Exercise 14

Samstag, 4. Juni 2022 19:37

a) PCA:

1. Centralization of datapoints
   ↳ calculate vector $\vec{\mu}$ of expected value of each feature
2. Calculate Covariance Matrix with entries
   $C_{ij} = E((x_i - E(x_i)) \cdot (x_j - E(x_j)))$
3. Calculate Eigenvalues and Eigenvectors of Cov. Matrix
   · sort Eigenvalues in descending order
4. Chose k biggest Eigenvalues
5. Transform dataset via $X' = XW$, W: Matrix of not rejected Eigenvectors
   $W = (v_1, v_2, \dots v_k)$

b) $\vec{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$ $\quad \mu_1 = \frac{1}{6}(1+3+1+2+3+2) = \underline{\underline{2}}$

$\mu_2 = \frac{1}{6}(1+0+3+0+1+1) = \underline{\underline{1}}$

$\Rightarrow \vec{\mu} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$

· $Cov(x,y) = \frac{1}{N-1} \sum_{i=1}^{N} (x_i - \bar{x})(y_i - \bar{y})$

$Cov(x,x) = \frac{1}{5}\left((1-2)^2 + (3-2)^2 + (1-2)^2 + (2-2)^2 + (3-2)^2 + (2-2)^2\right)$

$= \frac{1}{5}(1 + 1 + 1 + 1) = \frac{4}{5}$

$Cov(y,y) = \frac{1}{5}(3(1-1)^2 + 2 \cdot (0-1)^2 + (3-1)^2) = \underline{\underline{1}}$

$Cov(x,y) = \frac{1}{5}((3-2)(0-1) + (1-2)(3-1))$

$= -\frac{3}{5}$

$\Rightarrow Cov = \begin{pmatrix} \frac{4}{5} & -\frac{3}{5} \\ -\frac{3}{5} & 1 \end{pmatrix}$

Eigenvalues:

$\det \begin{vmatrix} \frac{4}{5} - \lambda & -\frac{3}{5} \\ -\frac{3}{5} & 1-\lambda \end{vmatrix} = 0$

$\Rightarrow (\frac{4}{5} - \lambda)(1-\lambda) - \frac{9}{25} = 0$

$\Leftrightarrow \lambda^2 - \frac{9}{5}\lambda + \frac{11}{25} = 0$

$$\Leftrightarrow \lambda^2 - \frac{9}{5}\lambda + \frac{11}{25} = 0$$

$$\Rightarrow \lambda_{1,2} = \frac{9}{10} \pm \sqrt{\left(\frac{9}{10}\right)^2 - \frac{11}{25}}$$

$$= \frac{9}{10} \pm \sqrt{\frac{81}{100} - \frac{44}{100}} = \frac{9}{10} \pm \sqrt{\frac{37}{100}} = \frac{9 \pm \sqrt{37}}{10}$$

$$\Rightarrow \underline{\lambda_1 = \frac{9-\sqrt{37}}{10}}, \quad \underline{\lambda_2 = \frac{9+\sqrt{37}}{10}}$$

Eigenvectors:

$$v_1 = \begin{pmatrix} (\sqrt{37}+1)/6 \\ 1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} (-\sqrt{37}+1)/6 \\ 1 \end{pmatrix}$$

Transformation:

· because $\lambda_1 < \lambda_2$, discard λ_1

$\Rightarrow$ Transformation $x' = x \cdot v_2$

$\Rightarrow$ New observables

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} \cdot v_2 = \frac{-\sqrt{37}+1}{6} + 1 \approx 0{,}15$$

$$\begin{pmatrix} 3 \\ 0 \end{pmatrix} \cdot v_2 = \frac{-\sqrt{37}+1}{2} \approx -2{,}54$$

$$\begin{pmatrix} 1 \\ 3 \end{pmatrix} \cdot v_2 = \frac{-\sqrt{37}+1}{6} + 3 \approx 2{,}15$$

$$\begin{pmatrix} 2 \\ 0 \end{pmatrix} \cdot v_2 = \frac{-\sqrt{37}+1}{3} \approx -1{,}69$$

$$\begin{pmatrix} 3 \\ 1 \end{pmatrix} v_2 = \frac{-\sqrt{37}+1}{2} + 1 \approx -1{,}54$$

$$\begin{pmatrix} 2 \\ 1 \end{pmatrix} v_2 = \frac{-\sqrt{37}+1}{3} + 1 \approx -0{,}69$$