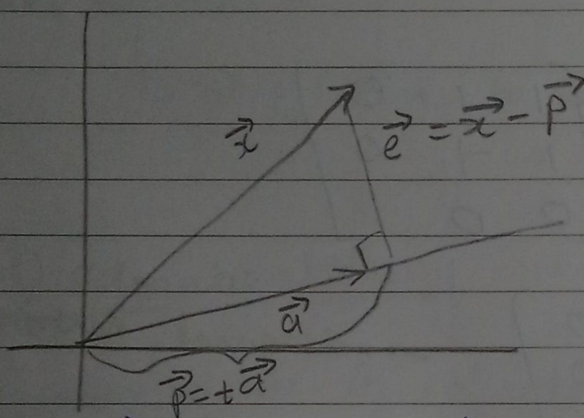
Dimensionality Reduction

- 'Most important features'
  High var and low Covariance
  (between independent variances)

- Orthogonal → dot product is zero
  [perpendicular]
- Normal → magnitude is one



$\vec{a}$ & $\vec{e}$ are orthogonal

$$\vec{a}\vec{e} = 0$$
$$\vec{a}(\vec{x} - \vec{p}) = 0$$
$$\vec{a}(\vec{x} - t\vec{a}) = 0$$
$$\vec{a}.\vec{x} - t\vec{a}.\vec{a} = 0$$
$$t = \frac{\vec{a}.\vec{x}}{\vec{a}.\vec{a}}$$

Projection vector $\vec{p} = t\vec{a}$

$$\vec{p} = \vec{a}\,\frac{\vec{a}^T\vec{x}}{\vec{a}^T\vec{a}} = \frac{\vec{a}\vec{a}^T\vec{x}}{\vec{a}^T\vec{a}}$$

— PCA
- Reduce dimensionality of a data set by finding a new set of variables smaller than the original set of values.

- 1^{st} PC is a maximum var fit
- 2^{nd} PC is perpendicular to first.

eg→

x_1	x_2
1	2
3	4
5	6

$E(x_1) = 3$, $E(x_2) = 4$

$m = 2$, $n = 3$

x_1	x_2
-2	-2
0	0
2	2
0	0

$$Cov(x_1, x_2) = \frac{\Sigma(x_1 - \bar{x}_1) \times (y_1 - \bar{y}_1)}{n-1}$$

$$= \frac{(-2 \times -2 + 0 + 2 \times 2)}{3-1}$$

$$\cancel{= \Sigma(-2 \times}$$

$$= 8/\cancel{3}2$$

	A	B
A	Cov(A,A) or Var(A)	Cov(B,A)
B	Cov(A,B)	Var(B)

$$Var(x_1) = \frac{\Sigma(x_1 - \bar{x}_1)^2}{n}$$

$$= \frac{(-2-0)^2 + 0 + (2-0)^2}{2}$$

$$= \frac{8}{\cancel{3}2}$$

$$Var(x_2) = \frac{8}{\cancel{3}2}$$

$\therefore$ Covariance Matrix → $\begin{bmatrix} 8/\cancel{3}2 & 8/\cancel{3}2 \\ 8/\cancel{3}2 & 8/\cancel{3}2 \end{bmatrix}$

$$|A - \lambda I| = 0$$

$$\left| \begin{bmatrix} \cancel{8/3}\,4 & \cancel{8/3}\,4 \\ \cancel{8/3}\,4 & \cancel{8/3}\,4 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \right| = 0$$

$$\begin{vmatrix} \cancel{8/3}\,4 - \lambda & \cancel{8/3}\,4 \\ \cancel{8/3}\,4 & \cancel{8/3}\,4 - \lambda \end{vmatrix} = 0 \qquad (\cancel{8/3}\,4 - \lambda)^2 - (\cancel{8/3}\,4)^2 = 0$$

$\left(\cancel{\tfrac{8}{3}}^{4}\right)^{2} + \lambda^{2} - 2 \times \tfrac{\cancel{8}^{4}}{3} \times \lambda - \cancel{\left(\tfrac{8}{3}^{4}\right)^{2}} = 0$

$\lambda^2 - \frac{\cancel{16}\,8}{3}\lambda = 0$

$\lambda^2 = \frac{\cancel{16}\,8}{3}\lambda$

$\lambda = \cancel{\frac{16}{3}}\,8$ or $\lambda = 0$

when $\lambda = \cancel{16/3}\,8$

$$\begin{bmatrix} 8/3 - 16/3 & 8/3 \\ 8/3 & 8/3 - 16/3 \end{bmatrix} V_1 = 0$$

$$\begin{bmatrix} -8/3 & 8/3 \\ 8/3 & -8/3 \end{bmatrix} \begin{bmatrix} V_{1,1} \\ V_{1,2} \end{bmatrix} = 0$$

$-\frac{8}{3} V_{1,1} + \frac{8}{3} V_{1,2} = 0$ ...... (1)

$\frac{8}{3} V_{1,1} - \frac{8}{3} V_{1,2} = 0$ ...... (2)

from (1) and (2)

$V_{1,1} = V_{1,2}$ $\qquad V_1 = k_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

when $\lambda = 0$

$$\begin{bmatrix} 8/3 & 8/3 \\ 8/3 & 8/3 \end{bmatrix} V_1 = 0$$

$$\begin{bmatrix} 8/3 & 8/3 \\ 8/3 & 8/3 \end{bmatrix} \begin{bmatrix} V_{1,1} \\ V_{1,2} \end{bmatrix}$$

$\frac{8}{3} V_{1,1} + \frac{8}{3} V_{1,2} = 0$

when $\lambda = 8$,

$$\begin{bmatrix} 4-8 & 4 \\ 4 & 4-8 \end{bmatrix} V_1 = 0$$

$$\begin{bmatrix} -4 & 4 \\ 4 & -4 \end{bmatrix} \begin{bmatrix} V_{1,1} \\ V_{1,2} \end{bmatrix} = 0$$

$$-4 V_{1,1} + 4 V_{1,2} = 0$$

$$4 V_{1,1} - 4 V_{1,2} = 0$$

$\therefore$ $V_{1,1} = V_{1,2}$

$$V_1 = k_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

when $\lambda = 0$

$$\begin{bmatrix} 4 & 4 \\ 4 & 4 \end{bmatrix} V_2 = 0$$

$$4 V_{2,1} + 4 V_{2,2} = 0$$

$$V_{2,1} = -V_{2,2}$$

$$V_2 = k_2 \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Arranging eigen values in descending order → 8, 0

$\therefore$ Proportion of variance explained by the 1st principal component is → $\dfrac{8}{8+0} = 1$

So $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ → significant eigen vector
[row feature vector] → $[1\ 1]$
transposed

Row Zero Mean Data → $\begin{bmatrix} -2 & 0 & 2 \\ -2 & 0 & 2 \end{bmatrix}$

Final Data $= [1\ 1] \begin{bmatrix} -2 & 0 & 2 \\ -2 & 0 & 2 \end{bmatrix}$

$= [-4\ \ 0\ \ 4]$

Q) Finding the Rank of a Matrix

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

$R_2 \to R_2 - R_1$

$\begin{bmatrix} 1 & 2 & 5 \\ 0 & 1 & -8 \\ 0 & 4 & 4 \end{bmatrix}$

$R_3 \to R_3 - 2R_1$

$\begin{bmatrix} 1 & 2 & 5 \\ 0 & 1 & -8 \\ 0 & 0 & -6 \end{bmatrix}$

This is Row Echlon Form

$\therefore$ $\boxed{\text{Rank} = 3}$ because there are 3 pivots 1, 1 & -6.

Q) Show that the equations

$x+y+z=6$
$x+2y+3z=14$
$x+4y+7z=30$ are consistent and solve them.

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 6 \\ 1 & 2 & 3 & 14 \\ 1 & 4 & 7 & 30 \end{array}\right]$$

$R_2 \to R_2 - R_1$
$R_3 \to R_3 - R_1$

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 6 \\ 0 & 1 & 2 & 8 \\ 0 & 3 & 6 & 24 \end{array}\right]$$

$R_3 \to R_3 - 3R_2$

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 6 \\ 0 & 1 & 2 & 8 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

$\therefore$ Rank $=2$ for both matrix & augmented matrix

Number of variables $=3$ $\therefore$ infinite solutions.

- If rank(A) = rank$([A,b])$ < Number of Variables [Then infinite solutions]
- If rank(A) = rank$([A,b])$ = Number of Variables [unique solution]
- If rank$(A) \neq$ rank$([A,b])$ $\to$ (No solution)

$x+y+z=6$ Let $x = c$ ↓ constant

$y+2z=8$

~~$\therefore y=$~~

$\therefore y = 8-2z$

$\therefore \cancel{x} c + 8 - 2z + z = 6$

$\therefore -z = 6-8-c$

$\therefore -z = -2-c$

$\therefore z = 2+c$

$\therefore \boxed{z = c+2}$

$\therefore y = 8-2z$

$= 8-(c+2)2$

$= 8-2c-4$

$\boxed{y = 4-2c}$

Note: Reduced Row Echlon Form (RREF)

$$REF = \begin{bmatrix} 1 & 2 & 5 \\ 0 & 1 & -8 \\ 0 & 0 & -6 \end{bmatrix}$$

$R_3 \rightarrow$

* Link between rank & determinant

- If rank[A] = rank[Augmented Matrix = Number of variables

[unique solution]

This will mean, that ~~~~ columns are linearly independent.

~~i.e. they are~~

$\therefore$ determinant will not be 0.

$\therefore$ $A^{-1} = \frac{1}{|A|} \times adj(A)$

$\therefore$ A^{-1} will exist.

- If there are infinitely many solutions

$\therefore$ determinant will be 0

$\therefore$ A^{-1} will not exist.

* Eigen Values & Eigen Vectors

$Ax = \lambda x$

when you apply linear transformation on eigen vectors it will give eigen value $\times$ eigen vector.

So eigen vectors can only scale, they won't transform. This scaling factors are eigen values.

$|A - \lambda I| = 0 \rightarrow$ ~~Characte~~ Characteristic Equation

Number of eigen values $\Rightarrow n$ for a $n \times n$ matrix eg→ $A = \begin{bmatrix} -5 & 2 \\ -7 & 4 \end{bmatrix}$

$|A-\lambda I| = \det\left(\begin{bmatrix} -5 & 2 \\ -7 & 4 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}\right)$

$= \det\left(\begin{bmatrix} -5-\lambda & 2 \\ -7 & 4-\lambda \end{bmatrix}\right)$

$= (-5-\lambda)(4-\lambda) - (-7)(2)$

$= -20 + 5\lambda - 4\lambda + \lambda^2 + 14$

$= \lambda^2 + \lambda - 6$

$= \lambda^2 + 3\lambda - 2\lambda - 6$

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

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

$\therefore \lambda = 2$ & $\lambda = -3$

$Ax = \lambda x$

when $\lambda = 2$

$\cancel{\begin{bmatrix} -5 & 2 \\ -7 & 4 \end{bmatrix}}$ $\begin{bmatrix} -5-2 & 2 \\ -7 & 4-2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0$

$\begin{bmatrix} -7 & 2 \\ -7 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

$-7x_1 + 2x_2 = 0$

$-7x_1 + 2x_2 = 0$

$\therefore x_1 = 2$ & $x_2 = +7$

~~$\therefore x_1 = 1, x_2 = 1$~~

$\therefore k\begin{bmatrix} 2 \\ 7 \end{bmatrix}$ ← solution

- Eigen values ~~&~~ are defined only in square matrix.
- Only square ~~and~~ and full ranked matrix is invertible.

when $\lambda = -3$

$$\begin{bmatrix} -5+3 & 2 \\ -7 & 4+3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} -2 & 2 \\ -7 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$-2x_1 + 2x_2 = 0$$
$$-7x_1 + 7x_2 = 0$$

$\therefore \quad k \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ <- solution

or

$$Ax = \lambda x$$

$$\begin{bmatrix} -5 & 2 \\ -7 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 2 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$A = \cancel{PD}$ eigen vector $\times$ eigen value $\times$ (eigen vector)^T

$= \begin{bmatrix} 2 & 1 \\ 7 & 1 \end{bmatrix}$

You can find eigenvectors by this also.

$$\begin{bmatrix} -5 & 2 \\ -7 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = -3 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

* Trace → Sum of Diagonals

of A $\begin{cases} \text{Trace} = \text{sum of eigen values} \\ \text{determinant} = \text{product of eigen values.} \end{cases}$

with this also we can also find eigen values.

* Singular Value Decomposition

$$A_{m\times n} = U_{m\times n} \Sigma_{m\times n} V^T_{n\times n}$$

where Σ → diagonal

$$\underbrace{AA^T}_{m\times m} = U \Sigma_{m\times n} \underbrace{V^T V}_{I_n} \Sigma^T_{n\times m} U^T$$

$$\underbrace{\Sigma_{m\times n} \times I_n \times \Sigma^T_{n\times m}}$$

This is a square diagonal matrix ($m\times m$) with first n rows & cols with σ_n^2

$$\begin{bmatrix} \sigma_1^2 & & & \\ & \ddots & & \\ & & \sigma_n^2 & \\ & & & \vdots \\ & & & 0 \end{bmatrix}$$

assuming $m > n$

sigma x identity x sigma

eg →

$$\begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \\ 0 & 0 \end{bmatrix} \underbrace{\begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \end{bmatrix}}_{\Sigma}$$

$$= \begin{bmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \sigma_2^2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Eigen vectors of a matrix ~~is~~ are orthogonal if they have different eigen values.

$$\therefore \underbrace{AA^T}_{m\times m} = U \begin{bmatrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_n^2 \\ 0 & \cdots & 0 \end{bmatrix} V^T$$

$$A_{m\times n} = U_{m\times m} \Sigma_{m\times n} V^T_{n\times n}$$

$$A^T = V \Sigma^T U^T$$

$$A^TA = V\Sigma^T U^T . U\Sigma V^T$$
$$= V\Sigma^T I_m \Sigma V^T$$

$$A^TA = V \begin{bmatrix} \sigma_1^2 & \cdots & \\ \vdots & \ddots & \\ & & \sigma_n^2 \end{bmatrix} V^T$$

$\sigma_i \rightarrow$ eigen value
$V \rightarrow$ eigen vector

eg→ $A = \begin{bmatrix} 1 & 0 & 1 \\ -2 & 1 & 0 \end{bmatrix}_{2\times 3}$ find value of V

Ans:

$$A^TA = \begin{bmatrix} 1 & -2 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ -2 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 5 & -2 & 1 \\ -2 & 1 & 0 \\ & & \end{bmatrix}$$

After eigen value decomposition.

$$A^TA = \begin{bmatrix} 5/\sqrt{30} & 0 & -1/\sqrt{6} \\ -2/\sqrt{30} & 1/\sqrt{5} & -2/\sqrt{6} \\ 1/\sqrt{30} & 2/\sqrt{5} & 1/\sqrt{6} \end{bmatrix} \begin{bmatrix} 6 & & \\ & 1 & \\ & & 0 \end{bmatrix} V^T$$

$$\Sigma = \begin{bmatrix} \sqrt{6} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}_{2\times 3}$$

$$\underset{2\times 2}{AA^T} = \begin{bmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ -2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix} \begin{bmatrix} 6 & \\ & 1 \end{bmatrix} V^T$$

$\Downarrow$

$$A = U \Sigma V^T$$

$$= U \begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_n \end{bmatrix} V^T$$

$$\sigma_1 \geq \sigma_2 \geq \cdots > \sigma_n$$

- Best Rank 1 approximation

$$A_{r_1} = U \begin{bmatrix} \sigma_1 & & & \\ & 0 & & \\ & & 0 \; 0 & \\ & & & 0 \end{bmatrix} V^T = \cancel{\sigma_1 u_1 v_1^T} \; \sigma_1 u_1 v_1^T$$

$$= [u_1 \; u_2 \; \ldots \; u_m] \begin{bmatrix} \sigma_1 & & & \\ & 0 & & \\ & & 0 \; 0 & \\ & & & 0 \end{bmatrix} \begin{bmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_n^T \end{bmatrix}$$

$$= [u_1\sigma_1 \; 0 \; 0 \; 0 \; \ldots] \begin{bmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_n^T \end{bmatrix} = \sigma_1 u_1 v_1^T$$

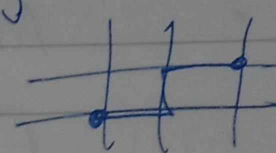
$$A_{\text{rank}\to r} = \sum_{i=1}^{r} \sigma_i u_i v_i^T$$

Rank 2 Approximation

eg $\to$ $A_{r_2} = [u_1\sigma_1 \; u_2\sigma_2 \; 0 \; 0] \begin{bmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_n^T \end{bmatrix} = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T$

Norm

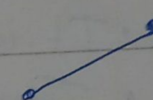
- $||x||_1 = \sum_{i=1}^{n} |x_i|$ → manhattan norm



$||[1 \ -1 \ 2]|| = 4$

- Euclidean norm

$||x||_2 = \sqrt{\sum_{i=1}^{n} x_i^2} = \sqrt{x^T x}$ → diagonal route



- Norm for a Matrix

$||A|| \geq 0$, $||A|| = 0$

$||\alpha A|| = |\alpha| \, ||A||$

$||A+B|| \leq ||A|| + ||B||$

$||AB|| \leq ||A|| \, ||B||$

- Max norm
- schatt en - norms

- Orthogonal Matrix
  - ~~A matrix~~ Square Matrix with only real numbers
  - $AA^T =$ Identity Matrix
  - $A^T = A^{-1}$
  - All identity matrices are orthogonal
  - Product of 2 orthogonal matrix ⇒ Orthogonal
  - Determinant of orthogonal matrix is ± 1
  - It is symmetric
  - Transpose & inverse is also orthogonal
  - Eigenvalues → ± 1

    Eigenvectors → orthogonal & real

- Row Matrix

  1 row and n columns

- Column Matrix

  ~~1~~ 1 col~~s~~ and n rows

- Square Matrix

  n x n ~~col~~

- Rectangular Matrix

  m x n, where m & n are not equal

- Diagonal Matrix

  n x n matrix, therefore square matrix

  All elements are 0 except the diagonals, some elements in the diagonal can also be 0.

- Scalar Matrix

  It is a diagonal matrix, with all elements in the diagonal same.

- Symmetric Matrix
- Transpose
  rows are interchanged to cols & viceversa

when $A = A^T$, then symmetric matrix

- Skew Symmetric Matrix
  $$A = -A^T$$

- Idempotent Matrix
  $$A^2 = A$$
  Eigen Values ~~of~~ $\Rightarrow$ 1 and 0

Kaleigh Hamilton Theorem $\rightarrow$ Every square matrix satisfies its characteristic equation.

$\lambda = \underline{\underline{1, 1}}, 2 \longrightarrow$ A.M = 1

A.M = 2

A.M $\rightarrow$ No. of times eigen Value gets repeated

G.M $\rightarrow$ $\underline{\underline{\text{dim.}}}$ of Eigenspace

* Diagonalising any Matrix

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

1) Find all eigen values
2) Find all eigen vectors

If $AM = GM$, then diagonalisable

AM $\downarrow$ Algebraic Multiplicity

GM $\downarrow$ Geometric Multiplicity

$\lambda = 1, -2, 3$

$\downarrow$

$[\;]$

Check only if eigenvalues are same.

If they are not same/

$$P^{-1}\underline{AP} = \left(\text{Diagonal Matrix}\right)$$

$$\left[\,[\;]\,[\;]\,[\;]\,\right]_{3\times 3}$$

where P is the matrix of eigenvectors

$\lambda = 1$

$(A - \lambda I)x = 0$

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}_{3\times 3} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$x_1 + 2x_2 + 3x_3 = 0$

$x_1 = -2t + 3k$

$x_2 = t$

$x_3 = k$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -2t + 3k \\ t \\ k \end{bmatrix}$$

$$= t\begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + k\begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}$$

$\dim(\quad) = 2$

$$t\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

$\downarrow$ P

→ Eigen Value Decomposition

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

$$|A - \lambda I| = \begin{bmatrix} 5 & 2 & 0 \\ 2 & 5 & 0 \\ 4 & -1 & 4 \end{bmatrix} - \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix}$$

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

$$\det = (5-\lambda)\left[(5-\lambda)(4-\lambda) - 0\right] - 2\left[2(4-\lambda) - 0\right] + 0$$

$$= (5-\lambda)\left[20 - 4\lambda - 5\lambda + \lambda^2\right] - 2\left[8 - 2\lambda\right]$$

$$= 100 - 45\lambda + 5\lambda^2 - 20\lambda + \cancel{4}9\lambda^2 - \lambda^3 - 16 + 4\lambda$$

$$= -\lambda^3 + 14\lambda^2 - 61\lambda + 84$$

<u>or</u>

$$\det = (4-\lambda)\left((5-\lambda)^2 - 4\right) = (4-\lambda)(25 + \lambda^2 - 10\lambda - 4)$$

$$= (4-\lambda)(21 + \lambda^2 - 10\lambda)$$

$\cancel{\therefore \lambda = 4}$ $\lambda = 4$ $\qquad \therefore \lambda = 3, 4, 7$

$$Ax = \lambda x$$

when $\lambda = 3$

$$\begin{bmatrix} 5 & 2 & 0 \\ 2 & 5 & 0 \\ 4 & -1 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 3 \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$5x_1 + 2x_2 = 3x_1 \qquad 2x_1 + 5x_2 = 3x_2$$

$$\boxed{x_1 = -x_2}$$

$4x_1 - x_2 + 4x_3 = 3x_3$

$4x_1 + x_1 + 4x_3 = 3x_3$

$5x_1 + 4x_3 = 3x_3$

$5x_1 = -x_3$

$x_1 = \frac{-1}{5} x_3$

$\therefore \boxed{x_3 = -5x_1}$

$\therefore$ Eigen vectors $\Rightarrow \begin{bmatrix} 1 \\ -1 \\ -5 \end{bmatrix}$

If $\lambda = 4$

$5x_1 + 2x_2 = 4x_1$

$2x_2 = -x_1$

$\boxed{\therefore x_2 = \frac{-1}{2} x_1}$

$4x_1 - x_2 + 4x_3 = 4x_3$

$\therefore x_3 = k$

$2x_1 + 5x_2 = 3x_2$

$2x_1 = -2x_2$

$-x_1 = x_2$

SVD

$$A_{m\times n} = \underset{m\times m}{U} \underset{m\times n}{\Sigma} V^T_{n\times n}$$

where, U & $V \Rightarrow$ orthonormal $\rightarrow$ [orthogonal & norm = 1]

$\therefore UU^T = I$, $VV^T = I$

$$AA^T = (U\Sigma V^T)((V^T)^T \Sigma^T U^T) \rightarrow \text{Since, } (AB)^T = B^T A^T$$

$$= U\Sigma \underbrace{V^T V}_{I \text{ because orthogonal}} \Sigma^T U^T$$

$$\boxed{AA^T = U\Sigma\Sigma^T U^T}$$ $\rightarrow$ In order to calculate U, we need to calculate AA^T

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

$$\boxed{A^T A = V\Sigma^T\Sigma V^T}$$ $\rightarrow$ In order to calculate V, we need to calculate $A^T A$

- Note: Since, A is a rectangular matrix, $\therefore AA^T$ & $A^T A$ will both be a square matrix.

- Note: $\Sigma_{m\times n}$ is a rectangular matrix, with diagonals containing square roots of eigen values of ~~$A^T A$~~ AA^T or $A^T A$.

eg $\rightarrow$ $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \end{bmatrix}$

— Steps

i) $A_{m\times n}$

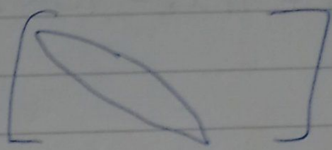
ii) A^T, AA^T, $A^T A$

iii) Calculate eigen values & eigen vectors of AA^T & $A^T A$

iv) $U \rightarrow$ eigen vector of AA^T

v) $V \rightarrow$ eigen vector of A^TA

V^T

vi) $\Sigma \rightarrow$ singular value matrix $\rightarrow$ 

Q) $A = \begin{bmatrix} 3 & 1 & 1 \\ -1 & 3 & 1 \end{bmatrix}_{2\times 3}$

$A^T = \begin{bmatrix} 3 & -1 \\ 1 & 3 \\ 1 & 1 \end{bmatrix}_{3\times 2}$

$AA^T = \cancel{\begin{bmatrix} 11 & \\ & \end{bmatrix}} \begin{bmatrix} 11 & 1 \\ 1 & 11 \end{bmatrix}$

$A^TA = \begin{bmatrix} 10 & 0 & 2 \\ 0 & 10 & 4 \\ 2 & 4 & 2 \end{bmatrix}$

Eigen Values of $AA^T \rightarrow \lambda^2 - ($sum of ~~diagonals~~ ~~roots~~ ~~Equations~~ roots$)\lambda +$ Product of roots

$\rightarrow \lambda^2 - (\text{trace})\lambda + \text{determinant}$

$= \lambda^2 - 22\lambda + 120$

$= \cancel{\lambda^2 - 11\lambda + 60}$

$= \cancel{\lambda^2 - 15\lambda + 4\lambda + 60}$

$= \cancel{\lambda(\lambda - 15) - 4(\lambda}$

$= \lambda^2 - 10\lambda - 12\lambda + 120$

$= \lambda(\lambda - 10) - 12(\lambda - 10)$

$= (\lambda - 12)(\lambda - 10)$

$\therefore \lambda = 10, 12.$

$\lambda_1 = 10$ A^TA $\begin{bmatrix} \sqrt{12} & 0 & 0 \\ 0 & \sqrt{10} & 0 \end{bmatrix}$ 2×3 $\begin{bmatrix} & & \\ & & \end{bmatrix}$ 2×3

$\lambda_2 = 12$ 2

$\lambda_3 = 0$ when $\lambda = 10$

$$\begin{bmatrix} 11 & 1 \\ 1 & 11 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 10 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$11x_1 + x_2 = 10x_1$ $\therefore \boxed{x_1 = -x_2}$

$x_1 + 11x_2 = 10x_2$ $\therefore \boxed{x_1 = -x_2}$

$\therefore$ Eigen vectors $= \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

when $\lambda = 12$

$$\begin{bmatrix} 11 & 1 \\ 1 & 11 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 12 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$11x_1 + x_2 = 12x_1$ $\therefore x_2 = x_1$ $\therefore \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$x_1 + 11x_2 = 12x_2$ $\therefore x_1 = x_2$

$\therefore$ eigen vectors $\Rightarrow$ $\cancel{\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}}$ $\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ ← First we consider higher eigen value while arranging

Eigen Values

$A^TA \rightarrow \begin{bmatrix} 10 & 0 & 2 \\ 0 & 10 & 4 \\ 2 & 4 & 2 \end{bmatrix}$

$$\lambda^3 - \underbrace{S_1}_{\text{Trace}}\lambda^2 + \underbrace{S_2}_{\text{sum of minors of diagonals}}\lambda - \underbrace{S_3}_{\text{determinant}} = 0$$

or

$|A - \lambda I| \rightarrow \begin{bmatrix} 10-\lambda & 0 & 2 \\ 0 & 10-\lambda & 4 \\ 2 & 4 & 2-\lambda \end{bmatrix} \rightarrow (10-\lambda)\left[(10-\lambda)(2-\lambda) - 16\right]$

$+ 2\left[0 - (10-\lambda)2\right]$

4

Page No.:
Date:

$$= (10-\lambda)^2(2-\lambda) - 16(10-\lambda) - 2(10-\lambda)2$$

$$= (10-\lambda)\left[(10-\lambda)(2-\lambda) - 16 - 4\right]$$

$$= (10-\lambda)\left[(20-10\lambda-2\lambda+\lambda^2) - 20\right]$$

$$= (10-\lambda)\left[20-12\lambda+\lambda^2-20\right]$$

$$= (10-\lambda)(\lambda^2-12\lambda) = (10-\lambda)(\lambda)(\lambda-12)$$

$$\therefore \lambda = 10, 0, 12$$

$$Ax = \lambda x$$

when $\lambda = 0$

$$\begin{bmatrix} 10 & 0 & 2 \\ 0 & 10 & 4 \\ 2 & 4 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$$

$10x_1 + 2x_3 = 0$ $\quad \therefore x_1 = \frac{-2x_3}{10} = -\frac{1}{5}x_3 \quad \therefore x_3 = -5x_1$

$10x_2 + 4x_3 = 0$ $\quad x_2 = \frac{-4x_3}{10} = -\frac{2}{5}x_3$

$2x_1 + 4x_2 + 2x_3 = 0$ $\quad \therefore 2x_1 = x_2$

$\therefore$ eigen vector $\rightarrow \begin{bmatrix} 1 \\ 2 \\ -5 \end{bmatrix}$

when $\lambda = 10$

$$\begin{cases} 10x_1 + 2x_3 = 10x_1 \\ 10x_2 + 4x_3 = 10x_2 \\ 2x_1 + 4x_2 + 2x_3 = 10x_3 \end{cases} \quad \therefore x_3 = 0$$

$2x_1 = -4x_2$

$\therefore x_1 = -2x_2$

$\therefore$ eigen vector $\rightarrow \begin{bmatrix} 1 \\ -1/2 \\ 0 \end{bmatrix}$

when $\lambda = 12$

$10x_1 + 2x_3 = 12x_1$ $\quad \therefore x_1 = x_3$

$10x_2 + 4x_3 = 12x_2$ $\quad \therefore x_2 = 2x_3$

$$\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

$\therefore$ Eigen Vectors $\begin{bmatrix} 1 & 1 & 1 \\ 2 & -1/2 & 2 \\ 1 & 0 & -5 \end{bmatrix} \leftarrow V$

$\therefore$

$$AA^T = \begin{bmatrix} 11 & \cancel{1}1 \\ 1 & 11 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}}_{U} \begin{bmatrix} 12 & \\ & 10 \end{bmatrix}$$

A

$$\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \sqrt{12} & 0 & 0 \\ 0 & \sqrt{10} & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 & 1 \\ & & \end{bmatrix}$$

σ_1, σ_2

$$\begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \; 5 \times 3$$

Norm

1) $\|A\|_1$ 1-norm $\rightarrow$ (Column norm)

$$\begin{bmatrix} 1 & 2 \\ -2 & 3 \\ 3 & 4 \end{bmatrix} \rightarrow \max\left[|1| + |-2| + |3|, |2| + |3| + |4|\right]$$

$=$

2) ∞-norm $\rightarrow$ [row norm]

$$\max(|1| + |2|, |-2| + |3|, |3| + |4|) =$$

3) 2-norm $\rightarrow$ [spectral norm]

$$\sqrt{\max(\text{Eigen Values of } A^TA)} \rightarrow \sigma_1(A)$$

Largest Singular Value

4) $$\|A\|_F = \sqrt{\text{trace}(A^TA)}$$

eg $\rightarrow$ $A = \begin{bmatrix} 0 & 1 & 1 \\ \sqrt{2} & 2 & 0 \\ 0 & 1 & 1 \end{bmatrix}$

Ans: 1 norm $\rightarrow \max(\sqrt{2}, 4, 2) = 4$

∞ norm $\rightarrow \max(2, 2+\sqrt{2}, 2) = 2+\sqrt{2}$

~~2 norm~~ $\rightarrow$

$$A^TA = \begin{bmatrix} 2 & 2\sqrt{2} & 2 \\ 2\sqrt{2} & 6 & 2 \\ 2 & 2 & 2 \end{bmatrix}$$

$\therefore \|A\|_F = \sqrt{10}$

Eigen values $\Rightarrow 2, 0, 8 \quad \therefore \|A\|_2 = \sqrt{8}$

Derivative of a Matrix

$\frac{d}{dx}(Ax)$

eg→ $Ax = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

$= \begin{bmatrix} x_1 + 2x_2 \\ 3x_1 + 4x_2 \end{bmatrix} \begin{matrix} \to f_1(x_1, x_2) \\ \to f_2(x_1, x_2) \end{matrix}$

$\frac{d}{dx}(Ax) = \begin{bmatrix} \frac{df_1}{dx_1} & \frac{df_2}{dx_1} \\ \frac{df_1}{dx_2} & \frac{df_2}{dx_2} \end{bmatrix} \begin{matrix} \to x_1 \\ \to x_2 \end{matrix}$

(differentiate function 1 & 2 w.r.t x_1)

(differentiate function 1 & 2 w.r.t x_2)

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

- ∴ if there are 3 functions & 4 variables, we will get 4×3 matrix.

* Derivative of $x^T A x$

$\underset{1\times2}{\begin{bmatrix} x_1 & x_2 \end{bmatrix}} \underset{2\times2}{\begin{bmatrix} a_{11} & a \\ a & a_{22} \end{bmatrix}} \underset{2\times1}{\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}}$

$= [x_1\ x_2] \begin{bmatrix} a_{11}x_1 + ax_2 \\ ax_1 + a_{22}x_2 \end{bmatrix} = \underset{1\times1}{\begin{bmatrix} a_{11}x_1^2 + 2ax_1x_2 + a_{22}x_2^2 \end{bmatrix}}$

↓

Here we just have one function.

$$\begin{bmatrix} \frac{df}{dx_1} \\ df/dx_2 \end{bmatrix} = \begin{bmatrix} 2a_{11}x_1 + 2ax_2 \\ 2a_{22}x_2 + 2ax_1 \end{bmatrix}$$

$$= 2 \begin{bmatrix} a_{11} & a \\ a & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$= 2Ax$$

$$\therefore \frac{d}{dx}(x^T A x) = 2Ax$$

In PCA, we will have a symmetric matrix as the covariance matrix is always symmetric.

* Lagrange Multipliers

$$u^T S u + \lambda(1 - u^T u)$$

$$\frac{d}{du}\left[u^T S u + \lambda(1 - u^T I u)\right]$$

$$0 = 2Su + \lambda(-2Iu)$$

$$\Rightarrow 2Su = 2\lambda I u$$

$$Su = \lambda I u$$

$$S\vec{u} = \lambda \vec{u}$$

$\downarrow$

eigenvector of S

& In PCA, S would be the covariance Matrix
That is why we find eigen vectors of S.

* Covariance Matrix

Apple	Banana	AB
1	1	1
3	-1	-3
-1	-1	1
$E(A) = 1$	$E(B) = -1/3$	$-1/3$

$$Cov(A,B) = E(AB) - E(A)E(B)$$
$$= \frac{-1}{3} - 1 \times \left(\frac{-1}{3}\right)$$
$$= \frac{-1}{3} + \frac{1}{3}$$

$$Cov(A,B) = \frac{\Sigma\left[0 + -2 \times -4/3 + -2 \times -4/3\right]}{3}$$
$$= \frac{-8/3 + 8/3}{3} = 0$$

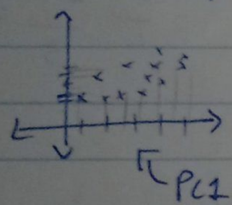
$$\frac{\Sigma(x_i - \bar{x})^2}{N} = \frac{\left((1-1)^2 + (3-1)^2 + (-1-1)^2\right)}{3}$$
$$= \frac{(4+4)}{3} = \frac{8}{3}$$

$$\frac{\Sigma(y_i - \bar{y})^2}{N} = \frac{\left((1+1/3)^2 + 2(-1+1/3)^2\right)}{3}$$
$$= \frac{\frac{16}{9} + 2 \times 4/9}{3} = \frac{\frac{24}{9}}{3} = \frac{8}{9}$$

$$\text{Covariance Matrix} = \begin{array}{c c} & \begin{array}{cc} A & B \end{array} \\ \begin{array}{c} A \\ B \end{array} & \begin{pmatrix} 8/3 & 0 \\ 0 & 8/9 \end{pmatrix} \end{array}$$

* Applications of PCA
1) Dimensionality Reduction
2) Data Visualisation
3) Feature Extraction.

eg-> $\{x_1, \dots x_N\}$ -> variables with dimension D
Goal is to map this data set in dim M where $M < D$.



we will take PC1 in such a way that the variation of the projections is maximum

PC1 is the eigen vectors of the covariance matrix which has the maximum eigen value.

Projections -> $(\vec{x}_i \cdot \vec{u}_1)\vec{u}_1 = (u_1^T x_i)\vec{u}$

Mean of Projections -> $u_1^T \bar{x} \vec{u}$

Variance of Projections -> $\frac{1}{N}\sum_{i=1}^{N}(u_1^T x_i - u^T \bar{x}\vec{u})^2$

$$= \frac{1}{N}\sum_{n=1}(u_1^T x_n - u_1^T \bar{x})^2$$

$$= \frac{1}{N}\sum_{n=1}^{N}[u_1^T(x_n - \bar{x})]^2$$

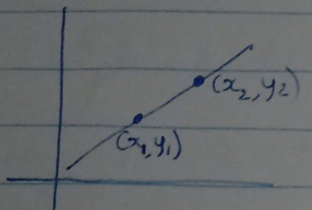
$$= \frac{1}{N}\sum_{n=1}^{N} u_1^T (x_n - \bar{x})(x_n - \bar{x})^T u_1$$

$$= u_1^T \left[\frac{1}{N}\sum_{n=1}^{N}(x_n - \bar{x})(x_n - \bar{x})^T\right] u_1$$

apsara

Date:

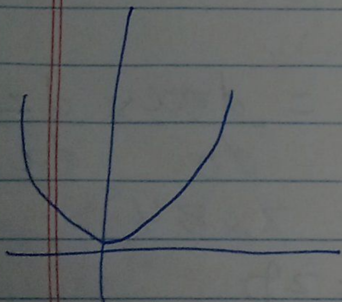
Derivative



Linear

Derivate $= \dfrac{y_2 - y_1}{x_2 - x_1}$

Derivative will be constant throughout.



$x_1 = 5$

$y_1 = x_1^2 = 25$

$x_2 = 5.0000001$

$y_2 = 25.0\ldots$

$y_2 = y_1 + \Delta y$

$y_1 + \Delta y = (x_1 + \Delta x)^2$ where $\Delta y = y_2 - y_1$

$\cancel{y_1} + \Delta y = \cancel{x_1^2} + 2\Delta x x_1 + (\Delta x)^2$

$\cancel{y_2 - y_1}$ $\Delta y = 2\Delta x x_1 + (\Delta x)^2$

Dividing throughout by Δx

$$\frac{\Delta y}{\Delta x} = \frac{2\Delta x x_1}{\Delta x} + \frac{(\Delta x)^2}{\Delta x} = 2x_1 + \Delta x$$

$\Delta x \downarrow$ this is very small so we ignore

$\therefore \dfrac{dy}{dx} = 2x \quad \because y = x^2$

Q) $y = \dfrac{1}{x^3}$

$y = x^{-3}$

$\dfrac{dy}{dx} = -3x^{-4}$

Q) $y = (x+5)(x-2)$

$y = x^2 \cancel{-10x} + 3x - 10$

$\therefore \dfrac{dy}{dx} = 2x + 3$

– Partial Derivatives $\rightarrow$ where we hold some variables as constant.

• Vector Space
• Basis

Date:

Chain Rule

$x = a^2 + 7b$

$y = c^3 + d$

$z = 4x + 3y$

$\frac{dz}{da} = ?$

$\frac{dz}{da} = \frac{dx}{da} \times \frac{dz}{dx}$

$= 2a \times 4$

$= 8a$

$\frac{dz}{db} = \frac{dx}{db} \times \frac{dz}{dx}$

$= 7 \times 4$

$= 28$

$\frac{dz}{dc} = 3c^2 \times 3$

$= 9c^2$

$\frac{dz}{dd} = 1 \times 3$

$= 3$

— Maxima & Minima

steps →

1) $y = f(x)$
2) $f'(x) = 0$ → Critical Point
3) $f''(x)$ -
   If $f''(x) > 0$ → local minima
   If $f''(x) < 0$ → local maxima

Note: After $f'(x) = 0$, we need to check for $f''(x)$ because if the slope is 0 it is not necessary that it will be local maxima or minima.

apsara

Date:

Q) $y = 4x^3 + 12x^2 + 12x + 10$

$$\frac{dy}{dx} = 12x^2 + 24x + 12 = 0$$

$$x^2 + 2x + 1 = 0$$
$$x^2 + x + x + 1 = 0$$
$$x(x+1) + 1(x+1)$$
$$(x+1)^2 = 0$$
$$\therefore x = -1$$

$$\frac{d^2y}{dx^2} = 24x + 24$$

$$\frac{d^2y}{dx^2} = 0$$

$f''(x) = 0 \rightarrow$ saddle point

Q) $y = 3x^4 + 4x^3 - 12x^2 + 12$

$$\frac{dy}{dx} = 12x^3 + 12x^2 - 24x = 0$$

$$= x^3 + x^2 - 2x = 0$$
$$= x(x^2 + x - 2)$$
$$\cancel{= x(x^2 + x} \quad x(x^2 + 2x - x - 2)$$
$$= x(x(x+2) - 1(x+2))$$
$$= x(x-1)(x+2)$$
$$\therefore x = 0, 1, -2$$

$$\frac{d^2y}{dx^2} = 36x^2 + 24x - 24$$

$$= 3x^2 + 2x - 2$$

when $x = 0 \rightarrow f''(x) = -24 \rightarrow$ ~~saddle point~~ local maxima

$x = 1 \rightarrow f''(x) = 3$

$x = -2 \rightarrow f''(x) = 6$

} local minima

Page No.

apsara

~~Higher Dimension~~ function in 2 variables:

Step i) $\frac{df}{dx}$, $\frac{df}{dy}$ Step ii) $\frac{d^2f}{dx^2}$, $\frac{d^2f}{dy^2}$, $\frac{d^2f}{dxdy}$

ii) $\frac{df}{dx} = \frac{df}{dy} = 0$ → stationary or critical points.

iii) $D(x,y) = \underbrace{\left(\frac{d^2f}{dx^2}\right)}_{r}\underbrace{\left(\frac{d^2f}{dy^2}\right)}_{s} - \underbrace{\left(\frac{d^2f}{dxdy}\right)^2}_{t}$

If $D > 0$ & $r > 0$ → local ~~maxima~~ minima
If $D > 0$ & $r < 0$ → local ~~minima~~ maxima
If $D < 0$ → saddle point. If $D = 0$ → some more info needed.

Q) $f(x,y) = 2x^2 + 3y^2 - 12x - 6y + 9$

Ans: i) $f'(x) = 4x - 12$, $f'(y) = +6y - 6$,
ii) $f''(x) = 4$, $f''(y) = +6$, $\frac{d^2f}{dxdy} = \frac{d}{dx}\left(\frac{df}{dy}\right) = 0$
iii) $D(x,y) = 4 \times (+6) - 0 = 24$
$\therefore D(x,y) > 0$
$r > 0$ $\therefore$ local ~~maxima~~ minima

$f'(x) = 4x - 12 = 0$ $\therefore x = 3$, $f'(y) = 6y - 6 = 0$
$\therefore y = 1$

$\therefore$ local ~~maxima~~ minima is at point $(3,1)$

Q) $f(x,y) = -\frac{1}{3}x^3 + \cancel{x - y^2}\; xy^2$

$f'(x) = -\frac{1}{3} \times 3x^2 + y^2 = y^2 - x^2$, $f'(y) = 2x$

Gradient Descent

Gradient → direction of the steepest descent/ascent
→ gradient is a vector whose components are scalar.

* Hessian → $\begin{bmatrix} \frac{d^2 f}{dx^2} & \frac{d^2 f}{dx dy} \\ \frac{d^2 f}{dy dx} & \frac{d^2 f}{dy^2} \end{bmatrix}$

Jacobian → $\begin{bmatrix} \frac{df_1}{dx_1} & \frac{df_1}{dx_2} & \cdots & \frac{df_1}{dx_n} \\ \frac{df_2}{dx_1} & \frac{df_2}{dx_2} & \cdots & \\ & & & \end{bmatrix}$

Transpose of Differentiation.