Convex function:

$$f[\lambda x_1 + (1-\lambda) x_2] \leq \lambda f(x_1) + (1-\lambda) f(x_2)$$

Strictly convex: $<$

Concave: $\geq$

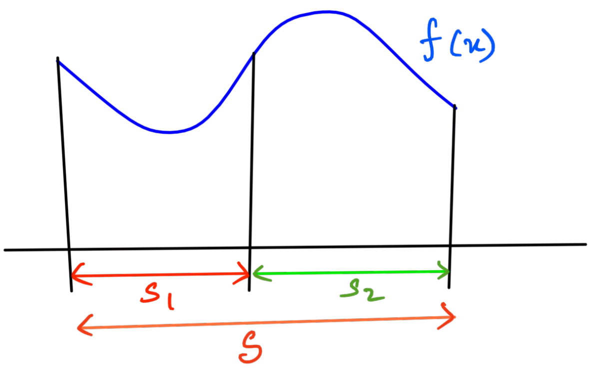
Strictly concave: $>$

Here, $\lambda \in (0,1)$ & $x_1, x_2 \in S$ (convex set)

If f is convex/(strictly convex), then
$-f$ is concave/(strictly concave)

Affine function $(Ax+b)$ is both Convex & Concave

Always associate function nature with set



$f(x)$ is —
* convex on S_1
* concave on S_2
* Neither concave/convex on S

If $f_1, \dots, f_n$ are convex

* Conic combination of function is convex

i.e. $$f(x) = \sum_{i=1}^{n} \alpha_i f_i(x) \quad ; \quad \alpha_i > 0$$

* $\max\{f_1, \dots, f_n$ is convex$\}$

- If $f: \mathbb{R}^n \to \mathbb{R}$ is convex

  Lower Level set, $S_\alpha = \{x \in S : f(x) \leq \alpha\}$ is a convex Set

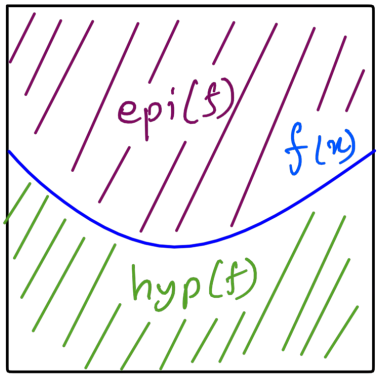
  The opposite is not true. e.g. $f(x) = \sqrt{|x|}$

  (Quasi convex function)

- Upper Level set, $S_\alpha = \{x \in S : f(x) \geq \alpha\}$

- Epigraph: $epi(f) = \{(x, y) : x \in S, y \in \mathbb{R}, y \geq f(x)\}$

  Hypograph: $hyp(f) = \{(x, y) : x \in S, y \in \mathbb{R}, y \leq f(x)\}$



- f is convex $\longrightarrow$ f is continuous? Yes

  $\longrightarrow$ f is differentiable? NO

  e.g. $f(x) = |x|$ @ $x = 0$

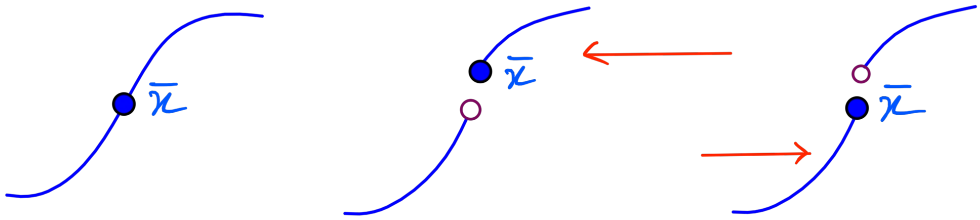
- f is continuous at $\bar{x}$ $\Rightarrow$

  $\forall \varepsilon > 0 \quad \exists \delta > 0 \qquad |f(x) - f(\bar{x})| < \varepsilon \qquad \forall x \in N_\delta(\bar{x})$

  $|x| < a \Rightarrow \begin{matrix} x < a \\ -x < a \end{matrix}$

$f(x) - f(\bar{x}) < \varepsilon$	$-f(x) + f(\bar{x}) < \varepsilon$
$\Rightarrow f(x) - \varepsilon < f(\bar{x})$	$\Rightarrow f(\bar{x}) < f(x) + \varepsilon$
Upper semi continuous	Lower semi continuous

Continuous at $\bar{x}$ | Upper Semi-Continuous at $\bar{x}$ | Lower Semi-Continuous at $\bar{x}$



The limit point ● $\Rightarrow$ Higher than the other limit point ○

Lower than the other limit point ○

$f(x) - \varepsilon < f(\bar{x})$

$\lim_{\varepsilon \to 0} f(x) = f(\bar{x})$

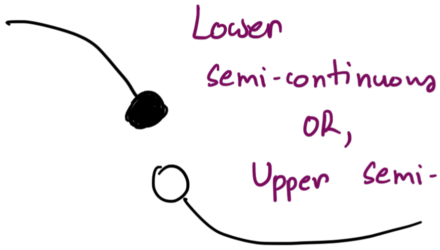
$f(\bar{x}) < f(x) + \varepsilon$

$f(\bar{x}) = \lim_{\varepsilon \to 0} f(x)$

$\Downarrow$

$\min_{x \in S} f(x)$ can be achieved

$\because f(\bar{x}) \leq f(x)$



Lower semi-continuous

OR,

Upper semi-continuous?

Answer: Upper semi-continuous (Dr. Cheng)

If f is convex on $S \Rightarrow$

f is continuous on $\text{Int}(S)$

Differentiability of f at $\bar{x} \in Int(S)$

f is differentiable at $\bar{x} \in Int(S)$

⊛ If $\exists\, \nabla f(\bar{x})$ such that

$$\lim_{x \to \bar{x}} \frac{f(x) - f(\bar{x}) - \nabla f(\bar{x})^T \cdot (x - \bar{x})}{\| x - \bar{x} \|} = 0 \quad \text{——} (1)$$

Here, $\| x - \bar{x} \|$ is a distance

$(x - \bar{x})$ is a vector

$\nabla f(\bar{x})$ is gradient vector at $\bar{x}$.

$$\nabla f(\bar{x}) = \left[\frac{\partial f(\bar{x})}{\partial x_1}, \frac{\partial f(\bar{x})}{\partial x_2}, \ldots\ldots, \frac{\partial f(\bar{x})}{\partial x_n} \right]^T$$

⊛ Let, (1) $\Rightarrow$

$$\frac{f(x) - f(\bar{x}) - \nabla f(\bar{x})^T \cdot (x - \bar{x})}{\| x - \bar{x} \|} = \alpha(\bar{x}, x - \bar{x})$$

$$\Rightarrow f(x) = f(\bar{x}) + \nabla f(\bar{x})^T (x - \bar{x}) + \| x - \bar{x} \| \cdot \alpha(\bar{x}, x - \bar{x}) \quad \text{——} (11)$$

$\forall x \in S$

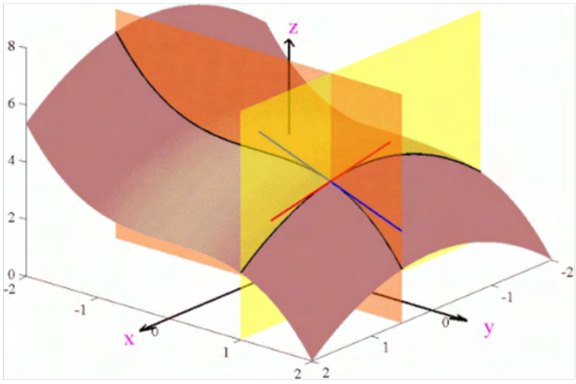
where, $\lim_{x \to 0} \alpha(\bar{x}, x - \bar{x}) = 0$

$\rightarrow$ First-order Taylor Series expansion of f at $\bar{x}$

f is differentiable on Open set $S' \subseteq S$ if

f is differentiable at each point in S'.

Partial derivative VS Directional derivative



Note, $f(x,y)$ is a surface not a curve

⊛ $\frac{\partial f}{\partial x} / \nabla_x f$ gives slope of the curve generated by the plane slicing $f(x,y)$ surface parallel to x-axis or vector $\vec{x}$. [Same explanation for $\frac{\partial f}{\partial y}$]

⊛ Directional derivative along vector $\vec{d}$, $\nabla_d f$ gives slope of the curve (Figure-1) generated by the plane slicing $f(x,y)$ surface parallel to vector $\vec{d}$; not necessarily a vector along x-axis or y-axis.

Directional derivative

⊛ Generally, $f'(\bar{x}, d) = \lim_{\lambda \to 0} \frac{f(\bar{x}+\lambda d) - f(\bar{x})}{\lambda}$

[Similar to Slope]

$\bar{x} + \lambda d$ = New point in direction d at a step λ from $\bar{x}$

$\lambda > 0$ but small

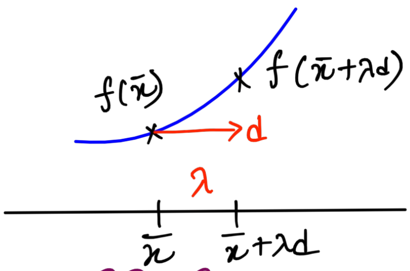


Figure-1

(*) If f is differentiable $\Rightarrow$

Directional derivative of vector $d(\neq 0)$ at $\bar{x}$,

$$f'(\bar{x}, d) = \lim_{\lambda \to 0} \frac{f(\bar{x}+\lambda d) - f(\bar{x})}{\lambda} = \nabla f(\bar{x})^T \cdot d \quad \text{—— (iii)}$$

(iii) can be derived from Taylor series expansion of (ii)

$$f(\bar{x}+\lambda d) = f(\bar{x}) + \nabla f(\bar{x})^T \cdot \lambda d + \lambda \|d\| \cdot \alpha(\bar{x}, \lambda d)$$

where, $\lim_{\lambda \to 0^+} \alpha(\bar{x}, \lambda d) = 0$

$[\because \lambda > 0]$

Here, in the expression $x - \bar{x} \equiv \lambda d$

- Recall: the slope a for the line $f = ax + b$ denotes the change of f in the direction of x-axis
- Directional derivatives characterize the local behavior of a function: it measures how f changes in the direction of d from $\bar{\mathbf{x}}$.

If f is CONVEX, directional derivative $f'(\bar{x}, d)$ exists at $\bar{x}$ along $d(\neq 0)$.

If f is CONVEX on CONVEX set S:

ξ is subgradient of f at $\bar{x} \in S$ if

$$f(x) \geq f(\bar{x}) + \xi^T (x - \bar{x}), \quad \forall x \in S$$

$\hookrightarrow$ $f(x)$ is lower bounded by linear function

$\leq$ If f is CONCAVE on CONVEX set S

$\hookrightarrow$ $f(x)$ is upper bounded by linear function

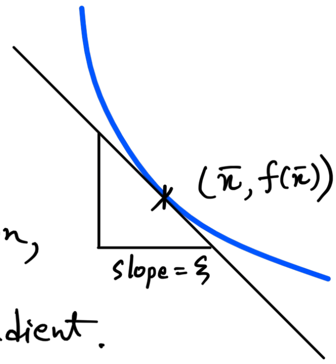
Analogous to: $f(x) = f(\bar{x}) + \nabla f(x)^T (x - \bar{x}) + \ldots$

Note: $f(x) \geq \underbrace{f(\bar{x}) + \xi^T (x-\bar{x})}$, $\forall x \in S$

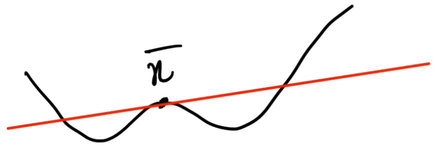
$\hookrightarrow$ Affine / Linear function with slope ξ and line passes through the point $(\bar{x}, f(\bar{x}))$

implies

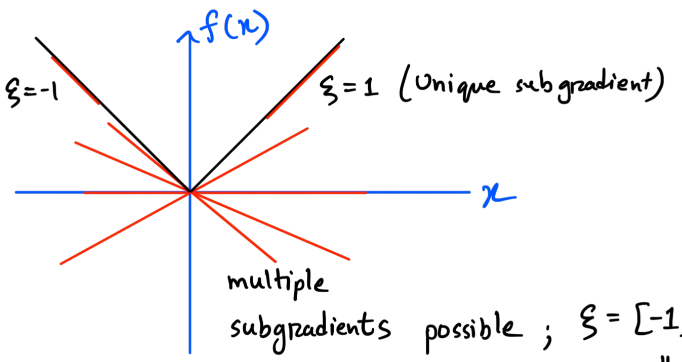
$f(x)$ is lower bounded by linear function



Conclusion: If $f(x)$ is bounded by a linear function, then the slope of the linear function is the subgradient.



No subgradient exists @ $\bar{x}$



$\Downarrow$

$\lambda(1) + (1-\lambda)(-1)$;
$\lambda \in [0,1]$

- "ξ" has the same dimension as "x"
- $H = \{(x, y) : y = f(\bar{x}) + \xi^T (x - \bar{x})\}$ corresponds to supporting Hyperplane to epigraph of f (CONVEX)
  $epi(f) = \{(x, y) : x \in S, y \in \mathbb{R}, y \geq f(x)\}$
- $H = \{(x, y) : y = f(\bar{x}) + \xi^T (x - \bar{x})\}$ corresponds to supporting Hyperplane to hypograph of f (CONCAVE)
  $hyp(f) = \{(x, y) : x \in S, y \in \mathbb{R}, y \leq f(x)\}$
- Collection of subgradients at $\bar{x}$ is a convex Set.
- Subgradient vector ξ corresponds to the slope of supporting Hyperplane

Alternatively, Hyperplane $H = \{(x,y) : y = f(\bar{x}) + \xi^{T}(x-\bar{x})\}$ supports epi(f) at $(\bar{x}, f(\bar{x}))$

$\circledast$ If f is Strictly CONVEX

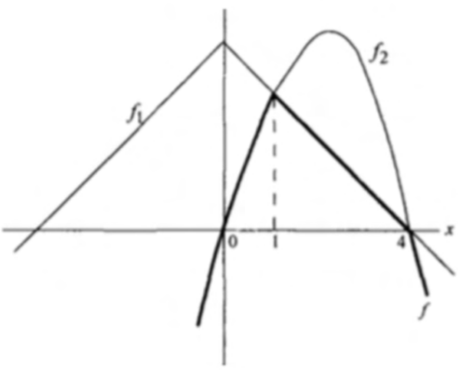
$$f(x) > f(\bar{x}) + \xi^{T}(x-\bar{x}) \quad ; \forall x \in S, \; x \neq \bar{x}$$

$\boxed{/}$ If f is convex on Convex Set, S

- We are not sure whether Gradient ∇f exists!
- We are alway sure the sub-Gradient ξ exists at $\forall x \in \text{int}(S)$

- An example: $f(x) = \min\{f_1(x), f_2(x)\}$, where $f_1(x) = 4 - |x|, x \in \mathbb{R}$ and $f_2(x) = 4 - (x-2)^2, x \in \mathbb{R}$.
- For $x \in [1, 4]$, $f_2(x) \geq f_1(x)$, then f can be represented by

$$f(x) = \begin{cases} 4 - x & 1 \leq x \leq 4, \\ 4 - (x-2)^2 & o/w. \end{cases}$$



- $x \in (1,4)$, $\xi = -1$
- $x < 1$ or $x > 4$, $\xi = (-2)(x-2)$
- at points $x = 1$ or $x = 4$, the sub gradients are not unique because many supporting hyperplanes exist
  at $x = 1$: $\xi = \lambda \nabla f_1(1) + (1-\lambda)\nabla f_2(1) = 2 - 3\lambda$ for $\lambda \in [0,1]$
  at $x = 4$: $\xi = \lambda \nabla f_1(4) + (1-\lambda)\nabla f_2(4) = -4 + 3\lambda$ for $\lambda \in [0,1]$

Existence of Subgradient

Let S be a nonempty convex set in $\mathbb{R}^n$, and let $f : S \to \mathbb{R}$ be convex. Then for $\bar{\mathbf{x}} \in int(S)$, there exists a vector ξ such that the hyperplane

$$H = \{(\mathbf{x}, y) : y = f(\bar{\mathbf{x}}) + \xi^T(\mathbf{x} - \bar{\mathbf{x}})\}$$

supports $epi(f)$ at $[\bar{\mathbf{x}}, f(\bar{\mathbf{x}})]$. In particular,

$$f(\mathbf{x}) \geq f(\bar{\mathbf{x}}) + \xi^T(\mathbf{x} - \bar{\mathbf{x}}), \forall \mathbf{x} \in S$$

that is, ξ is a subgradient of f at $\bar{\mathbf{x}}$. – Idea of Proof??

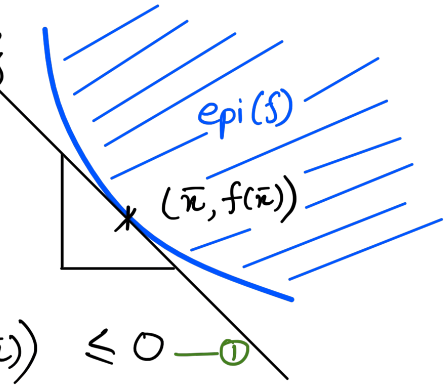
The proof will be based on "Supporting Hyperplane Theorem".
(SPT)

Supporting Hyperplane Theorem

Let S be a nonempty convex set in $\mathbb{R}^n$ and let $\bar{\mathbf{x}} \in \partial(S)$. There exists a hyperplane that supports S at $\bar{\mathbf{x}}$; that is, there exists a nonzero vector p such that $p^T(\mathbf{x} - \bar{\mathbf{x}}) \leq 0$ for each $\mathbf{x} \in cl(S)$.
– Idea of Proof:??

However, in SPT a Set is involved but we need to deal with a Function. How can we relate a Set with a Function ? ⇒ Epigraph!

We know; $epi(f) = \{(x, y) : x \in S, \ y \geq f(x)\}$



Here, $(\bar{x}, f(\bar{x})) \in \partial(epi(f))$

A/c SPT $\Rightarrow \exists\, (P, P_0)$ such that

$\forall\, (x, y) \in epi(f) \longrightarrow \underbrace{p^T(x - \bar{x})} + \underbrace{P_0\,(y - f(\bar{x}))} \leq 0$ — ①

Regular portion in SPT

This is an augmented portion since epi (f) has n+1 dimensions

Now, $P_0 \xrightarrow{?}$ ≤ 0 ✓ / ≥ 0 / $= 0$ $\Rightarrow$ since $(y - f(\bar{x})) \geq 0$ from epi (f)

where $y \geq f(x)$

Again, now, $P_0 \xrightarrow{?} \begin{matrix} =0 \\ <0 \checkmark \end{matrix}$

Proof by contradiction:

Let $P_0 = 0 \overset{①}{\Rightarrow} p^T(x-\bar{x}) \leq 0$ —— ②

We have $\bar{x} \in \text{int}(S) \Rightarrow \exists \varepsilon > 0 \quad N_\varepsilon(\bar{x}) \subseteq S$

$\Rightarrow \forall d \,;\, \|d\| < \varepsilon \,;\, \bar{x}+d \in S$

$② \Rightarrow \quad p^T(\bar{x}+d-\bar{x}) \leq 0 \,;\, \forall d \,;\, \|d\| < \varepsilon$

$\Rightarrow p^T d \leq 0 \,;\, \forall d \,;\, \|d\| < \varepsilon$

Let, $d = \frac{p}{\|p\|} \cdot \frac{\varepsilon}{2} \quad \Rightarrow p^T \frac{p}{\|p\|} \cdot \frac{\varepsilon}{2} \leq 0$

$\Rightarrow \frac{\|p\|^2}{\|p\|} \cdot \frac{\varepsilon}{2} \leq 0$

$\Rightarrow \|p\| \leq 0$

$\Rightarrow \|p\| = 0 \quad \because \|p\| \nless 0$

$\therefore p = 0$ if $P_0 = 0$; which violates $(p, P_0) \neq 0$

Therefore, $P_0 \neq 0$.

Now, $① \Rightarrow \quad p^T(x-\bar{x}) + P_0(y - f(\bar{x})) \leq 0$

$\div P_0 \quad \Rightarrow \frac{p^T}{P_0}(x-\bar{x}) + y - f(\bar{x}) \geq 0 \qquad [\because P_0 < 0]$

$\Rightarrow y \geq f(\bar{x}) - \frac{p^T}{P_0}(x-\bar{x})$

$\Rightarrow y \geq f(\bar{x}) + \xi^T(x-\bar{x})$; Let $\xi = -\frac{p}{P_0}$

$\forall (x,y) \in \text{epi}(f)$ —— ③

③ is valid for $\forall (x,y) \in epi(f)$

Now, from definition of $epi(f)$ $\quad (x, f(x)) \in epi(f)$ $\quad \because \; y \geq f(x)$

$\therefore$ In ③ we can replace y by $f(x)$ $\Rightarrow$ $f(x) \geq f(\bar{x}) + \xi^T (x - \bar{x})$

[Proved]

S is CONVEX ; f is CONVEX on S

$\Rightarrow$ Subgradient ξ^T of f exists at $\bar{x} \in Int(S)$

✗ $\Leftarrow$

f is CONVEX on $int(S)$ $\Leftarrow$ ✓

S is CONVEX ; Subgradient ξ^T exists at $\forall \bar{x} \in Int(S)$

$\Rightarrow f$ is CONVEX on $Int(S)$; NOT (S) $\Rightarrow$ It will be true for "S" if S is Open

Theorem 3.2.6

Let S be a nonempty convex set in $\mathbb{R}^n$, and let $f : S \to \mathbb{R}$. Suppose that for each $\bar{\mathbf{x}} \in int(S)$, there exists a subgradient vector ξ such that

$$f(\mathbf{x}) \geq f(\bar{\mathbf{x}}) + \xi^T (\mathbf{x} - \bar{\mathbf{x}}), \forall \mathbf{x} \in S.$$

Then, f is convex on $int(S)$. NOT "S"

Proof: Given, S is convex

Let, $x_1, x_2 \in int(S) \Rightarrow \exists \bar{x} \in int(S)$ such that $\bar{x} = \lambda x_1 + (1-\lambda) x_2$

For $\bar{x} \in int(S)$, we can write $\Rightarrow f(x_1) \geq f(\bar{x}) + \xi^T (x_1 - \bar{x})$ ——— ①

$$f(x_2) \geq f(\bar{x}) + \xi^T (x_2 - \bar{x}) \quad \text{——— ②}$$

λ ① $+ (1-\lambda)$ ② $\Rightarrow \lambda f(x_1) + (1-\lambda) f(x_2) \geq (\cancel{\lambda} + 1 - \cancel{\lambda}) f(\bar{x}) +$

$$\xi^T \left(\lambda x_1 - \lambda \cancel{\bar{x}} + (1-\lambda) x_2 - \bar{x} + \lambda \cancel{\bar{x}}\right)$$

$$\geq f(\bar{x}) + \xi^T (\cancel{\bar{x}} - \cancel{\bar{x}})$$

$$\geq f(\bar{x}) = f(\lambda x_1 + (1-\lambda) x_2)$$

$\therefore f$ is convex on $int(S)$.

S is CONVEX ; f is CONVEX on S and Differentiable at $\bar{x} \in Int(S)$

$\Rightarrow$ Collection of subgradients at $\bar{x}$ is Singleton Set

i.e. subgradient, ξ = gradient, $\{\nabla f(\bar{x})\}$

Lemma 3.3.2

Let S be a nonempty convex set in $\mathbb{R}^n$, and let $f : S \to \mathbb{R}$ be convex. Suppose that f is differentiable at $\bar{\mathbf{x}} \in int(S)$. Then the collection of subgradients of f at $\bar{\mathbf{x}}$ is the singleton set $\{\nabla f(\bar{\mathbf{x}})\}$.

Proof: Let, $\bar{x} \in int(S) \Rightarrow f(x) \geq f(\bar{x}) + \xi^T(x-\bar{x}) \; ; \forall x \in S$ —— ①

From Taylor Series : $f(x) = f(\bar{x}) + \nabla f(x)^T (x-\bar{x}) + \|x-\bar{x}\| \alpha(\bar{x}, x-\bar{x})$ —— ②

where, $\lim_{x \to \bar{x}} \alpha(\bar{x}, x-\bar{x}) \to 0$

① & ② $\cancel{f(\bar{x})} + \nabla f(\bar{x})^T (x-\bar{x}) + \|x-\bar{x}\| \alpha(\bar{x}, x-\bar{x}) \geq \cancel{f(\bar{x})} + \xi^T(x-\bar{x})$

$(\nabla f(\bar{x}) - \xi)^T (x-\bar{x}) + \|x-\bar{x}\| \alpha(\bar{x}, x-\bar{x}) \geq 0$ —— ③

$\because x \in int(S) \Rightarrow \forall d \;\; \exists \delta > 0 \; ; \; \bar{x} + \lambda d \in S \; ; \; \lambda \in (0, \delta)$

We can replace x by $\bar{x} + \lambda d$ in ③ $\Rightarrow$

$(\nabla f(\bar{x}) - \xi)^T \lambda d + \lambda \|d\| \alpha(\bar{x}, \lambda d) \geq 0$

$\div \lambda \Rightarrow (\nabla f(\bar{x}) - \xi)^T d + \|d\| \alpha(\bar{x}, \lambda d) \geq 0$

Taking limits, $\lim_{\lambda \to 0^+} (\nabla f(\bar{x}) - \xi)^T d \geq 0$ —— ④

Let, $d = -(\nabla f(\bar{x}) - \xi)$; ④ $\Rightarrow -(\nabla f(\bar{x}) - \xi)^T (\nabla f(\bar{x}) - \xi) \geq 0$

↳ Since it's valid for any d

$\Rightarrow -\|\nabla f(\bar{x}) - \xi\|^2 \geq 0$

$\Rightarrow \nabla f(\bar{x}) - \xi = 0 \Rightarrow \boxed{\nabla f(\bar{x}) = \xi}$

$\therefore$ The sub-gradient is the Gradient

Differentiable Convex Functions

$$\min f(\mathbf{x}), s.t. \ \mathbf{x} \in X$$

1. If f is convex on X, then any given point $\bar{\mathbf{x}}$, the affine function $f(\bar{\mathbf{x}}) + \nabla f(\bar{\mathbf{x}})^T(\mathbf{x} - \bar{\mathbf{x}})$ bounds f from below.

   $$\min f(\bar{\mathbf{x}}) + \nabla f(\bar{\mathbf{x}})^T(\mathbf{x} - \bar{\mathbf{x}}), s.t. \ \mathbf{x} \in X$$

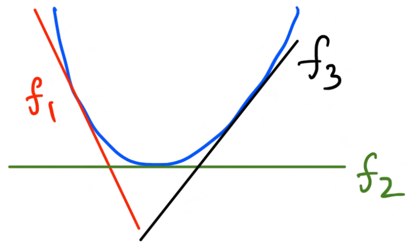
   yields a **lower bound** on the optimum value of the original optimization problem.
2. Outer linearization of the feasible region X

Steps 1 and 2 are used in many successive approximation algorithms for nonlinear optimization problems.

Here, $f_1(x) \leq f(x)$

$f_2(x) \leq f(x)$

$f_3(x) \leq f(x)$



☑ S is CONVEX & OPEN; f is Differentiable on S:

f is convex on S $\boxed{\text{iff}}$ for any $\bar{x} \in \cancel{\text{int}}(S)$ we have

Condition-1: $$f(x) \geq f(\bar{x}) + \nabla f(\bar{x})^T (x-\bar{x}); \ \forall x \in S$$

$= \xi$ since Diff.

Strictly CONVEX f: $>$ $\quad \forall x \neq \bar{x}; \ \bar{x} \in \text{int}(S)$

Condition-2: $$[\nabla f(x^{(1)}) - \nabla f(x^{(2)})](x^{(1)} - x^{(2)}) \geq 0$$

Strictly CONVEX f: $>$

$\forall x^{(1)}, x^{(2)} \in \cancel{\text{Int}}(S)$

The above two conditions are difficult/not practical to check.

For twice differentiable f, we can just check the Hessian, H $\Rightarrow$ if PSD/PD, f is convex.

Condition for twice differentiability:

$\exists\, \nabla f(\bar{x})$ and $\exists\, H(\bar{x})$ such that

$$f(x) = f(\bar{x}) + \nabla f(\bar{x})(x-\bar{x}) + \frac{1}{2}(x-\bar{x})^T H(\bar{x})(x-\bar{x}) + \|x-\bar{x}\|^2 \cdot \alpha(\bar{x}, x-\bar{x});$$

where, $\lim_{x\to\bar{x}} \alpha(\bar{x}, x-\bar{x}) = 0 \qquad \forall x \in S$

2nd order Taylor Series expansion of f.

Alternative 2nd order Taylor Series expansion of f:

Let, $\hat{x} = \lambda x + (1-\lambda)\bar{x}$; $\lambda \in (0,1)$

$$f(x) = f(\bar{x}) + \nabla f(\bar{x})(x-\bar{x}) + \frac{1}{2}(x-\bar{x})^T H(\hat{x})(x-\bar{x})$$

$$H(x) = \left[H_{ij}(x)\right]_{\substack{i=1\dots n\\ j=1\dots n}} = \left(\frac{\partial^2 f(x)}{\partial x_i x_j}\right)_{\substack{i=1\dots n\\ j=1\dots n}}$$

Let, $H = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$; H is $\Rightarrow$ $\qquad \forall z \neq 0$ vector

PSD if:	$a \geq 0$;	$c \geq 0$;	$ac - b^2 \geq 0$	$z^T M z \geq 0$
PD if:	$>$	$>$	$>$	$z^T M z > 0$;
NSD if:	$\leq$	$\leq$	$\geq$	$z^T M z \leq 0$
ND if:	$<$	$<$	$>$	$z^T M z < 0$

Set of all PSD matrices is a CONE; in fact CONVEX CONE.

$S = \{M : M \text{ is PSD}\}$ $\because$ $z^T M z \geq 0 \;\; \forall z \neq 0 \Rightarrow z^T \lambda M z \geq 0 \;\; \therefore \lambda M \in S$

For $H = \begin{bmatrix} a & b & c \\ b & d & e \\ c & e & f \end{bmatrix} \xrightarrow[\text{elemination}]{\text{G.J.}} \begin{bmatrix} a & b & c \\ 0 & d' & e' \\ 0 & e' & f' \end{bmatrix}$ PSD Check: $a \geq 0$, $\begin{bmatrix} d' & e' \\ e' & f' \end{bmatrix}$ is PSD.

$$f(x) = x_1^2 + 2x_1x_2 + x_2^2$$

$$H(x) = \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix} = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}$$

$$H_{11} = \frac{\partial^2 f(x)}{\partial x_1^2} = 2 \qquad H_{12} = \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} = 2$$

$$H_{21} = \frac{\partial^2 f(x)}{\partial x_2 \partial x_1} = 2 \qquad H_{22} = \frac{\partial^2 f(x)}{\partial x_2^2} = 2$$

$$f(x) = 2x_1^2 + 3x_2^2 + 3x_3^2 + 4x_1x_2 + 2x_1x_3 + 2x_2x_3$$

$$H(x) = \begin{pmatrix} H_{11} & H_{12} & H_{13} \\ H_{21} & H_{22} & H_{23} \\ H_{31} & H_{32} & H_{33} \end{pmatrix} = \begin{pmatrix} 4 & 4 & 2 \\ 4 & 6 & 2 \\ 2 & 2 & 6 \end{pmatrix} \xrightarrow[R_3 - \frac{1}{2}R_1]{R_2 - 1R_1}$$

$H(x)$ PSD

$$= \begin{bmatrix} \overset{>0}{4} & 4 & 2 \\ 0 & 2 & 0 \\ 0 & 0 & 5 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 0 \\ 0 & 5 \end{bmatrix} \rightarrow \text{PSD}$$

f is CONVEX on (OPEN) S $\Longleftrightarrow$ $\forall \bar{x} \in S\ \exists \xi : f(x) \geq f(\bar{x}) + \xi^T(x-\bar{x});\ \forall x \in S$

$\Updownarrow$ f is differentiable:

$f(x) \geq f(\bar{x}) + \nabla f(\bar{x})^T(x-\bar{x});\ \forall x \in S$

$[\nabla f(x^{(1)}) - \nabla f(x^{(2)})][x^{(1)} - x^{(2)}] \geq 0 \quad \forall x^{(1)}, x^{(2)} \in S$

$\Updownarrow$ f is twice differentiable:

$H(\bar{x})$ is PSD; $\forall \bar{x} \in S$

Theorem 3.3.7

Let S be a nonempty open convex set in $\mathbb{R}^n$, and let $f : S \to \mathbb{R}$ be twice differentiable on S. Then f is convex iff the Hessian matrix is positive semidefinite at each point in S.

– Idea of Proof???

$(\Rightarrow)$ $f(x)$ is convex & diff. on S $\Rightarrow$ $H(x)$ is PSD $\forall \bar{x} \in S$

$\forall d \neq 0 \quad \exists \lambda > 0 : \quad \bar{x} + \lambda d \in S$

We have: $\quad f(\bar{x}+\lambda d) \geq f(\bar{x}) + \nabla f(\bar{x})^T(\lambda d)$ —— ①

From Taylor series:

$$f(\bar{x}+\lambda d) = f(\bar{x}) + \nabla f(\bar{x})^T(\lambda d) + \frac{1}{2}(\lambda d)^T H(\bar{x})(\lambda d) + \|\lambda d\|^2 \alpha(\bar{x}, \lambda d)$$ —— ②

① & ② $\Rightarrow$ $\frac{1}{2}(\lambda d)^T H(\bar{x})(\lambda d) + \|\lambda d\|^2 \alpha(\bar{x}, \lambda d) \geq 0$

$\div \lambda^2 \Rightarrow \frac{1}{2} d^T H(\bar{x}) d + \|d\|^2 \alpha(\bar{x}, \lambda d) \geq 0$

$\Rightarrow \lim_{\lambda \to 0^+} \frac{1}{2} d^T H(\bar{x}) d + \|d\|^2 \alpha(\bar{x}, \lambda d) \geq 0$ ($\alpha \to 0$)

$\Rightarrow d^T H(\bar{x}) d \geq 0 \quad \therefore$ From definition $H(\bar{x})$ is $\boxed{\text{PSD}}$

$(\Leftarrow)$ $f(x)$ is convex & diff. on S $\Leftarrow$ $H(x)$ is PSD $\forall \bar{x} \in S$

We need to use the Alternative 2nd order Taylor Series expansion:

$$f(x) = f(\bar{x}) + \nabla f(\bar{x})^T (x-\bar{x}) + \frac{1}{2}(x-\bar{x})^T H(\hat{x})(x-\bar{x}) \quad ; \text{ where, } \hat{x} = \lambda x + (1-\lambda)\bar{x}, \ \lambda \in (0,1)$$

③

Since S is open and convex $\hat{x} \in S$

$\Rightarrow H(\hat{x})$ is PSD

$$\Rightarrow \frac{1}{2}\underbrace{(x-\bar{x})^T}_{\text{vectors}} H(\hat{x}) \underbrace{(x-\bar{x})}_{\text{vectors}} \geq 0$$

③ $\Rightarrow$ $f(x) \geq f(\bar{x}) + \nabla f(\bar{x})^T (x-\bar{x})$ $\Rightarrow$ $f(x)$ is convex & diff. on S

Optimal Solutions

$$\min f(\mathbf{x}) : \text{subject to } \mathbf{x} \in S \subseteq \mathbb{R}^n$$

- $\mathbf{x} \in S$ is a feasible solution
- If $\bar{\mathbf{x}} \in S$ and $f(\mathbf{x}) \geq f(\bar{\mathbf{x}})$ for each $\mathbf{x} \in S$, $\bar{\mathbf{x}}$ is an optimal solution, a global optimal solution, or simply a solution to the problem.
- The collections of optimal solutions are alternative optimal solutions
- If $\bar{\mathbf{x}} \in S$ and if there exists an ϵ-neighborhood $N_\epsilon(\bar{\mathbf{x}})$ around $\bar{\mathbf{x}}$ such that $f(\mathbf{x}) \geq f(\bar{\mathbf{x}})$ for each $\mathbf{x} \in S \cap N_\epsilon(\bar{\mathbf{x}})$, $\bar{\mathbf{x}}$ is a local optimal solution.
- Strictly local optimal solution, strong or isolated local optimal solution (self-study)

Global Optimal Sol^n, $\bar{x}$: $f(\bar{x}) \leq f(x)$; $\forall x \in S$

Local Optimal Sol^n, $\bar{x}$: $f(\bar{x}) \leq f(x)$; $\forall x \in S \cap N_\varepsilon(\bar{x})$

$\min \{ f(x) : x \in S \} \Rightarrow$ Convex Programming Problem (CPP)

if S is CONVEX &

f is CONVEX on S

For CPP: ⊛ Local optimal = Global optimal

⊛ $\exists$ Unique Global optimal iff f is Strictly convex i.e. $\nabla^2 f(x)$ or H is PD

Proof: [Local optimal = Global optimal]

Let $\bar{x}$ is local optima $\Rightarrow$ $f(\bar{x}) \leq f(x)$; $\forall x \in S \cap N_\varepsilon(\bar{x})$ — ①

Let $\bar{x}$ is NOT global optima $\Rightarrow$ $\exists \hat{x} \in S$: $f(\hat{x}) < f(\bar{x})$ — ②

Let, $\tilde{x} = \lambda \bar{x} + (1-\lambda)\hat{x}$; $\therefore$ By convexity of function f

$$f(\tilde{x}) \leq \lambda f(\bar{x}) + (1-\lambda) f(\hat{x}) < \lambda f(\bar{x}) + (1-\lambda) f(\bar{x}) = (\cancel{\lambda} + 1 - \cancel{\lambda}) f(\bar{x})$$

$$\Rightarrow f(\lambda \bar{x} + (1-\lambda)\hat{x}) < f(\bar{x}) \quad \text{— ③}$$

$\tilde{x} = \lambda \bar{x} + (1-\lambda)\hat{x}$ will be within $N_\varepsilon(\bar{x})$ under the following condn:

$$\| \lambda \bar{x} + (1-\lambda)\hat{x} - \bar{x} \| < \varepsilon$$

$$\Rightarrow \| (1-\lambda)(\hat{x} - \bar{x}) \| < \varepsilon$$

$$\Rightarrow 1 - \lambda < \frac{\varepsilon}{\|\hat{x} - \bar{x}\|} \Rightarrow \boxed{\lambda > 1 - \frac{\varepsilon}{\|\hat{x} - \bar{x}\|}}$$

With such λ values $\tilde{x} \in N_\varepsilon(\bar{x}) \cap S$ and ③ $\Rightarrow$ we have $f(\bar{x}) > f(\tilde{x})$

which contradicts ①. Therefore, assumption ② can't be true.

Existence of Global Solution:

- $\bar{x} \in S$ is optimal solution of CPP **iff** f has subgradient ξ at $\bar{x}$ such that:

$$\xi^T(x-\bar{x}) \geq 0 \; ; \; \forall x \in S \quad \bigstar$$

CPP

- Additionally if S is OPEN (e.g. $S=\mathbb{R}^n$): (Unconstrained Problem) $\xi^T = 0$.

- Additionally if f is Differentiable:

CPP ⊕ Diff. f

$$\nabla f(\bar{x})^T (x-\bar{x}) \geq 0 \; ; \; \forall x \in S$$

- Additionally if S is OPEN (e.g. $S=\mathbb{R}^n$): (Unconstrained Problem) $\nabla f(\bar{x}) = 0 \quad \bigstar$

Idea for ★: We know: $f(x) \geq f(\bar{x}) + \xi \; (x-\bar{x})$

($\Rightarrow$) $\bar{x}$ is optimal solution $\Rightarrow f(x) \geq f(\bar{x}) \Rightarrow \xi^T(x-\bar{x}) \geq 0$

($\Leftarrow$) $\xi^T(x-\bar{x}) \geq 0 \Rightarrow f(x) \geq f(\bar{x}) \; ; \; \forall x \in S \Rightarrow \bar{x}$ is optimal solution

Idea for ★: We know: $\underbrace{f(x) \geq f(\bar{x}) + \nabla f(x)^T(x-\bar{x})}_{①}$ $\because f$ is diff.

$\because S$ is OPEN, $\forall d \neq 0 \quad \exists \lambda > 0 \quad \bar{x}+\lambda d \in S$

$\because \bar{x}$ is optimal, ① $\Rightarrow \quad \nabla f^T(\bar{x})\,(x-\bar{x}) \geq 0$

Let, $x = \bar{x}+\lambda d \Rightarrow \quad \nabla f^T(\bar{x})\,(\lambda d) \geq 0$

Let, $d = -\nabla f(x) \Rightarrow \quad -\lambda \nabla f^T(\bar{x}) \cdot \nabla f(\bar{x}) \geq 0$

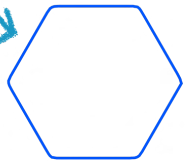
$$\Rightarrow -\|\nabla f(\bar{x})\|^2 \geq 0$$

$$\boxed{\therefore \; \nabla f(\bar{x}) = 0}$$

- $\max\{f(x) : x \in S\} \Rightarrow$ Convex Programming Problem (CPP)

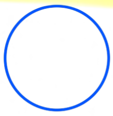
  if S is CONVEX &

  f is CONCAVE on S

- S is POLYTOPE (Compact Polyhedral) 

  f is CONVEX

  $\max\{f(x) : x \in S\} \Rightarrow$ optimal sol^{n}s on Extreme Points — (A)

  (*) If S is Compact (but not polytope) then optimal sol^{n}s on Boundary. 

  (*) NOT Applicable for minimization problem

Proof of (A): Since the feasible region S is a polytope, it can be represented as a convex hull:

$$S = \left\{ x : x = \sum_{i=1}^{m} \lambda_i x^{(i)} \; ; \; \lambda_i \geq 0 \quad \sum_{i=1}^{m} \lambda_i = 1 \right\}$$

where, $x^{(i)}$s are the 'm' extreme points.

By convexity:
$$f\left(\sum_{i=1}^{m} \lambda_i x^{(i)}\right) \leq \sum_{i=1}^{m} \lambda_i f(x^{(i)})$$

$$\leq \max_{i=1}^{m} f(x^{(i)}) \quad \left[\because \text{max. } \& \; \lambda_i \geq 0, \; \sum \lambda_i = 1\right]$$

$$= f(x^{(k)});$$

Let $x^{(k)}$ be the optimal sol^n which is an extreme point.

☒ CPP has nice property : Local optimal = Global optimal

In CPP ⇒

$\begin{cases} f \text{ is convex function} \\ S \text{ is convex set ; comprised of } g_i(x) \end{cases}$

If $g_i(x)$ is convex $\Rightarrow$ S is convex

⊛ But it may be difficult to prove $g_i(x)$ is convex

So we can consider $g_i(x)$ to be Quasiconvex and still have the nice property of CPP.

$g_i(x) \leq 0$ is lower α-level set and for

Quasiconvex $g_i(x) \Leftrightarrow g_i(x) \leq 0$ (α-level set) is lower still convex

Also, $g_i(x)$ convex $\underset{\not\Leftarrow}{\Rightarrow}$ $g_i(x)$ Quasiconvex

⊛ Similarly we can relax f to be Strictly Quasiconvex instead of being convex and still have the nice property of CPP.

⊛ Similarly we can relax f to be Quasiconvex for the $\max\{f(x) : x \in S\}$ problem with polytope S and have optimal sol^{n}s @ extreme points.

Quasiconvex : $f(\lambda x_1 + (1-\lambda)x_2) \leq \max\{f(x_1), f(x_2)\}$

$\forall \lambda \in (0,1)$