Lecture 1

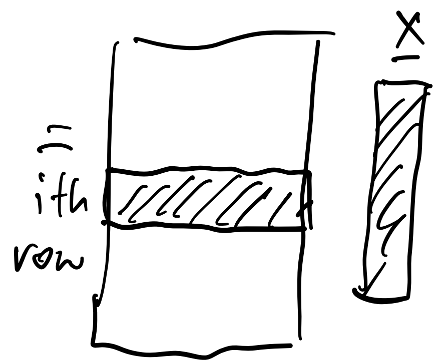
$\mathbb{R}^m$ $\quad \underline{x} = (x_1, \dots, x_m)$

$\mathbb{R}^{n \times m}$ $\quad n \times m$ matrices $\quad A = (A_{ij})$

$A \in \mathbb{R}^{n \times m}$ $\quad \underline{x} \in \mathbb{R}^m$ $\quad A \cdot \underline{x} \in \mathbb{R}^n$

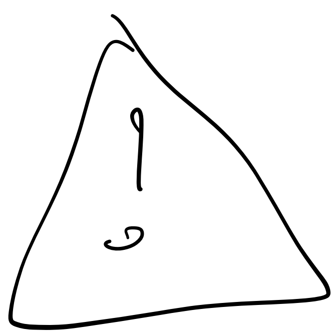
$\boxed{A \cdot \underline{x}}$

$(A\underline{x})_i =$



$$A = \begin{bmatrix} | & & | \\ \underline{a}_1 & \cdots & \underline{a}_m \\ | & & | \end{bmatrix} \qquad y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

$$A \cdot \underline{x} = \sum_{i=1}^{m} x_i \cdot \underline{a}_i$$



eg. $\underline{X} \in \mathbb{R}^{n \times m}$ $\quad \underline{X} = \begin{pmatrix} - & x_1 & - \\ & & \\ - & x_n & - \end{pmatrix}$

MINIMIZE $\beta \in \mathbb{R}^m$

$$\sum_{i=1}^{m} (y_i - x_i^T \cdot \beta)^2 = \| y - X\beta \|^2$$

$\boxed{A \cdot B}$ $\quad A \in \mathbb{R}^{n \times m} \quad B \in \mathbb{R}^{m \times p}$

$$(AB)_{ij} = \sum_{k=1}^{m} A_{ik} B_{kj}$$

$$A = \begin{pmatrix} | & & | \\ a_1 & \cdots & a_m \\ | & & | \end{pmatrix} \quad \text{columns}$$

$$B = \begin{pmatrix} — & b_1 & — \\ & \vdots & \\ — & b_m & — \end{pmatrix} \quad \text{rows}$$

$$A \cdot B = \sum_{k=1}^{m} \underset{n \times 1}{a_k} \cdot \underset{1 \times p}{b_k^T}$$

$n \times p$ rank 1 is of the form

$\underline{x}\, y^T \qquad \underline{x} \in \mathbb{R}^n \quad y \in \mathbb{R}^p$

e.g.

DATA $\underline{x}_1, \dots, \underline{x}_n \in \mathbb{R}^m$

DATA MATRIX

$$\boxed{\underset{\sim}{X} = \begin{bmatrix} — & x_1 & — \\ — & x_2 & — \\ & \vdots & \\ — & x_n & — \end{bmatrix} \in \mathbb{R}^{n \times m}}$$

$$\underbrace{\underset{\sim}{X}^T \underline{X}}_{m \times m} = \sum_{i=1}^{n} x_i \cdot x_i^T$$

$\left.\begin{array}{l}\text{columns of } X^T \\ \text{rows of } X\end{array}\right\}$ x_i's

say the data come some mean zero distrib. with covariance Σ then

$$\frac{1}{n} X^T X = \frac{1}{n} \sum_{i=1}^{n} x_i x_i^T \xrightarrow{P} \Sigma$$

LLN

(consistent estimator)

Singular Value Decomposition

$A \in \mathbb{R}^{n \times m}$

orthogonal matrices

$U \in O(m)$ if $UU^T = I_m$

$m \times m$

$U^T = U^{-1}$

so $U^T U = I_m$

$\det(UU^T) = \det(U)^2$

$\parallel$

1

the rows form orthonorm basis, the columns too

SVD $(n \geq m)$

$\exists \; U \in O(n), \; V \in O(m)$

s.t

$$\underset{n\times m}{A} = \underset{n\times n}{U} \cdot \underset{n\times m}{D} \cdot \underset{m\times m}{V^T}$$

with

$$D = \begin{bmatrix} \sigma_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_m \\ \hline 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{bmatrix} \qquad D_{ii} = \sigma_i \quad i = 1, \dots, m$$

where

$$\sigma_1 \geq \dots \geq \sigma_m \geq 0$$

σ_i = SINGULAR VALUES

Ex: u_i – columns of U

v_i – columns of V

show

$$\boxed{A \cdot v_i = \sigma_i \cdot u_i}$$

$$A \cdot v_i = U D V^T \cdot v_i$$

$\llcorner e_i$ ith canonic. unit vector

$$\boxed{A = \underbrace{U D}\underbrace{V^T} = \sum_{i=1}^{m} \sigma_i \cdot u_i v_i^T}$$

eigenvectors

$A \in \mathbb{R}^{m \times m}$ (square)

if $\exists \underline{v} \neq 0$ st

$A \cdot \underline{v} = \lambda \cdot \underline{v}$ for some λ

then $\underline{v}$ eigenvector

λ eigenvalue.

$\boxed{S^m}$ = $m \times m$ symmetric matrices

<u>Theorem (Spectral Thm)</u>

$A \in S^m$ $(A = A^T)$

there exists $U \in O(m)$

and diagonal $\Lambda \in \mathbb{R}^{m \times m}$

$$A = U \Lambda U^T$$

$\Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_m)$

$\lambda_i \in \mathbb{R}$

claim: the columns of U (u_i) are eigenvecto of A with eigenvalue λ_i.

$$A = \sum_{i=1}^{m} \lambda_i u_i u_i^T$$

quadratic forms

$A \in S^m$

$q_A(\underline{x}) = x^T A x$

A is positive semi-definite (PSD) if $q_A(x) \geq 0 \quad \forall \underline{x}$

A is positive-definite (PD) if $q_A(x) > 0 \quad \forall x \neq 0$

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

$S = (S_{ij}) \in \mathbb{R}^{n \times n}$ random matrix

$$\mathbb{E}X = (\mathbb{E}X_i)_{i=1,\dots,m}$$

$$\mathbb{E}S = (\mathbb{E}S_{ij})_{i,j}$$

$$\mu = \mathbb{E}X \in \mathbb{R}^m$$ mean vector

$$\Sigma = \operatorname{Var}(X)$$ covariance matrix

$$:= \mathbb{E}\left[(X-\mu)(X-\mu)^T\right]$$

$$= \left(\mathbb{E}[(X_i-\mu_i)(X_j-\mu_j)]\right)_{i,j}$$

$$\Sigma_{ii} = \mathbb{E}(X_i-\mu_i)^2 = \operatorname{Var}(X_i)$$

$$\Sigma_{ij} = \mathbb{E}(x_i - \mu_i)(x_j - \mu_j)$$
$$i \neq j$$
$$= \mathrm{Cov}(x_i, x_j)$$