~~Maxima / minima~~

Maxima / minima of function of Many variables.

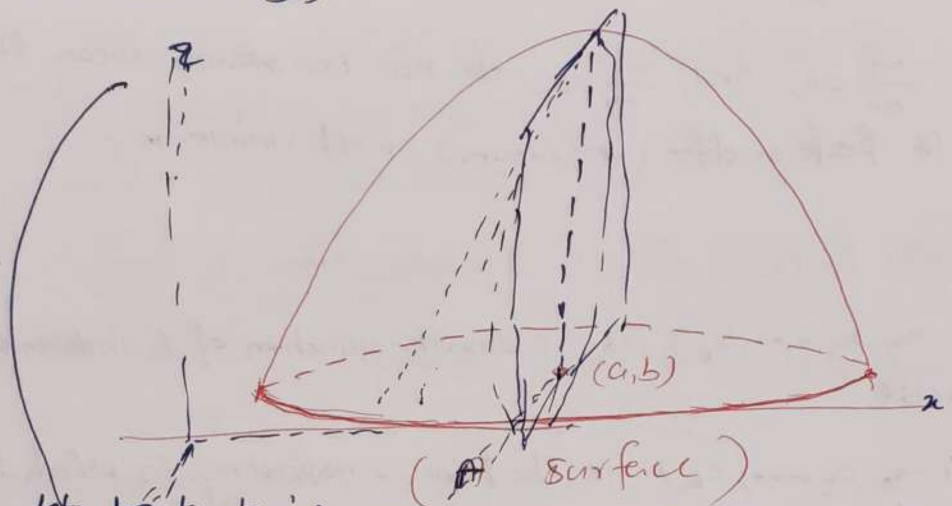
Case I:- function of two variable

Maxima Minima

Function of two variable :—

Let a function of two variables. $z = f(x,y)$

This when plotted gives you a surface like a mountain for $Z > 0$ ~~and valley for $Z < 0$~~.



— Let peak point occurs to the point $(x,y) = (a,b)$ on the xy-plane

Y

— Walk up the mountain along the curve, where the plane $x = a$ cuts the surface.

$\Rightarrow$ Reach the peak over the point (a,b).

— Consider the curve where the plane $y = b$ cuts the mountain $\Rightarrow$ walk up along the curve $\Rightarrow$ reach ~~the~~ over the peak (a,b) also.

$\Rightarrow$ ~~At~~ At the peak over the point (a,b), the tangent plane to the surface is parallel to the (x,y)-plane.

— For $x = b$ (fixed), the curves on the plane $x = b$ passing through peaks, the the curves have sharp corner at peak

$\hookrightarrow$ In this case 'f' is not differentiable with respect to x for $y = b$.

— Similarly for $x = a$, ~~there~~ there may be sharp corner at the peak along y-direction $\Rightarrow \frac{\partial f}{\partial y}$ does not exist.

[one can walk to the peak through many routes]

— When function is smooth

$$\frac{\partial f}{\partial x} = 0 \text{ and } \frac{\partial f}{\partial y} = 0 \quad \text{—— (1)}$$

corresponding to a local maximum or local minimum.

Note :- $\frac{\partial f}{\partial x} = 0$, and $\frac{\partial f}{\partial y} = 0$ do not necessarily mean that there is ~~peak or dip~~ peak (maximum) or dip (minimum).

Case 2 :- For function of k independent variable :—

— $f(x_1, x_2, x_3, \cdots, x_k)$ is a smooth function of k independent variable

— $(c_1, c_2, c_3, \cdots, c_k)$ is the point corresponding to which there is ~~a to~~ a local maximum or minimum, then

$$\frac{\partial f}{\partial x_1} = 0, \cdots, \frac{\partial f}{\partial x_k} = 0 \quad \text{at } (x_1, x_2, \cdots, x_k) = (c_1, c_2, \cdots, c_k)$$

$$D' f_c = 0 \quad \text{—— (2)}$$

— Now expanding $f(x_1, x_2, \cdots, x_k)$ in the neighbourhood of $c = (c_1, c_2, c_3, \cdots, c_k)$, we get

$\cancel{f(c_1+h_1, c_2+h_2, \cdots, c_k+h_k) = f(c_1, c_2, \cdots, c_k) +}$

$$f(c+h) = f(c) + \frac{1}{1!} h' D f_c + \frac{1}{2!} h' (D D' f_c) h + \cdots$$

$$\Rightarrow \quad f(c+h) - f(c) = h' D f_c + \frac{1}{2!} h' (D D' f_c) h + \cdots$$

i.e. (neighbourhood value) − (value at the peak or dip) = R.H.S.

Here $h = [h_1, h_2, \cdots, h_k]$ $\qquad h' D f_c = 0$

The leading term on the right side is:

$$\frac{1}{2!} h' (D D' f_c) h = \frac{1}{2!} [h_1, h_2, \cdots, h_k] \begin{bmatrix} \frac{\partial^2 f_c}{\partial x_1^2} & \cdots & \frac{\partial^2 f_c}{\partial x_1 \partial x_k} \\ \vdots & \cdots & \vdots \\ \frac{\partial^2 f_c}{\partial x_k \partial x_1} & \cdots & \frac{\partial^2 f_c}{\partial x_k^2} \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_k \end{bmatrix} \quad \text{—— (3)}$$

— This ~~term~~ determine the sign of whole R.H.S. when $h_1, h_2, \dots, h_k$ are infinitesimally small, where f_c means, after taking the derivatives evaluated at $x = c = (c_1, c_2, \dots, c_k)$

Maximum (local)

— If the ③ remains negative for all $h_1, h_2, \dots, h_k$ in the neighborhood of $c = (c_1, c_2, \dots, c_k)$ then from left side

Value of f at c ($= f_c$) is bigger than the neighborhood value ~~or f at c is local~~

$\Leftrightarrow \quad f(c+h) - f(c) < 0 \quad \forall h.$

$\Rightarrow$ ~~f~~ point 'c' is local maximum.

Local minimum (

~~f(c+h)~~ R.H.S ③ is positive

for all $h_1, h_2, \dots, h_k$ in the neighbourhood of $c = (c_1, c_2, \dots, c_k)$ then c correspond to a minimum.

i.e. $f(c+h) - f(c) > 0 \quad \forall h.$

$\Rightarrow$ Thus local maxima/minima is determine by the quadratic form $h'Ah$ being negative definite or positive definite where

$$A = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_k} \\ \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_k \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_k^2} \end{bmatrix} \quad \text{(4)}$$

; f_c = f evaluated at $c = (c_1, c_2, \dots, c_k)$

Procedure :— For all ~~smooth~~ smooth function compute $\frac{\partial f}{\partial x_i}$, $i = 1, 2, \dots, k$ and equate them ~~to~~ zero i.e.

$$\frac{\partial f}{\partial x_1} = 0, \dots, \frac{\partial f}{\partial x_k} = 0, \text{ or } \frac{\partial f}{\partial X} = \underset{\text{zero vector}}{0} \quad \text{(5)}$$

- Solve equation (5) for critical points.

(1) There may be not any solution. Then there is no critical point

(2) There may be unique solution. Then there is only one critical point

(3) There may be ~~sever~~ several solutions.
↳ Then there are several critical points

Let $c = (c_1, c_2, \cdots, c_n)$ one such solution.
↳ Does the poin c is critical point (local minima/local maxima) or something else?

⇓ For checking this

compute matrix A given in (4)

so check whether A is positive definite, negative definite or indefinite.

- If A is positive definite then c is minimum ~~maxima~~ point
- If A is negative definite the c corresponds to ~~minima~~ maxima
- If A is semi definite or indefinite there is no local maxima or minima at point c
  ↳ In this case we say that c is saddle points.

Note :- How to check definiteness of matrix :-

$$A = (a_{ij}) = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \cdots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{bmatrix}$$

(i):— Basic condition is that definiteness can be defined only on symmetric matrices with real and hermition matrices when in the complex domain.

i.e $a_{ij} = a_{ji}$; ~~when~~ for all i ~~any~~ and j, when it is real.

(ii) Positive Definite :- All leading minors of A are positive. This means:

$$a_{11} > 0, \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \cdots, |A| > 0 \quad \text{—— (1)}$$

where $|(\cdot)|$ means determinant of ~~A~~ $(\cdot)$.

Equivalent condition of (1) that all eigen value of A are positive. Another equivalent condition is that

$h'Ah > 0$ for all non-null h.

(iii) Negative Definiteness :— — A is positive definite, then A is negative definite.

- All leading minors are alternatively negative and positive.

or

- All the odd order minors are ~~neg~~ negative and the all the even order minors are positive. i.e.

$$a_{11} < 0, \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \cdots \quad \text{—— (2)}$$

- An equivalent condition is that all eigen values are negative
- Anothe equivalent condition is that $h'Ah < 0$ for all non-null h.

Now for $k=2$ the condition is reduced to following:—

For critical point $C = (c_1, c_2)$ to correspond to a maximum

$$\frac{\partial^2 f_c}{\partial x_1^2} < 0,$$

$$\begin{vmatrix} \frac{\partial^2 f_c}{\partial x_1^2} & \frac{\partial^2 f_c}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f_c}{\partial x_2 \partial x_1} & \frac{\partial^2 f_c}{\partial x_2^2} \end{vmatrix} = \frac{\partial^2 f_c}{\partial x_1^2} \cdot \frac{\partial^2 f_c}{\partial x_2^2} - \frac{\partial^2 f_c}{\partial x_1 \partial x_2} \cdot \frac{\partial f_c}{\partial x_2 \partial x_1}$$

$$= \frac{\partial^2 f_c}{\partial x_1^2} \cdot \frac{\partial^2 f_c}{\partial x_2^2} - \left(\frac{\partial^2 f_c}{\partial x_1 \partial x_2}\right)^2$$

$$> 0 \quad \text{(Negative Definite)}$$

$$\Rightarrow \text{maximum}$$

and for a minimum

$$\left.\begin{array}{l} \frac{\partial^2 f_c}{\partial x_1^2} > 0 \\ \frac{\partial^2 f_c}{\partial x_1^2} \cdot \frac{\partial^2 f_c}{\partial x_2^2} - \frac{\partial^2 f_c}{\partial x_1 \partial x_2} \cdot \frac{\partial^2 f_c}{\partial x_2 \partial x_1} > 0 \end{array}\right\} \rightarrow \text{(Negative Definite)} \Rightarrow \cancel{\text{minimum}}$$

$$\frac{\partial^2 f_c}{\partial x_1^2} \cdot \frac{\partial^2 f_c}{\partial x_2^2} - \left(\frac{\partial^2 f_c}{\partial x_1 \partial x_2}\right)^2 > 0. \Rightarrow \text{minimum.}$$

— If any of these condition is ~~not~~ violated then C is saddle point

Example:— check for the local maxima / minima of the following functions:

(i) $f(x,y) = x^2 + 5y^2 - 3xy - 2y + 7x + 7$

(ii) $f(x,y) = -2x^2 \bullet 3y^2 + 4xy - x + 2y + 8$

(iii) $f(x,y) = 2x^2 - 4y^2 + xy - 3x + 2y - 8$

(iv) $f(x,y) = x^2 + y^2 - 2xy - 7x + 3y - 10$

Solution :- (1) $f(x,y) = x^2+5y^2-3xy-2y+7x+7$

consider $\frac{\partial f}{\partial x}=0$, $\frac{\partial f}{\partial y}=0$, this gives:

$$\frac{\partial f}{\partial x} = 2x-3y+7=0 \quad \text{——— (1)}$$

$$\frac{\partial f}{\partial y} = -3x+10y-2=0 \quad \text{——— (2)}$$

Solving (1) and (2) we have the following

$$x=-\frac{64}{11}, \quad y=-\frac{17}{11}$$

Thus critical point $C=\left(-\frac{64}{11}, -\frac{17}{11}\right)$

C corresponds to maxima or minima or saddle point?

Now consider the second derivatives. Thus from (1) and (2)

$$\frac{\partial^2 f}{\partial x^2}=2, \quad \frac{\partial^2 f}{\partial x \partial y}=-3, \quad \frac{\partial^2 f}{\partial y^2}=10$$

The matrix of second order derivatives is then:

$$A=\begin{bmatrix} 2 & -3 \\ -3 & 10 \end{bmatrix}$$

$$2>0, \quad \begin{vmatrix} 2 & -3 \\ -3 & 10 \end{vmatrix} = \cancel{20-}(2)(10)-(-3)(-3)$$
$$= 20-9$$
$$\begin{vmatrix} 2 & -3 \\ -3 & 10 \end{vmatrix} = 11>0$$

Hence A is positive definite.

$\therefore$ critical point $C=\left(-\frac{64}{11}, -\frac{17}{11}\right)$ is local minimum.

(ii) $\frac{\partial f}{\partial x}=0 \Rightarrow -4x+4y-1=0 \Rightarrow 4x-4y=-1$ ——— (1)

$\frac{\partial f}{\partial y}=0 \Rightarrow 4x-6y+2=0 \Rightarrow 4x-6y=-2$ ——— (2)

Solving (1) and (2) we have critical point

$$x=\frac{1}{4}, \quad y=\frac{1}{2}$$

$\therefore$ critical point $C=\left(\frac{1}{4}, \frac{1}{2}\right)$. C corresponds to local minima, maxima or saddle point

Now evaluate second order derivative at ~~this~~ this point

$$\frac{\partial^2 f}{\partial x^2} = -4, \quad \frac{\partial^2}{\partial y^2} = -6, \quad \frac{\partial^2 f}{\partial x \partial y} = 4$$

$$\frac{\partial^2 f}{\partial y \partial x} = 4$$

The matrix of second order derivatives:

$$A = \begin{bmatrix} -4 & 4 \\ 4 & -6 \end{bmatrix}$$

$\because$ $-4 < 0$, $\begin{vmatrix} -4 & 4 \\ 4 & -6 \end{vmatrix} = (-4)(-6) - (4)(4)$

$= 24 - 16$

$= 8 > 0$

Thus A is negative definite.

Thus critical point is corresponds maximum.

(iii) $f(x,y) = 2x^2 - 4y^2 + xy - 3x + 2y - 8$

$$\frac{\partial f}{\partial x} = 0 \Rightarrow 4x + y - 3 = 0 \Rightarrow 4x + 3y = 3 \quad \text{—— (1)}$$

$$\frac{\partial f}{\partial y} = 0 \Rightarrow x - 8y + 2 = 0 \Rightarrow x - 8y = -2 \quad \text{—— (2)}$$

Solving (1) and (2) we get critical point

$$C' = (x,y) = \left(-\frac{22}{31}, \frac{5}{31}\right)$$

Now find second derivative

$$\frac{\partial^2 f}{\partial x^2} = 4, \quad \frac{\partial^2 f}{\partial y^2} = -8, \quad \frac{\partial^2 f}{\partial x \partial y} = 1, \quad \frac{\partial^2 f}{\partial y \partial x} = 1$$

$$A = \begin{bmatrix} 4 & 1 \\ 1 & -8 \end{bmatrix}, \quad 4 > 0, \quad \begin{vmatrix} 4 & 1 \\ 1 & -8 \end{vmatrix} = (4)(-8) - (1)(1)$$

$= -32 - 1$

$= -33 < 0$

Therefore, matrix is indefinite. Therefore point is a saddle point

Ex

~~Examples:~~

(iv) $f(x,y) = x^2+y^2-2xy-7x+3y-10$

$$\frac{\partial f}{\partial x} = 2x+0-2y-7=0 \Rightarrow 2x-2y=7 \quad \text{---} (1)$$

$$\frac{\partial f}{\partial y} = 2y-2x+3=0 \Rightarrow 2x-2y=3 \quad \text{---} (2)$$

There is no solution. Equations (1) and (2) represent two parallel line. $\Rightarrow$ There ~~was~~ is no critical point.

$\hookrightarrow$ No finite local maximum or local minimum.

Example:- A rectangular box is to be made with iron sheets to hold ~~75~~ 125 cubic meters of sand. What should be the length, width and height so that minimum amount of metal sheet used?

Answer:- Let x, y, z be the length, width, and depth of the box and let V represent the volume:

$$V = xyz = 1\cancel{75}25 \text{ cm}^3$$

The total surface area A is such that:

$$A = 2xy+2yz+2xz$$

Hence we minimize A subject to the condition $xyz = 75$.
Substituting for yz and xz we have:

$$A = 2xy + \frac{(2)(125)}{x} + \frac{(2)(125)}{y} \quad \text{---} (1)$$

Here $\frac{\partial A}{\partial x} = 0 \Rightarrow 2y - \frac{250}{x^2} = 0$

$$\Rightarrow \frac{2x^2y-250}{x^2} = 0 \Rightarrow x^2y = 125 \quad \text{---} (2)$$

$$\frac{\partial A}{\partial y} = 0 \Rightarrow 2x - \frac{150}{y^2} = 0 \Rightarrow xy^2 = 125 \quad \text{---} (3)$$

From (2) and (3) we have $125x = 125y \Rightarrow x = y$

From the symmetry of the problem

~~$x=y=z \Rightarrow x=(75)^{1/3} \Rightarrow 5\ x=y=z=$~~

$\because \; x = y = z \Rightarrow x^3 = 125$

$\therefore \; x = 5 = y = z.$

Thus critical point $c = (x, y, z) = (5, 5, 5)$

$\frac{\partial A}{\partial x} = 2y + 2z \Rightarrow \frac{\partial A_c}{\partial x} = 20, \quad \frac{\partial^2 A_c}{\partial x^2} = 0$

$\frac{\partial A}{\partial y} = 2x + 2z \Rightarrow \frac{\partial A_c}{\partial x} = 20, \quad \frac{\partial^2 A_c}{\partial y^2} = 0$

$\frac{\partial A}{\partial z} = 2y + 2x \Rightarrow \frac{\partial A_c}{\partial z} = 20; \quad \frac{\partial^2 A_c}{\partial z^2} = 0$

$\frac{\partial^2 A}{\partial x \partial y} = 2; \quad \frac{\partial^2 A}{\partial x \partial z} = 2;$

$\frac{\partial^2 A}{\partial y \partial x} = 2; \quad \frac{\partial^2 A}{\partial y \partial z} = 2;$

$\frac{\partial^2 A}{\partial z \partial x} = 2; \quad \frac{\partial^2}{\partial z \partial y} = 2;$

$$\nabla^2 A = \begin{bmatrix} \frac{\partial^2 A}{\partial x^2} & \frac{\partial^2 A}{\partial x \partial y} & \frac{\partial^2 A}{\partial x \partial z} \\ \frac{\partial^2 A}{\partial y \partial x} & \frac{\partial^2 A}{\partial y^2} & \frac{\partial^2 A}{\partial y \partial z} \\ \frac{\partial^2 A}{\partial z \partial x} & \frac{\partial^2 A}{\partial z \partial y} & \frac{\partial^2 y}{\partial z^2} \end{bmatrix} = \begin{bmatrix} 0 & 2 & 2 \\ 2 & 0 & 2 \\ 2 & 2 & 0 \end{bmatrix}$$

$0 \ngtr 0$

The critical point $(x,y,z) = (5,5,5)$ corresponds to a minimum for A because from equation (i) when x or y tends to infinitesimally small A explodes to huge quantity. Thus a minimum area

$$A = 2(xy + yz + xz) = 2(25+25+25) = 150 \text{ sq.m} \equiv 150 m^2$$

Note:- Find the value of A near point $C = (5,5,5)$ subject to $xyz = \cancel{75}\ 125$?

Problem For Exercise:

(1) Check for maxima / ~~and~~ minima for the following

(i) $f(x,y) = x^2+2y^2-5x+6y-8$

(ii) $f(x,y) = 2x^2+5y^2-2x+3y+5$

(iii) $f(x,y) = \dfrac{x-y}{x+y}$

(iv) $f(x,y) = \dfrac{xy}{x^2+y^2}$

(v) $f(x,y) = \sin x + \cos(x+y)$

(2) For the following data on time ~~time~~ t and the corresponding weight gain fit a linear model $w = a+bt$ by the ~~following~~ method of least squares:

Data:					
t	0	1	2	3	4
w	10	11	14	17	18

Hint:-

Model: $w_j = a + bt_j + e_j$; $j = 1, 2, 3, \dots, n$.

Error sum of squares:

$$\sum_{j=1}^{n} e_j^2 = \sum_{j=1}^{n} (w_j - a - bt_j)^2.$$

Method of least square $\Rightarrow$ minimize the error sum of squares. $\Rightarrow$ Estimate a and b and then consider the model $w = \hat{a} + \hat{b}t$, where $\hat{a}$ and $\hat{b}$ are least square estimates.

$$\boxed{\frac{\partial}{\partial a}\left(\sum_{j=1}^{n} e_j^2\right) = 0, \quad \frac{\partial}{\partial b}\left(\sum_{j=1}^{n} e_j^2\right) = 0}$$

Formulate vector matrix formulation

$$w_j = a + bt_j + e_j \,; \quad j = 1, 2, 3, \cdots, n.$$

Take $\bar{W} = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ \vdots \\ w_n \end{bmatrix}$, $\quad \bar{\theta} = \begin{bmatrix} a \\ b \end{bmatrix}$,

$\bar{H} = \begin{bmatrix} 1 & t_1 \\ 1 & t_2 \\ 1 & t_3 \\ \vdots & \vdots \\ 1 & t_n \end{bmatrix}$ $\quad \bar{e} = \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ \vdots \\ e_n \end{bmatrix}$

$\bar{W} = \bar{H}\bar{\theta} + \bar{e}$ —— observation model

$$J(\bar{\theta}) = \bar{e}^T\bar{e} = (\bar{W} - \bar{H}\bar{\theta})^T(\bar{W} - \bar{H}\bar{\theta}) = tr\left[(\bar{W} - \bar{H}\bar{\theta})(\bar{W} - \bar{H}\bar{\theta})^T\right]$$

$$\frac{\partial J(\bar{\theta})}{\partial \bar{\theta}} = -2\bar{H}^T(\bar{W} - \bar{H}\bar{\theta})$$

$\vdots$

Exercise 2 :

observations are made on the water level of Ganga ~~reiver~~ river at assi ghat of Varanasi on every ~~mond~~ Monday morning during the rainy season. If $x_1, x_2, \dots, x_n$ denotes the water levels, then what should be the value of a if:

(i): $\sum_{i=1}^{n} (x_i - a)^2$ is a minimum.

(ii): $\sum_{i=1}^{n} |x_i - a|$ is a minimum.

What are the maximum possible values in (i) and (ii), where $|(\cdot)|$ denotes the magnitude of $(\cdot)$?

Exercise 3:

Fit the model $w = a_0 + a_1 t + a_2 t^2$ to the same data in exercise-2 by the method of least squares. Estimate the values of a_0, a_1 and a_2.

Exercise 4 :

Evaluate the maximum and minimum values of the function

$$f(x,y) = x - 2y + 7, \text{ for } 0 \leq x \leq 2, \; -1 \leq y \leq 3$$

Maxima/Minima Subject to constraints:

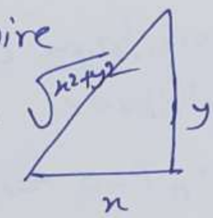
Some example

(1) For given sheet of card board make a box such that its volume is maximum.

i.e. lets side of cardboard mask are x, y and z

$V = xyz$; subject to $2(xy + yz + xz) = \text{fixed}$.

(ii) For given length of wire to fence the maximum area of ~~rectangular Shape triangular~~ right angle triangular area.



i.e. ~~mini~~ maximize the are $A = \frac{1}{2}xy$

for given ~~x^2~~ $x + y + \sqrt{x^2+y^2} = \text{fixed}$.

(iii) Which is the point on a given plane having the shortest distance from the origin. Let the given plane be

$$2x + 5y - z = 7$$

At any ~~arb~~ arbitrary point (x, y, z) on this plane; the distance from origin

$$\sqrt{(x-0)^2 + (y-0)^2 + (z-0)^2} = x^2 + y^2 + z^2$$

Subject to the condition $2x + 5y - z = 7$.

→ Lagrangian :–

— Convinient method is to solve the maxima / minima problem subject to one or more constraints is to use the method of ~~La~~ Lagrangian multipliers.

- Suppose function $f(x_1, x_2, \ldots, x_n)$ is maximized / minimized under the constraint condition $h(x_1, x_2, \ldots, x_k) = 0$, then take the function

$$g(x_1, x_2, \ldots, x_k) = f(x_1, x_2, \ldots, x_k) + \lambda h(x_1, x_2, \ldots, x_k) \quad \text{—— (1)}$$

where λ : An arbitrary constant, called Lagrangian multiplier.

$\because h(x_1, x_2, \ldots, x_k) = 0 \Rightarrow \lambda h(x_1, x_2, \ldots, x_k) = 0.$

$\therefore$ Adding $\lambda h(x_1, x_2, \ldots, x_k)$ does not affect the $f(x_1, x_2, \ldots, x_k)$

↳ Treat λ as an additional variable and then ~~take~~

$\dfrac{\partial g}{\partial \lambda} = 0$ produces the restriction

Consider the equation

$$\boxed{\frac{\partial g}{\partial x_1} = 0, \;\ldots,\; \frac{\partial g}{\partial x_k} = 0, \;\frac{\partial g}{\partial \lambda} = 0} \quad \text{—— (2)}$$

Solve the equation (2) to obtain the critical point

Note :- Lagrange multiplier is only helpfull to find critical points.
- one can not take second order derivative w.r.t λ also and also consider the matrix of all second order derivative

because if $(a_1, a_2, \ldots, a_n)$ is critical point then $(a_1, a_2, \ldots, a_n, \lambda)$ is not critical point

$\Rightarrow$ Thus further taking second order derivative is meaning less

$\Leftarrow$ With the help of critical point one has to check for maxima / minima independently of λ

Example:- Optimize

$f(x,y) = x^2+y^2+xy+4$ subject to $x+2y-3=0$.

Soln:- Since constraint is simple

Put $x = -2y+3$ in $f(x,y)$

for $f = (3-2y)^2+y^2+(3-2y)y+4$

This is a function of one variable. Hence, consider the equation:

$$\frac{\partial f}{\partial y} = 0 \Rightarrow -2(3-2y)+2y+3-4y=0$$

$$\Rightarrow \boxed{y = \frac{3}{2}}$$

Substitution in $x = 3-2y = 3-2\left(\frac{3}{2}\right) = 0$

$\therefore x = 0$

Hence the critical point $(x,y) = \left(0, \frac{3}{2}\right)$

Now taking second order derivative

$$\frac{\partial^2 f}{\partial y^2} = 2 > 0 \Rightarrow \text{Hence critical point is minima}$$

$$\therefore f_{min} = \frac{25}{4} = 6.25.$$

Method for Lagrangian Multipliers:

$$g(x,y) = f(x,y) - \lambda(x+2y-3) \quad [\text{lagrange multiplier } \rightarrow \lambda]$$

$$g(x,y) = x^2+y^2+xy-4-\lambda(x+2y-3)$$

$$\frac{\partial g}{\partial x} = 0 \Rightarrow 2x+y-\lambda = 0 \text{ ——— (1)}$$

$$\frac{\partial g}{\partial y} = 0 \Rightarrow x+2y-2\lambda = 0 \text{ ——— (2)}$$

$$\frac{\partial g}{\partial \lambda} = 0 \Rightarrow x+2y-3 = 0 \text{ ——— (3)}$$

From (1) and (2), equating λ we have

$$\cancel{4x+2y = x+2y} \Rightarrow$$

$$4x+2y = x+2y \Rightarrow x = 0$$

Putting in (3) we have $y = \frac{3}{2}$

$\therefore$ Critical point $(0, \frac{3}{2})$

[Notes:- (1) Same result was obtained through direct Substitution]

The value of 'f' at this point $(x,y) = (0, 3/2)$ is ~~24~~ $\frac{25}{4}$

~~(2)~~ $\cancel{\frac{\partial^2 g}{\partial x^2}}$.

(2) λ is used we can not ~~make~~ form matrix of second ~~order~~ partial derivative

When a non-relevant variable λ is present that matrix of second order partial derivatives will always be indefinite.

How to check the critical point is maximal/minima?

Compute the function value near critical point and see that this value is greater than value of ~~at critic.~~ function at critical point or less than..

⇒ Now consider the point $(0.1, \frac{3.1}{2})$ the value of function at this point

$$f(x,y) = x^2+y^2+xy+4$$
$$= (0.1)^2+\left(\frac{3.1}{2}\right)^2+(3.1)(0.1)+4$$
$$= 6.5675$$
$$> 6.25$$

Hence $f(0, \frac{3}{2}) < f(0.1, \frac{3.1}{2})$

Therefore $(x,y) = (0, \frac{3}{2})$ corresponds to local minimum point.

Note :- Here substitution is possible, because the constraint was a linear one ⇒ Thus we can solve for one variable from the constraint and substitute in the function. If the constraint was a complicated one this process would ~~have~~ not have been feasible.

⇓

But this method of Lagrangian multiplier enables us to proceed for finding critical points. ⇒ But ~~of~~ trade back is after evaluating the critical point we can not proceed with Lagrangian multiplier to check for definiteness of the matrix of second order partial derivatives.

Example:— Constrained Maxima:—

Consider the distance from the origin to an arbitrary point on a plane. Let $x-2y+z=3$ be the plane. What is the shortest distance to the plane.

Solution:— The distance from a point (x,y,z) to the origin $(0,0,0)$ is $\sqrt{x^2+y^2+z^2}$.

Minimizing $\sqrt{x^2+y^2+z^2}$ is equivalent to minimizing $x^2+y^2+z^2$ subject to condition $x-2y+z=3$. The general problem is to minimize $X'X$ subject to $X'a=b$ where $X'=(x_1,x_2,x_3,\ldots,x_n)$, $a'=(a_1,a_2,\ldots,a_n)$. In our example. $a'=$ coefficient vector $=(1,-2,1)$, $b=3$, $X'=(x,y,z)$. Let -2λ be a ~~lag~~ Lagrangian multiplier

Then let

$$g(X) = X'X - 2\lambda(X'a-b)$$

Then $\frac{\partial g}{\partial X} = \overset{\text{vector}}{0} \Rightarrow 2X - 2\lambda a = 0$

$$\Rightarrow \quad X = \lambda a, \quad \text{—(1)}$$

$$\frac{\partial g}{\partial \lambda} = 0 \Rightarrow -2(X'a-b)=0$$

$$\Rightarrow \quad X'a = b \quad \text{—(2)}$$

Now

$$x = \lambda a \Rightarrow x'a = \lambda a'a \quad \text{—— (3)}$$

Now using (2) in (3) we have

$$b = \lambda a'a$$

$$\Rightarrow \lambda = \frac{b}{a'a} \quad \text{—— (4)}$$

and $x'x = \lambda x'a = \lambda b$ [using (1) and (2)]

$$x'x = \frac{b}{a'a} b = \frac{b^2}{a'a} \text{ [using (4)]} \quad \text{—— (5)}$$

Thus the minimum is attained at $x = \lambda a = \left(\frac{b}{a'a}\right) a$ and the minimum value is $\frac{b^2}{a'a}$. Then

$$\min_{x'a=b} x'x \Rightarrow x = \left(\frac{b}{a'a}\right) a \text{ and } x'x = \frac{b^2}{a'a} \quad \text{—— (6)}$$

For example, $b = 3$, $a'a = [1, -2, 1]\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} = (1)^2 + (-2)^2 + (1)^2$

$a'a = 6$

$$\boxed{\therefore x'x = \frac{b^2}{a'a} = \frac{b^2}{a'a} = \frac{9}{6} = \frac{3}{2}}$$

$$x = \left(\frac{b}{a'a}\right) a = \frac{\cancel{3}^{1}}{\cancel{6}_{2}} \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$$

$$\boxed{\therefore x = \begin{bmatrix} 1/2 \\ -1 \\ 1/2 \end{bmatrix}}$$

Example :— Optimization of quadratic form:—

Optimize the quardatic form $X'AX$, $A=A'$ subject to the ~~quo~~ quadratic constraint $X'X=1$. Generally, this is equivalent to inscribing an ellipsoid in a ~~hyper~~ hypersphere when A is positive definite.

Solution:— Consider the Lagrangian multiplier λ.

Consider

$$g(X) = X'AX - \lambda(X'X - 1)$$

Then

$$\frac{\partial g}{\partial X} = \cancel{2}(A+A')X - 2\lambda X$$
$$= 2AX - 2\lambda X$$
$$\cancel{= 2X'AX - 2\lambda X}$$

$$\frac{\partial g}{\partial X} = 0 \Rightarrow \cancel{X'AX = \lambda X} \quad AX = \lambda X \quad \text{—(1)}$$

$$\frac{\partial g}{\partial \lambda} = -(X'X - 1), \quad \frac{\partial g}{\partial \lambda} = 0 \Rightarrow X'X = 1 \quad \text{—(2)}$$

$$AX = \lambda X \Rightarrow (A - \lambda I)X = 0$$

$\hookrightarrow$ ~~From~~ For this equation to have non-null solution for X, $X \neq 0$, the coefficient matrix must be singular which means its determinant must be zero. That is:

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

This means that λ is an eigenvalue of A.

From ① we have

$$X'AX = \lambda X'X = \lambda, \text{ since } X'X = 1$$

[using ①]

→ Thus the maximum value of $X'AX$ is the largest eigenvalue of A and the minimum value is the smallest eigen value of A. Hence the general result is the following:

~~max X'X~~

$$\max_{X'X=1} X'AX \Rightarrow \lambda_k = \text{largest eigenvalue of } A \quad \text{——③}$$

$$\min_{X'X=1} X'AX \Rightarrow \lambda_1 = \text{smallest eigenvalue of } A. \quad \text{——④}$$

Note :- Equation ④ is the basis for the topic called "Principal components Analysis". which is technique used for variable selection process.

For example :- For studying health condition of a person a doctor may make measurements on ~~pulse~~

① → Pulse
② → Blood pressure
③ → Blood Sugar
④ → cholestrol
... so on.

Suppose that he had taken ~~such~~ measurements on 1000 such items which included height, length of right arm, length of right arm, length of left foot, diameter of the skull and so on.

Definitely a lot of variables are ~~irre~~ irrelevant for ~~minoton~~ monitoring ~~health~~ health.

— He would like to select, say 5 most "important" variables

↳ one way of acheiving this is to use the method of " Principal Components Analysis".

Exercise:

Ex1: Check for maxima/minima for the following functions

(i) $f(x) = 2x^2 + 5xy + y^2 - 2x + 5y + 3$

subject to $2x + y = 5$

(ii) $f(x) = -3x^2 - 4y^2 - 2xy + 5x - 3y + 5$

subject to $x - 2y = 4$

(iii) $f(x) = 5x^2 - y^2 + 2xy - x + 2y - 7$

subject to $x + y = 7$

(iv) $f(x) = x_1^2 + x_2^2 + x_3^2 + 2x_1x_2$

subject $x_1 + x_2 + x_3 = 4$

(v) $f(x) = 4x_1^2 + x_2^2 + 5x_3^2 - 2x_1x_2$

subject to $x_1^2 + x_2^2 + x_3^2 = 2$; $x_1 \geq 0$, $x_2 \geq 0$, $x_3 \geq 0$

(vi) $f(x) = 2x_1^2 + 5x_2^2 + x_3^2 + 2x_1x_2 + x_1x_3$

subject $2x_1^2 + 4x_2^2 + x_3^2 = 1$

Ex 2:— Optimize the quadratic form $X'AX$, $A = A'$ subject to the following constraints:-

(i) $X'B = c$, X, B is $k\times 1$, c is 1×1; B and c are known

i.e. $X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix}$, $B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_k \end{bmatrix}$

$$X'B = [x_1, x_2, \dots, x_k] \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_k \end{bmatrix} = x_1 b_1 + x_2 b_2 + \dots + x_k b_k$$
$$= b_1 x_1 + b_2 x_2 + \dots + b_k x_k$$
$$= B'X$$

$$\therefore X'B = B'X = c$$

(ii) $X'X = 1$, X is $k\times 1 \Rightarrow [x_1, x_2, \dots, x_k] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix} = 1$

$$x_1^2 + x_2^2 + \dots + x_k^2 = 1$$

(iii) $X'BX = 1$, $\underbrace{B = B'}_{\text{Symmetric } k\times k}$, X is $k\times 1$

Ex 3:— Optimize the bilinear form $X'AY$, where X, A, and Y are $p\times 1$, $p\times q$ and $q\times 1$, $q \geq p$ and A is rank of p, subject to the conditions $X'BX = 1$, $Y'CY = 1$, where $B = B'$, $p\times p$, and positive definite, $C = C'$, $q\times q$, and positive definite. [This is the theoretical basis for a topic called Canonical Correlation Analysis.]

Ex. 4:— Write down the following quadratic form explicitly or open up the quadratic form $X'AX$, $A = A'$, where

(i) $A = \begin{bmatrix} 2 & 1 \\ 1 & -3 \end{bmatrix}$; (ii) $A = \begin{bmatrix} -1 & 1 & 0 \\ 1 & 2 & 3 \\ 0 & 3 & 4 \end{bmatrix}$

(iii) $A = \begin{bmatrix} 2 & -1 & 1 & 3 \\ -1 & 5 & -1 & 7 \\ 1 & -1 & 0 & -1 \\ 3 & 7 & -1 & 8 \end{bmatrix}$

Ex ~~5~~6: Write down the following bilinear form $X'AY$ where:

(i) $A = [1, -1, 2, 4]$; (ii) $A = \begin{bmatrix} 1 & -1 & 3 \\ 2 & 1 & -1 \end{bmatrix}$

(iii) $A = \begin{bmatrix} 2 & 0 & 1 & -1 \\ -1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 0 \end{bmatrix}$

Some Mathematical Preliminaries :-

Revisit of Linear form, Linear expression, Quadratic form, Quadratic expression :

$-x_1 \equiv$ a linear form (LF)

$x_2 + 5 \equiv$ a linear expression (LE)

$2x_1 - x_2 + x_3 \equiv$ a linear form

$5x_1 + x_2 \equiv$ a linear form

$2x_1 - x_5 + x_{10} - 6 \equiv$ a linear expression

maximum degree of any term is 1 (LE)

Degree of each term is one (LF)

Quadratic form and Quadratic Expression :-

↳ Each term has degree two

↳ maximum degree is two.

$x_1^2 + x_2^2 + \cdots + x_p^2 \equiv$ a quadratic form

$x_1^2 + x_1x_2 + x_2^2 \equiv$ a quadratic form

$x_5^2 \equiv$ a quadratic form

$2x_1^2 + x_1x_2 - 5x_2 + 7 \equiv$ a quadratic expression

$x_1^2 + x_1x_2 + x_2^2 - 2x_1 - 5x_2 \equiv$ a quadratic expression

$x_4^5 + 7 \equiv$ a quadratic ~~ex~~ expression

→ A general linear form can be written as:

$$u = a'X = X'a = a_1x_1 + a_2x_2 + \cdots + a_px_p$$

$$a = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{bmatrix}; \quad X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix}$$

— $u_1 = 2x_1 - x_2 + x_3$; $X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_3 \end{bmatrix}$; $a = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}$

Quadratic form:-

A general quadratic form will of the following form where a_{ij}'s are constants and x_j; $j=1,3\cdots,p$ are distinct real scalar variables:

$$
\begin{aligned}
& a_{11}x_1^2 + a_{12}x_1x_2 + \cdots + a_{1p}x_1x_p \\
&+ a_{21}x_2x_1 + a_{22}x_2^2 + \cdots + a_{2p}x_2x_p \\
&+ a_{31}x_3x_1 + a_{32}x_3x_2 + \cdots + a_{3p}x_3x_p \\
&+ \cdots \\
&+ a_{p1}x_px_1 + a_{p2}x_px_2 + \cdots + a_{pp}x_p^2 \\
&= X'AX; \quad X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix}; \quad A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \cdots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pp} \end{bmatrix} \quad \text{—— (1)}
\end{aligned}
$$

$A = [a_{ij}]_{p\times p}$

For example:

$$
\begin{aligned}
Q_1 &\equiv 2x_1^2 + 3x_2^2 + 6x_1x_2 \\
&\equiv [x_1 \;\; x_2] \begin{bmatrix} 2 & 6 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad \text{—— (2)} \\
&\equiv \underset{X'}{[x_1 \;\; x_2]} \underset{A}{\begin{bmatrix} 2 & 3 \\ 3 & 3 \end{bmatrix}} \underset{X}{\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}} \\
&\equiv X'AX; \quad \Rightarrow \underset{\text{symmetric}}{A = A'} \quad (3)
\end{aligned}
$$

- Note from (1) the detailed form it may be noted that since $x_ix_j = x_jx_i$ the coefficient of x_ix_j is $a_{ij}+a_{ji}$

e.g. coefficient of x_1x_2 is $a_{12}+a_{21}$

$\Downarrow$

And hence one can distribute equally as $\frac{1}{2}(a_{ij}+a_{ji})$ in the (i,j)th and (j,i)th position.

$\Rightarrow$ Without loss of generallity the matrix A in the quadratic form $X'AX$ can be assumed to be symmetric

Definiteness of Quadratic forms :-

Consider Quadratic form

$u = X'AX$, where X is $p \times 1$ vector of scalar variable. and A is $p \times p$ matrix of constant (free from X).

— What will be the value of quadratic form $u = X'AX$ for all possible non-null vector X.

- u ~~may~~ remains always positive
- u remains always negative
- May positive for some X and negative some X.

Definition 1.1 :- Definiteness :- Consider the quadratic form $X'AX$ in real scalar elements in X and real elements in A, where the constant matrix A is assumed to be symmetric without any loss of ~~generality~~. generality.

For all non-null X !

(i) $X'AX > 0$ means that $X'AX$ and the matrix A are positive definiteness

(ii) ~~$X'AX < 0$~~ $X'AX \geq 0$ means that $X'AX$ and A are positive semi-definite

(iii) $X'AX < 0$ means that $X'AX$ and A are negative definite

(iv) $X'AX \leq 0$ means that $X'AX$ and A are negative semi-definite.

(v) If $X'AX$ is positive for some non-null X and $X'AX$ is negative for some non-null X then $X'AX$ and symmetric matrix A is indefinite

Note:- Some essential property for positive definiteness of real symmetric matrix:-

- consider quadratic form $X'AX$ with $X' = [1.0\ 0 \cdots 0]$

Then $X'AX = a_{11}$ and hence $a_{11} > 0$ ~~if A is~~ if the symmetric matrix A is positive definite.

- By similar argument

 $a_{ii} > 0$, $i = 1, 2, 3, \ldots, p$ for $p \times p$ symmetric matrix A to be positive definite.

- From properties of matrices we know that a symmetric matrix A can be written as:

$$A = PDP', \quad P'P = I, \quad PP' = I, \quad D = \text{diag}(\lambda_1, \lambda_2 \cdots \lambda_p) \quad \text{—(4)}$$

where - P is orthonormal matrix

 - D is diagonal matrix with eigen values of A as the diagonal elements.

clearly $X'P = [0, \cdots, 1, 0, \cdots, 0]$ with the ith element 1. and all other elements zeros, then on the right of (4) we have that $\lambda_i > 0$ if A is positive definite

- Thus for a positive definite matrix all eigenvalues are positive $\Rightarrow$ Determinant = product of eigenvalues must be positive also

- In symmetric matrix A all leading submatrices are also symmetric and hence all the leading minors must also positive if A is positive definite.

Hence following property:-

(1) Property (1):- Positive Definiteness:-

If a $p \times p$ real symmetric matrix A is positive definite then (i) all the eigenvalues of A are positive (ii) all leading minors of A are positive, and conversely, If a $p \times p$ matrix A is real symmetric and if (a) all eigenvalues of A are positive then A is positive definite, (b) if all leading minors are positive then A is positive definite.

Note:- Symmetry is basic requirement of a matrix A to be positive definite

e.g. Consider $A = \begin{bmatrix} 1 & -5 \\ 0 & 2 \end{bmatrix}$, Here eigenvalue $\lambda_1 = 1, 2 > 0$, But $X' = [1, 1]$, $X'AX = 1-5+2 = -2 < 0$

Hence A is not symmetric then eigenvalues being positive or leading minors being positive need not imply that the matrix is positive definite.

Ex:- Check for the definiteness of the following matrices:

$$A_1 = \begin{bmatrix} 2 & 1 & 0 \\ -1 & 5 & 1 \\ 0 & 1 & 4 \end{bmatrix}; \quad A_2 = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 5 & 1 \\ 0 & 1 & 4 \end{bmatrix}$$

Soln:- A_1 is not symmetric $\Rightarrow$ no need to check definiteness of A_1

$A_2 = A_2'$ $\therefore$ A_2 is symmetric real.

Hence we check leading diagonal:

$$2 > 0, \quad \begin{vmatrix} 2 & 1 \\ 1 & 5 \end{vmatrix} = 10-1 = 9 > 0, \quad \begin{vmatrix} 2 & 1 & 0 \\ 1 & 5 & 1 \\ 0 & 1 & 4 \end{vmatrix} = 34 > 0$$

And hence A_2 is positive ~~definiteness~~ definite and we may write

$$A_2 > 0$$

Note:- Simplest way to check for definiteness of a matrix is to check the leading minors, since minor computation is easy whereas the eigenvalues will be difficult to compute if the order of the real symmetric matrix is ≥ 3.

Property (2) :— Negative definiteness and Indefiniteness

Let A be real symmetric matrix and thereby the quadratic form $X'AX$, is (a) negative definite then (i) all the eigenvalues of A are negative (ii) the leading minors of A are ~~alter~~ alternately negative and positive (all the odd order minors are ~~native negative~~) negative and all even order are positive)

(b) Indefinite then (i) the eigenvalues of A are all positive and negative and at least one positive and at least one negative (ii) All leading minors are either positive or negative and at least one negative and one positive.

Conversally

If A is $p \times p$ real symmetric and if (i) or (ii) of (a) holds then A is negative definite and if (i) and (ii) of (b) holds then A is indefinite.

When A, thereby the quadratic form $X'AX$, is negative ~~de~~ definite we may write it as $A < 0$

Note:– Take ~~$X' = e$~~

$$X' = [1, 0, 0, \cdots, 0], [0, 1, 0, \cdots, 0], \cdots, [0, 0, \cdots, 1]$$

it is clear that

$X'AX$ or A is negative definite then all diagonal elements of A are negative, i.e., $A = [a_{ij}]$ then $a_{ii} < 0$, $i = 1, \cdots, p$. Also, from the representation

$$A = P'DP, \quad PP' = I, \quad P'P = I, \quad D = \text{diag}(\lambda_1, \lambda_2, \cdots, \lambda_p)$$

where $\lambda_1, \lambda_2, \cdots, \lambda_p$ are the eigenvalues of A we have $\lambda_j < 0$, $j = 1, \cdots, p$ if $X'AX$ or A is negative definite.

~~Then~~

$\Rightarrow$ Then the product of the eigenvalues will be negative for an odd number of eigenvalues and positive for an even number of eigenvalues.

Note :- Pepole usually ~~not~~ misinterpret this property by simply checking the leading minors or eigenvalues without checking for symmetry of the matrix to start with.

e.g. $A = \begin{bmatrix} -1 & 6 \\ 0 & -2 \end{bmatrix}$ then the eigenvalues are -1 and -2, both negative, and one should not conclude that the matrix is negative definite.

Take for example, $X' = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ then $X'AX = 3 > 0$ and hence A is not negative definite. Hence without the matrix being symmetric, when real, we ~~cannot~~ cannot talk about definiteness of matrices.

Property (3): Semi-definiteness:-

If the quadratic form $X'AX$, thereby the $p \times p$ symmetric matrix A, is (a) positive semi-definite then (i) all the eigenvalues of A are positive or zero and atleast one zero; (ii) all the leading minors are positive or zero with atleast one ~~zero~~ of them zero; (b) negative semi-definite then (i) all the eigenvalues of A are negative or zero with atleast one of them zero, (ii) all the leading minors are alternatively negative and positive with atleast one of them zero.

– The converse also hold if the $p \times p$ real matrix A is symmetric and if the properties (i) or (ii) hold in the cases (a), (b) above.

Ex:- check the definiteness of the following matrices:

$$A_1 = \begin{bmatrix} 2 & 0 \\ 0 & -5 \end{bmatrix}; \quad A_2 = \begin{bmatrix} 3 & 2 & 4 \\ 2 & 2 & 1 \\ 4 & 1 & 5 \end{bmatrix}$$

Solution:- A_1 is symmetric and has eigenvalues $\lambda_1 = 2$, $\lambda_2 = -5$, one positive and one negative and hence the matrix indefinite.

A_2 is real symmetric: consider leading minors of A_2:

$$3 > 0, \quad \begin{vmatrix} 3 & 2 \\ 2 & 2 \end{vmatrix} = 2 > 0, \quad \begin{vmatrix} 3 & 2 & 4 \\ 2 & 2 & 1 \\ 4 & 1 & 5 \end{vmatrix} = -9 < 0$$

Hence A_2 is indefinite.

Note :- — Diagonal elements of real symmetric matrix A, must be all positive if $A > 0$, All positive or zero if $A \geq 0$, all negative if $A < 0$, all negative or zero if $A \leq 0$

— If diagonal elements of a real symmetric matrix are positive and negative then one can conclude that the matrix is indefinite.

Hermitian Definitness :

Elements of the matrix is complex, then the definiteness can be defined in terms of a hermitian matrix.

Definitions — Hermitian Matrix and Skew

Hermitian Matrix :- Consider the matrix A such that $A = A^*$; where $A^* \equiv$ conjugate transpose of A., then A is called Hermitian ~~transpose~~. matrix

If $A^* = -A$ then A is called skew hermitian.

The conjugate will be denoted by a bar, conjugate of A is $\bar{A}$ and hence

$$\boxed{A^* = \bar{A}'}$$

For real case $\bar{A} = A$ and hence $A^* = A'$.

Note :- ① When the matrix hermitian, then its diagonal element must be real.

② When matrix is skew hermitian then its diagonal element must be ~~complex~~ purely imaginary or zero

$$A_1 = \begin{bmatrix} 1+i & -1 & 3+2i \\ 5 & 2+3i & -3i \\ \sqrt{2}-4i & -6 & 1+i \end{bmatrix}$$

$$\bar{A}_1 = \begin{bmatrix} 1-i & -1 & 3-2i \\ 5 & 2-3i & 3i \\ \sqrt{2}+4i & -6 & 1-i \end{bmatrix}.$$

$$A_1^* \; \cancel{\bar{A}_1^*} = \cancel{\begin{bmatrix} 1+i & -1 & 3+2i \\ 5 & 2+3i & -3i \\ \sqrt{2}-4i & -6 & 1+i \end{bmatrix}} \quad \cancel{A_1^*} = \bar{A}$$

$$A_1^* = \bar{A}' = \begin{bmatrix} 1-i & 5 & \sqrt{2}+4i \\ -1 & 2-3i & -6 \\ 3-2i & 3i & 1-i \end{bmatrix}$$

$$A_2 = \begin{bmatrix} 2 & 1+i & 2-i \\ 1-i & 5 & 3+2i \\ 2+i & 3-2i & 7 \end{bmatrix} \Rightarrow A_2^* = A_2$$
(Hermitian)

$$A_3 = \begin{bmatrix} 2i & 1 & 1-i \\ -1 & 0 & 2 \\ -1-i & -2 & i \end{bmatrix} \Rightarrow A_2^* = -A$$
(Skew Hermitian)

Note: ① Eigen values of ~~a~~ hermitian matrix are all real.

② ~~When in ($X'AX$), the x and A are in complex~~
~~In complex domain then the quadratic form is called~~
~~Hermitian form and is defined as follows:~~

② When the element of vector and matrix are in the complex domain then the quadratic form in the complex domain is called a hermitian ~~matrix~~ form and it is defined as follows:

Definition ~~(1)~~ Hermitian form:— Let X be $p\times 1$ and A a $p\times p$ matrix of a constants, where in X and A the elements could be in the complex domain. Consider the function X^*AX, where without loss of generality we can assume A to be hermitian. Then

$u = X^*AX$, where $A = A^*$, is called a hermitian form, corresponding to quadratic form in the real case.

Definition (2):— Definiteness in Hermitian Form:—

For all non-null X if a hermitian form $X^*AX > 0$ then the hermitian form and the matrix A are called ~~positiv~~ positive definite, if $X^*AX \geq 0$ (positive semi-definite), $X^*AX < 0$ (negative definite), $X^*AX \leq 0$ (negative semi-definite) and if it is none of the above then it is called indefinite.

Property (4):— The eigenvalues of a hermitian matrix, as well as of real symmetric matrix, are all real and the ~~eigen values~~ eigenvalues of a skew hermitian matrix, as well as a real skew symmetric matrix, are purely imaginary or zero.

Proof (left for Exercise):— From definition of eigen values for a square matrix A:

$$AX = \lambda X, \quad X \neq 0 \qquad \text{--- (1)}$$

where λ is a scalar, X is $n\times 1$ vector ~~and~~ and A is $p\times q$ ~~$p\times q$~~ matrix.

Premultiply (1) by X^*, the conjugate transpose of X (in the real case by the transpose X'), then

$$X^* A X = \lambda X^* X, \quad \text{—— (2)}$$

Take conjugate transpose of (1) we have

$$X^* A^* = \bar{\lambda} X^* \quad \text{—— (3)}$$

~~Postmultiply (3) by~~

Postmultiply (3) by X, and $A^* = A$ for hermitian A, Hence

$$X^* A X = \bar{\lambda} X^* X \quad \text{—— (4)}$$

Now subtracting (2) from (4) we have:

$$[\bar{\lambda} - \lambda] X^* X = 0$$

$$\Rightarrow \lambda = \bar{\lambda} \qquad [\text{ since } X^* X > 0]$$
$$X \neq 0$$

This shows that λ must be real

Skew Hermitian Case:— If A is skew hermitian then

$$A^* = -A$$

Then equation (4) becomes.

$$-X^* A X = \bar{\lambda} X^* X \quad \text{—— (5)}$$

Adding (2) and (5) we have

$$[\lambda + \bar{\lambda}] X^* X = 0$$

$$\Rightarrow \lambda + \bar{\lambda} = 0$$

$$\Rightarrow \cancel{\lambda = -\bar{\lambda}} \ \lambda \text{ is purely imaginary or } \lambda = 0$$

Note:- If A is a hermitian matrix and if all its eigen-values are positive then it is a positive definite matrix and the corresponding hermitian form will be positive definite.

:Exercise:

(1):- Write the following linear forms in matrix notation or write in the form $a'X = Xa'$ where a and X are $p\times1$ and a is a constant ~~vect~~ vector. (i): $2x_1 + x_2 - x_3 + x_4$
(ii): $x_1 + x_2 + \cdots + x_p$ (iii) $x_1 - x_4 + x_5$

(2):- Write following quadratic forms in the matrix notation as $X'AX$ where A is symmetric: (i) $x_1^2 + 3x_1x_2 - x_3^2 + 2x_2x_3$; (ii) $x_1^2 + x_2^2 + x_3^2 + \cdots + x_p^2$
(iii) $x_1^2 + 2x_2^2 + 3x_3^2 + \cdots + px_p^2$; (iv) $2x_2^2 + 5x_2^3 + 3x_1x_2 + x_3^3 - x_2x_3$; (v) $2x_1^2 - x_2^2 + 5x_3^2 - 3x_1x_2$

(3) Check the definiteness of the quadratic forms in exercise (2).

(4) show that the set of quadratic forms is closed under scalar multiplication and addition.

(5) Taking $X'AX$ for $A = A'$ and for all non-null X remaining >0, ≥ 0, <0, ≤ 0 and none of the above as the definition for definiteness of the symmetric matrix A then show that

(i) rules regarding all leading minors follow from this definition;

(2) the rules regarding all eigenvalues follow from this definition.

(6) If the definition of definiteness is that the real matrix A is symmetric, $A = A'$, and that all eigen values are (i) positive (positive definite),
(ii) positive and zero (positive semi-definite)
(iii) negative (negative definite)
(iv) negative ~~defini~~ and zero (negative semi-definite)
(v) none of the above (indefinite), ~~then show that~~

Then show that
(i) The properties regarding the quadratic form $X'AX$ for all non-null X,
(ii) Properties regarding all leading minors will follow from this definition.

(7) If the definition of definiteness of a real symmetric matrix A is given in terms of the leading minors then show that
(i) Properties regarding the eigen values,
(ii) Properties regarding quadratic forms $X'AX$ for all non-null X, follow from this definition

(8) Write down three examples of a 3×3, (i) Hermitian matrix A, (ii) a skew hermitian matrix A, (iii) Hermitian forms corresponding to your example in (i)

(9) Check for the definiteness of the following matrices:

$$A_1 = \begin{bmatrix} 1 & 1+i \\ 1-i & 5 \end{bmatrix}; \quad A_2 = \begin{bmatrix} 1 & 1+i \\ 1-i & -4 \end{bmatrix}$$

(10) Check for the definiteness of the matrix

$$A = \begin{bmatrix} 1 & 1+i \\ 1-i & 2 \end{bmatrix}$$

(11) Are the following statements true or false, give reasons when true ~~and~~ and give counter example when false;

(i): For all $X \neq 0$ if $X'AX > 0$ for a given real A then all eigenvalues of A are positive;

(ii) $AX = \lambda X \Rightarrow X'AX = \lambda X'X$. If $X'AX > 0$ for all $X \neq 0$, then $\lambda > 0$;

(iii) $X'AX = X'\left(\frac{A+A'}{2}\right)X = X'BX$, $B = B'$. If $X'AX > 0$ for all $X \neq 0$ then the eigenvalues of B are positive and since the eigenvalues of A and A' are the same the ~~eigen~~ eigenvalue of A are positive;

(iv) If $A = \begin{bmatrix} 2 & -3 \\ 3 & 5 \end{bmatrix}$ then $X'AX > 0$ for all $X \neq 0$ which means that the eigenvalues of A are positive

(v) If ~~eigen values~~ eigenvalues of A are positive then A is positive definite in the sense $X'AX > 0$ for all $X \neq 0$

(vi): If all leading minors of A are positive then A is positive definite in the sense $X'AX > 0$ for all $X \neq 0$

(vii): If A is $p \times p$ positive definite, that is, $A = A' > 0$ and if B is $p \times p$ positive definite, $B > 0$, then $cA + dB$ is positive definite ~~in the sense $X'AX > 0$ for all $X \neq 0$~~ where c and d are real scalar quantities.

~~(viii)~~

(viii): If A is positive definite then $I+A$ and $I-A$ are positive definite.

(ix) For all real A if $X'AX = 0$ for all $X \neq 0$ then $A = 0$.

(x) For all A, where the elements could be complex also, if $X^*AX = 0$ for all $X \neq 0$ then $A = 0$

(12) If A is $p \times p$ matrix, can be written as $A = BB'$ where B is a real $p \times q$ matrix, $q \geq p$ then show that:

(i): $A = A'$; (ii): $A \geq 0$ (positive semi-definite);

(iii): $A > 0$ if the rank of B is p, and conversely, for any $p \times q$, $q \geq p$ real matrix B, ~~$BB' = A$~~ $BB' = A$ is either positive semi-definite or positive definite (non-negative definite)

Vector and Matrix Differential Operators:

→ Differential operators in the form of matrices and vectors ⇒ This makes many computation are simplified. Consider following

- Derivative of scalar function w.r.t. a vector or matrix.
- Vector function w.r.t. a vector.

Definition :- (Vector differential operator):

Consider a $p \times 1$ vector of real scalar independent variables $x_1, x_2, \dots, x_p$ and the vector of partial derivatives and a real valued scalar function f of $x_1, x_2, x_3, \dots, x_p$. We define and denote them as follows:

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix}; \quad \frac{\partial}{\partial X} = \begin{bmatrix} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \\ \vdots \\ \frac{\partial}{\partial x_p} \end{bmatrix}, \quad \frac{\partial f}{\partial X} = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_p} \end{bmatrix}$$

Ex :- Evaluate the resultant of the operator $\frac{\partial}{\partial X}$ operating on (1): $u_1 = 2x_1 - x_2 + 5x_3 - 8$; (2) $u_2 = x_1^2 + x_2^2 + \dots + x_p^2$; (3) $u_3 = x_1^2 - 2x_1x_2 + x_2^2$

Solution :- (1) $u_1 = 2x_1 - x_2 + 5x_3 - 8$ is linear form.

$$\frac{\partial u_1}{\partial x_1} = 2, \quad \frac{\partial u_1}{\partial x_2} = -1, \quad \frac{\partial u_1}{\partial x_3} = 5$$

$$\therefore \frac{\partial u_1}{\partial X} = \begin{bmatrix} \frac{\partial u_1}{\partial x_1} \\ \frac{\partial u_2}{\partial x_2} \\ \frac{\partial u_3}{\partial x_3} \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ 5 \end{bmatrix}$$

(2) $$\frac{\partial u_2}{\partial X} = \begin{bmatrix} 2x_1 \\ 2x_2 \\ \vdots \\ 2x_p \end{bmatrix} = 2\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix}$$

$$\frac{\partial u_2}{\partial X} = 2X$$

(3) $$\frac{\partial u_3}{\partial X} = \begin{bmatrix} 2x_1 - 2x_2 \\ -2x_1 + 2x_2 \end{bmatrix}$$

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

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

$$\boxed{\therefore \frac{\partial u_3}{\partial X} = 2AX}$$

~~Appl~~: Find the derivative for general linear form and quadratic form :-

Result ① Consider the linear form ~~the a'x~~ $u = a'X = X'a$, where $a' = [a_1, a_2, \ldots, a_p]$ and $X' = [x_1, x_2, \ldots, x_p]$ with $a_1, \ldots, a_p$ being constants and $x_1, \ldots, x_p$ being functionally independent real scalar variables. Then:

$$\frac{\partial u}{\partial X} = a = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{bmatrix}$$

Proof is trivial do yourself

Result (2) Let X be $p\times 1$ vector of independent real scalar variables and let $u_1 = X'X = x_1^2 + x_2^2 + \cdots + x_p^2$ and $u_2 = X'AX$. Then

$$\frac{\partial u_1}{\partial X} = \frac{\partial X'X}{\partial X} = 2X$$

$$\frac{\partial u_2}{\partial X} = \frac{\partial (X'AX)}{\partial X} = (A + A')X \text{ for general } A$$

$$= 2AX \text{ ; if } A = A'$$

i.e. A is symmetric

Proof:- (Do yourself)

Note :- Transpose of vector differential operator be denoted as:

$$\left[\frac{\partial}{\partial X}\right]' = \frac{\partial}{\partial X'} = \left[\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_p}\right]$$

Then $\frac{\partial}{\partial X \partial X'}$ is the matrix of second order partial derivatives given by:

$$\frac{\partial}{\partial X} \cdot \frac{\partial y}{\partial X'} = \frac{\partial}{\partial X}\left[\frac{\partial y}{\partial x_1}, \frac{\partial y}{\partial x_2}, \ldots, \frac{\partial y}{\partial x_p}\right]$$

$$= \begin{bmatrix} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \\ \vdots \\ \frac{\partial}{\partial x_p} \end{bmatrix} \left[\frac{\partial y}{\partial x_1}, \frac{\partial y}{\partial x_2}, \ldots, \frac{\partial y}{\partial x_p}\right]$$

$$= \begin{bmatrix} \frac{\partial^2 y}{\partial x_1^2} & \frac{\partial^2 y}{\partial x_1 \partial x_2} & \frac{\partial^2 y}{\partial x_1 \partial x_3} & \cdots & \frac{\partial^2 y}{\partial x_1 \partial x_p} \\ \frac{\partial^2 y}{\partial x_2 \partial x_1} & \frac{\partial^2 y}{\partial x_2^2} & \frac{\partial^2 y}{\partial x_2 \partial x_3} & \cdots & \frac{\partial^2 y}{\partial x_2 \partial x_p} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \frac{\partial^2 y}{\partial x_p \partial x_1} & \frac{\partial^2 y}{\partial x_p \partial x_2} & \frac{\partial^2 y}{\partial x_p \partial x_3} & \cdots & \frac{\partial^2 y}{\partial^2 x_p} \end{bmatrix}$$

$$\frac{\partial}{\partial X}\frac{\partial}{\partial X'} = \begin{bmatrix} \frac{\partial^2}{\partial x_1^2} & \frac{\partial^2}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2}{\partial x_1 \partial x_p} \\ \frac{\partial^2}{\partial x_2 \partial x_1} & \frac{\partial^2}{\partial x_2^2} & \cdots & \frac{\partial^2}{\partial x_2 \partial x_p} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial^2}{\partial x_p \partial x_1} & \frac{\partial^2}{\partial x_p \partial x_1} & \cdots & \frac{\partial^2}{\partial x_p^2} \end{bmatrix}$$

What will be the effect of this matrix of second order differential operators operating on a quadratic form?

$$\frac{\partial}{\partial x}\frac{\partial}{\partial x'} X'AX = \frac{\partial}{\partial x}\left[(A+A')x\right]' \Rightarrow \frac{\partial}{\partial x}\left[\frac{\partial}{\partial x}X'AX\right]'$$

$$= \frac{\partial}{\partial x}\left[X'(A+A')'\right]$$

$$= \frac{\partial}{\partial x}\left[X'(A+A')\right] \rightarrow \text{Linear form}$$

$$\frac{\partial}{\partial x}\frac{\partial}{\partial x'}(X'AX) = A+A' \quad \text{for general } A$$

$$= 2A \quad , \text{ when } A=A' \text{ i.e. } A \text{ is symmetric.}$$

Example:- Check for maxima/minima of the function

$$f(x_1, x_2) = 2x_1^2 - 2x_1x_2 + 3x_2^2 - 5x_1 - 2x_2 + 8$$

Solution:- In vector-matrix form the $f(x_1, x_2)$ can be written as:

$$\cancel{f(x_1,x_2) =} \begin{bmatrix} 2 & -1 \\ -1 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$f(x_1, x_2) = [x_1, x_2]\begin{bmatrix} 2 & -1 \\ -1 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + [-5 \; -2]\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + 8$$

$$f(x) = X'AX + BX + 8 .$$

where $X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, $A = \begin{bmatrix} 2 & -1 \\ -1 & 3 \end{bmatrix}$, $B = \cancel{[-5, -2]} \triangleq \begin{bmatrix} -5 \\ -2 \end{bmatrix}$

$$\frac{\partial f}{\partial x} = 0 , \quad \frac{\partial f}{\partial x} = \frac{\partial}{\partial x}(X'AX) + \frac{\partial}{\partial x}(BX) + \frac{\partial}{\partial x}(8)$$

$$= \frac{\partial}{\partial x}(X'AX) + \frac{\partial}{\partial x}(BX) + 0$$

$$= (A+A')X + B$$

$$= 2AX + B \quad [\text{Here since } A'=A]$$

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

or, $\begin{bmatrix} 4 & -2 \\ -2 & 6 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} -5 \\ -2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

or, $\begin{bmatrix} 4 & -2 \\ -2 & 6 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = -\begin{bmatrix} -5 \\ -2 \end{bmatrix}$

or, $2AX = -B$

$\therefore \boxed{X = -\frac{1}{2}A^{-1}B} = \begin{bmatrix} \frac{3}{2} \\ \frac{9}{10} \end{bmatrix}$

$\therefore x_1 = \frac{3}{2}, \quad x_2 = \frac{9}{10}$

Now check second derivative

$$\frac{\partial}{\partial X}\frac{\partial}{\partial X'}f(X) = 2A = \begin{bmatrix} 4 & -2 \\ -2 & 6 \end{bmatrix}$$

Check leading minor.

$$4>0, \quad \begin{vmatrix} 4 & -2 \\ -2 & 6 \end{vmatrix} = 24-4 = 20>0$$

$\therefore$ Therefor matrix corresponding $\frac{\partial}{\partial X}\frac{\partial}{\partial X'}f(X)$ is positive definite $\Rightarrow$ critical point correspond to minima.

Ex:- Let $M(T)$ denote the moment generating function of vector random variable $X' = (x_1, x_2, \dots, x_p)$ then $\frac{\partial M(T)}{\partial T}\Big|_{T=0} = E(X) = \mu$ mean value of vector and $\frac{\partial}{\partial T}\frac{\partial}{\partial T'}M(T)\Big|_{T=0} = \Sigma + \mu\mu'$ where Σ is the covarience matrix. If the $p \times 1$ vector X has a multivariate normal density, $X \sim N_p(\mu, \Sigma)$ with the parameter μ and Σ then its moment generating function is:

$$M(T) = \exp\left\{\mu' T + \frac{1}{2}T'\Sigma T\right\}$$

Evaluate the mean value vector and covariance matrix of X.

Derivative with respect to a matrix

Definition:– Let $X = [x_{ij}]$ be a $p \times p$ matrix of distinct real scalar variables and let $f(X)$ be a real valued scalar function of X. Then the derivative of f with respect to the matrix is defined as follows:

$$\frac{\partial f}{\partial X} = \left[\frac{\partial f}{\partial x_{ij}}\right]$$

That is, the matrix is formed by differentiating f partially with respect to each and every element in X and placing at the respective position.

For example:, $u_1 = tr(X) = x_{11} + \cdots + x_{pp}$ is scalar function, $u_2 = |X|$ = determinant of X is scalar function, $u_3 = 2|X| + 6$ is scalar function.

Derivative w.r.t matrix for some well known scal function :–

Result:– Let $X = [x_{ij}]$ be a $p \times p$ matrix of functionally independent real scalar variables x_{ij}'s. Let $A = [a_{ij}]$ be $p \times p$ matrix of constants. Then:

$$f(X) = tr(X) \Rightarrow \frac{\partial f}{\partial X} = I \quad \text{—(i)}$$

$$f(X) = tr(AX) \Rightarrow \frac{\partial f}{\partial X} = A'; \text{ if } X \text{ is not symmetric} \quad \text{—(ii)}$$

$$= A + A' - diag(A); \text{ if } X \text{ is symmetric} \quad \text{—(iii)}$$

$$= 2A - diag(A); \text{ if } X = X' \text{ and } A = A' \quad \text{—(iv)}$$

where $diag(A)$: diagonal matrix formed with the diagonal elements of A.

Proof:- $tr(X)$ = sum of diagonal element of X

$$= x_{11} + x_{22} + \cdots + x_{pp}$$

$$\therefore f(X) = tr(X) = x_{11} + x_{22} + \cdots + x_{pp}$$

$\frac{\partial f}{\partial X}$ = the derivatives with respect to leading diagonal is 1 and other derivatives produces zero.

$$= \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & & & & \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} = I.$$

$$f(X) = tr(AX) = a_{11}x_{11} + a_{12}x_{21} + \cdots + a_{1p}x_{p1}$$
$$+ a_{21}x_{12} + a_{22}x_{22} + \cdots + a_{2p}x_{p2}$$
$$+ \cdots$$
$$+ a_{p1}x_{p1} + a_{p2}x_{p2} + \cdots + a_{pp}x_{pp}$$

Thus partial derivative w.r.t x_{ij} produces ~~a_{ij}~~ a_{ji} for all ij. Hence

$$\frac{\partial}{\partial X} f(X) = \frac{\partial}{\partial X} tr(AX) = \begin{bmatrix} a_{11} & a_{21} & a_{31} & \cdots & a_{p1} \\ a_{12} & a_{22} & a_{23} & \cdots & a_{p2} \\ \vdots & & & & \\ a_{1p} & a_{2p} & a_{3p} & \cdots & a_{pp} \end{bmatrix}$$

$$= A'$$

[In this case all element of X are distinct]

when $X = X' \Rightarrow x_{ij} = x_{ji} \quad \forall i,j$

$$\frac{\partial}{\partial x_{ij}} tr(AX) = (a_{ij} + a_{ji}) \quad \text{for } i \neq j$$
$$= a_{ii} \quad \text{for } i = j$$

Thus diagonal elements comes only once and the off-diagonal elements are the sum of the element comming from $A+A'$

Hence if we write $A+A'$ then once the diagonal element must be subtracted, then answer is
$A+A' - \text{diag}(A)$

When A is symmetric then: $A+A' = 2A$

and $A+A' - \text{diag}(A) = 2A - \text{diag}(A)$

$\therefore \frac{\partial}{\partial X} tr(AX) = \cancel{2A - d} (A+A') - \text{diag}(A)$, if X is not symmetric

$= (A+A') - \text{diag}(A)$; when X is symmetric

$= 2A - \text{diag}(A)$; when $A=A'$ and $X=X'$

Result :- Let $f(X)$ be a real-valued scalar function of the matrix A. Then

$$f(X) = tr(A'X) = tr(X'A), \quad X: p\times q, \; A: p\times q$$

$$\Rightarrow \frac{\partial f}{\partial X} = A$$

$f(X) = a'Xb$, $X: p\times p$, $a: p\times 1$, $b: p\times 1$

$$\Rightarrow \frac{\partial f}{\partial X} = ab'$$

where a and b are vectors of constant

Note:- For any two matrices X and A where product defined $tr(AX) = tr(XA) = tr(AX)' = tr(X'A')$.