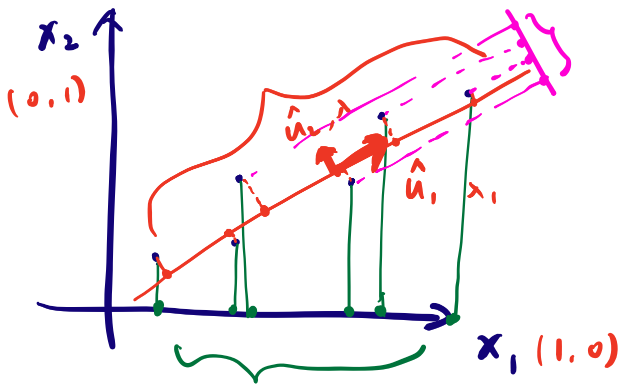
Nov 7, 2024 PCA




Goal: maximize the variance in low dimension space

$\hat{u}$ direction

$\underline{z_i = x_i^T \hat{u}}$

PCA: first normalize $X = X - \mu$,

mean$(X) = 0$

mean$(z) = 0$

variance $J = \frac{1}{N}\sum_{i=1}^{N}(z_i - 0)^2$

$$\max \underline{J} = \frac{1}{N}\sum_{i=1}^{N} \hat{u}^T x_i x_i^T \hat{u}$$

$$= \frac{1}{N}\hat{u}^T \underline{\sum_{i=1}^{N} x_i x_i^T} \cdot \hat{u}$$

Sample variance $S = \frac{1}{N}\sum_{i=1}^{N} x_i x_i^T$

$$= \hat{u}^T S \hat{u}$$

$\max\ \hat{u}^T S \hat{u}$

s.t. $\hat{u}^T \hat{u} = 1$

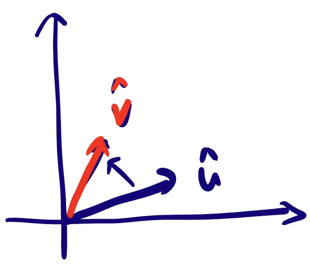
$$\max\ L = \underline{\hat{u}^T S \hat{u}} - \lambda(\hat{u}^T \hat{u} - 1)$$

Set $\frac{\partial L}{\partial \hat{u}} = 0$ Set $\frac{\partial L}{\partial \lambda} = 0$

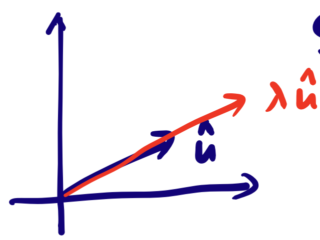
$$2S\hat{u} - 2\lambda\hat{u} = 0 \qquad \hat{u}^T\hat{u} = 1$$

$\boxed{S\hat{u} = \lambda\hat{u}}$ eigen decomposition of $S$

$S_{D \times D}$




$S\hat{u} = \hat{v}$




$S\hat{u} = \lambda\hat{u}$

eigen decomposition

$S\hat{u} = \lambda\hat{u}$

$\lambda$ : eigenvalues

$\hat{u}$ : eigenvectors

First PC : longest eigenvalue $\lambda_1$ , $\hat{u}_1$

Second PC : Second longest eigenvalue $\lambda_2$ $\hat{u}_2$

$$J = \frac{1}{N}\sum_{i=1}^{N} \delta_i^2$$

$$= \hat{u}_1^T \underbrace{S\hat{u}_1}_{\lambda\hat{u}_1} \quad \text{first one PC}$$

$$= \lambda_1$$

$$J = \sum_{i=1}^{L} \lambda_i \qquad \text{first } L \text{ PC} \quad L \leq D$$

$W$ eigenvectors of $S = \frac{1}{N}\sum_{i=1}^{N} x_i x_i^T$

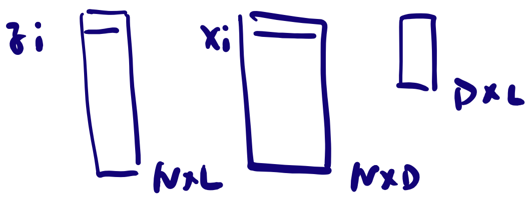
$$W = \begin{bmatrix} \hat{u}_1 & \hat{u}_2 & \hat{u}_L & \cdots & \hat{u}_D \end{bmatrix} \quad (\hat{u}_1: D\times 1)$$

$$\underline{\lambda_1 > \lambda_2\ \lambda_L} \cdots \lambda_D$$

Choose top $L$ eigenvectors of $S$

$W_{D\times L}$

$$\underset{N\times L}{Z} = \underset{N\times D}{X} \cdot W_{D\times L}$$




$$\underset{N\times D}{\hat{X}} = Z_{N\times L} \cdot W^T_{L\times D}$$

$$\hat{X} \sim X$$