Optimization

Fourth Class

2020 - 2021

By

Dr. Jawad Mahmoud Jassim

Dept. of Math.

Education College for Pure Sciences

University of Basrah

Iraq

Chapter One

Basic Concepts

Lecture 3

10: Function and Differential

Definition (8): (Gradient Vector)

Let f: $R^n \rightarrow R$ have first partial derivatives in R^n. The gradient vector of f at X in R^n is defined as

$$\nabla f(X) = [\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \frac{\partial f}{\partial x_3}, \cdots, \frac{\partial f}{\partial x_n}]^T$$

We use notation $g(X)$ for the gradient vector of f at X.

Definition (9): (Hessian Matrix)

Let f: $R^n \rightarrow R$ $have\ second\ partial\ derivatives\ in\ R^n$. The Hessian matrix $G(X)$ of f at X is defined as $G(X) = \left[\frac{\partial^2 f}{\partial x_i\, \partial x_j}\right], i, j = 1, 2, 3, \cdots, n.$

Note (9):

If f has continuous second partial derivatives in R^n then $\frac{\partial^2 f}{\partial x_i\,\partial x_j} = \frac{\partial^2 f}{\partial x_j\,\partial x_i}$, $i, j = 1, 2, 3, \cdots, n$. So that $G(X)$ is symmetric and only $\frac{n(n+1)}{2}$ second derivatives need be calculated in order to evaluate $G(X)$.

Example (1):

Find the gradient vector and Hessian matrix for the function $f\colon R^2 \to R$ defined by $f(X) = 3x_1^2 + 4x_1x_2 - 4x_2^2$.

Solution:

The gradient vector is defined as: $\nabla f(X) = [\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}]^T = [6x_1 + 4x_2, 4x_1 - 8x_2]^T$.

The Hessian matrix is defined as: $G(X) = \left[\frac{\partial^2 f}{\partial x_i\,\partial x_j}\right], i, j = 1, 2.$

$$\frac{\partial^2 f}{\partial x_1^2} = 6, \frac{\partial^2 f}{\partial x_1\,\partial x_2} = 4, \frac{\partial^2 f}{\partial x_2\,\partial x_1} = 4, \frac{\partial^2 f}{\partial x_2^2} = -8.$$

$$G(X) = \begin{bmatrix} 6 & 4 \\ 4 & -8 \end{bmatrix}.$$

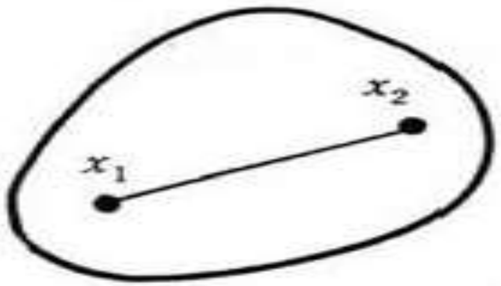
<u>11: Convex Sets and Convex Functions</u>

<u>Definition (10): (Convex Set)</u>

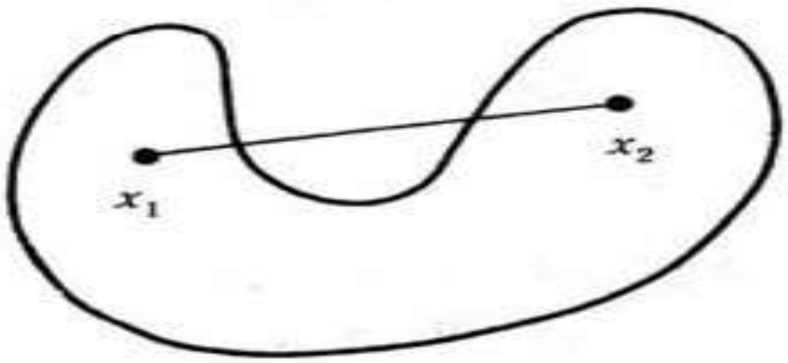
Let the set $D \subset R^n$. If for $any\ X, Y\ in\ D$, we have $\alpha X + (1-\alpha)Y\ in\ D\ for$ $\alpha \epsilon [0,1]$, $then\ D\ is\ said\ to\ be\ a$ $convex\ set$.

<u>Note (10):</u>

Definition (10) indicates, in geometry, that for any two points $X, Y\ in\ D$, the line segment joining $X\ and\ Y$ is entirely contained in D.



Convex Set



Nonconvex Set

Note (11):

$X=\alpha x_1+(1-\alpha)x_2$, where $\alpha \epsilon [0,1]$ is called convex combination of x_1 and x_2.

Note (12):

$X=\sum_{i=1}^{m} \alpha_i x_i$ *is called a convex combination* of $x_1, x_2, x_3, \cdots, x_m$, where $\sum_{i=1}^{m} \alpha_i = 1,\ \alpha_i \geq 0 , i = 1,2,3,\cdots,m.$

Lemma (1):

The set $D \subset R^n$ is convex if and only if for any $x_1, x_2, x_3, \cdots, x_n$ *in* D, $\sum_{i=1}^{m} \alpha_i x_i$ *in* D, *where* $\sum_{i=1}^{m} \alpha_i = 1,\ \alpha_i \geq 0 , i = 1,2,3,\cdots,m.$

Proof: (H.W.)

Theorem (1):

Let D_1 *and* D_2 be two convex *sets in* R^n. Then:

1: $D_1 \cap D_2$ is convex set.

2: $D_1 \mp D_2 = \{ x_1 \mp x_2 : x_1 \epsilon D_1 , x_2 \epsilon D_2 \}$ is convex set.

Proof: (H.W.)

Example (2):

Show that the set $H = \{ X \epsilon R^n : p^T X = c \}$ is convex, where $p \epsilon R^n$ is a nonzero vector and c is a scalar.

Solution:

Let $X_1, X_2 \epsilon H \; and \; \alpha \epsilon [0, 1]$.

$\therefore p^T X_1 = c \;\; and \; p^T X_2 = c$.

We want to prove that $\alpha X_1 + (1 - \alpha) X_2 \epsilon H$.

$\therefore p^T [\alpha X_1 + (1 - \alpha) X_2] = \alpha p^T X_1 + (1 - \alpha) p^T X_2 = \alpha c + (1 - \alpha) c = c$.

$\therefore \alpha X_1 + (1 - \alpha) X_2 \epsilon H$.

$\therefore$ H is convex set.

Example (3):

Show that the set $S = \{X \epsilon R^n : X = X_0 + \lambda d, \lambda \geq 0\}$ is a convex set, where $d \epsilon R^n\ is\ a\ nonzero\ vector\ and\ X_0$ is a fixed point.

Solution:

Let $X_1, X_2\ \epsilon\ S\ and\ \alpha\ \epsilon\ [0, 1]$.

$\therefore X_1 = X_0 + \lambda_1 d\ and\ X_2 = X_0 + \lambda_2 d\ , where\ \lambda_1, \lambda_2 \geq 0$.

We want to prove that $\alpha X_1 + (1 - \alpha) X_2\ \epsilon\ S$.

$$\begin{aligned} \therefore \alpha X_1 + (1 - \alpha) X_2 &= \alpha[X_0 + \lambda_1 d] + (1 - \alpha)[X_0 + \lambda_2 d\ ] \\ &= \alpha X_0 + \alpha \lambda_1 d + X_0 + \lambda_2 d - \alpha X_0 - \alpha \lambda_2 d \\ &= X_0 + (\alpha \lambda_1 + (1 - \alpha) \lambda_2) d \end{aligned}$$

Since $\alpha \lambda_1 + (1 - \alpha) \lambda_2 \geq 0$

$\therefore\ \alpha \lambda_1 + (1 - \alpha) \lambda_2 \epsilon S$

$\therefore\ S$ is a convex set.