→ Things to cover :

① Matrix factorizations

② Solving System of linear eq^{n}s

- errors and floating point width
- complexity

③ Optimization

④ Project (groups of 4)

- direct application
- analysis of or improvement of some algorithm } paper

Class 01

Fields : every non zero element has a multiplicative identity

e.g., $\mathbb{R}, \mathbb{Q}, \mathbb{C}$, finite fields

$\{\bar{0}, \bar{1}, \ldots, \overline{p-1}\}$; p is a prime

\# A set, V together with addition '+' is a vector space over $\mathbb{F}$ (a field) is a vector space vector space over if

- $a, b \in V \Rightarrow a+b \in V$

- $a+b=b+a$
- $\exists\, 0$ s.t. $a+0=0+a$
- $\forall a\in V, \exists -a$ s.t $a+(-a)=0$
- $(a+b)+c=a+(b+c)$
- $\forall \alpha\in\mathbb{F}$ & $a\in V$, $\alpha.a\in V$
- $\alpha(a+b)=\alpha a+\alpha b=(a+b)\alpha$

$\mathbb{F}=\mathbb{R}/\mathbb{C}$

$V=\mathbb{R}^n, n\geq 1 \;/\; \mathbb{C}^n, n\geq 1$

$\mathbb{F}^n=\{(a_1,\dots,a_n)\mid a_i\in\mathbb{F}\ \forall i\}$

Subspace : A vector subspace : Given a

V. space V over $\mathbb{F}$. A subspace of V is a subset $U\subseteq V$ s.t. U is a v. space over $\mathbb{F}$ in its own right

Linear Transformations

Given 2 v. spaces $V/_{\mathbb{F}}$ & $W/_{\mathbb{F}}$ A map $f: V\to W$ is a linear map if $\forall v, v'\in V$, $\alpha,\beta\in\mathbb{F}$,

$$f(\alpha v + \beta v') = \alpha f(v) + \beta f(v')$$

or

$$\begin{cases} f(v+v') = f(v) + f(v') \\ f(\alpha v) = \alpha f(v) \end{cases}$$

E.g., $V = \mathbb{R} = W$

(i) $1 : \mathbb{R} \to \mathbb{R}$

$a \mapsto a$

(ii) $c \in \mathbb{R}$ is fixed

$a \mapsto c.a$

$(a+b) \mapsto c(a+b)$

(iii) more examples: $x \to \int_0^x f(x)$?

$0 \in V$, $f(0) = ?$ where f is the zero map.

The zero map : $v \mapsto 0 \ \forall v \in V$, $0 \in W$

is a linear map

$f(0) = f(0+v) = 0.$

$f(-v) = -f(v)$

$f(0) = f(v + (-v)) = f(v) - f(v) = 0$

Defn A linear map $f : V \to W$ is called an isomorphism if ① f is a bijection (at the level of sets)

② f^{-1} (exists) is also linear

E.g., $f: a \mapsto c.a \; ; \; c \neq 0$

$f^{-1}: a \mapsto \frac{1}{c} a$

Defn : An $m \times n$ matrix over $\mathbb{F}$ is an array of mn elements in $\mathbb{F}$ which has m rows and n columns

E.g.,

$$\begin{array}{c} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & \ddots & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}_{m \times n} m \downarrow \\ n \rightarrow \end{array}$$

$= (a_{ij})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$

$=$ a tuple of 'rows' or 'columns'

$M_{m \times n}(\mathbb{F}) =$ the set of all $m \times n$ matrices is a vector space

Let $A = (a_{ij}) \in M_{m\times n}(\mathbb{F})$ & fix it

For $v \in \mathbb{F}^n$, $v = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$

Define $A.v = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & & \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$

$$= \begin{bmatrix} \sum_{j=1}^{n} a_{1j} v_j \\ \sum_{j=1}^{n} a_{2j} v_j \\ \vdots \\ \sum_{j=1}^{n} a_{mj} v_j \end{bmatrix}_{m \times 1} \in \mathbb{F}^m$$

we get a map. claim: This map is Linear

Matrix Vector Multiplication A

$$\# \quad A(v + v') = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} v_1 + v_1' \\ \vdots \\ v_n + v_n' \end{bmatrix}$$

$$= \begin{bmatrix} \sum a_{1j}(v_j + v_j') \\ \vdots \\ \sum a_{mj}(v_j + v_j') \end{bmatrix}$$

$$= \begin{bmatrix} \cdot \end{bmatrix} + \begin{bmatrix} \cdot \end{bmatrix}$$

$$A(\alpha v) = \alpha(Av)$$

If $A \in M_{m \times n}(\mathbb{F})$ then

$$v \mapsto Av$$

is a linear map from $\mathbb{F}^n \to \mathbb{F}^m$

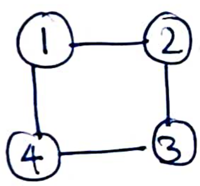
Where are matrices ?

1. Computer images are stored as matrices
2. Data matrix, consisting of observations and variables.

e.g., $\rightarrow$ locations

$\downarrow$ days $\quad a_{ij}$ = rainfall in location j on day i

3. Adjacency matrix of a network



	1	2	3	4
1	0	1	0	1
2	1	0	1	0
3	0	1	0	1
4	1	0	1	0

Lec - 2

→ $\mathbb{F}$ is a field ($\mathbb{R}/\mathbb{C}$)

→ $M_{m\times n}(\mathbb{F})$ = the set of all $m \times n$ matrices with entries in $\mathbb{F}$

→ $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & \ddots & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}_{m\times n} = (a_{ij})$ $\quad 1 \le i \le m$, $1 \le j \le n$

→ $\mathbb{F}^n = \{ (a_1, \ldots, a_n) \mid a_i \in \mathbb{F} \ \forall i \}$ is a vector space

or

$$= \left\{ \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \middle| \ a_i \in \mathbb{F} \ \forall i \right\}$$

→ Matrix vector multiplication :

$Av \in \mathbb{F}^m$

$A : \mathbb{F}^n \longrightarrow \mathbb{F}^m$

$v \longmapsto Av$

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \ddots & & \\ a_{m1} & \cdots & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^{n} a_{1j} v_j \\ \vdots \\ \sum_{j=1}^{n} a_{mj} v_j \end{bmatrix} \in \mathbb{F}^m$$

$A : v \longmapsto Av$ is linear

E.g., $A_{10\times 12} : \mathbb{F}^{12} \to \mathbb{F}^{10}$

Q. Given a linear map, $f : \mathbb{F}^n \to \mathbb{F}^m$, does there exist a matrix $A_f \in M_{m\times n}(\mathbb{F})$ s.t $f(v) = A_f \cdot v$ $\forall\, v \in \mathbb{F}^n$?

Check: $M_{m\times n}(\mathbb{F})$ is a vector space over $\mathbb{F}$

$$A = (a_{ij})_{\substack{1\le i\le m\\ 1\le j\le n}} \qquad B = (b_{ij})_{\substack{1\le i\le m\\ 1\le j\le n}}$$

(i) $(A+B) = (a_{ij} + b_{ij})$

(ii) $\alpha \in \mathbb{F}$

$\alpha A = (\alpha a_{ij})$

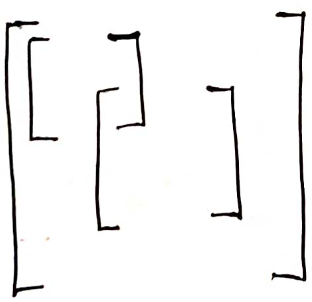
(iii) $O = (0)$, zero matrix

→ Square matrix if $m = n$
tall matrix if $m > n$
wide matrix if $m < n$

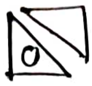
Column representation of matrices : $A = [A_{*1}, \ldots A_{*n}]$

Row representation : $A = \begin{bmatrix} A_{1*} \\ \vdots \\ A_{m*} \end{bmatrix}$

Submatrix :



→ The Identity matrix of order k : $Id_{k \times k}$ (square)

→ Upper Triangular matrix  : $a_{ij} = 0$ if $i > j$

Lower Triangular matrix  : $a_{ij} = 0$ if $i < j$

→ Density of a matrix :

$$nnz(A) = \left| \{ a_{ij} \neq 0 \mid a_{ij} \in A \} \right|$$

$$\text{density}(A) = \frac{nnz(A)}{mn}$$

zero matrix → $0 \leq den(A) \leq 1$

→ Sparse matrix has density close to 0.

$den(I_{k \times k}) = \frac{1}{k}$

$$den(UT_{strict}(k \times k)) = \frac{(k^2 + k)/2}{k^2} = \frac{k+1}{2k}$$

→ <u>Transpose</u>

$A_{m \times n} = (a_{ij})$; $A^T_{n \times m} = (a_{ji})$

$(A^T)^T = A$

→ <u>Symmetric</u> if $A = A^T$ and is A is square

Vector Spaces and matrices over $\mathbb{C}$

$z = a + ib$

$\bar{z} = a - ib$

Let V be a vector space over $\mathbb{C}$ (say, $V = \mathbb{C}^n$)

A complex inner product is a fn

$\langle , \rangle : V \times V \rightarrow \mathbb{C}$

defined by $\langle v, w \rangle = \sum v_i \overline{w_i} = w^* v$ where

w^* is a complex conjugate of a vector

Norm on V : $\|\cdot\| : V \rightarrow \mathbb{R}$ satisfying

$\|v\| > 0$

$\|v\| = 0 \Leftrightarrow v = 0$

$\|\alpha v\| = |\alpha| \|v\|$

$\|u + v\| \leq \|u\| + \|v\|$

→ Some Norms :

1) $\|v\|_1 = \sum |v_i|$

2) $\|v\|_2 = \sqrt{\langle v, v \rangle}$

$= \sqrt{v^* v}$

3) $\|v\|_\infty = \max |v_i|$

→ <u>Conjugate Transpose</u>

$A \in M_{m\times n}(\mathbb{C})$

$A^* = (\overline{a_{ji}})$

A square matrix is called Hermitian if $A = A^*$

$A = \begin{bmatrix} 1 & 2+3i \\ 4-6i & i \end{bmatrix}$, then $A^* = \begin{bmatrix} 1 & 4+6i \\ 2-3i & -i \end{bmatrix}$

<u>Lec - 3.</u>

→ AB is defined $\not\Rightarrow$ BA is defined

→ AB and BA is defined $\not\Rightarrow$ AB = BA

Lets focus on $M_n(\mathbb{F})$

$$A \cdot \begin{bmatrix} 1 & 0 & \cdots & \\ 0 & 0 & \ddots & \\ 0 & \vdots & & \\ \vdots & & & \\ 0 & & & \end{bmatrix} = \begin{bmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ a_{21} & 0 & \ddots & & \\ \vdots & \vdots & & \ddots & \\ a_{n1} & 0 & & & 0 \end{bmatrix}$$

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & \cdots & & 0 \\ \vdots & & & \\ 0 & \cdots & & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ 0 & \ddots & & \\ \vdots & & \ddots & \\ 0 & & & \end{bmatrix} A$$

$AB \neq BA$

$M_n(\mathbb{F})$ is

1. closed under matrix multiplication
2. not commutative
3. associative : $A(BC) = (AB)C$
4. distributive :

$$A(B+C) = AB + AC$$
$$(A+B)C = AC + BC$$

$AI_n = A = I_n A$

If $\exists$ A, B s.t. $AB = I$

$$AB = I$$
$$\Rightarrow BAB = BI = B$$
$$\Rightarrow (BA)B = B = (I)B$$
$$\Rightarrow BA = I$$

$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}^{-1}$ do not exist : $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix}$

→ A matrix $A \in M_n(\mathbb{N})$ is called invertible if $\exists$ $B \in M_n$ s.t. $AB = I_n = BA$. B is unique and is denoted by A^{-1}

$\rightarrow A : \mathbb{F}^n \rightarrow \mathbb{F}^n$

$$v \rightarrow Av$$

Note: If A is invertible, then the linear map $v \mapsto Av$ is a bijection and A^{-1} (i.e. $w \mapsto A^{-1}w$) is also linear. Such maps are called isomorphisms

$$\mathbb{F}^n \longrightarrow \mathbb{F}^n \longrightarrow \mathbb{F}^n$$
$$v \longmapsto Av \longmapsto A^{-1}(Av)$$

Determinants

$$\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$$

$$\det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = a_{11}(a_{22} \cdot a_{33} - a_{23} * a_{32}) - a_{12}(a_{21} a_{33} - a_{31} a_{23}) + a_{13}(a_{21} a_{32} - a_{31} a_{22})$$

Properties:

1) $\det(A) = \det(A^T)$
2) det do not change under the op $R_i \rightarrow R_i + cR_j$
3) B is obtained from A by $R_i \rightarrow cR_i$, then
$$\det(B) = c \det(A)$$
4) $\det A = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma)\, a_{\sigma(1)1} \cdots a_{\sigma(n)n}$

5> $\det(I_n) = 1$

6> $\det(O_n) = 0$

7> $\det(AB) = \det(BA) = \det(A).\det(B)$

If A is invertible, then

8> $\det(AA^{-1}) = \det(A).\det(A^{-1})$

$= \det(A).(\det(A))^{-1}$

Trace of a matrix

$tr(A)$ = The sum of its diagonal entries

9> $tr(AB) = tr(BA)$

10> $tr(A+B) = tr(A) + tr(B)$

Eigenvalues.

The characteristic polynomial of a matrix A is $\det(A_{n\times n} - x I_{n\times n})$

i.e $\det \begin{bmatrix} a_{11}-x & a_{12} & \cdots & \\ a_{21} & a_{22}-x & \cdots & \\ \vdots & \vdots & \ddots & \\ & & & a_{nn}-x \end{bmatrix}$

eigenvalues of A are the roots of the char poly of A : $\lambda_1, \lambda_2, \ldots, \lambda_n$ s.t

$\det(A - \lambda_i I) = 0$ for $i = 1, 2, \ldots, n$

Recall $A : \mathbb{F}^n \to \mathbb{F}^n$ is a linear map.

$$(v_1, v_2, \ldots, v_n) \mapsto (w_1, w_2, \ldots, w_n)$$

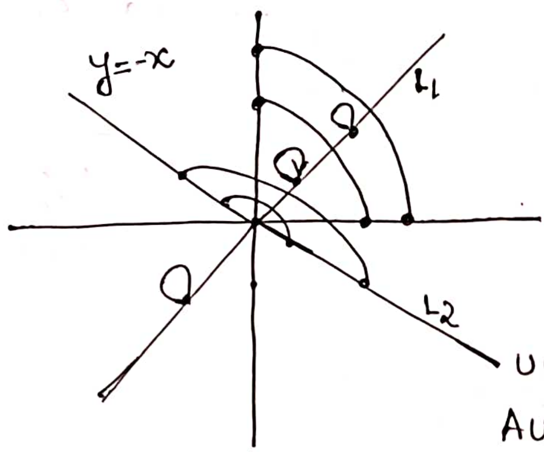
If $n=1$, $w = av$ for $a \neq 0$

$$w_i = a v_i$$

A scalar, λ is called an e-value of A if $\exists\, v \in \mathbb{F}^n \setminus \{\vec{0}\}$ s.t. $Av = \lambda v$; $\text{span}(v) = \{cv \mid c \in \mathbb{R}\}$

e.g., $n = 2$, $A : \mathbb{R}^2 \to \mathbb{R}^2$

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$



$A(L_1) = L_1$
$v \in L_1$
$Av = v$

$u \in L_2$
$Au = (-1)u$

$(0,b) \mapsto (b,0)$

$$\begin{pmatrix} a \\ b \end{pmatrix} \longmapsto \begin{pmatrix} b \\ a \end{pmatrix}$$

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} b \\ a \end{bmatrix}$$

$$\det \left[\begin{bmatrix} -x & 1 \\ 1 & -x \end{bmatrix}\right] = x^2 - 1 = 0 \Rightarrow x = \pm 1$$

Any matrix 'A' is <u>diagonalizable</u> if $\exists$ an invertible matrix P s.t. $P^{-1}AP = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & 0 \\ \vdots & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix} = \text{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$

<u>Spectrum of A</u>

= the set of e.values of $A = \{\lambda_1(A), \ldots, \lambda_n(A)\}$

Spectral Radius of A $\rho(A)$

$$\rho(A) = \max \{ |\lambda_i(A)| \mid \lambda_i(A) \in \text{spectrum}(A) \}$$

\# $\det(A) = \lambda_1(A) \cdot \lambda_2(A) \cdots \lambda_n(A)$

$\mathrm{tr}(A) = \lambda_1(A) + \lambda_2(A) + \cdots + \lambda_n(A)$

For $n=2$ case.

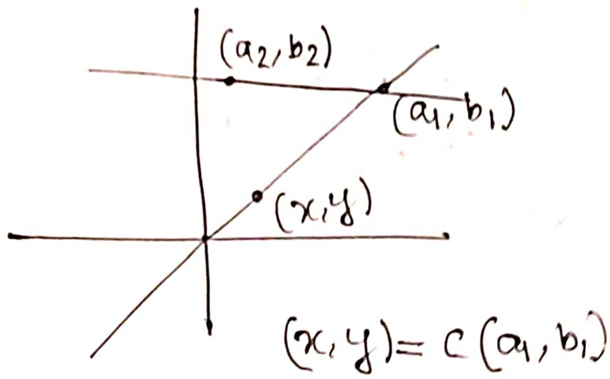
$$\begin{vmatrix} a-x & b \\ c & d-x \end{vmatrix} = (a-x)(d-x) - bc$$

$$= x^2 - \underbrace{(a+d)}_{\text{trace}}x + \underbrace{(ad-bc)}_{\det}$$

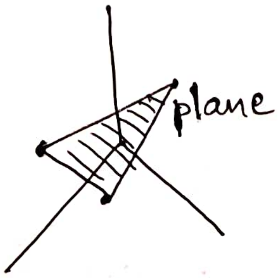
Linear Independence

Consider $\mathbb{R}$, $a, b \in \mathbb{R}\setminus\{0\}$, $\exists\, c \in \mathbb{R}$ s.t. $a = bc$

Consider $\mathbb{R}^2$,



Consider $\mathbb{R}^3$,



k vectors $v_1, v_2, \ldots, v_k$ in $\mathbb{R}^n$ are linearly independent if any linear combination

$$c_1 v_1 + c_2 v_2 + \cdots + c_k v_k = 0$$

$$\Rightarrow c_1 = c_2 = \cdots c_k = 0$$

Otherwise (i.e. atleast one c_i's are 0), they are called linearly dependent (or redundant)

$$c_1 v_1 + c_2 v_2 = 0$$

$$\Rightarrow v_1 = \left(-\frac{c_2}{c_1}\right) v_2$$

$\mathbb{F} = \mathbb{C}/\mathbb{R}$, $A^* =$ conjugate transpose of A

$$A(a_{ij}) \Rightarrow A^* = (a_{ij})^* = (\overline{a_{ji}})$$

For reals, $A^* = A^t$

Two vectors $v, w \in \mathbb{F}^n$ are orthogonal if $v^* w = 0 = \langle v, w \rangle$ written $v \perp w$

$$\boxed{\langle v, w \rangle = |v| . |w| \cos\theta}$$

A set S of vectors in $\mathbb{F}^n$ is called orthonormal if $x, y \in S$ & $x \neq y \Rightarrow x \perp y$ &

$$\|x\|_2 = \sqrt{\langle x, x \rangle} = 1$$

An $n \times n$ matrix of reals is orthogonal $\rightarrow O$ if its cols ($\equiv$ rows) form an orthonormal set of vectors.

If O is orthogonal, then (i) $O^t O = O O^t = I_n$ or $\boxed{O^{-1} = O^t}$

(ii) $\|Ov\| = \|v\| \; \forall \; v \in \mathbb{R}^n$

e.g., $O = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$, $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

A complex matrix A is called unitary if

$$AA^* = A^*A = I$$

Real counterpart → Orthogonal

Fact: (i) e. values of unitary matrices have modulus 1

(ii) e. values of orthogonal matrices ~~have~~ are +1, −1.

Singular values of A

They are the positive square roots of the e. values of A^*A (or A^tA if A is real)

For reals, $(A^tA)v = \lambda v \quad , v \neq 0$

$$\Rightarrow v^t(A^tA)v = v^t \lambda v$$

$$\Rightarrow (v^tA^t)Av = \lambda(v^tv)$$

$$\Rightarrow (Av)^t Av = \lambda \|v\|^2$$

$$\Rightarrow \|Av\|^2 = \lambda \|v\|^2$$

$$\Rightarrow \lambda > 0$$

Lec-04 — Motivation for Linear Algebra

Input : Samples v

Output : Computed Classification, $w = F(v)$

easy: $F(v) = Av$ (i.e. F is linear)

↳ entries of A are called "weights of F"

harder: $F(v) = \underline{A}v + \underline{b}$ (affine)

both are to be learned.

Optimisation step : Choose A and b which minimise the total loss of all training data at every iteration.

$$\min \|F(v) - (Av + b)\|^2$$

e.g., $Ax = b$ ← disease info

x ← risk factors

A ← health info of patients

① Solve $Ax = b$

② $Av = \lambda v$

③ $Av = \sigma v$

④ $\min\limits_{x \in S} \dfrac{\|Ax\|^2}{\|x\|^2}$

⑤ Factorisation of A. e.g., CR, LU, QR, $U\Sigma V^t$ (S.V.D.), $B^T B$ (Cholesky)

① Solving $Ax = b$ or

$$\boxed{A_{m\times n}\, x_{n\times 1} = b_{m\times 1}}$$

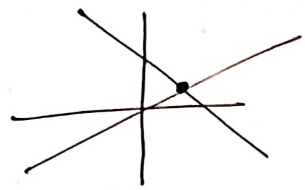
(system of m linear eq^{n}s in n variables)

$n = 1$, $ax = b$, $a \neq 0$

$\Rightarrow x = b/a$

$n = 2$

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \Rightarrow \begin{cases} a_{11} x_1 + a_{12} x_2 = b_1 \\ a_{21} x_1 + a_{22} x_2 = b_2 \end{cases}$$



A is $n \times n$ and is invertible

$$A x = b$$

$$\Rightarrow (A^{-1} A) x = (A^{-1} b)$$

$$\Rightarrow \quad x = (A^{-1} b)$$

(i) too restrictive,
(ii) computationally expensive

LU Decomposition.

$$U = \begin{bmatrix} U_{11} & & * \\ & \ddots & \\ 0 & & U_{nn} \end{bmatrix}, \quad L = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ * & & 1 \end{bmatrix}$$

$A = LU$ (assume)

Then $\quad A x = b$

$$\Rightarrow LU x = b$$

$$\Rightarrow \quad U x = L^{-1} b \quad (L^{-1} \text{ is easy to find})$$

$$\Rightarrow \quad U x = c \quad ; \quad x = L^{-1} b.$$

$$\Rightarrow \begin{cases} U_{nn} x_n = c_n \\ U_{n-1,n-1} \cdot x_{n-1} + U_{n-1,n} x_n = c_{n-1} \\ \cdots \end{cases}$$

V is a vector space over $\mathbb{R}$

① **Subspace** : A subset $U \subset V$ which is a v. space in its own right.

$(\equiv 0 \in U,\ a, b \in U,\ \alpha \in \mathbb{R} \Rightarrow a + \alpha b \in U)$

② **Linear combination of vectors** :

$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k\ ,\ \alpha_i \in \mathbb{R}\ \forall i$

③ Vectors $v_1, v_2, \ldots, v_k$ are L.I. if for any lin. comb. $\sum \alpha_i v_i = 0 \Rightarrow \alpha_i = 0\ \forall i$. Otherwise it is L.D.

④ Let $S = \{v_1, v_2, \ldots, v_k\}$, then

$\text{span}(S) = \{\sum \alpha_i v_i \text{ for all possible values of } \alpha_i\text{'s}\}$

$\text{span}(v) = \{\alpha v \mid \alpha \in \mathbb{R}\}$

verify : span(S) is a subspace of V.

⑤ A finite set $\mathcal{B} = \{u_1, u_2, \ldots, u_n\}$ of V is a basis of V if $\text{span}(\mathcal{B}) = V$. If $\text{Span}(\mathcal{B}) = V$, $u_1, u_2, \ldots, u_n$ are linearly independent. $\dim V = |\mathcal{B}| = \# \mathcal{B}$

⑥ Basis of $V = \mathbb{R}^n = \{e_1, e_2, \ldots, e_n\}$ where $e_i = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \leftarrow i^{th}$ pos

Standard Basis.

⑦ $A_{m \times n} = [(a_{ij})]$. The column space, col(A)

$= \text{span}\{A_{*1}, A_{*2}, \ldots, A_{*n}\}$

is the subspace of $\mathbb{R}^m$ spanned by columns of A.

Sly, $\text{row}(A) = \text{span}\{A_{1*}, A_{2*}, \ldots, A_{m*}\} \subseteq \mathbb{R}^n$

\# $$AX = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_m \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_m \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}x_1 \\ \vdots \\ a_{m1}x_1 \end{bmatrix} + \begin{bmatrix} a_{12}x_2 \\ \vdots \\ a_{m2}x_2 \end{bmatrix} + \cdots + \begin{bmatrix} a_{1n}x_m \\ \vdots \\ a_{mn}x_m \end{bmatrix}$$

$$= x_1 A_{*1} + x_2 A_{*2} + \cdots + x_m A_{*n}$$

= A L.C. of columns of A.

The system $Ax = b$ has a solution $\Longleftrightarrow b \in Col(A)$

v. imp:

$$A = CR \quad \text{or} \quad A_{m\times n} = C_{m\times r} R_{r\times n}$$

Look at A_{*1}. If it is non-zero, then $C_{*1} = A_{*1}$

If $A_{*2} \neq \alpha A_{*1}$, then $C_{*2} = A_{*2}$. else if $A_{*3} \neq \alpha A_{*1}$ then $C_{*2} = A_{*3}$. If $A_{*1}; A_{*2}, A_{*3}$ are L.I then $C_{*3} = A_{*3}$

$r = \#$ linearly independent columns of A.

E.g., $$A = \begin{bmatrix} 1 & 3 & 8 \\ 1 & 2 & 6 \\ 0 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \end{bmatrix}$$

Linearly independent cols of A) = # cols of C

$$\begin{array}{ccc} \text{Linearly independent cols of A} & = & \text{\# cols of C} \\ \| & & \| \\ \text{Linearly independent rows of A} & = & \text{\# rows of R} \\ \| & & // \\ \dim \text{Col}(A) & = & \dim \text{Row}(A) \\ & & \| \\ & & \text{The rank of } A = \text{rank } A. \end{array}$$

Lec-05 : LU Decomposition

Column space of A

$= \{ Ax \mid x \in \mathbb{R}^n \}$

[= the set of all l.c.s of cols of A]

$= \{ x_1 A_1 + x_2 A_2 + \cdots + x_n A_n \mid x_i \in A \}$

To find a set of vectors that are linearly independent and span col(A). The number of vectors of any such set is fixed. It is called the dimension of that space.

Column Rank of A = dim(~~Row~~ Column (A))
Row Rank of A = dim(Row(A))

<u>Thm</u>: col Rank (A) = Row rank (A)

Q. Calculate rank(A)?
It will help find if the answer also gives us a basis of col (A)

Q. Find the dim of span $\left\{ \begin{pmatrix} 10 \\ 7 \\ 8 \\ 7 \end{pmatrix}, \begin{pmatrix} 7 \\ 5 \\ 6 \\ 5 \end{pmatrix}, \begin{pmatrix} 8 \\ 6 \\ 10 \\ 9 \end{pmatrix}, \begin{pmatrix} 7 \\ 5 \\ 9 \\ 10 \end{pmatrix} \right\}$

Cor: rank(A) = rank(A^T)

$$\begin{bmatrix} 10 & 7 & 8 & 7 \\ 7 & 5 & 6 & 5 \\ 8 & 6 & 10 & 9 \\ 7 & 5 & 9 & 10 \end{bmatrix} \xrightarrow[R_2 \to R_2 - \frac{7}{10}R_1 \dots]{R_2 \to \frac{7}{10}R_1 \dots} \begin{bmatrix} 10 & 7 & 8 & 7 \\ 0 & * & * & * \\ 0 & 0 & * & * \\ 0 & 0 & 0 & * \end{bmatrix}$$

Raw operations changes the look of the rows but the property of linear dependency / independency doesn't change.

If there are dependencies atleast 1 of them will become a zero vector.

No. of non zero rows = Rank.

$$\begin{bmatrix} 10 & 7 & 8 & 7 \\ 7 & 5 & 6 & 5 \\ 8 & 6 & 10 & 9 \\ 7 & 5 & 9 & 10 \end{bmatrix} \xrightarrow[\substack{R_3 \to R_3 - \frac{8}{10}R_1 \\ R_4 \to R_4 - \frac{7}{10}R_1}]{R_2 \to R_2 - \frac{7}{10}R_1} \begin{bmatrix} 10 & 7 & 8 & 7 \\ 0 & 1/10 & 19/5 & 7/2 \\ 0 & 2/5 & 18/5 & 17/5 \\ 0 & 1/10 & 17/5 & 51/10 \end{bmatrix}$$

$\downarrow$

$(7\,5\,6\,5) - \frac{7}{10}(10\,7\,8\,7)$

$(8\,6\,10\,9) - \frac{8}{10}(10\,7\,8\,7)$

$(7\,5\,9\,10) - \frac{7}{10}(10\,7\,8\,7)$

new R_i = old $R_i + cR_j$, $c \neq 0$

$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \xrightarrow{R_2 \to R_2 - \frac{a_{21}}{a_{11}}} \begin{bmatrix} a_{11} & a_{12} \\ 0 & a_{22} - \frac{a_{21}}{a_{11}} a_{12} \end{bmatrix}$ $= R_2$

$$\begin{bmatrix} a_{11} & a_{12} \\ 0 & \frac{a_{11}a_{22} - a_{12}a_{21}}{a_{11}} \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 \\ -\frac{a_{21}}{a_{11}} & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

$R_i \leftarrow R_i + cR_j$ is equivalent to matrix multiplication

$j<i$ XA where

$$X = \begin{bmatrix} 1 & & & 0 \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & \cdots & & \\ -\frac{a_{21}}{a_{11}} & \ddots & & 0 \\ -\frac{a_{31}}{a_{11}} & 0 & \ddots & \\ \vdots & & & \\ -\frac{a_{m1}}{a_{11}} & & & \end{bmatrix} A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} \\ 0 & & \vdots & \vdots \\ \vdots & & & \\ 0 & a_{m2}^{(1)} & & a_{nn}^{(1)} \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 1 & 0 & 0 & \cdots & \\ 0 & 1 & 0 & \cdots & \\ \vdots & -a_{32}^{(1)}/a_{22} & 1 & \cdots & \\ \vdots & \vdots & & \ddots & \\ 0 & -a_{n2}/a_{22} & & & 1 \end{bmatrix} A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} \\ 0 & 0 & & \vdots \\ \vdots & \vdots & * & \\ 0 & 0 & & \end{bmatrix}$$

P.T.O

An $n \times n$ matrix L is called an unit. lower triangular if

- $l_{i,i} = 1$; $i = 1$ to n
- $l_{ij} = 0$; $j > i$

Facts :

① If A is a Unit Lower Triangular, then $\det(A) = 1$

② A^{-1} is also unit Lower Triangular (ULT)

③ A, B ULT $\Rightarrow$ AB is also ULT.

④ $\underbrace{L_n(\cdots(L_3(L_2(L_1}_{ULT}(A)))\cdots)$ ← Gaussian Elimination

$$= \begin{bmatrix} a_{11} & \cdots & & a_{1n} \\ 0 & a_{22}^{(1)} & \cdots & a_{21}^{(1)} \\ 0 & 0 & a_{33}^{(1)} \cdots & a_{31}^{(1)} \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

$L^{-1}A = U$

$\Rightarrow \boxed{A = LU}$

L06 – LU Factorization

→ A is a $n \times n$ invertible

Gaussian Elimination

$$A \xrightarrow[R_i \leftarrow R_i + cR_j]{i > j} A'$$

$\equiv E_{(ij)} A = A'$ where

Recall:

1. L is unit lower triangular
2. $\det(L) = 1$
3. U is upper triangular
4. $\det(U) = \det(L)$

$E_{(ij)}$ is the matrix which has 1's on the diagonal, c in $(i,j)^{th}$ position and 0 elsewhere.

$\equiv$ performing $R_i \leftarrow R_i + cR_j$ on $(Id_{n \times n})$.

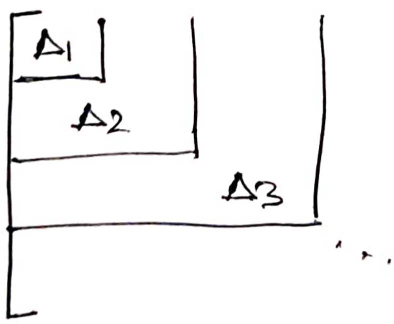
Defn: Principal Submatrices of A.

For every $k \in \{1, 2, \dots, n\}$ the k^{th} Principal Submatrix, denoted by $\Delta_k(A)$ is the $k \times k$ submatrix of A containing 1^{st} k rows and k columns.

$$\Delta_1(A) = \begin{bmatrix} a_{11} \end{bmatrix}$$

$$\Delta_2(A) = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$ and so on.. $\Delta_n(A) = A$

$A =$



Def 2: The k^{th} pivot of A

Recall:

$$E_1 A = \begin{bmatrix} a_{11} & \cdots & & a_{1n} \\ 0 & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} \\ 0 & 0 & & \\ \vdots & \vdots & & \\ 0 & 0 & \cdots & a_{nn}^{(1)} \end{bmatrix}$$

The k^{th} pivot of A is the entry $a_{kk}^{(k-1)}$ in the matrix

$E_k(E_{k-1}(\cdots(E_1 A))\cdots)$

$$E_2(E_1 A) = \begin{bmatrix} a_{11} & \cdots & & & a_{1n} \\ 0 & a_{22}^{(1)} & \cdots & & a_{2n}^{(1)} \\ 0 & 0 & a_{33}^{(2)} & \cdots & a_{3n}^{(2)} \\ 0 & 0 & 0 & & \\ \vdots & \vdots & \vdots & & \end{bmatrix}$$

Prop^n (1) Let A be an $n \times n$ invertible matrix. Then,

$$A = LU \iff \text{Each } \Delta_k(A) \text{ is invertible for } k = 1 \text{ to } n$$

E.g., $\begin{bmatrix} 0 & 1 \\ 4 & 6 \end{bmatrix}$ doesn't have LU factorization

Suppose we want to solve $Ax = b$

$$Ax = b$$
$$\Rightarrow L(Ux) = b$$
$$\Rightarrow Ly = b \quad \text{where } y = Ux$$

$$\begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ l_{21} & 1 & 0 & \cdots & 0 \\ \vdots & & \ddots & & \\ & & & & \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

then
$$y_1 = b_1$$
$$l_{21} y_1 + y_2 = b_2$$
$$\cdots$$

Then solve $Ux = y$.

Backward / Forward Substitution

 # Fact: If B is a Δ^{lar} matrix, then $\det(B) = \prod_{i=1}^{n} b_{ii}$

Proof of Prop ①

Suppose $A = LU$

$$\det(A) = \det(L)\det(U)$$
$$= 1\left(a_{11} \cdot a_{22}^{(1)} a_{33}^{(2)} \cdots a_{mm}^{(m-1)}\right)$$
$$\neq 0 \longrightarrow ①$$

$\Rightarrow$ each pivot is non-zero.

$$\begin{array}{c} \overbrace{\hspace{2em}}^{k} \quad \overbrace{\hspace{1em}}^{n-k} \\ \begin{bmatrix} \Delta_k(A) & A_2 \\ A_3 & A_4 \end{bmatrix} \end{array} = \begin{bmatrix} L_1 & 0 \\ L_3 & L_4 \end{bmatrix} \begin{bmatrix} U_1 & U_2 \\ 0 & U_4 \end{bmatrix}$$

$$RHS = \begin{bmatrix} L_1 U_1 & L_1 U_2 \\ L_3 U_1 & L_3 U_2 + L_4 U_4 \end{bmatrix}$$

$\Rightarrow \Delta_k A = L_1 U_1$

$\Rightarrow \det(\Delta_k A) \neq 0$ (from ①) $\Rightarrow \Delta_k(A)$ is invertible

Converse : Each $\Delta_k(A)$ is invertible.

$$\Delta_1(A) = [a_{11}], \quad a_{11} \neq 0$$

$$\therefore \exists\, E_1 \text{ s.t. } E_1 A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} \\ 0 & 0 & \ddots & * \\ \vdots & \vdots & & \end{bmatrix} = A_2$$

Infact,

$$E_1 = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ -a_{21}/a_{11} & 1 & \cdots & 0 \\ \vdots & & \ddots & \\ -a_{n1}/a_{11} & 0 & \cdots & 1 \end{bmatrix}$$

Assume $E_{k-1}E_{k-2}\cdots E_1 A = A_k$

$$= \begin{bmatrix} a_{11} & \cdots & \cdots & a_{1n} \\ 0 & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} \\ 0 & 0 & \ddots & * \\ \vdots & \vdots & & \\ 0 & 0 & \cdots & \end{bmatrix}$$

Let $L' = E_{k-1}\cdots E_1$, it is unit lower Δ^{lar}

Then $L'A = A_k$

$$LHS = \begin{bmatrix} L_1 & 0 \\ L_3 & L_4 \end{bmatrix} \begin{bmatrix} \Delta_k(A) & A_2 \\ A_3 & A_4 \end{bmatrix}$$

$$= \begin{bmatrix} L_1\Delta_k(A) & L_1A_2 \\ L_3\Delta_k(A) + L_4A_3 & L_3A_2 + L_4A_4 \end{bmatrix}$$

$$RHS = \begin{bmatrix} U_1 & * \\ 0 & * \end{bmatrix}$$

$\therefore \; L_1(\Delta_k(A)) = U_1$

$\Rightarrow$ the $(k,k)^{th}$ entry of U_1 is non-zero

$\Rightarrow$ We can perform Gaussian Elimination on $L'A$ to bring zeros in k^{th} column below row k.

Hence, proved

$$U = \begin{bmatrix} U_{11} & & * \\ & \ddots & \\ 0 & & U_{m} \end{bmatrix} = \begin{bmatrix} U_1 & & 0 \\ & \ddots & \\ 0 & & U_{nn} \end{bmatrix} \begin{bmatrix} 1 & & * \\ & 1 \ddots & \\ 0 & & 1 \end{bmatrix}$$

$$= DU'$$

$A = LU = LDU'$ ← unit upper Δ^{lar}

D ↑ diagonal matrix

L ↑ unit lower Δ^{lar}

#

$$A = \begin{bmatrix} 0 & 1 \\ 4 & 6 \end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_2} \begin{bmatrix} 4 & 6 \\ 0 & 1 \end{bmatrix} = B = I_2 B$$

$$\underset{(PA)}{\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 4 & 6 \end{bmatrix}} = \underset{(LU)}{I_2 B}$$

$\boxed{A = LU \text{ GE without pivoting}}$

Suppose A is invertible but k^{th} pivot position is 0 for some k

Defn ③ : A matrix P is called a permutation matrix if it is obtained by permuting rows/cols of I_n

Propn ② : Let A be $n \times n$ invertible matrix. Then $\exists$ a permutation matrix P s.t. $A_k(PA)$ is invertible for $k = 1$ to n.

Cor: $PA = LU$

Eg., $A = \begin{bmatrix} 0 & 1 & 1 \\ 2 & 2 & 3 \\ 5 & 1 & 0 \end{bmatrix}$. Find P, L, U s.t $PA = LU$.

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 1 \\ 2 & 2 & 3 \\ 5 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 5/2 & -4 & 1 \end{bmatrix} \begin{bmatrix} 2 & 2 & 3 \\ 0 & 1 & 1 \\ 0 & 0 & -3 \end{bmatrix}$$

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 1 \\ 2 & 2 & 3 \\ 5 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 2 & 3 \\ 0 & 1 & 1 \\ 5 & 1 & 0 \end{bmatrix}$$

Lec-07

Prop^n: A is a $n\times n$ invertible matrix. Then $\exists$ a permutation matrix P s.t. $\Delta_K(PA)$ is invertible for $k = 1$ to n ; $[n] = \{1, 2, 3, \dots, n\}$

Def^n: A matrix P is an $n\times n$ permutation matrix if $\exists$ a bijection $f: [n] \to [n]$ s.t. $P_{*f(i)} = e_i$

where $e_i = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \leftarrow i^{th}$ pos.

$\det(P) = \pm 1$

E.g.,

$Id = [e_1 | e_2 | \cdots | e_n]$

$P = [e_{f(1)} | e_{f(2)} | \cdots | e_{f(n)}]$

if $f = \begin{pmatrix} 1 & 2 & 3 \\ \downarrow & \downarrow & \downarrow \\ 2 & 1 & 3 \end{pmatrix} \to P_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

if $f = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \to P_1 = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$

\# $A_{*j} \to j^{th}$ column of A

$A_{j*} \to j^{th}$ row of A

\# $|\det(A)| = |\det(PA)|$

Proof of Prop.

'By induction on n'

If $n=1$ or 2 are clear

Suppose the statement is true for all invertible matrices upto order $n-1$

A is $n\times n$ invertible

A' is $n\times(n-1)$ rows obtained by removing the last col. Then $\operatorname{rank}(A') = n-1$

$\Rightarrow A'$ has $(n-1)$ lin. independent rows.

Choose a perm. matrix P_1 s.t. $P_1 A'$ has linearly indep. We've that $P_1 A$ is invertible & $\Delta_{n-1}(P_1 A)$ is also invertible with size $(n-1)\times(n-1)$

Slly, $\exists\ P_2'$ of order $n-1$ s.t. $P_2'(\Delta_{n-1}(P_1 A))$ has invertible sub matrices. Extend P_2' to P_2 by adding e_n.

$$P_2 = \begin{bmatrix} \begin{bmatrix} P_2' \end{bmatrix} & \begin{matrix} 0 \\ 0 \\ \vdots \end{matrix} \\ 0 & 1 \end{bmatrix}$$

Consider $P_2 P_1 A$. continue if needed.

Cor: A is $n\times n$ invertible $\Rightarrow PA = LU$

E.g., If $n=3$, $E_2 P_2 E_1 P_1 A = U = \begin{bmatrix} u_{11} & & * \\ 0 & u_{22} & \\ 0 & 0 & u_{33} \end{bmatrix}$

Let $E_1' = P_2 E_1 P_2^{-1}$

$E_2' = E_2$ then

Note that is unit lower Δ lar. always.

$$\underbrace{E_2' E_1'}_{=L^{-1}} \underbrace{P_2 P_1}_{=P} A = E_2 P_2 E_1 (P_2^{-1} P_2) P_1 A = E_2 P_2 E_1 P_1 A = U$$

$\therefore$ $P_2 E_1 P_2^{-1}$ is also unit lower Δ^{lar}

so that $L^{-1}P = U$ so that $P = LU$.

In general,

$$E_k' = P_{m-1} \cdots P_{k+1} E_k P_{k+1}^{-1} \cdots P_{m-1}^{-1}$$

$$E_{m-1}(P_{m-1}(E_{m-2}(P_{m-2}(\cdots(E_1 P_1 A)\cdots) \Rightarrow L^{-1}PA = U$$

LU Factorisation

1. Check if $a_{11} \neq 0$. If yes, then calculate L1
2. If $a_{11} = 0$, then choose $j \geq 2$ s.t. $a_{1j} \neq 0$ (this is the pivot) & exchange row 1 & j.
3. Calculate L1 for $P_1 A$
4. $A_1 := L_1 P_1 A = \begin{bmatrix} * & \cdots & * \\ 0 & & \\ \vdots & * & \\ 0 & & \end{bmatrix}$

Pseudocode for GE without pivoting

Input: $A_{m \times m}$ and Output: L & U

Init: $U = A$, $L = I$

```
for k = 1 to m-1
    for j = k+1 to m
        l_jk = u_jk / u_kk
        for i = k to m
            u_ji = u_ji - l_jk u_ki
        end for
    end for
end for
```

$$\begin{bmatrix} U_{11} & U_{12} & U_{13} \\ U_{21} & U_{22} & U_{23} \\ U_{31} & U_{32} & U_{33} \end{bmatrix}$$

$\underline{k=1}$: $l_{21} : U_{21}/U_{11}$

$U_{21} \leftarrow U_{21} - l_{21} U_{11}$

$U_{22} \leftarrow U_{22} - l_{21} U_{12}$

$U_{23} \leftarrow U_{23} - l_{21} U_{13}$

$l_{31} = U_{31}/U_{11}$

$U_{31} \leftarrow U_{31} - l_{31} U_{11}$

$U_{32} \leftarrow \cdots$

$U_{33} \leftarrow \cdots$

$\underline{k=2}$

$l_{32} = \dfrac{U_{32}}{U_{22}}$

GE for Partial Pivoting (GEPP)

$U = A, \; L = I, \; P = I$

~~for~~ ~~for~~ $k=1$ to $m-1$

  select $i \geq k$ to maximise $|U_{ik}|$

  exchange row i & k ($U_{kj} \leftrightarrow U_{ij}$ for $j=1$ to $k-1$)

  $P_{*i} \longleftrightarrow P_{*k}$

  for $j = k+1$ to m

    $l_{jk} = U_{jk}/U_{kk}$

    for $i = k$ to m

      $U_{ji} = U_{ji} - l_{jk} U_{ki}$

    end

  end

end

Operation Count of GE.

flop : floating point operation

$+, -, *, \div, \sqrt{\ }$

assignment do not count to flops.

$\underline{\text{divisions in } k^{th} \text{ col}}$:

$= m-k$

$\underline{1^{st}\text{ col}}$: Calculate $l_{21}, \dots, l_{m1}$: $m-1$ divisions +
$(m-1)$ rows & 2 ~~rows~~ flops for ~~each~~ $m-1$ entries in that row.

$\underline{k^{th}\text{ col}}$: $(m-k)$ divisions + $2(m-k)$ ops for $m-k$ rows

$$\text{Total operation count} = \sum_{k=1}^{m-1}(m-k) + \sum_{k=1}^{m-1} 2(m-k)^2$$

$$= \left(\sum_{1}^{m-1} m-k\right) + 2 \cdot \sum_{1}^{m-1}(m-k)^2$$

$$= \frac{m(m-1)}{2} + 2 \cdot \frac{(m-1)m(2m-1)}{6}$$

$$= \frac{4m^3 - m^2 - m}{6}$$

$$= \frac{2}{3}m^3 + O(m^2)$$

$$\sim \frac{2}{3}m^3 \text{ flops where } A_{m\times m}$$

Lec-08

\# $\underbrace{E_3 E_2 E_1}_{\downarrow \; L} \underbrace{P_3 P_2 P_1}_{P} A = U$

(why is this unit lower Δ^{lar} ?)

\# Operation count of LU factorisation $= O\left(\frac{2}{3}m^3\right)$
same as (GE)

Solving $Ax=b$, A assumed invertible

Method 1 : Sol^n using GE

Step 1: $Ax=b \rightarrow L'Ax = L'b$
$= Ux = L'b$

(multiplying by a unit lower Δ^{lar} matrix)
there are $m+(m-1)+\cdots+1 = \frac{m(m+1)}{2}$ multiplications
and $(m-1)+(m-2)+\cdots+1 = \frac{(m-1)m}{2}$ additions.

Step 2:
$$\begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1m} \\ 0 & u_{22} & \cdots & u_{2m} \\ & 0 & & \vdots \\ & & & u_{mm} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

$$x_K = \left(b_K - \sum_{j=K+1}^{m} u_{Kj}\, x_j\right) \Big/ u_{KK}$$

backward substitution, m divisions, $1+2+\cdots+(m-1)$ subtracti[ons] $= \frac{(m-1)m}{2}$
, $1+2+\cdots+(m-1) = \frac{(m-1)m}{2}$ multiplications

Total operation Count (OC) $= \frac{m(m-1)}{2} \times 2 + m = (2m^2 - m)$

Order $= O(\frac{2}{3}m^3 + 2m^2 - m) + \frac{m(m+1)}{2}$

Method 2

~~Step 2~~: Using L and U : $LUx = b$.

Step 1:

Step 2: Forward Substitution. First solve $Ly = b$;

then solve $Ly = b$

Total operation count $= 2m^2 - m$

Order $= O(\frac{2}{3}m^3 + 2m^2 - m)$

\# Method 2 is slightly better than Method 1.

Instability of Gaussian Elimination (w/o pivoting)

$$A = \begin{bmatrix} 10^{-20} & 1 \\ 1 & 1 \end{bmatrix} \quad R_2 \to R_2 - 10^{20} R_1 \quad \begin{bmatrix} 10^{-20} & 1 \\ 0 & 1-10^{20} \end{bmatrix}$$

$$L = \begin{bmatrix} 1 & 0 \\ 10^{20} & 1 \end{bmatrix}, \quad U = \begin{bmatrix} 10^{-20} & 1 \\ 0 & 1-10^{20} \end{bmatrix}$$

In practise, $\tilde{L} = \begin{bmatrix} 1 & 0 \\ -10^{20} & 0 \end{bmatrix}$, $\tilde{U} = \begin{bmatrix} 10^{-20} & 1 \\ 0 & -10^{20} \end{bmatrix}$

$\tilde{L}\tilde{U}$

$$\begin{bmatrix} 10^{-20} & 1 \\ -1 & 0 \end{bmatrix}$$

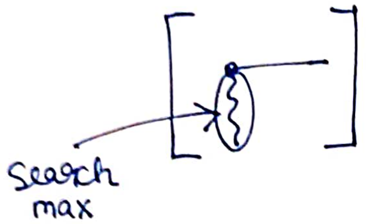
or

$$\begin{bmatrix} 1 & 0 \\ 10^{20} & 1 \end{bmatrix} \cdot \quad = \begin{bmatrix} 0 & 1 \\ 0 & -10^{20} \end{bmatrix}$$

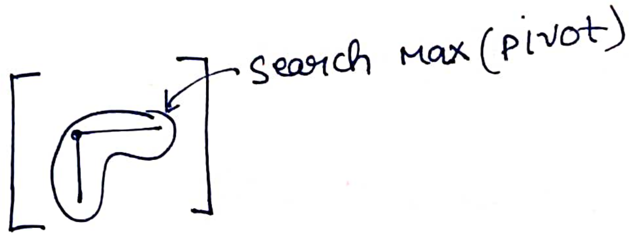
Pivotting helps deal with precision errors!

1) <u>Partial Pivotting</u>

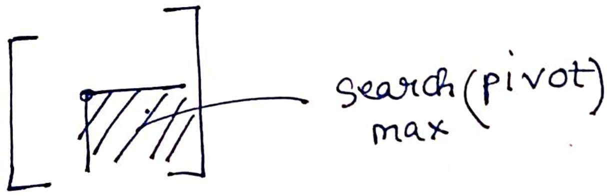
Select $i \geq k$ s.t. $|U_{ik}|$ is max among (k,k) to (n,k) entries and exchange $R_i \leftrightarrow R_k$.



2.) <u>Rook Pivotting</u>:



3.) <u>Complete Pivotting</u>:



\# In Rook/Complete Pivotting, we get,

$$PAQ = LU$$

\# <u>Tridiagonal Matrix</u>

$$\begin{bmatrix} b_1 & c_1 & 0 & 0 & 0 \\ a_2 & b_2 & c_2 & 0 & 0 \\ 0 & a_3 & b_3 & c_3 & 0 \\ 0 & 0 & a_3 & b_3 & c_3 \\ 0 & 0 & 0 & a_3 & b_3 \end{bmatrix}$$

$$\text{Density} = \frac{3n-2}{n^2}$$

For large n, Density $\to 0$

L and U will have a 'nice' structure.

Lec - 09

#An $A_{n\times n}$ matrix is called +ve definite if $\forall x \in \mathbb{R}^n \setminus \{0\}$

$$x^t A x > 0$$

$$\equiv \langle x, Ax\rangle \equiv \langle Ax, x\rangle > 0$$

$$\begin{bmatrix} x_1 & x_2 \end{bmatrix}\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22}\end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 & x_2 \end{bmatrix}\begin{bmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{bmatrix}$$

$$= a_{11}x_1^2 + a_{12}x_1x_2 + a_{21}x_1x_2 + a_{22}x_2^2$$

In general, $\sum_{i,j} a_{ij} x_i x_j > 0$

If A is +ve def.

#Suppose $Ax = 0 \Rightarrow \langle x, Ax\rangle = \langle x, 0\rangle = 0 \Rightarrow x = 0$

$\Rightarrow$ Cols of A are L.I.

$\Rightarrow$ A^{-1} exists

#2 Let λ be an eigen value of A and v be the corresponding eigen vector

$$v^t A v = v^t . \lambda v$$
$$= \lambda \|v\|^2 > 0$$

$\Rightarrow \lambda > 0$

$\Rightarrow \det(A) > 0$

#3. If A is p.d. then $a_{ii} > 0$

$$e_i^T A e_i = a_{ii} > 0$$

Prop^n: If A is real, symmetric, PD, then each $\Delta_k(A)$ is also SPD. (and hence LU decomposible)

Pf: $\Delta_k(A) = \begin{bmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \\ a_{k1} & \cdots & a_{kk} \end{bmatrix}$

Clearly it is symmetric.

To prove: Given $y \in \mathbb{R}^k \setminus \{0\}$, $y^t \Delta_k(A) y > 0$

$$x = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_k \\ 0 \\ 0 \end{bmatrix} \begin{matrix} \\ \\ \\ \\ \big\} n-k \\ \end{matrix} \in \mathbb{R}^n \setminus \{0\}$$

$$y^t \Delta_k(A) y = x^t A x > 0$$

$$\underbrace{[\, y \,|\, x \,]}_{k \quad n-k} \begin{bmatrix} \Delta_k(A) & A_2 \\ A_3 & A_4 \end{bmatrix} \begin{bmatrix} y \\ x' \end{bmatrix}$$

$$= [\, y^t \,|\, x'^t \,] \begin{bmatrix} \Delta_k(A) y + A_2 x' \\ A_3 y + A_4 x' \end{bmatrix}$$

$$= y^t \Delta_k(A) y + y^t A_2 x' + \cdots$$

A is SPD $\Rightarrow$ $A = LU$

Let $\Lambda = \text{diagonal}(u_{11}, u_{12}, \dots, u_{nn})$, each $u_{ii} > 0$ (verify!)

$$U' = \begin{bmatrix} 1 & & * \\ & \ddots & \\ 0 & & 1 \end{bmatrix}$$

Then $A = L\Lambda U'$

$\overset{A=}{\Rightarrow} A^t = (U')^t \Lambda^t L^t$

$= (U')^t \Lambda L^t$

$\Rightarrow U' = L^t$ ($\because A = A^t$)

Therefore, $A = L\Lambda L^T = L\sqrt{\Lambda} \cdot \sqrt{\Lambda} L^T$ where $\sqrt{\Lambda} = \text{diag}(\sqrt{u_{11}}, \dots, \sqrt{u_{nn}})$

$= (L\sqrt{\Lambda}) \cdot (L\sqrt{\Lambda})^T$

$= BB^T$ where $B = L\sqrt{\Lambda}$

Thm: (Cholesky factorization) Let $A_{n\times n}$ be an $SPD_{n\times n}$. Then $\exists$ a lower Δlar matrix B s.t. $A = BB^T$. Furthermore, B can be chosen so that its diagonal entries are +ve, in which case, B is unique.

Proof: Induction on n

Case $n = 1$: $[a] = [\sqrt{a}][\sqrt{a}]$

Suppose it is true for size upto $n-1$

$$A = \left[\begin{array}{c|c} a_{11}^{1\times 1} & w^T \\ \hline \underset{(n-1\times 1)}{w} & e_{n-1\times n-1} \end{array}\right] \quad ; \quad \alpha = \sqrt{a_{11}}$$

$$= \underbrace{\begin{bmatrix} \alpha & 0 \\ \frac{w}{\alpha} & I_{n-1} \end{bmatrix}}_{B_1} \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & \left(e - \frac{ww^T}{a_{11}}\right) \end{bmatrix}}_{A_1} \underbrace{\begin{bmatrix} \alpha & \frac{w^T}{\alpha} \\ 0 & I_{n-1} \end{bmatrix}}_{B_1^T}$$

Verify : $e - \frac{ww^T}{a_{11}}$ is P.D.

Then $e - \frac{ww^T}{a_{11}} = LL^T$ b/c induction and

$$A = \begin{bmatrix} \alpha & 0 \\ \frac{w}{\alpha} & I_{n-1} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & LL^T \end{bmatrix} \begin{bmatrix} \alpha & \frac{w^T}{\alpha} \\ 0 & I_{n-1} \end{bmatrix}$$

$$= \ldots$$

$$= \begin{bmatrix} \alpha & 0 \\ \frac{w}{\alpha} & L \end{bmatrix} \begin{bmatrix} \alpha & w^T/\alpha \\ 0 & L^T \end{bmatrix}$$

$$= BB^T.$$

Proof of uniqueness.

Let $A = B_1 B_1^T = B_2 B_2^T$

$\Delta_i = \text{diag}(\text{diag } B_i)$; $i = 1, 2, \ldots$

$$A = \underbrace{(B_1 \Delta_1^{-1})}_{\text{U.L.T}}\underbrace{(\Delta_1 B^T)}_{\text{U.T}}$$

= LU decomposition

$$\left.\begin{array}{l} B_1 \Delta_1^{-1} = B_2 \Delta_2^{-1} \\ \text{So } (\Delta_1 B_1)^T = (\Delta_2 B_2)^T \end{array}\right\} \quad (B_i \Delta_i)_{ji} = (B_i)_{ji}^2$$

$$\Delta_1 = \Delta_2$$

$$\Rightarrow B_1 = B_2$$

e.g.,

$$A = \begin{bmatrix} b_{11} & 0 & 0 \\ b_{21} & b_{22} & 0 \\ b_{31} & b_{32} & b_{33} \end{bmatrix} \begin{bmatrix} b_{11} & b_{21} & b_{31} \\ 0 & b_{22} & b_{32} \\ 0 & 0 & b_{33} \end{bmatrix}$$

$$a_{11} = b_{11}^2 \Rightarrow b_{11} = \sqrt{a_{11}}$$

$$a_{22} = b_{11}^2 + b_{22}^2 \Rightarrow b_{22} = \sqrt{a_{22} - b_{21}^2}$$

$$b_{33} = \sqrt{a_{33} - (b_{31}^2 + b_{32}^2)}$$

$$\vdots$$

$$\boxed{b_{kk} = \sqrt{a_{kk} - \left(\sum_{i=1}^{k-1} b_{ki}^2\right)}} \leftarrow \text{Diagonal entries.}$$

$$a_{21} = b_{21} b_{11} = a_{21}/b_{11} \; (= a_{12})$$

$$a_{31} = b_{31} b_{11} = a_{31}/b_{11}$$

$$a_{32} = b_{31} b_{21} + b_{32} b_{22}$$

$$\Rightarrow b_{32} = \frac{a_{32} - b_{31} b_{21}}{b_{22}}$$

$i \neq j$

$$b_{ij} = \frac{a_{ij} - \sum_{k=1}^{j-1} b_{ik} b_{jk}}{b_{jj}}$$

Cholesky (A)

$A = BB^T$

for $j = 1$ to $n-1$:

if $j == 1$:

$B_{11} = \sqrt{A_{11}}$

for $i = 2$ to n :

$B_{i1} = A_{i1}/B_{11}$

else

$B_{jj} = \sqrt{A_{jj} - \sum_{k=1}^{j-1} A_{jk}}$

$B_{ij} = \sum B_{ik} B_{jk}$

end

Operation Count :

$\sim \frac{n^3}{3} \rightarrow$

Lec-10 : Cholesky Decomposition

A is SPD (works also for complex Hermitian PD, $A = A^*$)

$\exists$ L which is lower Δ^{lar} with +ve diagonal s.t

$A = LL^T$

Pseudocode : $L = A$

for $k = 1$ to m:

$\quad l_{kk} = \left(a_{kk} - \sum_{j=1}^{k-1} l_{kj}^2\right)^{1/2}$

$\quad$ for $j = k+1$ to m :

$\quad\quad l_{jk} = \left(a_{jk} - \sum_{i=1}^{k+1} l_{ji} l_{ki}\right) / l_{kk}$

$\quad$ end

end

Operation Count (O.C.)

inner for loop : ~~1 division~~, $m-j+1$ multiplications and divisions
div, subtractions

Outer for loop: 1 square root, " "

$$O.C. = \sum_{k=1}^{m} \sum_{j=k+1}^{m} 2(m-j)+3$$

$$\approx 2m - 2 \cdot \sum_{k=1}^{m} \sum_{j=k+1}^{m} j$$

$$\approx \sum_{k=1}^{m} k^2 + k$$

$$= \frac{1}{3} m^3 + O(m^2)$$

Solving $Ax = b$

$\Rightarrow LL^T x = b \approx O(m^3)$

1. $\rightarrow$ $Ly = b$ (forward substitution) $\approx O(m^2)$

2. $\rightarrow$ $L^T x = y$ (back substitution) $\approx O(m^2)$

O.C. $\approx O(m^3)$ but all others are $O(m^2)$

Binary Representation

$$39 = 32 + 4 + 2 + 1$$

$$= 1\times 2^5 + 0\times 2^4 + 0\times 2^3 + 1\times 2^2 + 1\times 2^1 + 1\times 2^0$$

$$= (100111)_2$$

1bit for the sign.

\# $+39 = (+100111)_2$

$-39 = (-100111)_2$

If the highest power of 2 that divides N is 'n' then n+2 bits :

sign			. . . . . .		
	2^n	2^{n-1}	. . . .	2^1	2^0

The system represents integers from $-(2^{K+1}-1)$ to $(2^{K+1}-1)$

What about fractions ?

① $9{\cdot}75 = 8 + 1 + 0{\cdot}5 + 0{\cdot}25$

$$= 1.2^3 + 0.2^2 + 0.2^1 + 1.2^0 + \overset{2^{-1}}{1\left(\frac{1}{2}\right)^1} + \overset{2^{-2}}{1.\left(\frac{1}{2}\right)^2}$$

$$= (1001.11)_2$$

② $\frac{1}{3} = 0.2^0 + 0\left(\frac{1}{2}\right)^1 + 1.\left(\frac{1}{2}\right)^2 + 0.\left(\frac{1}{2}\right)^3 + 1.\left(\frac{1}{2}\right)^4 + \ldots$

$$= 0{\cdot}10101010\ldots$$

If we have k bits after the binary point and l bits before it then we can represent nos between 0 to 2^l with increments of $\frac{1}{2^k}$. Other numbers are not represented! So one has to come up with "rounding off" strategies.

If $k=1$, then $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$ in our system will be

$0{\cdot}1 \times 0{\cdot}1 = 0{\cdot}0$

Floating point representation

Scientific Representation/notation: (e.g., 6.022×10^{23})

$$\underbrace{\pm}_{\text{sign.}} \underbrace{\left(d_0 + d_1 b^{-1} + d_2 b^{-2} + \cdots + d_{p-1} b^{1-p}\right)}_{\text{mantissa.}} \times b^{e}$$

$e \leftarrow$ exponent

b: base

- $b \in [L, U]$; $L, U \in \mathbb{N}$, $L \leq U$
- $p :=$ precision of the representation
- $0 \leq d_i \leq b-1$; $i = 1$ to $p-1$
  $1 \leq d_0 \leq b-1$

Example ① Binary ; $(b = 2)$

$L = -1$, $U = 1$, $p = 3$. $\Rightarrow d_0 = 1$

$\boxed{\cdot} \rightarrow 0, 1$ or -1

$\pm 1.\square\square \times 2$

mantissa ↓ \ exp →	2^{-1}	2^{0}	2^{1}
$(1)_{10} = (1.00)_2$	$(0.500)_{10}$	$(1.000)_{10}$	$(2.000)_{10}$
$(1.25)_{10} = (1.01)_2$	$(0.625)_{10}$	$(1.250)_{10}$	$(2.500)_{10}$
$(1.50)_{10} = (1.10)_2$	$(0.750)_{10}$	$(1.500)_{10}$	$(3.000)_{10}$
$(1.75)_{10} = (1.11)_2$	$(0.875)_{10}$	$(1.750)_{10}$	$(3.500)_{10}$

precision $= 3 \Rightarrow$ upto 3 decimal places.

We can represent these 12 no's with this system.

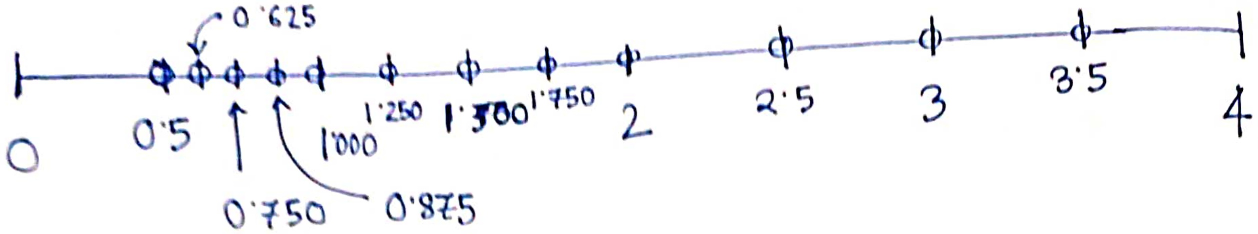
① But what would 0.8 be?

$|0.800 - 0.750| = 0.050$ ✓

$|0.800 - 0.875| = 0.075$ ✗

② 0.250? 2.000 or 2.500?

On number line,



Gaps increase as we move from $0 \to 4$

Lec-11

Defn: The floating point reprn of a real number, x with base β and exponent e and precision p.

$$sgn(x)\left(d_0 + d_1\beta^{-1} + \cdots + d_{p-1}\beta^{1-p}\right)\beta^e$$

$$= sgn(x)\left(\sum_{i=0}^{p-1} d_i \beta^{-i}\right)\beta^e ; \quad d_1, \ldots, d_{p-1} \in \{0, \ldots, \beta-1\}$$

A representation is normalised if $1 \le d_0 \le \beta-1$ and non-normalised otherwise.

Ex. (i) $x = \frac{11}{2}$

$\beta = 10, \quad p = 2$

$\pm (d_0 . d_1) 10^e$

$\therefore \frac{11}{2} = 5.5 \times 10^0$

(ii) $\beta = 2$

$(d_0 . d_1 d_2 \cdots d_{p-1}) \times 2^e$

$\frac{11}{2} = (1.2^0 + 0.2^{-1} + 1.2^{-2} + 1.2^{-3}).2^2$

$= (1.011)_2 \times 2^2$

$(= 101.1_2)$

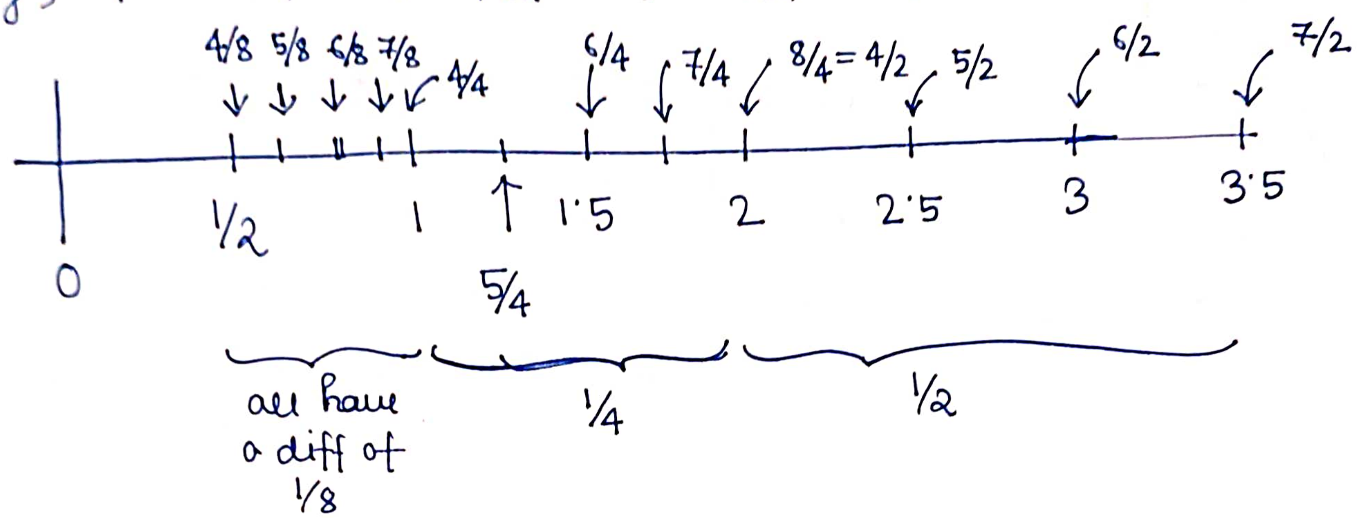
(iii) $x = \pi$, $p = 2$, $\beta = 10$

$\pi \rightsquigarrow 3.1 \times 10^0$

<u>Defn</u>: A floating number system w/ base β, exp e, precision p is a finite set

$$F = \{0\} \cup \left\{ \pm \left(\sum_{i=0}^{p-1} d_i \beta^{-i} \right) \beta^e \right\} \qquad \begin{matrix} 1 \le d_0 \le \beta - 1 \\ 0 \le d_i \le \beta - 1 \\ e_{min} \le e \le e_{max} \end{matrix}$$

E.g., $\beta = 2$, $p = 3$, $e_{min} = -1$, $e_{max} = 1$: $1.d_1 d_2 \times 2^e$



Consider the mantissa,

$$\sum_{i=0}^{p-1} d_i \beta^{-i} \le \sum_{i=0}^{p-1} (\beta - 1) \beta^{i}$$

$$= (\beta - 1) \left[1 + \beta^{-1} + \cdots + \beta^{1-p} \right]$$

$$= (\beta - 1) \left[1 + \frac{1}{\beta} + \cdots + \left(\frac{1}{\beta}\right)^{p-1} \right]$$

$$= (\beta - 1) \left[\frac{1 - (1/\beta)^p}{1 - 1/\beta} \right]$$

$$= \cancel{(\beta - 1)} \left[\frac{(\beta^p - 1)/\beta^p}{\cancel{(\beta - 1)}/\beta} \right] = \frac{\beta^p - 1}{\beta^{p-1}} = \beta(1 - \beta^{-p}) < \beta$$

$$\Rightarrow 1 \leq \sum d_i \beta^{-i} < \beta \quad ; \bar{x} \in F \setminus \{0\}$$

$$\bar{X}_{max} = \left(\sum (\beta - 1) \beta^{i}\right) \beta^{e_{max}} = (1 - \beta^{-p}) \beta^{e_{max}+1}$$

$$|\bar{X}_{min}| = \beta^{e_{min}}$$

The distance between 1 and the next ~~largest~~ consecutive fl. point no is β^{1-p}.

Defn: The machine epsilon / unit roundoff of F is

$$\varepsilon_{mach} = \frac{1}{2} \beta^{(1-p)}$$

The gap between consecutive fpn's in $[1, \beta]$ is $\beta^{-(p-1)}$

" " " " " $[\beta^{e}, \beta^{e+1}]$ is $\beta^{-(p-1)} \cdot \beta^{e}$

Defn: Floating pt. representation is a map

$$fl : \mathbb{R} \to F$$

$$x \mapsto fl(x) = \text{'closest' no in } F \text{ to } x.$$

$\downarrow$

called rounding of x.

E.g., $\beta = 10$, $p = 4$ $(d_0 . d_1 d_2 d_3 \times 10^{e})$

$$fl(10\pi) = fl(3.141592 \cdots \times 10^{1})$$

In F, we have,

$$F = \{ \ldots 3.140, 3.141, 3.142, 3.143, \ldots \} \times 10^{1}$$

$\uparrow$

10π

$$fl(x) := \begin{cases} sgn(x)\left(\sum_{i=0}^{\beta-1} d_i \beta^{-i}\right)\beta^e & \text{if } d_p < \frac{\beta}{2} \\ sgn(x)\left(\sum_{i=0}^{\beta-1} d_i \beta^{-i} + \beta^{-(p-1)}\right)\beta^e & \text{if } d_p \geq \frac{\beta}{2} \end{cases}$$

with ∞ precision full, i.e.

$$x = \left(\sum_{i=0}^{\infty} d_i \beta^{-i}\right)\beta^e$$

E.g., $\beta = 10,\ p = 3$

Relative & Absolute error

① $fl(1.234 \times 10^{-1}) = 1.234 \times 10^{-1} \rightarrow |fl(1.234\times10^{-1}) - (1.234\times10^{-1})| = 0.0004 < 5\times10^{-3}$

② $fl(1.235 \times 10^{-1}) = 1.24 \times 10^{-1}$

rel. error $= \left(\frac{4\times10^{-4}}{1.234\times10^{-1}}\right) \approx 3.2\times10^{-3} < 5\times10^{-2}$

③ $fl(1.295 \times 10^{-1}) = 1.30 \times 10^{-1}$

Thm: If x is a real number within the range of F and $|x| \in [\beta^e, \beta^{e+1}]$. Then the absolute error $|f(x) - x| \leq \frac{1}{2}\beta^{e(1-p)}$, The relative error is

$$\frac{|fl(x) - x|}{|x|} \leq \frac{1}{2}\beta^{1-\beta}$$

E.g., $(fl(10\pi) - 10\pi) =$ Abs. error $\approx 4.1 \times 10^{-3} < 5\times10^{-3}$

$\frac{fl(10\pi) - 10\pi}{10\pi} =$ Rel. error $\approx 1.3.. \times 10^{-4} < 5\times10^{-3}$

Floating pt arithmetic

⊛ is a binary operation F, ⊕ in F analogue of $+$ in $\mathbb{R}$.

$x \circledast y$ in $\mathbb{R} \equiv fl(fl(x) \circledast fl(y))$ in F

$= \bar{x} \circledast \bar{y},\ \bar{x}, \bar{y} \in F$

E.g., $\beta = 10,\ P = 3,\ e \in [-2, 2]$

① $x = 1.03 \times 10^{2}$, $y = 7.89 \times 10^{-1}$, $\bar{x} = x,\ \bar{y} = y$

$$\bar{x} + \bar{y} = 1.030 \times 10^{2} + 7.89 \times 10^{-1}$$

$$= 1.030 \times 10^{2} + 0.00789 \times 10^{2}$$

$$= 1.03789 \times 10^{2}$$

$$\bar{x} \oplus \bar{y} = 1.04 \times 10^{2}$$

② $z = 7.89 \times 10^{-2} = 0.000789 \times 10^{2}$

$$x \oplus z = 1.030 \times 10^{2} + 0.000789 \times 10^{2}$$

$$= 1.030 \times 10^{2}$$

$$= x$$

③ $(3.01 \times 10^{6})(4.56 \times 10^{15})$; $e \in [-17, 17]$

$$= (3.01 \times 4.56) \times 10^{21}$$

$$= 13.7256 \times 10^{21}$$

$= 1.37 \times 10^{22}$ ← 'out of bounds' error

H.W. Check $\dfrac{|fl(\bar{x} \otimes \bar{y}) - (\bar{x} \otimes \bar{y})|}{|\bar{x} \otimes \bar{y}|} \leq \varepsilon_{mach}$

$\forall x \in \mathbb{R}\ \exists\ \varepsilon$ s.t. $|\varepsilon| \leq \varepsilon_{mach}$ & $fl(x) = x(1+\varepsilon)$

$fl(\bar{x} \otimes \bar{y}) = (\bar{x} \otimes \bar{y})(1+\delta)$; $|\delta| \leq \varepsilon_{mach}$.

I.E.E.E. standards for F.P.A

1	e	p
sgn	exponent	mantissa.

$= (-1)^{\delta} \times 1.M \times 2^{E-(\ )}$

1 + e + p many bits → 34 bits (single precision)
→ 64 bits (double precision)

	single	double
mantissa (p)	23	52
exp (e)	8	11
sign (1)	1	1
total	32	64
e_{max}	+127	1023
e_{min}	−127	−1023
smallest fl.p.n	$2^{-126} \approx 10^{-38}$	$2^{-1022} \approx 10^{-308}$
largest fl.p.n	$2^{127} \approx 10^{38}$	$2^{1023} \approx 10^{308}$
ϵ_{mach}	$2^{-31} \approx 10^{-8}$	$2^{-63} \approx 10^{-16}$

Lec-11 Error Analysis / Perturbation Theory

- Attributes of a numerical algorithm
- Efficiency
  - time complexity (how fast?)
  - space (how much memory?)
- Accuracy
  - how ~~much~~ to measure error & accuracy?
  - given a tolerance, ϵ, can the algorithm find the soln within ϵ?
  - can you get a reasonable estimate of the error?

- Stability
  - How does a small change in the input affect the output?
  - How to control error propagation?
  - How much intervention is needed?

- How general your algorithm is?

- Implementation

Error Analysis

How much the solution of a problem is perturbed if the input is slightly perturbed?

$$f: \underset{\text{Data space}}{X} \xrightarrow{\text{cts}} \underset{\text{Soln space}}{Y}$$

e.g., given x, find $\sqrt{x}$. Define a problem to be well conditioned if a small change in x yields a small change in $f(x)$.

Say f is differentiable. Then

$$|f(x+\delta x) - f(x)| \approx |\cancel{f(x)} + \delta(x) f'(x) - \cancel{f(x)}|$$
$$= |\delta x| \, |f'(x)|$$

Even though $\delta x \to 0$, what if $f'(x) \to \infty$?

<u>Defn</u>: The absolute condition number of f at x:

$$\boxed{\hat{K} = \lim_{\substack{f, x \\ \delta \to 0}} \sup_{\|\delta x\| < \delta} \frac{\|\delta f\|}{\|\delta x\|}}$$

$\underline{\text{Def}^n}$: A problem is ill-conditioned at x if $\hat{\kappa}$ is large.

$\underline{\text{Def}^n}$: The relative condition number

$$\kappa = \lim_{\delta \to 0} \sup_{\|\delta x\| < \delta} \left(\frac{\|\delta f\|}{\|f(x)\|} \Big/ \frac{\|\delta x\|}{\|x\|} \right)$$

If f is one-variable differentiable, then,

$$\kappa = \frac{|f'(x)|}{|f(x)|} |x|$$

— Solving $x^2 + bx + c = 0$

Lets say $f(b,c) \longmapsto \sqrt{b^2 - 4c}$

$$\nabla f = f'\Big|_{(b,c)} = \left[\frac{1}{\cancel{2}\sqrt{b^2-4c}} \cancel{2}b , \frac{1}{\cancel{2}\sqrt{b^2-4c}} (\overset{-2}{\cancel{-4}}) \right]$$

$$\Rightarrow \|f'\|^2 = \frac{b^2}{b^2-4c} + \frac{4}{b^2-4c} = \frac{b^2+4}{b^2-4c}$$

$$\Rightarrow \|f'\| = \sqrt{\frac{b^2+4}{b^2-4c}} = \hat{\kappa}$$

<u>Rounding Errors.</u>

Suppose input is x, the ideal output is $f(x)$ but the actual output is $\text{alg}(x)$.

An algorithm is called backward stable if $\forall x \; \exists$ a small δx s.t.

$$\text{alg}(x) = f(x + \delta x) \; ; \quad \delta x : \text{backward error}$$

$$\text{error} = |\text{algo}(x) - f(x)| = |f(x+\delta x) - f(x)|$$
$$\approx |f'(x)| \cdot |\delta x|$$

An algorithm is backward stable if $|\delta x|$ is always small & $|f'(x)|$ is not too large.

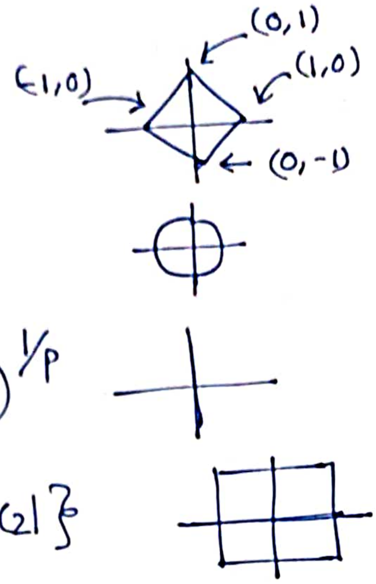
Norms.

$$\|x\|_1 = |x_1| + |x_2|$$

$$\|x\|_2 = \sqrt{x_1^2 + x_2^2}$$

$$\|x\|_{p<\infty} = (x_1^p + x_2^p)^{1/p}$$

$$\|x\|_\infty = \max\{|x_1|, |x_2|\}$$



If A is SPD, then

$$\langle x, y\rangle_A := x^t A y \;;\; \|x\|_2 = \sqrt{\langle x, x\rangle}$$

Defn: A matrix norm on $M_{m\times n}(\mathbb{R})$ is a function

$$\|\cdot\| : M_{m\times n}(\mathbb{R}) \to \mathbb{R}_{\geq 0} \text{ s.t.}$$

1) $\|A\| = 0 \iff A = 0_{m\times n}$

2) $\|\alpha A\| = |\alpha| \cdot \|A\|$

3) $\|A + B\| \leq \|A\| + \|B\|$

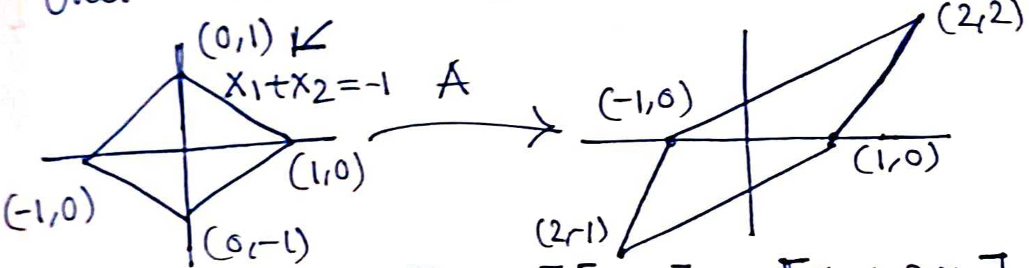
Operator Norm on $M_{m\times n}(\mathbb{R})$

$$\|A\| := \sup_{x\in\mathbb{R}^n\setminus\{0\}} \frac{\|Ax\|_{\mathbb{R}^m}}{\|x\|_{\mathbb{R}^n}}$$

$$= \sup_{\|y\|=1} \|Ay\|$$

Examples: $A = \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix}$; $A : \mathbb{R}^2 \to \mathbb{R}^2$. Consider $\|\cdot\|_1$ on $\mathbb{R}^2$. Then what is $\|A\|_1$?

unit circle on $\|\cdot\|_1$ norm



$$Ax = \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1+2x_2 \\ 2x_2 \end{bmatrix}$$

$$\left\|A\begin{pmatrix}0\\1\end{pmatrix}\right\|_1 = \left\|\begin{pmatrix}2\\2\end{pmatrix}\right\| = 4$$

$$\|Ax\|_1 = |x_1+2x_2| + |2x_2|$$

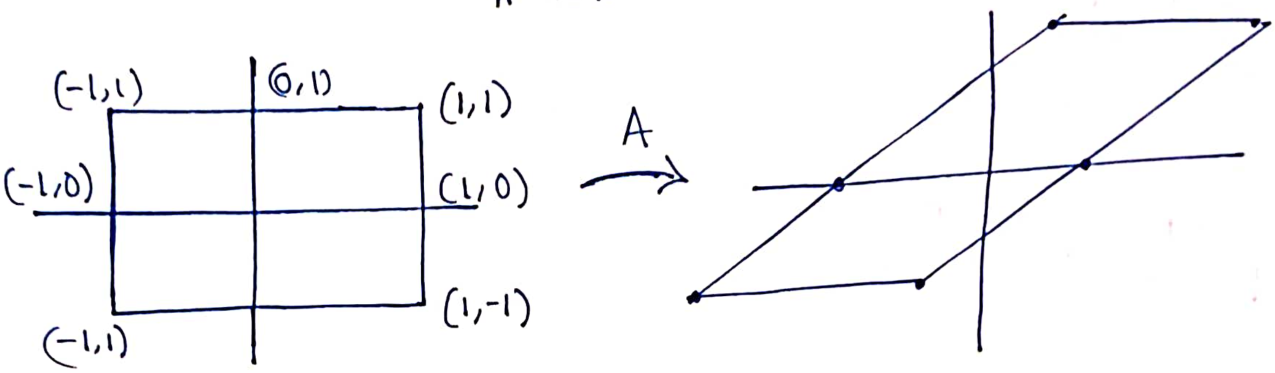
$$= |x_1| + 4(x_2)$$

$$= (|x_1|+|x_2|) + 3|x_2|$$

$$= 1 + 3|x_2| \leq 4 \quad \Rightarrow \sup = 4.$$

$\Rightarrow \|A\|_1 = 4 = $ abs(max col. sums)

— $\|A\|_\infty = \sup_{\|x\|_\infty = 1} \|Ax\|_\infty$



$\|A\|_\infty = 3 =$ abs (Max row sum)

— Not so trivial but $\|A\|_2 =$ largest e.v. of $A^T A$.

$\boxed{\text{Lec 12}}$

The operator norm on matrices.

$$\|A\|_p := \sup_{\|x\|_p = 1} \|Ax\|_p$$

Thm: ① $\|A\|_1 = \max_j \{ \|A_{*j}\|_1 \}$

$$= \max_j \left(\sum_i \{ |a_{ij}| \} \right)$$ (absolute max of col sums)

② $\|A\|_\infty = \max_i \{ \|A_{i*}\|_1 \}$

$$= \max_i \{ \sum_j |a_{ij}| \}$$ (absolute max of row sums)

③ $\|A\|_2 = \sup_{x \neq 0} \dfrac{x^t A^t A x}{x^t x}$

$=$ the largest e.value of A^tA.

* Abs error = | Computed error − true error |

* relative error = $\dfrac{\text{Abs error}}{\text{true error}}$

* <u>To compute $\sqrt{2}$</u>

You employ a method and you get 1·4

① Forward error = $|1{\cdot}4 - \sqrt{2}|$ (= absolute error)
(F.E.)

② Backward error = $|1{\cdot}4^2 - (\sqrt{2})^2| = 0{\cdot}04$

<u>Defn</u>: The Backward error of an algorithm (or an appx. soln to a numerical problem) is the amount by which the problem statement would have to change to make the appx. soln exact.

A problem / an algorithm is called "well conditioned" if

$$|BE| < \varepsilon_1 \Rightarrow |FE| < \varepsilon_2$$

Otherwise it is called ill conditioned.

<u>E.g.,</u> Solve : $ax = b$; $a \neq 0$

$x_0 = b/a$

x_c = computed soln

$FE = |x_c - x_0| = |\delta x|$ $\qquad \because b = a x_0$

$BE = |b - a x_c| = |a|\,|x_0 - x_c| = |a|\, FE$

$$\therefore \quad FE = \frac{1}{|a|} BE$$

The problem is well conditioned only when $|a| \gg 1$

$$\text{Rel. } FE = \frac{FE}{|x_0|}$$

$$= \frac{FE}{|b|/|a|}$$

$$= \frac{1}{|b|}(|a| FE)$$

$$= \frac{BE}{|b|}$$

<u>Defn</u>: The absolute condition number of a problem/ algorithm/ method.

$\hat{K} :=$ the ratio $\frac{BE}{FE}$ for small changes in the problem statement

$$= \lim_{\epsilon \to 0} \sup_{\|\delta x\| < \epsilon} \frac{\|\delta f\|}{\|\delta x\|}$$

Relative condition number

$$K = \lim_{\epsilon \to 0} \sup_{\|\delta x\| < \epsilon} \frac{\left(\frac{\|\delta f\|}{\|f\|}\right)}{\left(\frac{\|\delta x\|}{\|x\|}\right)}$$

<u>E.g.</u>, given x, compute $\sqrt{x}$; $x > 0$

$$\hat{K}(x) = \left|\frac{1}{2\sqrt{x}}\right|$$

$$K(x) = \frac{1}{2}$$

Consider the matrix vector multiplication problem

$A_{n \times n}$ is fixed, A^{-1} exists.

given $x \in \mathbb{R}^n$, Compute $Ax \in \mathbb{R}^m$.

$$\boxed{\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1+2x_2 \\ 2x_1+3x_2 \end{bmatrix}}$$

$$K = \sup_{\|\delta x\| < \epsilon} \frac{\dfrac{\|A(x+\delta x) - Ax\|}{\|Ax\|}}{\dfrac{\|\delta x\|}{\|x\|}}$$

$$= \sup \frac{\dfrac{\|A(\delta x)\|}{\|Ax\|}}{\dfrac{\|\delta x\|}{\|x\|}}$$

$$= \sup_{\|\delta x\| < \epsilon} \frac{\dfrac{\|A \delta x\|}{\|\delta x\|}}{\dfrac{\|Ax\|}{\|x\|}}$$

$$\text{RHS} = \frac{\|x\|}{\|Ax\|}$$

$$= \frac{\|A\|}{\|Ax\|} \cdot \|A^{-1}Ax\|$$

$$\leq \frac{\|A\| \, \|A^{-1}\| \, \cancel{\|Ax\|}}{\cancel{\|Ax\|}}$$

$$\leq K \leq \|A\| \, \|A^{-1}\|$$

* Try to show $K = \|A\| \, \|A^{-1}\|$

* Consider the problem of solving $Ax=b$ for invertible A w.r.t ① changes in A and x

$$A \longmapsto A^{-1}b.$$

$$(A+\delta A)(x+\delta x) = b.$$

$$\Rightarrow \quad A\delta x + \delta A x + \underbrace{\delta A\,\delta x}_{\text{v. small}} = 0$$

$$\Rightarrow \quad A\delta x + \delta A = 0.$$

$$\Rightarrow \quad \delta x = -A^{-1}\delta A x$$

$$\Rightarrow \quad \|\delta x\| = \|-A^{-1}\delta A x\|$$

$$\leq \|A^{-1}\| . \|\delta A\| . \|x\|$$

$$\Rightarrow \quad \frac{\|\delta x\|}{\|x\|} \leq \|A^{-1}\|\,\|\delta A\| = \|A^{-1}\|\left(\frac{\|\delta A\|}{\|A\|}\right).\|A\|$$

$$\Rightarrow \quad \frac{\left(\frac{\|\delta x\|}{\|x\|}\right) \leftarrow \text{rel. change in O/t}}{\left(\frac{\|\delta A\|}{\|A\|}\right) \leftarrow \text{rel. change in input}} \leq \|A^{-1}\|\,\|A\|$$

$$\Rightarrow \quad \kappa \leq \|A\| . \|A^{-1}\|$$

② changes in b. to x

$$b \longmapsto A^{-1}b$$

$$A(x+\delta x) = (b+\delta b)$$

$$\Rightarrow \quad \delta x = A^{-1}\delta b.$$

$$\Rightarrow \quad \|\delta x\| \leq \|A^{-1}\|\,\|\delta b\| \rightarrow (*1)$$

$$\|b\| \leq \|A\| . \|x\| \rightarrow (*2)$$

(*1) and (*2)

$$\Rightarrow \quad \frac{\|\delta x\|}{\|x\|} \le \|A\| \, \|A^{-1}\| \, \frac{\|\delta b\|}{\|b\|} .$$

$$\Rightarrow \quad \frac{1}{\mathcal{K}} \le \|A^{-1}\| \, \|A\|$$

$$\mathcal{K} = \sup_{\|\delta x\| < \epsilon} \frac{\dfrac{\|A(x+\delta x) - Ax\|}{\|Ax\|}}{\dfrac{\|\delta x\|}{\|x\|}}$$

$\vdots$

to get $\mathcal{K} \le \|A\| \cdot \|A^{-1}\|$

$$\Rightarrow \mathcal{K} = \|A\| \, \|A^{-1}\|$$

Defn: The p-condition number of a square matrix,

$$\mathcal{K}_p(A) = \|A\|_p \, \|A^{-1}\|_p$$

Thm: If $b \neq 0$, then

$$\frac{\|\delta x\|}{\|x\|} \le \mathcal{K}(A) \frac{\|\delta b\|}{\|b\|}$$

rel. change in output

rel. change in input

$f : (x_1, x_2) \longmapsto x_1 - x_2$

$\tilde{f} : (x_1, x_2) \longmapsto fl(x_1) \ominus fl(x_2)$

$fl(x_i) = x_i(1 + \varepsilon_i) \, ; \, |\varepsilon_i| < \varepsilon_m$

$$fl(x_1) \ominus fl(x_2) = \left[x_1(1+\varepsilon_1) - x_2(1+\varepsilon_2)\right](1+\varepsilon_3)$$

$$\parallel$$

$$\left(fl\left(fl(x_1) - fl(x_2)\right)\right) = x_1(1+\varepsilon_1) - x_2(1+\varepsilon_2)$$

$$|\varepsilon_1|,\ |\varepsilon_2| < 2\varepsilon_m$$

$$= \tilde{x}_1 - \tilde{x}_2$$

$$\Rightarrow \tilde{f}(x_1, x_2) = f(\tilde{x}_1, \tilde{x}_2)$$

$$\frac{|\tilde{x}_1 - x_1|}{|x_1|} = \varepsilon_4 < \theta(\varepsilon_m)$$

\# Compute $x \div x = 1$

$$\tilde{f}(x) = fl(x) \oplus\!\!\!\!\div\, fl(x)$$

$$= \left(x(1+\varepsilon_1) \div x(1+\varepsilon_1)\right)(1+\varepsilon_2)$$

$$= 1 + \varepsilon_2 \neq 1$$

\# Suppose, $fl(A) = A + \delta A$

$fl(b) = b + \delta b$

$$\delta A = (a_{ij}\, \varepsilon_{ij})$$

$$fl(a_{ij}) = a_{ij}(1+\varepsilon_{ij})$$

$$fl(b_i) = b_i(1+\varepsilon_i)$$

If

$$\|\delta A\|_\infty \leq \varepsilon_m \|A\|_\infty$$

$$\|\delta b\|_\infty \leq \varepsilon_m \|b\|_\infty$$

If $\hat{x}$ is the computed soln
and x_0 is the actual soln

Then,
$$\frac{\|\hat{x}-x_0\|_\infty}{\|x_0\|_\infty} \leq \left(\frac{K_\infty(A)}{1-\varepsilon_m K_\infty(A)}\right)(2\,\varepsilon_m)$$

$$g_n(A) = \frac{\max\limits_{i,j,k} |a_{ij}^{(k)}|}{\max\limits_{i,j} |a_{ij}|} = \text{The growth factor of a matrix; } A_{n\times n}$$

$$\text{rel FE} \leq O(n^3) \cdot g_n\, \varepsilon_m\, K_\infty(A)$$

Relative FE $\leq \Theta(n^3)\, g_n(A) \cdot \varepsilon_m\, K_\infty(A)$ → condition no.

$$A_{m\times n}\, x_{n\times 1} = b_{m\times 1}$$

07/03/24

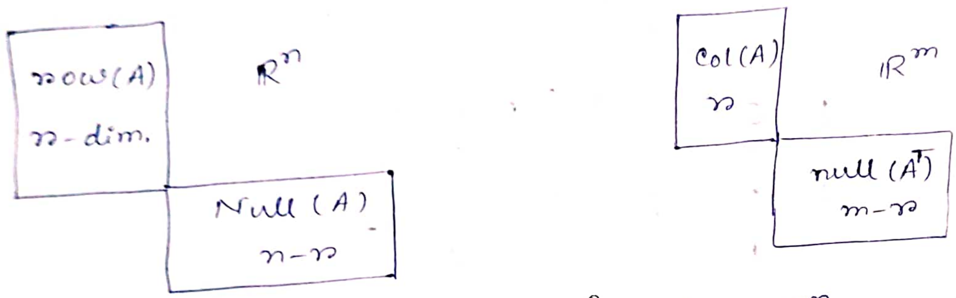
$A : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ linear map

$\text{Col}(A) = \{ Ax \mid x \in \mathbb{R}^n \} \subseteq \mathbb{R}^m$

$\text{Row}(A) = \{ A^T y \mid y \in \mathbb{R}^m \} \subseteq \mathbb{R}^n$

The null space $\text{Null}(A) = \{ x \in \mathbb{R}^n \mid Ax = 0 \}$

The left null space, $\text{Null}(A^T) = \{ y \in \mathbb{R}^m \mid A^T y = 0 \}$



The rank-nullity Theorem : $A : \mathbb{R}^n \to \mathbb{R}^m$

row-rank (A) + nullity (A) = n

col-rank (A) + nullity (A^T) = m

Thm. : Suppose $Ax = b$ has a sol^n. Then the set of all solutions is of the form x_0 + null (A), where x_0 is a particular sol^n.

Thm : If A is invertible, then $Ax = b$ has a unique sol^n.

Eg

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ +1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

↓ Row-echelon form of A

$A \xrightarrow{R_2 - R_1} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = E_1 A = U$

$L = E_1^{-1}$

$= \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}^{-1}$

$= \begin{bmatrix} 1 & & \end{bmatrix}$

- All zero rows are at the bottom.
- The leading entry of every non-zero row is on the right of the leading row of every row above.

Def^n : The leading entry of every a non-zero row is the left most non-zero entry in that row.

Thm. : If $A_{n\times m}$ then $\exists$ a permutation matrix $P_{n\times n}$, a unit lower Δ^{r} $L_{n\times n}$ & a row-echelon form matrix $U_{n\times m}$ such that

$$PA = LU$$

e.g.

① $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$ $P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

$A \xrightarrow[R_3-7R_1]{R_2-4R_1} \begin{bmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & -6 & -12 \end{bmatrix} = E_1A \xrightarrow{R_3-2.R_2} \underset{=U}{[\quad]}$ where $E_1 = \begin{bmatrix} 1 & 0 & 0 \\ -4 & 1 & 0 \\ -7 & 0 & 1 \end{bmatrix}$

$-6-(-3).2$

$A = \underset{L}{\begin{bmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ 7 & 2 & 1 \end{bmatrix}} \underset{U}{\begin{bmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & 0 & 0 \end{bmatrix}}$

$= \underset{L}{\begin{bmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ 7 & 2 & 1 \end{bmatrix}} \underset{D}{\begin{bmatrix} 1 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 0 \end{bmatrix}} \underset{U'}{\begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 9 \end{bmatrix}}$

This can take any real no.

② $A = \begin{bmatrix} 1 & 0 & 2 & 1 & 5 \\ 1 & 1 & 5 & 2 & 7 \\ 1 & 2 & 8 & 4 & 11 \end{bmatrix}_{3\times 5}$

$\frac{1}{2}$ $\frac{3}{6}$ $\frac{1}{3}$ $\frac{2}{6}$

$= \underset{3\times 3}{\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 2 & 1 \end{bmatrix}} \begin{bmatrix} 1 & 0 & 2 & 1 & 5 \\ 0 & 1 & 3 & 1 & 2 \\ 0 & 0 & 0 & 1 & 2 \end{bmatrix}_{3\times 5}$

Reduced row-echelon form :

→ The matrix in in reference, the leading entry in each non-zero row is 1.

→ Each column containing lead 1, has all other entries 0.

<u>Proposition</u>: Rank (A) = Rank (ref (A)) = Rank (rref (A))

= No. of pivots in ref (A)

= No. of leading 1's in rref (A)

<u>Proposition</u>: $Ax = b$ has a solⁿ <=> there is no pivots in the last column of rref $([A \mid b])$

Suppose $Ax = b$ has a solⁿ.

x_j is called a pivot variable if j-th column contains a leading 1.

Otherwise x_j is called a free variable.

<u>Solve</u>:

$$x_1 - x_2 + x_3 + x_4 - 2x_5 = (-1)$$
$$-2x_1 + 2x_2 - x_3 + x_5 = 2$$
$$x_1 - x_2 + 2x_3 + 3x_4 - 5x_5 = -1.$$

$$[A \mid b] = \left[\begin{array}{ccccc|c} +1 & -1 & 1 & 1 & -2 & -1 \\ -2 & 2 & -1 & 0 & 1 & 2 \\ 1 & -1 & 2 & 3 & -5 & -1 \end{array}\right]$$

$$= \left[\begin{array}{ccccc|c} +1 & -1 & 1 & 1 & -2 & -1 \\ 0 & 0 & 1 & 2 & -3 & 0 \\ 0 & 0 & 1 & 2 & -3 & 0 \end{array}\right]$$

$R_2 \to R_2 - (-2).R_1$

$R_3 \to R_3 - (1)R_1$

$$= \left[\begin{array}{ccccc|c} -1 & -1 & 1 & 1 & -2 & -1 \\ 0 & 0 & 1 & 2 & -3 & 0 \\ 0 & & & & & \end{array}\right]$$

~~$R_2 \leftrightarrow R_3$~~

$R_2 \to R_2 - R_3$

$$= \left[\begin{array}{ccccc|c} +1 & -1 & 1 & 1 & -2 & -1 \\ 0 & 0 & 1 & 4 & -7 & -2 \\ 0 & 0 & 1 & 2 & -3 & 0 \end{array}\right]$$

& $R_2 \leftrightarrow R_3$

$R \to R - R$

$$\begin{array}{cccccc|c} & s & & t & u & & \\ 1 & -1 & 0 & -1 & 1 & & -1 \\ 0 & 0 & 1 & 2 & -3 & & 0 \\ 0 & 0 & 0 & 0 & 0 & & 0 \\ \uparrow & & \uparrow & & & & \end{array}$$

rank $(A) =$ rank $([A|b]) = 2 < 5$

$Ax=b$ sol^n

$x_1 - s - t + u = -1$

$\Rightarrow x_1 = s + t - u - 1$

$x_3 = 3u - 2t$

The sol^n space is $\left\{ \begin{pmatrix} s+t-u-1 \\ s \\ 3u-2t \\ t \\ u \end{pmatrix} \middle| \; s, t, u \in \mathbb{R} \right\}$ ✓

✓

$$\underbrace{\begin{pmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + s\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + t\begin{pmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{pmatrix} + u\begin{pmatrix} -1 \\ 0 \\ 3 \\ 0 \\ 1 \end{pmatrix}}$$

null space

08/03/24

QR Factorization :

$A = QR \rightarrow$ upper triangular matrix

✓ The col^s form an orthonormal basis of col (A)

$A_{m \times n}$ $m \geq n$ & rank $(A) = n$

$$\| u - v \|^2 = (u-v)^T (u-v)$$
$$= (u^T - v^T)(u - v)$$
$$= u^T u - u^T v - v^T u + v^T v$$
$$= \| u \|^2 + \| v \|^2$$

Recall : Two vectors u, v are orthogonal if $\langle u, v \rangle = u^T v = 0$

Def^n: A set $\{v_1, v_2, \dots, v_k\}$ is an orthonormal basis of its span if

$$v_i^T v_j = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$$

Def^n: Two Subspaces $U, V \subseteq \mathbb{R}^n$ are orthogonal if $u^T v = 0$, $\forall u \in U$, $\forall v \in V$

Given $A_{m \times n} : \mathbb{R}^n \longrightarrow \mathbb{R}^m$

$$\text{row}(A) \perp \text{null}(A)$$
$$\text{col}(A) \perp \text{null}(A^T)$$

Remark: Later we'll see SVD which finds an orthonormal basis $\{v_1, v_2, \dots, v_r\}$ of row(A) & $\{u_1, u_2, \dots, u_r\}$ of col(A) such that $Av_i = \sigma_i u_i$, $v_i \in \mathbb{R}^n$ & $u_i \in \mathbb{R}^m$ for some $\sigma_i \in \mathbb{R}$

$\sigma_i \downarrow$ Singular values of A.

☑ $Q_{m \times n}$ is called orthogonal if $Q^T Q = I$

→ Q^T is called left inverse of Q.

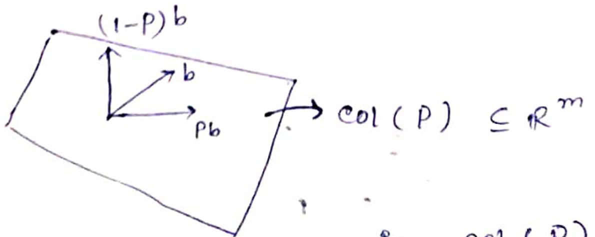
→ Cols of Q are orthonormal.

→ $P : QQ^T$

$$P^2 = (QQ^T)(QQ^T) = Q(Q^TQ)Q^T = QQ^T = P$$

Such matrices are called Projection Matrices.

☑ If $b \in \mathbb{R}^m$ then $Pb \; (= Q^T Q b)$ is the orthogonal projection of b on col(P).



Unique vector in col(P) closest to b !!

- If $m = n$ then $Q^T = Q^{-1}$
- If Q_1, Q_2 are orthonormal then $Q_1 Q_2$ is orthonormal

e.g. $A = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$, $\quad B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

$B^T B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I.$

$A^T A = I$

Generalized form of orthogonal matrix :

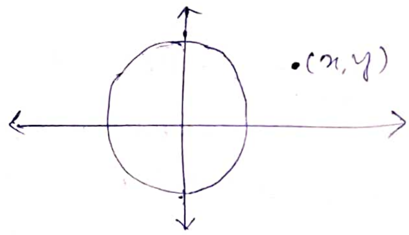
$$\left\{ \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \;\middle|\; \theta \in \mathbb{R} \right\} \cup \left\{ \begin{bmatrix} -\sin\theta & \cos\theta \\ \cos\theta & \sin\theta \end{bmatrix} \;\middle|\; \theta \in \mathbb{R} \right\}.$$

▪ $\|Qx\|^2 = (Qx)^T (Qx)$

$= (x^T Q^T)(Qx) = x^T (Q^T Q) x = x^T x = \|x\|^2$

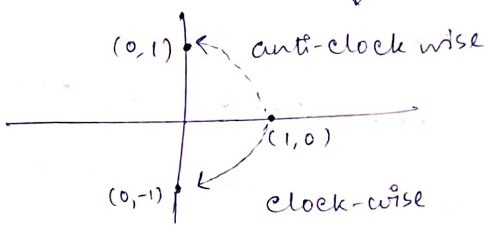
so for any x, $\|Qx\|^2 = \|x\|^2$ when Q orthogonal

▪ $\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} u \\ v \end{bmatrix}$



$\theta = \pi/2 \quad \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$

$\theta = -\pi/2 \quad \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$



▪ Suppose $A = \tilde{Q} \tilde{R}$ $\quad A_{m\times n}$, $m \geq n$, rank$(A) = n$.

$\tilde{Q}_{m\times m}$ is orthogonal

$\tilde{R}_{m\times n} = \begin{bmatrix} \begin{matrix} 0 & * \end{matrix} \\ O \end{bmatrix} \to R_{n\times n}$ is non-singular

$$m \begin{bmatrix} \\ A \\ \\ \end{bmatrix}_{n} = m \begin{bmatrix} \\ \tilde{Q} \\ \\ \end{bmatrix}_{m} \begin{bmatrix} 0 \quad * \\ \hline O \quad \tilde{R} \end{bmatrix} \begin{matrix} n \\ m-n \end{matrix}$$
$$\qquad\qquad\qquad\qquad\qquad\qquad n$$

$\tilde{Q} = [\, Q_{m\times n} \quad Q'_{m\times \overline{m-n}} \,]$

$\tilde{R} = \begin{bmatrix} R_{n\times n} \\ \hline O_{\overline{m-n}\times n} \end{bmatrix}$

$A = \tilde{Q}\tilde{R} = QR + Q'O = QR$ where Q has n orthonormal cols & R is $n\times n$ upper Δ^r matrix & invertible.

$$Q = \begin{bmatrix} q_{11} & q_{12} & \cdots & q_{1n} \\ q_{21} & q_{22} & \cdots & q_{2n} \\ \vdots & & & \\ q_{m1} & q_{m2} & \cdots & q_{mn} \end{bmatrix} \qquad R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & \cdots & r_{2n} \\ \vdots & & & \\ 0 & 0 & \cdots & r_{nn} \end{bmatrix}$$

$$QR = \begin{bmatrix} q_{11}r_{11} & q_{11}r_{12} + q_{12}r_{22} & \cdots & \sum_{i=1}^{n} q_{1i}r_{in} \\ q_{21}r_{11} & q_{21}r_{12} + q_{22}r_{22} & \cdots & \sum q_{2i}r_{in} \\ \vdots & \vdots & & \vdots \\ q_{m1}r_{11} & q_{m1}r_{12} + q_{m2}r_{22} & \cdots & \sum q_{mi}r_{in} \end{bmatrix}$$

$$= [\, a_1 \quad a_2 \quad \cdots \quad a_n \,]$$

$\Rightarrow$ Each column of A is a linear combination of cols of Q.

$\Rightarrow$ Cols Q form an orthonormal basis of col(A)

Consider $Ax = b$

Col$(A) \subseteq \mathbb{R}^m$

Case I: $b \in$ col(A) $\Rightarrow Ax = b$ has a unique solⁿ.

$Ax = b$

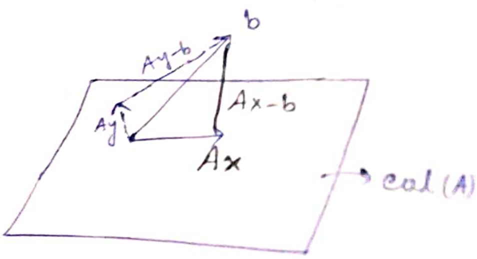
$\Rightarrow QRx = b$

$\Rightarrow Q^TQRx = Q^Tb$

$\Rightarrow Rx = Q^Tb$

$\Rightarrow \boxed{x = R^{-1}Q^Tb}$

Case II: $Ax = b$, & $b \notin$ col(A) has no solⁿ.



If b is not in col(A), we find some vector Ax s.t. $Ax - b$ has the least norm.

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad \| Ax - b \|_2$$

$$\| Ax - b \|_2^2 = \| b - Ax \|_2^2$$

$$= \| b - \tilde{Q}\tilde{R}x \|^2$$

$$= \| \tilde{Q}\tilde{Q}^T b - \tilde{Q}\tilde{R}x \|^2$$

$$= \| \tilde{Q}(\tilde{Q}^T b - \tilde{R}x) \|^2$$

$$= \| \tilde{Q}^T b - \tilde{R}x \|^2 \qquad \left[\because \| \tilde{Q} \|^2 = Q^T Q = I \right]$$

$$= \left\| \begin{bmatrix} Q^T b \\ Q'^T b \end{bmatrix} - \begin{bmatrix} Rx \\ 0 \end{bmatrix} \right\|^2$$

$$= \left\| \begin{bmatrix} Q^T b - Rx \\ Q'^T b \end{bmatrix} \right\|^2$$

~~$Ax = b$~~
~~$QRx = b$~~

$$= \begin{bmatrix} Q^T b - Rx \\ Q'^T b \end{bmatrix}^T \begin{bmatrix} Q^T b - Rx \\ Q'^T b \end{bmatrix}$$

$$= \begin{bmatrix} (Q^T b - Rx)^T & (Q'^T b)^T \end{bmatrix} \begin{bmatrix} Q^T b - Rx \\ Q'^T b \end{bmatrix}$$

$$= \| Q^T b - Rx \|^2 + \| Q'^T b \|^2$$

This is x independent

$$\underset{x \in \mathbb{R}^n}{\text{Minimize}} \quad \| \underset{n\times 1}{Q^T b} - \underset{n\times 1}{Rx} \|^2$$

$$\underset{m\times n}{A} \underset{n\times 1}{x} = \underset{m\times 1}{b}$$

$R_{n\times n}$ is invertible.

$$\underset{n\times m}{Q^T} \underset{m\times 1}{b}$$

$$Rx = Q^T b \Rightarrow x = R^{-1} Q^T b$$

Unique Sol^n.

$$R_{n\times n} x_{n\times 1}$$

Here we just find $\hat{x}$ s.t.

$$\| A\hat{x} - b \|^2 < \| Ax - b \|^2 \text{ for all } x \in \mathbb{R}^n \setminus \{\hat{x}\}$$

12/03/24

$A = Q_{m\times n} R_{n\times n}$ s.t. $Q^T Q = I_{n\times n}$ & R is an upper Δ^r & invertible.

$$R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & \cdots & r_{2n} \\ \vdots & & & \\ 0 & 0 & \cdots & r_{nn} \end{bmatrix} = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{bmatrix} \begin{matrix} R_1 \\ R_2 \\ \vdots \\ R_n \end{matrix}$$

$$= [q_1 | q_2 | \cdots | q_n] \begin{bmatrix} r_1^T \\ r_2^T \\ \vdots \\ r_n^T \end{bmatrix}$$

$$= \left(q_1 r_1^T + q_2 r_2^T + \cdots + q_n r_n^T \right)$$

$$= [q_1 r_{11} | q_1 r_{12} | \cdots | q_1 r_{1n}] + [0 | q_2 r_{22} | \cdots | q_2 r_{2n}] + \cdots + [0|0|\cdots|q_n r_{nn}]$$

$$\boxed{A = [a_1 | a_2 | \cdots | a_n]}$$

$$= [(r_{11} q_1) | (r_{12} q_1 + q_2 r_{22}) | + \cdots + | (q_1 r_{1n} + q_2 r_{2n} + \cdots + q_n r_{nn})]$$

If we have such a factorization then the kth column of A is $r_{1k} q_1 + r_{2k} q_2 + \cdots + r_{kk} q_k$.

We want, $\| q_i \| = 1$, $\forall i = 1$ to n

$\langle q_i, q_j \rangle = 0$, $i \neq j$

$q_j \in \text{Span}\{a_1, a_2, \ldots, a_{j-1}\}$

$q_j \perp \text{span}\{q_1, q_2, \ldots, q_{j-1}\}$

Since Q is an orthogonal, its columns $q_1, q_2, \ldots$ are orthonormal, so $\| q_1 \| = 1$.

Hence $\| a_1 \| = | r_{11} |$

$a_1 = r_{11} q_1$

$\Rightarrow q_1 = \frac{a_1}{r_{11}}$

and

$a_2 = r_{12} q_1 + r_{22} q_2$

$\Rightarrow q_2 = \frac{a_2 - q_1 r_{12}}{r_{22}}$

$\Rightarrow a_2 = r_{12}\left(\frac{a_1}{\|a_1\|}\right) + r_{22} q_2$

$$\boxed{\therefore q_1 = \frac{a_1}{\| a_1 \|}}$$

$\| q_1 \| = \frac{\| a_1 \|}{| r_{11} |} = 1.$

$\Rightarrow r_{11} = \| a_1 \|$

Q. Given two non-zero vectors a, b. which multiple of a is closet to b?

$\rightarrow$ $\min\limits_{\alpha \in \mathbb{R}^+} \| b - \alpha a \|_2^2$

$\| b - \alpha a \|^2 = \langle b - \alpha a, b - \alpha a \rangle$

$$= \alpha^2 a^T a - \alpha a^T b - \alpha b^T a + b^T b$$

$$= f(\alpha) \quad (\text{say})$$

$$f'(\alpha) = 2\alpha a^T a - a^T b - b^T a$$

$$= 2\alpha a^T a - 2 a^T b = 0$$

To minimize $f'(\alpha) = 0$

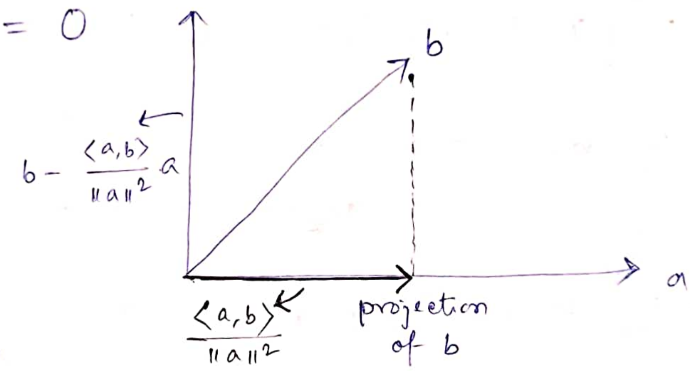
$$\Rightarrow \alpha = (a^T b)(a^T a)^{-1}$$

$$= \frac{a^T b}{a^T a}$$

$$= \frac{\langle a, b\rangle}{\|a\|^2} \rightarrow \text{Global min.}$$

$\therefore \frac{\langle a,b\rangle}{\|a\|^2} \cdot a$ is closet to b.

$$\left\langle a,\ b - \frac{\langle a,b\rangle}{\|a\|^2} \cdot a \right\rangle = 0$$

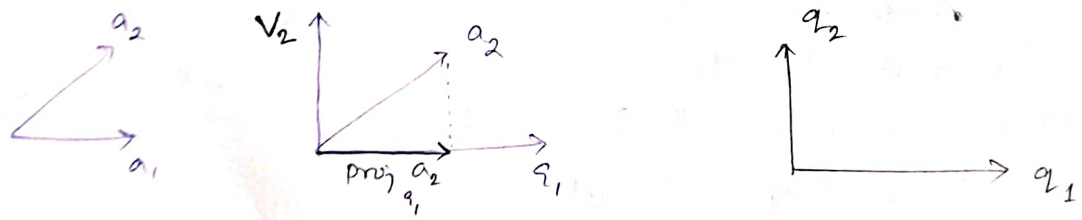


Projection of $\vec{b}$ on $\vec{a}$

$$= \frac{\langle a,b\rangle}{\|a\|^2} \cdot \vec{a}$$

If $\|a\| = 1$ then projection of $\vec{b}$ on $\vec{a}$ $= \langle a,b\rangle \cdot \vec{a}$

In general, projection of $\vec{b}$ on Span $\{a_1, a_2, \dots, a_n\} = S$ is $\text{Proj}_S\, b = \langle a_1, b\rangle \vec{a}_1 + \langle a_2, b\rangle \vec{a}_2 + \dots + \langle a_n, b\rangle \vec{a}_n$



$$a_1 = \overbrace{\|a_1\|}^{r_{11}} q_1$$

$$a_2 = \underbrace{\langle a_2, q_1\rangle}_{r_{12}} q_1 + \underbrace{\|a_2 - \langle a_2, q_1\rangle q_1\|}_{r_{22}} q_2$$

Write $V_3 = a_3 - \langle a_3, q_1\rangle q_1 + \langle a_3, q_2\rangle q_2$

$$a_3 = \underbrace{\langle a_3, q_1\rangle}_{r_{13}} q_1 + \underbrace{\langle a_3, q_2\rangle}_{r_{23}} q_2 + \underbrace{\|V_3\|}_{r_{33}} q_3 \quad \text{so on.}$$

Classical Gram-Schmidt

Given $A_{m\times n}$, $m \geq n$, rank$(A) = n$.

$A = [a_1 \mid a_2 \mid \dots \mid a_n]$

for $j = 1$ to n:

$\quad v_j = a_j$

$\quad$ for $i = 1$ to $j-1$:

$\qquad r_{ij} = \langle q_i, a_j \rangle$

$\qquad v_j = v_j - r_{ij} \cdot q_i$

$\quad$ end

$\quad r_{jj} = \| v_j \|_2$

$\quad q_j = \frac{1}{r_{jj}} v_j$

end.

e.g. (1) $A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}$

$q_1 = a_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$

$a_2 = r_{12} q_1 + r_{22} q_2$

$= 1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \sqrt{2} \, \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \rightarrow V_2 = a_2 - r_{12} q_1$

$r_{12} = \langle q_1, a_2 \rangle = \langle a_1, a_2 \rangle = 1.$

$V_2 = a_2 - r_{12} q_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - 1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$

$\rightarrow r_{22} = \| V_2 \| = \sqrt{2}$

$\rightarrow q_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$

$\rightarrow r_{13} = \langle q_1, a_3 \rangle = 1$

$r_{23} = \langle q_2, a_3 \rangle = \frac{1}{\sqrt{2}}$

$V_3 = a_3 - r_{13} q_1 - r_{23} q_2$

$= \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} - 1 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$

$= \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$

$= \begin{pmatrix} 0 \\ 1/2 \\ -1/2 \end{pmatrix}$

$r_{33} = \| V_3 \| = \frac{1}{\sqrt{2}}$

$q_3 = \sqrt{2} \begin{pmatrix} 0 \\ 1/2 \\ -1/2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix}$

$$\therefore A = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/\sqrt{2} & 1/\sqrt{2} \\ 0 & 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}}_{Q} \underbrace{\begin{bmatrix} 1 & 1 & 1 \\ 0 & \sqrt{2} & 1/\sqrt{2} \\ 0 & 0 & 1/\sqrt{2} \end{bmatrix}}_{R}$$

(2) let $\varepsilon << 0$ be very small (+)ve no: s.t.

$fl(\varepsilon^2) = 0$

$$a_1 = \begin{pmatrix} 1 \\ \varepsilon \\ 0 \\ 0 \end{pmatrix} \quad a_2 = \begin{pmatrix} 1 \\ 0 \\ \varepsilon \\ 0 \end{pmatrix} \quad a_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \varepsilon \end{pmatrix}$$

$\rightarrow r_{11} = \| v_1 \| = \| a_1 \| = \sqrt{1+\varepsilon^2} \approx 1$

\# $q_1 = a_1$

$\rightarrow r_{12} = \langle q_1, a_2 \rangle = 1$, $r_{22} = \| v_2 \| = \sqrt{2\varepsilon^2} \approx 0$
$\approx \sqrt{2}\,\varepsilon$.

$v_2 = a_2 - r_{12} \cdot q_1$

$\rightarrow r_{13} = \langle q_1, a_3 \rangle = 1$.

$$= \begin{pmatrix} 1 \\ 0 \\ \varepsilon \\ 0 \end{pmatrix} - 1 \begin{pmatrix} 1 \\ \varepsilon \\ 0 \\ 0 \end{pmatrix}$$

$\rightarrow r_{23} = \langle q_2, a_3 \rangle = 0$

$\rightarrow r_{33} = \| v_3 \| = \sqrt{2\varepsilon^2} = \sqrt{2}\,\varepsilon$

$$= \begin{pmatrix} 0 \\ -\varepsilon \\ \varepsilon \\ 0 \end{pmatrix}.$$

$v_3 = a_3 - r_{13}\, q_1 - r_{23} \cdot q_2$

\# $q_2 = \dfrac{v_2}{r_{22}} = \dfrac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \end{pmatrix}$

$$= \begin{pmatrix} 1 \\ 0 \\ 0 \\ \varepsilon \end{pmatrix} - \begin{pmatrix} 1 \\ \varepsilon \\ 0 \\ 0 \end{pmatrix} - 0$$

$$= \begin{pmatrix} 0 \\ -\varepsilon \\ 0 \\ \varepsilon \end{pmatrix}$$

\# $q_3 = \dfrac{v_3}{r_{33}} = \dfrac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ -1 \\ 0 \\ 1 \end{pmatrix}$

$$Q = \begin{bmatrix} 1 & 0 & 0 \\ \varepsilon & -1/\sqrt{2} & -1/\sqrt{2} \\ 0 & 1/\sqrt{2} & 0 \\ 0 & 0 & 1/\sqrt{2} \end{bmatrix} \quad R = \begin{bmatrix} 1 & 1 & 1 \\ 0 & \sqrt{2}\,\varepsilon & 0 \\ 0 & 0 & \sqrt{2}\,\varepsilon \end{bmatrix}$$

Given $\{a_1, a_2, \dots, a_n\}$ be linearly independent vectors in $\mathbb{R}^m$. 14/03/24

$$q_1 = \frac{a_1}{\| a_1 \|}$$

$$v_k = a_k - \overbrace{\langle q_1, a_k \rangle}^{q_1^T a_k} q_1 - \dots - \overbrace{\langle q_{k-1}, a_k \rangle}^{q_{k-1}^T a_k} q_{k-1}$$

$$q_k = \frac{v_k}{\| v_k \|} = \frac{v_k}{r_{kk}}$$

$(I - \hat{Q}\hat{Q}^T) a_k$

$\hat{Q}\hat{Q}^T = \underbrace{[q_1 | q_2 | \dots | q_{k-1}]}_{m \times k-1} \underbrace{\begin{bmatrix} q_1^T \\ q_2^T \\ \vdots \\ q_{k-1}^T \end{bmatrix}}_{k-1 \times m} \underbrace{a_k}_{m\times 1}$

$= \langle q_1 . a_k \rangle q_1$

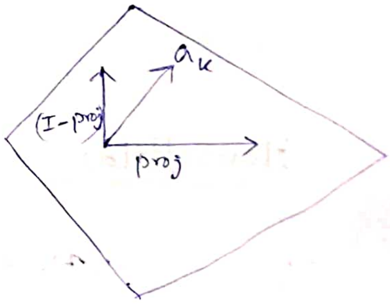
Def^n : $k-1 = 1$ to $n-1$

$$\hat{Q}_{k-1} = [q_1 | q_2 | \dots | q_{k-1}]$$

$$P_k = I - \hat{Q}_{k-1} \hat{Q}^T_{k-1}$$

Compact form

$$\boxed{\begin{aligned} v_k &= P_k . a_k \\ q_k &= \frac{P_k a_k}{\| P_k a_k \|} \quad \text{for } k = 1 \text{ to } n \end{aligned}}$$



Verify :

① $P_{1q} = I - \underbrace{q}_{m\times 1} q^T$ where q is a norm 1 vector.

rank $(P_{1q}) = m - 1$

② $P_k = P_{1q_{k-1}} \dots P_{1q_2} P_{1q_1}$.

③ OC (Gram-Smidth) $\simeq 2mn^2$

Classical

```
for j = 1 to n
    v_j = a_j
    for i = 1 to j-1
        r_ij = q_i^T . a_j
        v_j = v_j - r_ij . q_i
    end
    r_jj = || v_j ||^2
    q_j = v_j / r_jj
end
```

Modified

```
for i = 1 to n
    v_i = a_i
end
for i = 1 to n
    r_ii = || v_i ||
    q_i = v_i / r_ii
    ~~end~~
    (nested) → for j = i+1 to n
        r_ij = q_i^T . v_j
        v_j = v_j - r_ij . q_i
    end
end
```

Combined

for $j = 1$ to n

$v_j = a_j$

for $i = 1$ to $j-1$ (nested)

$$r_{ij} = \begin{cases} q_i^T a_j & \text{classical} \\ q_i^T v_j & \text{modified} \end{cases}$$

$$v_j = v_j - r_{ij} \cdot q_i$$

end

$r_{jj} = \| v_j \|$

$$q_j = \frac{v_j}{r_{jj}}$$

end.

Householder's Method :

$H_k \cdots H_2 H_1 A = \tilde{R}$

Each H_i is orthogonal & zero-introducing.

$$\tilde{R} = \left[\begin{array}{c} \begin{matrix} & * \\ 0 & \end{matrix} \\ \hline 0 \end{array}\right] = \left[\begin{matrix} & * \\ 0 & \end{matrix}\right]$$

$$A = [\ *\ ] \xrightarrow{H_1 A} \left[\begin{array}{c|ccc} * & * & \cdots & \\ \hline 0 & & & \\ 0 & & * & \\ \vdots & & & \\ 0 & & & \end{array}\right] \xrightarrow{H_2 H_1 A} \left[\begin{array}{cc|c} * & * & * \\ \hline 0 & * & \\ \vdots & 0 & * \\ & \vdots & \\ 0 & 0 & \end{array}\right] \xrightarrow{H_3 H_2 H_1 A} \left[\begin{array}{cccc} * & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & 0 & * \cdots & * \\ \hline & & 0 & \\ 0 & 0 & \vdots & * \end{array}\right]$$

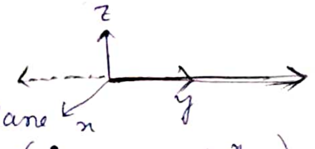
and so on.

Def^n of Reflection Matrix :

$H_{m \times m} : \mathbb{R}^m \longrightarrow \mathbb{R}^m$ if $\exists$ an $(n-1)$ dimensional subspace U with $Hu = u$

$\forall u \in U$ & $Hv = -v$ $\forall v \in U^T$

e.g. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$



Any points on the xz-plane remains same
But the points along y axis ... other side
$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \longmapsto \begin{pmatrix} x \\ -y \\ z \end{pmatrix}$$

☑ ☑ $\|u\| = 1$

$$H_u = I - 2uu^T$$

then $H_u(u) = (I - 2uu^T)(u) = u - 2u(\underline{u^T u})$

$$= u - 2u$$

$$= -u$$

Let $v \in U^{\perp}$

$$(H_u)(v) = (I - 2uu^T)(v) = v - 2u(\underline{u^T v})$$

$$= v$$

<u>QR Using HouseHolders reflection :</u> <u>21/03/24</u>

A has full col rank.

let $v \in \mathbb{R}^m \setminus \{0\}$ $\quad u = \frac{v}{\|v\|}$

<u>Define</u> $H_v = I - 2\frac{vv^T}{v^Tv} = I - 2uu^T$

$$H_v^T = (I - 2uu^T)^T = I - 2(u^T)^T u^T = I - 2uu^T = H_v$$

$\therefore H_v$ is asymmetric matrix.

$$H_v^T H_v = H_v H_v^T$$

$$= (I - 2uu^T)(I - 2uu^T) \qquad \boxed{A^TA = I}$$

$$= I - 4uu^T + 4u\underline{u^T u}u^T$$

$$= I - 4uu^T + 4uu^T$$

$$= I$$

$\therefore H_v$ is orthogonal.

☑ $H_v v = (I - 2uu^T).v = Iv - 2uu^T.v$

$$= v - \frac{2vv^T}{v^Tv}.v$$

$$= -v$$

$\Rightarrow H_v$ acts on $\times(-1)$ on $\text{span}\{v\}$

☑

$w \in v^{\perp}$ (i.e, the hyperplane $\perp^r$ to v)

$H_v(w) = (I - 2uu^T)w = w - 2uu^T.w = w - 2u(u^Tw)$

$\Rightarrow H_v$ acts an identity on the hyperplane $v^\perp$

From the both cases, H_v is a reflection matrix.

Q. Given $x, y \in \mathbb{R}^m \setminus \{0\}$ s.t. $\|x\| = \|y\|$

Does $\exists\, v$ s.t. $H_v(x) = y$?

$\Rightarrow$ Let $v = x - y$

$$H_v(x) = \left(I - 2\frac{(x-y)(x-y)^T}{(x-y)^T(x-y)} \right) x$$

$$= x - 2\frac{(x-y)(x-y)^T}{x^Tx - 2x^Ty + y^Ty} x$$

$$= x - 2\frac{(x-y)(x^Tx - y^Tx)}{2(x^Tx - x^Ty)}$$

$\left[\because \|x\| = \|y\| \quad x^T = y^T \right]$

$\& \; x^Ty = y^Tx$

$$= y$$

so possible.

☑ Given $x \in \mathbb{R}^m \setminus \{0\}$

Let $v = x - (\pm \|x\|) e_1$ $\qquad x = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}$, $H_v(x) = \begin{pmatrix} \pm\|x\| \\ 0 \\ \vdots \\ 0 \end{pmatrix}$

$\Leftrightarrow \; y = (\pm \|x\|) e_1$

$H_v(x) = \pm \|x\| e_1$

☑ $A_{m\times 3}$ $\quad [a_1 \mid a_2 \mid a_3]$ $\qquad$ rank $(A) = 3$.

$v_1 = a_1 - \|a_1\| e_1 \quad \in \mathbb{R}^m$

$H_1 = H_v$

$$H_1 A = \left[\begin{array}{c|c|c} \|a_1\| & * & * \\ \hline 0 & & \\ \vdots & b_{m-1\times 1} & * \\ 0 & & \end{array}\right]$$

$v_2 = b - \|b\| e_1 \quad \in \mathbb{R}^{m-1}$

$$H_2 = \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & \\ \vdots & H_{v_2} \\ 0 & \end{array}\right]$$

$$H_2 H_1 A = \left[\begin{array}{cc|c} \|a_1\| & + & * \\ 0 & \|b\| & * \\ \hline 0 & 0 & \\ \vdots & \vdots & C_{m-2\times 1} \\ 0 & 0 & \end{array}\right]$$

$$v_3 = c - \|c\| e_1 \in \mathbb{R}^{m-2}$$

$$H_3 = \left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & \\ \hline 0 & & H_{v_3} \end{array}\right]$$

$$H_3 H_2 H_1 A = \underbrace{\begin{bmatrix} \|a\| & + & * \\ 0 & \|b\| & * \\ 0 & 0 & \|e\| \\ \vdots & \vdots & \vdots \\ 0 & 0 & 0 \end{bmatrix}}_{R}$$

$\therefore\ H_3 H_2 H_1 = Q^T$

so $Q^T A = R \quad \Rightarrow A = Q_{m\times m} R_{m\times 3}$

<u>Algorithm</u>:

Input : $A_{m\times n}$

for $K = 1$ to n :

$x = A_{(K:m),K}$

$v_K = sgn(x) \cdot \|x\| e_1 + x$

$v_K = \dfrac{v_K}{\|v_K\|}$

$A_{(K:m),(K:n)} = A_{((K:m),(K:n))} - 2 v_K \left(v_K^T A_{((K:m),(K:n))}\right)$

end

<u>Def^n</u>: let $x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}$

$sgn(x) = sgn(x_1)$

$Ax = b$

$\Rightarrow QRx = b$

$\Rightarrow Rx = Q^T b$

$x = R^{-1} Q^T b$

" b is a col space of A " is independent.

$\min \|Ax - b\|$

$\equiv Rx = Q^T b$

Calculate $Q^T b$

for $k = 1$ to n :
$b_{k:m} = b_{k:m} - 2 v_k \langle v_k, b_{k:m} \rangle$

end

$H_n \ldots H_2 H_1 A x = H_n \ldots H_2 H_1 b$ $\quad (Rx = Q^T b)$

Operation Counts

$O(2mn^2 - \frac{2}{3} n^3)$

QR by Given's rotation OP. count $= O(3mn^2 - n^3)$

Full QR, $A = QR$ (Householders)

Reduced QR, $A = QR$ (Gram-Smidth)

Def^n : Given an $n \times n$ matrix (over $\mathbb{R}$ / $\mathbb{C}$)
a (real / complex) numbers λ is an eigen value of A if $\exists$ $x \in \mathbb{R}^n \setminus \{0\}$ s.t. $Ax = \lambda x$

$\left(\equiv x \longmapsto \text{span}\{x\} \right)$

Here x is called eigen vector of A corresponding to λ.

$A : \mathbb{R}^n \to \mathbb{R}^n$

Suppose A has distinct non-zero e-values $\lambda_1, \lambda_2, \ldots \lambda_n$ & $Ax_1 = \lambda x_1$

$\{x_1, x_2, \ldots, x_n\}$ are linearly independent

$\Rightarrow \{x_1, x_2, \ldots, x_n\}$ is a basis of $\mathbb{R}^n$.

Note : $\quad Ax = \lambda x$

$$
\begin{aligned}
A^2 x = A(Ax) &= A(\lambda x) \\
&= \lambda (Ax) \\
&= \lambda(\lambda x) \\
&= \lambda^2 x
\end{aligned}
$$

In general $A^k x = \lambda^k x$

If $\lambda \neq 0$ & A^{-1} exists $\quad A^{-1} x = \lambda^{-1} x$

$Ax = \lambda x$

$\Rightarrow x = A^{-1}\lambda x = \lambda (A^{-1}x)$

$\Rightarrow A^{-1}x = \lambda^{-1}x$

Claim : If A has n distinct non-zero e-values $\lambda_1, \dots, \lambda_n$ then corresponding e-vectors are linearly independent

$c_1x_1 + c_2x_2 + \dots + c_nx_n = 0$ $\quad c_i = 0, \forall i$

$A(c_1x_1 + \dots + c_nx_n) = A\,0 = 0$

☑ Assume $x_1, x_2, \dots, x_n$ are linearly dependent.

Let $\{x_1, x_2, \dots, x_k\}$ be a minimal dependent sub-set s.t.

$\Rightarrow$ we have $c_1, c_2, \dots, c_k$ not all zero

$c_1x_1 + c_2x_2 + \dots + c_kx_k = 0$

$A\left(\sum_{i=1}^{k} c_ix_i\right) = 0$

$c_1\lambda_1 x_1 + x_2 \dots + \lambda_n x_n = 0$

$\Rightarrow \sum_{i=1}^{k} c_i\lambda_i x_i = 0$

$\Rightarrow \sum_{i=2}^{k} c_i(\lambda_i - \lambda_1)x_i = 0$

$\Rightarrow \{x_2, x_3, \dots, x_k\}$ is a dependent subset, which is a contradiction.

26/03/24

$A_{n\times n} \in M_{n\times n}(\mathbb{R})$

$Av = \lambda v, \quad v \neq 0$

$Iv = v \longrightarrow$ Here 1 is an eigen-value which is counted n times.

$Av = \lambda v = \lambda I v$

$\Rightarrow (A - \lambda I)v = 0 \longrightarrow$ A system of homogeneous linear eqn.

$v \in \mathbb{R}^n \setminus \{0\}$ & $A - \lambda I$ is a matrix

$\Rightarrow v \in \ker(A - \lambda I)$

$\Rightarrow \ker(A - \lambda I)$ is non-trivial

$\Leftrightarrow \dim(\ker(A - \lambda I)) > 0$

$\Leftrightarrow$ As a linear map (i.e. $A - \lambda I : \mathbb{R}^n \longrightarrow \mathbb{R}^n$)

$A - \lambda I$ is not injective [$\because$ Ker is not trivial]

$\Rightarrow A - \lambda I$ is not invertible

$\Rightarrow |A - \lambda I| = 0$

$$\Rightarrow \det \begin{bmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{bmatrix} = 0$$

The co-factor expansion gives us a degree n monic poly in λ, with real co-efficients

$$\chi_A(\lambda) = \lambda^n + c_{n-1}\lambda^{n-1} + \cdots + c_1\lambda + c_0 = 0.$$

✓ The characteristic poly. of A

Theorem: A degree n polynomial has exactly n roots counted upto multiplicity.

Note: $\chi_A(\lambda = A) = A^n + c_{n-1}A^{n-1} + \cdots + c_1 A + c_0 I = 0$

The minimal polynomial of A is the minimum degree poly. satisfied by A & it divides any other poly that A satisfies.

$\chi_A(x) = (x - \lambda_1)(x - \lambda_2) \cdots (x - \lambda_n)$ where λ_i's are e-values of A.

Constant term : $(-1)^n \lambda_1 \lambda_2 \cdots \lambda_n = \det(A)$

Coefficient of degree 1 term $:= -(\lambda_1 + \lambda_2 + \cdots + \lambda_n)$
$= -\text{trace}(A)$

e.g. $A_{2\times 2}$

$= \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

$$|A - \lambda I| = \begin{vmatrix} a-\lambda & b \\ c & d-\lambda \end{vmatrix} = 0$$

$$\Rightarrow \lambda^2 - \underbrace{(a+d)}_{tr(A)}\lambda + \underbrace{ad - bc}_{\det(A)} = 0$$

e.g. $A_{2\times 2} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$

$\chi_A(\lambda) = \det \begin{bmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{bmatrix} = 0 \quad \Rightarrow \lambda^2 - 4\lambda + 3 = 0$

$\Rightarrow (\lambda-3)(\lambda-1) = 0.$

Let $\lambda = 1$

$(A - I)\,v = 0$

$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

$\Rightarrow x + y = 0$

$\Rightarrow x = -y$

$\Rightarrow \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ is an eigen vector.

We can say that, The Span$\left\{\begin{pmatrix} 1 \\ -1 \end{pmatrix}\right\}$ is the eigen space corresponding to $\lambda = 1$.

$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ -1 \end{pmatrix} = 1\begin{pmatrix} 1 \\ -1 \end{pmatrix}$

Let $\lambda = 3$

$(A - 3I)\,v = 0 \quad \Rightarrow \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

$\Rightarrow -x + y = 0 \;\&\; x - y = 0$

$\Rightarrow x = y$

$\Rightarrow \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ is an eigen vector.

Similarly, The Span $\left\{\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right\}$ is the eigen space corresponding to $\lambda = 3$.

$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \end{pmatrix} = 3\begin{pmatrix} 1 \\ 1 \end{pmatrix}$

Note $\left\{\underset{u_1}{\begin{pmatrix} 1 \\ 1 \end{pmatrix}}, \underset{u_2}{\begin{pmatrix} 1 \\ -1 \end{pmatrix}}\right\}$ is a basis of $\mathbb{R}^2$. & $A : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$

$\mathbb{R}^2 = \{c_1 u_1 + c_2 u_2 \mid c_1, c_2 \in \mathbb{R}\}$

$A(c_1 u_1 + c_2 u_2) = c_1 A u_1 + c_2 A u_2 = c_1 \underline{u_1} + c_2 \underline{3u_2}$

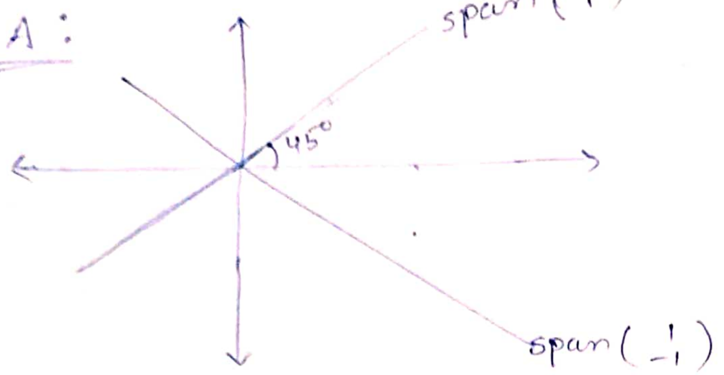
$\therefore$ In $\{u_1, u_2\}$ basis, the linear map has the following representation $\begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$

$\boxed{\text{H}}$ $Q = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ $|Q| = 1.$

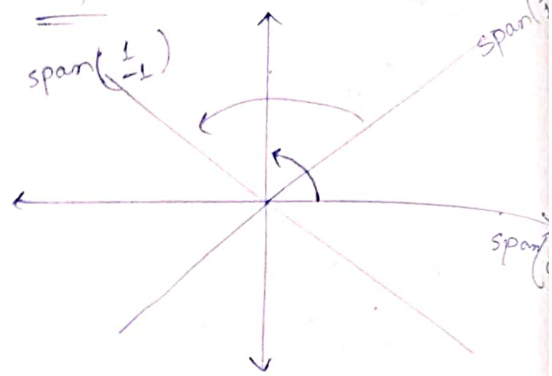
$|Q - \lambda I| = \begin{vmatrix} -\lambda & -1 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 + 1 = 0$

$Qv = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} -v_2 \\ v_1 \end{pmatrix}$

$\underline{A:}$



$\underline{Q:}$



$Q: \quad \text{span}\begin{pmatrix}1\\0\end{pmatrix} \rightleftarrows \text{span}\begin{pmatrix}0\\1\end{pmatrix}$

$\text{span}\begin{pmatrix}1\\1\end{pmatrix} \rightleftarrows \text{span}\begin{pmatrix}1\\-1\end{pmatrix}$

It moves every subspace of $\mathbb{R}^2$, Hence Q cannot have real eigen values.

Q has e-values $\pm i$.

$\underline{\lambda = i}$ $\quad \left[\begin{array}{cc|c} -i & -1 & 0 \\ 1 & -i & 0 \end{array}\right] \xrightarrow{R_2 - iR_1} \left[\begin{array}{cc|c} -i & -1 & 0 \\ 0 & 0 & 0 \end{array}\right]$

The eigen-vector $\begin{pmatrix}1\\-i\end{pmatrix}$

$Ax = \lambda$
$(A - \lambda I)x$

$\underline{\lambda = (-i)}$ $\begin{pmatrix} i & -1 \\ 1 & i \end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix}$ $\qquad$ eigen-vector $\begin{pmatrix}1\\i\end{pmatrix}$

$(Q - \lambda I)x$

$\Rightarrow x_1 i - x_2 = 0$

$\Rightarrow x_1 + i x_2 = 0$

$\underline{\text{e.g.}}$ ① $A = \begin{bmatrix} 8 & 3 \\ 2 & 7 \end{bmatrix}$ $\qquad \lambda_1 = 10, \quad \lambda_2 = 5$

$v_1 = \begin{bmatrix}3\\2\end{bmatrix}, \quad v_2 = \begin{bmatrix}1\\-1\end{bmatrix}$

They are l. i. but not orthogonal

② $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$

$\chi_A(\lambda) = \lambda^2 = 0 \Rightarrow \lambda = 0, 0$

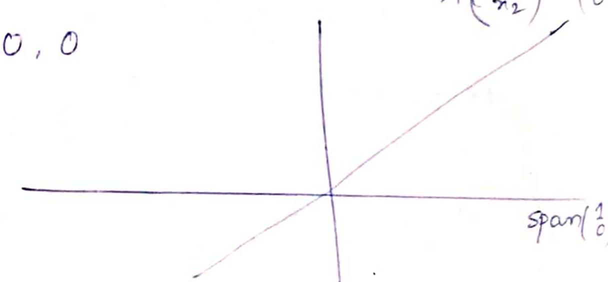
$(A - 0I)x = \begin{pmatrix}0\\0\end{pmatrix}$

$\Rightarrow \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix}$

$\Rightarrow x_2 = 0$

The only e-vector is $\begin{pmatrix}1\\0\end{pmatrix}$.

$A\begin{pmatrix}x_1\\x_2\end{pmatrix} = \begin{pmatrix}x_2\\0\end{pmatrix}$



Every 1-dim subspace is mapped to span$\left\{\begin{pmatrix}1\\0\end{pmatrix}\right\}$

$\boxed{1}$ $Av = \lambda v$

P is an invertible matrix

$(PAP^{-1})(Pv) = PAv = P(\lambda v) = \lambda(Pv)$

$\Rightarrow \lambda$ is an e-value of PAP^{-1} with Pv as corresponding eigen vector.

We say that A & PAP^{-1} is similar

$\{PAP^{-1} \mid P$ lies over all $n \times n$ invertible matrix $\}$

$\boxed{28/03/24}$

Suppose A has n linearly independent eigen-vectors $v_1, v_2, \dots, v_n$

$$P = [v_1 \mid v_2 \mid \dots \mid v_n]$$

$$AP = [Av_1 \mid Av_2 \mid \dots \mid Av_n]$$

$$= [\lambda_1 v_1 \mid \lambda_2 v_2 \mid \dots \mid \lambda_n v_n]$$

$$= P \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix}$$

$AP = P\Lambda \Rightarrow A = P\Lambda P^{-1}$

$\Rightarrow A$ is similar to a diagonal matrix, with e-values on the diagonal.

$$A^k = P\Lambda^k P^{-1}$$

Geometric Multiplicity of an e-value $(GM(\lambda))$: dim Null $(A - \lambda I)$.

Algebraic " " $(AM(\lambda))$: Multiplicity of λ as a root of $\chi_A(x)$

In general, $AM(\lambda) \geq GM(\lambda)$.

e.g. $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ $\lambda = 0, 0$ $\cancel{\chi_A(\lambda)} = \boxed{\lambda^2 = 0}$

$\chi_A(x) = x^2$

$AM(0) = 2$

$GM(0) = 1$

The Spectral Thm : Let S be a sym. $n \times n$ real matrix. Then all the n eigen values of S are real numbers, the n eigen vectors can be choosen to be orthogonal ($\Rightarrow$ eigen vectors are l. i.)

$$S = Q\Lambda Q^T \quad \text{where} \quad QQ^T = Q^TQ = I$$

Theorem: For a sym. $n \times n$ matrix S, TFAE

① S is PD

② $\text{Spec}(S) \subseteq \mathbb{R}_{>0}$ (all e-values are non-neg.)

③ $\exists A$ with n linearly independent columns s.t. $S = A^T A$

④ All the leading principal minors are (+)ve.

⑤ All the pivots in G.E. are (+)ve.

Proof: ① => ②

Let S be P.D. & $\lambda \in \text{Spec}(S)$

$\Rightarrow Sx = \lambda x$; x is corresponding e-vector to λ

$$\Rightarrow x^T S x = x^T \lambda x$$
$$= \lambda \|x\|^2 > 0$$

$$\Rightarrow \lambda > 0.$$
$$\Rightarrow \text{Spec}(S) \subseteq \mathbb{R}_{>0}$$

② => ① Since S is sym., it has n orthogonal eigen-vectors which form a basis for $\mathbb{R}^n$.

$$S = Q \Lambda Q^T, \quad Q = [v_1 | v_2 | \cdots | v_n]$$

Any $x \in \mathbb{R}^n \setminus \{0\}$ can be written as $x = \sum_i c_i v_i$

$$x^T S x = \left(\sum c_i v_i\right)^T S\left(\sum c_i v_i\right)$$
$$= \left(\sum c_i v_i\right)^T \left(\sum \lambda_i c_i v_i\right)$$
$$= \sum v_i^T v_i \, c_i^2 \lambda_i = \sum c_i^2 \lambda_i \|v_i\|^2 > 0.$$

③ => ① $S = A^T A$

$$\Rightarrow x^T S x = x^T A^T A x = (Ax)^T Ax = \|Ax\|^2 > 0$$

Since A has linearly ind. columns, $x \neq 0$ so $Ax \neq 0$

① => ③ Choleskey Factorization.

Digression: The 2×2 case $S = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$ sym.

$a > 0$, $ac - b^2 > 0 \longrightarrow$ PD

eigen-values $= \frac{1}{2}\left[(a+c) \pm \sqrt{(a+1)^2 - 4(ac-b^2)}\right] >$

Pivots : $a > 0$, $\frac{ac-b^2}{a} > 0$

The associated quadratic form $ax^2 + 2bxy + cy^2 > 0$

$F : \mathbb{R}^2 \to \mathbb{R}$ has a global min. at (x_0, y_0)

$\frac{\partial F}{\partial x} = \frac{\partial F}{\partial y} = 0$ at (x_0, y_0)

The matrix $\begin{bmatrix} \frac{\partial^2 F}{\partial x^2} & \frac{\partial^2 F}{\partial x \partial y} \\ \frac{\partial^2 F}{\partial y \partial x} & \frac{\partial^2 F}{\partial y^2} \end{bmatrix}$ is P.D.

$A \in M_{m \times n}(\mathbb{R})$

$F_A : \mathbb{R}^n \setminus \{0\} \longrightarrow \mathbb{R}_{\geq 0}$

$F_A(v) = \frac{\|Av\|_2^2}{\|v\|_2^2} = \frac{v^T A^T A v}{v^T v}$

$\boxed{\begin{array}{l} \alpha \in \mathbb{R} \setminus \{0\} \\ F_A(\alpha v) = F_A(v) \end{array}}$

WLOG $\|v\| = 1$

$F_A(v) = \|Av\|^2$, $\forall v$ s.t. $\|v\| = 1$.

$= v^T A^T A v$

Use Lagrange multipliers,

$\mathcal{L}(\lambda, v) = F_A(v) - \lambda(v^T v - 1)$

$\frac{\partial \mathcal{L}}{\partial \lambda} = -(v^T v - 1) = 0$ and $\frac{\partial \mathcal{L}}{\partial v} = A^T A v - \lambda v = 0$.

The critical pt. are $(A^T A) v = \lambda v$

$v^T A^T A v \geq 0$

$\Rightarrow$ eigen values are $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n > 0$

For $i = 1$ to n : Define $u_i = A v_i$

$\lambda_i u_i = \lambda_i A v_i = A(\lambda_i v_i)$

$= A(A^T A v_i)$

$= (AA^T)(A v_i) = (AA^T)(u_i)$

$\Rightarrow \lambda_1, \lambda_2, \ldots, \lambda_n$ are e-values of AA^T with corresponding e-vectors $Av_1, \ldots, Av_n$

$\|u_i\|_2 = \frac{\|Av_i\|_2}{\cancel{\|v_i\|_2}} = \sqrt{\|Av_i\|^2} = \sqrt{v_i^T A^T A v_i} = \sqrt{v_i^T \lambda_i v_i}$

$= \sqrt{\lambda_i}\, \|v_i\|_2$

$A_{m\times n}$, $m \geq n$, rank $(A) \leq n$ 02/04/24

TFAE

1. $S_{n\times n}$ is symmetric (+)ve semi-definite
$(x^T S x \geq 0 \ , \ \forall x \in \mathbb{R}^n \setminus \{0\})$

2. The e-values are non-negative
Spec $(S) \subset \mathbb{R}_{\geq 0}$
$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$

$Sv = \lambda v$
$\lambda \|v\|^2 = v^T S v \geq 0$
$\Rightarrow \lambda \geq 0$

3. $\exists$ $A_{m\times n}$ s.t. $S = A^T A$

$A^T = n \times m$
$(A^T A)_{n\times n}$

$x^T S x = \|Ax\|^2 \geq 0$

$\boxed{}$ The real quantity $v^T A^T A v$ & $\|v\|_2 = 1$ is maximized when v is an e-vector of $A^T A$.

$\boxed{}$ $(A^T A)_{n\times n}$ is symmetric (+)ve semi-definite
$A^T A$ has n linearly independent e-vectors $v_1, v_2, \ldots, v_n$
& corr. e-values $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$

$\boxed{}$ $f_A : \mathbb{R}^n \longrightarrow \mathbb{R}$
$v \longmapsto v^T A^T A v$

$f_A(v_i) = v_i^T (A^T A v_i) = v_i^T \lambda_i v_i = \lambda_i \|v_i\|^2$

$\boxed{}$ Define for $i = 1$ to n, $u_i = Av_i$

$$\begin{aligned}
\lambda_i u_i = \lambda_i A v_i &= A(\lambda_i v_i) \\
&= A(A^T A v_i) \\
&= (AA^T)(Av_i) \\
&= (AA^T) u_i
\end{aligned}$$

$\Rightarrow \lambda_i$ is an e-value of AA^T with $u_i (\equiv Av_i)$ as the corresponding e-vector.

So $\lambda_1, \lambda_2, \ldots, \lambda_n$ are e-values of AA^T corres. e-vectors $Av_1, Av_2, \ldots, Av_n$

$\|u_i\|_2 = \|Av_i\|$

$= \sqrt{\|Av_i\|^2}$

$= \sqrt{v_i^T A^T A v_i}$

$= \sqrt{v_i^T \lambda_i v_i} = \sqrt{\lambda_i}\, \|v_i\|_2$

Note : If u is an e-vector of AA^T then $A^T A\, A^T u$ is either 0 or an e-vector of $A^T A$.

Let $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_k > 0$

$\lambda_{k+1} = \lambda_{k+2} = \dots = \lambda_n = 0$

for $i = 1$ to k $\qquad AA^T := u \qquad \underline{v = A^T u}$

$A^T A := v$

$A^T A v_i = \lambda_i v_i$

$AA^T u_i = \lambda_i u_i$

let $Av_i = \sqrt{\lambda_i}\, u_i$ $\quad$ & $\|u_i\| = \|v_i\| = 1$

$\tilde{U}_{m\times k} = [u_1 | u_2 | \dots | u_k]$ $\qquad \tilde{V}_{n\times k} = [v_1 | v_2 | \dots | v_k]$

e_i be the ith unit vector in $\mathbb{R}^k$

$\underset{k\times m\ \ m\times n\ \ n\times k}{\tilde{U}^T A \tilde{V}}\, e_{i\,k\times 1} = \tilde{U}^T A v_i$

$= \frac{1}{\lambda_i} \tilde{U}^T A (\lambda_i v_i)$

$= \frac{1}{\lambda_i} \tilde{U}^T A (A^T A v_i)$

$= \frac{1}{\lambda_i} \tilde{U}^T (AA^T)(Av_i)$

$= 1/\lambda_i\ \tilde{U}^T (\underline{AA^T})(\sqrt{\lambda_i} u_i)$

$= \cancel{\lambda_i} \sqrt{\lambda_i}\, \tilde{U}^T u_i$

$= \sqrt{\lambda_i}\, e_i$ $\qquad \tilde{\Sigma} = \begin{bmatrix} \sqrt{\lambda_1} & 0 \\ 0 & \sqrt{\lambda_k} \end{bmatrix}$

$\tilde{U}^T A \tilde{V} = \tilde{\Sigma}$

nullity $(A^T A) = n - k$

Choose an orthonormal basis $\{v_{k+1}, v_{k+2}, \dots v_n\}$ of Null $(A^T A)$ & form $V_{n\times n} = [\tilde{V} | v_{k+1} | \dots | v_n]$

Similarly, $U_{m\times m} = [\tilde{U} | \underbrace{u_{k+1} | u_{k+2} | \dots | u_m}_{\text{Null }(AA^T)}]$

$$\Sigma_{m\times n} = \begin{array}{cc} & \begin{bmatrix} \sqrt{\lambda_1} & & & 0 \\ & \ddots & \sqrt{\lambda_k} & \\ 0 & & & 0 \end{bmatrix} \\ m-k & \\ & \underbrace{}_{n-k} \end{array}$$

Then $U^T A V = \Sigma \Rightarrow \boxed{A = U \Sigma V^T}$ SVD

Left (U), right (V^T)

U & V are orthogonal matrices.

$\sigma_i = \sqrt{\lambda_i}$ is the ith singular value.

U is orthogonal s.t. the first k columns $u_1, u_2, \dots, u_k$ form an orthonormal basis of col (A).

& $u_{k+1}, \dots, u_n$ form an orthonormal basis of null (A^T)

V is orthogonal matrix s.t. $v_1, v_2, \dots, v_k$ ~~form~~ span row (A) & $v_{k+1}, \dots, v_n$ span Null (A)

$Av_i = \sigma_i u_i$ $i = 1$ to k

$Av_i = 0$ $i = k+1$ to n.

$A^T u_i = \sigma_i v_i$ $i = 1$ to k

$A^T u_i = 0$ $i = k+1$ to n.

Rank (A)
= No. of non-zero singular values.

Singular values are for arb. matrices, not necessarily sq. matrix, but e-values are for sq. matrices.

If rank $= n$, singular values $= n$
But e-values $\leq n$, if the matrix is sq. matrix.

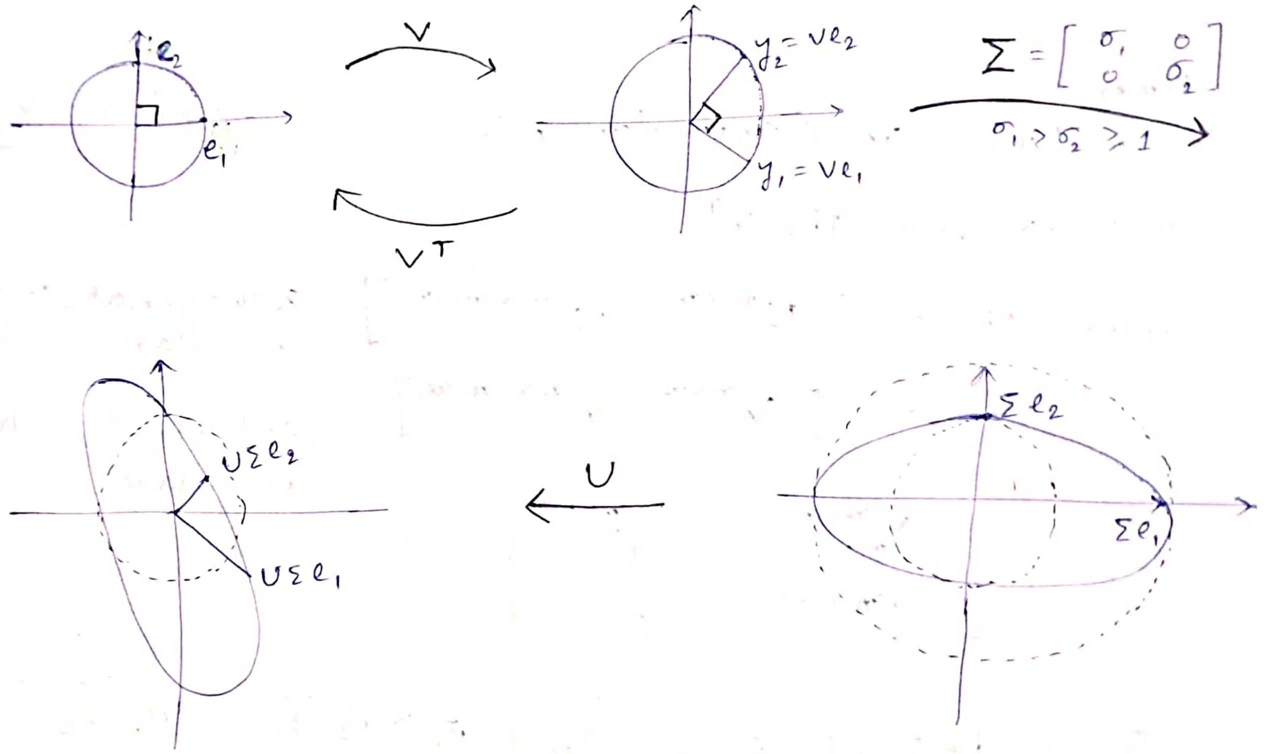
Geometry of the Mapping $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ using SVD

Take $m = n = 2$ $\qquad A_{2\times 2} = U \Sigma V^T$ $\qquad$ rank $(A) > 0$
(1 or 2)

A (Unit Circle)

Choose x s.t. $\|x\| = 1$

$Ax = U\Sigma V^T x = (U(\Sigma (V^T x)))$



** If rank $= 1$, a unit circle becomes a st. line of length 1.

e.g.

$A = \begin{bmatrix} 3 & 0 \\ 4 & 5 \end{bmatrix}$ $\qquad A^T A = \begin{bmatrix} 25 & 20 \\ 20 & 25 \end{bmatrix}$ $\quad$ symm. (+)ve semi definite

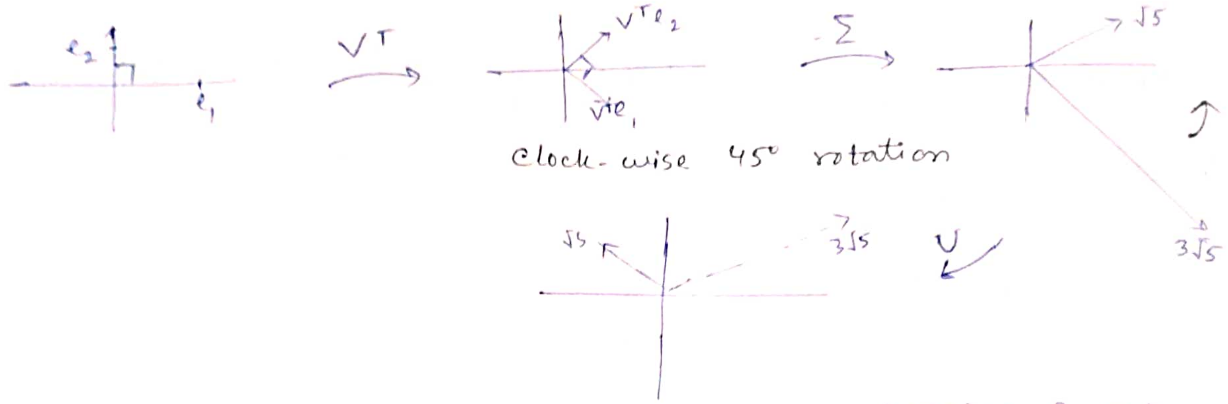
$\sigma_1^2 = 45$

$\sigma_2^2 = 5$

$V = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ $\qquad U = \frac{1}{\sqrt{10}} \begin{bmatrix} 1 & -3 \\ 3 & 1 \end{bmatrix}$

$A = U \begin{bmatrix} 3\sqrt{5} & 0 \\ 0 & \sqrt{5} \end{bmatrix} V^T$

$$V^T = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$



rotating ellipse (major & minor axes are NOT co-ordinate axes)

04/04/24

$A_{m \times n}$

$\text{rank}(A) = r \le \min\{m, n\}$

SVD $\Rightarrow$ $A = U \Sigma V^T$

where $U = [\,U^{m \times r} \quad U^{m \times m-r}\,]$ orthogonal, left singulars

$V = [\,V^{n \times r} \quad V^{n \times n-r}\,]$ orthogonal, right singulars

$$\Sigma = \begin{bmatrix} \Sigma_r^{r \times r} & 0^{r \times n-r} \\ 0^{m-r \times r} & 0^{m-r \times n-r} \end{bmatrix}$$

$\Sigma_r = \text{diag}(\sigma_1, \sigma_2, \ldots, \sigma_r)$ $\quad \sigma_1 \ge \sigma_2 \ge \ldots \ge \sigma_r > 0$

$$A = [\,U_r \quad U_{m-r}\,] \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_r^T \\ V_{n-r} \end{bmatrix}$$

$A = U_r \Sigma_r V_r^T$ $\rightarrow$ Reduced / Compact SVD

- $\text{Col}(U_r) = \text{Col}(A)$
- $\text{Col}(V_r) = \text{Col}(A^T) = \text{row}(A)$
- $\text{Col}(U_{m-r}) = \text{Null}(A^T)$
- $\text{Col}(V_{n-r}) = \text{Null}(A)$

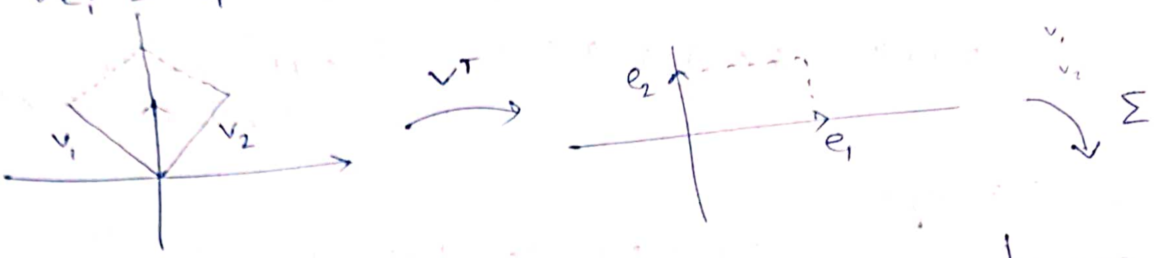
$AA^T = U \underset{m \times n}{\Sigma} \underset{n \times m}{\Sigma^T} U^T$ & $A^T A = V \Sigma^T \Sigma V^T$

Geometry : $A : \mathbb{R}^2 \to \mathbb{R}^2$

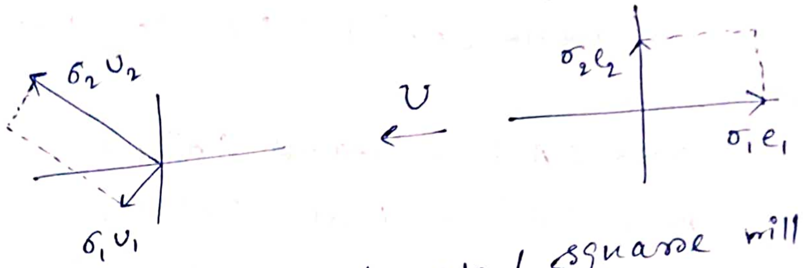
$A = U \Sigma V^T \qquad V = [v_1 \;\; v_2]$

$e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \qquad U = [u_1 \;\; u_2]$

$V e_i = v_i \quad \& \quad V^T v_i = e_i \qquad [\because \|v_i\| = 1]$



$\boxed{\sigma_1 > \sigma_2 \geq 1}$



Wheather the size of the rectangle / square will shrink or expand will depend on the σ_i's .

(1) $A_{m \times n}$, $B_{n \times n}$ full rank

$\Rightarrow$ rank (AB) = rank (A)

Proof :- $v \in col(A)$ $\qquad (x \in \mathbb{R}^n)$

$\exists\, x$ s.t. $Ax = v$

$\exists\, y$ s.t. $x = By$

Since B^{-1} exists , so $B^{-1}x = y$

$\therefore Ax = A(By) = (AB)y = \cancel{(AB)(B^{-1}x)}$

$\Rightarrow v \in col(AB)$

Conversely let $v \in col(AB)$

$\therefore v = ABx = A(Bx) = Ay$

Since $B_{n \times n}$ has full rank, $\exists\, y$ s.t. $Bx = y$

$\Rightarrow v \in col(A)$

$\therefore col(A) = col(AB)$

so A & AB have same rank.

(2) Similarly , if $B_{n \times n}$ has full rank , then rank (BA) = rank (B)

(3) rank (AA^T) = rank (A^TA)

Proof : By Spectral Thm., $A^TA = Q\Lambda Q^T$ where $QQ^T = Q^TQ = I$

$\Rightarrow A^TAQ = Q\Lambda$

rank (A^TAQ) = rank (A^TA)

rank $(Q\Lambda)$ = rank (Λ) = No. of non-zero e-values of A^TA

Similarly,

rank (AA^T) = No. of non-zero e-values of AA^T

rank (A) = rank (A^TA)

Null (A) = Null (A^TA)

rank (A) + Null (A) = n

rank (A^TA) + Null (A^TA) = n

$A_{m\times n}$, rank $(A) = r \le \min\{m,n\}$

A^TA, AA^T are symmetric positive semi-definite of rank r.

Both have r exactly, strictly (+)ve e-values.

rank (A) = No. of non-zero singular values.

e.g. (1) $A = \begin{bmatrix} 3 & 0 \\ 4 & 5 \end{bmatrix}$

$$A = \frac{1}{\sqrt{10}}\begin{bmatrix} 1 & -3 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} 3\sqrt{5} & 0 \\ 0 & \sqrt{5} \end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & +1 \\ -1 & 1 \end{bmatrix}$$

$U \quad \Sigma \quad V^T$

(2) $A = \begin{bmatrix} 4 & 4 \\ -3 & 3 \end{bmatrix}$

$$AA^T = \begin{bmatrix} 4 & 4 \\ -3 & 3 \end{bmatrix}\begin{bmatrix} 4 & -3 \\ 4 & 3 \end{bmatrix} = \begin{bmatrix} 32 & 0 \\ 0 & 18 \end{bmatrix}$$

$\sigma_1 = \sqrt{32} = 4\sqrt{2}$

$\sigma_2 = \sqrt{18} = 3\sqrt{2}$

$A = U\Sigma V^T$

U: e-vectors of AA^T V^T: e-vectors of A^TA

$U = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

$$A^T = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 4\sqrt{2} & 0 \\ 0 & 3\sqrt{2} \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 4\sqrt{2} & 0 \\ 0 & 3\sqrt{2} \end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

$$A = U\Sigma V^T$$
$$AV = U\Sigma$$
$$Av_i = u_i\sigma_i$$
$$u_i = \frac{1}{\sigma_i}Av_i$$

$$A^T = V\Sigma^T U^T$$
$$A^T U = V\Sigma^T$$
$$A^T u_i = v_i\sigma_i$$
$$v_i = \frac{1}{\sigma_i}A^T u_i$$

(3) $A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

$$A^TA = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = AA^T$$

$$= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2$$

$$A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \cos\theta & -\sin\theta \\ -\sin\theta & -\cos\theta \end{pmatrix}$$

(1) $\|A\|_2 = \sigma_1(A)$ $\qquad \sigma_1(A) = \sigma_{max}(A)$

L.H.S. $\sup\limits_{x\neq 0} \dfrac{\|Ax\|_2}{\|x\|_2}$

$$= \sup_{x\neq 0}\frac{\|U\Sigma V^T x\|_2}{\|x\|_2}$$

$$= \sup_{\|y\|\neq 0}\frac{\|\Sigma y\|_2}{\|Vy\|_2}$$

Let $V^Tx = y$
$\Rightarrow x = Vy$

$$= \sup_{\|y\|\neq 0}\frac{\|\Sigma y\|_2}{\|y\|_2}$$

$$= \sup_{y\neq 0}\frac{\sqrt{\sum_i \sigma_i^2 y_i^2}}{\sqrt{\sum_i y_i^2}}$$

$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0$

$\sigma_1^2 \geq \sigma_2^2 \geq \dots \geq \sigma_n^2 \geq 0$

$\therefore \sum \sigma_i^2 y_i^2 \leq \sigma_1^2(\sum_i y_i^2)$

$$\leq \sup_{y\neq 0}\frac{\sigma_1(\sum y_i^2)^{1/2}}{\sqrt{\sum_i y_i^2}}$$

$$\Rightarrow \|A\|_2 \leq \sigma_1$$

Consider $y = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$

$$\left\| \Sigma \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \right\|_2 = \sigma_1$$

$$\Rightarrow \| A \|_2 = \sigma_1$$

$\boxed{+}$ Let $A_{n \times n}$ be invertible

$$\Rightarrow \| A^{-1} \|_2 = \frac{1}{\sigma_{min}(A)}$$

$$\| A^{-1} \|_2 = \| (U \Sigma V^T)^{-1} \|_2$$

$$= \| \Sigma^{-1} \|_2$$

$$= \left\| \begin{bmatrix} 1/\sigma_1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & \ddots & \\ 0 & & & 1/\sigma_n \end{bmatrix} \right\|_2$$

$$= \frac{1}{\sigma_n(A)}$$

$$\boxed{\kappa(A) = \frac{\sigma_{max}(A)}{\sigma_{min}(A)}}$$

05/04/24

The Frobenius Norm:

$$\| A \|_F = \sqrt{\underbrace{\sum_{i,j} a_{ij}^2}_{\text{2-norm in } \mathbb{R}^{mn}}} = \sqrt{tr(A^T A)} = \sqrt{tr(AA^T)}$$

$$\| A \|_F = \sqrt{\sum_i \sigma_i^2}$$

$$(A^T A)_{ii} = \langle (A_{*i})^T, A_{*i} \rangle$$

$$= \sum_{k=1}^{m} a_{ik} a_{ik}$$

$$\therefore tr(A^T A) = \sum_{i=1}^{n} \sum_{k=1}^{m} a_{ik}^2 = \| A \|_F^2$$

$$= \sum_i \lambda_i (A^T A)$$

$$= \sum_i \sigma_i^2 (A^T A)$$

If $U_{m\times m}$ & $V_{n\times n}$ are orthogonal then

$$\| U A_{m\times n} \|_F = \| A \|_F = \| AV \|_F$$

$$\| UA \|_F^2 = tr\left((UA)^T (UA) \right)$$
$$= tr(A^T A) = \| A \|_F^2$$

$$\| A \|_F = \| U \Sigma V^T \|_F = \| \Sigma \|_F = \sqrt{\sum_{i,j} \Sigma_{ij}^2}$$
$$= \sqrt{\sum_i \sigma_i^2}$$

Now $\sigma_1^2 \geq \sigma_2^2 \geq \cdots \geq \sigma_n^2$

$$\sigma_1^2 \leq \sum_i \sigma_i^2 \leq n\sigma_1^2$$

$$\Rightarrow \sigma_1 \leq \| A \|_F \leq \sqrt{n}\, \sigma_1$$

$A_{m \times n} : \mathbb{R}^n \to \mathbb{R}^m$

$\text{rank}(A) = r$

In general. Image $(A) = \text{col}(A)$

$A|_{\text{row}(A)} : \mathbb{R}^{\text{rank}(A)} \xrightarrow{\cong} \mathbb{R}^{\text{rank}(A)}$

$\therefore A|_{\text{row}(A)}$ is an invertible map.

We want a map $f : \mathbb{R}^m \to \mathbb{R}^n$ s.t. the matrix of $f|_{\text{col}(A)}$ is the inverse of $A|_{\text{row}(A)}$.

The matrix of f will be denoted by A^+

$x \in \text{row}(A)$

$A^+(Ax) = x$ & $y \in \text{col}(A)$

$A(A^+y) = y$

$$\boxed{\text{Null}(A^+) = \text{Null}(A^T)}$$

__Moore - Penrose :__ (generalized / pseudo inverse)

- If A^{-1} exists then $A^+ = A^{-1}$
- $A_{m \times n}$ then A^+ is $n \times m$
- $(AA^+)A = A$
- $(A^+A)A^+ = A^+$
- $(AA^+)^T = AA^+$
- $(A^+A)^T = (A^+A)$

09/04/24

If A is invertible then $A^+ = A^{-1}$

$\Rightarrow$ If cols of A are linearly independent then

$A^+ = (A^TA)^{-1}A^T$

If rows of A are linearly independent then

$$A^+ = A^T (A A^T)^{-1}$$

$$\Sigma_r = \text{diag}(\sigma_1, \sigma_2, \ldots, \sigma_r)$$

$$\Sigma_r^{-1} = \text{diag}\left(\frac{1}{\sigma_1}, \frac{1}{\sigma_2}, \ldots, \frac{1}{\sigma_r}\right)$$

$$A_{m\times n} = \underset{m\times m}{U} \underset{m\times n}{\Sigma} \underset{n\times n}{V^T} \quad \& \quad \text{rank}(A) = r$$

where $\Sigma = \begin{bmatrix} \Sigma_r & O_{r\times n-r} \\ O_{n-r\times r} & O_{(n-r)\times(n-r)} \end{bmatrix}$

verify : $\Sigma^+ = \begin{bmatrix} \Sigma_r^{-1} & O_{r\times(n-r)} \\ O_{r\times n-r} & O_{(n-r)\times(n-r)} \end{bmatrix}$

$$\Rightarrow A^+ = V \Sigma^+ U^T$$

Compact SVD

$$A = U_r \Sigma_r V_r^T$$

$$A^+ = V_r \Sigma_r^+ U_r^T$$

$$AA^+ = (U_r U_r^T)_m = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}_{m\times m}$$

$\Rightarrow AA^+$ is projection on col(A)

$\Rightarrow A^+A$ " " " col(A^T) or row(A).

Revisit LSP :

Recall : $m > n$

we want to minimize $\| Ax - b \|$

The minimizer $\hat{x}$ satisfying $A\hat{x}$ = proj of b on col(A)

$$A\hat{x} = AA^+b$$

$$\Rightarrow \hat{x} = A^+ b$$

rank$(A) = n \Rightarrow \hat{x} = (A^TA)^{-1}A^Tb$

Minimal Norm LSP : $m < n$

Consider $Ax = b \longrightarrow$ (i)

It has infinitely many sol^n. We want the one with the least norm.

$$b = Ax$$
$$= U_r \Sigma_r V_r^T x$$
$$U_r^T b = \Sigma_r V_r^T x$$
$$\Sigma_r^{-1} U_r^T b = V_r^T x \longrightarrow \text{(ii)}$$

Any x that satisfies (2) solves (i)

where $A = U\Sigma V^T$ full SVD

$V = [V_r \quad V_{n-r}]$

$$\|x\| = \|V^T x\|$$
$$= \left\| \begin{bmatrix} V_r^T \\ V_{n-r}^T \end{bmatrix} x \right\|$$
$$= \left\| \begin{bmatrix} \overset{\text{constant}}{V_r^T x} \\ V_{n-r}^T x \end{bmatrix} \right\| \longrightarrow \text{(iii)}$$

Eq^n (iii) is minimized when $V_{n-r}^T x = 0$

$$V^T x = \begin{bmatrix} V_r^T x \\ V_{n-r}^T x \end{bmatrix} = \begin{bmatrix} \Sigma_r^{-1} U_r^T b \\ 0 \end{bmatrix}$$

$$\Rightarrow VV^T x = V \begin{bmatrix} \Sigma_r^{-1} U_r^T b \\ 0 \end{bmatrix}$$

$$\Rightarrow x = [V_r \quad V_{n-r}] \begin{bmatrix} \Sigma_r^{-1} U_r^T b \\ 0 \end{bmatrix}$$

$$= \underbrace{V_r \Sigma_r^{-1} U_r^T}_{\text{compact SVD of } A^+} b$$

$$= A^+ b$$

Note: $\hat{x} = A^{+} b$ is the unique solⁿ with min. norm.

[$Ax = b$ has many solⁿ]

$V_{n-r}^{T} \hat{x} = 0$

$\Rightarrow \hat{x}$ is $\perp$ to $col(V_{n-r})$

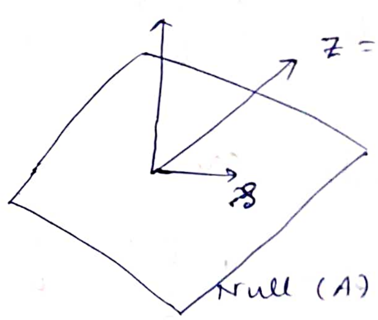
$\Rightarrow \hat{x} \perp Null(A)$

$U \rightarrow col(A)$
$Null(A^T)$

$V \rightarrow col(A^T)$
$\Rightarrow row(A)$
$Null(A)$

if $Z = \hat{x} + s$ where $s \in Null(A)$ then Z is a solⁿ to $Ax = b$ & $\|Z\| > \|\hat{x}\|$

$$\boxed{\begin{array}{l} rank(A) = m \\ \hat{x} = A^T (AA^T)^{-1} b \end{array}}$$



e.g. $A = \begin{bmatrix} 4 & 4 \\ 3 & 3 \end{bmatrix}$

$A^T A = \begin{bmatrix} 4 & 3 \\ 4 & 3 \end{bmatrix} \begin{bmatrix} 4 & 4 \\ 3 & 3 \end{bmatrix} = \begin{bmatrix} 25 & 25 \\ 25 & 25 \end{bmatrix}$

$\chi(A^T A, \lambda) \rightarrow$ characteristic poly.

$\Rightarrow (25 - \lambda)^2 - 25^2 = 0$

$\Rightarrow (25 - \lambda)^2 = 25^2$

$\Rightarrow 25 - \lambda = \pm 25$

$\Rightarrow \lambda = 0$ or 50 say $\lambda_1 = 50$, $\lambda_2 = 0$

v_1 is the eigen-vector corresponding to λ_1

$\left[\begin{array}{cc|c} -25 & 25 & v_1 \\ +25 & -25 & v_2 \end{array}\right] = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

$\therefore v_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$v_2 \in Null(A) = Span\left\{ \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}$

$v_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

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

$\begin{bmatrix} -25v_1 + 25v_2 \\ 25v_1 \cancel{+0} - 25v_2 \end{bmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ $-25v_1 + 2$

$v_1 = v_2$

$-25v_1 + 25v_2 = 0$

$\boxed{v_1 = v_2}$ $\begin{bmatrix} c \\ c \end{bmatrix}$

$R^{2\times 2}$ $c\begin{pmatrix} 1 \\ 1 \end{pmatrix} c$ c c $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

2

$$AV_1 = \sigma_1 u_1$$

$$u_1 = \frac{1}{\sqrt{50}} \begin{bmatrix} 4 & 4 \\ 3 & 3 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$= \frac{1}{10} \begin{bmatrix} 8 \\ 6 \end{bmatrix}$$

$$= \frac{1}{5} \begin{bmatrix} 4 \\ 3 \end{bmatrix}$$

$$u_2 \in \text{Null}(A^T) = \text{Span}\left\{ \begin{bmatrix} -3 \\ 4 \end{bmatrix} \right\}$$

$$u_2 = \frac{1}{5} \begin{bmatrix} -3 \\ 4 \end{bmatrix}$$

$$U = \frac{1}{5} \begin{bmatrix} 4 & -3 \\ 3 & 4 \end{bmatrix}$$

Full SVD

$$A = \frac{1}{5} \begin{bmatrix} 4 & -3 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 5\sqrt{2} & 0 \\ 0 & 0 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

$$A^+ = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \frac{1}{5\sqrt{2}} & 0 \\ 0 & 0 \end{bmatrix} \frac{1}{5} \begin{bmatrix} 4 & 3 \\ -3 & 4 \end{bmatrix}$$

Compact SVD

$$A = \frac{1}{5} \begin{bmatrix} 4 \\ 3 \end{bmatrix} \begin{bmatrix} 5\sqrt{2} \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 4 \\ 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \end{bmatrix}$$

e R

$$A^+ = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} \frac{1}{5\sqrt{2}} \end{bmatrix} \frac{1}{5} \begin{bmatrix} 4 & 3 \end{bmatrix}$$

$$= \frac{1}{50} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} 4 & 3 \end{bmatrix}$$

$$= \frac{1}{50} \begin{bmatrix} 4 & 3 \\ 4 & 3 \end{bmatrix}$$

(2) $A = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

$A^T A = \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = 5$

$\sigma_1^2 = 5$

$v_1 = 1$

$[V] = [1]$

$u_1 = \frac{1}{\sqrt{5}} A v_1 = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

$\text{Null}(A^T) = \text{Span}\left\{ \begin{bmatrix} -2 \\ 1 \end{bmatrix} \right\}$

$u_2 = \frac{1}{\sqrt{5}} \begin{bmatrix} -2 \\ 1 \end{bmatrix}$

$U = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix}$

$A = \frac{1}{\sqrt{5}} \underbrace{\begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix}}_{2\times2} \underbrace{\begin{bmatrix} \sqrt{5} \\ 0 \end{bmatrix}}_{2\times1} \begin{bmatrix} 1 \end{bmatrix} = \begin{bmatrix} 1 & \cancel{-2} \\ 2 & \cancel{1} \end{bmatrix} \frac{1}{\sqrt{5}} \begin{bmatrix} \sqrt{5} \\ 2\sqrt{5} \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

$A^+ = \begin{bmatrix} 1 \end{bmatrix} \underbrace{\begin{bmatrix} 1/\sqrt{5} & 0 \end{bmatrix}}_{1\times2} \frac{1}{\sqrt{5}} \underbrace{\begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}}_{2\times2} = \begin{bmatrix} 1/5 & 2/5 \end{bmatrix}$

(3) $A = \begin{bmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{bmatrix}$

$A^T A = \underbrace{\begin{bmatrix} 1 & -2 & 2 \\ -1 & 2 & -2 \end{bmatrix}}_{2\times3} \underbrace{\begin{bmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{bmatrix}}_{3\times2} = \begin{bmatrix} 9 & -9 \\ -9 & 9 \end{bmatrix}$

$\chi(A^T A, \lambda) = (9-\lambda)^2 - 81 = 0$

$\Rightarrow (9-\lambda) = \pm 9$

$\Rightarrow \lambda = 0 \text{ or } 18$ $\qquad \lambda_1 = 18, \quad \lambda_2 = 0$

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

$u_1 = \frac{1}{\sqrt{18}} \begin{bmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} -1 \\ 2 \\ -2 \end{bmatrix}$

Nullity $(A^T) = 2$

$$u_2 = \frac{1}{3}\begin{bmatrix} \sqrt{8} \\ 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$$

$$u_3 = \begin{bmatrix} 0 \\ 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

$$A = \underbrace{\begin{bmatrix} -1/3 & \sqrt{8}/3 & 0 \\ 2/3 & 1/3\sqrt{2} & 1/\sqrt{2} \\ -2/3 & -1/3\sqrt{2} & 1/\sqrt{2} \end{bmatrix}}_{3\times 3} \underbrace{\begin{bmatrix} 3/\sqrt{2} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}}_{3\times 2} \cdot \frac{1}{\sqrt{2}} \underbrace{\begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}}_{2\times 2}$$

$$A^+ = \frac{1}{\sqrt{2}} \underbrace{\begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}}_{2\times 2} \underbrace{\begin{bmatrix} \sqrt{2}/3 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}}_{2\times 3} \underbrace{\begin{bmatrix} -1/3 & 2/3 & -2/3 \\ \sqrt{8}/3 & 1/3\sqrt{2} & -1/3\sqrt{2} \\ 0 & 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}}_{3\times 3}$$