Optimalizace

9. SVD

Tomáš Kroupa Tomáš Werner

2021 LS

Fakulta elektrotechnická
ČVUT v Praze

Motivace

Low rank approximation pro symetrickou matici $\mathbf{A} \in \mathbb{R}^{n \times n}$

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

- $\mathbf{A} = \lambda_1 \mathbf{v}_1 \mathbf{v}_1^\mathsf{T} + \cdots + \lambda_n \mathbf{v}_n \mathbf{v}_n^\mathsf{T}$, předpokládáme $\lambda_1 \leq \cdots \leq \lambda_n$
- Díky symetrii matice $\mathbf{A}$ je řešení úlohy v elegantním tvaru

$$\mathbf{B} = \lambda_{n-k+1} \mathbf{v}_{n-k+1} \mathbf{v}_{n-k+1}^\mathsf{T} + \cdots + \lambda_n \mathbf{v}_n \mathbf{v}_n^\mathsf{T}$$

- Hledáme zobecnění spektrálního rozkladu umožňující podobné vyjádření řešení pro libovolnou matici $\mathbf{A} \in \mathbb{R}^{m \times n}$

Singulární rozklad matice

Věta (Existence SVD)

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

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

kde diagonální matice $\mathbf{S} \in \mathbb{R}^{m \times n}$ má na diagonále $p := \min\{m, n\}$ singulárních čísel $s_1, \ldots, s_p \geq 0$ a ortogonální matice

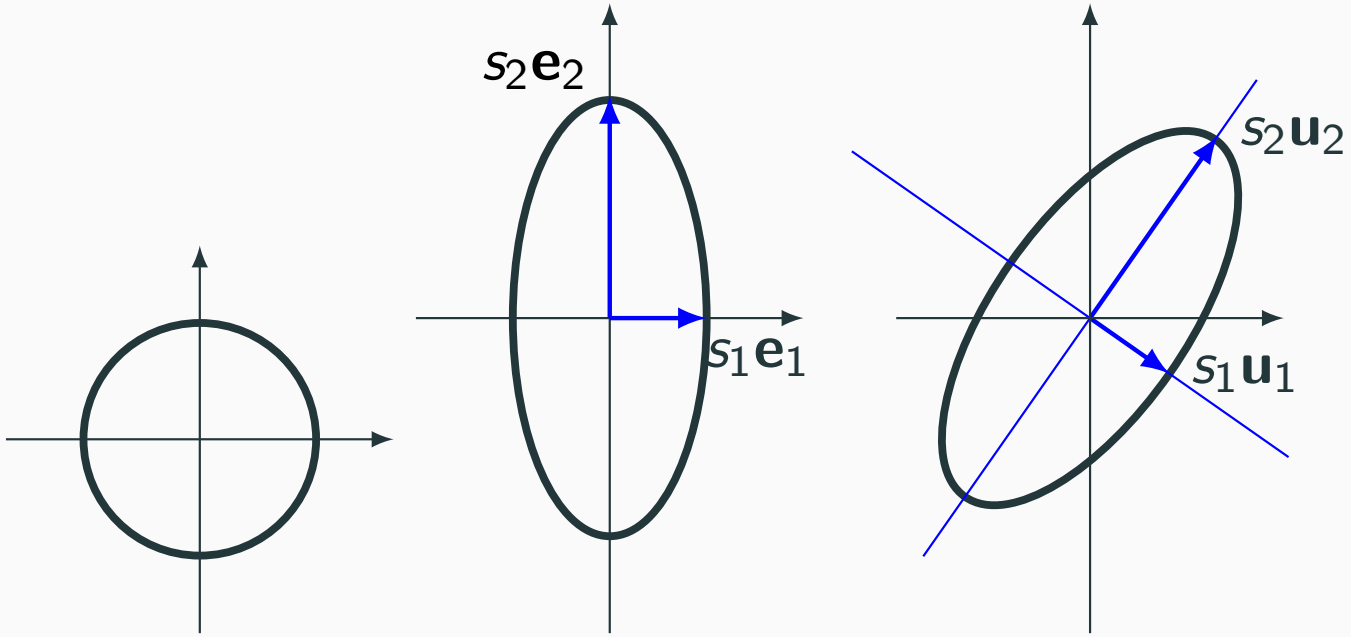
$$\mathbf{U} = \begin{bmatrix} \mathbf{u}_1 \cdots \mathbf{u}_m \end{bmatrix} \in \mathbb{R}^{m \times m}, \qquad \mathbf{V} = \begin{bmatrix} \mathbf{v}_1 \cdots \mathbf{v}_n \end{bmatrix} \in \mathbb{R}^{n \times n}$$

mají ve sloupcích levé/pravé singulární vektory.

Singulární čísla řadíme sestupně: $s_1 \geq \cdots \geq s_p$

Lineární transformace s regulární maticí A jednotkové kružnice

$$\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^\mathsf{T} = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 \end{bmatrix} \begin{bmatrix} s_1 & 0 \\ 0 & s_2 \end{bmatrix} \begin{bmatrix} \mathbf{v}_1^\mathsf{T} \\ \mathbf{v}_2^\mathsf{T} \end{bmatrix} = s_1\mathbf{u}_1\mathbf{v}_1^\mathsf{T} + s_2\mathbf{u}_2\mathbf{v}_2^\mathsf{T}$$



Obrazem jednotkové kružnice je natočená elipsa s délkami os s_1,s_2.

Různé verze SVD

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

Plné SVD

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

Redukované SVD

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

Rank-minimální SVD

$\mathbf{U} \in \mathbb{R}^{m\times r}$, $\mathbf{S} \in \mathbb{R}^{r\times r}$, $\mathbf{V} \in \mathbb{R}^{n\times r}$, kde $r := \operatorname{rank}\mathbf{S}$

Tedy r je počet nenulových singulárních čísel a $r = \operatorname{rank}\mathbf{A}$.

SVD symetrické matice

Spektrální rozklad symetrické matice

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

Seřaďme diagonálu $\boldsymbol{\Lambda}$ takto: $|\lambda_1| \geq \cdots \geq |\lambda_n|$

- Když je $\mathbf{A}$ pozitivně semidefinitní, pak $\mathbf{A} = \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^\mathsf{T}$ je SVD
- Pokud není, pak dostaneme SVD $\mathbf{A} = \mathbf{V}'\mathbf{S}\mathbf{V}^\mathsf{T}$ tak, že položíme $\mathbf{S} = |\boldsymbol{\Lambda}|$ a $\mathbf{V}'$ vznikne z $\mathbf{V}$ násobením -1 sloupců odpovídajích záporným vlastním číslům

Jak souvisí singulární čísla s vlastními čísly?

Spektrální rozklad matic $\mathbf{AA}^\mathsf{T}$ a $\mathbf{A}^\mathsf{T}\mathbf{A}$ pomocí SVD

$$\mathbf{AA}^\mathsf{T} = \mathbf{USV}^\mathsf{T}\mathbf{VS}^\mathsf{T}\mathbf{U}^\mathsf{T} = \mathbf{USS}^\mathsf{T}\mathbf{U}^\mathsf{T}$$
$$\mathbf{A}^\mathsf{T}\mathbf{A} = \mathbf{VS}^\mathsf{T}\mathbf{U}^\mathsf{T}\mathbf{USV}^\mathsf{T} = \mathbf{VS}^\mathsf{T}\mathbf{SV}^\mathsf{T}$$

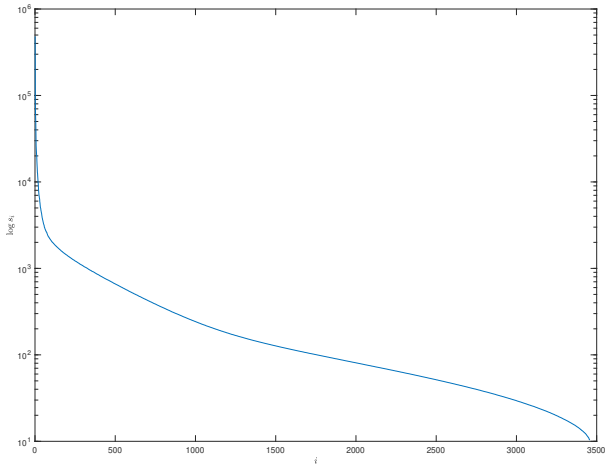
- Matice $\mathbf{AA}^\mathsf{T}$ a $\mathbf{A}^\mathsf{T}\mathbf{A}$ mají stejná nenulová vlastní čísla
- Nenulová singulární čísla matice $\mathbf{A}$ jsou odmocniny nenulových vlastních čísel matice $\mathbf{AA}^\mathsf{T}$
- Levé singulární vektory jsou vlastní vektory matice $\mathbf{AA}^\mathsf{T}$
- Pravé singulární vektory jsou vlastní vektory matice $\mathbf{A}^\mathsf{T}\mathbf{A}$

Singulární čísla některých matic

- Ortogonální matice má všechna singulární čísla rovna 1
- Hilbertova matice $\mathbf{A} = \left[\frac{1}{i+j-1}\right]_{ij}$ řádu n je regulární a např. pro $n = 100$ platí

$$s_1 = 2.1827, \ldots, s_{100} \approx 10^{-17}$$

- Černobílý obrázek psa je matice 3456×4608 s plnou řádkovou hodností, její singulární čísla jsou v grafu s $\log_{10}$ stupnicí



Low rank approximation pomocí redukovaného SVD

$$\mathbf{A} \in \mathbb{R}^{m\times n}, \quad \mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^{\mathsf{T}}, \quad s_1 \geq \cdots \geq s_p, \quad p = \min\{m, n\}$$

Věta (Eckart-Young)

Nechť $k \leq p$. Řešením úlohy

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

je matice

$$\mathbf{B}_k = \mathbf{U}\mathbf{S}_k\mathbf{V}^{\mathsf{T}} = s_1\mathbf{u}_1\mathbf{v}_1^{\mathsf{T}} + \cdots + s_k\mathbf{u}_k\mathbf{v}_k^{\mathsf{T}},$$

kde $\mathbf{S}_k$ vznikne z $\mathbf{S}$ vynulováním singulárních čísel $s_{k+1}, \ldots, s_p$.

Měříme kvalitu low rank approximation

Tvrzení

Pro každou matici $\mathbf{A} \in \mathbb{R}^{m \times n}$ hodnosti r platí

$$\|\mathbf{A}\| = \sqrt{s_1^2 + \cdots + s_r^2}.$$

Relativní chybu aproximace maticí $\mathbf{B}_k$ hodnosti nejvýše k lze zjistit ze singulárních čísel matice $\mathbf{A}$:

- Pro $k = 1, \ldots, r-1$ dostaneme

$$\frac{\|\mathbf{A} - \mathbf{B}_k\|}{\|\mathbf{A}\|} = \sqrt{\frac{\sum_{i=k+1}^{r} s_i^2}{s_1^2 + \cdots + s_r^2}}$$

- Pro $r \leq k \leq p$ triviálně platí $\mathbf{B}_k = \mathbf{A}$ a chyba je 0

Komprese obrázku psa pomocí SVD

$k = 5$, chyba 17%, $5 \times (3456 + 4608)$

$k = 20$, chyba 9%, $20 \times (3456 + 4608)$

$k = 200$, chyba 4%, $200 \times (3456 + 4608)$

Originál 3456×4608

Jak počítat SVD?

Teoreticky SVD spočítáme ze spektrálního rozkladu:

- Nenulová singulární čísla matice $\mathbf{A}$ jsou odmocniny nenulových vlastních čísel matice $\mathbf{A}\mathbf{A}^\mathsf{T}$
- Levé singulární vektory jsou vlastní vektory matice $\mathbf{A}\mathbf{A}^\mathsf{T}$
- Pravé singulární vektory jsou vlastní vektory matice $\mathbf{A}^\mathsf{T}\mathbf{A}$

- Prakticky použitelné jen v některých situacích (např. $m \gg n$)
- Je vhodnější používat specializované *SVD algoritmy*, které se vyhýbají explicitnímu výpočtu matic $\mathbf{A}\mathbf{A}^\mathsf{T}$ a $\mathbf{A}^\mathsf{T}\mathbf{A}$

Výběrová kovarianční matice

Matice $\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 \cdots \mathbf{a}_n \end{bmatrix} \in \mathbb{R}^{m \times n}$ má datové vektory $\mathbf{a}_i$ ve sloupcích a předpokládáme, že výběrové průměry veličin jsou nulové, $\overline{\mathbf{a}} = \mathbf{0}$.

Definice

Výběrová kovarianční matice je matice $\frac{1}{n}\mathbf{A}\mathbf{A}^\mathsf{T}$. Na diagonále má *výběrové rozptyly* a jinde *výběrové kovariance* veličin i a j:

$$[c_{ij}] = \frac{1}{n}\mathbf{A}\mathbf{A}^\mathsf{T}, \qquad c_{ii} = \frac{1}{n}\sum_{\ell=1}^{n} a_{\ell i}^2, \qquad c_{ij} = \frac{1}{n}\sum_{\ell=1}^{n} a_{\ell i} a_{\ell j}$$

Stejnou interpretaci má matice $\frac{1}{n}\mathbf{A}^\mathsf{T}\mathbf{A}$, pokud jsou normalizované datové vektory uspořádány do řádků matice $\mathbf{A} \in \mathbb{R}^{n \times m}$.

Příklad (1)

3 recenzenti posuzují 4 různé filmy (body $0, \ldots, 5$). Po normalizaci průměrným hodnocením filmů dostaneme $\mathbf{A}$ a kovarianční matici:

$$\begin{bmatrix} 5 & 4 & 1 \\ 5 & 5 & 0 \\ 0 & 0 & 5 \\ 1 & 0 & 4 \end{bmatrix} \qquad \mathbf{A} = \begin{bmatrix} 1.67 & 0.67 & -2.33 \\ 1.67 & 1.67 & -3.33 \\ -1.67 & -1.67 & 3.33 \\ -0.67 & -1.67 & 2.33 \end{bmatrix}$$

$$\tfrac{1}{3}\mathbf{A}\mathbf{A}^{\mathsf{T}} = \begin{bmatrix} 2.89 & 3.89 & -3.89 & -2.56 \\ 3.89 & 5.56 & -5.56 & -3.89 \\ -3.89 & -5.56 & 5.56 & 3.89 \\ -2.56 & -3.89 & 3.89 & 2.89 \end{bmatrix}$$

Hodnoty kovariancí naznačují, že existují $k = 2$ typy filmů. Snažme se je najít pomocí PCA.

Příklad (2)

- SVD matice $\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^\mathsf{T}$: dostaneme vlastní vektory matice $\mathbf{A}\mathbf{A}^\mathsf{T}$ z matice $\mathbf{U} = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 & \mathbf{u}_4 \end{bmatrix}$, jsou seřazeny podle singulárních čísel $s_1 = 7.05$, $s_2 = 1$, $s_3 = s_4 = 0$
- PCA pro $k = 2$: klademe $\mathbf{Y} := \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 \end{bmatrix}$ a matice souřadnic datových vektorů promítnutých do $\mathbb{R}^2$ je

$$\mathbf{C} = \mathbf{Y}^\mathsf{T}\mathbf{A} = \begin{bmatrix} -2.88 & -2.88 & 5.74 \\ -0.71 & 0.71 & 0 \end{bmatrix}$$

- Chyba aproximace je nulová díky $s_3 = s_4 = 0$ a proto jsou data pouze *dvourozměrná*

Reference

- T. Werner. *Optimalizace* (kapitola 7). Elektronická skripta. FEL ČVUT, 2020.
- M.P. Deisenroth et al. *Mathematics for Machine Learning*. `https://mml-book.github.io`