PCA
Analýza Hlavních Komponent

Jiřina Valdaufová

Podzim 2008

Obsah

- Matematika
- PCA – orientace objektů
- PCA – konverze barevného obrázku

Průměr

- Vzoreček

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

- Matlab
  - mean()
- Příklad
  - X1 = [2 2 5]
    - 3.000
  - X2 = [1 2 5]
    - 2.667

Rozptyl

- Vzoreček

$$s = \sqrt{\frac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}{(n-1)}}$$

- Matlab
  - var()
- Jiná forma zápisu

$$s^2 = \frac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}{(n-1)}$$

- Příklad
  - X1 = [2 2 5]
    - 3.000

Kovariance

- Rozptyl

$$\mathrm{var}(X) = \frac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)\left(X_i - \overline{X}\right)}{(n-1)}$$

- Kovariance

$$\mathrm{cov}(X) = \frac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)\left(Y_i - \overline{Y}\right)}{(n-1)}$$

- Příklad
  - X = [2 2 5], Y = [3 2 4]
    - 1.500

Kovarianční Matice

- Vzoreček

$$C^{n\times x} = \left(c_{i,j}\right), c_{i,j} = \operatorname{cov}\left(Dim_i, Dim_j\right)$$

- Matlab
  - cov()
- Příklad
  - X = [2 2 5]
  - Y = [3 2 4]

    C = [3.000 1.500
    1.500 1.000]

$$C = \begin{pmatrix} \operatorname{cov}(x,x) & \operatorname{cov}(x,y) \\ \operatorname{cov}(y,x) & \operatorname{cov}(y,y) \end{pmatrix}$$

Vlastní Čísla A Vektory

- Je-li A čtvercová matice řádu *n* a existuje-li číslo *λ* a vektor $u \in R_n, u \neq 0$, pro které platí

  $$A\vec{u} = \lambda\vec{u}$$

  pak *λ* se nazývá vlastní číslo matice *A* a *u* vlastním vektorem matice *A*

- Charakteristická rovnice $(A - \lambda E)u = 0$

  má nenulové řešení, pokud $\det(A - \lambda E) = 0$

Vlastní Čísla A Vektory

- Matlab
  - eig()
- Příklad

$$A = \begin{pmatrix} 3 & -1 \\ 2 & 0 \end{pmatrix}$$

$$\begin{vmatrix} 3-\lambda & -1 \\ 2 & -\lambda \end{vmatrix} = 0$$

$$\lambda^2 - 3\lambda + 2 = 0$$

$$\lambda_1 = 2, \lambda_2 = 1$$

$$\begin{pmatrix} 3-\lambda_1 & -1 \\ 2 & -\lambda_1 \end{pmatrix} \cdot \vec{u}_1 = 0$$

$$\begin{pmatrix} 1 & -1 \\ 2 & -2 \end{pmatrix} \cdot \vec{u}_1 = 0$$

$$\vec{u}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} 3-\lambda_2 & -1 \\ 2 & -\lambda_2 \end{pmatrix} \cdot \vec{u}_2 = 0$$

$$\begin{pmatrix} 2 & -1 \\ 2 & -1 \end{pmatrix} \cdot \vec{u}_2 = 0$$

$$\vec{u}_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

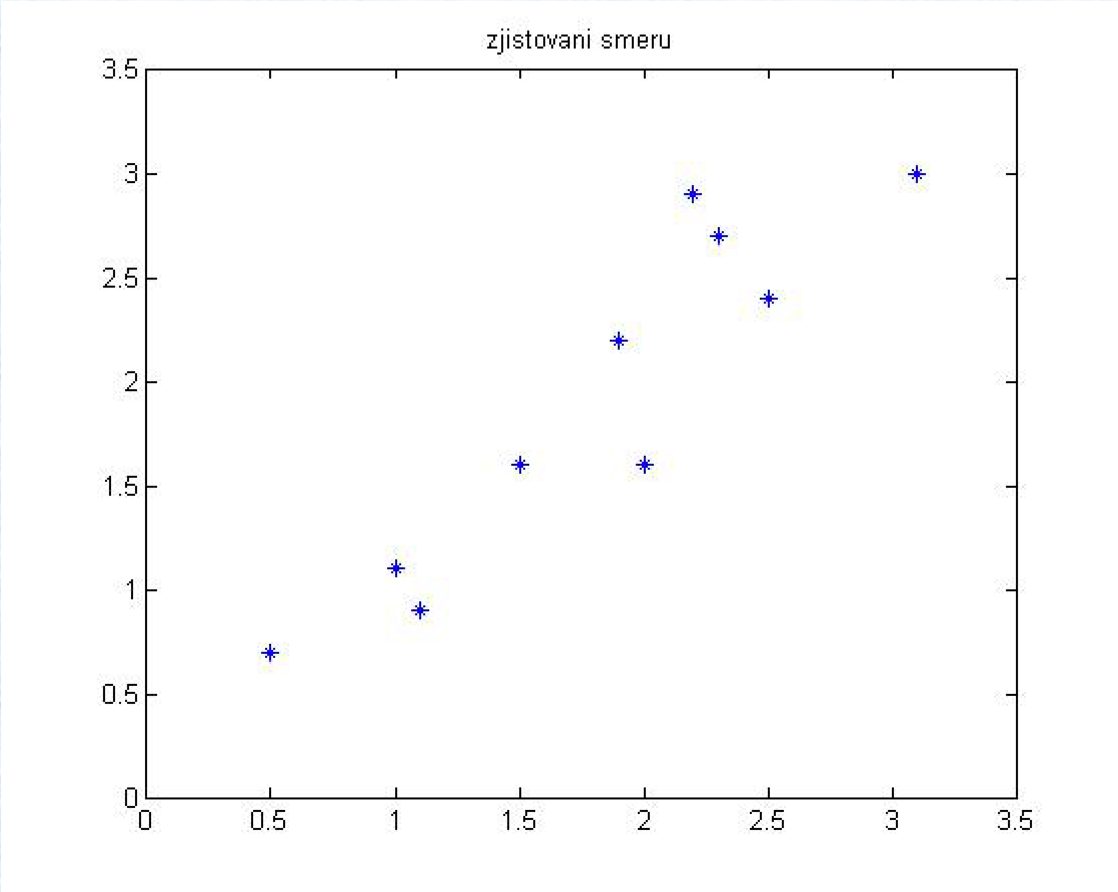
PCA

- Princip $Y = XP$
  - *X* je centrovaná matice (n,d)
  - Y je matice výstupních dat (n,d)
  - P je matice (n,n) vlastních vektorů kovarianční matice C_X definované vztahem

    $$C_X = P\Lambda P^T$$

    - Λ je diagonální matice vlastních čísel
    - P je ortonormální, tj. platí $P^T P = E_d$
      - E_d je jednotková matice (d,d)

PCA – orientace - příklad



PCA – orientace – to do

PCA – konverze – to do

Shrnutí

- Pro PCA potřebujeme
  - Průměr
  - Rozptyl
  - Kovarianční matici
  - Vlastní čísla a vektory
- Princip PCA

$$Y = XP$$

Literatura

- <http://www.mathworks.com>
- Smith L. I., A tutorial on Principal Component Analysis
- Mudrová M., Procházka A., Kolínová M., Signal Feature Using Component Analysis Methods
- Mudrová M., Procházka A., Principal Component Analysis in Image Processing