Thursday: 6pm - 7pm
↳ extra class

LECTURE - 8 : 15th Jan. 2025

Some Properties of Convex Functions

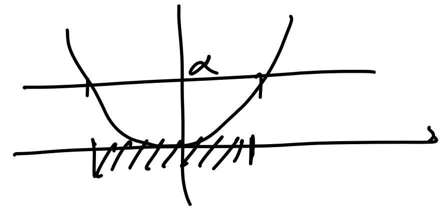
The following properties are true for convex functions.

- If $f : \mathbb{R}^n \to \mathbb{R}$ is a convex function, then it is continuous over the interior of $\texttt{dom}(f)$. Moreover, f is Lipschitz over every compact subset of the interior of $\texttt{dom}(f)$.

- If f and g are strictly convex functions, then $f + g$ is strictly convex as well.

- If f is a strongly convex function and g is a convex function, then $f + g$ is strongly convex as well.

Level-set Characterization



Definition 15. *For any $\alpha \in \mathbb{R}$, the level set of function $f : \mathbb{R}^n \to \bar{\mathbb{R}}$ at level α is defined as*

$$\underline{\underline{\mathtt{lev}_\alpha(f)}} := \{x \in \mathtt{dom}(f) | f(x) \le \alpha\}.$$

To establish Proposition 15, pick $x, y \in \mathrm{lev}_\alpha(f)$.

$\Rightarrow f(x) \le \alpha,\ f(y) \le \alpha$

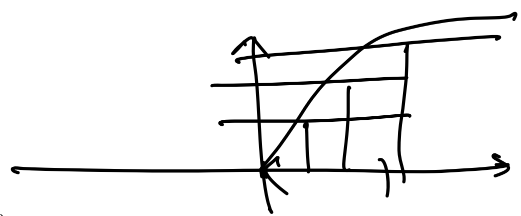
we need to show $\lambda x + (1-\lambda) y \in \mathrm{lev}_\alpha(f)$ when $\lambda \in [0,1]$.

Proposition 15. *If a function f is a convex function, then* **every** *level set of f is a convex set.*

we evaluate $\underline{f(\lambda x + (1-\lambda) y)} \le \lambda f(x) + (1-\lambda) f(y)$

$\le \lambda \alpha + (1-\lambda)\alpha = \alpha$

Therefore $\mathrm{lev}_\alpha(f)$ is a convex set.



In other words, if we can find some α for which $\mathtt{lev}_\alpha(f)$ is not a convex set, the function f is not a convex function.

[Ex: $f(x) = \sqrt{x}$, over $x > 0$.]

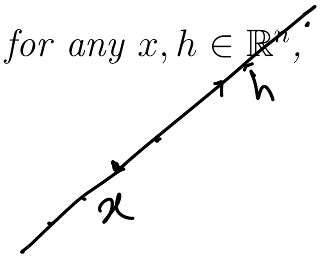
Converse is not true. A function is called quasi-convex if its domain and all level sets are convex sets.

HW: Give an example of a function which is quasi-convex but not convex.

Restriction of a Convex Function on a Line

Proposition 16. *If a function f is convex if and only if for any $x, h \in \mathbb{R}^n$, the function $\phi(t) = f(x + th)$ is a convex function on $\mathbb{R}$.*

s.t. $x + th \in \mathrm{dom}(f)$

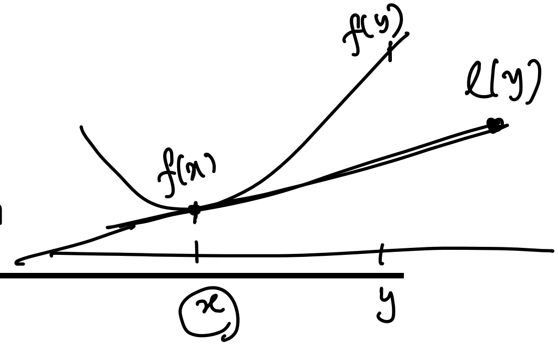


If we know how to check convexity of functions defined on $\mathbb{R}$, then we can check convexity of functions defined on $\mathbb{R}^n$.

$\ell(y) := f(x) + \nabla f(x)^T (y-x)$,

$$\begin{cases} \ell(x) = f(x) \\ \nabla \ell(y) = \nabla f(x) \ \forall y. \end{cases}$$

First Order Condition



Proposition 17. *If a function f is differentiable, then it is convex if and only if* $\mathtt{dom}(f)$ *is a convex set and for any $x, y \in \mathtt{dom}(f)$, we have*

$$f(y) \geq f(x) + \nabla f(x)^\top (y-x). = \ell(y) \qquad : \text{FOC}$$

A global lower bound on the function can be obtained at any point based on local information $(f(x), \nabla f(x))$.

part 1: Suppose FOC is satisfied for f. Show that f is a convex function.

Let $\bar{x}, \bar{y} \in \mathrm{dom}(f)$, $\lambda \in [0,1]$, $\bar{z} = \lambda \bar{x} + (1-\lambda)\bar{y} \in \mathrm{dom}(f)$

$f(\bar{z}) \leq \lambda f(\bar{x}) + (1-\lambda) f(\bar{y})$

From the FOC, we have

$$[f(\bar{x}) \geq f(\bar{z}) + \nabla f(\bar{z})^T (\bar{x} - \bar{z})] \times \lambda$$

$$(+)\ [f(\bar{y}) \geq f(\bar{z}) + \nabla f(\bar{z})^T (\bar{y} - \bar{z})] \times (1-\lambda).$$

$$\lambda f(\bar{x}) + (1-\lambda) f(\bar{y}) \geq (\lambda f(\bar{z}) + (1-\lambda) f(\bar{z})) + \nabla f(\bar{z})^T \times [\lambda \bar{x} - \lambda \bar{z} + (1-\lambda)\bar{y} - (1-\lambda)\bar{z}]$$

$$= f(\bar{z}) + \nabla f(\bar{z})^T [\underbrace{\lambda \bar{x} - \cancel{\lambda \bar{z}} + (1-\lambda)\bar{y} - \bar{z} + \cancel{\lambda \bar{z}}}_{=0})$$

$\Rightarrow f$ is a convex function.

part 2: Let f be a convex function

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

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

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

$$\Rightarrow \lim_{\lambda \to 0} (\text{LHS}) = \nabla f(y)^T (x-y) \leq f(x) - f(y)$$

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

Second Order Condition

Proposition 18. *If a function f is twice differentiable, then it is convex if and only if* $\texttt{dom}(f)$ *is a convex set and $\nabla^2 f(y) \succeq 0$ for every $y \in$* $\texttt{dom}(f)$.

(positive -semidefinite)

$\nabla^2 f(y) - mI \succeq 0$

- The function f is strongly convex if and only if $\nabla^2 f(y) \succeq mI$ for some $m > 0$ for every $y \in \texttt{dom}(f)$. Here I is the identity matrix of appropriate dimension.
- If $\nabla^2 f(y) \succ 0$ for every $y \in \texttt{dom}(f)$, then the function is strictly convex. The converse is not true.
- f is concave if and only if $\nabla^2 f(y) \preceq 0$ for every $y \in \texttt{dom}(f)$.

$f(x) = -\log(x)$, $f'(x) = \frac{-1}{x}$, $f''(x) = \frac{1}{x^2} > 0$

$\{x > 0\}$

Example: $f(x) = x^2, -\log(x), ||Ax - b||^2$, and so on.

$f''(x) = 2 > 0 \Rightarrow x^2$ is a convex function

$v^T A^T A v = ||Av||_2^2 \geq 0$

$f_3(x) = ||Ax - b||_2^2$, $\nabla^2 f_3(x) = 2A^T A$

↪ always positive semidefinite.

↪ always convex

↪ for strict convexity,

A must have full column rank $\Leftrightarrow$ nullspace$(A) = \emptyset$.

Examples of Convex Functions

$x \in \mathbb{R}^n, \; y \in \mathbb{R}$ $\qquad \nabla^2 f_1 \in \mathbb{R}^{(n+1)\times(n+1)}$

- $f_1(x, y) = \frac{x^\top x}{y}$ if $y > 0$ and $+\infty$ if $y \le 0$ (square to linear function).
- $f_2(x) = \log(\sum_{i=1}^n e^{x_i})$ (log-sum-exp function).
- $f_3(X) = -\log \det(X)$ where $X \in \mathbb{S}_n^{++}$ (log-determinant function).

$$\nabla f_1(x,y) = \begin{bmatrix} \frac{2x}{y} \\ -\frac{x^T x}{y^2} \end{bmatrix}, \quad \underbrace{[\,v_1^T \;\vdots\; v_2^T\,] \; \nabla^2 f_1(x,y) = \begin{bmatrix} \frac{2}{y} I & -\frac{2}{y^2} x \\ -\frac{1}{y^2} 2x^T & \frac{2}{y^3} x^T x \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}}$$

$v = [\,v_1 \;\vdots\; v_2\,], \quad v_1 \in \mathbb{R}^n, \; v_2 \in \mathbb{R}$

$$v^T \nabla^2 f_1(x,y) v = [\,v_1^T \;\vdots\; v_2^T\,] \begin{bmatrix} \frac{2}{y} I v_1 - \frac{2}{y^2} x v_2 \\ -\frac{1}{y^2} 2x^T v_1 + \frac{2}{y^3} x^T x v_2 \end{bmatrix} = \frac{2}{y} v_1^T v_1 - \frac{2}{y^2} v_1^T x v_2 - \frac{2}{y^2} v_2^T x^T v_1 + \frac{2}{y^3} x^T x v_2^T v_2$$

$$= \frac{2}{y} v_1^T v_1 - \frac{4}{y^2} v_2 v_1^T x + \frac{2}{y^3} (v_2)^2 x^T x$$

$$= \frac{2}{y^3} \left[y^2 v_1^T v_1 + (v_2)^2 x^T x - 2y v_2 v_1^T x \right]$$

$$= \frac{2}{y^3} \left[\| y v_1 - v_2 x \|_2^2 \right] \geq 0 .$$

Hence f_1 is a convex function.

Examples of Convex Functions

$f(X) = -\log |X|$, $X \in S^n_{++}$, let $X = X^{1/2} X^{1/2}$, $\underline{X^{1/2} \succ 0}$

$\underline{g(t)} = f(\underline{X + tV})$, $t \in \mathbb{R}$, $V \in S^n$, $\underline{X + tV} \in S^n_{++}$

$= -\log |X + tV|$

$= -\log | X^{1/2} (I + t X^{-1/2} V X^{-1/2}) X^{1/2} |$

$= -\log \left[\underline{|X^{1/2}|} \; \underline{|I + t X^{-1/2} V X^{-1/2}|} \; \underline{|X^{1/2}|} \right]$

$= -\log \left[|X| \cdot |I + t X^{-1/2} V X^{-1/2}| \right]$

$= -\log |X| - \log \left[|\underbrace{I + t X^{-1/2} V X^{-1/2}}_{\text{positive definite matrix.}}| \right]$.

$\text{eig}\left(I + t \underline{X^{-1/2} V X^{-1/2}}\right) = 1 + t\lambda_i$,

where λ_i are eigenvalues of $X^{-1/2} V X^{-1/2}$.

$= -\log |X| - \log \left(\prod_{i=1}^{n} (1 + t\lambda_i) \right)$

$\underline{g(t) = -\log |X| + \sum_{i=1}^{n} \underline{-\log (1 + t\lambda_i)}}$.

$g_i(t) = -\underline{\log(1 + tc)}$

$g_i'(t) = \dfrac{-1}{1 + tc} \cdot c$

$g_i''(t) = \dfrac{1}{(1 + tc)^2} \cdot c^2 > 0$

$\Rightarrow g(t)$ is a convex function

$\Rightarrow f(X) = -\log |X|$ is a convex function.

$\text{dom } g = \bigcap_{i \in I} \text{dom}(f_i)$. Since $\text{dom}(f_i)$ is a convex set for all i, and intersection of convex sets is a convex set, dom g is also a convex set.

Convexity Preserving Operations

Ex: $x^2 + 5e^x + (\quad) + (\quad)$

Proposition 19 (Conic Combination). *Let $\{f_i(x)\}_{i \in I}$ be a collection of convex functions and let $\alpha_i \geq 0$ for all $i \in I$. Then, $g(x) := \sum_{i \in I} \alpha_i f_i(x)$ is a convex function.*

$$g(\lambda x_1 + (1-\lambda)x_2) = \sum_{i \in I} \alpha_i f_i(\lambda x_1 + (1-\lambda)x_2)$$

$$\leq \sum_{i \in I} \alpha_i \left(\lambda f_i(x_1) + (1-\lambda) f_i(x_2)\right)$$

$$= \lambda g(x_1) + (1-\lambda) g(x_2) \quad = \lambda \sum_{i \in I} \alpha_i f_i(x_1) + (1-\lambda) \sum_{i \in I} \alpha_i f_i(x_2)$$

Proposition 20 (Affine Composition). *If $f : \mathbb{R}^m \to \mathbb{R}$ is a convex function, then $g(x) := f(Ax + b)$ is also a convex function where $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$.*

Ex: $e^x \rightarrow e^{a^\top x + b}$

$\|x\|_2^2 \rightarrow \|Ax+b\|_2^2$

$x \in \text{dom } g \iff Ax + b \in \text{dom } f$

dom g is inverse affine map of dom f, hence convex.

$$g(\lambda x_1 + (1-\lambda) x_2)$$

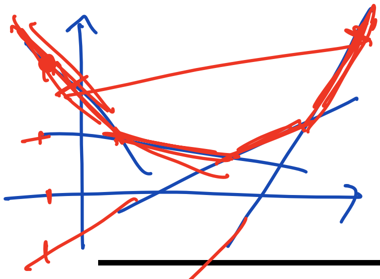
$$= f(A(\lambda x_1 + (1-\lambda) x_2) + b)$$

$$= f(\lambda(Ax_1 + b) + (1-\lambda)(Ax_2 + b))$$

$$\leq \lambda f(Ax_1 + b) + (1-\lambda) f(Ax_2 + b)$$

$$= \lambda g(x_1) + (1-\lambda) g(x_2).$$

Example: $g(x) = ||Ax + b||^2, h(x) = -\sum_{i=1}^n \log(a_i^\top x + b)$.



Convexity Preserving Operations

Proposition 21 (Pointwise Maximum). *Let $\{f_i(x)\}_{i\in I}$ is a collection of convex functions, then $g(x) := \max_{i\in I} f_i(x)$ is a convex function.*

The set I need not be a finite set.

$\text{epi}\, g = \bigcap_{i\in I} \text{epi}\, f_i$

$(x,t) \in \text{epi}\, g \iff g(x) \le t$

$\iff \max_{i\in I} f_i(x) \le t$

$\iff f_i(x) \le t \quad \forall i$

$\iff (x,t) \in \text{epi}\, f_i \quad \forall i$

Since intersection of convex sets is convex, epi g is a convex set

$\Rightarrow$ g is a convex function

Ex: $\underline{e^{\max_{i\in\{1,\dots,n\}} (a_i^T x + b)}} \underline{+ \|Gx - h\|_2^2}$

Example: Largest singular value of a matrix X $\quad f : \mathbb{R}^{n\times n} \to \mathbb{R}$

$$f(X) = \sigma_{\max}(X) = \max_{v:||v||_2=1} ||Xv||_2.$$

Proposition 22 (Pointwise Supremum). *Let $f(x, \omega)$ is convex in x for any $\omega \in \Omega$, then $g(x) := \sup_{\omega\in\Omega} f(x, \omega)$ is convex in x.*

Proposition 23 (Partial Minimization). *If $f(x,y)$ is* (strictly) *convex in (x,y), and Y is a convex set, then $g(x) := \inf_{y\in Y} f(x,y)$ is a convex function.* (strictly)

Example: Schur Complement Lemma

Suppose we are given a matrix $S = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \in S^{(n+m)\times(n+m)}$, and $A \succeq 0$. Then

S is positive semidefinite if and only if $A - BC^{-1}B^T \succeq 0$.

or $S \succ 0 \iff C \succ 0$ & $C - B^T A^{-1} B \succ 0$.

$$S \succeq 0 \iff [x^T ; y^T] \begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \geq 0$$

$$\Rightarrow [x^T \; y^T] \begin{bmatrix} Ax + By \\ B^T x + Cy \end{bmatrix} \geq 0$$

$$\Rightarrow \underbrace{x^T A x + 2x^T B y + y^T C y}_{f(x,y)} \geq 0. \quad \forall (x,y)$$

$$f(x,y) = x^T A x + 2x^T B y + y^T C y$$

$$g(x) = \inf_{y \in \mathbb{R}^n} f(x,y)$$

$$\nabla_y f(x,y) = 2B^T x + 2Cy = 0$$

$$\Rightarrow y^* = -C^{-1}B^T x$$

$$g(x) = x^T A x + 2x^T B(-C^{-1}B^T x) + x^T B C^{-1} B^T x$$

$$= x^T \left[A - BC^{-1}B^T \right] x.$$

If $S \succ 0$, then $f(x,y)$ is strictly convex in both (x,y). Under partial minimization, $g(x)$ also needs to be strictly convex, $\Rightarrow A - BC^{-1}B^T \succ 0$.

Convexity Preserving Operations

Proposition 24 (Scalar Composition)**.** *Let $f(x) := h(g(x))$ where $g : \mathbb{R} \to \mathbb{R}$ and $h : \mathbb{R} \to \mathbb{R}$. Then, f is convex if*

- *g is convex and h is convex non-decreasing*
- *g is concave and h is convex non-increasing.*

f is concave if

- *g is convex and h is concave non-increasing*
- *g is concave and h is concave non-decreasing.*

The conditions do not necessarily hold in the reverse direction.

$$f'(x) = h'(g(x))\, g'(x)$$

$$f''(x) = \underbrace{h''(g(x))}_{+ve} \underbrace{(g'(x))^2}_{+ve} + \underbrace{h'(g(x))}_{\geq 0} \underbrace{g''(x)}_{+ve}$$

Let us prove case-3: we need to show

$$f(\lambda x + (1-\lambda) y) \geq \lambda f(x) + (1-\lambda) f(y) \quad \forall\, x, y \in \text{dom},\ \lambda \in [0,1].$$

LHS:
$$h(g(\lambda x + (1-\lambda) y)) \geq h[\lambda g(x) + (1-\lambda) g(y)]$$
$$\geq \lambda h(g(x)) + (1-\lambda) h(g(y))$$
$$= \lambda f(x) + (1-\lambda) f(y).$$

Convexity Preserving Operations

Proposition 25 (Vector Composition). *Let $\{g_i\}_{i \in \{1,2,\ldots k\}}$ are convex functions on $\mathbb{R}^n$, and $h : \mathbb{R}^k \to \mathbb{R}$ is convex and non-decreasing in each argument, then the function $f(x) = h(g(x))$ is convex.*

Other scalar composition rules can also be directly extended to the vector case.

$g_i : \mathbb{R}^n \to \mathbb{R}$

$h : \mathbb{R}^k \to \mathbb{R}$

$f : \mathbb{R}^n \to \mathbb{R}$.

$$h(g(x)) = h\left(\begin{bmatrix} g_1(x) \\ \vdots \\ g_k(x) \end{bmatrix}\right)$$

Examples:

- If g is convex, then $e^{g(x)}$ is also convex.
- If g is concave and positive, then $\frac{1}{g(x)}$ is convex.
- If g_i are convex, then $\log(\sum_{i=1}^{k} e^{g_i(x)})$ is convex.

LECTURE 10 : 16th Jan

Recall: Optimization Problem

An optimization problem can be stated (in abstract form) as

$$\min_{x \in X} f(x), \tag{2}$$

where

- x decision variable, often a vector in $\mathbb{R}^n$
- X set of feasible solutions, often a subset of $\mathbb{R}^n$
- $f : \mathbb{R}^n \to \mathbb{R}$ cost function

Goal:

- Find $x^* \in X$ that minimizes the cost function, i.e., $f(x^*) \leq f(x)$ for every $x \in X$.
- Optimal value: $f^* := \inf_{x \in X} f(x)$
- Optimal solution: $x^* \in X$ if $f(x^*) = f^*$.

Often, we write optimization problems in standard form as:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f(x) \\ \text{subject to} \quad & g_i(x) \leq 0, \qquad i \in \{1, 2, \ldots, m\} \\ & h_j(x) = 0, \qquad j \in \{1, 2, \ldots, p\}. \end{aligned}$$

$X = \{x \in \mathbb{R}^n \mid g_i(x) \leq 0,\ h_j(x) = 0,\ i \in \{1, \ldots m\},\ j \in \{1, 2, \ldots p\}\}$

Recall

- The problem is infeasible when X is an empty set. In this case, $f^* := +\infty$.

- The problem is unbounded when $f^* = -\infty$.

Definition 16. *Recall that*

- *a feasible solution $x^* \in X$ is a global optimum if $f(x^*) \leq f(x)$ for all $x \in X$. In this case, $f^* = f(x^*)$,*
- *the set of global optima:* $\operatorname{argmin}_{x \in X} f(x) := \{z \in X | f(z) = f^*\}$,
- *a feasible solution $x^* \in X$ is a local optimum if $f(x^*) \leq f(x)$ for all $x \in B(x^*, r)$ for some $r > 0$.*

Theorem: Weierstrass Theorem

If the cost function f is continuous and the feasible region X is compact (closed and bounded), then (at least one global) optimal solution x^* exists.

Feasibility Problem

Goal: Find $x \in \mathbb{R}^n$ which satisfies a collection of inequality and equality constraints.

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & 0 \\ \text{subject to} \quad & g_i(x) \leq 0, \qquad i \in \{1, 2, \ldots, m\} \\ & h_j(x) = 0, \qquad j \in \{1, 2, \ldots, p\}. \end{aligned}$$

$f^* = 0$ if a feasible solution exists. Otherwise, $f^* = +\infty$.

Convex Optimization Problems

An optimization problem in abstract form

$$\min_{x \in X} f(x), \tag{3}$$

is convex when the feasibility set X is a convex set and the cost function $f(x)$ is a convex function.

An optimization problem in standard form

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f(x) \\ \text{subject to} \quad & g_i(x) \leq 0, \qquad i \in \{1, 2, \ldots, m\} \\ & \underline{h_j(x) = 0,} \qquad j \in \{1, 2, \ldots, p\}, \end{aligned}$$

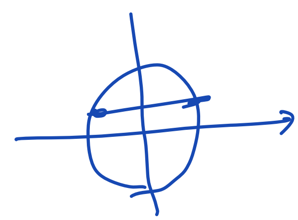
is convex when

- f and g_i are convex functions.
- h_j **are affine functions.**

$h_j(x) \leq 0$

$-h_j(x) \leq 0$

$\{x \in \mathbb{R}^2 | \underline{x_1^2 + x_2^2} = 1\}$ is not a convex set.



Simple Examples of Convex Functions

① $\min_{x \in \mathbb{R}} \; (5-x)^2$ s.t. $x \geq 0$ — convex,

$\max f(x) = \min (-f(x))$

② $\max_{x \in \mathbb{R}} \; \log x$ s.t. $x > 5$ — convex, as $-\log x$ is a convex function.

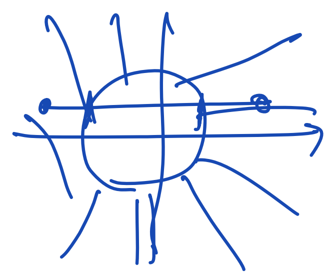
③ $\min_{x \in \mathbb{R}^n} \left[\max_{i \in \{1,2..K\}} \left[a_i^T x + b_i \right] \right] =: g(x)$. is a convex function. Hence convex opt. problem.

④ $\min_{x \in \mathbb{R}} \; -e^{-x}$ s.t. $1 < x < 5$ → not a convex opt. problem.

$f(x) = -e^{-x}$, $f'(x) = e^{-x}$, $f''(x) = -e^{-x} < 0$

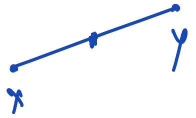
⑤ $\min_{x \in \mathbb{R}} \; x^2$ s.t. $\log(x) \leq 2$ → $\log(x) \leq 2 \Leftrightarrow \underline{x \leq e^2}$, convex.

⑥ $\min_{x \in \mathbb{R}^n} \; c^T x$ s.t. $\|x\|_2^2 \geq 5$ → not a convex opt. problem.



part 2 : let x, y be both optimal solutions. , $f(x) = f(y) = f^*$

$f(\lambda x + (1-\lambda) y) \leq \lambda f(x) + (1-\lambda) f(y) = f^*$



since f^* is the optimal value,

1. Local Optimum is Global

$f(\lambda x + (1-\lambda) y) = f^*$

Hence $\lambda x + (1-\lambda) y \in X_{opt}$

Theorem: Local and Global Optima

Consider the optimization problem $\min_{x \in X} f_0(x)$. If f_0 is a convex function and X is a convex set, then any locally optimal solution is also globally optimal. Moreover, the set of optimal solutions $X_{\text{opt}} := \operatorname{argmin}_{x \in X} f_0(x)$ is a convex set.

proof : Let x^* be a locally optimal solution.

we want to show $f_0(y) \geq f_0(x^*)$ $\forall y \in X$.

Since X is a convex set,

$\lambda x^* + (1-\lambda) y \in X$.

$$f_0(\lambda x^* + (1-\lambda) y) \leq \lambda f_0(x^*) + (1-\lambda) f(y)$$

$$\Rightarrow f_0(\lambda x^* + (1-\lambda) y) - f_0(x^*) \leq (1-\lambda) f(y) + \lambda f_0(x^*) - f_0(x^*)$$

$$= [1-\lambda] [f(y) - f_0(x^*)] \quad \forall \lambda.$$

Let x^* be locally optimal over $B(x^*, r)$.

then, we can choose $\bar{\lambda}$ sufficiently close to 1 s.t.

$$\bar{\lambda} x^* + (1-\bar{\lambda}) y \in B(x^*, r)$$

then, $$f_0(x^*) \leq f_0(\bar{\lambda} x^* + (1-\bar{\lambda}) y)$$

then $$0 \leq f_0(\bar{\lambda} x^* + (1-\bar{\lambda}) y) - f_0(x^*) \leq [1-\bar{\lambda}] [f_0(y) - f_0(x^*)]$$

$$\Rightarrow \boxed{f_0(x^*) \leq f_0(y)} \quad \forall y \in X.$$

2. Uniqueness under Strict Convexity

Theorem: Uniqueness under Strict Convexity

Consider the optimization problem $\min_{x \in X} f_0(x)$. If f_0 is a strictly convex function and X is a convex set, and $x^\star$ is an optimal solution to the problem, then, $x^\star$ is the unique optimal solution, i.e., $X_{\texttt{opt}} := \{x^\star\}$.

Suppose $x, y \in X_{opt}$, $x \neq y$.

$$f(\lambda x + (1-\lambda) y) < \lambda f(x) + (1-\lambda) f(y)$$

$$= \lambda f^* + (1-\lambda) f^* = f^* .$$

which violates the optimality of x and y.

Hence $X_{opt} = \{x^*\}$.

3. Necessary and Sufficient Optimality Condition

Theorem: Necessary and Sufficient Optimality Condition

Consider the optimization problem $\min_{x \in X} f_0(x)$ where f_0 is a convex and differentiable function, and X is a convex set. Then,

$$x^\star \quad \text{is optimal} \iff \underline{\nabla f_0(x^\star)^\top (y - x^\star) \geq 0, \qquad \forall y \in X.}$$

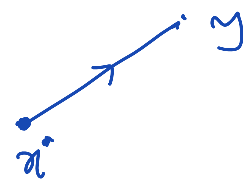
<u>Sufficiency</u>: Suppose x^* satisfies $\nabla f_0(x^*)^T (y - x^*) \geq 0 \quad \forall y \in X$.

<u>First order condition</u>:

$$f_0(y) \geq f_0(x^*) + \underbrace{\nabla f_0(x^*)^T (y - x^*)}_{\geq 0} \qquad \forall\, \underline{y, x^* \in X}$$

$$f_0(y) \geq f_0(x^*) \qquad \forall\, y \in X$$

<u>Necessity</u>: We are given that x^* is an optimal solution.



$$\phi(t) = f_0(x^* + t(y - x^*))$$

$$\phi'(t) = \nabla f_0(x^* + t(y - x^*))^T (y - x^*)$$

$$\phi'(t)\big|_{t=0} = \underline{\nabla f_0(x^*)^T (y - x^*)}.$$

Since x^* is an optimal solution, $\phi'(t)\big|_{t=0} \geq 0$. $\underline{\forall y \in X}$.

LECTURE 11 : 16th JAN.

Equivalent Optimization Problems

Consider the following two optimization problems:

$$\min_{x \in X} f(x). \tag{4}$$

$$\min_{y \in Y} g(y). \tag{5}$$

The above problems are equivalent if

- Given an optimal solution x^* of (4), we can find an optimal solution y^* of (5), and
- given an optimal solution y^* of (5), we can find an optimal solution x^* of (4).

Equivalence: Maximization

(A) $\min_{x \in X} \underline{f(x)}$ (B) $\max_{x \in X} \underline{-f(x)}$.

Let $x^* \in X_{opt}^{A}$.

i.e. $f(x^*) \leq f(x) \ \forall x \in X$

$\Rightarrow -f(x^*) \geq -f(x) \ \forall x \in X \Rightarrow x^*$ is the optimal solution of (B).

Equivalence: Epigraph Form

(A) $\min_{x \in X} f(x)$

(B) $\min_{x \in X, t \in \mathbb{R}} t$ s.t. $f(x) \le t$

Let $x^* \in X^A_{opt}$, then $(x^*, f(x^*))$ is optimal solution of B.

Proof by contradiction: Suppose we have $(\bar{x}, \bar{t}) \in X^B_{opt}$,

$\bar{x} \neq x^*$, $\bar{t} = f(x^*)$

$\bar{t} < f(x^*)$

Since $(\bar{x}, \bar{t})$ are feasible to (B), then

$\bar{x} \in X$

$f(\bar{x}) \le \bar{t}$

$\Rightarrow f(\bar{x}) \le \bar{t} < f(x^*)$

which contradicts optimality of x^* for (A).

Hence, $(x^*, f(x^*))$ is optimal solution of (B).

Equivalence: Slack Variables

(A) $\min_{x \in \mathbb{R}^n} f(x)$
s.t. $g_1(x) \le 0$
$g_2(x) \le 0$
$\vdots$
$g_K(x) \le 0$

(B) $\min_{\substack{x \in \mathbb{R}^n \\ s \in \mathbb{R}^K}} f(x)$
s.t. $g_1(x) + s_1 = 0$
$g_2(x) + s_2 = 0$
$\vdots$
$g_k(x) + s_K = 0$
$s_1, s_2 \ldots s_K \geq 0$.

Can we construct optimal sol^n of (B)?

If $\bar{x} \in X^A$, is $(\bar{x}, \bar{s}) \in X^B$?, yes: define $\bar{s}_i = \begin{cases} 0 & \text{if } g_i(\bar{x}) = 0 \\ -g_i(\bar{x}) & \text{if } g_i(\bar{x}) < 0 \end{cases}$

Similarly. If $(\hat{x}, \hat{s}) \in X^B$, $\hat{x} \in X^A$.

Hence $x^* \in X^A_{opt} \Rightarrow (x^*, s^*) \in X^B_{opt}$ & so on.

Equivalence: From Equality to Inequality Constraints

(A) $\min_{x} \quad f(x)$

s.t. $\left[\begin{array}{l} h_1(x)=0 \\ h_2(x)=0 \\ \vdots \\ h_p(x)=0 \end{array}\right.$

$\left[\begin{array}{l} h_1(x)\le 0 \\ h_1(x)\ge 0 \end{array}\right.$

$(=)$

(B) $\min_{x} \quad f(x)$

s.t. $h_i(x) \le 0$

$-h_i(x) \le 0$, $\forall\, i=1,2 \cdot p$

Equivalence: From Constrained to Unconstrained

(A) $\min_{x} f(x)$

s.t. $x \in X$

(B) $\min_{x \in \mathbb{R}^n} \underline{\underline{f(x) + I_X(x)}}$

<u>optimal solutions of both problems are identical.</u>

Equivalence: Scalar Multipliers and Constant Terms

(A) $\min_{x \in X} f(x)$

(B) $\min_{x \in X} af(x) + b,\quad a > 0,\ b \in \mathbb{R}$

optimal value is different, but optimal solution is unchanged.

(A) $\min_{x \in \mathbb{R}^n} f(x)$

s.t. $g_i(x) \le 0$

$h_j(x) = 0$

(B) $\min_{x \in \mathbb{R}^n} f(x)$

s.t. $\alpha_i g_i(x) \le 0$

$\beta_j h_j(x) = 0,$

where $\alpha_i > 0,\ \underline{\beta_j \in \mathbb{R} \setminus \{0\}}$

Both (A) & (B) have identical optimal solutions, optimal values, and feasibility sets.

Equivalence: Monotone Transformations

(A) $\min_{x \in \mathbb{R}^n} f(x)$

s.t. $g_i(x) \leq 0$.

(B) $\min_{x \in X} \phi(f(x))$,

s.t. $\phi_i(g_i(x)) \leq 0 \quad \forall i$

Both have the same optimal solution.

ϕ_i: strictly monotonically increasing function.

$$\min_{x \in \mathbb{R}^n} \alpha_0 x_1^{a_0^1} x_2^{a_0^2} \cdots x_n^{a_0^n}$$

s.t. $\alpha_j x_1^{a_j^1} x_2^{a_j^2} \cdots x_n^{a_j^n} \leq \beta_j$

$x_i > 0$

define $y_i = \log(x_i)$

$$\log\left(\alpha_0 x_1^{a_0^1} x_2^{a_0^2} \cdots x_n^{a_0^n}\right) = \log(\alpha_0) + a_0^1 \log(x_1) + \cdots + a_0^n \log(x_n)$$

$$= \log(\alpha_0) + a_0^1 y_1 + \cdots + a_0^n y_n$$

$$= \log(\alpha_0) + a_0^T y$$

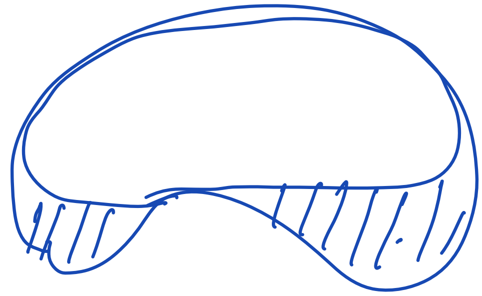
$\Rightarrow \log(\qquad\qquad) \leq \log(\beta_j)$

$\Rightarrow \log(\alpha_j) + (a_j)^T y \leq \log(\beta_j)$

Inner Approximation

$\min_{x \in X} f_0(x) : (A)$

$\min_{x \in X'} f_0(x) : (B).$



If $X' \subseteq X$, then we say (B) is an inner approximation of (A).

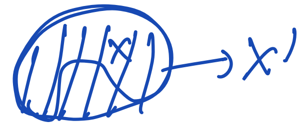
If $X' \subseteq X$, then $x \in X' \Rightarrow x \in X$

If $x^* \in X^B_{opt}$, $f(x^*) \leq f(x) \; \forall \; x \in X'$

But there may be points in $X^A \setminus X^B$ that have smaller optimal value that $f(x^*)$.

Hence $\boxed{f^*_B \geq f^*_A}$

Relaxation and Soft Constraints



$$\min_{x \in X} f(x) \quad (A)$$

$$\min_{x \in X'} f(x) \quad (B)$$

If $X' \supseteq X$, then we say (B) is a relaxation of (A).

In this case $f_B^* \leq f_A^*$.

$$\min_{x \in X} \; f(x) + \underline{\sum_{i=1}^{K} M_i \varepsilon_i^2}$$

$$\text{s.t.} \quad \left. \begin{aligned} & g_1(x) \leq \varepsilon_1, \\ & g_2(x) \leq \varepsilon_2 \\ & \vdots \\ & g_K(x) \leq \varepsilon_K \\ & \varepsilon_i \geq 0. \end{aligned} \right]$$

called soft constraints.