Lecture 13 PCA

$$\log p(x)$$

$$\log p(x) \geq ELBO$$

I like

$$\begin{array}{c c}
\begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \\ \\ \\ \end{bmatrix} & \begin{array}{l} a \\ b \\ c \\ \vdots \\ \text{alike} \\ \vdots \end{array}
\end{array}$$

dictionary

$\vec{x} = (x_1, x_2 \cdots x_m)$

regression SVM

EM

1024 dim

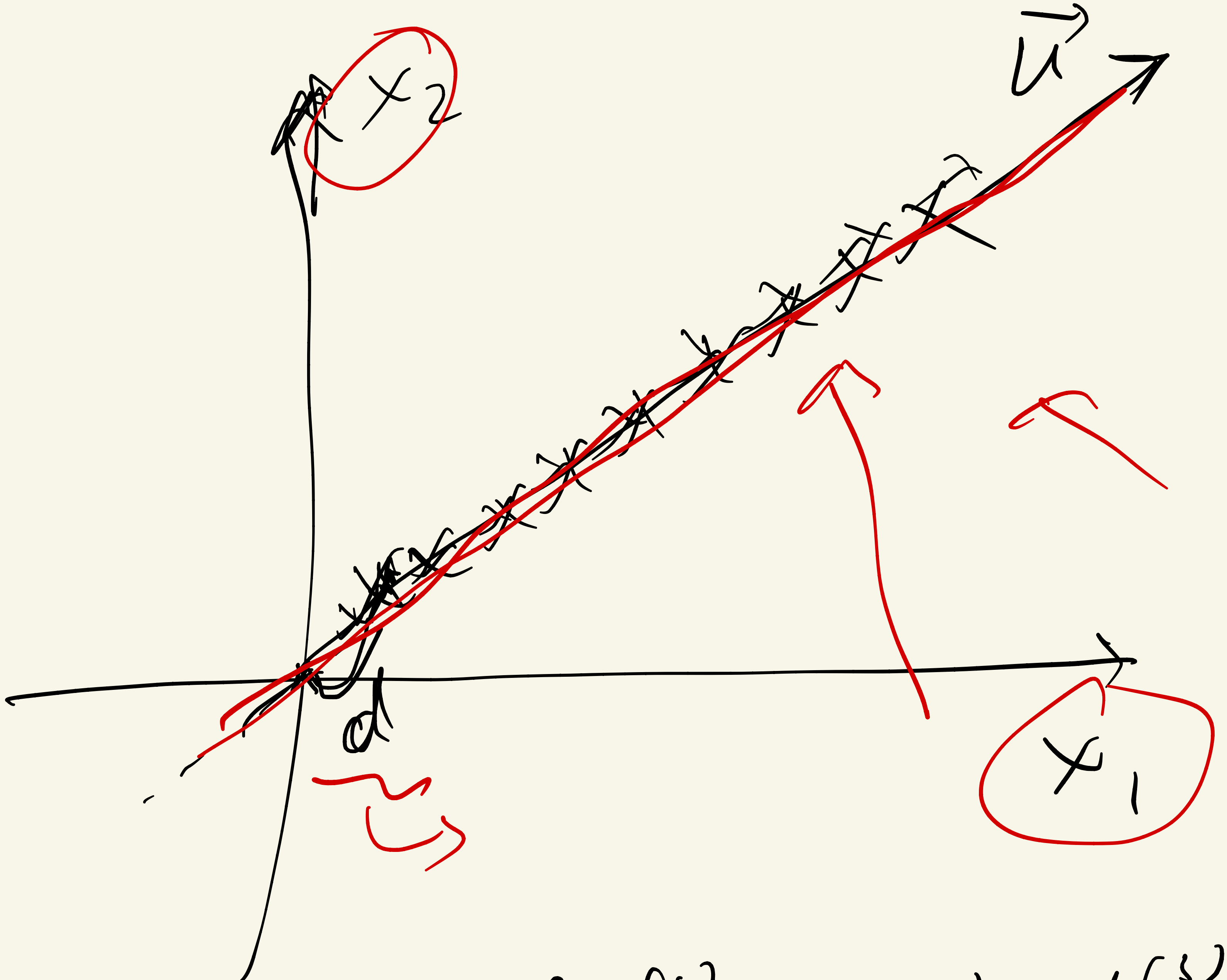
text, date, sender email.

email agent

GLM

$$\underbrace{\theta^T x}_{\theta}$$

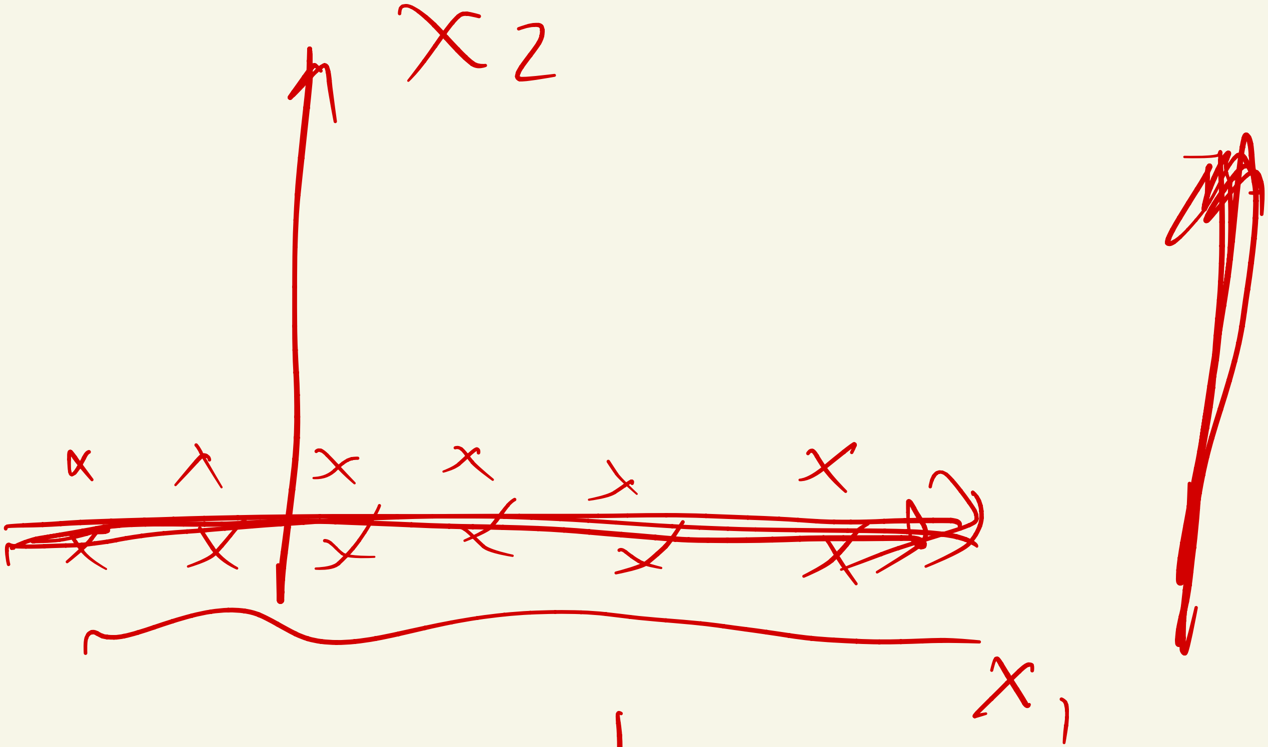
$$\underbrace{\theta_1}_{5} x_1 + \underbrace{\theta_2}_{0.01} x_2 + \underbrace{\theta_3}_{3} x_3$$



$\vec{u}$ $\quad$ $d^{(1)}$ $\quad$ $d^{(2)}$ $\quad$ $d^{(3)}$

$\underbrace{[x_1^{(1)}, x_2^{(1)}]}_{n \times 2}$ $\quad$ $[x_1^{(2)}, x_2^{(2)}]$

$2 \times n$ $\quad$ $n \times 2 \rightarrow \vec{u}$

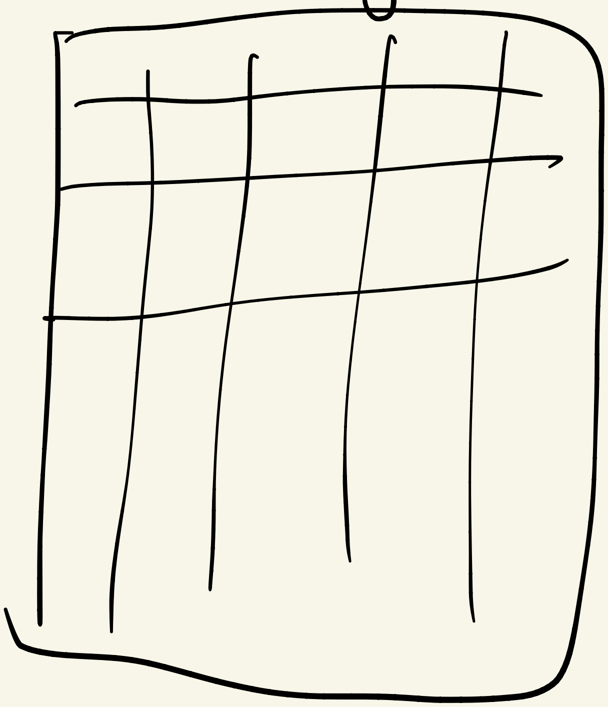


$$\theta_1 x_1 + \underline{\theta_2} x_2$$

$$\underbrace{\theta_2 = 0.01}$$

human faces

image



Pixels

→ 256 x 256

male / femal

ethinity Asia white
black

GMM

$P(z=1)$

$P(z=2)$

$P(z)$

$P(z) = 1/2$

$P(z=3)$

$2+6$

2   3

z

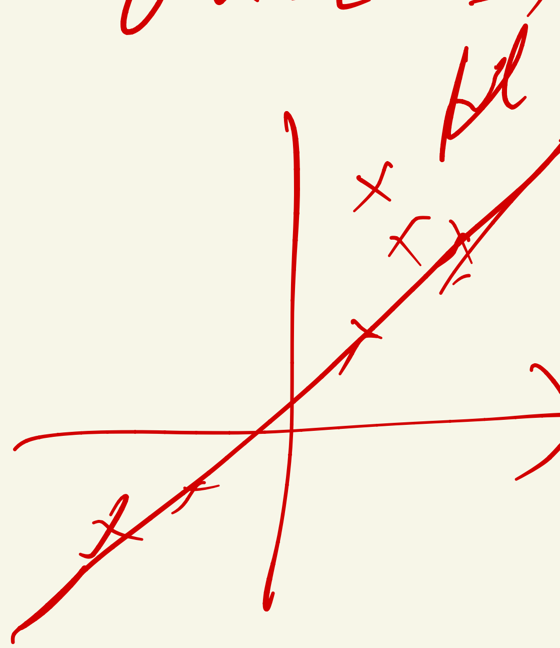
(μ_1, Σ_1) (μ_2, Σ_2)

(μ_3, Σ_3) $6 \times d$

x

$x \in R^d$

$\boxed{\text{d dimension}}$



$x \sim N(\mu, \sigma^2)$

$x = \mu + \sigma \xi$

$\xi \in N(0, 1)$

noise

unsupervised

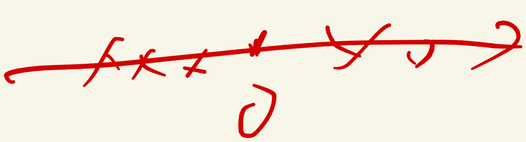
$\vec{u}$

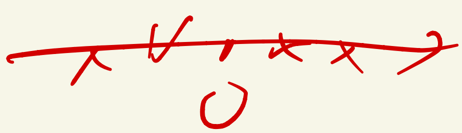
[price, size, years]

price	size	years
1000000	100	1-10

0.5

1 -0.5





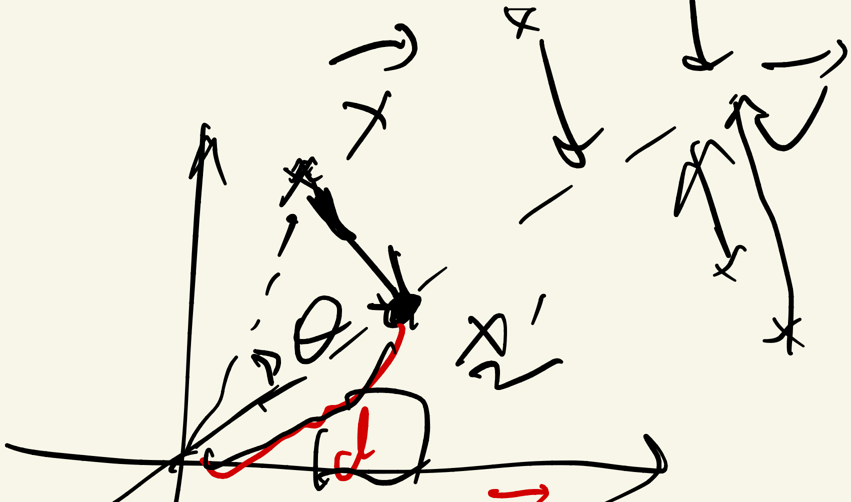
$x \in R^d$

map $x \to$ 1-dim

$\vec{v}$

after projection, variance $\to$ large

Variance ?



$\|\vec{v}\| = 1$

$x' = \langle \vec{x}, \vec{v} \rangle \vec{v}$

$d = \langle \vec{x}, \vec{v} \rangle$

$\|\vec{x'}\| = \|\vec{x}\| \cos\theta$

$$\cos\theta = \frac{\langle \vec{x}, \vec{v} \rangle}{\|\vec{x}\| \, \|\vec{v}\|}$$

$$\mathrm{Var} = E[(x - \underbrace{\mathrm{mean}(x)}_{\Downarrow\ 0})^2]$$

$$\underbrace{E[d^2]}$$

$$\underbrace{X \in \mathbb{R}^{n \times d}} \quad \begin{bmatrix} x_1^T \\ x_2^T \\ x_3^T \\ \vdots \end{bmatrix}$$

$$\underbrace{E[\langle \vec{x}, \vec{v} \rangle^2]}$$

$$\|$$

$$\boxed{\frac{1}{n}\sum_{i=1}^{n} (v^T x_i)^2 = \underline{v^T X X^T V}}$$

$$\max_{v} \; v^T X X^T v \quad \text{s.t. } v^T v = 1$$

$$\|v\| = 1$$

Lagrange multipolier

$$v^T X X^T v - \lambda (v^T v - 1)$$

$$\boxed{(XX^T - \lambda I)\, v = 0} \qquad v^T v - 1 = 0$$

$$v \qquad \|v\| = 1$$

non-zero $\quad v \neq 0$

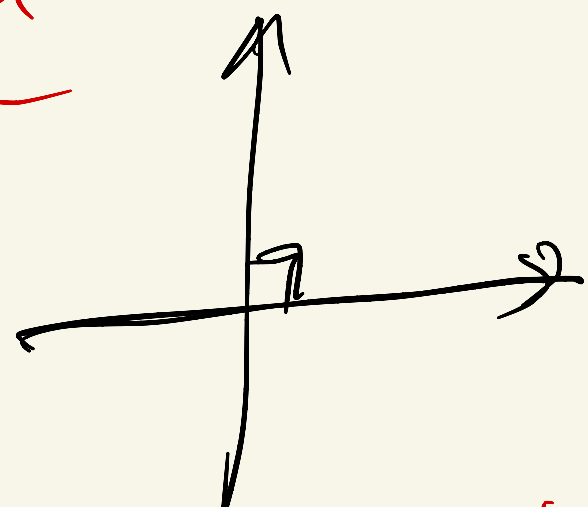
$$\det |XX^T - \lambda I| = 0$$

$x \in R^d \longrightarrow 1$ dim

$\boxed{x \in R^d \longrightarrow R^k}$

$k < d$

$\boxed{XX^T}$



top k orthogonal directions that maxim variance

Variance $\qquad XX^T v = \lambda v$

$= v^T X X^T v$

$= \lambda \underbrace{v^T v} = \underline{\lambda}$

$X \in R^D$

$\underline{X} \in R^D \quad \underline{V_1, V_2 \cdots V_K}$

$[\langle X, V_1 \rangle, \langle X, V_2 \rangle \cdots \langle X, V_K \rangle]$

$X' \in R^K$

$$\min_{v} \frac{1}{n}\sum_{i=1}^{n} \|x_i - v^T x_i\|^2$$

$$R^d \longrightarrow R^k$$

how to select $\underline{k}$

$\underline{1000,\ \lambda_i}$

$k=5$

$\lambda_5\ \lambda_6$

15

$$E(x-\mu)^2$$

direction of curve