Introduction to Optimization

CO 255

Ricardo Fukasawa

LaTeXed by Sibelius Peng

Preface

Disclaimer Much of the information on this set of notes is transcribed directly/indirectly from the lectures of CO 255 during Winter 2020 as well as other related resources. I do not make any warranties about the completeness, reliability and accuracy of this set of notes. Use at your own risk.

For any questions, send me an email via `https://notes.sibeliusp.com/contact/`.

You can find my notes for other courses on `https://notes.sibeliusp.com/`.

Sibelius Peng

Contents

0

Info

Ricardo: MC 5036. OH: M 1:30 - 3pm
TA: Adam Brown: MC 5462. OH: F 10-11am

Books (not required)

- Intro to Linear Opt. Bertsimas
- Int Programming. Conforti

Grading

- assns: 20% (≈ 5)
- mid: 30% (Feb 11 in class)
- final: 50%

1

Introduction

Given a set S, and a function $f : S \to \mathbb{R}$. An optimization problem is:

$$\begin{aligned} &\max f(x) \\ &\underbrace{s.t.}_{\text{subject to}} \quad x \in S \end{aligned} \qquad \text{(OPT)}$$

- S **feasible region**
- A point $\overline{x} \in S$ is a **feasible solution**
- $f(x)$ is **objective function**

(OPT) means: "Find a feasible solution x^* such that $f(x) \leq f(x^*), \forall x \in S$"

- Such x^* is an **optimal solution**
- $f(x^*)$ is **optimal value**

Other ways to write (OPT):

$$\max\{f(x), x \in S\}$$

$$\max_{x \in S} f(x)$$

Analogous problem

$$\begin{aligned} &\min f(x) \\ &s.t. \quad x \in S \end{aligned}$$

> **Note**
>
> $$\begin{aligned} &\max f(x) \\ &s.t. \quad x \in S \end{aligned} = -1 \begin{pmatrix} \min -f(x) \\ s.t. \quad x \in S \end{pmatrix}$$

Problem x^* may not exist

a) Problem is unbounded:

$$\forall M \in \mathbb{R}, \exists \overline{x} \in S, \ s.t. \ f(\overline{x}) > M$$

b) $S = \phi$, i.e. (OPT) is **INFEASIBLE**

c) There may not exist x^* achieving supremum.

> Example:
>
> $$\begin{array}{ll} \max & x \\ \text{s.t.} & x < 1 \end{array}$$

supremum

$$\sup\{f(x) : x \in S\} = \begin{cases} +\infty & \text{if OPT unbounded} \\ -\infty & \text{if } S = \varnothing \\ \min\{x : x \geq f(x), \forall x \in S\} & \text{otherwise} \end{cases}$$

always exist and are well-defined

infimum

$$\inf\{f(x) : x \in S\} = -1 \cdot \sup\{-f(x) : x \in S\}$$

From this point on, we will abuse notation and say $\max\{f(x) : x \in S\}$ is $\sup\{f(x) : x \in S\}$.

One way to specify that I want an opt. sol. (if exists) is

$$x^* \in \text{argmax}\{f(x) : x \in S\}$$

2 Linear Optimization (Programming) (LP)

$$S = \{x \in \mathbb{R}^n : Ax \leq b\}$$

where $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$ and $f(x) = c^T x, c \in \mathbb{R}^n$.

$$\downarrow$$

$$\begin{aligned} &\max c^T x \\ &s.t.\ Ax \leq b \end{aligned} \qquad (LP)$$

> **Note**
>
> $$A = \begin{pmatrix} | & & | \\ A_1 & \cdots & A_n \\ | & & | \end{pmatrix} \qquad A = \begin{pmatrix} - & a_1^T & - \\ & \vdots & \\ - & a_m^T & - \end{pmatrix}$$

Clarifying

$$u, v \in \mathbb{R}^n, \qquad u \leq v \iff u_j \leq v_j, \forall j \in 1, \dots, n$$

> **Note**
>
> $u \nleq v$ is not the same as $u > v$
>
> $$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \nleq \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

> Example:
>
> $$\begin{array}{llrl} \max & 2x_1 + & 0.5x_2 & \\ s.t. & x_1 & & \leq 2 \\ & x_1 + & x_2 & \leq 2 \\ & x & & \geq 0 \end{array}$$

- Strict ineq. not allowed

halfspace, hyperplane, polyhedron

Let $h \in \mathbb{R}^n, h_0 \in \mathbb{R}$.

$\{x \in \mathbb{R}^n : h^T \leq h_0\}$ is a **halfspace**.

$\{x \in \mathbb{R}^n : h^T = h_0\}$ is a **hyperplane**.

$Ax \leq b$ is a **polyhedron** (i.e. intersection of finitely many halfspaces).

Example:

n products, m resources. Producing $j \in \{1, \ldots, n\}$ given c_j profit/unit and consumes a_{ij} units of resource i, $\forall i \in \{1, \ldots, m\}$. There are b_i units available $\forall i \in \{1, \ldots, m\}$.

$$\begin{aligned} \max \sum_{j=1}^{n} c_j x_j & \\ s.t. \quad \sum_{j=1}^{n} a_{ij} x_j \leq b_i, & \qquad \forall i = 1, \ldots, m \\ x \geq 0 & \end{aligned}$$

which is an LP.

2.1 Determining Feasibility

Given a polyhedron

$$P = \{x \in \mathbb{R}^n : Ax \leq b\}$$

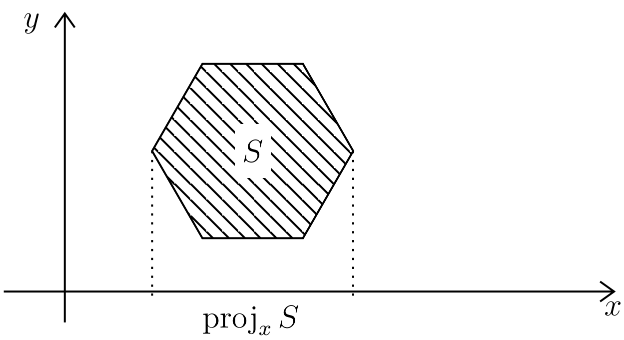
either find $\overline{x} \in P$ or show $P = \varnothing$.

Idea In 1-d, easy. $\rightarrow$ Reduce problem in dimension n to one in dimension $n-1$.

Notation Let $S = \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^p : Ax + Gy \leq b\}$, then

$$\operatorname{proj}_x S := \{x \in \mathbb{R}^n : \exists y \text{ so that } (x, y) \in S\}$$

is the (orthogonal) *projection* if S onto x.



We will find if $P = \varnothing$ by looking at $\text{proj}_{x_1,\dots,x_{n-1}}$ (P)

2.2 Fourier-Motzkin Elimination

Call a_{ij} entries of A. Let

$$
\begin{aligned}
M &:= \{1, 2, \dots, m\} \\
M^+ &:= \{i \in M : a_{in} > 0\} \\
M^- &:= \{i \in M : a_{in} < 0\} \\
M^0 &:= \{i \in M : a_{in} = 0\}
\end{aligned}
$$

For $i \in M^+$ (1):

$$a_i^T \le b_i \iff \sum_{j=1}^{n} a_{ij} x_j \le b_i \iff \sum_{j=1}^{n-1} \frac{a_{ij}}{a_{in}} x_j + x_n \le \frac{b_i}{a_{in}}, \qquad \forall i \in M^+$$

For $i \in M^-$ (2):

$$a_i^T x \le b_i \iff \sum_{j=1}^{n-1} \frac{a_{ij}}{a_{in}} x_j - x_n \le \frac{b_i}{-a_{in}}, \qquad \forall i \in M^-$$

For $i \in M^0$ (3):

$$a_i^T x \le b_i \iff \sum_{j=1}^{n-1} a_{ij} x_j \le b_i, \qquad \forall i \in M^0$$

$$P = \{x \in \mathbb{R}^n : (1)(2)(3)\}$$

Define (4):

$$\sum_{j=1}^{n-1} \left(\frac{a_{ij}}{a_{in}} - \frac{a_{kj}}{a_{kn}} \right) x_j \le \frac{b_i}{a_{in}} - \frac{b_i}{a_{kn}}, \qquad \forall i \in M^+, \forall k \in M^-$$

Theorem 2.1

$$(\overline{x}_1, \dots, \overline{x}_{n-1}) \text{ satisfies (3), (4)} \iff \exists \overline{x}_n : (\overline{x}_1, \dots, \overline{x}_n) \in P$$

Proof:

$\Longleftarrow$ If $(\overline{x}_1, \dots, \overline{x}_n)$ satisfies (1), (2), (3) then $(\overline{x}_1, \dots, \overline{x}_{n-1})$ satisfies (3) and adding (1), (2) $\implies$ $(\overline{x}_1, \dots, \overline{x}_{n-1})$ satisfies (4)

$\implies$ If $(\overline{x}_1, \dots, \overline{x}_{n-1})$ satisfies (4)

$$\sum_{j=1}^{n-1} \frac{a_{ij}}{a_{in}} \overline{x}_j - \frac{b_i}{a_{in}} \le \sum_{j=1}^{n-1} \frac{a_{kj}}{a_{kn}} \overline{x}_j - \frac{b_k}{a_{kn}}, \quad \forall i \in M^+, k \in M^-$$

Let

$$\overline{x}_n := \max_{i \in M^+} \left\{ \sum_{j=1}^{n-1} \frac{a_{ij}}{a_{in}} \overline{x}_j - \frac{b_i}{a_{in}} \right\}$$

$$\implies \sum_{j=1}^{n-1} \frac{a_{ij}}{a_{in}} \overline{x}_j - \frac{b_i}{a_{in}} \le -\overline{x}_n, \quad \forall i \in M^+$$

and

$$-\overline{x}_n \le \sum_{j=1}^{n-1} \frac{a_{kj}}{a_{kn}} \overline{x}_j - \frac{b_k}{a_{kn}}, \qquad \forall k \in M^-$$

$$\implies (\overline{x}_1, \dots, \overline{x}_n) \in P$$

□

Note

Proof assumes M^+, M^- are nonempty. But statement holds regardless.

(if M^+ or $M^- = \varnothing$ then (4) yields no constraints)

Fourier-MotzKin

- $A^n = A, b^n = b$
- given A^i, b^i obtain A^{i-1}, b^{i-1} (A^{i-1} has one less column than A^i column than A^i) by applying the steps described

  $$P_i := \{x \in \mathbb{R}^i : A^i x \le b^i\}$$

  then

  $$P_{i-1} = \operatorname{proj}_{x_1, \dots, x_{i-1}} P_i$$

  and $P_{i-1} = \varnothing \iff P_i = \varnothing$.
- Keep applying projection until $i = 1$.

$$P_0 = \varnothing \iff P_n = P = \varnothing$$

Let

$$P_i^n = P_i \times \mathbb{R}^{n-i} = \{x \in \mathbb{R}^n (A^i, 0) x \le b^i\}$$

not hard to see $P_i^n = \varnothing \iff P_i = \varnothing$

Notice that

$$P_0 = \varnothing \iff P_0^n = \varnothing, P_0^n = \{0 \le b^0\}$$

Example:

$$P_2 = \left\{ x \in \mathbb{R}^2 : \begin{array}{ccc} x_1 & +x_2 & \leq 1 \\ -x_1 & & \leq 0 \\ & -x_2 & \leq -2 \\ -3x_1 & -3x_2 & \leq -6 \end{array} \right\}$$

draw the graph, clearly empty

M^+: $\frac{1}{2}x_1 + x_2 \leq \frac{1}{2}$

M^-: $-x_2 \leq -2 \qquad -x_1 - x_2 \leq -2$

M^0: $-x_1 \leq 0$

$$P_1 = \left\{ x_1 \in \mathbb{R} : \begin{array}{cc} -x_1 & \leq 0 \\ \frac{1}{2}x_1 & \leq -\frac{3}{2} \\ -\frac{1}{2}x_1 & \leq -\frac{3}{2} \end{array} \right\}$$

M^+: $x_1 \leq -3$

M^-: $-x_1 \leq 0$ and $-x_1 \leq -3$

$$P_0^2 = \left\{ x \in \mathbb{R}^2 : \begin{array}{c} 0 \leq -3 \\ 0 \leq -6 \end{array} \right\} = \varnothing$$

Here $b^0 = \binom{-3}{-6}$

Remark:
Inequality in P_i^n:

- All inequalities are obtained by a nonnegative combination of inequality in P_{i+1}^n
  $\implies$ all nonnegative combination of inequalities in P.
- If all A, b are rational then so are all A^i, b^i
- If $b = 0, b_i = 0, \forall i$

Theorem 2.2: Farkas' Lemma

$$P = \{x \in \mathbb{R}^n : Ax \leq b\} = \varnothing \iff \exists u \in \mathbb{R}^m : \begin{array}{l} u^T A = 0 \\ u^T b < 0 \\ u \geq 0 \end{array}$$

Proof:

$(\Longleftarrow)$ Suppose $\overline{x}$ satisfies $A\overline{x} \leq b$.

$$0 = u^T A\overline{x} \leq u^T b < 0$$

which is impossible.

$(\Longrightarrow)$ If $P = \varnothing$. Apply Fourier-Motzkin until we get

$$P_0^n = \varnothing = \{x \in \mathbb{R}^n : 0x \leq b^0\}$$

i.e. there exists j for which $b_j^0 < 0$.

If we look at corresponding constraint in P_0^n is

$$0^T x \leq b_j^0$$

which can be obtained by a vector u such that $u^T A = 0, u^T b = b_j^0, u \geq 0$.

□

Farkas' Lemma (alternate statement)

Exactly one of the following has a solution:

a) $Ax \leq b$

b) $\begin{aligned} u^T A &= 0 \\ u^T b &< 0 \\ u &\geq 0 \end{aligned}$

Farkas' Lemma (Different Form)

Exactly one of the following has a solution:

a) $\begin{aligned} Ax &= b \\ x &\geq 0 \end{aligned}$

b) $\begin{aligned} u^T A &\geq 0 \\ u^T b &< 0 \end{aligned}$

Proof:

(Sketch)

$$P = \left\{ x : \begin{aligned} Ax &= b \\ x &\geq 0 \end{aligned} \right\} = \left\{ x : \underbrace{\begin{pmatrix} A \\ -A \\ -I \end{pmatrix}}_{A'} x \leq \underbrace{\begin{pmatrix} b \\ -b \\ -0 \end{pmatrix}}_{b'} \right\}$$

Apply original Farkas' Lemma to get $P = \varnothing \iff \exists u_1 \in \mathbb{R}^m, u_2 \in \mathbb{R}^m, v \in \mathbb{R}^n$:

$$\begin{aligned} u_1^T A - u_2^T A - v = 0 \\ u_1^T b - u_2^T b < 0 \\ u_1, u_2, v \geq 0 \end{aligned}$$

Let $u = (u_2 - u_2)$

$$u^T A - v = 0 \implies u^T A \geq 0, \quad u^T b < 0$$

$\square$

Consider a linear programming (LP):

$$\begin{aligned} \max\ & c^T x \\ s.t.\ & Ax \leq b \end{aligned} \qquad (LP)$$

Theorem 2.3: Fundamental Theorem of Linear Programming

(LP) has exactly one of 3 outcomes:

a) Infeasible

b) Unbounded

c) There exists an optimal solution.

Proof:
Let's assume a), b) don't hold.

If $n = 1$, then (LP) has an optimal solution. (Why?)

Else, define

$$\begin{aligned} \max z& \\ s.t.\ \ z - c^T x &\leq 0 \\ Ax &\leq b \end{aligned} \qquad (LP')$$

(LP') is also not in case a) or b). (Why?)

Also if (x^*, z^*) is an optimal solution to (LP'), then x^* is an optimal solution to (LP). (Why?)

Apply Fourier-Motzkin to

$$\left\{ (x,z) : \begin{aligned} z - c^T x \leq 0 \\ Ax \leq b \end{aligned} \right\}$$

Until we are left with a polyhedron

$$\{z \in \mathbb{R} : A'z \leq b'\}$$

Now $\max z \qquad s.t\ \ A'z \leq b'$ is not cases a) or b). (Why?)

→ can get an optimal solution z^* to such problem. Apply Fourier-Motzkin back to get (x^*, z^*) optimal solution to (LP'). (Why?) □

2.3 Certifying Optimality

$$\begin{aligned} \max\ & c^T x \\ s.t\ & Ax \le b \end{aligned} \qquad (LP)$$

and let $\overline{x} \in P = \{x : Ax \le b\}$

Question Can we certify that $\overline{x}$ is optimal?

Example:

$$\begin{aligned} \max\ & 2x_1 + x_2 \\ & x_1 + 2x_2 \le 2 \\ s.t.\ & x_1 + x_2 \le 2 \\ & x_1 - x_2 \le 0.5 \end{aligned}$$

Consider $\overline{x} = (0, 1)^T$ is clearly NOT optimal.

$x^* = (1, 0.5)^T$ and $c^T x^* = 2.5$. Any feasible solution satisfies

$$\begin{array}{rll} x_1 + 2x_2 & \le 2 & \times 1/3 \\ x_1 + x_2 & \le 2 & \times 1 \\ +\ \ x_1 - x_2 & \le 0.5 & \times 2/3 \\ \hline 2x_1 + x_2 & \le 3 & \end{array}$$

Instead do $1 \times 1st$ constraint $+ 1 \times 3rd$ constraint $\implies 2x_1 + x_2 \le 2.5$

In general:

$$\begin{array}{rll} x_1 + 2x_2 & \le 2 & \times y_1 \\ x_1 + x_2 & \le 2 & \times y_2 \\ +\ \ x_1 - x_2 & \le 0.5 & \times y_3 \\ \hline \end{array}$$

$$(y_1 + y_2 + y_3)x_1 + (2y_1 + y_2 - y_3)x_2 \le 2y_1 + 2y_2 + 0.5y_3$$

As long as $y_1, y_2, y_3 \ge 0$ and

$$\begin{aligned} y_1 + y_2 + y_3 &= 2 \\ 2y_1 + y_2 - y_3 &= 1 \end{aligned}$$

This leads to the following linear program:

$$\begin{aligned} \min\ & 2y_1 + 2y_2 + 0.5y_3 \\ & y_1 + y_2 + y_3 = 2 \\ s.t.\ & 2y_1 + y_2 - y_3 = 1 \\ & y_1, y_2, y_3 \ge 0 \end{aligned}$$

This is called the dual LP.

In general:

$$\begin{aligned} \max\ & c^T x \\ s.t.\ & Ax \leq b \end{aligned} \qquad (P)$$

Dual of (P)

$$\begin{aligned} \min\ & b^T y \\ s.t.\ & \begin{aligned} y^T A &= c^T \\ y &\geq 0 \end{aligned} \end{aligned} \qquad (D)$$

Remark:
We call (P) primal LP.

Theorem 2.4: Weak Duality

Let $\overline{x}$ feasible for (P), $\overline{y}$ feasible for (D). Then $c^T x \leq b^T y$.

Proof:

$$c^T \overline{x} = \overline{y}^T (A\overline{x}) \leq \overline{y}^T b$$

where we used $A\overline{x} \leq b$ and $\overline{y} \geq 0$. $\square$

Corrollary 2.5

Several results:

- If (P) is unbounded then (D) is infeasible.
- If (D) is unbounded then (P) is infeasible.

Note
(P) and (D) can both be infeasible.

- If $\overline{x}$ is feasible for (P) $\overline{y}$ feasible for (D) $c^T\overline{x} = b^T\overline{y}$, then $\overline{x}$ optimal for (P), $\overline{y}$ optimal for (D).

Theorem 2.6: Strong Duality

x^* is optimal for (P) $\iff$ $\exists y^*$ feasible for (D) such that $c^T x^* = b^T y^*$.

Proof:
($\Longleftarrow$) ✓

($\Longrightarrow$) Is (D) infeasible?

$$\text{Suppose } \left\{ y \in \mathbb{R}^n : \begin{aligned} A^T y &= c \\ y &\geq 0 \end{aligned} \right\} = \varnothing$$

(Alternate version of Farkas' Lemma) $\exists u : \begin{matrix} u^T A \geq 0 \\ u^T c < 0 \end{matrix} \iff \exists d : \begin{matrix} Ad \leq 0 \\ c^T d > 0 \end{matrix}$

Take look at $x' = x^* + d$, then

$$\begin{aligned} Ax' &= Ax^* + Ad \leq b \\ c^T x' &= c^T x^* + c^T d > c^T x^* \end{aligned}$$

Contradiction. Thus (D) has an optimal solution y^*.

Now let $\gamma = b^T y^*$, and let $\theta := \left\{ x \in \mathbb{R}^n : \begin{matrix} Ax \leq b \\ -c^T x \leq -\gamma \end{matrix} \right\}$.

If $\theta = \varnothing$, by Farkas'

$$\exists \begin{pmatrix} \overline{y} \\ \overline{\lambda} \end{pmatrix} : \begin{cases} \begin{pmatrix} \overline{y} \\ \overline{\lambda} \end{pmatrix}^T \begin{pmatrix} A \\ -c^T \end{pmatrix} = 0 \\ \begin{pmatrix} \overline{y} \\ \overline{\lambda} \end{pmatrix}^T \begin{pmatrix} b \\ -\gamma \end{pmatrix} < 0 \\ \begin{pmatrix} \overline{y} \\ \overline{\lambda} \end{pmatrix} \geq 0 \end{cases} \iff \begin{matrix} A^T \overline{y} = c\overline{\lambda} \\ b^T \overline{y} < \gamma \overline{\lambda} \\ \overline{y} \geq 0 \\ \overline{\lambda} \geq 0 \end{matrix}$$

<u>Case 1:</u> $\overline{\lambda} > 0$.

Let $y' = \frac{\overline{y}}{\overline{\lambda}}$. Then we have

$$A^T y' = A^T \frac{\overline{y}}{\overline{\lambda}} = c \quad \text{and} \quad b^T y' = b^T \frac{\overline{y}}{\overline{\lambda}} < \gamma \quad \text{and} \quad y' = \frac{\overline{y}}{\overline{\lambda}} \geq 0$$

Contradicts optimality of y^*.

<u>Case 2:</u> $\overline{\lambda} = 0$. Then $\begin{matrix} A^T y = 0 \\ b^T y < 0 \\ \overline{y} \geq 0 \end{matrix}$

Now we can do the same thing previously. Let $y' = y^* + \overline{y}$, then

$$A^T y' = A^T y^* + A^T \overline{y} = c$$

and

$$\begin{aligned} y' &= y^* + \overline{y} \geq 0 \\ b^T y' &= b^T y^* + b^T \overline{y} < b^T y^* \end{aligned}$$

Contradicts optimality of y^*.

Thus $\theta \neq \varnothing$.

Let $\overline{x} \in \theta$,

$$c^T x^* \underbrace{\leq}_{\text{weak duality}} b^T y^* = \gamma \underbrace{\leq}_{\overline{x}\in\theta} c^T \overline{x} \leq c^T x^*$$

where the last inequality is because $\overline{x}$ feasible for (P), x^* optimal for (P).

□

2.4 Possible Outcomes

See here.

2.5 Duals of generic LPs

$$\begin{array}{ll}
\max & 2x_1 + 3x_2 - 4x_3 \\
\text{s.t.} & \begin{array}{llll}
x_1 & & +7x_3 & \leq 5 \\
 & 2x_2 & -x_3 & \geq 3 \\
x_1 & & +x_3 & = 8 \\
 & x_2 & & \leq 6 \\
x_1 & & & \geq 0 \\
 & x_2 & & \leq 0
\end{array}
\end{array}$$

$$\begin{array}{ll}
\max & (2, 3, -4)x \\
\text{s.t.} & \begin{pmatrix} 1 & 0 & 7 \\ 0 & -2 & 1 \\ 1 & 0 & 1 \\ -1 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} x \leq \begin{pmatrix} 5 \\ -3 \\ 8 \\ -8 \\ 6 \\ 0 \\ 0 \end{pmatrix}
\end{array}$$

and dual

$$\begin{array}{ll}
\min & (5, -3, 8, -8, 6, 0, 0)y \\
\text{s.t.} & \begin{pmatrix} 1 & 0 & 1 & -1 & 0 & -1 & 0 \\ 0 & -2 & 0 & 0 & 1 & 0 & 1 \\ 7 & 1 & 1 & -1 & 0 & 0 & 0 \end{pmatrix} y = \begin{pmatrix} 2 \\ 3 \\ -4 \end{pmatrix} \quad \text{and} \quad y \geq 0
\end{array} \qquad (D_1)$$

$$\begin{array}{ll}
\min & (5, -3, 8, -8, 6)y \\
\text{s.t.} & \begin{pmatrix} 1 & 0 & 1 & -1 & 0 \\ 0 & -2 & 0 & 0 & 1 \\ 7 & 1 & 1 & -1 & 0 \end{pmatrix} y \begin{array}{c} \geq \\ \leq \\ = \end{array} \begin{pmatrix} 2 \\ 3 \\ -4 \end{pmatrix} \quad \text{and} \quad y \geq 0
\end{array} \qquad (D_2)$$

Claim $(y_1^*, \dots, y_5^*)$ is optimal for (D_2) $\iff$ $(y_1^*, \dots, y_5^*, y_6^*, y_7^*)$ optimal for (D_1) with

$$\begin{aligned} y_6^* &= y_1^* + y_3^* - y_4^* - 2 \\ y_7^* &= 3 - (-2y_2^* + y_5^*) \end{aligned}$$

$$\begin{array}{ll} \min & (5,3,8,6)y \\ \text{s.t.} & \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 2 & 0 & 1 \\ 7 & -1 & 1 & 0 \end{pmatrix} y \begin{array}{c} \geq \\ \leq \\ = \end{array} \begin{pmatrix} 2 \\ 3 \\ -4 \end{pmatrix} \quad \text{and} \quad y_1 \geq 0, y_2 \leq 0 \qquad y_4 \geq 0 \end{array} \qquad (D_3)$$

Claim Opt value of (D_2) and (D_3) are same.

In general

$$\begin{array}{ll} \max & c^T x \\ \text{s.t.} & \begin{array}{r} Ax?b \\ x?0 \end{array} \end{array} \quad (P) \quad \Bigg| \quad \begin{array}{ll} \min & b^T y \\ \text{s.t.} & \begin{array}{r} A^T y?c \\ y?0 \end{array} \end{array} \quad (D)$$

2.5.1 Cheat Sheet

Here or

Primal (max)		Dual (min)	
Constraint	$\leq$	≥ 0	Variable
	$\geq$	≤ 0	
	$=$	free	
Variable	$\geq$	≥ 0	Constraint
	$\leq$	≤ 0	
	free	$=$	

Remark:
This is not symmetric... The way you can remember it is by thinking natural variables in real life, like you cannot have negative number of cars and so on...

Q What if you start with a minimization LP as primal?

Example:

$$\begin{array}{ll} \min & x_1 - x_2 \\ \text{s.t.} & \begin{array}{l} 2x_1 + 3x_2 \leq 5 \\ x_1 - x_2 \geq 3 \\ x_1 + 5x_2 = 7 \\ x_1 \geq 0, x_2 \leq 0 \end{array} \end{array} \qquad \text{(P)}$$

Rewrite as:

$$-1 \times \begin{pmatrix} \max & -x_1 + x_2 \\ \downarrow & \\ \text{s.t.} & \dots \end{pmatrix}$$

Will lead to finding dual:

$$\begin{array}{ll} \max & 5y_1 + 3y_2 + 7y_3 \\ \downarrow & \\ & 2y_1 + y_2 \leq 1 \\ \text{s.t.} & 3y_1 - y_2 + 5y_3 \geq -1 \\ & y_1 \leq 0, y_2 \geq 0, y_3 \text{ free} \end{array}$$

Also

- Weak duality holds.
  If $\overline{x}$ feasible for (P), $\overline{y}$ feasible for (D), then $c^T\overline{x} \geq b^T\overline{y}$.
- Strong duality holds

Note
The dual of the dual of (P) is (P).

Example:
Given a simple undirected graph $G = (V, E)$. $M \subseteq E$ is a *matching* if every vertex $v \in V$ is incident to ≤ 1 edge in M.

See examples of matching in CO 342 or MATH 249.

Max cardinality matching

Find matching M with largest $|M|$.

Define $x_e = \begin{cases} 1, & \text{if } e \in M \\ 0, & \text{otherwise} \end{cases}$.

$$\begin{array}{ll} \max & \sum_{e \in E} x_e \\ \downarrow & \\ \text{s.t.} & \begin{array}{ll} \sum_{e \in \delta(v)} x_e \leq 1, & \forall v \in V \\ 0 \leq x_e, & \forall e \in E \end{array} \end{array}$$

where $\delta(v)$ = set of edges in E incident to v.

$$\begin{array}{lll} \min & \sum_{v \in V} y_v & \\ \downarrow & & \\ \text{s.t.} & y_u + y_v \geq 1, & \forall e = uv \in E \\ & y \geq 0 & \end{array}$$

2.6 Other interpretations of dual

Example:

			Resources	
		Per unit Profit	Per unit consumption A	B
Product	1	5	2	3
	2	3	4	1
Available Resources			15	10

$$\begin{array}{ll} \max & 5x_1 + 3x_2 \\ \downarrow & \\ & 2x_1 + 4x_2 \leq 15 \\ \text{s.t.} & 3x_1 + x_2 \leq 10 \\ & x \geq 0 \end{array}$$

Suppose somebody wants to buy A, B from me. What is the lowest price I should ask?

Let y_A, y_B be prices:

$$\begin{array}{ll} \min & 15y_A + 10y_B \\ \downarrow & \\ & 2y_A + 3y_B \geq 5 \\ \text{s.t.} & 4y_A + y_B \geq 3 \\ & y \geq 0 \end{array}$$

Example: Zero-Sum

Alice, Bob play game. A: m choices. B: n choices. Alice play i, Bob plays j, Bob pays Alice M_{ij} dollars.

		Alice R	P	S
Bob	R	0	1	-1
	P	-1	0	1
	S	1	-1	0

Zero-sum: Amount won by Alice - Amount won by Bob = 0

Let $y \in \mathbb{R}^m_+$, Alice's probability distribution.
Let $x \in \mathbb{R}^n_+$, Bob's probability distribution.

Expected Amount Bob pays Alice:

$$\sum_{i=1}^{m} \sum_{j=1}^{n} y_i M_{ij} x_j = y^T M_x$$

$$P = \left\{ x \in \mathbb{R}^n : \begin{array}{l} \sum x_j = 1 \\ x \geq 0 \end{array} \right\}$$

$$Q = \left\{ y \in \mathbb{R}^m : \begin{array}{l} \sum y_i = 1 \\ y \geq 0 \end{array} \right\}$$

Alice wants $\max\limits_{y \in Q} \left\{ \min\limits_{x \in P} \ y^T M_x \right\}$. Bob wants $\min\limits_{x \in P} \left\{ \max\limits_{y \in Q} \ y^T M_x \right\}$.

Suppose $\overline{y} \in Q$ is fixed. Bob's problem is

$$\min_{x \in P} \ \overline{y}^T M_x \ = \ \begin{array}{ll} \min & \displaystyle\sum_{j=1}^n \left(\sum_{i=1}^m M_{ij} \overline{y}_i \right) x_j \\ \downarrow & \\ \text{s.t.} & \begin{array}{l} \displaystyle\sum_{j=1}^n x_j = 1 \\ x \geq 0 \end{array} \end{array}$$

This is equivalent to picking smallest number in

$$\left\{ \sum_{i=1}^m M_{ij} \overline{y}_i \right\}_{j=1}^n$$

$$\implies \max_{y \in Q} \min_{x \in P} y^T M_x = \max_{y \in Q} \left\{ \begin{array}{ll} \max & u \\ \downarrow & \\ \text{s.t.} & \quad u \leq y^T M e_j, \quad \forall j = 1, \ldots, n \end{array} \right\}$$

$$= \begin{array}{ll} \max & u \\ \downarrow & \\ \text{s.t.} & \begin{array}{l} u \leq y^T M e_j, \quad \forall j = 1, \ldots, n \\ y^T = 1 \\ y \geq 0 \end{array} \end{array}$$

Similarly Bob's problem:

$$\begin{array}{ll} \min & v \\ \downarrow & \\ \text{s.t.} & \begin{array}{l} v \geq e_i^T M_x, \quad \forall i = 1, \ldots, m \\ x^T = 1 \\ x \geq 0 \end{array} \end{array}$$

There are x^*, y^* for which strategy values match $\to$ Nash's Equilibrium.

Now get back to Farkas' Lemma Theorem 2.2. [1]

Proof:

$$\begin{array}{ll} \max & 0^T x \\ \downarrow & \\ \text{s.t.} & \quad Ax \leq b \end{array} \tag{P}$$

[1] Rephrase it a little bit: Exactly one of the two has a solution (i) $Ax \leq b$ (ii) $u^T \ldots$.

$$\begin{array}{ll} \min & b^T u \\ \downarrow & \\ \text{s.t.} & \begin{array}{l} u^T A = 0 \\ u \geq 0 \end{array} \end{array} \tag{D}$$

(D) is always feasible ($u = 0$).

If $\exists \overline{x} : A\overline{x} \leq b$, $\overline{x}$ optimal for (P) $\implies$ optimal for (D) has value 0.
$\implies \nexists u$ satisfying (i).

And the converse is also true. □

2.7 Complementary Slackness (C.S.)

Let x^*, y^* be feasible for primal and dual respectively.

C.S.

Complementary Slackness:

i) Either $x_j^* = 0$ or corresponding dual constraint is tight at y^*, $\forall j = 1, \ldots, n$.

ii) Either $y_i^* = 0$ or corresponding primal constraint is tight at x^*, $\forall i = 1, \ldots, m$.

Example:

$$\begin{array}{ll} \min & x_1 - x_2 \\ \downarrow & \\ \text{s.t.} & \begin{array}{l} 2x_1 + 3x_2 \leq 5 \\ x_1 - x_2 \geq 3 \\ x_1 + 5x_2 = 7 \\ x_1 \geq 0, x_2 \leq 0 \end{array} \end{array} \tag{P}$$

$$\begin{array}{ll} \max & 5y_1 + 3y_2 + 7y_3 \\ \downarrow & \\ \text{s.t.} & \begin{array}{l} 2y_1 + y_2 + y_3 \leq 1 \\ 3y_1 - y_2 + 5y_3 \geq -1 \\ y_1 \leq 0, y_2 \geq 0 \end{array} \end{array} \tag{D}$$

i) $x_1^* = 0$ OR $2y_1^* + y_2^* + y_3^* = 1$
$x_2^* = 0$ OR $3y_1^* - y_2^* + 5y_3^* = -1$

ii) $y_1^* = 0$ OR $2x_1^* + 3x_2^* = 5$
$y_2^* = 0$ OR $x_1^* - x_2^* = 3$
$y_3^* = 0$ OR $x_1^* + 5x_2^* = 7$

Theorem 2.7

Let x^*, y^* be feasible for primal/dual respectively. TFAE[a]

a) x^* opt for primal AND y^* opt. for dual

b) Obj. value of x^* = Obj. value of y^*

c) x^*, y^* satisfy C.S.

[a] the following are equivalent

Proof:

a) $\iff$ b) done.

b) $\iff$ c) Proof for

$$\begin{array}{lll} \max & c^T x & \\ \downarrow & & \\ \text{s.t.} & \begin{array}{l} Ax \le b \\ x \ge 0 \end{array} \end{array} \qquad \begin{array}{lll} \min & b^T y & \\ \downarrow & & \\ \text{s.t.} & \begin{array}{l} A^T y \ge c \\ y \ge 0 \end{array} \end{array}$$

Note

$$A^T y \ge c \iff \sum_{i=1}^m a_{ij} y_i \ge c_j, \quad \forall j = 1, \dots, n$$

$$\begin{aligned} c^T x^* &= \sum_{j=1}^n c_j x^* \\ &\le \sum_{j=1}^n \left(\sum_{i=1}^m a_{ij} y_i^* \right) x_j^* \\ &= \sum_{i=1}^m \left(\sum_{j=1}^n a_{ij} x_i^* \right) y_i^* \\ &\le \sum_{i=1}^m b_i y_i^* = b^T y^* \end{aligned}$$

where first and second inequalities come from $x \ge 0, y \ge 0$ respectively.

(b) $c^T x^* = b^T y^* \iff$ C.S. holds. (Just play with some strict inequality conditions)

□

Example:

$$\begin{array}{lll} \max & x_1 + x_2 \\ \downarrow & \\ \text{s.t.} & x_1 + x_2 \le 1 \end{array} \qquad \begin{array}{ll} \min & y \\ \downarrow & \\ & y = 1 \\ \text{s.t.} & y = 1 \\ & y \ge 0 \end{array}$$

Consider a pair $x^* = (0,0), y^* = 1$ which violates CS.

2.7.1 Geometric Interpretation of C.S.

$$\begin{array}{lll} \max & c^T x \\ \downarrow \\ \text{s.t.} & Ax \le b \end{array} \qquad \begin{array}{lll} \min & c^T y \\ \downarrow \\ \text{s.t.} & A^T y = c \\ & y \ge 0 \end{array}$$

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

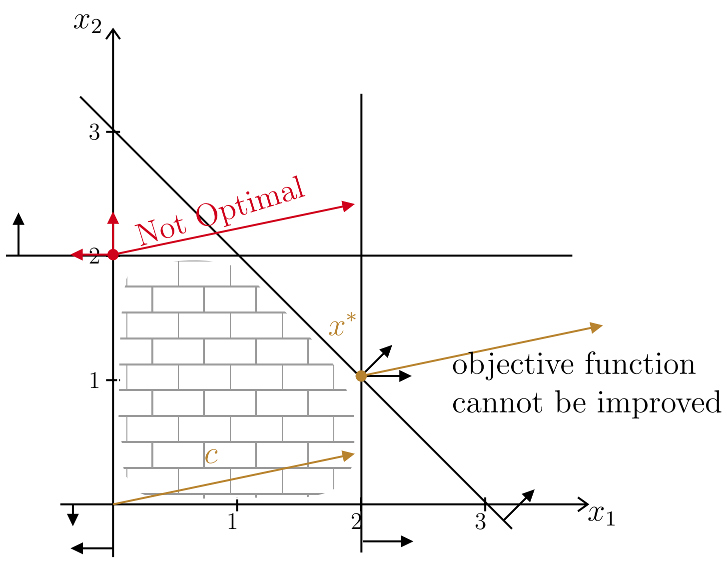
C.S says $a_i^T x^* = b_i$ or $y_i^* = 0$.

$$A^T y = c \implies \begin{pmatrix} | & | & & | \\ a_1 & a_2 & \cdots & a_m \\ | & | & & | \end{pmatrix} y = c \implies \sum_{i=1}^m a_i y_i = c$$

C.S. says c is a nonnegative combination of tight constraint at x^*.

Example:

$$\begin{array}{lll} \max & 2x_1 + 0.5x_2 \\ \downarrow \\ \text{s.t.} & x_1 \le 2 \\ & x_2 \le 2 \\ & x_1 + x_2 \le 3 \\ & x_1, x_2 \ge 0 \end{array}$$



Theorem 2.8

$$\begin{array}{ll} \max & c^T x \\ \downarrow & \\ \text{s.t.} & Ax \leq b \end{array} \qquad \text{(P)}$$

is unbounded iff (P) is feasible and $\exists d \in \mathbb{R}^n : \begin{array}{l} c^T d > 0 \\ Ad \leq 0 \end{array}$.

Proof:

$\Longrightarrow$) Let $\overline{x}$ feasible for (P), $\overline{x} + \lambda d$ is also feasible for (P) $\forall \lambda \geq 0$.

$c^T(\overline{x} + \lambda d)$ can be made arbitrary large.

$\Longleftarrow$) Hard exercise but doable.

$\square$

2.8 Geometry of Polyhedra

line segment

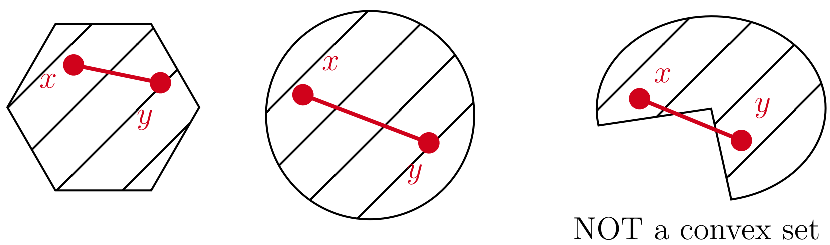
$\overline{x}, \overline{y} \in \mathbb{R}^n$ the line segment between $\overline{x}, \overline{y}$ is

$$\left\{ x \in \mathbb{R}^n : \begin{array}{l} x = \lambda \overline{x} + (1-\lambda)\overline{y} \\ \text{for some } \lambda \in [0,1] \end{array} \right\}$$

convex set

S is a convex set if $\forall x, y \in S$, line segment between x, y is contained in S.

Example:



NOT a convex set

Polyhedra are convex sets. $P = \{x : Ax \leq b\}$. $\overline{x}, \overline{y} \in P$ then

$$A(\underbrace{\lambda}_{\geq 0} \overline{x} + \underbrace{(1-\lambda)}_{\geq 0} \overline{y}) \leq \lambda b + (1-\lambda) b = b$$

convex combination

Given $x^1, \dots, x^k \in \mathbb{R}^n$. We say $\overline{x}$ is a convex combination of $x^1, \dots, x^k$ if $\exists \lambda$:

$$\overline{x} = \sum_{i=1}^{k} \lambda_i x^i$$
$$\sum_{i=1}^{k} \lambda_i = 1$$
$$\lambda \geq 0$$

Optimal solution seems to be happen at "corners".

Let P be a polyhedron $P = \{x \in \mathbb{R}^n : Ax \leq b\}$.

vertex

$\overline{x}$ is a vertex of P if $\exists c$: $\overline{x}$ is unique optimal solution to

$$\begin{array}{ll} \max & c^T x \\ \text{s.t.} & Ax \leq b \end{array}$$

extreme point

$\overline{x}$ is an extreme point of P if $\nexists u, v \in P \setminus \{\overline{x}\}$ such that $\overline{x}$ is in lien segment between u, v.

basic feasible solution

$\overline{x} \in P$ os a basic feasible solution of P if there are n linearly independent tight constraints at $\overline{x}$.

Note
Constraints

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

are linearly independent if $\{a_i\}_{i=1}^m$ are linearly independent.

Theorem 2.9

Let $\overline{x} \in P$. TFAE:

a) $\overline{x}$ is a vertex of P.

b) $\overline{x}$ is a basic feasible solution of P.

c) $\overline{x}$ is a extreme point of P.

Proof:

a) $\implies$ c) Suppose $\exists u, v \in P \setminus \{\overline{x}\}$ such that

$$\overline{x} = \lambda u + (1-\lambda)v$$

for some $\lambda \in (0,1)$. Consider c for which $\overline{x}$ is an optimal solution to

$$\begin{array}{ll} \max & c^T x \\ \text{s.t.} & x \in P \end{array}$$

$$\implies \begin{array}{l} c^T\overline{x} \geq c^T u \\ c^T\overline{x} \geq c^T v \end{array}$$

and

$$c^T\overline{x} = \underbrace{\lambda}_{\geq 0} c^T u + \underbrace{(1-\lambda)}_{\geq 0} c^T v \leq \lambda c^T\overline{x} + (1-\lambda)c^T\overline{x} = c^T\overline{x}$$

$$\implies c^T u = c^T v = c^T\overline{x}$$

$\implies$ $\overline{x}$ NOT a vertex.

c) $\implies$ b) Suppose $\overline{x}$ is not a BFS. Let $i \subseteq \{1, \ldots, m\}$ be the index set of tight constraint at $\overline{x}$. Consider

$$a_i^T d = 0, \quad \forall i \in I \tag{$*$}$$

But since $\overline{x}$ not BFS, $\exists \overline{d} \neq 0$ satisfying $(*)$.[a]

$$x(\epsilon) = \overline{x} + \epsilon\overline{d}$$

$$a_i^T x(\epsilon) = a_i^T\overline{x} \leq b_i, \quad \forall i \in I$$
$$a_i^T x(\epsilon) = \underbrace{a_i^T\overline{x}}_{<b_i} + \epsilon a_i^T d \leq b_i, \quad \forall i \notin I$$

which is satisfied if $|\epsilon|$ is small enough.

$x(\epsilon) \in P$ if $|\epsilon|$ is small enough.

But then

$$\overline{x} = \frac{1}{2}x(\epsilon) + \frac{1}{2}x(-\epsilon)$$

b) $\implies$ a) Let $I \subseteq \{1, \ldots, m\}$ index set of tight constraint at $\overline{x}$.

Define

$$c := \sum_{i \in I} a_i$$

Then $\forall x \in P$

$$c^T x = \sum_{i \in I} a_i^T x \leq \sum_{i \in I} b_i$$

And

$$c^T\overline{x} = \sum_{i \in I} a_i^T\overline{x} = \sum_{i \in I} b_i$$

$\implies \overline{x}$ is optimal solution to

$$\begin{array}{ll} \max & c^T x \\ \text{s.t.} & x \in P \end{array} \tag{**}$$

If $x' \in P$ is optimal solution to $(**)$, then

$$a_i^T x' = b_i, \qquad \forall i \in I \tag{***}$$

But since there are n linear independent constraints in I, $\overline{x}$ is unique solution to $(***)$. $\implies x' = \overline{x}$.

□

[a] by Rank-Nullity Theorem.

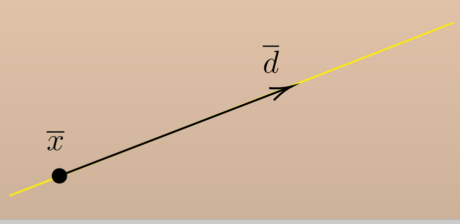
Q When does P have extreme points?

line

Let $\overline{x}, \overline{d} \in \mathbb{R}^n$, $\overline{d} \neq 0$. The set

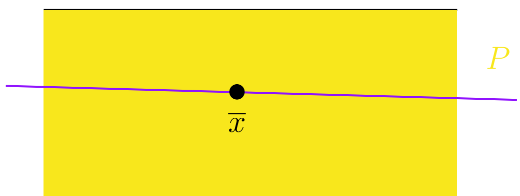
$$\{x \in \mathbb{R}^n : x = \overline{x} + \lambda d \text{ for some } \lambda \in \mathbb{R}\}$$

is called a line.



We say a polyhedron P has a line if $\exists \overline{x}, \overline{d}$ has a line if $\exists \overline{x}, \overline{d}$ s.t. $\overline{x} \in P, \overline{d} \neq 0$ and

$$\{x \in \mathbb{R} : x = \overline{x} + \lambda \overline{d} \text{ for some } \lambda \in \mathbb{R}\} \subseteq P$$



Proposition 2.10

$P = \{x \in \mathbb{R}^n : Ax \leq b\}$ has a line iff $P \neq \varnothing$ and $\exists \overline{d} \neq 0$ such that $A\overline{d} = 0$

$\iff P \neq \varnothing$ and $\operatorname{rank}(A) < n$

Proof:
Exercise. □

Theorem 2.11

$P = \{x \in \mathbb{R}^n : Ax \leq b\}$ has an extreme point

$\iff$ $P \neq \varnothing$ and P has no lines.

Proof:
Exercise. □

pointed polyhedron

A non-empty polyhedron is called pointed if it has no lines.

Note
not pointed does not imply bounded. For example, in $\mathbb{R}^2$, $x \geq 0$ and $y \geq 0$.

Theorem 2.12

Let $P \neq \varnothing$ pointed polyhedron. If $\begin{array}{ll} \max & c^T x \\ \text{s.t.} & x \in P \end{array}$ (LP) has an optimal solution, it has an optimal solution that is an extreme point.

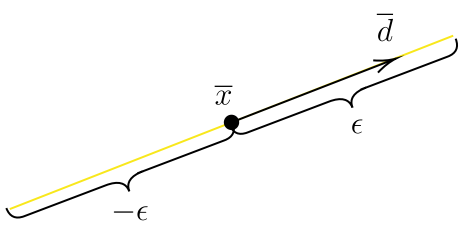
Proof:
Let $\overline{x}$ be an optimal solution to (LP) with largest number of linear independent tight constraints.

Suppose there are $\leq n-1$ linear independent tight constraints at $\overline{x}$.

Pick $\overline{d} \neq 0$ such that $a_i^T \overline{d} = 0, \forall i \in I$, where I is the index set of tight constraints. By the exact same argument as before, $\overline{x} \pm \epsilon\overline{d} \in P$ for ϵ small enough. But

$$c^T(\overline{x} \pm \epsilon\overline{d}) = c^T\overline{x} \pm \epsilon c^T\overline{d}$$

$\implies c^T\overline{d} = 0$
$\implies c^T d(\overline{x} \pm \epsilon d) = c^T\overline{x}$



Since P is pointed, $\exists\overline{\epsilon}$ for which

$$\overline{x} \pm \overline{\epsilon} \in P$$

and one of them not in P if $|\epsilon| > \overline{\epsilon}$. That can only happen if

$$a_k^T(\overline{x} + \overline{\epsilon}\overline{d}) = b_k \qquad \text{or} \qquad a_k^T(\overline{x} - \overline{\epsilon}\overline{d}) = b_k$$

for some $k \notin I$.

$\implies a_k^T \overline{d} \neq 0, \implies a_k$ is linear independent from $\{a_i\}_{i \in I}$ since non-zero cannot be linear combination of zeros. Contradiction to choice of $\overline{x}$. $\square$

2.9 Simplex Algorithm

Standard Equality Form

A linear program is in Standard Equality Form (SEF) if it is of the form

$$\begin{array}{ll} \max & c^T x \\ \downarrow & \\ \text{s.t.} & \begin{array}{l} Ax = b \\ x \geq 0 \end{array} \end{array}$$

Proposition 2.13

Given any linear program, there exists an equivalent LP in SEF.

Index