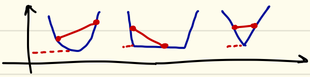
Convex Optimization :

* Convex function : $f : \mathbb{R}^n \longrightarrow \mathbb{R}$, $x \longmapsto f(x)$ is a convex function iff $\forall x, y \in \mathbb{R}^n$, $\forall \alpha \in [0,1]$: $f(\alpha x + (1-\alpha) y) \leq \alpha f(x) + (1-\alpha) f(y)$

Another way to check convexity (if f is twice differentiable) is to check if $\nabla^2 f = \frac{\partial^2 f(x)}{\partial x^2} \succcurlyeq 0 \quad \forall x$ (PSD)

→ Convex functions have a unique global minima (min value of f)



The line connecting any two points in domain of x is on or above f.

However, there may be multiple points x that achieve the global min value (all equally good)

Ex) Any norm is convex : $f(x) = \|x\|_2$: $\forall x, y$, $\forall \alpha \in [0,1]$:

$$f(\alpha x + (1-\alpha) y) \underset{\text{triang. eq.}}{\leq} f(\alpha x) + f((1-\alpha) y) \underset{\text{Homogeneity}}{=} |\alpha| f(x) + |1-\alpha| f(y) \overset{\alpha \geq 0}{=} \alpha f(x) + (1-\alpha) f(y) \checkmark$$

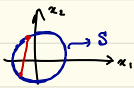
Ex) $f(x) = \frac{1}{2} \| Ax - y \|_2^2 \longrightarrow \frac{\partial f}{\partial x} = A^T (Ax - y) \longrightarrow \frac{\partial^2 f}{\partial x^2} = \nabla^2 f = A^T A \succcurlyeq 0$ PSD ✓

* Convex Set : A set $S \subseteq \mathbb{R}^n$ is convex iff $\forall x, y \in S$, $\forall \alpha \in [0,1]$ we have $\alpha x + (1-\alpha) y \in S$.

S : convex $\qquad$ S : nonconvex

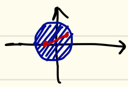
For any two point in the set, the line connecting them must be fully in the set.

Ex) $S = \{ x : \|x\|_2 = 1 \}$



non-convex set

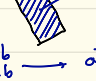
Ex) $S = \{ x : \|x\|_2 \leq 1 \}$



Convex set

Similarly for any norm $f(x)$: $S = \{ x : f(x) \leq c \}$ ($c \geq 0$) is convex and $S = \{ x : f(x) = c \}$ ($c \geq 0$) is nonconvex.

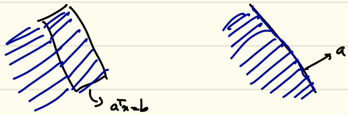
Ex) Given $a \in \mathbb{R}^n$, $b \in \mathbb{R}$: $S = \{ x : a^T x = b \}$



Convex

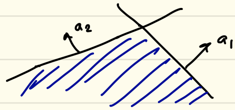
$a^T x_1 = b$, $a^T x_2 = b \longrightarrow a^T (x_1 - x_2) = 0$

Ex) $S = \{x : a^T x \leq b\}$: convex



Ex) $S = \{x : \underset{m \times n}{A}\, \underset{n \times 1}{x} \leq \underset{m \times 1}{d}\}$

$$\begin{pmatrix} a_1^T x \leq d_1 \\ \vdots \\ a_m^T x \leq d_m \end{pmatrix}$$



→ intersection of convex sets is always convex.

→ How about the union of convex sets ?

$S_1 \cap S_2$: convex $\quad$ $S_1 \cup S_2$: non convex

* Convex Optimization :

→ minimizing a convex function wrt $x \in \mathbb{R}^n$ when set of constraints on x are also convex.

$$\begin{cases} \min f(x) \rightsquigarrow \text{convex function} \\ \text{s.t. } h(x) = 0 \rightsquigarrow \text{convex set} \\ \quad\;\; g(x) \leq 0 \rightsquigarrow \text{convex set} \end{cases}$$

There are fast methods to solve this in polynomial time
CVX , Grad. descent , Projected Grad. descent .

* If we were solving $\min_x f(x)$ → $\left.\frac{\partial f}{\partial x}\right|_{x^*} = 0$ → to obtain minimizer

* How about $\begin{cases} \min f(x) \\ \text{s.t. } h_i(x) = 0 \quad i = 1, \dots, p \end{cases}$? ⟹ Using Lagrangian function

$$L(x, \alpha_1, \dots, \alpha_p) = f(x) + \sum_{i=1}^{p} \alpha_i h_i(x)$$

To obtain optimal solution : solve $\begin{cases} \left.\frac{\partial L}{\partial x}\right|_{(x^*, \alpha^*)} = 0 \\ \left.\frac{\partial L}{\partial \alpha_i}\right|_{(x^*, \alpha^*)} = 0 \quad \forall i = 1, \dots, p \end{cases}$ ⟹ $\begin{cases} \frac{\partial f(x^*)}{\partial x} + \sum_i \alpha_i^* \frac{\partial h_i(x^*)}{\partial x} = 0 \\ h_i(x^*) = 0 \, , \; \forall i = 1, \dots, p \end{cases}$

Ex) $\begin{cases} \min \ \frac{1}{2}\|x\|_2^2 & \longrightarrow \text{Convex Function} \\ \text{s.t.} \ \ a^T x = b & \longrightarrow \text{convex constraint} \end{cases}$

$$L(x,\alpha) = \frac{1}{2}\|x\|_2^2 + \alpha(a^T x - b)$$

$$\left.\frac{\partial L}{\partial x}\right|_{(x^*,\alpha^*)} = x^* + \alpha^* a = 0$$

$$\left.\frac{\partial L}{\partial \alpha}\right|_{(x^*,\alpha^*)} = a^T x^* - b = 0$$

$$\hookrightarrow \begin{cases} x^* + \alpha^* a = 0 \\ a^T x^* - b = 0 \end{cases} \longrightarrow (x^*, \alpha^*)\,?$$

$$x^* = -\alpha^* a \rightsquigarrow a^T(-\alpha^* a) - b = -\alpha^* a^T a - b = 0 \longrightarrow \alpha^* = -\frac{b}{\|a\|_2^2}$$

$$\longrightarrow x^* = -\alpha^* a = \boxed{\frac{b}{\|a\|_2^2} a}$$

$$\begin{cases} f(x^*) = \frac{1}{2}\|x^*\|_2^2 = \frac{b^2}{\|a\|^4}\|a\|^2 = \frac{b^2}{\|a\|_2^2} = \left(\frac{b}{\|a\|_2}\right)^2 \\ x^* = \frac{b}{\|a\|_2^2} a \end{cases}$$

How about the case where we have inequality constraints:

$$\begin{cases} \min \ f(x) \\ \text{s.t.} \ \ h_i(x) = 0, \ i = 1, \dots, p \\ \qquad g_j(x) \leq 0, \ j = 1, \dots, m \end{cases}$$

* Generalized Lagrangian Function: $L(x, \vec{\alpha}, \vec{\beta}) = f(x) + \sum_{i=1}^{p} \alpha_i h_i(x) + \sum_{j=1}^{m} \beta_j g_j(x)$; $\beta_j \geq 0, \forall j$; $\alpha_i \in \mathbb{R}, \forall i$