Optimalizace

7. Spektrální rozklad

Tomáš Kroupa Tomáš Werner

2023 LS

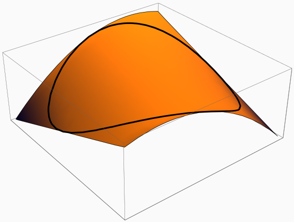
Fakulta elektrotechnická
ČVUT v Praze

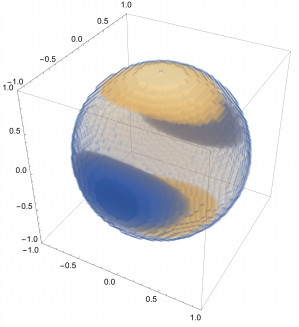
Proč nás zajímají vlastní čísla/vektory v optimalizaci?

- Pro symetrickou matici $\mathbf{A} \in \mathbb{R}^{n \times n}$ často řešíme úlohy

$$\min_{\mathbf{x} \in \mathbb{R}^n} \{\mathbf{x}^T \mathbf{A} \mathbf{x} \mid \|\mathbf{x}\| = 1\} \quad \text{a} \quad \max_{\mathbf{x} \in \mathbb{R}^n} \{\mathbf{x}^T \mathbf{A} \mathbf{x} \mid \|\mathbf{x}\| = 1\}$$

- Ukážeme, že minimum/maximum je vlastní vektor odpovídající nejmenšímu/největšímu vlastnímu číslu matice $\mathbf{A}$





Vlastní čísla a vlastní 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.

λ je vlastní číslo

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

$\mathbf{v}$ je vlastní vektor

$$\mathbf{v} \in \operatorname{null}(\mathbf{A} - \lambda \mathbf{I})$$

Spektrum matice je množina všech jejích vlastních čísel.

Diagonalizovatelné matice

Definujme

$$\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n) \qquad \text{a} \qquad \mathbf{V} = [\mathbf{v}_1 \ldots \mathbf{v}_n].$$

- Platí $\mathbf{AV} = \mathbf{V}\boldsymbol{\Lambda}$.
- Matice $\mathbf{A}$ je diagonalizovatelná pokud je $\mathbf{V}$ regulární.

Spektrální rozklad

Pro diagonalizovatelnou matici $\mathbf{A}$ platí

$$\mathbf{A} = \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^{-1} \qquad \text{a} \qquad \boldsymbol{\Lambda} = \mathbf{V}^{-1}\mathbf{AV}.$$

Spektrální rozklad symetrické matice

Věta

Nechť $\mathbf{A} \in \mathbb{R}^{n \times n}$ je symetrická. Potom je každé vlastní číslo matice $\mathbf{A}$ reálné a existuje ortonormální množina vlastních vektorů $\mathbf{v}_1, \ldots, \mathbf{v}_n$ matice $\mathbf{A}$.

Tedy pro reálnou symetrickou matici $\mathbf{A}$ je matice $\mathbf{V} = [\mathbf{v}_1 \ldots \mathbf{v}_n]$ dokonce ortogonální a platí

$$\mathbf{A} = \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^\mathsf{T} = \lambda_1 \mathbf{v}_1 \mathbf{v}_1^\mathsf{T} + \cdots + \lambda_n \mathbf{v}_n \mathbf{v}_n^\mathsf{T}.$$

Spektrální rozklad symetrické matice – geometricky

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

Interpretace

1. Matice $\mathbf{A}$ vyjadřuje škálování maticí $\boldsymbol{\Lambda}$ v souřadném systému popsaném bází $\mathbf{v}_1, \ldots, \mathbf{v}_n$
2. Matice $\mathbf{A}$ je součtem škálovaných ortogonálních projektorů $\lambda_i \mathbf{v}_i \mathbf{v}_i^\mathsf{T}$ na vzájemně kolmé přímky o směrech $\mathbf{v}_i$

Jak počítat vlastní čísla?

QR algoritmus (Francis, 1961) je základem soudobých efektivních metod na výpočet vlastních čísel matice $\mathbf{A}$.

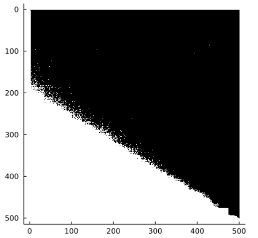
1. $\mathbf{A}_0 := \mathbf{A}$, $i := 0$
2. Dokud není splněna ukončovací podmínka:
   2.1 Sestroj QR rozklad, $\mathbf{A}_i = \mathbf{QR}$
   2.2 $\mathbf{A}_{i+1} := \mathbf{RQ}$
   2.3 $i := i + 1$

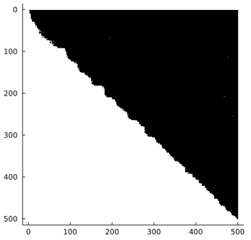
Tvrzení

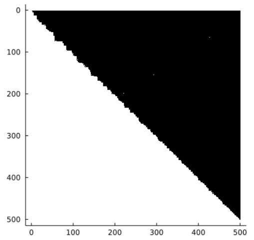
Matice $\mathbf{A}_0, \mathbf{A}_1, \ldots$ mají stejné spektrum.

QR algoritmus – ukázka

- Posloupnost $\mathbf{A}_0, \mathbf{A}_1, \ldots$ konverguje téměř ve všech případech k blokově horní trojúhelníkové matici s bloky o velikosti 1 a 2
- Pro náhodnou matici $\mathbf{A}$ řádu 500 dostaneme $\mathbf{A}_{50}, \mathbf{A}_{500}, \mathbf{A}_{2000}$:







Pozitivně semidefinitní matice

Matice $\mathbf{A} \in \mathbb{R}^{n \times n}$ je pozitivně semidefinitní, pokud

$$\mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x} \geq 0 \quad \text{pro každý } \mathbf{x} \in \mathbb{R}^n.$$

Tvrzení

Pro symetrickou matici $\mathbf{A} \in \mathbb{R}^{n \times n}$ jsou tato tvrzení ekvivalentní.

1. $\mathbf{A}$ pozitivně semidefinitní.
2. Vlastní čísla matice $\mathbf{A}$ jsou nezáporná.
3. Existuje matice $\mathbf{B} \in \mathbb{R}^{n \times n}$ taková, že $\mathbf{A} = \mathbf{B}^\mathsf{T}\mathbf{B}$.
4. Všechny hlavní minory matice $\mathbf{A}$ jsou nezáporné.

Pozitivně definitní matice

Matice $\mathbf{A} \in \mathbb{R}^{n \times n}$ je pozitivně definitní, pokud

$$\mathbf{x}^\mathsf{T} \mathbf{A} \mathbf{x} > 0 \quad \text{pro každý } \mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}.$$

Tvrzení

Pro symetrickou matici $\mathbf{A} \in \mathbb{R}^{n \times n}$ jsou tato tvrzení ekvivalentní.

1. $\mathbf{A}$ pozitivně definitní.
2. Vlastní čísla matice $\mathbf{A}$ jsou kladná.
3. Existuje regulární matice $\mathbf{B} \in \mathbb{R}^{n \times n}$ taková, že $\mathbf{A} = \mathbf{B}^\mathsf{T} \mathbf{B}$.
4. Všechny vůdčí hlavní minory matice $\mathbf{A}$ jsou kladné.

Negativně semidefinitní a indefinitní matice

Matice $\mathbf{A} \in \mathbb{R}^{n \times n}$ je

- negativně semidefinitní, pokud $\mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x} \leq 0$ pro každé $\mathbf{x} \in \mathbb{R}^n$,
- negativně definitní, pokud $\mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x} < 0$ pro každé $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$,
- indefinitní, existuje-li $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ tak, že $\mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x} < 0 < \mathbf{y}^\mathsf{T}\mathbf{A}\mathbf{y}$.

Pozorování

$\mathbf{A}$ je negativně definitní, právě když $-\mathbf{A}$ je pozitivně definitní.

Choleského rozklad

Věta

Nechť $\mathbf{A} \in \mathbb{R}^{n \times n}$ je symetrická. Je-li $\mathbf{A}$ pozitivně semidefinitní, potom existuje horní trojúhelníková matice $\mathbf{R} \in \mathbb{R}^{n \times n}$ tak, že

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

Je-li $\mathbf{A}$ pozitivně definitní, je taková matice $\mathbf{R}$ jediná.

Aplikace pro symetrickou pozitivně definitní matici $\mathbf{A}$:

- Řešení soustavy $\mathbf{A}\mathbf{x} = \mathbf{b}$
- Invertování matice $\mathbf{A}$

Kvadratické formy

Kvadratická forma je homogenní polynom $f\colon \mathbb{R}^n \to \mathbb{R}$ stupně 2,

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

Poznámka

Pro každou kvadratickou formu s maticí $\mathbf{A} \in \mathbb{R}^{n\times n}$ je matice $\frac{\mathbf{A}+\mathbf{A}^\mathsf{T}}{2}$ symetrická a platí $\mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x} = \mathbf{x}^\mathsf{T}\frac{\mathbf{A}+\mathbf{A}^\mathsf{T}}{2}\mathbf{x}$.

Extrémy kvadratické formy

Tvrzení

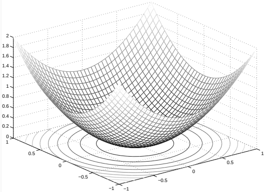
Uvažujme kvadratickou formu $f(\mathbf{x}) = \mathbf{x}^\top \mathbf{A}\mathbf{x}$, kde $\mathbf{A} \in \mathbb{R}^{n\times n}$.

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

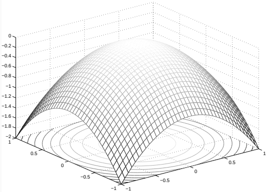
Analogicky pro negativně semidefinitní matice/maximum.

Příklad: kvadratické formy s diagonální maticí pro $n = 2$

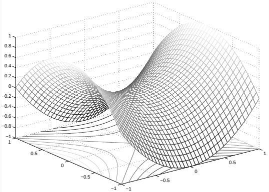
$$g(\mathbf{y}) = \mathbf{y}^\mathsf{T} \mathbf{\Lambda} \mathbf{y} = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2$$



$$g(\mathbf{y}) = y_1^2 + y_2^2 \qquad \mathbf{\Lambda} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$



$$g(\mathbf{y}) = -y_1^2 - y_2^2 \qquad \mathbf{\Lambda} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$



$$g(\mathbf{y}) = y_1^2 - y_2^2 \qquad \mathbf{\Lambda} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

Diagonalizace kvadratické formy

Tvrzení

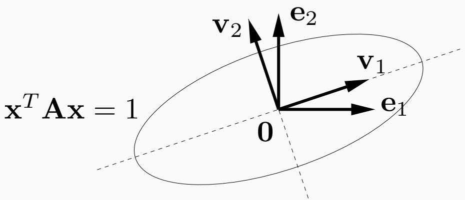
Pro každou kvadratickou formu f se symetrickou maticí $\mathbf{A}$ existuje kvadratická forma g s diagonální maticí $\mathbf{\Lambda}$ tak, že $f(\mathbf{x}) = g(\mathbf{V}^\top \mathbf{x})$, kde $\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^\top$ je spektrální rozklad.

- Forma f má jednodušší popis v novém souřadném systému tvořeném ortonomální bází ze sloupců $\mathbf{V}$
- Typ kvadratické formy poznáme podle znamének na diagonále

Příklad: vrstevnice kvadratické formy pro $n = 2$

Elipsa

- Vrstevnice výšky 1 kvadratické formy $f(\mathbf{x}) = \mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x}$ s pozitivně definitní maticí $\mathbf{A} \in \mathbb{R}^2$ je pootočená elipsa
- Elipsa má osy ve směru vlastních vektorů $\mathbf{v}_1$ a $\mathbf{v}_2$
- Délky poloos elipsy jsou $\frac{1}{\sqrt{\lambda_1}}$ a $\frac{1}{\sqrt{\lambda_2}}$



Kvadratická funkce a její vrstevnice

Kvadratická funkce je polynom $f\colon \mathbb{R}^n \to \mathbb{R}$ druhého stupně,

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

kde $\mathbf{A} \in \mathbb{R}^{n\times n}$ je symetrická, $\mathbf{b} \in \mathbb{R}^n$ a $c \in \mathbb{R}$.

Kvadrika je vrstevnice výšky 0 kvadratické funkce, tj. množina

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

Speciální případy

- Elipsoid (**A** je pozitivně definitní)
- Kuželosečka (pro $n = 2$)