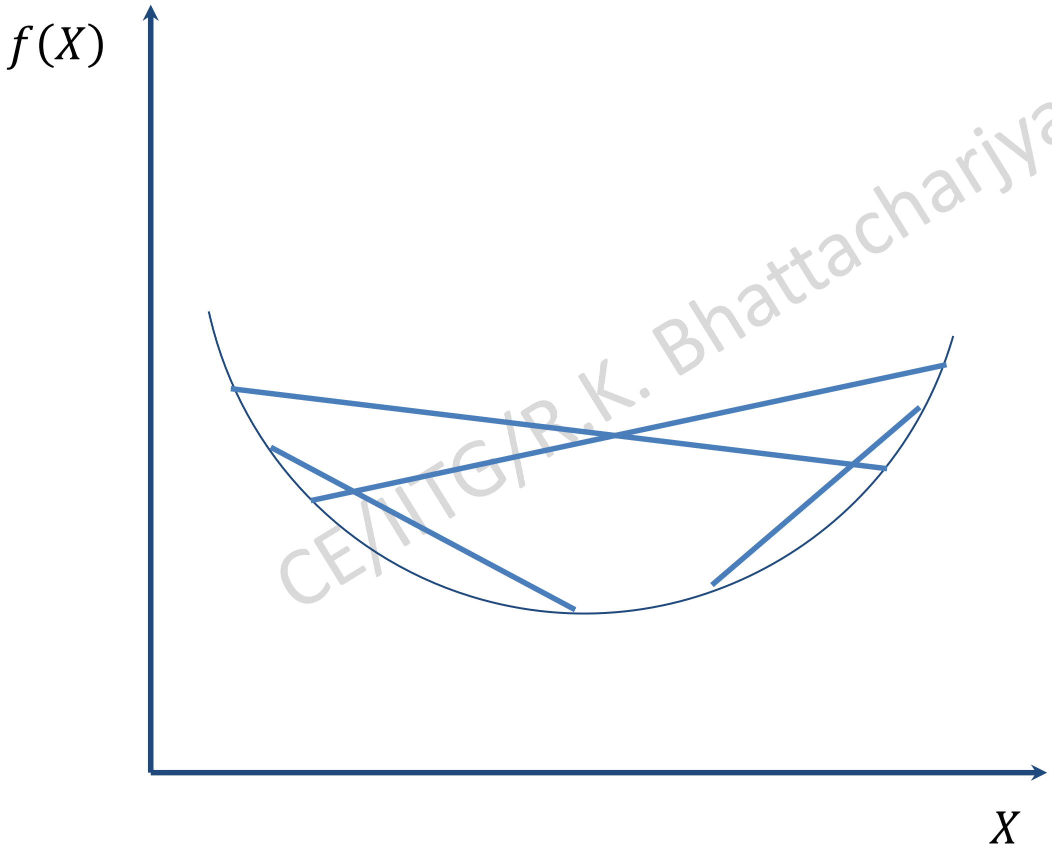
Convex Function

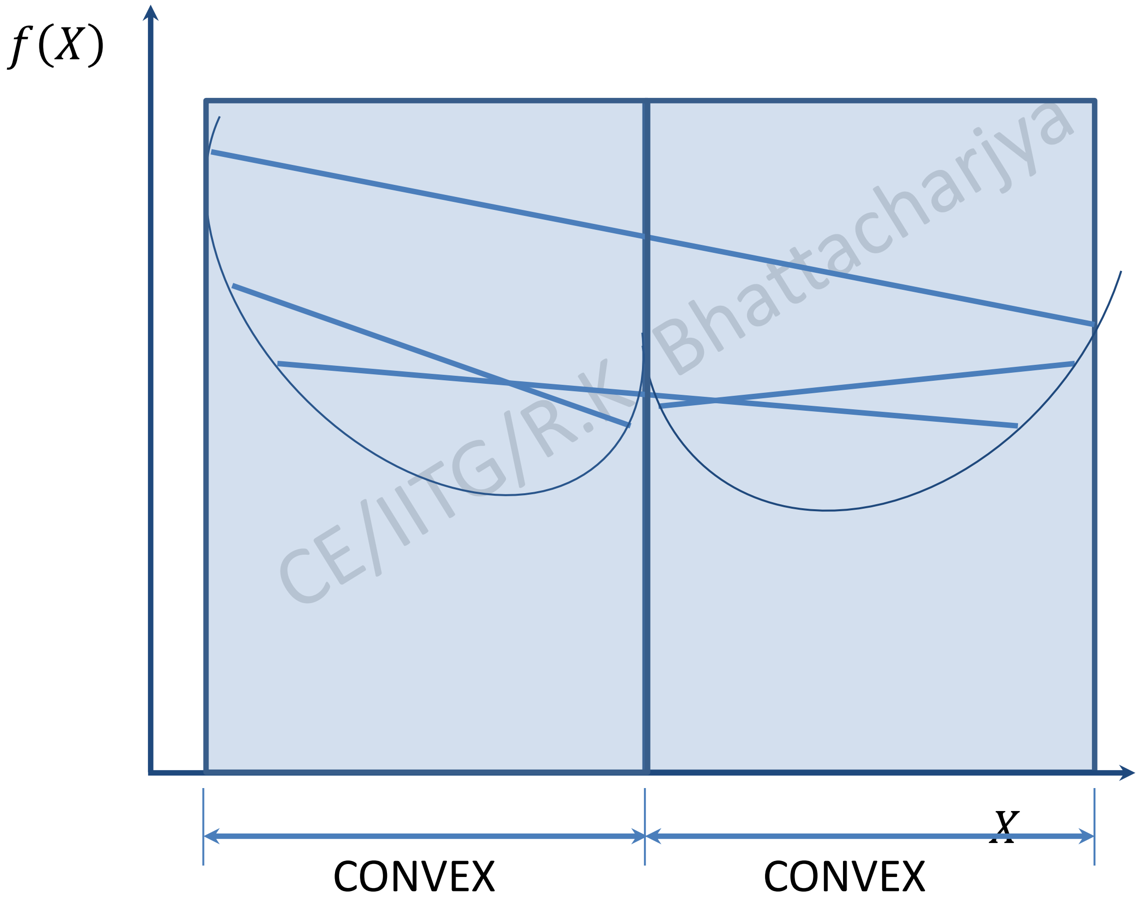
Prof. (Dr.) Rajib Kumar Bhattacharjya

Professor, Department of Civil Engineering
Indian Institute of Technology Guwahati, India
Room No. 005, M Block
Email: rkbc@iitg.ernet.in, Ph. No 2428

CONVEX FUNCTION



CONVEX FUNCTION

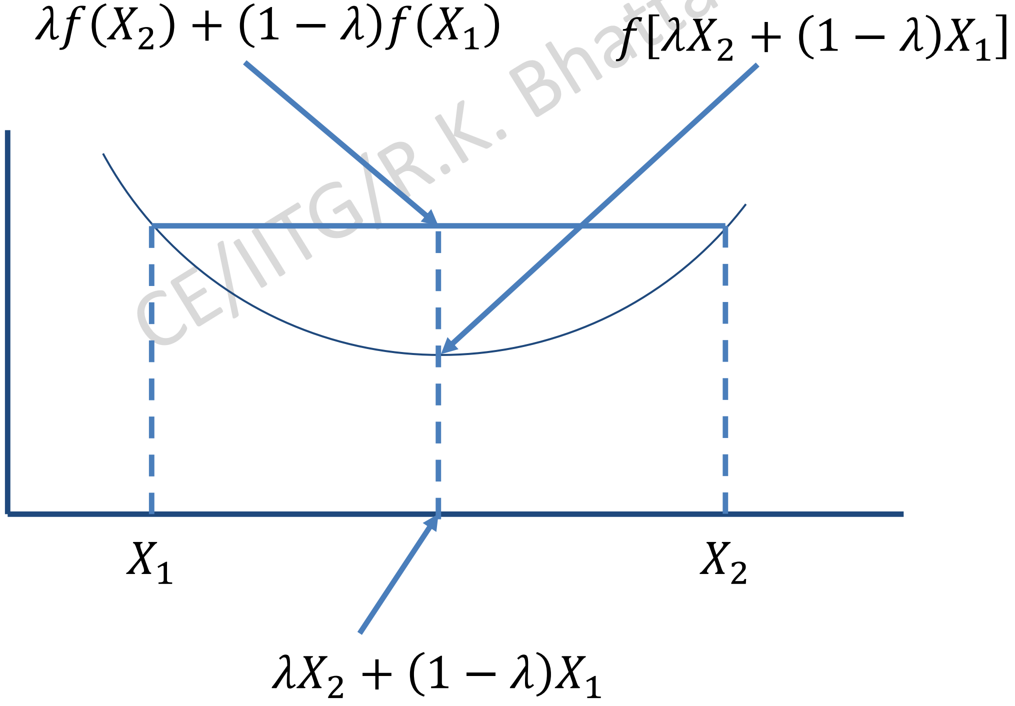


CONVEX FUNCTION

A function $f(X)$ is said to be convex if for any pair of points $X_1 = [x_1^1, x_2^1, x_3^1, \ldots, x_n^1]^T$ and $X_2 = [x_1^2, x_2^2, x_3^2, \ldots, x_n^2]^T$ and all λ where $0 \leq \lambda \leq 1$

$$f[\lambda X_2 + (1-\lambda)X_1] \leq \lambda f(X_2) + (1-\lambda) f(X_1)$$

That is, if the segment joining the two points lies entirely above or on the graph of $f(X)$

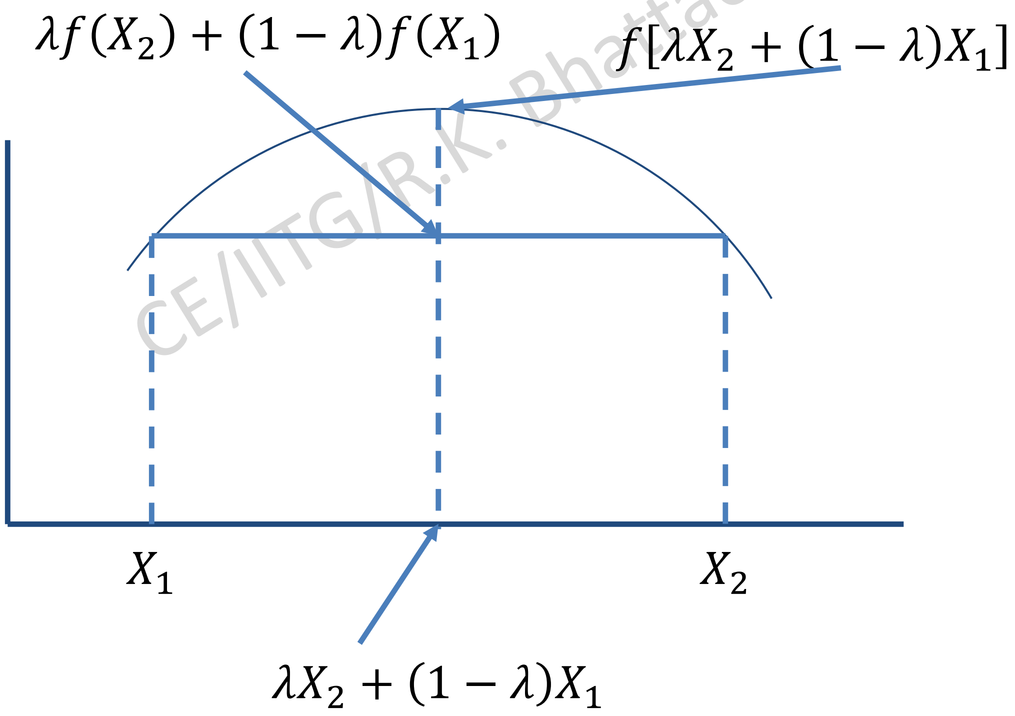


CONCAVE FUNCTION

A function $f(X)$ is said to be convex if for any pair of points $X_1 = [x_1^1, x_2^1, x_3^1, \dots, x_n^1]^T$ and $X_2 = [x_1^2, x_2^2, x_3^2, \dots, x_n^2]^T$ and all λ where $0 \le \lambda \le 1$

$$f[\lambda X_2 + (1-\lambda)X_1] \ge \lambda f(X_2) + (1-\lambda)f(X_1)$$

That is, if the segment joining the two points lies entirely above or on the graph of $f(X)$



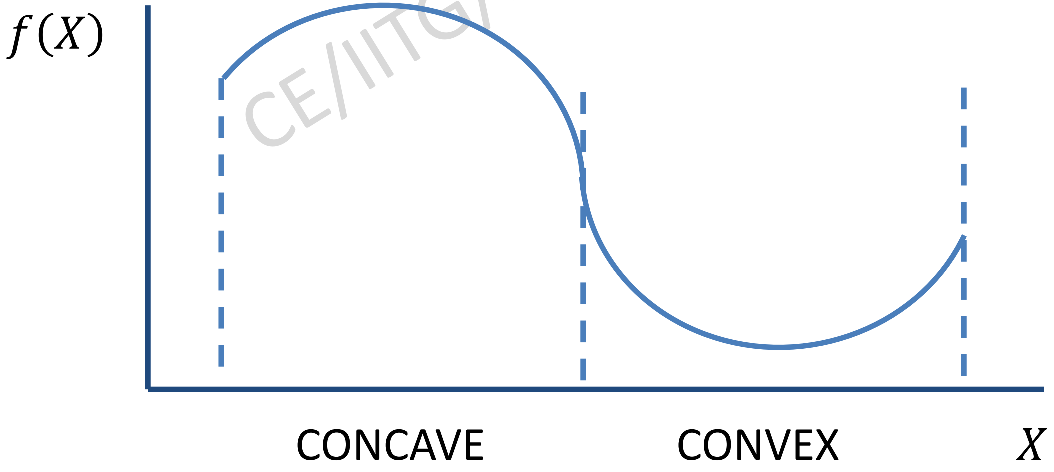
A function $f(X)$ will be called strictly convex if

$$f[\lambda X_2 + (1-\lambda)X_1] < \lambda f(X_2) + (1-\lambda)f(X_1)$$

A function $f(X)$ will be called strictly concave if

$$f[\lambda X_2 + (1-\lambda)] > \lambda f(X_2) + (1-\lambda)f(X_1)$$

Further a function may be convex within a region and concave elsewhere



Theorem 1: A function $f(X)$ is convex if for any two points X_1 and X_2, we have

$$f(X_2) \geq f(X_1) + \nabla f^T(X_1)(X_2 - X_1)$$

Proof: If $f(X)$ is convex, we have

$$f[\lambda X_2 + (1-\lambda)X_1] \leq \lambda f(X_2) + (1-\lambda)f(X_1)$$

$$f[X_1 + \lambda(X_2 - X_1)] \leq f(X_1) + \lambda[f(X_2) - f(X_1)]$$

$$\lambda[f(X_2) - f(X_1)] \geq f[X_1 + \lambda(X_2 - X_1)] - f(X_1)$$

$$[f(X_2) - f(X_1)] \geq \frac{f[X_1 + \lambda(X_2 - X_1)] - f(X_1)}{\lambda(X_2 - X_1)}(X_2 - X_1)$$

By defining $\Delta X = \lambda(X_2 - X_1)$

$$[f(X_2) - f(X_1)] \geq \frac{f[X_1 + \lambda(X_2 - X_1)] - f(X_1)}{\Delta X}(X_2 - X_1)$$

By taking limit as $\Delta X \to 0$

$$[f(X_2) - f(X_1)] \geq \nabla f^T(X_1)(X_2 - X_1)$$

$$f(X_2) \geq f(X_1) + \nabla f^T(X_1)(X_2 - X_1)$$

Theorem 2: A function $f(X)$ is convex if Hessian matrix H(X) is positive semi definite.

Proof: From the Taylor's series

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

Let X^*=X_1, $X^* + h = X_2$ and $h = (X_2 - X_1)$

We have

$$f(X_2) = f(X_1) + \nabla f^T(X_1)(X_2 - X_1) + \frac{1}{2!}(X_2 - X_1)H(X_2 - X_1)^T$$

$$f(X_2) - f(X_1) = \nabla f^T(X_1)(X_2 - X_1) + \frac{1}{2!}(X_2 - X_1)H(X_2 - X_1)^T$$

Now $\quad f(X_2) - f(X_1) \geq \nabla f^T(X_1)(X_2 - X_1)$

if $\quad (X_2 - X_1)H(X_2 - X_1)^T \geq 0$

That is H should be positive semi definite

Theorem 3: A local minimum of a convex function $f(X)$ is a global minimum

Proof: Suppose there exist two different local minima, say X_1 and X_2, for the function $f(X)$.

Let $\quad f(X_2) < f(X_1)$

Since $f(X)$ is convex between X_1 and X_2 we have

$$f(X_2) \geq f(X_1) + \nabla f^T(X_1)(X_2 - X_1)$$

$$f(X_2) - f(X_1) \geq \nabla f^T(X_1)(X_2 - X_1)$$

Or $\quad f^T(X_1)(X_2 - X_1) \leq 0$

Or $\quad f^T(X_1)S \leq 0 \quad$ Where $S = (X_2 - X_1)$

This is the condition of descent direction

As such X_1 is not an optimal points and function value will reduce if you go along the direction S

Convex optimization problem

Standard form

Minimize $f(X)$

Subject to $g(X) \leq 0$

$h(X) = 0$

The problem will be convex, if

$g(X)$ is a convex function

$h(X)$ is a affine function

$$h(X) = AX + B$$

$h(X) = 0$ can be written as

$h(X) \leq 0 \quad and \quad -h(X) \leq 0$

If $h(X) \leq 0$ is convex, then $-h(X) \leq 0$ is concave

Hence only way that $h(X) = 0$ will be convex is that $h(X)$ to be affine