# *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% ($\approx 5$)
- 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:

$$\begin{array}{ll} \max & f(x) \\ \underbrace{\text{s.t.}}_{\text{subject to}} & x \in S \end{array} \tag{OPT}$$

- $S$ **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

$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & x \in S \end{array}$$

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

**Problem** $x^*$ may not exist

a) Problem is unbounded:

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

b) $S = \varnothing$, 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

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

# 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 $f(x) = c^T x, c \in \mathbb{R}^n$.

$$\downarrow$$

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

**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**

$u \not\leq v$ is not the same as $u > v$

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \not\leq \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Example:

$$\begin{array}{lccr} \max & 2x_1 + & 0.5x_2 & \\ \text{s.t.} & x_1 & & \leq 2 \\ & x_1 + & x_2 & \leq 2 \\ & x & & \geq 0 \end{array}$$

- Strict ineq. not allowed

**halfspace, hyperplane, polyhedron**

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

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

$\{x \in \mathbb{R}^n : h^T x = h_0\}$ 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 $\forall i \in \{1, \ldots, m\}$.

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

which is an LP.

## 2.1 Determining Feasibility

Given a polyhedron

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

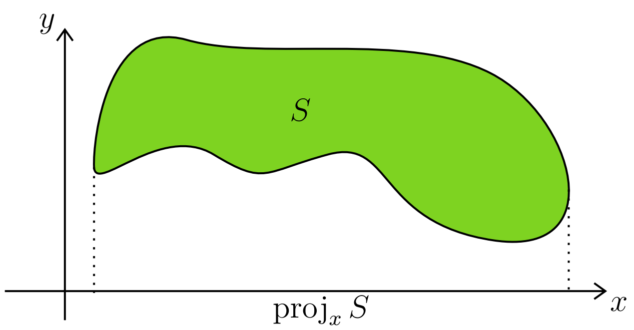
either find $\overline{x} \in P$ or show $P = \varnothing$.

**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 $i \in M^+$:

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

For $i \in M^-$

$$a_i^T x \le b_i \iff \sum_{j=1}^{n-1} \frac{a_{ij}}{a_{in}} x_j - x_n \le \frac{b_i}{-a_{in}}, \qquad \forall i \in M^- \tag{2}$$

For $i \in M^0$

$$a_i^T x \le b_i \iff \sum_{j=1}^{n-1} a_{ij} x_j \le b_i, \qquad \forall i \in M^0 \tag{3}$$

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

Define

$$\sum_{j=1}^{n-1} \left( \frac{a_{ij}}{a_{in}} - \frac{a_{kj}}{a_{kn}} \right) x_j \le \frac{b_i}{a_{in}} - \frac{b_i}{a_{kn}}, \qquad \forall i \in M^+, \forall k \in M^- \tag{4}$$

**Theorem 2.1**

$(\overline{x}_1, \dots, \overline{x}_{n-1})$ 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)

$\Longrightarrow$ 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

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

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

and

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

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

□

**Note**

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

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

**Algorithm 1:** Fourier-Motzkin

1. $A^n = A, b^n = b$
2. 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

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

then

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

3. Keep applying projection until $i = 1$.

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

Let

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

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

Notice that

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

Example:

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

draw the graph, clearly empty

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

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

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

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

$M^+$: $x_1 \leq -3$

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

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

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

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

which is impossible.

( $\implies$ ) 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 $u^TA = 0, u^Tb = b^0_j, u \geq 0$.

□

**Farkas' Lemma (alternate statement)**

Exactly one of the following has a solution:

a) $Ax \leq b$

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

**Farkas' Lemma (Different Form)**

Exactly one of the following has a solution:

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

b) $\begin{aligned} u^TA &\geq 0 \\ u^Tb &< 0 \end{aligned}$

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 $P = \varnothing \iff \exists u_1 \in \mathbb{R}^m, u_2 \in \mathbb{R}^m, v \in \mathbb{R}^n$:

$$\begin{aligned} u_1^TA - u_2^TA - v &= 0 \\ u_1^Tb - u_2^Tb &< 0 \\ u_1, u_2, v &\geq 0 \end{aligned}$$

Let $u = (u_2 - u_2)$

$$u^TA - v = 0 \implies u^TA \geq 0, \quad u^Tb < 0$$

□

Consider a linear programming (LP):

$$\begin{aligned} \max \quad & c^Tx \\ \text{s.t.} \quad & Ax \leq b \end{aligned} \tag{LP}$$

**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

$$\begin{array}{ll} \max & z \\ \text{s.t.} & \begin{array}{l} z - c^T x \leq 0 \\ Ax \leq b \end{array} \end{array} \tag{LP'}$$

(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

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

Until we are left with a polyhedron

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

Now $\begin{array}{ll} \max & z \\ \text{s.t.} & A'z \leq b' \end{array}$ is not cases a) or b). (Why?)

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

## 2.3 Certifying Optimality

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

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

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

Example:

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

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

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

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

In general:

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

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

As long as $y_1, y_2, y_3 \geq 0$ 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:

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

This is called the dual LP.

In general:

$$\begin{aligned} \max\quad & c^T x \\ \text{s.t.}\quad & Ax \leq b \end{aligned} \tag{P}$$

Dual of (P)

$$\begin{aligned} \min\quad & b^T y \\ \text{s.t.}\quad & y^T A = c^T \\ & y \geq 0 \end{aligned} \tag{D}$$

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 $\overline{y} \geq 0$. □

**Corollary 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?

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

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

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 $\theta := \left\{ x \in \mathbb{R}^n : \begin{array}{r} Ax \leq b \\ -c^T x \leq -\gamma \end{array} \right\}$.

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

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

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

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

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

Contradicts optimality of $y^*$.

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

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

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

and

$$y' = y^* + \overline{y} \geq 0$$
$$b^Ty' = b^Ty^* + b^T\overline{y} < b^Ty^*$$

Contradicts optimality of $y^*$.

Thus $\theta \neq \varnothing$.

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

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

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

$\square$

## 2.4 Possible Outcomes

See here.

## 2.5 Duals of generic LPs

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

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

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**

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

### 2.5.1 Cheat Sheet

Here or

<table>
<tr><th colspan="2">Primal (max)</th><th colspan="2">Dual (min)</th></tr>
<tr><td>Constraint</td><td>$\leq$<br>$\geq$<br>$=$</td><td>$\geq 0$<br>$\leq 0$<br>free</td><td>Variable</td></tr>
<tr><td>Variable</td><td>$\geq$<br>$\leq$<br>free</td><td>$\geq 0$<br>$\leq 0$<br>$=$</td><td>Constraint</td></tr>
</table>

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:

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

Rewrite as:

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

Will lead to finding dual:

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

**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 $\leq 1$ edge in $M$.

See examples of matching in CO 342 or MATH 249.

### Max cardinality matching

Find matching $M$ with largest $|M|$.

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

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

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

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

## 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 |
| | R | 0 | 1 | -1 |
| Bob | 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 $\max_{y \in Q} \left\{ \min_{x \in P} \; y^T M_x \right\}$. Bob wants $\min_{x \in P} \left\{ \max_{y \in Q} \; y^T M_x \right\}$.

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

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

This is equivalent to picking smallest number in

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

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

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

Similarly Bob's problem:

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

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

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

Proof:

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

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

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

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

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 (ii).

And the converse is also true. □

## 2.7 Complementary Slackness (C.S.)

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

**Complementary Slackness**

Abbreviated as C.S.

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:

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

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

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

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

**Note**

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

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

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

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

□

Example:

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

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

### 2.7.1 Geometric Interpretation of C.S.

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

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

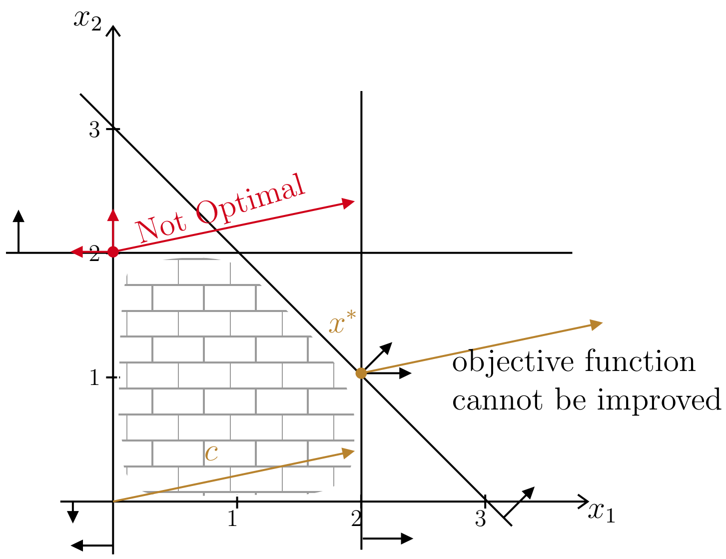
C.S says $a_i^T x^* = b_i$ or $y_i^* = 0$.

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

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

Example:

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



**Theorem 2.8**

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

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

Proof:

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

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

$\Longleftarrow$ ) Hard exercise but doable.

□

## 2.8 Geometry of Polyhedra

**line segment**

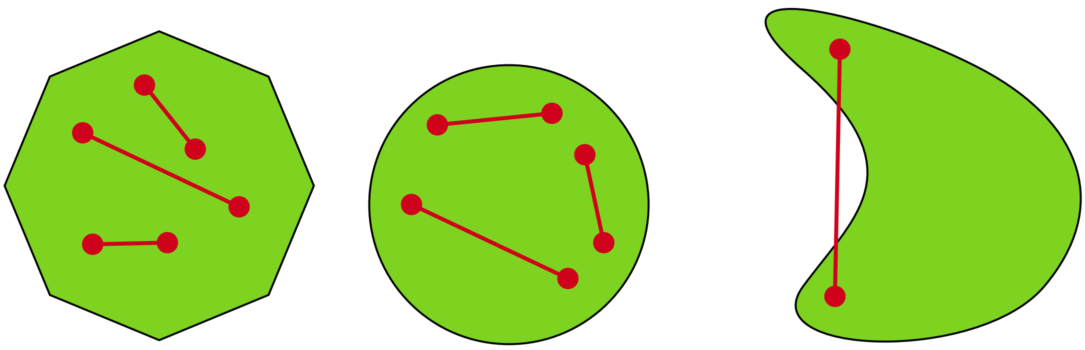
$\overline{x}, \overline{y} \in \mathbb{R}^n$ the line segment between $\overline{x}, \overline{y}$ is

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

**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. $P = \{x : Ax \le b\}$. $\overline{x}, \overline{y} \in P$ then

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

**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 $\exists \lambda$:

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

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 line segment between $u, v$.

**basic feasible solution**

$\overline{x} \in P$ is 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.

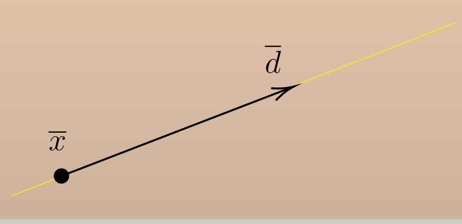
**Q** When does $P$ have extreme points?

**line**

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

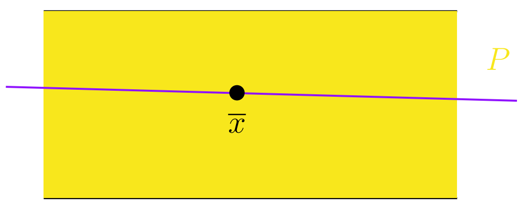
$$\{x \in \mathbb{R}^n : x = \overline{x} + \lambda d \text{ for some } \lambda \in \mathbb{R}\}$$

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.

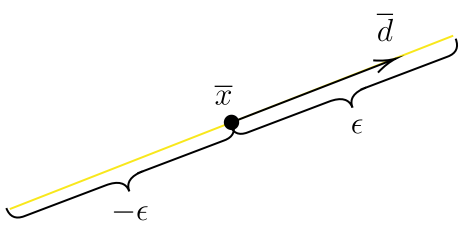
Proof:
Let $\overline{x}$ be an optimal solution to (LP) with largest number of linear independent tight constraints.

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 $\epsilon$ 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

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

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.

Example:

$$\begin{array}{ll} \max & x_1 + 2x_2 + x_3 \\ \downarrow & \\ \text{s.t.} & \begin{array}{l} 3x_1 + x_2 \leq 5 \\ -x_1 + x_3 \geq 6 \\ x_1 \leq 0, x_3 \geq 0 \end{array} \end{array} \tag{P1}$$

$x_1' = -x_1 \geq 0$ and
$x_2 = x_2^+ - x_2^-$ where $x_2^+ \geq 0, x_2^- \geq 0$

We introduce

$$s_1 = 5 - 3x_1 - x_2 \geq 0, \qquad s_2 = -x_1 + x_3 - 6 \geq 0$$

Then

$$\begin{array}{ll} \max & -x_1' + 2x_2^+ - 2x_2^- + x_3 \\ \downarrow & \\ \text{s.t.} & \begin{array}{l} -3x_1' + 2x_2^+ - x_2^- + s_1 = 5 \\ x_1' + x_3 - s_2 = 6 \\ x_1', x_2^+, x_2^-, x_3, s_1, s_2 \geq 0 \end{array} \end{array} \tag{P2}$$

$x$ feasible for (P1) $\iff$ $(x_1', x_2^+, x_2^-, x_3, s_1, s_2)$ feasible for (P2) and they have same cost.

**Assumption** $A \in \mathbb{R}^{m \times n} \rightarrow \text{rank}(A) = m$. This is WLOG. Since if

$$a_i = \sum_{k \neq i} \lambda_k a_k$$

Either

$$b_i \neq \sum_{k\neq i} \lambda_k b_k$$

in which case (SEF) is infeasible. Or $a_i^T x = b_i$ is redundant. So it can be removed from (SEF).

> **Note**
> $\{x : Ax = b, x \geq 0\}$ is *pointed* polyhedron (if nonempty).

**Structure of BFS** Any feasible solution has $m$ linear independent tight constraints $(n-m)$ extra tight constraint must come from $x_j \geq 0$.

Let $B \subseteq \{1, \dots, n\}$ such that $|B| = m$ and $A_B$[2] is invertible.

$N = \{1, \dots, n\} \setminus B$. $x_N = 0$, i.e. $x_j = 0, \forall j \in N$.

Feasible solutions obtained this way are precisely BFS.

Example:

$$\begin{aligned} &\max && \begin{pmatrix} 3 & 2 & 1 & 4 \end{pmatrix} x \\ &\downarrow && \\ &\text{s.t.} && \begin{pmatrix} 1 & 2 & -1 & 0 \\ 2 & 1 & 0 & 1 \end{pmatrix} x = \begin{pmatrix} 5 \\ 7 \end{pmatrix} \\ & && x \geq 0 \end{aligned}$$

If we pick

$$\begin{aligned} B &= \{1, 2\} & A_B &= \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \\ N &= \{3, 4\} & A_N &= \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \\ C_B &= (3 \quad 2)^T & C_N &= (1 \quad 4)^T \end{aligned}$$

$$x_B = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad x_N = \begin{pmatrix} x_3 \\ x_4 \end{pmatrix}$$

---

$$B = \{1, 3\}, B = \{2, 4\}, A_B = \begin{pmatrix} 1 & -1 \\ 2 & 0 \end{pmatrix}, A_N = \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix}$$

$$C_B = \begin{pmatrix} 3 \\ 1 \end{pmatrix}, C_N = \begin{pmatrix} 2 \\ 4 \end{pmatrix}, x_B = \begin{pmatrix} x_1 \\ x_3 \end{pmatrix}, x_N = \begin{pmatrix} x_2 \\ x_4 \end{pmatrix}$$

If we set $x_N = 0$ (for $B = \{1, 3\}$) we are left with

$$\begin{pmatrix} 1 & -1 \\ 2 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_3 \end{pmatrix} = \begin{pmatrix} 5 \\ 7 \end{pmatrix}$$

This has a unique solution $x_1 = 3.5, x_3 = -1.5$, but not feasible.

[2] $A_B$ is submatrix obtained by picking columns of $A$ indexed by $B$. Such $B$ is called a basis.

If we pick $B = \{1, 2\}$

$$\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 5 \\ 7 \end{pmatrix}$$

$\underbrace{x_3 = x_4}_{x_N} = 0,\ x_1 = 3, x_2 = 1$, which is feasible.

In general,

$$Ax = b \iff A_B x_B + \overbrace{\cancel{A_N x_N}}^{0} = b$$

has unique solution $x_b = A_B^{-1} b$.

For any basis $B$, the corresponding *basic solution* is

$$\begin{pmatrix} x_B \\ x_N \end{pmatrix} = \begin{pmatrix} A_B^{-1} b \\ 0 \end{pmatrix}$$

If $A_B^{-1} b \geq 0$, then it is a *BFS*.

### 2.9.1 Canonical Form

Let $B$ be a feasible basis (i.e. corresponding basis solution is feasible).

$$\begin{aligned} Ax = b &\iff A_B x_B + A_N x_N = b \\ &\iff x_B + A_B^{-1} A_N x_N = A_B^{-1} b \end{aligned}$$

Now let's take a look at objective.

$$\begin{aligned} c^T x &= c_B^T x_B + c_N^T x_N - c_B^T (x_B + A_B^{-1} A_N x_N - A_B^{-1} b) \\ &= (c_N^T - c_B^T A_B^{-1} A_N) x_N + c_B^T A_B^{-1} b \end{aligned}$$

Thus (SEF) is said to be in canonical form for $B$ if it is written as

$$\begin{aligned} \max\ \ & \overbrace{(c_N^T - c_B^T A_B^{-1} A_N)}^{\bar{c}_N^T \to \text{Reduced costs}} x_N + c_B^T A_B^{-1} b \\ \downarrow \\ \text{s.t.}\ \ & x_B + A_B^{-1} A_N x_N = A_B^{-1} b \\ & x_B, x_N \geq 0 \end{aligned}$$

Example:

Back to our previous example...

$B = \{1, 2\}$. Rewriting in canonical form for $B$:

$$A_B^{-1} = \begin{pmatrix} -1/3 & 2/3 \\ 2/3 & -1/3 \end{pmatrix}$$

$$A_B A = \begin{pmatrix} 1 & 0 & 1/3 & -2/3 \\ 0 & 1 & 2/3 & -1/3 \end{pmatrix}$$

$$c_B^T A_B^{-1} A_N = \begin{pmatrix}3 & 2\end{pmatrix} \begin{pmatrix}1/3 & -2/3 \\ 2/3 & -1/3\end{pmatrix} = \begin{pmatrix}7/3 & -8/3\end{pmatrix}$$

$$c_N^T - c_B^T A_B^{-1} A_N = \begin{pmatrix}-4/3 & 4/3\end{pmatrix}$$

Then

$$\begin{array}{ll} \max & \begin{pmatrix}0 & 0 & -4/3 & 4/3\end{pmatrix} x + 11 \\ \downarrow & \\ \text{s.t.} & \begin{pmatrix}1 & 0 & 1/3 & -2/3 \\ 0 & 1 & 2/3 & -1/3\end{pmatrix} x = \begin{pmatrix}3 \\ 1\end{pmatrix} \\ & x \geq 0 \end{array}$$

is in canonical form for $B = \{1, 2\}$.

Example:

$$\begin{array}{ll} \max & \begin{pmatrix}1 & 3 & -2 & 0 & 0\end{pmatrix} x \underbrace{+0}_{\text{obj. value}} \\ \downarrow & \\ \text{s.t.} & \begin{pmatrix}1 & 1 & 1 & 1 & 0 \\ 1 & -1 & 3 & 0 & 1\end{pmatrix} x = \begin{pmatrix}4 \\ 1\end{pmatrix} \\ & x \geq 0 \end{array} \tag{LP}$$

Canonical form for $B = \{4, 5\}$.

Corresponding BFS $\begin{matrix} x_4 = 4 \\ x_5 = 1 \end{matrix}, \quad x_j = 0, \forall j \in N$

$x = \begin{pmatrix}0 & 0 & 0 & 4 & 1\end{pmatrix}^T$

Objective value $= 0$

If increase $x_1$ or $x_2$. Objective function increases.

Let's try to increase $x_1$ from $0 \to \theta$. (Keep $x_2 = x_3 = 0$)

$$\begin{aligned} \theta + x_4 = 4 &\iff x_4 = 4 - \theta \\ \theta + x_5 = 1 &\iff x_5 = 1 - \theta \end{aligned}$$

New objective: $0 + \theta$. However, we have

$$\left.\begin{aligned} x_4 \geq 0 &\implies \theta \leq 4 \\ x_5 \geq 0 &\implies \theta \leq 1 \end{aligned}\right. \implies \text{Increase } x_1 \text{ by } 1$$

$x_5$ will be $0 \to \begin{matrix} x_1 \text{ enters basis} \\ x_5 \text{ leaves basis} \end{matrix}$. Then new basis $B = \{1, 4\}$.

---

Rewriting (LP) in canonical form for $B = \{1, 4\}$.

$$\begin{array}{ll} \max & \begin{pmatrix} 0 & 4 & -5 & 0 & -1 \end{pmatrix} x + \underbrace{1}_{\text{obj. value}} \\ \downarrow & \\ \text{s.t.} & \begin{pmatrix} 1 & -1 & 3 & 0 & 1 \\ 0 & 2 & -2 & 1 & -1 \end{pmatrix} x = \begin{pmatrix} 1 \\ 3 \end{pmatrix} \\ & x \geq 0 \end{array}$$

Corresponding BFS:

$$x = \begin{pmatrix} 1 & 0 & 0 & 3 & 0 \end{pmatrix}^T$$

Obj. value = 1

---

Pick $j \in N$: $\bar{c}_j > 0$ $(j = 2)$

Increase $x_2$ to $\theta$, keep $x_3 = x_5 = 0$

$$\begin{aligned} x_1 - \theta = 1 &\iff x_1 = 1 + \theta \\ x_4 + 2\theta = 3 &\iff x_4 = 3 - 2\theta \end{aligned}$$

and

$$\begin{aligned} x_1 \geq 0 &\implies \theta \geq -1 \\ x_4 \geq 0 &\implies \theta \leq \frac{3}{2} \end{aligned}$$

Set $\theta \leftarrow \frac{3}{2} \rightarrow \begin{array}{l} x_2 \text{ enters basis} \\ x_4 \text{ leaves basis} \end{array}$

New basis $B = \{1, 2\}$.

---

(LP) in canonical form for $B = \{1, 2\}$.

$$\begin{array}{ll} \max & \begin{pmatrix} 0 & 0 & -1 & -2 & 1 \end{pmatrix} x + 7 \\ \downarrow & \\ \text{s.t.} & \begin{pmatrix} 1 & 0 & 2 & 0.5 & 0.5 \\ 0 & 1 & -1 & 0.5 & -0.5 \end{pmatrix} x = \begin{pmatrix} 2.5 \\ 1.5 \end{pmatrix} \\ & x \geq 0 \end{array}$$

Corresponding BFS:

$$x = \begin{pmatrix} 2.5 & 1.5 & 0 & 0 & 0 \end{pmatrix}^T$$

Obj. value = 7

---

Find $j \in N$, $\bar{c}_j > 0$ $(j = 5)$

$$\begin{array}{l} x_1 = 2.5 - 0.5\theta \geq 0 \\ x_2 = 1.5 + 0.5\theta \geq 0 \end{array} \implies \begin{array}{l} \theta \leq 5 \\ \theta \geq -3 \end{array} \rightarrow \begin{array}{l} x_1 \text{ leaves basis} \\ x_5 \text{ enters basis} \end{array}$$

New basis $B = \{2, 5\}$

---

(LP) in canonical form for $B = \{2, 5\}$

$$\begin{array}{ll} \max & \begin{pmatrix} -2 & 0 & -5 & -3 & 0 \end{pmatrix} x + 12 \\ \downarrow & \\ \text{s.t.} & \begin{pmatrix} 1 & 1 & 1 & 1 & 0 \\ 2 & 0 & 4 & 1 & 1 \end{pmatrix} x = \begin{pmatrix} 4 \\ 5 \end{pmatrix} \\ & x \geq 0 \end{array}$$

$$\left.\begin{array}{l} \text{BFS } x = \begin{pmatrix} 0 & 4 & 0 & 0 & 5 \end{pmatrix}^T \\ \\ \text{Obj. value} = 12. \end{array}\right\} \text{Optimal Solution}$$

### 2.9.2 Iteration of simplex

**Algorithm 2:** Iteration of simplex

1. Start with feasible basis $B$
2. Rewrite LP in canonical form for $B$
3. Pick $j \in N : \bar{c}_j > 0$ ($x_j$ enters basis)
4. Let $\bar{b} = A_B^{-1} b$, $\overline{A}_N = A_B^{-1} A_N$
   Find largest $\theta$ so that $\bar{b} - \theta \overline{A}_j \geq 0$.
   Corresponding basic variable that becomes 0 (say $x_k$) leaves basis.
5. $B \leftarrow B \setminus \{k\} \cup \{j\}$. Iterate.

If problem has optimal solution AND $\theta$ is always $> 0$, simplex finishes.

**Note**
If at current BFS we have a basic variable $= 0$, we may have $\theta = 0$. $\rightarrow$ May lead to cycling. (i.e. return to current basis in future iteration)

**Bland's Rule**

If there are multiple choices of entering or leaving variables, always pick lowest index variable.

Using Bland's Rule avoids cycling

**Observations** If $\bar{c}_N \leq 0$, then the (LP) obj. value in canonical form is

$$\underbrace{\bar{c}_N^T}_{\leq 0} \underbrace{x_N}_{\geq 0} + c_B^T A_B^{-1} b \leq c_B^T A_B^{-1} b$$

For any feasible solution $\implies$ Current BFS is optimal

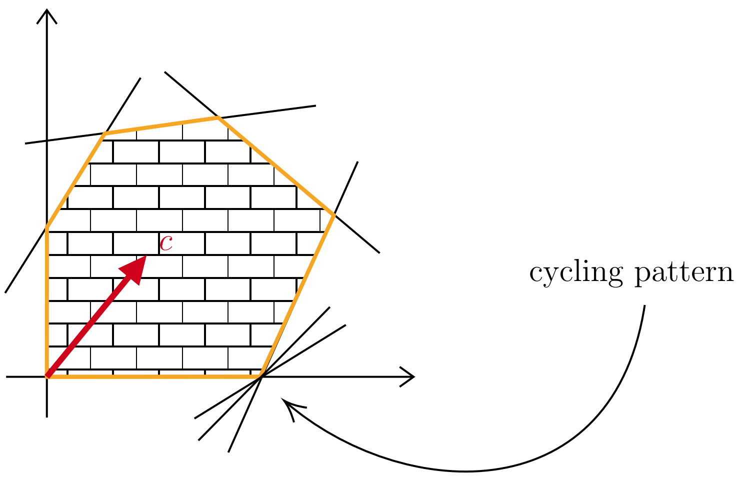



Figure 2.1: Simplex method

Original LP

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

Dual

$$\begin{array}{ll} \min & b^T y \\ \downarrow & \\ \text{s.t.} & A^T y \geq c \end{array} \iff \begin{array}{ll} \min & y^T b \\ \downarrow & \\ \text{s.t.} & y^T A \geq c^T \end{array}$$

$$\iff \begin{array}{ll} \min & y^T b \\ \downarrow & \\ \text{s.t.} & \begin{array}{l} y^T A_B \geq c_B^T \\ y^T A_N \geq c_N^T \end{array} \end{array}$$

If satisfies C.S with BFS corresponding to $B$

$$\begin{array}{rl} & y^T A_B = c_B^T \\ \implies & \quad y^T = c_B^T A_B^{-1} \iff c_B^T A_B^{-1} A_N \geq c_N^T \iff \bar{c}_N \leq 0 \\ & y_T A_N \geq c_N^T \end{array}$$

### 2.9.3 Mechanics of Simplex

Example: 1

$$\begin{array}{ll} \max & \begin{array}{ccccc} \text{enters basis} & & & j & \\ \uparrow & & & \uparrow & \\ \end{array} \\ & \begin{pmatrix} 1 & 3 & -2 & 0 & 0 \end{pmatrix} x \\ \downarrow & \\ \text{s.t.} & \overset{\text{pivot}}{} \begin{pmatrix} 1 & 1 & 1 & 1 & 0 \\ 1 & -1 & 3 & 0 & 1 \end{pmatrix} x = \begin{pmatrix} 4 \\ 1 \end{pmatrix} \quad \text{row } \ell \\ & x \geq 0 \end{array}$$

For $\theta$

$$\theta \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \begin{pmatrix} x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 4 \\ 1 \end{pmatrix}$$

and we have

$$\begin{pmatrix} x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 4-\theta \\ 1-\theta \end{pmatrix} \geq 0 \implies \begin{matrix} \theta \leq 4 \\ \boxed{\theta \leq 1} \end{matrix}$$

We are actually picking $\min \left\{ \frac{4}{1}, \frac{1}{1} \right\}$

Pick, out of all rows $\min \left\{ \frac{\bar{b}_i}{\bar{a}_{ij}} \right\}$ where $j$ is entering variable.

Then now in row $\ell$ (second row here). Make row operations so that pivot element become 1, all others in col $j$ becomes 0.

→ Row 2 $\times 1$

→ Subtract tow 2 from row 1

→ subtract row 2 from objective function (with RHS multiplied by $-1$)

$$\begin{aligned}
\max \quad & \begin{pmatrix} 0 & \overset{\overset{j}{\uparrow}}{4} & -5 & 0 & -1 \end{pmatrix} x + 1 \\
\downarrow \quad & \\
\text{s.t.} \quad & \begin{pmatrix} 0 & \overset{\text{pivot}}{2} & -2 & 1 & -1 \\ 1 & -1 & 3 & 0 & 1 \end{pmatrix} x = \begin{pmatrix} 3 \\ 1 \end{pmatrix} \qquad \text{row } \ell \\
& x \geq 0
\end{aligned}$$

$$\begin{aligned}
2\theta + x_4 = 3 &\iff x_4 = 3 - 2\theta \geq 0 \implies \theta \leq \frac{3}{2} \\
-\theta + x_1 = 1 &\iff x_1 = \theta + 1 \geq 0 \implies \theta \geq -1
\end{aligned}$$

where we are finding $\min\limits_{\bar{a}_{ij}>0} \left\{ \dfrac{\bar{b}_i}{\bar{a}_{ij}} \right\}$. Now follow the similar procedure, we have

$$\begin{aligned}
\max \quad & \begin{pmatrix} 0 & 0 & -1 & -2 & 1 \end{pmatrix} x + 7 \\
\downarrow \quad & \\
\text{s.t.} \quad & \begin{pmatrix} 0 & 1 & -1 & 0.5 & -0.5 \\ 1 & 0 & 2 & 0.5 & 0.5 \end{pmatrix} x = \begin{pmatrix} 1.5 \\ 2.5 \end{pmatrix}
\end{aligned}$$

**In general** Pick $j \in N : \bar{c}_j > 0$.

Let $\ell = \operatorname{argmin}\limits_{\bar{a}_{ij}>0} \left\{ \dfrac{\bar{b}_i}{\bar{a}_{ij}} \right\}$ (Ratio Test)

- Multiply row $\ell$ by $\dfrac{1}{\bar{a}_{\ell j}}$

- Add $-\dfrac{\bar{a}_{ij}}{\bar{a}_{\ell j}}$ times row $\ell$ to row $i \neq \ell$.

- Add $-\dfrac{\overline{c}_j \cdot \overline{a}_{\ell k}}{\overline{a}_{\ell j}}$ to variable coeff in objective. $\forall k \in 1, \dots, n$
- Add $\dfrac{\overline{b}_\ell \cdot \overline{c}_j}{\overline{a}_{ij}}$ to objective value in objective function

Example: 2

$$
\begin{array}{ll}
\max & \begin{array}{l} \;\; j \\ \;\;\uparrow \\ \begin{pmatrix} 2 & 1 & 1 & 0 & 0 \end{pmatrix} x \end{array} \\
\downarrow & \\
\text{s.t.} & \text{pivot} \; \begin{pmatrix} 1 & 2 & -1 & 1 & 0 \\ 2 & -2 & -1 & 0 & 1 \end{pmatrix} x = \begin{pmatrix} 2 \\ 3 \end{pmatrix} \qquad \text{row } \ell \\
& x \geq 0
\end{array}
$$

**Ratio Test** $\min \left\{ \frac{2}{1}, \frac{3}{2} \right\} = 1.5$. $\ell = 2$. ($x_2$ enters, $x_5$ leaves)

$$
\begin{array}{ll}
\max & \begin{array}{l} \qquad\;\; j \\ \qquad\;\;\uparrow \\ \begin{pmatrix} 0 & 3 & 2 & 0 & -1 \end{pmatrix} x + 3 \end{array} \\
\downarrow & \\
\text{s.t.} & \begin{pmatrix} 0 & 3 & -0.5 & 1 & -0.5 \\ 1 & -1 & -0.5 & 0 & 0.5 \end{pmatrix} x = \begin{pmatrix} 0.5 \\ 1.5 \end{pmatrix} \\
& x \geq 0
\end{array}
$$

If we increase $x_3 \to \theta$ and keep $x_2 = x_5 = 0$

$$
\begin{array}{l} -0.5\theta + x_4 = 0.5 \\ -0.5\theta + x_1 = 1.5 \end{array} \implies \begin{array}{l} x_1 = 1.5 + 0.5\theta \\ x_4 = 0.5 + 0.5\theta \end{array} \to \text{Problem is unbounded!}
$$

**In general** Let $B$ be a basis

$$
\begin{array}{ll}
\max & \overline{c}_N^T x_N \\
\downarrow & \\
\text{s.t.} & x_B + \overline{A}_N x_N = \overline{b} \\
& x_B, x_N \geq 0
\end{array}
$$

Found $j : \overline{c}_j > 0$ AND $\overline{A}_j \leq 0$.

Construct $d \in \mathbb{R}^n$ to reflect what we are trying to do when we increase $x_j \to \theta$.

Right now, we are at BFS:

$$
\begin{pmatrix} x_B \\ x_N \end{pmatrix} = \begin{pmatrix} A_B^{-1} b \\ 0 \end{pmatrix}
$$

We want:

$$
\begin{pmatrix} x_B \\ x_N \end{pmatrix} = \begin{pmatrix} A_B^{-1} b \\ 0 \end{pmatrix} + \theta \begin{pmatrix} d_B \\ d_N \end{pmatrix}
$$

where $d_N = \left(\begin{array}{c} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \\ 0 \end{array}\right)\begin{array}{l} \\ \\ \\ \leftarrow j \\ \\ \\ \\ \end{array} = e_j$ and $d_B = -\overline{A}_j = -A_B^{-1}A_j$.

Found $d$: $d \geq 0$, then

$$Ad = A_B d_B + A_N d_N = -A_B A_B^{-1} A_j + A_j = 0$$

and

$$c^T d = c_B^T d_B + c_N^T d_N = -c_B^T A_B^{-1} A_j + c_j = \overline{c}_j > 0$$

i.e.,

$$\left.\begin{aligned} c^T d &> 0 \\ Ad &= 0 \\ d &\geq 0 \end{aligned}\right. \implies \text{Problem is unbounded}$$

*But wait, how to find an initial BFS?*

Given

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

where $b \geq 0$.

Construct auxiliary

$$\begin{array}{ll} \max & -e^T w \\ \downarrow & \\ \text{s.t.} & \begin{array}{l} Ax + Iw = b \\ x, w \geq 0 \end{array} \end{array} \tag{AUX}$$

**Note**

- (AUX) is feasible ($x = 0, w = b$)
- (AUX) is bounded $-e^T w \leq 0$

So (AUX) has an optimal solution.

**Proposition 2.14**

(AUX) has optimal value 0 iff (LP) is feasible.

Proof:

If optimal solution $(x^*, w^*)$ has value 0, then $w^* = 0$ so $Ax^* + I0 = b$

$\implies$ $x^*$ is feasible for (LP)

If $x$ is feasible for (LP) then $(x, 0)$ has value 0 in (AUX).

Moreover, if optimal value of (AUX) is $< 0$, then we can use the dual for a certificate.

$$\begin{array}{ll} \min & y^T b \\ \downarrow & \\ \text{s.t.} & \begin{array}{l} y^T A \geq 0 \\ y \geq -e \end{array} \end{array} \tag{DAUX}$$

$y^*$ optimal ${y^*}^T b < 0$ and ${y^*}^T A \geq 0$

$\implies y^*$ satisfies $\{x : Ax = b,\ x \geq 0\} = \varnothing$ $\square$

### 2.9.4 Two Stage Simplex

#### Phase 1

- write (AUX)
- solve (AUX) with BFS corresponding to $w$
- if opt value $< 0$, get certificate $y^*$ (LP) is infeasible
- opt value 0, BFS $x$ where $w = 0$

#### Phase 2

- simplex with $x$ as initial BFS

Example: 1

$$\begin{array}{ll} \max & \begin{pmatrix} 2 & 1 & 3 \end{pmatrix} x \\ \downarrow & \\ \text{s.t.} & \begin{pmatrix} 2 & 1 & 0 \\ 1 & 1 & 2 \end{pmatrix} x \begin{array}{cc} \leq & -1 \\ \geq & 3 \end{array} \\ & x \geq 0 \end{array}$$

$$\begin{array}{ll} \max & \begin{pmatrix} 2 & 1 & 3 & 0 & 0 \end{pmatrix} x \\ \downarrow & \\ \text{s.t.} & \begin{pmatrix} -2 & -1 & 0 & -1 & 0 \\ 1 & 1 & 2 & 0 & -1 \end{pmatrix} x = \begin{pmatrix} 1 \\ 3 \end{pmatrix} \\ & x \geq 0 \end{array} \tag{SEF}$$

$$\begin{array}{ll} \max & \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & -1 & -1 \end{pmatrix} x \\ \downarrow & \\ \text{s.t.} & \begin{pmatrix} -2 & -1 & 0 & -1 & 0 & 1 & 0 \\ 1 & 1 & 2 & 0 & -1 & 0 & 1 \end{pmatrix} x = \begin{pmatrix} 1 \\ 3 \end{pmatrix} \\ & x \geq 0 \end{array} \tag{AUX}$$

canonical form: $B = \{6, 7\}$

$$\begin{array}{ll} \max & \begin{pmatrix} -1 & 0 & 2 & -1 & -1 & 0 & 0 \end{pmatrix} x - 4 \\ \downarrow & \\ \text{s.t.} & \begin{pmatrix} -2 & -1 & 0 & -1 & 0 & 1 & 0 \\ 1 & 1 & 2 & 0 & -1 & 0 & 1 \end{pmatrix} x = \begin{pmatrix} 1 \\ 3 \end{pmatrix} \\ & x \geq 0 \end{array}$$

add 3 to the basis

$\min \left( \frac{b_i}{a_{i3}} \right) = \frac{3}{2}$

7 leaves the basis.

canonical form for $B = \{3, 6\}$

$$\begin{array}{ll} \max & \begin{pmatrix} -2 & -1 & 0 & -1 & 0 & 0 & -1 \end{pmatrix} x - 1 \\ \downarrow & \\ \text{s.t.} & \begin{pmatrix} -2 & -1 & 0 & -1 & 0 & 1 & 0 \\ 1/2 & 1/2 & 1 & 0 & -1/2 & 0 & 1/2 \end{pmatrix} x = \begin{pmatrix} 1 \\ 3/2 \end{pmatrix} \end{array}$$

$x^* = \begin{pmatrix} 0 & 0 & \frac{3}{2} & 0 & 0 & 1 & 0 \end{pmatrix}$

certificate of infeasibility

$$\begin{aligned} y^T &= c_B^T A_B^{-1} \\ &= \begin{pmatrix} 0 & -1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 2 & 0 \end{pmatrix}^{-1} \\ &= \begin{pmatrix} 0 & -1 \end{pmatrix} \begin{pmatrix} 0 & 1/2 \\ 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} -1 & 0 \end{pmatrix} \end{aligned}$$

Example: 2

$$\begin{array}{ll} \max & \begin{pmatrix} 1 & 0 & 2 \end{pmatrix} x \\ \downarrow & \\ \text{s.t.} & \begin{pmatrix} 2 & 1 & 1 \\ -1 & -1 & -2 \end{pmatrix} x = \begin{pmatrix} 7 \\ -5 \end{pmatrix} \\ & x \geq 0 \end{array}$$

in SEF.

$$\begin{array}{ll} \max & \begin{pmatrix} 1 & 0 & 2 \end{pmatrix} x \\ \downarrow & \\ \text{s.t.} & \begin{pmatrix} 2 & 1 & 1 \\ 1 & 1 & 2 \end{pmatrix} x = \begin{pmatrix} 7 \\ 5 \end{pmatrix} \end{array}$$

$$\begin{array}{ll} \max & \begin{pmatrix} 0 & 0 & 0 & -1 & -1 \end{pmatrix} x \\ \downarrow & \\ \text{s.t.} & \begin{pmatrix} 2 & 1 & 1 & 1 & 0 \\ 1 & 1 & 2 & 0 & 1 \end{pmatrix} x = \begin{pmatrix} 7 \\ 5 \end{pmatrix} \end{array} \tag{AUX}$$

canonical form $B = \{4, 5\}$

$$\begin{array}{ll} \max & \begin{pmatrix} 3 & 2 & 3 & 0 & 0 \end{pmatrix} x - 12 \\ \downarrow & \\ \text{s.t.} & \begin{pmatrix} 2 & 1 & 1 & 1 & 0 \\ 1 & 1 & 2 & 0 & 1 \end{pmatrix} x = \begin{pmatrix} 7 \\ 5 \end{pmatrix} \\ & x \geq 0 \end{array}$$

1 enters basis $x + \theta d \quad d = \begin{pmatrix} 1 & 0 & 0 & -2 & -1 \end{pmatrix}^T$

$\min \left( \frac{b_i}{a_{i1}} \right) = \frac{7}{2}$

4 leaves the basis

$$\begin{array}{ll} \max & \begin{pmatrix} 0 & 1/2 & 3/2 & -3/2 & 0 \end{pmatrix} x - 3/2 \\ \downarrow & \\ \text{s.t.} & \begin{pmatrix} 1 & 1/2 & 1/2 & 1/2 & 0 \\ 0 & 1/2 & 3/2 & -1/2 & 1 \end{pmatrix} x = \begin{pmatrix} 7/2 \\ 3/2 \end{pmatrix} \\ & x \geq 0 \end{array}$$

2 enters the basis

$\min \left( \frac{b_i}{a_{i2}} \right) = \frac{3/2}{1/2}$

5 leaves the basis

$$\begin{array}{ll} \max & \begin{pmatrix} 0 & 0 & 0 & -1 & -1 \end{pmatrix} x + 0 \\ \downarrow & \\ \text{s.t.} & \begin{pmatrix} 1 & 0 & -1 & 1 & -1 \\ 0 & 1 & 3 & -1 & 2 \end{pmatrix} x = \begin{pmatrix} 2 \\ 3 \end{pmatrix} \\ & x \geq 0 \end{array}$$

Thus $x = \begin{pmatrix} 2 & 3 & 0 & 0 & 0 \end{pmatrix}$ is optimal for (AUX)

Forget (AUX). Start Simplex with $x = \begin{pmatrix} 2 & 3 & 0 \end{pmatrix}$ as initial BFS.

Now return to SEF.

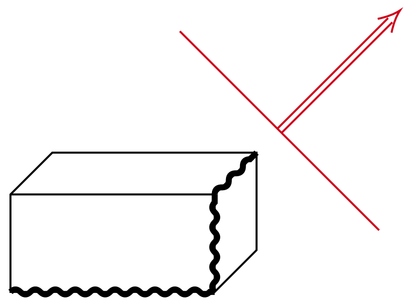
$$\begin{array}{ll} \max & \begin{pmatrix} 1 & 0 & 2 \end{pmatrix} x \\ \downarrow & \\ \text{s.t.} & \begin{pmatrix} 2 & 1 & 1 \\ 1 & 1 & 2 \end{pmatrix} x = \begin{pmatrix} 7 \\ 5 \end{pmatrix} \\ & x \geq 0 \end{array} \tag{SEF}$$

canonical form for $B = \{1, 2\}$

$$\begin{array}{ll} \max & \begin{pmatrix} 0 & 0 & 3 \end{pmatrix} x + 2 \\ \downarrow & \\ \text{s.t.} & \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 3 \end{pmatrix} x = \begin{pmatrix} 2 \\ 3 \end{pmatrix} \end{array}$$

*How long does simplex take?*

At each pivot, we move from an extreme point to another.



Every pivot rule has a bad example.

Sprelman & Teng (2001): bad examples are pathological. Small changes become good examples.

**Polynomial Hirsch Conjecture**

~~Polynomially many vertex for bounded Polyhedral.~~

Let $G$ be the graph of a $d$-polytope with $n$ facets. Then the diameter of $G$ is bounded above by a polynomial of $d$ and $n$.

*or*

The (combinatorial) diameter of a polytope of dimension $d$ with $n$ facets cannot be greater than $n - d$.

Remark:

Here we call combinatorial diameter of a polytope the maximum number of steps needed to go from one vertex to another, where a step consists in traversing an edge.

What this conjecture tells us is that it will take only finitely many edges from initial BFS to optimal one.

There's one counterexample: 43-dimensional polytope with 86 facets and diameter (at least) 44.

## 2.10 Ellipsoid Algorithm

**Feasibility** Given polyhedron $P$, find $\overline{x} \in P$ or show $P = \varnothing$.

Fourier-Motzkin & simplex solve this problem.

**Aside** Given an algorithm an input $I$ to it,

$$\text{size}(I) = \# \text{ of bits needed to represent } I.$$

Example:

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

Assume $c \in \mathbb{Q}^n, A \in \mathbb{Q}^{m \times n}, b \in \mathbb{Q}^n$.

By scaling, we may assume $c \in \mathbb{Z}^n, A \in \mathbb{Z}^{m \times n}, b \in \mathbb{Z}^m$.

Let $\alpha = \max\{\|c\|_\infty, \|A\|_\infty, \|b\|_\infty\}$.

Size of input to LP $\approx (n + n, m + m) \log(\alpha)$

**Efficient Algorithm** # of operations to solve an instance of size $k$ are bounded by a polynomial on $k$.

Thus Simplex & FM NOT Efficient.

**Goal** Derive an efficient alg.

If you have an efficient algorithm to solve feasibility for any polyhedron $P$, can be used to solve LP.

## Option 1

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

Assume I know $L \leq \text{OPT} \leq U$.

**Algorithm 3:** Option 1

**while** *Repeat* **do**
  $V = \dfrac{L+U}{2}$
  $P' = \left\{ x : \begin{array}{l} Ax \leq b \\ c^T x \geq V \end{array} \right\}$
  **if** $P' == \varnothing$ **then**
    $U \leftarrow V$
  **else**
    $L \leftarrow V$
  **end**
**end**

## Option 2

Is the following nonempty?

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

### 2.10.1 Ellipsoid

**Ball** $B(z, R) := \{x \in \mathbb{R}^n : \|x - z\| \leq R\}$

**Unit Ball** $B := B(0, 1)$

Apply an affine map to $B$.

$f(x) = A(x - b)$ where $b \in \mathbb{R}^n, A \in \mathbb{R}^{n \times n}$ invertible

$$f(B) := \{x \in \mathbb{R}^n : \|f(x)\| \leq 1\} = \{x \in \mathbb{R}^n : \|A(x - b)\| \leq 1\}$$

Sets of this form are **Ellipsoid**. Denoted $E(A, b)$.

#### Idea

- Suppose I know $P \subseteq B(0, R)$
- Also, suppose either $P = \varnothing$ OR $\operatorname{Vol} P \geq \epsilon > 0$.

**Algorithm 4:** Ellipsoid Algorithm

$E \leftarrow E(M, z)$, where $P \subseteq E(M, z)$.
**while** $\operatorname{Vol}(E) \geq \epsilon$ **do**
  **if** $z \in P$ **then**
    STOP
  **else**
    • Find $\alpha^T x \leq \alpha_0$ so that $\alpha^T x \leq \alpha_0, \forall x \in P$ and $\alpha^T z > \alpha_0$
    • Find $E(M', z')$ such that $E \cap \{x : \alpha^T x \leq \alpha_0\} \subseteq E(M', z')$ and volume of $E(M', z')$ is much lower than $E$
    • $E \leftarrow E(M', z')$
  **end**
**end**

> **Note**
> At any point $P \subseteq E$.
>
> The reason why we choose ellipsoid instead of ball is that it can actually shrink "thinner" than ball.

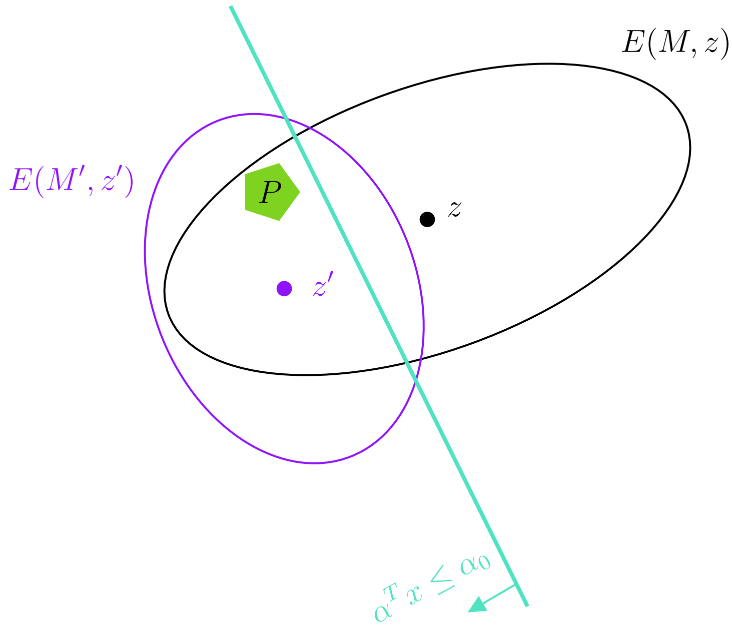



Figure 2.2: Ellipsoid Algorithm

**Lemma 2.15**

There exists $E(M', z')$ that can be computed in polynomial time such that

$$\frac{\mathrm{Vol}(E(M', z'))}{\mathrm{Vol}(E(M, z))} \leq e^{-\frac{1}{2n+2}}$$

## Number of While Loop Iterations

If $B(0, R)$ initial ellipsoid, then $\mathrm{Vol}(B(0, R)) \leq (2R)^n$. After $k(2n+2)$ iterations, $\mathrm{Vol}(E) \leq e^{-k}(2R)^n$.

We want

$$e^{-k}(2R)^n < \epsilon \implies -k + n\ln(2R) < \ln(\epsilon) \implies k \geq \lceil n\ln(2R) - \ln(\epsilon) \rceil$$

Alg stops after $\lceil n\ln(2R) - \ln(\epsilon) \rceil (2n+2)$ iterations.

We only used that

$$z \notin P \iff \begin{array}{l} \exists \alpha^T x \leq \alpha_0 \text{ such that} \\ \alpha^T \overline{x} \leq \alpha_0, \forall \overline{x} \in P \\ \alpha^T z > \alpha_0 \end{array}$$

**Theorem 2.16: Separating Hyperplane**

Let $C$ be a closed, convex set, $z \in \mathbb{R}^n$. Then $z \notin C \iff \exists$ a hyperplane $\alpha^T x \leq \alpha_0$ separating $z$ and $C$.

Is runtime polynomial?

- $\ln(R)$ is polynomial in input size $\to$ NOT a problem
- Finding a separating hyperplane: can be done in polynomial time.

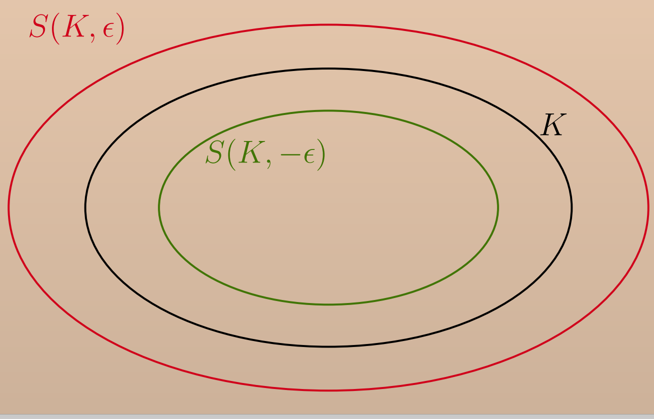
## 2.11 Grötchel-Lovász-Schrijver (GLS)

**$S(K, \pm\epsilon)$**

Let $K \subseteq \mathbb{R}^n$ be closed bounded convex set.

$$S(K, \epsilon) := \{x : \|x - y\| \le \epsilon, \text{ for some } y \in K\}$$

$$S(K, -\epsilon) := \{x : S(x, \epsilon) \subseteq K\}$$




### 2.11.1 3 problems

- Optimization

  Given $K \subseteq \mathbb{R}^n$, $c \in \mathbb{Q}^n$.

  Find $x^* \in K$ such that
  $$c^T x^* \ge c^T x, \forall x \in K$$
  or determine $K = \varnothing$.

- Separation

  Given $K \subseteq \mathbb{R}^n$, $w \in \mathbb{R}^n$.

  Determine if $w \in K$ or find $\alpha$:
  $$\|\alpha\|_\infty = 1 \qquad \alpha^T x < \alpha^T w, \forall x \in K$$

- FEASIBILITY

  Given $K \subseteq \mathbb{R}^n$.

  Find $\overline{x} \in K$ or determine $K = \varnothing$.

Feas $\leq_p$ Opt. (i.e. if we can solve opt efficiently, we can solve feas efficiently)

Weaker version...

- WEAK OPTIMIZATION

  Give $K \subseteq \mathbb{R}^n, c \in \mathbb{Q}^n, \epsilon > 0$

  Find $x^* \in S(K, \epsilon)$ such that

  $$c^T x \leq c^T x^* + \epsilon, \qquad \forall x \in S(K, -\epsilon)$$

  or determine $S(K, -\epsilon) = \varnothing$

- WEAK SEPARATION

  Given $K \subseteq \mathbb{R}^n, w \in \mathbb{R}^n, \epsilon > 0$.

  Determine if $w \in S(K, \epsilon)$ or find $\alpha$:

  $$\|\alpha\|_\infty = 1 \qquad \alpha^T x < \alpha^T w + \epsilon, \forall x \in S(K, -\epsilon)$$

- WEAK FEASIBILITY

  Given $K \subseteq \mathbb{R}^n$.

  Determine $S(K, -\epsilon) = \epsilon$ or find $\overline{x} \in S(K, \epsilon)$

W-Feas $\leq_p$ W-Opt.

Ellipsoid gives us: W-Feas $\leq_p$ W-Sep.

- Grötchel-Lovász-Schrijver (GLS) have shown that

  W-SEP, W-Feas, W-OPT are polynomially equivalent.

In particular, for rational polyhedra[3] (even unbounded) then OPT, FEAS, SEP are polynomially equivalent.

Khachiyan ('80) used ellipsoid to give polytime algorithm for LPs.

### 2.11.2 Consequence of GLS

**Example** TSP: **complete** graph $G = (V, E)$

[3] $\{x \in \mathbb{R}^n : Ax \leq b\}$ where $A \in \mathbb{Q}^{m \times n}, b \in \mathbb{Q}^m$

Edge costs $c_e, \forall e \in E$.

Find a tour visiting every vertex exactly once of min cost.

**IP formulation** $x_e = \begin{cases} 1, & \text{if } e \text{ is in tour} \\ 0, & \text{otherwise} \end{cases}$

$$\begin{array}{ll} \min & \sum_{e \in E} c_e x_e \\ \downarrow & \\ \text{s.t.} & \sum_{e \in \delta(v)} x_e = 2, \quad \forall v \in V \end{array}$$

In general, $\delta(S) = \left\{ uv \in E : \begin{array}{c} u \in S \\ v \notin S \end{array} \right\}$ where $S \subseteq V$.

**Subtour elimination** $\displaystyle\sum_{e \in \delta(S)} x_e \geq 2, \quad \forall \varnothing \subsetneq S \subsetneq V$

$$\begin{array}{lll} \min & \sum_{e \in E} c_e x_e & \\ \downarrow & & \\ & \sum_{e \in \delta(v)} x_e = 2, & \forall v \in V \\ \text{s.t.} & \sum_{e \in \delta(S)} x_e \geq 2, & \forall \varnothing \subsetneq S \subsetneq V \\ & x_e \in \{0,1\}, & \forall e \in E \end{array}$$

**LP-relaxation** Replace $x_e \in \{0,1\}$ by $0 \leq x_e \leq 1, \forall e \in E$.

Can I solve the LP in polynomial time on # vertices/edges?

**Separation/Feasibility** Given $\overline{x}_e, \forall e \in E$. Can I know if $\overline{x}_e$ if feasible for LP in time polynomial in # vertices?

If YES, GLS tells we can also solve OPT.

In polytime (in # vertices) I can check $\begin{cases} \sum_{e \in \delta(v)} \overline{x}_e = 2, & \forall v \in V \\ 0 \leq \overline{x}_e \leq 1, & \forall e \in E \end{cases}$

**Min-Cut problem** Given $G = (V, E), w_e \geq 0$. Find $\displaystyle\sum_{e \in \delta(S)} w_e$

Problem can be solved in polytime in # vertices.

Then we solve mincut with $w_e = \overline{x}_e$. If optimal value is $\geq 2$, then $\overline{x}$ feasible for LP. Otherwise found $S : \displaystyle\sum_{e \in \delta(S)} \overline{x}_e < 2$.

3

# Integer Programming

An integer program is a problem of the form:

$$\begin{array}{ll} \max & c^T x \\ \downarrow & \\ \text{s.t.} & Ax \le b \\ & x_i \in \mathbb{Z}, \forall j \in I \end{array}$$

where $\varnothing \neq I \subseteq \{1, \ldots, n\}$.

If $I = \{1, \ldots, n\}$, it's pure IP. Otherwise, Mixed IP (MIP).

If all variables are constrained to be in $\{0, 1\}$, it's a Binary IP.

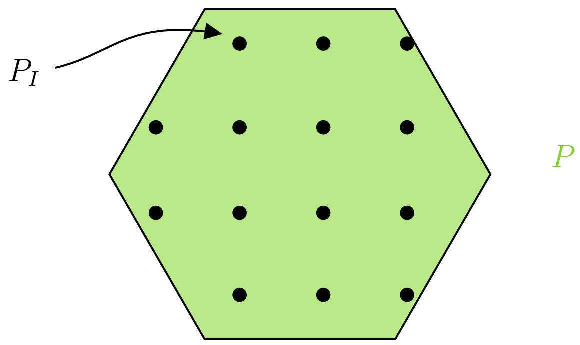
**Key Assumption: All data is rational** ($A \in \mathbb{Q}^{m \times n}, b \in \mathbb{Q}^m$) i.e, $Ax \le b$ is a rational polyhedron.

Let $P = \{x \in \mathbb{R}^n : Ax \le b\}$, $P_I = P \cap \{x_j \in \mathbb{Z} : j \in I\}$.

**Theorem 3.1**

$\text{conv}(P_I)$ is a polyhedron.

From now on, assume we have a pure IP.



**recession cone**

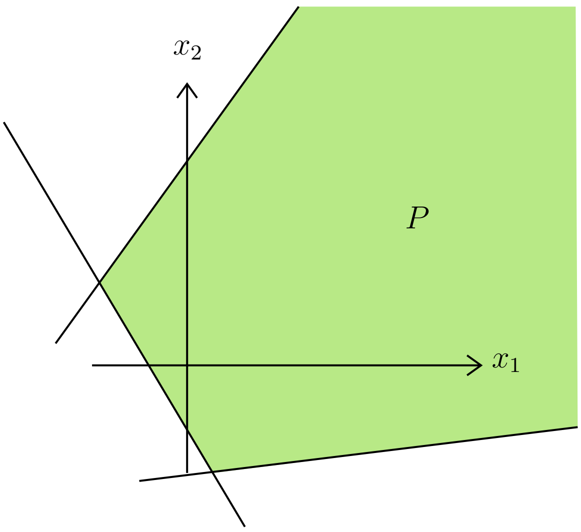
Let $P$ be a polyhedron. Its recession cone is

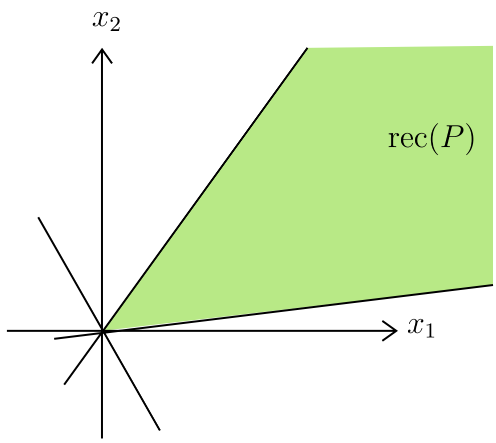
$$\operatorname{rec}(P) := \left\{ r \in \mathbb{R}^n : \begin{array}{l} \forall \overline{x} \in P \\ \forall \lambda \geq 0 \\ \overline{x} + \lambda r \in P \end{array} \right\}$$

**Lemma 3.2**

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

$$\underbrace{\operatorname{rec}(P)}_{R_1} = \underbrace{r \in \mathbb{R}^n : Ar \leq 0}_{R_2}$$




Proof:

$R_2 \subseteq R_1$) Let $\overline{x} \in P, \lambda \geq 0, r \in R_2$

$$A(\overline{x} + \lambda r) = A\overline{x} + \lambda A r \leq b \implies \overline{x} + \lambda r \in P \implies r \in R_1$$

$R_1 \subseteq R_2$) Let $r \notin R_2$, i.e., $\exists i : a_i^T r > 0$

Let $\overline{x} \in P$, it is clear $\exists \lambda > 0 : a_i^T(\overline{x} + \lambda r) > b_i \implies r \notin R_1$.

$\square$

**Theorem 3.3**

$P \neq \varnothing$ is a bounded polyhedron

$\iff P = \operatorname{conv}(x^1, \ldots, x^k)$ for some vectors $x^1, \ldots, x^k \in \mathbb{R}^n$.

*conv*$(x^1, \ldots, x^k)$ is smallest convex set containing $x^1, \ldots, x^k \iff$ set of all finite combinations of $x^1, \ldots, x^k$.

Proof:

$\Leftarrow$) $P = \left\{ x \in \mathbb{R}^n : \begin{array}{r} x = \sum_{i=1}^k \lambda_i x^i \\ \sum_{i=1}^k \lambda_i = 1 \\ \lambda \geq 0 \end{array} \right\}$

$P' = \left\{ (x, \lambda) \in \mathbb{R}^n \times \mathbb{R}^k : \begin{array}{r} x = \sum_{i=1}^k \lambda_i x^i \\ \sum_{i=1}^k \lambda_i = 1 \\ \lambda \geq 0 \end{array} \right\}$ is a bounded polyhedron.

$P = \text{proj}_x P'$ which is a bounded polyhedron.

$\Rightarrow$) $P$ bounded $\implies$ $P$ has no lines.

Let $x^1, \ldots, x^k$ be extreme points. Want to show $P = conv(x^1, \ldots, x^k)$

$P \supseteq conv(x^1, \ldots, x^k)$ follows since $P$ is a convex set containing $x^1, \ldots, x^k$.

Suppose $\exists \overline{x} \in P \setminus conv(x^1, \ldots, x^k)$

Consider

$$\begin{array}{llll} \min & 0^T \lambda & & \\ \downarrow & & & \\ & \sum_{i=1}^k \lambda_i x^i & = \overline{x} & \alpha \in \mathbb{R}^n \\ \text{s.t.} & \sum_{i=1}^k \lambda_i & = 1 & \alpha_0 \in \mathbb{R} \\ & \lambda & \geq 0 & \end{array} \tag{1}$$

and its dual

$$\begin{array}{ll} \max & \alpha^T \overline{x} + \alpha_0 \\ \text{s.t.} & \alpha^T x^i + \alpha_0 \leq 0, \quad \forall i = 1, \ldots, k \end{array} \tag{2}$$

$(\alpha, \alpha_0) = (0, 0)$ feasible for (2). By assumption, (1) is infeasible.

Let $(\overline{\alpha}, \overline{\alpha}_0)$ be such that $\overline{\alpha}^T \overline{x} + \overline{\alpha}_0 > 0$

Now consider

$$\begin{array}{ll} \max & \overline{\alpha}^T x + \overline{\alpha}_0 \\ \text{s.t.} & x \in P \end{array} \tag{3}$$

(3) has optimal solution since $P \neq \varnothing$ bounded and its has an optimal extreme point, i.e., $\overline{\alpha}^T x^i + \overline{\alpha}_0$ is optimal value. But by (2)

$$\overline{\alpha}^T x^i + \overline{\alpha}_0 \leq 0 < \overline{\alpha}^T \overline{x} + \overline{\alpha}_0$$

Contradiction.

$\square$

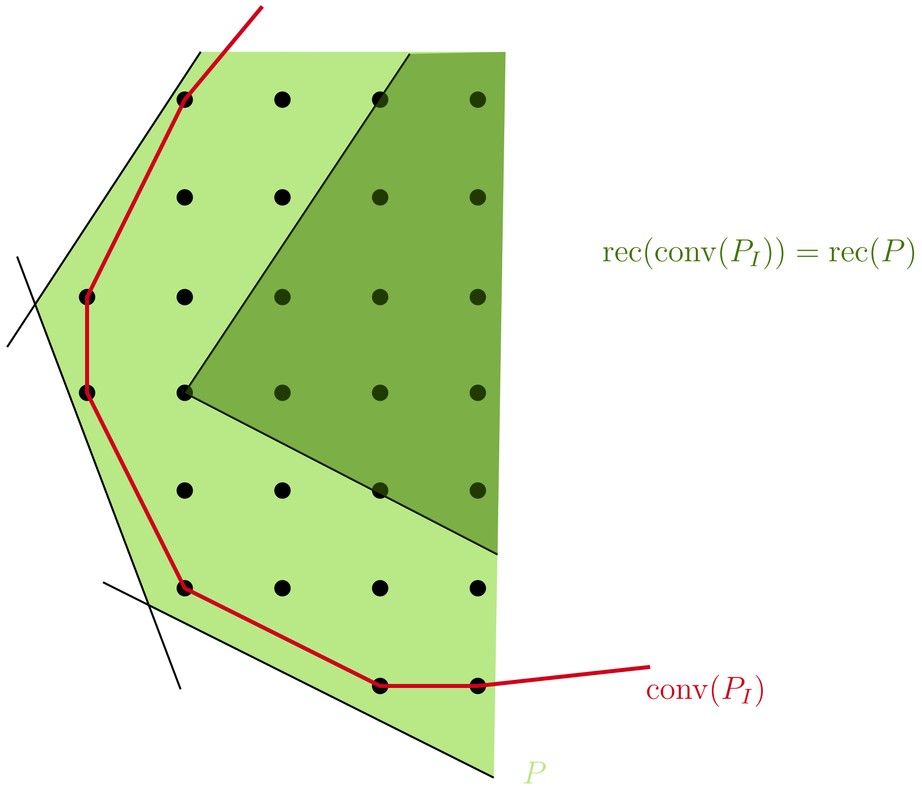
Back to IP...

**Theorem 3.4**

If $P$ is a rational polyhedron, then $\text{conv}(P_I)$ is also a rational polyhedron ($P_I = P \cap \mathbb{Z}^n$). Moreover, if $P_I \neq \varnothing$, $\text{rec}(\text{conv}(P_I)) = \text{rec}(P)$.

Proof:
Done if $P$ is bounded ($\{0\}$).

Skipped for unbounded $P$.




$\square$

**Theorem 3.5**

$$\begin{array}{ll} \max & c^T x \\ \text{s.t.} & x \in P_I \end{array} = \begin{array}{ll} \max & c^T x \\ \text{s.t.} & \text{conv}(P_I) \end{array}$$

**Note**

1. Using Fund Thm of LP. I know IP is either infeas., unbounded, or $\exists$ opt. sol.

2. If $P_I \neq \varnothing$, then unboundedness can be detected by checking if $\begin{array}{ll} \max & c^T x \\ \text{s.t.} & x \in P \end{array}$ is unbounded. Since $\begin{array}{ll} \max & c^T x \\ \text{s.t.} & x \in P \end{array}$ unbounded iff $P \neq \varnothing$ and $\exists r : \begin{array}{c} c^T r > 0 \\ Ar \leq 0 \end{array}$.

$P_I \neq \varnothing \implies P \neq \varnothing$. But then this implies $\begin{array}{ll} \max & c^T x \\ \text{s.t.} & x \in conv(P_I) \end{array}$ unbounded.

Proof:
WMA (we may assume) $P_I \neq \varnothing$.

Let $z_1 = \begin{array}{ll} \max & c^T x \\ \text{s.t.} & x \in P_I \end{array}$, $z_2 = \begin{array}{ll} \max & c^T x \\ \text{s.t.} & x \in conv(P_I) \end{array}$.

Since $P_I \subseteq conv(P_I) \implies z_1 \leq z_2$.

Now let $x^* \in conv(P_I) \implies \begin{array}{r} x^* = \sum_{i=1}^k \lambda_i x^i \\ \sum_{i=1}^k \lambda_i = 1 \\ \lambda \geq 0 \end{array}$ for $x^1, \ldots, x^k \in P_I$.

$\implies \exists i : c^T x^i \geq c^T x^*$ since otherwise

$$c^T x^* = \sum_{i=1}^k \lambda_i (c^T x^*) > \sum_{i=1}^k \lambda_i (c^T x^i) = c^T \left( \sum_{i=1}^k \lambda_i x^i \right) = c^T x^*$$

contradiction $\implies z_1 \geq z_2$. □

**Corollary 3.6**

If $P \neq \varnothing$ and pointed. Then $\text{conv}(P_I)$ is pointed and any extreme point of $\text{conv}(P_I)$ is integral.

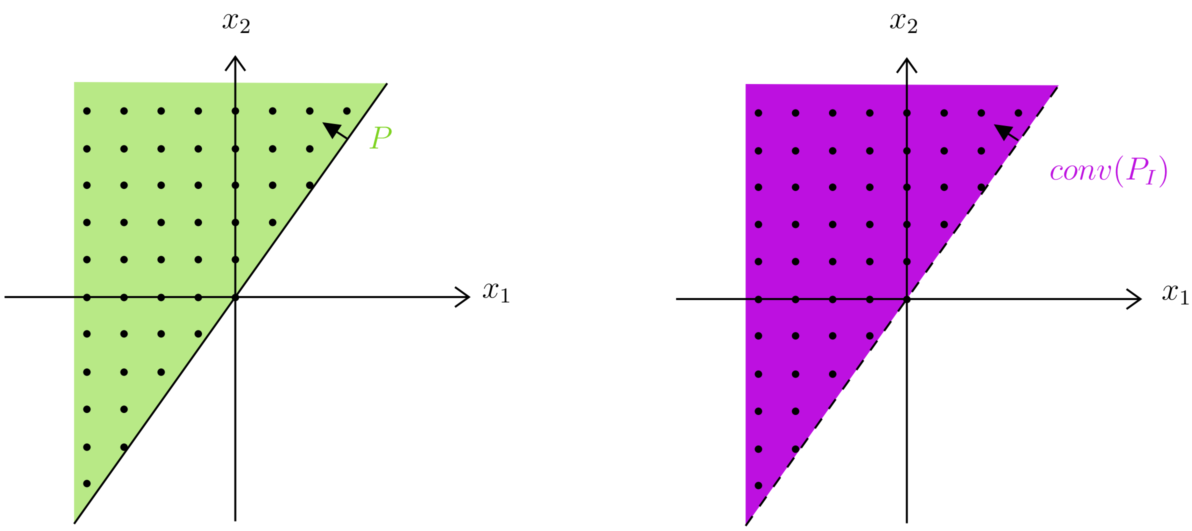
Proof:
$\text{rec}(P) = \text{rec}(conv(P_I))$ implies $conv(P_I)$ pointed.

Let $x^*$ be extreme point of $conv(P_I)$. Let $c$ be such that $x^*$ is unique optimal solution to $\begin{array}{ll} \max & c^T x \\ \text{s.t.} & x \in conv(P_I) \end{array}$.

By theorem, $\exists \overline{x} \in P_I : c^T \overline{x} = c^T x^*$.

By uniqueness of $x^*$, $\overline{x} = x^*$, then $x^*$ is integral. □

**Note**
$P = \{x \in \mathbb{R}^2 : x_2 \geq \sqrt{2} x_1\}$




$conv(P_I)$ is not even closed (dotted line plus $(0,0)$), NOT a polyhedron.

## 3.1 Cutting Plane Algorithm

$$\begin{array}{ll} \max & c^T x \\ \text{s.t.} & x \in P_I := P \cap \mathbb{Z}^n \end{array} \tag{IP}$$

where $P$ is rational polyhedron.

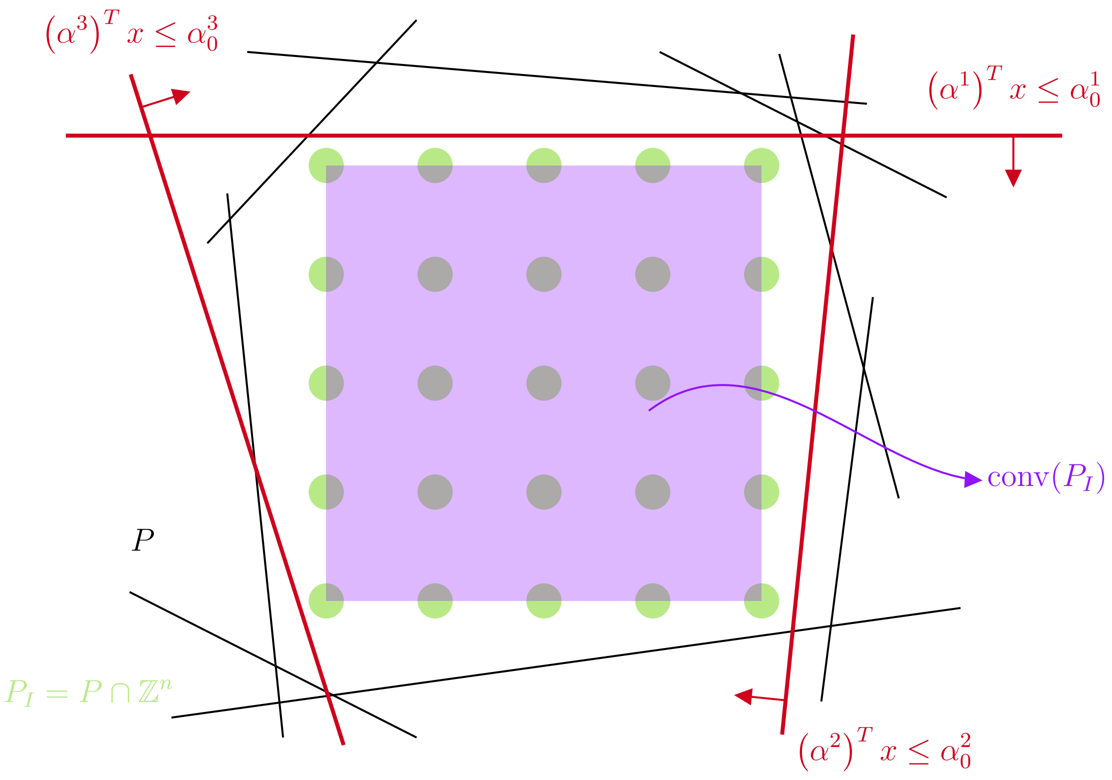
We know it can be solved by solving $\begin{array}{ll} \max & c^T x \\ \text{s.t.} & \text{conv}(P_I) \end{array}$

**Problem** Hard to compute $\text{conv}(P_I)$.

$\text{conv}(P_I)$ is smallest convex set containing $P_I$. $P$ is **a** convex set containing $P_I$.

### Idea

- Start with $P$
- Iteratively make $P$ "closer" to $\text{conv}(P_I)$




**Idea 2** Want to know only part of $\mathrm{conv}(P_I)$ that is in the "direction I am optimizing".

**LP relaxation**

The LP you obtain from (IP) after dropping integrality, i.e.,

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

**valid ineq**

An ineq $\alpha^T x \leq \alpha_0$ is valid for $S \subseteq \mathbb{R}^n$ if $\forall \overline{x} \in S$: $\alpha^T \overline{x} \leq \alpha_0$.

**Assumption** LP relaxation has an optimal solution.

If $P = \varnothing$, then $P_I = \varnothing$. If LP relaxation is unbounded, either $P_I = \varnothing$ or (IP) is unbounded.

**Algorithm 5:** Cutting Plane Algorithm

$R \leftarrow P$
**do**
    Let $x^*$ be optimal solution to $\begin{array}{ll} \max & c^T x \\ \text{s.t.} & x \in R \end{array}$
    **if** $x^*$ *is integral* **then**
        `STOP // ` $x^*$ ` is opt sol for (IP)`
    **else**
        Find valid ineq $\alpha^T x \leq \alpha_0$ for $\text{conv}(P_I)$ s.t. $\alpha^T x^* > \alpha_0$
        $R \leftarrow R \cap \{x : \alpha^T x \leq \alpha_0\}$
    **end**
**while** $R \neq \varnothing$;
Declare (IP) infeasible

Issues...

1. $\alpha, \alpha_0$ must be rational
2. Finiteness?
3. How to find $\alpha, \alpha_0$?

**Note**

Any any point $P_I \subseteq \text{conv}(P_I) \subseteq R \subseteq P$.

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

If $x^* \in \mathbb{Z}^n$, then $x^* \in P_I$.

$$\implies \begin{array}{ll} \max & c^T x \\ \text{s.t.} & x \in P_I \end{array} \geq c^T x^* \implies x^* \text{ is optimal for } P_I$$

To solve the issues, impose $x^*$ being an opt. BFS of $\begin{array}{ll} \max & c^T x \\ \text{s.t.} & x \in R \end{array}$

**Proposition 3.7**

Let $R$ be a pointed rational polyhedron such that $R \cap \mathbb{Z}^n = P_I$. Let $x^*$ be a BFS of $R$.

Then $x^*$ is integral $\iff$ $x^* \in \text{conv}(P_I)$

Proof:
Exercise. $\square$

How to find valid ineq for $\text{conv}(P_I)$ $\alpha_T x \leq \alpha_0$ s.t. $\alpha^T x^* > \alpha_0$?

Call such ineq. a **CUTTING PLANE** or a **CUT** separating $\text{conv}(P_I)$ and $x^*$.

**Assumption** $R = \left\{ x \in \mathbb{R}^n : \begin{array}{l} Ax = b \\ x \geq 0 \end{array} \right\}.$

$$\begin{array}{ll} \max & c^T x \\ \downarrow & \\ \text{s.t.} & \begin{array}{l} Ax = b \\ x \geq 0 \end{array} \end{array} \tag{1}$$

Let $B$ be opt. basis.

$$(1) \quad \iff \quad \begin{array}{ll} \max & \overline{c}_N^T x_N + c_B^T A_B^{-1} b \\ \downarrow & \\ \text{s.t.} & \begin{array}{l} x_B + \overbrace{A_B^{-1} A_N}^{\overline{A}_N} x_N = \overbrace{A_B^{-1} b}^{\bar{b}} \\ x \geq 0 \end{array} \end{array}$$

$$x^* \text{ is integral} \iff A_B^{-1} b \in \mathbb{Z}^m$$

If $x^*$ is not integral, then $\exists i \in \{1, \ldots, m\} : (A_B^{-1} b)_i \notin \mathbb{Z}$.

Look at constraint

$$x_i + \sum_{j \in N} \overline{a}_{ij} x_j = \overline{b}_i$$

is valid for $P_I$ since it is valid for $R$.

$$x_i + \sum_{j \in N} \lfloor \overline{a}_{ij} \rfloor x_j \leq \overline{b}_i$$

is valid for $P_I$ since it is valid for $R$.

Since $\lfloor \overline{a}_{ij} \rfloor \leq \overline{a}_{ij}$ and $x_j \geq 0 \implies \lfloor \overline{a}_{ij} \rfloor x_j \leq \overline{a}_{ij} x_j$.

Since LHS is integer $\forall x \in P_I$,

$$x_i + \sum_{j \in N} \lfloor \overline{a}_{ij} \rfloor x_j \leq \lfloor \overline{b}_i \rfloor \tag{$\star$}$$

is valid for $P_I$.

> **Note**
>
> For $x^*$, $\quad x_j^* = 0, \forall j \in N \ x_i^* = \overline{b}_i$.
>
> Thus
>
> $$x_i^* + \sum_{j \in N} \lfloor \overline{a}_{ij} \rfloor x_j^* = \overline{b}_i > \lfloor \overline{b}_i \rfloor$$

($\star$) is the cut we wanted. Called a Chvátal-Gomory (CG) cut.

**Algorithm 6:** Cutting Plane Algorithm (Correct)

$R \leftarrow P$ // ($P$ pointed)
**do**
  Let $x^*$ be optimal BFS solution to $\begin{array}{ll} \max & c^T x \\ \text{s.t.} & x \in R \end{array}$
  **if** $x^*$ *is integral* **then**
    STOP // $x^*$ is opt sol for (IP)
  **else**
    Find valid ineq $\alpha^T x \leq \alpha_0$ for $\operatorname{conv}(P_I)$ s.t. $\alpha^T x^* > \alpha_0$
    $R \leftarrow R \cap \{x : \alpha^T x \leq \alpha_0\}$
  **end**
**while** $R \neq \varnothing$;
Declare (IP) infeasible

**Theorem 3.8**

The cutting plane algorithm using CG cuts terminates in finitely many iterations (for pure IPs).

Proof:
SKIPPED. $\square$

Example:

$$\begin{array}{ll} \max & \begin{pmatrix} 1 & 3 & -2 & 0 & 0 \end{pmatrix} x \\ \downarrow & \\ \text{s.t.} & \begin{pmatrix} 1 & 2 & 1 & 1 & 0 \\ 1 & -1 & 3 & 0 & 1 \end{pmatrix} x = \begin{pmatrix} 3 \\ 1 \end{pmatrix} \\ & x \geq 0, \qquad x \in \mathbb{Z}^5 \end{array}$$

Opt basis for LP relaxation: $B = \{2, 5\}$.

In canonical form:

$$\begin{array}{ll} \max & \begin{pmatrix} -0.5 & 0 & -3.5 & -1.5 & 0 \end{pmatrix} x + 4.5 \\ \downarrow & \\ \text{s.t.} & \begin{pmatrix} 0.5 & 1 & 0.5 & 0.5 & 0 \\ 1.5 & 0 & 3.5 & 0.5 & 1 \end{pmatrix} x = \begin{pmatrix} 1.5 \\ 2.5 \end{pmatrix} \\ & x \geq 0 \end{array}$$

and $x^* = \begin{pmatrix} 0 & 1.5 & 0 & 0 & 2.5 \end{pmatrix}^T$

**CG-cut**:

$$\begin{array}{rl} 0x_1 + x_2 + 0x_3 + 0x_4 + 0x_5 \leq 1 \iff x_2 \leq 1 & \text{From 1st constraint} \\ x_1 + 3x_3 + x_5 \leq 2 & \text{CG-cut from 2nd constraint} \end{array}$$

Can add both to $R$.

**New LP**

$$\begin{array}{ll} \max & \begin{pmatrix}1 & 3 & -2 & 0 & 0\end{pmatrix} x \\ \downarrow & \\ \text{s.t.} & \begin{pmatrix}1 & 2 & 1 & 1 & 0 \\ 1 & -1 & 3 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 3 & 0 & 1\end{pmatrix} x \begin{array}{c} = \\ = \\ \leq \\ \leq \end{array} \begin{pmatrix}3 \\ 1 \\ 1 \\ 2\end{pmatrix} \\ & x \geq 0 \end{array}$$

Add $x_6, x_7 \geq 0$ convert to SEF, where

$$x_2 + x_6 = 1, \qquad x_1 + 3x_3 + x_5 + x_7 = 2$$

If $x_1, \ldots, x_5 \in \mathbb{Z}$, then $x_6, x_7 \in \mathbb{Z}$.

New Opt for LP:

$$x^T = \begin{pmatrix}1 & 1 & 0 & 0 & 1 & 0 & 0\end{pmatrix}$$

So opt sol to original LP is $\begin{pmatrix}1 & 1 & 0 & 0 & 1\end{pmatrix}$.

## 3.2 Total Unimodularity

**totally unimodular**

A matrix $U$ is called totally unimodular (TU) if all its square submatrices have determinant in $\{-1, 0, 1\}$.

Example:

$\begin{pmatrix}\boxed{2} & 0 & 0 \\ 0 & 0 & 0\end{pmatrix}$ is not TU.

$\begin{pmatrix}\boxed{1} & 1 & \boxed{-1} & 0 \\ 0 & 0 & 0 & 0 \\ \boxed{1} & 0 & \boxed{1} & 1\end{pmatrix}$ is NOT TU.

**Note**
Square submatrices are obtained by deleting rows/columns.

$\begin{pmatrix}0 & -1 & 1 \\ 1 & 1 & -1\end{pmatrix}$ is TU.

**Theorem 3.9**

If $A \in \mathbb{Z}^{m \times n}$ is TU and $b \in \mathbb{Z}^m$ then every BFS of $P = \left\{ x \in \mathbb{R}^n : \begin{array}{c} Ax = b \\ x \geq 0 \end{array} \right\}$ is integral.

Recall

**Cramer's Rule**

If $D$ is $n \times n$ invertible, then unique solution to $Dx = b$ is given by

$$x_i = \frac{\det D(i)}{\det D}$$

where $D(i)$ is $D$ replacing $i$-th column with $b$.

Example:

$$\begin{pmatrix} 1 & -1 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

Solution

$$x_1 = \frac{\det \begin{pmatrix} 2 & -1 \\ 1 & 3 \end{pmatrix}}{\det \begin{pmatrix} 1 & -1 \\ 0 & 3 \end{pmatrix}} = \frac{7}{3}, \qquad x_2 = \frac{\det \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}}{\det \begin{pmatrix} 1 & -1 \\ 0 & 3 \end{pmatrix}} = \frac{1}{3}$$

Proof:

Let $x^*$ be a BFS of $\left\{ x : \begin{array}{c} Ax = b \\ x \geq 0 \end{array} \right\}$, $B$ corresponding basis.

Then $x_B^* = A_B^{-1} b, x_N^* = 0$

**Note** $x_B^*$ is unique solution to $A_B x_B = b$

$\implies$ By Cramer's rule,

$$x_i^* = \frac{\det A_B(i)}{\det A_B} \in \mathbb{Z}$$

since $\det A_B(i) \in \mathbb{Z}$ and by TU, $\det A_B \in \{1, -1\}$ which cannot be 0 since invertible. $\square$

**Note**

Result remains true if $P = \{x : Ax \leq b\}$ or $P = \left\{ x : \begin{array}{c} Ax \leq b \\ x \geq 0 \end{array} \right\}$

**integral**

We say a polyhedron is integral if all its extreme points are integral.

**Lemma 3.10**

*$P$ is an integral polyhedron iff $P = \operatorname{conv}(P \cap \mathbb{Z}^n)$.*

Proof:
Exercise. $\square$

**Lemma 3.11**

Let $A \in \mathbb{Z}^{m \times n}$ TU.

Then applying any of the following operations on $A$ yields a TU matrix.

a) Delete row/column

b) Multiply row/column by $-1$

c) Permute rows/columns

d) Transpose

e) Duplicate row/column

f) Add a row/column with at most one nonzero entry, which is in $\{+1, -1\}$.

Proof:

a) ✓

b)-d) Potentially changes signs of det.

e) Only can create new submatrices if row and its duplicate are in it. But that has det $= 0$.

f) Recall

**Laplace formula**

$D$ square:

$$D = \begin{pmatrix} & | & \\ -- & d_{ij} & -- \\ & | & \end{pmatrix}$$

Let $M_{ij}$ be the matrix obtained by deleting row $i$, column $j$.

Then for any row $i$ of $D$:

$$\det(D) = \sum_j (-1)^{i+j} d_{ij} \det(M_{ij})$$

For any column $j$:

$$\det(D) = \sum_i (-1)^{i+j} d_{ij} \det(M_{ij})$$

$$A' = \left(\begin{array}{c|c} 1 & \\ 0 & \\ 0 & \\ \vdots & A \\ \vdots & \\ \vdots & \\ 0 & \end{array}\right)$$

Let $D$ be square submatrix of $A'$. If $D$ does not contain first col, then $\det(D) \in \{\pm 1, 0\}$ since $A$ is TU.

If $D$ does not contain first row, but contains first column, then $\det(D) = 0$.

Else,

$$D = \left(\begin{array}{c|ccccc} 1 & \times & \times & \times & \times & \times \\ \hline 0 & & & & & \\ \vdots & & & \overline{D} & & \\ 0 & & & & & \\ 0 & & & & & \end{array}\right)$$

By Laplace formula: $|\det(D)| = |\det\big(\overline{D}\big)| \in \{0, 1\}$.

$\square$

**Application 1** Suppose $A$ is TU $\in \mathbb{Z}^{m\times n}$. If $b \in \mathbb{Z}^m$ and $\ell, u \in \mathbb{Z}^n$, then

$$P = \left\{ x \in \mathbb{R} : \begin{array}{c} Ax \leq b \\ \ell \leq x \leq u \end{array} \right\}$$

is integer polyhedron.

$$P = \left\{ x \in \mathbb{R}^n : \underbrace{\begin{pmatrix} A \\ I \\ -I \end{pmatrix}}_{A'} x \leq \underbrace{\begin{pmatrix} b \\ u \\ -\ell \end{pmatrix}}_{b'} \right\}$$

$b'$ integral, $A'$ TU $\implies P$ is integral

**Application 2** $A \in \mathbb{Z}^{m\times n}$ TU, $b \in \mathbb{Z}^m$, $c \in \mathbb{Z}^n$, then

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

have integral opt solutions (if both are feasible).

## 3.3 Sufficient condition for TU

**Lemma 3.12**

Let $A \in \mathbb{Z}^{m\times n}$ with entries $\{-1,0,1\}$. If $A$ has:

- At most two nonzeros per column, AND
- There exists a partition $I_1, I_2$ of its rows such that, for every column:
  i) Nonzero entries of same sign lie in different partitions
  ii) Nonzero entries of opposite signs lie in same partition.

Then $A$ is TU.

Example:

$$A = \left(\begin{array}{ccccccc} 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & -1 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ \hline 1 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 1 \end{array}\right)$$

above the line: $I_1$; below: $I_2$. $A$ is TU.

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Line 1 and line 3: $I_1$; Line 2 and 4: $I_2$. $A$ is TU.

Proof:
Suppose Lemma is False. Let $M$ be a minimal counterexample, i.e.,

- $M$ is not TU,
- $M$ satisfies conditions of Lemma,
- Any submatrix of $M$ is TU.

Then $M$ itself is a square matrix with $\det(M) \notin \{-1,0,1\}$ and all its submatrix have det $\in \{-1,0,1\}$.

If $M$ has $\leq 1$ nonzero in a column, then $M$ is obtained by adding a column with at most 1 nonzero to a TU matrix $\implies$ $M$ is TU (By Lemma 3.11).

Thus, we may assume all columns of $M$ has exactly two nonzero elements.

$$M = \begin{pmatrix} - & M_1^T & - \\ & \vdots & \\ - & M_m^T & - \end{pmatrix}$$

Consider:

$$\sum_{i \in I_1} M_i - \sum_{i \in I_2} M_i = 0$$

since i) and ii) hold. Then this means $\{M_i\}_{i=1}^m$ are **not** linearly independent, which implies $\det(M) = 0$. $\square$

Example:

Given $G = (V, E)$ undirected simple graph.

$G$ is bipartite if $V = \underbrace{V_1 \dot\cup V_2}_{\text{disjoint union}}$ and $\forall u, v \in E$ has $u \in V_1, v \in V_2$.

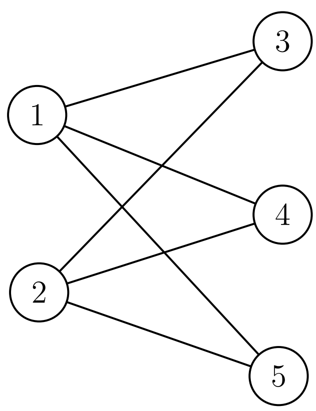
$M \subseteq E$ is a matching if $|M \cap \delta(v)| \leq 1, \forall v \in V$ where $\delta(v) := \{e \in E : v \text{ is an endpoint of } e\}$.

Given $G$ bipartite. **Goal**: Find max carnality matching.

Let $x_e \in \{0, 1\}$ and $x_e = \begin{cases} 1, & \text{if } e \in M \\ 0, & \text{if } e \notin M \end{cases}.$

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

Let's now take a look at example.



$$x = \begin{pmatrix} x_{13} & x_{14} & x_{15} & x_{23} & x_{24} & x_{25} \end{pmatrix}^T$$

$$\begin{array}{ll}
\max & \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix} x \\
\downarrow & \\
\text{s.t.} & \left(\begin{array}{cccccc} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ \hline 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right) x \leq \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \quad \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ \text{vertex} \end{array} \\
& x \in \{0,1\}^E
\end{array}$$

In general:

- $I_1 \to$ constraints correspond to $V_1$
- $I_2 \to$ constraints correspond to $V_2$

If we look at a column $x_{uv}$, it will have a 1 in row of $u$ a 1 in row of $v$, 0 everywhere else.

$\to$ Bipartite $\implies$ Lemma is satisfied $\implies$ (1) can be solved via LP.

Let (2) be LP relaxation of (1) without $x_e \leq 1, \forall e \in E$, otherwise the first constraint is violated.

$$\begin{array}{ll}
\max & \sum_{e \in E} x_e \\
\downarrow & \\
\text{s.t.} & \sum_{e \in \delta(v)} x_e \leq 1, \quad \forall c \in V \\
& x \geq 0
\end{array} \tag{2}$$

Let us write the dual of (2)

$$\begin{array}{ll}
\min & \sum_{v \in V} y_v \\
\downarrow & \\
\text{s.t.} & y_u + y_v \geq 1, \quad \forall uv \in E \\
& y \geq 0
\end{array} \tag{3}$$

and add integral constraints,

$$\begin{array}{ll}
\min & \sum_{v \in V} y_v \\
\downarrow & \\
\text{s.t.} & y_u + y_v \geq 1, \quad \forall uv \in E \\
& y \in \{0,1\}^V
\end{array} \tag{4}$$

Let $z_i$ be the optimal value for (i) then

$$z_1 \leq z_2 = z_3 \leq z_4$$

$$G \text{ bipartite } \implies \begin{array}{l} z_1 = z_2 \\ z_3 = z_4 \end{array}$$

**Vertex Cover**: such that $\forall e \in E, |e \cap U| \geq 1$. **Problem**: Finding smallest vertex cover.

**König's Theorem**

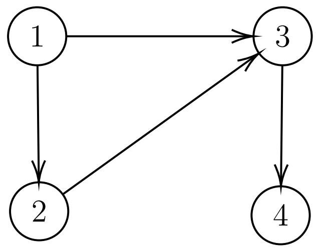
In bipartite graph $G$, size of largest matching = size of smallest vertex cover.

Example:

Consider a directed graph $D = (V, A)$.

Incidence matrix of $D$ has one row per vertex, one column per arc.

For $v \in V$, $(w, y) \in A$, then $a_{ve} = \begin{cases} -1, & \text{if } v = w \\ 1, & \text{if } v = y \\ 0, & \text{otherwise} \end{cases}$




$$\begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{pmatrix} -1 & -1 & 0 & 0 \\ 1 & 0 & 1 & -1 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

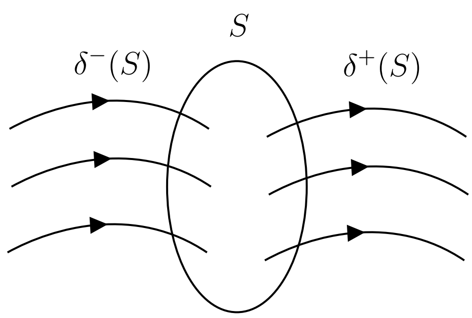
$I_1$ = everything, $I_2 = \varnothing \implies$ Matrix is TU

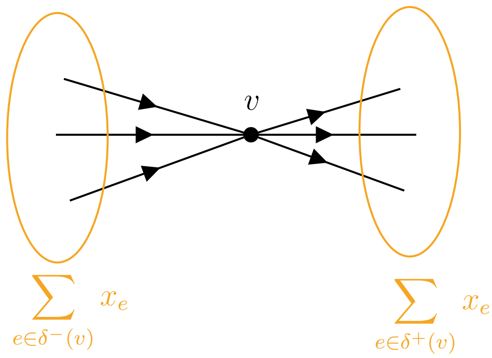
**Max Flow**: Given $D = (V, A)$, $s, t \in V (s \neq t)$. An $s$-$t$ flow is a nonnegative vector $x \in \mathbb{R}^A$, where

$$\sum_{e \in \delta^-(v)} x_e - \sum_{e \in \delta^+(v)} x_e = 0, \quad \forall v \in V \setminus \{s, t\}$$

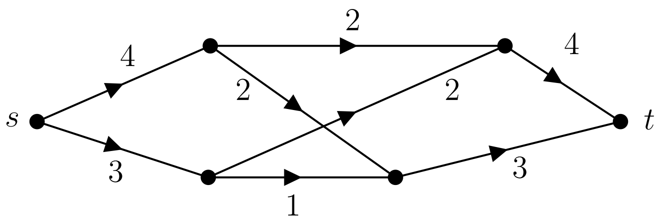
where

$$\delta^-(S) = \left\{ (u, v) \in A : \begin{array}{c} u \notin S \\ v \in S \end{array} \right\} \qquad \text{and} \qquad \delta^+(S) = \left\{ (u, v) \in A : \begin{array}{c} u \in S \\ v \notin S \end{array} \right\}$$






**Goal**: Find a flow maximizing $\sum_{e \in \delta^+(S)} x_e$




also $0 \le x_e \le c_e, \forall e \in A$ where $c_e$ is some capacity constraint.

TU $\implies$ max flow is integral if $c_e \in \mathbb{Z}, \forall e \in A$.

**Theorem 3.13**

An $m \times n$ integral matrix $A$ is TU iff for every subset $R \subseteq \{1, \dots, m\}$, there exists a partition of $R$ into $R_1, R_2$ (that is, $R_1 \cup R_2 = R$ and $R_1 \cap R_2 = \varnothing$) such that

$$\sum_{i \in R_1} a_{ij} - \sum_{i \in R_2} a_{ij} \in \{-1, 0, 1\}, \forall j = 1, \dots, n$$

**Note**

Careful that in the previous result that we had seen, we just needed to partition the original rows into two such sets.

This result says that if I pick ANY SUBSET of rows, I must be able to do the same.

Skipped branch-and-bound, Minimum Cost Perfect Matching in Bipartite Graphs... due to one week suspension

# 4 Nonlinear Programming

The general form: Let $f, g_1, \ldots, g_m : \mathbb{R}^m \to \mathbb{R}$.

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & g_i(x) \leq 0, \quad \forall i = 1, \ldots, m \end{aligned} \tag{NLP}$$

Note that this is minimization problem with "$\leq$" constraints.

**Example: Linear Programs**

$f(x) := c^T x$ and $g_i(x) := a_i^T x - b_i$. These give us

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & a_i^T x \leq b_i, \quad \forall i = 1, \ldots, m \end{aligned}$$

**Example: Binary integer program**

Let $f(x) := c^T x$, $g_1(x) := x_1(1 - x_1)$ and $g_2(x) := -x_1(1 - x_1)$. These give us

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & x_1(1 - x_1) = 0 \end{aligned}$$

where the constraint is equivalent to $x_1 \in \{0, 1\}$. Extend it to

$$\begin{aligned} \min \quad & c^T x \\ \downarrow & \\ \text{s.t.} \quad & \begin{aligned} & Ax \leq b \\ & x \in \{0, 1\}^n \end{aligned} \end{aligned}$$

## 4.1 Convex functions

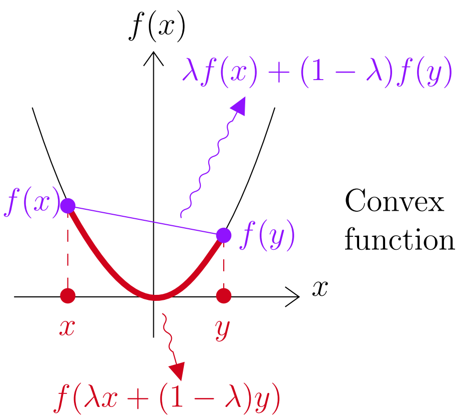
**convex functions**

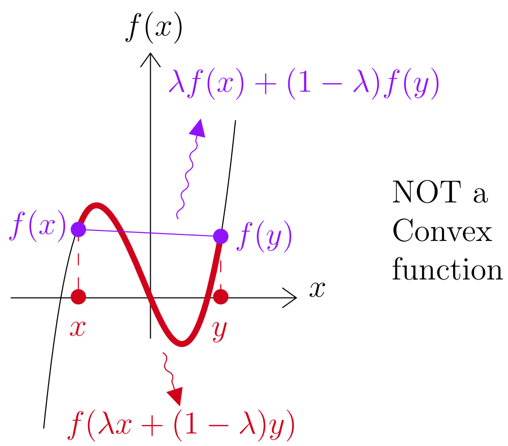
Let $S \subseteq \mathbb{R}^n$ be a convex set. The function $f : S \to \mathbb{R}^n$ is a convex function if $\forall x, y \in S, \forall \lambda \in [0, 1]$,

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

Example:

Here we let $S = \mathbb{R}$.




A **convex NLP** is one of the form:

$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & g_i(x) \leq 0, \quad \forall i = 1, \ldots, m \end{array} \tag{CVX}$$

where $f, g_1, \ldots, g_m \colon \mathbb{R}^n \to \mathbb{R}$ are convex functions.

**Note**

It is important that constraints are $\leq$ and that the objective is a minimization problem.

**Proposition 4.1**

If $g : \mathbb{R}^n \to \mathbb{R}$ is a convex function, then $S = \{x \in \mathbb{R}^n : g(x) \leq 0\}$ is a convex set.

Proof:
Let $x, y \in S$, i.e., $g(x) \leq 0, g(y) \leq 0$. Now we want to prove $\lambda x + (1-\lambda)y \in S$.

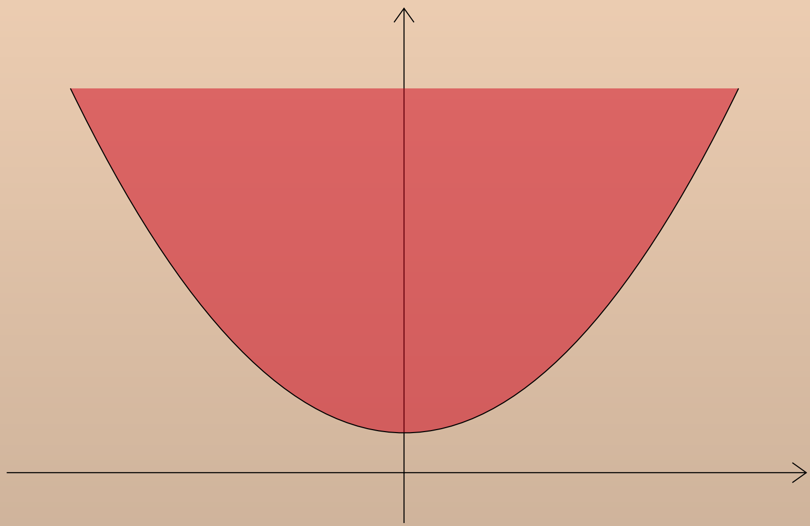
$$\begin{aligned} g(\lambda x + (1-\lambda)y) &\leq \lambda g(x) + (1-\lambda)g(y) \quad \text{since } g \text{ is a convex function} \\ &\leq 0 \end{aligned}$$

where the last ineq is from $\begin{aligned} g(x) \leq 0, \lambda \geq 0 \\ g(y) \leq 0, (1-\lambda) \geq 0 \end{aligned}$.

This implies $\lambda x + (1-\lambda)y \in S, \quad \forall \lambda \in [0,1]$. □

**epigraph**

$$\text{epi}(f) = \{(x,y) : y \geq f(x)\}$$



$f$ is convex $\iff$ epi$(f)$ is convex.

## 4.2 Gradients & Hessian

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a twice differentiable function.

The **gradient** of $f$ at $\overline{x}$ is the vector

$$\nabla f(\overline{x}) = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix}$$

The **Hessian** of $f$ at $\overline{x}$ is the $n \times n$ symmetric matrix

$$\nabla^2 f(\overline{x})$$

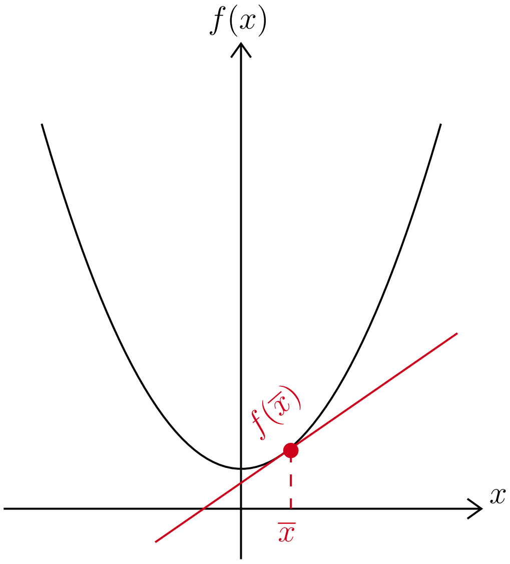
where the element is defined as

$$\left[\nabla^2 f(\overline{x})\right]_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_i}$$

Example:
$f(x) = x_1^2 x_2 + 2x_1 + 3$. Then

$$\nabla f(x) = \begin{pmatrix} 2x_1x_2 + 2 \\ x_1^2 \end{pmatrix} \quad \text{and} \quad \nabla^2 f(x) = \begin{pmatrix} 2x_2 & 2x_1 \\ 2x_1 & 0 \end{pmatrix}$$

Now looking at 1-D convex functions, two key properties stand out:




- second derivative is $\geq 0$ (at any point $\overline{x}$)
- value of $f$ is above tangent line at $\overline{x}$

Translating:

- $f''(x) \geq 0, \forall x$
- $f(x) \geq f(\overline{x}) + f'(\overline{x})(x - \overline{x}), \forall x, \overline{x}$

**Theorem 4.2**

Let $S \subseteq \mathbb{R}$ be a convex set. Let $S \to \mathbb{R}$ be twice differentiable. TFAE:

a) $f$ is convex on $S$

b) $f(x) \geq f(\overline{x}) + f'(\overline{x})(x - \overline{x}), \forall x, \overline{x} \in S$

c) $(f'(x) - f'(\overline{x}))(x - \overline{x}) \geq 0, \forall x, \overline{x} \in S$

d) $f''(x) \geq 0, \forall x \in S$.

*What is the generalization of b), c), d) to $f : \mathbb{R}^n \to \mathbb{R}$?*

b): $f(x) \geq f(\overline{x}) + \nabla f(\overline{x})^T(x - \overline{x}), \quad \forall x, \overline{x} \in S$.

c): $(\nabla f(x) - \nabla f(\overline{x}))^T(x - \overline{x}) \geq 0, \quad \forall x, \overline{x} \in S$.

d): $\nabla^2 f(x)$ is Positive Semidefinite (PSD), $\forall x \in S$.

**Note**

A symmetric $n \times n$ matrix $Q$ is said to be **positive semidefinite** if $\forall y \in \mathbb{R}^n$,

$$y^T Q y \geq 0$$

Denoted as $Q \succeq 0$.

$Q$ is said to be **positive definite** (PD) if $\forall y \in \mathbb{R}^n, y \neq 0$,

$$y^T Q y > 0$$

Denoted as $Q \succ 0$.

**Theorem 4.3**

Let $S \subseteq \mathbb{R}^n$ be a convex set. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a continuous twice differentiable function. TFAE:

a) $f$ is convex on $S$

b) $f(x) \geq f(\overline{x}) + \nabla f(\overline{x})^T(x - \overline{x}), \quad \forall x, \overline{x} \in S$

c) $(\nabla f(x) - \nabla f(\overline{x}))^T(x - \overline{x}) \geq 0, \quad \forall x, \overline{x} \in S$

d) $\nabla^2 f(x) \succeq 0$, $\forall x \in S$.

Example:

$f(x) = \|x\|^2 = \sum_{j=1}^{n} x_j^2$

$$\nabla f(x) = \begin{pmatrix} 2x_1 \\ \vdots \\ 2x_n \end{pmatrix} \quad \text{and} \quad \nabla^2 f(x) = 2I$$

Now

$$y^T \nabla^2 f(x) y = 2y^T I y = 2y^T y = 2\|y\|^2 \geq 0$$

$\implies \nabla^2 f(x) \succeq 0, \forall x \implies f(x)$ is convex.

Example:

$f(x) = \frac{1}{2}x^T x Q x + d^T x + p$ where $Q$ is PSD.

$$f(x) = \sum_{j=1}^{n} \frac{x_j^2}{2} g_{jj} + \frac{1}{2} \sum_{i=1}^{n} \sum_{j>i} 2x_i x_j q_{ij} + \sum_{j=1}^{n} x_j d_j + p$$

$$\nabla f(x) = \begin{pmatrix} \frac{2x_1}{2}q_{11} + \sum_{j=2}^{n} x_j q_{ij} + d_1 \\ \vdots \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^{n} x_j q_{ij} + d_1 \\ \vdots \end{pmatrix} = Qx + d$$

$\nabla^2 f(x) = Q \succeq 0 \implies f$ is convex.

## 4.3 Local vs. Global optimality

Consider an NLP

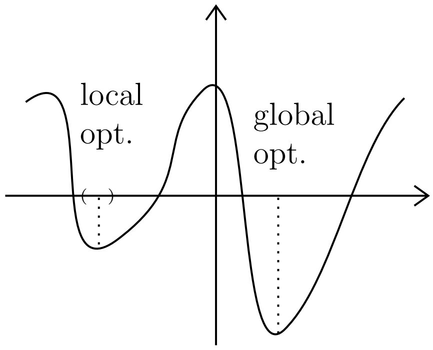
$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & g_i(x) \leq 0, \quad \forall i = 1, \dots, m \end{array} \tag{NLP}$$

Let $S$ be its feasible region. $x^* \in S$ is said to be a **local optimum** if $\exists R > 0$ so that

$$f(x^*) \leq f(x), \quad \forall x \in B(x^*, R) \cap S.$$

$x^*$ is said to be a **global optimum** if

$$f(x^*) \leq f(x), \quad \forall x \in S.$$




**Proposition 4.4**

If (NLP) is a convex program, then

$$x^* \text{ is a local optimum} \iff x^* \text{ is a global optimum.}$$

Proof:

($\Leftarrow$) Trivial.

($\Rightarrow$) Suppose $x^*$ is a local optimum. But suppose $\exists \overline{x} \in S$: $f(x^*) > f(\overline{x})$.

Consider $x(\lambda) = \lambda \overline{x} + (1-\lambda)x^*$.

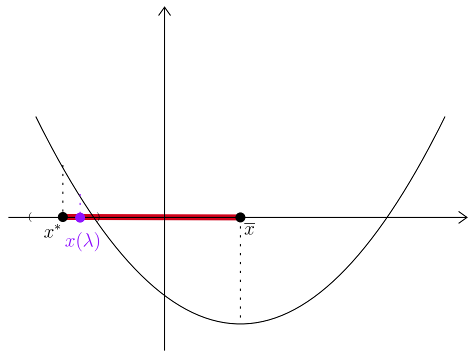
Since (NLP) is a convex program, $S$ is a convex set, therefore $x(\lambda) \in S, \forall \lambda \in [0,1]$. Since $f$ is a convex function, we have

$$f(x(\lambda)) = f(\lambda\overline{x} + (1-\lambda)x^*) \leq \lambda f(\overline{x}) + (1-\lambda)f(x^*)$$

Also, for any $\lambda > 0$, we have $\lambda f(\overline{x}) < \lambda f(x^*)$. Therefore,

$$f(x(\lambda)) < \lambda f(x^*) + (1-\lambda)f(x^*) = f(x^*), \ \forall \lambda \in (0,1]$$

Therefore, $\forall R > 0, \exists \lambda$ such that $x(\lambda) \in B(x^*, R) \cap S$. Contradicts local optimality of $x^*$.




$\square$

**Note**
This does not require differentiability.

### 4.3.1 Characterizing Optimality

The previous proposition suggests that only local information is needed for determining optimality.

*Can we characterize optimality based on local info?*

**Proposition 4.5**

Consider a convex optimization problem where $f$ is differentiable. Let $S$ be the feasible set. The $x^*$ is global optimal iff

$$\nabla f(x^*)^T(x - x^*) \geq 0, \quad \forall x \in S.$$

Proof:

($\Leftarrow$) From convexity of $f$

$$f(x) \geq f(x^*) + \underbrace{\nabla f(x^*)^T(x - x^*)}_{\geq 0} \geq f(x^*), \qquad \forall x \in S$$

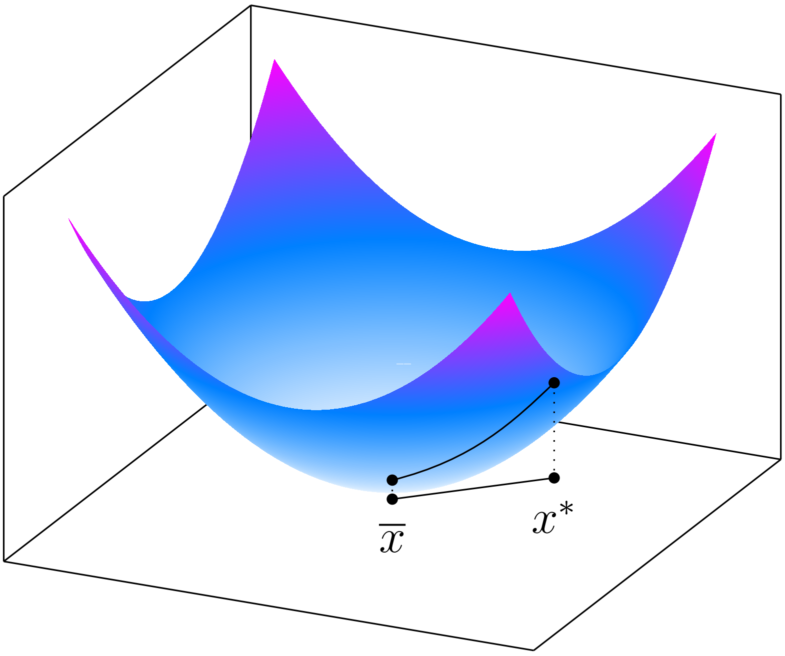
($\Rightarrow$) Sketch idea:

Suppose $\exists \overline{x} \in S : \nabla f(x^*)^T < 0$

Define $g(\lambda) := f(\lambda \overline{x} + (1 - \lambda)x^*)$

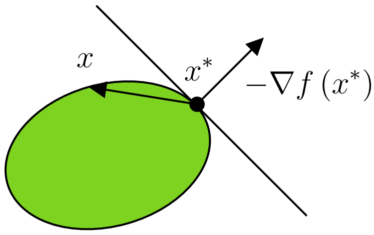
Can be argued that $g'(0) = \nabla f(x^*)^T(\overline{x} - x^*) < 0$.

For small $\lambda$, $g(\lambda) < g(0) = f(x^*)$. Therefore, $x^*$ is not optimal.




□

**Intuition** Going from $x^*$ in the direction towards another $x$ feasible takes us in the opposite direction that we want to go (opposite to the gradient).




**Corollary 4.6**

If $f : \mathbb{R}^n \to \mathbb{R}$ is convex, differentiable then $x^*$ is optimal to

$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & x \in \mathbb{R}^n \end{array}$$

iff $\nabla f(x^*) = 0$.

Proof:

($\Leftarrow$) Follows from previous proposition.

($\Rightarrow$) Suppose $\nabla f(x^*) \neq 0$. Let $y = -\nabla f(x^*) + x^*$.

$$\nabla f(x^*)^T(y - x^*) = -\nabla f(x^*)^T \nabla f(x^*) = -\|\nabla f(x^*)\|^2 \leq 0$$

$\implies$ $x^*$ is not optimal from previous proposition.

□

## 4.4 Lagrangian Duality

Consider a general NLP

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & g_i(x) \le 0, \quad \forall i = 1, \dots, m \end{aligned} \tag{NLP}$$

(that is NOT necessarily convex)

**Lagrangian**

The Lagrangian of (NLP) is the following function $L : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$,

$$L(x, \lambda) := f(x) + \sum_{i=1}^{m} \lambda_i g_i(x)$$

$\lambda_i$ are called **Lagrangian multipliers** associated to $g_i$ constraints.

Intuitively, we associate a penalty term $\lambda_i$ that would steer us away from points with $g_i \gg 0$, if we try to minimize $L(x, \lambda)$. We can restate the previous result as a generalization of LP weak duality.

**Proposition 4.7**

If $\overline{x} \in S$ and $\lambda \ge 0$, then $L(\overline{x}, \lambda) \le f(\overline{x})$.

Proof:

$$L(\overline{x}, \lambda) = f(\overline{x}) + \overbrace{\sum_{i=1}^{m} \underbrace{\lambda_i}_{\ge 0} \underbrace{g_i(\overline{x})}_{\le 0}}^{\le 0} \le f(\overline{x})$$

□

Now let $\ell(\lambda) = \min_{x \in \mathbb{R}^n} L(x, \lambda)$.

It follows that, $\forall \lambda \ge 0$, $\ell(\lambda) \le z^*$ where $x^*$ is optimal value of (NLP).

Thus we get a lower bound for any $\lambda \ge 0$.

As in LP duality, we are interested in the best possible lower bound.

So we want

$$\begin{aligned} \max \quad & \ell(\lambda) \\ \text{s.t.} \quad & \lambda \ge 0 \end{aligned} \tag{LD}$$

This is called the **Lagrangian dual** problem.

**Proposition 4.8: Weak duality**

If $\overline{x} \in S$ and $\lambda \geq 0$, then $\ell(\lambda) \leq f(\overline{x})$.

Example:

$$\begin{array}{ll} \min & c^T x \\ \text{s.t.} & Ax \leq b \iff Ax - b \leq 0 \end{array}$$

Then $f(x) = c^T x, g_i(x) = a_i^T x - b_i, \quad \forall i = 1, \ldots m$

$$\begin{aligned} L(x,\lambda) &= f(x) + \sum_{i=1}^{m} \lambda_i g_i(x) \\ &= c^T x + \sum_{i=1}^{m} \lambda_i (a_i^T x - b_i) \\ &= \left( c^T + \sum_{i=1}^{m} \lambda_i a_i^T \right) x - \sum_{i=1}^{m} \lambda_i b_i \end{aligned}$$

Then

$$\begin{aligned} \ell(\lambda) &= \min_{x \in \mathbb{R}^n} L(x,\lambda) \\ &= \begin{array}{ll} \min & (c^T + \sum_{i=1}^m \lambda_i a_i^T)x - \sum_{i=1}^m \lambda_i b_i \\ \text{s.t.} & x \in \mathbb{R}^n \end{array} \\ &= \begin{cases} -\infty, & \text{if } \left(c^T + \sum_{i=1}^m \lambda_i a_i^T\right) \neq 0 \\ -\sum_{i=1}^m \lambda_i b_i, & \text{if } \left(c^T + \sum_{i=1}^m \lambda_i a_i^T\right) = 0 \end{cases} \end{aligned}$$

Then

$$\begin{array}{ll} \max & \ell(\lambda) \\ \downarrow & \\ \text{s.t.} & \lambda \geq 0 \end{array} = \begin{array}{ll} \max & -\sum_{i=1}^m \lambda_i b_i \\ \downarrow & \\ \text{s.t.} & c^T + \sum_{i=1}^m \lambda_i a_i^T = 0 \\ & \lambda \geq 0 \end{array} \overset{y=-\lambda}{=} \begin{array}{ll} \max & b^T y \\ \downarrow & \\ \text{s.t.} & y^T A = c^T \\ & y \leq 0 \end{array}$$

Example:

$$\begin{array}{ll} \min & (x_1 - 1)^2 + (x_2 - 1)^2 \\ \downarrow & \\ \text{s.t.} & x_1 + 2x_2 - 1 \leq 0 \\ & 2x_1 + x_2 - 1 \leq 0 \end{array}$$

$$L(x,\lambda) = (x_1 - 1)^2 + (x_2 - 1)^2 + \lambda_1(x_1 + 2x_2 - 1) + \lambda_2(2x_1 + x_2 - 1)$$

Check: $L(x, \lambda)$ is a convex function (for a fixed $\lambda$ it is a convex function of $x$)

Now for $\ell(\lambda) = \min_{x \in \mathbb{R}^n} L(x, \lambda)$ is achieved when $\nabla_x L(x, \lambda) = 0$

$$\begin{pmatrix} 2(x_1 - 1) + \lambda_1 + 2\lambda_2 \\ 2(x_2 - 1) + 2\lambda_1 + \lambda_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \implies \begin{aligned} x_1^* &= \frac{-\lambda_1 - 2\lambda_2}{2} + 1 \\ x_2^* &= \frac{-2\lambda_1 - \lambda_2}{2} + 1 \end{aligned}$$

$$
\begin{aligned}
L(x^*, \lambda) &= \left(\frac{-\lambda_1 - 2\lambda_2}{2}\right)^2 + \left(\frac{-2\lambda_1 - \lambda_2}{2}\right)^2 + \lambda_1\left(\frac{-\lambda_1 - 2\lambda_2}{2} + 1 - 2\lambda_1 - \lambda_2 + 2 - 1\right) \\
&\qquad + \lambda_2\left(-\lambda_1 - 2\lambda_2 + 2 + \frac{(-2\lambda_1 - \lambda_2)}{2} + 1 - 1\right) \\
&= -1.25\lambda_1^2 - 1.25\lambda_2^2 - 2\lambda_1\lambda_2 + 2\lambda_1 + 2\lambda_2 \\
&=: \ell(\lambda)
\end{aligned}
$$

$$
\begin{array}{ll} \max & \ell(\lambda) \\ \text{s.t.} & \lambda \geq 0 \end{array} \quad = \quad \begin{array}{ll} \max & L(x^*, \lambda) \\ \text{s.t.} & \lambda \geq 0 \end{array}
$$

If we set $\nabla_\lambda L(x^*, \lambda) = 0$, we get $\lambda^* = \left(\frac{4}{9}, \frac{4}{9}\right)$ with objective value

$$
\ell(\lambda^*) = -2.5 \times \left(\frac{4}{9}\right)^2 - 2\left(\frac{4}{9}\right)^2 + 4 \times \frac{4}{9} = \frac{8}{9}
$$

And note that $x^* = \left(\frac{1}{3}, \frac{1}{3}\right)$ gives $f(x^*) = \frac{8}{9}$, which gives optimal solution.

## 4.5 Karush-Kuhn-Tucker Optimality Conditions

### Lagrangean dual for problems with equality constraints

For problems of the form,

$$
\begin{array}{ll}
\min & f(x) \\
\downarrow & \\
\text{s.t.} & \begin{array}{ll} g_i(x) \leq 0, & \forall i = 1, \dots, m \\ h_i(x) = 0, & \forall i = 1, \dots, p \end{array}
\end{array}
\tag{NLP}
$$

We can define

$$
L(x, \lambda, \nu) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x) + \sum_{i=1}^{p} \nu_i h_i(x)
$$

Here the Lagrangean dual:

$$
\begin{array}{ll} \max & \ell(\lambda, \nu) \\ \text{s.t.} & \lambda \geq 0, \nu \in \mathbb{R}^p \end{array}
$$

where $\ell(\lambda, \nu) = \min_{x \in \mathbb{R}^n} L(x, \lambda, \nu)$. Weak duality still holds for $\lambda \geq 0, \nu \in \mathbb{R}^p$.

**Note**
If $f, g_i$ are convex, $\forall i = 1, \dots, m$ and $h_i(x)$ are affine functions, then (NLP) is a convex program.

**Note**
Weak Duality holds regardless if $g_i, h_i$ are convex.

**Example: Least square solutions of linear equations**

Suppose we want to find, out of all possible solutions to $Ax = b$, the one with smallest norm.

$$\begin{array}{ll} \min & x^T x \\ \text{s.t.} & Ax = b \end{array}$$

Lagrangian: $L(x, \nu) = x^T x + \nu^T (Ax - b)$.

Then $\ell(\nu) = \min_{x \in \mathbb{R}^n} L(x, \nu)$.

$$\nabla_x L(x, \nu) = 0 \implies 2x + A^T \nu = 0 \implies x = -\frac{A^T \nu}{2}$$

$$\begin{aligned} \implies \ell(\nu) &= \frac{\nu^T A A^T \nu}{4} - \frac{\nu^T A A^T \nu}{2} - b^T \nu \\ &= -\frac{\nu^T A A^T \nu}{4} - b^T \nu \\ &\leq \begin{array}{ll} \min & x^T x \\ \text{s.t.} & Ax = b \end{array} \end{aligned}$$

*When does Strong Duality Hold?*

This is hard to characterize in general, but there are some easily checkable sufficient conditions.

Let

$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & g_i(x) \leq 0, \quad \forall i = 1, \ldots, m \end{array} \tag{CVX}$$

where $f, g_i$ are convex $\forall i = 1, \ldots, m$.

**Slater's Condition**

$\exists \overline{x} : g_i(\overline{x}) < 0, \quad \forall i = 1, \ldots, m.$

That is, there exists a point in the relative interior of the feasible region.

**Theorem 4.9**

If Slater's condition holds for (CVX), then $\exists \lambda^* \geq 0$ such that

$$\ell(\lambda^*) = \min_{x \in \mathbb{R}^n} L(x, \lambda^*) = \boxed{\begin{array}{ll} \min & f(x) \\ \text{s.t.} & g_i(x) \leq 0, \quad \forall i = 1, \ldots, m \end{array}}$$

Recall that this was abuse of notation and it is not clear that $\exists x^*$ achieving inf.

i.e.,

$$\max_{\lambda \geq 0} \ell(\lambda) = \begin{array}{ll} \min & f(x) \\ \text{s.t.} & g_i(x) \leq 0, \quad \forall i = 1, \ldots, m \end{array}$$

and the max is attained at $\lambda^*$.

Proof:

SKIPPED. $\square$

To derive optimality conditions, suppose we have $\lambda^*, x^*$ opti. for dual/primal.

$$\ell(\lambda^*) = \min_{x \in \mathbb{R}^n} f(x) + \sum_{i=1}^m \lambda_i^* g_i(x) \leq f(x^*) + \sum_{i=1}^m \lambda_i^* g_i(x^*) \underset{\substack{\uparrow \\ \lambda^* \geq 0,\, g_i(x^*) \leq 0}}{\leq} f(x^*)$$

Now if we want strong duality to hold, i.e., we want $\ell(\lambda^*) = f(x^*)$ then all above inequalities must hold at equality.

The first inequality holding as equality implies $x^*$ is a minimizer of $L(x, \lambda^*)$ for all $x \in \mathbb{R}^n$.

$$L(x, \lambda) = f(x) + \sum_{i=1}^m \lambda_i g_i(x) \implies \nabla_x L(x^*, \lambda^*) = 0 \implies \nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla g_i(x^*) = 0$$

The second inequality holding as equality means a complementary slackness-type condition, i.e., $\lambda_i^* g_i(x^*) = 0 \iff \lambda_i^* = 0 \quad \text{or} \quad g_i(x^*) = 0$.

Formally, these are the so-called **Karush-Kuhn-Tucker (KKT)** optimality conditions:

**KKT conditions**

i) $g_i(x^*) \leq 0, \ \forall i = 1, \ldots, m$

ii) $\lambda^* \geq 0$

iii) $\lambda_i^* g_i(x^*) = 0, \ \forall i = 1, \ldots, m$

iv) $\nabla f(x^*) + \sum_{i=1}^m \lambda_i^* g_i(x^*) = 0$

**Theorem 4.10: Necessary opt. conditions**

Consider

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & g_i(x) \leq 0, \quad \forall i = 1, \ldots, m \end{aligned} \tag{NLP}$$

where $f, g_i$ are differentiable, $\forall i = 1, \ldots, m$.

If $x^*, \lambda^*$ are optimal to the (NLP) and its Lagrangean dual, respectively, such that $f(x^*) = L(x^*, \lambda^*) = \ell(\lambda^*)$, then KKT conditions hold.

Proof:
Follows from above discussion. □

**Theorem 4.11: Sufficient opt. conditions**

Assume that, in addition, the functions $g_i$ are convex, $\forall i = 1, \ldots, m$, $f$ is convex. Then if $x^*, \lambda^*$ satisfy KKT conditions, $x^*, \lambda^*$ are optimal for (NLP) and its Lagrangean dual, and $f(x^*) = \ell(\lambda^*) = L(x^*, \lambda^*)$.

Proof:
Follows similar to necessity proof, using the fact that $L(x, \lambda)$ is a convex function and thus $\nabla_x L(x^*, \lambda^*) = 0 \implies x^*$ is a minimizer of $L(x, \lambda^*)$ over $x \in \mathbb{R}^n$. $\square$

**Note**
For problems of the form:

$$\begin{array}{ll} \min & f(x) \\ \downarrow & \\ \text{s.t.} & \begin{array}{l} g_i(x) \leq, \quad \forall i = 1, \ldots, m \\ h_i(x) = 0, \quad \forall i = 1, \ldots, p \end{array} \end{array} \tag{NLP-EQ}$$

the KKT conditions are:

**KKT**

i) $g_i(x^*) \leq 0, \; \forall i = 1, \ldots, m$

ii) $h_i(x^*) = 0, \; \forall i = 1, \ldots, p$

iii) $\lambda^* \geq 0$

iv) $\lambda_i^* g_i(x^*) = 0, \; \forall i = 1, \ldots, m$

v) $\nabla f(x^*) + \sum_{i=1}^m \lambda_i^* g_i(x^*) + \sum_{i=1}^p \nu_i \nabla h_i(x^*) = 0$

With equality constraint:

- If $x^*$ opt for (NLP-EQ), $(\lambda^*, \nu^*)$ opt for its lag. dual and $f(x^*) = \ell(\lambda^*, \nu^*)$ then KKT holds.

- If $f, g_1, \ldots, g_m$ are convex and $h_1, \ldots, h_p$ are affine functions, then $x^*, \lambda^*, \nu^*$ satisfying KKT $\implies x^*$ opt for (NLP-EQ), $\lambda^*, \nu^*$ opt for its Lag. dual and $f(x^*) = \ell(\lambda^*, \nu^*)$.

*Where is Slater's condition needed in convex programs?*

Example:

$$\begin{array}{ll} \min & x \\ \text{s.t.} & x^2 \leq 0 \end{array}$$

is a convex program with unique feasible solution $x = 0 \implies$ Slater's condition does not hold.

Now $x = 0$ is optimal. But $\nabla f(x^*) + \sum_{i=1}^m \lambda_i \nabla g_i(x^*) = 1 + 0 = 1 \neq 0$.

**Note**
$L(x, \lambda) = x + \lambda x^2$ and

$$\ell(\lambda) = \min_{x \in \mathbb{R}} x + \lambda x^2 = \begin{cases} -\infty, & \text{if } \lambda = 0 \\ -\frac{1}{2\lambda}, & \text{if } \lambda > 0 \end{cases}$$

This problem violates Slater's condition and $\nexists x^*, \lambda^*$ achieving strong duality.