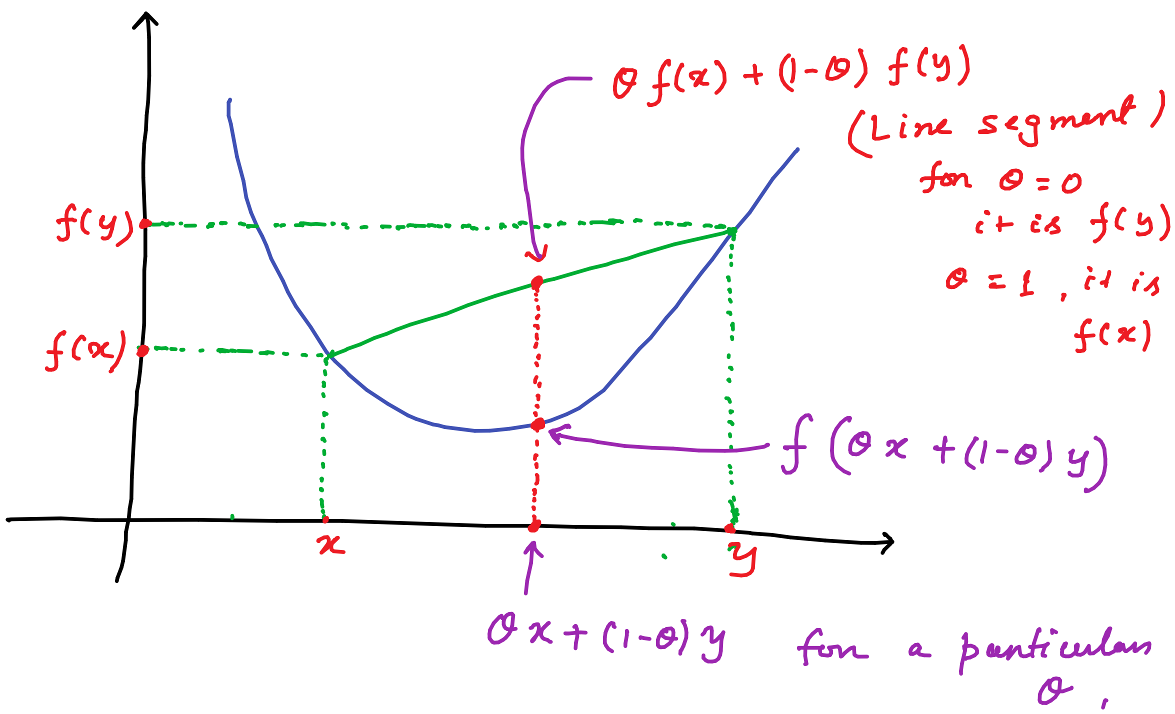
CONVEX SETS & CONVEX FUNCTIONS

Part-2

→ **Convex Function :**

Let $S \subseteq \mathbb{R}^n$ be a convex set. A real-valued function $f : S \to \mathbb{R}$ is said to be "convex" if :

$$\boxed{f(\underbrace{\theta x + (1-\theta) y}_{\text{line segment}}) \leq \theta f(x) + (1-\theta) f(y) \quad \text{for all } x, y \in S \ \& \ \theta \in [0, 1]}$$



→ A realvalued function $f : S \to \mathbb{R}$ where $S \subseteq \mathbb{R}^n$ is "concave" if $-f$ is convex.

- The convexity of domain set S is required to qualify the function f to be convex.

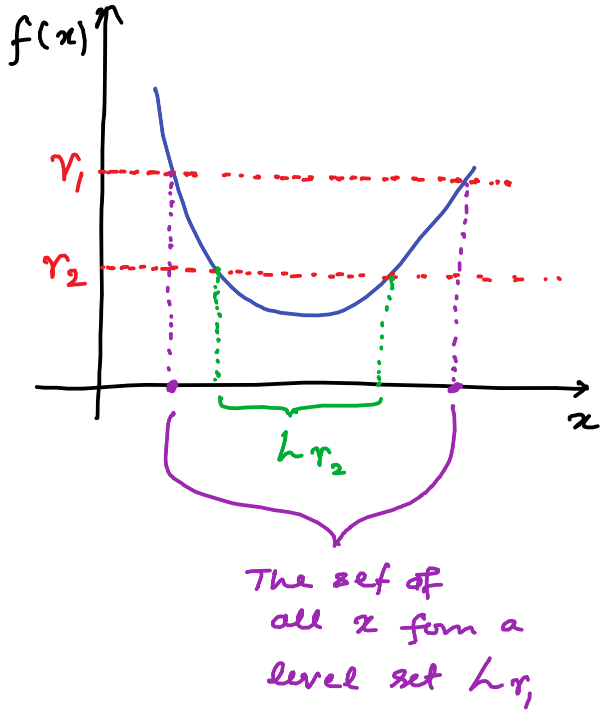
→ Level Sets :

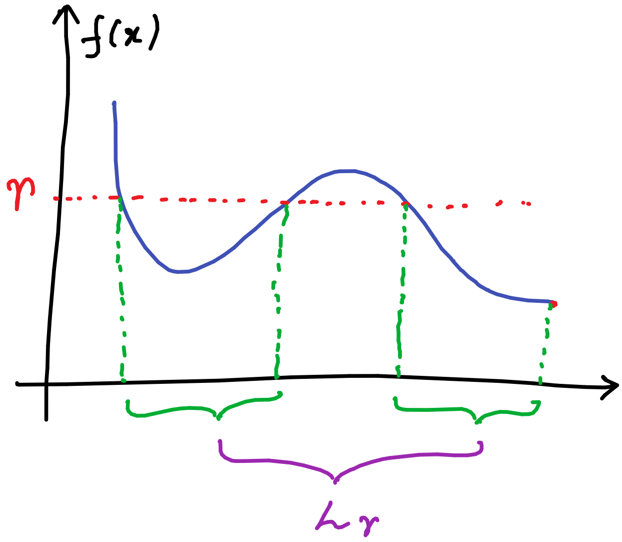
For a given function $f : X \to \mathbb{R}$ with $X \subseteq \mathbb{R}^n$, and a scalar r, the sets :

1. $\mathcal{L}_r := \{ x \mid f(x) \leq r \}$, $x \in X$
2. $\mathcal{L}_r := \{ x \mid f(x) < r \}$

are the level sets of f.

- If f is a convex function, then all of its level sets are convex for all scalars r.





- For $X \subseteq \mathbb{R}^2$, the level sets are curves in $\mathbb{R}^2$, & $X \subseteq \mathbb{R}^3$, L_r are surfaces in $\mathbb{R}^3$.

→ <u>Extended Real-Valued Convex Functions</u>

Sometimes the real-valued convex functions can take values $-\infty, \infty$ at some point x in its domain. Such functions are referred to as <u>extended real-valued convex functions</u> where we allow following arithmatic operations:

(i) $\infty + \infty = \infty$

(ii) $\infty \cdot 0 = 0$

(iii) $\alpha \cdot \infty = \infty$ for $\alpha > 0$ & $-\infty$ for $-\alpha$.

(iv) $-\infty - \infty = -\infty$

Notations for such extended real-valued functions:

$$f : X \longrightarrow [-\infty, \infty] = \mathbb{R} \cup \{\pm\infty\}$$

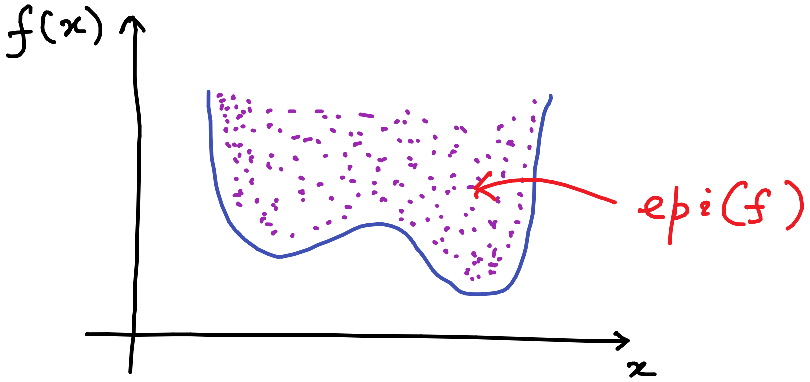
$$f : X \longrightarrow (-\infty, \infty]$$

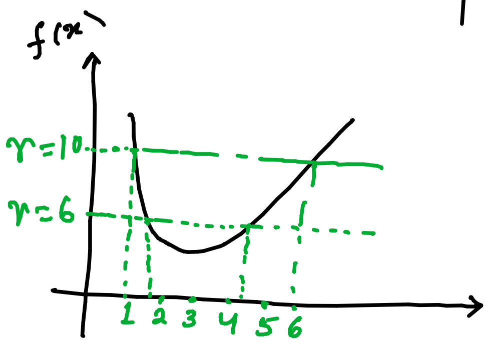
where $X \subseteq \mathbb{R}^n$.

→ Epigraph

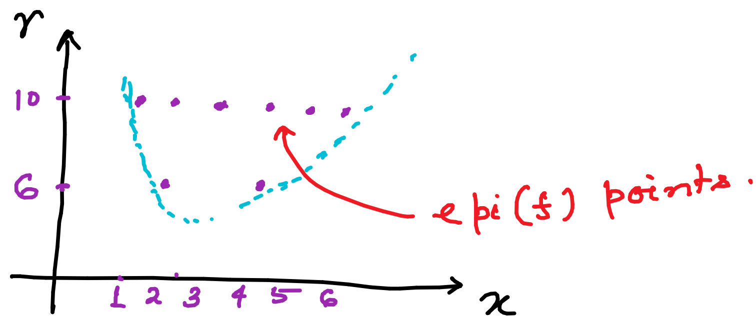
For a function $f : S \rightarrow [-\infty, \infty]$ with $S \subseteq \mathbb{R}^n$, its epigraph is a subset of $\mathbb{R}^{n+1}$, which is defined as:

$$\boxed{epi(f) := \left\{ \begin{bmatrix} x \\ \gamma \end{bmatrix} \in \mathbb{R}^{n+1} \;\middle|\; x \in S \text{ \& } \gamma \in \mathbb{R},\ \underbrace{f(x) \le \gamma}_{\text{Level set}} \right\}}$$





→



- Effective domain of $f : S \to [\pm\infty, \infty]$ is the set :

$$\text{dom}(f) := \{ x \in S \mid f(x) < \infty \}$$

- Note that the scalar r used in $\text{epi}(f)$ is a real number, & it can not take values $-\infty, \infty$. Hence, even if the function f takes value ∞ at $\tilde{x}$ i.e. $f(\tilde{x}) = \infty$, the point $\begin{bmatrix} \tilde{x} \\ \infty \end{bmatrix}$ does not belong to $\text{epi}(f)$.

- Observe that the effective domain $\text{dom}(f)$ is the projection of $\text{epi}(f)$ on $\mathbb{R}^n$.

- If f is restricted to the set $\{x \mid f(x) < \infty\}$ its epigraph $\text{epi}(f)$ will remain unaltered.

→ Proper Function

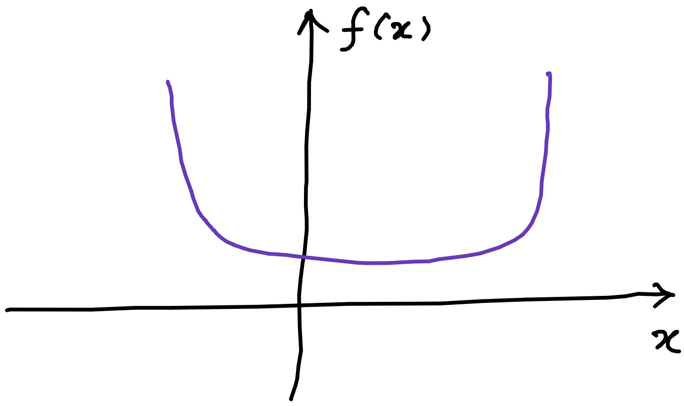
An extended real valued function

$f : X \to (-\infty, \infty]$ is said to be "proper",

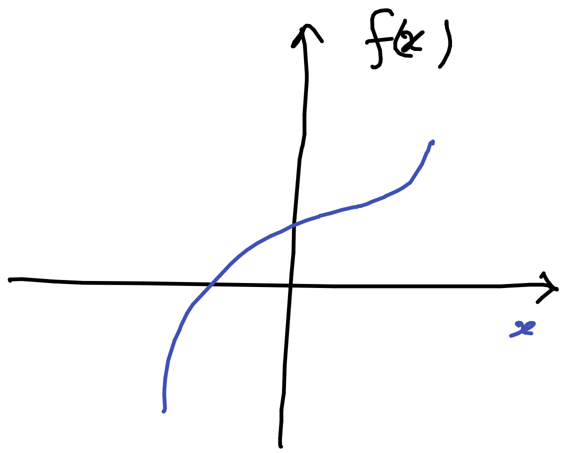
if following statements hold :

i) $f(x) < \infty$ for atleast one $x \in X$, and

ii) $f(x) > -\infty$ for all $x \in X$



Proper function



Not - proper
or
Improper
funt?.

→ Let $S \subseteq \mathbb{R}^n$ be a convex set. Then,

an extended real-valued function

$f : S \to (-\infty, \infty]$ is said to be convex

if its epigraph $(epi(f))$ is a convex

subset of $\mathbb{R}^{n+1}$.

- According to the above definition, convexity of an extended real-valued function implies :

  (i) its effective domain $\mathrm{dom}(f)$ is a convex set, &

  (ii) its level sets:

$$\left.\begin{array}{l} \{x \in S \mid f(x) \leq r\} \\ \{x \in S \mid f(x) < r\} \end{array}\right\} \text{ are convex for all } r.$$

→ Facts

1. A linear function is a convex function.

2. Any vector norm is a convex function.

3. Let f_i, for $i = 1, 2 \cdots n$ be convex.

   Then, the function

$$\boxed{h(x) = \sum_{i=1}^{n} \omega_i f_i \quad \text{with} \quad \omega_i > 0}$$

   is also a convex function.

4. Let $A \in \mathbb{R}^{n \times m}$ be a given matrix. Let $f(x)$ be a convex function. Then the function $\boxed{g(x) := f(Ax)}$ is also convex.

→ Results :

Let a function $f : \mathbb{R}^n \to (-\infty, \infty]$. Then, the following statements are equivalent :

i) The level set: $L_r = \{ x \in \mathbb{R}^n \mid f(x) \leq r \}$ is closed for every scalar r.

ii) f is lower semi-continuous over $\mathbb{R}^n$.

iii) $epi(f)$ is closed set.

→ Characterization of Differentiable Convex Functions :

Let $S_c \subset \mathbb{R}^n$ be a convex set. Let $f : \mathbb{R}^n \to \mathbb{R}$ be differentiable over S_c. Then,

i) the function f is convex over S_c if and only if

$$\boxed{f(y) \geq f(x) + (y-x)^T \nabla f(x) \quad \text{for all } x, y \in S_c.}$$

(*)

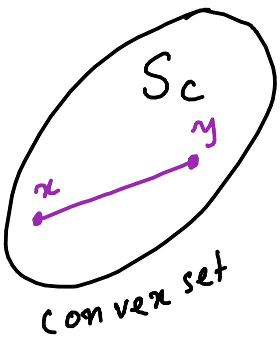
(ii) if the inequality ⊛ is strict for $x \neq y$ then f is strictly convex over S_c $\left[f(\theta x + (1-\theta) y) < \theta f(x) + (1-\theta) f(y) \text{ for all } x, y \cdot \theta \in [0,1] \right]$.

Proof

First assume that $f : \mathbb{R}^n \to \mathbb{R}$ is convex over S_c. Then, we show that the inequality ⊛ holds.

From the convex set S_c, choose any two points x, y:

i.e. $x \in S_c$ & $y \in S_c$



Then, according to the definition of convexity, the point:

$$\boxed{\begin{array}{l} \theta y + (1-\theta) x \in S_c, \quad \forall\ \theta \in [0,1] \\ \Rightarrow\ x + \theta (y - x) \in S_c \end{array}}$$

The directional derivative of f at x in the direction of $d := y - x$ is:

$$\lim_{\alpha \to 0} \frac{f(x + \alpha d) - f(x)}{\alpha} = f'(x, d)$$

Since f is differentiable at x, according to previous result (Gateaux differentiability), we have:

$$f'(x, d) = \nabla f(x)^T d \quad , \quad \forall\, d \in S_c$$

$$= d^T \nabla f(x)$$

change in the direction of d at pt x.

Further, since the function f is convex, it satisfies:

$$f(x + \theta(y - x)) = f(\theta y + (1-\theta) x)$$

$$\leq \theta f(y) + (1 - \theta) f(x), \quad \forall\, \theta \in [0,1]$$

$$= \theta f(y) + f(x) - \theta f(x)$$

$$\Rightarrow \frac{f(x + \theta(y - x)) - f(x)}{\theta} \leq f(y) - f(x)$$

$$\forall\, \theta \in [0,1]$$

$$\Rightarrow \frac{f(x+\theta d) - f(x)}{\theta} \leq f(y) - f(x) \quad \forall \theta \in [0,1]$$

(Using $d = y - x$)

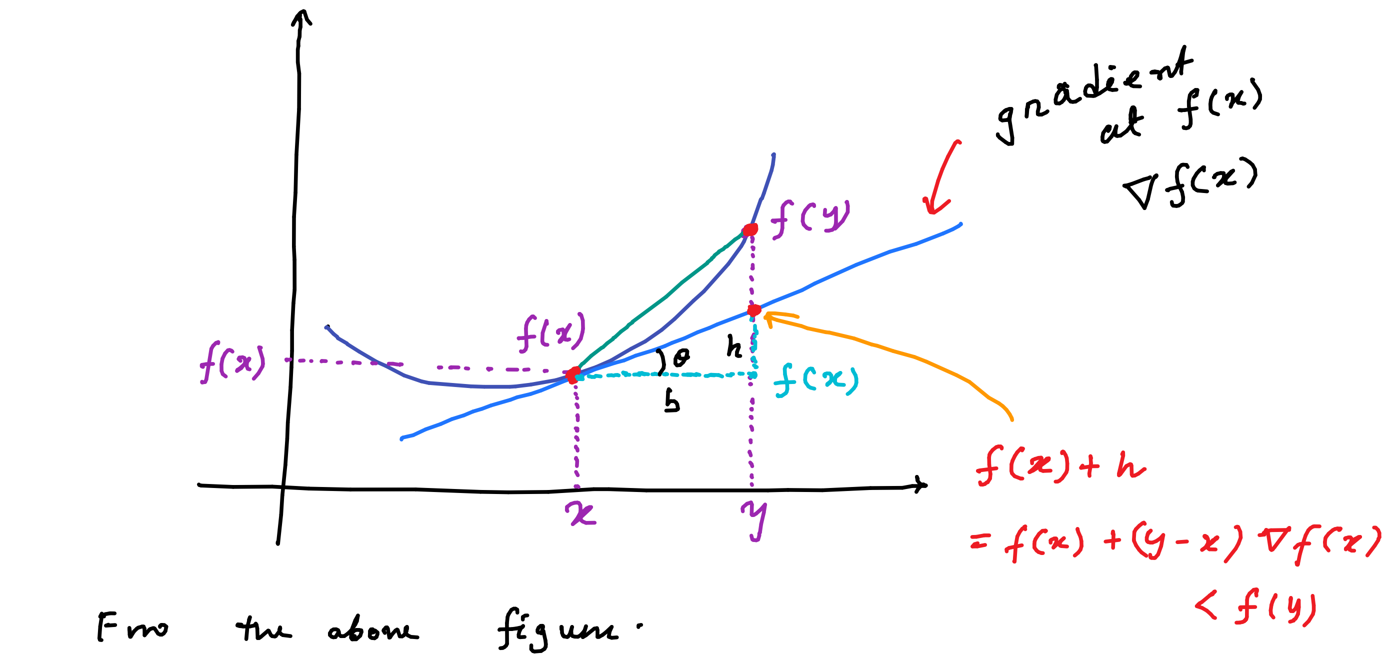
Then by taking limit $\theta \to 0$, we have:

$$f'(x, d) \leq f(y) - f(x)$$

$$\Rightarrow d^T \nabla f(x) \leq f(y) - f(x)$$

$$\boxed{\Rightarrow f(x) + (y-x)^T \nabla f(x) \leq f(y)}$$

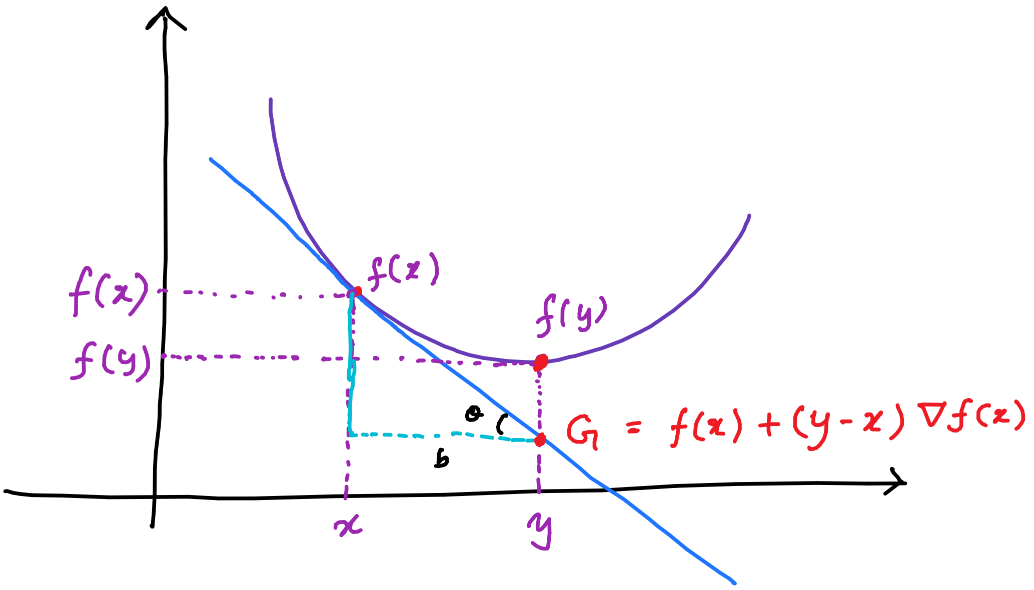
Graphical representation:



From the above figure:

$$\frac{f(x) + h - f(x)}{y - x} = \nabla f(x) \leftarrow \text{slope/gradient}$$

$$\Rightarrow f(x) + h = f(x) + (y-x)\nabla f(x)$$



From the figure:

$$\frac{G - f(x)}{y - x} = \nabla f(x)$$

$$\Rightarrow \quad G = f(x) + (y-x)\nabla f(x) < f(y)$$

- - - - - - - - - - - -

We now prove the reverse part of the 1^{st} statement.

Assume that following inequality holds:

$$f(y) \geq f(x) + (y-x)^T \nabla f(x) \quad \cdots\cdots (*)$$

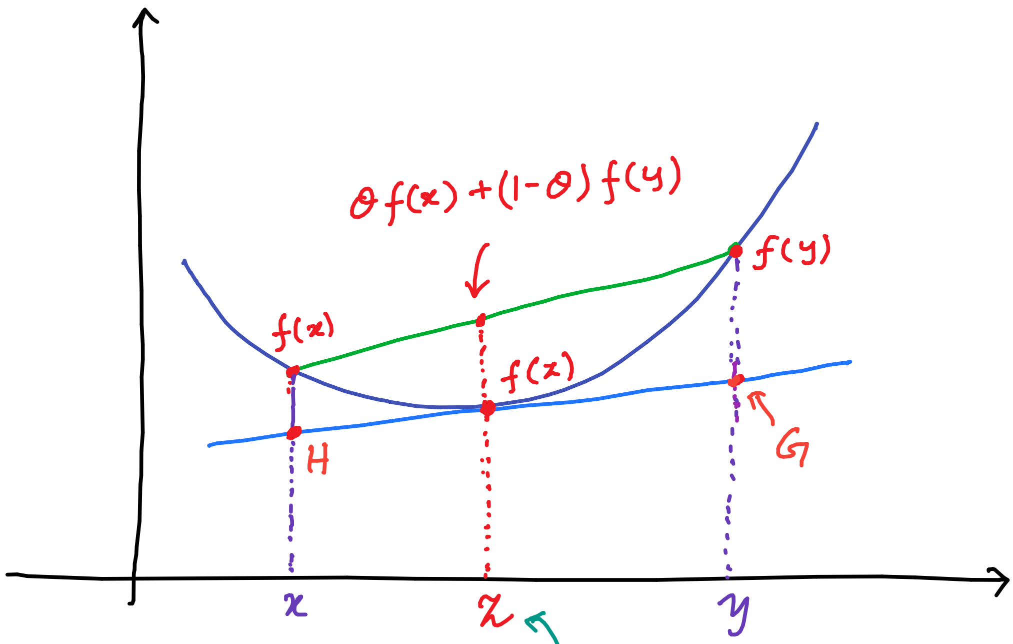
Then we show that f is convex over S_c.

Let $x \in S_c$ & $y \in S_c$ be two points.

For some $\theta \in [0,1]$, define a point:

$$z := \theta x + (1-\theta) y \text{, which belongs to } S_c.$$

Now applying the inequality ⊛ twice at points x & y w.r.t. z (as shown in fig below):



$\theta x + (1-\theta) y$

$$G = f(z) + (y-z) \nabla f(z)$$

$$H = f(z) + (x-z) \nabla f(z)$$

we have:

$$f(x) \geq f(z) + (x-z)^T \nabla f(z) \quad \cdots\cdots \quad (1)$$

$$f(y) \geq f(z) + (y-z)^T \nabla f(z) \quad \cdots\cdots \quad (2)$$

Multiply θ with (1) & $(1-\theta)$ with (2), & add them. Then:

$$\theta f(x) + (1-\theta) f(y) \geq \theta f(z) + (1-\theta) f(z)$$

$$+ \theta (x-z)^T \nabla f(z) + (1-\theta)(y-z)^T \nabla f(z)$$

$$= f(z) + \left[\underbrace{(\theta x + (1-\theta) y)}_{z} - z\right]^T \nabla f(z)$$

$$= f(z) = f(\theta x + (1-\theta) y)$$

Hence, by defination, f is convex.

- Statement - 2 proof is similar. ▨

→ Results for twice differentiable Convex functions:

Let S_c be a convex subset of $\mathbb{R}^n$. Let $f : \mathbb{R}^n \to \mathbb{R}$ be twice continuously differentiable over $\mathbb{R}^n$. Then following statements hold.

1. If the Hessian matrix $\nabla^2 f(x)$ is positive semidefinite for all $x \in S_c$, then f is convex over S_c.

$$\boxed{\nabla^2 f(x) \geq 0, \ \forall x \in S_c \Rightarrow f \text{ is convex}}$$

2. If $\nabla^2 f(x)$ is positive-definite for all $x \in S_c$, then f is strictly convex over S_c.

$$\boxed{\nabla^2 f(x) \succ 0 \ , \ \forall x \in S_c \Rightarrow f \text{ is strictly convex}}$$

3. If S_c is open and f is convex, then $\nabla^2 f(x)$ is positive semi-definite for all $x \in S_c$.

$$\boxed{\begin{array}{c} S_c \text{ is open } \& \ f \text{ is convex} \\ \Downarrow \\ \nabla^2 f(x) \succeq 0 \ , \ \forall x \in S_c \end{array}}$$

4. The quadratic function $f(x) = x^T Q x$, when Q is symmetric, is convex if and only if $Q \succeq 0$ (P.S.D.). Further f is strictly convex if and only if $Q \succ 0$ (P.D.).

Proof : Assignment .