Non-linear Programming

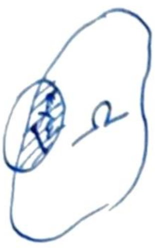
constraints
optimal conditions > need not be linear

Let f be real-valued function defined on $\Omega \in \mathbb{R}^n$
$\Rightarrow$ minimize or maximize $f(x)$ subject to
$x \in \Omega$ (f is not linear)
$\hookrightarrow$ (defined need not be by linear)

Def: A point $x^* \in \Omega$ is called a local minimum (maximum) of f, if $\exists\, \varepsilon > 0$ s.t.

$$\boxed{\begin{array}{l} \forall x \in \Omega \text{ and } \|x - x^*\| < \varepsilon \\ \Rightarrow f(x^*) \leq f(x) \ (\text{or } f(x^*) \geq f(x)) \end{array}}$$

ball center

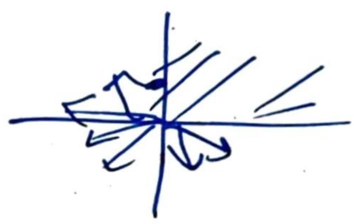


$\Omega \cap \underset{\varepsilon}{B}(x^*) \Rightarrow \{\forall x : \|x - x^*\| < \varepsilon\}$

Def: A point $x^* \in \Omega$ is called global minimum of f, if $f(x^*) \leq f(x)$ $\forall x \in \Omega$

Def: A vector $d \in \mathbb{R}^n$, $d \neq 0$ is said to be feasible direction at $x^* \in \Omega$, if $\exists\, c > 0$ s.t.

$\forall t$, $0 \leq t \leq c$,
$x^* + td \in \Omega$



Ex: $\Omega = \{[x_1, x_2]^T : x_1 \geq 0, x_2 \geq 0\}$

at $[0,0]$, $d = [d_1, d_2]^T$ is a feasible direction, then $d_1 \geq 0$ and $d_2 \geq 0$

at $[0, \frac{1}{2}]$, $d_1 \geq 0$ and ~~$d_2 \leq 0$~~ $d_2 \in \mathbb{R}$

at $[\frac{1}{2}, 0]$, $d_1 \in \mathbb{R}$ and $d_2 \geq 0$.

at $[k_1, k_2]$ $k_1, k_2 > 0 \rightarrow d_1, d_2 \in \mathbb{R}$ $d \in \mathbb{R}^2$

First order necessary conditions for a local minimum

(linear approx)

Remark 1:

If x^* is interior point of Ω, then any $d \in \mathbb{R}^n$ is feasible direction.

Theorem1: Let $f: \Omega \to \mathbb{R}$ be continuously differentiable function.

(partial derivatives are continuous)

$$f(x_1, x_2) = x_1^2 - 3x_1x_2 + 2x_2^2$$
$$= x_1^2 - 2x_1x_2 - x_1x_2 + 2x_2^2$$
$$= x_1(x_1 - 2x_2) - x_2(x_1 - 2x_2)$$
$$= (x_1 - x_2)(x_1 - 2x_2)$$

$$\frac{\partial f}{\partial x_1} = 2x_1 - 3x_2$$
$$\frac{\partial f}{\partial x_2} = -3x_1 + 4x_2$$

> continuous functions (from Ω to $\mathbb{R}$) over Ω, so, f is continuously differentiable.

⇒ If x^* is local minimum point, then for any feasible direction d at x^*

$$\nabla f(x^*) d \geq 0$$

(converse need not be true)

$$\nabla f(\underline{x}) = [2x_1 - 3x_2, -3x_1 + 4x_2]$$

$$(\nabla f(x^*))_i = \left(\frac{\partial f}{\partial x_i}\right)(x^*)$$

Theorem2: $f: \Omega \to \mathbb{R}$ be continuously differentiable function. Let x^* be interior point of Ω. x^* is local minimum then

$$\nabla f(x^*) = 0$$

$$\left(\begin{array}{l} \nabla f(x^*) d \geq 0 \\ \nabla f(x^*)(-d) \geq 0 \\ \Rightarrow \nabla f(x^*) d \leq 0 \\ \Rightarrow \nabla f(x^*) = 0 \end{array}\right)$$

as $d \neq 0$.

all $d \in \mathbb{R}^n$ are feasible directions for interior points.

ex: $f(x_1, x_2) = x_1^3 - x_1^2 x_2 + 2x_2^2$

$$\Omega = \{[x_1, x_2]^T : x_2 \geq 0, x_1 \geq 0\}$$

$$\frac{\partial f}{\partial x_1} = 3x_1^2 - 2x_1x_2$$

$$\frac{\partial f}{\partial x_2} = -x_1^2 + 4x_2$$

$$\nabla f(x) = [0, 0]$$

$$\Rightarrow x_1 = 0 \quad x_2 = 0$$
$$x_1 = 6 \quad x_2 = 9$$

$$x_1^2 = 4x_2$$
$$\cancel{12}_6 x_2 = \cancel{2} x_1 \cancel{x_2}$$
$$x_1 = 6$$
$$x_2 = 9$$

⇒ both these satisfy ~~nece~~ first order necessary conditions for local minimum.

$\begin{bmatrix}0\\0\end{bmatrix}$ is local minimum of f

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

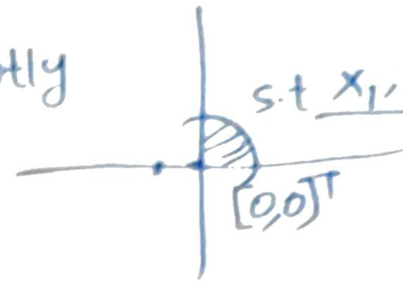
$f(0,0) = 0.$

$f(x_1,x_2) < 0 \Rightarrow x_2 = 0$ and $x_1 < 0$ → outside feasible region

$\cancel{x_1 < x_2}$ $\Rightarrow x_2 \neq 0,\ |x_2| \geq |x_1|$ } $f(x_1, x_2)$ can be <0.

$\cancel{x_1 < 2x_2}$

for x_1, x_2 sufficiently small say $x_1, x_2 < 1$



$$x_1^2 \underbrace{(x_1 - x_2)}_{<1} + 2x_2^2 \geq 0$$

$$x_1^2 < x_2^2 < 1$$

$\Rightarrow \begin{bmatrix}0\\0\end{bmatrix}$ is local minimum (when ε chosen <1)

$\Rightarrow$ If $\Omega = \mathbb{R}^2 \Rightarrow \begin{bmatrix}0\\0\end{bmatrix}$ is not local minimum, left points (no matter how close) give $f < 0$.

ex: minimize $f(x_1, x_2) = x_1^2 - x_1 + x_2 + x_1 x_2$

$\Omega = \{[x_1, x_2]^T,\ x_1 \geq 0,\ x_2 \geq 0\}$

f is continuosly differentiable

$[\frac{1}{2}, 0]$ satisfies first order necessary conditions for local minimum.

$$\Rightarrow \frac{\partial f}{\partial x_1} = 2x_1 - 1 + x_2$$

$$\frac{\partial f}{\partial x_2} = 1 + x_1$$

$[\frac{1}{2}, 0]$

$d_1 \in \mathbb{R}$

$d_2 \geq 0$

$$\nabla f(x) = [2x_1 - 1 + x_2,\ 1 + x_1]$$

$$\nabla f(x^*) = [0,\ 3/2]$$

$$x^* = [\tfrac{1}{2}, 0]$$

$\nabla f(x^*) d$ $\forall$ feasible directions

$$\Rightarrow [0, \tfrac{3}{2}] \begin{bmatrix}d_1\\d_2\end{bmatrix} = \tfrac{3}{2} d_2 \geq 0,\ \forall\ d \in D_{x^*}$$

$$D_{x^*} = \left\{ \begin{bmatrix}d_1\\d_2\end{bmatrix} : d_1 \in \mathbb{R},\ d_2 \geq 0 \right\}$$

↳ set of all feasible directions at x^*

$f(\tfrac{1}{2}, 0) = \tfrac{1}{4} - \tfrac{1}{2} = -\tfrac{1}{4}$

$$= \underbrace{x_1^2 - x_1}_{\text{min at } \frac{1}{2}} + \underbrace{x_2 + x_1 x_2}_{\text{positive, at best } 0.}$$

$\left[\begin{array}{l} 2x_1 - 1 = 0 \\ \to x_1 = 1/2 \\ 2 > 0 \text{ (Min)} \end{array}\right]$

↳ Global Minimum $\begin{bmatrix} \frac{1}{2} \\ 0 \end{bmatrix}$

Theorem 1 proof:

$x^* + td$

$\left(\begin{array}{l} 0 \le t \le c \\ d \text{ is feasible direction.} \end{array} \right.$

Ω

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

(Compare $g(0), g(t)$ for small t)

$$g(t) = g(0) + t g'(0) + o(t)$$

$\lim\limits_{t \to 0} \dfrac{o(t)}{t} = 0.$

(Taylor's expansion).

$g(t) = f(\underbrace{x^* + td}_{\in \Omega})$

d and x^* are fixed

$\left\{ \because \dfrac{o(t)}{t} = \dfrac{t^2 + t^3}{t} \to 0 \text{ as } t \to 0 \right\}$

$$f(x^* + td) = f(x^*) + t g'(0) + o(t).$$

$$= f(x^*) + t \nabla f(x^*) d + o(t)$$

$$= f(x^*) + t \left[\nabla f(x^*) d + \frac{o(t)}{t} \right]$$

for t small, sign is deterimened by $\nabla f(x^*) d$.

for minimum local, $\nabla f(x^*) d \ge 0$.

[if $\nabla f(x^*) d < 0$. then we can choose t small $\Rightarrow$ contradicts x^* being local minimum]

$$\boxed{\begin{aligned} g'(0) &= \lim_{h \to 0} \frac{g(h) - g(0)}{h} \\ &= \lim_{h \to 0} \frac{f(x^* + hd) - f(x^*)}{h} \\ &= \text{directional derivative of } f \text{ along } d. \text{ at } x^* \\ &= \nabla f(x^*) d \qquad d = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \end{aligned}}$$

along one of the coordinates

Second-order necessary conditions for local minimum

Theorem 3:

$f: \Omega \to \mathbb{R}$ be a twice continuously differentiable (second-order partial derivates exist and are continuous) $\left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)$

If x^* is a local minimum of f, then for any feasible direction d at x^*

1. $\nabla f(x^*) d \geq 0$.
2. If $\nabla f(x^*) d = 0$, then $d^T \nabla^2 f(x^*) d \geq 0$

$\nabla^2 f \to$ Hessian maxtrix of f

$$(\nabla^2 f)_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}$$

Theorem 4: If x^* is interior point. If it's local minimum then

1. $\nabla f(x^*) d = 0$
2. $d^T \nabla^2 f(x^*) d \geq 0$, $\forall d \in \mathbb{R}^n$ ($\nabla^2 f(x^*)$ is positive semidefinite)

ex

$$f(x_1, x_2) = x_1^3 - x_1^2 x_2 + 2x_2^2$$

$$\Omega = \{[x_1, x_2]^T : x_1 \geq 0, x_2 \geq 0\}$$

$$\nabla f(x) = [3x_1^2 - 2x_1 x_2, -x_1^2 + 4x_2] \left\{ = [0,0] \to \begin{matrix} [0,0]^T \\ [6,9]^T \end{matrix} \right\}$$

may be more

$$\nabla^2 f(x) = \begin{bmatrix} 6x_1 - 2x_2 & -2x_1 \\ -2x_1 & 4 \end{bmatrix}$$

$d^T \nabla^2 f(x^*) d \geq 0 \quad \forall d$ for interior x^*

$x^* = \begin{bmatrix} 6 \\ 9 \end{bmatrix}$ is interior.

$$\Rightarrow \begin{bmatrix} d_1 & d_2 \end{bmatrix} \begin{bmatrix} 18 & -12 \\ -12 & 4 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$$

$d = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \Rightarrow d^T \nabla^2 f(x^*) d \geq 0 \ \forall d \in \mathbb{R}^2$ (not satisfied)

$(= -2)$

$\Rightarrow$ Not a candidate for local minimum first-order necessary conditions are necessary but not sufficient to be local minimum $\left(\begin{bmatrix}6\\9\end{bmatrix}\right.$ is not local minimum$\left.\right)$

$x^* = \begin{bmatrix}0\\0\end{bmatrix} \to \begin{matrix} d_1 \geq 0 \\ d_2 \geq 0 \end{matrix}$

$\nabla^2 f(x^*) = \begin{bmatrix} 0 & 0 \\ 0 & 4 \end{bmatrix}$

$d_i a_{ij} d_j$

$$\begin{bmatrix} d_1 & d_2 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} = \begin{bmatrix} 0 & 4d_2 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} = 4d_2^2 \geq 0.$$

$\Rightarrow$ So, It satisfies second-order necessary conditions.
so, $\begin{bmatrix}0\\0\end{bmatrix}$ is local minimum.

$\Omega = \mathbb{R}^2 \Rightarrow \begin{bmatrix}0\\0\end{bmatrix}$ satisfies both first-order necessary & second-order necessary, but $\begin{bmatrix}0\\0\end{bmatrix}$ is not local minimum.

<u>ex</u> :

$$f(x_1, x_2) = x_1^2 - x_1 + x_2 + x_1 x_2$$
$$x_1 \geq 0 \quad x_2 \geq 0$$

f is twice continuosly differentiable.

$x^* = \begin{bmatrix} 1/2 \\ 0 \end{bmatrix}$

$\nabla f(x_1, x_2) = \begin{bmatrix} 2x_1 - 1 + x_2 , & 1 + x_1 \end{bmatrix}$ $\qquad \nabla f(x^*) = \begin{bmatrix} 0, 3/2 \end{bmatrix}$

$\nabla f(x^*) d = 3/2 \, d_2$
$d_2 \geq 0$

first-order satisfied

$\nabla^2 f(x_1, x_2) = \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix}$

$d_1 \in \mathbb{R}, \; d_2 \geq 0.$

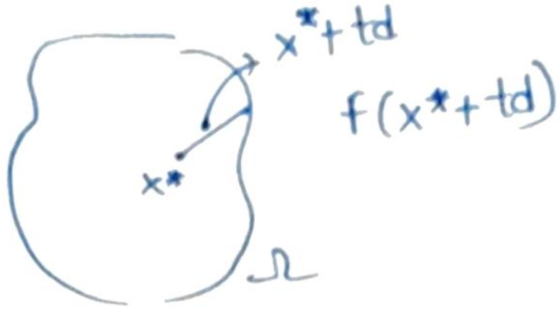
$$\Rightarrow \quad d^T \nabla^2 f(x^*) d = \begin{bmatrix} d_1 & d_2 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$$

If $\nabla f(x^*) d = 0.$
$\Rightarrow d_2 = 0.$

$$= \begin{bmatrix} 2d_1 + d_2 , & d_1 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$$
$$= 2d_1^2 + d_1 d_2 + d_1 d_2$$
$$= 2d_1^2 + 2d_1 d_2$$
$$= 2d_1 (d_1 + d_2)$$
$$= 2d_1^2 \geq 0.$$

$\Rightarrow$ second order also satisfied.

$x^* + td$,
$0 \leq t \leq c$, $c > 0$
feasible



$$f(x^* + td) = g(t)$$

$$\Rightarrow \quad g(t) = g(0) + t g'(0) + \frac{t^2}{2} g''(0) + o(t^2)$$

$$\lim_{t \to 0} \frac{o(t^2)}{t^2} = 0.$$

$$g'(0) = \nabla f(x^*) d$$

$$g'(t) = \lim_{h \to 0} \frac{g(t+h) - g(t)}{h}$$

$$= \nabla f(x^* + td)\, d$$

$$g'(t) = \sum_{i=1}^{n} d_i \left(\frac{\partial f}{\partial x_i}\right)(x^* + td)$$

$$g'(t) = \sum_{i=1}^{n} d_i\, h_i(t)$$

$$g''(t) = \sum_{i=1}^{n} d_i\, h_i'(t)$$

$$h_i'(t) = \lim_{k \to 0} \frac{h_i(t+k) - h_i(t)}{k}$$

$$= \nabla\left(\frac{\partial f}{\partial x_i}\right)(x^* + td) \cdot d$$

$$\Rightarrow \cancel{g''(t) =} = \sum_{j=1}^{n} d_j \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right)(x^* + td).$$

$$g''(t) = \sum_{i=1}^{n} d_i \left[\sum_{j=1}^{n} d_j \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right)(x^* + td)\right]$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} \underbrace{\frac{\partial^2 f(x^* + td)}{\partial x_i \partial x_j}}_{a_{ij}} d_j\, d_i$$

$$g''(0) = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(x^*)\, d_j\, d_i$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}\, d_i\, d_j$$

$= d^T A d$

$= d^2 \nabla^2 f(x^*) d$

(Hessian matrix is symmetric, as f is twice continuosly differentiable)

$g(t) = g(0) + t g'(0) + \frac{t^2}{2} g''(0) + o(t^2)$

$\Rightarrow f(x^* + td) = f(x^*) + t \nabla f(x^*) d + \frac{t^2}{2} d^T \nabla^2 f(x^*) d + o(t^2)$.

If $\nabla f(x^*) d = 0$.

$\Rightarrow f(x^* + td) = f(x^*) + \frac{t^2}{2} d^T \nabla^2 f(x^*) d + o(t^2)$

$= f(x^*) + \frac{t^2}{2} \left[d^T \nabla^2 f(x^*) d + \frac{2 o(t^2)}{t^2} \right]$

as $t \to 0 \Rightarrow \frac{o(t^2)}{t^2} \to 0$.

$\Rightarrow$ sufficient small t, sign is decided by $d^T \nabla^2 f(x^*) d$, so

~~if $d^T \nabla^2 f(x^*) d \geq 0 \Rightarrow x^*$ may be local minimum~~

(If < 0, t can be chosen sufficiently small to contradict x^* being local minimum).

So, $\boxed{d^T \nabla^2 f(x^*) d \geq 0}$

$\nabla^2 f(x^*) \to$ positive semidifinite.

Def : Real symmetric matrix A is said to be positive semidefinite $x^T A x \geq 0 \ \forall x \in \mathbb{R}^n$

Def: positive definite if $x^T A x > 0 \ \forall x \neq 0$.

eigen values non-neg $\leftarrow H = \begin{bmatrix} 0 & 0 \\ 0 & 4 \end{bmatrix}$ $\quad d^T H d = 4 d_2^2 \geq 0$

$d = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \neq 0 \quad = 0$

Theorem:

$\lambda \in \mathbb{C}$ is eigen value of $n \times n$ A, if $\exists \ x \in \mathbb{C}^n$, $x \neq 0$ such that $Ax = \lambda x$

$|\lambda I - A|$ upper triangular

$\Rightarrow (\lambda - a_{11})(\lambda - a_{22})(\lambda - a_{33}) = 0$.

$a_{11}, a_{22}, a_{33} \to$ eigen values

1. A is positive semidefinite
2. All eigen values are non-negative.
3. All principal minors are non-negative.

} equivalent

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

$\alpha \subseteq \{1,2,3\}$
$\beta \subseteq \{1,2,3\}$

$\alpha = \{1,2\}$ $\quad \beta = \{3\}$

$$A_{\alpha\beta} = \begin{bmatrix} a_{13} \\ a_{23} \end{bmatrix}$$

$\alpha = \{1,2\}$ $\quad \beta = \{1,3\}$

$$A_{\alpha\beta} = \begin{bmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{bmatrix}$$

When $\alpha = \beta$, $A_{\alpha\alpha}$ is principal submatrix of A.

$\alpha = \{1\}$ $\quad A_{\alpha\alpha} = [a_{11}]$ $\quad |A_{\alpha\alpha}| = a_{11} \rightarrow$ principal minor A

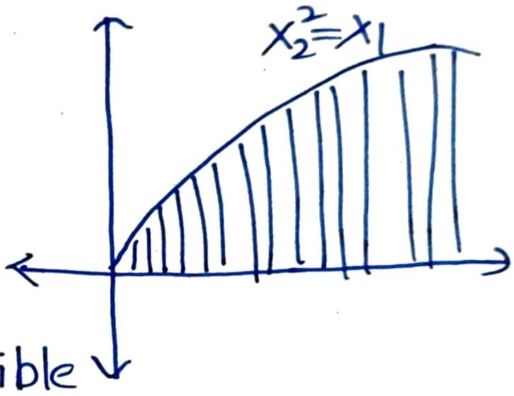
$\alpha = \{1,2\}$ $\quad A_{\alpha\alpha} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ $\quad |A_{\alpha\alpha}| = a_{11}a_{22} - a_{12}a_{21}$

$\alpha = \{1,3\} = A_{\alpha\alpha} = \begin{bmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{bmatrix}$

$f(x^* + td) = f(x^*) + t\,\nabla f(x^*)d +$ ← 1st order approx
$>0 \quad \times \quad o(t)$

⇒ ex: $f(x_1, x_2) = x_1 - 2x_2^2$ $\quad \Omega = \{ [x_1, x_2]^T : x_1 \geq 0, x_2^2 \leq x_1 \}$

$$\nabla f(x_1, x_2) = \begin{bmatrix} 1 \\ -4x_2 \end{bmatrix}$$



$x^* = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow \nabla f(x^*)^t \begin{bmatrix} 1 \\ 0 \end{bmatrix} d$
$= d_1$

$d_1 > 0$ at $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ (for any feasible direction).

more
sufficient conditions for local minimum:

Theorem: let $f : \Omega \rightarrow \mathbb{R}$ be a twice continuously differentiable.
let x^* be interior point
1. $\nabla f(x^*) = 0$
2. $\nabla^2 f(x^*)$ is positive definite

⇒ then x^* is a local minimum of f.

$$f(x^* + td) = f(x^*) + t\,\nabla f(x^*)d + \frac{t^2}{2} d^T \nabla^2 f(x^*) d + o(t^2)$$

Theorem:

If A is symmetric,

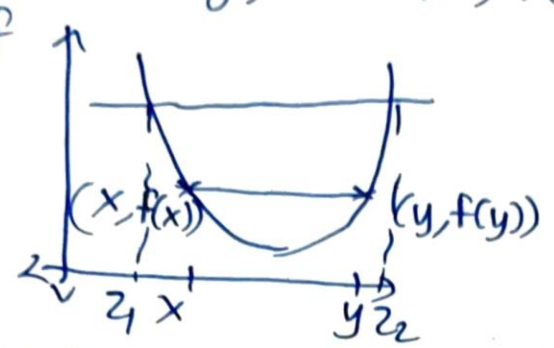
1. A is positive definite
2. All eigenvalues are positive.
3. All principal minors are positive.

$\Rightarrow$ non-singular (non-zero determinant) positive semidefinite — positive definite.

Def: A real valued function f defined on a convex set $\Omega \subseteq \mathbb{R}^n$ is said to be convex function on Ω, if $\forall\ x, y \in \Omega$ and all $0 \leq \alpha \leq 1$

$$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y)$$

(line joining x, y is above f graph)



Def: function is concave, if $-f$ is convex function.

Theorem1: If f is a convex function on Ω (convex set).

Then the set $S = \{x \in \Omega : f(x) \leq c\}$ is a convex set $\boxed{\forall \text{ real } c}$.

$[z_1, z_2]$ is convex set above.

$\Rightarrow$ **Proof:** (empty set is also convex)

$\Rightarrow \quad S = \{x \in \Omega : f(x) \leq c\} \quad c \in \mathbb{R}$

let $S \neq \emptyset$, let $x_1, x_2 \in S$ $\quad f(x_1) \leq c$

$\qquad f(x_2) \leq c$

$\Rightarrow \quad \alpha f(x_1) + (1-\alpha) f(x_2) \leq c \quad 0 \leq \alpha \leq 1$

$f(\alpha x_1 + (1-\alpha)x_2) \leq$

$\Rightarrow \quad \alpha x_1 + (1-\alpha) x_2 \in S \quad \forall \alpha$

$\Rightarrow \quad S$ is convex set.

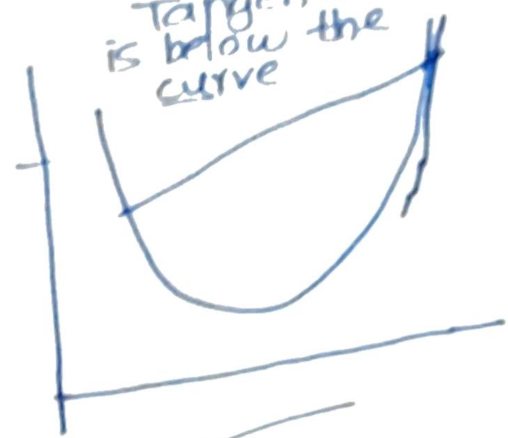
Theorem 2:

Let f be a continously differentiable function defined on convex set, $\Omega \subseteq \mathbb{R}^n$, then f is convex iff

$$f(y) \geq f(x) + \nabla f(x)(y-x) \quad \forall x, y \in \Omega$$

Tangent plane at x.

$\Rightarrow$ approximation from this of value of function (first order)



Proof:

Convex May be discontinuo-s, shoot up on boundary

"$\Rightarrow$" f is a convex on Ω

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

$\Rightarrow$ $f(x + \alpha(y-x)) - f(x) \leq \alpha(f(y) - f(x))$

for $\alpha > 0$.

$\Rightarrow \lim_{\alpha \to 0} \dfrac{f(x + \alpha(y-x)) - f(x)}{\alpha} \leq f(y) - f(x)$.

$\Rightarrow$ (directional derivative at x along direction $y-x$)

$\Rightarrow$ $\nabla f(x)(y-x) + f(x) \leq f(y)$.

$\alpha = 0 \quad y = x$ ✓

"$\Leftarrow$" Let f satisfy $f(y) \geq f(x) + \nabla f(x)(y-x) \quad \forall x, y \in \Omega$

fix some $x, y \in \Omega$

$z = \alpha x + (1-\alpha) y \qquad 0 \leq \alpha \leq 1$.

$\alpha f(x) \geq \alpha f(z) + \alpha \nabla f(z)(x - z)$.

$f(y) \geq f(z) + \nabla f(z)(y-z)$.

$\Rightarrow (1-\alpha) f(y) \geq (1-\alpha) f(z) + (1-\alpha) \nabla f(z)(y-z)$.

$\Rightarrow \alpha f(x) + (1-\alpha) f(y) \geq f(z) + \nabla f(z)(\cancel{\alpha x + (1-\alpha) y}) \cancel{- \nabla f(z) z}$

$\Rightarrow \quad f(z) \leq \alpha f(x) + (1-\alpha) f(y)$

$\Rightarrow \quad f(\alpha x + (1-\alpha) y) \leq \alpha f(x) + (1-\alpha) f(y)$

f is convex

Theorem 3: Let f be twice continuously differentiable, on convex set Ω (Ω has atleast one interior point). ~~If~~ f convex $\Leftrightarrow$ $\forall x \in \Omega$ $\nabla^2 f(x)$ is positive semidefinite.

ex:
$$f(x_1, x_2) = x_1^3 - x_1^2 x_2 + 2x_2^2$$
$$\Omega = \{ [x_1\ x_2]^T : x_1 \geq 0, x_2 \geq 0 \}$$
$$\Rightarrow \quad \nabla^2 f(x_1, x_2) = \begin{bmatrix} 6x_1 - 2x_2 & -2x_1 \\ -2x_1 & 4 \end{bmatrix}$$
$$x_1 = 1 \quad x_2 = 3 \Rightarrow \begin{bmatrix} 0 & -2 \\ -2 & 4 \end{bmatrix}$$

det $= -4 \Rightarrow$ so not positive semidefinite

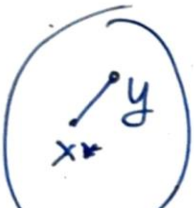
$\Rightarrow f$ is not convex.

Theorem 4: $f : \Omega \to \mathbb{R}$ is continuously differentiable. If f is convex on Ω, then x^* is a global minimum of $f \Leftrightarrow$ $\forall$ feasible directions d at x^*, $\nabla f(x^*) d \geq 0$

Corollary: x^* is interior, f convex, global minimum $\Leftrightarrow \nabla f(x^*) = 0$.

Proof: "$\Leftarrow$" $\nabla f(x^*) d \geq 0$ $\forall$ feasible directions d at x^*.

To show x^* is global minimum.



$x^* + \lambda (y - x^*)$ $\quad \forall y$ in Ω

$\Rightarrow y - x^*$ is a feasible direction at x^*.

$\Rightarrow \nabla f(x^*)(y - x^*) \geq 0. \ \forall y \in \Omega$

$f(y) \geq f(x^*) + \underbrace{\nabla f(x^*)(y - x^*)}_{\geq 0}$ (f is convex) $\forall y$

$\Rightarrow$ So, x^* is ~~local~~ global minimum

"$\Rightarrow$" Any local minimum satisfies $\nabla f(x^*) d \geq 0$.

$\Rightarrow$ So, local minimum is global minimum for continuously differentiable convex function.

$\Rightarrow$ linear is both convex & concave

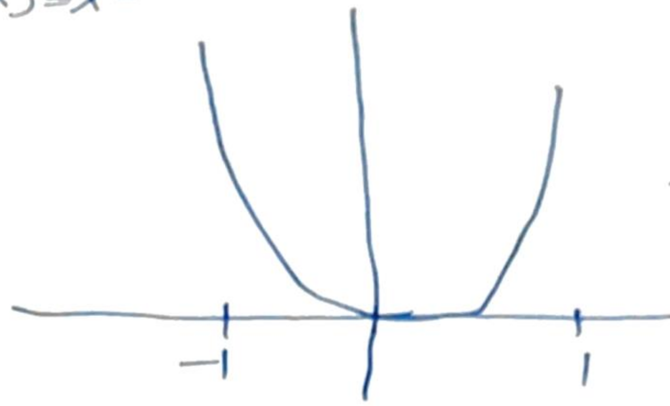
minimizing convex is like maximizing concave (necessary become sufficient).

Theorem 5: f is convex function.

1. set of minimums is convex set.
2. x^* is local minimum $\Rightarrow$ it is global minimum

Remark: Theorem 5 doesnot hold when maximizing a convex function.

ex: $f(x)=x^2$



max $f(x)=x^2$

$-1 \leq x \leq 1$

Set of all optimal solutions $=\{-1,1\}$

not convex set.

$-1 \leq x \leq 2$

$\Rightarrow$ local maximum is not global maximum (-1).