# CO 250 Spring 2022:
## Lecture Notes

Lecture notes taken, unless otherwise specified, by myself during section 001 of the Spring 2022 offering of CO 250, taught by Martin Pei.

Notes are adapted into this format from hard copies/Obsidian.

## Lectures

# Chapter 1

# Intro

## 1.1 Linear programs

> **Example 1.1.1.** Suppose we are selling apples and bananas at a stand. Apples sell for \$2 per kilogram, and bananas sell for \$1.5 per kilogram. Our stand holds up to 75 kilograms of fruits. Also, there are only 4 square metres of shelf space. Each kilogram of apples/bananas takes up roughly 0.08/0.05 square metres of shelf space, respectively. How much of each fruit should we stock to maximize the total sales?

*Solution.* Let $x_1$, $x_2$ be weight of apples, bananas (kg). Define objective function max $2x_1 + 1.5x_2$. Add constraints $x_1 + x_2 \leq 75$ for weight, $0.08x_1 + 0.05x_2 \leq 4$ for shelf space, and $x_1, x_2 \geq 0$ for common sense.

Summarize as a <u>linear program</u>:

$$
\begin{aligned}
\max \quad & 2x_1 + 1.5x_2 \\
\text{subject to (s.t.)} \quad & x_1 + x_2 \leq 75 \\
& 0.08x_1 + 0.05x_2 \leq 4 \\
& x_1, x_2 \geq 0
\end{aligned}
$$

Trial and error:

- $(x_1, x_2) = (30, 20)$ satisfies constraints (<u>feasible</u>) with <u>objective value</u> 90
- $(x_1, x_2) = (31, 20)$ feasible with objective value 92
- $(x_1, x_2) = (50, 0)$ feasible with objective value 100
- $(x_1, x_2) = (8\frac{1}{3}, 66\frac{2}{3})$ feasible with objective value $116\frac{2}{3}$ (claim without proof that this is <u>optimal</u>)

N.B.: we take domain to be $\mathbb{R}$ since we can take fractional parts of a kilogram of fruit

Plot feasible solutions:
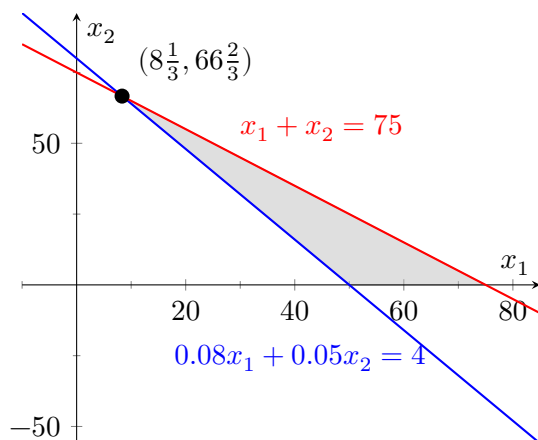
Bound by convex region defined by axes, $x_1 + x_2 = 23$, and $0.08x_1 + 0.05x_2 = 4$ to give optimal solution at vertex $\qquad\square$

**Course overview**

- Formulation/modelling: create mathematical programs from problems
- Solving linear programs: use simplex method to optimize
- Geometric interpretation: conceptualize linear programs and simplex method
- Integer programs: linear programs defined over $\mathbb{Z}$
- Nonlinear programs: convex functions

---

**Definition 1.1.2** (optimization problem)

Given a set of <u>feasible points</u> $A \subseteq \mathbb{R}^n$ and $f : A \to \mathbb{R}$, find some $x \in A$ that minimizes or maximizes the <u>objective value</u> $f(x)$.

Composed of <u>decision variables</u> $\mathbf{x} \in \mathbb{R}^n$, the <u>objective function</u> $\max f(\mathbf{x})$ or $\min f(\mathbf{x})$, and some <u>constraints</u> of the form $g_i(\mathbf{x}) \le b_i$

---

**Definition 1.1.3** (affine function)

Function of the form $f(\mathbf{x}) = \mathbf{a}^T\mathbf{x} + b = a_1x_1 + \cdots + a_nx_n + b$ for constants $\mathbf{a}$ and $b$

---

**Definition 1.1.4** (linear function)

Affine function with $b = 0$

---

**Definition 1.1.5** (linear program)

An optimization problem with affine objective function $f(\mathbf{x})$ and finitely many linear constraint functions $g_i(x) \ge b_i$ (or $\le b_i$ or $= b_i$) with constant $\mathbf{b}$.

N.B.: constraints cannot be strict inequalities

---

## 1.2 LP Formulation

**Example 1.2.1.** A company makes 4 types of products, each requiring time on two different machines and two types of labour. The amount of machine time and labour needed to produce one unit of each product along with its sale price are summarized in the following table.

| Product | Machine 1 | Machine 2 | Skilled labour | Unskilled labour | Unit sale price |
|---------|-----------|-----------|----------------|------------------|-----------------|
| 1 | 11 | 4 | 8 | 7 | 300 |
| 2 | 7 | 6 | 5 | 8 | 260 |
| 3 | 6 | 5 | 5 | 7 | 220 |
| 4 | 5 | 4 | 6 | 4 | 180 |

Each month, the company can use up to 700 hours on machine 1, and 500 hours on machine 2, with no cost. The company can hire up to 600 hours of skilled labour at $8 per hour, and up to 650 hours of unskilled labour at $6 per hour. How should the company operate to maximize their monthly profit?

*Solution.* Let $\mathbf{x} \in \mathbb{R}^4$ be number of units of products, $y_s$ and $y_u$ be hours of labour hired

Let the objective function be max $300x_1 + 260x_2 + 220x_3 + 180x_4 - 8y_s - 6y_u$ (unit sale revenue net of labour costs)

Let the constraints be $11x_1 + 7x_2 + 6x_3 + 5x_4 \leq 700$ (machine 1), $4x_1 + 6x_2 + 5x_3 + 4x_4 \leq 500$ (machine 2), $8x_1 + 5x_2 + 5x_3 + 6x_4 = y_s$, $7x_1 + 8x_2 + 7x_3 + 4x_4 = y_u$ (defining $y_s$ and $y_u$), $y_s \leq 600$, $y_u \leq 650$ (labour), and $x_1, x_2, x_3, x_4, y_s, y_u \geq 0$ (non-negativity) □

**Example 1.2.2.** A certain company provides heading oil for the local commnity. They have historical data that helps them predict demand for heating oil in the next four months: 5000, 8000, 9000, 6000 (litres/month)

At the beginning of each month, they can purchase oil from the supplier at the current market rate. The projected rates are given: 0.75, 0.72, 0.92, 0.90 ($/litre)

There is a storage tank that holds up to 4000 litres of oil, and at the start of month 1, it contains 2000 litres. How should the company buy the required oil to minimize the total money spent?

*Solution.* Let $x_i$ be the amount of oil purchased in the $i$th month, and $y_i$ be the amount of oil in the storage tank at the start of month $i$.

Then, we want to minimize $0.75x_1 + 0.72x_2 + 0.92x_3 + 0.9x_4$.

The storage tank constrains us by $y_i \leq 4000$ and the problem gives $y_1 = 2000$.

Non-negativity gives $x_i, y_i \geq 0$.

For each demand $d_i$, we have $x_i + y_i = d_i + y_{i+1}$.

Then, we can write:

$$\begin{aligned}
\min \quad & 0.75x_1 + 0.72x_2 + 0.92x_3 + 0.9x_4 \\
\text{s.t.} \quad & y_1, y_2, y_3, y_4 \leq 4000 \\
& y_1 = 2000 \\
& x_1 + y_1 = 5000 + y_2 \\
& x_2 + y_2 = 8000 + y_3 \\
& x_3 + y_3 = 9000 + y_4 \\
& x_4 + y_4 = 6000 \\
& x_1, x_2, x_3, x_4, y_1, y_2, y_3, y_4 \geq 0
\end{aligned}$$

$\square$

**Example 1.2.3.** Instead of minimizing the total money spent, suppose we do not have much money to spend each month, and we want to reduce the maximum amount spent in a month.

*Solution.* Let $M = \max\{0.75x_1, 0.72x_2, 0.92x_3, 0.9x_4\}$.

Since $M$ is not linear, we cannot simply put $\min M$ in an LP.

Instead, define $m$ with constraints $m \geq 0.75x_1$, $m \geq 0.72x_2$, $m \geq 0.92x_3$, $m \geq 0.9x_4$.

Since we are doing $\min m$, we are guaranteed that the optimal solution will give $m = M$ (if $m$ is not $M$, we can make $m$ smaller). $\square$

**Example 1.2.4.** Given a set of data points $\{(x_i, y_i) : i = 1, \ldots, n\}$ on the plane. Find a line $y = ax + b$ that "best fits" this set of data points.

*Solution.* Define "best fit" as minimizing total vertical distance between points and the line.

That is, we must minimize $\sum |ax_i + b - y_i|$, but that is not affine.

Define instead the errors $e_i$ associated with the point $i$.

We want to constrain $e_i = |ax_i + b - y_i|$, which we can do with $e_i \geq ax_i + b - y_i$ and $e_i \geq y_i - ax_i - b$ since $|x| = \max\{x, -x\}$.

Then, we can use $\min \sum e_i$ as above to get the final LP:

$$\begin{aligned}
\min \quad & \sum e_i \\
\text{s.t.} \quad & e_i \geq ax_i + b - y_i \\
& e_i \geq y_i - ax_i - b
\end{aligned}$$

**N.B.:** since $e_i$, $e_j$ do not share constraints when $i \neq j$, $\min \sum e_i$ is equivalent to $\min e_1, \ldots, \min e_n$. $\square$

**Exercise 1.2.5.** Modify this to find the best fit parabola. Is this an LP?

*Solution.* Yes, since considering the error function $ax_i^2 + bx_i + c - y_i$ is still linear with respect to the variables for optimization $a$, $b$, and $c$. $\square$

## 1.3  Formulating IPs

> **Example 1.3.1.** Consider the job application process where a company has 3 positions available, and there are 4 applicants for these jobs. For each applicant and position, the company assigns a number indicating how well the applicant is suited for the position. The goal is to hire a different applicant for each position to maximize the total suitability.
>
> $$M = \text{Positions} \begin{pmatrix} 3 & 5 & 2 & 4 \\ 3 & 1 & 4 & 3 \\ 1 & 4 & 2 & 3 \end{pmatrix} \overset{\text{Candidates}}{}$$

*Lecture 3*
*May 9*

*Solution.* Want: For each position, who gets that position

Define: Create binary variable $x_{ij}$ for each position $i$ and candidates $j$. Let $x_{ij} = 1$ if position $i$ given to candidate $j$, and 0 otherwise

Objective function: $\max \sum \sum M_{ij} x_{ij}$

Constraints: $\sum_j x_{ij} = 1$ for each $i$ (each position filled by exactly one candidate), and $\sum_i x_{ij} \leq 1$ for each $j$ (each candidate takes at most one position), $x_{ij} \geq 0$, $x_{ij} \leq 1$, $x_{ij} \in \mathbb{Z}$ (integrality)

$$
\begin{aligned}
\max \quad & \sum_{i=0}^{3} \sum_{j=0}^{4} M_{ij} x_{ij} \\
\text{s.t.} \quad & \sum_{j=0}^{4} x_{ij} = 1 && i = 1, \dots, 3 \\
& \sum_{i=0}^{3} x_{ij} \leq 1 && j = 1, \dots, 4 \\
& 0 \leq x_{ij} \leq 1, x_{ij} \in \mathbb{Z}
\end{aligned}
$$

$\square$

> **Notation.** We define $x \in \{0, 1\}$ to mean the constraints $0 \leq x \leq 1$ and $x \in \mathbb{Z}$

> **Example 1.3.2** (Knapsack problem)**.** There are 4 types of items that you can put into your backpack. You can take any integer number of units of any item. However, you can only carry a maximum of 40 pounds. Each unit of item you take is also worth a certain amount of money. The goal is to maximize the total value of the items you carry.
>
> | Item | A | B | C | D |
> |---|---|---|---|---|
> | Weight (lbs) | 1 | 7 | 3 | 2 |
> | Value ($) | 10 | 50 | 20 | 15 |

*Solution.* Let $x_i$, $i = A, B, C, D$ be the number of units of $i$ packed

Objective function: $\max 10 x_A + 50 x_B + 20 x_C + 15 x_D$

Constraints: $x_A + 7x_B + 3x_C + 2x_D \le 40$ (weight limit), $x_i \ge 0$, $x_i \in \mathbb{Z}$ (integrality)

$$\begin{aligned}
\max \quad & 10x_A + 50x_B + 20x_C + 15x_D \\
\text{s.t.} \quad & x_A + 7x_B + 3x_C + 2x_D \le 40 \\
& x_A, x_B, x_C, x_D \ge 0 \\
& x_A, x_B, x_C, x_D \in \mathbb{Z}
\end{aligned}$$

$\square$

**Example 1.3.3.** Suppose we are allowed to take A only if we take at least one unit of B.

*Solution.* Want: if $x_B = 0$, then we must have $x_A = 0$. If $x_B \ge 1$, no restriction on $A$.

Equivalently, add the constraint $x_A \le x_B \max x_A = 40x_B$. When $x_B = 0$, the RHS goes to 0 and constrains $x_A = 0$. Otherwise, since $x_B \ge 1$, $40x_B \ge 40$ which is the maximum value of $x_A$, so there are effectively no constraints on $x_A$. $\square$

**Example 1.3.4.** Suppose we want the following conditions to hold:

1. We carry at least 5 units of items A and/or B; or
2. We carry at least 7 units of items C and/or D.

*Solution.* Define a binary variable $y$. Want: $y = \begin{cases} 1 & \text{condition 1 is true} \\ 0 & \text{condition 2 is true} \end{cases}$

If $y = 1$, then $x_A + x_B \ge 5$; if $y = 0$, no restrictions on $x_A$, $x_B$. We can implement this by adding the constraint $x_A + x_B \ge 5y$, since $y = 0$ will send the RHS to 0

If $y = 0$, then $x_C + x_D \ge 7$; if $y = 1$, no restrictions on $x_C$, $x_D$. Similarly implement with $x_C + x_D \ge 7(1 - y)$, since $y = 1$ will send the RHS to 0

Notice that setting $y$ does not force the other condition *not* to hold, i.e., this implements an inclusive or.

In summary: $x_A + x_B \ge 5y$, $x_C + x_D \ge 7(1 - y)$, and $y \in \{0, 1\}$

**N.B.:** When feeding these constraints into an algorithm, ensure that the constraints are truly linear, i.e., move variables to one side. For example, $x_A + x_B - 5y \ge 0$ $\square$

**Exercise 1.3.5.** Implement an exclusive or of these two conditions

**Example 1.3.6.** Suppose that the value of item A is \$10 for the first 5 units, but any more units beyond that has value \$5.

*Solution.* Separate $x_A$ into two variables $x_{A1}$ for first five units and $x_{A2}$ for remainder. Then, we have $x_A = x_{A1} + x_{A2}$ and change the objective function to $10x_{A1} + 5x_{A2} + 50x_B + 20x_C + 15x_C$.

We can create a constraint to force $x_{A2}$ only to go up when $x_{A1}$ is 5 with $x_{A2} \le (x_{A1} - 4) \max x_{A2}$ which will work in tandem with the non-negativity constraint. This is not actually necessary since

the maximum will always fill $x_{A1}$ before $x_{A2}$ because it is worth more (i.e. trading one $x_{A2}$ for $x_{A1}$ will increase the objective function by 5)

In summary: change the objective function and add the constraints $x_A = x_{A1} + x_{A2}$, $x_{A1} \leq 5$, $x_{A1}, x_{A2} \geq 0$, $x_{A1}, x_{A2} \in \mathbb{Z}$ □

*Lecture 4*
*May 11*

> **Notation** (vector notation). Write $\mathbb{1} := (1, \dots, 1)^T$ and $x \leq y$ if $x_i \leq y_i$ for all $i$.
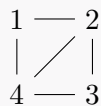
> **Definition 1.3.7** (graph)
>
> $G = (V, E)$ consists of a set of objects $V$ (<u>vertices</u>) and a set of unordered pairs of vertices $E$ (<u>edges</u>).
>
> We restrict graphs by disallowing empty graphs, redundant edges, directed edges, or self-connections.

> **Example 1.3.8.** $G = (V, E)$ by $V = \{1, 2, 3, 4\}$, $E = \{12, 23, 34, 41, 24\}$
>
> $$
> \begin{array}{ccc}
> 1 & \!\!-\!\!\!-\!\! & 2 \\
> | & \diagup & | \\
> 4 & \!\!-\!\!\!-\!\! & 3
> \end{array}
> $$

> **Definition 1.3.9** (incidence relation)
>
> For an edge $e = uv$, $e$ is incident to $u$ and $v$. $\delta(v)$ is the set of all edges incident to $v$.
>
> The <u>incidence matrix</u> $B \in \{0, 1\}^{|V| \times |E|}$ has rows indexed by $V$, columns by $E$, and $B_{ve} = 1$ when $e \in \delta(v)$ and 0 otherwise.

> **Example 1.3.10.** For Example 1.3.8, $B = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{ccccc} e_1 & e_2 & e_3 & e_4 & e_5 \end{array} \\ \left( \begin{array}{ccccc} 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \end{array} \right) \end{array}.$

> **Remark 1.3.11.** Each column has exactly two ones, so $B\mathbb{1} = 2\mathbb{1}$.

> **Definition 1.3.12** (matching)
>
> $M \subseteq E$ where each vertex is incident with exactly zero or one edge in $M$ (i.e., $|M \cap \delta(v)| \leq 1$ for all $v \in V$)

**Example 1.3.13.** For Example 1.3.8, $\{e_1, e_3\}$ and $\{e_5\}$ are matchings but $\{e_1, e_5\}$ is not since 2 is incident to both edges

**Example 1.3.14** (maximum-weight matching)**.** Given graph $G = (V, E)$ and weights $w_e$ for each $e \in E$. Find a matching in $G$ with the maximum edge weight, i.e., maximize $\sum_{e \in M} w_e$.

*Solution.* Define a vector $x \in \{0, 1\}^{|E|}$ by $x_e = 1$ if $e \in M$ and 0 otherwise. Then, the objective function is $\max w^T x$. To ensure each node appears only once, add constraints $\sum_{e \in \delta(v)} x_e \leq 1$ for each $v \in V$. This is equivalent to taking the incidence matrix $A$ and saying $Ax \leq \mathbb{1}$ This gives us the integer program
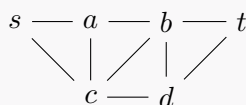
$$\max w^T x$$
$$\text{s.t. } Ax \leq \mathbb{1}$$
$$x_e \in \{0, 1\} \quad e \in E$$

$\square$

**Definition 1.3.15** ($v_1, v_k$-path)

Sequence of edges $v_1 v_2, v_2 v_3, \ldots, v_{k-1} v_k$ such that $v_1, \ldots, v_k$ are distinct

**Example 1.3.16.** Consider graph $(\{s, t, a, b, c, d\}, \{sa, sc, ab, ac, bd, bt, cb, cd, dt\})$.



Then, $sa, ab, bt$ and $sc, cb, bd, dt$ are $s,t$-paths but $sa, ab, bc, cd, cb, bt$ is not since $b$ is visited twice.

**Problem 1.3.17** (shortest path)

Given graph $G = (V, E)$, vertices $s$ and $t$, and positive weights $w_e$ for each $e \in E$. Find an $s,t$-path $P$ with the minimum edge weight, i.e., minimize $\sum_{e \in P} w_e$.

Define a vector $x \in \{0, 1\}^{|E|}$ by $x_e = 1$ if $e \in P$ and 0 otherwise.

The objective function is $\min w^T x$.
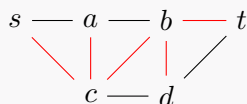
Need to constrain $x$ into a path: use cuts.

> **Definition 1.3.18** (cut)
>
> The <u>cut induced by vertices $W$</u> is the set $\delta(W)$ of all edges with exactly one endpoint in $W$. Formally, $\delta(W) = \{uv \in E : u \in W, v \notin W\}$.
>
> An <u>$s,t$-cut</u> $\delta(W)$ has $s \in W$ and $t \notin W$.

> **Example 1.3.19.** In Example 1.3.16, $W = \{s, a, b\}$ induces the cut $\delta(W) = \{sc, ac, bc, bd, bt\}$
>
> $$s \longrightarrow a \longrightarrow b \longrightarrow t$$
> $$c \longrightarrow d$$

> **Proposition 1.3.20**
>
> Notice that the edges in an $s,t$-cut separate $s$ from $t$, so an $s,t$-path must use at least one edge from every $s,t$-cut (formal proof in graph theory course)

We get a constraint $\sum_{e \in \delta(W)} x_e \geq 1$ for all $s,t$-cuts $\delta(W)$ (that is, for all $W \subset V$ with $s \in W$ and $t \notin W$)

> **Proposition 1.3.21**
>
> If a set of edges intersects every $s,t$-cut, then it contains an $s,t$-path

Minimizing the edge weights will ensure that the extraneous edges are optimized away and the $s,t$-path remains so long as the edge weights are all positive.

This gives us a final IP of

$$\min \ w^T x$$
$$\text{s.t.} \sum_{e \in \delta(W)} x_e \geq 1 \qquad \qquad \delta(W) \text{ an } s,t\text{-cut}$$
$$x_e \geq 0, x_e \in \mathbb{Z} \qquad \qquad e \in E$$

## 1.4   Formulating NLPs

> **Definition 1.4.1** (non-linear program)
>
> A program of the general form $\min f(x)$ subject to $g_i(x) \leq 0$ for some arbitrary functions $f : \mathbb{R}^n \to \mathbb{R}$, $g_i : \mathbb{R}^n \to \mathbb{R}$ with no restrictions

**Example 1.4.2.** Among all the points $x$ that satisfy $Ax \leq b$, find one that is closest to the target point $\bar{x}$

*Solution.* We can take the norm and minimize $\|x - \bar{x}\| = \sqrt{\sum(x_i - \bar{x}_i)^2}$. This gives us the non-linear program $\min \|x - \bar{x}\|$, s.t. $Ax \leq b$. $\qquad\square$

*Lecture 6*
*May 18*

Since the definition for NLP has no constraints on $f$ and $g_i$, a LP is an NLP.

The integrality constraint makes IPs not NLPs. To get around this, use a periodic function like $\sin \theta = 0$ which permits values $\theta = k\pi$ for integer $k$, so $x \in \mathbb{Z} \Leftrightarrow \sin x\pi = 0$. Using this makes IPs into NLPs.

If we can solve NLPs, we can also solve LPs and IPs.

## 1.5   LP outcomes

An algorithm that solves LPs should produce:

- The optimal solution (or that no solution exists)
- Certificate of correctness that reduces complexity of verification

> **Definition 1.5.1** (infeasibility)
>
> No feasible solutions exist.

**Example 1.5.2.** $\max x$ s.t. $x \leq 2$ and $x \geq 3$. Obviously, no $x$ exists.

**Example 1.5.3.** $\max (3, 1, -7, 4)x$ s.t. $\begin{pmatrix} -5 & 4 & 3 & -1 \\ 2 & 1 & -5 & 3 \\ -1 & -3 & 1 & -2 \end{pmatrix} x = \begin{pmatrix} 3 \\ -2 \\ 1 \end{pmatrix}$ and $x \geq \mathbb{0}$.

Taking $-2R_1 - 3R_2 - 4R_3$, we get $8x_1 + x_2 + 5x_3 + x_4 = -4$ but each entry in $x$ must be non-negative, so this is impossible.

Formally, we can let $y = (-2, -3, -4)^T$, then multiply on the left by $y^T$ to give us the same equation as $(8, 1, 5, 1)x = -4$.

Then, $y$ is the <u>certificate of infeasibility</u>.

> **Proposition 1.5.4**
>
> The system $Ax = b$, $x \geq \mathbb{0}$ is infeasible if there exists a vector $y$ such that $y^T A \geq \mathbb{0}$ but $y^T b < 0$

*Proof.* Suppose the system is feasible with $x$ as the feasible solution. Then, $Ax = b$ and $x \geq \mathbb{0}$. However, $y^T Ax \geq 0$ and $y^T b < 0$. Contradiction. $\qquad\square$

The converse is also true. Proof will come later as Farkas' Lemma.

**Definition 1.5.5** (unboundedness)

Infinitely better feasible solutions exist.

Formally, a max (resp. min) LP is <u>unbounded</u> if there exists a series of feasible solutions $x(t)$ with the objective value of $x(t)$ approaching $+\infty$ (resp. $-\infty$) as $t \to \infty$.

**Example 1.5.6.** $\max x$ s.t. $x \geq 1$: there is no best solution (cf. strict inequalities)

**Example 1.5.7.** $\max (-1, 2, -3, 4)x$ s.t. $\begin{pmatrix} 3 & 0 & 2 & -5 \\ -2 & 3 & -4 & 4 \end{pmatrix} x = \begin{pmatrix} 4 \\ 1 \end{pmatrix}$ and $x \geq \mathbb{0}$.

Consider $\bar{x} = (3, 1, 0, 1)^T$ and $d = (0, 4, 5, 2)^T$. Define $x(t) = \bar{x} + td$ and consider $t$ from $0 \to \infty$. We must show feasibility and unboundedness.

Obviously, $x(t) \geq \mathbb{0}$ since $\bar{x}, d \geq \mathbb{0}$ and $t \geq 0$.

Notice $Ax(t) = A\bar{x} + tAd = (4, 1)^T + t(0, 0)^T = b$. That is, $\bar{x}$ solves $Ax = b$ and $d$ lies in the kernel of $A$.

The objective value $c^T x(t) = c^T(\bar{x} + td) = c^T\bar{x} + tc^Td = 3 + t$ clearly goes to $+\infty$ as $t \to \infty$.

Then, $(\bar{x}, d)$ is a certificate of unboundedness for the LP.

**Proposition 1.5.8**

The LP $\max\{c^T x : Ax = b, x \geq \mathbb{0}\}$ is unbounded if there exist vectors $\bar{x}$ and $d$ such that $\bar{x}, d \geq \mathbb{0}$, $Ad = \mathbb{0}$, $c^Td > 0$.

# Chapter 2

# Solving LPs

## 2.1 Preparation

**Example 2.1.1.** $\max (0, -2, -3, 0)x + 7$ subject to $\begin{pmatrix} 1 & 3 & -5 & 0 \\ 0 & -1 & 2 & 1 \end{pmatrix} x = \begin{pmatrix} 6 \\ 9 \end{pmatrix}$ and $x \geq \mathbb{0}$

*Solution.* Trivial solution: $\bar{x} = (6, 0, 0, 9)^\top$ gives objective value 7

Claim: $\bar{x}$ is optimal

Proof: the term $(0, -2, -3, 0)x \leq 0$ since $x \geq \mathbb{0}$. Its highest value is then 0, so the highest objective value is 7. Since the objective value of $\bar{x}$ is 7, it is optimal. $\qquad \square$

> **Theorem 2.1.2** (Fundamental Theorem of Linear Programming)
>
> For a linear program $P$, exactly one of the following holds:
>
> - $P$ is infeasible
> - $P$ is unbounded
> - $P$ has an optimal solution

This does not apply to non-linear programs: e.g., $\max x$ subject to $x < 1$. This NLP is feasible (consider $\bar{x} = 0$) and bounded ($x < 1$), but has no optimal solution (given $\bar{x}$ a solution, $\frac{\bar{x}+1}{2}$ is a better solution)

> **Definition 2.1.3** (standard equality form)
>
> A linear program of the form $\max\{c^\top x + \bar{z} : Ax = b, x \geq \mathbb{0}\}$
>
> Requires maximization, equality constraint, and non-negative variables

Simplex requires SEF, so must convert LPs into equivalent SEF LP.

> **Definition 2.1.4** (equivalence)
>
> $P$ and $P'$ are equivalent if (1) $P$ infeasible iff $P'$ infeasible, (2) $P$ unbounded iff $P'$ unbounded, and (3) optimal solutions of $P$ can be constructed from $P'$ and vice versa

Find an equivalent SEF by:

- Given a minimization LP min $f(x)$, just take max $-f(x)$
- Given an inequality $x \leq k$, define a slack variable $x + x' = k$ with $x' \geq 0$
- Given a free variable $x$, define two non-negative variables $x^+$ and $x^-$ so that we can replace $x = x^+ - x^-$

**Example 2.1.5.** Find an equivalent LP in SEF

$$\min (-1, 2, -3)x$$
$$\text{s.t. } \begin{pmatrix} 1 & 5 & 3 \\ 2 & -1 & 2 \\ 1 & 2 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \begin{matrix} \leq \\ \geq \\ = \end{matrix} \begin{pmatrix} 5 \\ 4 \\ 2 \end{pmatrix}$$
$$x_1, x_2 \geq 0$$

*Solution.* Switch min to max with new objective function $(1, -2, 3)x$

Divide $x_3 = x_3^+ - x_3^-$ giving $\begin{pmatrix} 1 & 5 & 3 & -3 \\ 2 & -1 & 2 & -2 \\ 1 & 2 & -1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3^+ \\ x_3^- \end{pmatrix}$

Add slack variables $x_4$, $x_5$ giving new rows $(1, 5, 3, 1, 0)x = 5$ and $(2, -1, 2, 0, 1)x = 4$

Let $x = (x_1, x_2, x_3^+, x_3^-, x_4, x_5)^\top$ and combine to get the SEF

$$\max \left\{ (1, -2, 3, -3, 0, 0)x : \begin{pmatrix} 1 & 5 & 3 & -3 & 1 & 0 \\ 2 & -1 & 2 & -2 & 0 & 1 \\ 1 & 2 & -1 & 1 & 0 & 0 \end{pmatrix} x = \begin{pmatrix} 5 \\ 4 \\ 2 \end{pmatrix}, x \geq \mathbb{0} \right\}$$

Suppose simplex solves this and gives optimal solution $(\frac{11}{4}, 0, \frac{3}{4}, 0, 0, 3)^\top$ with optimal value 5. Then, the original LP is solved by $(\frac{11}{4}, 0, \frac{3}{4} - 0)^\top = (\frac{11}{4}, 0, \frac{3}{4})$ with optimal value $-5$.  □

*Lecture 8*
*May 25*

**Example 2.1.6.** $\max(3, -2, 0, 0, 0)x$ such that $x \geq \mathbb{0}$ and $\begin{pmatrix} 4 & -1 & 1 & 0 & 0 \\ 3 & -3 & 0 & 1 & 0 \\ -2 & 2 & 0 & 0 & 1 \end{pmatrix} x = \begin{pmatrix} 8 \\ 9 \\ 1 \end{pmatrix}$

*Solution.* Feasible solution: $\bar{x} = (0, 0, 8, 9, 1)^\top$ with objective value $(3, -2, 0, 0, 0)(0, 0, 8, 9, 1)^\top = 0$

To increase objective value, must increase $x_1$ (the only one with a positive coefficient in the objective function)

Change $x_1 \mapsto t$, keep $x_2 = 0$, and set $x_3 \mapsto 8 - 4t$, $x_4 \mapsto 9 - 3t$, and $x_5 \mapsto 1 + 2t$ to maintain feasibility

But we still have non-negativity, giving us $t \leq 2$, i.e., $\bar{\bar{x}} = (2, 0, 0, 3, 5)^\top$                                    □

This example worked because we had (1) the identity matrix embedded in columns, (2) those columns have zero coefficients in the objective function. Equivalent strategies will exist if the matrix's rows are independent.

> **Notation.** Given a matrix $A$, notate the column $j = 1, \ldots, n$ of $A$ by $A_j$ and then the submatrix formed by columns $J \subseteq \{1, \ldots, n\}$ by $A_J$

> **Proposition 2.1.7**
>
> TFAE: $B \subseteq \{1, \ldots, n\}$ is a basis; $A_B$ is invertible; and $|B| = m$ and the columns $A_B$ are linearly independent

> **Definition 2.1.8**
>
> Given a basis $B$, a <u>non-basis</u> $N = \{1, \ldots, n\} \setminus B$.
>
> The variables $x_B$ where $B$ a basis are <u>basic variables</u>; conversely, $x_N$ are <u>non-basic variables</u>. This lets us write $Ax = A_B x_B + A_N x_N$.
>
> A <u>basic solution</u> (with respect to $B$) is a solution $x$ where $x_N = 0$.

To find a basic solution $x_B$, multiply both sides of the constraint $Ax = b$ on the left by $A_B^{-1}$. Then, $A_B^{-1} Ax = A_B^{-1} b$ and we can read the basic solution $x_B = A_B^{-1} b$.

> **Definition 2.1.9** (basic feasible solution)
>
> A value $\bar{x}$ such that $\bar{x}_B \geq \mathbb{0}$.

> **Definition 2.1.10** (canonical form)
>
> SEF $\max\{c^\top x + \bar{z} : Ax = b, x \geq \mathbb{0}\}$ if $A_B = I$ and $c_B = \mathbb{0}$

*Lecture 9*
*May 30*

> **Example 2.1.11.** Write the canonical form for $B = \{2, 3\}$ of $\max(3, -1, 4, 0, -1)x + 2$ s.t. $x \geq \mathbb{0}$ and $\begin{pmatrix} 2 & 1 & -2 & 5 & -4 \\ -1 & 0 & -1 & -3 & 2 \end{pmatrix} x = \begin{pmatrix} 3 \\ -2 \end{pmatrix}$

*Solution.* Multiply the constraint on the left by $A_B^{-1} = \begin{pmatrix} 1 & -2 \\ 0 & -1 \end{pmatrix}$ to get new constraint $\begin{pmatrix} 4 & 1 & 0 & 11 & -8 \\ 1 & 0 & 1 & 3 & -2 \end{pmatrix} x = \begin{pmatrix} 7 \\ 2 \end{pmatrix}$ which is equivalent

Then, we need to set the entries in the objective function vector $c_B = \mathbb{0}$

To find an expression for $-c_B^\top x_B$ which can cancel, take $-c_B^\top Ax = -c_B b$ to get $(0, 1, -4, -1, 0)x = -1 \implies (0, 1, -4, -1, 0)x + 1 = 0$

Add this to $c^\top x + \bar{z}$ to get the new objective function $\max(3, 0, 0, -1, -1)x + 3$ □

In general, to convert from SEF to canonical form given a basis $B$:

1. Amend the constraint: $Ax = b \mapsto A_B^{-1} Ax = A_B^{-1} b$
2. Calculate $y = (A_B^{-1})^\top c_B$ (this will be used later as a certificate)
3. Amend the objective function: $c^\top x + \bar{z} \mapsto (c^\top - y^\top A)x + (\bar{z} - y^\top b)$

## 2.2 Simplex

Main idea: go between feasible bases, attempting to raise the objective value

1. Start with a feasible basis $B$.

   Use "2-phase simplex" to find one if there is not a trivial feasible basis

2. Convert the LP to canonical form with respect to $B$.

   If $c_N \leq \mathbb{0}$, then the optimal solution is the basic feasible solution. **Stop**.

3. Pick a non-basic entering variable $x_k$ where $c_k > 0$.

   (Bland's rule) Pick the variable with lowest index $k$.

   If $A_k \leq \mathbb{0}$, then the linear program is unbounded. **Stop**.

   Increase $x_k = t$ while reducing $x_B$ to increase the objective value, i.e., maximize $t$ subject to $x_B = b - A_k t$ and $x \geq \mathbb{0}$. That is, $t = \min\{\frac{b_i}{A_{ik}} : A_{ik} > 0\}$

4. Pick a leaving variable $x_\ell$.

   Let $x_\ell$ be the index used for the minimum above. Then, $x_\ell = 0$ after setting $x_k = t$ making $x_\ell$ non-basic.

   Repeat step 2 with the new basis $B = B \cup \{k\} \smallsetminus \{\ell\}$

**Example 2.2.1.** Maximize $(0, 2, 1, 0, 0, 0)x$ such that $x \geq \mathbb{0}$ and

$$\begin{pmatrix} 1 & -1 & 1 & 1 & 0 & 0 \\ -2 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & -2 & 0 & 0 & 1 \end{pmatrix} x = \begin{pmatrix} 5 \\ 3 \\ 5 \end{pmatrix}$$

*Solution.* Perform simplex:

1. Entering variable $x_2$ since $c_2 = 2 > 0$

   Set $x_2 = t = \min\{-, \frac{3}{1}, \frac{5}{1}\} = 3$

   Leaving variable is $x_5$, new basis $B = \{2, 4, 6\}$

   New canonical form $\max(4, 0, 1, 0, -2, 0)x + 6$ with $x \geq \mathbb{0}$ and

   $$\begin{pmatrix} -2 & 1 & 0 & 0 & 1 & 0 \\ -1 & 0 & 1 & 1 & 1 & 0 \\ 2 & 0 & 2 & 0 & -1 & 1 \end{pmatrix} x = \begin{pmatrix} 3 \\ 8 \\ 2 \end{pmatrix}$$

2. Entering variable $x_1$. Set $x_1 = t = \min\{-, -, \frac{2}{2}\} = 1$.

   Leaving variable $x_6$. New basis $B = \{1, 2, 4\}$

   New canonical form $\max(0, 0, 5, 0, 0, -2)x + 10$ with $x \geq \mathbb{0}$ and

   $$\begin{pmatrix} 1 & 0 & -1 & 0 & -\frac{1}{2} & \frac{1}{2} \\ 0 & 1 & -2 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & \frac{1}{2} & \frac{1}{2} \end{pmatrix} x = \begin{pmatrix} 1 \\ 5 \\ 9 \end{pmatrix}$$

3. Entering variable $x_3$. Set $x_3 = t = \min\{-, -, -\} = \infty$. No bound on $t$, so this LP is unbounded.

Iteration 3 entered $x_3$, and $A_3 \leq 0$. Then, $t$ had no bounds, so the LP is unbounded. Find a certificate of unboundedness:

*Lecture 10*
*June 1*

Current basic feasible solution: $\bar{x} = (1, 5, 0, 9, 0, 0)^\top$

Set $x_3 = t$. New solution is $x(t) = (1 + t, 5 + 2t, t, 9, 0, 0)^\top = \bar{x} + td$ where $d = (1, 2, 1, 0, 0, 0)^\top$.

Then, $x(t)$ is feasible for all $t \geq 0$ and $c^\top x(t) = 5t - 10 \to \infty$. This is all we need to show that $(\bar{x}, t)$ form a certificate of unboundedness.                                              $\square$

Recall: simplex iterates over feasible bases to optimize one variable at a time and set coefficients to zero, eventually finding an optimal solution or showing unboundedness

Does simplex always terminate? Intuitively it should, since each iteration increases the objective value. However, this is not true in general

> **Definition 2.2.2** (degenerate iteration)
>
> Iteration where the objective value is held constant. That is, only the basis is changed.

> **Example 2.2.3.** Maximize $(1, 0, 0, 0)x$ subject to $x \geq \mathbb{0}$ and $\begin{pmatrix} 3 & 2 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{pmatrix} x = (0, 2)^\top$.
>
> The start basis is clearly $\{1, 4\}$ The only entering variable is $x_1$. Let $t = x_1$ and $t = \min\{\frac{2}{1}, \frac{0}{3}\} = 0$ We pick $x_4$ for a leaving variable, giving $B' = \{1, 4\}$, which is the same.

> **Definition 2.2.4** (cycling)
>
> Repeating bases with the same basic feasible solution.

This happens when there is a series of degenerate iterations that repeat bases already used. Bland's Rule, when applied, prevents cycling and guarantees simplex terminates. That is, when there is a choice for the entering or leaving variable, pick the one with the lowest index.

**Tableau method** (easier to perform by hand)  Use a matrix to encode the entire state of the algorithm, allowing row reductions to replace taking inverses

**Example 2.2.5.** Maximize $z = (3, -2, 0, 0, 0)x + 0$ subject to $x \geq \mathbb{0}$ and

$$\begin{pmatrix} 4 & -1 & 1 & 0 & 0 \\ 3 & -3 & 0 & 1 & 0 \\ -2 & 2 & 0 & 0 & 1 \end{pmatrix} x = \begin{pmatrix} 8 \\ 9 \\ 1 \end{pmatrix}$$

*Solution.* Create a tableau $\left( \begin{array}{c|c|c} 1 & -c & \bar{z} \\ \hline \mathbb{0} & A & B \end{array} \right) = \left( \begin{array}{c|ccccc|c} 1 & -3 & 2 & 0 & 0 & 0 & 0 \\ \hline 0 & 4 & -1 & 1 & 0 & 0 & 8 \\ 0 & 3 & -3 & 0 & 1 & 0 & 9 \\ 0 & -2 & 2 & 0 & 0 & 1 & 1 \end{array} \right)$

Entering variable: look for negative coefficient in objective row $x_k$ (in this case $x_1$)

Leaving variable: look at ratio between $b$ and $A_k$ to find row with $x_\ell$ (in this case $A^1$)

Take the entry in column of $x_k$ and the row of $x_\ell$ (in this case $A_{11}$). Use row operations to make it 1 and the rest of the column 0.

In this case $\frac{1}{4}R_2$, $R_1 + 3R_2$, $R_3 - 3R_2$, $R_4 + 2R_2$, to yield $\left( \begin{array}{c|ccccc|c} 1 & 0 & \frac{5}{4} & \frac{3}{4} & 0 & 0 & 6 \\ \hline 0 & 1 & -\frac{1}{4} & \frac{1}{4} & 0 & 0 & 2 \\ 0 & 0 & -\frac{9}{4} & -\frac{3}{4} & 1 & 0 & 3 \\ 0 & 0 & \frac{3}{2} & \frac{1}{2} & 0 & 1 & 5 \end{array} \right)$

Since top row is all positive, this means that $c$ is all negative, so we have an optimal solution with optimal objective value 6 $\qquad\qquad\square$

Since tableau method involves iterative dividing of decimals, computers will lose precision after each iteration. With canonical form, calculations are always done from the original LP, so no compounded error.

If no obvious basic feasible solution to start simplex, do two-stage simplex.

## 2.3   Two-phase simplex

Main idea: Form an auxiliary LP with an obvious BFS, then solve with simplex to help find a BFS or prove that no BFS exists

- Negate constraints with negative right-hand sides
- Arbitrarily add auxiliary variables to append an identity matrix to $A$
- Apply simplex to the new LP: there is a solution to the original LP if and only if this LP has solutions with the auxiliary variables set to zero

**Example 2.3.1.** Maximize $(-3, 2, 0, 0, 0)x$ subject to $x \geq \mathbb{0}$ and

$$\begin{pmatrix} -1 & 1 & 1 & 0 & 0 \\ -1 & -2 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{pmatrix} x = \begin{pmatrix} -1 \\ -4 \\ 4 \end{pmatrix}$$

*Solution.* Notice that $B = \{3, 4, 5\}$ is a basis. However, the trivial basic solution $(0, 0, -1, -4, 4)^\top$

is not feasible.

Negate constraints to get $\begin{pmatrix} 1 & -1 & -1 & 0 & 0 \\ 1 & 2 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{pmatrix} x = \begin{pmatrix} 1 \\ 4 \\ 4 \end{pmatrix}$.

Add auxiliary variables to get $\begin{pmatrix} 1 & -1 & -1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 2 & 0 & -1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \end{pmatrix} x = \begin{pmatrix} 1 \\ 4 \\ 4 \end{pmatrix}$.

This LP has a BFS $(0, 0, 0, 0, 0, 1, 4, 4)^\top$.

Start with $B = \{6, 7, 8\}$ with BFS $\bar{x} = (0, 0, 0, 0, 0, 1, 4, 4)^\top$.

Minimize the sum of auxiliary variables: $\min (0, 0, 0, 0, 0, 1, 1, 1)x$ subject to same constraints.

Same as saying $\max (0, 0, 0, 0, 0, -1, -1, -1)x$, which we can solve with simplex.

If optimal value is 0, then auxiliary variables must be 0. If optimal value is negative, no feasible solution exists.

Based on canonical form, multiply constraint on left by $(1, 1, 1)$ to get $(2, 1, -1, -1, 1, 1, 1, 1)x - 9 = 0$.

Add to objective function to get new objective function $(2, 1, -1, -1, 1, 0, 0, 0)x - 9$.

Run simplex to get basis $\{1, 3, 5\}$ and BFS $(4, 0, 3, 0, 4, 0, 0, 0)^\top$.

Now that we have a BFS $(4, 0, 3, 0, 4)^\top$, run simplex. □

If auxiliary LP does not have a solution, then $y = A_B^{-T} x_B$ is a certificate of infeasibility for the original LP.

Optimal solution must exist: it is feasible by construction and not unbounded because the objective value cannot exceed 0.

Therefore, two-phase simplex will always work.

## 2.4   Geometry of simplex

Consider $\max (1, 1)x$ subject to $x \geq \mathbb{0}$ and $\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 2 & 1 \end{pmatrix} x \leq \begin{pmatrix} 4 \\ 3 \\ 4 \end{pmatrix}$.

Equivalently, $\max (1, 1, 0, 0, 0)x$ subject to $x \geq \mathbb{0}$ and $\begin{pmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 2 & 1 & 0 & 0 & 1 \end{pmatrix} x = \begin{pmatrix} 4 \\ 3 \\ 4 \end{pmatrix}$.

Equalities define lines, inequalities define halfspaces:

- Non-negativity bounds $(x_1, x_2) \in \mathbb{R}^2$ to the first quadrant.
- Constraints bound $x_1 \leq 4$, $x_2 \leq 3$, and $x_1 + 2x_2 \leq 8$.

All feasible solutions are thus contained in the convex polygon bounded by $(0, 0)$, $(0, 3)$, $(2, 3)$,

$(4, 2)$, $(4, 0)$.

Consider contours of the objective function $x_1 + x_2 = k$ starting at $k = 0$. Imagine dragging the objective function line "up" along the normal vector $(1, 1)^\top$ maintaining the slope. Last point $(4, 2)^\top$ which touches the highest contour $k = 6$ is the optimal solution.

Notice that $(4, 2)^\top$ is generated by $x_1 = 4$ and $x_1 + 2x_2 = 8$, i.e., setting constraints (1) and (3) to equality. Then, in SEF, $(4, 2)^\top \mapsto (4, 2, 0, 1, 0)^\top$ is a BFS.

Likewise, consider the extreme point $(4, 0)^\top$ generated by $x_1 = 4$ and $x_2 = 0$, i.e., constraints (1) and (5) to equality. In SEF, $(4, 0)^\top \mapsto (4, 0, 0, 3, 4)^\top$ which is a BFS.

We can make some observations.

> **Definition 2.4.1** (boundary)
>
> A point $p \in S$ is on the <u>boundary of $S$</u> if every open neighbourhood of $p$ contains a point in $S$ and a point not in $S$.

> **Remark 2.4.2.** Optimal solutions must always be on the boundary.

> **Definition 2.4.3** (tight constraint)
>
> An inequality satisfied with equality by a feasible solution $\bar{x}$.

> **Theorem 2.4.4**
>
> If an optimal solution exists and the rows of $A$ are linearly independent, then there is at least one optimal "corner".

*Proof.* Setting inequality to equality means setting a slack variable in SEF to 0. Every corner point is defined by $n$ linear equalities in $\mathbb{R}^n$, i.e., $n$ (or more) tight constraints. Therefore, a corner point's canonical form solution will have $n$ slack variables set to zero, creating a BFS.                  $\square$

Geometrically, simplex walks along extreme points (BFSs) until it finds an optimal solution. Two-phase simplex is required when the origin is not available as a corner to start from. Each iteration picks an adjacent extreme point, because it swaps one tight constraint for another one and maintains the other $n - 1$ constraints.

> **Definition 2.4.5** (degenerate)
>
> $\bar{x}$ with more than $n$ tight constraints, i.e., some basic variables are 0.

This means an iteration may bounce endlessly between degenerate BFSs. Degenerate iterations are avoided by Bland's rule.

Finally, simplex requires that the feasible region be convex. It always is because it is the convex hull generated by the extreme points.

Formalize some basic geometric ideas from last lecture.

---

**Definition 2.4.6** (geometry)

A <u>hyperplane</u> is an equation of the form $a^\top x = \beta$ with <u>normal vector</u> $a$.

A <u>halfspace</u> is an inequality of the form $a^\top x \le \beta$ with <u>normal vector</u> $a$. Note: the halfspace is always on the opposite side of the normal vector.

A <u>polyhedron</u> is a set of points $\{x \in \mathbb{R}^n : Ax \le b\}$, i.e., the intersection of $m$ halfspaces for a polytope with $m$ facets.

---

**Definition 2.4.7** (convexity)

Subset $S \subseteq \mathbb{R}^n$ such that for all $x^1, x^2 \in S$ and $\lambda \in [0, 1]$, $\lambda x^1 + (1 - \lambda)x^2 \in S$. This expression is a <u>convex combination</u>.

---

**Proposition 2.4.8**

A halfspce is convex.

---

*Proof.* Let $H = \{x \in \mathbb{R}^n : a^\top x \le \beta\}$ be a halfspace and $x^1, x^2 \in H$. Then, with arbitrary $\lambda \in [0, 1]$, let $x = \lambda x^1 + (1 - \lambda)x^2$.

We have $a^\top x = \lambda a^\top x^1 + (1 - \lambda)a^\top x^2$. But since $x^1, x^2 \in H$, we get $a^\top x \le \lambda\beta + (1 - \lambda)\beta = \beta$ as desired. $\qquad\square$

---

**Proposition 2.4.9**

The intersection of convex sets is convex.

---

*Proof.* If two points are in the intersection, they are in each convex set. The line segment joining them is then also in each convex set, meaning it is in the intersection. $\qquad\square$

---

**Corollary 2.4.10.** A polyhedron is convex.

---

**Definition 2.4.11** (extreme point)

Point $\bar{x} \in C$ a convex set such that $x$ is not a convex combination of two distinct points in $C$ distinct from $\bar{x}$.

---

**Theorem 2.4.12** (characterization of extreme points in polyhedra)

For $\bar{x} \in P = \{x \in \mathbb{R}^n : Ax \le b\}$ a polyhedron where $A^= x = b^=$ is the set of tight constraints for $\bar{x}$, $\bar{x}$ is an extreme point of $P$ if and only if $\text{rank}(A^=) = n$.

*Proof.* Proceed by the contrapositive.

($\Leftarrow$) Suppose $\bar{x}$ is not an extreme point. Then, there exist distinct $x^1, x^2 \in P$ such that $\bar{x} = \lambda x^2 + (1 - \lambda)x^2$ for some $0 < \lambda < 1$.

We have
$$A^=\bar{x} = A^=(\lambda x^1 + (1 - \lambda)x^2) = \lambda(A^=x^1) + (1 - \lambda)(A^=x^2) \leq b^=.$$

That is, there is equality throughout the line $A^=x^1 = A^=x^2 = b^=$.

Then, since $A^=x = b^=$ has at least three solutions $(x^1, x^2, \bar{x})$, we cannot have $\text{rank}(A^=) = n$.

($\Rightarrow$) Suppose $\text{rank}(A^=) < n$. Then, there exists non-zero $d$ such that $A^=d = \mathbb{0}$.

Let $x^1 = \bar{x} + \varepsilon d$ and $x^2 = \bar{x} - \varepsilon d$ for some small positive $\varepsilon$. By construction, $\bar{x}$ is properly contained in the line segment from $x^1$ to $x^2$.

Then, we can write
$$A^=x^1 = A^=(\bar{x} + \varepsilon d) = A^=\bar{x} + \varepsilon A^=d = A^=\bar{x} = b^=$$

which means $x^1$ (and likewise $x^2$) satisfy the tight constraints.

For other non-tight constraints $a^\top x \leq \beta$, we have

$$a^\top \bar{x} + \varepsilon(a^\top d) < \beta + \varepsilon(a^\top d) \leq \beta$$

for sufficiently small $\varepsilon$. Therefore, $x^1$ and $x^2$ are in $P$ and $\bar{x}$ is not an extreme point. $\qquad \square$

> **Theorem 2.4.13** (characterization of extreme points in SEF)
>
> Let $P = \{x \in \mathbb{R}^n : Ax = b, x \geq \mathbb{0}\}$ where $A$ has full row rank and let $\bar{x} \in P$.
>
> Then, $\bar{x}$ is an extreme point of $P$ if and only if $\bar{x}$ is a BFS of $Ax = b$.

*Proof.* Construct a polyhedron $\left( \dfrac{A}{\dfrac{-A}{-I}} \right) x \leq \left( \dfrac{b}{\dfrac{-b}{\mathbb{0}}} \right) x.$

Suppose $A\bar{x} = b$. By thm. 2.4.12, $\bar{x}$ and extreme point if and only if the tight constraints have rank $n$. The top two submatrices are always tight, so we automatically get rank at least $m$. We need at least $n - m$ tight constraints in $-I\bar{x} \leq \mathbb{0}$. That is, $n - m$ entries of $\bar{x}$ are zero.

Then, $\bar{x}$ is a basic feasible solution with $m$ basic variables and $n - m$ non-basic variables. $\qquad \square$

# Chapter 4

# Duality

## 4.1 Strong and weak duality

**Example 4.1.1.** Find $\max\left\{(4, -9, 2, 4)x : \begin{pmatrix} 1 & -4 & 3 & 0 \\ -1 & 7 & -5 & 1 \end{pmatrix} x = \begin{pmatrix} 7 \\ 3 \end{pmatrix}, x \geq \mathbb{0}\right\}.$

*Solution.* Notice that $(6, 2, 3, 10)^\top$ is feasible. Objective value 52.

$(7, 0, 0, 10)^\top$ is feasible. Objective value 68. Claim this is optimal.

Notice that:

$$
\begin{aligned}
68 &= 7 \cdot 8 + 3 \cdot 4 \\
&= 8(x_1 - 4x_2 + 3x_3) + 4(-x_1 + 7x_2 - 5x_3 + x_4) \\
&= 4x_1 - 4x_2 + 4x_3 + 4x_4 \\
&\geq 4x_1 - 9x_2 + 2x_3 + 4x_4 = c^\top x
\end{aligned}
$$

Since $c^\top x \leq 68$, ny feasible solution has objective value at most 68. As $(7, 0, 0, 10)^\top$ is one such solution, it is optimal. □

In general, consider $y$. Then, $y^\top b = y^\top A x \overset{?}{\geq} c^\top x$. This works when $y^\top A \geq c^\top$ or equivalently $A^\top y \geq c^\top$. Then, $y^\top b$ is an upper bound on the objective value. If $y^\top b$ is the lowest upper bound, it is the optimal objective value.

To find $y$, we can solve the dual LP.

> **Definition 4.1.2** (dual LP)
>
> Given an LP (P) in SEF $\max\{c^\top x : Ax = b, x \geq \mathbb{0}\}$, construct the dual (D) $\min\{b^\top y : A^\top y \geq c\}$

Each constraint $a_i x_i = b_i$ in (P) corresponds to a variable $y_i$ in (D). Each variable $x_j$ in (P) corresponds to a constraint $(a^j)^\top y_j \leq c_j$ in (D).

Notice that the feasible objective values of the dual are at most the feasible objective values of the primal.

Then, the highest (optimal) value of the primal is bounded above by the lowest (optimal) value of the dual. In fact, they meet only at optimality.

> **Theorem 4.1.3** (weak duality)
>
> For an LP in SEF (P) $\max\{c^\top x : Ax = b, x \geq \mathbb{0}\}$ and its dual (D) $\min\{b^\top y : A^\top y \geq c\}$, if $\bar{x}$ and $\bar{y}$ are feasible, then $c^\top \bar{x} \leq b^\top \bar{y}$.
>
> Moreover, if $c^\top \bar{x} = b^\top \bar{y}$, then $\bar{x}$ and $\bar{y}$ are optimal.

*Proof.* We have $b^\top \bar{y} = (A\bar{x})^\top \bar{y} = \bar{x}^\top A^\top \bar{y} \geq \bar{x}^\top c = c^\top \bar{x}$, as desired.

Suppose $c^\top \bar{x} = b^\top \bar{y}$. Let $x$ be arbitrary. Then, $c^\top x \leq b^\top \bar{y} = c^\top \bar{x}$. This is what it means for $\bar{x}$ to be optimal. Argue symmetrically for $y$.  □

> **Corollary 4.1.4.** To certify $\bar{x}$ is optimal for (P), it suffices to show that: $\bar{x}$ is feasible, $\bar{y}$ is feasible, and $c^\top \bar{x} = b^\top \bar{y}$.

> **Corollary 4.1.5.** If (P) is unbounded, then (D) must be infeasible. Similarly, if (D) is unbounded, then (P) is infeasible. Finally, if both (P) and (D) are feasible, then they must have optimal solutions.

*Proof.* Any feasible solution to (D) is an upper bound to the objective values of (P). Likewise in the other direction.  □

> **Theorem 4.1.6** (strong duality)
>
> If (P) has an optimal solution, then (D) has an optimal solution with the same optimal values.

*Proof.* Perform simplex on (P) to get a basic optimal solution $\bar{x}$ with basis $B$. The objective function in canonical form for $B$ is $(c^\top - \bar{y}^\top A)x + b^\top \bar{y}$ where $\bar{y} = A_B^{-\top} c_B$.

Since $B$ is optimal, $c^\top - \bar{y}^\top A \leq \mathbb{0}^\top$ which gives $A^\top y \geq c$, i.e., $\bar{y}$ is feasible. Also, the objective value of $\bar{x}$ is $c^\top x = b^\top y$.  □

*Lecture 16*
*June 22*

Summarize the weak duality theorem results:

| Dual \ Primal | Optimal | Unbounded | Infeasible |
|---|---|---|---|
| Optimal | ✓ | ✗ | ✗ |
| Unbounded | ✗ | ✗ | ✓ |
| Infeasible | ✗ | ✓ | ✓ (edge case) |

> **Proposition 4.1.7** (dual of inequality)
>
> Dual of $\max\{c^\top x : Ax \le b, x \ge \mathbb{0}\}$ is $\min\{b^\top y : A^\top y \ge c, y \ge \mathbb{0}\}$.

*Proof.* Turn into SEF by adding slack variables: $\max\{c^\top x + \mathbb{0}^\top x' : Ax + Ix' = b, x \ge \mathbb{0}\}$.

Get the dual of that which is $\min\{b^\top y : A^\top y \ge c, y \ge \mathbb{0}\}$.

Likewise, the dual of $\max\{c^\top x : Ax \ge b, x \ge \mathbb{0}\}$ is $\min\{b^\top y : A^\top y \ge c, y \le \mathbb{0}\}$ $\qquad\square$

> **Proposition 4.1.8** (dual with free variables)
>
> Dual of $\max\{c^\top x : Ax = b\}$ is $\min\{b^\top y : A^\top y = c\}$.

*Proof.* Turn into SEF: $\max\{c^\top x^+ - c^\top x^- : Ax^+ - Ax^- = b, x^+ \ge \mathbb{0}, x^- \ge \mathbb{0}\}$.

Dual is $\min\{b^\top y : A^\top y \ge c, -A^\top y \ge -c\} = \min\{b^\top y : A^\top y = c\}$.

In fact, having $x \ge \mathbb{0}$ will give $A^\top y \ge c$ and $x \le \mathbb{0}$ will give $A^\top y \le c$. $\qquad\square$

> **Proposition 4.1.9** (dual of min)
>
> Dual of the minimization LP (P*) is the maximization LP (D*) whose dual is (P*).

*Proof.* From max to min: constraints (swapped) $\to$ variables, variables $\to$ constraints.

From min to max: constraints $\to$ variables, variables (swapped) $\to$ constraints. $\qquad\square$

> **Example 4.1.10.** Find the dual: Maximize $(3, 1, 4, 1)x$ such that $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \\ 2 & 2 & 2 & 2 \end{pmatrix} x \begin{matrix} \ge \\ \le \\ = \end{matrix} \begin{pmatrix} 3 \\ 4 \\ 5 \end{pmatrix}$
>
> with $x_1, x_2 \ge 0$, $x_3 \le 0$, and $x_4$ free.

*Solution.* Minimize $(3, 4, 5)y$ such that:

- Dual constraints max $\to$ min kept: $\begin{pmatrix} 1 & 4 & 2 \\ 2 & 3 & 2 \\ 3 & 2 & 2 \\ 4 & 1 & 2 \end{pmatrix} y \begin{matrix} \ge \\ \ge \\ \le \\ = \end{matrix} \begin{pmatrix} 3 \\ 1 \\ 4 \\ 1 \end{pmatrix}$.
- Dual variables max $\to$ min flipped: $y_1 \le 0$, $y_2 \ge 0$, $y_3$ free.

$\qquad\square$

> **Theorem 4.1.11** (general weak duality)
>
> Let (P) be a maximization LP with dual (D). If $\bar{x}$ and $\bar{y}$ are feasible for (P) and (D), respectively, then the objective value of $\bar{x}$ in (P) is at most the objective value of $\bar{y}$ in (D).
>
> If they are the same value, then they are both optimal.

> **Theorem 4.1.12** (general strong duality)
>
> Let (P) be an LP with dual (D). If (P) has an optimal solution, then (D) has an optimal solution with the same optimal value.

## 4.2   Complementary slackness

> **Definition 4.2.1** (complementary slackness conditions)
>
> For $(A^\top y)^\top x \geq c^\top x$ to hold with equality, we have $A_i^\top y x_i = c_i x_i$. Equivalently, either $A_i^\top y = c_i$ or $x_i = 0$.

CS conditions are satisfied by feasible $\bar{x}$ and $\bar{y}$ if and only if $\bar{x}$ and $\bar{y}$ are optimal.

For an LP in SEF, the CS conditions are "either $x_i = 0$ or the $i$-th dual constraint is tight". Equivalently, "if $x_i > 0$, then the $i$-th dual constraint is tight" and "if the $i$-th dual constraint is not tight, $x_i = 0$".

In general, given primal (P) and dual (D), their CS conditions are: (1) either $x_i = 0$ or the $i$-th constraint of (D) is tight and (2) either $y_j = 0$ or the $j$-th constraint of (P) is tight. If there are equality constraints, we can ignore that set of CS conditions since all constraints are tight.

> **Theorem 4.2.2** (complementary slackness)
>
> Let (P) and (D) be a primal-dual pair with feasible solutions $\bar{x}$ and $\bar{y}$, respectively. Then, $\bar{x}$ and $\bar{y}$ are optimal if and only if all CS conditions hold.

To show $\bar{x}$ is optimal for (P), we can provide $\bar{y}$ and check that (1) $\bar{x}$ is feasible, (2) $\bar{y}$ is feasible, (3) CS conditions.

**Example 4.2.3.** Find an optimal solution for $\max \left\{ (1, -2)x : \begin{pmatrix} 1 & -1 \\ 2 & 1 \\ 0 & 1 \end{pmatrix} x \begin{matrix} = \\ \geq \\ \leq \end{matrix} \begin{pmatrix} 2 \\ 1 \\ 4 \end{pmatrix}, x_1 \geq 0 \right\}$

with dual $\min \left\{ (2, 1, 4)y : \begin{pmatrix} 1 & 2 & 0 \\ -1 & 1 & 1 \end{pmatrix} y \begin{matrix} \geq \\ = \end{matrix} \begin{pmatrix} 1 \\ -2 \end{pmatrix} : y_2 \leq 0, y_3 \geq 0 \right\}$

*Solution.* CS conditions:

- $x_1 = 0$ or $y_1 + 2y_2 = 1$
- (always true) $x_2 = 0$ or $\boxed{-y_1 + y_2 + y_3 = -2}$
- (always true) $y_1 = 0$ or $\boxed{x_1 - x_2 = 2}$
- $y_2 = 0$ or $2x_1 + x_2 = 1$
- $y_3 = 0$ or $x_2 = 4$

Suppose we think that $\bar{x} = (1, -1)^\top$ is optimal. By inspection, $\bar{x}$ feasible. Since $\bar{x}_1 \neq 0$, we require $\bar{y}_1 + 2\bar{y}_2 = 1$. The second condition is satisfied. Also, $\bar{x}_2 \neq 4$, so $\bar{y}_3 = 0$.

Adding on the requirement that $\bar{y}$ feasible, we can solve for $\bar{y} = (\frac{5}{3}, -\frac{1}{3}, 0)^\top$.

Then, $\bar{y}$ is feasible and satisfies CS conditions, so $\bar{x}$ and $\bar{y}$ are optimal. □

If we try the above example with $\bar{x} = (4, 2)^\top$, we get CS conditions implying $\bar{y} = (1, 0, 0)^\top$, but this is not feasible. Therefore, $\bar{x}$ is not optimal.

The above example with simplex gives first

$$\max \left\{ (3, -2, 0, 0, 0)x : \begin{pmatrix} 4 & -1 & 1 & 0 & 0 \\ 3 & -3 & 0 & 1 & 0 \\ -2 & 2 & 0 & 0 & 1 \end{pmatrix} x = \begin{pmatrix} 8 \\ 9 \\ 1 \end{pmatrix}, x \geq \mathbb{0} \right\}$$

and ends with

$$\max \left\{ (0, -\frac{5}{4}, -\frac{1}{4}, 0, 0)x + 6 : \begin{pmatrix} 1 & -\frac{1}{4} & \frac{1}{4} & 0 & 0 \\ 0 & -\frac{9}{4} & -\frac{3}{4} & 1 & 0 \\ 0 & \frac{3}{2} & \frac{1}{2} & 0 & 1 \end{pmatrix} x = \begin{pmatrix} 2 \\ 3 \\ 5 \end{pmatrix}, x \geq \mathbb{0} \right\}$$

For each variable $i$, either $x_i = 0$ (it is non-basic) or $c_i = 0$ (it is basic). Recall: objective coefficients are $c' = c - y^\top A$ where $y = A_B^{-\top} c_B$ is the dual solution.

Then, $c_i' = 0$ implies $c_i - y^\top A_i = 0$ or $y^\top A_i = c_i$ or $A_i^\top y = c_i$. That is, the $i^{\text{th}}$ constraint in the dual is tight.

That is, either $x_i = 0$ or the $i$-th constraint is tight, so all simplex BFS's satisfy the CS conditions.

Also, when picking $c_e > 0$ for the next iteration, $A_e^\top y < c_e$, i.e., we picked an infeasible $y$ to the dual (since the dual constraint is $Ay \leq c$)

That is, whenever positive coefficients exist, $y$ is infeasible for the dual. Once the coefficients are non-positive, $y$ is feasible and we have an optimal solution.

## 4.3  Geometry of the dual

**Example 4.3.1.** For arbitrary $c$, $\max \left\{ c^\top x : \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 2 \\ -1 & 0 \\ 0 & -1 \end{pmatrix} x \leq \begin{pmatrix} 4 \\ 3 \\ 8 \\ 0 \\ 0 \end{pmatrix} \right\}$.

*Solution.* (imagine polyhedron bounded by $(0, 0)$, $(0, 3)$, $(2, 3)$, $(4, 2)$, and $(4, 0)$)

$(4, 2)^\top$ is the optimal solution whenever the slope is between the two constraints defining $(4, 2)^\top$, i.e., $-\infty \leq m \leq -\frac{1}{2}$.

Formally, any $c$ in the <u>cone</u> of $(1, 0)^\top$ and $(1, 2)^\top$ (the tight constraints). □

**Definition 4.3.2** (cone (generated by $x$))
The set $\{\lambda^\top x : \lambda \geq \mathbb{0}\}$

> **Lemma 4.3.3** (cone of tight constraints)
>
> $c$ is in the cone of tight constraints for $\bar{x}$ if and only if $\bar{x}$ is optimal for $c$.

**Example 4.3.4.** Consider the primal-dual pair $\max\left\{(c_1, c_2)x : \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 2 \\ -1 & 0 \\ 0 & -1 \end{pmatrix} x \le \begin{pmatrix} 4 \\ 3 \\ 8 \\ 0 \\ 0 \end{pmatrix}\right\}$ and

$$\min\left\{(4, 3, 8, 0, 0)y : \begin{pmatrix} 1 & 0 & 1 & -1 & 0 \\ 0 & 1 & 2 & 0 & -1 \end{pmatrix} y = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}, y \ge \mathbb{0}\right\}$$

*Solution.* Suppose $c$ makes $\bar{x} = (4, 2)^\top$ optimal. Let $y$ be the correponding optimal dual solution. The dual constraints are

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} y_1 + \begin{pmatrix} 0 \\ 1 \end{pmatrix} y_2 + \begin{pmatrix} 1 \\ 2 \end{pmatrix} y_3 + \begin{pmatrix} -1 \\ 0 \end{pmatrix} y_4 + \begin{pmatrix} 0 \\ -1 \end{pmatrix} y_5 = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$$

CS conditions: either $y_j = 0$ or the $j$-th primal constraint is tight. Here, constraints 1 and 3 are tight. Then, $y_2 = y_4 = y_5 = 0$ and we have

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} y_1 + \begin{pmatrix} 1 \\ 2 \end{pmatrix} y_3 = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$$

with $y_1, y_3 \ge 0$.

That is, $c$ is in the cone formed by the tight constraints for $\bar{x}$. $\qquad\square$

> **Lemma 4.3.5** (Farkas' lemma)
>
> Given $A$ and $b$, either there exists $x$ such that $Ax = b$ and $x \ge \mathbb{0}$, or there exists $y$ such that $y^\top A \ge \mathbb{0}^\top$ and $y^\top b < 0$.
>
> Equivalently, either a solution or certificate of infeasibility exists.

*Proof.* First, notice that having a certificate of infeasibility proves that there cannot be a solution to $Ax = b$ with $x \ge \mathbb{0}$. Thus, both cannot hold.

Now, suppose the first condition is false. Consider $\max\{\mathbb{0}^\top x : Ax = b, x \ge \mathbb{0}\}$.

This is infeasible by supposition. The dual is $\min\{b^\top y : A^\top y \ge \mathbb{0}\}$ which is trivially feasible since it admits $y = \mathbb{0}$, so it must be unbounded.

Then, there must exist a solution $\bar{y}$ with negative objective value, i.e., $b^\top \bar{y} < 0$ and $A^\top y \ge \mathbb{0}$, as desired. $\qquad\square$

Geometrically, consider the cone of column vectors. That is, all $b$ such that $Ax = b$, $x \ge \mathbb{0}$.

If $b$ is in the cone, we are done. If $b$ is not in the cone, then there must exist a hyperplane with normal $y$ between the cone and $b$. Such a $y$ will give by definition $y^\top b < 0$ and $y^\top A > \mathbb{0}^\top$.

## 4.4 Shortest paths, revisited

Recall the formulation for shortest path problems: $\min\{c^\top x : \sum_{e \in \delta(S)} x_e \geq 1, x_e \geq 0, x_e \in \mathbb{Z}\}$

Since there are $O(2^n)$ constraints and it is an IP, this is hard to solve.

Instead, ignore integrality and take the dual: $\max\{\mathbb{1}^\top y : \sum_{e \in \delta(S)} y_S \leq c, y \geq \mathbb{0}\}$.

There are dual variables $y_S$ corresponding to each $s, t$-cut and dual constraints for each edge. The objective function is the sum of all $y$ variables.

Dual constraints: One constraint for each edge $e$, coefficient of $y_S$ is 1 if $e$ is in the cut $\delta(S)$. That is, sum of $y$ values is at most the length of $e$

Interpret a dual solution $y$ as assigning a value $y_S$ to each cut so that the "sum of cuts" across an edge is at most the edge's weight.

Since an $s, t$-path must cross all $s, t$-cuts, this means that the "sum of cuts" is at most the sum of weights of the edges in the path, i.e., the sum of cuts is at most the length of the path.

Equivalently, by weak duality, the objective value of the primal (length of the path) is at least the objective value of the dual (sum of all cuts)

Imagine $y_S$ as the width of a barrier that a path must cross, so that a complete path must pass all barriers. This is the width assignment.

Since we are maximizing, the barriers will expand to fill the path.

Take the dual and consider CS conditions:

- If $e$ is in the path, then the sum of width of all cuts containing $e$ is equal to its length and $e$ is tight
- If $y_S$ has positive width, then exactly one edge in $\delta(S)$ is in the path

Therefore, we can construct an algorithm to raise the $y$-value of $\delta(U)$ to create a tight edge.

**Example 4.4.1.** $E = \{sa, sb, ab, at, bt\}$ with weights $w = \{50, 30, 30, 40, 70\}$.

*Solution.* Walk through our "algorithm":

- Initialize $y = \mathbb{0}$, $U = \{s\}$, $T = \varnothing$

- Iteration 1: $U = \{s\}$, $\delta(U) = \{sa, ab\}$. Can increase by 30 until $sb$ becomes tight, so set $y_{\{s\}} = 30$. Append $b$ to $U$.

- Iteration 2: $U = \{s, b\}$, $\delta(U) = \{sa, ab, bt\}$. Can increase by 20 until $sa$ becomes tight, so set $y_{\{s,b\}} = 20$. Append $a$ to $U$.

- Iteration 3: $U = \{s, b, a\}$, $\delta(U) = \{at, bt\}$. Can increase by 40 until $at$ becomes tight, so set $y_{\{s,b,a\}} = 40$. Append $t$ to $U$.

Tight edges $T = \{sb, sa, at\}$ so select path $x_{sa} = x_{at} = 1$ with $y_{\{s\}} = 30$, $y_{\{s_b\}} = 20$, $y_{\{s,a,b\}} = 40$. $\qquad\square$

Now, formalize the algorithm.

---

**Definition 4.4.2** (slack)

The length of $e$ minus width of all cuts through $e$, i.e., $\text{slack}(e) = c_e - \sum_{\{S : e \in \delta(S)\}} y_S$

---

**Algorithm 1** Shortest path algorithm

1: Set $y = \mathbb{0}$, $U = \{s\}$, $T = \varnothing$
2: **while** $t \notin U$ **do**
3:     Calculate slack of all edges in $\delta(U)$
4:     Pick edge $uv$ with minimum slack where $u \in U$, $v \notin U$
5:     Set $y_U := \text{slack}(uv)$
6:     $U := U \cup \{v\}$, $T := T \cup \{uv\}$
7: Output an $s,t$-path in $T$. Use $y$ as an optimal dual solution.

---

**Example 4.4.3.** $V = \{s, a, b, c, t\}$, $E = \{sa, sb, sc, ab, bc, at, bt, ct\}$ with $w = \{20, 50, 30, 20, 10, 50, 40, 40\}$.

*Solution.* Perform the algorithm:

1. $U = \{s\}$, $T = \varnothing$, $\delta(U) = \{sa, sb, sc\}$.

   Slacks are 20, 50, and 30, so set $y_{\{s\}} = 20$ to get tight $sa$.

2. $U = \{s, a\}$, $T = \{sa\}$, $\delta(U) = \{at, ab, sb, sc\}$.

   Slacks are 50, 20, 50, and 10, so set $y_{\{s,a\}} = 10$ to get tight $sc$.

3. $U = \{s, a, c\}$, $T = \{sa, sc\}$, $\delta(U) = \{at, ab, sb, bc, ct\}$.

   Slacks are 40, 10, 20, 10, 40, so set $y_{\{s,a,c\}} = 10$ to get tight $ab$ and $bc$. Arbitrarily pick $ab$.

4. $U = \{s, a, b, c\}$, $T = \{sa, sc, ab\}$, $\delta(U) = \{at, bt, ct\}$.

   Slacks are 30, 40, and 30, so set $y_{\{s,a,b,c\}} = 30$ to get tight $at$ and $ct$. Arbitrarily pick $ct$.

5. $U = \{s, a, b, c, t\}$, $T = \{sa, sc, ab, ct\}$.

   Conclude with path $\{sc, ct\}$.

   $\square$

Claim (proof omitted) that the algorithm produces a path that satisfies the second CS condition (i.e., every cut with positive width crosses only one edge in the path)

Sketch: To cross a cut more than once, must go back and forth. To go back, must select an edge going "backwards", i.e., selecting a node already in $U$ to append to $U$. Contradiction.

# Chapter 6

# Solving IPs

We cannot apply fundamental theorem of LPs to IPs, since some IPs may be feasible, bounded, and still lack an optimal solution.

> **Example 6.0.1.** $\max\{x_1 - \sqrt{2}x_2 : x_1 - \sqrt{2}x_2 \leq 0, x \geq 1, x \in \mathbb{Z}^2\}$ has a feasible solution $x = (1,1)^\top$, is bounded since $(1, -\sqrt{2})^\top x \leq 0$, but has no optimal solution.

> **Definition 6.0.2** (convex hull)
>
> The smallest convex set that contains all points in a set. Formally, $\text{conv}(x^1, \ldots, x^k) = \{\sum \lambda_i x^i : \sum \lambda_i = 1, \lambda \geq \mathbb{0}\}$

> **Theorem 6.0.3** (Fundamental Theorem of Integer Programming)
>
> Given a polyhedron $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ with $A, b \in \mathbf{M}(\mathbb{Q})$, the convex hull $\text{conv}(\mathbb{Z}^n \cap P)$ is another polyhedron $Q = \{x \in \mathbb{R}^n : A'x \leq b'\}$ with $A', b' \in \mathbf{M}(\mathbb{Q})$.

This reduces solving an IP to solving a related LP, since the extreme points of $Q$ are just integer points in $P$.

> **Theorem 6.0.4**
>
> Given (IP) $\max\{c^\top x : Ax \leq b : x \in \mathbb{Z}^n\}$, let $\{x : A'x \leq b'\}$ be the convex hull of solutions of (IP). Let (LP) be $\max\{c^\top x : A'x \leq b'\}$. Then:
>
> - (IP) infeasible if and only if (LP) infeasible
> - (IP) unbounded if and only if (LP) unbounded
> - Every optimal solution of (IP) is an optimal solution of (LP)
> - Every optimal solution of (LP) that is an extreme point is an optimal solution of (IP)

Not a great idea in general because finding the convex hull is hard.

We need to find feasible solutions in order to define the convex hull anyways and the number of constraints to define convex hull is $O(2^n)$ in the original constraints.

There is no finite convex hull if coefficients are irrational.

---

**Definition 6.0.5** (LP relaxation)

An IP without the integer constraint.

---

**Example 6.0.6.** Find $\max\left\{\begin{pmatrix} 0 & 1 \\ 3 & -1 \\ -5.5 & -4 \end{pmatrix} x \leq \begin{pmatrix} 5.5 \\ 10.5 \\ -22 \end{pmatrix} : x \in \mathbb{Z}^3 \right\}$

---

*Solution.* The optimal solution of the relaxation is $(\frac{128}{35}, \frac{33}{70})^\top$.

To make a cutting plane, need (1) current non-integral optimal solution outside of halfspace and (2) all integral feasible solutions inside halfspace.

For example, shrink the feasible region with $x_2 \geq 2$ to get new optimal solution $(\frac{28}{11}, 2)^\top$ which is "better"

To find cutting plane, note that simplex BFS $\bar{x} = (\frac{128}{35}, \frac{33}{70}, \frac{176}{35}, 0, 0)^t$ with objective value $-\frac{289}{70}$.

Take floor of coefficients on LHS, change $=$ to $\leq$: $x_1 + \lfloor \frac{8}{35} \rfloor x_4 + \lfloor -\frac{2}{35} \rfloor x_5 = x_1 - x_5 \leq \frac{128}{35}$

Since $x \geq \mathbb{0}$, LHS only decreases and all solutions are preserved.

Take floor of RHS constant: $x_1 - x_5 \leq \lfloor \frac{128}{35} \rfloor \leq 3$.

Since $x$ integral, coefficients integral, we can take floor without losing integral solutions. This also eliminates $\bar{x}$ since it satisfied the original constraint with equality.

Then, we have $x_1 - x_5 \leq 3$ a cutting plane.                                    $\square$

*Lecture 22*
*July 13*

Summarize the algorithm:

---
**Algorithm 2** Cutting plane algorithm
---
1: Relax the IP by removing the integrality constraint
2: Run simplex on the resulting LP to get non-integral BFS $\bar{x}$
3: Select a constraint $\sum a_i x_i = b_j$ where $b_j \notin \mathbb{Z}$
4: Construct cutting plane $\sum \lfloor a_i \rfloor x_i \leq \lfloor b_j \rfloor$
5: Add slack variable to turn into equality and append to LP (will have to run either two-phase simplex or dual simplex to find a new BFS)

---

# Chapter 7

# Solving NLPs

> **Definition 7.0.1** (non-linear program)
>
> $\min\{f(x) : g_i(x) \leq 0\}$ where $f, g_i : \mathbb{R}^n \to \mathbb{R}$

IPs are just NLPs, so general NLP problem is even harder. The feasible regions of NLPs can be disjoint and just all around awful.

To solve, we try to limit functions to being convex. Limiting the feasible region to be convex to removes some of the awfulness. Local optimal solutions of convex regions are still globally optimal. We can apply simplex to some sort of reduction of the region.

We construct halfspaces of non-linear tight constraints based on gradients. For example, $g(x) \leq 0$ becomes $\nabla g(\bar{x}) \leq 0$.

> **Definition 7.0.2** (convex function)
>
> For all $a, b \in \mathbb{R}^n$ and $0 \leq \lambda \leq 1$, $f(\lambda a + (1 - \lambda)b) \leq \lambda f(a) + (1 - \lambda)f(b)$. Equivalently, the line between points $f(a)$ and $f(b)$ lies above the curve.

**Example 7.0.3.** Show $f : \mathbb{R} \to \mathbb{R} : x \mapsto x^2$ is convex.

*Proof.* Let $a, b \in \mathbb{R}$ and $0 \leq \lambda \leq 1$.

Then, $f(\lambda a + (1 - \lambda)b) = (\lambda a + (1 - \lambda)b)^2 = \lambda^2 a^2 + 2\lambda(1 - \lambda)ab + (1 - \lambda)^2 b^2$.

But $(a - b)^2 = a^2 - 2ab + b^2 \geq 0$, so $2ab \geq a^2 + b^2$.

So we have $f(\lambda a + (1 - \lambda)b) \leq \lambda^2 a^2 + \lambda(1 - \lambda)(a^2 + b^2) + (1 - \lambda)^2 b^2 = \lambda a^2 + (1 - \lambda)b^2 = \lambda f(a) + (1 - \lambda)f(b)$ as desired $\qquad \square$

**Lemma 7.0.4**

The following are convex functions:

- $kx^n$ for even $n$ and positive $k$
- Affine functions $c^\top x + \bar{z}$
- Positive linear combinations of convex functions $\sum a_i f_i(x)$ with $a \geq \mathbb{0}$

**Definition 7.0.5** (epigraph)

The set $\{(x, r)^\top : x \in \mathbb{R}^n, r \in \mathbb{R}, r \geq f(x)\}$ of points "above" the function.

**Lemma 7.0.6**

$f$ is convex function if and only if the epigraph of $f$ is a convex set.

Since constraints in NLPs have the form $g_i(x) \leq 0$, we can use convexity.

**Proposition 7.0.7**

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

*Proof.* Let $a, b \in S$ and $0 \leq \lambda \leq 1$. Then $g(a) \leq 0$ and $g(b) \leq 0$. Since $g$ is convex, $g(\lambda a + (1-\lambda)b) = \lambda g(a) + (1-\lambda)g(b) \leq \lambda \cdot 0 + (1-\lambda) \cdot 0 = 0$ so the convex combination is in $S$ and $S$ is convex. $\square$

In a convex NLP (i.e. all constraints $g_i$ are convex), the feasible region is the intersection of convex sets, which is itself convex.

**Definition 7.0.8** (relaxation)

Enlargement of a feasible region, i.e., if $R$ feasible region and $R'$ relaxation, $R \subset R'$.

If optimal solution $\bar{x} \in R'$, then it must be optimal for $R$.

Similar to IPs, we define halfspaces that relax the feasible region.

**Definition 7.0.9** (subgradient)

$s \in \mathbb{R}^n$ normal vector to the boundary at a point, i.e., $g(\bar{x}) + s^\top(x - \bar{x}) \leq g(x)$ for all $x \in \mathbb{R}^n$

Since the LHS is just an affine function, this is a hyperplane beneath $g(x)$ with equality at $x = \bar{x}$

**Definition 7.0.10** (supporting halfspace)

$\{s^\top x \leq 0 : x \in \mathbb{R}^n\} \supseteq \{g(x) \leq 0 : x \in \mathbb{R}^n\}$ for subgradient $s$

Note: Supporting halfspaces only work with convex sets

**Example 7.0.11.** Consider $g : \mathbb{R}^2 \to \mathbb{R} : x \mapsto \frac{1}{4}x_1^2 + \frac{1}{4}x_2^2 - 1$ with $\bar{x} = (2, 0)^\top$.

$s = (2, 0)^\top$ is a (unique) subgradient. The plane is $s^\top(x - \bar{x}) = 2x_1 - 4 = 0$.

**Example 7.0.12.** Consider $g : \mathbb{R}^2 \to \mathbb{R} : x \mapsto |x_1| + |x_2| - 2$ with $\bar{x} = (2, 0)^\top$.

Multiple subgradients exist because $\bar{x}$ is on a sharp edge.

One is $s = (1, 0)^\top$ with $s^\top(x - \bar{x}) = x_1 - 2 = 0$. Another is $s = (1, \frac{1}{2})^t$ with $s^\top(x - \bar{x}) = x_1 + \frac{1}{2}x_2 - 2 = 0$.

In fact, $s = (1, t)^\top$ is a subgradient for $-1 \le t \le 1$.

**Definition 7.0.13** (supporting halfspace)

For convex set $C \subseteq \mathbb{R}^n$ and $\bar{x} \in C$, $F = \{x \in \mathbb{R}^n : s^\top x \le \beta\}$ is a supporting halfspace of $C$ at $\bar{x}$ if $C \subseteq F$ and $s^\top \bar{x} = \beta$ ($\bar{x}$ is on the boundary of the halfspace)

Note: subgradients are normal vectors of supporting halfspaces.

**Theorem 7.0.14**

For convex function $g : \mathbb{R}^n \to \mathbb{R}$, $g(\bar{x}) = 0$ with $s$ a subgradient of $g$ at $\bar{x}$ and $C = \{g(x) \le 0\}$, $F = \{s^\top(x - \bar{x}) \le 0\}$ is a supporting halfspace of $C$ at $\bar{x}$

*Proof.* ($C \subseteq F$) If $g(x') \le 0$, then by definition of subgradients, $\underbrace{g(\bar{x})}_{=0} + s^\top(x' - \bar{x}) \le g(x') \le 0$. That is, $x' \in F$.

(boundary) Plug in $x = \bar{x}$ in $F$, $x^\top(\bar{x} - \bar{x}) = 0$. $\qquad\square$

**Definition 7.0.15** (gradient)

For differentiable function $f : \mathbb{R}^n \to \mathbb{R}$, $\nabla f(x) = (\frac{\partial f}{\partial x_1}, ..., \frac{\partial f}{\partial x_n})^\top$.

**Lemma 7.0.16**

$\nabla f(\bar{x})$ is a subgradient of $f$ at

**Example 7.0.17.** $\min\{-x_1 - x_2 : x_1^2 - 2x_2 - 4 \le 0, -7x_1 + x_2^2 - 8 \le 0, -x_1 - x_2 + 1 \le 0\}$

*Solution.* Propose $\bar{x} = (4, 6)^\top$ is optimal.

Replace tight constraints (1) and (2) by supporting halfspaces:

- $g_1(x) = x_1^2 - 2x_2 - 4$ so $\nabla g_1(x) = (2x_1, -2)^\top$ and $\nabla g_1(\bar{x}) = (8, -2)^\top$.

  Supporting halfspace $8x_1 - 2x_2 - 20 \leq 0$.

- $g_2(x) = -7x_1 + x_2^2 - 8$, so $\nabla g_2(x) = (-7, 2x_2)^\top$ and $\nabla g_1(\bar{x}) = (-7, 12)^\top$.

  Supporting halfspace $-7x_1 + 12x_2 - 44 \leq 0$.

Consider the LP $\max \left\{ (1,1)x : \begin{pmatrix} 8 & -2 \\ -7 & 12 \end{pmatrix} x \leq \begin{pmatrix} 20 \\ 44 \end{pmatrix} \right\}$. Optimal if $(1,1)^\top$ in cone of tight constraints, which it is.

Since LP is a relaxation and it is optimal, $\bar{x}$ is optimal for the original NLP. $\qquad \square$

**Proposition 7.0.18**

For a convex NLP $\min\{c^\top x : g_i(x) \leq 0\}$ and a feasible solution $\bar{x}$, if $-c$ is in the cone generated by the subgradients at $\bar{x}$ of the tight constraints for $\bar{x}$, then $\bar{x}$ is optimal.

The converse is almost true. Need extra condition and differentiability

**Definition 7.0.19** (Slater point)

$x'$ where $g_i(x') < 0$, i.e., points in the interior of feasible region.

**Theorem 7.0.20** (Karush–Kuhn–Tucker)

For NLP $\min\{f(x) : g_i(x) \leq 0\}$ and feasible point $\bar{x}$, if $f$, $g_i$ are all convex and differentiable at $\bar{x}$ and a Slater point exists, then $\bar{x}$ is optimal if and only if $-\nabla f(\bar{x})$ is in the cone generated by $\nabla g_i(\bar{x})$ over tight constraints $g_i$

# List of Named Results

# Index of Defined Terms