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 8
Solved Problems

Problem (1):

Let the function $f: R^2 \to R$ is defined by

$f(X) = x_1 e^{x_2} + x_1 x_2$. Does f has any critical points?

Solution:

1: We find the gradient vector as

$$g(X) = \left[\frac{\partial f}{\partial x_1} \quad \frac{\partial f}{\partial x_2}\right]^T = [e^{x_2} + x_2 \quad x_1 e^{x_2} + x_1]^T.$$

2: Put $g(X) = 0$

Since $e^{x_2} + x_2 \neq 0 \; for \; all \; x_2$ in R ((because $e^{x_2} > 0$))

Then we conclude that the function

$f(X) = x_1 e^{x_2} + x_1 x_2$ has no critical points.

Problem (2):

Show that the matrix $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}, where\ a > 0\ is$ positive definite if $ac > b^2$.

Solution:

Let $Y \neq 0\ in\ R^2$.

We want to prove $Y^T AY > 0$.

$$Y^T AY = [y_1 \quad y_2] \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

$$= [ay_1 + by_2 \quad by_1 + ay_2] \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = ay_1^2 + by_1y_2 + by_1y_2 + cy_2^2$$

$$= ay_1^2 + 2by_1y_2 + cy_2^2 = a[y_1^2 + \frac{2b}{a}y_1y_2 + \frac{c}{a}y_2^2]$$

$$= a[y_1^2 + \frac{2b}{a}y_1y_2 + \frac{b^2}{a^2}y_2^2 - \frac{b^2}{a^2}y_2^2 + \frac{c}{a}y_2^2]$$

$$= a[y_1^2 + \frac{2b}{a} y_1 y_2 + \frac{b^2}{a^2} y_2^2 - \frac{b^2}{a^2} y_2^2 + \frac{c}{a} y_2^2]$$

$$= a[\left(y_1 + \frac{b}{a} y_2\right)^2 + \frac{acy_2^2 - b^2 y_2^2}{a^2}]$$

$$= a[\left(y_1 + \frac{b}{a} y_2\right)^2 + \frac{(ac - b^2) y_2^2}{a^2}$$

Then

$$Y^T AY = a[\left(y_1 + \frac{b}{a} y_2\right)^2 + \frac{(ac - b^2) y_2^2}{a^2}]$$

Since $ac > b^2 \rightarrow ac - b^2 > 0.$

Hence $a[\left(y_1 + \frac{b}{a} y_2\right)^2 + \frac{(ac - b^2) y_2^2}{a^2}] > 0$

Then , we have $Y^T AY > 0$.

Thus the matrix A is positive definite.

Problem (3):

Let the function $f\colon R^2 \to R$ *is defined by* $f(X) = x_1^2 e^{x_2^2} + x_2^2$.
Show that the point $X^* = [0, 0]^T$ is the strong global minimizer of f *over* R^2.

Solution:

By Theorem (7):

X^* is critical point **+** $G(X^*)$ is positive definite matrix

⇓

X^* is a local minimizer point

By Theorem (5)

The function f $\textit{is strictly convex}$

⇕

$G(X^*)$ $\textit{is positive definite matrix}$

By Theorem (12)

If The function f $\textit{is strictly convex}$

⇓

The point X^* is unique global minimizer

1: We find the gradient vector as

$$g(X) = \left[\frac{\partial f}{\partial x_1} \quad , \frac{\partial f}{\partial x_2}\right]^T = \left[2x_1 e^{x_2^2} \quad , 2x_1^2 x_2 e^{x_2^2} + 2x_2\right]^T$$

2: We find the critical points as follows:

Put $g(X) = 0$

$$\therefore 2x_1 e^{x_2^2} = 0 \rightarrow x_1 = 0 .$$

And

$$2x_1^2 x_2 e^{x_2^2} + 2x_2 = 0 \rightarrow x_2 = 0 .$$

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

3: We find the Hessian matrix as follows:

$$\frac{\partial^2 f}{\partial x_1^2} = 2e^{x_2^2}, \quad \frac{\partial^2 f}{\partial x_1 \partial x_2} = 4x_1 x_2 e^{x_2^2} = \frac{\partial^2 f}{\partial x_2 \partial x_1},$$

$$\frac{\partial^2 f}{\partial x_2^2} = 4x_1^2 x_2^2 e^{x_2^2} + 2x_1^2 e^{x_2^2} + 2.$$

$$\therefore G(X^*) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

4: We prove $G(X^*)$ is positive definite matrix as follows:

Let $Y \neq 0\ in\ R^2$. We want to prove $Y^T G(X^*) Y > 0$.

$$Y^T G(X^*) Y = \begin{bmatrix} y_1 & y_2 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 2y_1 & 2y_2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

$$= 2y_1^2 + 2y_2^2 > 0, for\ Y \neq 0 .$$

∴By Theorem (7), we conclude that X^* is a local minimizer point.

Now, by Theorem (5), we conclude that the function $f\ is\ strictly\ convex\ function$.

Finally, by Theorem (12), we conclude that the point X^* is the unique global minimizer for the function f .

Problem (4):

Let $f: R^2 \to R$ *is defined by* $f(X) = 3x_1^2 + 4x_1x_2 - 4x_2^2$.

Show that the point $X^* = [0, 0]^T$ is a saddle point for f.

Solution:

1: We find the gradient of the function as follows:

$$g(X) = \left[\frac{\partial f}{\partial x_1} \quad , \frac{\partial f}{\partial x_2}\right]^T = [6x_1 + 4x_2 \quad 4x_1 - 8x_2]^T$$

2: We find the critical points as follows:

Put $g(X) = 0.$

$\therefore 6x_1 + 4x_2 = 0$ ……………………………….……….…. (1)

$4x_1 - 8x_2 = 0$ ………………………………………….…. (2)

Now, from (1) and (2), we get $x_1 = 0$ and $x_2 = 0$.

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

3: We find the Hessian matrix as follows:

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

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

4: Determine whether the Hessian matrix is positive definite or not.

Let $Y \neq 0\ in\ R^2$.

$$\therefore Y^T G(X^*) Y = \begin{bmatrix} y_1 & y_2 \end{bmatrix} \begin{bmatrix} 6 & 4 \\ 4 & -8 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

$$= \begin{bmatrix} 6y_1 + 4y_2 & 4y_1 - 8y_2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

$$= 6y_1^2 + 4y_1 y_2 + 4y_1 y_2 - 8y_2^2$$

$$\therefore Y^T G(X^*) Y = 6y_1^2 + 8y_1y_2 - 8y_2^2 = 6\left[y_1^2 + \frac{8}{6}y_1y_2 - \frac{8}{6}y_2^2\right]$$

$$= 6\left[y_1^2 + \frac{4}{3}y_1y_2 - \frac{4}{3}y_2^2\right]$$

$$= 6\left[y_1^2 + \frac{4}{3}y_1y_2 + \frac{4}{9}y_2^2 - \frac{4}{9}y_2^2 - \frac{4}{3}y_2^2\right]$$

$$= 6[(y_1 + \frac{2}{3}y_2)^2 - \frac{16}{9}y_2^2]$$

∴The last expression is not positive everywhere, for example if we take

$Y = [1,0]^T$, **then** $Y^T G(X^*) Y = 6 > 0$, ***while if we take***

$Y = [-\frac{2}{3}, 1]^T$, **then** $Y^T G(X^*) Y = -\frac{32}{3} < 0$.

∴$X^* = [0,0]^T$ **is not local minimizer for the function** f.

Now, we must prove $X^* = [0,0]^T$ **is not local maximizer.**

Let $h(X) = -f(X)$.

$\therefore$The Hessian matrix of the function $h(X)$ at X^* is given as

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

Now, determine whether $H(X^*)$ is positive definite or not.

Let $V \neq 0$.

$$\therefore V^T H(X^*) \mathrm{V} = \begin{bmatrix} v_1 & v_2 \end{bmatrix} \begin{bmatrix} -6 & -4 \\ -4 & 8 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

$$= \begin{bmatrix} -6v_1 - 4v_2 & , -4v_1 + 8v_2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

$$= -6v_1^2 - 4v_1v_2 - 4v_1v_2 + 8v_2^2 = -6v_1^2 - 8v_1v_2 + 8v_2^2$$

$$= -6\left[v_1^2 + \frac{8}{6}v_1v_2 - \frac{8}{6}v_2^2\right] = -6[(v_1 + \frac{2}{3}v_2)^2 - \frac{16}{9}v_2^2].$$

$\therefore$The expression term is not positive everywhere for example if we take

$V = [1, 0]^T$, **then** $V^T H(X^*) V = -6 < 0$, $\boldsymbol{while\ if\ we\ take}$
$V = [-\frac{2}{3}, 1]^T$, **then** $V^T G(X^*) V = \frac{32}{3} > 0$.
$\therefore X^* = [0, 0]^T$ **is not local minimizer for the function** $h = -f$.
$\therefore X^* = [0, 0]^T$ **is not local maximizer for the function** f
$\therefore X^* = [0, 0]^T$ **is not local maximizer and not local maximizer for the function** f.
Thus, we conclude that the point $X^* = [0, 0]^T$ **is a saddle point for the function** f.