Lecture 12: 29th Jan

Slack Variables and Epigraph Form

The following two optimization problems are equivalent:

$$\begin{aligned} \min_{x} \quad & \sum_{i=1}^{k} \phi_i(x) \\ \text{s.t.} \quad & x \in \mathcal{X}, \end{aligned} \tag{A}$$

$$\begin{aligned} \min_{x,t} \quad & \sum_{i=1}^{k} t_i \\ \text{s.t.} \quad & \left.\begin{aligned} & x \in \mathcal{X}, \\ & \phi_i(x) \le t_i, \end{aligned}\right\} i \in \{1, 2, \dots, k\}. \end{aligned} \tag{B}$$

If x is a feasible solution of (A) $\iff \bar{x} \in \mathcal{X}$,

then $(\bar{x}, \bar{t})$ where $\underline{\bar{t}_i = \phi_i(\bar{x})}$ is a feasible solution (B).

Change of Variables

Consider an optimization problem:

$$
\begin{aligned}
\min_{x} \quad & f_o(x) \\
\text{s.t.} \quad & f_i(x) \le 0, \qquad i \in \{1, 2, \ldots, k\} \\
& h_i(x) = 0, \qquad i \in \{1, 2, \ldots, p\}.
\end{aligned}
$$

Let $F : \mathcal{X} \to \mathcal{Y}$ be an invertible mapping with $y = F(x)$. Then, we have the following equivalent optimization problem with respect to y:

$y = F(x) \iff x = F^{-1}(y).$

$$
\begin{aligned}
\min_{y} \quad & f_o(F^{-1}(y)) \\
\text{s.t.} \quad & f_i(F^{-1}(y)) \le 0 \\
& h_i(F^{-1}(y)) = 0
\end{aligned}
$$

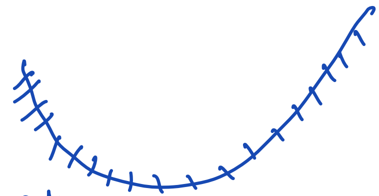
When is convexity preserved? If $y = Ax + b$ with A being an invertible matrix, then the convexity of the optimization problem is preserved under the transformation.

Substituting Equality with Inequality Constraints

Consider an optimization problem:

$$p^* = \min_{x \in \mathcal{X}} \quad f(x)$$
$$\text{s.t.} \quad b(x) = u.$$

(A)

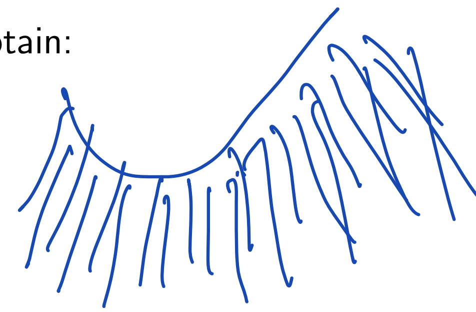


If we replace the equality constraint with an inequality, we obtain:

In general $g^* \leq p^*$.

$$g^* = \min_{x \in \mathcal{X}} \quad f(x)$$
$$\text{s.t.} \quad b(x) \leq u.$$

(B)



Proposition. $p^* = g^*$ under the following conditions: (i) f is non-increasing over $\mathcal{X}$, (ii) b is non-decreasing over $\mathcal{X}$, and (iii) the optimal values are attained for both problems.

proof: Suppose $p^* \neq g^*$. Then $g^* < p^*$.

Let x_p^* be the optimal solution of (A), $f(x_p^*) = p^*$.

Let x_g^* be the " " of (B), $f(x_g^*) = g^*$.

Since $g^* < p^* \Rightarrow f(x_g^*) < f(x_p^*)$

$\Rightarrow x_g^* \geq x_p^*$

$\Rightarrow b(x_g^*) \geq b(x_p^*) = u$

$\Rightarrow b(x_g^*) \geq u$

$\Rightarrow b(x_g^*) = u$

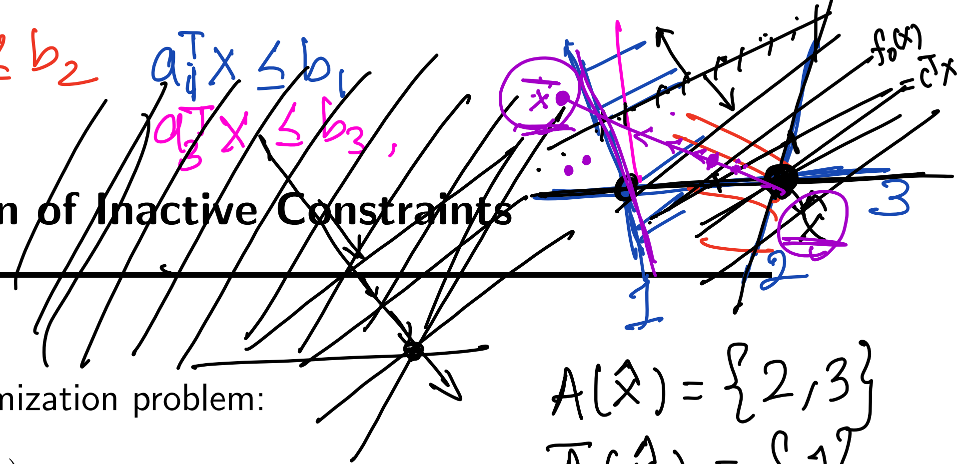
$\Rightarrow x_g^*$ is feasible to (A).

However, $f(x_g^*) < f(x_p^*)$, hence x_p^* is not an optimal solution of (A), which is a contradiction. ∎

$$\left.\begin{aligned} a_2^T \hat{x} &= b_2 \\ a_3^T \hat{x} &= b_3 \end{aligned}\right\}$$

$a_2^T x \le b_2$ $a_1^T x \le b_1$ $a_3^T x \le b_3$,

Elimination of Inactive Constraints

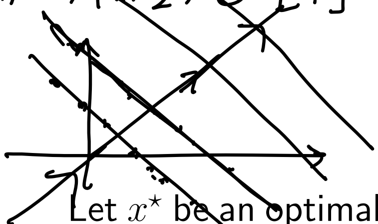


$\mathcal{A}(\hat{x}) = \{2, 3\}$

$\bar{\mathcal{A}}(\hat{x}) = \{1\}$.

Consider an ~~optimization~~ optimization problem:

$f_0(x) = x_1 + x_2$, $c = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$



$$\begin{aligned} \min_{x} \quad & f_o(x) \\ \text{s.t.} \quad & f_i(x) \le 0, \qquad i \in \{1, 2, \ldots, m\}, \\ & Ax = b. \end{aligned} \tag{1}$$

Let $x^\star$ be an optimal solution. We define the set of active and inactive constraints as follows:

$$\mathcal{A}(x^\star) := \{i \in \{1, 2, \ldots, m\} : f_i(x^\star) = 0\}, \text{ : set of active constraints}$$
$$\bar{\mathcal{A}}(x^\star) := \{i \in \{1, 2, \ldots, m\} : f_i(x^\star) < 0\}.$$

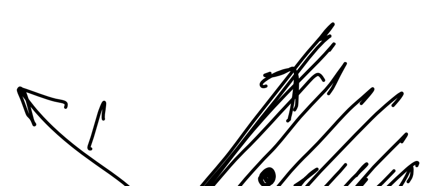
The following proposition says that when the problem is convex, we can remove the inactive constraints without changing the optimal solution.

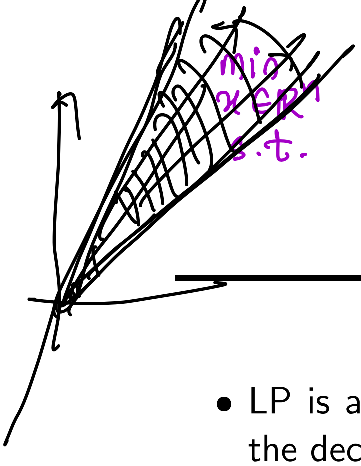
Let (1) be a convex optimization problem.

HW

Proposition. Let $x^\star$ be an optimal solution of (1). Then, $x^\star$ is also an optimal solution of

$$\begin{aligned} \min_{x} \quad & f_o(x) \\ \text{s.t.} \quad & f_i(x) \le 0, \qquad i \in \mathcal{A}(x^\star), \\ & Ax = b. \end{aligned} \tag{2}$$

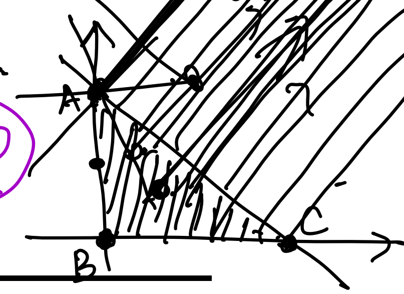




$\min_{x \in \mathbb{R}^n} c^T x$ s.t. $Ax \le b$

$\{x_1 \geq 0, x_2 \geq 0\}$

$[-1\ \ 0]x \le 0, \ [0, -1]x \le 0$



Linear Programming (LP)

$[1\ \ 1]x \le 1$

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

- A polyhedron is intersection of finitely many half-spaces:

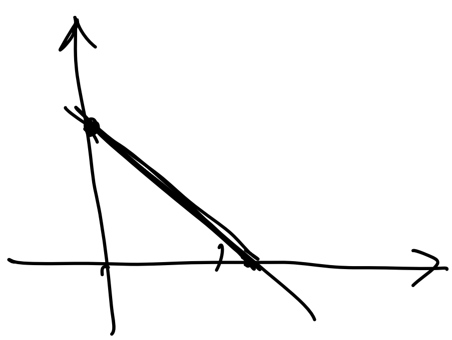
$$H = \{x \in \mathbb{R}^n | Ax \le b\} = \{x \in \mathbb{R}^n | a_i^\top x \le b_i, i = 1, 2, \dots, m\}.$$

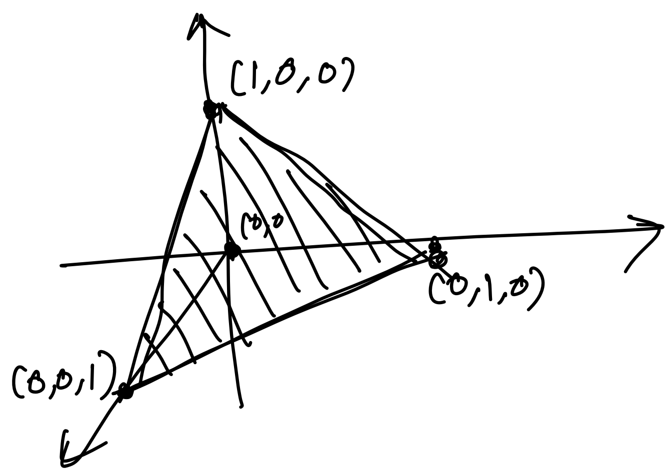
- Any polyhedron can also be represented as the Minkowski sum of the convex hull of a finite number of **extreme points** $\{v_1, v_2, \dots, v_k\}$ and cone generated by a finite number of **extreme rays** $\{r_1, r_2, \dots, r_p\}$, i.e.,

$$x \in H \iff x = \sum_{i=1}^{k} \lambda_i v_i + \sum_{j=1}^{p} \mu_j r_j, \mu_j \ge 0, \lambda_i \ge 0, \sum_{i=1}^{k} \lambda_i = 1.$$

- If a polyhedron is bounded, it is called a polytope, and the set of extreme directions is empty.

- Example: The probability simplex $= \{x \in \mathbb{R}^n \mid x_i \geq 0, \sum_{i=1}^{n} x_i = 1\}$





Linear Programming in Standard Forms

LP in standard equality form:

$$\min_{x \in \mathbb{R}^n} \ c^\top x \quad \text{s.t.} \ \bar{A}x \le \bar{b} \iff$$

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

$$\Updownarrow$$

LP in standard inequality form:

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

$$\begin{aligned} \min_{x\in\mathbb{R}^n} \quad & c^\top x \\ \text{s.t.} \quad & Ax \le b \\ & -Ax \le -b \\ & -x \le 0 \end{aligned}$$

We can easily go from one form to the other. Any LP can be represented in each of the above standard forms.

Example:

$$\begin{aligned} \min_{x\in\mathbb{R}^2} \quad & 3x_1 + 1.5x_2 \\ \text{s.t.} \quad & -1 \le x_1 \le 2, \\ & 0 \le x_2 \le 3. \end{aligned}$$

$$\bar{A} = \begin{bmatrix} A \\ -A \\ -I \end{bmatrix}, \quad \bar{b} = \begin{bmatrix} b \\ -b \\ 0 \end{bmatrix}$$

$Ax \le b.$

$$\Updownarrow$$

$Ax + s = b, \ s \ge 0$, $x = x^+ - x^-, \ x^+, x^- \ge 0$

$$x = \begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 0 \\ 1 \end{bmatrix} - \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}$$

$A[x^+ - x^-] + s = b,$

$x^+ \ge 0, x^- \ge 0, s \ge 0$

$$\begin{aligned} \min_{x^+, x^-, s} \quad & c^\top x^+ - c^\top x^- \\ \text{s.t.} \quad & A[x^+ - x^-] + s = b \\ & x^+ \ge 0, \ x^- \ge 0, \ s \ge 0 \end{aligned}$$

Lecture - 13 : 30th Jan.

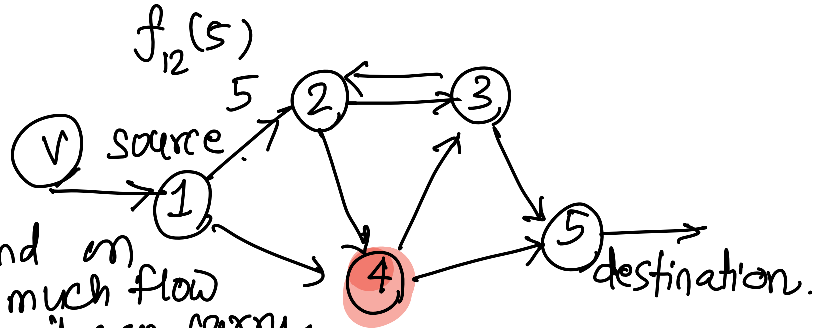
Network Flows

Consider a directed graph $G = (V, E)$

$V = \{v_1, v_2 \ldots v_m\}$ set of vertices

$E = \{(v_1, v_2), (v_3, v_4) \ldots \}$, $|E| = n$

$E = \{ (1,2), (2,3), (3,2),$
$(1,4), (2,4), (4,3),$
$(4,5), (3,5) \}$



For edge e,
- u_e: upper bound on how much flow it can carry.
- l_e: lower bound
- $f_e(\cdot)$: cost function for edge e.

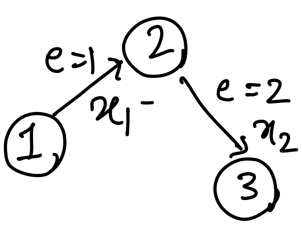
decision variables: $x \in \mathbb{R}^n$, x_e: flow through edge e

cost function: $\sum_{e=1}^{n} f_e(x_e)$

constraints: $l_e \leq x_e \leq u_e \;\; \forall e \in E \iff$ $l \leq x \leq u$, $Bx = \beta = \begin{bmatrix} V \\ 0 \\ \vdots \\ -V \end{bmatrix}$

define incidence matrix $B \in \mathbb{R}^{m \times n}$

ith row of B corresponds to node v_i



$$[B]_{ie} = \begin{cases} +1 & \text{if } e \text{ is an outgoing edge from } i \\ -1 & \text{if } e \text{ is an incoming edge to } i \\ 0 & \text{otherwise.} \end{cases}$$

$$B = \begin{bmatrix} +1 & 0 \\ -1 & +1 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ -x_1 + x_2 \\ -x_2 \end{bmatrix} = \begin{bmatrix} V \\ 0 \\ \vdots \\ 0 \\ -V \end{bmatrix}$$

$Bx \in \mathbb{R}^m$, $[Bx]_i = \sum_{e \in O_i} x_e - \sum_{e \in I_i} x_e$

The minimum cost flow problem can now be written as:

$$\boxed{\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & \sum_{e=1}^{n} f_e(x_e) \\ \text{s.t.} \quad & l \leq x \leq u \\ & \underline{Bx} = \beta . \end{aligned}}$$

If $f_e(\cdot)$ are convex f^ns, then the problem is convex.

Now, let us pose the problem of maximizing the amount of flow that can be sent from source vertex to destination vertex.

$$\max_{x,\, v} \quad v = \underline{c^T \begin{bmatrix} x \\ v \end{bmatrix}}, \quad \underline{c^T = [0 \ldots 0 \,\vdots\, 1]}$$

$$\text{s.t.} \quad l \leq x \leq u$$

$$\boxed{Bx = v\alpha},$$

$$\Updownarrow$$

$$\boxed{Bx - \alpha v = 0}$$

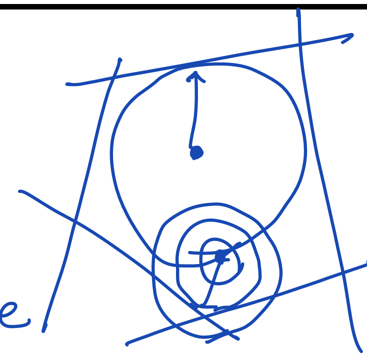
$$\Updownarrow$$

$$\boxed{\begin{bmatrix} B \,\vdots\, -\alpha \end{bmatrix} \begin{bmatrix} x \\ v \end{bmatrix} = 0}$$

$$\alpha = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ \vdots \\ -1 \end{bmatrix}$$

Chebyshev Center of a Polyhedron

For a polyhedron, its chebyshev center is the point that is farthest away from its boundary.



Alternatively, it is the center of the largest ball that is entirely contained inside the polyhedron.

$H = \{x \mid Ax \le b\}$

Problem: For a given H, find its chebyshev center.

decision variables: (y, r), $y \in \mathbb{R}^n$: center, $r \in \mathbb{R}$: radius.

cost function: $-r$

constraints: $y \in H$, $r \geq 0$, if $z \in B(y, r)$, then $Az \le b$.

if $\|z - y\|_2 \le r$, then $Az \le b$.

Let $z - y = v$

$\Rightarrow z = v + y$,

$\Leftrightarrow \|v\|_2 \le r$.

Essentially, we have $Ay + Av \le b$ for every $\|v\|_2 \le r$

The optimization problem can be written as:

Linear program

$$\begin{aligned} \min_{y, r} \quad & -r \\ \text{s.t.} \quad & Ay + Av \le b \quad \forall \|v\|_2 \le r \\ & r \geq 0 \end{aligned}$$

$a_i^T y + r\|a_i\|_2 \le b_i$, $\forall i = \{1, 2, \ldots m\}$

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

$a_i^T y + a_i^T v \le b_i$ $\quad \forall \|v\|_2 \le r$. $\forall i \in \{1, 2, \ldots m\}$

$\Updownarrow$

$a_i^T y + \sup_{\|v\|_2 \le r} a_i^T v \le b_i$, $i \in \{1, 2 .. m\}$

$a_i^T \left[a_i \frac{r}{\|a_i\|_2} \right] = r \frac{\|a_i\|_2^2}{\|a_i\|_2} = r\|a_i\|_2$ the vector that attains the supremum $a_i \cdot \frac{r}{\|a_i\|_2}$

Fundamental Theorem of Linear Programming

optimal value is $-\infty$,

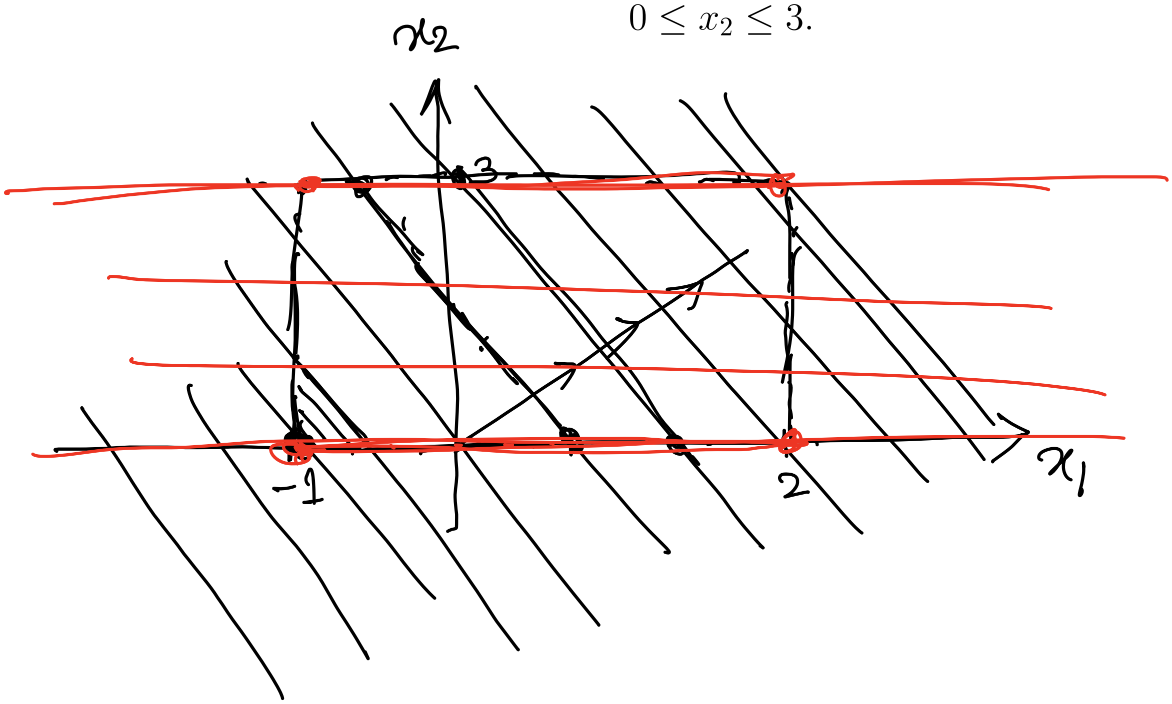
Theorem 1

A linear programming problem is either infeasible, or unbounded or has an optimal solution.

- An intersection of a polytope with a supporting hyperplane is called a face of the polytope. Face includes vertex, edge or facet as special cases.
- The optimal solution lies on a face of the polytope, and is unique if it is on a vertex.
- Example:

$c = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

$$\min_{x \in \mathbb{R}^2} \quad 3x_1 + 1.5x_2$$
$$\text{s.t.} -1 \leq x_1 \leq 2,$$
$$0 \leq x_2 \leq 3.$$



Obtaining a lower bound on the cost function

Consider a LP in the standard equality form:

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

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

$a_1^T x = b_1 \; : \; y_1$

$a_2^T x = b_2 \; : \; y_2$

$\vdots$

$a_m^T x = b_m \; : \; y_m$

$y_i \in \mathbb{R}$, $y \in \mathbb{R}^m$

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

$$y^T A x = y^T b.$$

if we can find $\underline{y \in \mathbb{R}^m}$ s.t. $\boxed{y^T A \leq c^T}$, then

$y^T A x \leq c^T x$ for $\underline{x \geq 0}$ & $Ax = b$.

$\Rightarrow \underline{y^T b} \leq c^T x$

In other words, if $y \in \mathbb{R}^m$ satisfies $y^T A \leq c^T$, then $y^T b \leq \underline{c^T x}$ for every x satisfying $Ax = b$, $x \geq 0$.

Now, let us try to find the "best" lower bound.

$$\boxed{\begin{aligned} \max_{y} \quad & y^T b \\ \text{s.t.} \quad & y^T A \leq c^T \end{aligned}}$$

Finding best possible lower bound

This happens to be another linear program:

$$\max_{y \in \mathbb{R}^m} \quad b^\top y \qquad \text{s.t.} \quad A^\top y \leq c. \tag{D}$$

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).

$A^T y \leq c$: n number of constraints.

$\Rightarrow x^T A^T y \leq x^T c$, $x \geq 0$

If we can find x s.t. $x \geq 0$, $Ax = b$,
then $b^T y \leq c^T x$.

The best upper bound on $b^T y$ is obtained by

$$\boxed{\begin{aligned} \min_{x} \quad & c^T x \\ \text{s.t.} \quad & x \geq 0 \\ & Ax = b \end{aligned}}$$

Lecture 14: 30th Jan

Properties

Theorem 2

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 $c^\top \bar{x} \geq b^\top \bar{y}$.
4. **Strong Duality**: Suppose both (P) and (D) are feasible. Show that for the respective optimal solutions $x^\star$ and $y^\star$, we must have $c^\top x^\star = b^\top y^\star$.

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

Lemma 1 (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:*

S*1.* $\{x \in \mathbb{R}^n \mid Ax = b, x \geq 0\}$

S*2.* $\{y \in \mathbb{R}^m \mid A^\top y \leq 0, b^\top y > 0\}$.

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

(1) Since (P) is infeasible, S_1 is an empty set. From Farka's lemma, we know that S_2 is nonempty. Let $\bar{y} \in S_2$, i.e., $A^\top \bar{y} \leq 0$, $b^\top \bar{y} > 0$

Since (D) is feasible, $\exists\ \hat{y}$ satisfying $A^\top \hat{y} \leq c$

$\hat{y}$ $\quad$ $\hat{y} + \lambda \bar{y}$, $\lambda \geq 0$.

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

$$A^\top(\hat{y} + \lambda \bar{y}) = \underbrace{A^\top \hat{y}}_{\leq c} + \lambda \underbrace{A^\top \bar{y}}_{\leq 0} \leq c.$$

$$b^\top(\hat{y} + \lambda \bar{y}) = b^\top \hat{y} + \lambda b^\top \bar{y}$$

As $\lambda \to \infty$, $\hat{y} + \lambda \bar{y}$ remains feasible to (D), and the value $b^T(\hat{y} + \lambda \bar{y}) \to \infty$. Hence (D) is

Proof an unbounded optimization problem.

---

(2) Suppose (D) is feasible.

$\Rightarrow \exists\ \bar{y}$ s.t. $A^T \bar{y} \leq c$.

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

$\Rightarrow b^T \bar{y} \leq f^*$ which is the optimal value of (P).

Hence $f^* > -\infty$, which contradicts the unboundedness of (P).

Thus, if (P) is unbounded, (D) does not have a feasible solution.

(3) already established.

(4) Let (x^*) be optimal soln of (P).

y^* be optimal soln of (D).

(P) $\min_{x \in \mathbb{R}^n} c^T x$ s.t. $Ax = b$, $x \geq 0$ (D) $\max_{y \in \mathbb{R}^m} b^T y$ s.t. $A^T y \leq c$

Then,

$$\left.\begin{array}{l} Ax^* = b,\ x^* \geq 0 \\ A^T y^* \leq c \quad \checkmark \\ c^T x^* \leq b^T y^* \end{array}\right] \text{should have a solution.}$$

We will express the above conditions as

$S_1 = \{\bar{x} \in \mathbb{R}^x \mid \bar{A}\bar{x} = \bar{b},\ \bar{x} \geq 0\}$

To argue that S_1 is non-empty, we need to show

$S_2 = \{\bar{z} \in \mathbb{R}^y \mid \bar{z}^T \bar{A} \leq 0,\ \bar{b}^T \bar{z} > 0\}$.

is an empty set.

Step-1: Express (*) in terms of S_1.

$$\begin{cases} Ax = b \\ A^T(y_+ - y_-) + s_1 = c \\ c^T x - b^T y_+ + b^T y_- + s_2 = 0 \end{cases}$$

Proof

$[x, y_+, y_-, s_1, s_2] \geq 0$

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

$\bar{A} \in \mathbb{R}^{(n+m+1)\times(2n+2m+1)}$

$$\overbrace{\begin{bmatrix} A & 0 & 0 & 0 & 0 \\ 0 & A^T & -A^T & I & 0 \\ c^T & -b^T & b^T & 0 & 1 \end{bmatrix}}^{\bar{A}} \begin{bmatrix} x \\ y_+ \\ y_- \\ s_1 \\ s_2 \end{bmatrix} = \overbrace{\begin{bmatrix} b \\ c \\ 0 \end{bmatrix}}^{\bar{b}}$$

Step-2: Assume on the contrary that $\bar{z} \in S_2$; $\bar{z} = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix}$.

$\underbrace{[z_1^T : z_2^T : z_3^T]}_{\in \mathbb{R}} \left[\quad \right] \leq 0$, $\bar{z}^T \bar{b} > 0$

$z_1^T A + z_3 c^T \leq 0$

$\boxed{z_1^T b + z_2^T c > 0}$

$$\left.\begin{aligned} z_2^T A^T - z_3 b^T &\leq 0 \\ -z_2^T A^T + z_3 b^T &\leq 0 \end{aligned}\right\} \Rightarrow z_2^T A^T = z_3 b^T \iff A z_2 = z_3 b$$

$z_2^T \leq 0$

$z_3 \leq 0$

Case (a): $z_3 < 0$. Claim 1 $\left(\frac{z_2}{+z_3}\right)$ is feasible to (P).

Since $z_2 \leq 0$, $z_3 < 0$,

Then: $\frac{z_2}{+z_3} \geq 0$.

$A\left(z_2/_{+z_3}\right) = b$

claim 2: $\frac{z_1}{-z_3}$ is feasible to (D).

$z_1^T A \leq -z_3 c^T$

$\Rightarrow \begin{pmatrix} z_1 \\ -z_3 \end{pmatrix}^T A \leq c^T$

Let us check whether weak duality holds for $\begin{pmatrix} z_2 \\ z_3 \end{pmatrix}$ & $\begin{pmatrix} z_1 \\ -z_3 \end{pmatrix}$.

$$c^T \begin{pmatrix} z_2 \\ z_3 \end{pmatrix} \geq b^T \begin{pmatrix} z_1 \\ -z_3 \end{pmatrix}$$

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

$$\Rightarrow c^T z_2 + b^T z_1 \leq 0$$

$\Rightarrow$ $c^T z_2 + b^T z_1 > 0$ is not satisfied.

case (b) : $z_3 = 0$

$$\begin{cases} z_1^T A \leq 0 \\ A z_2 = 0, \; z_2 \leq 0, \\ z_1^T b + z_2^T c > 0 \end{cases}$$

Let $\tilde{x}$ be a feasible to (P)

$A\tilde{x} = b$, $\tilde{x} \geq 0$

$\tilde{x} - \lambda z_2$, $\lambda \geq 0$, are feasible to (P).

$c^T(\tilde{x} - \lambda z_2) \geq b^T \tilde{y}$

$\Rightarrow c^T \tilde{x} - \lambda c^T z_2 \geq b^T \tilde{y}$ for $\tilde{y}$ feasible to (D) $\forall \lambda \geq 0$ $\Rightarrow$ $c^T z_2 < 0$

Let $\tilde{y}$ be a feasible sol^n to (D).

$\Rightarrow A^T \tilde{y} \leq c \quad \Rightarrow A^T(\tilde{y} + \lambda z_1) \leq c \quad \forall \lambda \geq 0$

weak duality implies $\quad c^T \tilde{x} \geq b^T(\tilde{y} + \lambda z_1) \quad \forall \lambda \geq 0$

$$\Rightarrow b^T z_1 < 0$$

Together $c^T z_2 + b^T z_1 < 0$, which is a contradiction.

$\Rightarrow$ The set S_2 is an empty set.

$\Rightarrow$ The set S_1 is non-empty, and any element of S_1 specifies a pair of (x, y) where x is feasible to (P) y is feasible to (D) & $c^T x = b^T y$.

$\Rightarrow$ both are optimal & Strong duality holds.

Lagrangian Function

Consider the following optimization problem in standard form:

$$\min_{x \in \mathbb{R}^n} \quad f(x)$$
$$\text{s.t.} \begin{cases} g_i(x) \le 0, i \in [m] := \underline{\{1, 2, \dots, m\}}, & : \lambda_i \ , \lambda \in \mathbb{R}^m \\ h_j(x) = 0, j \in [p]. & : \mu_j \ , \mu \in \mathbb{R}^p \end{cases}$$

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

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

where

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

Lower Bound Property:

Lemma 2. *If $\bar{x}$ is feasible and $\bar{\lambda} \ge 0$, then $f(\bar{x}) \ge \underline{L(\bar{x}, \bar{\lambda}, \bar{\mu})}$.*

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

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

or $$f(\bar{x}) \ge \underline{L(\bar{x}, \bar{\lambda}, \bar{\mu})} \ge \underline{\inf_{x \in \mathbb{R}^n} L(x, \bar{\lambda}, \bar{\mu})} =: d(\bar{\lambda}, \bar{\mu})$$

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

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

- $d(\lambda, \mu)$ 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.

Recall: pointwise supremum of convex functions is convex.

Thus, it can be easily shown that pointwise infimum of concave functions is concave.

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

For a fixed value of x, $L(x, \lambda, \mu)$ is affine in (λ, μ).

$\Rightarrow L(x, \lambda, \mu)$ is concave in (λ, μ) for a fixed x.

Thus, $\inf_x L(x, \lambda, \mu)$ is concave in (λ, μ).

Lagrangian Dual Optimization Problem

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

$$\begin{cases} \max\limits_{\lambda \in \mathbb{R}^m, \mu \in \mathbb{R}^p} & d(\lambda, \mu) \\ \text{s.t.} & \lambda \geq 0, \\ & (\lambda, \mu) \in \texttt{dom}(d). \end{cases}$$

- The above is a convex optimization problem since $d(\lambda, \mu)$ is concave in λ and μ irrespective of whether the original problem is convex or not.
- Let the optimal value of the dual optimization problem be denoted d^*.
- $\delta^* = f^* - d^*$ is called the duality gap.
- 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 1: Lagrangian Dual of LP

$$\min_{x \in \mathbb{R}^n} \quad c^\top x$$
$$\text{s.t.} \quad Ax = b, \cancel{x \geq 0}. \quad : \mu$$
$$-x \leq 0 \quad : \lambda$$

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

$$L(x,\lambda,\mu) = c^T x + \lambda^T(-x) + \mu^T[Ax - b]$$
$$= [c^T - \lambda^T + \mu^T A]x - \mu^T b$$

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

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

Dual optimization problem:

$$\left[\begin{array}{ll} \max_{\lambda,\mu} & -\mu^T b \\ \text{s.t.} & \lambda \geq 0 \\ & c^T - \lambda^T + \mu^T A = 0 \end{array}\right.$$

$$c^T + \mu^T A = \lambda^T \geq 0$$

$$y = -\mu.$$

$$c^T \geq -\mu^T A$$
$$= y^T A$$

$$\boxed{\begin{array}{ll} \max_{y} & b^T y \\ \text{s.t.} & y^T A \leq c^T \end{array}}$$

Example 2: Least Norm Solution

Least norm solution:

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

Find L and d.

$0 \leq x \leq 1$

Example 3

optimal solution : 1
optimal value : -1

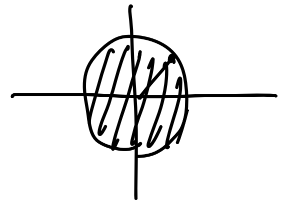
$$\min_{x\in\mathbb{R}} \quad -x^2$$
$$\text{s.t.} \quad x-1 \leq 0, \quad -x \leq 0.$$

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

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

$$d(\lambda_1,\lambda_2) = \inf_x L(x,\lambda_1,\lambda_2)$$
$$= \inf_{x\in\mathbb{R}} \left[-x^2 + (\lambda_1-\lambda_2)x - \lambda_1\right]$$
$$= -\infty$$

- strong duality does not hold.



Example 4

optimal value = -1

$$\max \quad x_1^2 + x_2^2$$

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

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

$$L([x_1, x_2], \lambda) = -x_1^2 - x_2^2 + \lambda(x_1^2 + x_2^2 - 1)$$

$$= (\lambda - 1)(x_1^2 + x_2^2) - \lambda$$

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

$$= \begin{cases} -\lambda, & \text{if } \lambda \geq 1 \\ -\infty, & \text{if } \lambda < 1. \end{cases}$$

<u>Dual problem</u>: $\max_{\lambda} \ -\lambda$ s.t. $\lambda \geq 1$, $\lambda \geq 0$ $\Rightarrow$ optimal solution: +1, optimal value: -1