

NL/28/10/25

→ Convex Sets, Convex Functions & Optimization

- Line passing through two points x_1, x_2



$$y = \theta x_1 + (1 - \theta) x_2 \quad \theta \in \mathbb{R}$$

- Line segment:

$$y = \theta x_1 + (1 - \theta) x_2$$

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



- Affine Combination of Points

$$\{x_1, x_2, \dots, x_n\}$$

$$y = \theta_1 x_1 + \theta_2 x_2 + \dots + \theta_n x_n$$

$$\sum_{i=1}^n \theta_i = 1$$

- Affine Set is collection of all affine points of a set.

NL/29/10/25

- Line
- Line segment

- Affine combination of points

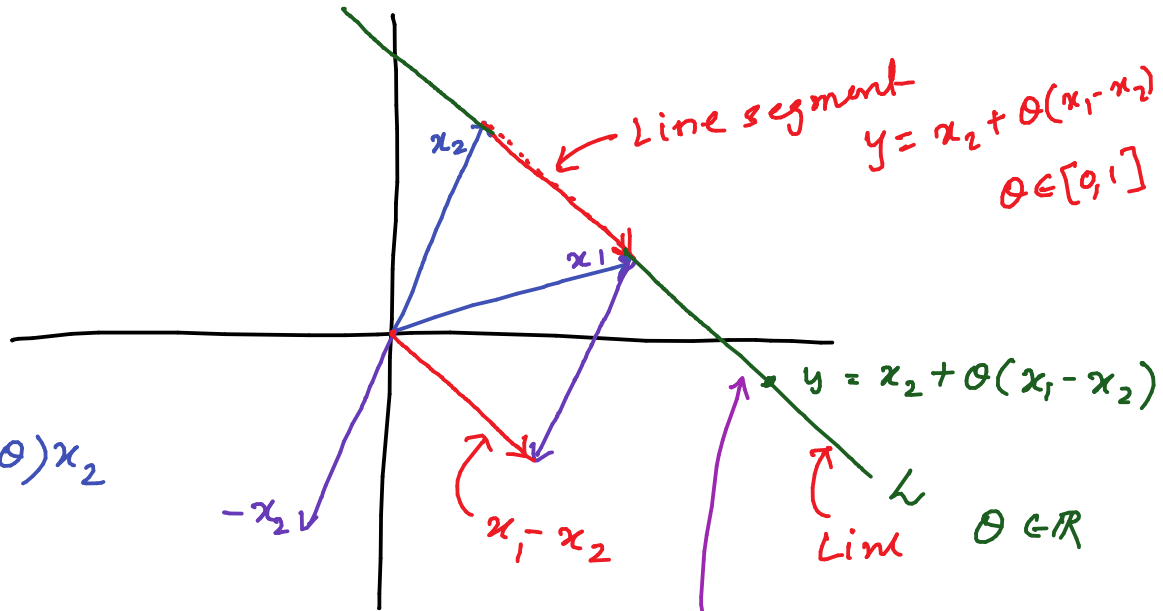
$$x_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$x_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$y = \theta x_1 + (1-\theta)x_2$$

$$\theta \in \mathbb{R}$$

$$= x_2 + \theta(x_1 - x_2)$$



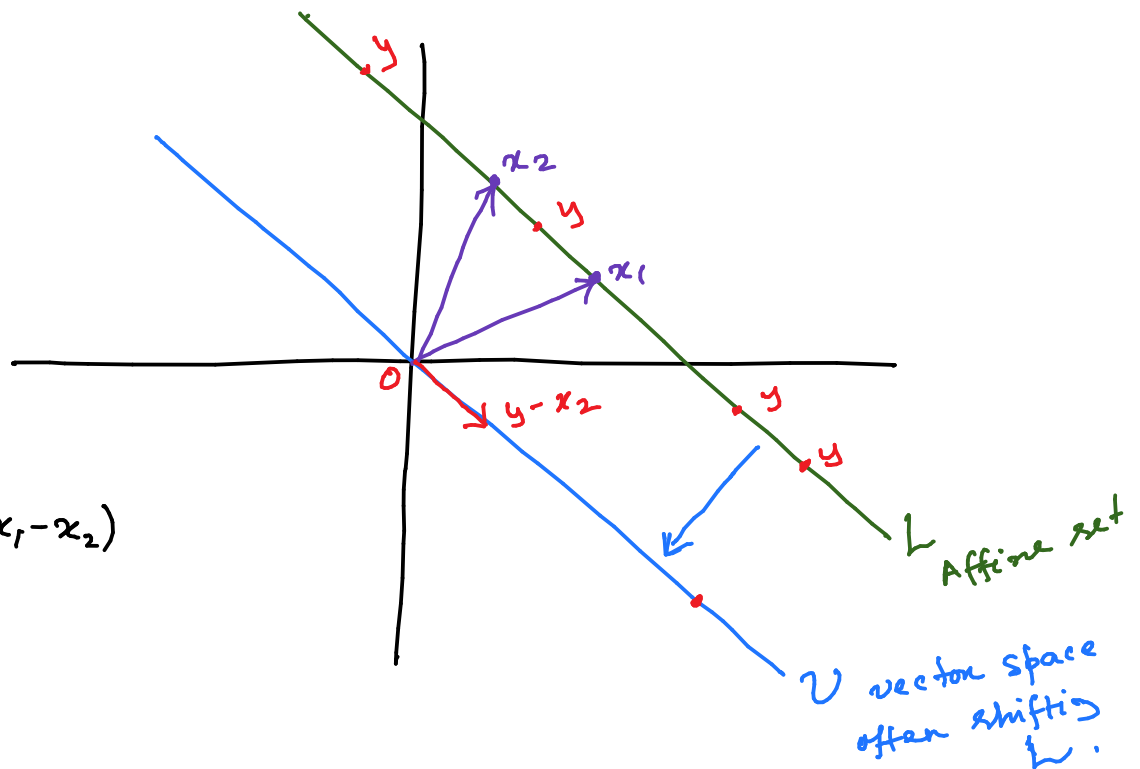
The points lying on this line are affine combination of points on L .

$$y = \sum_{i=1}^k \theta_i x_i \quad \text{for } x_i \in S \quad \& \quad \sum_{i=1}^k \theta_i = 1$$

↑
 affine combination of points

$$x_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$x_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$



$$y = x_2 + \theta(x_1 - x_2)$$

$$v = y - x_2$$

$\text{If } S_a \text{ is an } \underbrace{\text{affine set}}_{\substack{\text{collection of} \\ \text{all affine} \\ \text{combinat}^n \text{ of} \\ \text{points}}} \quad \exists x_0 \in S_a.$

$$V := S_a - x_0 = \{ y - x_0 \mid y \in S_a \}$$

V is a
vector space!

$$S_a = \{ v_i + x_0 \mid v_i \in V \}$$

affine set on affine space
 $\uparrow$
 depends on V

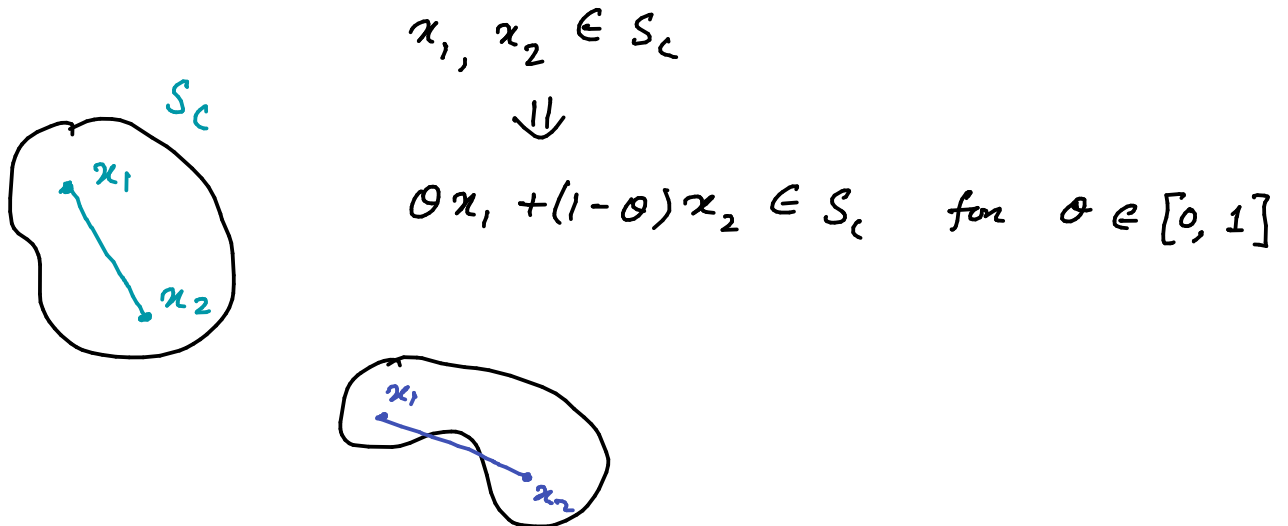
- Affine Hull of a set S_a :

$$\text{aff}(S_a) \quad \{x_1, x_2 \dots x_k\} \in S_a$$

$$\text{aff}(S_a) = \left\{ y \mid y = \sum_{i=1}^k \theta_i x_i, \quad x_i \in S_a, \quad \sum_{i=1}^k \theta_i = 1 \right\}$$

- Convex Set :

A set S_c is a convex set if the line segment between any two points in S_c lies entirely in S_c .



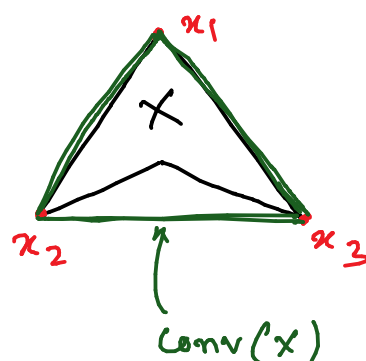
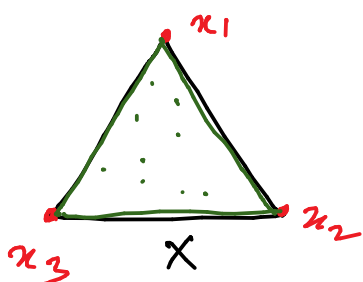
- Convex combination of points

$$y = \theta_1 x_1 + \theta_2 x_2 + \dots + \theta_k x_k$$

$$\theta_i \geq 0 \quad \& \quad \sum_{i=1}^k \theta_i = 1$$

- Convex Hull of a set X :

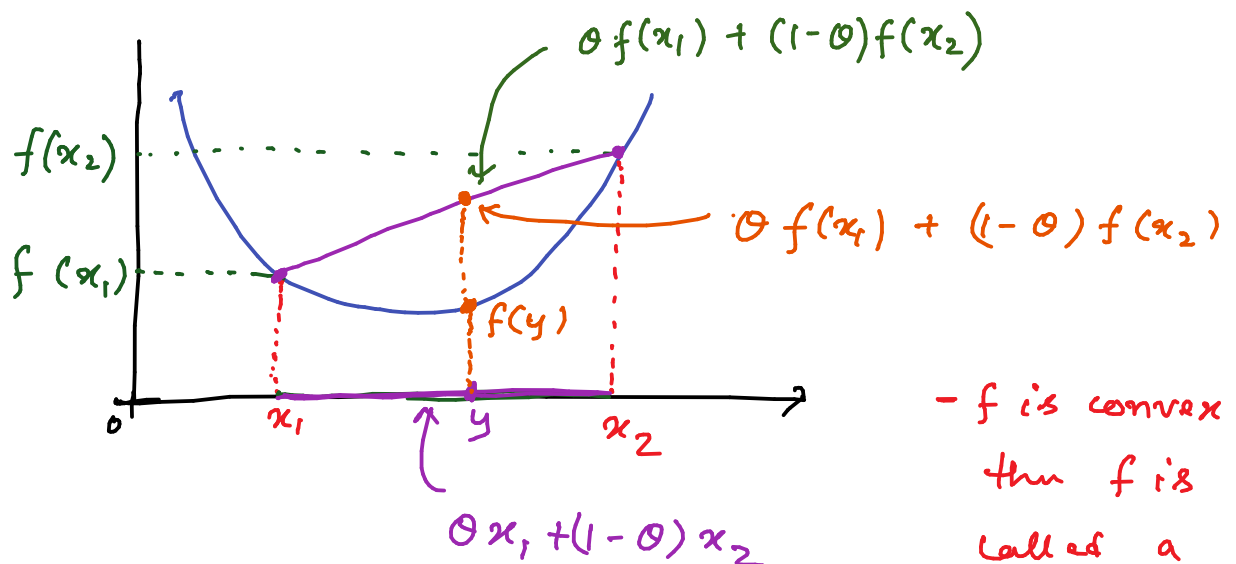
$$\text{conv}(X) := \left\{ y = \sum_{i=1}^k \theta_i x_i \mid x_i \in X, \theta_i \geq 0, \sum_{i=1}^k \theta_i = 1 \right\}$$



- The convex hull of a set X is always a convex set (does not depend on the convexity of X)
- The intersection of any collection of convex sets is convex.
- The convex hull of a compact set is compact.
- Convex Function:

A real valued funⁿ $f: X \rightarrow \mathbb{R}$, where $X \subseteq \mathbb{R}^n$ is a convex set, is said to be a convex function if:

$$f(\underbrace{\theta x_1 + (1-\theta)x_2}_y) \leq \theta f(x_1) + (1-\theta)f(x_2) \quad \forall x_1, x_2 \in X \quad \theta \in [0, 1]$$



- f is convex
then f is
called a
concave
funⁿ.

→ Conditions to test Convexity of a funⁿ

- First order Condition

Let $f : X \rightarrow \mathbb{R}$, where $X \subseteq \mathbb{R}^n$ is an open set, be a real valued function. Let f be differentiable in X . Then, f is convex if and only if:

(i) the domain set X is convex, and

(ii) $f(y) \geq f(x) + [\nabla f(x)]^T (y-x)$... (*)

holds for all $x, y \in X$.

1st order Taylor series approximation

y can be thought of as $y = x + \epsilon d$

where $\epsilon > 0$ is very small number

& $d \in \mathbb{R}^n$ is the direction.

- The inequality in (*) says that the local test can be used to infer the function nature (convexity) globally.

- Second Order Condition :

Now assume that $f: X \rightarrow \mathbb{R}$ be twice continuously differentiable, i.e. its

Hessian $\nabla^2 f(x)$ exists at each point in X .

Then, f is convex if and only if :

(i) X is convex, and

(ii) $\nabla^2 f(x) \succeq 0 \rightarrow$ Symmetric positive semi-definite for all $x \in X$.

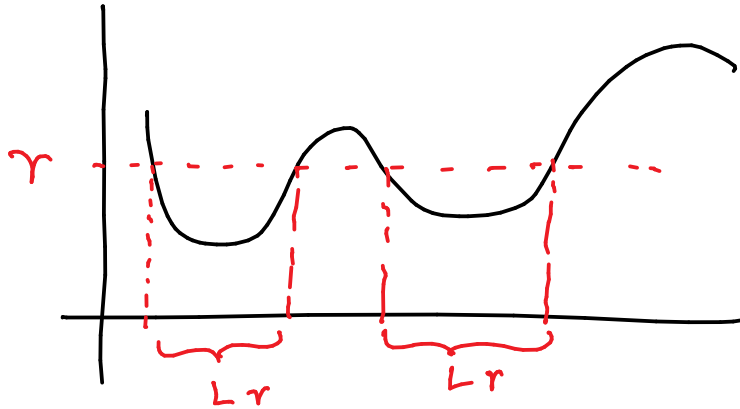
Can geometrically be interpreted as the graph at x has positive (upward) curvature.

- f is concave if and only if

$\left\{ \begin{array}{l} \text{(i) } X \text{ is convex and} \\ \text{(ii) } \nabla^2 f(x) \preceq 0, \forall x \in X. \end{array} \right.$

- Level sets of f .

$$L_r = \{x \mid f(x) \leq r\} \quad \text{when } r \text{ is a scalar.}$$



→ $f : X \rightarrow \mathbb{R}$ with $X \subseteq \mathbb{R}^n$ is convex

if

- (i) X is convex set and
- (ii) $f(\theta x_1 + (1-\theta)x_2) \leq \theta f(x_1) + (1-\theta)f(x_2)$

for all $\theta \in [0, 1]$ and $x_1, x_2 \in X$.

→

Optimizat' :

$$\min f_0(x)$$

$$x \in S$$

↑ constraint set

Constraint sets are characterized

$$S \begin{cases} f_i(x) \leq 0 \\ h_i(x) = 0 \end{cases}$$

Optimizat'

Unconstrained

$$\min f_0(x)$$

$$x \in \mathbb{R}^n$$

Constrained

$$\min f_0(x)$$

$$x \in S$$

$$S \subset \mathbb{R}^n$$

The constraint set

$$S = \left\{ x \mid x \in \left(\bigcap_{i=1}^m (\text{dom } f_i) \cap \left(\bigcap_{j=1}^k (\text{dom } h_j) \right) \right) \right\}$$

$$\min x_1 + x_2$$

$$f_i: \mathbb{R}^2 \rightarrow \mathbb{R}$$

$$s.t. \rightarrow x_1^2 + x_2^2 \geq 0$$

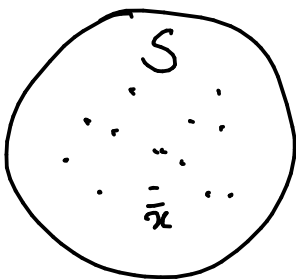
$$f_1 \rightarrow x_1^2 \geq 0$$

$$f_2 \rightarrow x_1 + 2x_2 = 0$$

$$h_1$$

NL/31/10/25

- A point $\bar{x}$ is feasible if



$$\begin{cases} f_i(\bar{x}) \leq 0 \\ h_j(\bar{x}) = 0 \end{cases}$$

- A point x_0 is said to be optimal

if

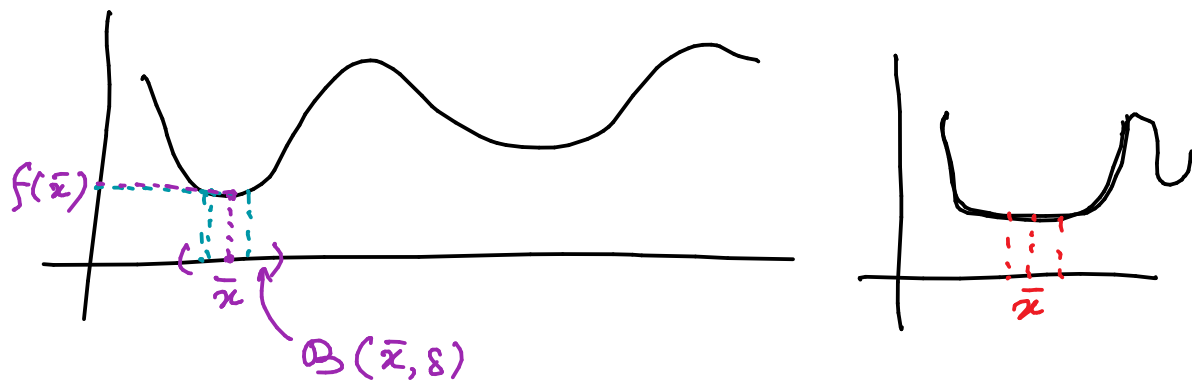
$$x_0 = \inf \left\{ f_0(x) \mid f_i(x) \leq 0, h_j(x) = 0 \right\}$$

$$\left. \begin{array}{l} \min f_0(x) \\ \text{s.t. } f_i(x) \leq 0 \\ h_j(x) = 0 \end{array} \right\} \equiv \left\{ \begin{array}{l} \max -f_0(x) \\ \text{s.t. } f_i(x) \leq 0 \\ h_j(x) = 0 \end{array} \right.$$

- Local minimum :

A feasible point $\bar{x}$ is local minimum of an O.P. if there exists a hypersphere $\mathcal{B}(\bar{x}, \delta)$ s.t.

$$f(\bar{x}) \leq f(x) \quad \forall x \in \mathcal{B}(\bar{x}, \delta).$$



- strictly local minimum if

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

- Global minimum :

$$\text{if } f(\bar{x}) \leq f(x) \quad \forall x \in \mathbb{R}^n$$

$$\begin{array}{lcl}
 \min f_0(x) & & \min f_0(x) \\
 \text{s.t.} & \left\{ \begin{array}{l} f_i(x) \leq 0 \\ h_j(x) = 0 \end{array} \right. & \rightarrow \left\{ \begin{array}{l} \boxed{r_i \geq 0} \\ f_i(x) + \underline{r_i} = 0 \\ h_j(x) = 0 \end{array} \right.
 \end{array}$$

r_i are called "slack variables".

→ Convex Optimization Problem

$$\begin{array}{l}
 \left\{ \begin{array}{l} \min f_0(x) \\ \text{s.t.} \quad f_i(x) \leq 0 \\ \quad \quad h_j(x) = 0 \end{array} \right. \\
 \swarrow \\
 \text{said to be convex O.P. if}
 \end{array}$$

(i) $f_0(x)$ is a convex function.

(ii) The search area S (constraint set) must be convex

↙ to ensure this

• $f_i(x)$ must be convex function.

• $h_i(x)$ must be in the following form: $a_j^T x - b_j = 0$

$$\begin{bmatrix} * & * & x & * \end{bmatrix}^T \\
 a_j$$

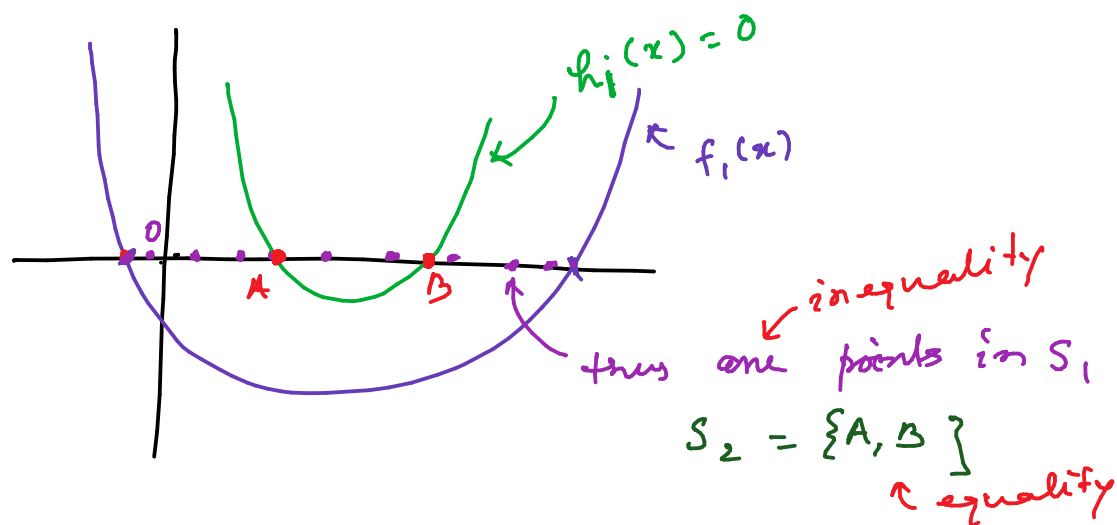
$$a_j^T x - b_j = 0$$

$$\begin{bmatrix} 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} - 6 = 0$$

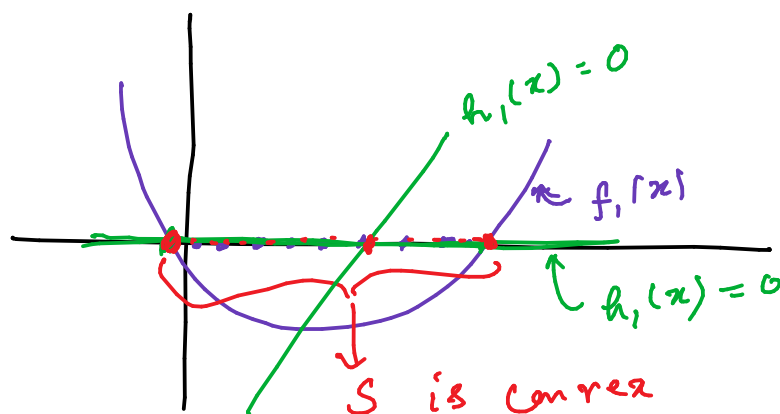
$a^T \quad x \quad b$

$$2x_1 + 3x_2 + 4x_3 - 6 = 0$$

$$h_j(x) = 0 \Rightarrow Ax - b = 0$$



$$S = S_1 \cap S_2 = \{A, B\} \leftarrow \text{Not a convex set!}$$



$$S_1 = \{x \mid f_i(x) \leq 0\} \leftarrow f_i \text{ must be convex funct?}$$

$$S = S_1 \cap S_2$$

$\uparrow$
 Now convex

$$S_2 = \{x \mid h_j(x) = 0\} \leftarrow h_j \text{ must be in the form of affine funct?}$$

• Optimality Conditions: [For Unconstrained Optimizant]

(i) Let $\bar{x}$ be a local minimum of $f: \mathbb{R}^n \rightarrow \mathbb{R}$.
Assume that f is continuously differentiable
in an open set S containing $\bar{x}$.

↓

(i) $\nabla f(\bar{x}) = 0$ (First order necessary condⁿ)

If f is twice continuously differentiable.

thn

$$\nabla^2 f(\bar{x}) \succeq 0$$

Hessian is a ↓

Positive semi-definite matrix

A matrix $M \in \mathbb{R}^{n \times n}$ is said to be
positive definite if $x^T M x > 0$.
 $\forall x \in \mathbb{R}^n$

* Read how to test a matrix is
positive definite or not.

positive semi-definite $x^T M x \succeq 0$
 $\forall x \in \mathbb{R}^n$

- Sufficient Optimality Condition
[Unconstrained]

Let $f: \mathbb{R}^n \rightarrow \mathbb{R}$ be twice continuously differentiable funtⁿ in an open set S .

Assume that a vector $\bar{x} \in S$ satisfies

- $\nabla f(\bar{x}) = 0$
- $\nabla^2 f(\bar{x}) \succ 0$

 $\leftarrow$ Positive definite.

$\Downarrow$

$\bar{x}$ is strictly local minimum of f .

$\rightarrow$ For Convex Optimization:

Let $f: S \rightarrow \mathbb{R}$ be a convex funtⁿ over a convex set S . Then

(i) A local minimum is also a global minimum.

(ii) Let S be an open set. Then, a necessary & sufficient condition for vector $\bar{x} \in S$ to be global minim^u is if $\nabla f(\bar{x}) = 0$.

NL/04/11/25

Unconstrained Optimization

$$x \in \mathbb{R}^n \quad f_0 : \mathbb{R}^n \rightarrow \mathbb{R}$$

(i)

$$\bar{x} \text{ is minimum} \Rightarrow \begin{cases} \nabla f(\bar{x}) = 0 & 1^{\text{st}} \text{ order necessary cond.} \\ \nabla^2 f(\bar{x}) \succeq 0 & 2^{\text{nd}} \text{ order necessary cond.} \end{cases}$$

(ii)

Let $\bar{x}$ satisfies

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

$$\text{2 } \nabla^2 f(\bar{x}) \succeq 0$$

$\Rightarrow \bar{x}$ is strictly

local minimum

Sufficient condition

• For convex Problems

$$f : S \rightarrow \mathbb{R}$$



is also
a convex function

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



convex set

$\bar{x}$ is global
minimum

$$\Leftrightarrow \nabla f(\bar{x}) = 0$$

$\bar{x}$ is local minimum if :

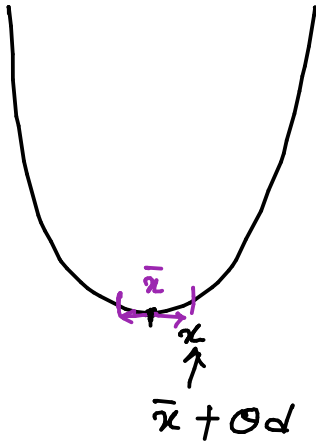
x can be written as $\bar{x} + \theta d$

$$f(x) \geq f(\bar{x}) \quad \forall x \in \mathcal{O}_B(\bar{x}, \delta)$$

local minimum

$$f(x) = f(\bar{x} + \theta d)$$

direct
small value
 $d \in \mathbb{R}^n$



Since $\bar{x}$ is local minimum
we have

$$f(\bar{x} + \theta d) \geq f(\bar{x})$$

$$\Rightarrow f(\bar{x} + \theta d) - f(\bar{x}) \geq 0$$

$$f'(\bar{x}, d) = \lim_{\theta \rightarrow 0} \frac{f(\bar{x} + \theta d) - f(\bar{x})}{\theta} \geq 0$$

↑
Directional derivative

$$f'(\bar{x}, d) = \nabla f(\bar{x})^T d$$

$\Downarrow$

(*)

$$\nabla f(\bar{x})^T d \geq 0 \quad \text{for } \forall d \in \mathbb{R}^n$$

The choice of d is free in $\mathbb{R}^n$

Since we are dealing with unconstrained optimization

- For unconstrained optimization,
the choice of 'd' is free
in $\mathbb{R}^n$.

$$\nabla f(\bar{x})^T d \geq 0 \quad \forall d \in \mathbb{R}^n$$

let us choose a 'direct' $d = -\nabla f(\bar{x})$

$$-\nabla f(\bar{x})^T \nabla f(\bar{x}) \geq 0$$

$$-\|\nabla f(\bar{x})\|_2^2 \geq 0 \Leftrightarrow \boxed{\nabla f(\bar{x}) = 0}$$

→ For Constrained Optimization:

$$\min f(x)$$

$$\text{s.t. } x \in S$$

↑
constraint set.

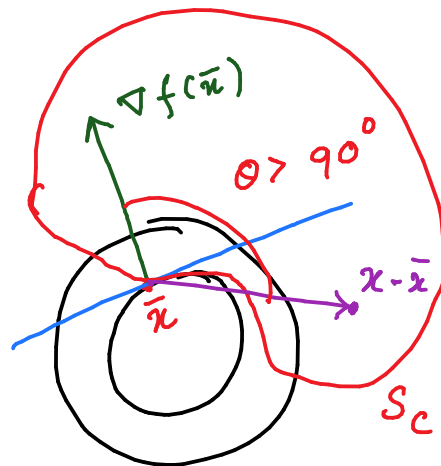
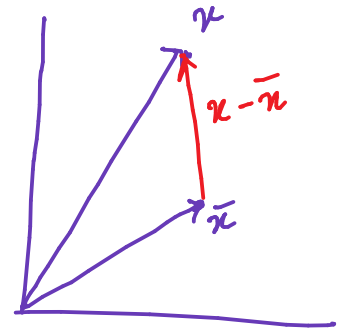
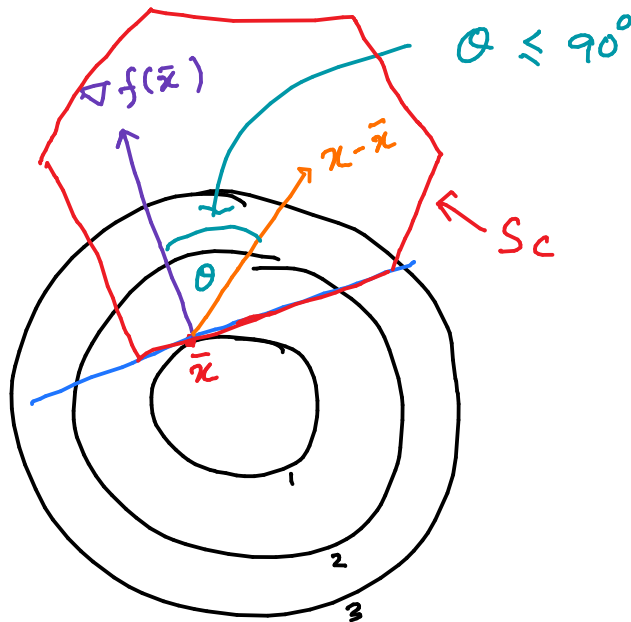
Optimality Condition

if $\bar{x}$ is local minimum of $\underline{f}$
over a convex set S_c , then

$$\boxed{\nabla f(\bar{x})^T (x - \bar{x}) \geq 0 \quad \forall x \in \underline{S_c}}$$

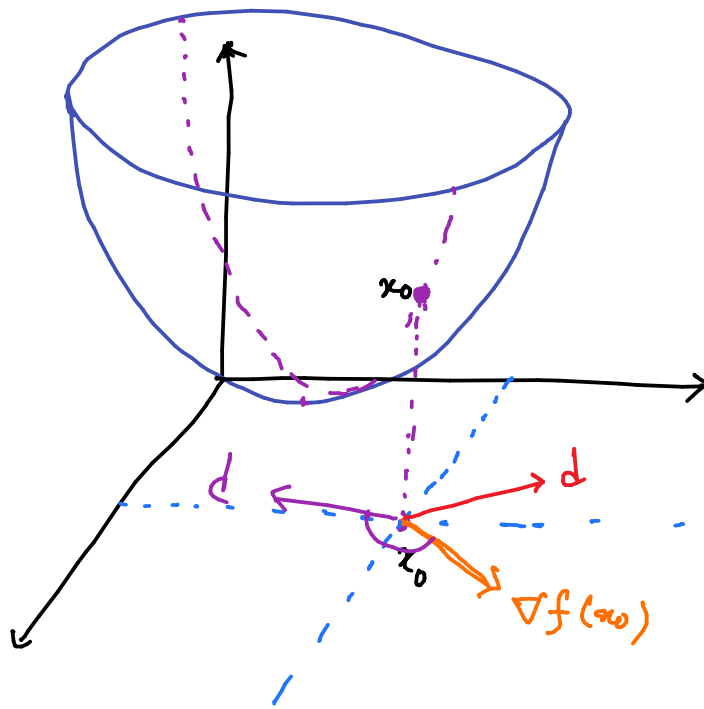
(ii) Assume that f is also a convex funt,
 then the following condⁿ is necessary
 & sufficient (S_c is convex) :

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



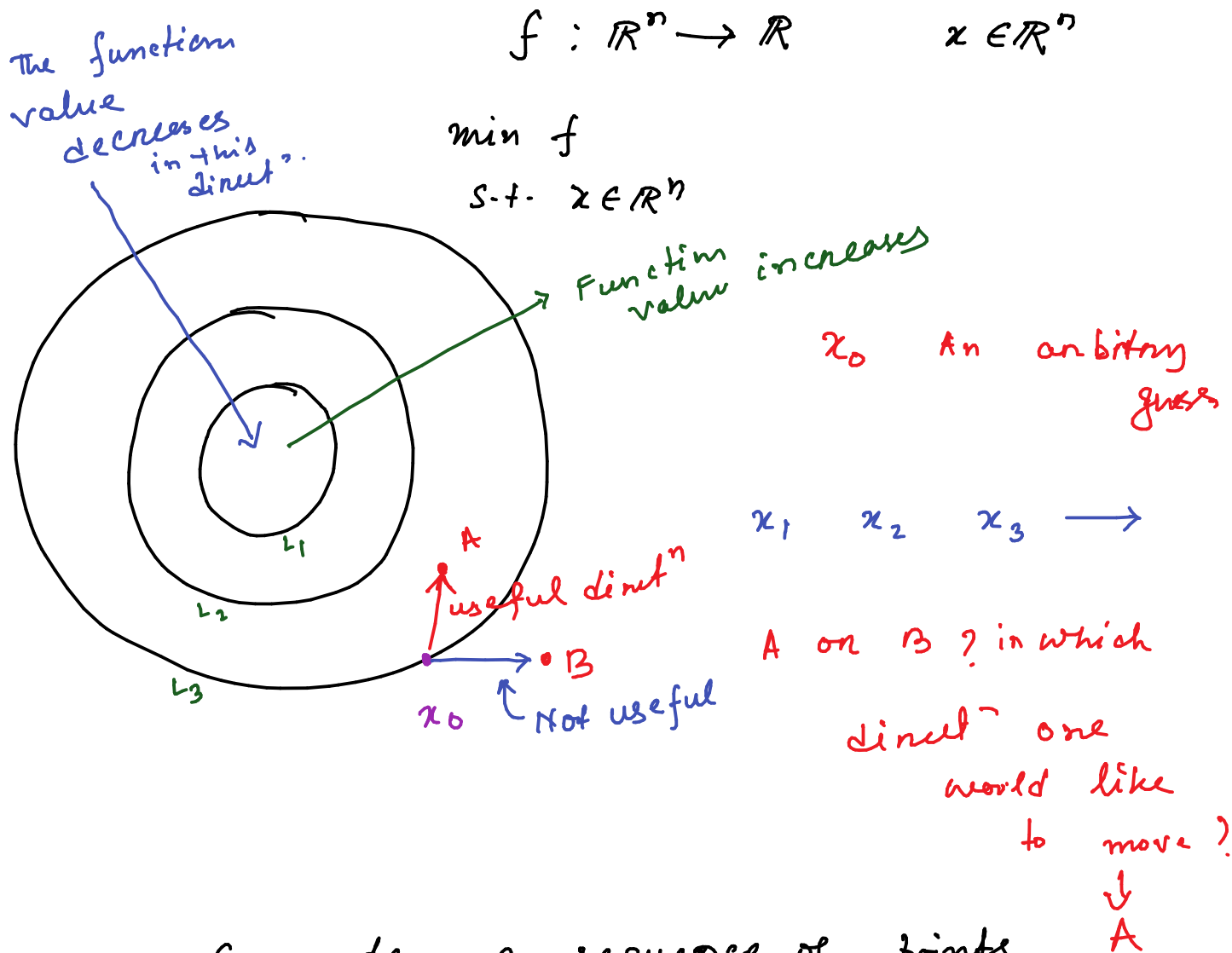
$\bar{x}$ is also called
 "stationary
 point."

→ Solving an Unconstrained Optimization



NL/06/11/25

- Search Algorithm for Unconstrained Optimization:



Generate a sequence of points

$k = 0, 1, 2 \dots$

$$x_{k+1} = x_k + \theta_k d_k$$

$$x_k \in \mathbb{R}^n$$

$$\theta \geq 0$$

$$d \in \mathbb{R}^n$$

$$x_1 = x_0 + \theta_0 d_0$$

$$\rightarrow x_0 \quad x_1 \quad x_2 \quad \dots$$

$$\rightarrow f(x_0) \quad f(x_1) \quad f(x_2) \quad \dots \quad f(x_k)$$

$$f(x_{k+1}) \leq f(x_k)$$

For minimization we need

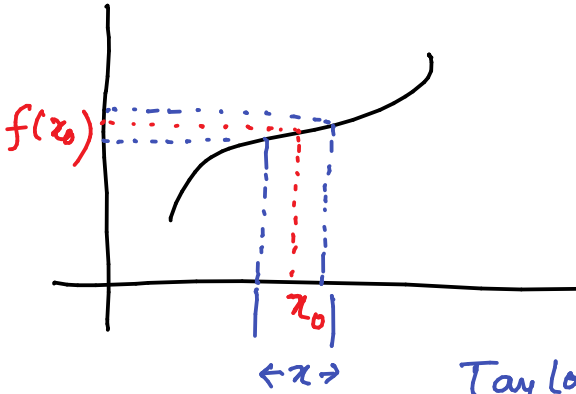
$$f(x_{k+1}) \leq f(x_k)$$

Verification
Stages

Reach

$$f(x_{k+1}) = f(x_k)$$

Minimum



Taylor series expansion of $f(x)$ at x_0 :

$$f(x) = f(x_0) + f'(x)|_{x_0} (x - x_0) + f''(x)|_{x_0} \frac{(x - x_0)^2}{2!} + \dots$$

$$f: \mathbb{R}^n \rightarrow \mathbb{R}$$

$$f(x) = f(x_0) + \nabla f(x_0)^T (x - x_0) + \text{higher order terms}$$

If $x - x_0$ is very small

H.O.T. can be neglected

$$f(x) \approx f(x_0) + \nabla f(x_0)^T (x - x_0)$$

$$x_{k+1} = x_k + \theta_k d_k$$

Starting point is x_0

$$f(x_1) = f(x_0 + \theta_0 d_0) \quad \leftarrow f(x)$$

$$\approx f(x_0) + \nabla f(x_0)^T (x_0 + \theta_0 d_0 - x_0)$$

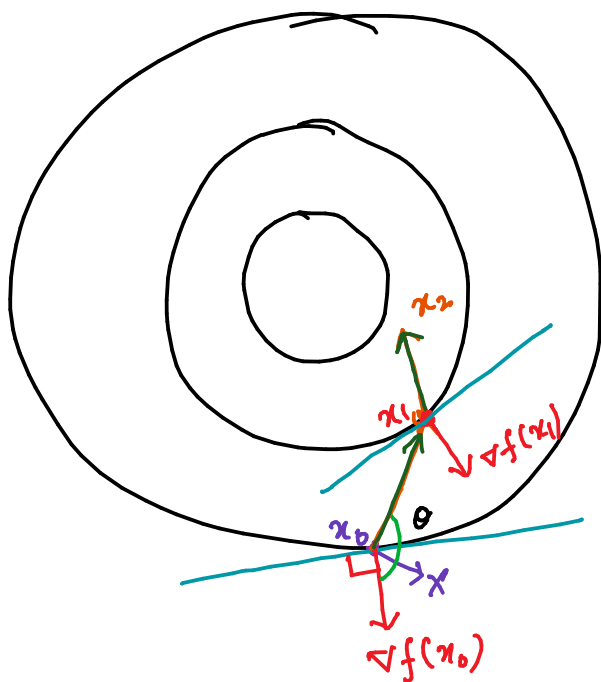
$$\approx f(x_0) + \nabla f(x_0)^T \theta_0 d_0$$

Assumed
 $\theta_0 \geq 0$

$$\approx f(x_0) + \theta_0 \underbrace{\nabla f(x_0)^T d_0}_{< 0}$$

To ensure $f(x_1) \leq f(x_0)$

$$\text{we need } \nabla f(x_0)^T d_0 \leq 0$$



For $\theta \geq 90^\circ$ } The gradient vector $\nabla f(x_0)$
 $\nabla f(x_0)^T d_0 \leq 0$ } $\exists d_0$ must form obtuse angle

Construct a sequence of points

$$x_{k+1} = x_k + \theta_k d_k$$

s.t. d_k must form

obtuse angle with } $\nabla f(x_k) d_k \leq 0$
 gradient vector

Free variables : θ_k $\leftarrow$ step size

Learning rate $\leftarrow \theta_k \leftarrow$ small, 'how small'?

$d_k \leftarrow \nabla f(x_k)^T d_k \leq 0$

d_k $\leftarrow$ step direct

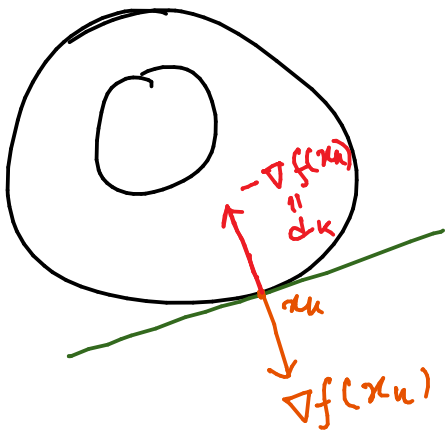
not unique

So which direct to choose?

The above Algorithm is called "Gradient descent" algorithm.

One particular choice of

$$d_k \text{ is } d_k = -\nabla f(x_k)$$



$$\begin{aligned} f(x_1) &\approx f(x_0) - \theta_0 \nabla f(x_0)^T \nabla f(x_0) \\ &= f(x_0) - \theta_0 \|\nabla f(x_0)\|_2^2 \end{aligned}$$

For this choice of d_k

the algorithm is called

"Steepest descent" Algorithm.