Institute of Computer Science
Academy of Sciences of the Czech Republic

Přehled některých důležitých vět z teorie matic

Jiří Rohn

Technical report No. 895

5. srpna 2003

Pod Vodárenskou věží 2, 182 07 Prague 8 phone: +420 2 688 42 44, fax: +420 2 858 57 89,
e-mail:e-mail:ics@cs.cas.cz

Institute of Computer Science
Academy of Sciences of the Czech Republic

Přehled některých důležitých vět z teorie matic

Jiří Rohn[1]

Technical report No. 895

5. srpna 2003

Abstrakt:

Tento report obsahuje přehled vybraných důležitých vět z teorie matic. Je psán formou „definice–věta–důkaz" a předpokládá předběžné znalosti přibližně na úrovni 1. semestru lineární algebry. Cílem bylo provést důkazy „až do posledního ε" u řady vět, jejichž důkazy se v literatuře většinou obcházejí nebo jsou prezentovány zjednodušenou formou (jako např. vlastnosti pseudoinverzní matice, Grevillův algoritmus, reálná forma Schurovy triangularizační věty, Courant-Fischerova věta, odmocnina z matice atp.). Autor byl inspirován knihami [1]–[9], ale všechny důkazy jsou přepracované jednak kvůli jednotnosti výkladu, jednak v důsledku výše uvedené snahy o jeho maximální jasnost. Výběr vět je subjektivní a jako takový nutně neúplný; je možné, že bude později rozšířen. V každém případě autor doufá, že tento text vyplní jistou mezeru v česky psané matematické literatuře.

Keywords:
Fundamentální podprostory, RREF tvar matice, Choleského rozklad, Householderova transformace, QR rozklad, Hessenbergův tvar, SVD rozklad, Moore-Penroseova inverze, Grevillův algoritmus, ortogonální projekce, soustavy lineárních rovnic, metoda nejmenších čtverců, Farkasova věta, metoda sdružených gradientů, Schurova triangularizační věta, unitární diagonalizovatelnost, spektrální rozklad, Courant-Fischerova věta, odmocnina z matice.

[1]Sepsání tohoto textu bylo podpořeno Grantovou agenturou České republiky z grantu 201/01/0343

Obsah

1 Fundamentální podprostory 3

2 RREF tvar matice 4

3 Algoritmus pro výpočet RREF tvaru 8

4 Pozitivní (semi)definitnost 10

5 Ortogonální matice 13

6 QR rozklad 15

7 Redukce na Hessenbergův tvar 16

8 SVD rozklad I 17

9 Moore-Penroseova inverze 21

10 Ortogonální projekce 27

11 Soustavy lineárních rovnic 28

12 Metoda nejmenších čtverců 29

13 Farkasova věta 30

14 Metoda sdružených gradientů 32

15 Vlastní čísla a vektory 34

16 Schurova triangularizační věta 37

17 Unitární diagonalizovatelnost 41

18 Hermitovské matice 42

19 SVD rozklad II 43

20 Vlastní čísla symetrických matic 44

1 Fundamentální podprostory

Definice Pro matici $A \in \mathbb{R}^{m\times n}$ definujeme sloupcový prostor

$$\mathcal{R}(A) = \{Ax;\ x \in \mathbb{R}^n\},$$

jádro

$$\mathcal{N}(A) = \{x;\ Ax = 0\}$$

a hodnost

$$\operatorname{rank}(A) = \dim \mathcal{R}(A).$$

Věta 1 (ortogonální rozklad) *Pro každou matici $A \in \mathbb{R}^{m\times n}$ platí:*
1) $\mathcal{R}(A)^\perp = \mathcal{N}(A^T)$,
2) $\mathcal{N}(A)^\perp = \mathcal{R}(A^T)$,
3) $\mathbb{R}^m = \mathcal{R}(A) \oplus \mathcal{N}(A^T)$,
4) $\mathbb{R}^n = \mathcal{R}(A^T) \oplus \mathcal{N}(A)$.

Důkaz 1) Protože $\mathcal{R}(A) = [A_{\bullet 1}, \dots, A_{\bullet n}]$, je $x \in \mathcal{R}(A)^\perp$ právě když $x^T A_{\bullet j} = 0$ pro $j = 1, \dots, n$, tj. $x^T A = 0^T$, což platí právě když $A^T x = 0$, tj. $x \in \mathcal{N}(A^T)$.

2) Aplikací 1) na matici A^T dostáváme $\mathcal{R}(A^T)^\perp = \mathcal{N}(A)$ a tedy $\mathcal{N}(A)^\perp = \mathcal{R}(A^T)$.

3) Podle věty o direktním součtu je $\mathbb{R}^m = \mathcal{R}(A) \oplus \mathcal{R}(A)^\perp = \mathcal{R}(A) \oplus \mathcal{N}(A^T)$.

4) Podobně $\mathbb{R}^n = \mathcal{N}(A) \oplus \mathcal{N}(A)^\perp = \mathcal{N}(A) \oplus \mathcal{R}(A^T) = \mathcal{R}(A^T) \oplus \mathcal{N}(A)$. $\square$

Věta 2 *Pro každou matici $A \in \mathbb{R}^{m\times n}$ platí:*
1) $A^T A$ *je čtvercová symetrická,*
2) $\mathcal{N}(A^T A) = \mathcal{N}(A)$,
3) $\mathcal{R}(A^T A) = \mathcal{R}(A^T)$,
4) $A^T A$ *je regulární právě když sloupce* A *jsou lineárně nezávislé,*
5) $\operatorname{rank}(A^T A) = \operatorname{rank}(A)$,
6) $A^T A = 0$ *implikuje* $A = 0$.

Důkaz 1) Protože $A \in \mathbb{R}^{m\times n}$ a $A^T \in \mathbb{R}^{n\times m}$, je $A^T A \in \mathbb{R}^{n\times n}$ čtvercová; je symetrická protože $(A^T A)^T = A^T (A^T)^T = A^T A$.

2) Je-li $x \in \mathcal{N}(A)$, potom $Ax = 0$, takže $A^T A x = 0$ a $x \in \mathcal{N}(A^T A)$. Naopak, je-li $x \in \mathcal{N}(A^T A)$, potom $A^T A x = 0$, tedy $x^T A^T A x = \|Ax\|^2 = 0$, z čehož plyne $Ax = 0$ a $x \in \mathcal{N}(A)$.

3) Podle věty 1 a tvrzení 2) je $\mathcal{R}(A^T A) = \mathcal{N}((A^T A)^T)^\perp = \mathcal{N}(A^T A)^\perp = \mathcal{N}(A)^\perp = \mathcal{R}(A^T)$.

4) Je-li $A^T A$ regulární, potom $\mathcal{N}(A^T A) = \{0\}$ a podle 2) je $\mathcal{N}(A) = \{0\}$, takže sloupce A jsou lineárně nezávislé; naopak, jsou-li sloupce A lineárně nezávislé, potom $\mathcal{N}(A^T A) = \mathcal{N}(A) = \{0\}$ a $A^T A$ je regulární.

5) Z tvrzení 3) plyne $\operatorname{rank}(A^T A) = \dim \mathcal{R}(A^T A) = \dim \mathcal{R}(A^T) = \operatorname{rank}(A^T) = \operatorname{rank}(A)$.

6) Je-li $A^T A = 0$, potom pro každé $j = 1, \dots, n$ je $(A^T A)_{jj} = \sum_i A_{ij}^2 = 0$, takže celý j-tý sloupec matice A je nulový. $\square$

Věta 3 (hodnostní rozklad) *Pro každou matici* $A \in \mathbb{R}^{m \times n}$ *hodnosti* $r \geq 1$ *existují matice* $B \in \mathbb{R}^{m \times r}$, $C \in \mathbb{R}^{r \times n}$ *takové, že platí*

$$A = BC,$$

přičemž $\operatorname{rank}(B) = \operatorname{rank}(C) = r$.

Důkaz Nechť $b_1, \ldots, b_r$ je libovolná báze $\mathcal{R}(A)$, potom každý sloupec $A_{\bullet j}$, $j = 1, \ldots, n$, lze psát právě jedním způsobem ve tvaru $A_{\bullet j} = \sum_{i=1}^{r} c_{ij} b_i$. Nechť B je matice o sloupcích $b_1, \ldots, b_r$ a $C = (c_{ij})$. Potom $B \in \mathbb{R}^{m \times r}$, $C \in \mathbb{R}^{r \times n}$, $A = BC$, $\operatorname{rank}(B) = r$ a $\operatorname{rank}(C) \leq r$. Přitom je

$$r = \operatorname{rank}(A) \leq \min\{\operatorname{rank}(B), \operatorname{rank}(C)\} = \operatorname{rank}(C) \leq r,$$

tedy $\operatorname{rank}(C) = r$ a věta je dokázána. □

2 RREF tvar matice

Definice Matice $A \in \mathbb{R}^{n \times n}$ se nazývá regulární jestliže její sloupce jsou lineárně nezávislé (jinými slovy, jestliže soustava $Ax = 0$ má jediné řešení $x = 0$); v opačném případě se nazývá singulární.

Věta 4 *Jsou-li* $A, B \in \mathbb{R}^{n \times n}$ *regulární, potom i* AB *je regulární.*

Důkaz Je-li $ABx = 0$, potom z regularity A plyne $Bx = 0$ a z regularity B plyne $x = 0$. □

Definice Matice $A \in \mathbb{R}^{m \times n}$ je v RREF[2] (odstupňovaném) tvaru jestliže existují $0 \leq r \leq m$ a $1 \leq k_1 < k_2 < \ldots < k_r \leq n$ tak, že platí:

1. $A_{i1} = \ldots = A_{i,k_i - 1} = 0$ a $A_{\bullet k_i} = e_i$ pro $i = 1, \ldots, r$,
2. $A_{i\bullet} = 0^T$ pro $i = r + 1, \ldots, m$.

Věta 5 (RREF rozklad) *Pro každou matici* $A \in \mathbb{R}^{m \times n}$ *existuje regulární matice* $P \in \mathbb{R}^{m \times m}$ *a matice* $A^R \in \mathbb{R}^{m \times n}$ *v RREF tvaru tak, že platí*

$$A = PA^R. \tag{1}$$

Matice A^R *je tímto rozkladem určena jednoznačně, stejně tak jako prvních* r *sloupců matice* P*, kde* r *je počet nenulových řádků* A^R.

Důkaz (a) *Existence.* Je-li $A = 0$, je A v RREF tvaru s $r = 0$, takže stačí volit $A^R = 0$ a $P = I$. Nechť tedy $A \neq 0$. Pro $j = 1, \ldots, n$ položme

$$S_j = [A_{\bullet 1}, \ldots, A_{\bullet j}], \tag{2}$$

[2] z angl. "reduced row-echelon form"

potom $S_n = \mathcal{R}(A)$, takže $\dim S_n = \dim \mathcal{R}(A) = \operatorname{rank}(A) = r \geq 1$. Pro $i = 1, \ldots, r$ definujme

$$k_i = \min\{j;\, \dim S_j = i\}. \tag{3}$$

Potom $1 \leq k_1 < \ldots < k_r \leq n$. Ukážeme, že sloupce $A_{\bullet k_1}, \ldots, A_{\bullet k_r}$ jsou lineárně nezávislé. Předpokládejme sporem, že platí $\sum_{i=1}^{r} \beta_i A_{\bullet k_i} = 0$, kde aspoň jedno β_i je nenulové. Položme $p = \max\{i;\, \beta_i \neq 0\}$, potom $A_{\bullet k_p} \in [A_{\bullet k_1}, \ldots, A_{\bullet k_{p-1}}] \subseteq S_{k_{p-1}}$. Je-li $k_{p-1} < j < k_p$, plyne z (3) že $A_{\bullet j} \in S_{k_{p-1}}$, celkem tedy $S_{k_p} \subseteq S_{k_{p-1}}$, takže $p = \dim S_{k_p} \leq \dim S_{k_{p-1}} = p-1$, což je spor. Tedy $A_{\bullet k_1}, \ldots, A_{\bullet k_r}$ jsou lineárně nezávislé a jelikož $\dim S_{k_i} = i$, tvoří $A_{\bullet k_1}, \ldots, A_{\bullet k_i}$ bázi S_{k_i} pro $i = 1, \ldots, r$ a navíc $A_{\bullet k_1}, \ldots, A_{\bullet k_r}$ tvoří bázi $\mathcal{R}(A)$, neboť $A_{\bullet j} \in S_{k_r}$ pro $j > k_r$ podle (3). Proto pro každé $j = 1, \ldots, n$ lze $A_{\bullet j}$ jednoznačně vyjádřit ve tvaru

$$A_{\bullet j} = \sum_{i=1}^{r} \alpha_{ij} A_{\bullet k_i}. \tag{4}$$

Položme $(A^R)_{ij} = \alpha_{ij}$ pro $i = 1, \ldots, r$ a $(A^R)_{ij} = 0$ pro $i = r+1, \ldots, m$. Potom A^R je v RREF tvaru. Pro každé $i = 1, \ldots, r$ a $1 \leq j < k_i$ je totiž $A_{\bullet j} \in S_{k_{i-1}}$ podle (3), takže $A_{\bullet j}$ je lineární kombinací $A_{\bullet k_1}, \ldots, A_{\bullet k_{i-1}}$ a tedy $(A^R)_{ij} = 0$. Dále $A_{\bullet k_i}$ lze jakožto prvek báze vyjádřit jednoznačně ve tvaru $A_{\bullet k_i} = \sum_{\ell \neq i} 0 \cdot A_{\bullet k_\ell} + 1 \cdot A_{\bullet k_i}$, z čehož plyne že $(A^R)_{\bullet k_i} = e_i$. Nakonec, $(A^R)_{ij} = 0$ pro $i = r+1, \ldots, m$ a $j = 1, \ldots, n$. Tedy A^R je v RREF tvaru. Položíme-li nyní $P_{\bullet i} = A_{\bullet k_i}$ pro $i = 1, \ldots, r$, jsou sloupce $P_{\bullet 1}, \ldots, P_{\bullet r}$ lineárně nezávislé a lze je doplnit na regulární matici P. Potom z (4) plyne

$$A_{\bullet j} = \sum_{i=1}^{r} (A^R)_{ij} P_{\bullet i} = \sum_{i=1}^{m} (A^R)_{ij} P_{\bullet i} = (PA^R)_{\bullet j}$$

pro každé $j = 1, \ldots, n$, tedy $A = PA^R$.

(b) *Jednoznačnost.* Nechť $A = PA^R$, kde P je regulární a A^R je v RREF tvaru s parametry $r, k_1, \ldots, k_r$. Je-li $A = 0$, potom za předpokladu $r \geq 1$ by bylo $A_{\bullet k_1} = P(A^R)_{\bullet k_1} = Pe_1 = P_{\bullet 1} \neq 0$, spor; tedy $r = 0$, z čehož plyne $A^R = 0$ a A^R je určena jednoznačně. Nechť tedy $A \neq 0$, takže $r \geq 1$. Potom pro každé $i = 1, \ldots, r$ je $A_{\bullet k_i} = (PA^R)_{\bullet k_i} = P(A^R)_{\bullet k_i} = Pe_i = P_{\bullet i}$, takže $A_{\bullet k_1}, \ldots, A_{\bullet k_r}$, jakožto sloupce regulární matice P, jsou lineárně nezávislé a z RREF tvaru matice A^R ($(A^R)_{i\bullet} = 0^T$ pro $i = r+1, \ldots, m$) plyne, že pro každé $j = 1, \ldots, n$ je podle (1)

$$A_{\bullet j} = \sum_{i=1}^{m} (A^R)_{ij} P_{\bullet i} = \sum_{i=1}^{r} (A^R)_{ij} P_{\bullet i} = \sum_{i=1}^{r} (A^R)_{ij} A_{\bullet k_i}, \tag{5}$$

tedy $A_{\bullet k_1}, \ldots, A_{\bullet k_r}$ generují $\mathcal{R}(A)$ a protože jsou lineárně nezávislé, tvoří bázi tohoto prostoru. Z toho plyne, že $r = \dim \mathcal{R}(A) = \operatorname{rank}(A)$, takže r je určeno maticí A jednoznačně. Dokážeme dále indukcí, že pro $i = 1, \ldots, r$ platí

$$k_i = \min\{j;\, \dim S_j = i\}, \tag{6}$$

kde S_j jsou podprostory definované v (2). To je zřejmé pro $i = 1$, neboť $A_{\bullet k_1}$ je první nenulový sloupec matice A (pro $1 \leq j < k_1$ je totiž $A_{\bullet j} = P(A^R)_{\bullet j} = P \cdot 0 = 0$ a

$A_{\bullet k_1} = Pe_1 = P_{\bullet 1} \neq 0$). Nechť tedy tvrzení (6) platí pro jisté $1 \leq i-1 < r$, takže

$$\dim S_{k_{i-1}} = i - 1. \tag{7}$$

Pro $k_{i-1} < j < k_i$ jsou nenulové prvky v j-tém sloupci matice A^R pouze v řádcích 1 až $i-1$, takže podle (5) je $A_{\bullet j} \in S_{k_{i-1}}$ a tedy $\dim S_j = i-1$ pro každé takové j. Protože sloupce $A_{\bullet k_1}, \ldots, A_{\bullet k_i}$ jsou lineárně nezávislé, $A_{\bullet k_1}, \ldots, A_{\bullet k_{i-1}} \in S_{k_{i-1}}$ a $\dim S_{k_{i-1}} = i-1$, tvoří $A_{\bullet k_1}, \ldots, A_{\bullet k_{i-1}}$ bázi $S_{k_{i-1}}$ a tedy $A_{\bullet k_i} \notin S_{k_{i-1}}$, takže $\dim S_{k_i} = 1 + \dim S_{k_{i-1}} = 1 + (i-1) = i$ podle indukčního předpokladu (7), čímž je tvrzení (6) indukcí dokázáno. Z rovnosti (6) vzhledem k (2) nyní plyne, že indexy $k_1, \ldots, k_r$, a tedy i sloupce $A_{\bullet k_1} = P_{\bullet 1}, \ldots, A_{\bullet k_r} = P_{\bullet r}$, jsou jednoznačně určeny maticí A. Ze vztahu (5) potom dostáváme, že pro každé $j = 1, \ldots, n$ jsou koeficienty $(A^R)_{ij}$, $i = 1, \ldots, r$, určeny jednoznačně jakožto souřadnice vektoru $A_{\bullet j}$ v bázi $A_{\bullet k_1}, \ldots, A_{\bullet k_r}$, a z definice RREF tvaru plyne, že $(A^R)_{ij} = 0$ pro $i = r+1, \ldots, m$. To znamená, že matice A^R je určena jednoznačně, stejně tak jako prvních r sloupců matice P. $\square$

Definice Matici A^R z věty 5, která je maticí A jednoznačně určena, nazýváme RREF tvarem matice A.

Věta 6 *Nechť A^R je RREF tvar matice $A \in \mathbb{R}^{m \times n}$. Potom platí:*

1) $r = \operatorname{rank}(A)$,
2) $k_i = \min\{j; \dim[A_{\bullet 1}, \ldots, A_{\bullet j}] = i\}$ pro $i = 1, \ldots, r$,
3) $A_{\bullet k_1}, \ldots, A_{\bullet k_i}$ je báze $[A_{\bullet 1}, A_{\bullet 2}, \ldots, A_{\bullet k_i}]$ pro $i = 1, \ldots, r$,
4) $A_{\bullet k_1}, \ldots, A_{\bullet k_r}$ je báze $\mathcal{R}(A)$,
5) $A_{\bullet j} = \sum_{i=1}^{r} (A^R)_{ij} A_{\bullet k_i}$ pro $j = 1, \ldots, n$.

Důkaz Všechny tyto vlastnosti byly dokázány v důkazu věty 5. $\square$

Věta 7 *Pro matici $(A\ B)$ rozdělenou na bloky platí*

$$(A\ B)^R = (A^R\ X)$$

pro jistou matici X.

Důkaz Nechť $(A\ B)^R = (C\ X)$, kde C, X jsou stejných typů jako A, B. Potom podle věty 5 je $(A\ B) = P(C\ X) = (PC\ PX)$ pro jistou regulární matici P, z čehož plyne $A = PC$, kde P je regulární a C je v RREF tvaru, takže podle věty 5 je $C = A^R$. $\square$

Věta 8 *Matice $A \in \mathbb{R}^{n \times n}$ je regulární právě když $A^R = I$.*

Důkaz Je-li $A^R = I$, potom $A = PA^R = P$, takže A je regulární. Naopak, nechť $A^R \neq I$, takže $k_i \neq i$ pro jisté i. Nechť $p = \min\{i; k_i \neq i\}$. Potom prvních $p-1$ sloupců A^R tvoří sloupce jednotkové matice a $k_p > p$, takže $(A^R)_{pp} = 0$. Definujme vektor $x \in \mathbb{R}^n$ předpisem $x_i = -(A^R)_{ip}$ pro $i = 1, \ldots, p-1$, $x_p = 1$ a $x_j = 0$ pro $j = p+1, \ldots, n$. Potom $A^R x = 0$, takže $Ax = PA^R x = 0$, přičemž $x \neq 0$ (neboť $x_p = 1$), tedy A je singulární. $\square$

Věta 9 *Ke každé regulární matici $A \in \mathbb{R}^{n\times n}$ existuje právě jedna matice $X \in \mathbb{R}^{n\times n}$ s vlastností*
$$AX = XA = I. \tag{8}$$
Naopak, existuje-li k $A \in \mathbb{R}^{n\times n}$ matice X s vlastností (8), potom A je regulární.

Důkaz Protože A je regulární, platí podle vět 7, 8
$$(A\ I)^R = (A^R\ X) = (I\ X)$$
pro jistou matici $X \in \mathbb{R}^{n\times n}$. Přitom je
$$(A\ I) = P(A\ I)^R = P(I\ X) = (P\ PX),$$
z čehož plyne $P = A$ a $AX = I$. Potom $X^T A^T = I$ a z toho plyne že i A^T je regulární, neboť $A^T x = 0$ implikuje $x = Ix = X^T A^T x = 0$. Na matici A^T můžeme tedy aplikovat předchozí výsledek, podle kterého existuje matice $\hat{X}$ s vlastností $A^T\hat{X} = I$. Potom je $\hat{X}^T A = I$ a z toho vzhledem k asociativnosti součinu matic dostáváme
$$\hat{X}^T = \hat{X}^T I = \hat{X}^T(AX) = (\hat{X}^T A)X = IX = X,$$
takže $\hat{X}^T = X$ a tedy $XA = I$. Tím je existence matice X s vlastností (8) dokázána. Pro důkaz jednoznačnosti předpokládejme, že jistá matice $\tilde{X} \in \mathbb{R}^{n\times n}$ splňuje $A\tilde{X} = \tilde{X}A = I$. Potom z rovnosti $A\tilde{X} = I$ plyne $(A\ I) = A(I\ \tilde{X})$, kde A je regulární a $(I\ \tilde{X})$ je v RREF tvaru, takže podle věty 5 je $(I\ \tilde{X}) = (A\ I)^R = (I\ X)$ a tedy $\tilde{X} = X$, což dokazuje jednoznačnost. Nakonec, existuje-li k dané $A \in \mathbb{R}^{n\times n}$ matice X s vlastností (8), potom z $Ax = 0$ plyne $x = XAx = 0$, takže A je regulární. □

Definice Matici X splňující (8) nazýváme inverzní maticí k matici A a značíme ji A^{-1}.

Věta 10 *Jestliže pro $A, X \in \mathbb{R}^{n\times n}$ platí $AX = I$, potom $X = A^{-1}$.*

Důkaz Toto tvrzení bylo dokázáno v důkazu věty 9. □

Věta 11 *Pro danou matici $A \in \mathbb{R}^{n\times n}$ nechť $(A\ I)^R = (C\ X)$. Potom platí:*

1) je-li $C = I$, je $X = A^{-1}$,
2) je-li $C \neq I$, potom A je singulární a nemá inverzní matici.

Důkaz 1) Je-li $C = I$, potom stejným postupem jako na začátku důkazu věty 9 dostáváme $AX = XA = I$ a tedy $X = A^{-1}$. 2) Je-li $C \neq I$, potom $A^R = C \neq I$, takže A je podle věty 8 singulární a podle věty 9 nemá inverzní matici. □

Věta 12 *Nechť $A, B \in \mathbb{R}^{n\times n}$ jsou regulární matice. Potom platí:*

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

3) $(\alpha A)^{-1} = \frac{1}{\alpha}A^{-1}$ *pro* $\alpha \neq 0$*,*
4) $(AB)^{-1} = B^{-1}A^{-1}$*.*

Důkaz V důkazu všech tvrzení použijeme větu 10, podle které z $AX = I$ plyne $X = A^{-1}$. 1) Podle definice inverzní matice je $A^{-1}A = I$, a z toho podle věty 10 plyne $A = (A^{-1})^{-1}$. 2) Z $A^{-1}A = I$ transpozicí dostáváme $A^T(A^{-1})^T = I$ a z toho $(A^{-1})^T = (A^T)^{-1}$. 3) Platí $\alpha A \cdot \frac{1}{\alpha}A^{-1} = AA^{-1} = I$, a tedy $\frac{1}{\alpha}A^{-1} = (\alpha A)^{-1}$. 4) Protože $ABB^{-1}A^{-1} = AA^{-1} = I$, dává věta 10, že $B^{-1}A^{-1} = (AB)^{-1}$. □

Věta 13 (Sherman-Morrison) *Nechť* $A \in \mathbb{R}^{n\times n}$ *je regulární a nechť* $b, c \in \mathbb{R}^n$*. Potom platí:*

1) je-li $c^TA^{-1}b \neq -1$*, je*

$$(A + bc^T)^{-1} = A^{-1} - \frac{1}{1+c^TA^{-1}b}A^{-1}bc^TA^{-1}, \tag{9}$$

2) je-li $c^TA^{-1}b = -1$*, je* $A + bc^T$ *singulární.*

Důkaz 1) Vynásobením dostáváme

$$\begin{aligned} & (A + bc^T)(A^{-1} - \tfrac{1}{1+c^TA^{-1}b}A^{-1}bc^TA^{-1}) \\ = \; & I - \tfrac{1}{1+c^TA^{-1}b}bc^TA^{-1} + bc^TA^{-1} - \tfrac{1}{1+c^TA^{-1}b}b(c^TA^{-1}b)c^TA^{-1} \\ = \; & I + (-\tfrac{1}{1+c^TA^{-1}b} + 1 - \tfrac{c^TA^{-1}b}{1+c^TA^{-1}b})bc^TA^{-1} = I, \end{aligned}$$

neboť výraz v poslední závorce je roven nule. Z toho podle věty 10 plyne (9).

2) Je-li $c^TA^{-1}b = -1$, potom $(A + bc^T)A^{-1}b = b + b(c^TA^{-1}b) = b - b = 0$, přičemž $A^{-1}b \neq 0$ vzhledem k tomu že $c^TA^{-1}b = -1$, čili $A + bc^T$ je singulární. □

3 Algoritmus pro výpočet RREF tvaru

Definice Následující tři operace nazýváme elementárními operacemi s maticí A:

1. vynásobení i-tého řádku číslem $\alpha \neq 0$ (tj. $A_{i\bullet} := \alpha A_{i\bullet}$),
2. vynásobení i-tého řádku číslem α a přičtení k j-tému řádku, $j \neq i$ (tj. $A_{j\bullet} := A_{j\bullet} + \alpha A_{i\bullet}$),
3. výměna i-tého a j-tého řádku, $i \neq j$ (značíme $A_{i\bullet} \leftrightarrow A_{j\bullet}$).

Věta 14 *Pro matici* $\tilde{A}$ *vzniklou z matice* $A \in \mathbb{R}^{m\times n}$ *provedením*

1) elementární operace $A_{i\bullet} := \alpha A_{i\bullet}$ *platí* $\tilde{A} = (I + (\alpha - 1)e_ie_i^T)A$*,*
2) elementární operace $A_{j\bullet} := A_{j\bullet} + \alpha A_{i\bullet}$ *platí* $\tilde{A} = (I + \alpha e_je_i^T)A$*,*
3) elementární operace $A_{i\bullet} \leftrightarrow A_{j\bullet}$ *platí* $\tilde{A} = (I + (e_i - e_j)(e_j - e_i)^T)A$*.*

Důkaz 1) Matice $\tilde{A} = (I + (\alpha - 1)e_i e_i^T)A = A + (\alpha - 1)e_i A_{i\bullet}$ vzniká z A přičtením matice jejíž všechny řádky kromě i-tého jsou nulové a i-tý je roven $(\alpha - 1)A_{i\bullet}$, takže $\tilde{A}$ se ve všech řádcích kromě i-tého shoduje s maticí A a její i-tý řádek je roven $\alpha A_{i\bullet}$, tedy $\tilde{A}$ je výsledkem elementární operace 1.

2) Podobně z $\tilde{A} = (I + \alpha e_j e_i^T)A = A + \alpha e_j A_{i\bullet}$ plyne, že všechny řádky matice $\tilde{A}$ kromě j-tého jsou shodné s příslušnými řádky matice A a j-tý řádek se nahradí řádkem $A_{j\bullet} + \alpha A_{i\bullet}$, takže je to výsledek elementární operace 2.

3) Nakonec z $\tilde{A} = (I + (e_i - e_j)(e_j - e_i)^T)A = A + e_i(A_{j\bullet} - A_{i\bullet}) + e_j(A_{i\bullet} - A_{j\bullet})$ plyne, že v $\tilde{A}$ je i-tý řádek matice A nahrazen j-tým řádkem a j-tý řádek je nahrazen i-tým řádkem, přičemž ostatní řádky se nemění, je to tedy výsledek elementární operace 3. $\square$

Věta 15 *Matice $\tilde{A}$ vzniklá z matice $A \in \mathbb{R}^{m\times n}$ provedením konečné posloupnosti elementárních operací je tvaru*
$$\tilde{A} = QA,$$
kde Q je jistá čtvercová regulární matice.

Důkaz Podle věty 14 je provedení kterékoliv ze tří elementárních operací ekvivalentní vynásobení matice A zleva jistou maticí tvaru
$$I + bc^T,$$
a tato matice je podle věty 13 regulární právě když $c^Tb \neq -1$. V případě elementární operace 1 máme $c^Tb = (\alpha - 1)e_i^T e_i = \alpha - 1 \neq -1$ (neboť $\alpha \neq 0$ podle definice), v případě elementární operace 2 je $c^Tb = \alpha e_i^T e_j = 0 \neq -1$ (neboť $i \neq j$), a v případě elementární operace 3 je $c^Tb = (e_j - e_i)^T(e_i - e_j) = -2 \neq -1$, takže všechny tyto matice jsou regulární. Výsledná matice $\tilde{A}$ vzniklá z A provedením posloupnosti elementárních operací je tedy tvaru
$$\tilde{A} = (I + b_p c_p^T) \cdot \ldots \cdot (I + b_1 c_1^T)A,$$
kde všechny matice $I + b_j c_j^T$, $j = 1, \ldots, p$, jsou regulární, tedy i
$$Q = (I + b_p c_p^T) \cdot \ldots \cdot (I + b_1 c_1^T)$$
je regulární (věta 4) a platí $\tilde{A} = QA$. $\square$

Algoritmus pro nalezení RREF tvaru matice A.

0. Polož $i := 1$.
1. Je-li $a_{\ell j} = 0$ pro každé $\ell \geq i$ a $1 \leq j \leq n$, polož $A^R := A$ a ukonči.
2. Jinak urči $k := \min\{j;\ a_{\ell j} \neq 0$ pro jisté $\ell \geq i\}$.
3. Nalezni $a_{\ell k} \neq 0$, $\ell \geq i$ a vyměň řádky $A_{i\bullet}$ a $A_{\ell\bullet}$.
4. Polož $A_{i\bullet} := \frac{1}{a_{ik}} A_{i\bullet}$.
5. Pro každé $\ell \neq i$ polož $\alpha := a_{\ell k}$ a $A_{\ell\bullet} := A_{\ell\bullet} - \alpha A_{i\bullet}$.
6. Je-li $i < m$, polož $i := i + 1$ a jdi na krok 1. Jinak polož $A^R := A$ a ukonči.

Věta 16 *Pro libovolnou výchozí matici $A \in \mathbb{R}^{m\times n}$ dává algoritmus její RREF tvar A^R.*

Důkaz Dokážeme nejprve, že matice $\tilde{A}$ vypočtená algoritmem je v RREF tvaru. Je-li $A = 0$, je matice v RREF tvaru (s $r = 0$) a algoritmus končí v kroku 1. Nechť tedy $A \neq 0$. Označme r poslední hodnotu indexu i, se kterou se provádí eliminace v kroku 5 algoritmu, a nechť pro každé $1 \leq i \leq r$ je k_i rovno hodnotě k vypočtené v kroku 2. Dokážeme indukcí podle i, že pro každé $1 \leq i \leq r$ po provedení eliminace s pivotem a_{ik_i} v kroku 5 je podmatice sestávající z prvních k_i sloupců matice v RREF tvaru, přičemž její řádky počínaje $(i+1)$-ním jsou nulové. To je zřejmé pro $i = 1$, neboť k_1 je index prvního nenulového sloupce původní matice, takže po provedení eliminace je podmatice sestávající z prvních k_1 sloupců evidentně v RREF tvaru a všechny její řádky počínaje druhým jsou nulové.

Nechť tedy tvrzení platí pro $1 \leq i-1 < r$, takže po provedení eliminace s pivotem $a_{i-1,k_{i-1}}$ je podmatice sestávající z prvních k_{i-1} sloupců v RREF tvaru a všechny její řádky počínaje i-tým jsou nulové. To znamená, že pro index k_i vypočtený v kroku 2 platí $k_{i-1} < k_i$ a při eliminaci s pivotem a_{ik_i} se prvních k_{i-1} sloupců nemění a navíc se vytvoří i-tý řádek a k_i-tý sloupec v RREF tvaru . To ukazuje, že v podmatici sestávající z prvních k_i sloupců je prvních i řádků v RREF tvaru a její zbývající řádky jsou vzhledem k výběru k_i v kroku 2 a k eliminaci nulové, takže celá podmatice je v RREF tvaru. Tím je tvrzení indukcí dokázáno. Po provedení poslední eliminace s hodnotou $i = r$ je buď $i = m$, nebo všechny řádky počínaje $(i+1)$-ním jsou nulové, takže výsledná matice $\tilde{A}$ je v RREF tvaru.

Protože matice se v průběhu algoritmu upravuje pouze elementárními operacemi 1, 2, 3 (v krocích 4, 5, 3), platí pro výslednou matici $\tilde{A}$ podle věty 15 $\tilde{A} = QA$, kde Q je jistá regulární matice. To znamená, že pro $P = Q^{-1}$ dostáváme $A = P\tilde{A}$, kde P je regulární a $\tilde{A}$ je v RREF tvaru, takže z jednoznačnosti tohoto rozkladu dokázané ve větě 5 plyne $\tilde{A} = A^R$. □

4 Pozitivní (semi)definitnost

Definice Symetrická matice $A \in \mathbb{R}^{n\times n}$ se nazývá pozitivně semidefinitní jestliže $x^TAx \geq 0$ pro každé $x \in \mathbb{R}^n$, a pozitivně definitní jestliže $x^TAx > 0$ pro každé $0 \neq x \in \mathbb{R}^n$.

Věta 17 *Pro každou matici $A \in \mathbb{R}^{m\times n}$ je A^TA pozitivně semidefinitní. Jsou-li sloupce A lineárně nezávislé, je A^TA pozitivně definitní.*

Důkaz A^TA je symetrická podle věty 2. Pro každé $x \in \mathbb{R}^n$ je $x^TA^TAx = (Ax)^T(Ax) = \|Ax\|_2^2 \geq 0$, tedy A je pozitivně semidefinitní. Jsou-li sloupce A lineárně nezávislé, potom z $x^TA^TAx = 0$ plyne $\|Ax\|_2 = 0$ a tedy $Ax = 0$, což znamená, že $x = 0$, tj. A^TA je pozitivně definitní. □

Věta 18 *Nechť A je pozitivně semidefinitní a nechť*

$$\tilde{x}^T A \tilde{x} = 0$$

pro jisté $\tilde{x}$. Potom

$$A\tilde{x} = 0.$$

Důkaz Vzhledem k pozitivní semidefinitnosti matice A platí pro každé $\lambda \in \mathbb{R}$

$$\begin{aligned} 0 &\le (\lambda\tilde{x} + A\tilde{x})^T A(\lambda\tilde{x} + A\tilde{x}) = \lambda^2 \tilde{x}^T A \tilde{x} + 2\lambda \|A\tilde{x}\|_2^2 + \tilde{x}^T A^3 \tilde{x} \\ &= 2\lambda\|A\tilde{x}\|_2^2 + \tilde{x}^T A^3 \tilde{x}, \end{aligned}$$

což je možné jedině když $\|A\tilde{x}\|_2 = 0$, tj. $A\tilde{x} = 0$. $\square$

Věta 19 *Matice*

$$\begin{pmatrix} \alpha & a^T \\ a & \tilde{A} \end{pmatrix}$$

je pozitivně definitní právě když $\alpha > 0$ a $\tilde{A} - \frac{1}{\alpha} a a^T$ je pozitivně definitní.

Důkaz Je-li A pozitivně definitní, potom $\alpha = e_1^T A e_1 > 0$ a pro každé $0 \ne x \in \mathbb{R}^{n-1}$ platí

$$x^T(\tilde{A} - \tfrac{1}{\alpha} a a^T)x = x^T \tilde{A} x - \tfrac{1}{\alpha}(a^T x)^2 = \begin{pmatrix} -\frac{1}{\alpha} a^T x \\ x \end{pmatrix}^T \begin{pmatrix} \alpha & a^T \\ a & \tilde{A} \end{pmatrix} \begin{pmatrix} -\frac{1}{\alpha} a^T x \\ x \end{pmatrix} > 0,$$

takže $\tilde{A} - \frac{1}{\alpha} a a^T$ je pozitivně definitní. Naopak, je-li $\alpha > 0$ a $\tilde{A} - \frac{1}{\alpha} a a^T$ je pozitivně definitní, potom pro každé $x \in \mathbb{R}^n$, píšeme-li ho ve tvaru $x = (\xi, x'^T)^T$, kde $x' \in \mathbb{R}^{n-1}$, platí

$$\begin{aligned} x^T A x &= \begin{pmatrix} \xi \\ x' \end{pmatrix}^T \begin{pmatrix} \alpha & a^T \\ a & \tilde{A} \end{pmatrix} \begin{pmatrix} \xi \\ x' \end{pmatrix} = \alpha\xi^2 + 2\xi a^T x' + x'^T \tilde{A} x' \\ &= (\sqrt{\alpha}\xi + \tfrac{1}{\sqrt{\alpha}} a^T x')^2 + x'^T(\tilde{A} - \tfrac{1}{\alpha} a a^T)x' \ge 0, \end{aligned}$$

takže A je pozitivně semidefinitní a $x^T A x = 0$ implikuje $x' = 0$ a $\xi = 0$, tedy A je pozitivně definitní. $\square$

Věta 20 (Choleského rozklad) *Symetrická matice A je pozitivně definitní právě když existuje dolní trojúhelníková matice L s kladnými diagonálními prvky taková, že*

$$A = LL^T.$$

Tato matice je určena jednoznačně.

Důkaz Je-li $A = LL^T$, potom pro každé x je $x^TAx = x^TLL^Tx = (L^Tx)^T(L^Tx) = \|L^Tx\|_2^2 \geq 0$, přičemž $x^TAx = 0$ implikuje $L^Tx = 0$ a tedy vzhledem ke kladnosti diagonálních koeficientů $x = 0$, takže A je pozitivně definitní. Důkaz opačné implikace provedeme indukcí podle řádu matice n. Pro $n = 1$ je $a_{11} > 0$, takže $L = (\sqrt{a_{11}})$. Nechť tedy tvrzení platí až do řádu $n - 1 \geq 1$ a nechť

$$A = \begin{pmatrix} \alpha & a^T \\ a & \tilde{A} \end{pmatrix} \in \mathbb{R}^{n\times n}$$

je pozitivně definitní. Potom podle věty 19 je $\alpha > 0$ a matice $\tilde{A} - \frac{1}{\alpha}aa^T \in \mathbb{R}^{(n-1)\times(n-1)}$ je pozitivně definitní, proto podle indukčního předpokladu existuje dolní trojúhelníková matice $\tilde{L} \in \mathbb{R}^{(n-1)\times(n-1)}$ s kladnými diagonálními prvky taková, že $\tilde{A} - \frac{1}{\alpha}aa^T = \tilde{L}\tilde{L}^T$. Položme nyní

$$L = \begin{pmatrix} \sqrt{\alpha} & 0^T \\ \frac{1}{\sqrt{\alpha}}a & \tilde{L} \end{pmatrix},$$

potom L je dolní trojúhelníková s kladnými diagonálními prvky a platí

$$LL^T = \begin{pmatrix} \sqrt{\alpha} & 0^T \\ \frac{1}{\sqrt{\alpha}}a & \tilde{L} \end{pmatrix} \begin{pmatrix} \sqrt{\alpha} & \frac{1}{\sqrt{\alpha}}a^T \\ 0 & \tilde{L}^T \end{pmatrix} = \begin{pmatrix} \alpha & a^T \\ a & \tilde{A} \end{pmatrix} = A,$$

což je hledaný rozklad. Pro důkaz jednoznačnosti předpokládejme, že $A = L_1L_1^T$ pro jistou dolní trojúhelníkovou matici

$$L_1 = \begin{pmatrix} \lambda & 0^T \\ \ell & \hat{L} \end{pmatrix}$$

s kladnými diagonálními prvky. Potom z rovnosti

$$A = \begin{pmatrix} \alpha & a^T \\ a & \tilde{A} \end{pmatrix} = L_1L_1^T = \begin{pmatrix} \lambda^2 & \lambda\ell^T \\ \lambda\ell & \ell\ell^T + \hat{L}\hat{L}^T \end{pmatrix}$$

plyne $\lambda = \sqrt{\alpha}$, $\ell = \frac{1}{\sqrt{\alpha}}a$ a $\hat{L}\hat{L}^T = \tilde{A} - \ell\ell^T = \tilde{A} - \frac{1}{\alpha}aa^T = \tilde{L}\tilde{L}^T$, takže podle indukčního předpokladu je $\hat{L} = \tilde{L}$ a tedy $L_1 = L$. Tím je důkaz indukcí proveden. □

Algoritmus

0. Polož $L := 0$ a $k := 1$.
1. Je-li $a_{kk} - \sum_{j=1}^{k-1} \ell_{kj}^2 \leq 0$, ukonči: A není pozitivně definitní.
2. Jinak vypočti

$$\begin{aligned} \ell_{kk} &:= \sqrt{a_{kk} - \sum_{j=1}^{k-1} \ell_{kj}^2} \\ \ell_{ik} &:= \frac{1}{\ell_{kk}}(a_{ik} - \sum_{j=1}^{k-1} \ell_{ij}\ell_{kj}) \end{aligned}$$

   $(i = k+1, \ldots, n)$.
3. Polož $k := k + 1$. Je-li $k \leq n$, jdi na krok 1. Jinak ukonči: A je pozitivně definitní a platí $A = LL^T$.

5 Ortogonální matice

Definice Matice $Q \in \mathbb{R}^{n\times n}$ se nazývá ortogonální jestliže $Q^TQ = I$.

Věta 21 *Následující tvrzení pro matici $Q \in \mathbb{R}^{n\times n}$ jsou ekvivalentní:*
(i) Q je ortogonální,
(ii) Q je regulární a $Q^{-1} = Q^T$,
(iii) $QQ^T = I$,
(iv) Q^T je ortogonální,
(v) řádky Q tvoří ortonormální bázi $\mathbb{R}^n$,
(vi) sloupce Q tvoří ortonormální bázi $\mathbb{R}^n$.

Důkaz Dokážeme (i)⇒(ii)⇒(iii)⇒(iv)⇒(i), (i)⇔(vi), (iv)⇔(v).
(i)⇒(ii): Z $Q^TQ = I$ plyne $Q^{-1} = Q^T$ a tedy Q je regulární.
(ii)⇒(iii): Protože $Q^{-1} = Q^T$, je $QQ^T = QQ^{-1} = I$.
(iii)⇒(iv): Protože $(Q^T)^TQ^T = QQ^T = I$, je Q^T ortogonální.
(iv)⇒(i): Z $QQ^T = I$ plyne $Q^T = Q^{-1}$ a tedy $Q^TQ = Q^{-1}Q = I$.
(i)⇔(vi): Q je ortogonální právě když $(Q^TQ)_{ij} = Q_{\bullet i} \cdot Q_{\bullet j} = I_{ij}$ pro všechna i, j, což je ekvivalentní tomu, že sloupce Q tvoří ortonormální systém; protože je jich n, tvoří ortonormální bázi $\mathbb{R}^n$.
(iv)⇔(v) je přepisem ekvivalence (i)⇔(vi) pro matici Q^T. □

Věta 22 *Je-li $Q \in \mathbb{R}^{n\times n}$ ortogonální, potom:*
1) $\|Q_{i\bullet}\|_2 = \|Q_{\bullet i}\|_2 = 1$ pro každé i,
2) $|Q_{ij}| \leq 1$ a $|(Q^{-1})_{ij}| \leq 1$ pro každé i, j,
3) $\|Qx\|_2 = \|x\|_2$ pro každé $x \in \mathbb{R}^n$.

Důkaz Z $Q^TQ = I$ plyne, že pro každé $x \in \mathbb{R}^n$ je $\|Qx\|_2^2 = x^TQ^TQx = x^Tx = \|x\|_2^2$, což dokazuje tvrzení 3). Pro $i, j = 1, \ldots, n$ odsud plyne $|Q_{ij}| \leq \|Q_{\bullet i}\|_2 = \|Qe_i\|_2 = \|e_i\|_2 = 1$, a aplikací tohoto výsledku na $Q^T = Q^{-1}$ dostáváme zbývající tvrzení. □

Věta 23 *Jsou-li $Q_1, Q_2 \in \mathbb{R}^{n\times n}$ ortogonální, je i Q_1Q_2 ortogonální.*

Důkaz Platí $(Q_1Q_2)^T(Q_1Q_2) = Q_2^TQ_1^TQ_1Q_2 = Q_2^TQ_2 = I$. □

Věta 24 (Householderova transformace) *Pro každý vektor $x \in \mathbb{R}^n$ takový, že $\|x\|_2 = 1$, je matice*

$$H(x) = I - 2xx^T$$

symetrická a ortogonální.

Důkaz $H(x)$ je symetrická protože $H(x)^T = I - 2xx^T = H(x)$. Dále je

$$H(x)^TH(x) = (I - 2xx^T)^T(I - 2xx^T) = (I - 2xx^T)^2 = I - 4xx^T + 4x(x^Tx)x^T = I,$$

takže $H(x)$ je ortogonální. □

Věta 25 *Pro každé dva vektory* $y, z \in \mathbb{R}^n$ *takové že* $y \neq z$ *a* $\|y\|_2 = \|z\|_2$*, platí*

$$y = H(\tfrac{y-z}{\|y-z\|_2})z,$$

jinými slovy každé dva různé vektory o stejné normě lze převést jeden na druhý Householderovou transformací.

Důkaz Platí

$$\begin{aligned} H(\tfrac{y-z}{\|y-z\|_2})z &= (I - 2\tfrac{y-z}{\|y-z\|_2} \cdot \tfrac{(y-z)^T}{\|y-z\|_2})z = z - \tfrac{2(y^T z - \|z\|_2^2)}{\|y-z\|_2^2}(y-z) \\ &= z + \tfrac{\|y\|_2^2 + \|z\|_2^2 - 2y^T z}{\|y-z\|_2^2}(y-z) = z + \tfrac{\|y-z\|_2^2}{\|y-z\|_2^2}(y-z) = y. \end{aligned}$$

□

Věta 26 *Pro každou matici* $A \in \mathbb{R}^{m\times n}$ *existuje ortogonální matice* $X \in \mathbb{R}^{m\times m}$ *taková, že*

$$(XA)_{\bullet 1} = \|A_{\bullet 1}\|_2 e_1.$$

Důkaz Je-li $A_{\bullet 1} = \|A_{\bullet 1}\|_2 e_1$, potom tvrzení platí s $X = I$. Nechť tedy $A_{\bullet 1} \neq \|A_{\bullet 1}\|_2 e_1$. Potom položíme-li $z = A_{\bullet 1}$, $y = \|A_{\bullet 1}\|_2 e_1$, je $y \neq z$ a $\|y\|_2 = \|A_{\bullet 1}\|_2 = \|z\|_2$, takže podle věty 25 pro matici

$$X = H(\tfrac{y-z}{\|y-z\|_2})$$

platí $(XA)_{\bullet 1} = XA_{\bullet 1} = Xz = y = \|A_{\bullet 1}\|_2 e_1$. □

Věta 27 *Pro každé* $x \neq e_1$, $\|x\|_2 = 1$, *je*

$$H(\tfrac{x-e_1}{\|x-e_1\|_2})$$

ortogonální matice, jejíž prvním sloupcem je x.

Důkaz Protože $\|x\|_2 = 1 = \|e_1\|_2$, je podle věty 25

$$x = H(\tfrac{x-e_1}{\|x-e_1\|_2})e_1,$$

z čehož plyne

$$x = (H(\tfrac{x-e_1}{\|x-e_1\|_2}))_{\bullet 1},$$

což je tvrzení věty. □

6 QR rozklad

Věta 28 (QR rozklad) *Pro každou matici $A \in \mathbb{R}^{m\times n}$ existuje ortogonální matice $Q \in \mathbb{R}^m$ a horní trojúhelníková matice $R \in \mathbb{R}^{m\times n}$ s nezápornými diagonálními prvky tak, že platí*

$$A = QR. \tag{10}$$

Důkaz Důkaz provedeme indukcí podle n. Je-li $n = 1$, potom podle věty 26 existuje ortogonální matice X taková, že $XA = XA_{\bullet 1} = \varrho e_1$, $\varrho \geq 0$. Potom $A = X^T \varrho e_1$, takže stačí položit $Q = X^T$, $R = \varrho e_1$. Nechť tedy tvrzení platí až do $n - 1 \geq 1$ a nechť $A \in \mathbb{R}^{m\times n}$. Podle téže věty 26 existuje ortogonální matice X taková, že XA je tvaru

$$XA = \begin{pmatrix} \varrho & r^T \\ 0 & \tilde{A} \end{pmatrix},$$

kde $\tilde{A} \in \mathbb{R}^{(m-1)\times(n-1)}$ a $\varrho \geq 0$. Podle indukčního předpokladu existuje ortogonální matice $\tilde{Q} \in \mathbb{R}^{(m-1)\times(m-1)}$ a horní trojúhelníková matice $\tilde{R}$ s nezápornými diagonálními prvky tak, že $\tilde{A} = \tilde{Q}\tilde{R}$. Potom

$$\begin{pmatrix} 1 & 0^T \\ 0 & \tilde{Q}^T \end{pmatrix} XA = \begin{pmatrix} 1 & 0^T \\ 0 & \tilde{Q}^T \end{pmatrix} \begin{pmatrix} \varrho & r^T \\ 0 & \tilde{A} \end{pmatrix} = \begin{pmatrix} \varrho & r^T \\ 0 & \tilde{Q}^T\tilde{A} \end{pmatrix} = \begin{pmatrix} \varrho & r^T \\ 0 & \tilde{R} \end{pmatrix},$$

takže matice

$$Q = X^T \begin{pmatrix} 1 & 0^T \\ 0 & \tilde{Q} \end{pmatrix}$$

je ortogonální,

$$R = \begin{pmatrix} \varrho & r^T \\ 0 & \tilde{R} \end{pmatrix}$$

je horní trojúhelníková matice s nezápornými diagonálními prvky a platí $A = QR$, čímž je tvrzení indukcí dokázáno. □

Věta 29 *Jsou-li sloupce matice $A \in \mathbb{R}^{m\times n}$ lineárně nezávislé, potom v rozkladu (10) jsou matice R a prvních n sloupců matice Q určeny jednoznačně, přičemž všechny diagonální prvky R jsou kladné.*

Důkaz Protože sloupce A jsou lineárně nezávislé, je $m \geq n$. Vzhledem k tomu, že R je horní trojúhelníková, musí být její řádky počínaje $(n+1)$-ním nulové, takže označíme-li $\hat{R}$ čtvercovou matici sestávající z prvních n řádků R a $\hat{Q}$ matici sestávající z prvních n sloupců Q, platí opět

$$A = \hat{Q}\hat{R}. \tag{11}$$

Přitom $\hat{Q}$ obecně není čtvercová, ale stále splňuje $\hat{Q}^T\hat{Q} = I$, takže z (11) plyne

$$A^TA = \hat{R}^T\hat{Q}^T\hat{Q}\hat{R} = \hat{R}^T\hat{R}.$$

Jelikož A má lineárně nezávislé sloupce, je A^TA pozitivně definitní (věta 17) a tedy regulární, z čehož plyne že $\hat{R}$ je regulární, takže její diagonální prvky jsou kladné.

Položíme-li nyní $L = \hat{R}^T$, potom L je dolní trojúhelníková s kladnými diagonálními prvky a splňuje

$$A^T A = LL^T,$$

což je Choleského rozklad matice $A^T A$, takže podle věty 20 je matice L určena jednoznačně, tedy i $\hat{R}$ je určena jednoznačně stejně tak jako $\hat{Q} = \hat{R}^{-1} A$. $\square$

7 Redukce na Hessenbergův tvar

Definice Říkáme, že matice $A \in \mathbb{R}^{n\times n}$ je v horním Hessenbergově tvaru jestliže platí $A_{ij} = 0$ pro $j < i - 1$ $(i, j = 1, \ldots, n)$.

Věta 30 *Pro každou matici $A \in \mathbb{R}^{n\times n}$, $n \geq 2$, existuje ortogonální matice $X \in \mathbb{R}^{n\times n}$ taková, že $X_{\bullet 1} = e_1$, $X_{1\bullet} = e_1^T$ a*

$$(XAX^T)_{\bullet 1} = A_{11} e_1 + \|A_{\bullet 1} - A_{11} e_1\|_2 e_2.$$

Důkaz Nechť $\tilde{A}$ je matice vzniklá z A vyškrtnutím prvního řádku. Podle věty 26 existuje ortogonální matice $\tilde{X} \in \mathbb{R}^{(n-1)\times(n-1)}$ taková, že $(\tilde{X}\tilde{A})_{\bullet 1} = \|\tilde{A}_{\bullet 1}\|_2 \tilde{e}_1$, kde $\tilde{e}_1 = (I_{n-1})_{\bullet 1}$. Položme

$$X = \begin{pmatrix} 1 & 0^T \\ 0 & \tilde{X} \end{pmatrix},$$

potom $X_{\bullet 1} = e_1$, $X_{1\bullet} = e_1^T$ a platí

$$\begin{aligned}
(XAX^T)_{\bullet 1} &= \begin{pmatrix} 1 & 0^T \\ 0 & \tilde{X} \end{pmatrix} \begin{pmatrix} A_{1\bullet} \\ \tilde{A} \end{pmatrix} \begin{pmatrix} 1 & 0^T \\ 0 & \tilde{X}^T \end{pmatrix} e_1 = \begin{pmatrix} 1 & 0^T \\ 0 & \tilde{X} \end{pmatrix} \begin{pmatrix} A_{1\bullet} \\ \tilde{A} \end{pmatrix} e_1 \\
&= \begin{pmatrix} 1 & 0^T \\ 0 & \tilde{X} \end{pmatrix} \begin{pmatrix} A_{11} \\ \tilde{A}_{\bullet 1} \end{pmatrix} = \begin{pmatrix} A_{11} \\ \|\tilde{A}_{\bullet 1}\|_2 \tilde{e}_1 \end{pmatrix} = A_{11} e_1 + \|A_{\bullet 1} - A_{11} e_1\|_2 e_2.
\end{aligned}$$

$\square$

Věta 31 (redukce na Hessenbergův tvar) *Ke každé čtvercové matici A existuje ortogonální matice Q a matice H v horním Hessenbergově tvaru tak, že platí*

$$A = QHQ^T.$$

Důkaz Důkaz provedeme indukcí podle řádu n matice A, přičemž navíc dokážeme, že Q lze volit tak, aby $Q^T e_1 = e_1$. Pro $n = 1$ tvrzení platí pro $H = (A_{11})$, $Q = (1)$. Nechť tedy tvrzení platí až do řádu $n - 1 \geq 1$ včetně. Podle věty 30 existuje ortogonální matice X taková, že $Xe_1 = e_1$ a $(XAX^T)_{\bullet 1} = \alpha e_1 + \beta e_2$ pro jistá $\alpha, \beta \in \mathbb{R}$, takže XAX^T je tvaru

$$XAX^T = \begin{pmatrix} \alpha & a^T \\ \beta \tilde{e}_1 & \tilde{A} \end{pmatrix},$$

kde $\tilde{e}_1 = (I_{n-1})_{\bullet 1}$. Podle indukčního předpokladu existuje ortogonální matice $\tilde{Q} \in \mathbb{R}^{(n-1)\times(n-1)}$ taková, že $\tilde{Q}^T\tilde{A}\tilde{Q} = \tilde{H}$ je v horním Hessenbergově tvaru a $\tilde{Q}^T\tilde{e}_1 = \tilde{e}_1$. Potom

$$\begin{aligned}
\begin{pmatrix} 1 & 0^T \\ 0 & \tilde{Q}^T \end{pmatrix} XAX^T \begin{pmatrix} 1 & 0^T \\ 0 & \tilde{Q} \end{pmatrix} &= \begin{pmatrix} 1 & 0^T \\ 0 & \tilde{Q}^T \end{pmatrix} \begin{pmatrix} \alpha & a^T \\ \beta\tilde{e}_1 & \tilde{A} \end{pmatrix} \begin{pmatrix} 1 & 0^T \\ 0 & \tilde{Q} \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0^T \\ 0 & \tilde{Q}^T \end{pmatrix} \begin{pmatrix} \alpha & a^T\tilde{Q} \\ \beta\tilde{e}_1 & \tilde{A}\tilde{Q} \end{pmatrix} \\
&= \begin{pmatrix} \alpha & a^T\tilde{Q} \\ \beta\tilde{Q}^T\tilde{e}_1 & \tilde{Q}^T\tilde{A}\tilde{Q} \end{pmatrix} = \begin{pmatrix} \alpha & a^T\tilde{Q} \\ \beta\tilde{e}_1 & \tilde{H} \end{pmatrix},
\end{aligned}$$

takže výsledná matice

$$H = \begin{pmatrix} \alpha & a^T\tilde{Q} \\ \beta\tilde{e}_1 & \tilde{H} \end{pmatrix}$$

je v horním Hessenbergově tvaru a platí

$$A = QHQ^T,$$

kde

$$Q = X^T \begin{pmatrix} 1 & 0^T \\ 0 & \tilde{Q} \end{pmatrix}$$

jakožto součin ortogonálních matic je ortogonální a nakonec

$$Q^T e_1 = \begin{pmatrix} 1 & 0^T \\ 0 & \tilde{Q}^T \end{pmatrix} X e_1 = \begin{pmatrix} 1 & 0^T \\ 0 & \tilde{Q}^T \end{pmatrix} e_1 = e_1.$$

□

Definice Říkáme, že matice $A \in \mathbb{R}^{n\times n}$ je třídiagonální jestliže platí $A_{ij} = 0$ pro $|i-j| > 1$ $(i, j = 1, \dots, n)$.

Věta 32 *Ke každé symetrické matici A existuje ortogonální matice Q a symetrická třídiagonální matice H tak, že platí*

$$A = QHQ^T.$$

Důkaz Podle věty 31 existuje ortogonální matice Q tak, že $H = Q^TAQ$ je v horním Hessenbergově tvaru. Vzhledem k symetrii A pak platí $H^T = Q^TA^TQ = Q^TAQ = H$, takže H je symetrická a tedy nutně třídiagonální. □

8 SVD rozklad I

Věta 33 *Nechť $A \in \mathbb{R}^{m\times n}$. Potom existuje $\tilde{y} \in \mathbb{R}^n$, $\|\tilde{y}\|_2 = 1$ pro které*

$$\|A\tilde{y}\|_2 = \max\{\|Ay\|_2;\ \|y\|_2 = 1\}$$

a pro každé $\tilde{y}$ s touto vlastností platí

$$A^TA\tilde{y} = \|A\tilde{y}\|_2^2\tilde{y}.$$

Důkaz Protože funkce $\|Ay\|_2$ je spojitá na kompaktní jednotkové sféře $\{y \in \mathbb{R}^n;\ \|y\|_2 = 1\}$, existuje podle Weierstrassovy věty vektor $\tilde{y}$, $\|\tilde{y}\|_2 = 1$, pro který

$$\|A\tilde{y}\|_2 = \max\{\|Ay\|_2;\ \|y\|_2 = 1\}.$$

Pro každé $x \in \mathbb{R}^n$, $x \neq 0$ je potom

$$\frac{\|Ax\|_2}{\|x\|_2} = \left\| A \frac{x}{\|x\|_2} \right\|_2 \leq \max\{\|Ay\|_2;\ \|y\|_2 = 1\} = \|A\tilde{y}\|_2,$$

takže

$$x^T A^T A x \leq \|A\tilde{y}\|_2^2 x^T x$$

a

$$x^T (\|A\tilde{y}\|_2^2 I - A^T A) x \geq 0$$

pro každé x, tedy matice $\|A\tilde{y}\|_2^2 I - A^T A$ je pozitivně semidefinitní a platí

$$\tilde{y}^T (\|A\tilde{y}\|_2^2 I - A^T A)\tilde{y} = \|A\tilde{y}\|_2^2 - \|A\tilde{y}\|_2^2 = 0,$$

takže podle věty 18 je

$$A^T A \tilde{y} = \|A\tilde{y}\|_2^2 \tilde{y}.$$

□

Věta 34 (SVD rozklad) *Nechť $A \in \mathbb{R}^{m\times n}$ a $q = \min\{m, n\}$. Potom existuje matice $\Sigma = (\sigma_{ij}) \in \mathbb{R}^{m\times n}$ splňující $\sigma_{ij} = 0$ pro $i \neq j$ a $\sigma_{11} \geq \sigma_{22} \geq \ldots \geq \sigma_{qq} \geq 0$ a ortogonální matice $X \in \mathbb{R}^{m\times m}$, $Y \in \mathbb{R}^{n\times n}$ takové, že platí*

$$A = X\Sigma Y^T.$$

Důkaz Důkaz provedeme indukcí podle n, přičemž navíc dokážeme, že Σ lze volit tak, aby

$$\sigma_{11} = \max\{\|Ay\|_2;\ \|y\|_2 = 1,\ y \in \mathbb{R}^n\}.$$

Pro $n = 1$ sestává A z jediného sloupce a. Je-li $a = 0$, potom tvrzení platí pro $\Sigma = 0$, $X = I_m$, $Y = (1)$ (matice 1×1); je-li $a \neq 0$, položme $\Sigma = \|a\|_2 e_1$, $\tilde{x} = \frac{a}{\|a\|_2}$, potom $\|\tilde{x}\|_2 = 1$, takže $\tilde{x}$ lze doplnit na ortogonální matici $X = (\tilde{x}\ \tilde{X})$ a zvolíme-li $Y = (1)$, platí $A = X\Sigma Y^T$ a $\sigma_{11} = \|a\|_2 = \max\{\|ay\|_2;\ \|y\|_2 = 1,\ y \in \mathbb{R}^1\}$.

Nechť tedy tvrzení platí až do $n - 1 \geq 1$ včetně, a nechť $A \in \mathbb{R}^{m\times n}$. Je-li $A = 0$, potom tvrzení platí pro $\Sigma = 0$, $X = I_m$, $Y = I_n$. Nechť tedy $A \neq 0$; položme

$$\sigma = \|A\tilde{y}\|_2 = \max\{\|Ay\|_2;\ \|y\|_2 = 1,\ y \in \mathbb{R}^n\},$$

kde $\|\tilde{y}\|_2 = 1$, potom $\sigma > 0$ a podle věty 33 je

$$A^T A \tilde{y} = \sigma^2 \tilde{y}.$$

Položme $\tilde{x} = \frac{1}{\sigma}A\tilde{y}$. Potom $\|\tilde{x}\|_2 = \|\tilde{y}\|_2 = 1$, takže $\tilde{x}$ lze doplnit na ortogonální matici $X_1 = (\tilde{x}\ \tilde{X})$ a podobně $\tilde{y}$ lze doplnit na ortogonální matici $Y_1 = (\tilde{y}\ \tilde{Y})$. Potom platí

$$X_1^T A Y_1 = \begin{pmatrix} \tilde{x}^T \\ \tilde{X}^T \end{pmatrix} A(\tilde{y}\ \tilde{Y}) = \begin{pmatrix} \tilde{x}^T A\tilde{y} & \tilde{x}^T A\tilde{Y} \\ \tilde{X}^T A\tilde{y} & \tilde{X}^T A\tilde{Y} \end{pmatrix},$$

přičemž

$$\tilde{x}^T A\tilde{y} = \tfrac{1}{\sigma}\|A\tilde{y}\|_2^2 = \tfrac{1}{\sigma}\sigma^2 = \sigma$$

a z ortogonality X_1, Y_1 plyne

$$\tilde{x}^T A\tilde{Y} = (A^T\tilde{x})^T\tilde{Y} = \tfrac{1}{\sigma}(A^T A\tilde{y})^T\tilde{Y} = \sigma\tilde{y}^T\tilde{Y} = 0^T,$$

$$\tilde{X}^T A\tilde{y} = \sigma\tilde{X}^T\tilde{x} = 0,$$

takže

$$X_1^T A Y_1 = \begin{pmatrix} \sigma & 0^T \\ 0 & \tilde{X}^T A\tilde{Y} \end{pmatrix}.$$

Protože matice $\tilde{X}^T A\tilde{Y}$ je typu $(m-1)\times(n-1)$, platí podle indukčního předpokladu

$$\tilde{X}^T A\tilde{Y} = \hat{X}\Sigma'\hat{Y}^T,$$

kde $\hat{X} \in \mathbb{R}^{(m-1)\times(m-1)}$ a $\hat{Y} \in \mathbb{R}^{(n-1)\times(n-1)}$ jsou ortogonální a $\Sigma' = (\sigma'_{ij}) \in \mathbb{R}^{(m-1)\times(n-1)}$ splňuje $\sigma'_{ij} = 0$ pro $i \neq j$, $\sigma'_{11} \geq \ldots \geq \sigma'_{q-1,q-1} \geq 0$, přičemž

$$\sigma'_{11} = \max\{\|\tilde{X}^T A\tilde{Y}z\|_2;\ \|z\|_2 = 1,\ z \in \mathbb{R}^{n-1}\}.$$

Položíme-li nyní

$$X_2 = \begin{pmatrix} 1 & 0^T \\ 0 & \hat{X} \end{pmatrix}, \quad Y_2 = \begin{pmatrix} 1 & 0^T \\ 0 & \hat{Y} \end{pmatrix},$$

potom X_2 i Y_2 jsou ortogonální a platí

$$\begin{aligned} X_2^T X_1^T A Y_1 Y_2 &= \begin{pmatrix} 1 & 0^T \\ 0 & \hat{X}^T \end{pmatrix} \begin{pmatrix} \sigma & 0^T \\ 0 & \tilde{X}^T A\tilde{Y} \end{pmatrix} \begin{pmatrix} 1 & 0^T \\ 0 & \hat{Y} \end{pmatrix} \\ &= \begin{pmatrix} \sigma & 0^T \\ 0 & \hat{X}^T\tilde{X}^T A\tilde{Y}\hat{Y} \end{pmatrix} = \begin{pmatrix} \sigma & 0^T \\ 0 & \Sigma' \end{pmatrix} = \Sigma. \end{aligned}$$

Odtud dostáváme

$$A = X\Sigma Y^T,$$

kde $X = X_1X_2$ a $Y = Y_1Y_2$ jsou ortogonální, Σ má nulové nediagonální prvky a platí $\sigma'_{11} \geq \ldots \geq \sigma'_{q-1,q-1}$. K dokončení důkazu je proto potřeba dokázat, že $\sigma \geq \sigma'_{11}$.

Protože $X_1 = (\tilde{x}\ \tilde{X})$ a $Y_1 = (\tilde{y}\ \tilde{Y})$ jsou ortogonální, je $\tilde{Y}^T\tilde{Y} = I_{n-1}$ a $X_1X_1^T = \tilde{x}\tilde{x}^T + \tilde{X}\tilde{X}^T = I_m$, takže $\tilde{X}\tilde{X}^T = I_m - \tilde{x}\tilde{x}^T$. K danému $z \in \mathbb{R}^{n-1}$, $\|z\|_2 = 1$, vezměme $y = \tilde{Y}z \in \mathbb{R}^n$. Potom $\|y\|_2^2 = z^T\tilde{Y}^T\tilde{Y}z = z^Tz = \|z\|_2^2 = 1$, takže $\|y\|_2 = 1$ a platí

$$\begin{aligned} \|\tilde{X}^T A\tilde{Y}z\|_2^2 &= \|\tilde{X}^T Ay\|_2^2 = y^T A^T\tilde{X}\tilde{X}^T Ay = y^T A^T(I - \tilde{x}\tilde{x}^T)Ay \\ &= \|Ay\|_2^2 - (\tilde{x}^T Ay)^2 \leq \|Ay\|_2^2, \end{aligned}$$

z čehož plyne

$$\sigma'_{11} = \max\{\|\tilde{X}^T A \tilde{Y} z\|_2;\ \|z\|_2 = 1,\ z \in \mathbb{R}^{n-1}\} \leq \max\{\|Ay\|_2;\ \|y\|_2 = 1,\ y \in \mathbb{R}^n\} = \sigma,$$

čímž je indukční krok proveden. $\square$

Věta 35 *Je-li*

$$A = X\Sigma Y^T \tag{12}$$

libovolný SVD rozklad matice $A \in \mathbb{R}^{m\times n}$, potom:

1) $\mathrm{rank}(A) = r$, *kde r je počet kladných prvků na diagonále Σ,*
2) $X_{\bullet 1}, \dots, X_{\bullet r}$ *tvoří ortonormální bázi sloupcového prostoru $\mathcal{R}(A)$,*
3) $Y_{\bullet r+1}, \dots, Y_{\bullet n}$ *tvoří ortonormální bázi prostoru $\mathcal{N}(A)$.*

Důkaz Matice Σ má v prvních r řádcích kladné diagonální prvky a posledních $m-r$ řádků nulových. Z (12) plyne, že pro každé $1 \leq j \leq r$ je

$$X_{\bullet j} = \tfrac{1}{\Sigma_{jj}} A Y_{\bullet j} \in \mathcal{R}(A)$$

a

$$A_{\bullet j} = \sum_{k=1}^{r} (\Sigma Y^T)_{kj} X_{\bullet k},$$

takže $X_{\bullet 1}, \dots, X_{\bullet r}$ generují $\mathcal{R}(A)$ a protože jsou lineárně nezávislé, tvoří jeho ortonormální bázi. Z toho ihned dostáváme, že $\mathrm{rank}(A) = \dim \mathcal{R}(A) = r$. Protože $A^T = Y\Sigma^T X^T$, tvoří podle právě dokázaného tvrzení vektory $Y_{\bullet 1}, \dots, Y_{\bullet r}$ ortonormální bázi $\mathcal{R}(A^T)$, takže vektory $Y_{\bullet r+1}, \dots, Y_{\bullet n}$ tvoří ortonormální bázi jeho ortogonálního doplňku $\mathcal{R}(A^T)^\perp = \mathcal{N}(A)$ (věta 1). $\square$

Věta 36 (polární rozklad) *Ke každé matici $A \in \mathbb{R}^{n\times n}$ existují pozitivně semidefinitní matice P, S a ortogonální matice Q tak, že platí*

$$A = PQ = QS. \tag{13}$$

Přitom matice P, S jsou určeny jednoznačně. Je-li A regulární, potom P, S jsou pozitivně definitní a i Q je určena jednoznačně.

Důkaz Nechť $A = X\Sigma Y^T$ je SVD rozklad matice A. Položme $P = X\Sigma X^T$, $S = Y\Sigma Y^T$, $Q = XY^T$. Potom vzhledem k ortogonalitě X, Y je

$$PQ = X\Sigma X^T X Y^T = X\Sigma Y^T = A,$$

$$QS = XY^T Y\Sigma Y^T = X\Sigma Y^T = A,$$

což dokazuje (13). Dále $Q = XY^T$ je ortogonální a pro každé $x \in \mathbb{R}^n$ je $x^T P x = \sum_{i=1}^n \Sigma_{ii}(X^T x)_i^2 \geq 0$, takže P je pozitivně semidefinitní, a stejným způsobem se totéž dokáže pro S. Přitom platí

$$P^2 = X\Sigma X^T X\Sigma X^T = X\Sigma^2 X^T = X\Sigma Y^T Y\Sigma X^T = AA^T,$$

$$S^2 = Y\Sigma Y^T Y\Sigma Y^T = Y\Sigma^2 Y^T = Y\Sigma X^T X\Sigma Y^T = A^T A.$$

Pro důkaz jednoznačnosti P, S se zde musíme odvolat na větu 70, kterou dokážeme později, podle které ke každé pozitivně semidefinitní matici C existuje právě jedna pozitivně semidefinitní matice B s vlastností $B^2 = C$. Aplikujeme-li tuto větu na rovnosti $P^2 = AA^T$, $S^2 = A^T A$, dostáváme jednoznačnost P a S. Je-li A regulární, potom z (13) plyne, že i P, S jsou regulární, tedy pozitivně definitní, a v tom případě je i $Q = P^{-1}A$ určena jednoznačně. □

9 Moore-Penroseova inverze

Věta 37 (Moore-Penrose) *Pro každou matici $A \in \mathbb{R}^{m\times n}$ existuje právě jedna matice $A^+ \in \mathbb{R}^{n\times m}$ s těmito vlastnostmi:*

1) $AA^+A = A$,
2) $A^+AA^+ = A^+$,
3) $(AA^+)^T = AA^+$,
4) $(A^+A)^T = A^+A$.

Důkaz Je-li $A = 0 \in \mathbb{R}^{m\times n}$, potom $A^+ = 0 \in \mathbb{R}^{n\times m}$ splňuje 1)-4). Nechť tedy $A \neq 0$, takže $r = \mathrm{rank}(A) \geq 1$. Nechť $A = BC$ je hodnostní rozklad matice A (věta 3), potom $B \in \mathbb{R}^{m\times r}$, $C \in \mathbb{R}^{r\times n}$ a $\mathrm{rank}(B) = \mathrm{rank}(C) = r$, tedy B má lineárně nezávislé sloupce a C má lineárně nezávislé řádky, takže matice B^TB a CC^T jsou regulární a jejich inverze existují. Položme nyní

$$A^+ = C^T(CC^T)^{-1}(B^TB)^{-1}B^T.$$

Ukážeme, že A^+ splňuje 1)-4).

1) $AA^+A = BCC^T(CC^T)^{-1}(B^TB)^{-1}B^TBC = BC = A$,

2) $A^+AA^+ = C^T(CC^T)^{-1}(B^TB)^{-1}B^TBCC^T(CC^T)^{-1}(B^TB)^{-1}B^T$
$= C^T(CC^T)^{-1}(B^TB)^{-1}B^T = A^+$,

3) $AA^+ = BCC^T(CC^T)^{-1}(B^TB)^{-1}B^T = B(B^TB)^{-1}B^T = (B(B^TB)^{-1}B^T)^T = (AA^+)^T$,

4) $A^+A = C^T(CC^T)^{-1}(B^TB)^{-1}B^TBC = C^T(CC^T)^{-1}C = (C^T(CC^T)^{-1}C)^T = (A^+A)^T$.

Pro důkaz jednoznačnosti předpokládejme, že jistá matice $A^\#$ splňuje 1)-4). Položme $D = A^+ - A^\#$. Z $AA^+A = A$ plyne $A^+AA^T = A^T$, podobně $A^\#AA^T = A^T$, odečtením $DAA^T = 0$, tudíž $DA(DA)^T = DAA^TD^T = 0$ a odtud podle věty 2, tvrzení 6), je $DA = 0$. Dále z $A^+AA^+ = A^+$ plyne $AA^+(A^+)^T = (A^+)^T$ a podobně $AA^\#(A^\#)^T = (A^\#)^T$, tedy $DD^T = D((A^+)^T - (A^\#)^T) = D(AA^+(A^+)^T - AA^\#(A^\#)^T) = DA(A^+(A^+)^T - A^\#(A^\#)^T) = 0$ a odtud opět podle věty 2 je $D = 0$, tj. $A^+ = A^\#$. □

Současně jsme dokázali tuto větu:

Věta 38 *Nechť $A = BC$ je libovolný hodnostní rozklad matice A. Potom platí*

$$A^+ = C^T(CC^T)^{-1}(B^TB)^{-1}B^T = C^T(B^TAC^T)^{-1}B^T.$$

Definice Matice A^+ splňující vlastnosti 1)-4) z věty 37 se nazývá pseudoinverzní matice (resp. Moore-Penroseova inverze).

Věta 39 *Je-li*

$$A = X \begin{pmatrix} S & 0 \\ 0 & 0 \end{pmatrix} Y^T$$

libovolný SVD rozklad matice $A \neq 0$, potom

$$A^+ = Y \begin{pmatrix} S^{-1} & 0 \\ 0 & 0 \end{pmatrix} X^T.$$

Důkaz Dokážeme, že A^+ splňuje 1)-4). Přitom použijeme častěji faktu, že $X^T X = I$, $Y^T Y = I$ a

$$\begin{pmatrix} S & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} S^{-1} & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} S^{-1} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} S & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}.$$

Platí

$$AA^+A = X \begin{pmatrix} S & 0 \\ 0 & 0 \end{pmatrix} Y^T Y \begin{pmatrix} S^{-1} & 0 \\ 0 & 0 \end{pmatrix} X^T X \begin{pmatrix} S & 0 \\ 0 & 0 \end{pmatrix} Y^T = X \begin{pmatrix} S & 0 \\ 0 & 0 \end{pmatrix} Y^T = A,$$

$$A^+AA^+ = Y \begin{pmatrix} S^{-1} & 0 \\ 0 & 0 \end{pmatrix} X^T X \begin{pmatrix} S & 0 \\ 0 & 0 \end{pmatrix} Y^T Y \begin{pmatrix} S^{-1} & 0 \\ 0 & 0 \end{pmatrix} X^T = Y \begin{pmatrix} S^{-1} & 0 \\ 0 & 0 \end{pmatrix} X^T = A^+,$$

čili A^+ splňuje 1) a 2), dále

$$AA^+ = X \begin{pmatrix} S & 0 \\ 0 & 0 \end{pmatrix} Y^T Y \begin{pmatrix} S^{-1} & 0 \\ 0 & 0 \end{pmatrix} X^T = X \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} X^T,$$

tedy AA^+ je symetrická a proto

$$(AA^+)^T = AA^+,$$

a nakonec

$$A^+A = Y \begin{pmatrix} S^{-1} & 0 \\ 0 & 0 \end{pmatrix} X^T X \begin{pmatrix} S & 0 \\ 0 & 0 \end{pmatrix} Y^T = Y \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} Y^T,$$

tedy A^+A je rovněž symetrická a proto

$$(A^+A)^T = A^+A,$$

takže A^+ splňuje 1)-4). □

Věta 40 *Moore-Penroseova inverze A^+ matice $A \in \mathbb{R}^{m \times n}$ má tyto vlastnosti:*

1) $A^+ = A^{-1}$ *je-li A čtvercová regulární,*
2) $A^+ = (A^T A)^{-1} A^T$ *je-li* $\operatorname{rank}(A) = n$,
3) $A^+ = A^T (AA^T)^{-1}$ *je-li* $\operatorname{rank}(A) = m$,

4) $(A^+)^+ = A$,
5) $(A^T)^+ = (A^+)^T$,
6) $A^T = A^T AA^+ = A^+ AA^T$,
7) $A^+ = (A^T A)^+ A^T = A^T (AA^T)^+$,
8) $(A^T A)^+ = A^+ (A^T)^+$, $(AA^T)^+ = (A^T)^+ A^+$,
9) $(A^T A)^+ A^T A = A^+ A$, $(AA^T)^+ AA^T = AA^+$,
10) $\mathcal{R}(A^+) = \mathcal{R}(A^T) = \mathcal{R}(A^+ A)$,
11) $\mathcal{N}(A^+) = \mathcal{N}(A^T) = \mathcal{N}(AA^+)$,
12) $\operatorname{rank}(A^+) = \operatorname{rank}(A) = \operatorname{rank}(A^+ A) = \operatorname{rank}(AA^+)$.

Důkaz Vlastnosti 1) a 2) dokážeme společně. Je-li $\operatorname{rank}(A) = n$, potom $A^T A$ je regulární a pro matici $A^+ = (A^T A)^{-1} A^T$ platí $A^+ A = (A^T A)^{-1} A^T A = I$; odtud dostáváme $AA^+ A = AI = A$, $A^+ AA^+ = IA^+ = A^+$, $(A^+ A)^T = I = A^+ A$ a $(AA^+)^T = (A(A^T A)^{-1} A^T)^T = A(A^T A)^{-1} A^T = AA^+$, čímž jsme ověřili, že A^+ má vlastnosti 1)-4) z definice pseudoinverzní matice a proto platí 2). Je-li A čtvercová regulární, je $\operatorname{rank}(A) = n$ a právě dokázaný výsledek dává $A^+ = (A^T A)^{-1} A^T = A^{-1} (A^T)^{-1} A^T = A^{-1}$, což je 1).

3) Je-li $\operatorname{rank}(A) = m$, potom AA^T je regulární a pro matici $A^+ = A^T (AA^T)^{-1}$ platí $AA^+ = I$, odtud $AA^+ A = IA = A$, $A^+ AA^+ = A^+ I = A^+$, $(AA^+)^T = I = AA^+$ a $(A^+ A)^T = (A^T (AA^T)^{-1} A)^T = A^T (AA^T)^{-1} A = A^+ A$, čímž jsme opět ověřili vlastnosti 1)-4) z definice.

Před důkazem dalších tvrzení 4)-12) si povšimněme, že všechna jsou splněna pro $A = 0$ a $A^+ = 0$. Můžeme proto v dalším předpokládat, že $A \neq 0$. V tom případě má A podle věty 34 SVD rozklad tvaru

$$A = X \begin{pmatrix} S & 0 \\ 0 & 0 \end{pmatrix} Y^T, \tag{14}$$

kde $S \in \mathbb{R}^{r \times r}$, $r \geq 1$, je diagonální matice s kladnými diagonálními prvky a podle věty 39 platí

$$A^+ = Y \begin{pmatrix} S^{-1} & 0 \\ 0 & 0 \end{pmatrix} X^T. \tag{15}$$

Tohoto explicitního tvaru pseudoinverzní matice použijeme k důkazu dalších tvrzení.

4) Podle (14), (15) a věty 39 je

$$(A^+)^+ = X \begin{pmatrix} S & 0 \\ 0 & 0 \end{pmatrix} Y^T = A.$$

5) Podle (14) je

$$A^T = Y \begin{pmatrix} S & 0 \\ 0 & 0 \end{pmatrix} X^T \tag{16}$$

a opět podle věty 39 je

$$(A^T)^+ = X \begin{pmatrix} S^{-1} & 0 \\ 0 & 0 \end{pmatrix} Y^T = (A^+)^T.$$

6) Platí $A^TAA^+ = A^T(AA^+)^T = (AA^+A)^T = A^T$, $A^+AA^T = (A^+A)^TA^T = (AA^+A)^T = A^T$.

7) Podle (14), (16) je

$$A^TA = Y\begin{pmatrix} S^2 & 0 \\ 0 & 0 \end{pmatrix}Y^T \tag{17}$$

a tedy

$$(A^TA)^+A^T = Y\begin{pmatrix} S^{-2} & 0 \\ 0 & 0 \end{pmatrix}Y^TY\begin{pmatrix} S & 0 \\ 0 & 0 \end{pmatrix}X^T = Y\begin{pmatrix} S^{-1} & 0 \\ 0 & 0 \end{pmatrix}X^T = A^+.$$

Aplikací tohoto výsledku na transponovanou matici dostáváme s použitím tvrzení 5)

$$A^+ = ((A^T)^+)^T = ((AA^T)^+A)^T = A^T((AA^T)^+)^T = A^T((AA^T)^T)^+ = A^T(AA^T)^+.$$

8) S použitím (15), (16), (17) dostáváme

$$A^+(A^T)^+ = Y\begin{pmatrix} S^{-1} & 0 \\ 0 & 0 \end{pmatrix}X^TX\begin{pmatrix} S^{-1} & 0 \\ 0 & 0 \end{pmatrix}Y^T = Y\begin{pmatrix} S^{-2} & 0 \\ 0 & 0 \end{pmatrix}Y^T = (A^TA)^+,$$

z čehož dosazením $A := A^T$ plyne druhá část tvrzení $(A^T)^+A^+ = (AA^T)^+$.

9) Podle (17), (15) je

$$(A^TA)^+A^TA = Y\begin{pmatrix} S^{-2} & 0 \\ 0 & 0 \end{pmatrix}Y^TY\begin{pmatrix} S^2 & 0 \\ 0 & 0 \end{pmatrix}Y^T = Y\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}Y^T,$$

$$A^+A = Y\begin{pmatrix} S^{-1} & 0 \\ 0 & 0 \end{pmatrix}X^TX\begin{pmatrix} S & 0 \\ 0 & 0 \end{pmatrix}Y^T = Y\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}Y^T, \tag{18}$$

z čehož plyne rovnost $(A^TA)^+A^TA = A^+A$. Druhá rovnost se dostane opět dosazením $A := A^T$.

10) Je-li $y \in \mathcal{R}(A^+)$, potom pro jisté x je $y = A^+x = A^T((AA^T)^+x) \in \mathcal{R}(A^T)$; je-li $y' \in \mathcal{R}(A^T)$, potom $y' = A^Tx' = A^+AA^Tx' \in \mathcal{R}(A^+A)$; je-li $y'' \in \mathcal{R}(A^+A)$, potom $y'' = A^+Ax'' \in \mathcal{R}(A^+)$ (použili jsme tvrzení 7) a 6)). Dokázali jsme tedy, že $\mathcal{R}(A^+) \subseteq \mathcal{R}(A^T) \subseteq \mathcal{R}(A^+A) \subseteq \mathcal{R}(A^+)$, takže všude platí rovnost.

11) Je-li $x \in \mathcal{N}(A^+)$, potom $A^+x = 0$ a podle 6) je $A^Tx = 0$, což dává $\mathcal{N}(A^+) \subseteq \mathcal{N}(A^T)$, podobně 7) dává $\mathcal{N}(A^T) \subseteq \mathcal{N}(AA^+)$, dále opět podle 6) je $\mathcal{N}(AA^+) \subseteq \mathcal{N}(A^T)$ a podle 7) $\mathcal{N}(A^T) \subseteq \mathcal{N}(A^+)$. Tím jsme dokázali $\mathcal{N}(A^+) \subseteq \mathcal{N}(A^T) \subseteq \mathcal{N}(AA^+) \subseteq \mathcal{N}(A^T) \subseteq \mathcal{N}(A^+)$, takže všude platí rovnost.

12) Jelikož hodnost matice se při násobení ortogonální (tj. regulární) maticí nemění, plyne z (14), (15), že $\mathrm{rank}(A) = r = \mathrm{rank}(A^+)$ a podobně z (18) plyne $r = \mathrm{rank}(A^+A)$ a aplikací této rovnosti na $A := A^T$ dostáváme nakonec $r = \mathrm{rank}((A^T)^+A^T) = \mathrm{rank}((AA^+)^T) = \mathrm{rank}(AA^+)$. □

Poznámka Na rozdíl od inverzní matice obecně neplatí $AA^+ = A^+A$, $(AB)^+ = B^+A^+$.

Věta 41 (Greville) *Nechť $A \in \mathbb{R}^{m\times n}$, $a \in \mathbb{R}^m$. Potom je*

$$(A\ a)^+ = \begin{pmatrix} A^+ - db^T \\ b^T \end{pmatrix},$$

kde

$$\begin{aligned} d &= A^+a, \\ c &= a - Ad, \\ b^T &= \begin{cases} \frac{c^T}{c^Tc} & \text{je} - \text{li } c \neq 0, \\ \frac{d^TA^+}{1+d^Td} & \text{je} - \text{li } c = 0. \end{cases} \end{aligned}$$

Důkaz Nejprve dokážeme některé pomocné vztahy. Především je

$$c^TA = (a - AA^+a)^TA = a^T(I - AA^+)A = a^T(A - AA^+A) = 0^T \tag{19}$$

podle vlastnosti 1) pseudoinverzní matice, a dále

$$c^T(c - a) = c^T(-Ad) = -(c^TA)d = 0 \tag{20}$$

podle (19). Nyní dokážeme, že

$$b^TA = \begin{cases} 0^T & \text{je} - \text{li } c \neq 0, \\ \frac{d^T}{1+d^Td} & \text{je} - \text{li } c = 0, \end{cases} \tag{21}$$

a

$$b^Ta = \begin{cases} 1 & \text{je} - \text{li } c \neq 0, \\ \frac{d^Td}{1+d^Td} & \text{je} - \text{li } c = 0. \end{cases} \tag{22}$$

Skutečně, je-li $c \neq 0$, potom

$$b^TA = \tfrac{c^TA}{c^Tc} = 0^T$$

podle (19), a pro $c = 0$ je

$$b^TA = \tfrac{d^TA^+A}{1+d^Td} = \tfrac{a^T(A^+)^TA^+A}{1+d^Td} = \tfrac{a^T(A^+AA^+)^T}{1+d^Td} = \tfrac{a^T(A^+)^T}{1+d^Td} = \tfrac{d^T}{1+d^Td},$$

čímž je (21) dokázáno. Dále pro $c \neq 0$ je

$$b^Ta = \tfrac{c^Ta}{c^Tc} = \tfrac{c^Tc}{c^Tc} = 1$$

podle (20), a pro $c = 0$ je

$$b^Ta = \tfrac{d^TA^+a}{1+d^Td} = \tfrac{d^Td}{1+d^Td},$$

což dokazuje (22). Nyní dokážeme, že matice $(A\ a)$ a $\begin{pmatrix} A^+ - db^T \\ b^T \end{pmatrix}$ mají vlastnosti 1)-4) z věty 37. Důkaz rozdělíme do dvou částí.

(a) Nechť $c \neq 0$, tj. $a \neq Ad = AA^+a$. Potom

$$\begin{pmatrix} A^+ - db^T \\ b^T \end{pmatrix} (A\ a) = \begin{pmatrix} A^+A - db^TA & A^+a - db^Ta \\ b^TA & b^Ta \end{pmatrix} = \begin{pmatrix} A^+A & 0 \\ 0^T & 1 \end{pmatrix}$$

podle (21), (22), takže matice je symetrická, což dokazuje vlastnost 4) z věty 37. Odtud dostáváme

$$\begin{aligned} \begin{pmatrix} A^+ - db^T \\ b^T \end{pmatrix} (A\ a) \begin{pmatrix} A^+ - db^T \\ b^T \end{pmatrix} &= \begin{pmatrix} A^+A & 0 \\ 0^T & 1 \end{pmatrix} \begin{pmatrix} A^+ - db^T \\ b^T \end{pmatrix} \\ &= \begin{pmatrix} A^+ - A^+AA^+ab^T \\ b^T \end{pmatrix} = \begin{pmatrix} A^+ - db^T \\ b^T \end{pmatrix}, \end{aligned}$$

což dokazuje vlastnost 2). Dále

$$(A\ a)\begin{pmatrix} A^+ - db^T \\ b^T \end{pmatrix}(A\ a) = (A\ a)\begin{pmatrix} A^+A & 0 \\ 0^T & 1 \end{pmatrix} = (AA^+A\ \ a) = (A\ a),$$

což je vlastnost 1), a nakonec

$$(A\ a)\begin{pmatrix} A^+ - db^T \\ b^T \end{pmatrix} = AA^+ + cb^T = AA^+ + \tfrac{cc^T}{c^Tc},$$

tedy matice je symetrická, což dokazuje vlastnost 3).

(b) Nechť $c = 0$, tj. $a = Ad = AA^+a$. Potom

$$(A\ a)\begin{pmatrix} A^+ - db^T \\ b^T \end{pmatrix} = AA^+ - Adb^T + ab^T = AA^+,$$

takže matice je symetrická, což dokazuje vlastnost 3). Odtud

$$(A\ a)\begin{pmatrix} A^+ - db^T \\ b^T \end{pmatrix}(A\ a) = AA^+(A\ a) = (A\ Ad) = (A\ a),$$

což je vlastnost 1), dále

$$\begin{pmatrix} A^+ - db^T \\ b^T \end{pmatrix}(A\ a)\begin{pmatrix} A^+ - db^T \\ b^T \end{pmatrix} = \begin{pmatrix} A^+ - db^T \\ b^T \end{pmatrix}AA^+ = \begin{pmatrix} A^+ - db^TAA^+ \\ b^TAA^+ \end{pmatrix}$$

$$= \begin{pmatrix} A^+ - \frac{dd^TA^+AA^+}{1+d^Td} \\ \frac{d^TA^+AA^+}{1+d^Td} \end{pmatrix} = \begin{pmatrix} A^+ - \frac{dd^TA^+}{1+d^Td} \\ \frac{d^TA^+}{1+d^Td} \end{pmatrix} = \begin{pmatrix} A^+ - db^T \\ b^T \end{pmatrix},$$

což je vlastnost 2), a nakonec

$$\begin{aligned}\begin{pmatrix} A^+ - db^T \\ b^T \end{pmatrix}(A\ a) &= \begin{pmatrix} A^+A - db^TA & A^+a - db^Ta \\ b^TA & b^Ta \end{pmatrix} \\ &= \begin{pmatrix} A^+A - \frac{dd^T}{1+d^Td} & \frac{d}{1+d^Td} \\ \frac{d^T}{1+d^Td} & \frac{d^Td}{1+d^Td} \end{pmatrix}\end{aligned}$$

podle (21), (22), tedy matice je symetrická, což dokazuje vlastnost 4).

V obou případech (a), (b) jsme ukázali, že matice $(A\ a)$ a $\begin{pmatrix} A^+ - db^T \\ b^T \end{pmatrix}$ splňují vlastnosti 1)-4), čímž je tvrzení věty dokázáno. □

Věta 42 *Je-li $a \in \mathbb{R}^m$, potom*

$$a^+ = \begin{cases} \frac{a^T}{a^Ta} & \text{je} - \text{li } a \neq 0, \\ 0^T & \text{je} - \text{li } a = 0. \end{cases}$$

Důkaz Tvrzení věty plyne přímým ověřením vlastností 1)-4) z věty 37. $\square$

Grevillův algoritmus pro výpočet A^+ v MATLABu:

```
function [X]=greville(A)
% computes the pseudoinverse X of A by Greville's algorithm
[m,n]=size(A); tol=1.0e-10;
d=A(:,1);
if all(abs(d)<tol*ones(m,1)), X=zeros(1,m); else X=d'/(d'*d); end
for j=2:n
  d=X*A(:,j); c=A(:,j)-A(:,1:(j-1))*d;
  if all(abs(c)<tol*ones(m,1)), bt=d'*X/(1+d'*d); else bt=c'/(c'*c); end
  X=[X-d*bt; bt];
end
```

10 Ortogonální projekce

Věta 43 *Nechť je dána matice $A \in \mathbb{R}^{m\times n}$. Potom pro libovolné $x \in \mathbb{R}^m$, $y \in \mathbb{R}^n$ platí:*

1) $x_{\mathcal{R}(A)} = AA^+x$,
2) $x_{\mathcal{N}(A^T)} = (I - AA^+)x$,
3) $y_{\mathcal{R}(A^T)} = A^+Ay$,
4) $y_{\mathcal{N}(A)} = (I - A^+A)y$.

Důkaz Z ortogonálního rozkladu $\mathbb{R}^m = \mathcal{R}(A) \oplus \mathcal{N}(A^T)$ (věta 1, tvrzení 3)) plyne, že každý vektor $x \in \mathbb{R}^m$ lze psát právě jedním způsobem ve tvaru $x = x' + x''$, kde $x' \in \mathcal{R}(A)$ a $x'' \in \mathcal{N}(A^T)$; potom je $x' = x_{\mathcal{R}(A)}$, $x'' = x_{\mathcal{N}(A^T)}$. Položme $x' = AA^+x$ a $x'' = (I - AA^+)x$, potom $x' = A(A^+x) \in \mathcal{R}(A)$, $A^Tx'' = (A^T - A^TAA^+)x = 0$ (věta 40, tvrzení 6)), tedy $x'' \in \mathcal{N}(A^T)$, a současně $x' + x'' = x$, takže x' a x'' tvoří ortogonální rozklad x a proto $x' = x_{\mathcal{R}(A)}$, $x'' = x_{\mathcal{N}(A^T)}$, což dokazuje tvrzení 1) a 2). Tvrzení 3), 4) se dokáží aplikací 1), 2) na matici A^T s využitím faktu, že $A^T(A^T)^+ = A^T(A^+)^T = (A^+A)^T = A^+A$ (podle tvrzení 5) věty 40 a definice A^+). $\square$

Věta 44 *Pro každou matici $A \in \mathbb{R}^{m\times n}$ platí*

1) $\mathcal{R}(A) = \mathcal{N}(I - AA^+)$,
2) $\mathcal{N}(A) = \mathcal{R}(I - A^+A)$.

Důkaz 1) Je-li $x \in \mathcal{R}(A)$, potom $x = Ay$ pro jisté $y \in \mathbb{R}^n$ a platí $AA^+x = AA^+Ay = Ay = x$, tedy $(I - AA^+)x = 0$, takže $x \in \mathcal{N}(I - AA^+)$. Naopak, je-li $x \in \mathcal{N}(I - AA^+)$, potom $(I - AA^+)x = 0$ a tedy $x = AA^+x = A(A^+x) \in \mathcal{R}(A)$.

2) Je-li $x \in \mathcal{N}(A)$, potom x se ortogonální projekcí na $\mathcal{N}(A)$ zobrazuje na sebe, proto podle věty 43, tvrzení 4), je $x = x_{\mathcal{N}(A)} = (I - A^+A)x \in \mathcal{R}(I - A^+A)$. Naopak, je-li $x \in \mathcal{R}(I-A^+A)$, potom $x = (I-A^+A)y$ pro jisté $y \in \mathbb{R}^n$ a platí $Ax = (A-AA^+A)y = 0$ (definice A^+), tedy $x \in \mathcal{N}(A)$. $\square$

Věta 45 *Je-li* $A = X\Sigma Y^T$ *libovolný SVD rozklad matice* $0 \neq A \in \mathbb{R}^{m\times n}$, *potom*

$$\begin{aligned} AA^+ &= \hat{X}\hat{X}^T, \\ I - A^+A &= \hat{Y}\hat{Y}^T, \end{aligned}$$

kde $\hat{X}$ *je matice sestávající z prvních* r *sloupců matice* X, $\hat{Y}$ *matice sestávající z posledních* $n-r$ *sloupců matice* Y *a* $r = \operatorname{rank}(A)$ *je počet kladných prvků na diagonále* Σ.

Důkaz Protože $A \neq 0$, má Σ tvar

$$\Sigma = \begin{pmatrix} S & 0 \\ 0 & 0 \end{pmatrix},$$

kde $S \in \mathbb{R}^{r\times r}$ je regulární a podle věty 39 platí

$$A^+ = Y \begin{pmatrix} S^{-1} & 0 \\ 0 & 0 \end{pmatrix} X^T.$$

Potom

$$AA^+ = X \begin{pmatrix} S & 0 \\ 0 & 0 \end{pmatrix} Y^T Y \begin{pmatrix} S^{-1} & 0 \\ 0 & 0 \end{pmatrix} X^T = X \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} X^T = \hat{X}\hat{X}^T$$

a podobně

$$I - A^+A = I - Y \begin{pmatrix} S^{-1} & 0 \\ 0 & 0 \end{pmatrix} X^T X \begin{pmatrix} S & 0 \\ 0 & 0 \end{pmatrix} Y^T = Y \begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix} Y^T = \hat{Y}\hat{Y}^T.$$

□

11 Soustavy lineárních rovnic

Věta 46 *Nechť* $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$. *Potom soustava*

$$Ax = b \tag{23}$$

má řešení právě když platí

$$AA^+b = b. \tag{24}$$

Je-li tato podmínka splněna, potom množina řešení $X(A,b)$ *soustavy (23) je popsána vztahy*

$$X(A,b) = A^+b + \mathcal{N}(A) = \{A^+b + (I - A^+A)y;\ y \in \mathbb{R}^n\}, \tag{25}$$

přičemž

$$\|A^+b\|_2 = \min\{\|x\|_2;\ x \in X(A,b)\} \tag{26}$$

a A^+b *je jediné řešení, ve kterém se tohoto minima nabývá.*

Důkaz Má-li soustava (23) řešení x_0, potom s využitím definičního vztahu $AA^+A = A$ dostáváme $AA^+b = AA^+Ax_0 = Ax_0 = b$, což je (24). Naopak, platí-li (24), potom $x = A^+b$ je řešením soustavy (23).

Nechť tedy platí (24). Potom $A^+b \in X(A,b)$ a pro každé $x \in X(A,b)$ je $A(x-A^+b) = b - b = 0$, tedy $x - A^+b \in \mathcal{N}(A)$ a $x \in A^+b + \mathcal{N}(A)$. Naopak, je-li $x \in A^+b + \mathcal{N}(A)$, potom $x = A^+b + y$ pro jisté $y \in \mathcal{N}(A)$ a tedy $Ax = AA^+b = b$, takže $x \in X(A,b)$. Tím jsme dokázali rovnost

$$X(A,b) = A^+b + \mathcal{N}(A).$$

Z popisu množiny $\mathcal{N}(A)$ v tvrzení 2) věty 44 nyní plyne

$$X(A,b) = A^+b + \mathcal{R}(I - A^+A) = \{A^+b + (I - A^+A)y;\ y \in \mathbb{R}^n\},$$

což je (25). Pro důkaz (26) zvolme libovolné $x \in X(A,b)$; to lze podle (25) psát ve tvaru $x = A^+b + y$, kde $y \in \mathcal{N}(A)$. Přitom podle tvrzení 10) věty 40 a tvrzení 2) věty 1 je $A^+b \in \mathcal{R}(A^+) = \mathcal{R}(A^T) = \mathcal{N}(A)^\perp$, z čehož plyne, že $(A^+b)^Ty = 0$ a tedy

$$\|x\|_2^2 = (A^+b + y)^T(A^+b + y) = \|A^+b\|_2^2 + \|y\|_2^2 = \|A^+b\|_2^2 + \|x - A^+b\|_2^2 \geq \|A^+b\|_2^2,$$

což dokazuje (26), a rovnost se nabývá právě když $\|x - A^+b\|_2 = 0$, tj. $x = A^+b$. □

Věta 47 *Soustava (23) má jediné řešení právě když platí*

$$AA^+b = b, \quad A^+A = I. \tag{27}$$

V tom případě je jediným řešením soustavy (23) vektor A^+b.

Důkaz Má-li soustava (23) jediné řešení, potom podle věty 46 platí (24) a z popisu (25) plyne, že musí být $I - A^+A = 0$. Naopak, platí-li (27), potom A^+b je řešením (23) a z (25) plyne, že je jediné. □

12 Metoda nejmenších čtverců

Věta 48 *Nechť $A \in \mathbb{R}^{m\times n}$ a $b \in \mathbb{R}^m$. Pro soustavu*

$$Ax = b \tag{28}$$

a jí odpovídající soustavu normálních rovnic

$$A^TAx = A^Tb \tag{29}$$

platí:

1) *soustava (29) má vždy řešení,*
2) *má-li soustava (28) řešení, potom obě soustavy mají stejnou množinu řešení.*

Důkaz 1) S použitím vlastností 7), 6) z věty 40 dostáváme $(A^TA)(A^TA)^+A^Tb = (A^TA)((A^TA)^+A^T)b = A^TAA^+b = A^Tb$, což je rovnost (24) z věty 46 pro soustavu (29). To znamená, že soustava (29) má řešení.

2) Má-li soustava (28) řešení, potom podle věty 46 je množina jejích řešení $X(A,b)$ popsaná vztahem

$$X(A,b) = \{A^+b + (I - A^+A)y;\ y \in \mathbb{R}^n\} \tag{30}$$

a podle téže věty s ohledem na již dokázanou existenci řešení soustavy (29) platí

$$X(A^TA, A^Tb) = \{(A^TA)^+A^Tb + (I - (A^TA)^+A^TA)y;\ y \in \mathbb{R}^n\}. \tag{31}$$

Avšak podle vlastností 7), 9) z věty 40 je

$$\begin{aligned}(A^TA)^+A^Tb &= A^+b,\\ (A^TA)^+A^TA &= A^+A\end{aligned}$$

a dosazením do (31) dostáváme z (30)

$$X(A^TA, A^Tb) = \{A^+b + (I - A^+A)y;\ y \in \mathbb{R}^n\} = X(A,b),$$

tj. soustavy (28) a (29) mají stejnou množinu řešení. □

Věta 49 *Následující tvrzení jsou ekvivalentní:*

1) x je řešením soustavy (29),

2) $Ax = b_{\mathcal{R}(A)}$, kde $b_{\mathcal{R}(A)}$ je ortogonální projekce b na sloupcový prostor $\mathcal{R}(A)$,

3) $\|Ax - b\|_2 = \min\{\|Ay - b\|_2;\ y \in \mathbb{R}^n\}$.

Důkaz Dokážeme 1)⇒2)⇒3)⇒2)⇒1).

1)⇒2): Je-li $A^TAx = A^Tb$, potom $A^Ty = 0$ pro $y = b - Ax$, takže $b = Ax + y$, kde $Ax \in \mathcal{R}(A)$ a $y \in \mathcal{N}(A^T)$. Podle věty 1, tvrzení 3) to znamená, že $b = Ax + y$ je ortogonální rozklad vektoru b, a tedy $Ax = b_{\mathcal{R}(A)}$.

2)⇒3): Protože $b_{\mathcal{R}(A)}$ jakožto ortogonální projekce vektoru b na prostor $\mathcal{R}(A)$ má ze všech bodů $\mathcal{R}(A)$ nejmenší vzdálenost od b, je $\|Ax - b\|_2 = \|b_{\mathcal{R}(A)} - b\|_2 = \min\{\|z - b\|_2;\ z \in \mathcal{R}(A)\} = \min\{\|Ay - b\|_2;\ y \in \mathbb{R}^n\}$.

3)⇒2): Platí-li 3), potom vektor $Ax \in \mathcal{R}(A)$ má ze všech vektorů v $\mathcal{R}(A)$ minimální vzdálenost od b, a takový bod je jak známo jednoznačně určen a roven $b_{\mathcal{R}(A)}$.

2)⇒1): Je-li $Ax = b_{\mathcal{R}(A)}$, potom podle věty 1, tvrzení 3) lze b psát ve tvaru $b = Ax+y$, kde $y \in \mathcal{N}(A^T)$. Potom $A^Tb = A^TAx + A^Ty = A^TAx$, takže x je řešením (29). □

13 Farkasova věta

Věta 50 (Farkas) *Nechť $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$. Potom soustava*

$$Ax = b \tag{32}$$

má nezáporné řešení právě když pro každý vektor $y \in \mathbb{R}^m$ takový, že $A^Ty \ge 0$, platí $b^Ty \ge 0$.

Důkaz a) Má-li soustava (32) řešení $x \geq 0$ a platí-li $A^T y \geq 0$ pro jistý vektor $y \in \mathbb{R}^m$, potom $b^T y = (Ax)^T y = x^T (A^T y) \geq 0$.

b) Pro důkaz opačné implikace předpokládejme, že soustava (32) *nemá* nezáporné řešení. Dokážeme, že potom existuje vektor $y \in \mathbb{R}^m$ takový, že $A^T y \geq 0$ a $b^T y < 0$, což je pro naše účely vhodnější psát ve sloupcovém tvaru $y^T A_{\bullet j} \geq 0$ $(j = 1, \ldots, n)$, $y^T b < 0$. Toto tvrzení dokážeme indukcí podle n.

b1) Je-li $n = 1$, potom A sestává z jediného sloupce a. Nechť $W = \{\alpha a;\ \alpha \in \mathbb{R}\}$ je podprostor generovaný vektorem a. Podle věty 1 existuje ortogonální rozklad vektoru b

$$b = b_W + b_{W^\perp},$$

kde $b_W \in W$ a $b_{W^\perp} \in W^\perp$. Rozlišíme dva případy. Je-li $b_{W^\perp} = 0$, potom $b \in W$, takže $b = \alpha a$ pro jisté $\alpha \in \mathbb{R}$. Jelikož podle předpokladu $Ax = b$ nemá nezáporné řešení, musí být $\alpha < 0$ a $a \neq 0$, takže položíme-li $y = a$, je $y^T a = \|a\|_2^2 \geq 0$ a $y^T b = \alpha \|a\|_2^2 < 0$ a tedy y je hledaný vektor. Je-li $b_{W^\perp} \neq 0$, položme $y = -b_{W^\perp}$, potom $y^T a = 0$ a $y^T b = -\|b_{W^\perp}\|_2^2 < 0$, takže y je opět hledaný vektor.

b2) Nechť tedy tvrzení platí pro $n - 1 \geq 1$ a nechť (32) nemá nezáporné řešení, přičemž $A \in \mathbb{R}^{m \times n}$. Potom ho nemá ani soustava

$$\sum_{j=1}^{n-1} A_{\bullet j} x_j = b$$

(jinak bychom pro $x_n = 0$ dostali nezáporné řešení (32)), proto podle indukčního předpokladu existuje $\overline{y} \in \mathbb{R}^m$ takové, že

$$\overline{y}^T A_{\bullet j} \geq 0 \qquad (j = 1, \ldots, n-1), \tag{33}$$
$$\overline{y}^T b < 0. \tag{34}$$

Je-li $\overline{y}^T A_{\bullet n} \geq 0$, je $\overline{y}$ hledaný vektor a jsme hotovi. Předpokládejme tedy, že

$$\overline{y}^T A_{\bullet n} < 0. \tag{35}$$

Položme

$$\begin{aligned} \alpha_j &= \overline{y}^T A_{\bullet j} \qquad (j = 1, \ldots, n), \\ \beta &= \overline{y}^T b, \end{aligned}$$

potom $\alpha_1 \geq 0, \ldots, \alpha_{n-1} \geq 0$, $\alpha_n < 0$, $\beta < 0$, a uvažujme soustavu

$$\sum_{j=1}^{n-1} (\alpha_n A_{\bullet j} - \alpha_j A_{\bullet n}) x_j = \alpha_n b - \beta A_{\bullet n}. \tag{36}$$

Kdyby tato soustava měla nezáporné řešení $x_1, \ldots, x_{n-1}$, potom bychom její úpravou dostali

$$\sum_{j=1}^{n-1} A_{\bullet j} x_j + A_{\bullet n} x_n = b, \tag{37}$$

kde

$$x_n = \frac{\beta - \sum_{j=1}^{n-1} \alpha_j x_j}{\alpha_n} > 0$$

vzhledem k (33), (34), (35), takže soustava (37), a tedy i (32), by měla nezáporné řešení $x_1, \ldots, x_n$ ve sporu s předpokladem. Proto soustava (36) nemá nezáporné řešení a tedy podle indukčního předpokladu existuje vektor $\tilde{y}$ takový, že

$$\begin{aligned} \tilde{y}^T(\alpha_n A_{\bullet j} - \alpha_j A_{\bullet n}) &\geq 0 \qquad (j = 1, \ldots, n-1), && (38) \\ \tilde{y}^T(\alpha_n b - \beta A_{\bullet n}) &< 0. && (39) \end{aligned}$$

Položíme nyní

$$y = \alpha_n \tilde{y} - (\tilde{y}^T A_{\bullet n})\overline{y}$$

a ukážeme, že y je hledaný vektor. Pro $j = 1, \ldots, n-1$ máme podle (38)

$$\begin{aligned} y^T A_{\bullet j} &= \alpha_n \tilde{y}^T A_{\bullet j} - (\tilde{y}^T A_{\bullet n})\overline{y}^T A_{\bullet j} \geq \alpha_j \tilde{y}^T A_{\bullet n} - (\tilde{y}^T A_{\bullet n})\overline{y}^T A_{\bullet j} \\ &= \alpha_j \tilde{y}^T A_{\bullet n} - (\tilde{y}^T A_{\bullet n})\alpha_j = 0, \qquad (40) \end{aligned}$$

pro $j = n$ dostáváme

$$y^T A_{\bullet n} = \alpha_n \tilde{y}^T A_{\bullet n} - (\tilde{y}^T A_{\bullet n})\overline{y}^T A_{\bullet n} = \alpha_n \tilde{y}^T A_{\bullet n} - (\tilde{y}^T A_{\bullet n})\alpha_n = 0, \qquad (41)$$

a nakonec podle (39)

$$y^T b = \alpha_n \tilde{y}^T b - (\tilde{y}^T A_{\bullet n})\overline{y}^T b < \beta \tilde{y}^T A_{\bullet n} - (\tilde{y}^T A_{\bullet n})\overline{y}^T b = \beta \tilde{y}^T A_{\bullet n} - (\tilde{y}^T A_{\bullet n})\beta = 0, \qquad (42)$$

takže z (40), (41), (42) plyne $A^T y \geq 0$ a $b^T y < 0$, čili y je hledaný vektor a důkaz indukcí je tím dokončen. □

14 Metoda sdružených gradientů

Algoritmus (řešení $Ax = b$ s pozitivně definitní maticí A).

0. Zvol x_0, polož $g_0 := Ax_0 - b$, $d_0 := -g_0$, $i := 0$.
1. Je-li $g_i = 0$, ukonči: x_i je řešením $Ax = b$.
2. Jinak vypočti

$$\begin{aligned} \alpha_i &:= -\frac{d_i^T g_i}{d_i^T A d_i}, \\ x_{i+1} &:= x_i + \alpha_i d_i, \\ g_{i+1} &:= Ax_{i+1} - b, \\ \beta_i &:= \frac{g_{i+1}^T g_{i+1}}{g_i^T g_i}, \\ d_{i+1} &:= -g_{i+1} + \beta_i d_i. \end{aligned}$$

3. Polož $i := i + 1$ a jdi na 1.

Věta 51 *Nechť A je pozitivně definitní. Potom pro každé $i \geq 0$ takové, že $g_i \neq 0$, je krok 2 proveditelný a platí*

$$\begin{aligned} g_{i+1}^T d_i &= 0, & & (43)\\ g_{i+1}^T (d_{i+1} + g_{i+1}) &= 0, & & (44)\\ g_{i+1}^T g_j &= 0 \qquad (0 \leq j < i+1), & & (45)\\ d_{i+1}^T A d_j &= 0 \qquad (0 \leq j < i+1). & & (46) \end{aligned}$$

Důkaz Využijeme častěji faktu, že z kroku 2 algoritmu vyplývá $g_{i+1} - g_i = A(x_{i+1} - x_i) = \alpha_i A d_i$, tedy

$$g_{i+1} = g_i + \alpha_i A d_i \tag{47}$$

pro každé $i \geq 0$.

Důkaz se provádí indukcí podle i. Pro $i = 0$ máme dokázat

$$\begin{aligned} g_1^T d_0 &= 0, & & (48)\\ g_1^T (d_1 + g_1) &= 0, & & (49)\\ g_1^T g_0 &= 0, & & (50)\\ d_1^T A d_0 &= 0. & & (51) \end{aligned}$$

Je-li $g_0 \neq 0$, potom $d_0 = -g_0 \neq 0$, takže $d_0^T A d_0 > 0$, čili ve vztazích pro α_0, β_0 je jmenovatel různý od nuly a krok 2 je proveditelný.

1. Z $\alpha_0 = -\frac{d_0^T g_0}{d_0^T A d_0}$ plyne $\alpha_0 > 0$ a $0 = d_0^T (g_0 + \alpha_0 A d_0) = d_0^T g_1 = g_1^T d_0$ podle (47), což je (48).
2. Vztah (49) plyne z toho, že $g_1^T (d_1 + g_1) = g_1^T \beta_0 d_0 = 0$ podle definice d_1 a (48).
3. Dále je $g_1^T g_0 = g_1^T (-d_0) = 0$ podle (48), což je (50).
4. $d_1^T A d_0 = (-g_1 + \beta_0 d_0)^T \frac{1}{\alpha_0}(g_1 - g_0) = \frac{1}{\alpha_0}(-g_1^T g_1 - \beta_0 d_0^T g_0) = \frac{1}{\alpha_0}(-g_1^T g_1 + \beta_0 g_0^T g_0) = 0$ podle definice d_1, (47), (48), (50) a definice β_0, což je (51).

Nechť tedy vzorce (43)–(46) platí pro $i - 1 \geq 0$, tj.

$$\begin{aligned} g_i^T d_{i-1} &= 0, & & (52)\\ g_i^T (d_i + g_i) &= 0, & & (53)\\ g_i^T g_j &= 0 \qquad (0 \leq j < i), & & (54)\\ d_i^T A d_j &= 0 \qquad (0 \leq j < i), & & (55) \end{aligned}$$

a nechť $g_i \neq 0$. Potom pro $d_i = 0$ by z $d_i = -g_i + \beta_{i-1} d_{i-1}$ (definice d_i) plynulo $g_i = \beta_{i-1} d_{i-1}$ a $g_i^T g_i = \beta_{i-1} g_i^T d_{i-1} = 0$ podle (52), tedy $g_i = 0$, spor. Proto $d_i \neq 0$, takže krok 2 je proveditelný.

1. Z $\alpha_i = -\frac{d_i^T g_i}{d_i^T A d_i}$ plyne opět $\alpha_i > 0$ (neboť $-d_i^T g_i = g_i^T g_i > 0$ podle (53) a $d_i^T A d_i > 0$ z pozitivní definitnosti) a přenásobením $0 = d_i^T g_i + \alpha_i d_i^T A d_i = d_i^T (g_i + \alpha_i A d_i) = d_i^T g_{i+1} = g_{i+1}^T d_i$ podle (47), což je (43).
2. Dále $g_{i+1}^T (d_{i+1} + g_{i+1}) = g_{i+1}^T \beta_i d_i = 0$ podle definice d_{i+1} a (43), což dává (44).
3. Dokážeme (45) rozborem tří případů:

a) Je-li $j = 0$, potom z (47) plyne $g_{i+1}^T g_0 = (g_i + \alpha_i A d_i)^T g_0 = g_i^T g_0 + \alpha_i d_i^T A g_0 = g_i^T g_0 - \alpha_i d_i^T A d_0 = 0$ podle (54), (55).

b) Je-li $0 < j < i$, je $g_{i+1}^T g_j = (g_i + \alpha_i A d_i)^T g_j = \alpha_i d_i^T A g_j = \alpha_i d_i^T A(-d_j + \beta_{j-1} d_{j-1}) = -\alpha_i d_i^T A d_j + \alpha_i \beta_{j-1} d_i^T A d_{j-1} = 0$ podle (47), (54), definice d_j a (55).

c) Nakonec pro $j = i$ je $g_{i+1}^T g_i = (g_i + \alpha_i A d_i)^T g_i = g_i^T g_i + \alpha_i d_i^T A(-d_i + \beta_{i-1} d_{i-1}) = -d_i^T g_i - \alpha_i d_i^T A d_i + \alpha_i \beta_{i-1} d_i^T A d_{i-1} = \alpha_i \beta_{i-1} d_i^T A d_{i-1} = 0$ podle (47), definice d_i, (53), definice α_i a (55), čímž jsme indukcí dokázali (45).

4. Dokážeme (46) rozborem dvou možných případů:

a) Pro $0 \le j < i$ je $d_{i+1}^T A d_j = (-g_{i+1} + \beta_i d_i)^T A d_j = -g_{i+1}^T A d_j = -g_{i+1}^T \frac{1}{\alpha_j}(g_{j+1} - g_j) = 0$ podle definice d_{i+1}, (55), (47) a (45).

b) Pro $j = i$ je $d_{i+1}^T A d_i = (-g_{i+1} + \beta_i d_i)^T \frac{1}{\alpha_i}(g_{i+1} - g_i) = \frac{1}{\alpha_i}(-g_{i+1}^T g_{i+1} + g_{i+1}^T g_i + \beta_i d_i^T g_{i+1} - \beta_i d_i^T g_i) = \frac{1}{\alpha_i}(-g_{i+1}^T g_{i+1} + \beta_i g_i^T g_i) = 0$ podle definice d_{i+1}, (47), (45), (43), (53) a definice β_i, což dává (46).

$\square$

Věta 52 *Nechť A je pozitivně definitní. Potom existuje $m \le n$ tak, že $g_m = 0$, tj. x_m je řešením $Ax = b$.*

Důkaz Jestliže $g_m = 0$ pro jisté $0 \le m \le n-1$, je tvrzení dokázáno. Předpokládejme tedy, že $g_0 \ne 0, \dots, g_{n-1} \ne 0$. Dokážeme, že potom $g_n = 0$. Vektory $g_0, \dots, g_{n-1}$ jsou nenulové ortogonální a tedy jsou lineárně nezávislé, takže tvoří bázi a g_n lze psát ve tvaru $g_n = \sum_{j=0}^{n-1} \gamma_j g_j$ a odtud podle (45) je $g_n^T g_n = \sum_{j=0}^{n-1} \gamma_j (g_n^T g_j) = 0$, tedy $g_n = 0$. $\square$

15 Vlastní čísla a vektory

Definice Nechť $A \in \mathbb{C}^{n \times n}$. Platí-li

$$Ax = \lambda x$$

pro jisté $\lambda \in \mathbb{C}$ a $0 \ne x \in \mathbb{C}^n$, potom λ se nazývá vlastním číslem matice A a vektor x vlastním vektorem příslušným k tomuto vlastnímu číslu.

Definice Matice $A \in \mathbb{C}^{n \times n}$ je podobná matici $B \in \mathbb{C}^{n \times n}$ jestliže platí $A = SBS^{-1}$ pro jistou regulární matici $S \in \mathbb{C}^{n \times n}$.

Věta 53 *Podobné matice mají stejná vlastní čísla.*

Důkaz Nechť $A = SBS^{-1}$ a nechť λ je vlastní číslo matice A, tj. $Ax = \lambda x$, $x \ne 0$. Potom z rovnosti $SBS^{-1}x = \lambda x$ dostáváme $B(S^{-1}x) = \lambda(S^{-1}x)$, kde $S^{-1}x \ne 0$ neboť S je regulární a $x \ne 0$, takže λ je vlastní číslo B. Dokázali jsme, že každé vlastní číslo A je vlastním číslem B. Použijeme-li vztahu $B = S^{-1}AS = S^{-1}A(S^{-1})^{-1}$, dostáváme aplikací předchozího výsledku, že každé vlastní číslo B je vlastním číslem A, takže obě matice mají stejná vlastní čísla. □

Definice Matice $A \in \mathbb{C}^{n\times n}$ se nazývá diagonalizovatelná jestliže je podobná diagonální matici.

Věta 54 *Matice $A \in \mathbb{C}^{n\times n}$ je diagonalizovatelná právě když má n lineárně nezávislých vlastních vektorů.*

Důkaz Je-li A diagonalizovatelná, potom $A = S\Lambda S^{-1}$ pro jistou regulární matici S a diagonální matici Λ. Podle věty 53 sestává diagonála matice Λ právě ze všech vlastních čísel matice A. Ze vztahu $AS = S\Lambda$ dostáváme potom

$$AS_{\bullet j} = \Lambda_{jj} S_{\bullet j},$$

takže $S_{\bullet j}$ je vlastní vektor příslušející vlastnímu číslu Λ_{jj} ($j = 1, \ldots, n$). Protože S je regulární, jsou její sloupce lineárně nezávislé, takže A má n lineárně nezávislých vlastních vektorů.

Naopak, nechť A má n lineárně nezávislých vlastních vektorů $x_1, \ldots, x_n$ příslušejících vlastním číslům $\lambda_1, \ldots, \lambda_n$. Položíme-li $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ a je-li S matice sestávající ze sloupců $x_1, \ldots, x_n$, potom S je regulární protože její sloupce jsou podle předpokladu lineárně nezávislé, a pro $j = 1, \ldots, n$ platí

$$(AS)_{\bullet j} = AS_{\bullet j} = Ax_j = \lambda_j x_j = (S\Lambda)_{\bullet j},$$

tedy $AS = S\Lambda$ a $A = S\Lambda S^{-1}$, takže A je diagonalizovatelná. □

Věta 55 *Jsou-li všechna vlastní čísla matice $A \in \mathbb{C}^{n\times n}$ navzájem různá, potom A je diagonalizovatelná.*

Důkaz Nechť A má n navzájem různých vlastních čísel $\lambda_1, \ldots, \lambda_n$ a nechť $x_1, \ldots, x_n$ jsou k nim příslušné vlastní vektory. Dokážeme, že $x_1, \ldots, x_n$ jsou lineárně nezávislé. Předpokládejme sporem, že jsou lineárně závislé. Potom mezi všemi jejich netriviálními lineárními kombinacemi, které se rovnají nule, existuje taková, která má *nejmenší* počet nenulových koeficientů. Nechť je to lineární kombinace

$$\alpha_{i_1} x_{i_1} + \ldots + \alpha_{i_m} x_{i_m} = 0, \tag{56}$$

kde jsme vzali pouze členy s nenulovými koeficienty, takže $\alpha_{i_j} \neq 0$ pro $j = 1, \ldots, m$ a $m \geq 2$, neboť všechny vektory x_j jsou nenulové. Přenásobením (56) maticí A dostáváme

$$A(\alpha_{i_1} x_{i_1} + \ldots + \alpha_{i_m} x_{i_m}) = \alpha_{i_1} \lambda_{i_1} x_{i_1} + \ldots + \alpha_{i_m} \lambda_{i_m} x_{i_m} = 0. \tag{57}$$

Vynásobíme-li nyní rovnici (56) číslem λ_{i_m} a odečteme ji od rovnice (57), dostáváme

$$\alpha_{i_1}(\lambda_{i_1} - \lambda_{i_m}) x_{i_1} + \ldots + \alpha_{i_{m-1}}(\lambda_{i_{m-1}} - \lambda_{i_m}) x_{i_{m-1}} = 0,$$

kde všechny koeficienty jsou podle předpokladu různosti vlastních čísel nenulové, čímž jsme dostali netriviální lineární kombinaci s $m-1$ nenulovými koeficienty, což je v rozporu s definicí m jakožto nejmenšího počtu nenulových koeficientů v netriviální lineární kombinaci. Tím jsme dokázali sporem, že vlastní vektory $x_1, \ldots, x_n$ jsou lineárně nezávislé a tedy podle věty 54 je matice A diagonalizovatelná. $\square$

Věta 56 *Nechť $A \in \mathbb{C}^{m\times n}$ a $B \in \mathbb{C}^{n\times m}$, $m \leq n$. Potom vlastními čísly matice BA jsou právě všechna vlastní čísla matice AB (počítaná v obou případech s jejich násobnostmi), plus dodatečných $n-m$ vlastních čísel matice BA rovných nule. Je-li $m = n$, potom AB a BA mají stejná vlastní čísla; je-li navíc jedna z matic A, B regulární, potom AB a BA jsou podobné.*

Důkaz Platí

$$\begin{pmatrix} AB & 0 \\ B & 0 \end{pmatrix}\begin{pmatrix} I & A \\ 0 & I \end{pmatrix} = \begin{pmatrix} AB & ABA \\ B & BA \end{pmatrix},$$

$$\begin{pmatrix} I & A \\ 0 & I \end{pmatrix}\begin{pmatrix} 0 & 0 \\ B & BA \end{pmatrix} = \begin{pmatrix} AB & ABA \\ B & BA \end{pmatrix},$$

kde všechny matice jsou z $\mathbb{C}^{(m+n)\times(m+n)}$. Protože matice

$$\begin{pmatrix} I & A \\ 0 & I \end{pmatrix}$$

je regulární, dostáváme

$$\begin{pmatrix} I & A \\ 0 & I \end{pmatrix}^{-1}\begin{pmatrix} AB & 0 \\ B & 0 \end{pmatrix}\begin{pmatrix} I & A \\ 0 & I \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ B & BA \end{pmatrix},$$

takže matice

$$\begin{pmatrix} AB & 0 \\ B & 0 \end{pmatrix} \quad \text{a} \quad \begin{pmatrix} 0 & 0 \\ B & BA \end{pmatrix}$$

jsou podobné a proto mají stejná vlastní čísla. Vlastní čísla první z nich jsou vlastní čísla AB plus n nul, vlastní čísla druhé z nich jsou vlastní čísla BA plus m nul. Z toho plyne první tvrzení věty. Druhé z něj plyne přímo, a v případě že např. A je regulární platí $AB = A(BA)A^{-1}$, takže AB a BA jsou podobné. $\square$

Věta 57 *Nechť pro matice $A, B \in \mathbb{C}^{n\times n}$ platí $AB = BA$. Potom ke každému vlastnímu číslu A existuje vlastní vektor, který je rovněž vlastním vektorem B (příslušejícím obecně k jinému vlastnímu číslu).*

Důkaz Nechť $Ax = \lambda x$, $\lambda \in \mathbb{C}$, $0 \neq x \in \mathbb{C}^n$. Uvažujme posloupnost vektorů $x, Bx, B^2x, \ldots$. Jelikož prostor $\mathbb{C}^n$ je n-rozměrný, musí existovat $k \leq n$ takové, že B^kx je lineární kombinací vektorů $x, Bx, B^2x, \ldots, B^{k-1}x$, tj.

$$B^kx = \sum_{j=0}^{k-1} \beta_j B^j x.$$

Potom platí

$$BX = XC, \tag{58}$$

kde X je matice o sloupcích $x, Bx, \dots, B^{k-1}x$ a

$$C = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 & \beta_0 \\ 1 & 0 & 0 & \dots & 0 & \beta_1 \\ 0 & 1 & 0 & \dots & 0 & \beta_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & \beta_{k-1} \end{pmatrix},$$

tj. $X \in \mathbb{R}^{n\times k}$ a $C \in \mathbb{C}^{k\times k}$. Nechť $\mu \in \mathbb{C}$ je libovolné vlastní číslo matice C a $y \in \mathbb{C}^n$ jemu odpovídající vlastní vektor. Z (58) potom plyne

$$BXy = XCy = X\mu y = \mu Xy,$$

takže $z = Xy$ je vlastní vektor B (sloupce X jsou podle konstrukce k lineárně nezávislé a $y \neq 0$, takže $Xy \neq 0$) a platí pro něj

$$z = \sum_{j=0}^{k-1} y_{j+1} B^j x.$$

Nyní, z $AB = BA$ plyne $AB^j = B^jA$ pro každé $j \geq 1$. Pro $j = 1$ je to předpoklad věty, a dále indukcí: z $AB^{j-1} = B^{j-1}A$ plyne $AB^j = AB^{j-1}B = B^{j-1}AB = B^{j-1}BA = B^jA$. Tedy

$$Az = \sum_{j=0}^{k-1} y_{j+1} AB^j x = \sum_{j=0}^{k-1} y_{j+1} B^j Ax = \sum_{j=0}^{k-1} y_{j+1} B^j \lambda x = \lambda \sum_{j=0}^{k-1} y_{j+1} B^j x = \lambda z,$$

takže z je vlastním vektorem jak A, tak B. □

16 Schurova triangularizační věta

Definice Matice $U \in \mathbb{C}^{n\times n}$ se nazývá unitární jestliže $U^*U = I$.

Věta 58 (Schurova triangularizační věta) *Ke každé matici $A \in \mathbb{C}^{n\times n}$ existuje unitární matice $U \in \mathbb{C}^{n\times n}$ a horní trojúhelníková matice $T \in \mathbb{C}^{n\times n}$ tak, že*

$$A = UTU^*,$$

přičemž diagonálu matice T tvoří právě všechna vlastní čísla matice A v libovolném předem daném pořadí. Je-li $A \in \mathbb{R}^{n\times n}$ a jsou-li všechna vlastní čísla matice A reálná, potom U lze zvolit reálnou ortogonální a T reálnou.

Důkaz Důkaz provedeme indukcí podle n. Pro $n = 1$ stačí položit $U = (1)$, $T = (A_{11})$. Nechť tedy tvrzení platí až do $n - 1 \geq 1$ a nechť $A \in \mathbb{C}^{n\times n}$. Nechť $\lambda_1, \ldots, \lambda_n$ jsou vlastní čísla matice A ve zvoleném pořadí a nechť x je vlastní vektor příslušný k λ_1, $\|x\|_2 = 1$. Doplňme x na unitární matici $U_1 = (x\ X)$, kde $X \in \mathbb{C}^{n\times(n-1)}$. Potom

$$\begin{aligned} U_1^* A U_1 &= \begin{pmatrix} x^* \\ X^* \end{pmatrix} A(x\ X) = \begin{pmatrix} x^*Ax & x^*AX \\ X^*Ax & X^*AX \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1 & x^*AX \\ \lambda_1 X^*x & X^*AX \end{pmatrix} = \begin{pmatrix} \lambda_1 & x^*AX \\ 0 & X^*AX \end{pmatrix}. \end{aligned}$$

Protože U_1 je unitární, je výsledná matice podobná A, takže má stejná vlastní čísla, proto matice X^*AX musí mít zbylá vlastní čísla $\lambda_2, \ldots, \lambda_n$. K této matici podle indukčního předpokladu existuje unitární matice $\tilde{U}$ taková, že $\tilde{U}^* X^* A X \tilde{U} = \tilde{T}$, kde $\tilde{T}$ je horní trojúhelníková matice s diagonálními prvky $\lambda_2, \ldots, \lambda_n$. Potom

$$\begin{aligned} \begin{pmatrix} 1 & 0^T \\ 0 & \tilde{U} \end{pmatrix}^* U_1^* A U_1 \begin{pmatrix} 1 & 0^T \\ 0 & \tilde{U} \end{pmatrix} &= \begin{pmatrix} 1 & 0^T \\ 0 & \tilde{U}^* \end{pmatrix} \begin{pmatrix} \lambda_1 & x^*AX \\ 0 & X^*AX \end{pmatrix} \begin{pmatrix} 1 & 0^T \\ 0 & \tilde{U} \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1 & x^*AX\tilde{U} \\ 0 & \tilde{U}^*X^*AX\tilde{U} \end{pmatrix} = \begin{pmatrix} \lambda_1 & x^*AX\tilde{U} \\ 0 & \tilde{T} \end{pmatrix}, \end{aligned}$$

takže položíme-li

$$U = U_1 \begin{pmatrix} 1 & 0^T \\ 0 & \tilde{U} \end{pmatrix},$$

je U, jakožto součin unitárních matic, unitární a platí

$$A = UTU^*,$$

kde

$$T = \begin{pmatrix} \lambda_1 & x^*AX\tilde{U} \\ 0 & \tilde{T} \end{pmatrix}$$

je horní trojúhelníková matice s diagonálními prvky $\lambda_1, \ldots, \lambda_n$. Je-li A reálná a má-li reálná vlastní čísla, potom matice U_1 i $\tilde{U}$ lze zvolit reálné ortogonální a tedy i výsledná matice U je reálná ortogonální a $T = U^T A U$ je reálná. □

Věta 59 (Schurova triangularizační věta, reálná forma) *Ke každé matici $A \in \mathbb{R}^{n\times n}$ existuje ortogonální matice $Q \in \mathbb{R}^{n\times n}$ taková, že*

$$A = QTQ^T,$$

kde $T \in \mathbb{R}^{n\times n}$ je blokově diagonální horní trojúhelníková matice, na jejíž diagonále stojí buď bloky 1×1 sestávající z reálných vlastních čísel A, nebo bloky 2×2 jejichž vlastní čísla jsou dvojice komplexně sdružených vlastních čísel A. Přitom Q je možno volit tak, že diagonální bloky mohou být seřazeny v libovolném předem daném pořadí.

Důkaz Důkaz provedeme indukcí podle n. Pro $n = 1$ má matice právě jedno reálné vlastní číslo A_{11} a stačí volit $Q = (1)$, $T = (A_{11})$. Nechť tedy tvrzení platí až do $n - 1 \geq 1$ a nechť $A \in \mathbb{R}^{n\times n}$. Jestliže první blok ve zvoleném pořadí je matice 1×1 sestávající z reálného vlastního čísla λ_1, postupujeme jako v důkazu věty 58. Jestliže první blok odpovídá dvojici komplexně sdružených vlastních čísel $\lambda_1 \pm \lambda_2 i$, $\lambda_1, \lambda_2 \in \mathbb{R}$, $\lambda_2 \neq 0$, zvolme vlastní vektor $x = x_1 + x_2 i \neq 0$ odpovídající vlastnímu číslu $\lambda = \lambda_1 + \lambda_2 i$ ($x_1, x_2 \in \mathbb{R}^n$), potom z $Ax = \lambda x$ plyne

$$\begin{aligned} Ax_1 &= \lambda_1 x_1 - \lambda_2 x_2, &(59)\\ Ax_2 &= \lambda_2 x_1 + \lambda_1 x_2. &(60)\end{aligned}$$

Přitom platí

$$\begin{aligned} x_1 &= \tfrac{1}{2}(x + \overline{x}), &(61)\\ x_2 &= \tfrac{1}{2i}(x - \overline{x}). &(62)\end{aligned}$$

Vektory x, $\overline{x}$ jsou lineárně nezávislé: kdyby byly lineárně závislé, existovalo by $\alpha \neq 0$ tak, že $\overline{x} = \alpha x$, z čehož by plynulo $\alpha\lambda x = \alpha A x = A(\alpha x) = A\overline{x} = \overline{\lambda}\overline{x} = \alpha\overline{\lambda}x$ a odtud $\lambda = \overline{\lambda}$ ve sporu s $\lambda_2 \neq 0$. Podle (61), (62) jsou tedy i x_1, x_2 lineárně nezávislé. Položme $X = (x_1\ x_2) \in \mathbb{R}^{n\times 2}$, potom podle věty 29 existuje matice $\hat{Q} \in \mathbb{R}^{n\times 2}$ s ortonormálními sloupci a regulární matice $R \in \mathbb{R}^{2\times 2}$ tak, že $X = \hat{Q}R$. Z (59), (60) plyne

$$AX = X \begin{pmatrix} \lambda_1 & \lambda_2 \\ -\lambda_2 & \lambda_1 \end{pmatrix},$$

tedy

$$A\hat{Q}R = \hat{Q}R \begin{pmatrix} \lambda_1 & \lambda_2 \\ -\lambda_2 & \lambda_1 \end{pmatrix}$$

a odtud

$$\hat{Q}^T A \hat{Q} = R \begin{pmatrix} \lambda_1 & \lambda_2 \\ -\lambda_2 & \lambda_1 \end{pmatrix} R^{-1}. \tag{63}$$

Doplníme-li nyní matici $\hat{Q} \in \mathbb{R}^{n\times 2}$ na ortogonální matici $\tilde{Q} = (\hat{Q}\ Q_1)$, potom

$$Q_1^T A \hat{Q} = Q_1^T \hat{Q} R \begin{pmatrix} \lambda_1 & \lambda_2 \\ -\lambda_2 & \lambda_1 \end{pmatrix} R^{-1} = 0$$

(neboť $Q_1^T \hat{Q} = 0$) a tedy

$$\tilde{Q}^T A \tilde{Q} = \begin{pmatrix} \hat{Q}^T \\ Q_1^T \end{pmatrix} A(\hat{Q}\ Q_1) = \begin{pmatrix} \hat{Q}^T A \hat{Q} & \hat{Q}^T A Q_1 \\ Q_1^T A \hat{Q} & Q_1^T A Q_1 \end{pmatrix} = \begin{pmatrix} \hat{Q}^T A \hat{Q} & \hat{Q}^T A Q_1 \\ 0 & Q_1^T A Q_1 \end{pmatrix},$$

přičemž levý horní blok $\hat{Q}^T A \hat{Q}$ je vzhledem k (63) podobný matici

$$\begin{pmatrix} \lambda_1 & \lambda_2 \\ -\lambda_2 & \lambda_1 \end{pmatrix},$$

jejíž vlastní čísla jsou $\lambda_1 \pm \lambda_2 i$. Na pravou dolní matici $Q_1^T A Q_1$ nyní aplikujeme indukční předpoklad tak jako v důkazu věty 58, čímž je tvrzení dokázáno. $\square$

Věta 60 *Nechť pro $A, B \in \mathbb{C}^{n \times n}$ platí $AB = BA$. Potom existuje unitární matice U a horní trojúhelníkové matice T_1, T_2 tak, že platí*

$$A = UT_1U^*, \tag{64}$$

$$B = UT_2U^*. \tag{65}$$

Jsou-li A, B reálné a jsou-li všechna jejich vlastní čísla reálná, potom U lze zvolit reálnou ortogonální a T_1, T_2 reálné.

Důkaz Důkaz provedeme indukcí podle n. Pro $n = 1$ je tvrzení zřejmé, neboť stačí položit $U = I_1$, $T_1 = A$, $T_2 = B$. Nechť tedy tvrzení platí až do $n - 1 \geq 1$ včetně a nechť $A, B \in \mathbb{C}^{n \times n}$, $AB = BA$. Podle věty 57 mají A a B společný vlastní vektor $x \in \mathbb{C}^n$, tj. platí $Ax = \lambda x$, $Bx = \mu x$, $\|x\|_2 = 1$. Použijeme nyní tento *společný* vlastní vektor tak, jak jsme to učinili v důkazu Schurovy triangularizační věty 58: doplňme x na unitární matici $U_1 = (x \ X)$, potom, jak je ukázáno v důkazu věty 58, platí

$$U_1^* A U_1 = \begin{pmatrix} \lambda & x^*AX \\ 0 & X^*AX \end{pmatrix},$$

$$U_1^* B U_1 = \begin{pmatrix} \mu & x^*BX \\ 0 & X^*BX \end{pmatrix},$$

přičemž $(X^*AX)(X^*BX) = X^*ABX = X^*BAX = (X^*BX)(X^*AX)$, takže matice X^*AX, X^*BX komutují a podle indukčního předpokladu existuje unitární matice $\tilde{U}$ taková, že

$$\tilde{U}^* X^* A X \tilde{U} = \tilde{T}_1,$$
$$\tilde{U}^* X^* B X \tilde{U} = \tilde{T}_2,$$

kde $\tilde{T}_1$, $\tilde{T}_2$ jsou horní trojúhelníkové matice. Potom dostáváme

$$\begin{pmatrix} 1 & 0^T \\ 0 & \tilde{U} \end{pmatrix}^* U_1^* A U_1 \begin{pmatrix} 1 & 0^T \\ 0 & \tilde{U} \end{pmatrix} = \begin{pmatrix} \lambda & x^*AX\tilde{U} \\ 0 & \tilde{T}_1 \end{pmatrix},$$

$$\begin{pmatrix} 1 & 0^T \\ 0 & \tilde{U} \end{pmatrix}^* U_1^* B U_1 \begin{pmatrix} 1 & 0^T \\ 0 & \tilde{U} \end{pmatrix} = \begin{pmatrix} \mu & x^*BX\tilde{U} \\ 0 & \tilde{T}_2 \end{pmatrix},$$

takže (64), (65) platí pro

$$U = U_1 \begin{pmatrix} 1 & 0^T \\ 0 & \tilde{U} \end{pmatrix}, \quad T_1 = \begin{pmatrix} \lambda & x^*AX\tilde{U} \\ 0 & \tilde{T}_1 \end{pmatrix}, \quad T_2 = \begin{pmatrix} \mu & x^*BX\tilde{U} \\ 0 & \tilde{T}_2 \end{pmatrix}.$$

Zbývající tvrzení pro reálné matice s reálnými vlastními čísly je zřejmé, protože celou konstrukci lze potom provést v reálných číslech. □

Věta 61 *Ke každé matici $A \in \mathbb{C}^{n \times n}$ a ke každému $\varepsilon > 0$ existuje diagonalizovatelná matice $A' \in \mathbb{C}^{n \times n}$ taková, že*

$$\|A - A'\|_F < \varepsilon.$$

Důkaz Podle Schurovy triangularizační věty 58 lze A rozložit ve tvaru $A = UTU^*$, kde U je unitární a T je horní trojúhelníková. Položme

$$\delta = \frac{\min(\varepsilon, K)}{2n^2},$$

kde $K = 1$ jsou-li všechny diagonální prvky matice T stejné, a

$$K = \min_{ij}\{|T_{ii} - T_{jj}|;\ T_{ii} \neq T_{jj}\}$$

jinak. Nechť Δ je diagonální matice s diagonálními prvky $\Delta_{ii} = i$ $(i = 1, \ldots, n)$. Ukážeme, že matice

$$A' = U(T + \delta\Delta)U^* \tag{66}$$

má požadovanou vlastnost. Především všechny diagonální prvky horní trojúhelníkové matice $T + \delta\Delta$ jsou navzájem různé. Kdyby totiž platilo $T_{kk} + \delta k = T_{\ell\ell} + \delta\ell$ pro jistá $k \neq \ell$, potom

$$0 \neq |T_{kk} - T_{\ell\ell}| = |k - \ell|\delta < n\delta < K = \min_{ij}\{|T_{ii} - T_{jj}|;\ T_{ii} \neq T_{jj}\},$$

což je spor. Proto A', která je podle (66) podobná matici $T + \delta\Delta$, má všechna vlastní čísla vzájemně různá a je tedy podle věty 55 diagonalizovatelná. Nakonec

$$\|A - A'\|_F = \|U\delta\Delta U^*\|_F = \delta\|\Delta\|_F = \delta\sqrt{1^2 + \ldots + n^2} \leq \delta n^2 < \varepsilon,$$

čímž je důkaz dokončen. □

17 Unitární diagonalizovatelnost

Definice Matice $A \in \mathbb{C}^{n\times n}$ se nazývá unitárně diagonalizovatelná jestliže existuje unitární matice $U \in \mathbb{C}^{n\times n}$ taková, že U^*AU je diagonální matice.

Definice Matice $A \in \mathbb{C}^{n\times n}$ se nazývá normální jestliže $A^*A = AA^*$.

Věta 62 *Normální horní trojúhelníková matice je diagonální.*

Důkaz Důkaz provedeme indukcí podle řádu n matice. Je-li $n = 1$, je tvrzení zřejmé. Nechť tedy tvrzení platí pro $n-1 \geq 1$ a nechť $T \in \mathbb{C}^{n\times n}$ je normální horní trojúhelníková matice. Potom T lze psát ve tvaru

$$T = \begin{pmatrix} \tau & t \\ 0 & \tilde{T} \end{pmatrix}, \tag{67}$$

kde $\tau \in \mathbb{C}$, $t \in \mathbb{C}^{1\times(n-1)}$ a $\tilde{T} \in \mathbb{C}^{(n-1)\times(n-1)}$ je horní trojúhelníková. Z (67) dostáváme

$$T^*T = \begin{pmatrix} \tau^* & 0^T \\ t^* & \tilde{T}^* \end{pmatrix} \begin{pmatrix} \tau & t \\ 0 & \tilde{T} \end{pmatrix} = \begin{pmatrix} |\tau|^2 & \tau^* t \\ \tau t^* & t^*t + \tilde{T}^*\tilde{T} \end{pmatrix},$$

$$TT^* = \begin{pmatrix} \tau & t \\ 0 & \tilde{T} \end{pmatrix} \begin{pmatrix} \tau^* & 0^T \\ t^* & \tilde{T}^* \end{pmatrix} = \begin{pmatrix} |\tau|^2 + \|t\|_2^2 & t\tilde{T}^* \\ \tilde{T}t^* & \tilde{T}\tilde{T}^* \end{pmatrix},$$

a vzhledem k tomu, že obě matice se rovnají, platí $|\tau|^2 = |\tau|^2 + \|t\|_2^2$, z čehož plyne $t = 0^T$, a rovněž $\tilde{T}^*\tilde{T} = \tilde{T}\tilde{T}^*$, tedy horní trojúhelníková matice $\tilde{T} \in \mathbb{C}^{(n-1)\times(n-1)}$ je normální a podle indukčího předpokladu je proto diagonální. Z toho plyne, že i matice T ve tvaru (67) je diagonální, čímž je tvrzení indukcí dokázáno. □

Věta 63 *Matice $A \in \mathbb{C}^{n\times n}$ je unitárně diagonalizovatelná právě když je normální.*

Důkaz a) Je-li A unitárně diagonalizovatelná, potom $A = U\Lambda U^*$ pro jistou unitární matici U a diagonální matici Λ. Protože diagonální matice Λ je evidentně normální, platí

$$A^*A = U\Lambda^*U^*U\Lambda U^* = U\Lambda^*\Lambda U^* = U\Lambda\Lambda^*U^* = (U\Lambda U^*)(U\Lambda^*U^*) = AA^*,$$

takže A je normální.

b) Naopak, nechť A je normální. Podle Schurovy triangularizační věty 58 má A rozklad

$$A = UTU^*, \tag{68}$$

kde U je unitární a T je horní trojúhelníková matice, takže $T = U^*AU$ a z normality A plyne

$$T^*T = U^*A^*AU = U^*AA^*U = TT^*,$$

proto T je normální a tedy podle věty 62 je diagonální. V rozkladu (68) je tedy U unitární a T diagonální, takže A je unitárně diagonalizovatelná. □

18 Hermitovské matice

Definice Matice $A \in \mathbb{C}^{n\times n}$ se nazývá hermitovská jestliže $A^* = A$.

Věta 64 *Hermitovská matice má všechna vlastní čísla reálná.*

Důkaz Nechť $\lambda \in \mathbb{C}$ je libovolné vlastní číslo hermitovské matice $A \in \mathbb{C}^{n\times n}$ a nechť $x \in \mathbb{C}^n$ je k němu příslušný vlastní vektor. Potom z rovnice $Ax = \lambda x$ přenásobením vektorem x^* dostáváme

$$x^*Ax = \lambda x^*x,$$

kde $(x^*Ax)^* = x^*A^*x = x^*Ax$, takže číslo x^*Ax je reálné stejně tak jako $x^*x = \|x\|_2^2 > 0$, z čehož plyne že λ je reálné. □

Značení Vlastní čísla hermitovské matice značíme $\lambda_1(A), \ldots, \lambda_n(A)$ a číslujeme je v pořadí $\lambda_1(A) \geq \ldots \geq \lambda_n(A)$.

Věta 65 (spektrální věta pro hermitovské matice) *Ke každé hermitovské matici $A \in \mathbb{C}^{n\times n}$ existuje unitární matice $U \in \mathbb{C}^{n\times n}$ taková, že platí*

$$A = U\Lambda U^*, \tag{69}$$

kde

$$\Lambda = \mathrm{diag}(\lambda_1(A), \ldots, \lambda_n(A)).$$

Přitom j-tý sloupec matice U je vlastní vektor příslušející k vlastnímu číslu $\lambda_j(A)$ $(j = 1, \ldots, n)$. Je-li $A \in \mathbb{R}^{n\times n}$, potom U lze volit reálnou ortogonální.

Důkaz Hermitovská matice A splňuje $A^*A = A^2 = AA^*$, je tedy normální a podle věty 63 má rozklad tvaru (69), kde U je unitární a Λ diagonální. Z důkazu téže věty plyne, že (69) je Schurův rozklad matice A, ve kterém podle Schurovy triangularizační věty 58 lze dosáhnout seřazení vlastních čísel matice A na diagonále Λ v libovolném předepsaném pořadí, tedy i v pořadí $\lambda_1(A), \ldots, \lambda_n(A)$, a v případě že A je reálná lze U volit reálnou ortogonální. Z (69) plyne $AU = U\Lambda$, tedy $AU_{\bullet j} = \lambda_j(A)U_{\bullet j}$, takže $U_{\bullet j}$ je vlastní vektor příslušející k $\lambda_j(A)$ $(j = 1, \ldots, n)$. $\square$

Věta 66 *Pro každou matici $A \in \mathbb{R}^{m\times n}$ je matice A^*A hermitovská a všechna její vlastní čísla jsou nezáporná.*

Důkaz A^*A je hermitovská neboť $(A^*A)^* = A^*A$. Je-li λ libovolné její vlastní číslo a x k němu příslušný vlastní vektor, potom z $A^*Ax = \lambda x$ plyne $x^*A^*Ax = \lambda x^*x$, kde $x^*A^*Ax = (Ax)^*(Ax) = \|Ax\|_2^2 \geq 0$ a $x^*x = \|x\|_2^2 > 0$, takže $\lambda \geq 0$. $\square$

19 SVD rozklad II

Věta 67 (SVD rozklad pro komplexní matice) *Nechť $A \in \mathbb{C}^{m\times n}$ a $q = \min\{m, n\}$. Potom existuje matice $\Sigma = (\sigma_{ij}) \in \mathbb{R}^{m\times n}$ splňující $\sigma_{ij} = 0$ pro $i \neq j$ a $\sigma_{11} \geq \sigma_{22} \geq \ldots \geq \sigma_{qq} \geq 0$ a unitární matice $U \in \mathbb{C}^{m\times m}$, $V \in \mathbb{C}^{n\times n}$ takové, že platí*

$$A = U\Sigma V^*.$$

Je-li $A \in \mathbb{R}^{m\times n}$, potom U, V lze volit reálné ortogonální.

Důkaz Matice A^*A je podle věty 66 hermitovská s nezápornými vlastními čísly a má podle věty 65 spektrální rozklad

$$A^*A = V\Lambda V^* \tag{70}$$

kde V je unitární a Λ je diagonální s diagonálními prvky $\lambda_1(A^*A) \geq \ldots \geq \lambda_r(A^*A) > 0 = \lambda_{r+1}(A^*A) = \ldots = \lambda_n(A^*A)$, kde jsme r označili index posledního kladného vlastního čísla. Je-li $r = 0$, potom podle (70) je $A^*A = 0$ a tedy $A = 0$ a stačí položit $\Sigma = 0$, $U = I$, $V = I$. Nechť tedy $r \geq 1$. Položme

$$S = \mathrm{diag}(\sqrt{\lambda_1(A^*A)}, \ldots, \sqrt{\lambda_r(A^*A)}),$$

potom $S \in \mathbb{R}^{r \times r}$ je regulární a

$$\Lambda = \begin{pmatrix} S^2 & 0 \\ 0 & 0 \end{pmatrix}.$$

Píšeme-li V ve tvaru $V = (V_1\ V_2)$, kde $V_1 \in \mathbb{C}^{n \times r}$ a $V_2 \in \mathbb{C}^{n \times (n-r)}$, potom z (70) plyne

$$V^*A^*AV = \begin{pmatrix} V_1^* \\ V_2^* \end{pmatrix} A^*A(V_1\ V_2) = \begin{pmatrix} V_1^*A^*AV_1 & V_1^*A^*AV_2 \\ V_2^*A^*AV_1 & V_2^*A^*AV_2 \end{pmatrix} = \Lambda = \begin{pmatrix} S^2 & 0 \\ 0 & 0 \end{pmatrix}.$$

Porovnáním bloků na místech $(2, 2)$ dostáváme

$$(AV_2)^*AV_2 = V_2^*A^*AV_2 = 0,$$

z čehož plyne

$$AV_2 = 0, \tag{71}$$

a porovnáním bloků $(1, 1)$ dostáváme

$$V_1^*A^*AV_1 = S^2$$

a odsud

$$(AV_1S^{-1})^*(AV_1S^{-1}) = S^{-1}V_1^*A^*AV_1S^{-1} = I. \tag{72}$$

Tedy matice

$$U_1 = AV_1S^{-1} \in \mathbb{C}^{m \times r}$$

má ortonormální sloupce. Doplníme-li ji na ortogonální matici $U = (U_1\ U_2) \in \mathbb{C}^{m \times m}$, potom

$$U_2^*AV_1 = U_2^*U_1S = 0 \tag{73}$$

a podle (71), (72), (73) platí

$$U^*AV = \begin{pmatrix} U_1^*AV_1 & U_1^*AV_2 \\ U_2^*AV_1 & U_2^*AV_2 \end{pmatrix} = \begin{pmatrix} S & 0 \\ 0 & 0 \end{pmatrix} = \Sigma,$$

tedy

$$A = U\Sigma V^*,$$

což je hledaný rozklad. Je-li A reálná, potom podle věty 65 lze matici V v rozkladu (70) volit reálnou ortogonální, potom i U_1 je reálná a lze ji doplnit na reálnou ortogonální matici U. $\square$

20 Vlastní čísla symetrických matic

Věta 68 (spektrální věta pro symetrické matice) *Ke každé symetrické matici $A \in \mathbb{R}^{n \times n}$ existuje ortogonální matice $Q \in \mathbb{R}^{n \times n}$ taková, že platí*

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

kde

$$\Lambda = \mathrm{diag}(\lambda_1(A), \ldots, \lambda_n(A)).$$

Přitom j-tý sloupec matice Q je vlastní vektor příslušející k vlastnímu číslu $\lambda_j(A)$ $(j = 1, \ldots, n)$.

Důkaz Symetrická matice je hermitovská, takže pro ni platí věta 65 v její reálné formě. □

Věta 69 (Courant-Fischer) *Pro každou symetrickou matici $A \in \mathbb{R}^{n\times n}$ platí*

$$\lambda_k(A) = \max_{\dim W = k} \min_{0 \neq x \in W} \frac{x^T A x}{x^T x} = \max_{\dim W = k} \min_{\substack{x\in W \\ \|x\|_2 = 1}} x^T A x \tag{74}$$

$(k = 1, \ldots, n)$, kde maximum se bere přes všechny podprostory $\mathbb{R}^n$ dimenze k.

Důkaz Podle věty 68 má A spektrální rozklad

$$A = Q\Lambda Q^T, \tag{75}$$

kde Q je ortogonální a $\Lambda = \mathrm{diag}(\lambda_1(A), \ldots, \lambda_n(A))$. Z (75) plyne

$$AQ_{\bullet j} = \lambda_j(A) Q_{\bullet j} \tag{76}$$

pro každé j. Uvažujme nyní libovolné $k \in \{1, \ldots, n\}$, a nechť $\hat{W}$ je podprostor generovaný sloupci $Q_{\bullet 1}, \ldots, Q_{\bullet k}$. Potom $\dim \hat{W} = k$ a každý vektor $0 \neq x \in \hat{W}$ lze psát ve tvaru $x = \sum_{j=1}^k \alpha_j Q_{\bullet j}$, takže s využitím ortogonality Q a vztahu (76) dostáváme

$$\begin{aligned} x^T A x &= (\sum_{j=1}^k \alpha_j Q_{\bullet j}) A (\sum_{j=1}^k \alpha_j Q_{\bullet j}) = (\sum_{j=1}^k \alpha_j Q_{\bullet j})(\sum_{j=1}^k \alpha_j \lambda_j(A) Q_{\bullet j}) \qquad (77) \\ &= \sum_{j=1}^k \alpha_j^2 \lambda_j(A) \geq (\min_{j=1,\ldots,k} \lambda_j(A)) \sum_{j=1}^k \alpha_j^2 = \lambda_k(A)(x^T x), \end{aligned}$$

což dává

$$\frac{x^T A x}{x^T x} \geq \lambda_k(A)$$

pro každé $0 \neq x \in \hat{W}$ a tedy

$$\min_{0 \neq x \in \hat{W}} \frac{x^T A x}{x^T x} \geq \lambda_k(A),$$

z čehož plyne

$$\max_{\dim W = k} \min_{0 \neq x \in W} \frac{x^T A x}{x^T x} \geq \lambda_k(A). \tag{78}$$

Pro důkaz opačné nerovnosti zvolme libovolný podprostor W prostoru $\mathbb{R}^n$ dimenze k. Protože podprostor $\tilde{W}$ generovaný sloupci $Q_{\bullet k}, \ldots, Q_{\bullet n}$ má dimenzi $n-k+1$, je součet dimenzí obou podprostorů roven $n+1>n$, z čehož plyne $\dim(W \cap \tilde{W}) \geq 1$ a existuje tedy vektor $0 \neq x \in W \cap \tilde{W}$ pro který, píšeme-li ho ve tvaru $x = \sum_{j=k}^{n} \alpha_j Q_{\bullet j}$, podobně jako v (77) dostáváme

$$\begin{aligned} x^T A x &= (\sum_{j=k}^{n} \alpha_j Q_{\bullet j}) A (\sum_{j=k}^{n} \alpha_j Q_{\bullet j}) = (\sum_{j=k}^{n} \alpha_j Q_{\bullet j})(\sum_{j=k}^{n} \alpha_j \lambda_j(A) Q_{\bullet j}) \\ &= \sum_{j=k}^{n} \alpha_j^2 \lambda_j(A) \leq (\max_{j=k,\ldots,n} \lambda_j(A)) \sum_{j=k}^{n} \alpha_j^2 = \lambda_k(A)(x^T x), \end{aligned}$$

tedy

$$\frac{x^T A x}{x^T x} \leq \lambda_k(A)$$

a proto

$$\min_{0 \neq x \in W} \frac{x^T A x}{x^T x} \leq \lambda_k(A).$$

Jelikož W byl libovolný podprostor dimenze k, dostáváme odsud

$$\max_{\dim W = k} \min_{0 \neq x \in W} \frac{x^T A x}{x^T x} \leq \lambda_k(A), \tag{79}$$

což je opačná nerovnost, takže z (78) a (79) plyne (74). Druhou rovnost v (74) dostaneme z toho, že pro každé $x \neq 0$ lze psát

$$\frac{x^T A x}{x^T x} = \left(\frac{x}{\|x\|_2}\right)^T A \left(\frac{x}{\|x\|_2}\right).$$

$\square$

Věta 70 *Ke každé pozitivně semidefinitní matici $A \in \mathbb{R}^{n \times n}$ a ke každému přirozenému číslu $k \geq 2$ existuje právě jedna pozitivě semidefinitní matice B taková, že*

$$B^k = A. \tag{80}$$

Důkaz Podle věty 68 má matice A spektrální rozklad

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

kde Q je ortogonální a $\Lambda = \operatorname{diag}(\lambda_1(A), \ldots, \lambda_n(A))$, kde $\lambda_1(A), \ldots, \lambda_n(A)$ jsou vlastní čísla matice A, která jsou vzhledem k její pozitivní semidefinitnosti nezáporná. Položme $\Lambda^{1/k} = \operatorname{diag}(\sqrt[k]{\lambda_1}, \ldots, \sqrt[k]{\lambda_n})$ a

$$B = Q \Lambda^{1/k} Q^T.$$

Potom B je pozitivně semidefinitní a platí $B^k = Q\Lambda^{1/k}Q^T \cdot \ldots \cdot Q\Lambda^{1/k}Q^T = Q(\Lambda^{1/k})^k Q^T = Q\Lambda Q^T = A$, takže B je řešením maticové rovnice (80).

Pro důkaz jednoznačnosti předpokládejme, že pozitivně semidefinitní matice C splňuje $C^k = A$. Sestrojme Lagrangeův interpolační polynom s uzly λ_i a hodnotami $\sqrt[k]{\lambda_i}$, pro který tedy platí $p(\lambda_i) = \sqrt[k]{\lambda_i}$ pro $i = 1, \ldots, n$, potom

$$p(A) = p(Q\Lambda Q^T) = Qp(\Lambda)Q^T = Q\mathrm{diag}(\sqrt[k]{\lambda_1}, \ldots, \sqrt[k]{\lambda_n})Q^T = B,$$

tedy $B = p(A) = p(C^k)$ a odtud $BC = p(C^k)C = Cp(C^k) = CB$. Tedy matice B, C komutují a obě mají nezáporná (tj. reálná) vlastní čísla, proto podle reálné části věty 60 existuje ortogonální matice $\tilde{Q}$ taková, že matice

$$T_1 = \tilde{Q}^T B\tilde{Q},$$

$$T_2 = \tilde{Q}^T C\tilde{Q}$$

jsou horní trojúhelníkové. Přitom ze symetrie B plyne $T_1^T = \tilde{Q}^T B^T\tilde{Q} = \tilde{Q}^T B\tilde{Q} = T_1$ a podobně $T_2^T = T_2$, takže obě matice T_1, T_2 jsou diagonální a platí

$$T_1^k = \tilde{Q}^T B^k\tilde{Q} = \tilde{Q}^T A\tilde{Q} = \tilde{Q}^T C^k\tilde{Q} = T_2^k, \tag{81}$$

přičemž diagonální prvky matic T_1, T_2 jsou tvořeny nezápornými vlastními čísly pozitivně semidefinitních matic B a C, takže (81) implikuje $T_1 = T_2$ a odtud

$$B = \tilde{Q}T_1\tilde{Q}^T = \tilde{Q}T_2\tilde{Q}^T = C.$$

Tím jsme dokázali, že pozitivně semidefinitní matice B s vlastností (80) je určena jednoznačně. $\square$

Reference

[1] J. W. Demmel, *Applied Numerical Linear Algebra*, SIAM, Philadelphia, 1997.

[2] G. H. Golub and C. F. van Loan, *Matrix Computations*, The Johns Hopkins University Press, Baltimore, 1996.

[3] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, 1985.

[4] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, 1991.

[5] M. Marcus and H. Minc, *A Survey of Matrix Theory and Matrix Inequalities*, Prindle, Weber & Schmidt, Boston, 1964.

[6] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra*, SIAM, Philadelphia, 2000.

[7] G. W. Stewart, *Matrix Algorithms, Volume I: Basic Decompositions*, SIAM, Philadelphia, 1998.

[8] G. W. Stewart and J. Sun, *Matrix Perturbation Theory*, Academic Press, San Diego, 1990.

[9] D. S. Watkins, *Fundamentals of Matrix Computations*, John Wiley & Sons, New York, 2002.