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 6

<u>Theorem (6):</u>

If

1: $\boldsymbol{f}$: $\boldsymbol{R^n \rightarrow R}$ ***has continuous first partial derivatives in a convex set*** $\boldsymbol{D \subset R^n}$.

2: $\boldsymbol{X^*}$ ***is a local minimizer or maximizer of*** $\boldsymbol{f}$ ***over*** $\boldsymbol{D}$.

Then $\boldsymbol{g(X^*) = 0}$ ***, where g is the gradient of*** $\boldsymbol{f}$.

<u>Note (14):</u>

The converse of Theorem (6) is not in general true.

<u>Note (15):</u>

If $\boldsymbol{X^*}$ **is a local minimizer of** $\boldsymbol{f}$ ***over D***, **then** $\boldsymbol{X^*}$ **is also critical point of** $\boldsymbol{f}$, **but the converse is not true. For example the function** $\boldsymbol{f: R \rightarrow R}$ **defined by** $\boldsymbol{f(x) = x^3, x \in R}$ ***has*** $\boldsymbol{x^* = 0}$ ***is*** **a critical point but it is not a local minimizer of** $\boldsymbol{f}$.

Definition (16):

Let $f: R^n \to R$ *has continuous first partial derivatives in a convex set* $D \subset R^n$. The point $X^* \epsilon D$ is called **a saddle point** if and only if X^* is a critical point of f but it is not a local minimizer or maximizer of f.

Theorem (7):

If

1: $f: R^n \to R$ *has continuous first* and second partial derivatives in an open convex set D *containing* X^*.

2: X^* is a critical point of f *in* D.

3: $G(X^*)$ is positive definite matrix, where G is the Hessian matrix of f.

Then X^* is a strict local minimizer of f *over* D.

Example (12):

Let f: $R^2 \rightarrow R$ *be defined by* $f(X) = x_1^2 + x_1x_2 + x_2^2$, $X \epsilon R^2$.

1: Find the gradient vector of f.

2: Find the critical points of f.

3: Find the Hessian matrix at critical points.

4: Determine whether the critical points are local minimizers or not.

Solution:

1: The gradient vector of f is defined by

$$g(X) = [\frac{\partial f}{\partial x_1} , \frac{\partial f}{\partial x_2}]^T.$$

$$\frac{\partial f}{\partial x_1} = 2x_1 + x_2 \ , \ \frac{\partial f}{\partial x_2} = x_1 + 2x_2 \ .$$

$$\therefore g(X) = [2x_1 + x_2, x_1 + 2x_2]^T \ .$$

2: Set $g(X) = 0$.

$$\therefore 2x_1 + x_2 = 0 \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots . \; (1)$$

And

$$x_1 + 2x_2 = 0 \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \quad (2)$$

Multiply (1) by 2, we get

$$4x_1 + 2x_2 = 0 \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots . . (3)$$

From (2) and (3) we get $x_1 = 0$ **.**

From (1) we get $x_2 = 0$.

$\therefore The\ critical\ point\ is\ [0\ ,0]^T$.

3: The Hessian matrix of f is defined by

$$G(X) = \left[\frac{\partial^2 f}{\partial x_i \partial x_j}\right] \; i, j = 1, 2.$$

$$\frac{\partial^2 f}{\partial x_1^2} = 2\ , \frac{\partial^2 f}{\partial x_1 \partial x_2} = 1\ , \frac{\partial^2 f}{\partial x_2 \partial x_1} = 1\ , \frac{\partial^2 f}{\partial x_2^2} = 2.$$

$$\therefore G(X) = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}.$$

4: By using Theorem (7), we must prove that $G(X)$ is positive definite matrix.

$Let\ X \in R^2$.

$$\therefore X^T G X = [x_1, x_2]\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = [2x_1 + x_2\,, x_1 + 2x_2]\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$= 2x_1^2 + x_2x_1 + x_1x_2 + 2x_2^2$$

$$= 2x_1^2 + 2x_1x_2 + 2x_2^2$$

$$= 2x_1^2 + 2x_1x_2 + 2x_2^2 = 2[x_1^2 + x_1x_2 + x_2^2]$$

$$= 2[x_1^2 + x_1x_2 + \frac{1}{4}x_2^2 + \frac{3}{4}x_2^2]$$

$$= 2\left[\left(x_1 + \frac{1}{2}x_2\right)^2 + \frac{3}{4}x_2^2\right] > 0, for\ X \neq 0\,.$$

$\therefore G(X)$ is positive definite matrix.

$\therefore$By Theorem (7), the critical point $[0, 0]^T$ is a strict local minimizer of f.

Example (13):

Let $f: R^2 \rightarrow R$ *be defined by* $f(X) = x_1^3 - 2x_1^2 x_2 + x_2^2$, $X \in R^2$.

1: Find the gradient vector of f.

2: Find the critical points of f.

3: Find the Hessian matrix at critical points.

4: Determine whether the critical points are local minimizers or not.

Solution:

1: The gradient vector of f is defined by

$$g(X) = [\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}]^T.$$

$$\frac{\partial f}{\partial x_1} = 3x_1^2 - 4x_1 x_2, \frac{\partial f}{\partial x_2} = -2x_1^2 + 2x_2.$$

$$\therefore g(X) = [3x_1^2 - 4x_1 x_2, -2x_1^2 + 2x_2]^T.$$

2: Set $g(X) = 0$.

$$3x_1^2 - 4x_1x_2 = 0 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots .. \ (1)$$

And

$$-2x_1^2 + 2x_2 = 0 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots .. \quad (2)$$

From (2) we get $x_2 = x_1^2 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots .. (3)$

From (3) and (1) we get

$$3x_1^2 - 4x_1^3 = 0 \ \rightarrow \ x_1^2\,[3 - 4x_1] = 0 \ \rightarrow \ x_1 = 0 \ or \ x_1 = \frac{3}{4}\,.$$

$$\therefore \ x_2 = 0 \ or \ x_2 = \frac{9}{16}\,.$$

$$\therefore The\ critical\ points\ are\ X^* = [0,0]^T \ and \ X^* = [\frac{3}{4}\ ,\frac{9}{16}]^T.$$

3: The Hessian matrix of f is defined by

$$G(X) = \left[\frac{\partial^2 f}{\partial x_i \partial x_j}\right] \; i, j = 1, 2.$$

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

$$\therefore G(X) = \begin{bmatrix} 6x_1 - 4x_2 & -4x_1 \\ -4x_1 & 2 \end{bmatrix}.$$

$$When\ X^* = [0, 0]^T, then\ G(X^*) = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}.$$

$$When\ X^* = [\frac{3}{4} \, , \frac{9}{16}]^T, then\ G(X^*) = \begin{bmatrix} \frac{9}{4} & -3 \\ -3 & 2 \end{bmatrix}.$$

4: We know that from Theorem (7) if the Hessian matrix is positive definite at the critical point, then the critical point is a strict local minimizer.

Let $X^* = [0,0]^T, then\ G(X^*) = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}.$

$Let\ V \epsilon R^2.$

$\therefore V^T G(X^*) V = [v_1\ , v_2] \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = [0, 2v_2] \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = 2v_2^2 > 0$

$for\ V \neq 0.$

$\therefore G(X^*) = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}\ is\ positive\ definite\ matrix.$

∴By Theorem (7), $X^* = [0,0]^T$ is a strict local minimizer of f.

Let $X^* = [\frac{3}{4}\ , \frac{9}{16}]^T\ and\ G(X^*) = \begin{bmatrix} \frac{9}{4} & -3 \\ -3 & 2 \end{bmatrix}.$

$Let\ V \epsilon R^2.$

$$\therefore V^T G(X^*) V = [v_1 , v_2] \begin{bmatrix} \frac{9}{4} & -3 \\ -3 & 2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = [\frac{9}{4} v_1 - 3v_2 , -3v_1 + 2v_2] \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

$$= \frac{9}{4} v_1^2 - 3v_2 v_1 - 3v_1 v_2 + 2v_2^2 = \frac{9}{4} v_1^2 - 6v_1 v_2 + 2v_2^2$$

$$= \frac{9}{4} [v_1^2 - \frac{24}{9} v_1 v_2 + \frac{8}{9} v_2^2] = \frac{9}{4} [v_1^2 - \frac{8}{3} v_1 v_2 + \frac{8}{9} v_2^2]$$

$$= \frac{9}{4} [v_1^2 - \frac{8}{3} v_1 v_2 + \frac{8}{9} v_2^2 + \frac{8}{9} v_2^2 - \frac{8}{9} v_2^2]$$

$$= \frac{9}{4} [v_1^2 - \frac{8}{3} v_1 v_2 + \frac{16}{9} v_2^2 - \frac{8}{9} v_2^2]$$

$$= \frac{9}{4} [\left(v_1 - \frac{4}{3} v_2\right)^2 - \frac{8}{9} v_2^2]$$

$Let\ V = [1, 0]^T \rightarrow V^T G(X^*) V = \frac{9}{4} > 0$. $Let\ V = [\frac{4}{3} , 1]^T \rightarrow V^T G(X^*) V = -2 < 0$.

Hence $G(X^*)$ is indefinite matrix and so nothing can be conclude about the nature of critical point $X^* = [\frac{3}{4} , \frac{9}{16}]^T$.