Lecture - 21 (Saturday)

~~For ne two~~

We take two wild guesses (a, b)

where $f(a) =$ positive

and $f(b) =$ negative

then we take $f\left(\frac{a+b}{2}\right)$

Depending on this we change the direction and reach to some good approximation

$$e_{k+1} \simeq e_k^2 \left(\frac{f''(x^*)}{2f'(x^*)}\right)$$

η

$$\boxed{e_k \simeq (e_0 \eta)^{(2^k)} \cdot \eta^{-1}}$$

$$|e_0| = |x_0 - x^*| \leq |x_0 - x^*_{LB/UB}|$$

(For convex function)

▷ Why zeroes ? ⟶ Local optimization

$f'(x) = 0$ with quadratic covergence rate $\frac{f'''(x^*)}{2f''(x^*)}$.

Gradient Descent Method

$$|f''(x)| \leq L \text{ for } x \in [a,b]$$

$$f'(x^*) = 0$$

1. Start with an initial guess.
   Define $\gamma \leftarrow L^{-1}$

   known bound
2. while $f'(x) \neq 0$
   and do $x \leftarrow x - \gamma f'(x)$ endwhile
3. return x

we go the opposite direction for minimum

(Analysis has not been written!!!)

Scalar and vector functions,

$f: R^n \longrightarrow R^m,\ n, m \geq 1$

1. If $m = 1$ (real valued function) Scalar
2. If $m > 1$ (vector ——"——)
3. If $n = 1$ and $m > 1$ (trajectories | projectile)

$\vec{x} \in \mathbb{R}^d,\ \|\vec{x}\|_2 = \sqrt{x_1^2 + \dots + x_d^2}$ (L_2- Norm)

$\vec{x}, \vec{y} \in R^d,\ d_2(\vec{x}, \vec{y}) = \|\vec{x} - \vec{y}\|_2$ (L_2-Distance)

$f: R^n \longrightarrow R^m \quad \vec{a} \in R^n,\ \vec{l} \in R^m$

$\lim_{\vec{x} \to \vec{a}} f(x) = \vec{l}$. if $\forall \epsilon > 0, \exists \delta > 0$

$d_2(f(\vec{x}), \vec{l}) \leq \epsilon$

whenever

$0 < d_2(\vec{x}, \vec{a}) \leq \delta$

f is continuous at $\vec{a}$ if $\lim_{\vec{x} \to \vec{a}} f(\vec{x}) = f(\vec{a})$

Norm function is continuous

$f: R^3 \to R$

$(x, y, z) \to \sqrt{x^2 + y^2 + z^2}$

$$\left.\begin{array}{l} f(\vec{x}) \to \vec{f} \\ g(\vec{x}) \to \vec{g} \end{array}\right\} \text{ when } \vec{x} \to \vec{a}$$

$$f(\vec{x}) \pm g(\vec{x}) \to \vec{f} \pm \vec{g} \text{ as } \vec{x} \to \vec{a}$$

① $\alpha f(\vec{x}) \to \alpha \vec{f} \quad \forall \alpha \in \mathbb{R}$

② $f(\vec{x}) \cdot g(\vec{x}) \longrightarrow \vec{f} \cdot \vec{g}$

$$\boxed{\vec{x} \cdot \vec{y} = \sum_{i}^{n} x_i y_i}$$

$\vec{x}, \vec{y} \in \mathbb{R}^n$

③

$f : R^n \to R^m$

$g : R^m \to R^p$

$h = g \cdot f : R^n \to R^p$ by $h(\vec{x}) = g(f(\vec{x}))$ for $x \in \mathbb{R}^n$

f is continuous at $\vec{a} \in \mathbb{R}^n$

g is continuous at $f(\vec{a}) \in \mathbb{R}^m$

Then h is continuous at $\vec{a}$.

Ex1.

$f_1(x,y) = \sin(x^2 y)$ → Continuous everywhere

$f_2(x,y) = \log_e(x^2+y^2)$ → Not continuous at $(0,0)$

$f_3(x,y) = \dfrac{e^{x+y}}{x+y}$ → Not continuous where $x+y=0$

$$f(x,y) = \begin{cases} \dfrac{xy}{x^2+y^2} & ; (x,y) \neq (0,0) \\ 0 & ; (x,y) = (0,0) \end{cases}$$

Continuous for 'x' alone

Continuous for 'y' alone

→ Not continuous for both 'x' and 'y'
for $(x=y)$ it is not continuous

Differentiability of Scalar functions

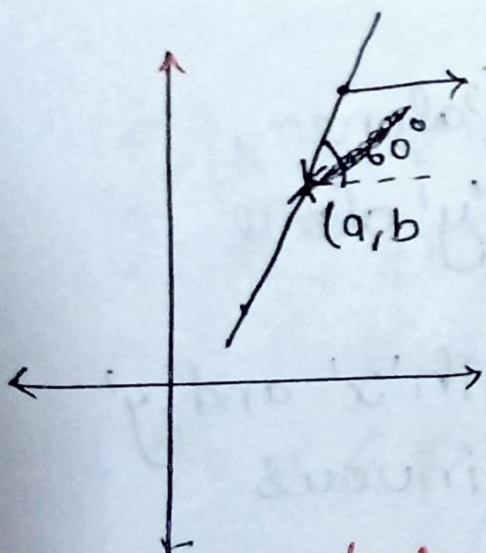
$f: R^n \to R$

$f(x) = 4x^3 - 5y^3$

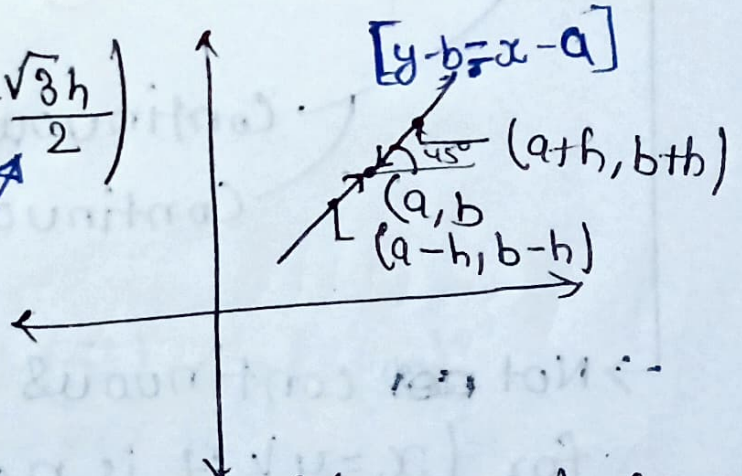
$\frac{\partial f}{\partial x} = 12x^2$ (Positive except 0) ≥ 0 | $\frac{\partial f}{\partial y} = -15y^2$ (Negative except 0) ≤ 0

cos θ

$\left(a + \frac{h}{2}, b + \frac{\sqrt{3}h}{2}\right)$



[y-b = x-a]



sin θ

$$\lim_{h \to 0} \frac{f\left(a+\frac{h}{2}, b+\frac{\sqrt{3}h}{2}\right) - f(a,b)}{h}$$

$$\frac{f(a+h, b+h) - f(a,b)}{h}$$

Gradient of f at $\vec{a}$ is the vector

$$\nabla f(\vec{a}) = \left(\frac{\partial f(\vec{a})}{\partial x_1}, \ldots, \frac{\partial f(\vec{a})}{\partial x_n}\right)$$

Directional derivatives

$f(\vec{a}, \vec{r}) = 0$

$\|\vec{y}\|_2 = 1$

$f'(\vec{a}, \vec{y})$ is directional ~~directive~~ derivative of f at $\vec{x}$

$\vec{y} = e_i$ along x_i

$f(\vec{a}, e_i) = \dfrac{\partial f(\vec{a})}{\partial x_i}$

⚠

▷ Existence of directional derivative $f'(\vec{a}, \vec{y})$ for each $\vec{y}$ does not guarantee f is continuous at $\vec{a}$

$$f(x, y) = \begin{cases} \dfrac{xy^2}{x^2 + y^4} & ; x \neq 0 \\ 0 & ; x = 0 \end{cases}$$

$\boxed{x = y^2}$ Not continuous

$$f(y^2, y) = \frac{y^4}{y^4 + y^4} = \frac{y^4}{2y^4} = \frac{1}{2}$$

f is not continuous at $(0,0)$

f is differentiable at $\vec{a}$ if for some $\delta > 0$ there exist $T_{\vec{a}} : \mathbb{R}^n \to \mathbb{R}$ and $E_{\vec{a}}(\vec{y})$ such that

$$f(\vec{a}+\vec{y}) = f(\vec{a}) + T_{\vec{a}}(\vec{y}) + \|\vec{y}\|_2 \cdot E_{\vec{a}}(\vec{y})$$

$$\forall \|\vec{y}\| < \delta$$

and $E_{\vec{a}}(\vec{y}) \to 0$ as $\|\vec{y}\| \to 0$

$T_{\vec{a}}$ is total derivative of f $\boxed{f'(\vec{a})}$

f is differentiable at $\vec{a}$

$\Rightarrow T_{\vec{a}}(\vec{y}) = f'(\vec{a}, \vec{y}) \;\forall \vec{y}$

$$T_{\vec{a}}(\vec{y}) = \nabla f(\vec{a}) \cdot \vec{y} = \sum_{i=1}^{n} \frac{\partial f(\vec{a})}{\partial x_i} y_i \;\forall \vec{y}$$

Taylor's first order formula

$$f(\vec{a}+\vec{y}) = f(\vec{a}) + \nabla f(\vec{a}) \cdot \vec{y} + \|y\| E_{\vec{a}}(\vec{y}),$$

$$\|y\| < \delta$$

$\|y\| = 1, \; f'(\vec{a}, \vec{y}) = \|\nabla f(\vec{a})\| \cdot \cos\theta$

θ angle between $f(\vec{a})$

$\vdots$

⚠ {Iske Baad Kya padhya No IDEA}

Lecture-22 (21st October '24)

- If A is invertible then $\det(A)$ can be written as Laplace Expansion either assume a row or a column

$$A = (a_{ij}) \in \mathbb{R}^{n\times n}$$

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} \det(A_{ij}) \text{ for any row 'i'}$$

or

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} \det(A_{ij}) \text{ for any column 'j'}$$

where the matrix $A_{ij} \in \mathbb{R}^{(n-1)\times(n-1)}$ and obtained by eliminating the 'ith' row and 'jth' column of A

A_{ij} is called 'MINOR' of A and $C_{ij} = (-1)^{i+j} A_{ij}$ is called the 'COFACTOR' of A.

Definitions:

(I) Two matrices $A, B \in \mathbb{R}^{n\times n}$ are 'SIMILAR' iff there exists an invertible matrix $S \in \mathbb{R}^{n\times n}$ where $B = S^{-1} A S$

(II) $\text{Trace}(A) = \sum_{i=1}^{n} a_{ii}$

Prove:

$\boxed{(a)}$ If $A, B \in \mathbb{R}^{n\times n}$ then $\det(AB) = \det(A)\cdot\det(B)$
$= \det(B)\cdot\det(A) = \det(BA)$

(b) A is invertible iff $\det(A) \neq 0$

(c) A is invertible then $\det(A^{-1}) = \dfrac{1}{\det(A)}$

(d) If A and B are similar.
then $\det(A) = \det(B)$

(e) $\det(A) = \det(A^T)\ \forall\ A \in \mathbb{R}^{n\times n}$

(f) $\det(\lambda A) = \lambda^n \cdot \det(A)$, $A \in \mathbb{R}^{n\times n}$

(g) $\det(A) \neq 0$ then rank of $(A) = n$

(h) $\text{Trace}(A+B) = \text{Trace}(A) + \text{trace}(B)$

(i) $\text{Trace}(\alpha A) = \alpha \cdot [\text{Trace}(A)]$

(j) $\text{Trace}(I) = n$, $I \in \mathbb{R}^{n\times n}$

$\boxed{(k)}$ $\text{Trace}(AB) = [\text{Trace}(A)] \cdot [\text{Trace}(B)]$

(l) If T is a triangular matrix $\in \mathbb{R}^{n\times n}$ (lower/upper)
then $\det(T) =$

- Characteristic polynomial
Eigenvector and Eigen value

Polynomial of degree 'n'

$$\sum_{i=0}^{n} a_i x^i = a_0 + a_1 x + a_2 x^2 \cdots a_n x^n = P_n(x)$$

Polynomial Equation
$P_n(x) = 0$ (when it is true → next page)

$P_n(x) = 0$ iff ~~q(n~~ $q_n(x) = 0$; where $q_n(x)$
=

Chacteristics polynomial

If $A \in \mathbb{R}^{n \times n}$
then $\det(A - \lambda I)$ is called as characteristics polynomial of degree 'n'.

$$\det(A - \lambda I) = \boxed{C_0} + C_1 \lambda + C_2 \lambda^2 + \dots \boxed{C_{n-1}} \lambda^{n-1} + \boxed{C_n} \lambda^n$$

$C_0 \rightarrow$ Determinant of A

$C_n \rightarrow (-1)^n$

$C_{n-1} \rightarrow (-1)^{n-1}$ Trace (A)

→ If $A \in \mathbb{R}^{n \times n}$, λ is an eigen ~~vect~~ value of A with an eigenvector $\vec{x} \mid \vec{x} \neq 0$ iff $A\vec{x} = \lambda \vec{x}$

Algebraic multiplicity (Algebra)
$\neq$
Geometric multiplicity
(Graph)

Proof: λ is an eigenvalue of A iff λ is a root of $\det(A - \lambda I^{n})= 0$

Note:

As polynomial λ can be =
① real
② real with multiplicity (A.M) ≥ 1
③ complex number along with its conjugate as a summation

$$A = \begin{bmatrix} 4 & 2 \\ 1 & 3 \end{bmatrix}$$

$$\det(A-\lambda I) = \det\left(\begin{bmatrix} 4-\lambda & 2 \\ 1 & 3-\lambda \end{bmatrix}\right)$$

$$= (4-\lambda)(3-\lambda) - 2$$

$$= 12 - 7\lambda + \lambda^2 - 2$$

$$= (\lambda-2)(\lambda-5)$$

The roots of $\det(A-\lambda I) = 0$

$$\Rightarrow \lambda = 2 \text{ and } \lambda = 5$$

Eigen vector

$$(A-\lambda I)\vec{x} = \vec{0}$$

$$\begin{bmatrix} 4-2 & 2 \\ 1 & 3-1 \end{bmatrix}\vec{x} = \vec{0} \Rightarrow \begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

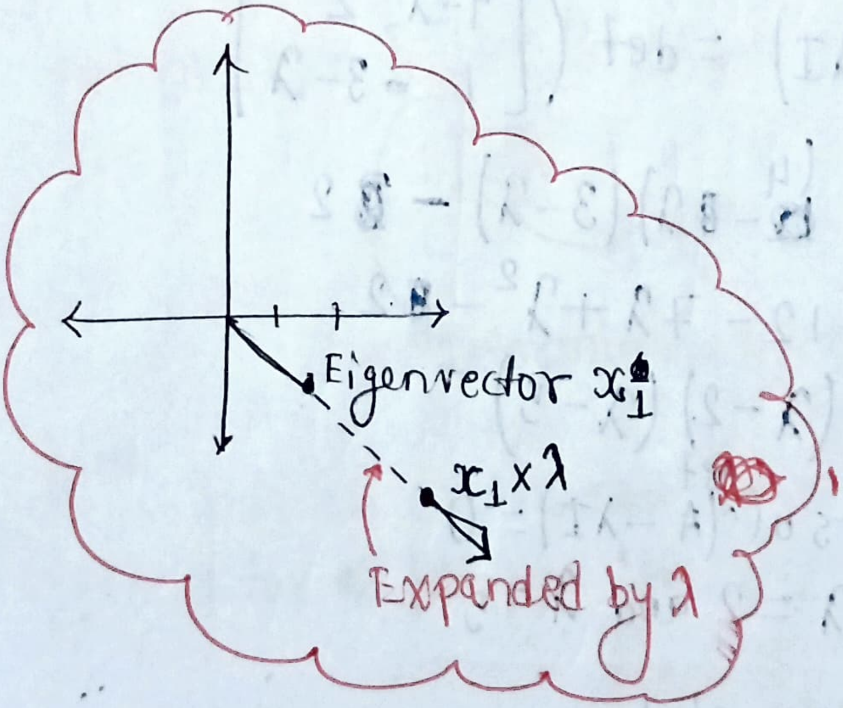
$$\Rightarrow \vec{x} = \begin{bmatrix} x_1 \\ -x_1 \end{bmatrix} \quad \vec{x}_1 = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$$

$\lambda = 5$

$$\begin{bmatrix}4-5 & 2\\ 1 & 3-5\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix} = \begin{bmatrix}0\\ 0\end{bmatrix}$$

$$\Rightarrow \begin{bmatrix}-1 & 2\\ 1 & -2\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix} = \vec{0}$$

$$\Rightarrow \vec{x} = \begin{bmatrix}2x_2\\ x_2\end{bmatrix} \quad \hat{\vec{x}} = \begin{bmatrix}2/\sqrt{5}\\ 1/\sqrt{5}\end{bmatrix}$$



Lecture - 23 (23rd October '24)

$$P_n(\lambda) = C_0 + C_1\lambda + \dots C_{n-1}\lambda^{n-1} + C_n\lambda^n$$

$$C_0 = \det(A), \quad C_{n-1} = (-1)^{n-1}\, \text{trace}(A)$$

Prove:

(a) $\det(A) = \prod_{i=1}^{n} \lambda_i$

(b) $\text{trace}(A) = \sum_{i=1}^{n} \lambda_i$

Always real

Definition:

- If $A \in \mathbb{R}^{n\times n}$ is a symmetric matrix then
  (a) the eigen values are real
  (b) there exists $\boxed{\text{an}}$ orthonormal basis ~~will be~~ ~~real~~ (eigen vector)

- Moreover if A is positive definite
  [In Addition to Symmetric ($\vec{x}^T A \vec{x} > 0, \vec{x} \neq 0$)]
  then eigenvalues are positive

Example:

$$A = \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & 2 \\ 2 & 2 & 3 \end{bmatrix}$$ find eigenvalues and eigen vectors.

Sol:

$$\det(A-\lambda I) = \begin{vmatrix} 3-\lambda & 2 & 2 \\ 2 & 3-\lambda & 2 \\ 2 & 2 & 3-\lambda \end{vmatrix}$$

$$= (-1)^{1+1}(3-\lambda)\begin{vmatrix} 3-\lambda & 2 \\ 2 & 2-\lambda \end{vmatrix} + (-1)^{1+2}\, 2\begin{vmatrix} 2 & 2 \\ 2 & 3-\lambda \end{vmatrix}$$

$$+ (-1)^{1+3}\, 2\begin{vmatrix} 2 & 3-\lambda \\ 2 & 2 \end{vmatrix}$$

$$= (3-\lambda)(\lambda^2-5\lambda+6) - (2)(6-2\lambda-4)$$

$$+ (2)(4-6-2\lambda)$$

$=$

$\Rightarrow \boxed{\lambda = 1, 1, 7}$

$[A - \lambda I]\vec{x} = \vec{0}, \vec{x} \neq \vec{0}$

$\lambda = 1$

$$\begin{bmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow \vec{x} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

$$\hat{\vec{x}} = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \end{bmatrix}, \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}, \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ -1/\sqrt{2} \end{bmatrix}, \begin{bmatrix} -1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix}$$

$\lambda = 7$

$$\begin{bmatrix} -4 & 2 & 2 \\ 2 & -4 & 2 \\ 2 & 2 & -4 \end{bmatrix} \begin{bmatrix} y_0 \\ y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

+

$$\Rightarrow \vec{y} = \begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix} \quad \hat{\vec{y}} = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$$

+

$$\text{Basis 1} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}$$

$$\text{Basis 2} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}$$

Diagonalizable:

A matrix is diagonalizable if there exists a diagonal matrix D ~~such that A=~~ and an invertible matrix S such that $\boxed{A = S^{-1} D S}$

Prove: A symmetric matric is diagonalizable

Proof: Let $P = [\vec{x}_1, \vec{x}_2 \dots \vec{x}_n]$ ← Also the orthonormal basis, be the eigenvectors of A

and let $D = \begin{bmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \ddots \\ & & & \lambda_n \end{bmatrix}$

listing the eigenvalues of A where λ_i is the eigenvalue with eigenvector $\vec{x}_i$

$$AP = A\left[\vec{x}_1, \vec{x}_2 \dots \vec{x}_n\right]$$

$$A\vec{x}_j = \lambda_j \vec{x}_j.$$

Therefore, $A\left[\vec{x}_1, \vec{x}_2 \dots \vec{x}_n\right] = \left[\vec{x}_1, \vec{x}_2 \dots \vec{x}_n\right] \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}$

$$\Rightarrow AP = PD.$$

$$\cancel{(AP)A^{-1}} \Rightarrow (AP)(P^{-1}) = (PD)P^{-1}$$

$$\Rightarrow A(PP^{-1}) = PDP^{-1}$$

$$\Rightarrow \boxed{A = PDP^{-1}}$$

P is nothing but eigenvector orthonormal basis

Orthonormal matrix

If A is symmetric then there exists a diagonal matrix D ~~exist~~ listing the eigenvalues that is similar to A.

Singular value Decomposition (SVD)

$$A = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 2 \end{bmatrix} \qquad A^T = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 2 \end{bmatrix}$$

$A^TA \in \mathbb{R}^{2\times 2}$ ← Least Square

Let $A \in \mathbb{R}^{m\times n}$ ~~then~~ with $\text{rank}(A) = r$, $r \in \{0 \dots \min(m,n)\}$ then ther exists orthogonal matrix U and V, $U \in \mathbb{R}^{m\times m}$, $V \in \mathbb{R}^{n\times n}$ and a diagonal matrix $\Sigma \in \mathbb{R}^{m\times n}$ where

$$\underset{m\times n}{A} = \underset{m\times m}{U}\,\underset{m\times n}{\Sigma}\,\underset{n\times n}{V}$$

and $\Sigma = \text{diagonal}(\sigma_1, \sigma_2 \dots \sigma_r, \dots 0)$,

with $\boxed{\sigma_1 \geq \sigma_2 \geq \sigma_3 \dots \geq \sigma_r \geq 0}$

$$\underset{\substack{m\times n \\ m=n}}{\Sigma} = \begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_r \end{bmatrix} \qquad \underset{\substack{m\times n \\ m>n}}{\Sigma} = \begin{bmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & \ddots \; \sigma_r \\ & 0 & \end{bmatrix}$$

$$\underset{\substack{m\times n \\ m<n}}{\Sigma} = \begin{bmatrix} \sigma_1 & & & & \\ & \sigma_2 & & & \\ & & \ddots & & \\ & & & \sigma_r & 0 \end{bmatrix}$$

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

Let $B = (A^T A)$. $\boxed{A \in \mathbb{R}^{m \times n}}$ $\boxed{A^T \in \mathbb{R}^{n \times m}}$

then $B \in \mathbb{R}^{n \times n}$ and B is <u>symmetric</u> Always symmetric

$B = A^T A$

$= (U \Sigma V)^T (U \Sigma V)$ $\quad [(AB)^T = B^T A^T]$

$= V^T \Sigma^T (U^T U) \Sigma V$

$= V^T \Sigma^T I \Sigma V$

$= V^T \hat{\Sigma} V$ $\quad$ where $\hat{\Sigma} = \Sigma^T \Sigma$

$= \text{diagonal}(\sigma_1^2, \sigma_2^2 \dots \sigma_r^2 \dots 0)$

$\Rightarrow B = V^{-1} \hat{\Sigma} V$ (V is orthogonal)

$\Rightarrow (\sigma_1^2, \sigma_2^2 \dots \sigma_r^2 \dots 0)$ are the eigenvalues of B.

<u>Summary:</u>

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

construct $B = A^T A \in \mathbb{R}^{n \times n}$

then there exists an ortho-normal basis V where

$B = V^{-1} \hat{\Sigma} V$

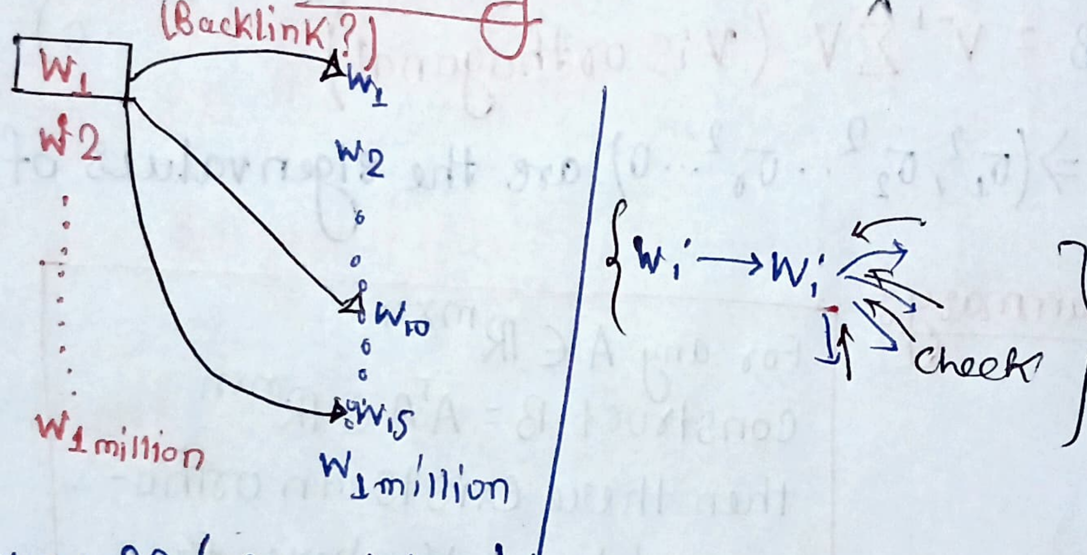
$\Sigma = \text{diagonal}(\sigma_1^2, \sigma_2^2 \dots \sigma_r^2 \dots 0)$

each σ_i^2 is an SVD and also an eigenvalue.

SVD is a versatile tool to address

(a) Data Compression
(b) Image Compression
(c) Noise reduction
(d) signal processing
(e) feature extraction
(f) dimensionality reduction
(g) Constraint optimization

Indexing



Lecture 23 (26 october)
Lecture 24 (28 october)
Lecture 25 (30 October)
Lecture 26 (2 November)
Lecture 27 (4 November)