

Optimalizace

4. Spektrální rozklad a kvadratické funkce

Tom Werner, Tom Kroupa

FEL ČVUT

Spektrální rozklad

Vlastní čísla a vektory

Nechť pro matici $\mathbf{A} \in \mathbb{R}^{n \times n}$, nenulový vektor $\mathbf{v} \in \mathbb{C}^n$ a $\lambda \in \mathbb{C}$ platí

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v}.$$

Pak λ je **vlastní číslo** matice \mathbf{A} a \mathbf{v} je **vlastní vektor** příslušný λ .

Můžeme přepsat jako

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{v} = \mathbf{0}$$

- λ je vlastní číslo, právě když je to kořen **charakteristického polynomu**:

$$p_{\mathbf{A}}(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I}) = \prod_{i=1}^n (\lambda_i - \lambda) = 0$$

Je to polynom stupně n , tedy \mathbf{A} má n vlastních čísel (počítaje násobnost).

- Množina vlastních vektorů příslušných jednomu vlastnímu číslu λ je

$$\text{null}(\mathbf{A} - \lambda\mathbf{I}) \setminus \{\mathbf{0}\}$$

Diagonalizovatelné matice

Nechť $\lambda_1, \dots, \lambda_n$ jsou vlastní čísla matice a $\mathbf{v}_1, \dots, \mathbf{v}_n$ jim příslušné vlastní vektory.

Soustavu

$$\mathbf{A}\mathbf{v}_i = \lambda_i\mathbf{v}_i, \quad i = 1, \dots, n$$

lze napsat jako

$$\mathbf{A}\mathbf{V} = \mathbf{V}\mathbf{\Lambda}$$

kde

$$\mathbf{V} = [\mathbf{v}_1 \ \cdots \ \mathbf{v}_n], \quad \mathbf{\Lambda} = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

- V Matlabu: $[\mathbf{V}, \mathbf{D}] = \text{eig}(\mathbf{A})$ (kde \mathbf{D} je $\mathbf{\Lambda}$)
- Jestliže \mathbf{V} je regulární, pak

$$\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}$$

$$\mathbf{\Lambda} = \mathbf{V}^{-1}\mathbf{A}\mathbf{V}$$

Tedy \mathbf{A} je **diagonalizovatelná** (= podobná diagonální matici).

Spektrální rozklad symetrické matice

Věta (Spektrální věta)

Nechť matice $\mathbf{A} \in \mathbb{R}^{n \times n}$ je symetrická ($\mathbf{A}^T = \mathbf{A}$). Pak:

- Všechna vlastní čísla matice \mathbf{A} jsou reálná.
- Existuje ortonormální množina n vlastních vektorů matice \mathbf{A} .

Některé pojmy z komplexní algebry:

- *Komplexně sdružené* číslo k číslu $z = a + bi \in \mathbb{C}$ je číslo $\bar{x} = a - bi$.
- *Absolutní hodnota* čísla $x \in \mathbb{C}$ je reálné číslo $|x| = (\bar{x}x)^{1/2} = (a^2 + b^2)^{1/2}$.
- *Adjungovaná* matice k matici $\mathbf{B} \in \mathbb{C}^{m \times n}$ je matice $\mathbf{B}^* = \bar{\mathbf{B}}^T \in \mathbb{C}^{n \times m}$.
- *Skalární součin* vektorů $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$ je $\mathbf{x}^* \mathbf{y} = \bar{x}_1 y_1 + \dots + \bar{x}_n y_n$. Platí $\mathbf{x}^* \mathbf{y} = \overline{\mathbf{y}^* \mathbf{x}}$.
- *Norma* vektoru $\mathbf{x} \in \mathbb{C}^n$ je $\|\mathbf{x}\| = (\mathbf{x}^* \mathbf{x})^{1/2} = (|x_1|^2 + \dots + |x_n|^2)^{1/2}$.

Dokážeme jen první tvrzení: Nechť $\mathbf{A}^* = \mathbf{A}$ a $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$. Pak

$$\lambda \|\mathbf{v}\|^2 = \lambda \mathbf{v}^* \mathbf{v} = \mathbf{v}^* \mathbf{A} \mathbf{v} = \mathbf{v}^* \mathbf{A}^* \mathbf{v} = (\mathbf{A} \mathbf{v})^* \mathbf{v} = (\lambda \mathbf{v})^* \mathbf{v} = \bar{\lambda} \mathbf{v}^* \mathbf{v} = \bar{\lambda} \|\mathbf{v}\|^2.$$

Když $\mathbf{v} \neq \mathbf{0}$, pak tedy $\lambda = \bar{\lambda}$, tedy $\lambda \in \mathbb{R}$.

Důsledek (spektrální rozklad symetrické matice)

Je-li $\mathbf{A} \in \mathbb{R}^{n \times n}$ symetrická, ve spektrálním rozkladu lze \mathbf{V} zvolit ortogonální a

$$\mathbf{A} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^{-1} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^T = \lambda_1 \mathbf{v}_1 \mathbf{v}_1^T + \dots + \lambda_n \mathbf{v}_n \mathbf{v}_n^T$$

Kvadratické formy

Kvadratická forma

Funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ je **kvadratická forma**, je-li to je homogenní polynom stupně 2. V maticové formě:

$$f(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x} = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j \quad \text{pro nějakou } \mathbf{A} \in \mathbb{R}^{n \times n}$$

Tvrzení

Pro každou kvadratickou formu $f: \mathbb{R}^n \rightarrow \mathbb{R}$ existuje *symetrická* matice $\mathbf{A} \in \mathbb{R}^{n \times n}$ taková, že $f(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x}$.

Důkaz: Pro každou čtvercovou \mathbf{A} je

$$\mathbf{A} = \underbrace{\frac{1}{2}(\mathbf{A} + \mathbf{A}^T)}_{\text{symetrická}} + \underbrace{\frac{1}{2}(\mathbf{A} - \mathbf{A}^T)}_{\text{antisymetrická}}.$$

Ale $\mathbf{x}^T(\mathbf{A} - \mathbf{A}^T)\mathbf{x} = \mathbf{x}^T \mathbf{A} \mathbf{x} - \mathbf{x}^T \mathbf{A}^T \mathbf{x} = \mathbf{x}^T \mathbf{A} \mathbf{x} - (\mathbf{x}^T \mathbf{A} \mathbf{x})^T = 0$ (protože $\mathbf{x}^T \mathbf{A} \mathbf{x}$ je skalár).

Tedy $\mathbf{x}^T \mathbf{A} \mathbf{x} = \frac{1}{2} \mathbf{x}^T (\mathbf{A} + \mathbf{A}^T) \mathbf{x}$.

Příklad:

$$f(x, y) = 2x^2 - 2xy + y^2 = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 2 & -2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

Definitnost kvadratické formy / její matice

Kvadratická forma f je

- **pozitivně semidefinitní**, když $f(\mathbf{x}) \geq 0$ pro každé \mathbf{x}
- **negativně semidefinitní**, když $f(\mathbf{x}) \leq 0$ pro každé \mathbf{x}
- **pozitivně definitní**, když $f(\mathbf{x}) > 0$ pro každé $\mathbf{x} \neq \mathbf{0}$
- **negativně definitní**, když $f(\mathbf{x}) < 0$ pro každé $\mathbf{x} \neq \mathbf{0}$
- **indefinitní**, když $f(\mathbf{x}) > 0$ a $f(\mathbf{y}) < 0$ pro nějaká \mathbf{x}, \mathbf{y}

Definitností čtvercové matice \mathbf{A} rozumíme definitnost kvadr. formy $f(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x}$.

Částečné uspořádání indukované definitností

Na množině $\mathbb{R}^{n \times n}$ zavedeme binární relaci \preceq takto:

$$\mathbf{A} \preceq \mathbf{B} \iff \mathbf{B} - \mathbf{A} \text{ je pozitivně semidefinitní}$$

Věta

Relace \preceq je částečné uspořádání (tedy reflexivní, transitivní a antisymetrická).

- Speciálně: $\mathbf{A} \preceq \mathbf{0}$ značí, že \mathbf{A} je pozitivně semidefinitní.
- \mathbf{A}, \mathbf{B} jsou neporovnatelné, právě když $\mathbf{A} - \mathbf{B}$ je indefinitní.

Extrémy kvadratické formy

Tvrzení

Nechť $f(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x}$ je kvadratická forma.

- Je-li f **pozitivně semidefinitní**, pak má v bodě $\mathbf{0}$ **minimum**.
- Je-li f **negativně semidefinitní**, pak má v bodě $\mathbf{0}$ **maximum**.
- Je-li f **pozitivně definitní**, pak má v bodě $\mathbf{0}$ **ostré minimum**.
- Je-li f **negativně definitní**, pak má v bodě $\mathbf{0}$ **ostré maximum**.
- Je-li f **indefinitní**, pak **nemá minimum ani maximum**.

Pozor: Je-li f např. pozitivně semidefinitní, pak může mít minimum i v jiných bodech než $\mathbf{0}$.

Diagonalizace kvadratické formy

Nechť $\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T$ je spektrální rozklad:

$$f(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x} = \underbrace{\mathbf{x}^T \mathbf{V}}_{\mathbf{y}^T} \mathbf{\Lambda} \underbrace{\mathbf{V}^T \mathbf{x}}_{\mathbf{y}} = \mathbf{y}^T \mathbf{\Lambda} \mathbf{y} = \lambda_1 y_1^2 + \dots + \lambda_n y_n^2 = g(\mathbf{y})$$

- Substituce $\mathbf{x} = \mathbf{V}\mathbf{y}$ převedla kvadr. formu f na diagonální kvadr. formu g .
- Transformace $\mathbf{x} = \mathbf{V}\mathbf{y}$ je isometrie a bijekce.

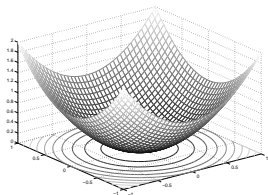
Z diagonální formy ihned vidíme:

Věta

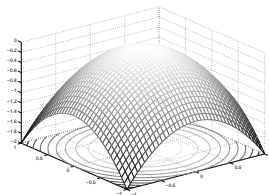
Symetrická matice je

- **pozitivně semidefinitní**, právě když má všechna vlastní čísla **nezáporná**
- **negativně semidefinitní**, právě když má všechna vlastní čísla **nekladná**
- **pozitivně definitní**, právě když má všechna vlastní čísla **kladná**
- **negativně definitní**, právě když má všechna vlastní čísla **záporná**
- **indefinitní**, právě když má **aspoň jedno kladné a aspoň jedno záporné** vl. číslo

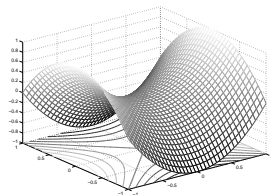
Příklad pro $n = 2$ proměnné



$$g(\mathbf{y}) = y_1^2 + y_2^2$$



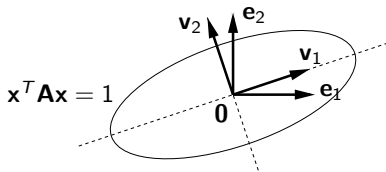
$$g(\mathbf{y}) = -y_1^2 - y_2^2$$



$$g(\mathbf{y}) = y_1^2 - y_2^2$$

Vrstevnice kvadratické formy f jsou kuželosečky se středem v počátku:

- Isometrie (rotace+zrcadlení) $\mathbf{V} = [\mathbf{v}_1 \ \mathbf{v}_2]$ určuje potočení elipsy: elipsa má hlavní osy ve směru vlastních vektorů \mathbf{v}_1 a \mathbf{v}_2 .
- Vlastní čísla λ_1, λ_2 určují typ kuželosečky.



Choleského rozklad

Věta

Pro každou symetrickou pozitivně semidefinitní \mathbf{A} existuje matice \mathbf{B} tak, že

$$\mathbf{A} = \mathbf{B}^T \mathbf{B}.$$

Důkaz: Položíme $\mathbf{B} = \mathbf{\Lambda}^{1/2} \mathbf{V}^T$, kde $\mathbf{A} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^T$ je spektrální rozklad \mathbf{A} .

Věta

Pro každou symetrickou pozitivně semidefinitní \mathbf{A} existuje **horní trojúhelníková** \mathbf{R} tak, že

$$\mathbf{A} = \mathbf{R}^T \mathbf{R}.$$

Důkaz: Uděláme QR rozklad $\mathbf{B} = \mathbf{Q} \mathbf{R}$. Pak $\mathbf{A} = \mathbf{B}^T \mathbf{B} = \mathbf{R}^T \mathbf{Q}^T \mathbf{Q} \mathbf{R} = \mathbf{R}^T \mathbf{R}$.

Tvrzení: Pro pozitivně definitní \mathbf{A} je \mathbf{R} určena jednoznačně.

Choleského rozklad: efektivní algoritmus na rozklad $\mathbf{A} = \mathbf{R}^T \mathbf{R}$.

- V Matlabu: `R=chol(A)`
- Příklad použití: řešení lineární soustavy $\mathbf{A} \mathbf{x} = \mathbf{b}$ kde \mathbf{A} je symetrická.

Různé charakterizace pozitivní (semi)definitnosti

Věta

Pro symetrickou matici \mathbf{A} jsou tato tvrzení ekvivalentní:

1. \mathbf{A} je pozitivně **definitní**.
2. Všechna vlastní čísla matice \mathbf{A} jsou **kladná**.
3. Existuje **čtvercová regulární** matice \mathbf{B} taková, že $\mathbf{A} = \mathbf{B}^T \mathbf{B}$.
4. Všechny **vůdčí** hlavní minory matice \mathbf{A} jsou **kladné**.

Věta

Pro symetrickou matici \mathbf{A} jsou tato tvrzení ekvivalentní:

1. \mathbf{A} je pozitivně **semidefinitní**.
2. Všechna vlastní čísla matice \mathbf{A} jsou **nezáporná**.
3. Existuje **čtvercová nebo úzká** matice \mathbf{B} taková, že $\mathbf{A} = \mathbf{B}^T \mathbf{B}$.
4. Všechny hlavní minory matice \mathbf{A} jsou **nezáporné**.

Kvadratické funkce

Kvadratická funkce

Kvadratická funkce je polynom $f: \mathbb{R}^n \rightarrow \mathbb{R}$ druhého stupně, tj.

$$f(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x} + \mathbf{b}^T \mathbf{x} + c$$

pro nějaké $\mathbf{A} \in \mathbb{R}^{n \times n}$, $\mathbf{b} \in \mathbb{R}^n$ a $c \in \mathbb{R}$.

Příklad pro $n = 2$:

$$f(x, y) = 2x^2 - 2xy + y^2 - 2y + 3 = \begin{bmatrix} x \\ y \end{bmatrix}^T \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} 0 \\ -2 \end{bmatrix}^T \begin{bmatrix} x \\ y \end{bmatrix} + 3$$

Doplnění kvadratické funkce na čtverec

Pro některé kvadratické lze nalézt $\mathbf{x}_0 \in \mathbb{R}^n$ a $y_0 \in \mathbb{R}$ tak, že

$$\begin{aligned}f(\mathbf{x}) &= \mathbf{x}^T \mathbf{A} \mathbf{x} + \mathbf{b}^T \mathbf{x} + c = (\mathbf{x} - \mathbf{x}_0)^T \mathbf{A} (\mathbf{x} - \mathbf{x}_0) + y_0 \\&= \mathbf{x}^T \mathbf{A} \mathbf{x} - \mathbf{x}^T \mathbf{A} \mathbf{x}_0 - \mathbf{x}_0^T \mathbf{A} \mathbf{x} + \mathbf{x}_0^T \mathbf{A} \mathbf{x}_0 + y_0 \\&= \mathbf{x}^T \mathbf{A} \mathbf{x} - \mathbf{x}_0^T (\mathbf{A} + \mathbf{A}^T) \mathbf{x} + \mathbf{x}_0^T \mathbf{A} \mathbf{x}_0 + y_0\end{aligned}$$

Porovnáme koeficienty u stejných monomů:

$$\mathbf{b} = -(\mathbf{A} + \mathbf{A}^T) \mathbf{x}_0 \quad (*)$$

$$c = \mathbf{x}_0^T \mathbf{A} \mathbf{x}_0 + y_0$$

- Rovnice (*) je vlastně stacionární podmínka $\nabla f(\mathbf{x}) = \mathbf{0}$.
- Doplnění na čtverec je možné, právě když rovnice (*) má řešení (tj. f má aspoň jeden stacionární bod).
- Druhá forma f je posunutá kvadratická forma.
Z ní ihned zjistíme, zda f má extrém a jaký.

Příklad: Tato kvadr. funkce doplnit na čtverec jde:

$$\begin{aligned} f(x, y) &= 2x^2 - 2xy + y^2 - 2y + 3 = \begin{bmatrix} x \\ y \end{bmatrix}^T \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} 0 \\ -2 \end{bmatrix}^T \begin{bmatrix} x \\ y \end{bmatrix} + 3 \\ &= \begin{bmatrix} x-1 \\ y-2 \end{bmatrix}^T \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} x-1 \\ y-2 \end{bmatrix} + 1 \\ &= 2(x-1)^2 - 2(x-1)(y-2) + (y-2)^2 + 1 \end{aligned}$$

tedy $\mathbf{x}_0 = (1, 2)$, $y_0 = 1$.

Příklad: A tato nejde:

$$f(x, y) = x^2 + y = \begin{bmatrix} x \\ y \end{bmatrix}^T \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix}^T \begin{bmatrix} x \\ y \end{bmatrix}$$

Soustava $\mathbf{b} = -(\mathbf{A} + \mathbf{A}^T)\mathbf{x}_0$, tj.

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix} = -2 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$$

nemá řešení.

Kvadriky

Množina

$$\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{x}^T \mathbf{A} \mathbf{x} + \mathbf{b}^T \mathbf{x} + c = 0 \}$$

kořenů (= vrstevnice výšky 0) kvadratické funkce se nazývá **kvadrika**.

Speciální případy:

- **elipsoid**: \mathbf{A} je pozitivně definitní
- **kuželosečka**: $n = 2$