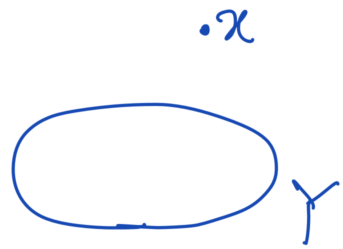
Convexity Preserving Operations - 6

Proposition 25 (Partial Minimization)**.** *If $f(x, y)$ is convex in (x, y), and Y is a convex set, then $g(x) := \inf_{y \in Y} f(x, y)$ is a convex function.*

$$\text{dist}(x, Y) = \inf_{y \in Y} \|x - y\|_2^2$$



convex function of x only when Y is a convex set.

HW For each of the convexity preserving operations disussed above, determine if these operations preserve concavity of functions.

Properties of concave functions: Let f be a concave function.

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

$$\Leftrightarrow \quad f(\lambda x + (1-\lambda)y) \geq \lambda f(x) + (1-\lambda) f(y)$$

$f(\sum_i \lambda_i x_i) \geq \sum_i \lambda_i f(x_i)$

(2) $f(y) \leq f(x) + \nabla f(x)^T (y - x)$

(3) $\nabla^2 f(y)$ is negative semidefinite.

Recall: Optimization Problem

An optimization problem can be stated 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 $\mathbb{R}^n$
  - often specified in terms of equality and inequality constraints $X := \big\{x \in \mathbb{R}^n | g_i(x) \leq 0, h_j(x) = 0, i \in \{1, 2, \ldots, m\}, j \in \{1, 2, \ldots, p\}\big\}$.
- $f : \mathbb{R}^n \to \mathbb{R}$ 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 $f(x^*) = f^*$.

Recall

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

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

Definition 15. *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: $\boxed{\operatorname{argmin}_{x \in X} f(x)} := \{z \in X | f(z) = f^*\}$,*
- *a feasible solution $x^* \in X$ is a local optimum if $\boxed{f(x^*) \leq f(x)}$ for all $\boxed{x \in B(x^*, r)}$ 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.

Abstract vs. Standard Form

An optimization problem can be stated in abstract form as

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

where $X := \left\{x \in \mathbb{R}^n | g_i(x) \leq 0, h_j(x) = 0, i \in \{1, 2, \ldots, m\}, j \in \{1, 2, \ldots, p\}\right\}$, or in "standard form" as

$$\neq \left[\begin{array}{lll} \min\limits_{x \in \mathbb{R}^n} & f(x) & \\ \text{subject to} & g_i(x) \leq 0, & i \in \{1, 2, \ldots, m\} \\ & h_j(x) = 0, & j \in \{1, 2, \ldots, p\}. \end{array}\right]$$

$$\left[\begin{array}{l} g_1(x) \leq 0 \\ g_2(x) \leq 0 \\ \vdots \\ g_m(x) \leq 0 \\ h_1(x) = 0 \\ h_2(x) = 0 \\ \vdots \\ h_p(x) = 0 \end{array}\right]$$

Feasibility Problem

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

$f(x) = 0$

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

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

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) $\min_{x \in X} g(x)$ (B) $\max_{x \in X} f(x)$

where $g(x) = -f(x)$

Suppose x^* is an optimal solⁿ of (A) $\Rightarrow g(x^*) \leq g(x) \ \forall x \in X$

We claim that x^* is optimal solⁿ of (B).

$$f(x^*) = -g(x^*) \geq -g(x) \ (\forall x \in X) \quad (\because x^* \text{ is optimal for (A)})$$
$$= f(x)$$

$$\Rightarrow \underline{f(x^*) \geq f(x) \quad \forall x \in X}$$

Now suppose $\bar{x}$ is optimal solⁿ of (B) $\Rightarrow f(\bar{x}) \geq f(x) \ \forall x \in X$

We claim that $\bar{x}$ is optimal solⁿ of (A).

$$g(\bar{x}) = -f(\bar{x}) \leq -f(x) = g(x) \quad \forall x \in X.$$

Equivalence: Epigraph Form

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

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

($\Rightarrow$) Let x^* be optimal for (A).
$\Rightarrow f(x^*) \le f(x) \;\forall x \in X$.

We claim (x^*, t^*) where $t^* = f(x^*)$ is optimal sol^n of (B).

Is (x^*, t^*) feasible for (B)? yes.

$f(x^*) \le t^*$ ✓
$x^* \in X$ ✓
$t^* \in \mathbb{R}$ ✓

Suppose (x^*, t^*) is not optimal for (B).

$\Rightarrow (\bar{x}, \bar{t})$ which is feasible for (B) & $\bar{t} < t^*$

$\Rightarrow$ $f(\bar{x}) \le \bar{t} < t^*$, $\bar{x} \in X$, $\bar{t} \in \mathbb{R}$ $\Longrightarrow$ $\exists\, \bar{x} \in X$ for which $f(\bar{x}) < f(x^*)$ which contradicts optimality of x^* in (A).

$\Rightarrow (x^*, t^*)$ is indeed optimal for (B).

Reverse direction

Suppose $(\hat{x}, \hat{t})$ is optimal sol^n of (B).

$\Rightarrow f(\hat{x}) \le \hat{t}$, $\hat{x} \in X$, $\hat{t} \in \mathbb{R}$

(+) $f(\hat{x}) = \hat{t}$, otherwise $(\hat{x}, \hat{t})$ is not optimal for (B).

Claim: $\hat{x}$ is optimal sol^n of (A).

Suppose not, & $\exists\, x'' \in X$ s.t. $f(x'') < f(\hat{x}) = \hat{t}$

then $(x'', f(x''))$ violates optimality of $(\hat{x}, \hat{t})$ is (B).

Equivalence: Slack Variables

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

s.t. $g_1(x) \leq 0$

$g_2(x) \leq 0$

$\vdots$

$g_K(x) \leq 0$

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

s.t. $g_1(x) + s_1 = 0$

$g_2(x) + s_2 = 0$

$\vdots$

$g_K(x) + s_K = 0$.

$s_1, s_2 \ldots s_K \geq 0$.

Let x^* be the optimal sol^n of (A).

Then (x^*, s^*), where $s_i^* = -g_i(x^*)$ is optimal sol^n of (B).

≥ 0

s_i's are called slack variables

Equivalence: From Equality to Inequality Constraints

(A)

$$\min_{x \in \mathbb{R}^n} f(x)$$
$$\text{s.t.} \quad \left.\begin{array}{l} h_1(x) = 0 \\ h_2(x) = 0 \\ \vdots \\ h_k(x) = 0 \end{array}\right] X_A$$

$$\min_{x \in \mathbb{R}^n} f(x)$$
$$\text{s.t.} \quad \left.\begin{array}{l} h_1(x) \leq 0 \\ -h_1(x) \leq 0 \\ h_2(x) \leq 0 \\ -h_2(x) \leq 0 \\ \vdots \\ h_k(x) \leq 0 \\ -h_k(x) \leq 0 \end{array}\right] X_B$$

(B)

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

It suffices to show that $X_A = X_B$ because cost function is the same in both (A) & (B) & the decision variable resides in the same space.

$$h_i(x) = 0 \iff \begin{array}{l} h_i(x) \leq 0 \\ h_i(x) \geq 0 \end{array} \iff \begin{array}{l} -h_i(x) \leq 0 \\ h_i(x) \leq 0 \end{array}$$

Equivalence: From Constrained to Unconstrained

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

I_X: indicator function of the set X

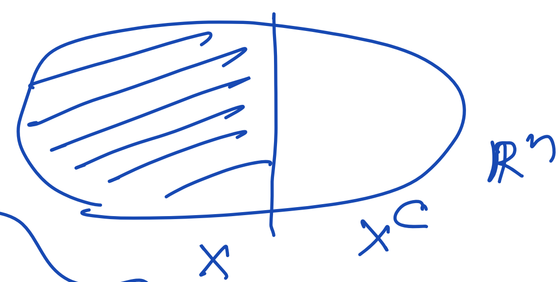
$$I_X(x) = \begin{cases} 0, & x \in X \\ \infty, & x \notin X \end{cases}$$

Let x^* be an optimal solution of (A).

$\Rightarrow f(x^*) \le f(x) \quad \forall x \in X.$

Claim: x^* is optimal sol^n of (B).

$f(x^*) \le f(x) \quad \forall x \in X$



when $x \in X^c$, $f(x) + I_X(x) = \infty$

$$\Rightarrow \boxed{f(x^*) + I_X(x^*) \le f(x) + I_X(x) \quad \forall x \in \mathbb{R}^n}$$

When $X = \emptyset$, neither problem has an optimal solution and $f^* = \infty$ for both (A) & (B).

Equivalence: Scalar Multipliers and Constant Terms

(A) $\min_{x \in X} f(x) \iff \min_{x \in X} af(x) + b, \; a > 0$

If x^* is optimal for (A), then

$$f(x^*) \le f(x) \quad \forall x \in X$$

$$\iff af(x^*) + b \le af(x) + b \quad \forall a > 0, \; b \in \mathbb{R}$$

$\iff x^*$ is optimal for (B).

(A)
$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f(x) \\ \text{s.t.} \quad & g_1(x) \le 0 \\ & g_2(x) \le 0 \\ & \vdots \\ & g_K(x) \le 0 \\ & h_1(x) = 0 \\ & \vdots \\ & h_m(x) = 0 \end{aligned} \iff \begin{aligned} \min_{x \in \mathbb{R}^n} \quad & \alpha_0 f(x) \\ \text{s.t.} \quad & \alpha_1 g_1(x) \le 0 \\ & \alpha_2 g_2(x) \le 0 \\ & \vdots \\ & \alpha_K g_K(x) \le 0 \\ & \beta_1 h_1(x) = 0 \\ & \vdots \\ & \beta_m h_m(x) = 0 \end{aligned}$$

$$\boxed{\alpha_i > 0, \; \beta_i \ne 0}$$

Equivalence: Monotone Transformations

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

when g is strictly monotonically increasing

$\underline{\text{e.g.:}}$ $\underline{g(x) = e^x}$

If x^* is optimal for (A)

then $f(x^*) \leq f(x) \quad \forall x \in X$

$\Leftrightarrow g(f(x^*)) \leq g(f(x)) \quad \forall x \in X$

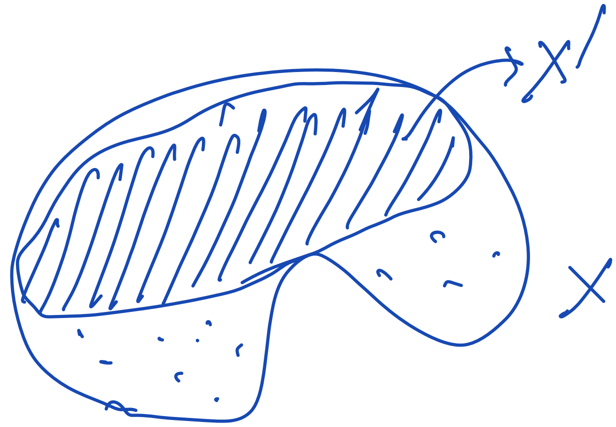
$\Leftrightarrow x^*$ is optimal for (B).

Inner Approximation

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

(B) is an inner approximation of (A) when

$$X' \subset X.$$



Let f_A^* be the optimal value of (A)

f_B^* " " " (B).

Then $f_A^* \le f_B^*$.

Optimal value of B gives an upper bound on the optimal value of A.

Relaxation and Soft Constraints

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

(B) is called a relaxation of (A) if $\underline{X' \supset X}$.

$\underline{f_A^* \geq f_B^*}$.

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

s.t. $g_1(x) \leq 0$

$\vdots$

$g_K(x) \leq 0$

$h_1(x) = 0$

$\vdots$

$h_p(x) = 0$

$\min_{x \in \mathbb{R}^n} f(x) + \lambda_1 \sum \varepsilon_i^2 + \lambda_2 \sum \delta_i^2$ (B)

s.t. $g_1(x) \leq \varepsilon_1$

$g_2(x) \leq \varepsilon_2$

$\vdots$

$g_K(x) \leq \varepsilon_K$

$-\delta_1 \leq h_1(x) \leq \delta_1$

$-\delta_2 \leq h_2(x) \leq \delta_2$

$\vdots$

$-\delta_p \leq h_p(x) \leq \delta_p$

(B) has <u>soft constraints</u> compared to (A).

λ_1, λ_2 are large constants to penalize large values of ε's & δ's

Convex Optimization Problems

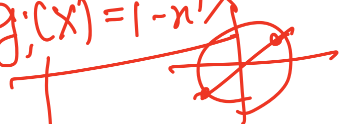
An optimization problem in abstract form

$$\min_{x \in X} f(x), \tag{6}$$

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

$x^T x \geq 1 \iff 1 - x^T x \leq 0$

$g_i(x) = 1 - x^T x$



An optimization problem in standard form

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

$x_1^2 + x_2^2 - 1 = 0$

is convex when

- f and g_i are convex functions.
- h_j **are affine functions.** $h_j(x) = a_j^T x + b_j$ for some $a_j \in \mathbb{R}^n$, $b_j \in \mathbb{R}$.

level set of function g_i at level 0

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

$\bar{X}_j = \{x \in \mathbb{R}^n \mid h_j(x) = 0\}$, $j = 1, 2 \ldots p$

$$X = \left(\bigcap_{i=1}^{m} X_i\right) \cap \left(\bigcap_{j=1}^{p} \bar{X}_j\right)$$

Since intersection of convex sets is a convex set, X is a convex set if each X_i and $\bar{X}_j$ are convex.

when g_i is a convex function, X_i is a convex set.

$$\boxed{h_j(x) = 0 \iff \begin{matrix} h_j(x) \leq 0 \\ h_j(x) \geq 0 \end{matrix}}$$

$\{x \in \mathbb{R}^n \mid h_j(x) \geq 0\}$ is a convex set when

h_j is a concave f^n.

⇒ $\overline{X}_j$ is a convex set when h_j is both convex & concave

⇓

h_j must be affine.

1. Local Optimum is Global

Ex: $\min_{x} f(x)$ s.t. $x^T x - 1 = 0$ ⇒ not a convex optimization problem since $\{x \mid x^T x = 1\}$ is not a convex set

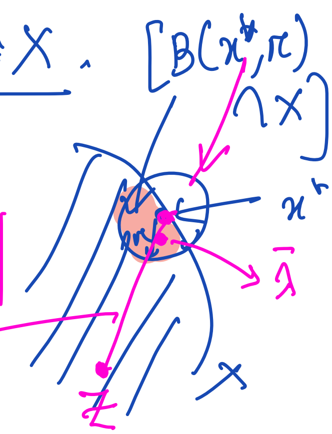
let $\min_{x \in X} f(x)$ be a convex optimization problem

⇒ f is a convex f^n
X is a convex set.

Let $x^* \in X$ be a locally optimal solution. ⇒

$f(x^*) \le f(x)$ $\forall\, x \in [B(x^*, r) \cap X]$ for radius $r > 0$.

we need to show $f(x^*) \le f(z)$ $\forall\, z \in X$.



we can choose $\bar{\lambda} \in [0, 1]$ which is sufficiently small such that

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

$\lambda z + (1 - \lambda) x^*$ [for $\lambda \in [0, 1]$] $\in X$

From the local optimality of x^*, we have

$$f(x^*) \le f(\bar{\lambda} z + (1 - \bar{\lambda}) x^*)$$
$$\le \bar{\lambda} f(z) + (1 - \bar{\lambda}) f(x^*)$$

(due to f being a convex function)

⇒ $\cancel{\bar{\lambda}} f(x^*) \le \cancel{\bar{\lambda}} f(z)$ $\forall\, z \in X$

⇒ x^* is a globally optimal solution.

2. Necessary and Sufficient Optimality Condition & let f be differentiable

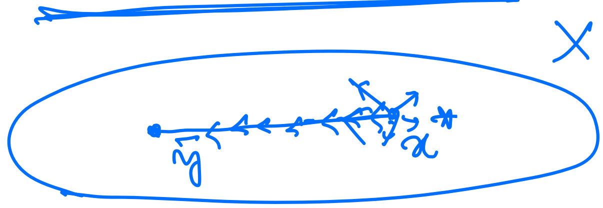
Let $\min_{x \in X} f(x)$ be a convex optimization problem;
$x^* \in X$ is a globally optimal solution if & only if
$$\nabla f(x^*)^T (y - x^*) \geq 0 \quad \forall\, y \in X.$$
The underlined condition is called the minimum principle.

proof:-
($\Leftarrow$) Suppose we have found x^* s.t. $\nabla f(x^*)^T (y - x^*) \geq 0$ $\forall y \in X$.

Let $y \in X$. Since f is a convex function,
$$f(y) \geq f(x^*) + \nabla f(x^*)^T (y - x^*)$$
$$\Rightarrow f(y) \geq f(x^*)$$
$\Rightarrow$ x^* is a globally optimal solution.

($\Rightarrow$) Let x^* be a globally optimal solution.
Suppose $\exists\, \bar{y} \in X$ for which $\nabla f(x^*)^T (\bar{y} - x^*) < 0$.

$$\phi(t) = f(x^* + t(\bar{y} - x^*)), \quad t \in \mathbb{R}$$



$$\phi'(t) = Df(x^* + t(\bar{y} - x^*))(\bar{y} - x^*)$$
$$= \nabla f(x^* + t(\bar{y} - x^*))^T (\bar{y} - x^*)$$
$$\phi'(t)\big|_{t=0} = \nabla f(x^*)^T (\bar{y} - x^*) < 0 \Rightarrow \exists\, \varepsilon \text{ close to } 0 \text{ s.t.}$$
$$f(x^* + \varepsilon(\bar{y} - x^*)) < f(x^*)$$
$\Rightarrow$ x^* is not globally

$\Rightarrow \nexists\ \bar{y} \in X$ s.t $\nabla f(x^*)^T(\bar{y} - x^*) < 0$ which is optimal a contradiction.

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

3. Set of Minimizers is a Convex Set

$\min_{x \in X} f(x)$, X is a convex set, f is a convex function.

$$X_{opt} = \{x \in X \mid f(x) = f^*\} \quad : \text{set of optimal solutions.}$$

$$= \arg\min_{x \in X} f(x).$$

Show that X_{opt} is a convex set.

Let $x, y \in X_{opt}$

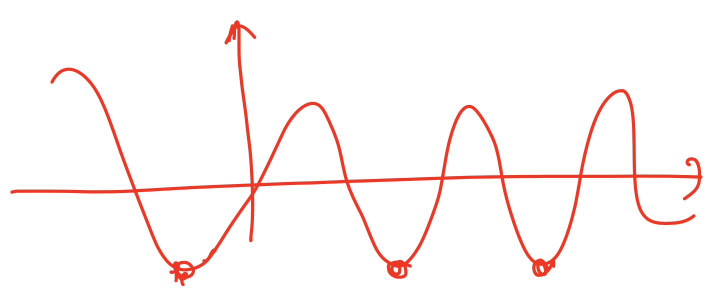
need to show $\lambda x + (1-\lambda) y \in X_{opt}$ $\forall \lambda \in [0,1]$.

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

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

Since f^* is the optimal value,

$$f(\lambda x + (1-\lambda) y) = f^* \Rightarrow \lambda x + (1-\lambda) y \in X_{opt}$$



$$\min_{x \in \mathbb{R}^n} \quad c^T x + d$$
$$\text{s.t.} \quad \boxed{a_i^T x \le b_i \quad , i = 1, 2 \ldots m} \iff \underline{Ax \le b}, \quad A \in \mathbb{R}^{m \times n}, \; b \in \mathbb{R}^m$$
$$\underline{g_j^T x = h_j} \quad , j = 1, 2 \ldots p$$

$$\Downarrow$$

$$Gx = h, \quad \underline{G \in \mathbb{R}^{p \times n}, \; h \in \mathbb{R}^p}$$

$$b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

$$A = \begin{bmatrix} - & a_1^T & - \\ - & a_2^T & - \\ & \vdots & \\ - & a_m^T & - \end{bmatrix}$$

In a more compact form,

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & c^T x \\ \text{s.t.} \quad & Ax \le b \\ & Gx = h \end{aligned}$$