Module - 2 Lecture Notes – 1

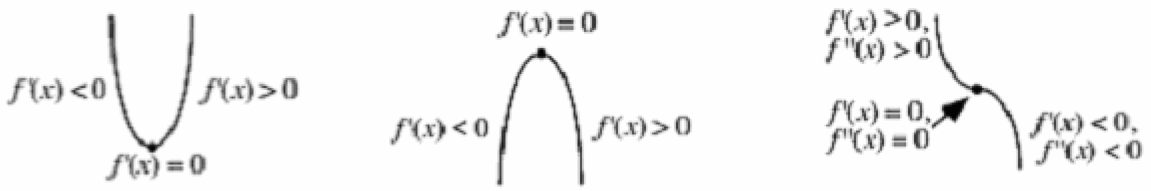
Stationary points: Functions of Single and Two Variables

Introduction

In this session, stationary points of a function are defined. The necessary and sufficient conditions for the relative maximum of a function of single or two variables are also discussed. The global optimum is also defined in comparison to the relative or local optimum.

Stationary points

For a continuous and differentiable function $f(x)$ a *stationary* point x^* is a point at which the slope of the function vanishes, i.e. $f'(x) = 0$ at $x = x^*$, where x^* belongs to its domain of definition.



minimum maximum inflection point

Fig. 1

A stationary point may be a minimum, maximum or an inflection point (Fig. 1).

Relative and Global Optimum

A function is said to have a *relative* or *local* minimum at $x = x^*$ if $f(x^*) \le f(x^* + h)$ for all sufficiently small positive and negative values of h, i.e. in the near vicinity of the point x^*. Similarly a point x^* is called a *relative* or *local* maximum if $f(x^*) \ge f(x^* + h)$ for all values of h sufficiently close to zero. A function is said to have a *global* or *absolute* minimum at $x = x^*$ if $f(x^*) \le f(x)$ for all x in the domain over which $f(x)$ is defined. Similarly, a function is said to have a *global* or *absolute* maximum at $x = x^*$ if $f(x^*) \geq f(x)$ for all x in the domain over which $f(x)$ is defined.

Figure 2 shows the global and local optimum points.

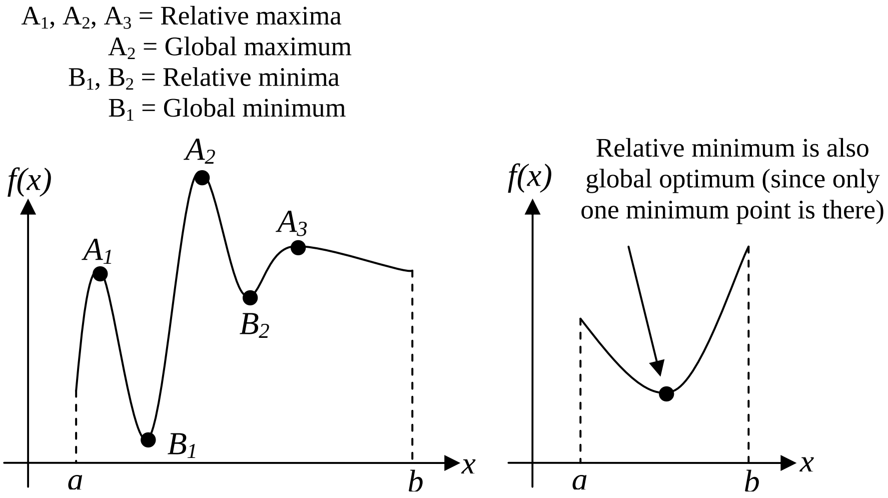


Fig. 2

Functions of a single variable

Consider the function $f(x)$ defined for $a \leq x \leq b$. To find the value of $x^* \in [a,b]$ such that $x = x^*$ maximizes $f(x)$ we need to solve a *single-variable optimization* problem. We have the following theorems to understand the necessary and sufficient conditions for the relative maximum of a function of a single variable.

Necessary condition: For a single variable function $f(x)$ defined for $x \in [a,b]$ which has a relative maximum at $x = x^*$, $x^* \in [a,b]$ if the derivative $f'(X) = df(x)/dx$ exists as a finite number at $x = x^*$ then $f'(x^*) = 0$. This can be understood from the following.

Proof.

Since $f\,'(x^*)$ is stated to exist, we have

$$f^{'}(x^*) = \lim_{h \to 0} \frac{f(x^*+h) - f(x^*)}{h} \tag{1}$$

From our earlier discussion on relative maxima we have $f(x^*) \geq f(x^*+h)$ for $h \to 0$. Hence

$$\frac{f(x^*+h) - f(x^*)}{h} \geq 0 \qquad h < 0 \tag{2}$$

$$\frac{f(x^*+h) - f(x^*)}{h} \leq 0 \qquad h > 0 \tag{3}$$

which implies for substantially small negative values of h we have $f(x^*) \geq 0$ and for substantially small positive values of h we have $f(x^*) \leq 0$. In order to satisfy both (2) and (3), $f(x^*)$ = 0. Hence this gives the necessary condition for a relative maxima at $x = x^*$ for $f(x)$.

It has to be kept in mind that the above theorem holds good for relative minimum as well. The theorem only considers a domain where the function is continuous and differentiable. It cannot indicate whether a maxima or minima exists at a point where the derivative fails to exist. This scenario is shown in Fig 3, where the slopes m_1 and m_2 at the point of a maxima are unequal, hence cannot be found as depicted by the theorem by failing for continuity. The theorem also does not consider if the maxima or minima occurs at the end point of the interval of definition, owing to the same reason that the function is not continuous, therefore not differentiable at the boundaries. The theorem does not say whether the function will have a maximum or minimum at every point where $f\,'(x) = 0$, since this condition $f\,'(x) = 0$ is for stationary points which include inflection points which do not mean a maxima or a minima. A point of inflection is shown already in Fig.1

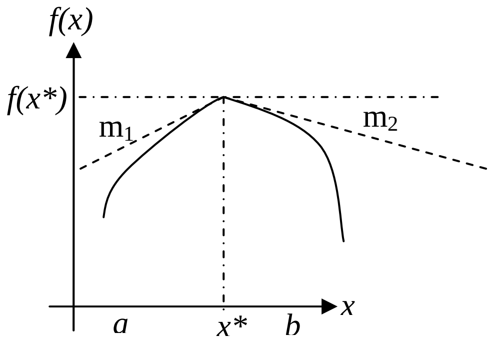


Fig. 3

Sufficient condition: For the same function stated above let $f'(x^*) = f''(x^*) = \ldots = f^{(n-1)}(x^*) = 0$, but $f^{(n)}(x^*) \neq 0$, then it can be said that $f(x^*)$ is (a) a minimum value of $f(x)$ if $f^{(n)}(x^*) > 0$ and n is even; (b) a maximum value of $f(x)$ if $f^{(n)}(x^*) < 0$ and n is even; (c) neither a maximum or a minimum if n is odd.

Proof

Applying the Taylor's theorem with remainder after n terms, we have

$$f(x^*+h) = f(x^*) + hf'(x^*) + \frac{h^2}{2!} f''(x^*) + \ldots + \frac{h^{n-1}}{(n-1)!} f^{(n-1)}(x^*) + \frac{h^n}{n!} f^n(x^*+\theta h) \qquad (4)$$

for $0<\theta<1$

since $f'(x^*) = f''(x^*) = \ldots = f^{(n-1)}(x^*) = 0$, (4) becomes

$$f(x^*+h) - f(x^*) = \frac{h^n}{n!} f^n(x^*+\theta h) \qquad (5)$$

As $f^{(n)}(x^*) \neq 0$, there exists an interval around x^* for every point x of which the nth derivative $f^{(n)}(x)$ has the same sign, viz., that of $f^{(n)}(x^*)$. Thus for every point $x^* + h$ of this interval, $f^{(n)}(x^* + h)$ has the sign of $f^{(n)}(x^*)$. When n is even $\frac{h^n}{n!}$ is positive irrespective of the sign of h, and hence $f(x^*+h) - f(x^*)$ will have the same sign as that of $f^{(n)}(x^*)$. Thus x^* will be a relative minimum if $f^{(n)}(x^*)$ is positive, with $f(x)$ convex around x^*, and a relative maximum if

$f^{(n)}(x^*)$ is negative, with $f(x)$ concave around x^*. When n is odd, $\frac{h^n}{n!}$ changes sign with the change in the sign of h and hence the point x^* is neither a maximum nor a minimum. In this case the point x^* is called a *point of inflection*.

Example 1.

Find the optimum value of the function $f(x) = x^2 + 3x - 5$ and also state if the function attains a maximum or a minimum.

Solution

$f'(x) = 2x + 3 = 0$ for maxima or minima.

or x^* = -3/2

$f''(x^*) = 2$ which is positive hence the point x^* = -3/2 is a point of minima and the function attains a minimum value of -29/4 at this point.

Example 2.

Find the optimum value of the function $f(x) = (x-2)^4$ and also state if the function attains a maximum or a minimum.

Solution

$f'(x) = 4(x-2)^3 = 0$ for maxima or minima.

or $x = x^*$ = 2 for maxima or minima.

$f''(x^*) = 12(x^*-2)^2 = 0$ at x^* = 2

$f'''(x^*) = 24(x^*-2) = 0$ at x^* = 2

$f''''(x^*) = 24$ at $x^* = 2$

Hence $f^n(x)$ is positive and n is even hence the point $x = x^* = 2$ is a point of minimum and the function attains a minimum value of 0 at this point.

Example 3.

Analyze the function $f(x) = 12x^5 - 45x^4 + 40x^3 + 5$ and classify the stationary points as maxima, minima and points of inflection.

Solution

$$f'(x) = 60x^4 - 180x^3 + 120x^2 = 0$$

$$=> x^4 - 3x^3 + 2x^2 = 0$$

or $\quad x = 0, 1, 2$

Consider the point $x = x^* = 0$

$$f''(x^*) = 240(x^*)^3 - 540(x^*)^2 + 240x^* = 0 \text{ at } x^* = 0$$

$$f'''(x^*) = 720(x^*)^2 - 1080x^* + 240 = 240 \text{ at } x^* = 0$$

Since the third derivative is non-zero, $x = x^* = 0$ is neither a point of maximum or minimum but it is a point of inflection.

Consider $x = x^* = 1$

$$f''(x^*) = 240(x^*)^3 - 540(x^*)^2 + 240x^* = -60 \text{ at } x^* = 1$$

Since the second derivative is negative the point $x = x^* = 1$ is a point of local maxima with a maximum value of $f(x) = 12 - 45 + 40 + 5 = 12$

Consider $x = x^* = 2$

$$f''(x^*) = 240(x^*)^3 - 540(x^*)^2 + 240x^* = 240 \text{ at } x^* = 2$$

Since the second derivative is positive, the point $x = x^* = 2$ is a point of local minima with a minimum value of $f(x) = -11$

Example 4.

The horse power generated by a Pelton wheel is proportional to *u(v-u)* where *u* is the velocity of the wheel, which is variable and *v* is the velocity of the jet which is fixed. Show that the efficiency of the Pelton wheel will be maximum at $u = v/2$.

Solution

$$f = \mathrm{K}.u(v-u)$$
$$\frac{\partial f}{\partial u} = 0 \Rightarrow \mathrm{K}v - 2\mathrm{K}u = 0$$
$$\text{or} \quad u = \frac{v}{2}$$

where K is a proportionality constant (assumed positive).

$$\left.\frac{\partial^2 f}{\partial u^2}\right|_{u=\frac{v}{2}} = -2\mathrm{K} \text{ which is negative.}$$

Hence, f is maximum at $u = \frac{v}{2}$

Functions of two variables

This concept may be easily extended to functions of multiple variables. Functions of two variables are best illustrated by contour maps, analogous to geographical maps. A contour is a line representing a constant value of $f(x)$ as shown in Fig.4. From this we can identify *maxima, minima* and *points of inflection.*

Necessary conditions

As can be seen in Fig. 4 and 5, perturbations from points of local minima in any direction result in an increase in the response function *f*(*x*), i.e. the slope of the function is zero at this point of local minima. Similarly, at *maxima* and *points of inflection* as the slope is zero, the first derivatives of the function with respect to the variables are zero.

Which gives us $\frac{\partial f}{\partial x_1} = 0; \frac{\partial f}{\partial x_2} = 0$ at the stationary points, i.e., the gradient vector of $f(\mathbf{X})$, $\Delta_x f$ at $\mathbf{X} = \mathbf{X}^* = [x_1, x_2]$ defined as follows, must equal zero:

$$\Delta_x f = \begin{bmatrix} \frac{\partial f}{\partial x_1}(\mathrm{X}^*) \\ \frac{\partial f}{\partial x_2}(\mathrm{X}^*) \end{bmatrix} = 0$$

This is the necessary condition.

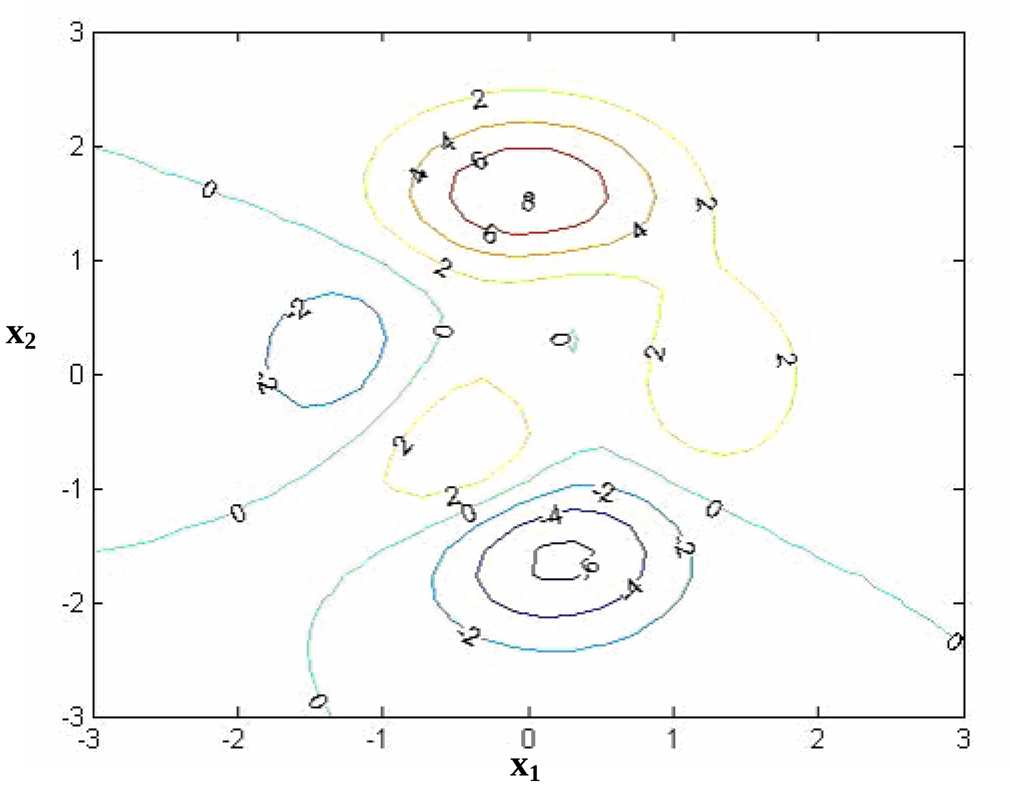


Fig. 4

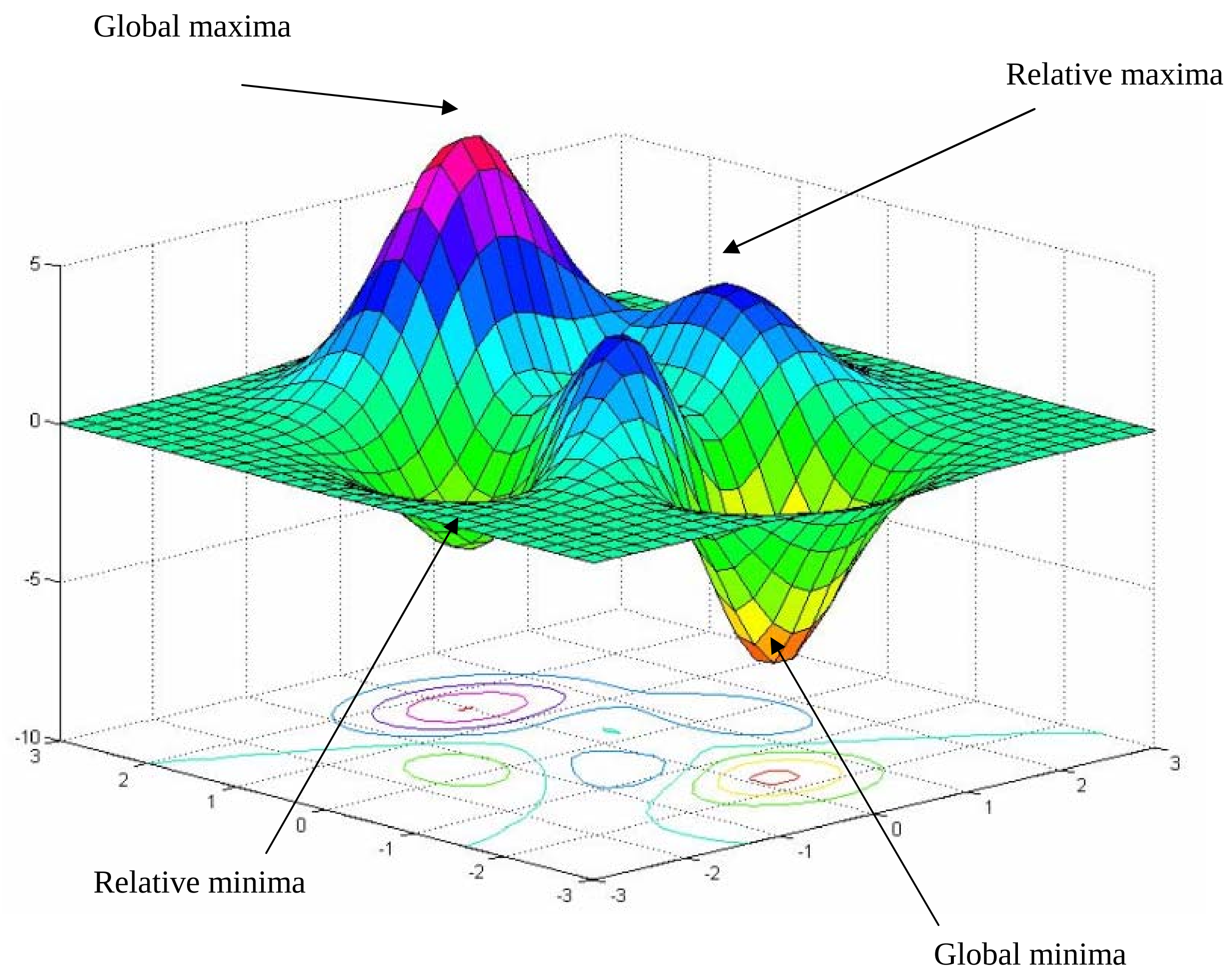


Fig. 5

Sufficient conditions

Consider the following second order derivatives:

$$\frac{\partial^2 f}{\partial x_1^2}; \frac{\partial^2 f}{\partial x_2^2}; \frac{\partial^2 f}{\partial x_1 \partial x_2}$$

The Hessian matrix defined by **H** is made using the above second order derivatives.

$$\mathbf{H} = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1^2} & \dfrac{\partial^2 f}{\partial x_1 \partial x_2} \\ \dfrac{\partial^2 f}{\partial x_1 \partial x_2} & \dfrac{\partial^2 f}{\partial x_2^2} \end{pmatrix}_{[x_1, x_2]}$$

a) If **H** is positive definite then the point **X** = [x_1, x_2] is a point of local minima.

b) If **H** is negative definite then the point **X** = [x_1, x_2] is a point of local maxima.

c) If **H** is neither then the point **X =** [x_1, x_2] is neither a point of maxima nor minima.

A square matrix is positive definite if all its eigen values are positive and it is negative definite if all its eigen values are negative. If some of the eigen values are positive and some negative then the matrix is neither positive definite or negative definite.

To calculate the eigen values λ of a square matrix then the following equation is solved.

$$|\mathbf{A} - \lambda \mathbf{I}| = 0$$

The above rules give the sufficient conditions for the optimization problem of two variables.

Optimization of multiple variable problems will be discussed in detail in lecture notes 3 (Module 2).

<u>Example 5.</u>

Locate the stationary points of $f(\mathbf{X})$ and classify them as relative maxima, relative minima or neither based on the rules discussed in the lecture.

$f(\mathbf{X}) = 2x_1^3/3 - 2x_1x_2 - 5x_1 + 2x_2^2 + 4x_2 + 5$

Solution

$$\Delta_x f = \begin{bmatrix} \frac{\partial f}{\partial x_1}(X^*) \\ \frac{\partial f}{\partial x_2}(X^*) \end{bmatrix} = \begin{bmatrix} 2x_1^2 - 2x_2 - 5 \\ -2x_1 + 4x_2 + 4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

From $\frac{\partial f}{\partial x_2}(\mathrm{X}) = 0$, $x_1 = 2x_2 + 2$

From $\frac{\partial f}{\partial x_1}(\mathrm{X}) = 0$

$$2(2x_2 + 2)^2 - 2x_2 - 5 = 0$$

$$8x_2^2 + 14x_2 + 3 = 0$$

$$(2x_2 + 3)(4x_2 + 1) = 0$$

$$x_2 = -3/2 \quad \text{or} \quad x_2 = -1/4$$

so the two stationary points are

$$X_1 = [-1,-3/2]$$

and

$$X_2 = [3/2,-1/4]$$

The Hessian of $f(\mathbf{X})$ is

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

$$\mathbf{H} = \begin{bmatrix} 4x_1 & -2 \\ -2 & 4 \end{bmatrix}$$

$$\left|\lambda\mathbf{I} - \mathbf{H}\right| = \begin{vmatrix} \lambda - 4x_1 & 2 \\ 2 & \lambda - 4 \end{vmatrix}$$

At X_1= [-1,-3/2],

$$\left|\lambda\mathbf{I} - \mathbf{H}\right| = \begin{vmatrix} \lambda + 4 & 2 \\ 2 & \lambda - 4 \end{vmatrix} = (\lambda + 4)(\lambda - 4) - 4 = 0$$

$$\lambda^2 - 16 - 4 = 0$$

$$\lambda^2 = 12$$

$$\lambda_1 = +\sqrt{12} \qquad \lambda_2 = -\sqrt{12}$$

Since one eigen value is positive and one negative, X_1 is neither a relative maximum nor a relative minimum.

At X_2 = [3/2,-1/4]

$$\left|\lambda\mathbf{I} - \mathbf{H}\right| = \begin{vmatrix} \lambda - 6 & 2 \\ 2 & \lambda - 4 \end{vmatrix} = (\lambda - 6)(\lambda - 4) - 4 = 0$$

$$\lambda^2 - 10\lambda + 20 = 0$$

$$\lambda_1 = 5 + \sqrt{5} \qquad \lambda_2 = 5 - \sqrt{5}$$

Since both the eigen values are positive, X_2 is a local minimum.

Minimum value of f(x) is -0.375.

Example 6

The ultimate strength attained by concrete is found to be based on a certain empirical relationship between the ratios of cement and concrete used. Our objective is to maximize strength attained by hardened concrete, given by $f(\mathbf{X}) = 20 + 2x_1 - x_1^2 + 6x_2 - 3x_2^2/2$, where x_1 and x_2 are variables based on cement and concrete ratios.

Solution

Given $f(\mathbf{X}) = 20 + 2x_1 - x_1^2 + 6x_2 - 3x_2^2/2$; where $\mathbf{X} = [x_1, x_2]$

The gradient vector $\Delta_x f = \begin{bmatrix} \frac{\partial f}{\partial x_1}(X^*) \\ \frac{\partial f}{\partial x_2}(X^*) \end{bmatrix} = \begin{bmatrix} 2 - 2x_1 \\ 6 - 3x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, to determine stationary point $\mathbf{X}^*$.

Solving we get $\mathbf{X}^* = [1,2]$

$$\frac{\partial^2 f}{\partial x_1^2} = -2; \frac{\partial^2 f}{\partial x_2^2} = -3; \frac{\partial^2 f}{\partial x_1 \partial x_2} = 0$$

$$\mathbf{H} = \begin{bmatrix} -2 & 0 \\ 0 & -3 \end{bmatrix}$$

$$|\lambda \mathbf{I} - \mathbf{H}| = \begin{vmatrix} \lambda + 2 & 0 \\ 0 & \lambda + 3 \end{vmatrix} = (\lambda + 2)(\lambda + 3) = 0$$

Here the values of λ do not depend on $\mathbf{X}$ and λ_1 = -2, λ_2 = -3. Since both the eigen values are negative, $f(\mathbf{X})$ is concave and the required ratio x_1:x_2 = 1:2 with a global maximum strength of $f(\mathbf{X})$ = 27 units.