Linear Programming (LP)

LP is a class of optimization problems where the cost function is linear in the decision variable and the feasibility set is a polyhedron.

LP in standard equality form:

$$
\begin{aligned}
\min_{x\in\mathbb{R}^n} \; & c^\top x \\
\text{s.t.} \; & \boxed{Ax = b}, \\
& \boxed{x \geq 0}.
\end{aligned}
\qquad \text{(P)}
$$

$x_+, x_- \in \mathbb{R}^n_{\geq 0}$

$x = x_+ - x_-$

$Ax \leq b$

$\Updownarrow$

$Ax_+ - Ax_- \leq b$

$$
\begin{aligned}
\min_{x_+, x_-, s} \; & c^\top x_+ - c^\top x_- \\
\text{s.t.} \; & Ax_+ - Ax_- + s = b \\
& x_+, x_-, s \geq 0
\end{aligned}
$$

Standard inequality form:

$$
\begin{aligned}
\min_{x\in\mathbb{R}^n} \; & c^\top x \\
\text{s.t.} \; & Ax \leq b
\end{aligned}
$$

Example:

$$
\begin{aligned}
\min_{x\in\mathbb{R}^2} \; & x_1 + x_2 \\
\text{s.t.} \; & x_1 \geq 0, \; x_2 \geq 0 \\
& 2x_1 + x_2 \leq 1
\end{aligned}
$$

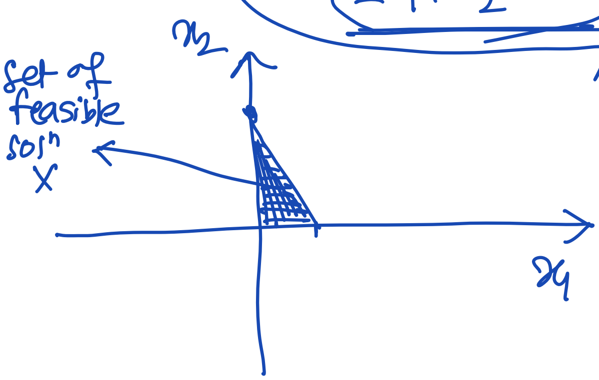
$2x_1 + x_2 \leq 1$

$-x_1 \leq 0$

$-x_2 \leq 0$

$A = \begin{bmatrix} 2 & 1 \\ -1 & 0 \\ 0 & -1 \end{bmatrix} \in \mathbb{R}^{3\times 2}$

$b = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \; c = [1, 1]$



$$
\begin{aligned}
\min_{x, s} \; & x_1 + x_2 \\
\text{s.t.} \; & 2x_1 + x_2 + s = 1 \\
& x_1 \geq 0, \; x_2 \geq 0, \; s \geq 0
\end{aligned}
$$

$c = [1 \;\; 1 \;\; 0]$

$A = [2 \;\; 1 \;\; 1], \; b = 1$

$x = \begin{bmatrix} x_1 \\ x_2 \\ s \end{bmatrix} \in \mathbb{R}^3_{\geq 0}$

Obtaining a lower bound on the cost function

Consider the problem

$$\min_{x \in \mathbb{R}^n} c^T x \quad \text{s.t. } Ax = b,\ x \geq 0 \qquad (P)$$

$$[y_1 \cdots y_m] Ax$$

$$y_1 a_1^T x = b_1 y_1$$
$$y_2 a_2^T x = b_2 y_2$$
$$\vdots$$
$$y_m a_m^T x = b_m y_m$$

(+)

$$y^T A x = \sum_{i=1}^{m} y_i a_i^T x = \sum_{i=1}^{m} b_i y_i = y^T b.$$

elementwise

If y is chosen such that $y^T A \leq c^T$, then

$$y^T b = y^T A x \leq c^T x$$

inequality is preserved because $x \geq 0$.

Whenever x is a feasible solution of (P) and $y \in \mathbb{R}^m$ satisfies $y^T A \leq c^T$, then $y^T b \leq c^T x$.

Suppose we find $\bar{y}$ s.t. $\bar{y}^T A \leq c^T$.

can we say $\bar{y}^T b$ is a lower bound on the optimal value of problem (P)?

Yes, we can say so because whenever x^* is optimal sol^n of (P), it is also feasible which implies $\bar{y}^T b \leq c^T x^*$.

What is the best possible lower bound?

$$\max_{\bar{y}} b^T \bar{y} \quad \text{s.t. } (\bar{y})^T A \leq c^T \iff A^T \bar{y} \leq c$$

$$\equiv \min_{\bar{y} \in \mathbb{R}^m} (-b)^T \bar{y}$$

Finding best possible lower bound

This happens to be another linear program:

$$\begin{aligned} &\max_{y\in\mathbb{R}^m} b^\top y \\ &\text{s.t.} A^\top y \le c. \end{aligned} \tag{D}$$

The above problem is referred to as the dual of problem (P).
A LP stated as above is called standard inequality form.
We can show that the dual of (D) is (P).

Any linear optimization problem belongs to one of three classes:
i) infeasible
ii) unbounded, or
iii) it has an optimal solution

Properties

Theorem 7

For the primal-dual pair of optimization problems stated above, the following are true.

1. If (P) is infeasible, and (D) is feasible, then (D) is unbounded.
2. If (P) is unbounded, then (D) is infeasible.
3. **Weak Duality**: For any feasible solution $\bar{x}$ and $\bar{y}$ of the respective problems, we always have $c^\top \bar{x} \geq b^\top \bar{y}$.

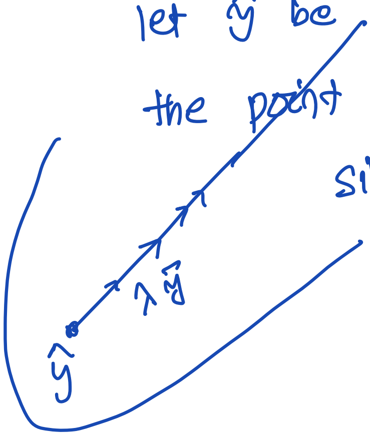
   Let both (P) & (D) be feasible $\Rightarrow$ both are bounded.
4. **Strong Duality**: Show that for the respective optimal solutions x^* and y^*, we must have $c^\top x^* = b^\top y^*$.

HW: Give an example of (P) and (D) where both are infeasible.

(1) $\{x \in \mathbb{R}^n \mid Ax = b, x \geq 0\}$ is empty $\Rightarrow \exists \bar{y}$ s.t. $A^T\bar{y} \leq 0$, $b^T\bar{y} > 0$

let $\hat{y}$ be feasible to (D) $\Rightarrow A^T\hat{y} \leq C$.

the point $\hat{y} + \lambda\bar{y}$ is feasible to (D) for every $\lambda \geq 0$.



since $A^T(\hat{y} + \lambda\bar{y}) = \underbrace{A^T\hat{y}}_{\leq C} + \underbrace{\lambda A^T\bar{y}}_{\leq 0} \leq C$

objective function $b^T\hat{y} + \lambda (b^T\bar{y})$ +ve

$\lambda \geq 0$

therefore $\lim_{\lambda \to \infty} b^T(\hat{y} + \lambda\bar{y}) = \infty \Rightarrow$ (D) is unbounded.

(2) proof by contradiction: suppose (D) is feasible $\Rightarrow \exists \bar{y} : A^T\bar{y} \leq C$

Whenever (D) is feasible, primal is bounded.

$\Rightarrow b^T\bar{y} \leq c^T x$ for every x that is feasible for (P)

$\Rightarrow$ primal is not unbounded.

Strong duality says that when x^* is optimal for (P), y^* is optimal for (D), then $c^T x^* = b^T y^*$

Proof

If strong duality holds, then $\exists\, x^*, y^*$ satisfying

$$\left.\begin{array}{l} Ax^* = b \\ x^* \geq 0 \\ A^T y^* \leq c \\ c^T x^* \leq b^T y^* \end{array}\right\}$$

Recall: Farka's lemma

(I) $\{x \in \mathbb{R}^n \mid Ax = b, x \geq 0\}$

(II) $\{y \in \mathbb{R}^m \mid y^T A \leq 0, b^T y > 0\}$

Exactly one of the above sets is empty.

$\{(x,y) \mid Ax = b, x \geq 0, A^T y \leq c, c^T x \leq b^T y\} \ldots (S_1)$

If S_1 is not an empty set, then strong duality holds.

To show that S_1 is not an empty set, we will express it as (I), & then show that (II) has no solution.

Step-1: Express (S_1) as (I). $y = y_+ - y_-$

$$\bar{x} = \begin{bmatrix} x \\ y_+ \\ y_- \\ s_1 \\ s_2 \end{bmatrix} \in \mathbb{R}^{n+2m+n+1} \geq 0$$

$$\left.\begin{array}{l} Ax = b \\ A^T y_+ - A^T y_- + s_1 = c \\ c^T x - b^T y_+ + b^T y_- + s_2 = 0 \end{array}\right] \Rightarrow \bar{A}\bar{x} = \bar{b}$$

$$\bar{A} = \begin{bmatrix} A & 0 & 0 & 0 & 0 \\ 0 & A^T & -A^T & I & 0 \\ c^T & -b^T & b^T & 0 & 1 \end{bmatrix}, \quad \bar{b} = \begin{bmatrix} b \\ c \\ 0 \end{bmatrix} \in \mathbb{R}^{n+m+1}$$

If S_1 is not empty, $\exists\, x_1, y_1$

s.t. $\begin{cases} Ax_1 = b,\ x_1 \geq 0 \Rightarrow x_1 \text{ is feasible for (P)} \\ A^T y_1 \leq c \Rightarrow y_1 \text{ is feasible for (D)} \end{cases}$

$c^T x_1 \leq b^T y_1$

From weak duality, $c^T x_1 \geq b^T y_1$

$\Rightarrow c^T x_1 = b^T y_1$

$\Rightarrow x_1$ & y_1 are respective optimal solutions

then $S_1 \equiv \{\bar{x} \geq 0, \bar{A}\bar{x} = \bar{b}\}$

Chybyshev Center of a polyhedron

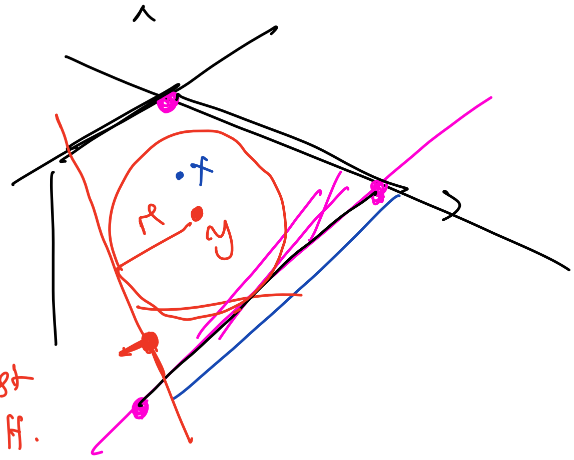
$H = \{ x \in \mathbb{R}^n \mid a_1^T x \le b_1,\ a_2^T x \le b_2,\ \dots,\ a_m^T x \le b_m \}$

Let H be a bounded set.

H can also be written as

$H = \mathrm{conv}(z_1, z_2 \dots z_K)$,

where $z_i \in \mathbb{R}^n$ is an extreme/corner point of H.



Goal: find $y \in \mathbb{R}^n$ which is at farthest distance from the boundary of H.

y is the center of the largest possible circle that we can fit in H.

Let us formulate an optimization problem to find y and r.

Decision variable: $y \in \mathbb{R}^n$, $r \in \mathbb{R}$

cost function: $-r$

constraints : $\{ x \mid \|x - y\| \le r \}$: circle.

For the circle to reside inside polyhedron, we need

$a_i^T x \le b_i \quad \forall i = 1, 2, \dots m$

$x = y + p$, then $a_i^T x \le b_i \iff a_i^T y + a_i^T p \le b_i$

$x - y = p \Rightarrow \|x - y\|_2 \le r \iff \|p\|_2 \le r$

$$\min_{y \in \mathbb{R}^n, r \in \mathbb{R}} \ -r$$
s.t. $a_i^T y + a_i^T p \le b_i \quad \forall \|p\|_2 \le r, \ \forall i = 1, 2, \dots m$

$\Updownarrow$

(LP)
$$\min_{y \in \mathbb{R}^n, r \in \mathbb{R}} \ -r$$
s.t. $a_i^T y + \sup_{\|p\|_2 \le r} a_i^T p \le b_i, \quad \forall i = 1, \dots m.$

$r \|a_i\|_2$

Farkas' Lemma

To prove the strong duality theorem, we will make use of an alternative form of Farka's lemma.

Lemma 2 (Farkas' Lemma). *Let $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$. Then, exactly one of the following sets must be empty:*

1. $\{x \in \mathbb{R}^n \mid Ax = b, x \geq 0\}$
2. $\{y \in \mathbb{R}^m \mid A^\top y \leq 0, b^\top y > 0\}$.

Lemma 3 (Alternative form of Farkas' Lemma). *Let $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$. Then, exactly one of the following sets must be empty:*

1. $\{x \in \mathbb{R}^n | Ax \leq b\}$
2. $\{y \in \mathbb{R}^m | y \geq 0, y^\top A = 0, y^\top b < 0\}$.

Suppose (II) has a solution (we will then arrive at a contradiction).

Case 1: let $z_3 < 0$. Case 2: on Blackboard

Let $\hat{x} = \frac{z_2}{z_3}$, $\hat{y} = \frac{z_1}{(-z_3)}$.

Since $z_2 \leq 0$, then $\hat{x} \geq 0$
Since $Az_2 = z_3 b$, then $A\hat{x} = b$ $\Rightarrow$ $\hat{x}$ is feasible to (P).

Since $A^T z_1 + z_3 c \leq 0$, then $A^T z_1 \leq (-z_3) c$

$\Rightarrow A^T \hat{y} \leq c \Rightarrow \hat{y}$ is feasible to (D)

From weak duality, $c^T \hat{x} \geq b^T \hat{y}$

$\Leftrightarrow c^T \frac{z_2}{z_3} \geq \frac{b^T z_1}{(-z_3)} \Leftrightarrow \frac{c^T z_2 + b^T z_1}{(+z_3)} \geq 0$

which is opposite of what we had for (II). $\Rightarrow$ $z_3 < 0$ is impossible. $\Leftrightarrow c^T z_2 + b^T z_1 \leq 0$

<u>Homework problem</u>: The function f is convex. Show that $g(x,t) = t f\left(\frac{x}{t}\right)$ is convex.

Let (x_1, t_1) & (x_2, t_2).

We need to show $g(\lambda x_1 + (1-\lambda)x_2, \lambda t_1 + (1-\lambda)t_2) \le \lambda g(x_1, t_1) + (1-\lambda) g(x_2, t_2)$. $\forall \lambda \in [0,1]$.

$$\text{LHS} = (\lambda t_1 + (1-\lambda)t_2)\, f\left(\frac{\lambda x_1 + (1-\lambda)x_2}{\lambda t_1 + (1-\lambda)t_2}\right)$$

$$= (\lambda t_1 + (1-\lambda)t_2)\, f\left[\underbrace{\frac{\lambda t_1}{\lambda t_1 + (1-\lambda)t_2}}_{=:\mu} \frac{x_1}{t_1} + \underbrace{\frac{(1-\lambda)t_2}{\lambda t_1 + (1-\lambda)t_2}}_{1-\mu.} \frac{x_2}{t_2}\right]$$

$$= (\lambda t_1 + (1-\lambda)t_2)\, f\left(\mu \frac{x_1}{t_1} + (1-\mu)\frac{x_2}{t_2}\right)$$

$$\le (\lambda t_1 + (1-\lambda)t_2)\left(\mu f\left(\frac{x_1}{t_1}\right) + (1-\mu) f\left(\frac{x_2}{t_2}\right)\right)$$

$$= \lambda\, t_1 f\left(\frac{x_1}{t_1}\right) + (1-\lambda) t_2 f\left(\frac{x_2}{t_2}\right)$$

$$= \lambda g(x_1, t_1) + (1-\lambda) g(x_2, t_2) \quad \blacksquare$$

Lagrangian Function

Consider the following optimization problem in standard form:

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

The Lagrangian function $L : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p : \mathbb{R}$ is defined as

$$L(x, \lambda, \mu) := f(x) + \sum_{i \in [m]} \lambda_i g_i(x) + \sum_{j \in [p]} \mu_j h_j(x),$$

where

- λ_i is the Lagrange multiplier associated with $g_i(x) \leq 0$
- μ_j is the Lagrange multiplier associated with $h_j(x) = 0$.

Lower Bound Property:

Lemma 4. *If $\bar{x}$ is feasible and $\bar{\lambda} \geq 0$, then* $\boxed{f(\bar{x}) \geq L(\bar{x}, \bar{\lambda}, \mu).}$

$$L(\bar{x}, \bar{\lambda}, \mu) = f(\bar{x}) + \underbrace{\sum_{i \in [m]} \underset{\geq 0}{\bar{\lambda}_i} \underset{\leq 0}{g_i(\bar{x})}}_{\leq 0} + \underbrace{\sum_{j \in [p]} \mu_j h_j(\bar{x})}_{= 0}$$

$$\Rightarrow \quad L(\bar{x}, \bar{\lambda}, \mu) \leq f(\bar{x}).$$

Lagrangian Dual

From the previous lemma, we know that if $\bar{x}$ is feasible and $\bar{\lambda} \geq 0$, then

$$f(\bar{x}) \geq L(\bar{x}, \bar{\lambda}, \mu) \geq \inf_x L(x, \bar{\lambda}, \mu) =: d(\bar{\lambda}, \mu).$$

where

$$d(\lambda, \mu) := \inf_x \left[f(x) + \sum_{i \in [m]} \lambda_i g_i(x) + \sum_{j \in [p]} \mu_i h_j(x) \right]. = \inf_{x \in \mathbb{R}^n} L(x, \lambda, \mu)$$

- $d(\lambda, \mu)$ requires solving an unconstrained optimization problem.
- Given any $\lambda \geq 0, \mu$, $d(\lambda, \mu) \leq f^*$ where f^* is the optimal value.
- $d(\lambda, \mu)$ may take value $-\infty$ for some choice of λ and μ.
- $d(\lambda, \mu)$ is concave in λ and μ.

If $\bar{x}$ is feasible, $\lambda \geq 0$ $f(\bar{x}) \geq d(\lambda, \mu)$

$\Rightarrow d(\lambda, \mu) \leq f^*$

If $f(x) = \sup_{y \in Y} g(x, y)$, and $g(x, y)$ is convex in x for every y, then $f(x)$ is convex in x $\rightarrow$ apply this to L

sup becomes inf for concave functions.

$L(x, \lambda, \mu)$ is affine in (λ, μ)

$\Rightarrow L(x, \lambda, \mu)$ is concave in (λ, μ)

$d(\lambda, \mu) = \inf_{x \in \mathbb{R}^n} L(x, \lambda, \mu)$ is concave in (λ, μ)

Lagrangian Dual Optimization Problem

Let us compute the best lower bound on f^*:

$$\begin{aligned} \max_{\lambda \in \mathbb{R}^m, \mu \in \mathbb{R}^p} \quad & d(\lambda, \mu) \\ \text{s.t.} \quad & \lambda \geq 0, \\ & (\lambda, \mu) \in \mathtt{dom}(d). \end{aligned}$$

- The above is a convex optimization problem since $d(\lambda, \mu)$ is concave in λ and μ.
- Let the optimal value be denoted d^*.

$$\max d(\lambda,\mu) \equiv \min \left(-d(\lambda,\mu)\right)$$

$$d(\lambda,\mu) = \inf_{x \in \mathbb{R}^n} \underbrace{\left[f(x) + \sum_{i=1}^{m} \lambda_i g_i(x) + \sum_{i=1}^{p} \mu_i h_i(x)\right]}_{\text{affine in } \lambda \text{ and } \mu}$$

$\Rightarrow$ concave in λ & μ.

Recall: $\sup_y f(x,y)$ is convex whenever f is convex in x for every y.

$\Updownarrow$

$\inf_y f(x,y)$ is concave whenever f is concave in x for every y.

proof as HW

Example 1: Lagrangian Dual of LP

$\min_{x \in \mathbb{R}^n} f(x)$
s.t. $g_i(x) \le 0$
$h_j(x) = 0$

$A \in \mathbb{R}^{m\times n}$

$$\min_{x\in\mathbb{R}^n} \quad c^\top x$$
$$\text{s.t.} \quad Ax = b, x \ge 0.$$

$\Rightarrow$ $f(x) = c^T x$
$g_i(x) \doteq -x \le 0$
$h_j(x) \doteq Ax = b.$

Find L, d and $\mathtt{dom}(d)$.

$$L(x,\lambda,\mu) = c^T x + \sum_{i\in[m]} \lambda_i(-x_i) + \sum_{j\in[p]} \mu_j (a_j^T x - b)$$

$\lambda \in \mathbb{R}^n$
$\mu \in \mathbb{R}^m$

$$= c^T x - \lambda^T x + \mu^T (Ax - b)$$

$$d(\lambda,\mu) = \inf_{x\in\mathbb{R}^n} \left[c^T x - \lambda^T x + \mu^T A x - \mu^T b \right]$$

$$= \inf_{x\in\mathbb{R}^n} \left[\left(\underline{c - \lambda + A^T \mu} \right)^T x - \mu^T b \right]$$

$$= \begin{cases} -\mu^T b & \text{if } \quad c - \lambda + A^T\mu = 0 \\ -\infty & , \text{ otherwise.} \end{cases}$$

Dual optimization problem:

$$\max_{\lambda \ge 0, \mu} \quad -\mu^T b$$
$$\text{s.t.} \quad A^T\mu + c = \lambda$$

$\Updownarrow$

$$\max_{\mu} \quad -\mu^T b$$
$$\text{s.t.} \quad \underline{A^T\mu + c \ge 0}$$

$z = -\mu.$

same as the dual derived for LP $\leftarrow$ $\max_{z} \; z^T b$ s.t. $A^T z \le c$ $\Longleftrightarrow$

Example 2: Least Norm Solution

Least norm solution:

$$\min_{x \in \mathbb{R}^n} \quad \frac{1}{2} x^\top x$$
$$\text{s.t.} \quad Ax = b.$$

Find L and d.

$$L(x, \mu) = \frac{1}{2} x^T x + \mu^T (Ax - b)$$

$$d(\mu) = \inf_{x \in \mathbb{R}^n} \left[\frac{1}{2} x^T x + \mu^T A x - \mu^T b \right]$$

$$\nabla L_x(x, \mu) = x + A^T \mu = 0 \Rightarrow \underline{x^* = -A^T \mu}$$

$\nabla^2 L_x(x, \mu) = I$ is positive

$$d(\mu) = L(x^*, \mu) = \frac{1}{2} (A^T \mu)^T (A^T \mu) + \mu^T A (-A^T \mu) - \mu^T b$$

$$= \frac{1}{2} \mu^T A A^T \mu - \mu^T A A^T \mu - \mu^T b$$

$$= -\frac{1}{2} \mu^T A A^T \mu - \mu^T b$$

Dual optimization problem: $\max_{\mu} d(\mu) = \min_{\mu} \left(\frac{1}{2} \mu^T \underline{A A^T} \mu + \mu^T b \right)$

the dual is an unconstrained convex optimization problem.

Weak and Strong Duality

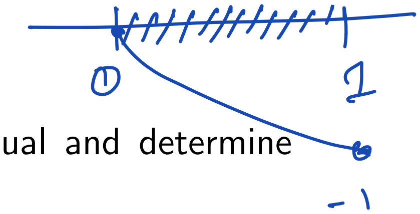
Weak Duality: $d^* \leq f^*$ always holds (even for non-convex problems).

Strong Duality: $d^* = f^*$ is guaranteed to hold for convex problems satisfying certain conditions, referred to as *constraint qualification* conditions.

Example 3

optimal value $= -1$ at $x^* = +1$

$$\min_{x\in\mathbb{R}} \quad -x^2$$
$$\text{s.t.} \quad x-1 \le 0, \quad -x \le 0.$$



Find the optimal value of the above problem, derive the dual and determine whether strong duality holds.

$$L(x,\lambda) = -x^2 + \lambda_1(x-1) - \lambda_2 x$$

$x \in \mathbb{R}$, $\lambda \in \mathbb{R}^2$

$$d(\lambda) = \inf_{x\in\mathbb{R}} \left[-x^2 + \lambda_1(x-1) - \lambda_2 x \right]$$

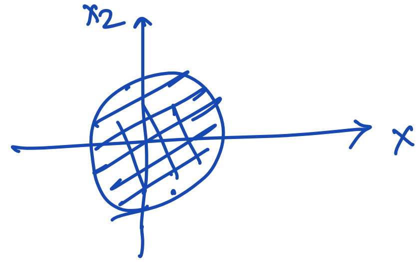
$$= -\infty$$

$\Rightarrow$ weak duality is satisfied.

$\Rightarrow$ strong duality is not satisfied.

feasibility set is a convex set.
but cost function is not convex.

Example 4



optimal value $= -1$

optimal sol^n: $\{x \in \mathbb{R}^n \mid x_1^2 + x_2^2 = 1\}$

$$\begin{aligned} \min_{x\in\mathbb{R}^2} \quad & -x_1^2 - x_2^2 \\ \text{s.t.} \quad & x_1^2 + x_2^2 - 1 \le 0. \end{aligned}$$

Find the optimal value of the above problem, derive the dual and determine whether strong duality holds.

$$\mathcal{L}(x,\lambda) = -x_1^2 - x_2^2 + \lambda(x_1^2 + x_2^2 - 1)$$

$$d(\lambda) = \inf_{x\in\mathbb{R}^2} \mathcal{L}(x,\lambda)$$

$$= \inf_{x\in\mathbb{R}^2} \left[\underline{(\lambda-1)}\, \underline{\underline{(x_1^2 + x_2^2)}} - \lambda \right]$$

$$= \begin{cases} -\lambda & \text{when } \lambda \ge 1 \\ -\infty & \text{when } \lambda < 1 \end{cases}$$

Dual optimization problem: $\max_{\lambda\ge 0} -\lambda$ s.t. $\lambda \ge 1$ $\iff$ $\min \lambda$ s.t. $\lambda \ge 1$

optimal dual solution is $\lambda^* = 1$

optimal value of dual is -1

Strong duality holds.

Karush Kuhn Tucker.

KKT Optimality Conditions

For the above primal and dual optimization problems, $\bar{x}, \bar{\lambda}$ and $\bar{\mu}$ are said to satisfy KKT optimality conditions if the following holds:

KKT conditions:

- **Primal Feasibility**: $g_i(\bar{x}) \leq 0, i \in [m], h_j(\bar{x}) = 0, j \in [p]$.
- **Dual Feasibility**: $\bar{\lambda} \geq 0$.
- **Complementary Slackness**: $\bar{\lambda}_i g_i(\bar{x}) = 0$ for all $i \in [m]$.
- **Stationarity**: $\nabla_x f(\bar{x}) + \sum_{i \in [m]} \bar{\lambda}_i \nabla_x g_i(\bar{x}) + \sum_{j \in [p]} \bar{\mu}_i \nabla_x h_j(\bar{x}) = 0$.

$$\nabla_x L(\bar{x}, \bar{\lambda}, \bar{\mu}) = 0$$

$\Rightarrow \bar{x}$ is a point which is a possible minimizer of $L(x, \bar{\lambda}, \bar{\mu})$.

$$\nabla_x^2 L(x, \bar{\lambda}, \bar{\mu}) = \underbrace{\nabla_x^2 f(x) + \sum_{i=1}^{m} \bar{\lambda}_i \nabla_x^2 g_i(x)}_{} + \cancel{\sum_{j=1}^{p} \bar{\mu}_j \nabla_x^2 h_j(x)} \to 0$$

If the problem is convex,

$\nabla_x^2 f$ is psd.

$\nabla_x^2 g_i$ is psd.

$\nabla_x^2 h_i = 0$

$\nabla_x^2 L(x, \bar{\lambda}, \bar{\mu})$ is psd.

Sufficient Condition for Optimality

Let $\bar{x}, \bar{\lambda}$ and $\bar{\mu}$ satisfy KKT conditions stated above. From primal and dual feasibility we have

$$d(\bar{\lambda}, \bar{\mu}) = \inf_x \left[f(x) + \sum_{i \in [m]} \bar{\lambda}_i g_i(x) + \sum_{j \in [p]} \bar{\mu}_j h_j(x) \right]$$

$$\le f(\bar{x}) + \sum_{i \in [m]} \underbrace{\bar{\lambda}_i g_i(\bar{x})}_{\le 0} + \sum_{j \in [p]} \bar{\mu}_i \underbrace{h_j(\bar{x})}_{=0} \le f(\bar{x}).$$

$\inf_x L(x,\bar{\lambda},\bar{\mu}) \le L(\bar{x},\bar{\lambda},\bar{\mu})$

Further, both inequalities hold with equality. $= 0$ due to complementary slackness

↳ holds with equality.

Thus, when the primal problem is convex, we have:

- $d(\bar{\lambda}, \bar{\mu}) = f(\bar{x})$ (strong duality)
- $\bar{x}$ is optimal solution of primal problem.
- $(\bar{\lambda}, \bar{\mu})$ are optimal solution of dual problem.

$$d(\bar{\lambda},\bar{\mu}) = \inf_x L(x,\bar{\lambda},\bar{\mu}) \le f(\bar{x})$$

$$d(\bar{\lambda},\bar{\mu}) = \inf_x L(x,\bar{\lambda},\bar{\mu}) = L(\bar{x},\bar{\lambda},\bar{\mu}) = f(\bar{x}).$$

Necessary and Sufficient Condition for Optimality

Theorem 8

Suppose the primal optimization problem is convex which satisfies Slater's constraint qualification condition: there exists $\bar{x} \in \mathtt{int}(\mathcal{D})$ in the domain of the optimization problem for which $g_i(\bar{x}) < 0$ for all $i \in [m]$ and $h_i(\bar{x}) = 0$ for all $i \in [p]$.

Then, strong duality holds. Equivalently, a feasible solution x^* is optimal if and only if there exist λ^*, μ^* such that (x^*, λ^*, μ^*) satisfy KKT optimality conditions.

Constraint qualification is required for the necessity part of the proof.

Sufficiency: shown in previous slide.

Necessity: If primal has an optimal sol^n x^*, then $\exists$ (λ^*, μ^*) s.t. (x^*, λ^*, μ^*) together satisfy KKT conditions.

Convex Theorem of the Alternative

Consider the following general form of optimization problem:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f(x) \\ \text{s.t.} \quad & g_i(x) \leq 0, i \in [m] := \{1, 2, \ldots, m\}, \end{aligned}$$

where f and g_i are convex functions.

Theorem 9

Let the constraint functions g_i satisfy **slater's** condition: there exists $\bar{x}$ such that $g_i(\bar{x}) < 0$ for all $i \in [m]$. Then, exactly one of the following two systems must be empty.

- $\{x \in \mathbb{R}^n | f(x) < 0, g_i(x) \leq 0, i \in [m]\}$ $= S_1$
- $\{\lambda \in \mathbb{R}^m | \inf_{x \in \mathbb{R}^n}[f(x) + \sum_{i \in [m]} \lambda_i g_i(x)] \geq 0$, $\lambda \geq 0 \} = S_2$

$f(x) \equiv f(x) - f(x^*)$

Proof: Blackboard.

Proof of Theorem 10:

Suppose x^* is an optimal solution. Find a set analogous to "S_1" which does not have a solution so that we can say that $\exists \lambda^* \in$ "S_2" satisfying $\inf_{x \in \mathbb{R}^n} [f(x) + \sum_{i=1}^m \lambda_i^* g_i(x)] \geq f(x^*)$.

$S_1' = \{x \in \mathbb{R}^n | f(x) - f(x^*) < 0, g_i(x) \leq 0, i \in [m]\}$

If x^* is optimal sol^n, then S_1' is an empty set.

$\Rightarrow S_2' = \{\lambda \in \mathbb{R}^m | \inf_{x \in \mathbb{R}^n} [f(x) - f(x^*) + \sum_i \lambda_i g_i(x)] \geq 0, \lambda \geq 0\}$ will be non-empty.

Let $\bar{\lambda} \in S_2' \Rightarrow \bar{\lambda} \geq 0, \inf_{x \in \mathbb{R}^n} [f(x) + \sum_i \bar{\lambda}_i g_i(x)] \geq f(x^*)$

For the converse:

Strong Duality Theorem

Consider the following general form of optimization problem:

$$\min_{x \in \mathbb{R}^n} \quad f(x)$$
$$\text{s.t.} \quad g_i(x) \le 0, i \in [m] := \{1, 2, \ldots, m\},$$

where f and g_i are convex functions satisfying Slater's condition.

$\exists \bar{x}$ s.t. $g_i(\bar{x}) < 0 \ \forall i$

Theorem 10

x^* is an optimal solution to the above problem if and only if there exists $\lambda^* \ge 0$ such that $\inf_{x \in \mathbb{R}^n}[f(x) + \sum_{i \in [m]} \lambda_i g_i(x)] \ge f(x^*)$.

Proof: Blackboard.

$\underbrace{f(x^*) \ge d(\lambda^*)}_{\text{weak duality.}} = \inf_x [f(x) + \sum_i \lambda_i^* g_i(x)] \ge f(x^*)$

$\Rightarrow d(\lambda^*) = f(x^*)$ & strong duality holds

Converse:

If $\lambda^* \ge 0$ satisfies

$d(\lambda^*) = \inf_{x \in \mathbb{R}^n} [f(x) + \sum_i \lambda_i^* g_i(x)] \ge f(x^*)$

$\Rightarrow \quad d(\lambda^*) \ge f(x^*)$

from weak duality $d(\lambda^*) \le f(x^*)$

$\Rightarrow \quad d(\lambda^*) = f(x^*) \quad \Rightarrow$ strong duality holds & x^* is optimal.

Other notions of constraint qualification

- If all the constraint functions $g_i(x)$ and $h_j(x)$ are affine, then constraint qualification holds.

- **Relaxed Slater Condition:** If some of the inequality constraints are affine, then they need not hold with strict inequality. It is sufficient to find $\bar{x} \in \texttt{relint}(\mathcal{D})$ such that $g_i(\bar{x}) < 0$ for all g_i that are not affine.

- **Linear Independence Constraint Qualification** holds at a feasible solution x^* if the vectors

$$\begin{aligned} &\nabla h_j(x^*), \quad j \in [p], \\ &\nabla g_i(x^*), \quad i \in \{k \in [m] | g_k(x^*) = 0\} \end{aligned}$$

  are linearly independent.