Optimalizace

5. PCA a SVD

Tom Werner, Tom Kroupa

FEL ČVUT

Úloha na nejmenší stopu

Stopa, skalární součin matic

Stopa čtvercové matice $\mathbf{A} = [a_{ij}] \in \mathbb{R}^{n\times n}$ je

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

Vlastnosti:

- $\operatorname{tr}(\mathbf{A}+\mathbf{B}) = \operatorname{tr}\mathbf{A} + \operatorname{tr}\mathbf{B}$
- $\operatorname{tr}(\alpha\mathbf{A}) = \alpha\operatorname{tr}\mathbf{A}$
- $\operatorname{tr}(\mathbf{A}^T) = \operatorname{tr}\mathbf{A}$
- $\operatorname{tr}(\mathbf{AB}) = \operatorname{tr}(\mathbf{BA})$ (kde $\mathbf{A},\mathbf{B}$ nemusí být čtvercové)
  Speciálně: $\operatorname{tr}(\mathbf{CAC}^{-1}) = \operatorname{tr}\mathbf{A}$
- $\operatorname{tr}\mathbf{A} = \lambda_1 + \cdots + \lambda_n$

Skalární součin a norma matic:

$$\langle\mathbf{A},\mathbf{B}\rangle = \sum_{i,j} a_{ij}b_{ij} = \operatorname{tr}(\mathbf{A}^T\mathbf{B})$$

$$\|\mathbf{A}\| = \sqrt{\sum_{i,j} a_{ij}^2} = \sqrt{\langle\mathbf{A},\mathbf{A}\rangle} = \sqrt{\operatorname{tr}(\mathbf{A}^T\mathbf{A})}$$

Úmluva o řazení vlastních čísel symetrické matice

Zopakujme spektrální rozklad symetrické matice $\mathbf{A} \in \mathbb{R}^{n \times n}$:

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

kde

$$\mathbf{V} = \begin{bmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_n \end{bmatrix}, \quad \mathbf{V}^T \mathbf{V} = \mathbf{I}, \qquad \boldsymbol{\Lambda} = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

Úmluva: Dále budeme předpokládat $\lambda_1 \leq \cdots \leq \lambda_n$.

Věta

Úloha

$$\min\{\, \mathbf{x}^T \mathbf{A}\mathbf{x} \mid \mathbf{x}^T\mathbf{x} = 1,\ \mathbf{x} \in \mathbb{R}^n \,\}$$

má minimální hodnotu λ_1 a ta se nabývá pro $\mathbf{x} = \mathbf{v}_1$.

Důkaz: Spektrální rozklad a substituce $\mathbf{x} = \mathbf{V}\mathbf{y}$:

$$\mathbf{x}^T\mathbf{A}\mathbf{x} = \mathbf{x}^T\mathbf{V}\mathbf{\Lambda}\mathbf{V}^T\mathbf{x} = \mathbf{y}^T\mathbf{\Lambda}\mathbf{y} = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2$$

Protože $\mathbf{V}$ je ortogonální, je

$$\mathbf{x}^T\mathbf{x} = \mathbf{y}^T\mathbf{V}^T\mathbf{V}\mathbf{y} = 1 \quad\Longleftrightarrow\quad \mathbf{y}^T\mathbf{y} = 1.$$

Tedy úloha je ekvivalentní úloze

$$\min\{\, \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2 \mid \mathbf{y} \in \mathbb{R}^n,\ y_1^2 + \cdots + y_n^2 = 1 \,\}$$

Optimum nastává pro $\mathbf{y} = (1, 0, \ldots, 0) = \mathbf{e}_1$, což odpovídá $\mathbf{x} = \mathbf{V}\mathbf{e}_1 = \mathbf{v}_1$ a $\mathbf{y}^T\mathbf{\Lambda}\mathbf{y} = \lambda_1$.

Úloha je ekvivalentní minimalizaci **Rayleighova kvocientu** $f(\mathbf{x}) = \dfrac{\mathbf{x}^T\mathbf{A}\mathbf{x}}{\mathbf{x}^T\mathbf{x}}$ na $\mathbb{R}^n \setminus \{\mathbf{0}\}$.

Úloha na nejmenší stopu

Věta

Nechť $k \leq n$. Úloha

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

má optimální hodnotu $\lambda_1 + \cdots + \lambda_k$ a ta se nabývá pro $\mathbf{X} = [\, \mathbf{v}_1 \ \cdots \ \mathbf{v}_k \,]$.

Interpretace: Jsou-li $\mathbf{x}_1, \ldots, \mathbf{x}_k \in \mathbb{R}^n$ sloupce matice $\mathbf{X}$, minimalizujeme

$$\operatorname{tr}(\mathbf{X}^T \mathbf{A} \mathbf{X}) = \mathbf{x}_1^T \mathbf{A} \mathbf{x}_1 + \cdots + \mathbf{x}_k^T \mathbf{A} \mathbf{x}_k$$

za podmínky, že $\mathbf{x}_1, \ldots, \mathbf{x}_k$ jsou ortonormální.

Myšlenka **důkazu**: dosazením $\mathbf{A} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^T$ a $\mathbf{X} = \mathbf{V} \mathbf{Y}$ opět převedeme na jednodušší úlohu

$$\min\{\, \operatorname{tr}(\mathbf{Y}^T \mathbf{\Lambda} \mathbf{Y}) \mid \mathbf{Y} \in \mathbb{R}^{n \times k},\ \mathbf{Y}^T \mathbf{Y} = \mathbf{I} \,\}.$$

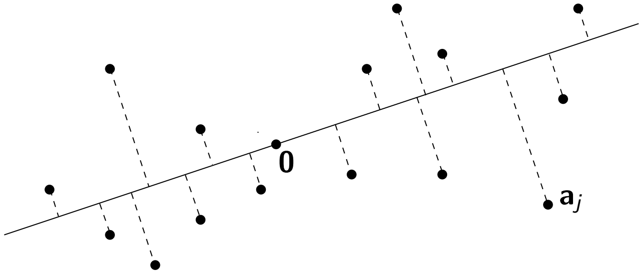
Tu vyřešíme úvahou.

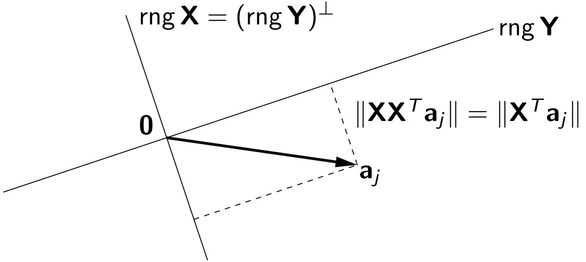
Proložení bodů podprostorem

Úloha na proložení bodů podprostorem

Dány body $\mathbf{a}_1, \ldots, \mathbf{a}_n \in \mathbb{R}^m$ a číslo $k \le m$.

Najdi podprostor dimenze k, který minimalizuje součet čtverců vzdáleností k bodům.





- Hledaný podprostor reprezentujme jako $\operatorname{rng} \mathbf{Y}$ kde $\mathbf{Y} \in \mathbb{R}^{m\times k}$.
- Jeho ortogonální doplněk je $(\operatorname{rng} \mathbf{Y})^\perp = \operatorname{rng} \mathbf{X}$ kde $\mathbf{X} \in \mathbb{R}^{m\times(m-k)}$ a $\mathbf{X}^T\mathbf{X} = \mathbf{I}$.
- Vzdálenost bodu $\mathbf{a}_j$ od podprostoru $\operatorname{rng} \mathbf{Y}$ je $\|\mathbf{X}\mathbf{X}^T\mathbf{a}_j\| = \|\mathbf{X}^T\mathbf{a}_j\|$
- Součet čtverců vzdáleností v bodům je

$$\|\mathbf{X}^T\mathbf{a}_1\|^2 + \cdots + \|\mathbf{X}^T\mathbf{a}_n\|^2 = \|\mathbf{X}^T\mathbf{A}\|^2 = \operatorname{tr}(\mathbf{X}^T\mathbf{A}\mathbf{A}^T\mathbf{X})$$

  kde $\mathbf{A} = [\,\mathbf{a}_1 \ \cdots \ \mathbf{a}_n\,] \in \mathbb{R}^{m\times n}$.

Řešení

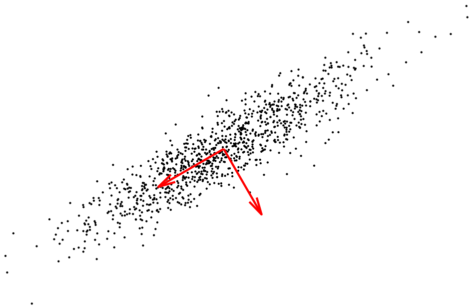
Tedy řešíme

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

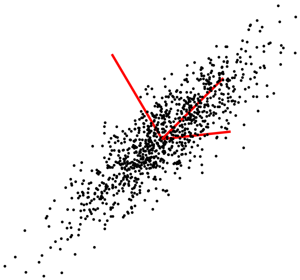
To je úloha na nejmenší stopu s maticí $\mathbf{A}\mathbf{A}^T$.

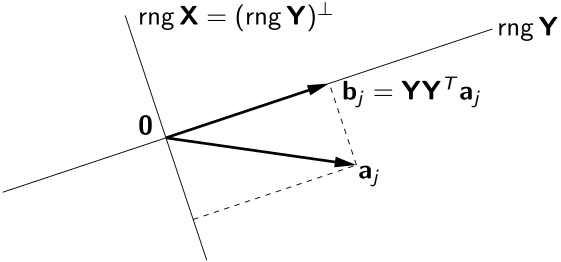
- Spočítej spektrální rozklad $\mathbf{A}\mathbf{A}^T = \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^T$ matice $\mathbf{A}\mathbf{A}^T$.
- Označ $\mathbf{X} \in \mathbb{R}^{m\times(m-k)}$ a $\mathbf{Y} \in \mathbb{R}^{m\times k}$ bloky matice $\mathbf{V} = \begin{bmatrix}\mathbf{X} & \mathbf{Y}\end{bmatrix} \in \mathbb{R}^{m\times m}$.
  Sloupce matice $\mathbf{Y}$ tvoří ortonormální bázi hledaného podprostoru.
  Sloupce matice $\mathbf{X}$ tvoří ortonormální bázi jeho ortogonálního doplňku.
- Optimální hodnota úlohy je $\lambda_1 + \cdots + \lambda_{m-k}$
  (všechna λ_i jsou nezáporná, protože $\mathbf{A}\mathbf{A}^T$ je pos. semidefinitní)

$\mathbf{A} \in \mathbb{R}^{2\times 1000}$, vl. čísla $(1,\ 0.04)$



$\mathbf{A} \in \mathbb{R}^{3\times 1000}$, vl. čísla $(1,\ 0.16,\ 0.04)$





Ortog. projekce bodů $\mathbf{a}_j$ na optimální podprostor:

$$\mathbf{b}_j = \mathbf{Y}\underbrace{\mathbf{Y}^T\mathbf{a}_j}_{\mathbf{c}_j} \qquad \text{neboli} \qquad \mathbf{B} = \mathbf{Y}\underbrace{\mathbf{Y}^T\mathbf{A}}_{\mathbf{C}}$$

kde $\mathbf{c}_j \in \mathbb{R}^k$ (sloupce $\mathbf{C}$) jsou souřadnice bodu $\mathbf{b}_j$ (sloupce $\mathbf{B}$) v ortonormální bázi $\mathbf{Y}$.

Použití:

- **Komprese:** matice $\mathbf{Y} \in \mathbb{R}^{m\times k}$ a $\mathbf{C} \in \mathbb{R}^{k\times n}$ zaberou méně místa než $\mathbf{A} \in \mathbb{R}^{m\times n}$
- **Redukce dimenze** (pro visualizaci, rozpoznávání): body $\mathbf{c}_j$ mají menší dimenzi než $\mathbf{a}_j$

Ekvivalentní formulace jako low-rank approximation

Úloha na nejbližší matici nižší hodnosti (low-rank approximation):

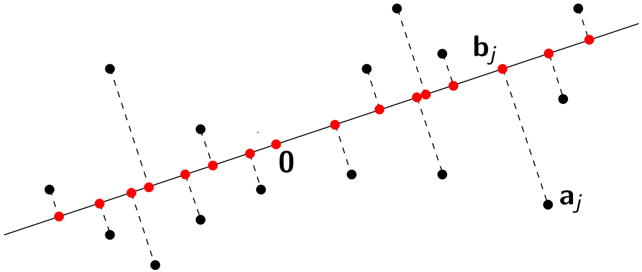
$$\min\{\, \|\mathbf{A} - \mathbf{B}\|^2 \mid \operatorname{rank} \mathbf{B} \leq k,\ \mathbf{B} \in \mathbb{R}^{m\times n} \,\}$$

Tvrzení

Tato úloha má stejnou opt. hodnotu jako úloha na proložení bodů podprostorem.

Její optimální řešení je $\mathbf{B} = \mathbf{Y}\mathbf{Y}^T\mathbf{A}$.

Důkaz: Hledáme body $[\,\mathbf{b}_1 \ \cdots \ \mathbf{b}_n\,] = \mathbf{B}$, které minimalizují $\|\mathbf{a}_1 - \mathbf{b}_1\|^2 + \cdots + \|\mathbf{a}_n - \mathbf{b}_n\|^2 = \|\mathbf{A} - \mathbf{B}\|^2$ za podmínky $\dim \operatorname{span}\{\mathbf{b}_1, \ldots, \mathbf{b}_n\} = \operatorname{rank} \mathbf{B} \leq k$.
Optimální body $\mathbf{b}_i$ jsou ortog. projekce bodů $\mathbf{a}_i$ na optimální podprostor.



Příklad na aproximaci matice

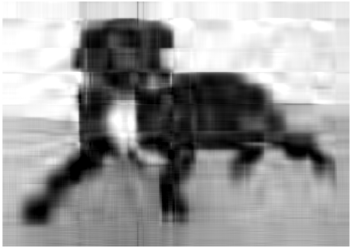
Matice $\mathbf{A} \in [0, 255]^{660 \times 940}$ je fotka:

Matice $\mathbf{B}$ hodnosti k nejbližší k matici $\mathbf{A}$:

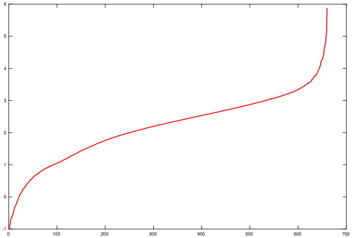
original

$k = 5, \quad \text{rms} \approx 31$

$k = 10, \quad \text{rms} \approx 20$



$\log \lambda_i$



$k = 50, \quad \text{rms} \approx 8$

$k = 100, \quad \text{rms} \approx 5$

Prokládáme afinním podprostorem

Tvrzení

Nechť $\mathbf{a}_1, \ldots, \mathbf{a}_n \in \mathbb{R}^m$ a $k \leq m$. Afinní podprostor dimenze k, který minimalizuje součet čtverců vzdáleností k bodům $\mathbf{a}_1, \ldots, \mathbf{a}_n$, prochází jejich **těžištěm**

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

Důkaz je ve skriptech.

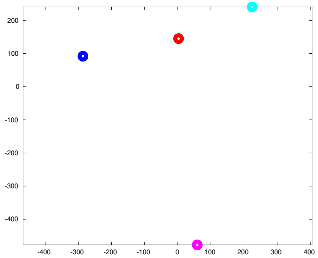
Proložení bodů afinním podprostorem dimenze $k \leq m$:

1. Zadané body posuneme tak, aby měly těžiště v počátku: $\mathbf{a}_i := \mathbf{a}_i - \overline{\mathbf{a}}$.
2. Posunuté body proložíme (lineárním) podprostorem dimenze k
3. Nalezený podprostor posuneme zpátky o $\overline{\mathbf{a}}$.

Příklad: Visualizace + rozpoznávání dat

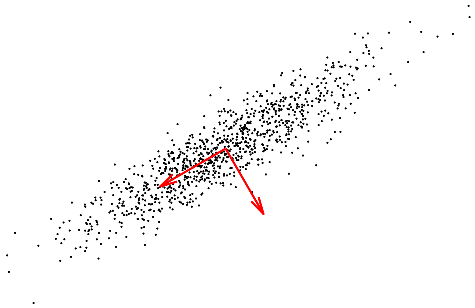
Average consumption of food types in grams per person per week in the UK
(https://setosa.io/ev/principal-component-analysis)

	England	N.Ireland	Scotland	Wales
Alcoholic drinks	375	135	458	475
Beverages	57	47	53	73
Carcase meat	245	267	242	227
Cereals	1472	1494	1462	1582
Cheese	105	66	103	103
Confectionery	54	41	62	64
Fats and oils	193	209	184	235
Fish	147	93	122	160
Fresh fruit	1102	674	957	1137
Fresh potatoes	720	1033	566	874
Fresh Veg	253	143	171	265
Other meat	685	586	750	803
Other Veg	488	355	418	570
Processed potatoes	198	187	220	203
Processed Veg	360	334	337	365
Soft drinks	1374	1506	1572	1256
Sugars	156	139	147	175



- 4 body v 17-rozměrném prostoru
- Odečteme těžiště.
- Proložíme podprostorem dimenze 2 a promítneme body do něj.
- Zobrazíme souřadnice bodů v ortonormální bázi $\mathbf{v}_{16}, \mathbf{v}_{17}$

Statistický pohled: Principal Component Analysis (PCA)



- $\mathbf{a}_1, \ldots, \mathbf{a}_n$ jsou vzorky náhodného vektoru $\mathbf{a}$
- Výběrová střední hodnota:

$$\mathrm{E}[\mathbf{a}] \approx \bar{\mathbf{a}} = \frac{1}{n} \sum_{j=1}^{n} \mathbf{a}_j$$

Označíme $\tilde{\mathbf{a}}_j = \mathbf{a}_j - \bar{\mathbf{a}}$, $\tilde{\mathbf{A}} = [\, \tilde{\mathbf{a}}_1 \; \cdots \; \tilde{\mathbf{a}}_n \,]$.

- Výběrová kovarianční matice:

$$\mathrm{E}[(\mathbf{a} - \mathrm{E}[\mathbf{a}])(\mathbf{a} - \mathrm{E}[\mathbf{a}])^T] \approx \frac{1}{n} \sum_{j=1}^{n} \tilde{\mathbf{a}}_j \tilde{\mathbf{a}}_j^T = \frac{1}{n} \tilde{\mathbf{A}} \tilde{\mathbf{A}}^T$$

Dekorelace:

- Když $\tilde{\mathbf{A}}\tilde{\mathbf{A}}^T = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T$, náhodný vektor $\mathbf{V}^T\mathbf{a}$ má diagonální kovarianční matici (jeho složky jsou nekorelované).
- Vlastní vektory (sloupce $\mathbf{V}$) se někdy nazývají **hlavní komponenty** dat.

Přibližné řešení přeurčených homogenních lineárních soustav

Kdy je homogenní lineární soustava $\mathbf{Ax} = \mathbf{0}$ 'přeurčená'?

- Má vždy triviální řešení $\mathbf{x} = \mathbf{0}$, tedy minimalizovat $\|\mathbf{Ax}\|^2$ nemá smysl.
- Smysluplná formulace:

  $$\min\{\, \|\mathbf{Ax}\|^2 \mid \mathbf{x}^T\mathbf{x} = 1,\ \mathbf{x} \in \mathbb{R}^n \,\}$$

  Řešení: vlastní vektor $\mathbf{A}^T\mathbf{A}$ příslušný nejmenšímu vlastnímu číslu.

- Obecnější formulace:

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

  kde k je předem známá dimenze prostoru řešení soustavy $\mathbf{Ax} = \mathbf{0}$.

Příklad: Přibližné proložení bodů kuželosečkou

Úloha: Najdi kuželosečku, která minimalizuje součet čtverců vzdáleností k daným bodům $(x_1, y_1), \dots, (x_m, y_m) \in \mathbb{R}^2$.

- Kuželosečka: $Q(x, y) = ax^2 + bxy + cy^2 + dx + ey + f = 0$
- Přesné řešení velmi těžké.
- Přibližná formulace:

$$\begin{aligned} \text{min} \quad & Q(x_1, y_1)^2 + \dots + Q(x_m, y_m)^2 \\ \text{za podmínek} \quad & a^2 + b^2 + c^2 + d^2 + e^2 + f^2 = 1 \\ & a, b, c, d, e, f \in \mathbb{R} \end{aligned}$$

Tedy: minimalizuj $\|\mathbf{A}\boldsymbol{\theta}\|$ za podmínky $\boldsymbol{\theta}^T\boldsymbol{\theta} = 1$ kde

$$\mathbf{A} = \begin{bmatrix} x_1^2 & x_1y_1 & y_1^2 & x_1 & y_1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_m^2 & x_my_m & y_m^2 & x_m & y_m & 1 \end{bmatrix}$$

$$\boldsymbol{\theta} = (a, b, c, d, e, f)$$

Singulární rozklad (SVD)

Singulární rozklad (Singular Value Decomposition, SVD)

Věta (existence SVD, redukovaná verze)

Každou matici $\mathbf{A} \in \mathbb{R}^{m \times n}$ lze rozložit jako

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

kde $p = \min\{m, n\}$, $\mathbf{S} \in \mathbb{R}^{p \times p}$ je diagonální a $\mathbf{U}, \mathbf{V}$ mají ortonormální sloupce.

- diagonální prvky $s_1 \geq \cdots \geq s_p \geq 0$ matice $\mathbf{S}$ jsou **singulární čísla**
- sloupce $\mathbf{u}_1, \ldots, \mathbf{u}_p$ matice $\mathbf{U} \in \mathbb{R}^{m \times p}$ jsou **levé singulární vektory**
- sloupce $\mathbf{v}_1, \ldots, \mathbf{v}_p$ matice $\mathbf{V} \in \mathbb{R}^{n \times p}$ jsou **pravé singulární vektory**

Plná verze SVD:

- Doplníme $\mathbf{U}$ přidáním $m - p$ sloupců na ortogonální matici $\mathbf{U} \in \mathbb{R}^{m \times m}$
- Doplníme $\mathbf{V}$ přidáním $n - p$ sloupců na ortogonální matici $\mathbf{V} \in \mathbb{R}^{n \times n}$
- Doplníme $\mathbf{S}$ přidáním $m - p$ nulových řádků a $n - p$ nulových sloupců na $\mathbf{S} \in \mathbb{R}^{m \times n}$

Souvislost SVD se spektrálním rozkladem

Tvrzení

Nenulová sing. čísla matice $\mathbf{A}$ jsou odmocniny nenulových vlast. čísel matice $\mathbf{A}^T\mathbf{A}$ příp. $\mathbf{A}\mathbf{A}^T$.

Důkaz: Jestliže $\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^T$, pak

$$\mathbf{A}^T\mathbf{A} = \mathbf{V}\mathbf{S}^T\mathbf{U}^T\mathbf{U}\mathbf{S}\mathbf{V}^T = \mathbf{V}\mathbf{S}^2\mathbf{V}^T$$
$$\mathbf{A}\mathbf{A}^T = \mathbf{U}\mathbf{S}\mathbf{V}^T\mathbf{V}\mathbf{S}^T\mathbf{U}^T = \mathbf{U}\mathbf{S}^2\mathbf{U}^T$$

To jsou spektrální rozklady matic $\mathbf{A}^T\mathbf{A}$ a $\mathbf{A}\mathbf{A}^T$.

Výpočet SVD pomocí spektrálního rozkladu:

- Spočteme $\mathbf{V}$ a $\mathbf{S}$ ze $\mathbf{A}^T\mathbf{A} = \mathbf{V}\mathbf{S}^2\mathbf{V}^T$.
- Pro nenulová sing. čísla s_j spočteme $\mathbf{u}_j = \mathbf{A}\mathbf{v}_j/s_j$ (důkaz ve skriptech).

Algoritmy na SVD se ale vyhýbají výpočtu součinu $\mathbf{A}^T\mathbf{A}$ nebo $\mathbf{A}\mathbf{A}^T$. V Matlabu:

- plná verze: `[U,S,V]=svd(A)`
- redukovaná verze: `[U,S,V]=svd(A,'econ')`

Nejbližší maticí nižší hodnosti z SVD

Zopakujme úlohu na nejbližší matici nižší hodnosti ($k \le p = \min\{m, n\}$):

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

Věta (Eckart-Young)

Nechť $\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^T = s_1\mathbf{u}_1\mathbf{v}_1^T + \cdots + s_p\mathbf{u}_p\mathbf{v}_p^T$. Opt. řešení úlohy je

$$\mathbf{B} = \mathbf{U}\mathbf{S}_k\mathbf{V}^T = s_1\mathbf{u}_1\mathbf{v}_1^T + \cdots + s_k\mathbf{u}_k\mathbf{v}_k^T \qquad \text{kde} \quad \mathbf{S}_k = \begin{bmatrix} \mathbf{I}_k & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{S}.$$

Myšlenka důkazu: Spočítej optimální $\mathbf{Y}$ ze spektrálního rozkladu $\mathbf{A}\mathbf{A}^T$ a dosaď do $\mathbf{B} = \mathbf{Y}\mathbf{Y}^T\mathbf{A}$.

Minimalizace lineární funkce matice s ortonormálními sloupci

Věta

Nechť $\mathbf{A} \in \mathbb{R}^{m \times n}$ kde $m \geq n$. Úloha

$$\max\{ \langle \mathbf{A}, \mathbf{X} \rangle \mid \mathbf{X}^T\mathbf{X} = \mathbf{I},\ \mathbf{X} \in \mathbb{R}^{n \times k} \}$$

nabývá optimum pro $\mathbf{X} = \mathbf{U}\mathbf{V}^T$ kde $\mathbf{U}\mathbf{S}\mathbf{V}^T = \mathbf{A}$ je redukované SVD matice $\mathbf{A}$.

Důkaz je ve skriptech.

Příklad: Nejbližší isometrie

Pro danou $\mathbf{A} \in \mathbb{R}^{m \times n}$ najděte nejbližší matici s ortonormálními sloupci:

$$\min\{\, \|\mathbf{X} - \mathbf{A}\|^2 \mid \mathbf{X}^T \mathbf{X} = \mathbf{I},\ \mathbf{X} \in \mathbb{R}^{m \times n} \,\}.$$

Řešení: Platí

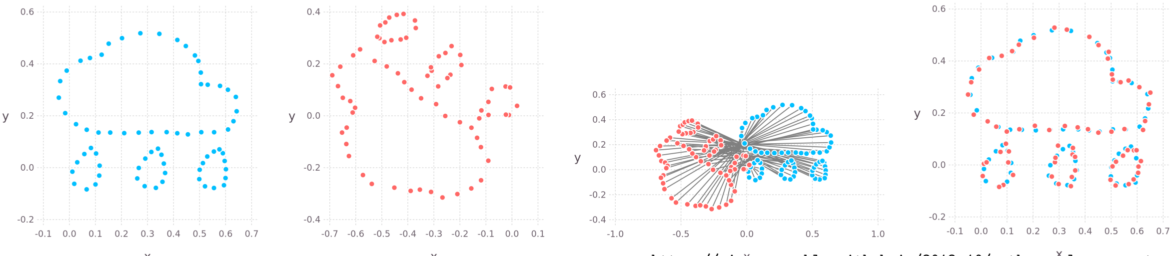
$$\|\mathbf{X} - \mathbf{A}\|^2 = \langle \mathbf{X} - \mathbf{A}, \mathbf{X} - \mathbf{A} \rangle = \|\mathbf{X}\|^2 - 2\langle \mathbf{A}, \mathbf{X} \rangle + \|\mathbf{A}\|^2$$

Výrazy $\|\mathbf{A}\|^2$ a $\|\mathbf{X}\|^2 = \operatorname{tr}(\mathbf{X}^T \mathbf{X}) = \operatorname{tr} \mathbf{I} = n$ nezávisí na $\mathbf{X}$.

Tedy stačí maximalizovat $\langle \mathbf{A}, \mathbf{X} \rangle$, čímž se úloha převedla na minulou úlohu.

Dány $\mathbf{A} = [\mathbf{a}_1 \ \cdots \ \mathbf{a}_k] \in \mathbb{R}^{n\times k}$ a $\mathbf{B} = [\mathbf{b}_1 \ \cdots \ \mathbf{b}_k] \in \mathbb{R}^{m\times k}$, kde $m \geq n$.
Najdi isometrii $\mathbf{X} \in \mathbb{R}^{m\times n}$, aby body $\mathbf{a}'_j = \mathbf{X}\mathbf{a}_j$ byly co nejblíže bodům $\mathbf{b}_j$:

$$\text{minimalizuj} \quad \sum_{j=1}^{k} \|\mathbf{a}'_j - \mathbf{b}_j\|^2 = \|\mathbf{X}\mathbf{A} - \mathbf{B}\|^2 \qquad \text{za podmínky} \quad \mathbf{X}^T\mathbf{X} = \mathbf{I}$$



https://simonensemble.github.io/2018-10/orthogonal-procrustes

Řešení: Platí

$$\|\mathbf{X}\mathbf{A} - \mathbf{B}\|^2 = \|\mathbf{X}\mathbf{A}\|^2 - 2\langle \mathbf{X}\mathbf{A}, \mathbf{B}\rangle + \|\mathbf{B}\|^2$$

Výrazy $\|\mathbf{X}\mathbf{A}\|^2 = \|\mathbf{A}\|^2$ a $\|\mathbf{B}\|^2$ nezávisejí na $\mathbf{X}$, tedy maximalizujeme $\langle \mathbf{X}\mathbf{A}, \mathbf{B}\rangle = \langle \mathbf{B}\mathbf{A}^T, \mathbf{X}\rangle$.