Bakalářská práce WIP

Benjamin de Silva

October 2023

Úvod

Motivace

1 Definice (hlavní zdroj - linalg skripta

V této kapitole si zavedeme základní definice potřebné pro práci s maticemi, operace na nich a různé typy matic, které budeme používat.

Definice 1.1 (Matice)**.** *Maticí* typu m × n nad tělesem **T** rozumíme obdelníkové schéma prvků z tělesa **T** s m řádky a n sloupci. Matice typu m × m se nazývá *čtvercová matice řádu m.* Matice typu m × 1 se nazýva *sloupcový aritmetický vektor* a matice typu 1 × m se nazývá *řádkový aritmetický vektor*. Zápisem $A = (a_{ij})_{m\times n}$ rozumíme matici typu $m \times n$ s prvkem $a_{ij} \in T$ na pozici (i,j).

V rámci této práce budeme uvažovat matice nad tělesem $T = \mathbb{C}$

Definice 1.2 (Typy matic)**.** Nechť A je čtvercová matice řádu n. Pak řekneme, že

- A je *nulová*, pokud $a_{ij} = 0, \forall i, j \in \{1, ..., n\}$,
- A je *jednotková*, pokud $a_{ij} = 0$ pro $i = j$ a $a_{ij} = 0$ jinak,
- A je *diagonální*, pokud $a_{ij} = 0$ pro $i \neq j$,
- A je *horní trojuhelníková*, pokud $a_{ij} = 0$ pro $i > j$,
- A je *dolní trojuhelníková*, pokud $a_{ij} = 0$ pro $i < j$,
- A je *permutační*, pokud má v každém řádku a sloupci právě jeden prvek 1 a zbytek jsou nuly,
- A je *symetrická*, pokud $a_{ij} = a_{ji}$,$\forall i, j \in \{1, ..., n\}$,
- A je *řídká*, pokud většina prvků matice je nulových.

Značení. Budeme značit jednotkovou matici $\mathbf{I_{m\times n}}$ a nulovou matici $\mathbf{0_{m\times n}}$, jsou-li obě matice typu $m \times n$. Pokud jsou matice čtvercové řádu n, tak značíme $\mathbf{I_n}, \mathbf{0_n}$ a v případě, že je to z kontextu zřejmé, tak index může být vynechán.

Definice 1.3 (Sčítání matic)**.** Nechť A,B jsou matice stejného typu m × n, pak definujeme jejich součet jako matici

$$A + B = (a_{ij} + b_{ij})_{m\times n}$$

Nechť A,B,C jsou matice stejného typu, pak operace sčítání matic má následující vlastnosti

- $(A + B) + C = A + (B + C)$,
- $A + 0 = A$,
- $A + (-A) = 0$,
- $A + B = B + A$.

Definice 1.4 (Násobení matice skalárem)**.** Nechť A je matice typu $m \times n$ a $t \in \mathbb{R}$. Pak definujeme t-násobek matice A jako matici

$$t \cdot A = tA = (t \cdot a_{ij})_{m \times n}$$

Nechť A,B jsou matice stejného typu a $s, t \in \mathbb{R}$. Pak platí následující vlastnosti

- $s(tA) = (st)A$,
- $1A = A$,
- $-A = (-1)A$,
- $(s + t)A = sA + tA$,
- $s(A + B) = sA + sB$.

Definice 1.5 (Transponovaná matice)**.** Nechť A je matice typu $m \times n$, pak transponovanou maticí k matici A rozumíme matici

$$A^T = (b_{ij})_{n \times m},$$

kde $b_{ij} = a_{ji}$ pro $i \in \{1, \ldots, n\}, j \in \{1, \ldots, m\}$.

Z definice symetrické matice a transponované matice plyne, že pro A symetrickou platí $A^T = A$.

Nechť A,B jsou matice stejného typu $m \times n$ a $s \in \mathbb{C}$, pak platí

- $(A^T)^T = A$
- $(A^T + B^T) = A^T + B^T$
- $(sA)^T = sA^T$

Značení. Někdy je vhodné nahlížet na matici jako na posloupnost sloupcových vektorů, tedy matici $A = (a_{ij})_{m \times n}$ můžeme zapsat jako $A = (\mathbf{a_1}|\mathbf{a_2}|...|\mathbf{a_n})$, kde $\mathbf{a_i} = (a_{1i}, a_{2i}, \ldots, a_{mi})^T, i \in \{1, \ldots, n\}$ je i-tý m-složkový sloupcový aritmetický vektor matice A,

Definice 1.6 (Součin matice a vektoru)**.** Nechť $A = (\mathbf{a_1}|\mathbf{a_2}|...|\mathbf{a_n})$ je matice typu $m \times n$ a $b = (b_1, b_2, \ldots, b_n)^T$ je aritmetický sloupcový vektor. Pak definujeme součin matice A s vektorem b jako

$$A \cdot b = b_1\mathbf{a_1} + b_2 \cdot \mathbf{a_2} + \ldots + b_n\mathbf{a_n},$$

což je m-složkový sloupcový vektor.

Definice 1.7 (Součin dvou matic)**.** Nechť A je matice typu $m \times n$ a $B = (\mathbf{b_1}, \ldots, \mathbf{b_k})$ je matice typu $n \times k$. Pak součin matic A a B je

$$AB = (A\mathbf{b_1}, A\mathbf{b_2}, \ldots, A\mathbf{b_k}),$$

což je matice typu $m \times k$.

Ze zavedených definic už můžeme začít řešit úlohu soustav lineárních rovnic. Mějme zadanou matici A typu $m \times n$ a $b = (\mathbf{b_1}, \dots, \mathbf{b_m})^T$ m-složkový aritmetický vektor. Pak hledáme $x = (\mathbf{x_1}, \dots, \mathbf{x_n})^T$ sloupcový vektor takový, že platí rovnost $\mathbf{Ax} = \mathbf{b}$. Každý takový vektor $\mathbf{x}$ nazveme řešením soustavy linearních rovnic matice A s pravou stranou $\mathbf{b}$.

Definice 1.8 (Inverzní matice)**.** Nechť A je matice typu $m \times n$. Pak

- matici X nazveme inverzní k A zprava, pokud platí $AX = I_m$
- matici X nazveme inverzní k A zleva, pokud platí $AX = I_n$
- je-li A čtvercová řádu n a matice X taková, že $AX = XA = I_n$, pak X nazveme inverzní k matici A, matici X značíme A^{-1} a matici A nazveme invertibilním

Definice 1.9 (Regulární matice)**.** Matici A řádu n nazveme regulární, pokud je zobrazení $f_A : \mathbb{C}^n \to \mathbb{C}^n, x \mapsto Ax$ určené maticí A bijektivní. Matice A, která není regulární, se nazývá singulární.

Pro čtvercovou matici A platí následující ekvivalence

- Matice A je regulární,
- Zobrazení f_A je na
- Zobrazení f_A je prosté
- Soustava $A\mathbf{x} = 0$ má právě jedno řešení $\mathbf{x} = 0$
- Matice A je invertovatelná
- Existuje matice X tak, že $AX = I_n$
- Existuje matice Y tak, že $YA = I_n$
- $rank(A) = n$

Definice 1.10 (Obor hodnot, jádro, hodnost)**.** Pro matici A typu $m \times n$ definujeme

$$\mathbf{R}(A) = \{A\mathbf{x} : x \in \mathbb{R}^n\}\ ,\ \mathbf{N}(A) = \{\mathbf{x} \in \mathbf{C}^n\}\ ,\ rank(A) \equiv dim(\mathbf{C}(A),$$

kde $R(A)$ nazveme *oborem hodnot* matice A, $\mathbf{N}(A)$ nazveme *jádrem* matice A a rank(A) *hodností* matice A.

2 Metody pro hledání nulového prostoru (zdroj-kniha se stromem

V této kapitole popíšeme různé metody pro hledání nulového prostoru matic. Mějme tedy nějakou matici A typu $m \times n$ řádu m, pak hledáme nějakou matici B typu $n \times (n-m)$ řádu $n-m$ tak, že platí $AB = 0$, tedy sloupce matice B tvoří bázi nulového prostoru matice A.

Pro každou metodu bude výhodné nejprve rozložit matici A a poté matici B spočítáme za použití získaného rozkladu. Abychom zachovali numerickou stabilitu a řídkost, tak je někdy nutné prohodit některé řádky a sloupce za pomocí permutačních matic, tedy ve skutečnosti hledáme rozklad matice $P_1 E P_2$.

2.1 Singulární rozklad

Singulární rozklad (zkráceně SVD - singular value decomposition) nám umožňuje určit hodnost či normu matice, ortogonální bázi oboru hodnot a nulového prostoru matice, nicméně má vysoké výpočetní nároky. Před zavedením singulárního rozkladu si nejprve zavedeme spektrální rozklad.

2.1.1 Spektrální rozklad

Věta 2.1. *Nechť $A \in \mathbb{C}^{m\times n}$, potom platí*

$$N(A) \bigoplus R(A^*) = \mathbb{C}^n, \;\; N(A) \perp R(A^*),$$

$$N(A^*) \bigoplus R(A) = \mathbb{C}^m, \;\; N(A^*) \perp R(A).$$

Věta 2.2 (Spektrální rozklad pro normální matice)**.** *Matice A je normální právě tehdy, když existuje unitární matice U a diagonální matice D tak, že*

$$U^* A U = D, \;\; tj. \;\; A = UDU^*$$

Předpokládejme, že A je hermitovská pozitivně semidefinitní řádu n s hodností r s nezápornými vlastními čísly $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_r > 0$, $\lambda_{r+1} = \ldots = \lambda_n = 0$. Podle věty 2.1 existuje rozklad matice A ve tvaru

$$A = Q\Lambda Q^*, \;\; Q^*Q = I, \;\; \Lambda = diag(\lambda_1, \ldots, \lambda_n).$$

Sloupce unitární matice $Q = (q_1, \ldots, q_n)$ jsou vlastní vektory matice A a platí

$$Aq_j = \lambda_j q_j, \;\; j = 1, \ldots, n$$

Vlastní vektory $q_1, q_2, \ldots, q_n$ tvoří ortonormální bázi prostoru $\mathbb{C}^n$.

Ze spektrálního rozkladu matice A vidíme bázi nulového prostoru a oboru hodnot matice A,

$$R(A) = span\{q_1, \ldots, q_r\}, \;\; N(A) = span\{q_{r+1}, \ldots, q_n\},$$

dále nám dává dyadický rozvoj matice A

$$A = \sum_{j=1}^{n} \lambda_j q_j q_j^* = \sum_{j=1}^{m} A_j, \;\; A_j = \lambda_j q_j q_j^*,$$

kde matice A_j jsou uspořádány podle spektrální (či Frobeniovy) normy matice a platí $\|A_j\| = \|A_j\|_F = \lambda_j$, tedy můžeme matici A aproximovat maticí nižší hodnosti

$$A^{(k)} = \sum_{j=1}^{k} A_j.$$

Pro odvození singulárního rozkladu budou důležité spektrální rozklady matic A^*A a AA^*. Matice A^*A a AA^* jsou čtvercové hermitovské pozitivně semidefinitní a můžeme uvažovat jejich spektrální rozklad.

Lemma 2.3. Pro komplexní matici $A \in \mathbb{C}^{m \times n}$ platí

$$N(A^*A) = N(A), \;\; R(A^*A) = R(A^*),$$

$$N(AA^*) = N(A^*), \;\; R(AA^*) = R(A).$$

Důsledek **Věty 2.1** a **Lemma 2.3** je, že

$$dim(R(AA^*)) = dim(R(A)) = r = dim(R(A^*)) = dim(R(A^*A)),$$

pro A hodnosti r.

Buď λ_j vlastní čísla a v_j ortonormální vlastní vektory matice A^*A, $j = 1, \ldots, n$. Uvažujme dále bez újmy na obecnosti

$$\lambda_1 \geq \ldots \geq \lambda_r > 0, \;\; \lambda_{r+1} = \lambda_n = 0,$$

pak pro $V = (v_1, \ldots, v_n)$ platí

$$V^*A^*AV = diag(\lambda_1, \ldots, \lambda_r, 0, \ldots, 0)$$

Dále z **Lemma 2.3** platí, že $v_1, \ldots, v_r$ resp. $v_{r+1}, \ldots, v_n$ tvoří ortonormální bázi

$$R(A^*A) = R(A^*) = span\{v_1, \ldots, v_r\},$$

resp.

$$N(A) = N(A^*A) = span\{v_{r+1}, \ldots, v_n\}.$$

Následující lemma nám dává vztah mezi ortonormálními vektory matic A^*A a AA^*.

Lemma 2.4. Uvažujme spektrální rozklad matice A^*A s vlastními čísly λ_j a v_j, $j = 1, \ldots, r$. Potom jsou vektory

$$u_j \equiv Av_j / \sqrt{\lambda_j}, \;\; j = 1, \ldots, r$$

ortonormální vlastní vektory AA^* a platí

$$AA^*v_j = \lambda_j u_j, \ \ \|u_j\| = 1, \ \ j = 1, \dots, r.$$

Z **Lemma 2.4** plyne, že vektory $u_j, \ j = 1, \dots, r$ jsou ortonormální vlastní vektory matice AA^*, které odpovídají právě m nenulovým vlastním číslům λ_j a platí

$$Av_j = \sqrt{\lambda_j} u_j, \ \ j = 1, \dots, r.$$

Analogie těchto vztahů platí i pro matici AA^*, tj. pokud $u_1, \dots, u_r$ tvoří ortonormální bázi oboru hodnot matice a $u_{r+1}, \dots, u_m$ tvoří ortonormální bázi doplňku oboru hodnot, pak platí

$$span\{u_1, \dots, u_r\} = R(AA^*) = R(A),$$

$$span\{u_{r+1}, \dots, u_m\} = N(AA^*) = N(A^*).$$

Z konstrukce plyne, že nenulová vlastní čísla matic A^*A a AA^* se rovnají včetně násobnosti.

Nyní můžeme popsat singulární rozklad.

Definice 2.5 (Singulární čísla)**.** Nechť $A \in \mathbb{C}^{m \times n}$. Odmocniny vlastních čísel matice A^*A nazveme singulárnímy čísly matice A,

$$\sigma_j = \sqrt{\lambda_j}, \ \ j = 1, \dots, r, \ \ r = rank(A).$$

Jelikož vlastní čísla λ_j jsou uspořádáná od největšího po nejmenší, tak platí

$$\sigma_1 \geq \ldots \geq \sigma_r > 0.$$

Tedy podle předchozích poznatků platí $Av_j = \sigma_j u_j, \ j = 1, \dots, r$ a můžeme psát

$$AV = U\Sigma,$$

kde

$$\Sigma = \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix} \in \mathbb{R}^{m \times n}, \ \ \Sigma_r = diag(\sigma_1, \dots, \sigma_r) \in \mathbb{R}^{r \times r}, \ \ ,$$

$$V = (v_1, \dots, v_n).U = (u_1, \dots, u_m).$$

Ekvivalentně

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

Vidíme tedy, že pokud bychom byli schopni spočítat singulární rozklad matice A, tak nulový prostor podle **Lemma 2.3** je roven lineárnímu obalu vektorů $v_{r+1}, \dots, v_n$.

Nyní pomocí SVD můžeme vyřešit úlohu hledání matice B splňující $AB = 0$ následovně

$$P_1 A P_2 = U\Sigma V,$$

$$[V_1, V_2] = V^T,$$

$$B = P_2 Q_2$$

SVD můžeme vypočítat za pomocí spektrálních rozkladů matic AA^* a A^*A, podmíněnost obou matic rovna $\kappa(A)^2$, tedy pokud by byla matice A špatně podmíněna, tak by mohla být singulární čísla spočtena nepřesně.

Věta 2.6. *Mějme matici $A \in \mathbb{R}^{m\times n}$ řádu m a její singulární čísla $\sigma_1 \geq \ldots \geq \sigma_m > 0$. Potom platí $\|A\| = \sigma_1$, $\|A\|_F = (\sigma_1^2 + \sigma_2^2 + \ldots + \sigma_m^2)^{1/2}$. Dále*

$$\|A\| \leq \|A\|_F \leq \sqrt{m}\,\|A\|.$$

Jelikož platí, že

$$\|A\| = \sigma_1,\ a\ \left\|A^-1\right\| = \sigma_m^-1,$$

tak dostáváme

$$\kappa(A) = \sigma_1/\sigma_n.$$

Další standardní metoda výpočtu SVD probíhá ve dvou krocích

1. Transformace matice A na bidiagonální tvar užitím unitárních transformací nebo Golub-Kahanovy iterační bidiagonalizace. Transformovaná matice A je ve tvaru

   $$B = PAQ,$$

   kde P,Q jsou unitární a B je bidiagonální.

2. Nalezení singulární rozkladu bidiagonální matice B pomocí iteračního procesu, který nuluje její naddiagonální prvky, například pomocí implicitního QR algoritmu.

Algoritmus (Rozklad SVD pomocí Golub-Kahanovy iterační bidiagonalizace)**.** Nejprve provedeme Golub-Kahanovu iterační bidiagonalizaci

Vstup $A,\ v$

$w_0 := 0$

$\delta_1 := 0$

$s_1 := v/\delta_1$

for $k = 1, 2, \ldots$

  $p := A^* s_k - \delta_k w_{k-1}$

  $\gamma_k := \|p\|$

  $w_k := p/\gamma_k$

  $q := Aw_k - \gamma_k s_k$

  $\delta_{k+1} := \|q\|$

  $s_{k+1} := q/\delta_{k+1}$

end

Iterace končí v k-tém kroce, pokud $\gamma_{k+1} = 0$ nebo $\delta_{k+1} = 0$. Pokud $\delta_{k+1} = 0$, pak platí

$$AW_k = S_k L_k,$$

kde

$$W_k \equiv [w_1, \ldots, w_k], \; L_k \equiv \begin{bmatrix} \gamma_1 & & & \\ \delta_2 & \gamma_2 & & \\ & \ddots & \ddots & \\ & & \delta_k & \gamma_k \end{bmatrix}, \; S_k = [s_1, \ldots, s_k].$$

Platí, že vektory $w_1, \ldots, w_{k+1}$ a $s_1, \ldots, s_{k+1}$ jsou ortonormální. Máme tedy rozklad A na bidiagonální matici

$$L_k = S_k^* A W_k$$

a W_k a S_k jsou ortonormální.

Výpočetní náročnost závisí na iteračním procesu, kterým počítáme singulární rozklad bidiagonální matice. Pokud bychom předpokládali, že máme pouze čtvercové matice řádu m, iterační proces rychle konverguje a nezvyšuje počet operací, pak dolní odhad na počet operací je $\frac{16}{3} m^3$.

2.2 LU rozklad

Jednou z nejdůležitejších metod pro řešení soustav lineárních rovnic je LU rozklad, který nám původní matici A rozloží na součin dolní trojuhelníkové matice a horní trojuhelníkové matice. LU rozklad je založen na Gaussově eliminaci následovně

$$A \longrightarrow A^{(1)} \equiv \begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ 0 & a_{2,2}^1 & \dots & a_{2,n}^1 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{m,2}^1 & \dots & a_{m,n}^1 \end{bmatrix} = M^{-1}A,$$

kde

$$M_1^{-1} = I - m_1 e_1^T = \begin{bmatrix} 1 & & & & \\ -\frac{a_{2,1}}{a_{1,1}} & 1 & & & \\ -\frac{a_{3,1}}{a_{1,1}} & & 1 & & \\ \vdots & & & \ddots & \\ -\frac{a_{m,1}}{a_{1,1}} & & & & 1 \end{bmatrix}, \quad m_1 = \begin{bmatrix} \frac{a_{2,1}}{a_{1,1}} \\ \frac{a_{3,1}}{a_{1,1}} \\ \vdots \frac{a_{m,1}}{a_{1,1}} \end{bmatrix}.$$

Jelikož A má hodnost m, tak permutováním zaručíme, že nikdy nedělíme 0, tedy LU rozklad lze provést. V i-tém kroku bude

$$A^{(i)} \equiv \begin{bmatrix} a_{1,1} & a_{1,2} & \dots & & & a_{1,n} \\ 0 & a_{2,2}^{(1)} & \dots & & & a_{2,n}^{(1)} \\ \vdots & 0 & \ddots & & & \\ & \vdots & & a_{i,i}^{(i-1)} & \dots & a_{i,n}^{(i-1)} \\ 0 & 0 & & a_{m,i}^{(i-1)} & \dots & a_{m,n}^{(i-1)} \end{bmatrix}, \quad m_i = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \frac{a_{i+1,i}}{a_{i,i}} \\ \frac{a_{i+2,i}}{a_{i,i}} \\ \vdots \\ \frac{a_{m,i}}{a_{i,i}} \end{bmatrix}.$$

Po m-1 krocích dostaneme

$$U \equiv A^{(m-1)} = M_{m-1}^{-1} \dots M_2^{-1} M_1^{-1} A, \quad L \equiv M_1 M_2 \dots M_{m-1},$$

kde U je horní trojuhelníková typu $m \times n$ a L je dolní trojuhelníková řádu m s jednotkami na diagonále a upravením rovnosti dostáváme

$$A = LU = L[U_1\ U_2], \quad U_1 = [u_1, \dots, u_m], \quad U_2 = [u_{m+1}, \dots, u_n].$$

Nyní můžeme LU rozklad použít k řešení naší úlohy $AB = 0$

$$P_1 A P_2 = L[U_1 U_2], J = U_1^{-1} U_2, B = P_2 \begin{bmatrix} -J \\ I_{n-m} \end{bmatrix}.$$

2.3 Gauss-Jordanova eliminace

2.4 Sloupcová eliminace

2.5 QR rozklad

2.6 LQ rozklad