Optimalizace

Spektrální rozklad a kvadratické funkce

Tomáš Kroupa Tomáš Werner

2024

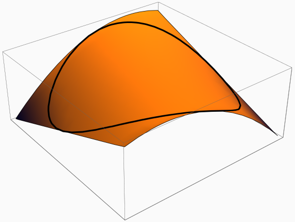
Fakulta elektrotechnická
ČVUT v Praze

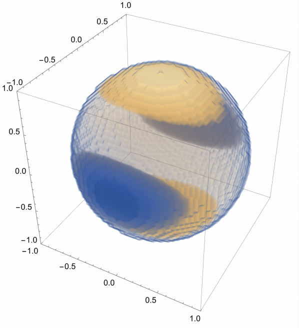
Vlastní čísla a vlastní vektory v optimalizaci

- Pro symetrickou matici $\mathbf{B} \in \mathbb{R}^{m\times m}$ řešíme nekonvexní úlohu

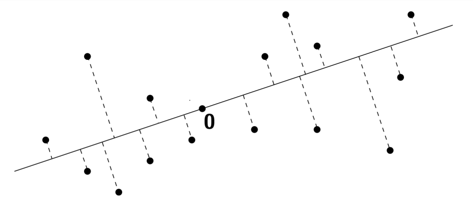
$$\max_{\mathbf{x}\in\mathbb{R}^m} \{\mathbf{x}^\mathsf{T}\mathbf{B}\mathbf{x} \mid \|\mathbf{x}\| = 1\}$$

- Ukážeme, že optimálním řešením je vlastní vektor odpovídající největšímu vlastnímu číslu matice $\mathbf{B}$





Příklad: proložení bodů lineárním podprostorem dimenze 1



- Předpoklad: data $\mathbf{a}_1, \ldots, \mathbf{a}_n \in \mathbb{R}^m$ leží přibližně na přímce směru $\mathbf{x} \in \mathbb{R}^m$ procházející počátkem, kde $\|\mathbf{x}\| = 1$
- Píšeme $\mathbf{A} = [\mathbf{a}_1 \ldots \mathbf{a}_n]$
- Lze ukázat, že neznámý směr $\mathbf{x}$ je optimálním řešením úlohy

$$\text{maximalizuj} \quad \mathbf{x}^\mathsf{T} \underbrace{\mathbf{A}\mathbf{A}^\mathsf{T}}_{\mathbf{B}} \mathbf{x} \quad \text{za podmínky } \|\mathbf{x}\| = 1,\ \mathbf{x} \in \mathbb{R}^m$$

Spektrální rozklad

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

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

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

Spektrální rozklad

Pro diagonalizovatelnou matici $\mathbf{A}$ platí

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

Spektrální rozklad symetrické matice

Věta

Symetrická matice $\mathbf{A} \in \mathbb{R}^{n \times n}$ má všechna vlastní čísla reálná a existuje ortonormální báze tvořena jejími vlastními vektory.

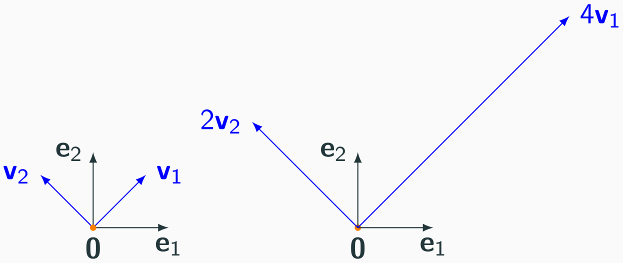
- Konvence: $\lambda_1 \geq \cdots \geq \lambda_n$
- Matice $\mathbf{V} = [\mathbf{v}_1 \ldots \mathbf{v}_n]$ je ortogonální
- Platí

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

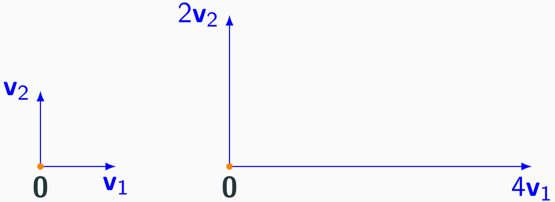
kde matice $\mathbf{v}_i \mathbf{v}_i^\mathsf{T}$ jsou ortogonální projektory na vzájemně kolmé přímky o směrech $\mathbf{v}_i$

Spektrální rozklad $\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^\top$ názorně

$$\mathbf{A} = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$$



$$\mathbf{\Lambda} = \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix}$$



Jak počítat vlastní čísla?

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

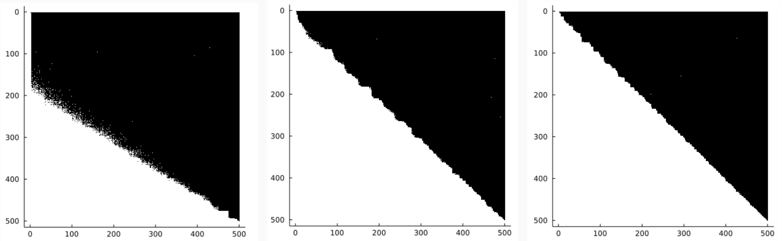
1. $\mathbf{A}_0 := \mathbf{A}$, $i := 0$
2. Dokud není splněna ukončovací podmínka:
   2.1 Sestroj QR rozklad, $\mathbf{A}_i = \mathbf{QR}$
   2.2 $\mathbf{A}_{i+1} := \mathbf{RQ}$
   2.3 $i := i + 1$

Tvrzení

Matice $\mathbf{A}_0, \mathbf{A}_1, \dots$ 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 $\mathbf{A}_{50}, \mathbf{A}_{500}, \mathbf{A}_{2000}$:



Kvadratické funkce a formy

Kvadratické funkce více proměnných

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

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

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

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

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

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 $\mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x} = \mathbf{x}^\mathsf{T}\frac{\mathbf{A}+\mathbf{A}^\mathsf{T}}{2}\mathbf{x}$.

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 dolní trojúhelníková matice $\mathbf{L} \in \mathbb{R}^{n \times n}$ tak, že

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

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

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

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

Extrémy kvadratické formy

Tvrzení

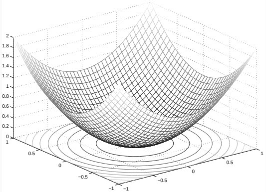
Uvažujme kvadratickou formu $f(\mathbf{x}) = \mathbf{x}^\top \mathbf{A}\mathbf{x}$, kde $\mathbf{A} \in \mathbb{R}^{n\times n}$.

- 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.

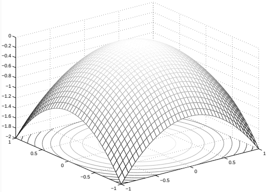
Analogicky pro negativně semidefinitní matice/maximum.

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

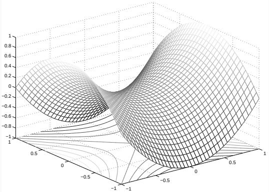
$$g(\mathbf{y}) = \mathbf{y}^\mathsf{T}\mathbf{\Lambda}\mathbf{y} = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2$$



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



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



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

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

$$f(\mathbf{x}) = g(\mathbf{V}^\mathsf{T}\mathbf{x}),$$

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

- Kvadratickou formu f diagonalizujeme substitucí $\mathbf{y} = \mathbf{V}^\mathsf{T}\mathbf{x}$
- Typ formy g poznáme podle znamének na diagonále $\mathbf{\Lambda}$

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 $\frac{1}{\sqrt{\lambda_2}}$