Optimization using Calculus

Stationary Points: Functions of Single and Two Variables

Objectives

- To define stationary points
- Look into the necessary and sufficient conditions for the relative maximum of a function of a single variable and for a function of two variables.
- To define the global optimum 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 function vanishes, i.e. $f'(x) = 0$ at $x = x^*$. x^* belongs to its domain of definition.
- A stationary point may be a minimum, maximum or an inflection point

Stationary points

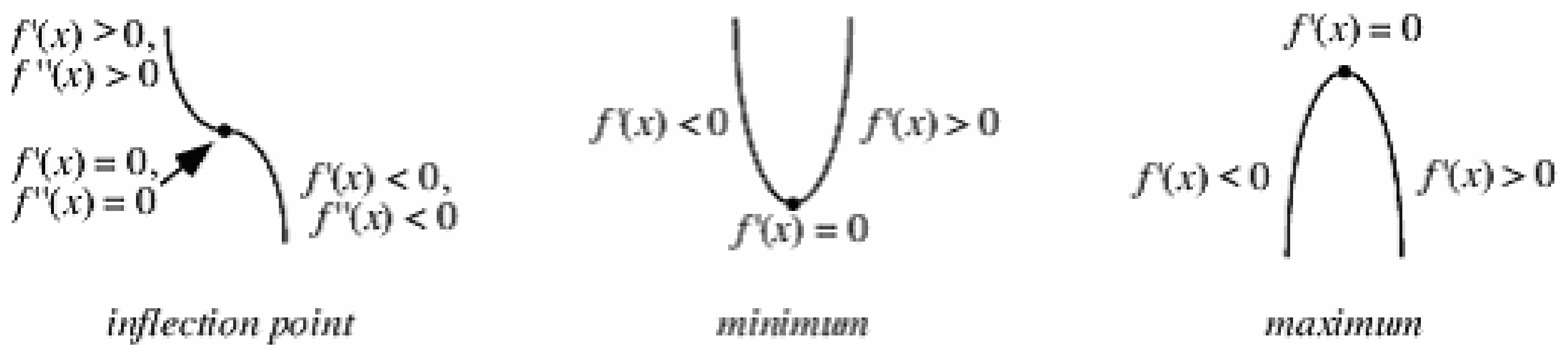


Figure showing the three types of stationary points (a) inflection point (b) minimum (c) maximum

Relative and Global Optimum

- A function is said to have a *relative* or *local* minimum at $x = x^*$ if $f(x^*) \leq f(x+h)$ 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 $f(x^*) \geq f(x+h)$ for all values of h sufficiently close to zero.
- A function is said to have a *global* or *absolute* minimum at $x = x^*$ if $f(x^*) \leq f(x)$ 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.

Relative and Global Optimum …contd.

A_1, A_2, A_3 = Relative maxima
A_2 = Global maximum
B_1, B_2 = Relative minima
B_1 = Global minimum

Relative minimum is also global optimum

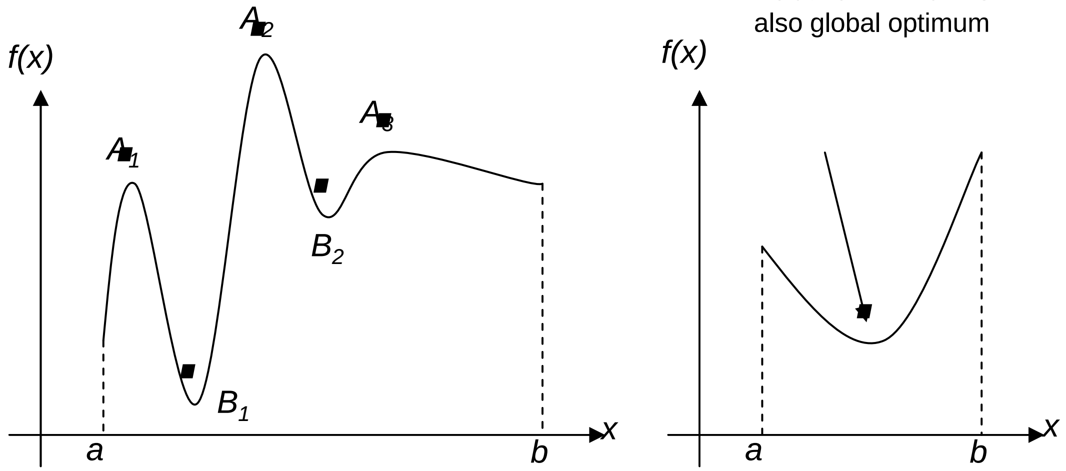


Fig. 2

Functions of a single variable

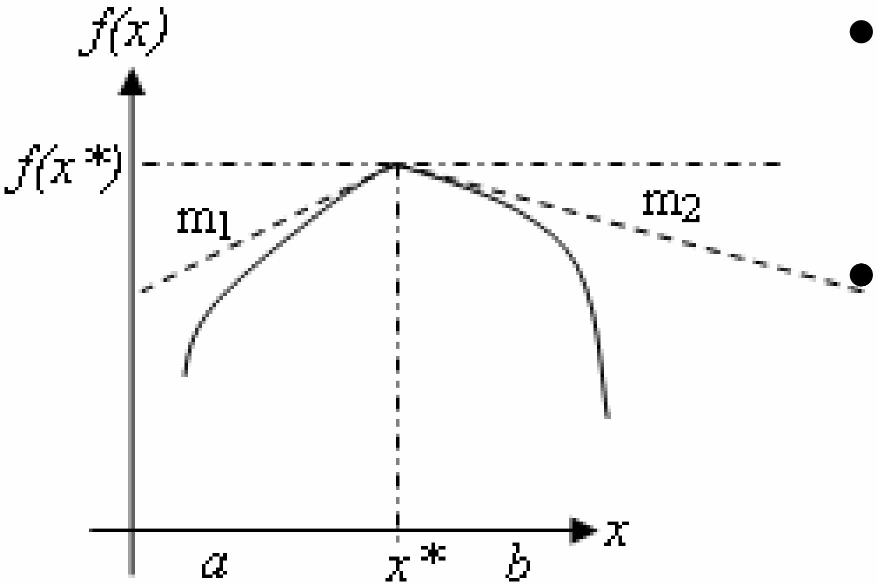
- Consider the function $f(x)$ defined for $a \le x \le b$
- To find the value of $x^* \in [a,b]$ such that x^* maximizes $f(x)$ we need to solve a *single-variable optimization* problem.
- We have the following theorems to understand the necessary and sufficient conditions for the relative maximum of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

Functions of a single variable …contd.

- **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

  $f^{`}(x^*) = 0.$
- We need to keep 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 derivative.
- It does not indicate the outcome if a maxima or minima exists at a point where the derivative fails to exist. This scenario is shown in the figure below, 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.

Functions of a single variable …contd.



Some Notes:

- The theorem does not consider if the maxima or minima occurs at the end point of the interval of definition.
- The theorem does not say that 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.

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

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.*

<u>*Solution*</u>

$f'(x) = 2x + 3 = 0$ 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:

$$f'(x) = 4(x-2)^3 = 0$$ 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:__

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

$$\Rightarrow x^4 - 3x^3 + 2x^2 = 0$$

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

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

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

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

Example 3 …contd.

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 $x = x^* = 1$

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

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 $x = x^* = 2$

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

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: $f = K.u(v-u)$

$$\frac{\partial f}{\partial u} = 0 \Rightarrow Kv - 2Ku = 0$$

or $u = \frac{v}{2}$

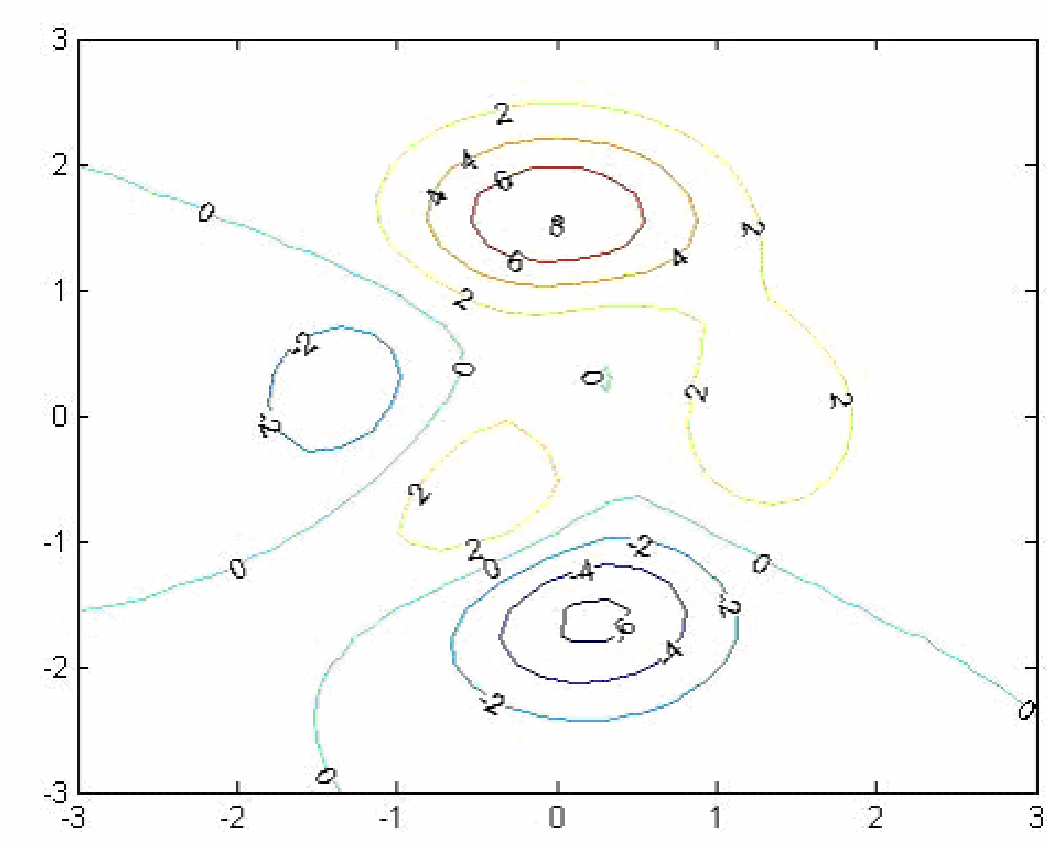
where K is a proportionality constant (assumed positive).

$\left.\frac{\partial^2 f}{\partial u^2}\right|_{u=\frac{v}{2}} = -2K$ which is negative. Hence, f is maximum at $u = \frac{v}{2}$

Functions of two variables

- The concept discussed for one variable functions 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 the following figure. From this we can identify *maxima*, *minima* and *points of inflection*.

A contour plot



Necessary conditions

- As can be seen in the above contour map, 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 derivative of the function with respect to the variables are zero.

Necessary conditions …contd.

- 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 $\boldsymbol{X} = \boldsymbol{X}^* = [x_1, x_2]$ 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.

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.

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

Sufficient conditions …contd.

- The value of determinant of the **H** is calculated and
  - if **H** is positive definite then the point $\boldsymbol{X} = [x_1, x_2]$ is a point of local minima.
  - if **H** is negative definite then the point $\boldsymbol{X} = [x_1, x_2]$ is a point of local maxima.
  - if **H** is neither then the point $\boldsymbol{X} = [x_1, x_2]$ is neither a point of maxima nor minima.

Example 5

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

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

<u>*Solution*</u>

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

Example 5 …contd.

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

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

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

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

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

So the two stationary points are

X_1 = [-1,-3/2] and X_2 = [3/2,-1/4]

Example 5 …contd.

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

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

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

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

Example 5 …contd.

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

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

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

Since both the eigen values are positive, X_{-2} is a local minimum.

Minimum value of *f(x)* is -0.375

Example 6

Maximize $f(\mathbf{X}) = 20 + 2x_1 - x_1^2 + 6x_2 - 3x_2^2 / 2$

$$\Delta_x f = \begin{bmatrix} \dfrac{\partial f}{\partial x_1}(X^*) \\ \dfrac{\partial f}{\partial x_2}(X^*) \end{bmatrix} = \begin{bmatrix} 2 - 2x_1 \\ 6 - 3x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow \boldsymbol{X}^* = [1,2]$$

$$\frac{\partial^2 f}{\partial x_1{}^2} = -2; \frac{\partial^2 f}{\partial x_2{}^2} = -3; \frac{\partial^2 f}{\partial x_1 \partial x_2} = 0 \; ; \qquad \mathbf{H} = \begin{bmatrix} -2 & 0 \\ 0 & -3 \end{bmatrix}$$

Example 6 …contd.

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

$$\lambda_1 = -2 \quad \textit{and} \quad \lambda_2 = -3$$

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 $f(\mathbf{X}) = 27$ MPa

Thank you