vlastní čísla  $A \otimes B$  jsou stejná jako vlastní čísla  $B \otimes A$ .
- (4) vlastní čísla matice  $A \otimes I_n + I_m \otimes B$  jsou  $\lambda_i + \mu_j \forall i, j$ .

*Důkaz.*

- (1) Symetrie  $A \otimes B$  plyne z věty 10.3(3). Kladnost (resp. nezápornost) vlastních čísel z věty 10.4, neboť násobíme kladná (resp. nezáporná) čísla.
- (2) Zřejmé.
- (3) Zřejmé.
- (4) Analogicky jako v důkazu věty 10.4 změnou báze prvního prostoru se matice  $A \otimes I_n + I_m \otimes B$  změní na  $S^{-1}AS \otimes I_n + S^{-1}S \otimes B$ . Tudíž, pokud  $S$  reprezentuje matici upravující  $A$  na Jordanovu normální formu  $J_A$ , tak dostaneme tvar  $J_A \otimes I_n + I_m \otimes B$ . Podobně pro matici  $B$ , takže ve výsledku dostáváme  $J_A \otimes I_n + I_m \otimes J_B$ . Tato je horní trojúhelníková matice a vlastní čísla  $A \otimes B$  jsou její čísla na diagonále. □

*Poznámka.* Matice  $A \otimes I_n + I_m \otimes B$  se nazývá Kroneckerův součet matic  $A, B$ .

Podobný vztah jako pro vlastní čísla  $A \otimes B$  platí i pro singulární čísla. To ukazuje, že  $\text{rank}(A \otimes B) = \text{rank}(A) \text{rank}(B)$ .

### 10.1.1 Maticové soustavy rovnic

Kroneckerův součin je užitečným pomocníkem při zacházení s maticovými soustavami lineárních rovnic.

Uvažujme například soustavu  $AX = B$ , kde  $A, B$  jsou dané matice a  $X$  je hledaná neznámá matice. Tuto soustavu můžeme přepsat do tvaru klasické soustavy lineárních rovnic  $(I \otimes A) \text{vec}(X) = \text{vec}(B)$ , kde  $\text{vec}(\cdot)$  je operátor, který z matice vytvoří jeden dlouhý vektor tak, že bere sloupce jeden za druhým:

$$\text{vec}(B) = (b_{11}, \dots, b_{m1}, b_{12}, \dots, b_{m2}, \dots, b_{1n}, \dots, b_{mn})^T.$$

Podobně maticovou soustavu  $AX + XC = B$  můžeme přepsat na tvar  $(I \otimes A + C^T \otimes I) \text{vec}(X) = \text{vec}(B)$ , nebo soustavu  $AXC = B$  můžeme přepsat na tvar  $(C^T \otimes A) \text{vec}(X) = \text{vec}(B)$ .

**Věta 10.6.** *Soustava  $AXC = B$  je ekvivalentní se soustavou  $(C^T \otimes A) \text{vec}(X) = \text{vec}(B)$ .*

*Důkaz.*  $(i, j)$ -tá rovnice v soustavě  $AXC = B$  je

$$\sum_{k, \ell} a_{ik} x_{k\ell} c_{\ell j} = A_{i*} X C_{*j} = (AXC)_{ij} = b_{ij}.$$

Na druhou stranu, rovnice v  $j$ -tém bloku soustavy  $(C^T \otimes A) \text{vec}(X) = \text{vec}(B)$  mají tvar

$$(c_{1j} A \quad \cdots \quad c_{nj} A) \text{vec}(X) = B_{*j},$$

neboli

$$\sum_{\ell} c_{\ell j} A X_{*\ell} = B_{*j}.$$

Nyní  $i$ -tá rovnice v tomto bloku jest

$$\sum_{\ell} c_{\ell j} A_{i*} X_{*\ell} = \sum_{k, \ell} c_{\ell j} a_{ik} x_{k\ell} = b_{ij}. \quad \square$$

Přepis maticové soustavy na klasický tvar je výhodný pro zjišťování různých vlastností. Pro řešení soustavy to nemusí být nejlepší postup, protože rozměr soustavy se zvětší. Pokud například matice v soustavě  $AX = B$  jsou čtvercové řádu  $n$ , pak  $(I \otimes A) \text{vec}(X) = \text{vec}(B)$  je soustava řádu  $n^2$ , čili velikost roste s druhou mocninou.

**Poznámka 10.7.** Existují i jiná explicitní řešení maticových soustav. Například řešení Sylvesterovy rovnice  $AX - XB = C$  lze vyjádřit pomocí maticové exponenciály ze sekce 9.2 jako

$$X = - \int_0^{\infty} e^{At} C e^{-Bt} dt.$$

Asi nejdůležitější použití Kroneckerova součinu a jeho vlastností je v teorii řízení pro řešení Sylvesterovy a Lyapunovy rovnice.

**Věta 10.8** (Sylvester, 1884, Rosenblum, 1956). *Nechť  $A, B \in \mathbb{R}^{n \times n}$  nemají společné vlastní číslo. Pak Sylvesterova rovnice  $AX - XB = C$  má právě jedno řešení pro libovolné  $C \in \mathbb{R}^{n \times n}$ .*

*Důkaz.* Rovnici lze přepsat do tvaru  $(I_n \otimes A - B^T \otimes I_n) \text{vec}(X) = \text{vec}(C)$ . Podle věty 10.5(4) je matice této soustavy regulární, neboť nemá nulové vlastní číslo. Tudíž je řešení jediné.  $\square$

**Věta 10.9** (Lyapunov, 1892). *Nechť  $A \in \mathbb{R}^{n \times n}$  má vlastní čísla pouze v záporné polorovině (tj., jejich reálná část je záporná). Pak Lyapunova rovnice  $AX + XA^T = -I_n$  má právě jedno řešení. Navíc je řešení symetrická, pozitivně definitní matice.*

*Důkaz.* Podle předpokladu matice  $A, -A^T$  nemají společné vlastní číslo, čili ze Sylvesterovy-Rosenblumovy věty 10.8 existuje jediné řešení. Transpozicí soustavy dostaneme  $X^T A^T + AX^T = -I_n$ , tudíž i  $X^T$  je řešením původní soustavy. Z jednoznačnosti pak  $X = X^T$ , tedy  $X$  je symetrická.

Důkaz pozitivní definitnosti matice  $X$  vynecháme pro obecný případ, dokážeme si speciální případ když  $A$  je symetrická. Podle předpokladu je  $A$  negativně definitní. Buď  $\lambda$  vlastní číslo  $X$  a  $v$  příslušný vlastní vektor. Pak

$$\begin{aligned} 0 > -v^T v &= v^T (AX + XA^T)v = v^T AXv + v^T XA^T v \\ &= v^T AXv + v^T AX^T v = 2v^T AXv = 2\lambda v^T Av. \end{aligned}$$

Protože  $v^T Av < 0$  podle Rayleigh–Ritzovy věty 2.4, tak musí  $\lambda > 0$ . □

Řešitelnost Lyapunovy rovnice je důležitá pro stabilitu v teorii řízení. Proto se také matici, jejíž vlastní čísla leží v záporné polovině, říká *stabilní*. Jednoduché vysvětlení můžeme získat, pokud budeme uvažovat soustavu lineárních diferenciálních rovnic řádu  $n$  ve tvaru  $x' = Ax$ . Víme (sekce 9.2), že řešení je tvaru  $x(t) = e^{At}c$ . Pokud je  $A$  stabilní, tak s rostoucím časem  $t$  bod  $x(t)$  konverguje. Pro jednorozměrný případ je to snadné nahlédnout přímo, vícerozměrný případ lze dokázat pomocí Lyapunovy věty.

Přehled vlastností Kroneckerova součinu a operátoru  $\text{vec}(\cdot)$  najdeme v Henderson and Searle [1981].

## Cvičení

10.1. Dokažte zbylé vlastnosti věty 10.3.

10.2. Pro vektory  $x, y \in \mathbb{R}^n$  vyjádřete  $x \otimes y^T$ .

10.3. Ukažte, že  $\text{trace}(A \otimes B) = \text{trace}(A) \text{trace}(B)$ .

10.4. Ukažte, že  $A \otimes B = 0$  právě tehdy, když  $A = 0$  nebo  $B = 0$ .

10.5. Ukažte pro matici  $A \in \mathbb{R}^{n \times n}$ :

(a)  $(A \otimes I_m)^k = A^k \otimes I_m$ ,

(b)  $e^{A \otimes I_m} = e^A \otimes e^{I_m}$ ,

10.6. Kdy všude platí  $A \otimes B = I_n$ ?

10.7. Ukažte:

(a)  $\|A \otimes B\| = \|A\| \cdot \|B\|$  pro maticové  $\ell_p$  normy  $\|\cdot\|_{\ell_p}$  a indukované 1- a  $\infty$ -normy  $\|\cdot\|_1$  a  $\|\cdot\|_\infty$ ,

(b)  $\|A \otimes B\| = \|A\| \cdot \|B\|$  pro spektrální a Frobeniovu normu,

(c)  $\|A \otimes I_m + I_m \otimes B\| \leq \|A\| + \|B\|$  pro maticové  $\ell_1$  a  $\ell_\infty$  normy a indukované 1- a  $\infty$ -normy.

10.8. Dokažte  $e^{I_n \otimes A} = I_n \otimes e^A$  pro  $A \in \mathbb{R}^{n \times n}$ .

## 10.2 Hadamardův součin

Hadamardův součin násobí matice po složkách tak, jako je sčítáme [Horn and Johnson, 1985; Horn and Johnson, 1991].

**Definice 10.10.** Buďte  $A, B \in \mathbb{R}^{m \times n}$ . Pak *Hadamardův součin* matic  $A, B$  je matice  $A \circ B \in \mathbb{R}^{m \times n}$  definována jako  $(A \circ B)_{ij} = a_{ij}b_{ij}$ .

Hadamardův součin se dříve také nazýval Schurův součin, právě podle autora nejvýznačnějšího výsledku této oblasti.

**Věta 10.11** (Schur, 1911). Buďte  $A, B \in \mathbb{R}^{n \times n}$ .

(1) Jsou-li  $A, B$  pozitivně semidefinitní, pak  $A \circ B$  je pozitivně semidefinitní.

(2) Jsou-li  $A, B$  pozitivně definitní, pak  $A \circ B$  je pozitivně definitní.

*Důkaz.*

- (1) Buď  $A = Q\Lambda Q^T$  spektrální rozklad matice  $A$ , tedy  $a_{ij} = \sum_{k=1}^n q_{ik}\lambda_k q_{jk}$ , kde  $\lambda_1, \dots, \lambda_n \geq 0$  jsou diagonální prvky matice  $\Lambda$  (čili vlastní čísla  $A$ ). Pro vektor  $x \neq 0$  upravme

$$\begin{aligned} x^T(A \circ B)x &= \sum_{i,j=1}^n x_i x_j a_{ij} b_{ij} = \sum_{i,j=1}^n x_i x_j b_{ij} \left( \sum_{k=1}^n q_{ik}\lambda_k q_{jk} \right) \\ &= \sum_{k=1}^n \lambda_k \sum_{i,j=1}^n b_{ij} (q_{ik}x_i)(q_{jk}x_j) = \sum_{k=1}^n \lambda_k \sum_{i,j=1}^n b_{ij} y_i^{(k)} y_j^{(k)} \\ &= \sum_{k=1}^n \lambda_k (y^{(k)})^T B y^{(k)} \geq 0, \end{aligned}$$

kde  $y^{(k)} := (q_{1k}x_1, \dots, q_{nk}x_n)^T$ ,  $k = 1, \dots, n$ .

- (2) Postupujeme analogicky jako v předchozím bodu. Rovnost  $x^T(A \circ B)x = 0$  nastane jen tehdy, když  $y^{(k)} = 0$  pro všechna  $k$ . Nyní

$$0 = \sum_{i,k=1}^n (y_i^{(k)})^2 = \sum_{i,k=1}^n (q_{ik}x_i)^2 = \sum_{i=1}^n x_i^2 \left( \sum_{k=1}^n q_{ik}^2 \right) = \sum_{i=1}^n x_i^2,$$

z čehož musí  $x = 0$ . □

Hadamardův součin má samozřejmě jiné vlastnosti než standardní maticový součin, například:

- Jsou-li  $A, B$  symetrické, pak  $A \circ B$  je symetrická.
- Jsou-li  $A, B$  regulární, pak  $A \circ B$  může být singulární:  $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ ,  $B = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ .

Z mnoha dalších vlastností uvedeme ještě jednu.

**Věta 10.12.** *Budte  $A, B \in \mathbb{R}^{n \times n}$ . Pak  $\|A \circ B\|_2 \leq \|A\|_2 \cdot \|B\|_2$ .*

*Důkaz.* Podle vět 3.2, 5.9, 5.13 pro čtvercovou matici  $C$  platí  $\|C\|_2^2 = \rho(C^T C) \leq \|C^T C\|_1 \leq \|C^T\|_1 \cdot \|C\|_1$ . Tedy speciálně pro  $C := A \circ B$  dostaneme  $\|A \circ B\|_2^2 \leq \|A^T \circ B^T\|_1 \cdot \|A \circ B\|_1$ . Nyní podle Cauchyho–Schwarzovy nerovnosti a Rayleighovy–Ritzovy věty 2.4 máme

$$\begin{aligned} \|A \circ B\|_1 &= \max_{j=1, \dots, n} \sum_{i=1}^n |a_{ij} b_{ij}| = \max_{j=1, \dots, n} \langle |A_{*j}|, |B_{*j}| \rangle \leq \max_{j=1, \dots, n} \|A_{*j}\|_2 \cdot \|B_{*j}\|_2 \\ &\leq \max_{j=1, \dots, n} \|A_{*j}\|_2 \max_{j=1, \dots, n} \|B_{*j}\|_2 = \max_{j=1, \dots, n} \sqrt{(A^T A)_{jj}} \max_{j=1, \dots, n} \sqrt{(B^T B)_{jj}} \\ &= \max_{j=1, \dots, n} \sqrt{e_j^T A^T A e_j} \max_{j=1, \dots, n} \sqrt{e_j^T B^T B e_j} \\ &\leq \sqrt{\lambda_1(A^T A)} \sqrt{\lambda_1(B^T B)} = \|A\|_2 \cdot \|B\|_2. \end{aligned}$$

Ze stejného důvodu i  $\|A^T \circ B^T\|_1 \leq \|A\|_2 \cdot \|B\|_2$ . □

**Příklad 10.13** (Komprese JPEG). Hadamardův součin se používá například při ztrátové kompresi metodou JPEG. Během fáze tzv. kvantizace se jistým způsobem upravená matice obrázku Hadamardově násobí (nebo přesněji řečeno dělí) kvantizační maticí, a pak se hodnoty zaokrouhlí. To provádí část komprimace (zaokrouhlením se mnoho hodnot vynuluje, což se použije ke kompresi) a zároveň ztrátu informace – podle daného komprimačního poměru se právě sestavují kvantizační matice. Při dekompresi se jednotlivé kroky provedou v opačném pořadí, čili i zde se použije Hadamardův součin pro násobení s kvantizační maticí. □

## Cvičení

- 10.9. Klasifikujete vlastnosti operace  $\circ$ .
- 10.10. Jaké třídy matic jsou uzavřené na operaci  $\circ$  a které ne?
- 10.11. Pro  $A \in \mathbb{R}^{n \times n}$  regulární dokažte  $(A^{-T} \circ A)e = e$ .
- 10.12. Pro matice  $A, B \in \mathbb{R}^{n \times n}$  ukažte  $\|A \circ B\|_F \leq \|A\|_F \cdot \|B\|_F$ .





# Kapitola 11

## Kdo nemá dosti pozitivní semidefinitnosti

Množina pozitivně semidefinitních matic má specifickou geometrickou strukturu, kterou v této kapitole probereme podrobněji. Označme jako  $\mathcal{S}_+$  množinu pozitivně semidefinitních matic z  $\mathbb{R}^{n \times n}$ .

**Lemma 11.1.** *Bud'  $A \in \mathcal{S}_+$  a  $x \in \mathbb{R}^n$ . Pokud  $x^T A x = 0$ , pak  $Ax = 0$ .*

*Důkaz.* Rovnost  $0 = x_i^T \sqrt{A} \sqrt{A} x_i = \|\sqrt{A} x_i\|_2^2$  implikuje  $\sqrt{A} x_i = 0$ , a tedy  $Ax = 0$ .  $\square$

**Věta 11.2.** *Množina  $\mathcal{S}_+$  tvoří konvexní kužel v prostoru symetrických matic v  $\mathbb{R}^{n \times n}$ . Vnitřek  $\mathcal{S}_+$  je tvořen pozitivně definitními maticemi a hranice  $\mathcal{S}_+$  singulárními maticemi. Extrémní směry kužele jsou tvořeny pozitivně semidefinitními maticemi hodnosti 1, tedy tvaru  $xx^T$  pro  $x \neq 0$ .*

*Důkaz.* Množina  $\mathcal{S}_+$  je uzavřená na součet a nezáporné násobky, proto tvoří konvexní kužel. Vnitřek je tvořen pozitivně definitními maticemi, protože s nimi tam leží i jejich okolí.

Každá matice  $A \in \mathcal{S}_+$  se dá podle spektrálního rozkladu vyjádřit jako  $A = Q\Lambda Q^T = \sum_{i=1}^n \lambda_i q_i q_i^T$ , kde  $q_i = Q_{*i}$ . Tudíž každá pozitivně semidefinitní matice je nezápornou kombinací matic tvaru  $xx^T$ .

Zbývá ukázat, že matice tvaru  $xx^T$  jsou extrémní, tj. neleží ve vnitřku úsečky spojující dvě jiné pozitivně semidefinitní matice. Pro spor předpokládejme, že  $xx^T = \alpha A + \beta B$ , kde  $A, B \in \mathcal{S}_+$ ,  $\alpha, \beta > 0$ . Bez újmy na obecnost nechť  $\|x\|_2 = 1$  a rozšířme jej na ortonormální bázi  $x, x_2, \dots, x_n$ . Pak pro každé  $i = 2, \dots, n$  je

$$0 = (x_i^T x)(x^T x_i) = x_i^T (xx^T) x_i = x_i^T (\alpha A + \beta B) x_i = \alpha x_i^T A x_i + \beta x_i^T B x_i \geq 0.$$

Tudíž

$$0 = x_i^T A x_i = x_i^T B x_i \geq 0.$$

z čehož  $0 = Ax_i = Bx_i$  podle Lemmatu 11.1. To znamená, že matice  $A, B$  mají hodnotu nanejvýš 1; stačí uvažovat případ, kdy je přesně 1. Protože  $\text{Ker}(A) = \text{Ker}(B) = \text{Ker}(xx^T)$ , platí rovnost i pro jejich ortogonální doplňky  $\mathcal{S}(A) = \mathcal{S}(B) = \mathcal{S}(xx^T) = \text{span}\{x\}$ . Tudíž matice  $A, B$  musí být kladné násobky matice  $xx^T$ .  $\square$

*Poznámka.* Pokud množinu  $\mathcal{S}_+$  řízeme nadrovinou, dostaneme geometrické těleso, nazývané spektrální edr (ang. spectrahedron).

Positivní (semi-)definitnost umožňuje zavést zajímavou maticovou relaci.

**Definice 11.3.** *Bud'  $A, B \in \mathbb{R}^{n \times n}$  symetrické. Pak  $A \succeq B$  pokud  $A - B$  je pozitivně semidefinitní, a  $A \succ B$  pokud  $A - B$  je pozitivně definitní.*

Nyní pozitivní semidefinitnost matice  $A$  lze značit jako  $A \succeq 0$ . Relace  $\succeq$  určuje částečné uspořádání a relace  $\succ$  určuje ostré uspořádání. Tato relace se nazývá také Löwnerovo uspořádání (podle amerického matematika českého původu Karla Löwnera).

**Věta 11.4.** *Relace  $\succeq$  představuje částečné uspořádání a relace  $\succ$  ostré uspořádání na třídě symetrických matic řádu  $n$ .*

*Důkaz. Reflexivita.* Platí triviálně  $A \succeq A$ .

*Transitivita.* Pokud  $A \succeq B \succeq C$ , tak matice  $A - B, B - C$  jsou pozitivně semidefinitní. Tudíž i jejich součet  $A - B + B - C = A - C$  je pozitivně semidefinitní.

*Anti-symetrie.* Pokud  $A \succeq B$  a zároveň  $B \succeq A$ , tak matice  $A - B$  má nezáporná a zároveň nekladná vlastní čísla, tudíž jsou všechna nulová a proto  $A - B = 0$ .

Analogicky pro relaci  $\succ$ . □

**Pozorování 11.5.** *Budte  $A, B, C, D \in \mathbb{R}^{n \times n}$  symetrické. Pak platí:*

- (1)  $A \succeq B \Rightarrow SAS^T \succeq SBS^T$  pro každou  $S \in \mathbb{R}^{n \times n}$ ,
- (2)  $A \succ B \Rightarrow SAS^T \succ SBS^T$  pro každou regulární  $S \in \mathbb{R}^{n \times n}$ ,
- (3)  $A \succeq B, C \succeq D \Rightarrow A + C \succeq B + D$ .

**Věta 11.6.** *Budte  $A, B \in \mathbb{R}^{n \times n}$  symetrické a  $A \succ 0$ . Pak  $A, B$  jsou současně diagonalizovatelné (podle teorie kvadratických forem), tj. existuje regulární  $U$  taková, že  $U^T A U$  a  $U^T B U$  jsou diagonální.*

*Důkaz.* Označme  $V := \sqrt{A}^{-1}$ , pak  $V^T A V = I_n$ . Matice  $V^T B V$  je symetrická, tudíž má spektrální rozklad  $V^T B V = Q \Lambda Q^T$ . Nyní stačí volit  $U := V Q$ , neboť  $U^T A U = Q^T V^T A V Q = Q^T I_n Q = I_n$  a  $U^T B U = Q^T V^T B V Q = Q^T Q \Lambda Q^T Q = \Lambda$ . □

**Věta 11.7.** *Budte  $A \succ 0, B \succeq 0$ . Pak*

- (1)  $A \succeq B$  právě tehdy, když  $\rho(A^{-1}B) \leq 1$ ,
- (2)  $A \succ B$  právě tehdy, když  $\rho(A^{-1}B) < 1$ .

*Důkaz.*

- (1) Podle důkazu věty 11.6 lze vyjádřit  $A = U U^T, B = U D U^T$ , kde  $D$  je diagonální matice s nezápornou diagonálou. Nyní  $A \succeq B$  právě tehdy, když  $U(I - D)U^T \succeq 0$ , neboli ekvivalentně  $I - D \succeq 0$ , což znamená  $d_{ii} \leq 1$ . Vlastní čísla matice  $A^{-1}B = U^{-T}U^{-1}U D U^T = U^{-T}D U^T$  jsou  $d_{11}, \dots, d_{nn}$  a tudíž platí  $\rho(A^{-1}B) \leq 1$  právě tehdy, když  $d_{ii} \leq 1$ .
- (2) Analogicky. □

**Věta 11.8.** *Budte  $A, B \succ 0$ . Pak  $A \succeq B \Leftrightarrow A^{-1} \preceq B^{-1}$ .*

*Důkaz.* Podle věty 11.7 je  $A \succeq B$  právě tehdy, když  $\rho(A^{-1}B) \leq 1$ . Protože  $\rho(A^{-1}B) = \rho(AA^{-1}BA^{-1}) = \rho((B^{-1})^{-1}A^{-1})$ , tak ekvivalentně podle té samé věty je  $A^{-1} \preceq B^{-1}$ . □

**Věta 11.9.** *Budte  $A, B \in \mathbb{R}^{n \times n}$  symetrické a  $A \succeq B$ . Pak  $\lambda_i(A) \geq \lambda_i(B)$  pro všechna  $i = 1, \dots, n$ .*

*Důkaz.* Podle Courantovy–Fischerovy věty 2.6 je

$$\lambda_i(A) = \max_{V \in \mathbb{R}^n: \dim V = i} \min_{x \in V: \|x\|_2 = 1} x^T A x \geq \max_{V \in \mathbb{R}^n: \dim V = i} \min_{x \in V: \|x\|_2 = 1} x^T B x = \lambda_i(B). \quad \square$$

Obrácená implikace obecně neplatí, protipříklad:  $A = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}, B = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}$ .

Z výše uvedených vlastností můžeme odvozovat další vlastnosti. Například  $A \succeq B \succeq 0$  implikuje  $\det(A) \geq \det(B)$ ,  $\text{trace}(A) \geq \text{trace}(B)$ , ...

**Věta 11.10** (Hadamardova nerovnost).

- (1) Je-li  $A \succ 0$ , pak  $\det(A) \leq \prod_{i=1}^n a_{ii}$ .
- (2) Je-li  $B \in \mathbb{R}^{n \times n}$ , pak  $|\det(B)| \leq \prod_{i=1}^n \|B_{*i}\|_2$ .

*Důkaz.*

- (1) Definujme  $D := \text{diag}(a_{11}^{-1/2}, \dots, a_{nn}^{-1/2})$ . Pak podmínka  $\det(A) \leq \prod_{i=1}^n a_{ii}$  je ekvivalentní s podmínkou  $\det(DAD) \leq 1$ , neboť  $\det(DAD) = \det(A) \det(D)^2 = \det(A) \prod_{i=1}^n a_{ii}^{-1}$ . Budte  $\lambda_1, \dots, \lambda_n$  vlastní čísla  $DAD$ , pak podle AG nerovnosti (mezi geometrickým a aritmetickým průměrem) je

$$\det(DAD) = \prod_{i=1}^n \lambda_i \leq \left(\frac{1}{n} \sum_{i=1}^n \lambda_i\right)^n = \left(\frac{1}{n} \text{trace}(DAD)\right)^n = 1.$$

- (2) Pro  $B$  singulární platí odhad triviálně, takže uvažujme  $B$  regulární. Definujme  $A := B^T B \succ 0$ . Pak nerovnost  $\det(A) \leq \prod_{i=1}^n a_{ii}$  má tvar  $\det(B)^2 = \det(B^T B) \leq \prod_{i=1}^n \|B_{*i}\|_2^2$ , z čehož odmocněním máme požadovanou nerovnost.  $\square$

Význam druhé nerovnosti lze dobře ilustrovat geometricky. Determinant je objem rovnoběžnostěny s hranami odpovídajícím sloupcům matice  $B$ . Pravá strana pak určuje objem kvádrů s hranami odpovídajícím sloupcům matice  $B$ . Z toho vidíme, že nerovnost se nabyde jako rovnost, pokud sloupce matice  $B$  jsou na sebe kolmé (například, je-li  $B$  ortogonální).

## Cvičení

11.1. Dokažte vlastnosti z pozorování 11.5.

11.2. Dokažte pro relaci  $\succeq$  a Hadamardův součin:

$$(a) A \succeq B, C \succeq 0 \Rightarrow A \circ C \succeq B \circ C,$$

$$(b) A \succeq B \succeq 0, C \succeq D \succeq 0 \Rightarrow A \circ C \succeq B \circ D \succeq 0,$$

11.3. Bez použití věty 11.9 dokažte  $A \succ B \Rightarrow \det(A) > \det(B)$ .

11.4. Odvoďte větu 11.8 přímo z věty 11.6 bez použití věty 11.7.

11.5. Připomeňme, že pro každý vektor  $x \in \mathbb{R}^n$  je  $\|x\|_\infty \leq \|x\|_2 \leq \|x\|_1$ . Odvoďte odhad pro determinant  $\det(B) \leq \prod_{i=1}^n \|B_{*i}\|_1$  na základě Gerschgorinových disků a ukažte, že je méně těsný, než ten z Hadamardovy nerovnosti (věta 11.10). Ukažte dále, že ještě těsnější odhad  $\det(B) \leq \prod_{i=1}^n \|B_{*i}\|_\infty$  již obecně neplatí.

11.6. Uvažujme symetrickou matici v blokovém tvaru

$$A = \begin{pmatrix} A_{11} & \dots & A_{1n} \\ \vdots & & \vdots \\ A_{1n}^T & \dots & A_{nn} \end{pmatrix},$$

kde diagonální bloky  $A_{11}, \dots, A_{nn}$  jsou čtvercové matice.

(a) Dokažte, že  $A \succ 0$  implikuje  $\det(A) \leq \prod_{i=1}^n \det(A_{ii})$ .

(b) Ukažte  $(A^2)_{11} \succeq (A_{11})^2$ .



# Kapitola 12

## Speciální matice

V řadě situací se setkáváme se speciálními typy matic a využíváme jejich specifických vlastností. Typy matic, jako například regulární, ortogonální, pozitivně definitní aj., již dobře známe a v následujícím se zaměříme na další důležité třídy matic.

### 12.1 M-matice

Základní zdroje: Horn and Johnson [1991]; Meyer [2000].

**Motivace k M-maticím.** Uvažujme čtvercovou soustavu lineárních rovnic  $Ax = b$ . Z  $i$ -té rovnice vyjádříme  $i$ -tou proměnnou  $x_i$

$$x_i = \frac{1}{a_{ii}} \left( b_i - \sum_{j \neq i} a_{ij} x_j \right), \quad i = 1, \dots, n.$$

Tento tvar nás přivádí k Jacobiho iterativní metodě na řešení soustavy. Jako počáteční vektor zvolíme libovolné  $x^0 \in \mathbb{R}^n$ . V  $k$ -té iteraci vyrobíme z vektoru  $x^{k-1}$  vektor  $x^k$  takto

$$x_i^k := \frac{1}{a_{ii}} \left( b_i - \sum_{j \neq i} a_{ij} x_j^{k-1} \right), \quad i = 1, \dots, n.$$

Za určitých předpokladů posloupnost  $x^0, x^1, \dots$  konverguje k řešení  $A^{-1}b$ .

Iteraci můžeme vyjádřit maticově. Rozložíme  $A = D + A'$  na diagonální a nediagonální část, tedy

$$D = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & a_{nn} \end{pmatrix}, \quad A' = \begin{pmatrix} 0 & a_{12} & \dots & a_{1n} \\ a_{21} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_{n-1,n} \\ a_{n1} & \dots & a_{n,n-1} & 0 \end{pmatrix}.$$

Pak rovnice  $Ax = b$  má tvar  $Dx + A'x = b$ , neboli  $Dx = b - A'x$ , a tedy  $k$ -tá iterace bude

$$x^k := D^{-1}(b - A'x^{k-1}) = -D^{-1}A'x^{k-1} + D^{-1}b.$$

Nic nás ale neomezuje rozložit matici  $A$  pouze podle diagonály. Uvažujme tedy obecný rozklad matice  $A = M - N$ , kde chceme, aby  $M$  byla snadno invertovatelná (tedy diagonální, trojúhelníková či jinak speciální). Pak  $k$ -tá iterace této variace má tvar

$$x^k := M^{-1}Nx^{k-1} + M^{-1}b.$$

Postačující podmínku pro konvergenci dává následující věta.

**Věta 12.1.** *Je-li  $\rho(M^{-1}N) < 1$ , pak  $A$  je regulární a  $x^k \rightarrow_{k \rightarrow \infty} A^{-1}b$  pro každý počáteční vektor  $x^0$  a pro každé  $b$ .*

*Důkaz.* Ze vztahu  $A = M - N$  vyjádříme  $M^{-1}A = I_n - M^{-1}N$ . Protože  $\rho(M^{-1}N) < 1$ , má matice  $I_n - M^{-1}N$  všechna vlastní čísla s kladnou reálnou částí, a tedy je regulární. Proto i matice  $A$  je regulární.

Označme  $x^* = A^{-1}b$  a uvažujme

$$\begin{aligned} x^1 - x^* &= M^{-1}Nx^0 + M^{-1}b - x^* = M^{-1}Nx^0 + M^{-1}(A - M)x^* = \\ &= M^{-1}Nx^0 - M^{-1}Nx^* = M^{-1}N(x^0 - x^*). \end{aligned}$$

Tedy pro libovolnou indukovanou maticovou normu platí

$$\|x^1 - x^*\| = \|M^{-1}N(x^0 - x^*)\| \leq \|M^{-1}N\| \cdot \|x^0 - x^*\|,$$

z čehož

$$\|x^k - x^*\| \leq \|M^{-1}N\|^k \cdot \|x^0 - x^*\|.$$

Podle věty 5.15 víme, že existuje indukovaná maticová norma taková, že  $\|M^{-1}N\| < 1$ . Tudíž vzdálenost  $x^k$  od řešení  $x^*$  ze zkracuje s geometrickou řadou.  $\square$

**Poznámka 12.2.** Věta platí i naopak. Je-li  $\rho(M^{-1}N) \geq 1$ , pak matice  $A$  nemusí být regulární, o čemž nás přesvědčí příklad  $M = N = I_n$ . Pokud je matice  $A$  regulární, i pak posloupnost  $x^0, x^1, \dots$  typicky nekonverguje pro všechna  $x^0$  a  $b$ . Necht'  $\rho(M^{-1}N)$  se nabyde pro reálné vlastní číslo  $|\lambda| \geq 1$  a necht' mu přísluší vlastní vektor  $v$ , tedy  $M^{-1}Nv = \lambda v$ . Potom pro volbu  $b := 0$  a  $x^0 := v$  posloupnost nekonverguje k řešení  $A^{-1}b = 0$ , ale skládá se z nezmenšujících se násobků vektoru  $v$ .

Pokud  $\rho(M^{-1}N)$  se nabyde jen pro komplexní vlastní číslo  $\lambda \pm i\nu$ , je postup trochu složitější. Podle věty 2.3 o reálném Schurově rozkladu existuje regulární  $S \in \mathbb{R}^{n \times n}$  a blokově horní trojúhelníková  $T \in \mathbb{R}^{n \times n}$  takové, že  $M^{-1}N = STS^{-1}$ . Navíc první blok matice  $T$  je tvaru  $B = \begin{pmatrix} \lambda & \nu \\ -\nu & \lambda \end{pmatrix}$ . Zvolme  $b := 0$  a  $x^0 := S(y^T, 0)^T = S(y_1, y_2, 0, \dots, 0)^T \neq 0$ . Nyní  $x^1 = M^{-1}Nx^0 = ST(y^T, 0)^T = S((By)^T, 0)^T$ . Vektory  $x^k$  jsou stejného tvaru jako lineární kombinace prvních dvou sloupců matice  $S$ . Nyní je důležité analyzovat vektor  $By$ . Snadno nahlédneme, že  $\|By\|_2 = \rho(B)\|y\|_2$ , tedy postupnými iteracemi se norma vektoru  $By$  zvětšuje, a tím i (globálně) norma  $x^k$ .

**Konečně M-matice.** M-matice představují mj. třídu matic, pro které Jacobiho metoda konverguje s tím, že za matici  $M$  stačí zvolit vhodný (tj. dostatečně velký) násobek jednotkové matice. Dokonce i každý jiný rozklad  $A = M - N$ , kde  $M$  je regulární a  $N$  nezáporná, splní svůj účel. Význam M-matic však překračuje tyto vlastnosti.

**Definice 12.3.** Matice  $A \in \mathbb{R}^{n \times n}$  je *M-maticí* pokud  $a_{ij} \leq 0$  pro  $i \neq j$  a  $A^{-1} \geq 0$ .

Zkratka z názvu pochází z roku 1937 od Alexandera Ostrowského a má připomínat Hermanna Minkowského, který ukázal první vlastnosti. M-matice se dají charakterizovat mnoha ekvivalentními podmínkami, ukazujícími, jak se na jednu vlastnost dá pohlížet z mnoha jiných úhlů pohledu [Horn and Johnson, 1991].

**Věta 12.4** (Ekvivalentní podmínky M-matice). *Bud'  $A \in \mathbb{R}^{n \times n}$  taková, že  $a_{ij} \leq 0$ . Pak následující podmínky jsou ekvivalentní:*

- (1)  $A$  je M-matice, tj.  $A^{-1} \geq 0$ ,
- (2) lze vyjádřit  $A = \alpha I_n - B$ , kde  $B \geq 0$ ,  $\rho(B) < \alpha$ ,
- (3) lze vyjádřit  $A = M - N$ , kde  $M^{-1} \geq 0$ ,  $N \geq 0$ ,  $\rho(M^{-1}N) < 1$ ,
- (4) reálné části vlastních čísel matice  $A$  jsou kladné,
- (5) reálná vlastní čísla matice  $A$  jsou kladná,
- (6) existuje LU rozklad  $A = LU$ , kde  $L, U$  jsou M-matice,
- (7) existuje  $x > 0$ :  $Ax > 0$ ,
- (8)  $A^{-1}e > 0$ ,
- (9) pokud  $Ax \geq 0$ , pak  $x \geq 0$ ,

(10) všechny vedoucí hlavní podmatice matice  $A$  mají kladný determinant.

*Důkaz.* Dokážeme vybrané vlastnosti.

**Vlastnost (2)**

(1)  $\Rightarrow$  (2): Definujme  $\alpha := \max_i |a|_{ii}$  a  $B := \alpha I_n - A \geq 0$ . Dále označme  $x := A^{-1}e > 0$ . Pak  $(\alpha I_n - B)x = AA^{-1}e = e > 0$ , z čehož  $Bx < \alpha x$ . Podle Perronovy teorie (věta 8.6) je  $\rho(B) < \alpha$ .

(2)  $\Rightarrow$  (1): Z věty 5.18 o Neumannových řadách je  $A^{-1} = (\alpha I_n - B)^{-1} = \frac{1}{\alpha}(I_n - \frac{1}{\alpha}B)^{-1} = \frac{1}{\alpha} \sum_{k=0}^{\infty} (\frac{1}{\alpha}B)^k \geq 0$ .

**Vlastnost (3)**

(1)  $\Rightarrow$  (3): Plyne z vlastnosti (2), stačí volit  $M := \alpha I_n$ ,  $N := B$ .

(3)  $\Rightarrow$  (1): Dle věty 5.18 o Neumannových řadách je  $A^{-1} = (M - N)^{-1} = (I_n - M^{-1}N)^{-1}M^{-1} = (\sum_{k=0}^{\infty} (M^{-1}N)^k) M^{-1} \geq 0$ .

**Vlastnost (4)**

(2)  $\Rightarrow$  (4): Vlastní čísla  $A$  jsou v zásadě vlastní čísla  $-B$  zvětšená o  $\alpha$ , tedy jsou napravo od počátku.

(4)  $\Rightarrow$  (2): Definujme  $\alpha := \max_i |a|_{ii}$  a  $B := \alpha I_n - A \geq 0$ . Matice  $B$  má reálné části vlastních čísel menší než  $\alpha$ . Podle Perronovy věty 8.1 je dominantní vlastní číslo matice  $B$  nezáporné, tedy  $\rho(B) < \alpha$ .

**Vlastnost (5)**

Analogicky.

**Vlastnost (6)**

(1)  $\Rightarrow$  (6): Při Gaussově eliminaci není potřeba permutovat řádky a stačí používat jen úpravu přičtení kladného násobku řádku s pivotem k řádku pod ním. M-maticovost se neztratí, tedy výsledná matice  $U$  bude M-maticí. Podobně pro  $L$ .

(6)  $\Rightarrow$  (1):  $A^{-1} = U^{-1}L^{-1} \geq 0$ .

**Vlastnost (7)**

(1)  $\Rightarrow$  (7): Stačí volit  $x := A^{-1}e > 0$ .

(7)  $\Rightarrow$  (2): Definujme  $\alpha := \max_i |a|_{ii}$  a  $B := \alpha I_n - A \geq 0$ . Pak  $Ax > 0$  má tvar  $Bx < \alpha x$ , tedy dle Perronovy teorie (věta 8.6) je  $\rho(B) < \alpha$ .

**Vlastnost (8)**

(1)  $\Rightarrow$  (8): Jasně z nezápornosti a regularity  $A^{-1}$ .

(8)  $\Rightarrow$  (7): Dosad'  $x := A^{-1}e > 0$ .

**Vlastnost (9)**

(1)  $\Rightarrow$  (9): Je-li  $Ax \geq 0$ , pak  $x = A^{-1}Ax \geq 0$ .

(9)  $\Rightarrow$  (1): Matice  $A$  musí být regulární, neboť jinak  $Ax = 0$  pro  $x \neq 0$  a jeden z vektorů  $\pm x$  není nezáporný. Nyní pro každé  $i = 1, \dots, n$  definuj  $x^i := A_{*i}^{-1} = A^{-1}e_i$ . Protože  $Ax^i = AA^{-1}e_i = e_i \geq 0$ , je z předpokladu  $0 \leq x^i = A_{*i}^{-1}$ . Všechny sloupce  $A^{-1}$  jsou tak nezáporné.

**Vlastnost (10)**

Viz Meyer [2000]. □

**Pozorování 12.5.** Symetrické M-matice jsou pozitivně definitní.

Ohledně umístění vlastních čísel M-matice lze odvodit ještě silnější výsledky. Pro  $n > 2$  se dá ukázat [Horn and Johnson, 1991], že úhel v počátku mezi reálnou osou a přímkou k vlastnímu číslu je menší než  $\frac{\pi}{2} - \frac{\pi}{n}$ .

**Příklad 12.6** (Laplacova matice grafu). Laplacova matice grafu  $G = (V, E)$  s  $n$  vrcholy je matice  $L \in \mathbb{R}^{n \times n}$  definovaná takto

$$L_{ij} = \begin{cases} \deg(i), \text{ stupeň vrcholu} & \text{pokud } i = j \\ -1 & \text{pokud } (i, j) \in E, \\ 0 & \text{pokud } (i, j) \notin E. \end{cases}$$

Matice  $L$  popisuje plně graf  $G$  a některé vlastnosti grafu z ní snadno zjistíme. Například počet koster grafu  $G$  je roven determinantu matice  $L$  po vyškrtnutí libovolného řádku a sloupce.

Laplacova matice sice není M-maticí podle definice (nazývá se *singulární M-matice*), ale má nekladné mimodiagonální prvky a podmínka (2) věty 12.4 platí s rovností ( $\rho(B) = \alpha$ ). Tudíž má matice  $L$  mnoho vlastností společných či podobných M-maticím. □

**Příklad 12.7** (Problém lineární komplementarity). Problém lineární komplementarity je úloha přípustnosti systému

$$y = Ax + b, \quad x, y \geq 0, \quad x^T y = 0,$$

kde  $A \in \mathbb{R}^{n \times n}$  a  $b \in \mathbb{R}^n$  jsou dané a  $x, y \in \mathbb{R}^n$  jsou proměnné. Omezení jsou lineární až na podmínku  $x^T y = 0$ . Této podmínce se říká podmínka komplementarity, protože ekvivalentně říká, že pro každé  $i$  je  $x_i = 0$  nebo  $y_i = 0$ .

Tato úloha je NP-těžká a objevuje se v podmínkách optimality v optimalizaci (kvadratické programování, přípustnost celočíselného programování) či k vyjádření Nashova ekvilibria bimaticových her.

Pokud  $A$  je M-matice, pak úloha lineární komplementarity má právě jedno řešení pro každé  $b \in \mathbb{R}^n$  (ve skutečnosti tuto vlastnost splňuje větší třída matic, tzv. P-matice). Navíc je řešení efektivně naležitelné.  $\square$

M-matice se objevují v řadě dalších situací. Mayer [2017] ukazuje příklady z diskretizace určitých typů diferenciálních rovnic.

## Cvičení

- 12.1. Ukažte, že hlavní podmatice M-matice jsou M-matice.
- 12.2. Ukažte ekvivalentní charakterizaci M-matic: existuje  $x \geq 0 : Ax > 0$ .
- 12.3. Buďte  $A, B$  dvě M-matice. Je  $\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$  také M-matice?
- 12.4. Rozhodněte, zda jsou (a případně za jakých podmínek jsou) M-matice uzavřené na:
  - (a) transpozici,
  - (b) součet,
  - (c) součin,
  - (d) shodnou permutaci řádků a sloupců.
- 12.5. Kdy je horní trojúhelníková matice M-maticí?
- 12.6. Ukažte, že vlastní číslo M-matice  $A$  s nejmenší reálnou částí je reálné.
- 12.7. Nechť  $A$  je M-maticí,  $B \geq A$ ,  $b_{ij} \leq 0$  pro  $i \neq j$ . Dokažte:
  - (a)  $B$  je M-maticí,
  - (b)  $A^{-1} \geq B^{-1}$ .
- 12.8. Ukažte, že Jacobiho metoda konverguje pokud je  $A$  diagonálně dominantní, tj.  $|a_{ii}| > \sum_{j \neq i} |a_{ij}|$  pro všechna  $i = 1, \dots, n$ .
- 12.9. Ukažte přímo, že problém lineární komplementarity má jednoznačné řešení a najděte jeho vyjádření pro případ, když  $A$  je M-matice a  $b \leq 0$ .



# Kapitola 13

## Další témata

### Aplikace SVD rozkladu

Těch je mnoho, například:

- *Lawrence Sirovich: A pattern analysis of the second Rehnquist U.S. Supreme Court*, 2003, <http://www.pnas.org/content/100/13/7432.abstract>  
Analýza nezávislosti hlasování soudců nejvyššího soudu za 2. předsednictví soudce Rehnquista.
- *Barry A. Cipra: Blockbuster Algorithm*, SIAM news, 2009, <https://www2.bc.edu/~baglivo/MT210/SVDBlockbuster.pdf>  
Cena Netflixu za milion dolarů o vylepšení aspoň o 10% jejich systém, kterým Netflix doporučuje zákazníkům filmy.

### Numerický range

Numerický range matice  $A \in \mathbb{C}^{n \times n}$  je množina komplexních čísel

$$F(A) := \{x^*Ax; x \in \mathbb{C}^n, x^*x = 1\}.$$

Protože tato množina zahrnuje vlastní čísla matice  $A$ , hodí se k jejich odhadu a k odhadu spektrálního poloměru  $\rho(A)$ . Pro symetrické resp. normální matice má  $F(A)$  specifický tvar.

Více: [Horn and Johnson, 1991, kap. 1]

### Zobecněná vlastní čísla

Mějme matice  $A, B \in \mathbb{R}^{n \times n}$ . Pak  $\lambda \in \mathbb{C}$  je vlastním číslem a  $0 \neq x \in \mathbb{C}^n$  příslušným vlastním vektorem páru  $(A, B)$  pokud  $Ax = \lambda Bx$ . Tedy pro  $B = I_n$  dostaneme klasické vlastní číslo a vektor.

Jaké jsou vlastnosti a jak najít zobecněná vlastní čísla? A jaká je motivace?

### Stabilní matice

Matice  $A \in \mathbb{R}^{n \times n}$  je (Hurwitzovsky) stabilní pokud reálné části všech vlastních čísel jsou záporné.

Motivace pochází z konvergence řešení diferenciálních rovnic. Uvažme kupříkladu diferenciální rovnici

$$x' = Ax.$$

Její řešení, vektorová funkce  $x(t)$  závislá na čase  $t$ , je  $x(t) = e^{At}x(0)$ . Pozice  $x(t)$  pak v čase dospěje do rovnovážného stavu, nebo k němu konverguje, právě tehdy, když  $A$  je stabilní.

Ekvivalentní charakterizace stability matice  $A$  je Lyapunova věta, která tvrdí, že  $A$  je stabilní právě tehdy, když soustava  $XA + A^T X = -I$  má jako řešení pozitivně definitní  $X$ .

## Iterativní metody na řešení soustav lineárních rovnic

Chceme řešit soustavu rovnic  $Ax = b$ . Rozložíme  $A = M - N$ , kde  $M$  se jednoduše invertuje (pokud  $M$  je diagonální s diagonálou rovnou diagonále  $A$ , pak dostaneme Jacobiho metodu). Soustavu pak přepíšeme  $x = M^{-1}(Nx + b)$ . To nás vede na následující iterativní algoritmus.

Zvolme  $x^0$ , položme  $k = 0$  a iterujme

$$\begin{aligned}x^{k+1} &:= M^{-1}(Nx^k + b), \\k &:= k + 1,\end{aligned}$$

Za jakých předpokladů algoritmus konverguje k řešení?

## Normální matice

Jak víme, matice je normální pokud její Schurův rozklad (věta 2.1) má diagonální matici. Ekvivalentní, (a překvapivá, snadno ověřitelná), charakterizace je, že  $A^T A = AA^T$ . Normální matice mají řadu pěkných vlastností – podobně jako symetrické matice, ale tvoří větší třídu matic.

Více: Horn and Johnson [1985]

## Netradiční maticový součin – Hadamardův

Hadamardův součin je definovaný podobně jako součet. To jest, pro matice  $A, B \in \mathbb{R}^{m \times n}$  definujeme  $A \circ B \in \mathbb{R}^{m \times n}$  jako

$$(A \circ B)_{ij} = a_{ij}b_{ij}, \quad i, j = 1, \dots, n.$$

Přestože Hadamardův součin nedosahuje takového významu jako klasický maticový součin, je to pomůcka, která se občas hodí. Jak jinak vyjádřit elegantní Schurovu větu tvrdící, že pokud matice  $A, B$  jsou pozitivně definitní, potom i  $A \circ B$  je pozitivně definitní?

Více: [Horn and Johnson, 1991, kap. 5]

## Teorie náhodných matic

Jaká je distribuce vlastních čísel náhodných matic?

[http://www-math.mit.edu/~edelman/publications/random\\_matrix\\_theory\\_innovative.pdf](http://www-math.mit.edu/~edelman/publications/random_matrix_theory_innovative.pdf)

## Totální nejmenší čtverce

Metoda nejmenších čtverců hledá přibližné řešení (typicky neřešitelné a přeuredené) soustavy  $Ax = b$  pomocí optimalizační úlohy

$$\min\{\|Ax - b\|_2; x \in \mathbb{R}^n\}$$

Tuto úlohu lze ekvivalentně vyjádřit jako

$$\min\{\|b'\|_2; Ax = b + b', x \in \mathbb{R}^n\}.$$

Jinými slovy, hledáme nejmenší (v eukleidovské normě) vektor  $b'$  takový, že když změním pravou stranu soustavy o tento vektor, soustava se stane řešitelnou.

Zde se nabízí přirozené rozšíření, a to hledat nejmenší změnu všech koeficientů, to jest matice  $(A | b)$ , tak, aby se soustava stala řešitelnou. Matematicky vyjádřeno jako optimalizační úloha:

$$\min\{\|(A' | b')\|; (A + A')x = b + b', x \in \mathbb{R}^n\}.$$

Toto se nazývá metoda totálních nejmenších čtverců a jako maticová norma používá Frobeniova či spektrální.

## Tenzory

O tenzorech skoro každý slyšel, ale málokdo přesně ví, co vlastně jsou.

## Nekonečné matice

# Značení

## Množiny

$\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$  množina přirozených, celých, racionálních a reálných čísel  
 $U + V$  součet množin,  $U + V = \{u + v; u \in U, v \in V\}$

## Matice a vektory

$\text{rank}(A)$  hodnost matice  $A$   
 $\text{trace}(A)$  stopa matice  $A$ ,  $\text{trace}(A) = \sum_i a_{ii}$   
 $\rho(A)$  spektrální poloměr matice  $A$ ,  $\rho(A) = \max\{|\lambda|; \lambda \text{ je vlastní číslo } A\}$   
 $\lambda_1(A) \geq \dots \geq \lambda_n(A)$  vlastní čísla symetrické matice  $A$ ,  
 $\sigma_1(A) \geq \dots \geq \sigma_n(A)$  singulární čísla matice  $A$ ,  
 $k(A)$  číslo podmíněnosti matice  $A$ , str. 35  
 $A^T$  transpozice matice  $A$   
 $A^*$  Hermitovská transpozice matice  $A$   
 $A^\dagger$  Mooreova–Penroseova pseudoinverze matice  $A$ , str. 21  
 $A^D$  Drazinova pseudoinverze matice  $A$ , str. 23  
 $A \otimes B$  Kroneckerův součin matic  $A, B$ , str. 59  
 $A \circ B$  Hadamardův součin matic  $A, B$ , str. 62  
 $A \geq B$  nezápornost matice  $A - B$ , tj.  $a_{ij} \geq b_{ij}$   
 $A > B$  kladnost matice  $A - B$ , tj.  $a_{ij} > b_{ij}$   
 $A \succeq B$  pozitivní semidefinitnost matice  $A - B$ , str. 65  
 $A \succ B$  pozitivní definitnost matice  $A - B$ , str. 65  
 $A_{i*}$   $i$ -tý řádek matice  $A$   
 $A_{*j}$   $j$ -tý sloupec matice  $A$   
 $\text{diag}(v)$  diagonální matice s diagonálními prvky  $v_1, \dots, v_n$   
 $0_n, 0$  nulová matice (všechny složky jsou rovny 0)  
 $1_n$  jedničková matice (všechny složky jsou rovny 1)  
 $I_n, I$  jednotková matice (diagonální s jedničkami na diagonále)  
 $e_i$  jednotkový vektor,  $e_i = I_{*i} = (0, \dots, 0, 1, \dots, 0)^T$   
 $e$  vektor ze samých jedniček,  $e = (1, \dots, 1)^T$

## Prostory

$\mathcal{S}(A)$  sloupcový prostor matice  $A$ ,  $\mathcal{S}(A) = \{Ax; x \in \mathbb{R}^n\}$   
 $\mathcal{R}(A)$  řádkový prostor matice  $A$ ,  $\mathcal{R}(A) = \mathcal{S}(A^T)$   
 $\text{Ker}(A)$  jádro matice  $A$ ,  $\text{Ker}(A) = \{x; Ax = 0\}$   
 $M^\perp$  ortogonální doplněk množiny vektorů  $M$ ,  $M^\perp := \{x \in V; \langle x, y \rangle = 0 \forall y \in M\}$



# Literatura

- A. Ben-Israel and T. N. Greville. *Generalized inverses. Theory and applications*. Springer, New York, 2nd edition, 2003. <http://www.maths.ed.ac.uk/~aar/papers/bengrev.pdf>. 21
- J. W. Demmel. The probability that a numerical analysis problem is difficult. *Math. Comput.*, 50(182):449–480, 1988. <https://pdfs.semanticscholar.org/d082/04cb86183544f1c20b1a2b861013675dfc74.pdf>. 36
- M. P. Drazin. Pseudo-inverses in associative rings and semigroups. *Am. Math. Monthly*, 65(7):506–514, 1958. <https://doi.org/10.2307/2308576>. 23
- E. Estrada and J. A. Rodríguez-Velázquez. Subgraph centrality in complex networks. *Phys. Rev. E*, 71(5):056103:1–9, 2005. <https://link.aps.org/doi/10.1103/PhysRevE.71.056103>. 56
- H. V. Henderson and S. R. Searle. The vec-permutation matrix, the vec operator and Kronecker products: a review. *Linear Multilinear Algebra*, 9(4):271–288, 1981. 62
- J. M. Hendrickx and A. Olshevsky. Matrix  $p$ -norms are NP-hard to approximate if  $p \neq 1, 2, \infty$ . *SIAM J. Matrix Anal. Appl.*, 31(5):2802–2812, 2010. 27
- N. J. Higham. *Accuracy and stability of numerical algorithms*. SIAM, Philadelphia, 1996. 27
- N. J. Higham. *Functions of matrices. Theory and computation*. SIAM, Philadelphia, 2008. 56
- M. Hladík. Lineární algebra (nejen) pro informatiky, 2017. elektronická skripta, [http://kam.mff.cuni.cz/~hladik/LA/text\\_la\\_upd.pdf](http://kam.mff.cuni.cz/~hladik/LA/text_la_upd.pdf). 7, 54
- R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, 1985. 3, 28, 39, 41, 47, 62, 74
- R. A. Horn and C. R. Johnson. *Topics in matrix analysis*. Cambridge University Press, 1991. 3, 54, 59, 62, 69, 70, 71, 73, 74
- P. D. Lax. *Linear Algebra and Its Applications*. Wiley, New York, 2nd edition, 2007. 43
- J. R. Magnus. On differentiating eigenvalues and eigenvectors. *Econometric Theory*, 1(2):179–191, 1985. [https://pure.uvt.nl/portal/files/649757/27745\\_6631.pdf](https://pure.uvt.nl/portal/files/649757/27745_6631.pdf). 44
- G. Mayer. *Interval Analysis and Automatic Result Verification*, volume 65 of *Studies in Mathematics*. De Gruyter, Berlin, 2017. 41, 72
- C. D. Meyer. *Matrix Analysis and Applied Linear Algebra*. SIAM, Philadelphia, 2000. <http://iuuk.mff.cuni.cz/~andrew/MatrixAnalysisAppliedLinearAlgebra.pdf>. 3, 41, 47, 54, 69, 71
- V. V. Prasolov. *Problems and theorems in linear algebra*. AMS, 1994. <http://www2.math.su.se/~mleites/books/prasolov-1994-problems.pdf>. 3, 33
- J. Rohn. *Lineární algebra a optimalizace*. Karolinum, Praha, 2004. <http://uivtx.cs.cas.cz/~rohn/other/lascript.pdf>. 3, 18

- 
- S. M. Rump. Verification methods: Rigorous results using floating-point arithmetic. *Acta Numer.*, 19: 287–449, 2010. <https://doi.org/10.1017/S096249291000005X>. 49
- G. Strang. *Linear Algebra and its Applications*. Thomson, USA, 3rd edition, 1988. [http://iuuk.mff.cuni.cz/~andrew/%5BStrang\\_G.%5D\\_Linear\\_algebra\\_and\\_its\\_applications%284%29.PDF](http://iuuk.mff.cuni.cz/~andrew/%5BStrang_G.%5D_Linear_algebra_and_its_applications%284%29.PDF). 3
- M. M. Wilde. *Quantum Information Theory*. Cambridge University Press, 2nd edition, 2017. preprint pojmenovaný *From Classical to Quantum Shannon Theory*: <https://arxiv.org/abs/1106.1445>. 31