Consider the optimization problem

$$\min f(x) \qquad \text{— (1)}$$
$$\text{subject to } x \in \Omega$$

① The function $f : \mathbb{R}^n \to \mathbb{R}$ that we wish to minimize is called objective function or cost function.

② $x = (x_1, x_2, \ldots x_n)$ are called the desicion variable

③ The set $\Omega \subseteq \mathbb{R}^n$ is called the feasible set or constrained set.

The above problem is a decision problem that involves finding a vector $x^*$ that results the smallest value of the objective function.

Note Maximization problem can be represented equivalently in the form of a minimization problem because maximization $f(x)$ is equivalent to minimizing $-f(-x)$.

Note 2 The above problem (1) is known as constrained optimization problem. If $\Omega = \mathbb{R}^n$, then

The above problem (1) is called unconstrained optimization problem.

**Defn** Suppose that $f: \mathbb{R}^n \to \mathbb{R}$ is a real valued function. A point $x^* \in \mathbb{R}^n$ is said to be a local minima of $f$ if there exists an $\varepsilon > 0$ such that $\boxed{f(x) \geq f(x^*)}$ for all $x$ satisfying $\|x - x^*\| < \varepsilon$

- A point $x^* \in \mathbb{R}^n$ is a global minima of $f$ if $f(x) \geq f(x^*)$ $\forall\, x \in \mathbb{R}^n$.

**Remark** An optimization problem is solved only when a global minimizer is found. However, it is difficult to find a global minima. Thus we have to find local minima in practice.

**Notation** Let $f: \mathbb{R}^n \to \mathbb{R}$ be a differentiable function. Then the first order derivative of $f$ is denoted by $Df$ is

$$Df = \left( \frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots \frac{\partial f}{\partial x_n} \right)$$

The gradient of $f$ is defined by

$$\nabla f = (Df)^T = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix}$$

Let $f$ be twice differentiable, that is, $\nabla f$ is differentiable and we write derivative of $\nabla f$ as

$$D^2 f = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_2 \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_1} \\ \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_2} \\ \vdots & & & \\ \frac{\partial^2 f}{\partial x_1 \partial x_n} & \cdots & & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

The matrix $D^2 f$ is called the Hessian matrix of $f$ at $x$ and denoted by $F(x)$.

First order necessary condition for local minima (FONC)

Let $f : \mathbb{R}^n \to \mathbb{R}$ be differentiable at $x^*$ and $x^*$ is a local minima of $f$, then $\nabla f(x^*) = 0$

Note $\nabla f(x^*) = 0$ is not a sufficient condition for $x^*$ to become a local minima.

Example Let $f : \mathbb{R}^2 \to \mathbb{R}$ by $f(x_1, x_2) = x_1^2 - x_2^2$

$$\nabla f(x_1, x_2) = 0 \implies (2x_1, -2x_2) = (0, 0)$$

$$x_1 = x_2 = 0$$

Suppose $x^* = (0,0)$ is a local minima. But

$$f\left(\frac{\varepsilon}{2}, 0\right) = \frac{\varepsilon^2}{4} > f(0,0)$$

and $f\left(0, \frac{\varepsilon}{2}\right) = -\frac{\varepsilon^2}{4} < f(0,0)$

Now, Both the points $\left(\frac{\varepsilon}{2}, 0\right)$ and $\left(0, \frac{\varepsilon}{2}\right)$ having distance from $x^* = (0,0)$ is less than $\varepsilon$ for any $\varepsilon > 0$.

$\therefore$ $x^* = (0,0)$ is not a local minima

**Second order necessary condition (SONC)**

Let $f : \mathbb{R}^n \to \mathbb{R}$ be twice differentiable at $x^*$ and $x^*$ is a local minima of $f$.

Then $\nabla f(x^*) = 0$ and $F(x^*)$ is positive semidefinite

where $F$ is the Hessian matrix.

[Note : A $n \times n$ matrix $F$ is said to PSD if
$x^T F x \geq 0 \quad \forall x \in \mathbb{R}^n$ ]

Prob 1 Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ by

$$f(x) = x^T \begin{pmatrix} 1 & 2 \\ 4 & 7 \end{pmatrix} x + x^T \begin{pmatrix} 3 \\ 5 \end{pmatrix} + 6$$

where $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$

ⓐ Find the gradient and Hessian matrix of $f$ at the point $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

(b) Find a point that satisfy FONC for f. Check that the same point satisfy the SONC or not.

Ans (a) $\nabla f\begin{pmatrix}1\\1\end{pmatrix} = (11, 25)$, $H\begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}2 & 6\\6 & 14\end{pmatrix}$

(b) $f(x_1, x_2) = x_1^2 + 6x_1x_2 + 7x_2^2 + 3x_1 + 5x_2 + 6$

$\nabla f(x_1, x_2) = (2x_1 + 6x_2 + 3, 6x_1 + 14x_2 + 5)$

$\nabla f(x_1, x_2) = (0, 0)$

$\Rightarrow \quad x_1 = 1.5, \; x_2 = -1$

$\therefore$ $(1.5, -1)$ satisfies the FONC.

$H(x_1, x_2) = \begin{pmatrix}2 & 6\\6 & 14\end{pmatrix}$

Let $d = \begin{bmatrix}d_1\\d_2\end{bmatrix}$, Now $d^T H d = [d_1 \; d_2]\begin{bmatrix}2d_1 + 6d_2\\6d_1 + 14d_2\end{bmatrix}$

$= 2d_1^2 + 6d_1d_2 + 6d_1d_2 + 14d_2^2$

$= d_1^2 + d_1^2 + 2.d_1.3d_2 + 9d_2^2$
$+ 6d_1d_2 + 5d_2^2$

take $[d_1 \; d_2] = [2, -1]$, then

$$d^T H d = 8 - 24 + 14 = -2 < 0$$

$\therefore$ the Hessian matrix is not PSD.

This implies that SONC is not satisfying.

Hence $x^* = (1.5, -1)$ is not a local minima.

$\therefore$ The function has no optimal soln,

Prob 2 Let $f: \mathbb{R}^2 \to \mathbb{R}$

$$f(x_1, x_2) = (x_1 - x_2)^4 + x_1^2 - x_2^2 - 2x_1 + 2x_2 + 1$$

Find all points satisfying FONC.

$(1, 1)$ $\quad X_1 = X_2$