Maticový počet-přednáška

1 Opakování základů lineární algebry

V této přednášce budeme pracovat pouze s tělesy reálných čísel $\mathbb{R}$ a komplexních čísel $\mathbb{C}$. Tato tělesa značíme souhrnně symbolem $\mathbb{F}$. Prvky tělesa nazýváme *skaláry*.
Lineární (nebo vektorový) prostor nad tělesem $\mathbb{F}$ je množina V (jejíž prvky nazýváme vektory), na které je definováno sčítání vektorů a násobení vektoru prvkem tělesa $\mathbb{F}$, přičemž pro tyto operace platí všechny přirozené axiomy.

Příklad 1.1. *V této přednášce budeme především pracovat s následujícími typy lineárních prostorů:*

- $\mathbb{R}$ *nad* $\mathbb{R}$, $\mathbb{C}$ *nad* $\mathbb{C}$, $\mathbb{C}$ *nad* $\mathbb{R}$,
- $$\mathbb{F}^n = \{\begin{pmatrix}\alpha_1\\ \vdots\\ \alpha_n\end{pmatrix}; \alpha_1, \dots \alpha_n \in \mathbb{F}\}$$
  je lineární prostor sloupcových vektorů délky n *nad* $\mathbb{F}$,
- $$\mathbb{F}^{n\times m} = \begin{pmatrix}\alpha_{1,1} & \cdots & \alpha_{1,m}\\ \vdots & \ddots & \vdots\\ \alpha_{n,1} & \cdots & \alpha_{n,m}\end{pmatrix}$$
  je lineární prostor matic o n *řádcích a* m *sloupcích nad* $\mathbb{F}$ *(takovým maticím také říkáme matice* typu $n \times m$*.*

Značení 1.2. *Pro matici* $A \in \mathbb{F}^{n\times m}$ *a* $i \in \{1, \dots, n\}$, $j \in \{1, \dots, m\}$ *značíme* $[A]_{i,j}$ *prvek matice* A *na* i*-tém řádku a* j*-tém sloupci.*

Zápis $A = (\vec{a_1}, \dots, \vec{a_n})$ *znamená, že vektory* $\vec{a_1}, \dots, \vec{a_n} \in \mathbb{F}^n$ *tvoří sloupce matice* A*.*

Definice 1.3. *Pro matice* $A \in \mathbb{F}^{m\times n}, B \in \mathbb{F}^{n\times l}$*, definujeme jejich* maticový součin *jako matici* $A \cdot B \in \mathbb{F}^{m\times l}$*, která má na pozici* (i, j) *prvek* $\sum_{k=1}^{n} a_{i,k} \cdot b_{k,j}$*.*

Poznámka 1.4. *Maticové násobení je asociativní, tj.* $(A \cdot B) \cdot C = A \cdot (B \cdot C)$*, kdykoli mají tyto součiny smysl, ale není komutativní, tj. existují matice* A, B *takové, že* $A \cdot B \neq B \cdot A$*.*

Definice 1.5.

- Transponovaná matice *k matici* $A \in \mathbb{F}^{n\times m}$ *je matice* $A^T \in \mathbb{F}^{m\times n}$ *splňující* $[A^T]_{i,j} = [A]_{j,i}$ *pro* $i \in \{1, \dots m\}, j \in \{1, \dots n\}$*. Jinak řečeno,* A^T *vznikne z* A *prohozením řádků za sloupce.*
- Hermitovsky sdružená matice *k matici* $A \in \mathbb{F}^{n\times m}$ *je matice* $A^* \in \mathbb{F}^{m\times n}$ *splňující* $[A^*]_{i,j} = \overline{[A]_{j,i}}$ *(komplexně sdružené číslo k* $[A]_{j,i}$*) pro* $i \in \{1, \dots m\}, j \in \{1, \dots n\}$*. Pro reálnou matici* A *je tedy* $A^* = A^T$*.*
- *Matice* $A \in \mathbb{F}^{n\times n}$ *je*
  - symetrická*, pokud* $A = A^T$*, například* $A = \begin{pmatrix}3 & -2\\ -2 & 3\end{pmatrix}$,
  - antisymetrická*, pokud* $A = -A^T$*, například* $A = \begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}$,
  - hermitovská*, pokud* $A = A^*$*, například* $A = \begin{pmatrix}1 & i\\ -i & 1\end{pmatrix}$,

– antihermitovská, *pokud* $A = -A^*$, *například* $A = \begin{pmatrix} 0 & 1+i \\ -1+i & 0 \end{pmatrix}$.

- Nulová matice řádu $m \times n$ *je matice* $O_{m\times n} \in \mathbb{F}^{m\times n}$, *která má všude nuly. Jde o nulový vektor v lineárním prostoru* $\mathbb{F}^{m\times n}$.
- Jednotková matice řádu n *je matice* $I_n \in \mathbb{F}^{n\times n}$, *která má na diagonále* 1, *všude jinde* 0. *Pro tuto matici platí* $A \cdot I_n = A$, $I_n \cdot B = B$, *kdykoli jsou tyto součiny definovány.*
- Inverzní matice *k matici* $A \in \mathbb{F}^{n\times n}$ *je matice* $A^{-1} \in \mathbb{F}^{n\times n}$ *splňující* $A \cdot A^{-1} = A^{-1} \cdot A = I_n$. *A je* regulární, *pokud existuje* A^{-1}, *jinak se A nazývá* singulární.

Poznámka 1.6. *Vektory lze chápat jako matice:* $\mathbb{F}^n \simeq \mathbb{F}^{n\times 1}$, $\{\vec{x}^T : \vec{x} \in \mathbb{F}^n\} \simeq \mathbb{F}^{1\times n}$.

Tvrzení 1.7.
- *Pro matice* $A \in \mathbb{F}^{m\times n}, B \in \mathbb{F}^{n\times l}$ *platí* $(A \cdot B)^T = B^T \cdot A^T$ *a* $(A \cdot B)^* = B^* \cdot A^*$.
- *Pro matice* $C, D \in \mathbb{F}^{m\times n}$ *a* $\alpha \in \mathbb{F}$ *platí*

$$(\alpha C)^T = \alpha C^T, (C+D)^T = C^T + D^T, (\alpha C)^* = \overline{\alpha} C^*, (C+D)^* = C^* + D^*$$

- *Jsou-li matice* $E, F \in \mathbb{F}^{n\times n}$ *regulární, potom* $E \cdot F$ *je také regulární, a platí* $(E \cdot F)^{-1} = F^{-1} \cdot E^{-1}$.

Definice 1.8. Permutace *množiny* M *je libovolná bijekce* $\pi \colon M \to M$. *Pro* $n \in \mathbb{N}$ *značíme* S_n *množinu všech permutací množiny* $\{1, \ldots, n\}$. Znaménko permutace π *je definováno jako*

$$sgn\pi = (-1)^{(\text{počet překřížení ve strunovém diagramu permutace } \pi)}.$$

Pro matici $A \in \mathbb{F}^{n\times n}$ *definujeme její* determinant *jako*

$$\det(A) = |A| = \sum_{\pi \in S_n} sgn\pi \cdot a_{1,\pi(1)} \cdot a_{2,\pi(2)} \cdot \ldots \cdot a_{n,\pi(n)}.$$

Věta 1.9 (Vlastnosti determinantu)**.** *Pro matice* $A, B \in \mathbb{F}^{n\times n}$ *platí:*

- A *je regulární* $\Leftrightarrow \det(A) \neq 0$.
- $\det(A \cdot B) = \det(A) \cdot \det(B)$ *(neplatí pro součet!),*
- $\det(A^T) = \det(A)$,
- $\det(A^*) = \overline{\det(A)}$,
- *pokud matice* C *vznikne z matice* A *prohozením dvou řádků, potom* $\det(C) = -\det(A)$,
- *pokud matice* C *vznikne z matice* A *vynásobením nějakého řádku skalárem* α, *potom* $\det(C) = \alpha \cdot \det(A)$,
- *pokud matice* C *vznikne z matice* A *přičtením násobku jednoho řádku k jinému řádku, potom* $\det(C) = \det(A)$,
- *pokud je matice* A *horní trojúhelníková (tj. má pod diagonálou samé nuly), potom* $\det(A)$ *je roven součinu prvků na diagonále* A,
- *(rozvoj determinantu podle řádku) pro* $i \in \{1, \ldots, n\}$ *platí*

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} \cdot [A]_{i,j} \cdot A^{i,j},$$

kde $A^{i,j}$ *je matice, která vznikne z matice* A *vynecháním* i*-tého řádku a* j*-tého sloupce.*

Značení 1.10 (blokový zápis matice)**.** *Pro $p, q \in \mathbb{N}$, a pro $k \in \{1, \ldots, p\}, l \in \{1, \ldots, q\}$, $m_k, n_l \in \mathbb{N}$ a $A_{k,l} \in \mathbb{F}^{m_k \times n_l}$. Potom zápis*

$$A = \begin{pmatrix} A_{1,1} & \cdots & A_{1,q} \\ \vdots & \ddots & \vdots \\ A_{p,1} & \cdots & A_{p,q} \end{pmatrix}$$

znaméná, že matice $A \in \mathbb{F}^{\sum_{k=1}^{p} m_k \times \sum_{l=1}^{q} n_l}$ vznikla poskládáním matic $(A_{k,l})$.

Příklad 1.11. $A = \begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$, $C = \begin{pmatrix} 0 & 0 \\ 0 & 1 \\ 5 & 0 \end{pmatrix}$, $D = I_3$, *pak*

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} 0 & 2 & 0 & 1 & 0 \\ 2 & 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 5 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Příklad 1.12. *Násobení blokových matic funguje analogicky jako klasické maticové násobení, pokud má násobení příslušných bloků smysl: Označme $E = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$, kde A, B, C, D jsou jako v předcházejícím příkladu. Potom*

$$E^2 = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \cdot \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} A^2 + B \cdot C & A \cdot B + B \cdot D \\ C \cdot A + D \cdot C & C \cdot B + D^2 \end{pmatrix}.$$

Po spočtení příslušných výrazů dostaneme

$$E^2 = \begin{pmatrix} 4 & 1 & 0 & 1 & -2 \\ -5 & 4 & 0 & 2 & -1 \\ 0 & 0 & 1 & 0 & 0 \\ 2 & 1 & 0 & 1 & -1 \\ 5 & 10 & 0 & 5 & 1 \end{pmatrix}.$$

Příklad 1.13. *Determinant horní blokové matice je roven součinu determinantů matic na diagonále: Uvažujme matici*

$$A = \begin{pmatrix} -2 & 9 & 8 & -4 & e^3 & 5 \\ 0 & 4 & -1 & 3 & 1 & 10^5 \\ 0 & -5 & 3 & -3 & 1 & \sqrt{2} \\ 0 & 0 & 0 & 3 & 2 & 0 \\ 0 & 0 & 0 & 2 & -1 & 5 \\ 0 & 0 & 0 & 0 & 0 & -2 \end{pmatrix}.$$

Pak

$$\det(A) = (-2) \cdot \begin{vmatrix} 4 & -1 \\ -5 & 3 \end{vmatrix} \cdot \begin{vmatrix} 3 & 2 \\ 2 & -1 \end{vmatrix} \cdot (-2) = -196.$$

Definice 1.14. *Nechť V je lineární prostor nad $\mathbb{F}$, $\vec{x}, \vec{y_1}, \ldots, \vec{y_n} \in V$, $\alpha_1, \ldots, \alpha_n \in \mathbb{F}$. Vektor $\vec{x}$ je* lineární kombinací *vektorů $\vec{y_1}, \ldots, \vec{y_n}$ s koeficienty $\alpha_1, \ldots, \alpha_n$, pokud $\vec{x} = \sum_{i=1}^{n} \alpha_i \vec{y_i}$.*

Definice 1.15. *Nechť V je lineární prostor nad $\mathbb{F}$. Množina $M \subseteq V$ je* lineárně nezávislá *(LN), pokud pro každé $n \in \mathbb{N}$, vektory $\vec{v_1}, \ldots, \vec{v_n} \in M$ a skaláry $\alpha_1, \ldots, \alpha_n \in \mathbb{F}$ platí implikace*

$$\sum_{i=1}^{n} \alpha_i \vec{v_i} = \vec{o} \Rightarrow \alpha_1 = \ldots = \alpha_n = 0,$$

jinak je lineárně závislá (LZ).

Definice 1.16. *Nechť V je lineární prostor nad $\mathbb{F}$.*

- lineární obal *množiny $M \subseteq V$ je množina všech lineárních kombinací prvků množiny M, značí se $span(M)$,*
- *množina $M \subseteq V$ je* lineární podprostor *(nebo* vektorový podprostor*) V, pokud $M = span(M)$, ekvivalentně, pokud pro všechna $\vec{x}, \vec{y} \in M$ a $\alpha \in \mathbb{F}$ platí $\vec{x} + \vec{y} \in M$ a $\alpha\vec{x} \in M$,*
- *množina $M \subseteq V$* generuje *V, pokud $span(M) = V$,*
- *množina $M \subseteq V$ je báze V, pokud generuje V a je LN,*
- *dimenze V je počet prvků libovolné báze V, značí se* $\dim(V)$.

Poznámka 1.17.

- *Většinou budeme pracovat s uspořádanými bázemi, tj. záleží na pořadí prvků v bázi.*
- *$\mathbb{R}$ je lineární podprostor $\mathbb{C}$ nad $\mathbb{R}$, ale není to lineární podprostor $\mathbb{C}$ nad $\mathbb{C}$.*
- *Dimenze $\mathbb{C}$ nad $\mathbb{C}$ je 1 (báze je například množina $\{1\}$), dimenze $\mathbb{C}$ nad $\mathbb{R}$ je 2 (báze je například množina $\{1, i\}$).*

Definice 1.18. *Pro $n \in \mathbb{N}$ a $i \in \{1, \dots, n\}$ značíme $\vec{e_i}$ vektor z $\mathbb{F}^n$, který má na i-té pozici 1, všude jinde nuly. $\vec{e_i}$ se nazývá i-tý kanonický vektor a seznam $E_n = (\vec{e_1}, \dots, \vec{e_n})$ se nazývá* kanonická báze $\mathbb{F}^n$.

Věta 1.19. *Nechť V je lineární prostor nad $\mathbb{F}$, $M, N \subseteq V$.*

- *Pokud M je LN a $N \subseteq M$, potom N je LN.*
- *Pokud M generuje V a $N \supseteq M$, potom N generuje V.*
- *Pokud M je LN a $\dim(V) < \infty$, potom existuje množina $\mathcal{B} \supseteq M$, která je báze V (Steinitz).*

Definice 1.20. *Nechť V je lineární prostor nad $\mathbb{F}$ a $B = (\vec{b_1}, \dots, \vec{b_n})$ je uspořádaná báze V. Potom pro každé $\vec{v} \in V$ existují jednoznačně určené skaláry $\alpha_1, \dots, \alpha_n \in \mathbb{F}$ takové, že platí $\vec{v} = \sum_{i=1}^{n} \alpha_i \vec{b_i}$. Značíme*

$$[\vec{v}]_{\mathcal{B}} = coord_{\mathcal{B}}(\vec{v}) = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix},$$

a vektor $[\vec{v}]_B$ nazýváme souřadnice *vektoru $\vec{v}$ vůči bázi $\mathcal{B}$.*

Definice 1.21. *Nechť V, W jsou lineární prostory nad $\mathbb{F}$, $f \colon V \to W$ je zobrazení. Řekneme, že f je* lineární, *pokud pro každé $\vec{x}, \vec{y} \in V$ a $\alpha \in \mathbb{F}$ platí*

$$f(\vec{x} + \vec{y}) = f(\vec{x}) + f(\vec{y}) \text{ a } f(\alpha\vec{x}) = \alpha f(\vec{x}).$$

Lineární zobrazení $f \colon V \to W$ se nazývá:

- monomorfismus*, pokud je prosté,*
- epimorfismus*, pokud je na,*
- isomorfismus*, pokud je bijekce, tj. prosté a na,*

Množina všech lineárních zobrazení $V \to W$ s přirozenými operacemi tvoří lineární prostor, který značíme $\mathcal{L}(V, W)$.

Poznámka 1.22.

- *Isomorfní lineární prostory jsou z hlediska lineární algebry stejné.*
- *Pokud $\dim(V) = n$, je V je isomorfní $\mathbb{F}^n$ (pokud $\mathcal{B}$ je libovolná báze V, potom zobrazení $\vec{v} \mapsto [\vec{v}]_{\mathcal{B}}$ je isomorfisums V na $\mathbb{F}^n$).*

- *Pokud* $\dim(V) = n$, $\dim(W) = m$, *potom* $\mathcal{L}(V, W)$ *je isomorfní* $\mathbb{F}^{m\times n}$ *(tj. každá matice přirozeně generuje lineární zobrazení, a každé lineární zobrazení lze reprezentovat maticí).*

Definice 1.23. *Nechť* V, W *jsou lineární prostory nad* $\mathbb{F}$, $f\colon V \to W$ *je lineární zobrazení.*

- Jádro *(*null space*)* f *je množina* $\ker(f) = N(f) = \{\vec{x} \in V\colon f(x) = \vec{o}\}$, *jde o lineární podprostor* V,
- Obraz *(*range space*)*f *je množina* $Im(f) = rng(f) = R(f) = \{f(\vec{x})\colon \vec{x} \in V\}$, *jde o lineární podprostor* W,
- defekt f *je* $def(f) = \dim(\ker(f))$,
- hodnost *(*rank*)* f *je* $rank(f) = \dim(rng(f))$.
- *Platí: Pokud* $\dim(V) < \infty$, *potom platí* $\dim(V) = def(f) + rank(f)$.

Příklad 1.24. *Pro libovolnou matici* $A \in \mathbb{F}^{m\times n}$ *je zobrazení* $f_A\colon \mathbb{F}^n \to \mathbb{F}^m$,

$$f_A(\vec{x}) = A \cdot \vec{x}, \quad \vec{x} \in \mathbb{F}^n,$$

lineární. Matici A *většinou identifikujeme se zobrazením* f_A, *a pro matici* A *tedy používáme pojmy jako jádro, obraz, hodnost, defekt, isomorfismus (mono, epi).*

Poznámka 1.25. *Pro libolovnou matici* $A \in \mathbb{F}^{m\times n}$ *platí, že* $rank(A)$ *je roven maximálnímu počtu LN řádků matice* A *(a je také roven maximálnímu počtu LN sloupců matice* A*).*

Věta 1.26. *Pro matici* $A \in \mathbb{F}^{n\times n}$ *jsou následující tvrzení ekvivalentní:*

(i) A *je regulární,*

(ii) $\det(A) \neq 0$,

(iii) A *má LN řádky,*

(iv) A *má LN sloupce,*

(v) A *je isomorfismus,*

(vi) A *je monomorfismus,*

(vii) A *je epimorfismus.*

Definice 1.27. *Nechť* V, W *jsou lineární prostory nad* $\mathbb{F}$, $\mathcal{B} = (\vec{b}_1, \dots, \vec{b}_n)$ *je uspořádaná báze* V, $\mathcal{C} = (\vec{c}_1, \dots, \vec{c}_n)$ *je uspořádaná báze* W, $f\colon V \to W$ *je lineární zobrazení. Potom* matice lineárního zobrazení f *vůči bázím* $\mathcal{B}$ *a* $\mathcal{C}$ *je matice, kterou značíme* $[f]_{\mathcal{B}\mapsto C}$, *definována jako*

$$[f]_{\mathcal{B}\mapsto C} = \big([f(\vec{b}_1)]_C, \dots, [f(\vec{b}_n)]_C$$

Tato matice splňuje, že pro každý vektor $\vec{v} \in V$ *platí*

$$[f(\vec{v})]_{\mathcal{C}} = [f]_{\mathcal{B}\mapsto C} \cdot [\vec{v}]_{\mathcal{B}}.$$

V případě, že $V = W$ *a* $f = id$, *značíme* $T_{\mathcal{B}\mapsto\mathcal{C}} = [id]_{\mathcal{B}\mapsto\mathcal{C}}$, *a tuto matici nazýváme* maticí transformace souřadnic z báze $\mathcal{B}$ do báze $\mathcal{C}$.

Tvrzení 1.28.

- *Matice transformace souřadnic z báze* $\mathcal{B}$ *do báze* $\mathcal{C}$ *je vždy regulární, a platí* $(T_{\mathcal{B}\mapsto\mathcal{C}})^{-1} = T_{\mathcal{C}\mapsto\mathcal{B}}$.
- *Jsou-li* $\mathcal{B}_1, \mathcal{B}_2$ *báze* V, $\mathcal{C}_1, \mathcal{C}_2$ *báze* W, *a* $f\colon V \to W$ *je lineární zobrazení, potom*

$$[f]_{\mathcal{B}_2\mapsto\mathcal{C}_2} = T_{\mathcal{C}_1\mapsto\mathcal{C}_2} \cdot [f]_{\mathcal{B}_1\mapsto\mathcal{C}_1} \cdot T_{\mathcal{B}_2\mapsto\mathcal{B}_1}.$$

Definice 1.29. Soustava m lineárních rovnic *nad* $\mathbb{F}$ *o* n *neznámých je systém tvaru* $A \cdot \vec{x} = \vec{b}$*, kde* $A \in \mathbb{F}^{m\times n}$ *je* matice soustavy*,* $\vec{b} \in \mathbb{F}^m$ *je* vektor pravých stran*, a* $\vec{x} \in \mathbb{F}^n$ *je* vektor proměnných*. Matice* $(A|\vec{b})$ *se nazývá* rozšířená matice soustavy*. Je-li* $\vec{b} = \vec{o}$*, soustava se nazývá* homogenní*.*

Věta 1.30 (Frobenius)**.** *Pro* $A \in \mathbb{F}^{m\times n}$ *a* $\vec{b} \in \mathbb{F}^m$ *jsou následující tvrzení ekvivalentní:*

(i) Soustava $(A|\vec{b})$ *má řešení,*

(ii) $\vec{b} \in rng(A)$*,*

(iii) $rank(A) = rank(A|\vec{b})$*.*

Navíc, pokud $\vec{x}_p$ *je nějaké řešení soustavy* $(A|\vec{b})$*, potom*

$$\{\vec{x}_p + \vec{x}_h \colon \vec{x}_h \in \ker(A)\}$$

je množina všech řešení soustavy $(A|\vec{b})$*.*

Věta 1.31. *Pro* $A \in \mathbb{F}^{n\times n}$ *jsou následující tvrzení ekvivalentní:*

(i) Soustava $(A|\vec{b})$ *má právě jedno řešení pro každé* $\vec{b} \in \mathbb{F}^n$*,*

(ii) A *je regulární.*

Pokud A *je regulární, pak jediné řešení soustavy* $(A|\vec{b})$ *je vektor* $\vec{x} = A^{-1} \cdot \vec{b}$*.*

2 Normy a maticové normy

Normy používáme k měření velikostí vektorů a matic. Maticové normy jsou normy na prostorech matic, které se chovají dobře vzhledem k maticovému násobení.

Definice 2.1. *Nechť* V *je lineární prostor nad* $\mathbb{F}$*. Zobrazení* $\|\cdot\| \colon V \to [0,\infty)$ *je* norma *na* V*, pokud pro všechna* $\vec{x}, \vec{y} \in V$ *a* $\alpha \in \mathbb{F}$ *platí:*

(N1) $\|\alpha\vec{x}\| = |\alpha| \, \|\vec{x}\|$ *(absolutní homogenita),*

(N2) $\|\vec{x} + \vec{y}\| \le \|\vec{x}\| + \|\vec{y}\|$ *(trojúhelníková nerovnost),*

(N3) $\|x\| = 0 \Leftrightarrow x = \vec{o}$*.*

Příklad 2.2. *Na* $\mathbb{F}^n$ *máme tyto základní normy:*

- $\|\vec{x}\|_1 = \sum_{i=1}^n |x_i|$*, (*$\ell_1$ *norma)*
- $\|\vec{x}\|_\infty = \max_{i=1,\dots,n} |x_i|$*, (*$\ell_\infty$ *norma)*
- *pro* $p \in (1,\infty)$*,* $\|\vec{x}\|_p = \left(\sum_{i=1}^n |x_i|^p\right)^{\frac{1}{p}}$ *(*ℓ_p *norma)-nejdůležitější je případ* $p = 2$*, kdy jde o klasickou Eukleidovskou vzdálenost, která je generovaná standardním skalárním součinem (bude v další kapitole).*

Důkaz. Pro ℓ_1 normu: Vezměme $\vec{x}, \vec{y} \in \mathbb{F}^n$ a $\alpha \in \mathbb{F}$. Potom platí

$$(N1)\ \|\alpha\vec{x}\|_1 = \sum_{i=1}^n |\alpha x_i| = \sum_{i=1}^n |\alpha| \, |x_i| = |\alpha| \sum_{i=1}^n |x_i| = |\alpha| \, \|\vec{x}\|_1 ,$$

$$(N2)\ \|\vec{x} + \vec{y}\|_1 = \sum_{i=1}^n |x_i + y_i| \le \sum_{i=1}^n |x_i| + |y_i| = \sum_{i=1}^n |x_i| + \sum_{i=1}^n |y_i| = \|\vec{x}\|_1 + \|\vec{y}\|_1 ,$$

$$(N3)\ \|\vec{x}\|_1 = \sum_{i=1}^n |x_i| = 0 \Leftrightarrow x_i = 0 \text{ pro každé } i \in \{1,\dots,n\} \Leftrightarrow \vec{x} = \vec{o}.$$

Pro ℓ_∞ normu jde o snadné cvičení. Pro ℓ_p normu je trojúhelníková nerovnost těžká, nazývá se *Minkowského nerovnost.* □

Příklad 2.3. *Pokud* $\vec{x} = \begin{pmatrix} 3 \\ 4-3i \\ 1 \end{pmatrix} \in \mathbb{C}^3$, *potom* $\|\vec{x}\|_1 = 9, \|\vec{x}\|_2 = \sqrt{35}$, $\|\vec{x}\|_\infty = 5$.

Poznámka 2.4. *Pro každé* $\vec{x} \in \mathbb{F}^n$ *platí* $\|\vec{x}\|_\infty \leq \|\vec{x}\|_2 \leq \|\vec{x}\|_1$.

Definice 2.5. *Nechť* V *je lineární prostor nad* $\mathbb{F}$, $\|\cdot\|$ *je norma na* V. *Řekneme, že posloupnost* $\{\vec{v}_n\}_{n\in\mathbb{N}}$ *prvků z* V konverguje *k vektoru* $\vec{v} \in V$, *pokud*

$$\lim_{n\to\infty} \|\vec{v} - \vec{v_n}\| = 0.$$

Definice 2.6. *Nechť* V *je lineární prostor nad* $\mathbb{F}$. *Řekneme, že normy* $\|\cdot\|_a, \|\cdot\|_b$ *na* V *jsou* ekvivalentní, *pokud existují konstanty* $C_1, C_2 \in (0,\infty)$ *takové, že pro každé* $\vec{v} \in V$ *platí*

$$C_1 \|\vec{v}\|_a \leq \|\vec{v}\|_b \leq C_2 \|\vec{v}\|_a .$$

Značíme $\|\cdot\|_a \simeq \|\cdot\|_b$.

Cvičení 2.7. *(a) Pokud* $\|\cdot\|_a \simeq \|\cdot\|_b$, *potom* $\|\cdot\|_b \simeq \|\cdot\|_a$.

(b) Pokud $\|\cdot\|_a \simeq \|\cdot\|_b$ *a* $\|\cdot\|_b \simeq \|\cdot\|_c$, *potom* $\|\cdot\|_a \simeq \|\cdot\|_c$.

Tvrzení 2.8. *Pokud normy* $\|\cdot\|_a, \|\cdot\|_b$ *jsou ekvivalentní, potom konvergence v normě* $\|\cdot\|_a$ *je ekvivalentní konvergenci v normě* $\|\cdot\|_b$.

Důkaz. Nechť $\{\vec{v}_n\}_{n\in\mathbb{N}}$ je posloupnost prvků z V a $\vec{v} \in V$.

Pokud $\{\vec{v}_n\}_{n\in\mathbb{N}}$ konverguje k $\vec{v}$ v normě $\|\cdot\|_a$, potom pro každé $n \in \mathbb{N}$ platí $0 \leq \|\vec{v} - \vec{v_n}\|_b \leq C_2 \|\vec{v} - \vec{v_n}\|_a$. Tedy, jelikož $\lim_{n\to\infty} C_2 \|\vec{v} - \vec{v_n}\|_a = 0$, platí také $\lim_{n\to\infty} \|\vec{v} - \vec{v_n}\|_b = 0$, a $\{\vec{v}_n\}_{n\in\mathbb{N}}$ konverguje k $\vec{v}$ v normě $\|\cdot\|_b$.

Pokud naopak $\{\vec{v}_n\}_{n\in\mathbb{N}}$ konverguje k $\vec{v}$ v normě $\|\cdot\|_b$, potom máme $0 \leq \|\vec{v} - \vec{v_n}\|_a \leq \frac{1}{C_1} \|\vec{v} - \vec{v_n}\|_b$ pro každé $n \in \mathbb{N}$, tedy analogicky $\{\vec{v}_n\}_{n\in\mathbb{N}}$ konverguje k $\vec{v}$ v normě $\|\cdot\|_a$. □

Věta 2.9 (Bolzano-Weierstrass)**.** *Nechť* $\{\vec{y}_n\}_{n\in\mathbb{N}}$ *je posloupnost v* $\mathbb{F}^n$, *která je omezená v* l_1*-normě, tj.* $\sup_{n\in\mathbb{N}} \|y_n\|_1 < \infty$. *Potom* $\{\vec{y}_n\}_{n\in\mathbb{N}}$ *má konvergentní podposloupnost.*

Věta 2.10. *Na prostoru konečné dimenze* V *jsou každé dvě normy jsou ekvivalentní.*

Důkaz. Jelikož V je isomorfní $\mathbb{F}^n$ pro $n = \dim(V)$, stačí to dokázat pro $\mathbb{F}^n$.

Stačí ukázat, že libovolná norma $\|\cdot\|$ na $\mathbb{F}^n$ je ekvivalentní normě $\|\cdot\|_1$, a použít cvičení 2.7(b). Vezměme tedy libovolnou normu $\|\cdot\|$ na $\mathbb{F}^n$, a chceme nalézt $C_1, C_2 > 0$ splňující

$$C_1 \|\vec{x}\|_1 \leq \|\vec{x}\| \leq C_2 \|\vec{x}\|_1$$

pro každé $\vec{x} \in \mathbb{F}^n$.
Na jednu stranu, pro libovolné $x \in \mathbb{F}^n$ máme

$$\|\vec{x}\| = \left\| \sum_{i=1}^n x_i \vec{e_i} \right\| \leq^{(N2)} \sum_{i=1}^n \|x_i \vec{e_i}\| =^{(N1)} \sum_{i=1}^n |x_i| \cdot \|\vec{e_i}\| \leq \max_{i=1,\dots,n} \|\vec{e_i}\| \cdot \sum_{i=1}^n |x_i| = \max_{i=1,\dots,n} \|\vec{e_i}\| \cdot \|\vec{x}\|_1 .$$

Označme tedy $C_2 = \max_{i=1,\dots,n} \|\vec{e}_i\|$.
Zbývá dokázat, že existuje $C_1 > 0$ takové, že pro každé $\vec{x} \in \mathbb{F}^n$ platí $\|\vec{x}\|_1 \leq \frac{1}{C_1} \|\vec{x}\|$. Důkaz provedem sporem. Předpokládejme, že pro každé $k \in \mathbb{N}$ existuje $\vec{x}_k \in \mathbb{F}^n$ takové, že $\|\vec{x}_k\|_1 > k \|\vec{x}_k\|$. Definujme $\vec{y}_k = \frac{\vec{x}_k}{\|\vec{x}_k\|_1}$. Potom pro každé $k \in \mathbb{N}$ platí

$$\|\vec{y}_k\|_1 = \left\| \frac{\vec{x}_k}{\|\vec{x}_k\|_1} \right\|_1 =^{(N1)} \frac{\|\vec{x}_k\|_1}{\|\vec{x}_k\|_1} = 1, \text{ a } \|\vec{y}_k\| = \left\| \frac{\vec{x}_k}{\|\vec{x}_k\|_1} \right\| =^{(N1)} \frac{\|\vec{x}_k\|}{\|\vec{x}_k\|_1} < \frac{1}{k}.$$

Nyní, dle Věty 2.9, existuje $\vec{y} \in \mathbb{F}^n$ a podposloupnost $\{\vec{y}_{k_j}\}_{j=1}^{\infty}$ takové, že $\lim_{j\to\infty} \left\|\vec{y} - \vec{y}_{k_j}\right\| = 0$.
Dále, jelikož pro každé $j \in \mathbb{N}$ platí

$$1 = \left\|\vec{y}_{k_j}\right\|_1 = \left\|(\vec{y}_{k_j} - \vec{y}) + \vec{y}\right\|_1 \leq^{(N2)} \left\|(\vec{y}_{k_j} - \vec{y})\right\|_1 + \|\vec{y}\|_1 \to^{j\to\infty} 0 + \|\vec{y}\|_1 ,$$

platí $\|\vec{y}\|_1 \geq 1$, tedy $\vec{y} \neq \vec{o}$. Na druhou stranu, pro každé $j \in \mathbb{N}$ máme

$$0 \leq \|\vec{y}\| = \left\|(\vec{y} - \vec{y}_{k_j}) + \vec{y}_{k_j}\right\| \leq \left\|\vec{y} - \vec{y}_{k_j}\right\| + \left\|\vec{y}_{k_j}\right\| < C_2 \cdot \left\|\vec{y} - \vec{y}_{k_j}\right\|_1 + \frac{1}{k_j} \to 0.$$

Tedy $\|\vec{y}\| = 0 \Rightarrow^{(N3)} \vec{y} = \vec{o}$, což je spor. □

Definice 2.11. Maticová norma *na* $\mathbb{F}^{n\times n}$ *je zobrazení* $\|\cdot\| : \mathbb{F}^{n\times n} \to [0,\infty)$ *splňující pro všechna* $A, B \in \mathbb{F}^{n\times n}$ *a* $\alpha \in \mathbb{F}$*:*

(M1) $\|\alpha A\| = |\alpha| \, \|A\|$,

(M2) $\|A + B\| \leq \|A\| + \|B\|$,

(M3) $\|A\| = 0 \Leftrightarrow A = O_{n\times n}$,

(M4) $\|A \cdot B\| \leq \|A\| \cdot \|B\|$.

Příklad 2.12. *Zobrazení* $\|A\|_{\max} = \max_{i,j=1,\ldots,n} |[A]_{i,j}|$ *je norma na* $\mathbb{F}^{n\times n}$*, ale ne maticová norma.*

Důkaz. Je to norma? Cvičení.
Není to maticová norma: $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$, pak $A^2 = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}$, tedy $\left\|A^2\right\|_{\max} = 2$, ale $\|A\|_{\max} \cdot \|A\|_{\max} = 1$, a tedy podmínka (M4) není splněna. □

Příklad 2.13. *Pro* $A \in \mathbb{F}^{n\times n}$ *definujme*

- $$\|A\|_F = \Big(\sum_{i=1}^{n}\sum_{j=1}^{n} |[A]_{i,j}|^2\Big)^{\frac{1}{2}}.$$

  $\|\cdot\|_F$ *je maticová norma na* $\mathbb{F}^{n\times n}$*, nazývá se* Frobeniova norma.

- $$\|A\|_2 = \max_{\vec{x}\in\mathbb{F}^n, \vec{x}\neq\vec{o}} \frac{\|A\cdot\vec{x}\|_2}{\|\vec{x}\|_2}.$$

  $\|\cdot\|_2$ *je maticová norma na* $\mathbb{F}^{n\times n}$*, nazývá se* spektrální norma.

Poznámka 2.14. *Dle definice platí* $\|A\vec{x}\|_2 \leq \|A\|_2 \cdot \|\vec{x}\|_2$ *pro* $A \in \mathbb{F}^{n\times n}$ *a* $\vec{x} \in \mathbb{F}^n$.

Definice 2.15. *Je-li* $\|\cdot\|$ *maticová norma na* $\mathbb{F}^{n\times n}$*, potom* číslo podmíněnosti *matice* A *(vzhledem k normě* $\|\cdot\|$*), je* $\kappa(A) = \|A\| \cdot \left\|A^{-1}\right\|$*, pokud* A *je regulární, jinak je to* ∞*. Pro normu* $\|\cdot\|_F$ *značíme číslo podmíněnosti symbolem* $\kappa_F(A)$*, pro normu* $\|\cdot\|_2$ *symbolem* $\kappa_2(A)$.

Značení 2.16. *Pro matici* $A \in \mathbb{F}^{m\times n}$*,* $i \in \{1,\ldots,m\}$ *a* $j \in \{1,\ldots,n\}$ *značíme* A_{i*} *i-tý řádek matice* A*, a* A_{*j} *j-tý sloupec matice* A.

Definice 2.17. Stopa (trace) *matice* $A \in \mathbb{F}^{n\times n}$ *je součet prvků na hlavní diagonále* A*, značíme* $tr(A)$.

Tvrzení 2.18. *Pro Frobeniovu maticovou normu platí následující tvrzení:*

(i) Pro každou matici $A \in \mathbb{F}^{n\times n}$*,*

$$\|A\|_F = \sqrt{\sum_{i=1}^{n} \|A_{i*}\|_2^2} = \sqrt{\sum_{j=1}^{n} \|A_{*j}\|_2^2} = \sqrt{tr(A^* \cdot A)} = \sqrt{tr(A \cdot A^*)}.$$

(ii) Pro každou matici $A \in \mathbb{F}^{n\times n}$ a vektor $\vec{x} \in \mathbb{F}^n$,

$$\|A\vec{x}\|_2 \le \|A\|_F \cdot \|\vec{x}\|_2 ,$$

tedy, Frobeniova norma $\|\cdot\|_F$ na $\mathbb{F}^{n\times n}$ a norma $\|\cdot\|_2$ na $\mathbb{F}^n$ jsou takzvaně kompatibilní.

Důkaz. V tvrzení (i) není moc co dokazovat, rozmyslete si jako cvičení.
Pro důkaz (ii), použijeme Cauchy-Schwartzovu nerovnost pro standardní skalární součin na $\mathbb{F}^n$ (která říká, že pro každé $\vec{u}, \vec{v} \in \mathbb{F}^n$ platí $|\langle \vec{u}, \vec{v}\rangle| \le \|\vec{u}\|_2 \cdot \|\vec{v}_2\|$), a dostaneme

$$\|A\vec{x}\|_2 = \sqrt{\sum_{i=1}^{n} |A_{i*} \cdot \vec{x}|^2} \le^{C.S.} \sqrt{\sum_{i=1}^{n} \|A_{i*}\|_2^2 \cdot \|\vec{x}\|_2^2} = \sqrt{\sum_{i=1}^{n} \|A_{i*}\|_2^2} \cdot \sqrt{\|\vec{x}\|_2^2} = \|A\|_F \cdot \|\vec{x}\|_2 .$$

□

Příklad 2.19. *Číslo podmíněnosti se používá například pro odhad stability soustavy lineárních rovnic $(A|\vec{b})$ vzhledem k změně vektoru $\vec{b}$ (chyba měření).*
Předpokládejme, že řešíme soustavu $(A|\vec{b})$ s regulární čtvercovou maticí A. Jak víme, tato soustava má jediné řešení, a to $\vec{x} = A^{-1}\cdot\vec{b}$. Přepodkládejme, že naše znalost vektoru pravých stran $\vec{b}$ je zatížena chybou $\Delta\vec{b} \in \mathbb{F}^n$. Řešení soustavy $(A|\vec{b} + \Delta\vec{b})$ můžeme napsat ve tvaru $\vec{x} + \Delta\vec{x} = A^{-1} \cdot (\vec{b} + \Delta\vec{b})$, chyba ve výsledku bude tedy rovna $\Delta\vec{x} = A^{-1} \cdot \Delta\vec{b}$.
Absolutní velikost chyby, měřená v ℓ_2 normě, je

$$\|\Delta\vec{x}\|_2 = \left\|A^{-1} \cdot \Delta\vec{b}\right\|_2 \le \left\|A^{-1}\right\|_2 \cdot \left\|\Delta\vec{b}\right\|_2 .$$

Nyní nás zajímá velikost relativní chyby $\frac{\|\Delta\vec{x}\|_2}{\|\vec{x}\|_2}$ v závislosti na relativní chybě měření $\frac{\|\Delta\vec{b}\|_2}{\|\vec{b}\|_2}$. Máme

$$\left\|\vec{b}\right\|_2 = \|A \cdot \vec{x}\|_2 \le \|A\|_2 \cdot \|\vec{x}\|_2 ,$$

a tedy dostaneme odhad

$$\frac{\|\Delta\vec{x}\|_2}{\|\vec{x}\|_2} \le \left\|A^{-1}\right\|_2 \cdot \left\|\Delta\vec{b}\right\|_2 \cdot \frac{1}{\|\vec{x}\|_2} \le \left\|A^{-1}\right\|_2 \cdot \left\|\Delta\vec{b}\right\|_2 \cdot \frac{\|A\|_2}{\left\|\vec{b}\right\|_2} = \kappa_2(A) \cdot \frac{\left\|\Delta\vec{b}\right\|_2}{\left\|\vec{b}\right\|_2}.$$

Pro toto odvození jsme použili pouze kompatibilitu maticové spektrální normy s vektorovou 2-normou. S použitím (ii) z Tvrzení 2.18 tedy zcela analogicky dostaneme odhad

$$\frac{\|\Delta\vec{x}\|_2}{\|\vec{x}\|_2} \le \kappa_F(A) \cdot \frac{\left\|\Delta\vec{b}\right\|_2}{\left\|\vec{b}\right\|_2}.$$

3 Skalární součiny

Skalární součiny používáme k měření úhlů mezi vektory.

Definice 3.1. *Nechť V je lineární prostor nad $\mathbb{F}$.* Skalární součin *na V je zobrazení $\langle\cdot,\cdot\rangle\colon V \times V \to \mathbb{F}$ splňující pro každé $\vec{x}, \vec{y} \in V$ a $\alpha \in \mathbb{F}$:*

(S1) $\langle\vec{x}, \vec{y}\rangle = \overline{\langle\vec{y}, \vec{x}\rangle}$ (sdružená symetrie),

(S2) $\langle\vec{x}, \cdot\rangle : \vec{y} \mapsto \langle\vec{x}, \vec{y}\rangle$ je lineární zobrazení V do $\mathbb{F}$ (linearita v druhé souřadnici),

(S3) $\langle \vec{x}, \vec{x} \rangle > 0$ *pro* $\vec{x} \neq \vec{o}$ *(pozitivní definitnost).*

Poznámka 3.2. *Pokud* $\langle \cdot, \cdot \rangle$ *je skalární součin na* V*, potom zobrazení* $\langle \cdot, \vec{y} \rangle : \vec{x} \mapsto \langle \vec{x}, \vec{y} \rangle$ *(v případě, kdy* $\mathbb{F} = \mathbb{C}$*) není lineární, ale je takzvaně* sdruženě lineární*, tj. pro každé* $\vec{x}_1, \vec{x}_2, \vec{y} \in V$ *a* $\alpha \in \mathbb{F}$ *máme*

$$\langle \vec{x}_1 + \vec{x}_2, \vec{y} \rangle =^{(S1)} \overline{\langle \vec{y}, \vec{x}_1 + \vec{x}_2 \rangle} =^{(S2)} \overline{\langle \vec{y}, \vec{x}_1 \rangle} + \overline{\langle \vec{y}, \vec{x}_2 \rangle} =^{(S1)} \langle \vec{x}_1, \vec{y} \rangle + \langle \vec{x}_2, \vec{y} \rangle,$$

a

$$\langle \alpha \vec{x}_1, \vec{y} \rangle =^{(S1)} = \overline{\langle \vec{y}, \alpha \vec{x}_1 \rangle} =^{(S2)} \overline{\alpha \cdot \langle \vec{y}, \vec{x}_1 \rangle} = \overline{\alpha} \cdot \overline{\langle \vec{y}, \vec{x}_1 \rangle} =^{(S1)} \overline{\alpha} \cdot \langle \vec{x}_1, \vec{y} \rangle.$$

Definice 3.3.

- *Zobrazení* $\langle \cdot, \cdot \rangle \colon \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ *definované předpisem*

$$\langle \vec{x}, \vec{y} \rangle \colon = \vec{x}^* \cdot \vec{y}$$

  se nazývá standardní skalární součin *na* $\mathbb{C}^n$*.*

- *Zobrazení* $\langle \cdot, \cdot \rangle \colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ *definované předpisem*

$$\langle \vec{x}, \vec{y} \rangle \colon = \vec{x}^T \cdot \vec{y}$$

  se nazývá standardní skalární součin *na* $\mathbb{R}^n$*.*

Definice 3.4. *Matice* $A \in \mathbb{F}^{n \times n}$ *se nazývá* pozitivně definitní *(PD), pokud je Hermitovská, tj.* $A = A^*$*, a pro každý nenulový vektor* $\vec{x} \in \mathbb{F}^n$ *platí* $\vec{x}^* \cdot A \cdot \vec{x} > 0$ *(několik dalších ekvivaletních definic pozitivně definitních matic uvidíme později).*

Věta 3.5. *Je-li* $A \in \mathbb{F}^{n \times n}$ *pozitivně definitní matice, potom zobrazení*

$$\langle \vec{x}, \vec{y} \rangle_A = \vec{x}^* \cdot A \cdot \vec{y}, \quad \vec{x}, \vec{y} \in \mathbb{F}^n,$$

je skalární součin na $\mathbb{F}^n$*.* A *se nazývá* Gramova matice *nebo* metrický tensor *skalárního součinu* $\langle \cdot, \cdot \rangle_A$*. Naopak, pro každý skalární součin* $\langle \cdot, \cdot \rangle$ *na* $\mathbb{F}^n$ *existuje právě jedna pozitivně definitiní matice* A *taková, že* $\langle \cdot, \cdot \rangle = \langle \cdot, \cdot \rangle_A$*.*

Důkaz. Nechť $A \in \mathbb{F}^{n \times n}$ je pozitivně definitní matice. Ověříme, že $\langle \cdot, \cdot \rangle_A$ je skalární součin. Pro $\vec{x}, \vec{y} \in \mathbb{F}^n$ máme

$$\overline{\langle \vec{y}, \vec{x} \rangle_A} = \overline{\vec{y}^* \cdot A \cdot \vec{x}} = \left(\vec{y}^* \cdot A \cdot \vec{x} \right)^* =^{1.7} \vec{x}^* \cdot A^* \cdot \vec{y} = \vec{x}^* \cdot A \cdot \vec{y} = \langle \vec{x}, \vec{y} \rangle_A,$$

a tedy máme ověřenou vlastnost (S1). Vlastnost (S2) plyne ihned z vlastností maticového násobení. Konečně, pro nenulové $\vec{x} \in \mathbb{F}^n$ máme

$$\langle \vec{x}, \vec{x} \rangle_A = \vec{x}^* \cdot A^* \cdot \vec{x} > 0$$

dle definice pozitivně definitní matice. Tedy $\langle \cdot, \cdot \rangle_A$ je skalární součin na $\mathbb{F}^n$.

Nechť nyní $\langle \cdot, \cdot \rangle$ je libovolný skalární součin na $\mathbb{F}^n$. Definujeme matici $A \in \mathbb{F}^{n \times n}$ předpisem

$$[A]_{i,j} = \langle \vec{e_i}, \vec{e_j} \rangle, \quad i, j \in \{1, \ldots, n\}.$$

Nejprve ověříme, že $\langle \cdot, \cdot \rangle = \langle \cdot, \cdot \rangle_A$: pro libovolné $\vec{x}, \vec{y} \in \mathbb{F}^n$ máme

$$\langle \vec{x}, \vec{y} \rangle =^{(S2)} \langle \sum_{i=1}^{n} x_i \cdot \vec{e_i}, \sum_{j=1}^{n} y_j \cdot \vec{e_j} \rangle =^{(S1),3.2} \sum_{i=1}^{n} \sum_{j=1}^{n} \overline{x_i} \cdot y_j \cdot \langle \vec{e_i}, \vec{e_j} \rangle = \vec{x}^* \cdot A \cdot \vec{y}.$$

Zbývá ověřit, že A je pozitivně definitní. Nejprve ukažme, že je hermitovská: pro $i, j \in \{1, \ldots, n\}$ máme

$$\overline{[A_{j,i}]} = \overline{\langle \vec{e_j}, \vec{e_i} \rangle} =^{(S1)} \langle \vec{e_i}, \vec{e_j} \rangle = [A_{i,j}].$$

Nakonec, pro nenulový vektor $\vec{x} \in \mathbb{F}^n$ máme

$$\vec{x}^* \cdot A \cdot \vec{x} = \langle \vec{x}, \vec{x} \rangle_A = \langle \vec{x}, \vec{x} \rangle >^{(S3)} 0,$$

čímž je důkaz ukončen. □

Věta 3.6. *Nechť V je lineární prostor nad $\mathbb{F}$ se skalárním součinem $\langle\cdot,\cdot\rangle$. Potom zobrazení*

$$\|\vec{x}\| = \sqrt{\langle\vec{x},\vec{x}\rangle}, \quad \vec{x} \in V,$$

definuje normu na V, která se nazývá norma indukovaná (nebo generovaná) skalárním součinem $\langle\cdot,\cdot\rangle$. *Navíc, pro všechna $\vec{x},\vec{y} \in V$ platí tzv.* Cauchy-Schwartzova nerovnost $|\langle\vec{x},\vec{y}\rangle| \leq \|\vec{x}\| \cdot \|\vec{y}\|$.

Příklad 3.7. *Standardní skalární součin na $\mathbb{F}^n$ generuje ℓ_2 normu (odsud pochází důvod, proč se ve skalárním součinu na $\mathbb{C}^n$ bere komplexní sdružení):*
Pro $\vec{x} \in \mathbb{F}^n$ máme

$$\sqrt{\langle\vec{x},\vec{x}\rangle} = \sqrt{\vec{x}^* \cdot \vec{x}} = \sqrt{\sum_{i=1}^{n} \overline{x_i} \cdot x_i} = \sqrt{\sum_{i=1}^{n} |x_i|^2} = \|\vec{x}\|_2 .$$

Definice 3.8. *Nechť V je lineární prostor nad $\mathbb{F}$, a $\|\cdot\|$ je norma na V indukovaná skalárním součinem $\langle\cdot,\cdot\rangle$.*

- *Pro nenulové vektory $\vec{x},\vec{y} \in V$ definujeme jejich* úhel $\phi \in [0, 2\pi)$ *jako*

$$\phi = \arccos \frac{\langle\vec{x},\vec{y}\rangle}{\|\vec{x}\| \cdot \|\vec{y}\|}.$$

- *Vektory $\vec{x},\vec{y} \in V$ jsou* kolmé, *pokud $\langle\vec{x},\vec{y}\rangle = 0$, značíme $\vec{x} \perp \vec{y}$.*
- *Množina $M \subseteq V$ je* ortogonální *(OG), pokud pro každé $\vec{x},\vec{y} \in M$, $\vec{x} \neq \vec{y}$, platí $\vec{x} \perp \vec{y}$.*
- *Množina $M \subseteq V$ je* ortonormální *(ON), pokud je ortogonální a všechny vektory v ní mají normu* 1.

Tvrzení 3.9. *Nechť V je lineární prostor nad $\mathbb{F}$, a $\|\cdot\|$ je norma na V indukovaná skalárním součinem $\langle\cdot,\cdot\rangle$. Potom pro každé $\vec{x},\vec{y} \in V$ platí:*

- *pokud $\vec{x} \perp \vec{y}$, potom $\|\vec{x}+\vec{y}\|^2 = \|\vec{x}\|^2 + \|\vec{y}\|^2$, (v případě ℓ_2 normy jde o klasickou* Pythagorovu větu*),*
- $\|\vec{x}+\vec{y}\|^2 + \|\vec{x}-\vec{y}\|^2 = 2\|\vec{x}\|^2 + 2\|\vec{y}\|^2$ *(tzv.* rovnoběžníkové pravidlo*).*

Důkaz. V případě, že $\vec{x} \perp \vec{y}$, máme

$$\begin{aligned}\|\vec{x}+\vec{y}\|^2 = \langle\vec{x}+\vec{y},\vec{x}+\vec{y}\rangle &=^{(S2),3.2} \langle\vec{x},\vec{x}\rangle + \langle\vec{x},\vec{y}\rangle + \langle\vec{y},\vec{x}\rangle + \langle\vec{y},\vec{y}\rangle = \\ &=^{(S1)} \langle\vec{x},\vec{x}\rangle + \langle\vec{x},\vec{y}\rangle + \overline{\langle\vec{x},\vec{y}\rangle} + \langle\vec{y},\vec{y}\rangle = \langle\vec{x},\vec{x}\rangle + \langle\vec{y},\vec{y}\rangle = \|\vec{x}\|^2 + \|\vec{y}\|^2 .\end{aligned}$$

Dále, pro libovolné $\vec{x},\vec{y} \in V$ platí:

$$\begin{aligned}\|\vec{x}+\vec{y}\|^2 + \|\vec{x}-\vec{y}\|^2 &= \langle\vec{x}+\vec{y},\vec{x}+\vec{y}\rangle + \langle\vec{x}-\vec{y},\vec{x}-\vec{y}\rangle =^{(S2),3.2} \\ &= \langle\vec{x},\vec{x}\rangle + \langle\vec{x},\vec{y}\rangle + \langle\vec{y},\vec{x}\rangle + \langle\vec{y},\vec{y}\rangle + \langle\vec{x},\vec{x}\rangle - \langle\vec{x},\vec{y}\rangle - \langle\vec{y},\vec{x}\rangle + \langle\vec{y},\vec{y}\rangle = \\ &= 2\langle\vec{x},\vec{x}\rangle + 2\langle\vec{y},\vec{y}\rangle = 2\|\vec{x}\|^2 + 2\|\vec{y}\|^2 .\end{aligned}$$

□

Poznámka 3.10. *Rovnoběžníkové pravidlo má důležitý teoretický význam: Platí, že norma $\|\cdot\|$ na lineárním prostoru V je indukovaná nějakým skalárním součinem právě tehdy, když pro ni platí rovnoběžníkové pravidlo (Fréchet-von Neuman).*

Tvrzení 3.11. *Pokud $M \subseteq V$ je ortogonální množina nenulových vektorů, potom je lineárně nezávislá.*

Důkaz. Ať $\{\vec{x}_1, \ldots, \vec{x}_n\}$ je libovolná konečná podmnožina M, a nechť skaláry $\alpha_1, \ldots, \alpha_n \in \mathbb{F}$ splňují

$$\sum_{i=1}^{n} \alpha_i \vec{x}_i = \vec{o}.$$

Chceme ukázat, že pro každé $j \in \{1, \ldots, n\}$, $\alpha_j = 0$. Zvolme tedy libovolné $j \in \{1, \ldots, n\}$. Potom máme

$$0 = \langle \vec{x}_j, \vec{o} \rangle = \langle \vec{x}_j, \sum_{i=1}^{n} \alpha_i \vec{x}_i \rangle =^{(S2)} \sum_{i=1}^{n} \alpha_i \langle \vec{x}_j, \vec{x}_i \rangle = \alpha_j \langle \vec{x}_j, \vec{x}_j \rangle.$$

Protože $\vec{x}_j \neq 0$, dle vlastnosti (S3) platí $\langle \vec{x}_j, \vec{x}_j \rangle > 0$. Tedy $\alpha_j = 0$, a tím pádem je množina M lineárně nezávislá. □

Na prostorech $\mathbb{F}^n$ v dalším vždy $\langle \cdot, \cdot \rangle$ značí standardní skalární součin a $\|\cdot\|$ ℓ_2 normu, pokud nebude uvedeno jinak.

Tvrzení 3.12. *Nechť U je lineární podprostor $\mathbb{F}^n$, $\mathcal{B} = (\vec{u}_1, \ldots, \vec{u}_k)$ je ON báze U. Potom pro každé $\vec{u} \in U$ platí*

$$\vec{u} = \sum_{i=1}^{k} \langle \vec{u}_i, \vec{u} \rangle \cdot \vec{u}_i.$$

Toto nazveme Fourierovým rozvojem *vektoru $\vec{u}$ vůči bázi $\mathcal{B}$.*

Důkaz. Jelikož $\mathcal{B}$ je báze U, pro každé $\vec{u} \in U$ platí $\vec{u} = \sum_{i=1}^{k} \alpha_i \vec{u}_i$, kde $[\vec{u}]_{\mathcal{B}} = (\alpha_1, \ldots, \alpha_k)^T$. Stačí tedy ukázat, že pro každé $j \in \{1, \ldots, k\}$ platí $\alpha_j = \langle \vec{u}_j, \vec{u} \rangle$: Máme

$$\langle \vec{u}_j, \vec{u} \rangle = \langle \vec{u}_j, \sum_{i=1}^{k} \alpha_i \vec{u}_i \rangle =^{(S2)} \sum_{i=1}^{k} \alpha_i \langle \vec{u}_j, \vec{u}_i \rangle = \alpha_j \|\vec{u}_j\|^2 = \alpha_j,$$

čímž je důkaz ukončen. □

Věta 3.13. *(Gram-Schmidt) Pokud V je lineární podprostor $\mathbb{F}^n$ a $\mathcal{B} = (\vec{v_1}, \ldots, \vec{v}_m)$ je báze V, potom existuje ON báze $\mathcal{C} = (\vec{u_1}, \ldots, \vec{u}_m)$ prostoru V taková, že pro každé $1 \leq k \leq m$,*

$$U_k := span(\vec{u}_1, \ldots, \vec{u}_k) = span(\vec{v}_1, \ldots, \vec{v}_k) =: V_k.$$

Navíc, báze $\mathcal{C}$ se dá nalézt podle rekurzivního vzorce:

- $\vec{u_1} = \frac{\vec{v_1}}{\|\vec{v_1}\|}$,
- $\vec{u}_k = \frac{\vec{v}_k - \sum_{i=1}^{k-1} \langle \vec{u}_i, \vec{v}_k \rangle \cdot \vec{u}_i}{\left\| \vec{v}_k - \sum_{i=1}^{k-1} \langle \vec{u}_i, \vec{v}_k \rangle \cdot \vec{u}_i \right\|}, \quad 1 < k \leq m.$

Důkaz. Volíme vektory $(\vec{u_1}, \ldots, \vec{u}_m)$ rekurzivně jako v tvrzení věty, a ukážeme, že to funguje. Indukcí ukážeme, že pro každé $1 \leq k \leq m$ je seznam $(\vec{u_1}, \ldots, \vec{u}_k)$ ON a platí $U_k = V_k$.

Pro $k = 1$ volíme $\vec{u_1} = \frac{\vec{v_1}}{\|\vec{v_1}\|}$, a tvrzení je zřejmé. Označme $\nu_1 = \|\vec{v}_1\|$.

Pro $1 < k < m$ předpokládejme, že jsme nalezli $(\vec{u_1}, \ldots, \vec{u}_{k-1})$, a definujme $\nu_k = \left\| \vec{v}_k - \sum_{i=1}^{k-1} \langle \vec{u}_i, \vec{v}_k \rangle \cdot \vec{u}_i \right\|$, a

$$\vec{u}_k = \frac{\vec{v}_k - \sum_{i=1}^{k-1} \langle \vec{u}_i, \vec{v}_k \rangle \cdot \vec{u}_i}{\nu_k}.$$

Potom $\|\vec{u}_k\| = 1$. Abychom ukázali, že seznam $(\vec{u_1}, \ldots, \vec{u}_k)$ je ON, stačí díky indukčnímu předpokladu ukázat, že pro každé $1 \leq j \leq k-1$,

$$\vec{u}_j \perp \vec{u}_k \Leftrightarrow \langle \vec{u}_j, \vec{v}_k - \sum_{i=1}^{k-1} \langle \vec{u}_i, \vec{v}_k \rangle \cdot \vec{u}_i \rangle = 0.$$

Máme:

$$\langle \vec{u}_j, \vec{v}_k - \sum_{i=1}^{k-1} \langle \vec{u}_i, \vec{v}_k \rangle \cdot \vec{u}_i \rangle =^{(S2)} \langle \vec{u}_j, \vec{v}_k \rangle - \sum_{i=1}^{k-1} \langle \vec{u}_i, \vec{v}_k \rangle \cdot \langle \vec{u}_j, \vec{u}_i \rangle =^{\text{indukční předpoklad}} \langle \vec{u}_j, \vec{v}_k \rangle - \langle \vec{u}_j, \vec{v}_k \rangle = 0.$$

Rovnost $U_k = V_k$: cvičení. □

Poznámka 3.14. *Pro vektor $\vec{u} \in U_m$ a $1 \leq k \leq m$ se vektor $\sum_{i=1}^{k} \langle \vec{u}_i, \vec{u} \rangle \cdot \vec{u}_i$ nazývá* ortogonální projekce *$\vec{u}$ na podprostor U_k, a $\vec{u} - \sum_{i=1}^{k} \langle \vec{u}_i, \vec{u} \rangle \cdot \vec{u}_i$ je* ortogonální rejekce *vektoru $\vec{u}$ podprostorem U_k.*

Věta 3.15. *(QR-rozklad) Nechť $A \in \mathbb{F}^{m \times n}$, $rank(A) = n$. Potom existují matice $Q \in \mathbb{F}^{m \times n}, R \in \mathbb{F}^{n \times n}$ takové, že:*

(1) Q má ON sloupce a $Im(Q) = Im(A)$,

(2) R je horní trojúhelníková a má kladná reálná čísla na hlavní diagonále,

(3) $A = Q \cdot R$.

Toto nazýváme QR-rozkladem *matice A.*

Důkaz. Užijeme Gram-Schmidtovu ortogonalizaci (Věta 3.13) na sloupce matice A (jelikož $rank(A) = n$, je $m \geq n$ a sloupce $(\vec{a}_1, \ldots, \vec{a}_n)$ matice A tvoří bázi $Im(A) \subseteq \mathbb{F}^m$). Výslednou ON bázi označíme $(\vec{q}_1, \ldots, \vec{q}_n)$. Platí tedy $\vec{q}_1 = \frac{\vec{a}_1}{\nu_1}$, a pro $1 < k \leq n$,

$$\vec{q}_k = \frac{\vec{a}_k - \sum_{i=1}^{k-1} \langle \vec{q}_i, \vec{a}_k \rangle \cdot \vec{q}_i}{\nu_k},$$

ekvivalentně platí $\vec{a}_1 = \nu_1 \cdot \vec{q}_1$, a

$$\vec{a_k} = \sum_{i=1}^{k-1} \langle \vec{q}_i, \vec{a}_k \rangle \cdot \vec{q}_i + \nu_k \vec{q}_k.$$

Maticově toto můžeme zapsat jako:

$$A = (\vec{a_1}, \ldots, \vec{a}_n) = (\vec{q}_1, \ldots, \vec{q}_n) \cdot \begin{pmatrix} \nu_1 & \langle \vec{q}_1, \vec{a}_2 \rangle & \langle \vec{q}_1, \vec{a}_3 \rangle & \ldots & \langle \vec{q}_1, \vec{a}_n \rangle \\ 0 & \nu_2 & \langle \vec{q}_2, \vec{a}_3 \rangle & \ldots & \langle \vec{q}_2, \vec{a}_n \rangle \\ 0 & 0 & \nu_3 & \ldots & \langle \vec{q}_3, \vec{a}_n \rangle \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \nu_n \end{pmatrix} = Q \cdot R.$$

Potom Q má ON sloupce, neboť $(\vec{q}_1, \ldots, \vec{q}_n)$ je ON množina, rovnost $Im(Q) = Im(A)$ je ekvivalentní rovnosti $span(\vec{q_1}, \ldots, \vec{q}_k) = span(\vec{a}_1, \ldots, \vec{a}_k)$, R je horní trojúhelníková, a na diagonále má čísla $\nu_1, \ldots, \nu_n$, která jsou kladná. □

Příklad 3.16. *(Aplikace QR-rozkladu na metodu nejmenších čvterců) Nechť $A \in \mathbb{R}^{m \times n}$, $rank(A) = n$, $\vec{b} \in \mathbb{R}^n$.* Metoda nejmenších čtverců *dává nejlepší možné přibližné řešení $\vec{x}$ soustavy $(A|\vec{b})$ (která žádné přesné řešení mít nemusí), ve smyslu*

$$\left\| A\vec{x} - \vec{b} \right\|_2 = \min_{\vec{y} \in \mathbb{R}^n} \left\| A\vec{y} - \vec{b} \right\|_2.$$

Toto přibližné řešení $\vec{x}$ je rovno přesnému řešení soustavy $(A^T \cdot A | A^T \vec{b})$. Je ale známo, že způsob hledání $\vec{x}$ pomocí výpočtu $A^T \cdot A$ a následné Gaussovy eliminace není příliš vhodný (je výpočetně náročný a numericky nestabilní). Mezi základní způsoby nalezení vektoru $\vec{x}$ patří QR-rozklad: pokud $A = Q \cdot R$, potom, jelikož Q má ON sloupce, je $Q^T \cdot Q = I_n$, a R^T je regulární (neboť je dolní trojúhelníková s pozitivními čísly na diagonále), a tedy

$$A^T \cdot A \cdot \vec{x} = A^T \cdot \vec{b} \Leftrightarrow R^T \cdot Q^T \cdot Q \cdot R \cdot \vec{x} = R^T \cdot Q^T \cdot \vec{b} \Leftrightarrow R^T \cdot R \cdot \vec{x} = R^T \cdot Q^T \cdot \vec{b} \Leftrightarrow R \cdot \vec{x} = Q^T \cdot \vec{b}.$$

Pokud známe QR-rozklad matice A, vektor $\vec{x}$ tedy můžeme hledat jako řešení soustavy $(R | Q^T \cdot \vec{b})$.

4 Projekce

Definice 4.1. *Nechť W je lineární prostor nad $\mathbb{F}$, U, V jsou lineární podprostory W. Řekneme, že W je* direktním součtem *U a V, píšeme $W = U \oplus V$, pokud platí, že $U \cap V = \vec{o}$ a pro každé $\vec{x} \in W$ existují $\vec{u} \in U, \vec{v} \in V$ takové, že $\vec{x} = \vec{u} + \vec{v}$. U a V se nazývají* komplementární podprostory *W.*
Pokud $W = U \oplus V$, potom každé $\vec{x} \in W$ lze pak jednoznačně zapsat ve tvaru $\vec{x} = \vec{u}_x + \vec{v}_x$ pro nějaká $\vec{u}_x \in U, \vec{v}_x \in V$. Lineární zobrazení $P_U^V \colon W \to U$, které vektoru $\vec{x}$ přiřadí vektor $\vec{u}_x$, nazýváme projekcí na U podle V. Projekce P_V^U se nazývá komplementární projekcí *k projekci P_U^V.*

Poznámka 4.2. *Všimněte si, že platí $(P_U^V)|_U = Id_U$, $(P_U^V)|_V = 0$, $U = Im(P_U^V)$ a $V = ker(P_U^V)$, a $P_V^U = Id - P_U^V$.*

Tvrzení 4.3. *Nechť $\mathbb{F}^n = U \oplus V$. Nechť $\mathcal{B}_U = (\vec{u}_1, \dots, \vec{u}_k), \mathcal{B}_V = (\vec{v}_1, \dots, \vec{v}_{n-k})$ jsou báze U, V. Označme B_U, B_V matice, které mají ve sloupcích vektory z bází $\mathcal{B}_U, \mathcal{B}_V$. Potom pro matici projekce P_U^V (vůči kanonické bázi $\mathbb{F}^n$) platí*

$$P_U^V = \begin{pmatrix} B_U & O_{n,n-k} \end{pmatrix} \cdot \begin{pmatrix} B_U & B_V \end{pmatrix}^{-1} = \begin{pmatrix} B_U & B_V \end{pmatrix} \cdot \begin{pmatrix} I_k & O_{k,n-k} \\ O_{n-k,k} & O_{n-k,n-k} \end{pmatrix} \cdot \begin{pmatrix} B_U & B_V \end{pmatrix}^{-1}.$$

Důkaz. Jelikož $\mathbb{F}^n = U \oplus V$, je $\mathcal{B}_U \cup \mathcal{B}_V$ báze $\mathbb{F}^n$ (cvičení). Tedy, matice $\begin{pmatrix} B_U & B_V \end{pmatrix}$ je regulární. Dále, dle definice projekce P_U^V platí $P_U^V(\vec{u}_i) = \vec{u}_i$ a $P_U^V(\vec{v}_j) = \vec{o}$ pro každé $i \in \{1, \dots, k\}, j \in \{1, \dots, n-k\}$. Tedy platí

$$P_U^V \cdot \begin{pmatrix} B_U & B_V \end{pmatrix} = \begin{pmatrix} P_U^V \cdot B_U & P_U^V \cdot B_V \end{pmatrix} = \begin{pmatrix} B_U & O_{n,n-k} \end{pmatrix},$$

a tedy

$$P_U^V = \begin{pmatrix} B_U & O_{n,n-k} \end{pmatrix} \cdot \begin{pmatrix} B_U & B_V \end{pmatrix}^{-1}.$$

Druhá rovnost plyne okamžitě z toho, jak funguje násobení blokových matic. □

Věta 4.4. *Pro matici $P \in \mathbb{F}^{n \times n}$ je ekvivalentní:*

(i) $Im(P) \oplus ker(P) = \mathbb{F}^n$ a P je matice projekce na $Im(P)$ podle $ker(P)$,

(ii) $P^2 = P$.

Důkaz. (i) ⇒ (ii): Plyne okamžitě z Tvrzení 4.3, neboť máme

$$P^2 = \begin{pmatrix} B_U & B_V \end{pmatrix} \cdot \begin{pmatrix} I_k & O_{k,n-k} \\ O_{n-k,k} & O_{n-k,n-k} \end{pmatrix} \cdot \begin{pmatrix} B_U & B_V \end{pmatrix}^{-1} \cdot \begin{pmatrix} B_U & B_V \end{pmatrix} \cdot \begin{pmatrix} I_k & O_{k,n-k} \\ O_{n-k,k} & O_{n-k,n-k} \end{pmatrix} \cdot \begin{pmatrix} B_U & B_V \end{pmatrix}^{-1} =$$
$$\begin{pmatrix} B_U & B_V \end{pmatrix} \cdot \begin{pmatrix} I_k & O_{k,n-k} \\ O_{n-k,k} & O_{n-k,n-k} \end{pmatrix} \cdot \begin{pmatrix} B_U & B_V \end{pmatrix}^{-1} = P.$$

(ii) ⇒ (i) Libovolné $\vec{x} \in \mathbb{F}^n$ lze zapsat ve tvaru

$$\vec{x} = P(\vec{x}) + (\vec{x} - P(\vec{x})).$$

Potom $P(\vec{x}) \in Im(P)$, a $\vec{x} - P(\vec{x}) \in ker(P)$, neboť

$$P(\vec{x} - P(\vec{x})) = P(\vec{x}) - P^2(\vec{x}) = P(\vec{x}) - P(\vec{x}) = \vec{o}.$$

Dále ukážeme, že $Im(P) \cap ker(P) = \vec{o}$: Pokud $\vec{y} \in Im(P) \cap ker(P)$, pak $\vec{y} = P(\vec{x})$ pro nějaké $\vec{x} \in \mathbb{F}^n$, a $P(\vec{y}) = \vec{o}$. Tedy

$$\vec{o} = P(\vec{y}) = P^2(\vec{x}) = P(\vec{x}) = \vec{y}.$$

Tedy $\mathbb{F}^n = Im(P) \oplus ker(P)$, a $\vec{x} = P(\vec{x}) + (\vec{x} - P(\vec{x}))$ je příslušný jednoznačný rozklad, což znamená, že $\vec{x} \mapsto P(\vec{x})$ je projekce na $Im(P)$ podle $ker(P)$. □

Definice 4.5. *Pro množinu $U \subseteq \mathbb{F}^n$ definujeme její* ortogonální doplněk *jako*

$$U^\perp = \{\vec{x} \in \mathbb{F}^n \colon \langle \vec{x}, \vec{u} \rangle = 0 \text{ pro každé } \vec{u} \in U\}.$$

Tvrzení 4.6. *Nechť $U, V \subseteq \mathbb{F}^n$. Potom platí:*

(i) $U \subseteq V \Leftrightarrow V^\perp \subseteq U^\perp$,

(ii) $U^\perp$ *je lineární podprostor* $\mathbb{F}^n$,

(iii) $U^\perp = span(U)^\perp$,

(iv) $\mathbb{F}^n = U^\perp \oplus span(U)$,

(v) $U^{\perp\perp} = span(U)$.

Důkaz. Cvičení. □

Definice 4.7. *Nechť U je lineární podprostor $\mathbb{F}^n$. Potom projekci $P_U^{U^\perp}$ nazveme* ortogonální projekcí *na U.*

Tvrzení 4.8. *Nechť U je lineární podprostor $\mathbb{F}^n$, $\mathcal{B}_U = (\vec{u}_1, \ldots, \vec{u}_k)$ je ortonormální báze U, a B_U je matice, která má ve sloupcích vektory z báze $\mathcal{B}_U$. Potom $P_U^{U^\perp} = B_U \cdot B_U^*$.*

Důkaz. Označme $V = U^\perp$. Nechť $\mathcal{B}_V$ je libovolná báze V a B_V je matice, která má ve sloupcích vektory z báze $\mathcal{B}_V$. Dle Tvrzení 4.3 víme, že $P_U^V = \begin{pmatrix} B_U & O_{n,n-k} \end{pmatrix} \cdot \begin{pmatrix} B_U & B_V \end{pmatrix}^{-1}$. Rovnost $P_U^V = B_U \cdot B_U^*$ je tedy ekvivalentní rovnosti

$$\begin{aligned} &\begin{pmatrix} B_U & O_{n,n-k} \end{pmatrix} \cdot \begin{pmatrix} B_U & B_V \end{pmatrix}^{-1} = B_U \cdot B_U^* \\ \Leftrightarrow &\begin{pmatrix} B_U & O_{n,n-k} \end{pmatrix} = B_U \cdot B_U^* \cdot \begin{pmatrix} B_U & B_V \end{pmatrix} = \begin{pmatrix} B_U \cdot B_U^* \cdot B_U & B_U \cdot B_U^* \cdot B_V \end{pmatrix}, \end{aligned}$$

která platí, neboť $B_U^* \cdot B_U = I_n$ dle ortonormality báze $\mathcal{B}_U$, a $B_U^* \cdot B_V = O_{n,n-k}$, neboť vektory z $\mathcal{B}_V$ jsou kolmé na vektory z $\mathcal{B}_U$. □

Poznámka 4.9. *Pro U jako výše a $\vec{x} \in \mathbb{F}^n$ platí*

$$P_U^{U^\perp}(\vec{x}) = B_U \cdot B_U^* \cdot \vec{x} = B_U \cdot \begin{pmatrix} \vec{u}_1^* \vec{x} \\ \vdots \\ \vec{u}_k^* \vec{x} \end{pmatrix} = B_U \cdot \begin{pmatrix} \langle \vec{u}_1, \vec{x} \rangle \\ \vdots \\ \langle \vec{u}_k, \vec{x} \rangle \end{pmatrix} = \sum_{i=1}^{k} \langle \vec{u}_i, \vec{x} \rangle \vec{u}_i,$$

tedy dostaneme vzoreček, který už známe z Tvrzení 3.12 a Věty 3.13.

5 Unitární a ortogonální matice, Householderova metoda

Definice 5.1.

- Unitární matice *je matice $U \in \mathbb{C}^{n \times n}$, jejíž sloupce tvoří ON bázi $\mathbb{C}^n$.*
- Ortogonální (nebo ortonormální) matice *je matice $P \in \mathbb{R}^{n \times n}$, jejíž sloupce tvoří ON bázi $\mathbb{R}^n$.*

Tvrzení 5.2. *Pro matici $U \in \mathbb{C}^{n \times n}$ jsou následující tvrzení ekvivalentní:*

(i) U *je unitární,*

(ii) *řádky* U *tvoří ON bázi* $\mathbb{C}^n$,

(iii) $U^* = U^{-1}$, *tj.* $U^* \cdot U = I_n$,

(iv) pro každé $\vec{x} \in \mathbb{C}^n$, $\|U \cdot \vec{x}\| = \|\vec{x}\|$.

Důkaz. Napišme U ve sloupcovém tvaru jako $U = (\vec{u}_1, \dots, \vec{u}_n)$.
Ekvivalence (i) $\Leftrightarrow$ (iii): U je unitární $\Leftrightarrow$ U má ON sloupce $\Leftrightarrow$ pro každé $i, j \in \{1, \dots, n\}$ platí, že

$$[U^* \cdot U]_{i,j} = \vec{u}_i^* \cdot \vec{u}_j = \langle \vec{u}_i, \vec{u}_j \rangle = \begin{cases} 1, & i = j, \\ 0, & i \neq j, \end{cases} \Leftrightarrow U^* \cdot U = I_n.$$

Dále, protože platí

$$U^* \cdot U = I_n \Leftrightarrow U^* = U^{-1} \Leftrightarrow U = (U^*)^* = (U^{-1})^* = (U^*)^{-1} \Leftrightarrow U \cdot U^* = I_n,$$

analogicky jako výše dostaneme ekvivalenci (ii) $\Leftrightarrow$ (iii).
(iii) $\Rightarrow$ (iv): Pokud platí (iii), potom pro každé $\vec{x} \in \mathbb{C}^n$ máme

$$\|U \cdot \vec{x}\|^2 = (U \cdot \vec{x})^* \cdot U\vec{x} = \vec{x}^* \cdot U^* \cdot U \cdot \vec{x} = \vec{x}^* \cdot \vec{x} = \|\vec{x}\|^2.$$

(iv) $\Rightarrow$ (i): Z (iv) plyne, že pro každé $\vec{x} \in \mathbb{C}^n$ platí

$$\vec{x}^* \cdot U^* \cdot U \cdot \vec{x} = \vec{x}^* \cdot \vec{x}. \qquad (*)$$

Aplikujeme $(*)$ ve třech různých případech.
Zaprvé, na vektor $\vec{x} = \vec{e_i}$ pro nějaké $i \in \{1, \dots, n\}$, a dostaneme

$$\vec{u_i}^* \cdot \vec{u_i} = \vec{e_i}^* \cdot U^* \cdot U \cdot \vec{e_i} = \vec{e_i}^* \vec{e_i} = 1.$$

Dále, aplikujeme $(*)$ na vektor $\vec{x} = \vec{e_j} + \vec{e_k}$ pro nějaká různá $j, k \in \{1, \dots, n\}$, a máme

$$\begin{aligned} &2 + \vec{u_k}^* \vec{u_j} + \vec{u_j}^* \vec{u_k} = \vec{u_j}^* \cdot \vec{u_j} + \vec{u_k}^* \vec{u_j} + \vec{u_j}^* \vec{u_k} + \vec{u_k}^* \cdot \vec{u_k} = \\ &= (\vec{u_j}^* + \vec{u_k}^*) \cdot (\vec{u_k} + \vec{u_k}) = (\vec{e_j} + \vec{e_k})^* \cdot U^* \cdot U \cdot (\vec{e_j} + \vec{e_k}) = (\vec{e_j} + \vec{e_k})^* \cdot (\vec{e_j} + \vec{e_k}) = 2, \end{aligned}$$

tedy

$$0 = \vec{u_k}^* \vec{u_j} + \vec{u_j}^* \vec{u_k} = \vec{u_k}^* \vec{u_j} + \overline{\vec{u_k}^* \vec{u_j}} = 2 \cdot Re(\vec{u_k}^* \vec{u_j}).$$

Nakonec, aplikujeme $(*)$ na vektor $\vec{x} = \vec{e_j} + i \cdot \vec{e_k}$ pro nějaká různá $j, k \in \{1, \dots, n\}$, a analogicky jako výše dostaneme $0 = Im(\vec{u_k}^* \vec{u_j})$.
Celkově tedy $\vec{u_k}^* \vec{u_j} = 0$ pro $i \neq j$, a tedy sloupce U jsou ortonormální. □

Tvrzení 5.3. *Pro matici $P \in \mathbb{R}^{n \times n}$ jsou následující tvrzení ekvivalentní:*

(i) P je ortogonální,

(ii) řádky P tvoří ON bázi $\mathbb{R}^n$,

(iii) $P^T = P^{-1}$, tj. $P^T \cdot P = I_n$,

(iv) pro každé $\vec{x} \in \mathbb{R}^n$, $\|P\vec{x}\| = \|\vec{x}\|$.

Důkaz. Důkaz je analogický jako v Tvrzení 5.2. □

Poznámka 5.4. *Jsou-li $(V_1, \|\cdot\|_a), (V_2, \|\cdot\|_b)$ lineární prostory s normou, potom lineární zobrazení $f \colon V_1 \to V_2$ je* isometrie*, pokud pro každé $\vec{x} \in V_1$, $\|f(\vec{x})\|_b = \|\vec{x}\|_a$. Dle tvrzení 5.2 a 5.3 jsou tedy všechny isometrie $\mathbb{C}^n \to \mathbb{C}^n$ (s ℓ_2 normou) reprezentovány unitárními maticemi, a všechny isometrie $\mathbb{R}^n \to \mathbb{R}^n$ (s ℓ_2 normou) jsou reprezentovány ortogonálními maticemi.*

Poznámka 5.5. *Isometrií mezi prostory $(\mathbb{F}^n, \|\cdot\|_p) \to (\mathbb{F}^n, \|\cdot\|_p)$, pro $p \neq 2$, je daleko méně, než v případě $p = 2$. Například, rotace v $\mathbb{R}^2$ o 45 stupňů je isometrie pro ℓ_2 normu, ale není to isometrie pro ℓ_p normu pro žádné $p \neq 2$ (to plyne například z tvaru jednotkové sféry).*

Důsledek 5.6. *Součin dvou unitárních matic v $\mathbb{C}^n$ je unitární matice. Součin dvou ortogonálních matic v $\mathbb{R}^n$ je ortogonální matice.*

Důkaz. Pokud $U, V \in \mathbb{C}^{n\times n}$ jsou unitární a $\vec{x} \in \mathbb{C}^n$ je libovolný vektor, potom dle Tvrzení 5.2, použitého zvlášť na matice U, V, platí

$$\|U \cdot V \cdot x\| = \|V \cdot x\| = \|x\|,$$

a tedy opět dle Tvrzení 5.2 je matice $U \cdot V$ unitární.
Pro ortogonální matice je to stejné. □

Definice 5.7. *Nechť $\vec{u} \in \mathbb{F}^n \setminus \{\vec{o}\}$.*

- *Linerární zobrazení dané maticí*

$$R_{\vec{u}} = I_n - 2\frac{\vec{u}\cdot\vec{u}^*}{\vec{u}^*\cdot\vec{u}} = I_n - 2\frac{\vec{u}\cdot\vec{u}^*}{\|\vec{u}\|^2}$$

  se nazývá elementární Householderova reflexe *vůči $\vec{u}^\perp$.*

Všimněte si, že v situaci z předcházející definice, je $I_n - \frac{\vec{u}\cdot\vec{u}^*}{\|\vec{u}\|^2}$ matice ortogonální projekce na $\vec{u}^\perp$.

Definice 5.8. *Matice $A \in \mathbb{F}^{n\times n}$ se nazývá* involuce, *pokud $A^2 = I_n$, tj. $A^{-1} = A$.*

Tvrzení 5.9. *Pro každé $\vec{u} \in \mathbb{F}^n$ je matice $R_{\vec{u}}$ involuce, unitární, a Hermitovská.*

Důkaz. Máme

$$R_{\vec{u}}^* = \left(I_n - 2\frac{\vec{u}\cdot\vec{u}^*}{\vec{u}^*\cdot\vec{u}}\right)^* = I_n^* - 2\frac{\vec{u}\cdot\vec{u}^*}{\vec{u}^*\cdot\vec{u}} = R_{\vec{u}}.$$

Dále, platí

$$\begin{aligned} R_u^2 &= \left(I_n - 2\frac{\vec{u}\cdot\vec{u}^*}{\vec{u}^*\cdot\vec{u}}\right)\cdot\left(I_n - 2\frac{\vec{u}\cdot\vec{u}^*}{\vec{u}^*\cdot\vec{u}}\right) = I_n^2 - 2\frac{\vec{u}\cdot\vec{u}^*}{\vec{u}^*\cdot\vec{u}} - 2\frac{\vec{u}\cdot\vec{u}^*}{\vec{u}^*\cdot\vec{u}} + 4\frac{\vec{u}\cdot\vec{u}^*\cdot\vec{u}\cdot\vec{u}^*}{\vec{u}^*\cdot\vec{u}\cdot\vec{u}^*\cdot\vec{u}} = \\ &= I_n - 4\frac{\vec{u}\cdot\vec{u}^*}{\vec{u}^*\cdot\vec{u}} + 4\frac{\vec{u}\cdot\vec{u}^*}{\vec{u}^*\cdot\vec{u}} = I_n. \end{aligned}$$

Tedy máme $R_{\vec{u}} = R_{\vec{u}}^* = R_{\vec{u}}^{-1}$, což jsme chtěli ukázat. □

Tvrzení 5.10. *Pro každé $\vec{x}, \vec{y} \in \mathbb{F}^n$ s $\|\vec{x}\| = \|\vec{y}\|$ existuje unitární matice U taková, že $U\cdot\vec{x} = \vec{y}$. Pokud $\vec{x}, \vec{y} \in \mathbb{R}^n$, lze nalézt ortogonální U.*

Myšlenka v případě $\vec{x}, \vec{y} \in \mathbb{R}^n$: stačí vzít Householderovu reflexi podle vektoru $(\vec{x}-\vec{y})^\perp$ (nakreslete si obrázek).

Důkaz. Nalezněme nejprve úhel $\theta \in [0, 2\pi)$ tak, že $e^{i\theta}\cdot\vec{y}^*\cdot\vec{x} \in \mathbb{R}$, a definujme

$$\vec{u} = e^{i\theta}\vec{x} - \vec{y}, \text{ a } U = e^{i\theta}R_{\vec{u}}.$$

Potom U dle Tvrzení 5.9 máme

$$U^* = \overline{e^{i\theta}}\cdot R_{\vec{u}}^* = e^{-i\theta}\cdot R_{\vec{u}}^{-1} = \frac{1}{e^{i\theta}}\cdot R_{\vec{u}}^{-1} = U^{-1},$$

tedy, U je unitární.
Dále, jelikož platí $\|\vec{x}\| = \|\vec{y}\|$, máme

$$\vec{x}^*\cdot\vec{x} = \|\vec{x}\|^2 = \|\vec{y}\|^2 = \vec{y}^*\cdot\vec{y}, \tag{1}$$

a, jelikož $e^{i\theta}\cdot\vec{y}^*\cdot\vec{x} \in \mathbb{R}$, je

$$e^{i\theta}\vec{y}^*\cdot\vec{x} = \overline{e^{i\theta}\vec{y}^*\cdot\vec{x}} = (e^{i\theta}\vec{y}^*\cdot\vec{x})^* = e^{-i\theta}\vec{x}^*\cdot\vec{y}. \tag{2}$$

Tedy

$$\begin{aligned} \vec{u}^*\cdot\vec{u} &= (e^{-i\theta}\vec{x}^* - \vec{y}^*)\cdot(e^{i\theta}\vec{x} - \vec{y}) = \vec{x}^*\cdot\vec{x} - e^{i\theta}\vec{y}^*\cdot\vec{x} - e^{-i\theta}\vec{x}^*\cdot\vec{y} + \vec{y}^*\cdot\vec{y} = \\ &=^{(1),(2)} 2(\vec{x}^*\cdot\vec{x} - e^{i\theta}\vec{y}^*\cdot\vec{x}) = 2(\vec{x}^* - e^{i\theta}\vec{y}^*)\cdot\vec{x} = 2e^{i\theta}(e^{-i\theta}\vec{x}^* - \vec{y}^*)\cdot\vec{x} = 2e^{i\theta}\vec{u}^*\cdot\vec{x}. \end{aligned} \tag{3}$$

Konečně, dostaneme

$$U \cdot \vec{x} = e^{i\theta} R_{\vec{u}} \cdot \vec{x} = e^{i\theta}\Big(\vec{x} - \frac{2\vec{u} \cdot \vec{u}^* \cdot \vec{x}}{\vec{u}^* \cdot \vec{u}}\Big) = e^{i\theta}\vec{x} - \frac{2e^{i\theta}\vec{u}^* \cdot \vec{x}}{\vec{u}^* \cdot \vec{u}} \cdot \vec{u} =^{(3)} e^{i\theta}\vec{x} - \vec{u} = \vec{y},$$

což jsme chtěli ukázat. □

Definice 5.11. *Matice $A \in \mathbb{F}^{m\times n}$ je v* horním lichoběžníkovém tvaru, *pokud $[A]_{i,j} = 0$ pro $j < i$, jinými slovy, pokud existuje horní trojúhelníková matice R tak, že*

- $A = \begin{pmatrix} R & B \end{pmatrix}$, *pro nějakou matici $B \in \mathbb{F}^{m\times n-m}$, pokud $m \leq n$, nebo*
- $A = \begin{pmatrix} R \\ O_{m-n,n} \end{pmatrix}$, *pokud $n < m$.*

Věta 5.12. *(Householderova redukce) Pro každou matici $A \in \mathbb{C}^{m\times n}$, existují unitární matice U a horní lichoběžníková matice T takové, že $U \cdot A = T$. Navíc:*

- *Pokud A je čtvercová, potom T je horní trojúhelníková, a*
- *pokud $A \in \mathbb{R}^{m\times n}$, potom lze nalézt U ortogonální.*

Důkaz. Nechť $\vec{x}$ je první sloupec matice A. Dle Tvrzení 5.10 nalezneme unitární matici $U_1 \in \mathbb{C}^{m\times m}$ tak, aby $U_1\vec{x} = \|\vec{x}\| \cdot \vec{e_1}$.
Potom $U_1 \cdot A = \begin{pmatrix} \|\vec{x}\| & * \\ O_{m-1,1} & B \end{pmatrix}$ pro nějakou matici $B \in \mathbb{C}^{(m-1)\times(n-1)}$. Nyní, opět dle Tvrzení 5.10 nalezneme unitární matici $\tilde{U}_2 \in \mathbb{C}^{(m-1)\times(m-1)}$ splňující $\tilde{U}_2 \cdot \vec{y} = \|\vec{y}\| \cdot \vec{e}_1$, kde $\vec{y}$ je první sloupec matice B. Definujme

$$U_2 = \begin{pmatrix} 1 & O_{1,m-1} \\ O_{m-1,1} & \tilde{U}_2 \end{pmatrix}.$$

Potom U_2 je unitární, a

$$U_2 \cdot U_1 \cdot A = U_2 \cdot \begin{pmatrix} \|\vec{x}\| & * \\ O_{m-1,1} & B \end{pmatrix} = \begin{pmatrix} \|\vec{x}\| & * \\ O_{m-1,1} & \tilde{U}_2 \cdot B \end{pmatrix} = \begin{pmatrix} \|\vec{x}\| & * & * \\ 0 & \|\vec{y}\| & * \\ O_{m-2,1} & O_{m-2,1} & C \end{pmatrix}$$

pro nějakou matici $C \in \mathbb{C}^{(m-2)\times(n-2)}$.

Pokračujeme analogicky dále, skončíme po $\min\{m-1, n-1\}$ krocích (pokud v některém kroku vyšel v prvním sloupci nulový vektor, neděláme nic, a přejdeme k dalšímu sloupci).

Potom, matice

$$U = U_{\min\{m-1,n-1\}} \cdot \ldots \cdot U_2 \cdot U_1$$

je unitární dle Důsledku 5.6, a $U \cdot A =: T$ je horní lichoběžníková. Je snadné si uvědomit, že pro čtvercovou matici A vyjde T horní trojúhelníková, a z Tvrzení 5.10 plyne, že pro reálnou matici A lze nalézt U ortogonální. □

Všimněte si, že pro regulární čtvercovou matici A z této věty dostaneme jiný důkaz existence QR-rozkladu (Věta 3.15) pro A, neboť v takovém případě je T horní tojúhelníková s kladnými čísly na diagonále, U^* má ON sloupce (dle Tvrzení 5.2), a máme

$$U \cdot A = T \Leftrightarrow A = U^* \cdot T.$$

6 Vlastní čísla a vlastní vektory, diagonalizace

Definice 6.1. *Nechť $A \in \mathbb{F}^{n\times n}$. Skalár $\lambda \in \mathbb{C}$ se nazývá* vlastní číslo *matice A, pokud existuje nenulový vektor $\vec{x} \in \mathbb{C}^n$ splňující $A\vec{x} = \lambda\vec{x}$, tj. $(A - \lambda I_n)\vec{x} = \vec{o}$. Vektor $\vec{x}$ se pak nazývá* vlastní vektor *A příslušný k vlastnímu číslu λ. Dále*

$$eigen(\lambda, A) = \{\vec{x} \in \mathbb{C}^n \colon A\vec{x} = \lambda\vec{x}\}$$

se nazývá vlastní podprostor *matice A příslušný k vlastnímu číslu λ.*

Množina všech vlastních čísel matice A se značí $\sigma(A)$ a nazývá se spektrum *matice A.*
Polynom $p_A \colon \mathbb{C} \to \mathbb{C}$ definovaný jako $p_A(x) = \det(A - xI_n)$ nazýváme charakteristický polynom *matice A.*

Pozorování 6.2. *Platí*

- $\lambda \in \sigma(A) \Leftrightarrow \exists \vec{x} \in \mathbb{F}^n, \vec{x} \neq \vec{o} \colon (A - \lambda I_n)\vec{x} = \vec{o} \Leftrightarrow ker(A - \lambda I_n) \neq \vec{o} \Leftrightarrow A - \lambda I_n$ *je singulární,*
- $0 \in \sigma(A) \Leftrightarrow A$ *je singulární,*
- $eigen(\lambda, A) = ker(A - \lambda I_n)$,
- $p_A(\lambda) = 0 \Leftrightarrow \det(A - \lambda I_n) = 0 \Leftrightarrow A - \lambda I_n$ *je singulární* $\Leftrightarrow \lambda \in \sigma(A)$.

Definice 6.3. *Pro $A \in \mathbb{F}^{n\times n}$ a $\lambda \in \mathbb{C}$ definujeme*

- algebraickou násobnost λ *jako násobnost λ jakožto kořene $p_A(x)$, značíme $an_A(\lambda)$,*
- geometrickou násobnost λ *jako $dim(eigen(\lambda, A))$, značíme $gn_A(\lambda)$.*

Definice 6.4.
- *Matice $A, B \in \mathbb{F}^{n\times n}$ jsou si* podobné *(zapisujeme $A \approx B$), pokud existuje regulární matice $T \in \mathbb{F}^{n\times n}$ splňující $B = T^{-1} \cdot A \cdot T \Leftrightarrow T \cdot B = A \cdot T \Leftrightarrow T \cdot B \cdot T^{-1} = A$.*
- *$A \in \mathbb{F}^{n\times n}$ je* diagonalizovatelná*, pokud je podobná nějaké diagonální matici.*

Připomeňte si, že podobné matice reprezentují stejné lineární zobrazení, jen vůči různým bázím (to plyne z Tvrzení 1.28).

Cvičení 6.5. *Pro matice $A, B, C \in \mathbb{F}^{n\times n}$ platí:*

- $A \approx A$,
- $A \approx B \Leftrightarrow B \approx A$,
- $A \approx B$ *a* $B \approx C \Rightarrow A \approx C$.

Tvrzení 6.6. *Pokud $A \approx B$, potom A a B mají stejný charakteristický polynom, a tedy i stejná vlastní čísla.*

Důkaz. Máme

$$p_B(\lambda) = \det(B - \lambda I_n) = \det(T^{-1} \cdot A \cdot T - \lambda I_n) = \det(T^{-1} \cdot A \cdot T - \lambda T^{-1} \cdot I_n \cdot T) =$$

$$= \det(T^{-1} \cdot (A - \lambda I_n)) \cdot T) = \det(T^{-1}) \cdot \det(A - \lambda I_n) \cdot \det(T) = \frac{1}{\det(T)} \cdot \det(A - \lambda I_n) \cdot \det(T) = p_A(\lambda).$$

□

Věta 6.7. *Pro $A \in \mathbb{F}^{n\times n}$ platí:*

(1)

$$\sum_{\lambda \in \sigma(A)} an_A(\lambda) = n.$$

(2) Pro každé $\lambda \in \sigma(A)$ platí $gn_A(\lambda) \leq an_A(\lambda)$.

Důkaz. Tvrzení (1) plyne z takzvané *základní věty algebry*, která říká, že každý komplexní polynom stupně n má v $\mathbb{C}$ přesně n kořenů (počítáno včetně násobností).
Pro důkaz (2), označme $r = gn_A(\lambda)$. Zvolme bázi $(\vec{b}_1, \dots, \vec{b}_r)$ prostoru $eigen(\lambda, A)$, a doplňme je dle Věty 1.19 na bázi $\mathcal{B} = (\vec{b}_1, \dots, \vec{b}_r, \vec{b}_{r+1}, \dots, \vec{b}_n)$ prostoru $\mathbb{F}^n$. Potom

$$[A]_{\mathcal{B}\mapsto\mathcal{B}} = \big([A\vec{b}_1]_{\mathcal{B}}, \dots, [A\vec{b}_r]_{\mathcal{B}}, [A\vec{b}_{r+1}]_{\mathcal{B}}, \dots, [A\vec{b}_n]_{\mathcal{B}}\big) = \big([\lambda\vec{b}_1]_{\mathcal{B}}, \dots, [\lambda\vec{b}_r]_{\mathcal{B}}, [A\vec{b}_{r+1}]_{\mathcal{B}}, \dots, [A\vec{b}_n]_{\mathcal{B}}\big) =$$
$$= \big(\lambda\vec{e}_1, \dots, \lambda\vec{e}_r, [A\vec{b}_{r+1}]_{\mathcal{B}}, \dots, [A\vec{b}_n]_{\mathcal{B}}\big) = \begin{pmatrix} \lambda I_r & C \\ O_{n-r,r} & D \end{pmatrix},$$

pro nějaké matice $C \in \mathbb{F}^{r\times n-r}$, $D \in \mathbb{F}^{n-r\times n-r}$. Potom, matice $[A]_{\mathcal{B}\mapsto\mathcal{B}}$ je podobná matici $A = [A]_{E_n\mapsto E_n}$ (Tvrzení 1.28), a tedy dle Tvrzení 6.6 platí

$$p_A(x) = \det([A]_{\mathcal{B}\mapsto\mathcal{B}} - xI_n) = \det\begin{pmatrix} (\lambda - x)I_r & C \\ O_{n-r,r} & D - xI_{n-r} \end{pmatrix} =^{1.13} (\lambda - x)^r \cdot \det(D - xI_{n-r}).$$

Tedy $am(\lambda) \geq r$.

□

Cvičení 6.8. *Nechť $V_1, \dots, V_n$ jsou lineárně nezávislé podmnožiny lineárního prostoru W. Přepodkládejme, že každá množina nenulových vektorů $\{\vec{v}_1, \dots, \vec{v}_n\}$ splňující $\vec{v}_i \in V_i$ pro každé $i = 1, \dots, n$, je lineárně nezávislá. Potom $\bigcup_{i=1}^n V_i$ je lineárně nezávislá.*

Tvrzení 6.9. *Nechť $A \in \mathbb{F}^{n\times n}$.*

(1) Každý systém vlastních vektorů $\{\vec{v}_1, \dots, \vec{v}_k\}$ matice A příslušným různým vlastním číslům $\{\lambda_1, \dots, \lambda_k\}$ je lineárně nezávislý.

(2) Pokud pro každé $\lambda \in \sigma(A)$, $\mathcal{B}_\lambda$ je báze $eigen(\lambda, A)$, potom $\bigcup_{\lambda\in\sigma(A)} \mathcal{B}_\lambda$ je lineárně nezávislá podmnožina $\mathbb{F}^n$.

(3) Pro každé $\lambda_1 \neq \lambda_2 \in \sigma(A)$ platí $eigen(\lambda_1, A) \cap eigen(\lambda_2, A) = \vec{o}$.

Důkaz. (1) Dokážeme indukcí. Pro $k = 1$, jelikož $\vec{v_1}$ je nenulový, tudíž $\{\vec{v_1}\}$ je LN.
Ať $\{\lambda_1, \dots, \lambda_{k+1}\}$ jsou různá vlastní čísla matice A a nechť $\{\vec{v}_1, \dots, \vec{v}_{k+1}\}$ jsou příslušné vlastní vektory. Dle indukčního předpokladu je množina $\{\vec{v}_1, \dots, \vec{v}_k\}$ lineárně nezávislá. Ať $\sum_{i=1}^{k+1} \alpha_i\vec{v}_i = \vec{o}$ pro skaláry $\alpha_1, \dots, \alpha_{k+1}$. Potom

$$\vec{o} = (A - \lambda_{k+1}I_n)(\sum_{i=1}^{k+1} \alpha_i\vec{v}_i) = \sum_{i=1}^{k+1} \alpha_i(A - \lambda_{k+1}I_n)\vec{v}_i = \sum_{i=1}^{k+1} \alpha_i(\lambda_i - \lambda_{k+1})\vec{v}_i = \sum_{i=1}^{k} \alpha_i(\lambda_i - \lambda_{k+1})\vec{v}_i.$$

Dle indukčního předpokladu platí pro každé $i = 1, \dots, k$, $\alpha_i(\lambda_i - \lambda_{k+1}) = 0$, a tedy $\alpha_i = 0$. Tedy z rovnosti $\sum_{i=1}^{k+1} \alpha_i\vec{v}_i = \vec{o}$ plyne $\alpha_{k+1}\vec{v}_{k+1} = \vec{o}$, a tedy $\alpha_{k+1} = 0$. Množina $\{\vec{v}_1, \dots, \vec{v}_{k+1}\}$ je tedy lineárně nezávislá.

(2) plyne ihned z (i) a Cvičení 6.8.

(3) plyne z (ii), neboť pokud by existoval nenulový vektor v průniku $eigen(\lambda_1, A) \cap eigen(\lambda_2, A)$, musel by být vygenerovaný bází obou prostorů $eigen(\lambda_1, A), eigen(\lambda_2, A)$ (detaily si rozmyslete). □

Věta 6.10. *Pro matici $A \in \mathbb{F}^{n\times n}$ jsou následující tvrzení ekvivalentní:*

(i) A je diagonalizovatelná,

(ii) existuje báze $\mathcal{B}$ prostoru $\mathbb{F}^n$ složená z vlastních vektorů matice A,

(iii) pro všechna $\lambda \in \sigma(A)$ platí $gn_A(\lambda) = an_A(\lambda)$.

Důkaz. Platí

$$\begin{aligned}(i) &\Leftrightarrow \text{existuje regulární } T \text{ a diagonální } D \text{ tak, že } A\cdot T = T\cdot D\\ &\Leftrightarrow \text{existuje regulární } T \text{ a diagonální } D, \forall j=1,\dots,n\colon A\cdot T_{*j} = [D]_{j,j}\cdot T_{*j}\\ &\Leftrightarrow \text{existuje } n \text{ LN vlastních vektorů matice } A \Leftrightarrow (ii)\end{aligned}$$

(ii) $\Rightarrow$ (iii) Nechť $\sigma(A)=\{\lambda_1,\dots,\lambda_r\}$. Pro $i=1,\dots,r$ označme $k_i = gn_A(\lambda_i)$. Potom, matice A má vůči bázi $\mathcal{B}$ tvar (podobně jako v důkazu Věty 6.7(ii))

$$[A]_{\mathcal{B}\mapsto\mathcal{B}} = Diag(\overbrace{\lambda_1,\dots,\lambda_1}^{k_1 \text{ krát}},\overbrace{\lambda_2,\dots,\lambda_2}^{k_2 \text{ krát}},\dots,\overbrace{\lambda_r,\dots,\lambda_r}^{k_r \text{ krát}}).$$

Tedy

$$p_A(x) =^{6.6} p_{[A]_{\mathcal{B}\mapsto\mathcal{B}}}(x) = (\lambda_1 - x)^{k_1}(\lambda_2 - x)^{k_2}\dots(\lambda_n - x)^{k_n},$$

a (iii) platí.
(iii) $\Rightarrow$ (ii) Nechť pro $\lambda\in\sigma(A)$, $\mathcal{B}_\lambda$ je báze $eigen(\lambda, A)$. Potom dle Tvrzení 6.9(2) je množina $\mathcal{B} = \bigcup_{\lambda\in\sigma(A)}\mathcal{B}_\lambda$ lineárně nezávislá. Navíc, platí

$$|\mathcal{B}| = \sum_{\lambda\in\sigma(A)}|B_\lambda| = \sum_{\lambda\in\sigma(A)} gn_A(\lambda) = \sum_{\lambda\in\sigma(A)} an_A(\lambda) =^{6.7(i)} n.$$

Tedy $\mathcal{B}$ je báze $\mathbb{F}^n$ složená z vlastních vektorů matice A. □

Důsledek 6.11. *Pokud $A\in\mathbb{F}^{n\times n}$ má n různých vlastních čísel, potom je diagonalizovatelná.*

Důkaz. Plyne ihned z Věty 6.10(iii). □

Vlastní čísla nesou řadu důležitých geomerticých, algebraických, i strukturálních informací o matici. Jelikož je ale hledání vlastních čísel výpočetně velmi náročné, hodí se mít nějaké metody, jak pozici vlastních čísel odhadnout. Jeden takový způsob dává následující Věta 6.13.

Definice 6.12. *Nechť $A\in\mathbb{F}^{n\times n}$. Pro $i\in\{1,\dots,n\}$ definujeme*

$$r_i = \sum_{1\le j\le n, j\neq i}|[A]_{i,j}|, \quad K_i = \{z\in\mathbb{C}\colon |z-[A]_{i,i}|\le r_i\},\ a\ \mathcal{G}_A = \bigcup_{i=1}^{n} K_i.$$

K_i se nazývá i-tý Geršgorinův kruh.

Věta 6.13. *(Geršgorin) Nechť $A\in\mathbb{F}^{n\times n}$. Potom $\sigma(A)\subseteq\mathcal{G}_A$. Navíc, je-li $J\subseteq\{1,\dots,n\}$ množina indexů taková, že*

$$\bigcup_{i\in J} K_i \cap \bigcup_{j\in\{1,\dots,n\}\setminus J} K_j = \emptyset,$$

potom $\bigcup_{i\in J} K_i$ obsahuje právě $|J|$ vlastních čísel (počítáno včetně násobností).

Poznámka 6.14.

- *Jelikož $\sigma(A)=\sigma(A^T)$, je $\sigma(A)\subseteq\mathcal{G}_A\cap\mathcal{G}_{A^T}$.*
- *Platí-li pro každé $i=1,\dots,n$,*

$$|[A]_{i,i}| > r_i = \sum_{1\le j\le n, j\neq i}|[A]_{i,j}|,$$

pak A je regulární (protože $0\notin\sigma(A)$). Protože $\sigma(A)=\sigma(A^T)$, platí analogické tvrzení i pro sloupce.

7 Unitární diagonalizace

Definice 7.1.
- *Matice $A, B \in \mathbb{F}^{n\times n}$ jsou si* unitárně podobné*, pokud existuje unitární matice $U \in \mathbb{F}^{n\times n}$ splňující $B = U^* \cdot A \cdot U \Leftrightarrow U \cdot B = A \cdot U \Leftrightarrow U \cdot B \cdot U^* = A$.*
- *Matice $A \in \mathbb{F}^{n\times n}$ je* unitárně diagonalizovatelná*, pokud je unitárně podobná nějaké diagonální matici.*

V této kapitole budeme chtít popsat, které matice jsou unitárně diagonalizovatelné.

Věta 7.2. *(Schurův rozklad) Každá matice $A \in \mathbb{F}^{n\times n}$ je unitárně podobná nějaké horní trojúhelníkové matici $T \in \mathbb{F}^{n\times n}$. Navíc, T má na diagonále vlastní čísla matice A.*

Důkaz. Nechť $\lambda_1 \in \sigma(A)$ a $\vec{x}_1 \in eigen(\lambda_1, A)$, $\|\vec{x}_1\|_2 = 1$. Dle Tvrzení 5.10 existuje unitární matice $X_1 \in \mathbb{C}^{n\times n}$ taková, že $X_1 \cdot \vec{e_1} = \vec{x}_1$, jinými slovy, $X_1 = \begin{pmatrix}\vec{x}_1 & V_1\end{pmatrix}$ pro nějakou matici $V_1 \in \mathbb{C}^{n\times(n-1)}$. Potom

$$A \cdot X_1 = \begin{pmatrix}\lambda_1\vec{x}_1 & A\cdot V_1\end{pmatrix}, \text{ a } X_1^* \cdot A \cdot X_1 = \begin{pmatrix}\vec{x}_1^* \\ V_1^*\end{pmatrix} \cdot \begin{pmatrix}\lambda_1\vec{x}_1 & A\cdot V_1\end{pmatrix} = \begin{pmatrix}\lambda_1 & \vec{x}_1^* \cdot A \cdot V_1 \\ O_{(n-1)\times 1} & V_1^* \cdot A \cdot V_1\end{pmatrix}.$$

Označme $A_1 = V_1^* \cdot A \cdot V_1$ a $U_1 = X_1$, $T_1 = (\lambda_1)$.
V kroku $1 < k < n$ předpokládejme, že

$$U_{k-1}^* \cdot A \cdot U_{k-1} = \begin{pmatrix}T_{k-1} & * \\ O_{(n-k+1)\times(k-1)} & A_{k-1}\end{pmatrix},$$

kde $U_{k-1} \in \mathbb{F}^{n\times n}$ je unitární, a $T_{k-1} \in \mathbb{F}^{(k-1)\times(k-1)}$ je horní trojúhelníková, pro nějakou matici $A_{k-1} \in \mathbb{F}^{(n-k+1)\times(n-k+1)}$.

Nechť $\lambda_k \in \sigma(A_{k-1})$, $\vec{x}_k \in eigen(\lambda_k, A_{k-1})$, $\|\vec{x}_k\|_2 = 1$. Opět dle Tvrzení 5.10 existuje unitární matice $\tilde{X}_k \in \mathbb{C}^{(n-k+1)\times(n-k+1)}$ taková, že $\tilde{X}_k = \begin{pmatrix}\vec{x}_k & V_k\end{pmatrix}$ pro nějakou matici $V_k \in \mathbb{C}^{(n-k+1)\times(n-k)}$. Potom analogicky jako výše máme

$$\tilde{X}_k^* \cdot A_{k-1} \cdot \tilde{X}_k = \begin{pmatrix}\lambda_k & * \\ O_{(n-k)\times 1} & V_k^* \cdot A_{k-1} \cdot V_k\end{pmatrix}.$$

Označme

$$A_k = V_k^* \cdot A_{k-1} \cdot V_k, \quad X_k = \begin{pmatrix}I_{k-1} & O_{(k-1)\times(n-k+1)} \\ O_{(n-k+1)\times(k-1)} & \tilde{X}_k\end{pmatrix}.$$

Pak

$$X_k^* \cdot U_{k-1}^* \cdot A \cdot U_{k-1} \cdot X_k = X_k^* \cdot \begin{pmatrix}T_{k-1} & * \\ O_{(n-k+1)\times(k-1)} & A_{k-1}\end{pmatrix} \cdot X_k = \begin{pmatrix}T_{k-1} & * \\ O_{(n-k+1)\times(k-1)} & \tilde{X}_k^* \cdot A_{k-1} \cdot \tilde{X}_k\end{pmatrix} =$$

$$= \begin{pmatrix}T_{k-1} & * \\ O_{(n-k+1)\times(k-1)} & \begin{pmatrix}\lambda_k & * \\ O_{(n-k)\times 1} & A_k\end{pmatrix}\end{pmatrix} = \begin{pmatrix}T_k & * \\ O_{(n-k)\times k} & A_k\end{pmatrix}$$

Označmíme-li $U_k = U_{k-1} \cdot X_k$, pak U je unitární dle Tvrzení 5.2. Tím končí k-tý krok.

Po $n-1$ krocích dostaneme požadovanou rovnost $U^* \cdot A \cdot U = T$.

Nakonec, jelikož matice A a T jsou si podobné, dle Tvrzení 6.6 platí $\sigma(A) = \sigma(T)$. A, jelikož T je horní trojúhelníková, je snadné si uvědomit, že má na diagonále svá vlastní čísla, tedy i vlastní čísla matice A. □

Tvrzení 7.3.
- *Pokud $U \in \mathbb{F}^{n\times n}$ je unitární matice, potom $\|U\|_2 = 1$.*
- *Pokud $D = Diag(d_1, \ldots, d_n) \in \mathbb{F}^{n\times n}$, potom $\|D\|_2 = \max_{j=1,\ldots,n} |d_j|$.*

Důkaz. Připomeňte si, že

$$\|A\|_2 = \max_{\vec{x}\in\mathbb{F}^n,\vec{x}\neq\vec{o}} \frac{\|A\cdot\vec{x}\|_2}{\|\vec{x}\|_2}.$$

Pokud U je unitární a $\vec{x}\in\mathbb{F}^n$, potom dle Tvrzení 5.2 platí $\|U\vec{x}\|_2 = \|\vec{x}\|_2$, tedy $\|U\|_2 = 1$.

Pokud $D = Diag(d_1,\dots,d_n)$, potom pro $\vec{x}\in\mathbb{F}^n\setminus\{\vec{o}\}$ máme

$$\|D\vec{x}\|_2 = \left\|\sum_{i=1}^n d_i x_i \vec{e_i}\right\|_2 = \sqrt{\sum_{i=1}^n |d_i|^2 |x_i|^2} \leq \sqrt{\max_{j=1,\dots,n}|d_j|^2 \sum_{i=1}^n |x_i|^2} = \max_{j=1,\dots,n}|d_j|\,\|\vec{x}\|_2,$$

tedy

$$\|D\|_2 = \max_{\vec{x}\in\mathbb{F}^n,\vec{x}\neq\vec{o}} \frac{\|D\cdot\vec{x}\|_2}{\|\vec{x}\|_2} \leq \max_{j=1,\dots,n}|d_j|.$$

Na druhou stranu, zvolme $i\in\{1,\dots n\}$ splňující $|d_i| = \max_{j=1,\dots,n}|d_j|$. Potom

$$\|D\vec{e_i}\|_2 = \sqrt{|d_i|^2} = |d_i|,$$

a tedy

$$\|D\|_2 = \max_{\vec{x}\in\mathbb{F}^n,\vec{x}\neq\vec{o}} \frac{\|D\cdot\vec{x}\|_2}{\|\vec{x}\|_2} \geq \frac{\|D\cdot\vec{e_i}\|_2}{\|\vec{e_i}\|_2} = |d_i| = \max_{j=1,\dots,n}|d_j|,$$

čímž je důkaz dokončen. □

Důsledek 7.4. *(Aproximace diagonalizovatelnou maticí) Pro každou matici $A\in\mathbb{F}^{n\times n}$ a libovolné $\epsilon>0$ existuje diagonalizovatelná matice B_ϵ splňující $\|A-B_\epsilon\|_2<\epsilon$.*

Důkaz. Dle Schurova rozkladu (Věta 7.2) nalezneme unitární matici U a horní trojúhelníkovou matici T tak, že $A = U\cdot T\cdot U^*$. Pro dané $\epsilon>0$ nalezneme $\epsilon_1,\dots,\epsilon_n\in\mathbb{R}$ tak, že čísla $[T]_{i,i}+\epsilon_i$, $i=1,\dots,n$, jsou navzájem různá, a $\max_{i=1,\dots,n}|\epsilon_i|<\epsilon$, a definujme matice

$$D_\epsilon = Diag(\epsilon_1,\dots,\epsilon_n),\ \text{a}\ B_\epsilon = U\cdot(T+D_\epsilon)\cdot U^* = A + U\cdot D_\epsilon\cdot U^*.$$

Potom vlastní čísla matice B_ϵ leží na diagonále matice $T+D_\epsilon$, a tedy má matice B n různých vlastních čísel, a je tedy diagonalizovatlená dle Důsledku 6.11. Navíc,

$$\|A-B_\epsilon\|_2 = \|U\cdot D_\epsilon\cdot U^*\|_2 \leq^{(M4)} \|U\|_2\cdot\|D_\epsilon\|_2\cdot\|U^*\|_2 =^{7.3} \max_{i=1,\dots,n}|\epsilon_i| < \epsilon.$$

□

Definice 7.5. *Matice $A\in\mathbb{F}^{n\times n}$ je* normální, *pokud $A\cdot A^* = A^*\cdot A$.*

Tvrzení 7.6. *(i) Následující matice jsou normální:*

- *hermitovské,*
- *antihermitovské,*
- *diagonální,*
- *unitární.*

(ii) Pokud matice A,B jsou si unitárně podobné a A je normální, potom B je normální.

Důkaz. (i) Pokud A je hermitovská, tj. $A = A^*$, potom $A \cdot A^* = A \cdot A = A^* \cdot A$, tj. A je normální. Pro antihermitovké analogicky, a pro diagonální je tvrzení zřejmé.
Pokud A je unitární, potom dle Tvrzení 5.2(iii) platí $A^* = A^{-1}$, a tedy $A \cdot A^* = A \cdot A^{-1} = I_n = A^{-1} \cdot A = A^* \cdot A$.
(ii) Nechť A je normální a $B = U^* \cdot A \cdot U$ pro nějakou unitární matici U. Potom $B^* = (U^* \cdot A \cdot U)^* =^{1.7} U^* \cdot A^* \cdot U$. Tedy

$$B \cdot B^* = U^* \cdot A \cdot U \cdot U^* \cdot A^* \cdot U =^{5.2(iii)} U^* \cdot A \cdot A^* \cdot U = U^* \cdot A^* \cdot A \cdot U =^{5.2(iii)} U^* \cdot A^* \cdot U \cdot U^* \cdot A \cdot U = B^* \cdot B,$$

a B je tedy normální. □

Lemma 7.7. *Nechť $A = (a_{i,j})_{i,j=1}^n \in \mathbb{F}^{n \times n}$ je normální horní trojúhelníková matice. Potom A je diagonální.*

Důkaz. Chceme ukázat, že pro $i = 1, \ldots, n$, i-tý řádek matice A je skalárním násobkem i-tého kanonického vektoru $\vec{e_i}$. Začneme s prvním řádkem. Jelikož A je normální, platí $[A \cdot A^*]_{1,1} = [A^* \cdot A]_{1,1}$, tedy

$$\sum_{j=1}^n |a_{1,j}|^2 = \sum_{j=1}^n a_{1,j} \cdot \overline{a_{1,j}} = [A \cdot A^*]_{1,1} = [A^* \cdot A]_{1,1} = \sum_{i=1}^n \overline{a_{i,1}} \cdot a_{i,1} = \sum_{i=1}^n |a_{i,1}|^2 =^{A \text{ je horní trojúhelníková}} |a_{1,1}|^2 .$$

Tedy

$$|a_{1,1}|^2 = \sum_{j=1}^n |a_{1,j}|^2 = |a_{1,1}|^2 + \sum_{j=2}^n |a_{1,j}|^2 .$$

Odtud plyne, že $a_{1,j} = 0$ pro každé $j > 1$, a tedy první řádek matice A je násobkem $\vec{e_1}$.
Pro ostatní řádky se použije indukce (cvičení). □

Věta 7.8. *Pro matici $A \in \mathbb{F}^{n \times n}$ jsou následující tvrzení ekvivalentní:*

(i) A je unitárně diagonalizovatelná,

(ii) A je normální,

(iii) existuje ON báze $\mathbb{C}^n$ tvořena vlastními vektory matice A.

Důkaz. (i) ⇒ (ii) Dle Tvrzení 7.6(i) je každá diagonální matice normální. A, dle Tvrzení 7.6(ii), pokud je A unitárně podobná normální matici, pak je normální.
(ii) ⇒ (i) Aplikujeme Schurův rozklad (Věta 7.2) na matici A, nalezneme tedy unitární matici U a horní trojúhelníkovou matici T tak, že $A = U^* \cdot T \cdot U$. Potom T je unitárně podobná matici A, tedy je normální dle Tvrzení 7.6(ii), a tedy je diagonální dle Lemma 7.7. A je tedy unitárně diagonalizovatelná.
Nakonec, platí

$$\begin{aligned}(i) &\Leftrightarrow \text{existuje unitární } U \text{ a diagonální } D \text{ tak, že } A \cdot U = U \cdot D \\ &\Leftrightarrow \text{existuje unitární } U \text{ a diagonální } D, \forall j = 1, \ldots, n\colon A \cdot U_{*j} = [D]_{j,j} \cdot U_{*j} \\ &\Leftrightarrow \text{existuje } n \text{ vlastních vektorů matice } A, \text{které tvoří ON množinu v } \mathbb{C}^n \Leftrightarrow (iii).\end{aligned}$$

□

8 Spektrální rozklad

Věta 8.1. *(Spektrální rozklad) Nechť $A \in \mathbb{F}^{n \times n}$ a $\sigma(A) = \{\lambda_1, \ldots, \lambda_k\}$. Potom A je diagonalizovatelná právě tehdy, když existují matice $G_i \in \mathbb{C}^{n \times n}, i = 1, \ldots, k$, s vlastnostmi:*

(i) $A = \sum_{i=1}^k \lambda_i G_i$ (toto vyjádření se nazývá spektrální rozklad *matice A),*

(ii) $I_n = \sum_{i=1}^k G_i$ (rozklad identity),

(iii) $G_i \cdot G_j = O_{n \times n}$ pro každé $i \neq j$,

(iv) pro každé $i \in \{1, \dots, n\}$, G_i je matice projekce na $ker(A - \lambda_i I_n)$ podle $Im(A - \lambda_i I_n)$.

Matice G_i se nazývají spektrální projekce *příslušné matici A.*
Pokud A je unitárně diagonalizovatelná, pak pro každé $i = 1, \dots, k$, $ker(A - \lambda_i I_n)^\perp = Im(A - \lambda I_n)$, a spektrální projekce A jsou ortogonální projekce na $ker(A - \lambda_i I_n)$.

Důkaz. Nechť $\sigma(A) = \{\lambda_1, \dots, \lambda_k\}$, a pro $j = 1, \dots, k$ označme $m_j = gm_A(\lambda_j)$.
Pokud A je diagonalizovatelná, dle Věty 6.10 existuje matice

$$D = Diag(\overbrace{\lambda_1, \dots, \lambda_1}^{m_1 \text{ krát}}, \overbrace{\lambda_2, \dots, \lambda_2}^{m_2 \text{ krát}}, \dots, \overbrace{\lambda_k, \dots, \lambda_k}^{m_j \text{ krát}})$$

a regulární matice $X = \begin{pmatrix} X_1 & X_2 & \dots & X_k \end{pmatrix}$ takové, že pro každé $j \in \{1, \dots, k\}$, $X_j \in \mathbb{C}^{n \times m_j}$, sloupce X_j tvoří bázi $eigen(\lambda_j, A)$, a $A = X \cdot D \cdot X^{-1}$. Označme

$$X^{-1} = Y = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_k \end{pmatrix},$$

kde $Y_j \in \mathbb{C}^{m_j \times n}$ pro každé $j \in \{1, \dots, k\}$.
Dále, definujme $G_i = X_i \cdot Y_i$ pro $i = 1, \dots, k$. Potom

$$A = X \cdot D \cdot X^{-1} = X \cdot D \cdot Y = \sum_{i=1}^{k} \lambda_i \cdot X_i \cdot Y_i = \sum_{i=1}^{k} \lambda_i \cdot G_i,$$

tedy platí (i).
Dále

$$I = X \cdot X^{-1} = X \cdot Y = \sum_{i=1}^{k} X_i \cdot Y_i = \sum_{i=1}^{k} G_i,$$

tedy platí (ii).
Také, jelikož $Y \cdot X = I_n$, dle vlastností blokového násobení matic máme

$$Y_i \cdot X_j = \begin{cases} I_{n_j}, & i = j, \\ O_{n_i \times n_j}, & i \neq j, \end{cases},$$

a tedy

$$G_i \cdot G_j = X_i \cdot Y_i \cdot X_i \cdot Y_i = \begin{cases} G_i, & i = j, \\ O_{n \times n}, & i \neq j, \end{cases},$$

tedy platí (iii), a také, jelikož $G_i^2 = G_i$, je G_i matice projekce (Věta 4.4).
Dále, ukážeme, že $Im(G_i) = ker(A - \lambda_i I_n)$:

$$Im(G_i) = Im(X_i \cdot Y_i) \subseteq Im(X_i) = Im(X_i \cdot I_{n_i}) = Im(X_i \cdot Y_i \cdot X_i) \subseteq Im(X_i \cdot Y_i) = Im(G_i),$$

a tedy $Im(G_i) = Im(X_i) = eigen(\lambda_i, A) = ker(A - \lambda I_n)$, neboť sloupce X_i tvoří bázi $eigen(\lambda_i, A)$.
Dále, chceme ukázat, že $ker(G_i) = Im(A - \lambda I_n)$.
Máme

$$G_i \cdot (A - \lambda_i I_n) = G_i \cdot (\sum_{j=1}^{n} \lambda_j G_j) - \sum_{j=1}^{k} \lambda_i G_j) = G_i \cdot (\sum_{j=1}^{k} (\lambda_j - \lambda_i) G_j) = G_i \cdot (\sum_{j=1, j \neq i}^{k} (\lambda_j - \lambda_i) G_j) = O_{n \times n},$$

tedy $Im(A - \lambda_i I_n) \subseteq ker(G_i)$. Navíc, dle Věty o dimenzi jádra a obrazu (viz 1.23) platí

$$\dim(Im(A - \lambda_i I_n)) = n - \dim(ker(A - \lambda_i I_n) = n - \dim(Im(G_i)) = \dim(ker(G_i)),$$

a tedy $ker(G_i) = Im(A - \lambda I_n)$, tedy platí (iv).
Pokud A je unitárně diagonalizovatelná, je matice X unitární, a tedy $Y = X^{-1} = X^*$. Zvolme $i \in \{1, \dots, k\}$, a označme $U = ker(A - \lambda_i I_n)$. Dle Věty 4.8 platí

$$P_U^{U^\perp} = X_i \cdot X_i^* = X_i \cdot Y_i = G_i = P_U^{Im(A-\lambda_i I_n)}.$$

Tedy G_i je ortogonální projekce na U, a $U^\perp = Im(A - \lambda_i I_n)$.

Nechť naopak existují matice $G_i, i = 1, \dots, k$, vlastnostmi (i),(ii),(iii),(iv). Ukážeme, že A je diagonalizovatelná. Dle (iv) pro každé $i \in \{1, \dots, k\}$ platí

$$\dim(Im(G_i)) = \dim(ker(A - \lambda_i I_n)) = gn_A(\lambda_i).$$

Dále, platí

$$\sum_{i=1}^{k} an_A(\lambda_i) \overset{\text{Věta 6.7(1)}}{=} n = \dim(Im(I_n)) =^{(ii)} \dim(Im(\sum_{i=1}^{k} G_i)) =^{(iii)} \sum_{i=1}^{k} \dim(Im(G_i)) = \sum_{i=1}^{k} gn_A(\lambda_i).$$

Jelikož pro každé i, $gn_A(\lambda_i) \leq an_A(\lambda_i)$ (Věta 6.7(2)), odtud plyne $gn_A(\lambda_i) = an_A(\lambda_i)$, a tedy A je diagonalizovatelná dle Věty 6.10. □

Následující věta říká, že spektrální projekce lze nalézt i bez znalosti vlastních vektorů.

Věta 8.2. *(tvar spektrálních projekcí) Nechť $A \in \mathbb{F}^{n\times n}$ je diagonalizovatelná a $\sigma(A) = \{\lambda_1, \dots, \lambda_k\}$. Potom pro $j \in \{1, \dots, k\}$ a matici spektrální projekce G_j z předchozí věty platí*

$$G_j = \frac{\prod\limits_{1\leq i\leq k, i\neq j} (A - \lambda_i I_n)}{\prod\limits_{1\leq i\leq k, i\neq j} (\lambda_j - \lambda_i)}.$$

Důkaz. Jelikož $A = \sum_{i=1}^{k} \lambda_i G_i$ a $I_n = \sum_{i=1}^{k} G_i$, platí pro $m \in \{1, \dots, k\}$,

$$A - \lambda_m I_n = \sum_{i=1}^{k} \lambda_i G_i - \sum_{i=1}^{k} \lambda_m G_i = \sum_{1\leq i\leq k, i\neq m} (\lambda_i - \lambda_m) G_i.$$

Tedy

$$\prod_{1\leq m\leq k, m\neq j} (A - \lambda_m I_n) = \prod_{1\leq m\leq k, m\neq j} \sum_{1\leq i\leq k, i\neq m} (\lambda_i - \lambda_m) G_i =$$

$$=^{(iii)} \prod_{1\leq m\leq k, m\neq j} (\lambda_j - \lambda_m) G_j^{k-1} = \big(\prod_{1\leq m\leq k, m\neq j} (\lambda_j - \lambda_m) \big) \cdot G_j.$$

Odtud plyne

$$G_j = \frac{\prod\limits_{1\leq m\leq k, m\neq j} (A - \lambda_m I_n)}{\prod\limits_{1\leq m\leq k, m\neq j} (\lambda_j - \lambda_m)},$$

což jsme chtěli ukázat. □

9 Pozitivně definitní matice

Připomeňte si, že matice $A \in \mathbb{F}^{n\times n}$ se nazývá pozitivně definitní, pokud je Hermitovská, a pro každý nenulový vektor $\vec{x} \in \mathbb{F}^n$ platí $\vec{x}^* \cdot A \cdot \vec{x} > 0$. Tyto matice jsou důležité například proto, že charakterizují skalární součiny na $\mathbb{F}^n$ (Věta 3.5). V této kapitole si dokážeme řadu ekvivalentních charakterizací těchto matic. Nejprve ale si bez důkazu ukážeme větu o tvaru koeficientů charakteristického polynomu.

Definice 9.1. *Nechť $A = (a_{i,j})_{i,j=1}^n \in \mathbb{F}^{n\times n}$, $k \in \{1,\dots,n\}$, a $i_1,\dots,i_k \in \{1,\dots,n\}$ jsou různé indexy. Potom skalár* $\det(B)$*, kde B vznikne z matice A vynecháním všech řádků i sloupců s indexy $i_1,\dots,i_k$, se nazývá* hlavní minor/subdeterminant *matice A řádu k. Determinanty matic tvaru $A_k = (a_{i,j})_{i,j=1}^k$ se navývají* vedoucí hlavní minory*.*

Věta 9.2. *Nechť $A \in \mathbb{F}^{n\times n}$ a $p_A(x) = \sum_{j=0}^n c_j \cdot x^j$. Potom pro $k \in \{1,\dots,n\}$ platí*

$$c_k = (-1)^k \cdot \sum\{\textit{hlavní minor } A \textit{ řádu } k\}.$$

Speciálně platí:

(1) $c_n = (-1)^n$,

(2) $c_{n-1} = (-1)^{n-1} \cdot tr(A) = (-1)^{n-1} \cdot \sum_{\lambda\in\sigma(A)} \lambda \cdot an_A(\lambda)$,

(3) $c_0 = \det(A) = \prod_{\lambda\in\sigma(A)} \lambda^{an_A(\lambda)}$.

Příklad 9.3. *Nalezněte charakteristický polynom matice* $A = \begin{pmatrix} 0 & 6 & 3 \\ -1 & 5 & 1 \\ -1 & 2 & 4 \end{pmatrix}$.

Důkaz. Platí:

$$c_0 = \det(A) = 27, \quad c_1 = (-1)\cdot\left(\begin{vmatrix} 5 & 1 \\ 2 & 4 \end{vmatrix} + \begin{vmatrix} 0 & 3 \\ -1 & 4 \end{vmatrix} + \begin{vmatrix} 0 & 6 \\ -1 & 5 \end{vmatrix}\right) = (-1)\cdot(18+3+6) = -27,$$

$$c_2 = tr(A) = 0+5+4 = 9, \quad c_3 = (-1)^3 = -1.$$

Tedy $p_A(x) = -x^3 + 9x^2 - 27x + 27$. □

Definice 9.4. *Matice $A = (a_{i,j})_{i,j=1}^n \in \mathbb{F}^{n\times n}$ se nazývá* pozitivně semidefinitní *(PSD), pokud je Hermitovská, a pro každý vektor $\vec{x} \in \mathbb{F}^n$ platí $\vec{x}^* \cdot A \cdot \vec{x} \geq 0$.*

Poznámka 9.5. *Pozitivně definitní a pozitivně semidefinitní matice se používají mimo jiné pro nalezení extrémů funkcí více proměnných.*

Tvrzení 9.6. *Nechť $A \in \mathbb{F}^{n\times n}$. Potom platí:*

(i) Je-li A Hermitovská, potom $\sigma(A) \subseteq \mathbb{R}$.

(ii) Je-li A antihermitovská, potom vlastní čísla A jsou imaginární nebo 0.

(iii) Je-li A unitární, každé vlastní číslo $\lambda \in \sigma(A)$ splňuje $|\lambda| = 1$.

(iv) Je-li A pozitivně definitní, potom $\sigma(A) \subseteq (0,\infty)$.

(v) Je-li A pozitivně semidefinitní, potom $\sigma(A) \subseteq [0,\infty)$.

Důkaz. Zvolme libovolné $\lambda \in \sigma(A)$ a $\vec{x} \in eigen(\lambda, A), \vec{x} \neq 0$.
(i) Je-li A Hermitovská, potom platí

$$0 = \vec{x}^* \cdot A \cdot \vec{x} - \vec{x}^* \cdot A^* \cdot \vec{x} = \vec{x}^* \cdot A \cdot \vec{x} - (A\cdot\vec{x})^* \cdot \vec{x} = \vec{x}^* \cdot \lambda\vec{x} - \overline{\lambda}\vec{x}^*\vec{x} = (\lambda - \overline{\lambda})\,\|\vec{x}\|_2^2\,.$$

Tedy, jelikož $\|\vec{x}\|_2 \neq 0$, $\lambda = \overline{\lambda}$.
(ii) Analogicky jako (i).
(iii) Pokud A je unitární, pak dle Tvrzení 5.2 platí

$$\|\vec{x}\|_2^2 = \|A \cdot \vec{x}\|_2^2 = (A \cdot \vec{x})^* \cdot A \cdot \vec{x} = \overline{\lambda} \cdot \vec{x}^* \cdot \lambda \cdot \vec{x} = |\lambda|^2 \cdot \|\vec{x}\|_2^2 .$$

Tedy, jelikož $\|\vec{x}\|_2 \neq 0$, $|\lambda| = 1$.
(iv) Pokud A je pozitivně definitní, pak

$$\lambda \|\vec{x}\|_2^2 = \lambda \cdot \vec{x}^* \cdot \vec{x} = \vec{x}^* \cdot \lambda\vec{x} = \vec{x}^* \cdot A \cdot \vec{x} > 0,$$

a tedy $\lambda = \frac{\vec{x}^* \cdot A \cdot \vec{x}}{\|\vec{x}\|_2^2} > 0$.
(v) Analogicky jako (iv). □

Poznámka 9.7. *Množiny Hermitovských, antihermitovských, pozitivně definitních, positivně semidefinitních, a unitárních matic, je někdy užitečné chapat jako maticové analogie následujích podmnožin komplexní roviny:* $\mathbb{R}, i \cdot \mathbb{R}, (0,\infty), [0,\infty), \{\lambda \in \mathbb{C} \colon |\lambda| = 1\}$ *(například, pokud máme rozhodnout, zda nějaká vlastnost platí pro pozitivně definitní matice, je dobré si nejdřív uvědomit, zda platí pro kladná reálná čísla).*

Tvrzení 9.8. *(1) Matice* $A = (a) \in \mathbb{F}^{1\times 1}$ *je pozitivně definitní (resp. pozitivně semidefinitní) právě tehdy, když* $a > 0$ *(resp.* $a \geq 0$*).*

(2) Pokud matice $A \in \mathbb{F}^{n\times n}$ *je pozitivně semidefinitní a* $B \in \mathbb{F}^{n\times n}$ *je libovolná matice, potom matice* $C = B^* \cdot A \cdot B$ *je pozitivně semidefinitní. Pokud* A *je pozitivně definitní a* B *je regulární, potom* C *je pozitivně definitní.*

(3) Pokud $A \in \mathbb{F}^{n\times n}$ *je blokově diagonální a každá z podmatic na její diagonále je pozitivně definitní, potom* A *je pozitivně definitní.*

(4) Každá pozitivně definitní matice je regulární.

(5) Pokud $A = \begin{pmatrix} B & C \\ C^* & D \end{pmatrix}$*, pro nějakou regulární Hermitovskou matici* B*, čtvercovou matici* D*, a nějakou matici* C*, potom existuje matice* M *a blokově diagonální matice* X *s maticí* B *na diagonále takové, že* $\det(M) = 1$ *a* $A = M^* \cdot X \cdot M$*.*

Důkaz. (1), (2), (3) cvičení.
(4) Plyne ihned z toho, že $0 \notin \sigma(A)$ dle Tvrzení 9.6(iv).
(5) Pišme (myšlenka je taková, jako kdybychom používali blokovou Gaussovu eliminaci na vynulování podmatic v matici A, které nejsou na diagonále)

$$A = \begin{pmatrix} B & C \\ C^* & D \end{pmatrix} = \begin{pmatrix} I & O \\ C^* \cdot B^{-1} & I \end{pmatrix} \cdot \begin{pmatrix} B & O \\ O & D - C^* \cdot B^{-1} \cdot C \end{pmatrix} \cdot \begin{pmatrix} I & B^{-1} \cdot C \\ 0 & I \end{pmatrix}.$$

Nyní, stačí označit $M = \begin{pmatrix} I & B^{-1} \cdot C \\ 0 & I \end{pmatrix}$. Potom M je horní trojúhelníková matice s jedničkami na diagonále, tedy $\det(M) = 1$. Dále, jelikož B je Hermitovská, je $(B^{-1})^* = (B^*)^{-1} = B^{-1}$, tedy B^{-1} je také Hermitovská. Tedy platí

$$M^* = \begin{pmatrix} I & B^{-1} \cdot C \\ 0 & I \end{pmatrix}^* = \begin{pmatrix} I & O \\ (B^{-1} \cdot C)^* & I \end{pmatrix} = \begin{pmatrix} I & O \\ C^* \cdot (B^{-1})^* & I \end{pmatrix} = \begin{pmatrix} I & O \\ C^* \cdot B^{-1} & I \end{pmatrix},$$

a tedy $A = M^* \cdot \begin{pmatrix} B & O \\ O & D - C^* \cdot B^{-1} \cdot C \end{pmatrix} \cdot M$. □

Věta 9.9. *Pro matici* $A = (a_{i,j})_{i,j=1}^n \in \mathbb{F}^{n\times n}$ *jsou následující tvrzení ekvivalentní:*

(i) A *je pozitivně definitní,*

(ii) A *je Hermitovská a* $\sigma(A) \subseteq (0, \infty)$*,*

(iii) existuje horní trojúhelníková matice $R \in \mathbb{F}^{n\times n}$ *s kladnými čísly na diagonále taková, že* $A = R^* \cdot R$ *(Choleského rozklad),*

(iv) existuje matice $R \in \mathbb{F}^{m\times n}$*,* $n \leq m$ *s lineárně nezávislými sloupci taková, že* $A = R^* \cdot R$*,*

(v) A *je Hermitovská a všechny její vedoucí hlavní minory jsou kladné (Sylvestrovo kritérium).*

Důkaz. (i) ⇒ (ii) To je dokázáno v Tvrzení 9.6(iv).
(ii) ⇒ (iii) Jelikož A je Hermitovská, je normální (Tvrzení 7.6), a tedy je unitárně diagonalizovatelná (Věta 7.8). Pišme tedy $A = U \cdot D \cdot U^*$, kde U je unitární a D je diagonální matice, jež má na diagonále vlastní čísla $\lambda_1, \ldots, \lambda_n$ matice A, zapsány včetně násobností. Jelikož $\sigma(A) \subseteq (0, \infty)$, můžeme definovat

$$D^{\frac{1}{2}} = Diag(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_n}).$$

Nyní, matice $D^{\frac{1}{2}} \cdot U^*$ je regulární, neboť je součinem dvou regulárních matic, můžeme na ni tedy aplikovat QR-rozklad (Věta 3.15), a napíšeme ji ve tvaru $D^{\frac{1}{2}} \cdot U^* = Q \cdot R$, kde Q je unitární matice, a R je horní trojúhelníková matice s kladnými čísly na diagonále. Celkem tedy

$$A = U \cdot D \cdot U^* = U \cdot D^{\frac{1}{2}} \cdot D^{\frac{1}{2}} \cdot U^* = (D^{\frac{1}{2}} \cdot U^*)^* \cdot D^{\frac{1}{2}} \cdot U^* = R^* \cdot Q^* \cdot Q \cdot R =^{Q \text{ je unitární}} R^* \cdot R.$$

(iii) ⇒ (iv) Plyne ihned z faktu, že horní trojúhelníková matice s kladnými čísly na diagonále je regulární, a regulární matice mají lineárně nezávislé sloupce (Věta 1.26).
(iv) ⇒ (i) Pokud $A = R^* \cdot R$, potom pro $\vec{x} \in \mathbb{F}^n$ máme

$$\vec{x}^* \cdot A \cdot \vec{x} = \vec{x}^* \cdot R^* \cdot R \cdot \vec{x} = (R \cdot \vec{x})^* \cdot (R \cdot \vec{x}) = \|R \cdot \vec{x}\|_2^2 \geq 0.$$

Navíc, jelikož R má lineárně nezávislé sloupce, jedná se o monomorfismus, tj. $ker(R) = \{\vec{o}\}$. Tedy,

$$\vec{x}^* \cdot A \cdot \vec{x} = \|R \cdot \vec{x}\|_2^2 = 0 \Leftrightarrow R \cdot \vec{x} = \vec{o} \Leftrightarrow \vec{x} = 0,$$

a tedy A je pozitivně definitní.
(i) ⇒ (v) Nejprve ukážeme, že z (i) plyne, že každá z matic $A_k = (a_{i,j})_{i,j=1}^k$ je pozitivně definitní. Zvolme tedy libovolné $k \in \{1, \ldots, n-1\}$. Zřejmě je každá z matic A_k Hermitovská. Dále, zvolme libovolný vektor $\vec{y} \in \mathbb{F}^k$, a nechť

$$\vec{x} = \begin{pmatrix} \vec{y} \\ O_{(n-k)\times 1} \end{pmatrix}.$$

Potom platí $\vec{y}^* \cdot A_k \cdot \vec{y} = \vec{x}^* \cdot \vec{A} \cdot \vec{x}$, a tedy pozitivní definitnost A implikuje pozitivní definitnost A_k. Tedy, dle již dokázané implikace (i) ⇒ (ii), platí $\sigma(A_k) \subseteq (0, \infty)$. Tedy, jelikož $\det(A_k)$ je součinem vlastních čísel matice A_k (Věta 9.2), je $\det(A_k) \geq 0$.
(v) ⇒ (i) Důkaz provedeme indukcí podle n. Důkaz pro $n = 1$ plyne z Tvrzení 9.8(1), neboť $a_{1,1} = \det(A) > 0$. Předpokládejme nyní, že tvrzení platí pro $n-1$, a nechť $A \in \mathbb{F}^{n\times n}$ splňuje podmínku (v). Jelikož A je Hermitovská, můžeme ji napsat ve tvaru

$$A = \begin{pmatrix} A_{n-1} & \vec{x} \\ \vec{x}^* & a_{n,n} \end{pmatrix},$$

pro nějakou matici $A_{n-1} \in \mathbb{F}^{(n-1)\times(n-1)}$ a $\vec{x} \in \mathbb{F}^{n-1}$. Potom také matice A_{n-1} má všechny hlavní minory kladné, a tedy je A_{n-1} je pozitivně definitní dle indukčího předpokladu. Dále, jelikož A_{n-1} je Hermitovská a regulární (Tvrzení 9.8(4)), dle Tvrzení 9.8(5) můžeme psát matici A ve tvaru

$$A = M^* \cdot X \cdot M,$$

kde $X = \begin{pmatrix} A_{n-1} & O \\ O & a \end{pmatrix}$ pro nějaký skalár $a \in \mathbb{F}$ a matici $M \in \mathbb{F}^{n\times n}$ s $\det(M) = 1$. Potom máme (na matici X můžeme aplikovat rozvoj dle posledního řádku nebo sloupce, nebo tvrzení o determinantech horní blokové matice)

$$\det(A) = \det(M^*) \cdot \det(X) \cdot \det(M) = \det(A_{n-1}) \cdot a,$$

a tedy $a = \frac{\det(A)}{\det(A_{n-1})} > 0$. Nyní, matice X je pozitivně definitní dle 9.8(3), a tedy matice A je pozitivně definitní dle 9.8(2), čímž je důkaz dokončen. □

Poznámka 9.10. *Choleského rozklad dává velmi rychlý a numericky stabilní způsob řešení soustavy rovnic $(A|\vec{b})$, pro A pozitivně definitní. Funguje tak, že místo přímého řešení soustavy $A \cdot \vec{x} = \vec{b}$ nejdříve vyřešíme soustavu $R^* \cdot \vec{y} = \vec{b}$, a následně soustavu $R \cdot \vec{x} = \vec{y}$. Potom platí $A \cdot \vec{x} = R^* \cdot R \cdot \vec{x} = R^* \vec{y} = \vec{b}$.*

Následující charakterizace pozitivně semidefinitních matice lze dokázat podobně jako Větu 9.9.

Věta 9.11. *Pro matici $A = \mathbb{F}^{n\times n}$ jsou následující tvrzení ekvivalentní:*

(i) A je pozitivně semidefinitní,

(ii) A je Hermitovská a $\sigma(A) \subseteq [0, \infty)$,

(iii) existuje matice $R \in \mathbb{F}^{m\times n}$, $n \leq m$ taková, že $A = R^* \cdot R$ a $rank(R) = rank(A)$,

(iv) A je Hermitovská a všechny její hlavní minory jsou nezáporné.

10 Singulární rozklad (SVD-singular value decomposition)

Definice 10.1. *Matice $A = (a_{i,j}) \in \mathbb{F}^{m\times n}$ je* obdélníková diagonální matice, *pokud $a_{i,j} = 0$ pro $i \neq j$. Píšeme $A = diag_{m\times n}((a_{1,1}, \ldots, a_{r,r})$ pro $r \leq \min(m, n)$. Pokud $r < \min(m, n)$, rozumí se tím, že zbylé diagonální prvky jsou nulové.*

Lemma 10.2. *Pro každou matici $A \in \mathbb{F}^{m\times n}$ platí $rank(A) = rank(A^* \cdot A)$.*

Důkaz. Stačí ukázat, že $\ker(A) = \ker(A^* \cdot A)$, tvrzení potom plyne z Věty o dimenzi jádra a obrazu, jelikož A a $A^* \cdot A$ mají stejné množství sloupců. Pokud $\vec{x} \in \ker(A)$, potom $A^* \cdot A \cdot \vec{x} = A^* \cdot \vec{o} = \vec{o}$, tedy $\vec{x} \in \ker(A^* \cdot A)$. Pokud naopak $\vec{x} \in \ker(A^* \cdot A)$, potom $0 = \vec{x}^* \cdot \vec{o} = \vec{x}^* \cdot A^* \cdot A \cdot \vec{x} = (A \cdot \vec{x})^* \cdot A \cdot \vec{x} = \|A\vec{x}\|_2^2$, a tedy $A\vec{x} = \vec{o}$, a $\vec{x} \in \ker(A)$. □

Věta 10.3. *(Singulární rozklad (SVD-singular value decomposition))*
Nechť $A \in \mathbb{F}^{m\times n}$, $rank(A) = r$. Potom existují ortonormální báze $\mathcal{B} = (\vec{v}_1, \ldots, \vec{v}_n)$, $\mathcal{C} = (\vec{u}_1, \ldots, \vec{u}_m)$ prostorů $\mathbb{F}^n$, respektive $\mathbb{F}^m$, takové, že

(1) $[A]_{\mathcal{B}\mapsto\mathcal{C}} = D = diag_{m\times n}(\sigma_1, \ldots, \sigma_r)$, kde $\sigma_1 \geq \sigma_2 \ldots \geq \sigma_r > 0$ (geometrická verze SVD),

(2) platí

$$A = U \cdot D \cdot V^*,$$

kde U, V jsou unitární matice $U = \begin{pmatrix} \vec{u}_1 & \ldots & \vec{u}_m \end{pmatrix}, V = \begin{pmatrix} \vec{v}_1 & \ldots & \vec{v}_m \end{pmatrix}$ (algebraická verze SVD), a

(3) platí $A = \sum_{i=1}^{r} \sigma_i \vec{u}_i \vec{v}_i^*$ (dyadický rozvoj SVD-zápis matice A jako lineární kombinace r matic hodnosti 1).

Skaláry $\sigma_1, \ldots, \sigma_r$ se nazývají singulární čísla (hodnoty) *matice A, a jsou to odmocniny nenulových vlastních čísel matice $A^* \cdot A$.*

Důkaz. Matice $A^* \cdot A \in \mathbb{F}^{n\times n}$ je dle Věty 9.11 pozitivně semidefinitní, tedy je Hermitovská (tedy normální dle Tvrzení 7.6(i)), a $\sigma(R^* \cdot R) \subseteq [0,\infty)$. Tedy, dle Věty 7.8 existuje ON báze $\mathcal{B} = (\vec{v}_1, \dots, \vec{v}_n)$ prostoru $\mathbb{F}^n$ tvořena vlastními vektory matice $A^* \cdot A$, a platí

$$A^* \cdot A = diag(\lambda_1, \dots, \lambda_n),$$

kde $\lambda_1 \geq \lambda_2 \dots \geq \lambda_n \geq 0$ jsou vlastní čísla $A^* \cdot A$, seřazená podle velikosti, a zapsaná včetně algebraických násobností. Jelikož $rank(A^* \cdot A) = rank(A) = r$ dle Lemma 10.2, jsou čísla λ_i nulová pro $i > r$ (jelikož vektory $\vec{v}_1, \dots, \vec{v}_n$ jsou LN a každý vektor $\vec{v}_i$ příslušný nenulovému vlastnímu číslu λ_i splňuje $\vec{v}_i = \lambda_i^{-1} A^* \cdot A \cdot \vec{v}_i$, a tedy leží v $Im(A^* \cdot A)$).
Pro $i = 1, \dots, r$ označme $\sigma_i = \sqrt{\lambda_i}$ a $\vec{u_i} = \sigma_i^{-1} \cdot A \cdot \vec{v}_i$. Pak pro $i, j \in \{1, \dots, r\}$ platí

$$\vec{u_i}^* \cdot \vec{u_j} = \sigma_i^{-1} \cdot (A \cdot \vec{v}_i)^* \cdot \sigma_j^{-1} \cdot A \cdot \vec{v}_j = \sigma_i^{-1}\sigma_j^{-1}(\vec{v_i}^* \cdot A^* \cdot A \cdot \vec{v}_j) = \sigma_i^{-1}\sigma_j^{-1}(\vec{v_i}^* \cdot \lambda_j \cdot \vec{v}_j) = \sigma_i^{-1}\sigma_j^{-1}\lambda_j(\vec{v_i}^* \cdot \vec{v}_j).$$

Tedy, pro $i \neq j$ jsou na sebe vektory $\vec{u_i}, \vec{u_j}$ kolmé, a $\vec{u_i}^* \cdot \vec{u_i} = \sigma_i^{-1}\sigma_i^{-1}\lambda_i = 1$. Seznam $(\vec{u_1}, \dots, \vec{u_r})$ je tedy ON, a tedy dle Vět 1.19 a 3.13 lze doplnit na ON bázi $(\vec{u_1}, \dots, \vec{u_m})$ prostoru $\mathbb{F}^m$ (dle Steinitzovy věty lze doplnit na bázi, a dle Gram-Schmidta lze tu doplněnou bázi mít ON).
Nyní, pro každé $i \in \{1, \dots, r\}$ platí $A \cdot \vec{v_i} = \sigma_i \vec{u_i}$, tedy $[A \cdot \vec{v_i}]_{\mathcal{C}} = \sigma_i \vec{e_i}$, a pro $i > r$ je

$$\|A \cdot \vec{v_i}\|_2^2 = \vec{v_i}^* \cdot A^* \cdot A \cdot \vec{v_i} = \vec{v_i}^* \cdot \lambda_i \cdot \vec{v_i} = 0,$$

tj. $A \cdot \vec{v_i} = \vec{o}$, a $[A \cdot \vec{v_i}]_C = \vec{o}$. Celkově má tedy matice A vzhledem k bázím $\mathcal{B}$ a $\mathcal{C}$ skutečně tvar

$$[A]_{\mathcal{B}\mapsto\mathcal{C}} = ([A \cdot \vec{v_1}]_{\mathcal{C}}, \dots, [A \cdot \vec{v_n}]_{\mathcal{C}}) = diag_{m\times n}(\sigma_1, \dots, \sigma_r) =: D,$$

tedy platí (1). Odtud plyne

$$A = [A]_{E_n\mapsto E_m} = T_{\mathcal{C}\mapsto E_m} \cdot [A]_{\mathcal{B}\mapsto\mathcal{C}} \cdot T_{E_n\mapsto\mathcal{B}} = T_{\mathcal{C}\mapsto E_m} \cdot [A]_{\mathcal{B}\mapsto\mathcal{C}} \cdot T^{-1}_{\mathcal{B}\mapsto E_n} = U \cdot D \cdot V^{-1} = U \cdot D \cdot V^*,$$

tedy platí (2). Bod (3) je jenom rozepsání rovnosti $A = (U \cdot D) \cdot V^*$ pomocí blokového násobení matic, kde matice $U \cdot D$ je zapsána blokově pomocí sloupců, a V^* je zapsána blokově pomocí řádků (podobně jako u spektrálního rozkladu-Věta 8.1). □

Singulární rozklad matice má řadu zajímavých teoretických i praktických aplikací, některé z nich si nyní ukážeme.

Nejprve si ukažme, že singulární hodnoty umožňují spočítat spektrální normu a číslo podmíněnosti matice:

Tvrzení 10.4. *Nechť $A \in \mathbb{F}^{n\times n}$, a $\sigma_1 \geq \sigma_2 \dots \geq \sigma_r > 0$ jsou singulární čísla A. Potom platí:*

(1) $\|A\|_2 = \sigma_1$,

(2) pokud A je regulární, potom $\left\|A^{-1}\right\|_2 = \frac{1}{\sigma_n}$,

(3) pokud A je regulární, potom $\kappa_2(A) = \frac{\sigma_1}{\sigma_n}$.

Důkaz. (1) Dle Věty 10.3, nechť $A = U \cdot D \cdot V^*$ je singulární rozklad matice A, a nechť $\vec{x} \in \mathbb{F}^n, \vec{x} \neq \vec{o}$. Označme

$$V^* \cdot \vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Potom, jelikož V^* je unitární matice, platí

$$\|\vec{x}\|_2 =^{5.2} \|V^* \cdot \vec{x}\|_2 = \sqrt{\sum_{i=1}^{n} |x_i|^2}.$$

Dále, platí

$$\|A\vec{x}\|_2 = \|U \cdot D \cdot V^*\vec{x}\|_2 =^{5.2} \|D \cdot V^*\vec{x}\|_2 = \sqrt{\sum_{i=1}^{r} \sigma_i^2 |x_i|^2} \leq \sqrt{\sum_{i=1}^{r} \sigma_1^2 |x_i|^2} \leq \sqrt{\sum_{i=1}^{n} \sigma_1^2 |x_i|^2} = \sigma_1 \|\vec{x}\|_2,$$

přičemž, pokud $x_i = 0$ pro $i > 1$, nastává rovnost. Tedy

$$\|A\|_2 = \max_{\vec{x}\in\mathbb{F}^n, \vec{x}\neq\vec{o}} \frac{\|A \cdot \vec{x}\|_2}{\|\vec{x}\|_2} \leq \max_{\vec{x}\in\mathbb{F}^n, \vec{x}\neq\vec{o}} \frac{\sigma_1 \|\vec{x}\|_2}{\|\vec{x}\|_2} = \sigma_1,$$

a rovnost nastane pro některé věktory $\vec{x}$.
(2) lze dokázat podobně jako (1), a (3) plyne ihned kombinací (1) a (2), neboť $\kappa_2(A) = \|A\|_2 \cdot \left\|A^{-1}\right\|_2$. □

Věta 10.5. *(aproximace maticí nižší hodnosti) Nechť $A \in \mathbb{F}^{m\times n}$, $rank(A) = r$. Potom pro každé $k \in \{1, \ldots, r-1\}$ platí*

$$\min\{\|A - B\|_2 : B \in \mathbb{F}^{m\times n}, rank(B) \leq k\} = \sigma_{k+1}.$$

Navíc, tohoto minima se nabývá pro matici $A_k = U \cdot diag_{m\times n}(\sigma_1, \ldots, \sigma_k) \cdot V^*$, kde $A = U \cdot D \cdot V^*$ je singulární rozklad matice A. Matice A_k se nazývá zkrácený SVD rozklad *matice A.*

Důkaz. Na jednu stranu, pro matici $A_k = U \cdot diag_{m\times n}(\sigma_1, \ldots, \sigma_k) \cdot V^*$ platí

$$\|A - A_k\|_2 = \|U \cdot diag_{m\times n}(0, \ldots, 0, \sigma_{k+1}, \ldots, \sigma_r) \cdot V^*\|_2 = \sigma_{k+1},$$

dle Tvrzení 10.4(1), neboť největší singulární číslo matice $A - A_k$ je σ_{k+1}. Zbývá ukázat, že pro každou matici $B \in \mathbb{F}^{m\times n}$ hodnosti nejvýše k je $\|A - B\|_2 \geq \sigma_{k+1}$. Nechť tedy B je taková matice. Potom $\dim(ker(B)) = def(B) = n - rank(B) \geq n - k$. Dále, jelikož V má LN sloupce, platí $dim(span(\vec{v_1}, \ldots, \vec{v}_{k+1})) = k+1$, a tedy, jelikož součet dimenzí těchto dvou podprostorů je větší než n, tyto dva podprostory se protínají, můžeme tedy zvolit nenulový vektor $\vec{x}$ v průniku těchto dvou podprostorů. Pišme $\vec{x} = \sum_{i=1}^{k+1} x_i\vec{v_i}$. Potom platí

$$\|\vec{x}\|_2 = \|V^* \cdot \vec{x}\|_2 = \left\|V^* \cdot \sum_{i=1}^{k+1} x_i\vec{v_i}\right\|_2 = \left\|\sum_{i=1}^{k+1} x_i V^* \cdot \vec{v_i}\right\|_2 =^{V \text{ má ON sloupce}} \left\|\sum_{i=1}^{k+1} x_i\vec{e_i}\right\|_2,$$

a

$$\|(A - B)\vec{x}\|_2 =^{\vec{x}\in\ker(B)} \|A\vec{x}\|_2 = \left\|U \cdot D \cdot V^* \cdot \sum_{i=1}^{k+1} x_i\vec{v_i}\right\|_2 =^{5.2} \left\|D \cdot \sum_{i=1}^{k+1} x_i V^*\vec{v_i}\right\|_2 =^{V \text{ má ON sloupce}}$$

$$= \left\|D \cdot \sum_{i=1}^{k+1} x_i\vec{e_i}\right\|_2 = \left\|\sum_{i=1}^{k+1} \sigma_i x_i\vec{e_i}\right\|_2 \geq \left\|\sum_{i=1}^{k+1} \sigma_{k+1} x_i\vec{e_i}\right\|_2 = \sigma_{k+1} \|\vec{x}\|_2.$$

Tedy

$$\|A - B\|_2 \geq \frac{\|(A - B)\vec{x}\|_2}{\|\vec{x}\|_2} \geq \sigma_{k+1}.$$

□

Poznámka 10.6. *Lze ukázat, že matice A_k z předchozí věty je nejlepší aproximace matice A o hodnosti nejvýše k i vzhledem k Frobeniově normě.*

Definice 10.7. *Nechť $A \in \mathbb{F}^{m\times n}, rank(A) = r$. Potom* Mooreova-Penroseova pseudoinverze *matice A je matice $A^+ \in \mathbb{F}^{n\times m}$, definována rovností $A^+ = V \cdot diag_{n\times m}(\sigma_1^{-1}, \ldots, \sigma_r^{-1}) \cdot U^*$, kde $A = U \cdot diag_{m\times n}(\sigma_1, \ldots, \sigma_r) \cdot V^*$ je singulární rozklad matice A.*

Pseudoinverze matice má řadu zajímavých vlastností, které si uvedeme bez důkazu. Bohužel je ale pseudoinverze také numericky velmi nestabilní, takže její význam je především teoretický.

Věta 10.8. *Nechť $A \in \mathbb{F}^{m\times n}$.*

(1) Pseudoinverze matice A je ekvivalentně definována pomocí následujících čtyř vlastností:

(i) $A \cdot A^+ \cdot A = A$,

(ii) $A^+ \cdot A \cdot A^+ = A^+$,

(iii) $(A^+ \cdot A)^* = A^+ \cdot A$,

(iv) $(A \cdot A^+)^* = A \cdot A^+$.

(2) Pokud A je regulární, potom $A^+ = A^{-1}$.

(3) Pokud soustava $(A|\vec{b})$ má řešení, potom $\vec{x} = A^+\vec{b}$ je řešení této soustavy s minimální ℓ_2 normou.

(4) Pokud soustava $(A|\vec{b})$ nemá řešení, potom $\vec{x} = A^+\vec{b}$ je přibližné řešení této soustavy metodou nejmenších čtverců s minimální ℓ_2 normou.

10.1 Geometrický význam SVD

Singulární rozklad matice říká, že libovolné lineární zobrazení $\mathbb{F}^n \to \mathbb{F}^n$, reprezentované maticí $A = U \cdot D \cdot V^* \in \mathbb{F}^{m\times n}$, lze reprezentovat jako kombinaci následujících kroků:

(i) isometrie prostoru $\mathbb{F}^n$ (tj. například rotace, reflexe, středová souměrnost, ...), daná maticí V^*,

(ii) změna měřítka ve směru souřadnicových os (s případnou degenerací některých os), daná maticí D,

(iii) isometrie prostoru $\mathbb{F}^m$, daná maticí U.

Pomocí singlární rozkladu lze také popsat obraz jednotkové sféry

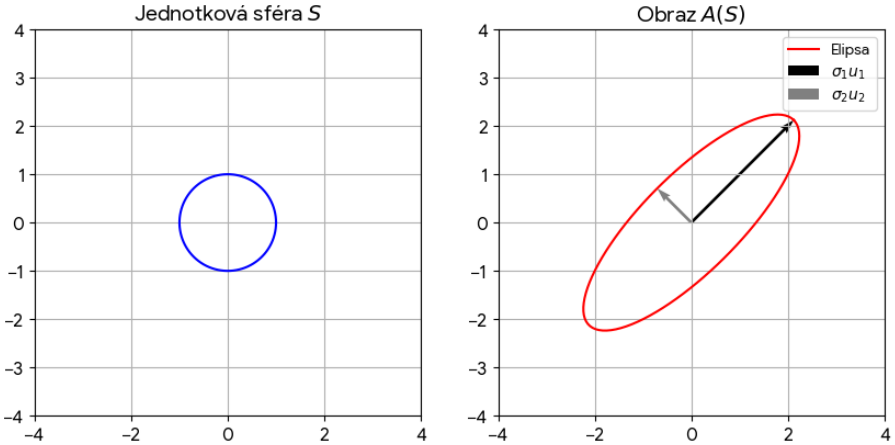
$$S_n = \{\vec{x} \in \mathbb{F}^n \colon \|\vec{x}\|_2 = 1\}$$

při zobrazení A. Zobrazení V^* je unitární, tedy isometrie pro ℓ_2 normu, a tedy $V^*(S_n) = S_n$, přičemž zobrazí vektor $\vec{v_i}$ na $\vec{e_i}$. Zobrazení D potom množinu S_n natáhne či smrští ve směru souřadnicových os (případně ještě ubere či přidá nějaké složky, v závislosti na $rank(A)$), tím vznikne zobecněný elipsoid (tj. dva body pro $r = 1$, elipsa pro $r = 2$, elipsoid pro $r = 3$, ...) s poloosami $\sigma_1\vec{e_1}, \ldots, \sigma_r\vec{e_r}$. Nakonec, zobrazení U převede isometricky tento elipsoid (tj. se zachováním délek vektorů), na elipsoid s poloosami $\sigma_1\vec{u_1}, \ldots, \sigma_r\vec{u_r}$. Pro regulární matici $A \in \mathbb{F}^{n\times n}$ je míra deformace jednotkové sféry měřena číslem $\kappa_2(A) = \frac{\sigma_1}{\sigma_n}$.

Příklad 10.9. *Matice* $\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ *má SVD rozklad*

$$A = U \cdot D \cdot V^* = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix} \cdot \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}.$$

Obraz jednotkové sféry je uveden na obrázku níže. Obrazem je elipsa, sloupce matice U udávají směry jejích poloos, singulární čísla udávají velikosti poloos.



Nyní, shrneme rozdíly mezi singulárními a vlastními čísly matice A:

- Singulární čísla existují vždy reálná a kladná pro každou (i obdélníkovou, i komplexní) matici, lze z nich odvodit celkovou deformaci jednotkové sféry zobrazením A, a tedy i spektrální normu, tj. berou v potaz změnu velikostí všech vektorů zobrazením A,
- Vlastní čísla existují pouze pro čtvercovou matici, mohou být komplexní i pro reálnou matici, udávají změnu velikosti vektorů pouze ve směrech, které zobrazení A zachovává, ingorují změnu velikostí vektorů, kterým zobrazení A mění směr, umožňují identifikovat směry, které matice zachovává, a tím pádem matici umocnit (pomocí diagonalizace nebo Jordanova tvaru matice, který probereme v příštích kapitolách).

10.2 SVD rozklad, komprese dat a filtrování šumu

Definice 10.10. *Nechť $A \in \mathbb{F}^{m\times n}$ je matice hodnosti r, $A = \sum_{i=1}^{r} \sigma_i \cdot \vec{u}_i \cdot \vec{v}_i^*$ Poměl*

$$\frac{\|A - A_k\|_F}{\|A\|_F} = \sqrt{\frac{\sigma_{k+1}^2 + \dots \sigma_r^2}{\sigma_1^2 + \dots \sigma_r^2}},$$

kde

$$\|A\|_F = \Big(\sum_{i=1}^{m}\sum_{j=1}^{n} |[A]_{i,j}|^2\Big)^{\frac{1}{2}} = \sigma_1^2 + \dots \sigma_r^2 \quad \textit{(cvičení)},$$

nazýváme relativní změna matice A při nahrazení maticí A_k.

V praxi, pokud $A \in \mathbb{F}^{m\times n}$ je matice hodnosti r nesoucí nějaká data, pak často bývá většina informace nesená maticí A_k (tj., relativní změna matice A při nahrazení maticí A_k je malá) pro k podstatně menší než r. Pokud tedy uchováme matici A_k, zachováme většinu informace o datech, přičemž potřebujeme podstatně méně paměti (všimněte si, že pro matici A potřebujeme znát $m \cdot n$ skalárů, zatímco pro $A_k = \sum_{i=1}^{k} \sigma_i \vec{u}_i \vec{v}_i^*$ $k \cdot (m + n + 1)$ skalárů).

Podobně lze SVD rozklad využít k odstranění náhodného šumu, například v obrázku nebo zvukovém signálu. Nechť $A \in \mathbb{F}^{m\times n}, rank(A) = r$, je matice nesoucí data znehodnocena nějakým náhodným šumem. SVD rozklad data rozloží na r nezávislých komponent $A = \sum_{i=1}^{r} \sigma_i Z_i$, kde $Z_i = \vec{u}_i \vec{v}_i^*$. Pokud je šum náhodný, každé $\sigma_i Z_i$ ho obsahuje přibližně stejné množství. Pokud jsou tedy hodnoty $\sigma_{k+1}, \dots, \sigma_r$ dost malé pro nějaké k, které je podstatně menší než r, může matice $A_k = \sum_{i=1}^{k} \sigma_i Z_i$ nést většinu informace z matice A, ale být zatížena malým množstvím šumu.

11 Jordanův tvar matice

V této kapitole probereme Jordanův tvar matice, jinými slovy, ukážeme, že každá čtvercová matice je podobná matici, která je "skoro diagonální". To umožňuje, mimo jiné, umocnit obecnou čtvercovou matici. Jordanův tvar lze tedy vnímat jako náhradu diagonalizace pro matice, které diagonalizovatelné nejsou.

11.1 Jordanův tvar a mocnina matice

Definice 11.1. Jordanova buňka *řádu (nebo velikosti)* $k \geq 1$ *příslušná prvku* $\lambda \in \mathbb{C}$ *je matice* $J_{\lambda,k} \in \mathbb{F}^{k\times k}$, *která má na diagonále* λ, *nad diagonálou* 1, *a všude jinde* 0, *tj.*

$$J_{\lambda,k} = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda & 1 & \dots & 0 & 0 \\ 0 & 0 & \lambda & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \lambda & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{pmatrix}.$$

Příklad 11.2. *Matice* $\begin{pmatrix} 5 & 1 \\ 0 & 5 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} i & 1 & 0 \\ 0 & i & 1 \\ 0 & 0 & i \end{pmatrix}, \begin{pmatrix} -8 \end{pmatrix}$ *jsou Jordanovy buňky* $J_{5,2}, J_{0,2}, J_{i,3}, J_{-8,1}$.

Definice 11.3. *Matice* $A \in \mathbb{F}^{n\times n}$ *je v* Jordanově kanonickém tvaru, *pokud je to blokově diagonální matice, a každý z jejích bloků je Jordanova buňka, tj. existují* $\lambda_1, \dots, \lambda_s \in \mathbb{C}$ *a* $k_1, \dots, k_s \in \mathbb{N}$ *takové, že*

$$A = diag(J_{\lambda_1,k_1}, \dots, J_{\lambda_s,k_s}) = \begin{pmatrix} J_{\lambda_1,k_1} & 0 & \cdots & 0 \\ 0 & J_{\lambda_2,k_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_{\lambda_s,k_s} \end{pmatrix}$$

Všimněte si, že matice je diagonální právě tehdy, když je v Jordanově tvaru a každá z jejích Jordanových buněk na diagonále má řád 1.

Příklad 11.4. *Matice*

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

je v Jordanově tvaru. Je složená z Jordanových buněk $J_{0,2}, J_{0,1}, J_{-2,3}$.

Nyní si ukážeme, jak umocnit matici v Jordanově tvaru. Jelikož mocnění blokově diagonálních matic znamená mocnění matic na diagonále, tj. pro

$$A = diag(J_{\lambda_1,k_1}, \dots, J_{\lambda_s,k_s}) = \begin{pmatrix} J_{\lambda_1,k_1} & 0 & \cdots & 0 \\ 0 & J_{\lambda_2,k_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_{\lambda_s,k_s} \end{pmatrix}$$

platí

$$A^m = diag(J^m_{\lambda_1,k_1}, \dots, J^m_{\lambda_s,k_s}) = \begin{pmatrix} J^m_{\lambda_1,k_1} & 0 & \cdots & 0 \\ 0 & J^m_{\lambda_2,k_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J^m_{\lambda_s,k_s} \end{pmatrix},$$

stačí nám naučit se mocnit Jordanovy buňky. Základním krokem je naučit se to pro Jordanovy buňky příslušné číslu 0. Aby se nám snáze pochopil důkaz následujícího tvrzení, nejdřív si mocnění vyzkoušíme na konkrétním příkladu.

Příklad 11.5. *Platí*

$$J_{0,4} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad J_{0,4}^2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad J_{0,4} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad J_{0,4} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Tvrzení 11.6. *Pro $k, m \in \mathbb{N}$ platí*

$$J_{0,k}^m = \begin{cases} \begin{pmatrix} \vec{e}_{m+1}^T \\ \vdots \\ \vec{e}_k^T \\ \vec{o}^T \\ \vdots \\ \vec{o}^T \end{pmatrix} = \begin{pmatrix} \vec{o} & \dots & \vec{o} & \vec{e}_1 & \vec{e}_2 & \dots & \vec{e}_{k-m} \end{pmatrix}, & m < k, \\ O_{k\times k}, & m \geq k. \end{cases}$$

Důkaz. Provedeme důkaz indukcí podle m.

Pro $m = 1$ tvrzení platí, neboť $J_{0,k} = \begin{pmatrix} \vec{e}_2^T \\ \vec{e}_3^T \\ \vdots \\ \vec{e}_{k-1}^T \\ \vec{o}^T \end{pmatrix} = \begin{pmatrix} \vec{o} & \vec{e}_1 & \vec{e}_2 & \dots & \vec{e}_{k-1} \end{pmatrix}$.

Platí-li tvrzení pro nějaké $m < k$, potom

$$J_{0,k}^{m+1} = J_{0,k}^m \cdot J_{0,k} = \begin{pmatrix} \vec{e}_{m+1}^T \\ \vdots \\ \vec{e}_k^T \\ \vec{o}^T \\ \\ \vdots \\ \vec{o}^T \end{pmatrix} \cdot \begin{pmatrix} \vec{o} & \vec{e}_1 & \vec{e}_2 & \dots & \vec{e}_{k-1} \end{pmatrix} =$$

$$= \begin{cases} \begin{pmatrix} \vec{o} & \dots & \vec{o} & \vec{e}_1 & \vec{e}_2 & \dots & \vec{e}_{k-m-1} \end{pmatrix}, & m+1 < k, \\ O_{k\times k}, & m+1 \geq k. \end{cases}$$

Pro $m \geq k$ je indukční krok snadný, neboť z výše uvedeného plyne, že $J_{0,k}^k = O_{k\times k}$. □

Nyní, abychom umocnili obecnou Jordanovu buňku, bude se nám hodit následující binomická věta pro matice.

Věta 11.7. *Pokud $A, B \in \mathbb{F}^{n\times n}$ spolu* komutují, *tj. $A \cdot B = B \cdot A$, potom pro každé $m \in \mathbb{N}$ platí*

$$(A+B)^m = \sum_{j=0}^{m} \binom{m}{j} \cdot A^{m-j} \cdot B^j.$$

Důkaz. Cvičení (trochu těžší). Lze dokázat indukcí podle m skoro stejně jako binomickou větu pro skaláry. □

V následujícím tvrzení používáme konvenci, že $\binom{m}{j} = 0$ pro $m < j$.

Tvrzení 11.8. *Pro $\lambda \in \mathbb{C} \setminus \{0\}$ a $k, m \in \mathbb{N}$ platí*

$$J^m_{\lambda,k} = \begin{pmatrix} \lambda^m & \binom{m}{1}\lambda^{m-1} & \binom{m}{2}\lambda^{m-2} & \cdots & \binom{m}{k-1}\lambda^{m-k+1} \\ 0 & \lambda^m & \binom{m}{1}\lambda^{m-1} & \cdots & \binom{m}{k-2}\lambda^{m-k+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda^m & \binom{m}{1}\lambda^{m-1} \\ 0 & 0 & \cdots & 0 & \lambda^m \end{pmatrix}$$

Důkaz. Důkaz plyne aplikací Věty 11.7 na matici $J_{\lambda,k} = \lambda \cdot I_k + J_{0,k}$, a použitím Tvrzení 11.6 (všimněte si, že matice $\lambda \cdot I_k$ a $J_{0,k}$ spolu komutují, protože matice $\lambda \cdot I_k$, jakožto skalární násobek jednotkové matice, komutuje se vším. □

Příklad 11.9. *Platí*

$$J_{3,4} = 3 \cdot I_4 + J_{0,4} = \begin{pmatrix} 3 & 1 & 0 & 0 \\ 0 & 3 & 1 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 3 \end{pmatrix},$$

$$J^2_{3,4} = 9 \cdot I_4^2 + 2 \cdot 3 \cdot I_n \cdot J_{0,4} + J^2_{0,4} = 9 \cdot I_4 + 6 \cdot J_{0,4} + J^2_{0,4} = \begin{pmatrix} 9 & 6 & 1 & 0 \\ 0 & 9 & 6 & 1 \\ 0 & 0 & 9 & 6 \\ 0 & 0 & 0 & 9 \end{pmatrix},$$

$$J^3_{3,4} = 27 \cdot I_4^3 + 3 \cdot 3^2 \cdot I_n^2 \cdot J_{0,4} + 3 \cdot 3 \cdot I_n \cdot J^2_{0,4} + J^3_{0,4} = 27 \cdot I_4 + 27 \cdot J_{0,4} + 9 \cdot J^2_{0,4} + J^3_{0,4} = \begin{pmatrix} 27 & 27 & 9 & 1 \\ 0 & 27 & 27 & 9 \\ 0 & 0 & 27 & 27 \\ 0 & 0 & 0 & 27 \end{pmatrix},$$

a pro $m > 3$,

$$J^m_{3,4} = \begin{pmatrix} 3^m & m3^{m-1} & \frac{m(m-1)}{2}3^{m-2} & \frac{m(m-1)(m-2)}{6}3^{m-3} \\ 0 & 3^m & m3^{m-1} & \frac{m(m-1)}{2}3^{m-2} \\ 0 & 0 & 3^m & m3^{m-1} \\ 0 & 0 & 0 & 3^m \end{pmatrix}.$$

Víme, že matice $A \in \mathbb{F}^{n \times n}$ je diagonalizovatelná právě tehdy, když existuje báze prostoru $\mathbb{F}^n$ tvořena vlastními vektory matice A (Věta 6.10). Pokud matice diagonalizovatelná není, převedeme ji do Jordanova tvaru tak, že místo toho budeme hledat takzvanou Jordanovu bázi prostoru $\mathbb{F}^n$.

Definice 11.10. *Báze $\mathcal{B}$ prostoru $\mathbb{F}^n$ se nazývá* Jordanova báze *matice $A \in \mathbb{F}^{n \times n}$, pokud $[A]_{\mathcal{B} \mapsto \mathcal{B}}$ je v Jordanově kanonickém tvaru.*

Nejprve si odvodíme, kdy má matice $A \in \mathbb{F}^{n \times n}$ vůči nějaké bázi $\mathcal{B} = (\vec{v}_1, \ldots, \vec{v}_n)$ Jordanův tvar, pro jednoduchost začneme se samotnou Jordanovou buňkou:

Pokud $[A]_{\mathcal{B} \mapsto \mathcal{B}} = J_{\lambda,n}$, potom

$$[A \cdot \vec{v}_1]_B = \begin{pmatrix} \lambda \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad [A \cdot \vec{v}_2]_B = \begin{pmatrix} 1 \\ \lambda \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \cdots, \quad [A \cdot \vec{v}_n]_B = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \lambda \end{pmatrix},$$

neboli

$$A \cdot \vec{v}_1 = \lambda \cdot \vec{v}_1, \quad A \cdot \vec{v}_2 = \lambda \vec{v}_2 + \vec{v}_1, \quad \cdots, \quad A \cdot \vec{v}_n = \lambda \cdot \vec{v}_n + \vec{v}_{n-1},$$

to jest

$$(A - \lambda I_n) \cdot \vec{v}_1 = \vec{o}, \quad (A - \lambda I_n) \cdot \vec{v}_2 = \vec{v}_1, \quad \cdots, \quad (A - \lambda I_n) \cdot \vec{v}_n = \vec{v}_{n-1},$$

schematicky

$$\vec{v}_n \overset{A-\lambda I_n}{\longrightarrow} \vec{v}_{n-1} \overset{A-\lambda I_n}{\longrightarrow} \cdots \overset{A-\lambda I_n}{\longrightarrow} \vec{v}_2 \overset{A-\lambda I_n}{\longrightarrow} \vec{v}_1 \overset{A-\lambda I_n}{\longrightarrow} \vec{o}.$$

Vidíme tedy, že λ je vlastní číslo matice A, a $\vec{v}_1$ je příslušný vlastní vektor.

Definice 11.11. *Je-li $A \in \mathbb{F}^{n\times n}$ a $\lambda \in \sigma(A)$, potom posloupnost vektorů $(\vec{v}_1, \ldots, \vec{v}_k)$ nazveme* Jordanovým řetězcem matice A délky k příslušný vlastnímu číslu λ s počátkem $\vec{v}_1$, *pokud platí $(A - \lambda I_n) \cdot \vec{v}_1 = \vec{o}$ a pro $1 < k \leq n$, $(A - \lambda I_n) \cdot \vec{v}_k = \vec{v}_{k-1}$. Vektorům $\vec{v}_i, i = 1, \ldots, k$ se říká* zobecněné vlastní vektory matice A.

Před touto definicí jsme tedy odvodili následující tvrzení.

Tvrzení 11.12. *Pokud $A \in \mathbb{F}^{n\times n}$ a $\mathcal{B} = (\vec{v}_1, \ldots, \vec{v}_n)$ je báze $\mathbb{F}^n$, potom $[A]_{\mathcal{B}\mapsto\mathcal{B}} = J_{\lambda,k}$ právě, tehdy $(\vec{v}_1, \ldots, \vec{v}_n)$ je Jordanův řetězec matice A délky n příslušný vlastnímu číslu λ.*

Odtud se už snadno dostane tvrzení pro obecný Jordanův tvar matice.

Tvrzení 11.13. *Pokud $A \in \mathbb{F}^{n\times n}$ a $\mathcal{B} = (\vec{v}_1, \ldots, \vec{v}_n)$ je báze $\mathbb{F}^n$, potom $[A]_{\mathcal{B}\mapsto\mathcal{B}} = diag(J_{\lambda_1,k_1}, \cdots, J_{\lambda_s,k_s})$ právě tehdy, když $\mathcal{B}$ je spojením s posloupností $\mathcal{B}_1, \ldots, \mathcal{B}_s$, kde pro každé $i = 1, \ldots, s$, $\mathcal{B}_i$ je Jordanův řetězec matice A délky k_i příslušný vlastnímu číslu λ_i.*
Tedy: báze $\mathcal{B}$ prostoru $\mathbb{F}^n$ je Jordanova báze A právě tehdy, když je tvořena zobecněnými vlastními vektory matice A.

Nyní budeme chtít ukázat, že pro každou komplexní čtvercovou matici existuje Jordanova báze, tedy, jinak řečeno, báze složená ze zobecněných vlastních vektorů této matice. Základem je nalézt tuto bázi pro nilpotentní matici, neboť, jak ukážeme později, každou matici lze rozložit na regulární a nilpotentní část. Nejprve si ale ukažme, že vektory z Jordanova řetězce jsou vždy lineárně nezávislé.

Tvrzení 11.14. *Pokud $A \in \mathbb{F}^{n\times n}$ a $(\vec{v}_1, \ldots, \vec{v}_k)$ je Jordanův řetězec matice A příslušný vlastnímu číslu λ takový, že $\vec{v}_1 \neq \vec{o}$, potom je seznam $(\vec{v}_1, \ldots, \vec{v}_k)$ lineárně nezávislý.*

Důkaz. Předpokládejme, že $c_1, \ldots, c_k \in \mathbb{F}$ jsou skaláry takové, že $\sum_{i=1}^{k} c_i \vec{v}_i = \vec{o}$. Aplikujeme na tuto rovnost matici $(A - \lambda I_n)^{k-1}$ zleva, a vynulují se nám všechny členy této sumy až na poslední, tj., dostaneme

$$(A - \lambda I_n)^{k-1} c_k \vec{v}_k = \vec{v}_1 = \vec{o},$$

a tedy $c_k = 0$.
Máme tedy $\sum_{i=1}^{k-1} c_i \vec{v}_i = \vec{o}$. Aplikujeme na tuto rovnost matici $(A - \lambda I_n)^{k-2}$ zleva, a analogicky jako výše dostaneme $c_{k-1} = 0$. Takto pokračujeme, až dostaneme $c_k = c_{k-1} = \ldots = c_1 = 0$. □

11.2 Jordanův tvar nilpotentní matice

Definice 11.15. *Matice $A \in \mathbb{F}^{n\times n}$ je* nilpotentní, *pokud existuje $k \geq 0$ takové, že $A^k = O_{n\times n}$. Pokud A je nilpotentní, potom* index nilpotence *A je nejmenší $k \geq 0$ takové, že $A^k = O_{n\times n}$, značíme ho $nil(A)$.*

Mocnění matic v horním trojúhelníkovém tvaru, které mají na diagonále samé nuly, funguje podobně, jako mocnění Jordanových buněk příslušných 0 v tom smyslu, že se nenulové prvky posouvají doprava nahoru.

Cvičení 11.16. *Každá horní trojúhleníková matice $T \in \mathbb{F}^{n\times n}$, která má na hlavní diagonále samé nuly, je nilpotentní. Navíc, $nil(T) \leq n$.*

Poznámka 11.17. *Každá nilpotentní matice je singulární (a naopak to neplatí).*

Důkaz. Pokud $A^k = O_{n\times n}$, potom $\det(A)^k = \det(A^k) = \det(O_{n\times n}) = 0$, tedy $\det(A) = 0$, a A je singulární. Naopak, například matice $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ je singulární, ale ne nilpotentní, neboť $A^k = A$ pro každé $k \in \mathbb{N}$. □

Na následující větě je vidět přesný rozdíl mezi singulárními a nilpotentními maticemi. Připomeňmě, že matice je singulární právě tehdy, když má alespoň jedno vlastní číslo nulové.

Věta 11.18. *Matice $A \in \mathbb{F}^{n\times n}$ je nilpotentní právě tehdy, když všechna její vlastní čísla jsou nulová. Navíc, je-li A nilpotentní, pak $nil(A) \leq n$.*

Důkaz. Nechť A je nilpotentní, $k = nil(A)$, a zvolme libovolné $\lambda \in \sigma(A)$, a nechť $\vec{x} \neq \vec{o}$ je příslušný vlastní vektor. Potom z rovnosti $A \cdot \vec{x} = \lambda\vec{x}$ plyne $A^k \cdot \vec{x} = \lambda^k\vec{x}$, a tedy $\vec{o} = A^k \cdot \vec{x} = \lambda^k\vec{x}$, a tedy $\lambda = 0$.

Nechť naopak všechna vlastní čísla matice A jsou nulová. Dle Schurova rozkladu (Věta 7.2) nalezneme unitární matici $U \in \mathbb{F}^{n\times n}$ a horní trojúhelníkovou matici $T \in \mathbb{F}^{n\times n}$ takovou, že $A = U \cdot T \cdot U^*$. Potom T má na diagonále vlastní čísla matice A, a tedy samé nuly. Tedy

$$A^n = (U \cdot T \cdot U^*)^n = U \cdot T^n \cdot U^* =^{\text{Cvičení 11.16}} U \cdot O_{n\times n} \cdot U^* = O_{n\times n}.$$

Tedy A je nilpotentní, a $nil(A) \leq n$. □

Nyní si ukážeme (bez důkazu, protože je dlouhý a těžký) návod, jak najít Jordanův tvar nilpotentní matice. Nejprve si uvedeme dvě pomocná tvrzení:

Lemma 11.19. *Pokud $A, B \in \mathbb{F}^{n\times n}$, potom $def(A \cdot B) = def(B) + \dim(\ker(A) \cap Im(B))$.*

Lemma 11.20. *Nechť $A \in \mathbb{F}^{n\times n}$ a $i \in \mathbb{N}$. Ať $X = (\vec{x}_1, \ldots, \vec{x}_k)$ je matice, která má ve sloupcích bázi prostoru $Im(A^i)$, a ať $(\vec{v}_1, \ldots, \vec{v}_s)$ je báze prostoru $\ker(A \cdot X)$. Potom seznam $(X \cdot \vec{v}_1, \ldots, X \cdot \vec{v}_s)$ je báze prostoru $Im(A^i) \cap \ker(A)$.*

Pro každou nilpotentní matici $A \in \mathbb{F}^{n\times n}$ lze nalézt Jordanovu bázi, tj. bázi, která vznikla spojením Jordanových řetězců matice A. Jelikož 0 je jediné vlastní číslo A, počátek všech těchto řetezců leží v $\ker(A)$. Navíc, všechny Jordanovy řetězce, a všechny Jordanovy buňky, jsou příslušné vlastnímu číslu 0. Hlavní myšlenka je následující:

- utvoříme prostory
$$W_i = \ker(A) \cap Im(A^i), \quad i = 0, 1, \ldots, k = nil(A),$$
splňující
$$\{\vec{o}\} = W_k \subseteq W_{k-1} \subseteq \cdots \subseteq W_1 \subseteq W_0 = \ker(A),$$
a prostor $\ker(A) \setminus \{\vec{o}\}$ rozdělíme na disjunktní sjednocení
$$ker(A) \setminus \{\vec{o}\} = ker(A) \setminus W_k = W_0 \setminus (W_0 \setminus W_1) \cup (W_1 \setminus W_2) \cup \cdots (W_{k-1} \setminus W_k) = \bigcup_{i=0}^{nil(A)-1} W_i \setminus W_{i+1}.$$
Nyní, budeme z těchto množin postupně tvořit Jordanovu bázi A, přičemž začneme s poslední množinou $W_{k-1} \setminus W_k = W_{k-1} \setminus \{\vec{o}\}$.
- Všimněte si, že seznam $(\vec{v}_1, \ldots, \vec{v}_k)$ je Jordanův řetězec matice A délky i příslušný vlastnímu číslu 0 právě tehdy, když $\vec{v}_1 \in W_{i-1}$. Tedy, vektory $\vec{v}$, nad kterými lze vztyčit Jordanův řetězec délky i, ale ne delší, leží v množině $W_{i-1} \setminus W_i$.
- Prostor W_{i-1} má dle Lemma 11.19 dimenzi $def(A \cdot A^i) - def(A^i) = def(A^{i+1}) - def(A^i)$.
- Počet Jordanových řetězců pro A délky přesně i (což je to samé, jako počet Jordanových buněk $J_{0,i}$ v Jordanově tvaru A), je roven počtu vektorů v bázi W_{i-1}, které nejsou v bázi W_i, je jich tedy
$$\dim(W_{i-1}) - \dim(W_i) = 2def(A^i) - def(A^{i+1}) - def(A^{i-1}).$$

Návod 11.21. *Pro nalezení Jordanova tvaru nilpotentní matice $A \in \mathbb{F}^{n\times n}$ postupujeme takto:*

(1) Označme $k = nil(A) \leq n$.

(2) Označíme $d_i = def(A^i), \quad i = 0, \ldots, k+1$. Poznamenejme, že $d_{k+1} = d_k$.

(3) Počet Jordanových buněk $J_{0,i}$ v Jordanově tvaru A je roven $2d_i - d_{i-1} - d_{i+1}, \quad i = 1, \ldots, k$.

(4) Jordanův tvar napíšeme jako blokově diagonální matici s Jordanovými buňkami na diagonále.

Návod 11.22. *Pro nalezení Jordanovy báze nilpotentní matice $A \in \mathbb{F}^{n\times n}$, máme-li již spočtena čísla $d_i = def(A^i), \quad i = 1, \ldots, k+1$, postupujeme takto:*

(1) Utvoříme prostory

$$W_i = \ker(A) \cap Im(A^i), \quad i = 0, 1, \ldots, k,$$

(2) Pomocí Lemma 11.20 (pokud je to nutné, často to jde i jednodušeji) nalezneme bázi prostoru W_{k-1}, a pro každý vektor $\vec{v}$ této báze sestrojíme Jordanův řetězec délky k s počátkem $\vec{v}$ (řešením soustav lineárních rovnic s maticí soustavy A).

(3) Pro $i \geq 1$, máme-li nalezenou bázi prostoru W_i, doplníme ji o

$$\dim(W_{i-1}) - \dim(W_i) = 2d_i - d_{i-1} - d_{i+1}$$

vektorů na bázi W_{i-1}, a pro každý vektor $\vec{v}$ této báze sestrojíme Jordanův řetězec délky $i-1$ s počátkem $\vec{v}$ (řešením soustav lineárních rovnic s maticí soustavy A).

(4) Takto postupujeme až k prostoru $W_0 = ker(A)$.

(5) Spojením všech nalezených Jordanových řetězců získáme Jordanovu bázi $\mathcal{J}$ matice A. Napíšeme-li ji do sloupců matice P, potom

$$P^{-1} \cdot A \cdot P = T_{E_n \mapsto \mathcal{J}} \cdot [A]_{E_n \mapsto E_n} \cdot T_{\mathcal{J} \mapsto E_n} = [A]_{\mathcal{J} \mapsto \mathcal{J}}$$

je v Jordanově tvaru.

11.3 Jordanův tvar obecné matice

Tvrzení 11.23. *Nechť $A \in \mathbb{F}^{n\times n}$. Potom*

$$\vec{o} = \ker(A^0) \subseteq \ker(A^1) \ldots \subseteq \ker(A^k) \subseteq \ker(A^{k+1}) \subseteq \ldots,$$

a

$$\mathbb{F}^n = Im(A^0) \subseteq Im(A^1) \supseteq \ldots Im(A^k) \supseteq Im(A^{k+1}) \supseteq \ldots.$$

Navíc, existuje $k_0 \leq n$ takové, že pro každé $k \geq k_0$ platí $\ker(A^k) = \ker(A^{k+1})$ *a* $Im(A^k) = Im(A^{k+1})$. *Nejmenší takové k_0 nazáme* index A, *značíme* $ind(A)$.

Důkaz. Protože $A^0 = I_n$, je $\ker(A^0) = \{\vec{o}\}$. Dále, ukážeme, že $\ker(A^k) \subseteq \ker(A^{k+1})$ pro každé $k \geq 0$: pokud $\vec{x} \in \ker(A^k)$, potom $A^k \cdot \vec{x} = \vec{o}$, a tedy $A^{k+1} \cdot \vec{x} = A \cdot (A^k \cdot \vec{x}) = A \cdot \vec{o} = \vec{o}$. Nyní, pro každé $k \geq 0$ musí platit $d_k := def(A^k) = \dim(\ker(A^k)) \leq \dim(\mathbb{F}^n) = n$. Máme tedy nekonečnou nelesající shora omezenou posloupnost přirozených čísel

$$0 = d_0 \leq d_1 \leq \ldots \leq d_k \leq d_{k+1} \leq \ldots \leq n.$$

Existuje tedy $k_1 \geq 0$ takové, že $d_k = d_{k_1}$ pro každé $k \geq k_1$. To znamená, že pro každé $k \geq k_0$ platí $\dim(\ker(A^k)) = \dim(\ker(A^{k+1}))$, což spolu s inkluzí $\ker(A^k) \subseteq \ker(A^{k+1})$ implikuje $\ker(A^k) = \ker(A^{k+1})$.

Analogicky se ukáže, že existuje $k_2 \geq 0$ takové, že pro každé $k \geq k_2$ platí $Im(A^k) = Im(A^{k+1})$.

Nakonec stačí zvolit $k_0 = \max\{k_1, k_2\}$. □

Poznámka 11.24. *Lze ukázat, že posloupnosti* $\{\ker(A^k)\}_{k=0}^{\infty}$, $\{Im(A^k)\}_{k=0}^{\infty}$, *se stabilizují ve stejném okamžiku, jinými slovy*

$$ind(A) = \min\{k \in \mathbb{N} : \ker(A^k) = \ker(A^{k+1})\} = \min\{k \in \mathbb{N} : Im(A^k) = Im(A^{k+1})\}.$$

Poznámka 11.25. *Všimněte si, že pro nilpotentní matici* A *je její index nilpotence* $nil(A)$ *roven* $ind(A)$, *neboť*

$$nil(A) = \min\{k \geq 0 \colon A^k = O_{n\times n}\} = \min\{k \geq 0 \colon Im(A^k) = \vec{o}\} = \min\{k \geq 0 \colon Im(A^k) = Im(A^{k+1})\} = ind(A).$$

Definice 11.26. *Pro matici* $A \in \mathbb{F}^{n\times n}$ *definujeme:*

- zobecněné jádro A *jako* $\ker_\infty(A) = \ker(A^{ind(A)})$,
- zobecněný obraz A *jako* $Im_\infty(A) = Im(A^{ind(A)})$.
- *Lineární podprostor* $V \subseteq \mathbb{F}^n$ *je* A-invariantní, *pokud pro každé* $\vec{x} \in V$ *platí* $A \cdot \vec{x} \in V$.

Připomeňte si Větu o dimenzi jádra a obrazu, která říká, že $\dim(\ker(A)) + \dim(Im(A)) = \dim(\mathbb{F}^n)$ pro každou matici $A \in \mathbb{F}^{n\times n}$. V následujícím tvrzení ukážeme, že pro zobecněné jádro a zobecněný obraz platí ještě něco více, a to, že $\mathbb{F}^n$ je jejich direktním součtem.

Tvrzení 11.27. *Nechť* $A \in \mathbb{F}^{n\times n}$. *Potom platí:*

(1) $\ker_\infty(A)$ *a* $Im_\infty(A)$ *jsou lineární* A*-invariantní podprostory* $\mathbb{F}^n$.

(2) Zobrazení

$$A_{nil} := A|_{\ker_\infty(A)} \colon \ker_\infty(A) \to \ker_\infty(A), \quad \vec{x} \mapsto A \cdot \vec{x},$$

je nilpotentní.

(3) Zobrazení

$$A_{invert} := A|_{Im_\infty(A)} \colon Im_\infty(A) \to Im_\infty(A), \quad \vec{x} \mapsto A \cdot \vec{x},$$

je invertibilní.

(4) Platí $\mathbb{F}^n = \ker_\infty(A) \oplus Im_\infty(A)$.

Důkaz. (1) Jelikož $\ker_\infty(A) = \ker(A^{ind(A)})$ a $Im_\infty(A) = Im(A^{ind(A)})$, jsou zobecněné jádro a zobecněný obraz lineární podprostory $\mathbb{F}^n$.

Ukážeme, že $\ker_\infty(A)$ je A-invariantní: Pokud $\vec{x} \in \ker_\infty(A)$, potom

$$A^{ind(A)} \cdot (A \cdot \vec{x}) = A \cdot (A^{ind(A)} \cdot \vec{x}) = A \cdot \vec{o} = \vec{o},$$

a tedy $A \cdot \vec{x} \in \ker_\infty(A)$.

Analogicky se ukáže, že $Im_\infty(A)$ je A-invariantní.

(2) Ukážeme, že $A_{nil}^{Ind(A)}$ je nulové zobrazení: Pokud $\vec{x} \in \ker_\infty(A) = \ker(A^{ind(A)})$, potom

$$A_{nil}^{Ind(A)} \cdot \vec{x} = A^{Ind(A)} \cdot \vec{x} = \vec{o}.$$

Tedy A_{nil} je nilpotentní.

(3) Ukážeme, že A_{invert} je isomorfismus. K tomu stačí ukázat, že jde o epimorfismus, dle Věty 1.26. Nechť tedy $\vec{x} \in Im_\infty(A) = Im(A^{ind(A)}) = Im(A^{ind(A)+1})$. Potom $\vec{x} = A^{ind(A)+1} \cdot \vec{y} = A \cdot (A^{ind(A)} \cdot \vec{y})$ pro nějaké $\vec{y} \in \mathbb{F}^n$. Tedy, jelikož $\vec{z} := A^{ind(A)} \cdot \vec{y} \in Im_\infty(A)$,

$$A_{invert} \cdot \vec{z} = A \cdot \vec{z} = A \cdot (A^{ind(A)} \cdot \vec{y}) = \vec{x},$$

je zobrazení A_{invert} epimorfismus.

(4) Nejprve ukážeme, že $\ker_\infty(A) \cap Im_\infty(A) = \{\vec{o}\}$. Zvolme tedy $\vec{x} \in \ker_\infty(A) \cap Im_\infty(A)$. Potom

$$A_{invert}^{ind(A)} \cdot \vec{x} = A^{ind(A)} \cdot \vec{x} = A_{nil}^{ind(A)} \cdot \vec{x} = \vec{o}.$$

Jelikož A_{invert} je isomorfsmus, $A_{invert}^{ind(A)}$ je také isomorfimus, tedy monomorfismus, a tedy $\vec{x} = \vec{o}$.
Dále, platí

$$n = def(A^{ind(A)}) + rank(A^{ind(A)}) = \dim(\ker_\infty(A)) + \dim(Im_\infty(A)),$$

což spolu s podmínkou $\ker_\infty(A) \cap Im_\infty(A) = \{\vec{o}\}$ implikuje $\mathbb{F}^n = \ker_\infty(A) \oplus Im_\infty(A)$.

□

Všimněte si, že v předchozím tvrzení jsme si ukázali toto: Pro libovolné lineární zobrazení $f\colon \mathbb{F}^n \to \mathbb{F}^n$ můžeme prostor $\mathbb{F}^n$ rozložit na lineární komplementární A-invariantní podprostory V, W tak, že $f|_V$ je isomorfimus, a $f|_W$ je nilpotentní, tj. "téměř nulové".

Lemma 11.28. *Nechť $A \in \mathbb{F}^{n\times n}$ a $\lambda \in \mathbb{F}$. Potom $A - \lambda I_n$ je nilpotentní právě tehdy, když $\sigma_A = \{\lambda\}$.*

Důkaz. Dle Věty 11.18 je $A - \lambda I_n$ je nilpotentní právě tehdy, když $\sigma(A - \lambda I_n) = \{0\}$. To je ekvivalentní s tím, že $\sigma_A = \{\lambda\}$. □

Definice 11.29. *Pro $A \in \mathbb{F}^{n\times n}$ a $\lambda \in \sigma(A)$ definujeme* index vlastního čísla λ matice A *jako $ind_\lambda(A) = ind(A - \lambda I_n)$ (což je rovno indexu nilpotence zobrazení $(A - \lambda I_n)_{nil}$).*

Nyní zformulujeme hlavní myšlenky existence Jordanova tvaru a Jordanovy báze obecné matice:
Nechť $A \in \mathbb{F}^{n\times n}$ a $\sigma(A) = \{\lambda_1, \dots, \lambda_p\}$, a pro $i = 1, \dots, p$ označme $m_i = an_A(\lambda_i)$. Lze ukázat, že prostory

$$V_i = \ker_\infty(A - \lambda_i I_n), \quad i = 1, \dots, n,$$

jsou komplementární, A-invariantní, dimenze m_i, stačí tedy nalézt části Jordanovy báze pro zobrazení $A|_{V_i}$, a nakonec je sjednotit. Tyto části Jordanovy báze nalezneme takto: platí

$$A|_{V_i} = (A - \lambda_i I_n)|_{V_i} + \lambda_i I_n|_{V_i} = (A - \lambda_i I_n)_{nil} + \lambda_i I_{m_i},$$

tedy, A se na V_i liší od nilpotentního zobrazení jen o skalární násobek identického zobrazení. Stačí tedy nalézt Jordanův tvar N_i a Jordanovu bázi $\mathcal{J}_i$ nilpotentního zobrazení $(A - \lambda_i I_n)_{nil}$ (což už umíme), a potom $B(\lambda_i) = N_i + \lambda_i I_{m_i}$ bude Jordanův tvar zobrazení A_{V_i}. Nakonec, $\mathcal{J} = \bigcup_{i=1}^p \mathcal{J}_i$ bude Jordanova báze A, a

$$[A]_{\mathcal{B}\mapsto\mathcal{B}} = diag(B(\lambda_1), \dots, B(\lambda_p)).$$

$B(\lambda_i)$ se nazývá *Jordanův segment* A příslušný vlastnímu číslu λ_i.
Nyní si ještě zformulujeme větu o existenci spektrálních projekcí pro obecnou matici, která je obdobou věty 8.1.

Věta 11.30. *Nechť $A \in \mathbb{F}^{n\times n}$, $\sigma(A) = \{\lambda_1, \dots, \lambda_p\}$, $A = P \cdot diag(B(\lambda_1), \dots, B(\lambda_p)) \cdot P^{-1}$ je Jordanův tvar A. Označme $m_i = an_A(\lambda_i)$, $k_i = ind_{\lambda_i}(A)$ $i = 1, \dots, p$ a*

$$P = \begin{pmatrix} P_1 & P_2 & \dots & P_p \end{pmatrix}, P^{-1} = \begin{pmatrix} S_1 \\ S_2 \\ \vdots \\ S_p \end{pmatrix},$$

kde $P_i \in \mathbb{F}^{n\times m_i}, S_i \in \mathbb{F}^{m_i\times n}$. Potom matice $G_i = P_i \cdot S_i$ je projekce na $\ker_\infty(A - \lambda_i I_n) = \ker((A - \lambda_i I_n)^{k_i})$ podle $Im_\infty(A - \lambda_i I_n) = Im((A - \lambda_i I_n)^{k_i})$, nazývá se spektrální projekce příslušná vlastnímu číslu λ_i.

Postup pro hledání Jordanova tvaru a Jordanovy báze obecné matice si shrneme v následujícím návodu.

Návod 11.31. *Pro nalezení Jordanova tvaru a Jordanovy báze obecné matice $A \in \mathbb{F}^{n\times n}$ postupujeme takto:*

(1) Nalezneme $\sigma(A) = \{\lambda_1, \dots, \lambda_p\}$, a pro každé $\lambda_j \in \sigma(A)$ zjistíme $m_j = an_A(\lambda_j)$.

(2) Jordanův tvar bude mít p Jordanových segmentů *$B(\lambda_j), j = 1, \dots, p$. Segment $B(\lambda_j)$ má rozměr $m_j \times m_j$. Počet buněk v segmentu $B(\lambda_j)$ je roven $gn_A(\lambda_j)$. Velikost největší buňky v segmentu $B(\lambda_j)$ je rovna $ind_{\lambda_j}(A)$.*

(3) Pro nalezení segmentu $B(\lambda_j)$: Utvoříme matici $A_j = A - \lambda_j I_n$, a nalezneme pro ni Jordanův tvar N_j tak jako v Návodu 11.21, s jediným rozdílem, že místo $k = nil(A_j)$ vezmeme $k = ind_{\lambda_j}(A) = nil((A_j)_{nil})$. Ten najdeme pomocí Poznámky 11.32, nebo tak, že postupně hledáme $ker(A_j^1), ker(A_j^2), \dots$, dokud neplatí $def(A_j^i) = \dim(ker_\infty(A_j)) = m_j$.

(4) Platí $B(\lambda_j) = N_j + \lambda_j I_{m_j}$.

(5) Pro nalezení j-té části Jordanovy báze použijeme Návod 11.22 na matici A_j. Nakonec všechny tyto nalezené části sjednotíme, a dostaneme Jordanovu bázi J matice A. Napíšeme-li ji do sloupců matice P, potom

$$P^{-1} \cdot A \cdot P = T_{E_n \mapsto \mathcal{J}} \cdot [A]_{E_n \mapsto E_n} \cdot T_{\mathcal{J} \mapsto E_n} = [A]_{\mathcal{J} \mapsto \mathcal{J}}$$

je v Jordanově tvaru.

Pro zjednodušení ručního počítání Jordanových tvarů malých matic je důležitá následující poznámka.

Poznámka 11.32. *Pro každý segment $B(\lambda)$ platí*

$$an_A(\lambda) = \sum_{i=1}^{gn_A(\lambda)} n_i,$$

kde n_i jsou velikosti Jordanových buněk v segmentu $B(\lambda)$.

Pokud tedy $gn_A(\lambda) = 1$, segment $B(\lambda_j)$ obsahuje pouze jednu buňku, a tedy platí $ind_{\lambda_j}(A) = an_A(\lambda)$ (protože velikost té jediné, a tedy největší, buňky, je rovna $ind_{\lambda_j}(A)$). Navíc, v tomto případě platí

$$\ker(A - \lambda I_n) \subseteq Im((A - \lambda I_n)^i), \quad i = 1, \dots, ind_\lambda(A).$$

12 Maticové funkce

V této kapitole se věnujeme konstrukci funkce čtvercových matic- co znamenají výrazy typu $e^A, \sin(A), \log(A), \sqrt{A}, ...$, a pro jaké matice $A \in \mathbb{F}^{n\times n}$ má taková definice smysl? Chceme při tom zachovat známé základní vlastnosti takových funkcí, například $e^{A+B} = e^A \cdot e^B, \sin^2(A) + \cos^2(A) = I_n, \cdots$.

Naivní pokus o takovou definici by mohl být po složkách, například pro funkci $f\colon \mathbb{C} \to \mathbb{C}$ definovat

$$f(\begin{pmatrix} 1 & 3 \\ -2 & 4 \end{pmatrix}) = \begin{pmatrix} f(1) & f(3) \\ f(-2) & f(4) \end{pmatrix}.$$

Takový přístup by nefungoval ani pro tak jednoduchou funkci, jako $f(x) = x^2$, protože namísto umocnění matice by člověk mocnil její složky.
Lepší pokus (který už opravdu pro řadu funkcí funguje), je použít definici mocninnou řadou: Pro polynomy je definovat maticové funkce snadné (jelikož víme, co znamená sčítání a násobení matic): Pokud $p(x) =$

$\sum_{k=0}^{n} a_k x^k$, potom můžeme jednoduše definovat $p(A) = \sum_{k=0}^{n} a_k A^k$. To nanznačuje, jak definovat $f(A)$, pokud lze funkci f rozvést do mocninné řady. Například, pro funkci

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} \text{ definujeme } e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!},$$

tato funkce se nazývá *maticová exponenciála.*
Obecně je s takovouto definicí několik problémů:

- Funguje jen pro funkce f, které lze rozvést do mocninné řady,
- může být těžké poznat, pro které matice A taková řada konverguje, a tedy, pro které matice A má taková definice smysl,
- jak explicitně vyjádřit složky $f(A)$? Tj. pokud $A = \begin{pmatrix} 1 & 3 \\ -2 & 4 \end{pmatrix}$, kolik je $[e^A]_{1,1}$?

12.1 Diagonalizovatelné matice

Pokud A je diagonalizovatelná, můžeme tyto problémy vyřešit alternativní definicí. Předpokládejme, že $A = T \cdot D \cdot T^{-1}$, kde $D = diag(\lambda_1, \dots, \lambda_n)$, potom

$$e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!} = \sum_{k=0}^{\infty} \frac{T \cdot D^k \cdot T^{-1}}{k!} = T \cdot \sum_{k=0}^{\infty} \frac{D^k}{k!} \cdot T^{-1} = T \cdot diag(\sum_{k=1}^{\infty} \frac{\lambda_1^k}{k!}, \dots, \frac{\lambda_n^k}{k!}) \cdot T^{-1} =$$
$$= T \cdot diag(e^{\lambda_1}, \dots, e^{\lambda_n}) \cdot T^{-1}.$$

Pro obecnou funkci f, definovanou na $\sigma(A) = \{\lambda_1, \dots, \lambda_n\}$, a diagonalizovatelnou matici A, můžeme tedy použít tuto definici:

Definice 12.1. *Nechť $A \in \mathbb{C}^{n \times n}$ je diagonalizovatelná matice, $\lambda_1, \dots, \lambda_n$ jsou vlastní čísla A zapsaná včetně algebraických násobností, a nechť f je komplexní funkce definována na $\sigma(A)$. Nechť $D = diag(\lambda_1, \dots, \lambda_n)$, a nechť regulární matice T splňuje $A = T \cdot D \cdot T^{-1}$. Potom definujeme*

$$f(A) = T \cdot f(D) \cdot T^{-1} = T \cdot diag(f(\lambda_1), \dots, f(\lambda_n)) \cdot T^{-1}.$$

Poznámka 12.2. *Jde něco udělat s předpokladem, že f je definována na $\sigma(A)$? Ne, a to ani pro matice velikosti 1, tj. skaláry- jelikož pro $A = (a_{1,1}) \in \mathbb{C}^{1\times 1}$, $\sigma(A) = \{a\}$, je funkce $f(A)$ definována právě tehdy, když je f definována na $\sigma(A) = \{a\}$.*

Tento přístup definice $f(A)$ řeší výše zmíněné tři problémy, ale zatím nevíme, zda je definice korektní, tj. zda nezávisí na volbě matice T, která není určená jednoznačně. S tím se vypořádáme v následující větě.

Věta 12.3. *Definice $f(A)$ v 12.1 nezávisí na volbě rozkladu $A = T \cdot D \cdot T^{-1}$. Je-li $\sigma(A) = \{\lambda_1, \dots, \lambda_k\}$, platí*

$$f(A) = \sum_{i=1}^{k} f(\lambda_i) G_i,$$

kde G_i jsou spektrální projekce z Věty 8.1.

Důkaz. Označme $m_j = gm_A(\lambda_j) = an_A(\lambda_j)$ pro $j \in \{1, \dots, k\}$. Nyní, pro libovolný rozklad $A = T \cdot D \cdot T^{-1}$ pišme jako v důkazu Věty 8.1

$$f(A) = T \cdot f(D) \cdot T^{-1} = \begin{pmatrix} X_1 & \cdots & X_k \end{pmatrix} \cdot diag(f(\lambda_1), \dots, f(\lambda_n)) \cdot \begin{pmatrix} Y_1 \\ \vdots \\ Y_k \end{pmatrix} = \sum_{i=1}^{k} f(\lambda_i) X_i \cdot Y_i = \sum_{i=1}^{k} f(\lambda_i) G_i, \tag{12.1}$$

kde pro $j \in \{1, \dots, k\}$, $X_j \in \mathbb{C}^{n \times m_j}$, sloupce X_j tvoří bázi $eigen(\lambda_j, A)$, a $Y_j \in \mathbb{C}^{m_j \times n}$.
Jelikož pro každé $i \in \{1, \dots, n\}$, G_i je matice projekce na $ker(A - \lambda_i I_n)$ podle $Im(A - \lambda_i I_n)$, je G_i jednoznačně určená maticí A (to plyne také z Věty 8.2), a tedy $f(A)$ nezávisí na volbě rozkladu $A = T \cdot D \cdot T^{-1}$. □

Podobně jako u příkladu exponenciální funkce výše, lze dokázat následující větu.

Věta 12.4. *Pokud $f(z) = \sum_{k=1}^{\infty} c_k \cdot (z - z_0)^k$ konverguje pro $|z - z_0| < r$ a pokud $|\lambda - z_0| < r$ pro každé $\lambda \in \sigma(A)$, potom řada $\sum_{k=1}^{\infty} c_k \cdot (A - z_0 I_n)^k$ koverguje, a pokud A je diagonalizovatelná, potom $f(A)$ z Definice 12.1 se rovná této řadě.*

Pro numerický výpočet $f(A)$ se někdy využívá takzvaný *Lagrangeův interpolační polynom.*

Věta 12.5. *Nechť $\alpha_1, \ldots, \alpha_k \in \mathbb{C}$ $\lambda_1, \ldots, \lambda_k$ jsou různá komplexní čísla. Potom existuje právě jeden polynom $L\colon \mathbb{C} \to \mathbb{C}$ takový, že $deg(L) \leq k - 1$ a $L(\lambda_i) = \alpha_i$ pro $i = 1, \ldots, k$. Tento polynom se nazývá* Lagrangeův interpolační polynom *pro systém bodů $\{(\lambda_i, \alpha_i)\}_{i=1}^k$, a splňuje*

$$L(z) = \sum_{j=1}^{k} \alpha_j \frac{\prod\limits_{1 \leq i \leq k, i \neq j} (z - \lambda_i)}{\prod\limits_{1 \leq i \leq k, i \neq j} (\lambda_j - \lambda_i)}, \quad z \in \mathbb{C}.$$

Důkaz. Pro $m = 1, \ldots, k$ platí

$$L(\lambda_m) = \alpha_m \frac{\prod\limits_{1 \leq i \leq k, i \neq m} (\lambda_m - \lambda_i)}{\prod\limits_{1 \leq i \leq k, i \neq m} (\lambda_m - \lambda_i)} + \sum_{j=1, j \neq m}^{k} \alpha_j \frac{\prod\limits_{1 \leq i \leq k, i \neq j} (\lambda_m - \lambda_i)}{\prod\limits_{1 \leq i \leq k, i \neq j} (\lambda_j - \lambda_i)} = \alpha_m + 0 = \alpha_m.$$

Jednoznačnost: Pokud L_1, L_2 jsou polynomy splňující předpoklady Věty, potom $p = L_1 - L_2$ je polynom stupně nejvýše $k - 1$, který má k různých kořenů, a tedy je nulový, což znamená, že $L_1 = L_2$. □

Všimněte si, že pokud $A \in \mathbb{C}^{n \times n}$ je diagonalizovatelná, $\sigma(A) = \{\lambda_1, \ldots, \lambda_k\}$, $f\colon \sigma(A) \to \mathbb{C}$, a L je Lagrangeův interpolační polynom pro systém bodů $\{(\lambda_i, f(\lambda_i))\}_{i=1}^k$, potom platí

$$f(A) = \sum_{j=1}^{k} f(\lambda_j) G_j =^{8.2} \sum_{j=1}^{k} f(\lambda_j) \frac{\prod\limits_{1 \leq i \leq k, i \neq j} (A - \lambda_i I_n)}{\prod\limits_{1 \leq i \leq k, i \neq j} (\lambda_j - \lambda_i)} = L(A).$$

Tedy: Hodnota $f(A)$ se dá vždy vyjádřit jako polynom matice A stupně nejvýše $k - 1$.

12.2 Obecné matice

Nyní budeme chtít myšlenky z předchozí části zobecnit na případ matic, které nejsou diagonalizovatelné. K tomu se nabízí využít Jordanův tvar matice. Pokud $A = P \cdot J \cdot P^{-1}$ je v Jordanově tvaru, chtěli bychom definovat $f(A) = P \cdot f(J) \cdot P^{-1}$ pro ty funkce f, pro které to dává smysl. Základním krokem je definovat $f(J_{\lambda,k})$, pro libovolnou Jordanovu buňku. K tomu může být dobré si připomenout, jak funguje mocnění Jordanovy buňky (Tvrzení 11.8)-pro funkci $f(z) = z^k$ by nám měl vyjít ten stejný vzorec.
Předpokládejme, že funkce f má na okolí λ rovoj do Taylorovy řady, tj. existuje $r > 0$ tak, že pro $|z - \lambda| < r$ platí

$$f(z) = f(\lambda) + f'(\lambda)(z - \lambda) + f''(\lambda)\frac{(z - \lambda)^2}{2!} + \ldots,$$

potom jako výše, definujeme

$$f(J_{\lambda,k}) = f(\lambda) I_k + f'(\lambda)(J_{\lambda,k} - \lambda I_k) + f''(\lambda)\frac{(J_{\lambda,k} - \lambda I_k)^2}{2!} + \ldots.$$

Jelikož je ale matice $J_{\lambda,k} - I_k = J_{0,k}$ nilpotentní řádu $k - 1$, je výše uvedená suma konečná, a stačí nám tedy předpokládat existenci derivací $f(\lambda), f'(\lambda), \ldots, f^{(k-1)}(\lambda)$. Navíc, jelikož víme, jak umocnit matici $J_{0,k}$ (viz Tvrzení 11.6), můžeme tuto sumu zapsat alternativním způsobem, jako v následující definici.

Definice 12.6. *Jestliže existují* $f(\lambda), f'(\lambda), \ldots, f^{(k-1)}(\lambda)$*, potom definujeme*

$$f(J_{\lambda,k}) = \sum_{i=0}^{k-1} f^{(i)}(\lambda)\frac{J_{0,k}^i}{i!} = \begin{pmatrix} f(\lambda) & f'(\lambda) & \frac{f''(\lambda)}{2!} & \ldots & \frac{f^{(k-1)}(\lambda)}{(k-1)!} \\ 0 & f(\lambda) & f'(\lambda) & \ldots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & f(\lambda) & f'(\lambda) \\ 0 & 0 & \cdots & 0 & f(\lambda) \end{pmatrix}$$

Nyní už můžeme definovat $f(A)$ pro libovolnou matici A, pokud f má konečné derivace dost vysokého řádu na spektru A. Připomeňte si, že pokud $B(\lambda)$ je Jordanův segment A, potom velikost jeho největší buňky je rovna $ind_\lambda(A)$.

Definice 12.7. *Nechť* $A \in \mathbb{C}^{n\times n}$*,* $\sigma(A) = \{\lambda_1, \ldots, \lambda_s\}$*, a* $k_j = ind_{\lambda_j}(A), \quad j = 1, \ldots, s.$

- *Řekneme, že komplexní funkce* f je definována (nebo existuje) v A*, pokud existují konečné derivace*

  $$f(\lambda_j), f'(\lambda_j), \ldots, f^{(k_j-1)}(\lambda_j)$$

  pro každé $j = 1, \ldots, s$*.*

- *Je-li* $A = P \cdot J \cdot P^{-1} = P \cdot diag(J_{\lambda_1,m_1}, \ldots, J_{\lambda_r,m_r}) \cdot P^{-1}$ *v Jordanově tvaru, potom definujeme*

  $$f(A) = P \cdot f(J) \cdot P^{-1} = P \cdot diag(f(J_{\lambda_1,m_1}), \ldots, f(J_{\lambda_r,m_r})) \cdot P^{-1},$$

  kde $f(J_{\lambda_i,k_i})$ *je definováno jako v 12.6.*

Nyní, podobně jako v diagonalizovatelném případě, lze ověřit, že tato definice $f(A)$ nezávisí na volbě rozkladu $A = P \cdot J \cdot P^{-1}$. Důkaz následující Věty je podobný (i když trochu komplikovanější) jako u Věty 12.3.

Věta 12.8. *(Spektrální rozklad* $f(A)$*) Nechť* $A \in \mathbb{C}^{n\times n}$*, a nechť komplexní funkce* f *je definována v* A*. Potom definice* $f(A)$ *v nezávisí na volbě rozkladu* $A = P \cdot J \cdot P^{-1}$*. Je-li* $\sigma(A) = \{\lambda_1, \ldots, \lambda_s\}$ *a* $k_i = ind_{\lambda_i}(A), \quad i = 1, \ldots, s$*, platí*

$$f(A) = \sum_{i=1}^{s}\sum_{j=0}^{k_i-1} f^{(j)}(\lambda_i)\frac{(A-\lambda_i I_n)^j}{j!}G_i,$$

kde G_i *jsou spektrální projekce z Věty 11.30.*
Označíme-li tedy $Z_{i,j} = \frac{(A-\lambda_i I_n)^j \cdot G_i}{j!}$*, potom*

$$f(A) = \sum_{i=1}^{s}\sum_{j=0}^{k_i-1} f^{(j)}(\lambda_i)Z_{i,j}.$$

Matice $Z_{i,j}$ *se nazývají* konstituenty, nebo komponentní matice $f(A)$.

Komponentní matice jsou tedy důležité z následujícího důvodu: Pokud máte matici $A \in \mathbb{C}^{n\times n}$, potom hodnotu libovolné funkce f, definované v A, lze vypočítat jako lineární kombinaci komponentních matic (a ty závisejí pouze na matici A, nikoli na funkci f). Na závěr této kapitoly si ještě ukážeme jednu techniku hledání těchto komponentních matic. Nejprve si ukažme příklad.

Příklad 12.9. *Popište* $f(A)$ *pro* $A = \begin{pmatrix} 6 & 2 & 8 \\ -2 & 2 & -2 \\ 0 & 0 & 2 \end{pmatrix}$*, pro* f *definovanou v* A*, bez hledání přechodové matice* P*.*

Důkaz. Platí $\sigma(A) = \{2, 4\}$, $an_A(2) = 1$, $an_A(2) = 2$, $gn_A(2) = 1$, $gn_A(2) = 1$, tedy $ind_2(A) = an_A(2) = 1$, $ind_4(A) = an_A(2) = 1$ (Poznámka 11.32). Tedy pro jakoukoli komplexní funkci definovanou v $2, 4$ takovou, že existuje vlastní $f^{'}(4)$, je $f(A)$ definováno, a platí

$$f(A) = f(2) \cdot G_1 + f(4) \cdot G_2 + f^{'}(4)(A - 4I_n)G_2.$$

Nyní, k nalezení spektrálních projekcí můžeme použít následující trik: dosadíme do rovnosti nahoře nějaké vhodně zvolenené funkce f.

Volba $f_1(z) = 1$ dává rovnost $I_3 = f_1(A) = G_1 + G_2$, a volba $f_2(z) = (z-4)^2$ dává $(A-4I_3)^2 = f_2(A) = 4G_1$. Tedy dostaneme jednoduchovu maticovou soustavu rovnic, jejíž řešení je evidentně $G_1 = \frac{(A-4I_3)^2}{4}$, $G_2 = I_3 - G_1$. Nyní můžeme snadno dopočítat G_1, G_2, a nebo rovnou psát

$$f(A) = f(2) \cdot \frac{(A - 4I_3)^2}{4} + f(4) \cdot \Big(I_3 - \frac{(A - 4I_3)^2}{4}\Big) + f^{'}(4)\Big(A - 4I_n - \frac{(A - 4I_3)^3}{4}\Big).$$

□

Toto lze použít pro libovolnou čtvercovou matici A, když člověk vezme dostatečně mnoho funkcí f_i: Například, pro $k = (\sum_{i=1}^{s} k_i) - 1$, volbou funkcí $f_p(z) = z^p$, $p = 0, 1, 2, \ldots, k$ (toto je jedna z možností, lze volit i jiné funkce f_p)v rovnosti

$$f(A) = \sum_{i=1}^{s} \sum_{j=0}^{k_i - 1} f^{(j)}(\lambda_i) Z_{i,j},$$

dostaneme rovnost

- $p = 0 \Rightarrow I_n = f_1(A) = \sum_{i=1}^{s} Z_{i,0}$,
- $p = 1 \Rightarrow A = f_1(A) = \sum_{i=1}^{s} \lambda_i Z_{i,0} + \sum_{i=1}^{s} Z_{i,1}$,
- $p = 2 \Rightarrow A^2 = f_2(A) = \sum_{i=1}^{s} \lambda_i^2 Z_{i,0} + \sum_{i=1}^{s} 2\lambda_i Z_{i,1} + \sum_{i=1}^{s} 2Z_{i,2}$,
- $\ldots$,

což lze maticově zapsat jako

$$\begin{pmatrix} 1 & \ldots & 1 & & & & & & & \\ \lambda_1 & \ldots & \lambda_s & 1 & \ldots & 1 & & & & \\ \lambda_1^2 & \ldots & \lambda_s^2 & 2\lambda_1 & \ldots & 2\lambda_s & 2 & \ldots & 2 & \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \\ \lambda_1^k & \ldots & \lambda_s^k & k\lambda_1^{k-1} & \ldots & k\lambda_s^{k-1} & \ldots & \ldots & \ldots & \ldots \end{pmatrix} \cdot \begin{pmatrix} Z_{1,0} \\ \vdots \\ Z_{s,0} \\ Z_{1,1} \\ \vdots \\ Z_{s,1} \\ \vdots \end{pmatrix} = \begin{pmatrix} I \\ A \\ A^2 \\ A^3 \\ \vdots \\ A^k \end{pmatrix}.$$

Toto se nazývá *zobecněná Vandermondeova soustava*, a lze ukázat, že má řešení.

Nakonec, podobně jako pro diagonalizovatlené matice lze ukázat, že hodnotu $f(A)$ lze vždy vyjádřit jako polynom matice A: pokud $A \in \mathbb{C}^{n \times n}$, $\sigma(A) = \{\lambda_1, \ldots, \lambda_s\}$, a f je definována v A, potom existuje jediný polynom H takový, že $deg(H) \leq (\sum_{i=1}^{s} ind_{\lambda_i}(A)) - 1$, a pro každé $i = 1, \ldots, s$ a $j = 1, \ldots, ind_{\lambda_i}(A)$ platí $H^{(j)}(\lambda_i) = f^{(j)}(\lambda_i)$, nazývá se *Hermiteův interpolační polynom*. Tento polynom tedy splňuje

$$f(A) = \sum_{i=1}^{s} \sum_{j=0}^{k_i - 1} f^{(j)}(\lambda_i) Z_{i,j} = \sum_{i=1}^{s} \sum_{j=0}^{k_i - 1} H^{(j)}(\lambda_i) Z_{i,j} = H(A).$$

13 Aplikace – Řešení soustav diferenciálních rovnic

V této kapitole si ukážeme, jak můžeme pomocí matic řešit soustavy diferenciálních rovnic.

Uvažujme následující soustavu diferenciálních rovnic. Cílem je pro dané skaláry $a_{i,j}$ a c_i, a funkce $f_i\colon \mathbb{R}\to\mathbb{R}$ nalézt funkce $u_i\colon \mathbb{R}\to\mathbb{R}$ splňující pro každé $t\in\mathbb{R}$:

$$\begin{aligned} u_1'(t) &= a_{1,1}u_1(t) + a_{1,2}u_2(t) + \ldots + a_{1,n}u_n(t) + f_1(t)\\ u_2'(t) &= a_{2,1}u_1(t) + a_{2,2}u_2(t) + \ldots + a_{2,n}u_n(t) + f_2(t)\\ &\vdots\\ u_n'(t) &= a_{n,1}u_n(t) + a_{n,2}u_2(t) + \ldots + a_{n,n}u_n(t) + f_n(t) \end{aligned}$$

splňující tzv. *počáteční podmínky* $u_1(0)=c_1, u_2(0)=c_2,\ldots,u_n(0)=c_n$. Maticově tuto soustavu můžeme zapsat jako $\vec{u}' = A\cdot\vec{u}+\vec{f}, \vec{u}(0)=\vec{c}$, kde

$$A=(a_{i,j})_{i,j=1}^n,\quad \vec{u}=\begin{pmatrix}u_1\\ \ldots\\ u_n\end{pmatrix},\quad \vec{c}=\begin{pmatrix}c_1\\ \ldots\\ c_n\end{pmatrix},\quad \text{a}\quad \vec{f}=\begin{pmatrix}f_1\\ \ldots\\ f_n\end{pmatrix},$$

přičemž derivování vektorů (nebo matic) funkcí chápeme tak, že se zderivuje každá složka vektoru (nebo matice). Tato soustava se nazývá *homogenní*, pokud $\vec{f}(t)=\vec{o}$ pro každé $t\in\mathbb{R}$, jinak se nazývá *nehomogenní*.

13.1 Homogenní soustavy

Připomeňte si, že jediným řešením jedné homogenní diferenciální rovnice $u'(t)=a\cdot u(t), u(0)=c$, je funkce $u(t)=e^{at}\cdot c$. Dá se tedy očekávat, že s řešením výše zmíněné soustavy bude nějak souviset funkce e^{tA}. Připomeňte si, že tuto funkci můžeme buďto vyjádřit jako mocninnou řadu

$$e^{t\cdot A}=\sum_{j=0}^{\infty}\frac{t^j}{j!}A^j,$$

nebo pomocí Jordanova tvaru či spektrálních projekcí: Je-li

$$A=P\cdot J\cdot P^{-1}=P\cdot diag(J_{\lambda_1,m_1},\ldots,J_{\lambda_r,m_r})\cdot P^{-1}$$

v Jordanově tvaru, potom dle definic 12.6 a 12.7, aplikovaných na komplexní funkci $z\mapsto e^{t\cdot z}$ platí

$$e^{t\cdot A}=P\cdot e^{t\cdot J}\cdot P^{-1}=P\cdot diag(e^{t\cdot J_{\lambda_1,m_1}},\ldots,e^{t\cdot J_{\lambda_r,m_r}})\cdot P^{-1},$$

kde

$$e^{t\cdot J_{\lambda_i,m_i}}=\begin{pmatrix} e^{\lambda t} & t\cdot e^{\lambda t} & \frac{t^2\cdot e^{\lambda t}}{2!} & \ldots & \frac{t^{k-1}\cdot e^{\lambda t}}{(k-1)!}\\ 0 & e^{\lambda t} & t\cdot e^{\lambda t} & \ldots & \vdots\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ 0 & \cdots & 0 & e^{\lambda t} & t\cdot e^{\lambda t}\\ 0 & 0 & \cdots & 0 & e^{\lambda t}\end{pmatrix},$$

a, pokud $\sigma(A)=\{\lambda_1,\ldots,\lambda_s\}$ a $k_i=ind_{\lambda_i}(A),\quad i=1,\ldots,s$, dle Věty 12.8 platí

$$e^{t\cdot A}=\sum_{i=1}^{s}\sum_{j=0}^{k_i-1}\frac{t^j\cdot e^{\lambda_i\cdot t}}{j!}\cdot(A-\lambda_i I_n)^j\cdot G_i,$$

kde G_i jsou spektrální projekce matice A.

Nejprve si uvedeme několik vlastností funkce $t\mapsto e^{t\cdot A}$.

Tvrzení 13.1. *Nechť $A \in \mathbb{C}^{n\times n}$ a $t \in \mathbb{R}$. Potom platí:*

(1) $\frac{d}{dt}e^{t\cdot A} = A \cdot e^{t\cdot A} = e^{t\cdot A} \cdot A$,

(2) Matice $e^{t\cdot A}$ je regulární a platí $(e^{t\cdot A})^{-1} = e^{-t\cdot A}$.

Důkaz. (1) Díky lokálně stejnoměrné konvergenci řady a spojitosti maticového násobení lze psát

$$\frac{d}{dt}e^{t\cdot A} = \frac{d}{dt}(\sum_{j=0}^{\infty}\frac{t^j}{j!}A^j) = \sum_{j=0}^{\infty}\frac{d}{dt}(\frac{t^j}{j!}A^j) = \sum_{j=1}^{\infty}\frac{t\cdot t^{j-1}}{j!}A^j = \sum_{j=1}^{\infty}\frac{t^{j-1}}{(j-1)!}A^j =$$

$$= A\cdot\sum_{j=1}^{\infty}\frac{t^{j-1}}{(j-1)!}A^{j-1} = A\cdot e^{t\cdot A}$$

Zároveň také $\sum_{j=1}^{\infty}\frac{t^{j-1}}{(j-1)!}A^j = (\sum_{j=1}^{\infty}\frac{t^{j-1}}{(j-1)!}A^{j-1})\cdot A = e^{t\cdot A}\cdot A$, což dokazuje $A\cdot e^{t\cdot A} = e^{t\cdot A}\cdot A$.

(2) Stačí ukázat, že funkce

$$F\colon \mathbb{R}\to\mathbb{C}^{n\times n},\quad F(t) = e^{t\cdot A}\cdot e^{t\cdot -A}$$

je konstantní, jelikož $F(0) = e^{O_{n\times n}}\cdot e^{O_{n\times n}} = I_n\cdot I_n = I_n$.

Platí

$$\frac{d}{dt}F(t) = A\cdot e^{t\cdot A}\cdot e^{t\cdot(-A)} + e^{t\cdot A}\cdot(-A)\cdot e^{t\cdot(-A)} =^{(1)} A\cdot e^{t\cdot A}\cdot e^{t\cdot(-A)} - A\cdot e^{t\cdot A}\cdot e^{t\cdot(-A)} = O_{n\times n}.$$

Tedy, každá složka matice $F(t)$ má nulovou derivaci, a je tedy konstantní. Tím pádem je i maticová funkce $F(t)$ konstantní, a tedy rovna I_n pro každé $t\in\mathbb{R}$.

□

Věta 13.2. *Nechť $A \in \mathbb{R}^{n\times n}$ a $\vec{c}\in\mathbb{R}^n$. Potom jediné řešení soustavy diferenciálních rovnic*

$$\vec{u}' = A\cdot\vec{u},\quad \vec{u}(0) = \vec{c},$$

je vektorová funkce $\vec{u}(t) = e^{t\cdot A}\cdot\vec{c},\quad t\in\mathbb{R}$.

Důkaz. Nejprve ukážeme, že funkce $\vec{u}(t) = e^{t\cdot A}\cdot\vec{c}$ je řešením dané soustavy. To je snadné: dle Tvrzení 13.1 platí

$$\vec{u}'(t) = A\cdot e^{t\cdot A}\cdot\vec{c} = A\cdot\vec{u}(t),\quad \text{a}\quad \vec{u}(0) = I_n\cdot\vec{c} = \vec{c}.$$

Nyní ukážeme jednoznačnost řešení. Předpokládejme, že vektorová funkce $\vec{v}\colon t\to\mathbb{C}^n$ je řešením dané soustavy. Potom

$$\frac{d}{dt}(e^{t\cdot(-A)}\cdot\vec{v}(t)) = -A\cdot e^{t\cdot(-A)}\cdot\vec{v}(t) + e^{t\cdot(-A)}\cdot\vec{v}'(t) =$$

$$= e^{t\cdot(-A)}\cdot -A\cdot\vec{v}(t) + e^{t\cdot(-A)}\cdot\vec{v}'(t) = e^{t\cdot(-A)}\cdot(\vec{v}'(t) - A\cdot\vec{v}(t)) = O_{n\times n}$$

Tedy $e^{t\cdot(-A)}\cdot\vec{v}(t)$ je konstantní vektor v $\mathbb{C}^n$, a navíc $e^{0\cdot(-A)}\cdot\vec{v}(0) = I_n\cdot\vec{c} = \vec{c}$.
Tedy pro každé $t\in\mathbb{R}$,

$$\vec{c} = e^{t\cdot(-A)}\cdot\vec{v}(t) = (e^{t\cdot A})^{-1}\cdot\vec{v}(t) \Leftrightarrow e^{t\cdot A}\cdot\vec{c} = \vec{v}(t),$$

čímž je důkaz dokončen. □

Při modelování fyzikálních systémů diferenciálními rovnicemi nás často místo přesného řešení zajímá pouze chování řešení v nekonečnu, tj. zda se systém ustálí v nějakém rovnovážném stavu nebo ne. Jelikož víme, že řešení soustavy $\vec{u}' = A\cdot\vec{u},\quad \vec{u}(0) = \vec{c}$ je vektorová funkce

$$\vec{u}(t) = e^{t\cdot A}\cdot\vec{c} = \sum_{i=1}^{s}\sum_{j=0}^{k_i-1}\frac{t^j\cdot e^{\lambda_i\cdot t}}{j!}\cdot(A-\lambda_i I_n)^j\cdot G_i\cdot\vec{c},$$

kde $\sigma(A) = \{\lambda_1, \ldots, \lambda_s\}$ a $k_i = ind_{\lambda_i}(A), \quad i = 1, \ldots, s$, a G_i jsou spektrální projekce matice A, můžeme o stabilitě řešení soustavy rozhodnout pouze na základě vlastních hodnot matice A: pro každé $i = 1, \ldots, s$, chování vektorové funkce $\sum_{j=0}^{k_i-1} \frac{t^j \cdot e^{\lambda_i \cdot t}}{j!} \cdot (A - \lambda_i I_n)^j \cdot G_i \cdot \vec{c}$ příslušné spektrální projekci G_i pro $t \to \infty$ je určeno funkcí $t^{k_i-1} \cdot e^{\lambda_i \cdot t}$ (protože u tohoto členu řady se t objevuje v nejvyšší mocnině).
Zvolme tedy libovolné $\lambda \in \sigma(A)$, nechť $k = ind_\lambda(A)$, a analyzujme chování funkce $t^{k-1} \cdot e^{\lambda \cdot t}$ pro $t \to \infty$. Pišme $\lambda = a + bi$, kde $a = Re(\lambda)$ a $b = Im(\lambda)$, potom

$$t^{k-1} \cdot e^{\lambda \cdot t} = t^{k-1} \cdot e^{a \cdot t} \cdot (\cos(bt) + i \sin(bt)) \to^{t \to \infty} \begin{cases} 0, & a < 0, \\ \text{neomezené}, & a > 0 \text{ nebo } (a = 0 \text{ a } k > 1), \\ \text{osciluje}, & a = 0, k = 1, \text{ a } b \neq 0, \\ 1, & a = b = 0 \text{ a } k = 1. \end{cases}$$

Tedy, pokud $Re(\lambda) < 0$ pro každé $\lambda \in \sigma(A)$, potom pro každé $\vec{c} \in \mathbb{R}^n$ řešení soustavy $\vec{u}' = A \cdot \vec{u}, \quad \vec{u}(0) = \vec{c}$, splňuje $\lim_{t \to \infty} \vec{u}(t) = \vec{o}$, v opačném případě mohou řešení divergovat, oscilovat, nebo být konstantní nenulové. Na základě toho je založena následující definice.

Definice 13.3. *Nechť $A \in \mathbb{R}^{n \times n}$. Potom soustava $\vec{u}' = A \cdot \vec{u}$ se nazývá:*

- stabilní *(nebo* asymptoticky stabilní*), pokud $Re(\lambda) < 0$ pro každé $\lambda \in \sigma(A)$,*
- nestabilní*, pokud existuje $\lambda \in \sigma(A)$ takové, že $Re(\lambda) > 0$ nebo ($Re(\lambda) = 0$ a zároveň $ind_\lambda(A) > 1$).*

Ve zbylém případě se soustava někdy nazývá semistabilní*.*

13.2 Nehomogenní soustavy

Připomeňte si, jak řešit nehomogenní diferenciální rovnici $u' = a \cdot u + f, \quad u(0) = c$ pomocí takzvané *variace konstant*. Ta spočívá v tom, že nejprve vyřešíme příslušnou homogenní rovnici $u' = a \cdot u, \quad u(0) = c$, která, jak víme, má řešení $u_h(t) = e^{at} \cdot c$. Nyní, variace konstant spočívá v tom, že pro původní rovnici nalezneme takzvané *partikulární řešení* tak, že v řešení homogenní rovnice uvažujeme funkci $c(t)$ namísto skaláru c (řešení původní rovnice je pak součtem řešení homogenní rovnice a partikulárního řešení). Pro nalezení partikulárního řešení tedy hledáme řešení ve tvaru $u_p(t) = e^{at} \cdot c(t)$. Tento výraz dosadíme do původní rovnice, a dostaneme rovnost

$$ae^{at} \cdot c(t) + e^{at} \cdot c'(t) = a \cdot e^{at} \cdot c(t) + f(t) \Rightarrow c'(t) = e^{-at} \cdot f(t),$$

z čehož plyne (až na konstantu) rovnost $c(t) = \int_0^t e^{-as} \cdot f(s) ds$. Partikulární řešení je tedy

$$u_p(t) = e^{at} \cdot c(t) = e^{at} \cdot \int_0^t e^{-as} \cdot f(s) ds,$$

a celkové řešení je

$$u(t) = u_h(t) + u_p(t) = e^{at} \cdot c + e^{at} \cdot \int_0^t e^{-as} \cdot f(s) ds.$$

Analogicky dostaneme následující větu.

Věta 13.4. *Nechť $A \in \mathbb{R}^{n \times n}$, $\vec{c} \in \mathbb{R}^n$, a $\vec{f}\colon \mathbb{R} \to \mathbb{R}^n$ je spojitá vektorová funkce. Potom jediné řešení soustavy diferenciálních rovnic*

$$\vec{u}' = A \cdot \vec{u} + \vec{f}, \quad \vec{u}(0) = \vec{c},$$

je vektorová funkce

$$\vec{u}(t) = \overbrace{e^{t \cdot A} \cdot \vec{c}}^{\text{řešení homogenní rovnice}} + \overbrace{e^{t \cdot A} \cdot \int_0^t e^{-s \cdot A} \cdot \vec{f}(s) ds}^{\text{partikulární řešení}}.$$

Důkaz. Nejprve ověříme, že daná vektorová funkce $\vec{u}$ je řešením dané soustavy. Máme

$$\vec{u}'(t) = A \cdot e^{t \cdot A} \cdot \vec{c} + A \cdot e^{t \cdot A} \cdot \int_0^t e^{-s \cdot A} \cdot \vec{f}(s) ds + e^{t \cdot A} \cdot e^{-t \cdot A} \cdot \vec{f}(t) = A \cdot \vec{u}(t) + \vec{f}(t),$$

a $\vec{u}(0) = \vec{c} + \vec{o} = \vec{c}$.
Nyní ukážeme jednoznačnost řešení. Předpokládejme, že vektorová funkce $\vec{v}\colon t \to \mathbb{R}^n$ je řešením dané soustavy. Potom

$$\frac{d}{dt}(e^{-t \cdot A} \cdot \vec{v}(t)) = -A \cdot e^{-t \cdot A} \cdot \vec{v}(t) + e^{-t \cdot A} \cdot \vec{v}'(t) =$$

$$=^{\text{Tvrzení13.1(1)}} e^{-t \cdot A} \cdot -A \cdot \vec{v}(t) + e^{-t \cdot A} \cdot \vec{v}'(t) = e^{-t \cdot A} \cdot (\vec{v}'(t) - A \cdot \vec{v}(t)) = e^{-t \cdot A} \cdot \vec{f}(t).$$

Tedy, zintegrováním obou stran rovnosti dostaneme

$$e^{-t \cdot A} \cdot \vec{v}(t) - e^{-0 \cdot A} \cdot \vec{v}(0) = \int_0^t e^{-s \cdot A} \cdot \vec{f}(s) ds,$$

tedy, jelikož $\vec{v}(0) = \vec{c}$,

$$e^{-t \cdot A} \cdot \vec{v}(t) = \vec{c} + \int_0^t e^{-s \cdot A} \cdot \vec{f}(s) ds,$$

a tedy, po přenásobení maticí $e^{t \cdot A}$ zleva máme

$$\vec{v}(t) = e^{t \cdot A} \cdot \vec{c} + e^{t \cdot A} \cdot \int_0^t e^{-s \cdot A} \cdot \vec{f}(s) ds,$$

což jsme chtěli dokázat. Důkaz je ukončen. □