Theorem:-

If $f(x)$ is differentiable function over the convex set $X \subseteq \mathbb{R}^n$ then $f(x)$ is convex over X iff

$$f(x^2) \geq f(x^1) + \nabla^T f(x^1)(x^2 - x^1)$$

for all $x^1, x^2 \in X$.

Proof:- Given - $f(x)$ is convex over X

For all $x^1, x^2 \in X$ and for all $\lambda \in (0,1]$

$$f(\lambda x^2 + (1-\lambda)x^1) \leq \lambda f(x^2) + (1-\lambda) f(x^1)$$

$$\Rightarrow \quad \frac{f[x^1 + \lambda(x^2 - x^1)] - f(x^1)}{\lambda} \leq f(x^2) - f(x^1)$$

Taking limit $\lambda \to 0$, it follows that

$$\left.\frac{df(x^1)}{d\lambda}\right|_{x^2 - x^1} \leq f(x^2) - f(x^1) \quad \text{—— (1)}$$

The directional derivative on LHS

$$\left.\frac{df(x^1)}{d\lambda}\right]_{x^2 - x^1} = \nabla^T f(x)(x^2 - x^1) \quad \text{—— (2)}$$

From (1) & (2) we have

$$\nabla^T f(x)(x^2 - x^1) \leq f(x^2) - f(x^1)$$

$$\Rightarrow \quad f(x^2) \geq f(x^1) + \nabla^T f(x)(x^2 - x^1)$$

$\Leftarrow$

<u>Given :-</u>

$$f(x^2) \geq f(x^1) + \nabla^T f(x^1)(x^2 - x^1)$$

for $x = \lambda x^2 + (1-\lambda)x^1 \in X$, $\lambda \in (0,1]$

$$f(x^2) \geq f(x) + \nabla^T f(x)(x^2 - x) \quad] \times \lambda$$

$$f(x^1) \geq f(x) + \nabla^T f(x)(x^1 - x) \quad] \times (1-\lambda)$$

$$\lambda f(x^2) + (1-\lambda) f(x^1) \geq f(x) + \nabla^T f(x) \underbrace{\left[\lambda(x^2 - x^1) + (1-\lambda)(x^1 - x)\right]}_{=0}$$

$$\left(\because \lambda(x^2 - x) + (1-\lambda)(x^1 - x) = 0 \right)$$

$\Rightarrow f(x) \leq \lambda f(x^2) + (1-\lambda) f(x^1)$, Hence f is convex. $\square$

Jacobian matrix

$$\left[\frac{\partial h}{\partial x}\right]_{ij} = \frac{\partial h_j}{\partial x_i}$$

i.e; $\left[\frac{\partial h}{\partial x}\right] = \left[\nabla h_1, \nabla h_2, \ldots, \nabla h_r\right]$

$$\nabla h_1 = \begin{pmatrix} \partial h_1/\partial x_1 \\ \partial h_1/\partial x_2 \\ \vdots \\ \partial h_1/\partial x_n \end{pmatrix}$$

$h = \left[h_1 \; h_2 \; \cdots \; h_r\right]$

or $(h_1, h_2, \ldots, h_r)$

Problem :- ——————————————————— (P)

Consider equality constraind problem

$$\min f(x)$$

such that $h_j(x) = 0, \quad j = 1, 2, \ldots, r < n$

In 1760 Lagrange tranformed this contrained problem to an unconstrained problem

By using Lagrange Multiplier λ_j

To formulate Lagrange funⁿ for $j = 1, 2, \ldots, r$

$$L(x, \lambda) = f(x) + \sum_{j=1}^{r} \lambda_j h_j(x) = f(x) + \lambda^T h(x)$$

Theorem

The equality constrained problem (P) for f & $h_j \in C^1$ and assume that the Jacobian matrix $\left[\frac{\partial h(x^*)}{\partial x}\right]$ is of rank 'r.'

The necessary cond^n for interior local minimum x^* of equality constrained problem (P) is.

x^* must coincide with the stationary point (x^*, λ^*) of Lagrange function L

OR: there exist a λ^* such that

$$\frac{\partial L}{\partial x_i}(x^*, \lambda^*) = 0 \quad , i = 1, 2, \ldots, n \quad \& \quad \frac{\partial L}{\partial \lambda_j}(x^*, \lambda^*) = 0$$

for $j = 1, 2, \ldots, r$

Proof: Given: (1) f & $h_j \in C^1$

it follows an interior local min at $x = x^*$

(2) $\left[\frac{\partial h(x^*)}{\partial x}\right]$ is of rank r.

To proove:- for $x = x^*$, $\exists\ \lambda^*$ such that

$$\frac{\partial L(x^*, \lambda^*)}{\partial x_i} = 0 , \quad i = 1, 2, \ldots, n ;$$

& $$\frac{\partial L}{\partial \lambda_j}(x^*, \lambda^*) = 0, \quad j = 1, 2, \ldots, r$$

As $x = x^*$ is local minimum at $x = x^*$

$$df = \nabla^T f(x^*)\, dx = 0 \quad \text{—— (1)}$$

$(\because df \geq 0$ for dx & for $-dx)$

$$dh_j = \nabla^T h_j(x^*)\, dx = 0, \quad j = 1, 2, \ldots, r. \quad (2)$$

consider the Lagrange fun^n

$$L(x, \lambda) = f(x) + \sum_{j=1}^{r} \lambda_j h_j(x)$$

The differential of L is given by

$$dL = df + \sum_{j=1}^{r} \lambda_j \, dh_j$$

$$\Rightarrow \quad dL = 0$$

$$\Rightarrow \quad \frac{\partial L}{\partial x_1} dx_1 + \frac{\partial L}{\partial x_2} dx_2 + \cdots + \frac{\partial L}{\partial x_n} dx_n = 0 \quad (3)$$

for all dx such that

$h(x) = 0$ i.e; $h(x^* + dx) = 0$

(as the constrained satisfied)

Choose Lagrange multipliers λ_j, $j = 1, 2, \ldots, r$ such that at x^*

$$\frac{\partial L}{\partial x_j}(x^*, \lambda) = \frac{\partial f}{\partial x_j}(x^*) + \left[\frac{\partial h}{\partial x_j}(x^*)\right]^T \lambda = 0 \quad (4)$$

$$j = 1, 2, \ldots, r.$$

The solution of this system provides the vector λ^*.

Here the 'r' variables, x_j, $j = 1, 2, \ldots, r$ may be any appropriate set of 'r' variables from the set x_i, $i = 1, 2, \ldots, n$.

A unique solution for λ^* exists as it is assumed that $\left[\frac{\partial h(x^*)}{\partial x}\right]$ is of rank 'r'.

Then the equation (3) reduces to

$$dL = \frac{\partial L}{\partial x_{r+1}}(x^*, \lambda^*)\, dx_{r+1} + \cdots + \frac{\partial L}{\partial x_n}(x^*, \lambda^*)\, dx_n = 0 \quad \text{(5)}$$

Again consider the constraints $h_j(X) = 0$, $j = 1, 2, \cdots, r$

If these equations are considered as a system of r-equations in the unknowns $x_1, x_2, \cdots, x_r$

these dependent unknowns can be solved for in terms of $x_{r+1}, x_{r+2}, \cdots, x_n$.

Hence the latter $(n-r)$ variables are the independent variables.

For any choice of these independent variables the other independent variables $x_1, x_2, \cdots, x_r$ are determined by solving $h(x) = [h_1(x), \cdots, h_r(x)]^T = [0, 0, \cdots 0]^T = \vec{0}$

In particular x_{r+1} to x_n may be varied one-by-one at x^* and it follows from $\boxed{5}$ that

$$\frac{\partial L}{\partial x_j}(x^*, \lambda^*) = 0 \quad , j = r+1, r+2, \ldots, n$$

and together with (4)

& constraints $h(x) = 0$.

it follows that the necessary conditions for an interior local minimum can be written as

$$\begin{cases} \dfrac{\partial L}{\partial x_i}(x^*, \lambda^*) = 0 & , \quad i = 1, 2, \ldots, n \\ \dfrac{\partial L}{\partial \lambda_j}(x^*, \lambda^*) = 0 & , \quad j = 1, 2, \ldots, r \end{cases}$$

$$\nabla_x L(x^*, \lambda^*) = 0 \ \& \ \nabla_\lambda L(x^*, \lambda^*) = 0$$