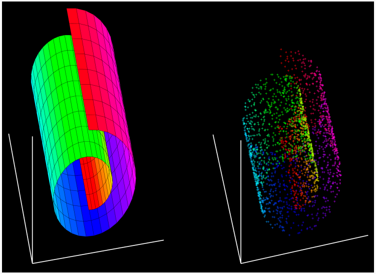
Optimalizace

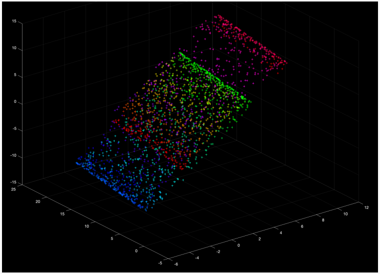
Aplikace spektrálního rozkladu

Tomáš Kroupa

FEL ČVUT

Motivace – redukce dimenze



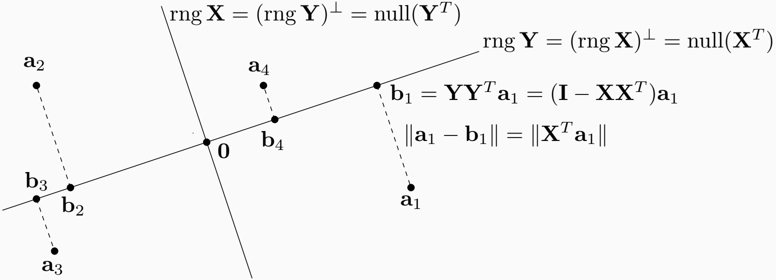


- Nalevo jsou zobrazeny původní datové vektory
- Napravo jsou jejich redukované obrazy v prostoru dimenze 2

Proložení bodů podprostorem

Úloha

K bodům $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{R}^n$ hledáme lineární podprostor $\operatorname{rng} \mathbf{Y}$ dimenze $k \leq n$ minimalizující součet čtverců vzdáleností.



Hledáme $\mathbf{X} \in \mathbb{R}^{n\times(n-k)}$ s ortonormálními sloupci minimalizující

$$\|\mathbf{X}^\mathsf{T}\mathbf{a}_1\|^2 + \cdots + \|\mathbf{X}^\mathsf{T}\mathbf{a}_m\|^2.$$

Stopa matice

Stopa čtvercové matice $\mathbf{A} \in \mathbb{R}^{n \times n}$ je číslo

$$\operatorname{tr} \mathbf{A} = a_{11} + \cdots + a_{nn}.$$

Vlastnosti

1. $\operatorname{tr}(\mathbf{A} + \mathbf{B}) = \operatorname{tr} \mathbf{A} + \operatorname{tr} \mathbf{B}$
2. $\operatorname{tr}(\alpha \mathbf{A}) = \alpha \operatorname{tr} \mathbf{A}$
3. $\operatorname{tr}(\mathbf{A}^\mathsf{T}) = \operatorname{tr} \mathbf{A}$
4. $\operatorname{tr}(\mathbf{AB}) = \operatorname{tr}(\mathbf{BA})$ pro každé $\mathbf{A} \in \mathbb{R}^{m \times n}$ a $\mathbf{B} \in \mathbb{R}^{n \times m}$

Stopy některých matic

Ortogonální projektor

Nechť $\mathbf{U} \in \mathbb{R}^{n \times k}$ je matice s ortonormálními sloupci. Ortogonální projektor $\mathbf{P} := \mathbf{U}\mathbf{U}^{\mathsf{T}}$ na podprostor $\operatorname{rng} \mathbf{U}$ dimenze k má stopu

$$\operatorname{tr} \mathbf{P} = k.$$

Stopy některých matic

Gramova matice

Datové vektory $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{R}^n$ uspořádejme do řádků matice

$$\mathbf{A} := \begin{bmatrix} \mathbf{a}_1 \cdots \mathbf{a}_m \end{bmatrix}^\mathsf{T} \in \mathbb{R}^{m \times n}.$$

Stopa matice $\mathbf{A}\mathbf{A}^\mathsf{T}$ je

$$\operatorname{tr}(\mathbf{A}\mathbf{A}^\mathsf{T}) = \|\mathbf{a}_1\|^2 + \cdots + \|\mathbf{a}_m\|^2.$$

Norma matice

Frobeniova norma matice $\mathbf{A} \in \mathbb{R}^{m \times n}$ je

$$\|\mathbf{A}\| := \sqrt{\operatorname{tr}(\mathbf{A}\mathbf{A}^\mathsf{T})} = \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2}.$$

Vzdálenost matic $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m \times n}$ je

$$\|\mathbf{A} - \mathbf{B}\|.$$

Úloha na nejmenší stopu

Věta

Nechť $\mathbf{A} \in \mathbb{R}^{n \times n}$ je symetrická s vlastními čísly $\lambda_1 \leq \cdots \leq \lambda_n$. Pro libovolné $k \leq n$ platí

$$\min \{\operatorname{tr}(\mathbf{X}^\mathsf{T}\mathbf{A}\mathbf{X}) \mid \mathbf{X} \in \mathbb{R}^{n \times k}, \mathbf{X}^\mathsf{T}\mathbf{X} = \mathbf{I}\} = \lambda_1 + \cdots + \lambda_k$$

a minima se nabývá pro $\mathbf{X} = \begin{bmatrix} \mathbf{v}_1 \cdots \mathbf{v}_k \end{bmatrix}$.

Proložení bodů podprostorem jako úloha na nejmenší stopu

V původní motivační úloze pro zadané $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{R}^n$ hledáme $\mathbf{X} \in \mathbb{R}^{n\times(n-k)}$ s ortonormálními sloupci minimalizující

$$\|\mathbf{X}^\mathsf{T}\mathbf{a}_1\|^2 + \cdots + \|\mathbf{X}^\mathsf{T}\mathbf{a}_m\|^2.$$

Úloha

Pro matici $\mathbf{A} := \begin{bmatrix}\mathbf{a}_1 \cdots \mathbf{a}_m\end{bmatrix}^\mathsf{T} \in \mathbb{R}^{m\times n}$ hledej

$$\min\left\{\|\mathbf{A}\mathbf{X}\|^2 \mid \mathbf{X} \in \mathbb{R}^{n\times(n-k)},\ \mathbf{X}^\mathsf{T}\mathbf{X} = \mathbf{I}\right\}.$$

Proložení bodů podprostorem – řešení

Úloha

$$\min \{\operatorname{tr}(\mathbf{X}^\top \mathbf{A}^\top \mathbf{A}\mathbf{X}) \mid \mathbf{X} \in \mathbb{R}^{n\times(n-k)}, \mathbf{X}^\top \mathbf{X} = \mathbf{I}\}$$

1. Pro matici $\mathbf{A}^\top \mathbf{A} \in \mathbb{R}^{n\times n}$ spočítáme spektrální rozklad $\mathbf{V}\mathbf{\Lambda}\mathbf{V}^\top$ a předpokládáme $\lambda_1 \leq \cdots \leq \lambda_n$
2. Označíme $\mathbf{V} = [\underbrace{\mathbf{v}_1 \cdots \mathbf{v}_{n-k}}_{\mathbf{X}^*} \underbrace{\mathbf{v}_{n-k+1} \cdots \mathbf{v}_n}_{\mathbf{Y}^*}]$
3. Ortonormální bázi hledaného podprostoru dimenze k nalezneme ve sloupcích matice $\mathbf{Y}^* \in \mathbb{R}^{n\times k}$

Nejbližší matice nižší hodnosti

Proložení bodů podprostorem – ekvivalentní formulace

Pro zadané $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{R}^n$ hledáme $\mathbf{b}_1, \ldots, \mathbf{b}_m \in \mathbb{R}^n$ ležící v podprostoru dimenze $k \leq n$ a minimalizující

$$\|\mathbf{a}_1 - \mathbf{b}_1\|^2 + \cdots + \|\mathbf{a}_m - \mathbf{b}_m\|^2.$$

$$\mathbf{A} := \begin{bmatrix}\mathbf{a}_1 \cdots \mathbf{a}_m\end{bmatrix}^\mathsf{T} \qquad \mathbf{B} := \begin{bmatrix}\mathbf{b}_1 \cdots \mathbf{b}_m\end{bmatrix}^\mathsf{T}$$

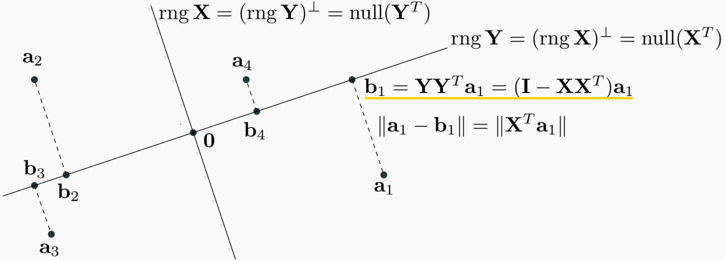
Low rank approximation

$$\min\left\{\|\mathbf{A} - \mathbf{B}\|^2 \mid \mathbf{B} \in \mathbb{R}^{m \times n}, \operatorname{rank} \mathbf{B} \leq k\right\}$$

Nejbližší matice nižší hodnosti – řešení

Nejbližší matice nižší hodnosti

$$\min \{\|\mathbf{A} - \mathbf{B}\|^2 \mid \mathbf{B} \in \mathbb{R}^{m\times n}, \operatorname{rank} \mathbf{B} \le k\}$$



- Optimální řešení $\mathbf{B}^* = \mathbf{A}\mathbf{Y}\mathbf{Y}^\mathsf{T} = \mathbf{A}(\mathbf{I} - \mathbf{X}\mathbf{X}^\mathsf{T})$
- Optimální hodnota $\|\mathbf{A} - \mathbf{B}^*\|^2 = \lambda_1 + \cdots + \lambda_{n-k}$

Prokládáme afinním podprostorem

Tvrzení

Nechť $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{R}^n$ a $k \leq n$. Afinní podprostor dimenze k, který minimalizuje součet čtverců vzáleností k bodům $\mathbf{a}_1, \ldots, \mathbf{a}_m$, obsahuje jejich těžiště

$$\overline{\mathbf{a}} = \frac{1}{m}(\mathbf{a}_1 + \cdots + \mathbf{a}_m).$$

1. Zadané body posuneme tak, aby měly těžiště v počátku:

$$\mathbf{a}_1 - \overline{\mathbf{a}}, \ldots, \mathbf{a}_m - \overline{\mathbf{a}}$$

2. Posunuté body proložíme lin. podprostorem Y dimenze k
3. Hledaný afinní podprostor je $Y + \overline{\mathbf{a}}$

Singulární rozklad matice

Věta (SVD)

Nechť $\mathbf{A} \in \mathbb{R}^{m \times n}$ a $p := \min\{m, n\}$. Potom existuje

- diagonální matice $\mathbf{S} \in \mathbb{R}^{m \times n}$ s diagon. prvky $s_1, \ldots, s_p \geq 0$,
- ortogonální matice $\mathbf{U} = \begin{bmatrix} \mathbf{u}_1 \cdots \mathbf{u}_m \end{bmatrix} \in \mathbb{R}^{m \times m}$ a $\mathbf{V} = \begin{bmatrix} \mathbf{v}_1 \cdots \mathbf{v}_n \end{bmatrix} \in \mathbb{R}^{n \times n}$ splňující

$$\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^\top = s_1 \mathbf{u}_1 \mathbf{v}_1^\top + \cdots + s_p \mathbf{u}_p \mathbf{v}_p^\top .$$

- Singulární čísla $s_1, \ldots, s_p$
- Levé a pravé singulární vektory $\mathbf{u}_1, \ldots, \mathbf{u}_m$, $\mathbf{v}_1, \ldots, \mathbf{v}_n$

Redukovaný SVD

$$\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^{\mathsf{T}} = s_1 \mathbf{u}_1 \mathbf{v}_1^{\mathsf{T}} + \cdots + s_p \mathbf{u}_p \mathbf{v}_p^{\mathsf{T}}, \qquad p = \min\{m, n\}$$

Plný SVD

$\mathbf{U} \in \mathbb{R}^{m \times m}$, $\mathbf{S} \in \mathbb{R}^{m \times n}$, $\mathbf{V} \in \mathbb{R}^{n \times n}$

Redukovaný SVD

$\mathbf{U} \in \mathbb{R}^{m \times p}$, $\mathbf{S} \in \mathbb{R}^{p \times p}$, $\mathbf{V} \in \mathbb{R}^{n \times p}$

Nejbližší matice nižší hodnosti pomocí SVD

Nechť $\mathbf{A} \in \mathbb{R}^{m \times n}$, $p := \min\{m, n\}$ a $\mathbf{USV}^\mathsf{T}$ je SVD matice $\mathbf{A}$, kde předpokládáme $s_1 \leq \cdots \leq s_p$.

Věta (Eckart-Young)

Řešením úlohy

$$\min\left\{\|\mathbf{A} - \mathbf{B}\|^2 \mid \mathbf{B} \in \mathbb{R}^{m \times n}, \operatorname{rank} \mathbf{B} \leq k\right\}$$

je matice $\mathbf{B}^* := \mathbf{US}'\mathbf{V}^\mathsf{T}$, kde matice $\mathbf{S}'$ se získá z matice $\mathbf{S}$ vynulováním prvků $s_1, \ldots, s_{p-k}$.

Reference

T. Werner. *Optimalizace* (kapitola 7). Elektronická skripta. FEL ČVUT, 2020.

http://www.gatsby.ucl.ac.uk/˜maneesh/dimred/