Introduction to Optimization

Rajib Kumar Bhattacharjya

Professor

Department of Civil Engineering

IIT Guwahati

Email: rkbc@iitg.ernet.in

Single variable optimization

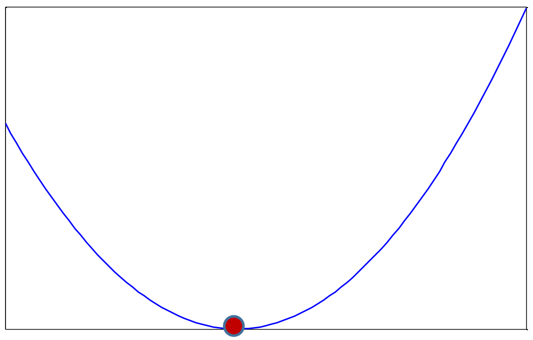
Objective function is defined as

Minimization/Maximization $f(x)$

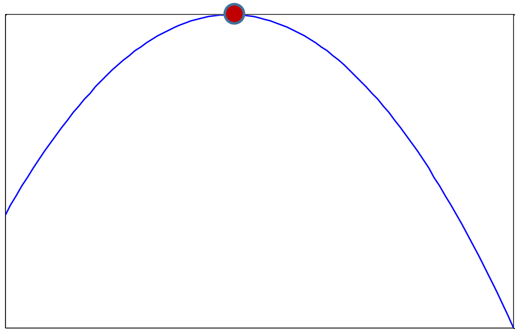
Single variable optimization

Stationary points

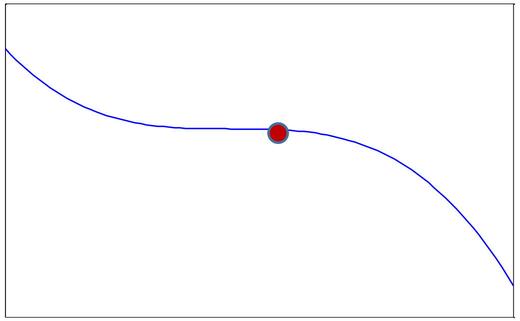
For a continuous and differentiable function $f(x)$, *a stationary point* x^* *is a point at which the* slope of the function is zero, *i.e.* $f'(x) = 0$ *at* $x = x^*$,



Minimum



Maximum



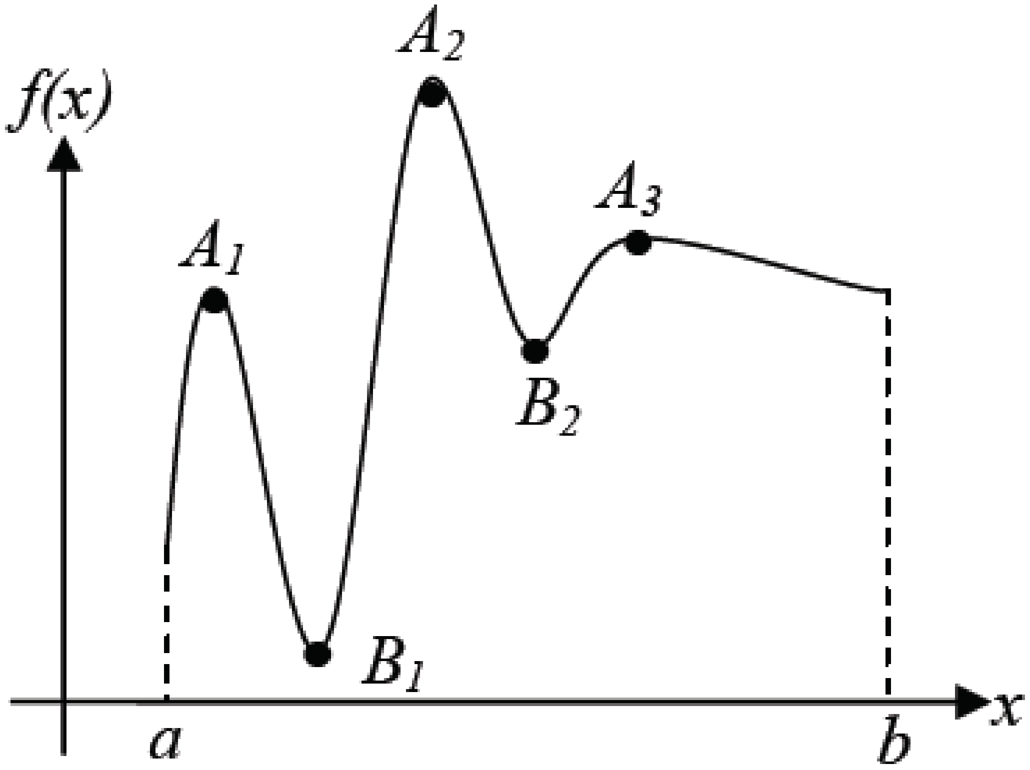
Inflection point

Global minimum and maximum

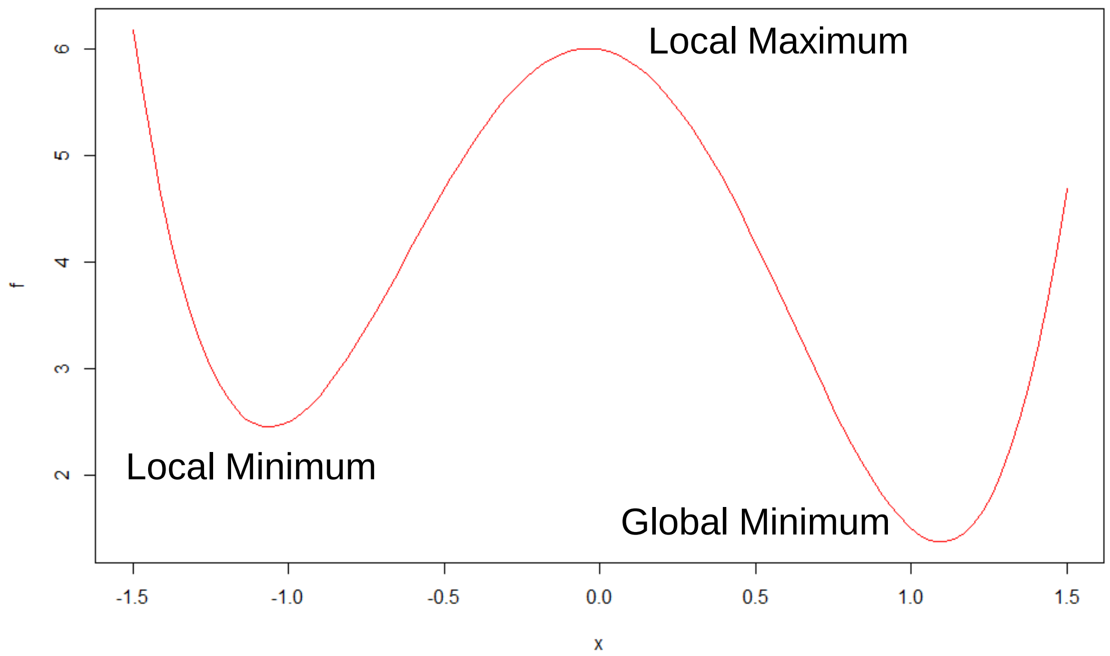
A function is said to have a *global or absolute minimum at* $x = x^*$ *if* $f(x^*) \leq f(x)$ *for all* x *in the domain over which* $f(x)$ *is defined.*

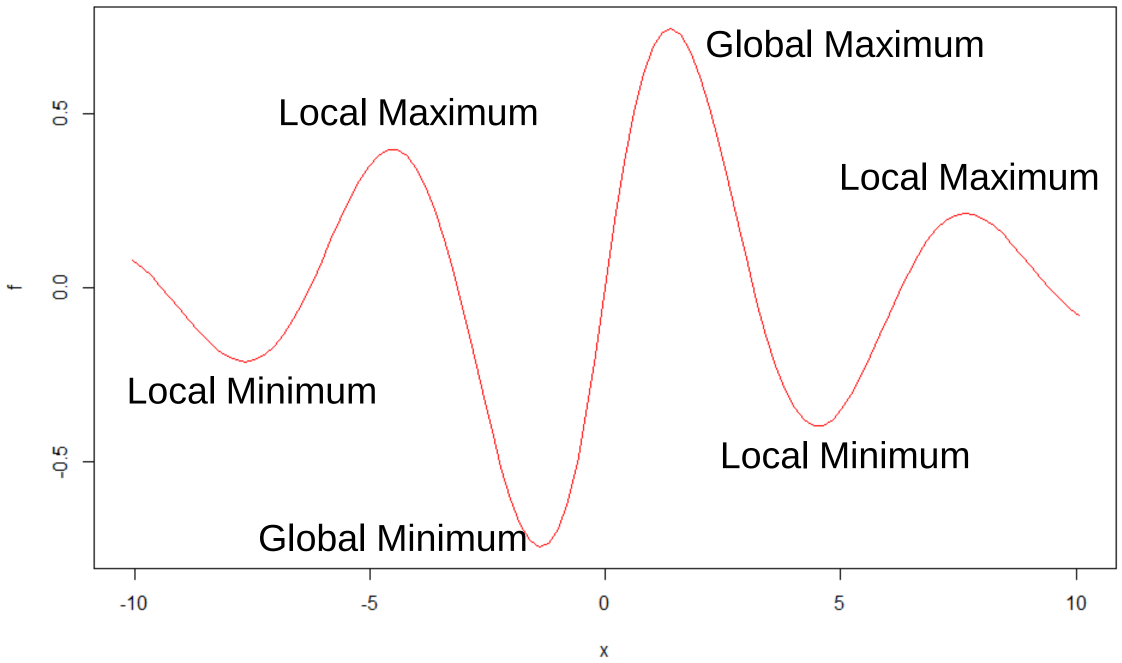
A function is said to have a *global or absolute maximum at* $x = x^*$ *if* $f(x^*) \geq f(x)$ *for all* x *in the domain over which* $f(x)$ *is defined.*

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



Global minimum and maximum





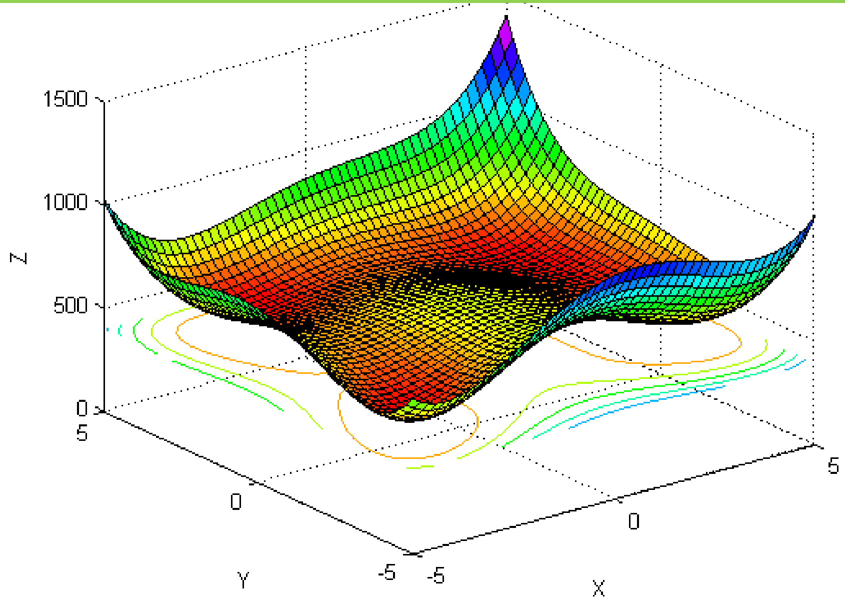
$$f = 2x^4 - 7x^2 - \frac{1}{2}x + 6$$

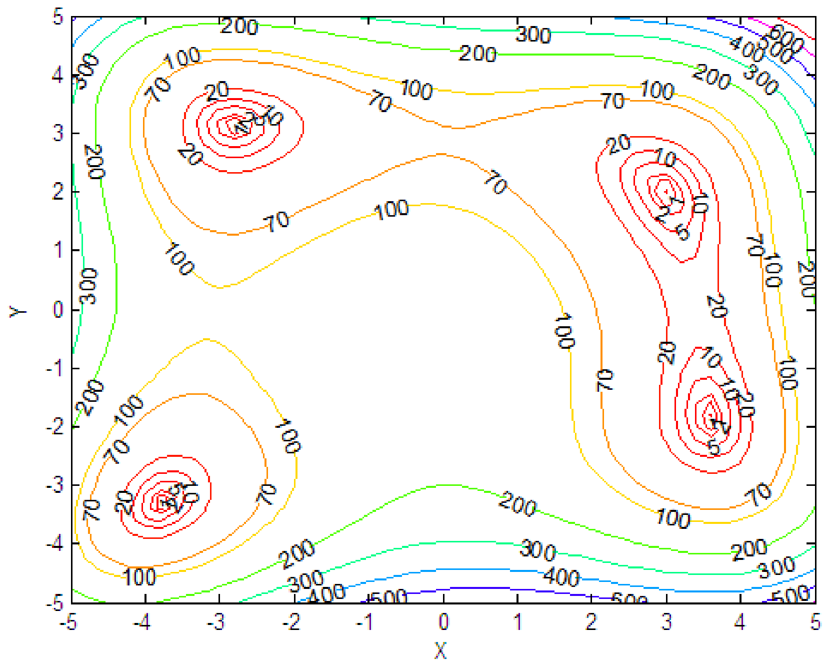
$$-1.5 \le x \le 1.5$$

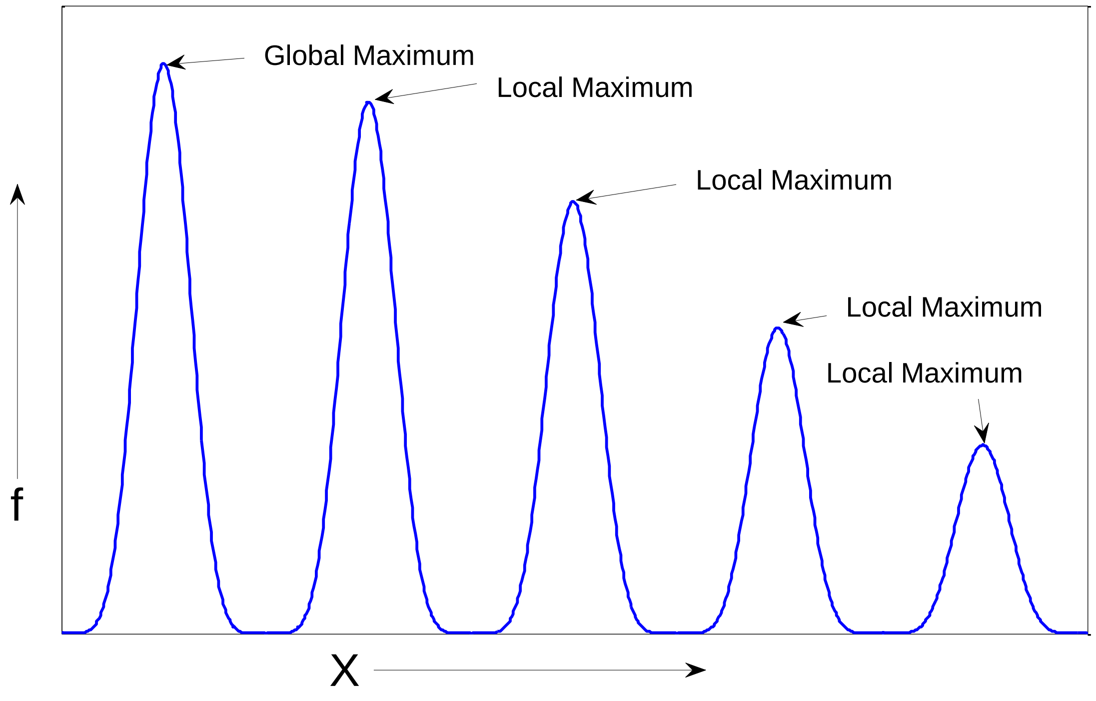
$$f = \sin(x) exp\left(-\left|\frac{x}{5}\right|\right)$$

$$-3.2\pi \le x \le 3.2\pi$$

Introduction to optimization







Necessary and sufficient conditions for optimality

Necessary condition

If a function $f(x)$ is defined in the interval $a \leq x \leq b$ and has a relative minimum at $x = x^*$, Where $a \leq x^* \leq b$ and if $f^{/}(x)$ exists as a finite number at $x = x^*$, then $f^{/}(x^*) = 0$

Proof

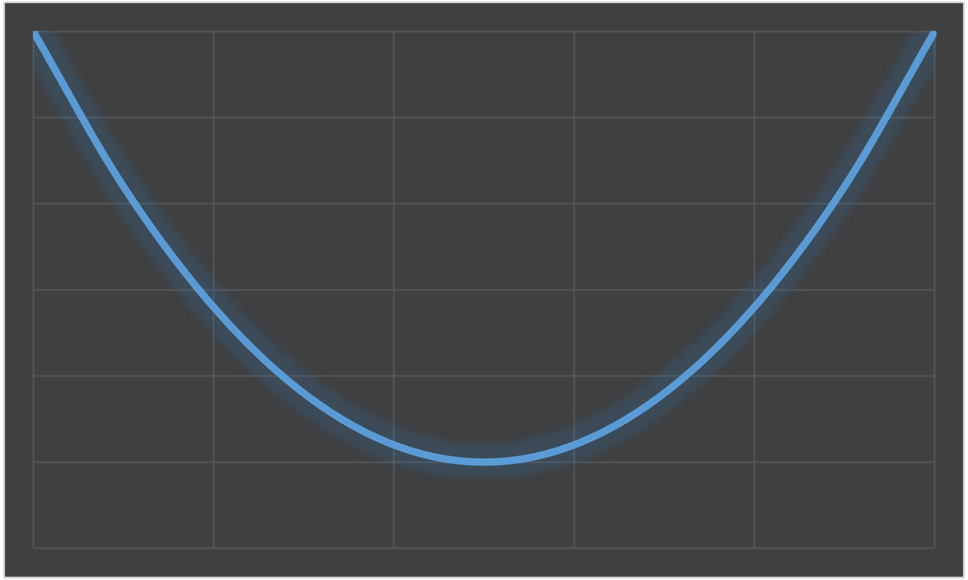
$$f^{/}(x^*) = \lim_{h \to 0} \frac{f(x^* + h) - f(x^*)}{h}$$

Since x^* is a relative minimum $\qquad f(x^*) \leq f(x^* + h)$

For all values of h sufficiently close to zero, hence

$$\frac{f(x^* + h) - f(x^*)}{h} \geq 0 \qquad \text{if } h \geq 0$$

$$\frac{f(x^* + h) - f(x^*)}{h} \leq 0 \qquad \text{if } h \leq 0$$



Necessary and sufficient conditions for optimality

Thus

$f^{/}(x^*) \geq 0$ If h tends to zero through +ve value

$f^{/}(x^*) \leq 0$ If h tends to zero through -ve value

Thus only way to satisfy both the conditions is to have

$$\boldsymbol{f^{/}(x^*) = 0}$$

Note:

- This theorem can be proved if x^* is a relative maximum
- Derivative must exist at x^*
- The theorem does not say what happens if a minimum or maximum occurs at an end point of the interval of the function
- It may be an inflection point also.

Sufficient conditions for optimality

Sufficient condition

Suppose at point x^*, the first derivative is zero and first nonzero higher derivative is denoted by n, then

1. *If n is odd, x^* is an inflection point*
2. *If n is even, x^* is a local optimum*
   - ✓ *If the derivative is positive, x^* is a local minimum*
   - ✓ *If the derivative is negative, x^* is a local maximum*

$$f'(x^*) = 0$$

$$f''(x^*) = 0$$

$$f^3(x^*) = 0$$

$$f^4(x^*) = 0$$

$$f^n(x^*) \neq 0$$

Sufficient conditions for optimality

Proof

Apply Taylor's series

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

Since $f'(x^*) = f''(x^*) = \cdots = f^{n-1}(x^*) = 0$

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

When n is even $\frac{h^n}{n!} \geq 0$

Thus if $f^n(x^*)$ is positive $f(x^* + h) - f(x^*)$ is positive **Hence it is local minimum**

Thus if $f^n(x^*)$ negative $f(x^* + h) - f(x^*)$ is negative **Hence it is local maximum**

When n is odd, $\left(\frac{h^n}{n!}\right)$ changes sign with the change in the sign of h.

Hence it is an inflection point

Sufficient conditions for optimality

Example

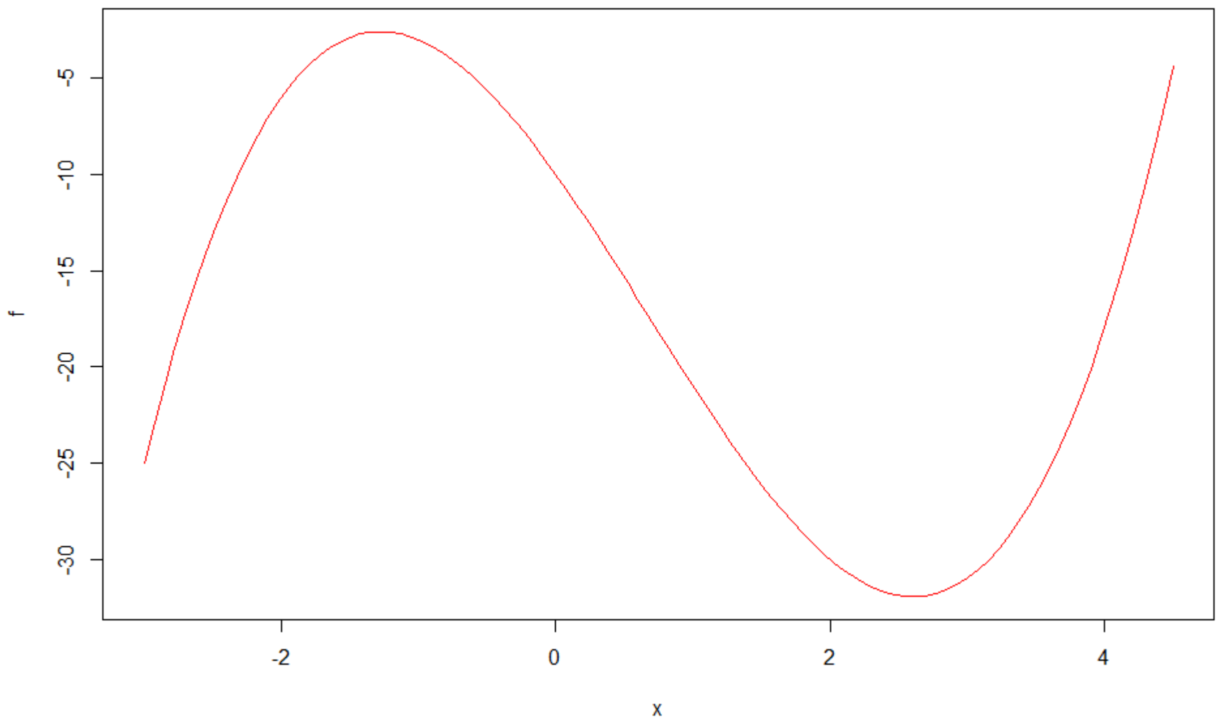
$$f(x) = x^3 - 10x - 2x^2 - 10$$

Apply necessary condition

$$f'(x) = 3x^2 - 10 - 4x = 0$$

Solving for x $\quad x^* = 2.61 \; and \; -1.28$

These two points are stationary points



Apply sufficient condition $\quad f''(x) = 6x - 4$

$f''(2.61) = 11.66$ *positive and n is even* $\quad f''(-1.28) = -11.68$ *negative and n is even*

$x^* = 2.61$ is a minimum point

$x^* = -1.28$ is a maximum point

Multivariable optimization without constraints

Minimize $f(X)$ Where $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$

Necessary condition for optimality

If $f(X)$ has an extreme point (maximum or minimum) at $X = X^*$ and if the first partial Derivatives of $f(X)$ exists at X^*, then

$$\frac{\partial f(X^*)}{\partial x_1} = \frac{\partial f(X^*)}{\partial x_2} = \cdots = \frac{\partial f(X^*)}{\partial x_n} = 0$$

Multivariable optimization without constraints

Sufficient condition for optimality

The sufficient condition for a stationary point X^* to be an extreme point is that the matrix of second partial derivatives of $f(X)$ evaluated at X^* is

(1) positive definite when X^* is a relative minimum
(2) negative definite when X^* is a relative maximum
(3) neither positive nor negative definite when X^* is neither a minimum nor a maximum

Proof Taylor series of two variable function

$$f(x+\Delta x, y+\Delta y) = f(x,y) + \Delta x \frac{\partial f}{\partial x} + \Delta y \frac{\partial f}{\partial y} + \frac{1}{2!}\left(\Delta x^2 \frac{\partial^2 f}{\partial x^2} + 2\Delta x \Delta y \frac{\partial^2 f}{\partial x \partial y} + \Delta y^2 \frac{\partial^2 f}{\partial y^2}\right) + \cdots$$

$$f(x+\Delta x, y+\Delta y) = f(x,y) + [\Delta x \quad \Delta y]\begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} + \frac{1}{2!}[\Delta x \quad \Delta y]\begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix}\begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} + \cdots$$

Multivariable optimization without constraints

$$f(X^* + h) = f(X^*) + h^T \nabla f(X^*) + \frac{1}{2!} h^T \boldsymbol{H} h + \cdots$$

Since X^* is a stationary point, the necessary condition gives that $\nabla f(X^*) = 0$

Thus

$$f(X^* + h) - f(X^*) = \frac{1}{2!} h^T \boldsymbol{H} h + \cdots$$

Now, X^* will be a minima, if $h^T \boldsymbol{H} h$ is positive

X^* will be a maxima, if $h^T \boldsymbol{H} h$ is negative

$h^T \boldsymbol{H} h$ will be positive if $\boldsymbol{H}$ is a positive definite matrix

$h^T H h$ will be negative if $\boldsymbol{H}$ is a negative definite matrix

A matrix H will be positive definite if all the eigenvalues are positive, *i.e.* all the λ values are positive which satisfies the following equation

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

Another test: Evaluation of determinants

$$A_1 = |a_{11}|$$

$$A_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$

$$A_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

$$A_n = \begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & a_{24} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & a_{n4} & \cdots & a_{nn} \end{vmatrix}$$

- ✓ A matrix $\boldsymbol{A}$ will be positive definite if any only if all the values $A_1, A_2, A_3, \ldots, A_n$ are positive.
- ✓ The matrix will be negative definite is and only if the sign of A_j is $(-1)^j$ for $j = 1, 2, 3, \ldots, n$

Multivariable optimization without constraints

Example

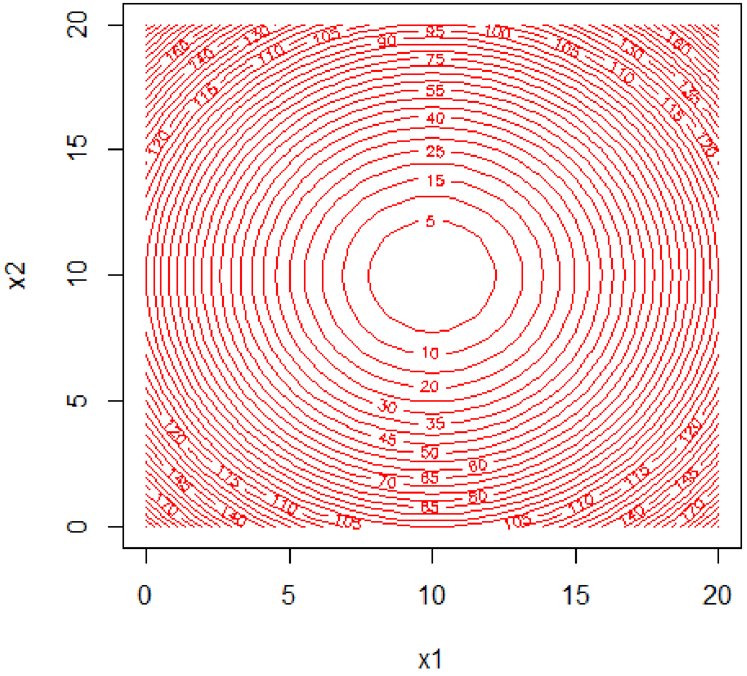
$$f(x_1, x_2) = (x_1 - 10)^2 + (x_2 - 10)^2$$

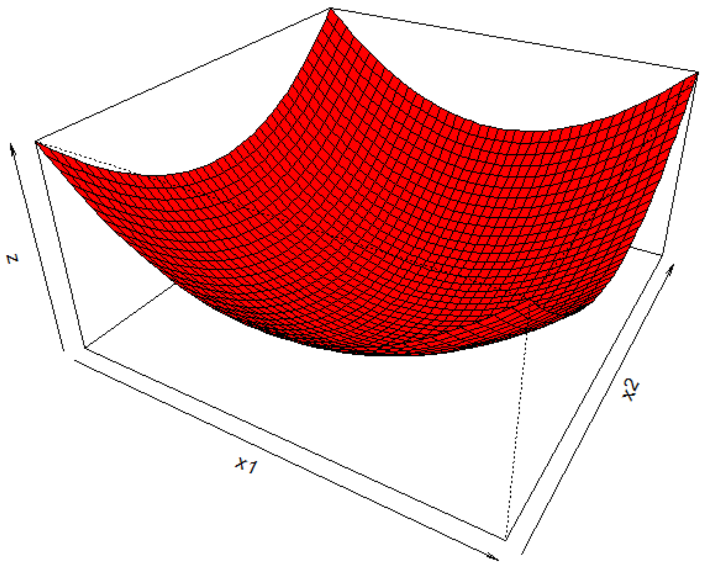
Necessary condition

$$\frac{\partial f}{\partial x_1} = 2(x_1 - 10) = 0 \Rightarrow x_1 = 10$$

$$\frac{\partial f}{\partial x_2} = 2(x_2 - 10) = 0 \Rightarrow x_2 = 10$$

$$X^* = \begin{pmatrix} 10 \\ 10 \end{pmatrix}$$





Sufficient condition

$$H = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

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

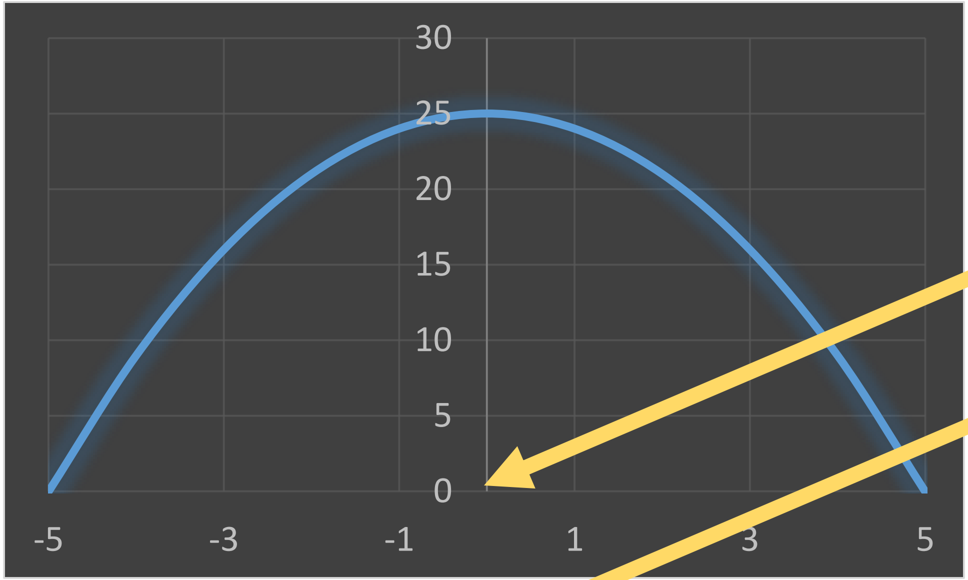
Thus the eigenvalues of the matrix H are 2, and 2.
The H is a positive definite matrix and the X^* is a relative minimum

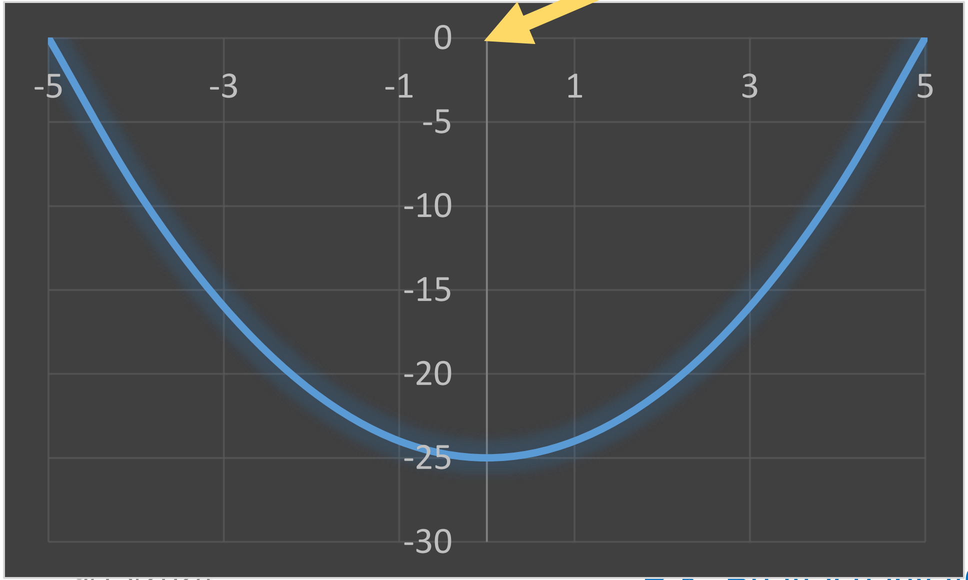
$$A_1 = |a_{11}| = 2$$

$$A_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \begin{vmatrix} 2 & 0 \\ 0 & 2 \end{vmatrix} = 4$$

Both A_1 and A_2 are positive
The H is a positive definite matrix
Hence X^* is a relative minimum

Unimodal and duality principle





Optimal solution $x^* = 0$

Optimal solution $x^* = 0$

Minimization $f(x)$ = Maximization $-f(x)$

Thank you

$$0 \le \lambda \le 1$$