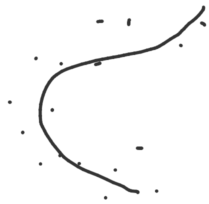
RDD - 07

29 March 2021 14:42

- o týždeň: sviatky (5.4.) → odpadá
- 12.4.: nebude naživo, iba video na stránke

Minule: Principal curves – krivka prechádzajúca stredom dát
– projection - expectation algoritmus



R-ko: food.txt

→ postupne:

cholesterol → tuky / kalórie → Bielkoviny → uhľovodíky

7 9 4

MNIST: zač. (vľavo): 0

koniec: 1

medzitým: zhluk 2/6

: 3

: zhluk 7/9/4

5, 8 - rozptýlené

ďalšie nelin. PCA: – nelineárna PCA pomocou neurónových sietí

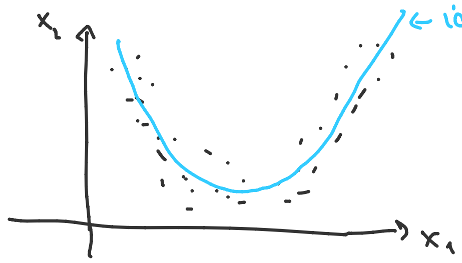
\ jadrová PCA

Jadrová PCA (Kernel PCA)

1.) nelineárne transformujeme dáta – do "ešte viac"-rozmerného priest.

2.) bežné PCA v tom priestore

Motivácia:



← ideálny nelineárny PC_1: $x_2 = ax_1^2 + bx_1 + c$

$x_2 - ax_1^2 - bx_1 - c = 0$

Vo svete: $\begin{pmatrix} x_1 \\ x_2 \\ x_1^2 \end{pmatrix}$ je ideálny PC_1:

$$(-b \; 1 \; -a)\begin{pmatrix} x_1 \\ x_2 \\ x_1^2 \end{pmatrix} + (-c) = 0$$

↳ lineárne: priamka

Formálne:

1.) transformácia: $\Phi : \mathbb{R}^k \to \mathbb{R}^p$

$(p >> k)$

$\Phi_1(x_i)$

$(p >> q)$

$$x_i \longmapsto \Phi(x_i) = \begin{pmatrix} \Phi_1(x_i) \\ \vdots \\ \Phi_p(x_i) \end{pmatrix}$$

$q \times 1$ $\quad$ $p \times 1$

– dokonca aj ∞ – rozmerné
(pre jednoduchosť $\mathbb{R}^p$, $p < \infty$)

2.) Spektrálny rozklad v $\mathbb{R}^p$:

Predp.: $\sum_{i=1}^{m} \Phi(x_i) = 0_p$: vycentrované v $\mathbb{R}^p$

$$S^* = \frac{1}{m} \sum_{i=1}^{m} \Phi(x_i) \cdot \Phi(x_i)^T = \frac{1}{m} \underset{p \times m}{\Phi^T} \underset{m \times p}{\Phi} \quad : \mathbb{R}^{p \times p}$$

$$\Phi = \begin{bmatrix} - \Phi(x_1)^T - \\ \vdots \\ - \Phi(x_m)^T - \end{bmatrix}_{m \times p}$$

– vl. h. a vektory: $S^* v_j = \lambda_j v_j \qquad j = 1, \dots, p$

$\|v_j\| = 1 \quad \forall j$

$r = \operatorname{rank}(S^*) = \operatorname{rank}(\Phi) < m$
$< p$

$\Rightarrow$ vl. hodnoty : $\lambda_1, \dots, \lambda_r, 0, \dots, 0$
v. vekt. : $v_1, \dots, v_r, v_{r+1}, \dots, v_p$

len tieto nás zaujímajú

Problém: $p \times p$, p – veľké (∞)

Riešenie: Jadrový trik: kl. myšlienka: $\boxed{\text{kladné vl. h. } A^T A = \text{kladné vl. h. } A A^T}$

$\Rightarrow \underbrace{\underset{p \times m}{\Phi^T} \underset{m \times p}{\Phi}}_{p \times p}$: rovnaké ⊕ vl. h. ako $\underbrace{\underset{m \times p}{\Phi} \underset{p \times m}{\Phi^T}}_{m \times m}$: oveľa menej!

$S^* = \frac{1}{m} \Phi^T \Phi \Rightarrow \Phi^T \Phi = m S^* \Rightarrow$ vl. h. : $\boxed{m \lambda_j \overset{\text{ozn.}}{=} \tilde{\lambda}_j}$
vl. vekt. : v_j

$\Rightarrow \underline{\Phi \Phi^T}$: vl. h. (kladné) : $\tilde{\lambda}_j \quad (j = 1, \dots, r)$
vl. vekt. : $\underline{b_j}$

Platí: $\boxed{v_j = \Phi^T b_j}$

Overme:

$$\underline{\Phi^T \Phi} \cdot (\Phi^T b_j) = \tilde{\lambda}_j \underbrace{\Phi^T b_j}_{v_j} = \tilde{\lambda}_j v_j$$

: otázka je $\Phi^T b_j$ vl. vektor $\Phi^T\Phi$

$$\underset{m \times m}{\Phi \Phi^T} = \begin{bmatrix} \Phi(x_1)^T \\ \vdots \\ \Phi(x_m)^T \end{bmatrix} \begin{bmatrix} \Phi(x_1) & \dots & \Phi(x_m) \end{bmatrix}$$

$$\underset{m\times m}{} \begin{bmatrix} \vdots \\ \Phi(x_m)^T \end{bmatrix} [\Phi(x_1), \dots, \Phi(x_m)]$$

: otázka je $\Phi^T b_j$ $\overset{v_j}{}$ vl. vektor $\Phi^T\Phi$

$$= \begin{bmatrix} \Phi(x_1)^T\Phi(x_1) & & \Phi(x_i)^T\Phi(x_j) \\ & \ddots & \\ & & \Phi(x_m)^T\Phi(x_m) \end{bmatrix} \overset{\text{ozn.}}{=} \underset{m\times m}{K}$$

$$K_{ij} = \Phi(x_i)^T\Phi(x_j) = \langle \Phi(x_i), \Phi(x_j) \rangle$$

– skal. súčin

↳ matica skal. súčinov

↳ jadro

$$\|v_j\| = 1 \Rightarrow 1 = v_j^T v_j = b_j^T \underbrace{\Phi\Phi^T}_{K} b_j = b_j^T \underbrace{K b_j}_{\tilde\lambda_j \cdot b_j} = \tilde\lambda_j \underbrace{b_j^T b_j}_{\|b_j\|^2} \Rightarrow \|b_j\| = \frac{1}{\tilde\lambda_j^{1/2}}$$

označme: vl. vektory K s dĺžkou 1: $\tilde b_j$

$$\Rightarrow b_j = \frac{1}{\tilde\lambda_j^{1/2}} \cdot \tilde b_j \qquad (j = 1, \dots, r)$$

Spekt. rozklad K: $\underset{m\times r}{\tilde B}\ \underset{r\times r}{\tilde\Lambda}\ \underset{r\times m}{\tilde B^T}$ $\qquad \tilde B = [\tilde b_1, \dots, \tilde b_r]$

$$K \tilde b_j = \tilde\lambda_j \tilde b_j \quad (j = 1, \dots, r) \qquad \tilde\Lambda = \begin{bmatrix} \tilde\lambda_1 & & 0 \\ & \ddots & \\ 0 & & \tilde\lambda_r \end{bmatrix}$$

Hľadáme: • v_j

$$v_j = \Phi^T b_j = \frac{1}{\tilde\lambda_j^{1/2}} \Phi^T \tilde b_j$$

• λ_j:

$$\lambda_j = \frac{1}{m} \tilde\lambda_j$$

→ Stačí mi spekt. rozklad K

→ Jadrový trik

skóre: súradnice $\Phi(x_1), \dots, \Phi(x_m)$ voči $v_1, \dots, v_r$

$$\Rightarrow \Phi(x_i) = V y_i \qquad / \cdot V^T$$

$$V = \underset{p\times r}{[v_1, \dots, v_r]}, \qquad \Lambda = \underset{r\times r}{\begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_r \end{bmatrix}}$$

$$S^* = V \Lambda V^T$$

$$V^T\Phi(x_i) = y_i \qquad i = 1, \dots, m$$

$$V^T V = I_r$$

$$v_j = \frac{1}{\sqrt{\tilde\lambda_j}} \Phi^T \tilde b_j$$

$$V^T [\Phi(x_1), \dots, \Phi(x_m)] = [y_1, \dots, y_m]$$

$V^TV = I_r$

$$V^T \underbrace{\left[\Phi(x_1), \dots, \Phi(x_n)\right]}_{\Phi^T} = \underbrace{\left[y_1, \dots, y_n\right]}_{Y^T}$$

$$v_j = \frac{1}{\sqrt{\tilde{\lambda}_j}} \Phi^T \tilde{b}_j$$

$$\Downarrow$$

$$\boxed{V = \Phi^T \tilde{B} \tilde{\Lambda}^{-1/2}}$$

$$V^T \Phi^T = Y^T$$

skóre: $\boxed{Y} = \Phi V \overset{\downarrow}{=} \underbrace{\Phi\Phi^T}_{K} \tilde{B} \tilde{\Lambda}^{-1/2} = K \tilde{B} \tilde{\Lambda}^{-1/2} = \tilde{B} \tilde{\Lambda} \tilde{\Lambda}^{-1/2}$

$K\tilde{B} = \tilde{B}\tilde{\Lambda}$

$$\boxed{= \tilde{B} \tilde{\Lambda}^{1/2}}$$

→ oraz stačí $K = \Phi\Phi^T$, potom spekt. rozklad: $K = \tilde{B}\tilde{\Lambda}\tilde{B}^T$

$\Rightarrow$ skóre: $\underset{n \times r}{Y} = \underset{n \times r}{\tilde{B}} \, \underset{r \times r}{\tilde{\Lambda}^{1/2}}$

Redukcia dimenzie: vezmem iba prvých $q < r$ vl. hodnôt a vektorov

$\Leftrightarrow$ vezmem prvých q stĺpcov Y

Jadrová PCA

Poznámky:

- ak dáta nie sú centrované vo Φ-priestore, tak akurát na miesto K vezmem centrované K:
$$\tilde{K} = \left(I_n - \tfrac{1}{n} J_n\right) K \left(I_n - \tfrac{1}{n} J_n\right)$$
J_n: $n \times n$ matica jednotiek

- pre všeob. priestor (aj ∞-rozmerný) – pracujeme so skal. súčinmi
$$: K_{ij} = \langle \Phi(x_i), \Phi(x_j) \rangle \quad (i, j = 1, \dots, n)$$

Jadrá:

v praxi nevolíme Φ, ale iba skal. súčiny (jadrá)

- gaussovské jadro: $K_{ij} = \langle \Phi(x_i), \Phi(x_j) \rangle = e^{-\frac{\|x_i - x_j\|^2}{2\sigma^2}}$ ← v R-ku trochu iný predpis

- laplaceovské: $K_{ij} = e^{-\frac{\|x_i - x_j\|}{\sigma}}$

- sigmoid: $K_{ij} = \tanh\left(a \langle x_i, x_j \rangle + b\right)$

  ↳ hyp. tangens: $\tanh(t) = \dfrac{e^t - e^{-t}}{e^t + e^{-t}}$

R-ko: gaussovské: $K_{ij} = e^{-\sigma_R \|x_i - x_j\|^2}$ $\qquad \sigma_R = \frac{1}{\dots}$, Rko: gaussovské: $K_{ij} = e^{-\sigma_R \|x_i - x_j\|^2}$ $\qquad \sigma_R = \dfrac{1}{2\sigma^2_{naša}}$

$\sigma_{naša} = 4 \Rightarrow \sigma_R = \dfrac{1}{2 \cdot 4^2}$

vstupom: $\sigma_{naša} = 1$

food → postupne (sprava ↺): ~~...~~

málo kalorické → bielkoviny → uhľovodíky → tuky/kalórie → cholesterol

(prehodené ~~...~~ uhľovodíky a bielkoviny oproti el. krivke)

MNIST – všetky v 1 zhluku

– zopár úplne mimo zhluku