Lecture 11

* PCA $X \in \mathbb{R}^{d \times n}$ Project so that we get a lesser dim. data. $p << d$

Centralized $X_c \leftarrow X - \mu$

let u_1 be vector.

Proj. of X_c on u_1

$Y = u_1^T X_c$

$\mu_y = 0$

$Var(y) = u_1^T S_{x_c} u_1$

for PCA: max. $u_1^T S_{x_c} u_1$

s.t. $\cancel{u_1^T u_1 = 1}$

max. $\underline{u_1^T S_{x_c} u_1}$

$\cancel{S_{x_c} u_1} =$

$S_{x_c} u_1 = 0$

$u_1 = 0$

max. $f - \lambda g$

max. $u_1^T S_{x_c} u_1 - \lambda (u_1^T u_1 - 1)$

$\lambda \rightarrow$ Lagrange mult.

$\cancel{\nabla_{u_1} \rightarrow 2S}$

$$\nabla_{u_1} \to 2 S_{x_c} u_1 - 2\lambda u_1 = 0$$

$$S_{x_c} u_1 = \lambda u_1$$

$u_1 \to$ eigenvector of S_{x_c}

$\lambda \to$ eigvalue.

$\cancel{u_1^T S}$

$$\max.\ u_1^T \underbrace{S_{x_c} u_1}_{\lambda u_1} = \lambda u_1^T u_1$$

$$= \lambda$$

$u_1 \to$ eigvector of S_{x_c}

corresponding to max. eigenvalue.

PCA matrix

$$U = [u_1 \ u_2 \ \cdots\cdots \ u_d]$$

$$u_i \in R^{d \times 1}$$

$$U \in R^{d \times d}$$

Encoding X_c

$$Y = U^T X_c \qquad X_c \in R^{d \times n}$$

$$Y \in R^{d \times n}$$

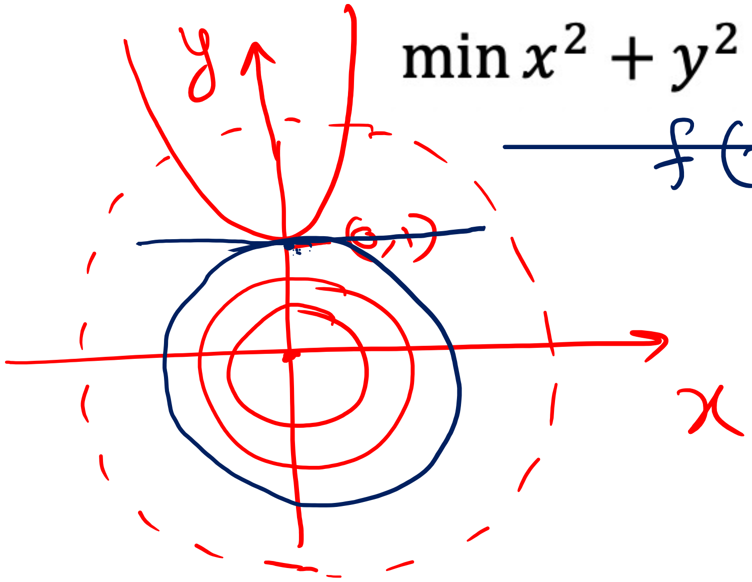
$$U U^T X_c = X_c$$

~~Take 1st~~

Take 1st p columns of U

$$U_p = [u_1, u_2 \cdots u_p]$$

$$\min x^2 + y^2 \ s.t.\ y = x^2 + 1$$



$$f(x,y) = x^2 + y^2$$

$$g(x,y) = y - x^2 - 1$$

at $(0,1)$

$$\nabla f = \begin{bmatrix} 0 \\ 2 \end{bmatrix} \qquad \nabla g = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$\nabla f(x,y) = \begin{bmatrix} 2x \\ 2y \end{bmatrix}$$

$$\nabla f = 2 \nabla g$$

$$\nabla f = \lambda \nabla g$$

$$\nabla g(x,y) = \begin{bmatrix} -2x \\ 1 \end{bmatrix}$$

$$\nabla (f - \lambda g) = 0$$

$U_p \in R^{d \times p}$

$Y = U_p^T X_c \in R^{p \times n}$

Var. across $u_1 \rightarrow \lambda_1$ (λ_{max})

$u_2 \rightarrow \lambda_2$ (2^{nd} max.)

$u_d \rightarrow \lambda_d$ (least)

Total var $\rightarrow \sum_{i=1}^{d} \lambda_i = \lambda_T$

$$\frac{\sum_{j=1}^{p} \lambda_j}{\lambda_T} = 90\% \quad = 95\%$$

If Sx is full rank, how many eigenvectors?
Which eigenvector will be u1?

- What if we minimize? What will be dir of u1?
- What if don’t have constraint?
- Are labels included in the formulation? Will it Impact classification?

$$X = \begin{bmatrix} 1 & 3 \\ 4 & 2 \end{bmatrix} \quad \mu = \begin{bmatrix} 2 \\ 3 \end{bmatrix} \quad X - \mu = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} = X_c$$

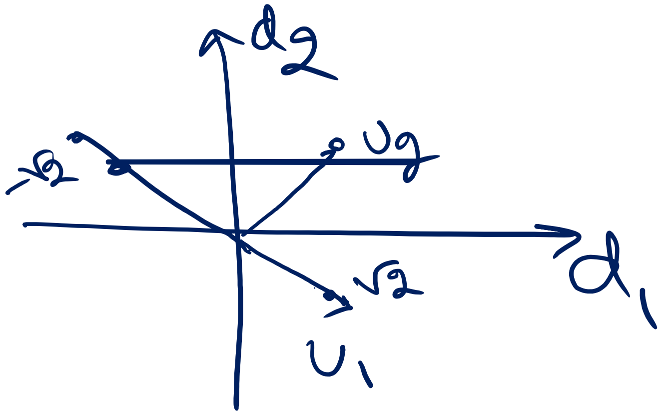
$$S = \frac{1}{2} \sum_{i=1}^{2} x_i x_i^T = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$

eigenvalues $\rightarrow$ $4, 0$

eigenvectors $\rightarrow$ $\left[\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right]'$ $\left[\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right]'$

$$U = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

$$Y = U^T X_c = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}^T \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} -\sqrt{2} & \sqrt{2} \\ 0 & 0 \end{bmatrix}$$

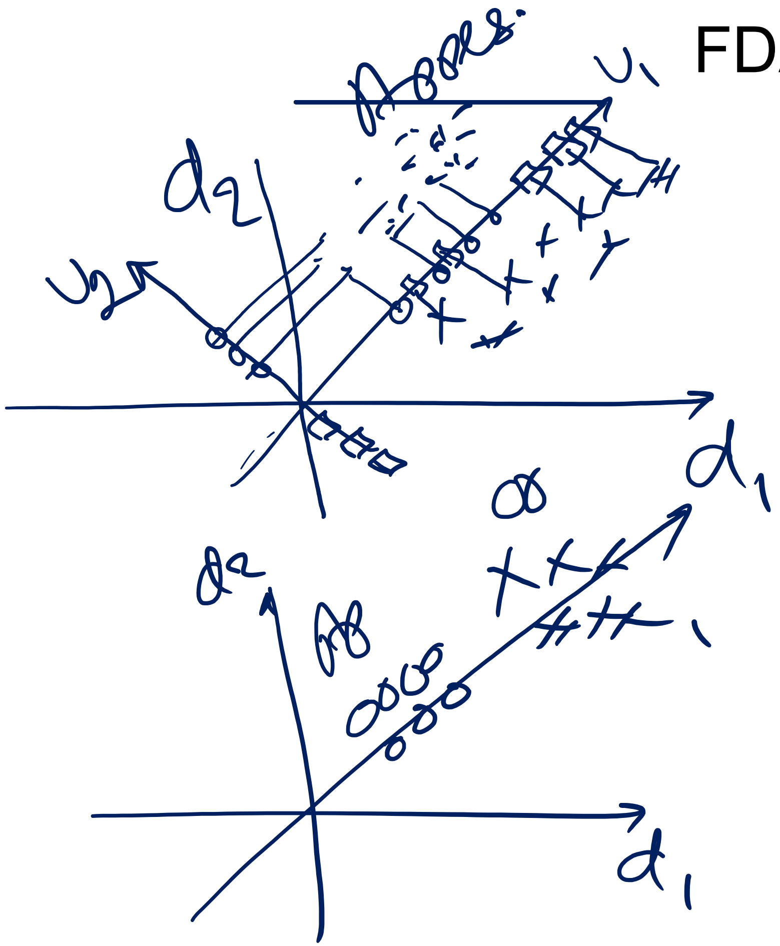


d_2
v_2
$-v_2$
v_2
d_1
v_1

$X_c \leftarrow \cancel{U \Sigma}\ U \Sigma V^T$ Eigendecomposition.

$\text{SVD}(X_c) = U$

FDA



$X^1 \in$ ~~goodest of Dept~~ Appled.

$X^2 \in$ Oranges.

~~test point~~ x_0

test point x_0 → App / orange.

$$X = [X^1 \; X^2] \in \mathbb{R}^{d \times 2n}$$