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

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

- 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

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

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

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

lze napsat jako

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

kde

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

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

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

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 $\mathbf{A}$.

Komplexně sdružené číslo k číslu $x = a + bi \in \mathbb{C}$ je číslo $\overline{x} = a - bi$.
Absolutní hodnota čísla $x \in \mathbb{C}$ je reálné číslo $|x| = (\overline{x}x)^{1/2} = (a^2 + b^2)^{1/2}$.
Adjungovaná matice ke komplexní matici $\mathbf{A}$ je matice $\mathbf{A}^* = \overline{\mathbf{A}}^T$.
Norma vektoru $\mathbf{x} \in \mathbb{C}^n$ je $\|\mathbf{x}\| = (\mathbf{x}^*\mathbf{x})^{1/2} = (|x_1|^2 + \cdots + |x_n|^2)^{1/2}$.

Důkaz prvního tvrzení: Nechť $\mathbf{A}^* = \mathbf{A}$ a $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$. Pak

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

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

Důkaz druhého tvrzení pro speciální případ různých vl. čísel: Jestliže $\lambda_i \neq \lambda_j$, pak

$$\lambda_j\mathbf{v}_i^T\mathbf{v}_j = \mathbf{v}_i^T(\mathbf{A}\mathbf{v}_j) = (\mathbf{A}\mathbf{v}_i)^T\mathbf{v}_j = \lambda_i\mathbf{v}_i^T\mathbf{v}_j.$$

Z toho $(\lambda_i - \lambda_j)\mathbf{v}_i^T\mathbf{v}_j = 0$, z toho $\mathbf{v}_i^T\mathbf{v}_j = 0$.

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

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

Kvadratické formy

Kvadratická forma

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

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

Tvrzení

Pro každou kvadratickou formu $f : \mathbb{R}^n \to \mathbb{R}$ 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

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

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

- **positivně semidefinitní**, když $f(\mathbf{x}) \geq 0$ pro každé $\mathbf{x}$
- **negativně semidefinitní**, když $f(\mathbf{x}) \leq 0$ pro každé $\mathbf{x}$
- **positivně 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:

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

Věta

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

- Speciálně: $\mathbf{A} \succeq \mathbf{0}$ značí, že $\mathbf{A}$ je positivně 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 positivně semidefinitní, pak má v bodě $\mathbf{0}$ minimum.
- Je-li f negativně semidefinitní, pak má v bodě $\mathbf{0}$ maximum.
- Je-li f positivně 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 (protože je neomezená shora i zdola).

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

Definitnost matice z hlavních minorů

Pro $\mathbf{A} \in \mathbb{R}^{n\times n}$ a neprázdnou množinu $I \subseteq \{1,\dots,n\}$ označme $\mathbf{A}_I = [a_{ij}]_{i,j\in I}$.

Příklady matic $\mathbf{A}_I$ pro $n = 3$:

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}}_{I=\{1,2,3\}} \quad \underbrace{\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}}_{I=\{1,2\}} \quad \underbrace{\begin{bmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{bmatrix}}_{I=\{1,3\}} \quad \underbrace{\begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix}}_{I=\{2,3\}}$$

- **Hlavní minor** matice $\mathbf{A}$ je číslo $\det \mathbf{A}_I$ pro nějakou neprázdnou $I \subseteq \{1,\dots,n\}$.
- **Vůdčí hlavní minor** je hlavní minor pro $I = \{1,\dots,k\}$ a nějaké $k \le n$.

Věta

Symetrická matice je

- positivně definitní, právě když všechny její vůdčí hlavní minory jsou kladné,
- positivně semidefinitní, právě když všechny její hlavní minory jsou nezáporné.

Věta umožňuje rozhodnout i negativní semidefinitnost a indefinitnost, neboť:

- $\mathbf{A}$ je negativně [semi]definitní, právě když $-\mathbf{A}$ je positivně [semi]definitní.
- Matice je indefinitní, právě když není ani positivně ani negativně semidefinitní.

Diagonalizace kvadratické formy

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

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

- 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.

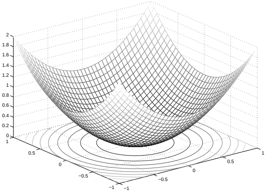
Z diagonální formy ihned vidíme:

Věta

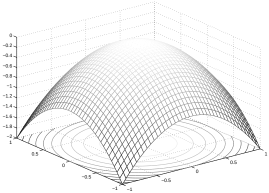
Symetrická matice je

- positivně semidefinitní, právě když má všechna vlastní čísla nezáporná
- negativně semidefinitní, právě když má všechna vlastní čísla nekladná
- positivně 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

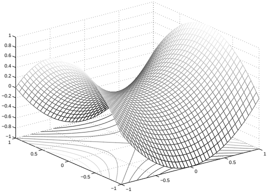
Příklad pro $n = 2$ proměnné



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



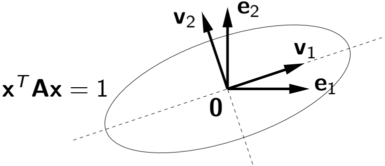
$g(\mathbf{y}) = -y_1^2 - y_2^2$



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

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 positivně 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 $\mathbf{A} = \mathbf{V\Lambda V}^T$ je spektrální rozklad $\mathbf{A}$.

Věta

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

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

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

Tvrzení: Pro positivně definitní $\mathbf{A} \in \mathbb{R}^{n\times n}$ 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 $\mathbf{Ax} = \mathbf{b}$ kde $\mathbf{A}$ je symetrická.

Různé charakterizace positivní (semi)definitnosti

Věta

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

1. $\mathbf{A}$ je positivně 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 positivně 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 $f\colon \mathbb{R}^n \to \mathbb{R}$ 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:

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

- 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 positivně definitní
- **kuželosečka**: $n = 2$