CS - 541 AI

09/27/2024

Last week: Random Projection

(dim reduction)

$f: \mathbb{R}^d \to \mathbb{R}^k \quad k < d$

$x \to M \cdot x$

$M: k \times d$

$M$: data independent

$M_{ij} \sim N(0, 1/k)$

Given $S = \{x_1, \ldots x_n\}$

$k \approx \Theta\left(\frac{1}{\epsilon^2} \cdot \log n\right) = 1$

$$1 - \epsilon \le \frac{\|M \cdot x\|}{\|x\|} \le 1 + \epsilon$$

$\epsilon \to 0$, projection preserves norm

$k = 1 \Rightarrow \epsilon = \sqrt{\log n}$

PCA: (principal component analysis)

$$X \rightarrow M \cdot X$$

M: data dependent

$x_1, \ldots x_n$

$$X = \begin{pmatrix} \uparrow & & \uparrow \\ x_1 & \cdots & x_n \\ \downarrow & & \downarrow \end{pmatrix}$$

$$M = f(X)$$

① SVD to solve PCA

What are applications of PCA?

→ compress image / text

- visualize data

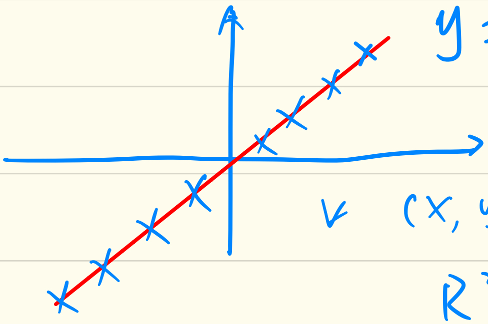
-

When we cannot apply PCA?

- when cannot compress.

orthogonal features

Fails



$y = x$

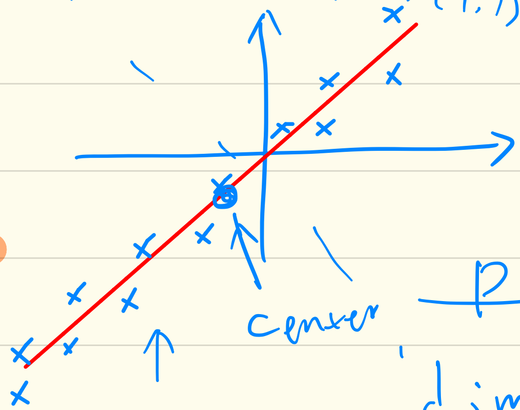
$(x, y) \to x$

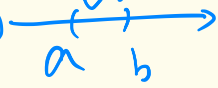
$R^2 \to R^1$

find subspace such that projection error is small.

$(1,1)$ $(2,2)$

$(3,3) = 3\cdot(1,1)$ $\quad 2\cdot(1,1)$




$\frac{1}{4}\cdot(b-a)^2$

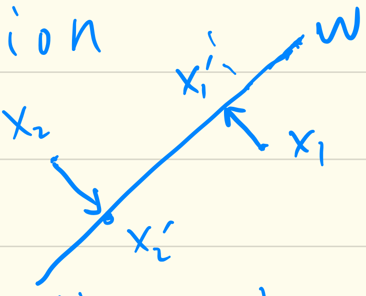
~~PCA~~ ?

dim red ?

or perform orthogonal projection

$X_1 \longrightarrow x_1' = \alpha_1 \cdot \boxed{v}$

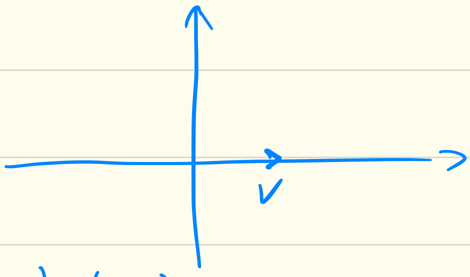
$X_2 \longrightarrow x_2' = \alpha_2 \cdot v$.




① find a direction (subspace) with largest ~~correlation~~ variance

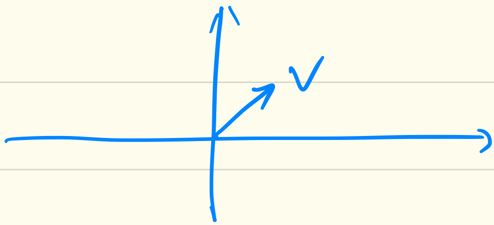
② find a low-dim subspace that best fits data

$v = (1, 0)$



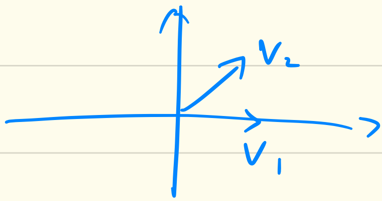

$\mathrm{sub}(v) = x - \text{axis}$

$v = (1, 1)$




$\mathrm{sub}(v) = \{(x, y) : y = x\}$

$\{v_1 = (1, 0),\ v_2 = (1, 1)\}$




$\mathrm{sub}(v_1, v_2) = R^2$

can be full rank

$$X = (X_1, \dots X_n)$$

$d \times n$

$$\approx (\alpha_1 \cdot V, \ \alpha_2 \cdot V, \ \dots, \alpha_n \cdot V)$$

$$= \underset{d \times 1}{V} \cdot \underset{1 \times n}{(\alpha_1, \alpha_2, \dots \alpha_n)}$$

$V$: $d \times 1$

$\alpha_i \in \mathbb{R}$.

$$= \underbrace{V}_{d \times 1} \cdot \underbrace{\alpha}_{1 \times n} \rightarrow \text{rank-1}$$

$$\approx \ d \times k \quad \underline{\underline{k \times n}} \rightarrow \text{rank} - k$$

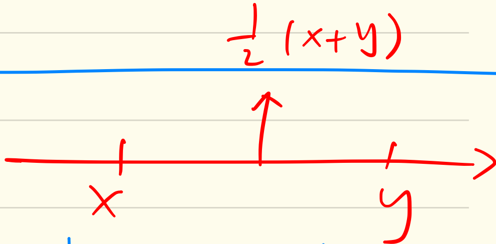
PCA: find low-rank approx to fit the data.

PCA (almost)

$$\begin{cases} \min\limits_{M \in \mathbb{R}^{d \times n}} \|M - X\|_F \\ \text{s.t. } \boxed{\operatorname{rank}(M) \le k} \end{cases}$$

$\{M : \operatorname{rank}(M) \le k\} \to$ non-convex set

$$\|M\|_F \triangleq \sqrt{\sum_{ij} M_{ij}^2} \quad (\approx \|v\|_2)$$




Solution:

History: 1930's - 1960's

convex programming

Solvable in poly-time

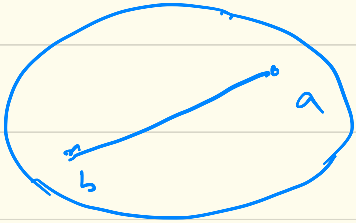
$\to$ poly(d, n)

k = 1

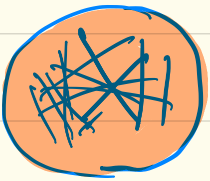
$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

rank = 2

# convex set

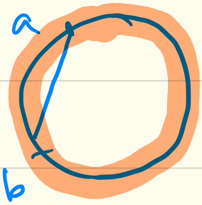



unit ball ✓

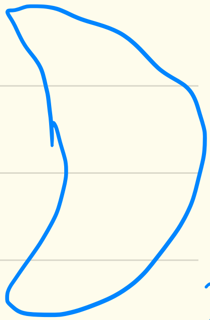


$x^2+y^2 \le 1$

circle ✗



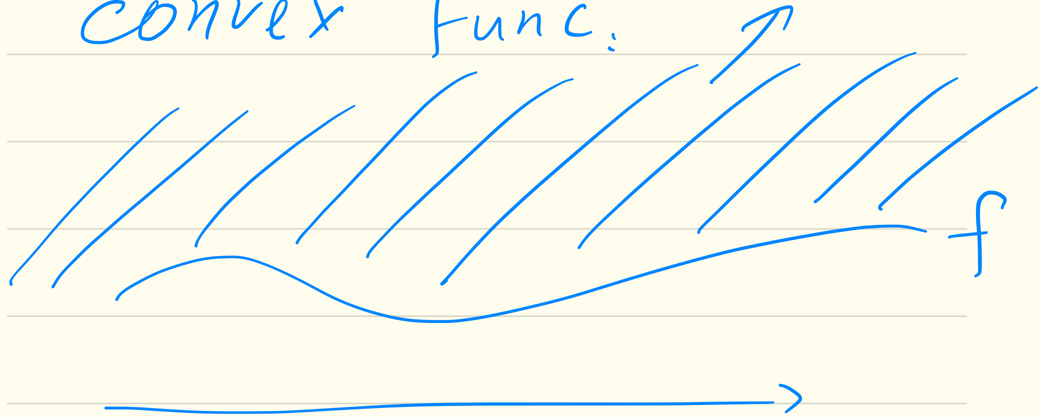

$(x,y):\ x^2+y^2=1$



✗

epi-graph

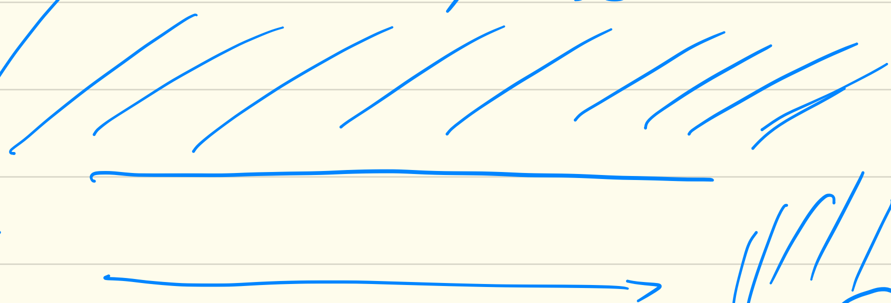
convex func.:



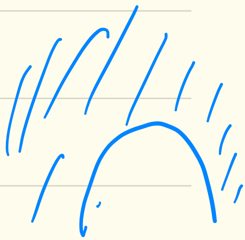
f is convex func.

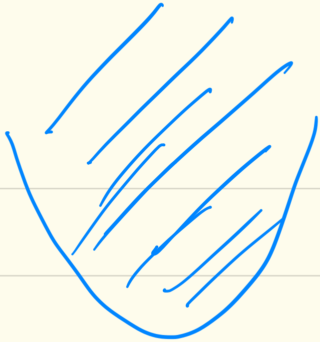
iff. epi-graph(f) is

convex set.

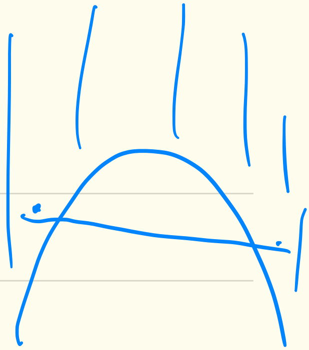


$= \{(x,y) : y \geq f(x)\}$





convex



not

$$\left\{\begin{array}{ll}\min\limits_{w} & f(w) \\ \text{s.t.} & g(w) \le 0\end{array}\right.$$

(standard expression)

convex program $\rightarrow$ solvable in poly-time.

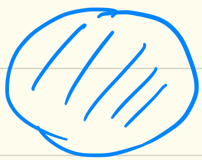
$\Longleftrightarrow$ $f$ is convex

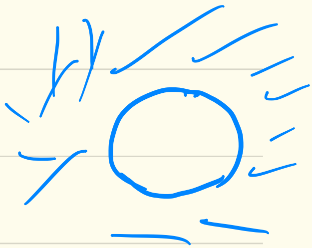
$\{w : g(w) \le 0\}$ is convex

$$\|w\|_2 \le 1$$

$$\|w\| > 1$$

$$g(w) = 1 - \|w\|$$





1930's

↓

1960's → general convex solvers

↓

$\epsilon \to 0$

$\epsilon q_0$ :

1990's → gradient descent

2000

↓

now → Stochastic GD (AI/ML)

non-convex

PCA ⊃ non-convex poly-time

$e^{100} q_0$

SVD → singular value decomposition.

Input $X : d \times n$

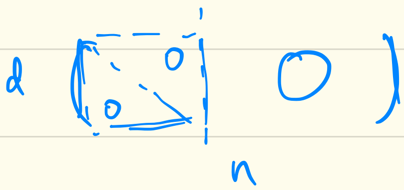
Ouput $U : d \times d$

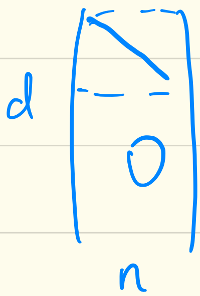
$S : d \times n$

$V : n \times n$

① $X = U S V^T$

② $S$ is diagonal




$d < n$

③ $U^T U = I_{d \times d}$

$V^T V = I_{n \times n}$

Claim: $\underbrace{U \text{ is subspace of data}}$

$$X = U \underbrace{\boxed{S\, V^T}}_{d\times n \quad n\times n} \to C$$

$$= U \cdot C \to d \times n$$

$$(\boxed{x_1} \cdots x_n) = \left( \underset{\downarrow}{\overset{\uparrow}{u_1}} \cdots \underset{\downarrow}{\overset{\uparrow}{u_d}} \right) \left( \begin{matrix} c_{11} \\ c_{21} \\ c_{31} \\ \vdots \\ c_{d1} \end{matrix} \right)$$

$R$

$$x_1 = u_1 \cdot c_{11} + u_2 \cdot c_{21} \cdots + u_d \cdot c_{d1}$$

$$= \sum_{j=1}^{d} c_{j1} \cdot \underline{u_j}$$

$$x_2 = \sum_{j=1}^{d} c_{j2}\, u_j$$

$\vdots$

$$S = \begin{pmatrix} s_1 & & 0 & \\ & \ddots & & 0 \\ 0 & & s_d & \end{pmatrix}$$

$$s_1 \geq s_2 \geq \cdots \geq s_d \geq 0$$

$$X = \underline{(u_1 \cdots u_d)} \underline{\left(\begin{array}{ccc|c} s_1 & & & \\ & \ddots & & 0 \\ & & s_d & \end{array}\right)} \cdot V^T$$

$$= \underline{(s_1 u_1, s_2 u_2, \cdots s_d u_d, 0, 0, \cdots 0)} \underline{\begin{pmatrix} \leftarrow v_1^T \rightarrow \\ \vdots \\ \leftarrow v_n^T \rightarrow \end{pmatrix}}$$

$$= \sum_{i=1}^{d} \underbrace{s_i}_{\mathbb{R}} \boxed{\underbrace{u_i}_{d\times 1} \cdot \underbrace{v_i^T}_{1\times n}} \rightarrow \text{rank-1 matrix}$$

$d\times n$

$d = 2$

$$X = s_1 \cdot \boxed{u_1 v_1^T} + s_2 u_2 v_2^T$$

$$= \boxed{10^6} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \cancel{10^{-6}} \cdot \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

$\text{rank}(X) = 2$

find rank-1 approx best fit $X$.

$$\begin{pmatrix} 10^6 & 0 \\ 0 & 10^{-6} \end{pmatrix}$$

$$\downarrow$$

$$\begin{pmatrix} 10^6 & 0 \\ 0 & 0 \end{pmatrix}$$

$$X = 10^{100}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & & 0 & \end{pmatrix} + 10^{-6} \; 10^{99}\begin{pmatrix} 10^{10000} & \\ & \end{pmatrix}$$

$$+ \; 10^{17}\begin{pmatrix} & \end{pmatrix} + \cancel{10^{-100}\begin{pmatrix} & \end{pmatrix}}$$

rank-k approx

keep $\sum_{i=1}^{k} s_i u_i v_i^T$

$= 1$

$U^T U = I \quad V^T V = I$

$\| u_i v_i^T \|_F^2$

$= tr( v_i u_i^T u_i v_i^T ) \cdots$

Input $X$ : $d \times n$.

Apply SVD $\rightarrow$ $U, S, V^T$

keep $\underbrace{(u_1 \ldots u_k)}$ $U' \rightarrow \boxed{d \times k}$

new subspace

dim - red :

$$x' = \sum_{i=1}^{k} \underline{\alpha_i}\, u_i$$

$$\underset{d \times n}{X} \longrightarrow \underset{k \times n}{X'}$$

$$\underset{d \times k}{} \quad \underset{k \times d}{(U')^T}$$

$$X \longrightarrow \boxed{(U')^T \cdot X}$$

error:

$$\min_{M} \|M - X\|_F^2$$

s.t. $\operatorname{rank}(M) \le k$.

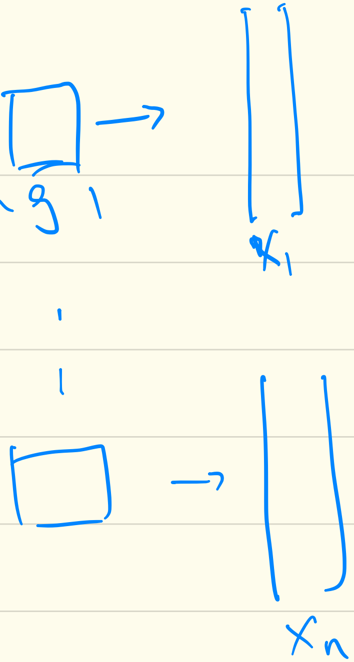
$$M = \sum_{i=1}^{k} s_i u_i v_i^T$$

$$X = \sum_{i=1}^{d} s_i u_i v_i^T$$

$$X - M = \sum_{i=k+1}^{d} s_i u_i v_i^T$$

$$\| X - M \|_F = \left\| \sum_{i=k+1}^{d} s_i u_i v_i^T \right\|_F$$

$$= \sqrt{\sum_{i=k+1}^{d} s_i^2} \quad \text{(fact)}$$




$X = USV^T$

$X \approx USV^T$

$X = (x_1 \; \cdots \; x_n)$

$d \times n$

$n < d$

$O(n^2 d)$

$n > d$

$O(d^2 n)$

$n \approx d$

$O(n^3)$