Lecture - 1

Convex Function and related Results

Ref :- Numerical Optimization with Applications. by S. Chandra, Jayadeva, and A. Mehra.

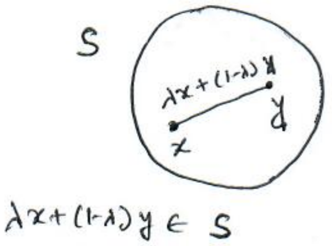
In earlier classes all the problems are constructed in presence of linearity structure. This problem gave us beautiful mathematical results as well as helped greatly in its algorithmic developement. However, most of the real life problems/applications lead to optimization problems which are non-linear in nature. Fortunately, most often this non-linearity is of 'parrabola' type, leading to the convexity structure which can also be exploited to study such non-linear optimization problems.

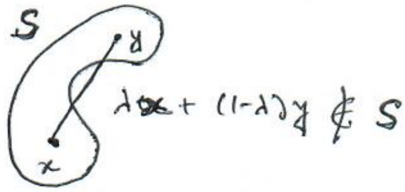
Our basic aim in this course is to understand the non-linear optimization problems such that convex optimization problems, i.e. those optimization problems which have the structure of 'convexity'. This problems are best understood in terms of the convexity/concavity of the objective and constraint functions.

We also study "quadratic" programming problems in the later part.

Def^n :- (Convex Set)

A set S is convex if for all x and y in S, the line segment connecting x and y is included in S.



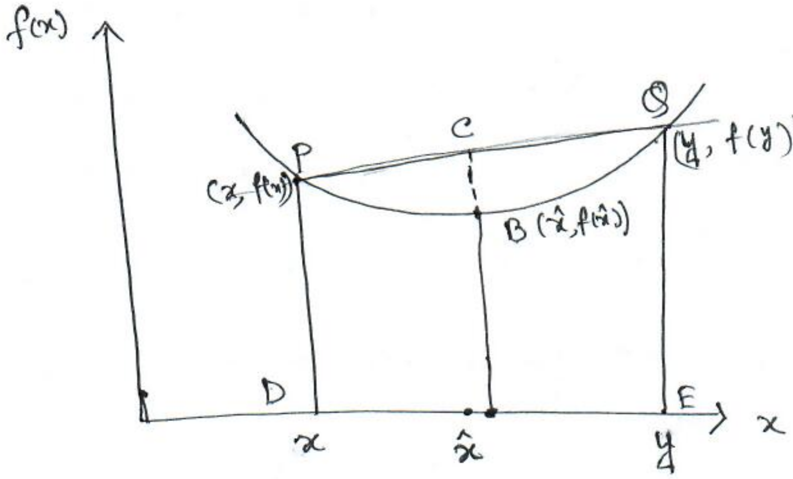


Def^n :- (Convex function)

Let $S \subseteq \mathbb{R}^n$ be a convex set and $f : S \to \mathbb{R}$.
Then f is called a convex function if

$\forall x, y \in S$ and $\forall\ 0 \le \lambda \le 1$, we have

$$f(\lambda x + (1-\lambda)y) \le \lambda f(x) + (1-\lambda) f(y).$$



Note :- A function f is convex if for any two points P & Q on the curve (graph of the function f), the line segment joining P and Q is always on or above the curve between P and Q but never below the curve.

Examples of convex function :- $f : \mathbb{R} \to \mathbb{R}$.

(i) $f(x) = x^2, \ x \in \mathbb{R}$

(ii) $f(x) = |x|, \ x \in \mathbb{R}$

(iii) $f(x) = e^{x}, \ x \in \mathbb{R}$

(iv) $f(x) = e^{-x}, \ x \in \mathbb{R}$

Exercise 1 :- Check the above functions convexity.

Remark 1 :- Let S be a convex set and $f : S \to \mathbb{R}$. Then f is called strictly convex function if $\forall x, y \in S, \forall\, 0 \le \lambda \le 1$,

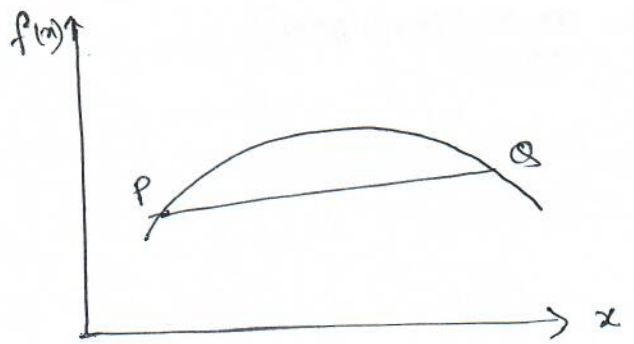
$$f(\lambda x + (1-\lambda) y) < \lambda f(x) + (1-\lambda) f(y).$$

Note :- Every strictly convex function is convex but the converse is NOT true.

Def^n :- (Concave Function)

Let $S \subseteq \mathbb{R}^n$ be a convex set and $f : S \to \mathbb{R}$. Then f is called a concave function if $\forall x, y \in S$ and $\forall\, 0 \le \lambda \le 1$, we have

$$f(\lambda x + (1-\lambda) y) \geq \lambda f(x) + (1-\lambda) f(y) \qquad (*)$$



Remark 2 :- If the inequality ⊛ is strict then the function f is called a strictly concave function.

Remark 3 :- If f is concave function iff $-f$ is a convex function.

Remark 4 :- If a function is both convex and concave, then it has to be a linear function.

Exercise 2 :- Give an example of a function which is neither convex nor concave.

Properties of Convex (or concave) functions :-

(1) A convex (or concave) function need not be differentiable.

Ex :- $f(x) = |x|$, $x \in \mathbb{R}$.

(2) A convex function need not even be continuous.

Ex :- $f(x) = \begin{cases} x^2, & -1 \leq x < 1 \\ 2, & x = 1 \end{cases}$

(3) Let f and g be two convex (or concave) functions over a convex set $S \subseteq \mathbb{R}^n$ then

$f+g$, αf where $\alpha > 0$, $h(x) = \max\limits_{x \in S} (f(x), g(x))$ or $h(x) = \min\limits_{x \in S} (f(x), g(x))$ also are convex (or concave).

Result :- If $f: \mathbb{R} \longrightarrow \mathbb{R}$, then f is a convex function on $\mathbb{R}$ (or $S \subseteq \mathbb{R}$) iff $\forall\, x, y \in S$, we have

$$(x-y)\,(f'(x) - f'(y)) \geq 0,$$

i.e. for $x \geq y$, $f'(x) \geq f'(y)$. Thus f' is an increasing fn. which is the well known definition of convexity for a real valued function of real variable.

Note :- We can state the above result for strictly convex and strictly concave functions with the obvious modification that all inequalities are strict.

Theorem :- Let $S \subseteq \mathbb{R}^n$ be an open convex set and $f : S \to \mathbb{R}$ be twice differentiable. Then f is a convex function on S iff the Hessian matrix $H_f(x)$ is positive semi-definite for all $x \in S$.

Recall the Hessian matrix $H_f(x)$ is defined as the $n \times n$ matrix of 2nd order partial derivatives

i.e. $H_f(x) = \left[\dfrac{\partial^2 f}{\partial x_i \partial x_j} \right]_{n \times n}$

- Hessian is always a symmetric matrix.
- The matrix being positive semi-definite means $y^T H_f(x)\, y \geq 0 \;\; \forall\, y \in \mathbb{R}^n$.

Remark 5 :- If f is concave function on S iff $H_f(x)$ is negative semi-definite $\forall\, x \in S$

i.e. $y^T H_f(x)\, y \leq 0 \;\; \forall\, y \in \mathbb{R}^n$.

Theorem :- Let $S \subseteq \mathbb{R}^n$ be an open convex set and $f : S \to \mathbb{R}$ be twice differentiable. If $H_f(x)$ is positive definite $\forall x \in S$, then f is a strictly convex function on S.

Note :- Similarly, we can define ~~sto~~ the result for strictly concave function.

Note :- Converse of the above Theorem is NOT true.
If f is a strictly convex function then, $H_f(x)$ may not be positive definite (though it is certainly positive semi definite)
For example, $f(x) = x^4$, $x \in \mathbb{R}$. is a strictly convex f^n. on $\mathbb{R}$. but $H_f(x) = 12x^2$ is not positive definite for $x=0$.

Example 1 :- Examine the convexity / strict-convexity of the functions (i) $f(x_1, x_2) = 2x_1^2 + x_2^2 + 4x_1x_2$
(ii) $f(x_1, x_2) = 4x_1^2 + x_2^2 + 4x_1x_2$.

$\Rightarrow$ (i) The Hessian matrix $H_f(x) = \begin{bmatrix} 4 & 2 \\ 2 & 2 \end{bmatrix}$

Which is positive definite. Hence f is a strictly convex function.

☒ In general if $f : \mathbb{R}^n \to \mathbb{R}$ is a quadratic form i.e. $f(x) = x^T Q x$, where Q is real symmetric. Then we can check $\nabla f(x) = 2Qx$ and $H_f(x) = 2Q$. Therefore, the nature of the given quadratic form $x^T Q x$ will depend upon the nature of the matrix Q.

☒ The easiest way to check if a matrix Q is positive definite / -ve definite / +ve-semi-definite / -ve-semi-definite / indefinite is to obtain the eigen values of Q. As Q is real symmetric matrix. Therefore, all it's eigen values $\lambda_1, \lambda_2, \ldots, \lambda_n$ are real.

Result:- (i) Q is positive definite iff all $\lambda_i > 0$

(ii) Q is positive semi-deffinite iff all $\lambda_i \geq 0$

(iii) Q is ~~positive~~ negative definite iff all $\lambda_i < 0$

(iv) Q is negative semi-definite iff all $\lambda_i \leq 0$

(v) Q is indefinite iff some $\lambda_i > 0$ and some $\lambda_i < 0$.