Principal component analysis

given a dataset $X \in \mathbb{R}^{3 \times 2}$, project the data using PCA

$$X = \begin{array}{cc} \text{feature A} & \text{feature B} \\ \end{array}$$

$$X = \begin{bmatrix} 2 & 4 \\ 1 & 3 \\ 0 & 0 \end{bmatrix}$$

#1 center X

$$\mu_A = \frac{2+1+0}{3} = 1 \qquad \mu_B = \frac{4+3+0}{3} = \frac{7}{3}$$

$$X_{center} = X - \mu = \begin{bmatrix} 1 & 5/3 \\ 0 & 2/3 \\ -1 & -7/3 \end{bmatrix}$$

#2 calculate covariance matrix $\Sigma = \frac{1}{n-1} X_{center}^T X_{center}$

$$\Sigma = \frac{1}{3-1} \cdot \begin{bmatrix} 1 & 5/3 \\ 0 & 2/3 \\ -1 & -7/3 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & -1 \\ 5/3 & 2/3 & -7/3 \end{bmatrix}$$

$$\Sigma = \frac{1}{2} \begin{bmatrix} 2 & 4 \\ 4 & 26/3 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 2 & 13/3 \end{bmatrix}$$

#3 compute eigenvalues & eigenvectors

(eigenvalue decomposition)

$\Sigma = \lambda \cdot v \rightarrow$ eigenvectors

λ ↙ eigenvalues

$\det(\Sigma - \lambda I) = 0$

$$\Sigma - \lambda I = \begin{bmatrix} 1-\lambda & 2 \\ 2 & 13/3 - \lambda \end{bmatrix}$$

$$\det(\Sigma - \lambda I) = (1-\lambda)\left(\frac{13}{3} - \lambda\right) - (2)(2)$$

$$= \frac{13}{3} + \lambda^2 - \frac{16}{3}\lambda - 4$$

$$= \lambda^2 - \frac{16}{3}\lambda + \frac{1}{3}$$

$$\Rightarrow \lambda_1 = \frac{-\sqrt{61}+8}{3} \quad \text{and} \quad \lambda_2 = \frac{\sqrt{61}+8}{3}$$

$= 0.06325$ $\qquad$ $= 5.27008$

(too long for him to ask us for v_1 and v_2 so im skipping to the end)

$$v_1 = \begin{bmatrix} \frac{-\sqrt{61}-5}{6} \\ 1 \end{bmatrix} \qquad v_2 = \begin{bmatrix} \frac{\sqrt{61}-5}{6} \\ 1 \end{bmatrix}$$

$$\Rightarrow \Sigma v_1 = \lambda_1 v_1 \quad \text{and} \quad \Sigma v_2 = \lambda_2 v_2$$

#4 choose k largest eigenvalues (lets just do k=1)

$$\lambda_2 = \frac{\sqrt{61}+8}{3} \quad \text{and} \quad v_2^T = \begin{bmatrix} \frac{\sqrt{61}-5}{6} & 1 \end{bmatrix}$$

#5 project original data to k-dimensional space

$$Y = \underset{\text{center}}{\overset{\overset{3\times 2}{\uparrow}}{X}} \cdot \overset{\overset{2\times 1}{\nearrow}}{W_k} \quad \text{s.t.} \quad Y \in \mathbb{R}^{\overbrace{n\times k}^{3\times 1}}$$

$$W_k \in \mathbb{R}^{d\times k} \qquad W_k = \begin{bmatrix} \overset{\overset{d\times 1}{\uparrow}}{v_1} & \overset{\overset{d\times 1}{\uparrow}}{v_2} & \cdots \end{bmatrix}$$

$$W_k = \begin{bmatrix} \frac{\sqrt{61}-5}{6} \\ 1 \end{bmatrix}$$

$$Y = \begin{bmatrix} 1 & 5/3 \\ 0 & 2/3 \\ -1 & -7/3 \end{bmatrix} \cdot \begin{bmatrix} \frac{\sqrt{61}-5}{6} \\ 1 \end{bmatrix}$$

$$Y = \begin{bmatrix} 2.13504 \\ 2/3 \\ -2.80171 \end{bmatrix}$$

→ new data w/ one feature

back to #3 but with singular value decomposition

singular values

$$X_{center} = U \Sigma V^T$$

left singular vectors

right singular vectors (PC's)

$X \in \mathbb{R}^{n \times d}$ $\quad U \in \mathbb{R}^{n \times n}$ $\quad \Sigma \in \mathbb{R}^{n \times d}$ $\quad V^T \in \mathbb{R}^{d \times d}$

from the same eigenvalue decomposition process

$$\Sigma = \begin{bmatrix} \sqrt{\lambda_1} & 0 \\ 0 & \sqrt{\lambda_2} \end{bmatrix} \qquad V = \begin{bmatrix} v_1 & v_2 \end{bmatrix}$$

PC's so essentially its the same thing

$$U = X_{center}^{-1} \Sigma V^T$$