# CO 250 - Introduction to Optimization

Cameron Roopnarine

Last updated: March 24, 2020

# Contents

# Chapter 1

# Introduction

2019-09-05

**EXAMPLE 1.0.1** (Manufacturing Tables and Chairs)**.**
PROCESS: raw materials → machine → labour → final products
RULES:

- Company has 30 workers and 40 machines available 40hrs/week.
- Manufacturing a table requires 2 machine-hours and 1 labour-hour.
- Manufacturing a chair requires 1 machine-hours and 3 labour-hours.
- Each manufacturer table yields \$10 of profit and each manufacturer chair yields \$15 of profit.

GOAL: The company wants to prepare a weekly production plan which maximizes total profit.

**Variables**

- $x_1$ := the number of tables manufactured per week
- $x_2$ := the number of chairs manufactured per week

**Objective function**

The total profit per week can be modelled by $10x_1 + 15x_2$, which is what we want to maximize.

**Constraints**

- Machine-hours used per week $\leq$ machine-hours available per week which can be modelled by $2x_1 + x_2 \leqslant 40 \times 40 = 1600$
- Labour-hours used per week $\leq$ labour-hours available per week which can be modelled by $x_1 + 3x_2 \leqslant 30 \times 40 = 1200$

We can then formulate the linear programming (LP) model:

$$\begin{aligned} \text{minimize} \quad & 10x_1 + 15x_2 \\ \text{subject to} \quad & 2x_1 + x_2 \leqslant 1600 \\ & x_1 + 3x_2 \leqslant 1200 \\ & x_1, x_2 \geqslant 0 \end{aligned} \tag{LP}$$

An *optimal* solution to the LP using an algorithm later in in this course is $\boldsymbol{x} := (720, 160)^\top$, which means that we want 720 tables, and 160 chairs.

**EXAMPLE 1.0.2** (A General Production Planning Problem)**.** There are resources $I := \{1, \ldots, m\}$ and products $J := \{1, \ldots, n\}$. There are $b_i$ units of resource $i$ available per week $\forall i \in I$. One unit of product $j$ yields $c_j$ of profit for $\forall j \in J$. Manufacturing one unit of product $j$ requires $a_{ij}$ units of resource $i$. We want to maximize the total profit of this manufacturing process. $x_j$ := amount of product $j$ manufactured per week.

$$\begin{aligned} &\text{maximize} && c_1x_1 + \cdots + c_nx_n = \sum_{j=1}^{n} c_jx_j \\ &\text{subject to} && \sum_{j=1}^{n} a_{ij}x_{ij} \leqslant b_i && \forall i \in \{1,\dots,m\} \\ & && x_j \geqslant 0 && \forall j \in \{1,\dots,n\} \end{aligned} \tag{LP}$$

**REMARK 1.0.3.** If $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ and $\boldsymbol{x} \leqslant \boldsymbol{y}$, then $x_1 \leqslant y_1, \dots, x_n \leqslant y_n$.

**REMARK 1.0.4.**

$$\boldsymbol{c} := \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}, \boldsymbol{x} := \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, A := \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}, \boldsymbol{b} := \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$$

Given $A, \boldsymbol{b}, \boldsymbol{c}$ with $\boldsymbol{x} \in \mathbb{R}^n$ as the variable vector, we realize that $\boldsymbol{c}^\top \boldsymbol{x} = \sum_{j=1}^{n} c_jx_j$ is exactly the model that we wanted to maximize in LP such that it satisfies $A\boldsymbol{x} \leqslant \boldsymbol{b}$, with $\boldsymbol{x} \geqslant \boldsymbol{0}$.

**DEFINITION 1.0.5.** Let $f : \mathbb{R}^n \to \mathbb{R}$. $f$ is an ***affine function*** if $f(\boldsymbol{x}) = \boldsymbol{a}^\top \boldsymbol{x} + \beta$ for some $\boldsymbol{a} \in \mathbb{R}^n$ and $\beta \in \mathbb{R}$.

**DEFINITION 1.0.6.** Let $f : \mathbb{R}^n \to \mathbb{R}$. $f$ is a ***linear function*** if $f$ is an affine function with $\beta = 0$.

**REMARK 1.0.7.** Every linear function is affine, but the converse is not true.

**DEFINITION 1.0.8.** A ***linear constraint*** is one of

$$\begin{aligned} f(\boldsymbol{x}) &\leqslant \beta \\ f(\boldsymbol{x}) &= \beta \\ f(\boldsymbol{x}) &\geqslant \beta \end{aligned}$$

where $f$ is a linear function and $\beta \in \mathbb{R}$.

**DEFINITION 1.0.9.** A ***linear program*** (LP) is the problem of minimizing or maximizing an affine function subject to a finite number of linear constraints.

---

2019-09-10

---

Recall the family of LP problems:

$$\max\{\boldsymbol{c}^\top \boldsymbol{x} : A\boldsymbol{x} \leqslant \boldsymbol{b},\ \boldsymbol{x} \geqslant \boldsymbol{0}\}$$

An assignment of values to all variables such that every constraint is satisfied is called a feasible solution. A feasible region is the set of all feasible solutions. An optimal solution is a feasible solution which has the best possible objective value among all feasible solutions. Note that an optimization problem may have many optimal solutions, but it may have one optimal value.

**EXAMPLE 1.0.10** (Extension of 1.0.1)**.** Suppose an entrepreneur offers at most 500 machine hours/week (rental) at \$2.5/hour. Can we incorporate this new situation into our mathematical model? Can it still be a LP? Yes. $x_3 :=$ the number of machine hours rented from the business person per week.

$$\begin{array}{ll} \text{maximize} & 10x_1 + 15x_2 - 2.5x_3 \\ \text{subject to} & 2x_1 + x_2 \leqslant 1600 + x_3 \\ & x_1 + 3x_2 \leqslant 1200 \\ & x_3 \leqslant 500 \\ & x_1, x_2, x_3 \geqslant 0 \end{array} \tag{LP}$$

**EXAMPLE 1.0.11** (Constraints on Ratios, Percentages and Proportions)**.** Suppose we are required to manufacture at least 10 tables and 80 chairs per week. Also

$$^{\#\text{of tables manufactured}}/_{\text{week}} \big/ ^{\#\text{of chairs manufactured}}/_{\text{week}} \geqslant 6$$

$$\left\{ \begin{array}{r} x_1 \geqslant 10 \\ x_2 \geqslant 80 \\ {}^{x_2}/_{x_1} \geqslant 6 \end{array} \right\} \iff \left\{ \begin{array}{r} x_1 \geqslant 10 \\ x_2 \geqslant 80 \\ x_2 \geqslant 6x_1 \end{array} \right\}$$

In general, suppose $f, g$ are affine functions

$$b_1 \leqslant {}^{f(\boldsymbol{x})}/_{g(\boldsymbol{x})} \leqslant b_2$$

provided that $g(\boldsymbol{x}) > 0$ for every feasible solution $\boldsymbol{x}$ we can equivalently write

$$\begin{aligned} f(\boldsymbol{x}) &\leqslant b_2 g(\boldsymbol{x}) \\ f(\boldsymbol{x}) &\geqslant b_1 g(\boldsymbol{x}) \end{aligned}$$

**EXAMPLE 1.0.12** (Multi-period, Multi-stage Optimization Problems)**.** Consider planning for multiple periods where in each period we want to decide how much to produce, how much to keep in stock (inventory) for the upcoming periods.

Variables

For all $i \in \{1, \dots, T\}$, where $T$ is the last period, we have:

- $s_i :=$ the amount of units sold in period $i$
- $p_i :=$ the amount of units purchased/manufactured of period $i$
- $t_i :=$ the amount of units in stock at the end of period $i$
- $t_0 :=$ the amount of units in stock at the beginning of the first period.

Key Constraints

$$\begin{aligned} p_i + t_{i-1} &= s_i + t_i && \forall i \in \{1, \dots, T\} \\ p_i, s_i, d_i &\geqslant 0 && \forall i \in \{1, \dots, T\} \end{aligned}$$

**REMARK 1.0.13.** Typically we have additional constraints on $s_i$, $p_i$, $t_i$, $t_0$.

**DEFINITION 1.0.14.** An ***integer program*** (IP) is obtained from linear program by requiring a non-empty subset of variables to be integers. If all variables are restricted to be integers, we call the integer program a ***Pure IP***, and if at least some variables may take real values, we call the integer program a ***Mixed IP***.

**REMARK 1.0.15.** In integer programs, some examples of constraints that can be used are:

- $x_i \in \mathbb{Z}$
- $x \in \mathbb{Z}^n$
- $x_i \in \{0, 1\}$
- $x_i$ is an integer
- $x_i \in \{0, 1\}^n$

**EXAMPLE 1.0.16** (Assignment Problem)**.** SPIT has a campus near the North Pole. They have three buildings named A, B, C, which need to be renovated to be served as one of a Library, Laboratory, or Gym (sometimes called functions). Each building must be assigned one activity, and each activity must be assigned one building. Renovation costs in millions of dollars are given:

| | Library | Laboratory | Gym |
|---|---|---|---|
| A | 10 | 60 | 20 |
| B | 60 | 70 | 50 |
| C | 20 | 60 | 40 |

Find an assignment of activities to buildings so that the total renovation cost is minimized.
**Solution.** Let us generalize to $n$ buildings and $n$ activities.

$$x_{ij} := \begin{cases} 1, \text{ if } i \text{ is assigned to activity } j \\ 0, \text{ otherwise} \end{cases} \qquad \forall i, j \in \{1, \dots, n\}$$

$$c_{ij} := \text{renovation cost for assigning activity } j \text{ to building } i$$

$$\begin{aligned} \text{minimize} \quad & \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} x_{ij} \\ \text{subject to} \quad & \sum_{i=1}^{n} x_{ij} = 1 \qquad \forall j \in \{1, \dots, n\} \\ & \sum_{j=1}^{n} x_{ij} = 1 \qquad \forall i \in \{1, \dots, n\} \\ & x_{ij} = \{0, 1\} \qquad \forall i, j \in \{1, \dots, n\} \end{aligned} \tag{IP}$$

- First constraint $\Longrightarrow$ every activity is assigned exactly one building
- Second constraint $\Longrightarrow$ every building is assigned exactly one activity
- Third constraint $\Longrightarrow$ $x_{ij}$ is a ***binary*** variable that takes values only $1$ or $0$. If we wanted an IP formulation, we would remove the constraint $x_{ij} = \{0, 1\}$ and add: $x_{ij} \geqslant 0$, $x_{ij} \leqslant 1$ and $x_{ij}$ integer.

Suppose $c_{ij} \in \mathbb{R}$ and consider the inequality version (if we don't assign **exactly** one item to another):

$$\begin{aligned}
\text{minimize} \quad & \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} x_{ij} & & \\
\text{subject to} \quad & \sum_{i=1}^{n} x_{ij} \leqslant 1 & & \forall j \in \{1, \dots, n\} \\
& \sum_{j=1}^{n} x_{ij} \leqslant 1 & & \forall i \in \{1, \dots, n\} \\
& x_{ij} = \{0, 1\} & & \forall i, j \in \{1, \dots, n\}
\end{aligned} \tag{IP}$$

We can generalize this class optimization problem further.

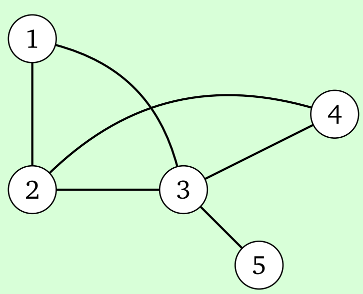
# Chapter 2

# Introduction to Graphs

**DEFINITION 2.0.1.** An ***undirected graph*** is a pair $G = (V, E)$, where $V$ is a finite set of elements called ***vertices***, and $E$ is a set of pairs of distinct vertices called ***edges***. All edges in an undirected graph are bidirectional.

**DEFINITION 2.0.2.** Let $G = (V, E)$ be a graph. Suppose $uv \in E$. $u$, $v$ are ***adjacent*** vertices. $u$, $v$ are the ***endpoints*** of the edge $uv$. The edge $uv$ is ***incident*** to vertices $u$ and $v$.

**EXAMPLE 2.0.3** (Undirected Graph)**.** Given $G :=$




we have

$$V = \{1, \ldots, 5\}$$
$$E = \{12, 13, 23, 24, 35, 34\}$$

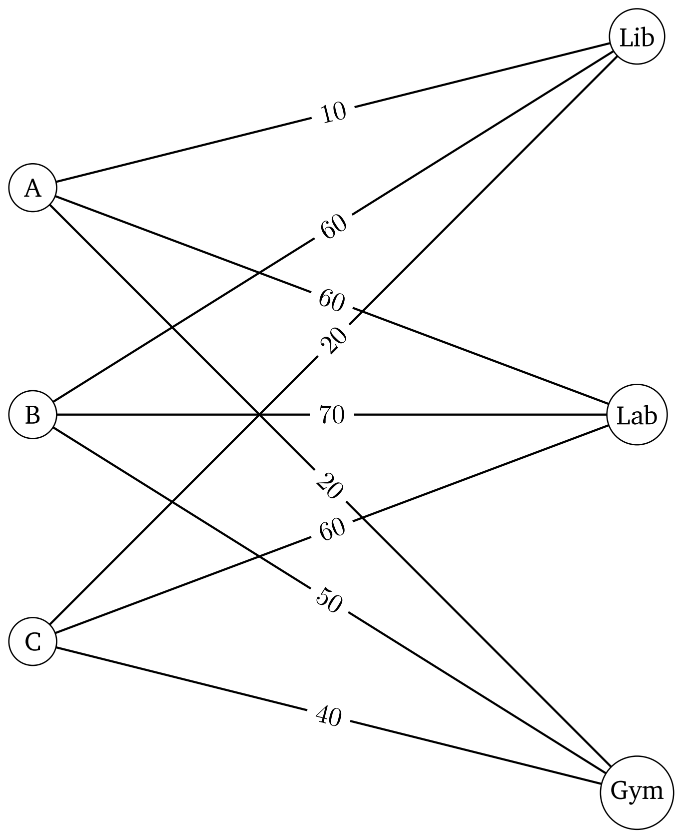
**DEFINITION 2.0.4.** Given a graph $G = (V, E)$, a ***matching*** $M$ in $G$ is a subset of edges in $G$ such that no two edges in $M$ share a common vertex.

**EXAMPLE 2.0.5** (Matching)**.** In the above example:

| Matching | Not a Matching |
|---|---|
| $M := \{12\}$ | $M := \{12, 25\}$ |
| $M := \emptyset$ | $M := \{67\}$ |
| $M := \{12, 35\}$ | |

**DEFINITION 2.0.6.** Given a graph $G = (V, E)$, if every vertex in $V$ of $G$ is an endpoint of an edge in $M$, we call the matching a ***perfect matching***.

The assignment problem is a special case of a *minimum cost perfect matching problem* or weighted graphs (in this case every edge is given a weight/cost $c_{ij}$)



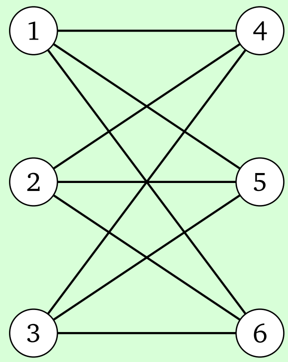

**REMARK 2.0.7.** In a perfect matching graph, there are $n^2$ edges, and $2n$ (an even number of) vertices.

---

2019-09-17

---

**EXAMPLE 2.0.8** (Minimum-Cost Perfect Matching Problem). Given an undirected graph $G = (V, E)$, and $c_e \in \mathbb{R}$, for every $e \in E$, we want to find a perfect matching in G with minimum total cost. The cost of matching $M$ is $\sum_{e \in M} c_e$. For each $v \in V$, $\delta(v) :=$ the set of edges incident to $v$. $G :=$




Examples of $\delta(v)$ in $G$:

- $\delta(1) = \{14, 15, 16\}$
- $\delta(5) = \{15, 25, 35\}$

$$x_e := \begin{cases} 1 & \text{if } e \text{ is chosen in the matching} \\ 0 & \text{otherwise} \end{cases}$$

$$\begin{aligned} \text{minimize} \quad & \sum_{e \in E} c_e x_e \\ \text{subject to} \quad & \sum_{e \in \delta(v)} x_e = 1 && \forall v \in V \\ & x_e \in \{0, 1\} && \forall e \in E \end{aligned} \tag{IP}$$

**DEFINITION 2.0.9.** A graph $G = (V, E)$ is ***bipartite*** if there exists a partition $V_1, V_2$ of $V$ where $V_1 \cup V_2 = V$ and $V_1 \cap V_2 = \emptyset$ such that

$$E \subseteq \{uv \mid u \in V_1,\, v \in V_2\}$$

Assignment problems are a special case of minimum cost perfect matching problems in bipartite graphs.

**THEOREM 2.0.10.** *A graph is bipartite if and only if it does not contain an odd cycle.*

*Proof.* Done in MATH 239. □

Given a situation where we have binary-valued variables

$$x_j := \begin{cases} 1 & \text{option } j \text{ is chosen} \\ 0 & \text{otherwise} \end{cases} \forall j \in \{1, \ldots, n\}$$

We solve how to formulate in an IP in the following conditions:

- **at most** $k$ options are chosen: $\sum_{j=1}^{n} x_j \leqslant k$
- **at least** $k$ options are chosen: $\sum_{j=1}^{n} x_j \geqslant k$

- **exactly** $k$ options are chosen: $\sum_{j=1}^{n} x_j = k$

We can also formulate many classes of the "OR" type constraint in IP problems.

**EXAMPLE 2.0.11** (Extension of 1.0.1).

$$\begin{array}{ll} \text{minimize} & 10x_1 + 15x_2 \\ \text{subject to} & 2x_1 + x_2 \leqslant 1600 \\ & x_1 + 3x_2 \leqslant 1200 \\ & x_1, x_2 \geqslant 0 \end{array} \tag{LP}$$

Suppose C&O is required to produce at least 10 tables per week or at least 80 chairs per week, or possibly both. $x_1 \geqslant 10$ or $x_2 \geqslant 80$ or both. We introduce a new binary-valued variable $z \in \{0, 1\}$.

$$z := \begin{cases} 1 & \text{if } x_1 \geqslant 10 \\ 0 & \text{if } x_2 \geqslant 80 \end{cases}$$

$$\{(x_1 \geqslant 10 \text{ OR } x_2 \geqslant 80) \text{ AND } (x_1 \geqslant 0 \text{ OR } x_2 \geqslant 0)\} \iff \left\{ \begin{array}{r} x_1 \geqslant 10z \\ x_2 \geqslant 80(1-z) \\ z \in \{0,1\} \\ x_1, x_2 \geqslant 0 \end{array} \right\}$$

**REMARK 2.0.12.** *Possibly both* means that you can choose either one of these conditions in the first OR above and it will be correct.

**EXAMPLE 2.0.13** (Extension of 2.0.11). Now, suppose C&O has a new condition every week. We must manufacture either exactly 3 chairs for every table or exactly 8 chairs for every table. Show how to incorporate this in an IP formulation

$$\{x_2 = 3x_1 \text{ OR } x_2 = 8x_1\} \iff \{(x_2 \leqslant 3x_1 \text{ AND } x_2 \geqslant 3x_1) \text{ OR } (x_2 \leqslant 8x_1 \text{ AND } x_2 \geqslant 8x_1)\}$$

Introduce a new binary-valued variable $z \in \{0, 1\}$.

$$z := \begin{cases} 1 & \text{if } x_2 = 3x_1 \\ 0 & \text{if } x_2 = 8x_1 \end{cases}$$

Existing constraints:

$$\left\{ \begin{array}{r} 2x_1 + x_2 \leqslant 1600 \\ x_1 + 3x_2 \leqslant 1200 \\ x_1, x_2 \geqslant 0 \end{array} \right\} \begin{array}{l} \implies x_1 \in [0, 800] \\ \implies x_2 \in [0, 500] \\ \phantom{x} \end{array}$$

So,

$$\begin{aligned} x_2 &\leqslant 3x_1 + 400(1-z) \\ x_2 &\geqslant 3x_1 - 2400(1-z) \\ x_2 &\leqslant 8x_1 + 400z \\ x_2 &\geqslant 8x_1 - 6400z \\ z &\in \{0,1\} \end{aligned}$$

**DEFINITION 2.0.14.** A *non-linear program* has the form

$$\min f(\boldsymbol{x})$$

subject to

$$\begin{aligned} g_1(\boldsymbol{x}) &\leqslant 0 \\ g_2(\boldsymbol{x}) &\leqslant 0 \\ &\vdots \\ g_m(\boldsymbol{x}) &\leqslant 0 \end{aligned}$$

where $f : \mathbb{R}^n \to \mathbb{R}$, $g_i : \mathbb{R}^n \to \mathbb{R}$, $\forall i \in \{1, \ldots, m\}$.

Every LP problem is a very special case of a NLP problem. IP problems can also be formulated as NLP problems.

**EXAMPLE 2.0.15** (Formulating LP Problems as NLP Problems)**.**

$$\begin{aligned} x_i \in \mathbb{Z} &\iff \sin(\pi x_i) = 0 \\ &\iff [\sin(\pi x_i)]^2 \leqslant 0 \end{aligned}$$

$$\begin{aligned} x_i \in \{0,1\} &\iff x_i(1 - x_i) = 0 \\ &\iff x_i^2(1 - x_i)^2 \leqslant 0 \end{aligned}$$

NLP problems have huge modelling power, as a result, one must understand the structure of the underlying problem and construct "good" NLP models that are amendable to analysis and solution techniques.

**EXAMPLE 2.0.16** (Portfolio Optimization)**.** There are $n$ stocks $1, \ldots, n$ to invest in. We have a budget of $B$ dollars. We have an expected return (for \$1 investment at the end of our planning horizon) of $\mu_1, \ldots, \mu_n$. We are also given $V \in \mathbb{R}^{n \times n}$, a variance covariance matrix so that if we invest in $x_1, \ldots, x_n$ dollars in $n$ stocks, $1, \ldots, n$ respectively, then the expected risk of such an investment is given by $\boldsymbol{x}^\top V \boldsymbol{x}$.

$$\sum_{i=1}^{n} \sum_{j=1}^{n} V_{ij} x_i x_j$$

$x_j :=$ amount of investment in stock $j$ in dollars.
Suppose we are also given a goal $G$ (a dollar amount we want as the value of our portfolio at the end of the planning horizon).
**Data**

- Budget (\$) $\to$ B
- Goal (\$) $\to$ G
- Expected return $\to (\mu_1, \ldots, \mu_n)^\top$
- Variance-covariance matrix $\to V \in \mathbb{R}^{n \times n}$

We want to minimize the risk of our portfolio while satisfying the budget and the goal constraints.
(NLP)

$$\min \boldsymbol{x}^\top V \boldsymbol{x}$$

subject to

$$\sum_{j=1}^{n} x_j \leqslant B$$
$$\sum_{j=1}^{n} \mu_j x_j \geqslant G$$
$$\boldsymbol{x} \geqslant \mathbf{0}$$

There are many variants of such models and extensions. For example, instead of a goal $G$, we may given an upper bound on the risk, say $R \in \mathbb{R}_{>0}$.
(NLP)

$$\max \sum_{j=1}^{n} \mu_j x_j$$

subject to

$$\sum_{j=1}^{n} x_j \leqslant B$$
$$\boldsymbol{x}^\top V \boldsymbol{x} \leqslant R$$
$$\boldsymbol{x} \geqslant \mathbf{0}$$

We can handle many more variants and extensions. Suppose investing in stock $j$ below $\ell_j$ dollars is not allowed. For diversity of our portfolio, we want to invest in at least 20 stocks, and for the sake of simplicity we want to invest in at most 150 stocks. We introduce a binary-valued variable $z_j$.

$$z_j := \begin{cases} 1, \text{ if we invest in stock } j \\ 0, \text{ otherwise} \end{cases} \qquad \forall j \in \{1, \ldots, n\}$$

Add these constraints:

$$\ell_j z_j \leqslant x_j \leqslant B z_j \qquad \forall j \in \{1, \ldots, n\}$$
$$20 \leqslant \sum_{j=1}^{n} z_j \leqslant 150 \qquad \forall j \in \{1, \ldots, n\}$$
$$z_j \in \{0, 1\} \qquad \forall j \in \{1, \ldots, n\}$$

# Chapter 3

# Solving Linear Programs

2019-09-24

## 3.1 Possible Outcomes

**DEFINITION 3.1.1.** Consider an LP with variables $x_1, \dots, x_n$. Then the assignment of values to all variables such that all constraints are satisfied, gives a *feasible solution*.
An optimization problem is called *feasible* if it has at least one feasible solution, otherwise it is called *infeasible*.

### 3.1.1 Infeasible Linear Programs

**EXAMPLE 3.1.2** (Infeasible LP). (LP)

$$\max x_1 + 2x_2 + 3x_3 + 4x_4 + 5x_5$$

subject to

$$\begin{array}{r} 1 \\ -2 \end{array} \underbrace{\begin{bmatrix} -3 & 2 & 7 & 1 & -7 \\ -2 & 1 & 2 & 0 & -4 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix}}_{\boldsymbol{x}} = \underbrace{\begin{bmatrix} 6 \\ 4 \end{bmatrix}}_{\boldsymbol{b}}$$

$$\boldsymbol{x} \geqslant \boldsymbol{0}$$

Let $\boldsymbol{y} := (1, -2)^\top$ and consider the facts

$$\begin{aligned} & A\boldsymbol{x} = \boldsymbol{b} \\ \implies & \boldsymbol{y}^\top A\boldsymbol{x} = \boldsymbol{y}^\top \boldsymbol{b} \\ \implies & \underbrace{\begin{bmatrix} 1 & 0 & 3 & 1 & 1 \end{bmatrix}}_{\geqslant \boldsymbol{0}^\top} \underbrace{\boldsymbol{x}}_{\geqslant \boldsymbol{0}} = \underbrace{6 - 8}_{<0} = -2 \end{aligned}$$

Therefore, since $\nexists$ any solution to $A\boldsymbol{x} = \boldsymbol{b}$, $\boldsymbol{x} \geqslant 0$ the LP is infeasible.

**THEOREM 3.1.3** (Infeasibility). *If* $\exists \boldsymbol{y} \in \mathbb{R}^m$ *such that*
*(1)* $\boldsymbol{y}^\top A \geqslant \boldsymbol{0}^\top$
*(2)* $\boldsymbol{y}^\top \boldsymbol{b} < 0$
*then, the LP*

$$\max\{\boldsymbol{c}^\top \boldsymbol{x} : A\boldsymbol{x} = \boldsymbol{b},\, \boldsymbol{x} \geqslant \boldsymbol{0}\}$$

*is infeasible. In particular, we call a vector* $\boldsymbol{y}$ *a* certificate of infeasibility.

*Proof.* Suppose there exists such a $\boldsymbol{y}$. Suppose for a contradiction that $\exists \bar{\boldsymbol{x}} \in \mathbb{R}^n$ (there is a feasible solution) such that

$$A\bar{\boldsymbol{x}} = \boldsymbol{b}, \bar{\boldsymbol{x}} \geqslant \boldsymbol{0}$$
$$A\bar{\boldsymbol{x}} = \boldsymbol{b} \implies \underbrace{\boldsymbol{y}^\top A}_{\geqslant \boldsymbol{0}^\top} \underbrace{\bar{\boldsymbol{x}}}_{\geqslant \boldsymbol{0}} = \underbrace{\boldsymbol{y}^\top \boldsymbol{b}}_{\not< 0}$$

a contradiction to (2). □

An optimization problem is called unbounded if $\forall M \in \mathbb{R}$, there exists a feasible solution of the optimization problem with the objective value strictly better than $M$.

### 3.1.2 Unbounded Linear Programs

**EXAMPLE 3.1.4** (Unbounded LP).

$$\max \begin{bmatrix} -1 & 3 & 0 & 0 & 1 \end{bmatrix} \boldsymbol{x}$$

subject to

$$\begin{bmatrix} -1 & 3 & -1 & 1 & 0 \\ -2 & 4 & 1 & 0 & 1 \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$
$$\boldsymbol{x} \geqslant \boldsymbol{0}$$

Consider

$$\tilde{\boldsymbol{x}} := \underbrace{\begin{bmatrix} 0 \\ 0 \\ 0 \\ 2 \\ 1 \end{bmatrix}}_{\boldsymbol{x}} + t \underbrace{\begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 2 \end{bmatrix}}_{\boldsymbol{d}}, \; t \geqslant 0$$

$$A\boldsymbol{x} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \bar{\boldsymbol{x}} \geqslant \boldsymbol{0}. \text{Therefore } \bar{\boldsymbol{x}} \text{ is a feasible solution.}$$

$$A\boldsymbol{d} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \boldsymbol{d} \geqslant \boldsymbol{0}.$$

$$A\tilde{\boldsymbol{x}} = A(\bar{\boldsymbol{x}} + t\boldsymbol{d}) = A\bar{\boldsymbol{x}} + t(A\boldsymbol{d}) = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$
$$\tilde{\boldsymbol{x}} = \bar{\boldsymbol{x}} + t\boldsymbol{d}$$

Therefore, $\tilde{\boldsymbol{x}}$ is a feasible solution $\forall t \geqslant 0$.
**Objective function value of $\tilde{\boldsymbol{x}}$:**

$$\begin{bmatrix} -1 & 3 & 0 & 0 & 1 \end{bmatrix} \left( \begin{bmatrix} 0 \\ 0 \\ 0 \\ 2 \\ 1 \end{bmatrix} + t \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 2 \end{bmatrix} \right) = 1 + t(-1+2) = 1 + t \to +\infty \text{ as } t \to +\infty$$

Therefore the LP is unbounded.

**THEOREM 3.1.5.** *If* $\exists \bar{\boldsymbol{x}} \in \mathbb{R}^n$ *such that*

$$A\bar{\boldsymbol{x}} = \boldsymbol{b}, \boldsymbol{x} \geqslant \boldsymbol{0}.$$

*and* $\exists \boldsymbol{d} \in \mathbb{R}^n$ *such that*
*(1)* $A\boldsymbol{d} = \boldsymbol{0}$
*(2)* $\boldsymbol{d} \geqslant \boldsymbol{0}$
*(3)* $\boldsymbol{c}^\top \boldsymbol{d} > 0$
*then, the LP*

$$\max\{\boldsymbol{c}^\top \boldsymbol{x} : A\boldsymbol{x} = \boldsymbol{b},\ \boldsymbol{x} \geqslant \boldsymbol{0}\}$$

*is unbounded. In particular, we call a pair of vectors* $\bar{\boldsymbol{x}}$*,* $\boldsymbol{d}$ *a* certificate of unboundedness.

*Proof.* Suppose there exists such $\boldsymbol{d}$. Consider

$$\tilde{\boldsymbol{x}} = \bar{\boldsymbol{x}} + t\boldsymbol{d}, t \geqslant 0$$

Then,

$$A\tilde{\boldsymbol{x}} = \underbrace{A\bar{\boldsymbol{x}}}_{\boldsymbol{b}} + t\underbrace{(A\boldsymbol{d})}_{\boldsymbol{0}} = \boldsymbol{b}$$

Therefore $\tilde{\boldsymbol{x}}$ is a feasible solution of the LP, $t \geqslant 0$. The objective value of the function is

$$\boldsymbol{c}^\top \tilde{\boldsymbol{x}} = \boldsymbol{c}^\top \bar{\boldsymbol{x}} + t\underbrace{(\boldsymbol{c}^\top \boldsymbol{d})}_{>\boldsymbol{0}} \to +\infty \text{ as } t \to +\infty$$

Therefore, the LP is unbounded. □

**REMARK 3.1.6.** If the LP is $\min$, then flip the equality for (3).

### 3.1.3 Linear Programs with Optimal Solutions

**EXAMPLE 3.1.7** (Optimal LP)**.**

$$\max 10x_1 + 15x_2$$

subject to

$$\begin{aligned} 2x_1 + x_2 + x_3 &= 1600 \\ x_1 + 3x_2 + x_4 &= 1200 \end{aligned}$$

$$\boldsymbol{x} \geqslant \boldsymbol{0}$$

Consider $\bar{\boldsymbol{x}} := (720, 160, 0, 0)^\top$ and $\boldsymbol{y} := (3, 4)^\top$.
Note that $A\bar{\boldsymbol{x}} = \boldsymbol{b}$, with $\bar{\boldsymbol{x}} \geqslant \boldsymbol{0}$, so $\bar{\boldsymbol{x}}$ is a feasible solution.
Also, $\boldsymbol{c}^\top \bar{\boldsymbol{x}} = 7200 + 2400 = 9600$. Every feasible solution satisfies

$$\begin{aligned} &A\boldsymbol{x} = \boldsymbol{b} \\ \implies &\boldsymbol{y}^\top A\boldsymbol{x} = \boldsymbol{y}^\top \boldsymbol{b} \end{aligned}$$

$$\boldsymbol{y}^\top A = \begin{bmatrix} 10 & 15 & 3 & 4 \end{bmatrix} \geqslant \begin{bmatrix} 10 & 15 & 0 & 0 \end{bmatrix} = \boldsymbol{c}^\top$$

$$\boldsymbol{y}^\top \boldsymbol{b} = 3 \times 1600 + 4 \times 1200 = 9600 = \boldsymbol{c}^\top \bar{\boldsymbol{x}}$$

Therefore $\bar{\boldsymbol{x}}$ is an optimal solution.

2019-09-26

**Summary of Outcomes**
(P)

$$\max\{\boldsymbol{c}^\top \boldsymbol{x} : A\boldsymbol{x} = \boldsymbol{b},\ \boldsymbol{x} \geqslant \boldsymbol{0}\}$$

- If there exists a vector $\boldsymbol{y}$ such that
  (1) $\boldsymbol{y}^\top A \geqslant \boldsymbol{0}^\top$
  (2) $\boldsymbol{y}^\top \boldsymbol{b} < 0$
  then (P) is infeasible. We call $\boldsymbol{y}$ a certificate of infeasibility.
- If there exists a feasible solution $\bar{\boldsymbol{x}}$ and a vector $\boldsymbol{d}$ such that:
  (1) $A\boldsymbol{d} = \boldsymbol{0}$
  (2) $\boldsymbol{d} \geqslant \boldsymbol{0}$
  (3) $\boldsymbol{c}^\top \boldsymbol{d} > 0$
  then (P) is unbounded. We call a pair of vectors $\bar{\boldsymbol{x}}, \boldsymbol{d}$ a certificate of unboundedness.
- If there exists a feasible solution $\bar{\boldsymbol{x}}$ and a vector $\bar{\boldsymbol{y}}$ such that:
  (1) $A^\top \bar{\boldsymbol{y}} \geqslant \boldsymbol{c}$
  (2) $\boldsymbol{c}^\top \bar{\boldsymbol{x}} = \bar{\boldsymbol{y}}^\top \boldsymbol{b}$
  then $\bar{\boldsymbol{x}}$ is an optimal solution of (P). We call $\bar{\boldsymbol{y}}$ a certificate of optimality.

## 3.2 Standard Equality Form

**DEFINITION 3.2.1.** An LP is said to be in *Standard Equality Form* (SEF) if it has the Form

$$\max\{\boldsymbol{c}^\top \boldsymbol{x} + \bar{z} : A\boldsymbol{x} = \boldsymbol{b},\ \boldsymbol{x} \geqslant \boldsymbol{0}\}$$

where $\bar{z}$ is a constant. In other words, it satisfies all of the conditions:
(1) It is a maximization problem
(2) All constraints are equations (other than non-negativity constraints)
(3) Every variable has a non-negativity constraint

Every LP can be converted to SEF. A pair of LP problems LP1 and LP2 are equivalent if they both have the same status (infeasible, unbounded, or optimal) and certificate of such a status for one problem can easily be converted into a certificate of the same type for the other LP.

Given an arbitrary LP problem,

- if the objective function is a minimization problem, then $\min \boldsymbol{c}^\top \boldsymbol{x} \to -(\max -\boldsymbol{c}^\top \boldsymbol{x})$

  **REMARK 3.2.2.** We often omit one negative sign from a TA on Piazza: "It's more just a convention of not putting $-$ before max when doing this and it's understood that the optimal value of one is the negative of the optimal value of the other"

- if there are constraints $\alpha \boldsymbol{x} \leqslant \alpha$, introduce a new non-negative *slack variable* $x_{n+1}$, $x_{n+1} \geqslant 0$.

- if some $x_j$ has no constraint on it, such variables are called *free variables* and we represent that free variable as a difference of two non-negative variables, $x_j = x_j^+ - x_j^-$, $x_j^+ \geqslant 0$, $x_j^- \geqslant 0$.

- if some $x_j < 0$ flip all signs correlating to $x_j$

**EXAMPLE 3.2.3** (Converting an LP to SEF)**.** (P)

$$\max 100x_1 + 200x_2$$

subject to

$$\begin{array}{rcrcl} x_1 & + & 2x_2 & \leqslant & 20 \\ 3x_1 & + & 4x_2 & \geqslant & 10 \end{array}$$

$$x_1 \geqslant 0$$

Converting into SEF we get (P′):

$$\max 100x_1 + 200(x_2^+ - x_2^-)$$

subject to

$$\begin{array}{rcrcrcrcl} x_1 & + & 2(x_2^+ - x_2^-) & + & x_3 & & & = & 20 \\ 3x_1 & + & 4(x_2^+ - x_2^-) & - & & & x_4 & = & 10 \end{array}$$

$$x_1, x_2^+, x_2^-, x_3, x_4 \geqslant 0$$

(P) and (P′) are equivalent.
Let $(\bar{x}_1, \bar{x}_2^+, \bar{x}_2^-, \bar{x}_3, \bar{x}_4)^\top$ be a feasible solution of (P′).
If

$$\hat{x}_1 := \bar{x}_1$$
$$\hat{x}_2 := \bar{x}_2^+ - \bar{x}_2^-$$

Then $(\hat{x}_1, \hat{x}_2)^\top$ is a feasible solution of (P).
Let $(\bar{x}_1, \bar{x}_2)^\top$ be a feasible solution of (P).
If

$$\bar{x}_3 := 20 - \bar{x}_1 - 2\bar{x}_2$$
$$\bar{x}_4 := 3\bar{x}_1 + 4\bar{x}_2 - 10$$

and if $\bar{x}_2 \geqslant 0$

$$\bar{x}_2^+ := \bar{x}_2$$
$$\bar{x}_2^- := 0$$

or $\bar{x}_2 < 0$

$$\bar{x}_2^+ := 0$$
$$\bar{x}_2^- := -\bar{x}_2$$

then $(\bar{x}_1, \bar{x}_2^+, \bar{x}_2^-, \bar{x}_3, \bar{x}_4)^\top$ is a feasible solution of (P′).

## 3.3 Bases and Caonical Forms

### 3.3.1 Bases

**DEFINITION 3.3.1.** Let $A \in \mathbb{R}^{m \times n}$, $B \subseteq \{1, \ldots, n\}$ such that $|B| = m$. If

$$A_B := \left[\begin{array}{c|c} a_i & i \in B \end{array}\right] \in \mathbb{R}^{m \times m}$$

where $A_B$ is non-singular (i.e. IMT holds), then $B$ is a *basis* of $A$. If $B$ is a basis of $A$, then $A_B$ is a basis for $\mathbb{R}^m$. We denote $N$ as the set that does not have the elements of $B$.

**DEFINITION 3.3.2.** A vector $\bar{\boldsymbol{x}}$ is a *basic solution* of $A\boldsymbol{x} = \boldsymbol{b}$ for a basis $B$ of $A$ if:
(1) $A\bar{\boldsymbol{x}} = \boldsymbol{b}$
(2) $\bar{\boldsymbol{x}}_{\boldsymbol{N}} = \boldsymbol{0}$
where $\bar{\boldsymbol{x}}_{\boldsymbol{N}}$ is the vector formed by the non-basic variables. That is, $N := \{1, \dots, n\} \setminus B$.

**DEFINITION 3.3.3.** A vector $\bar{\boldsymbol{x}}$ is a *basic feasible solution* of $\{A\boldsymbol{x} = \boldsymbol{b},\ \boldsymbol{x} \geqslant \boldsymbol{0}\}$ if it is a basic solution of $A\boldsymbol{x} = \boldsymbol{b}$ determined by a basis $B$ of $A$ that also satisfies $\bar{\boldsymbol{x}} \geqslant \boldsymbol{0}$. Thus $\bar{\boldsymbol{x}}$ satisfies $A\bar{\boldsymbol{x}} = \boldsymbol{b}$, $\bar{\boldsymbol{x}}_{\boldsymbol{N}} = \boldsymbol{0}$, and $\bar{\boldsymbol{x}} \geqslant \boldsymbol{0}$.

**EXAMPLE 3.3.4** (Bases of A)**.**

$$A := \begin{bmatrix} 1 & 0 & 2 & -1 & 1 \\ 0 & 1 & 2 & -1 & 2 \end{bmatrix}, \boldsymbol{b} := \begin{bmatrix} 2 \\ 5 \end{bmatrix}$$

Bases of A: $\{1, 2\}$, $\{2, 3\}$, $\{1, 4\}$.
Not a bases of A: $\emptyset$, $\{1\}$, $\{1, 2, 3\}$, $\{3, 4\}$.
To find the basic solution determined by $B := \{1, 4\}$, solve

$$\begin{bmatrix} 1 & -1 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_4 \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \end{bmatrix}$$

and we get $\bar{\boldsymbol{x}} = (-3, 0, 0, -5, 0)^\top$.

## 2019-10-01

Let $A \in \mathbb{R}^{m \times n}, \boldsymbol{b} \in \mathbb{R}^m, \boldsymbol{c} \in \mathbb{R}^n$. Consider (P)

$$\max \boldsymbol{c}^\top \boldsymbol{x}$$

subject to

$$A\boldsymbol{x} = \boldsymbol{b}$$

$$\boldsymbol{x} \geqslant \boldsymbol{0}$$

Suppose we are given $\tilde{\boldsymbol{x}} \in \mathbb{R}^n$ such that, $A\tilde{\boldsymbol{x}} = \boldsymbol{b}$, $\tilde{\boldsymbol{x}} \geqslant \boldsymbol{0}$ and $\boldsymbol{y} \in \mathbb{R}^m$ such that $A^\top \boldsymbol{y} \geqslant \boldsymbol{c}$, $\boldsymbol{y}^\top \boldsymbol{b} = \boldsymbol{c}^\top \bar{\boldsymbol{x}}$ with objective function value $= \boldsymbol{c}^\top \bar{\boldsymbol{x}}$.

Computing $\boldsymbol{c}^\top \bar{\boldsymbol{x}}$ we get

$$\begin{aligned} \boldsymbol{c}^\top \bar{\boldsymbol{x}} &= \boldsymbol{y}^\top \boldsymbol{b} \\ &= \boldsymbol{y}^\top (A\tilde{\boldsymbol{x}}) \\ &= \underbrace{(\boldsymbol{y}^\top A)}_{\geqslant \boldsymbol{c}^\top} \underbrace{\tilde{\boldsymbol{x}}}_{\geqslant \boldsymbol{0}} \\ &\geqslant \boldsymbol{c}^\top \tilde{\boldsymbol{x}} \end{aligned}$$

Since $\bar{\boldsymbol{x}}$ achieves the objective value of $\boldsymbol{c}^\top \bar{\boldsymbol{x}}$ and for every feasible solution the objective value is at most $\boldsymbol{c}^\top \bar{\boldsymbol{x}}$, $\bar{\boldsymbol{x}}$ is an optimal solution of (P).

### 3.3.2 Canonical Forms

**DEFINITION 3.3.5.** Consider the following LP in SEF:
(P)

$$\max \boldsymbol{c}^\top \boldsymbol{x} + \bar{z}$$

subject to

$$\begin{aligned} A\boldsymbol{x} &= \boldsymbol{b} \\ \boldsymbol{x} &\geqslant \boldsymbol{0} \end{aligned}$$

We say (P) is in *canonical form* for a basis $B$ of $A$ if
(C1) $A_B$ is an identity matrix
(C2) $\boldsymbol{c_B} = \boldsymbol{0}$

Now,

$$\begin{aligned} A\boldsymbol{x} &= \sum_{j=1}^{n} \boldsymbol{a_j} x_j \\ &= \sum_{j\in B}^{n} \boldsymbol{a_j} x_j + \sum_{j\in N}^{n} \boldsymbol{a_j} x_j \\ &= A_B \boldsymbol{x_B} + A_N \boldsymbol{x_N} \end{aligned}$$

Since $B$ is a basis of $A$, $A_B$ is non-singular,

$$\begin{aligned} & A\boldsymbol{x} = \boldsymbol{b} \\ \iff\ & A_B^{-1} A\boldsymbol{x} = A_B^{-1}\boldsymbol{b} \\ \iff\ & A_B^{-1}(A_B\boldsymbol{x_B} + A_N\boldsymbol{x_N}) = A_B^{-1}\boldsymbol{b} \\ \iff\ & (\underbrace{A_B^{-1}A_B}_{I}\boldsymbol{x_B}) + (A_B^{-1}A_N\boldsymbol{x_N}) = A_B^{-1}\boldsymbol{b} \\ \iff\ & \boldsymbol{x_B} = A_B^{-1}\boldsymbol{b} - (A_B^{-1}A_N\boldsymbol{x_N}) \end{aligned}$$

Consider (C2). For any $\boldsymbol{y} := (y_1, \ldots, y_m)^\top$ the equation

$$\boldsymbol{y}^\top A\boldsymbol{x} = \boldsymbol{y}^\top \boldsymbol{b}$$

can be written as

$$0 = \boldsymbol{y}^\top \boldsymbol{b} - \boldsymbol{y}^\top A\boldsymbol{x}$$

Since this equation holds for every feasible solution, we can add this constraint to the objective function which is now:

$$\max \boldsymbol{c}^\top \boldsymbol{x} + \bar{z} + \boldsymbol{y}^\top \boldsymbol{b} - \boldsymbol{y}^\top A\boldsymbol{x} \implies \max(\boldsymbol{c}^\top - \boldsymbol{y}^\top A)\boldsymbol{x} + \boldsymbol{y}^\top \boldsymbol{b} + \bar{z}$$

Let $\bar{\boldsymbol{c}}^\top := \boldsymbol{c}^\top - \boldsymbol{y}^\top A$. For (C2) to be satisfied we need $\bar{\boldsymbol{c}}_{\boldsymbol{B}} = \boldsymbol{0}$, so we need to choose $\boldsymbol{y}$ accordingly, such as

$$\bar{\boldsymbol{c}}_{\boldsymbol{B}}^\top = \boldsymbol{c}_{\boldsymbol{B}}^\top - \boldsymbol{y}^\top A_B = \boldsymbol{0}^\top$$

equivalently,

$$\boldsymbol{y}^\top A_B = \boldsymbol{c}_{\boldsymbol{B}}^\top \implies \boldsymbol{y}^\top = \boldsymbol{c}_{\boldsymbol{B}}^\top A_B^{-1}$$

We have shown the following:

**THEOREM 3.3.6** (Canonical Form). *Suppose an LP*

$$\max\{\boldsymbol{c}^\top \boldsymbol{x} + \bar{z} : A\boldsymbol{x} = \boldsymbol{b},\, \boldsymbol{x} \geqslant \boldsymbol{0}\}$$

*and a basis $B$ of $A$ are given. Then*

$$\max(\boldsymbol{c}^\top - \boldsymbol{y}^\top A)\boldsymbol{x} + \boldsymbol{y}^\top \boldsymbol{b} + \bar{z}$$

*subject to*

$$A_B^{-1} A\boldsymbol{x} = A_B^{-1}\boldsymbol{b}$$
$$\boldsymbol{x} \geqslant \boldsymbol{0}$$

*where $\boldsymbol{y}^\top = \boldsymbol{c}_{\boldsymbol{B}}^\top A_B^{-1}$, is an equivalent LP in canonical form for the basis $B$ of $A$.*

The canonical form is useful because it:

- allows us to simply read a basic solution
- gives us easy ways to move in the feasible region to improve the current basic feasible solution
- gives us a way to obtain optimality certificates if $\boldsymbol{c}^\top - \boldsymbol{y}^\top A \leqslant \boldsymbol{0}^\top$

**EXAMPLE 3.3.7** (Canonical Form). (P)

$$\max \begin{bmatrix} 0 & 0 & -4 & 1 & 0 \end{bmatrix} \boldsymbol{x}$$

subject to

$$\begin{bmatrix} 1 & 0 & -1 & 1 & 0 \\ 0 & 1 & 2 & -1 & 0 \\ 0 & 0 & -3 & 2 & 1 \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} 4 \\ 2 \\ 6 \end{bmatrix}$$
$$\boldsymbol{x} \geqslant \boldsymbol{0}$$

$B := \{1, 2, 5\}$ is a basis of $A$. Thus, the basic solution corresponding to the basis is

$$\bar{\boldsymbol{x}} := (4, 2, 0, 0, 6)^\top$$

$c_3 = -4$, increasing the value of $x_3$ from $0$ will decrease the objective value by $-4$ units
$c_4 = 1$, we want to increase the value of $x_4$, so

$$\begin{bmatrix} x_1 \\ x_2 \\ x_5 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \\ 6 \end{bmatrix} - x_4 \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} \geqslant 0$$

Let $t$ denote the maximum value we can assign to $x_4$ and stay feasible.
So, $t = \min\{4/1, \_, 6/2\} = 3$

---

2019-10-03

---

**EXAMPLE 3.3.8** (Previous Example Coninuted). So, the new basic feasible solution is $\bar{\boldsymbol{x}} := (1, 5, 0, 3, 0)^\top$ determined by the basis $B := \{1, 2, 5\} \cup \{4\} \setminus \{5\} = \{1, 2, 4\}$. Note that we exclude $\{5\}$ since the index of which $t$ achieved the minimum was at $6/2$, i.e. index $5$ (row $x_5$). The canonical form determined by the new basis is

$$\max \begin{bmatrix} 0 & 0 & -5/2 & 0 & -1/2 \end{bmatrix} \boldsymbol{x} + 3$$

subject to

$$\begin{bmatrix} 1 & 0 & 1/2 & 0 & -1/2 \\ 0 & 1 & 1/2 & 0 & 1/2 \\ 0 & 0 & -3/2 & 1 & 1/2 \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} 1 \\ 5 \\ 3 \end{bmatrix}$$

**REMARK 3.3.9.** $\bar{\boldsymbol{x}}$ is the optimal solution with optimal value $3$.

**REMARK 3.3.10.** How did we arrive to this LP? Using the formulae in Proposition 8.2. If you didn't want to calculate $A_B^{-1}$, then follow the below instructions.

**REMARK 3.3.11.** The following was not taught in class or the textbook. This method can be confusing and not intuitive.

**EXAMPLE 3.3.12** (Canonical Form Without Computing the Inverse)**.** Write

$$A := \left[\begin{array}{ccccc|c} 1 & 0 & -1 & 1 & 0 & 4 \\ 0 & 1 & 2 & -1 & 0 & 2 \\ 0 & 0 & -3 & 2 & 1 & 6 \end{array}\right] \to \left[\begin{array}{ccccc|c} 1 & 0 & 1/2 & 0 & -1/2 & 1 \\ 0 & 1 & 1/2 & 0 & 1/2 & 5 \\ 0 & 0 & -3/2 & 1 & 1/2 & 3 \end{array}\right] \begin{array}{l} -x_1 \\ -x_2 \\ -x_4 \end{array}$$

and row reduce $A$ to make fourth column get a leading one as seen above. The row-reduced matrix and the augment are your new constraints.
The objective function is tricky, we want a 0 in the fourth column of our $\boldsymbol{c}^\top$. Also, we denote $x_1, x_2, x_4$ as the rows of the matrix respectively as seen above. Using $x_4$ (which is our row-reduced $A$), we get

$$(-1)\left(\begin{bmatrix} 0 & 0 & -3/2 & 1 & 1/2 \end{bmatrix} \boldsymbol{x} - 3\right) + \left(\begin{bmatrix} 0 & 0 & -4 & 1 & 0 \end{bmatrix} \boldsymbol{x}\right)$$

The $-3$ right after the first matrix was the row of $\boldsymbol{b}$. General form:

$$c([\text{Row}_i(A)]\boldsymbol{x} - \boldsymbol{b_i}) + \text{original objective function}$$

where $c$ is a constant.

## 3.4 The Simplex Algorithm

**Algorithm 1:** Simplex Algorithm

**Input :** $A \in \mathbb{R}^{m \times n}$, $\boldsymbol{b} \in \mathbb{R}^m$, $\boldsymbol{c} \in \mathbb{R}^n$ such that we have linear program (P): $\max\left\{\boldsymbol{c}^\top \boldsymbol{x},\ A\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geqslant \boldsymbol{0}\right\}$, and a feasible basis $B$.

**Output:** An optimal solution $\bar{\boldsymbol{x}}$ of (P) or a certificate proving that the (P) is unbounded.

1 Compute the canonical form for the basis $B$. Let $\bar{\boldsymbol{x}}$ be the basic feasible solution for $B$.
2 If $\boldsymbol{c_N} \leqslant \boldsymbol{0}$, then stop ($\bar{\boldsymbol{x}}$ is optimal).
3 Choose $k \in N$ such that $c_k > 0$.
4 If $\boldsymbol{a_k} \leqslant \boldsymbol{0}$, then stop (the LP is unbounded).
5 Let $r$ be any index $i$ where the following minimum is attained:

$$t = \min\left\{\frac{b_i}{a_{ik}} : a_{ik} > 0\right\}$$

6 Let $\ell$ be the $r^{th}$ basis element.
7 Set $B := B \cup \{k\} \setminus \{\ell\}$.
8 Go to step 1.

**THEOREM 3.4.1** (Bland's Rule). *Throughout the Simplex iterations with $t = 0$, in Step 3, among all $j \in N$, with $c_j > 0$, choose $k := \min\{j \in N : c_j > 0\}$; also in Step 5, define $t$ as before and choose the smallest $r \in B$ with $a_{rk} > 0$, and ${}^{b_r}/_{a_{rk}} = t$.*

### 3.4.1 An Example with an Optimal Solution

**EXAMPLE 3.4.2** (Simplex Algorithm with Bland's Rule). Solve (P)

$$\max \begin{bmatrix}0 & 3 & 1 & 0\end{bmatrix} \boldsymbol{x}$$

subject to

$$\begin{bmatrix}1 & 2 & -2 & 0\\ 0 & 1 & 3 & 1\end{bmatrix} \boldsymbol{x} = \begin{bmatrix}2\\ 5\end{bmatrix}$$
$$\boldsymbol{x} \geqslant \mathbf{0}$$

using the Simplex Algorithm with Bland's Rule. Give a certificate of optimality or unboundedness for the problem, and verify it.
**Solution.**
**Iteration 1**
Useful values computed:

$$A_B = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} \Rightarrow A_B^{-1} = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$$
$$\boldsymbol{y}^\top = \boldsymbol{c}_{\boldsymbol{B}}^\top A_B^{-1} = \begin{bmatrix}0 & 0\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} = \begin{bmatrix}0\\ 0\end{bmatrix}$$

1. The LP is already in canonical form determined by $B = \{1, 4\}$. Let $\bar{\boldsymbol{x}} := (2, 0, 0, 5)^\top$ be the basic feasible solution for $B$.
2. $c_{\{2,3\}} \not\leq \mathbf{0}$, so $\bar{\boldsymbol{x}}$ is not optimal.
3. Using Bland's Rule we choose $k = 2 \in N$ which enters the basis as $c_2 \geqslant 0$.
4. $a_2 = \begin{bmatrix}2\\ 1\end{bmatrix} \not\leq \mathbf{0}$, so the LP is not unbounded.
5. $\begin{bmatrix}x_1\\ x_4\end{bmatrix} = \begin{bmatrix}2\\ 5\end{bmatrix} - t\begin{bmatrix}2\\ 1\end{bmatrix} \geqslant 0$ so

$$t = \min\left\{\frac{2}{2}, \frac{5}{1}\right\}$$

Minimum is attained at index $1$ ($x_1$). Let $r = 1$ be the index which attains the smallest value of $t$.
6. Let $1$ be the $1^{st}$ basis element.
7. Set $B := \{1, 4\} \cup \{2\} \setminus \{1\} = \{2, 4\}$
**Iteration 2**
Useful values computed:

$$A_B = \begin{bmatrix}2 & 0\\ 1 & 1\end{bmatrix} \Rightarrow A_B^{-1} = \begin{bmatrix}1/2 & 0\\ -1/2 & 1\end{bmatrix}$$
$$\boldsymbol{y}^\top = \boldsymbol{c}_{\boldsymbol{B}}^\top A_B^{-1} = \begin{bmatrix}3 & 0\end{bmatrix}\begin{bmatrix}1/2 & 0\\ -1/2 & 1\end{bmatrix} = \begin{bmatrix}3/2\\ 0\end{bmatrix}$$

1. Canonical form determined by $B = \{2, 4\}$ is

$$\max \begin{bmatrix}-3/2 & 0 & 4 & 0\end{bmatrix} + 3$$

subject to

$$\begin{bmatrix}1/2 & 1 & -1 & 0\\ -1/2 & 0 & 4 & 1\end{bmatrix} \boldsymbol{x} = \begin{bmatrix}1\\ 4\end{bmatrix}$$

$$\boldsymbol{x} \geqslant \mathbf{0}$$

Let $\bar{\boldsymbol{x}} := (0,1,0,4)^\top$ be the basic feasible solution.
2. $c_{\{1,3\}} \nleq \mathbf{0}$, so $\bar{\boldsymbol{x}}$ is not optimal.
3. Using Bland's Rule we choose $k = 3 \in N$ which enters the basis as $c_3 \geqslant 0$.
4. $a_3 = \begin{bmatrix} -1 \\ 4 \end{bmatrix} \nleq \mathbf{0}$, so the LP is not unbounded.
5. $\begin{bmatrix} x_2 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \end{bmatrix} - t \begin{bmatrix} -1 \\ 4 \end{bmatrix} \geqslant 0$ so

$$t = \min \left\{ \_, \frac{4}{4} \right\}$$

Minimum is attained at index 2 ($x_4$). Let $r = 2$ be the index which attains the smallest value of $t$.
6. Let 4 be the $2^{nd}$ basis element.
7. Set $B := \{1,4\} \cup \{3\} \setminus \{4\} = \{2,3\}$
**Iteration 3**
Useful values computed:

$$A_B = \begin{bmatrix} 2 & -2 \\ 1 & 3 \end{bmatrix} \Rightarrow A_B^{-1} = \begin{bmatrix} 3/8 & 1/4 \\ -1/8 & 1/4 \end{bmatrix}$$

$$\boldsymbol{y}^\top = \boldsymbol{c}_{\boldsymbol{B}}^\top A_B^{-1} = \begin{bmatrix} 3 & 1 \end{bmatrix} \begin{bmatrix} 3/8 & 1/4 \\ -1/8 & 1/4 \end{bmatrix} = \begin{bmatrix} 1 & 1 \end{bmatrix}$$

1. Canonical form determined by $B = \{2,3\}$ is

$$\max \begin{bmatrix} -1 & 0 & 0 & -1 \end{bmatrix} + 7$$

subject to

$$\begin{bmatrix} 3/8 & 1 & 0 & 1/4 \\ -1/8 & 0 & 1 & 1/4 \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$\boldsymbol{x} \geqslant \mathbf{0}$$

Let $\bar{\boldsymbol{x}} := (0,2,1,0)^\top$ be the basic feasible solution.
2. $c_{\{1,4\}} \leqslant \mathbf{0}$, stop $\bar{\boldsymbol{x}}$ is optimal.
The certificate of optimality is $\bar{\boldsymbol{y}} = (1,1)^\top$.
To verify that $\bar{\boldsymbol{y}} = (1,1)^\top$ is the certificate of optimality. We compute

$$A^\top \bar{\boldsymbol{y}} = \begin{bmatrix} 1 \\ 3 \\ 1 \\ 1 \end{bmatrix} \geqslant \begin{bmatrix} 0 \\ 3 \\ 1 \\ 0 \end{bmatrix} = \boldsymbol{c}$$

and

$$\boldsymbol{c}^\top \bar{\boldsymbol{x}} = \begin{bmatrix} 0 & 3 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 2 \\ 1 \\ 0 \end{bmatrix} = 7 = \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 5 \end{bmatrix} = \bar{\boldsymbol{y}}^\top \boldsymbol{b}$$

**REMARK 3.4.3.** This was obviously not done in class (in fact it's a textbook question!). It can be verified that $\bar{\boldsymbol{y}}$ is indeed the certificate of optimality by using the summary of outcomes as seen above.

### 3.4.2 Convergence of Simplex Algorithm

In each iteration, we choose $k \in N$ such that $c_k > 0$. Then, we compute $t = \min\{{}^{b_i}/_{A_{ik}} \mid A_{ik} > 0\}$. Then, throughout the rest of the Simplex iterations, we never see the same basis again. There are at most $\binom{n}{m}$ bases of A. Therefore, if $t > 0$ in each iteration, the Simplex Algorithm will terminate in at most $\binom{n}{m}$ iterations. The only way the algorithm will not terminate is when $t = 0$ for all iterations (after some # of iterations). If our choices for $k$ and $l$ are deterministic and consistent in this case if we repeat a basis we call it a *cycle*.

**THEOREM 3.4.4.** *The Simplex Algorithm starting from a basic feasible solution with Bland's Rule terminates.*

### 3.4.3 Implementation of the Simplex Algorithm in "Big Data"

In a given iteration of the Simplex Algorithm, what information do we need to execute the algorithm?

We have the original data ($A$, $\boldsymbol{b}$, $\boldsymbol{c}$) and we have the current $B$, $\bar{\boldsymbol{x}}$, $\bar{\boldsymbol{v}}$.

Pick any $k \in N$ such that $\bar{c_k} \geqslant 0$. $\bar{c_k} = c_k - \bar{\boldsymbol{y}}^\top A\boldsymbol{x}$ (where $\bar{\boldsymbol{y}}^\top = c_B A_B^{-1}$).

Then to compute $t$, we need $t = \min\{{}^{b_i}/_{A_{ik}} \mid A_{ik} > 0\}$.

So, we need $\bar{A}_k$: $\bar{A}_k = A_B^{-1} A_k$ and note that $\bar{\boldsymbol{x_N}} = \boldsymbol{0}$, $\bar{\boldsymbol{x_B}} = \bar{\boldsymbol{b}}$

We solve linear systems $A_B^\top \boldsymbol{y} = \boldsymbol{c_B}$ and $A_B \boldsymbol{d_B} = A_k$.

In implementations, we typically express $A_B$ or $A_B^{-1}$ as a product of elementary matrices.

In practice, good implementations of the Simplex Algorithm terminates after $2m$ to ${}^{n}/_{2}$ iterations. Each iteration is very fast.

It is an open problem whether there exists a variant of Simplex Algorithm which is guaranteed to terminate in at most $pn^q$ iterations for LP problems in SEF with $n$ variables, where $p$, $q$ are constants.

---

2019-10-10

---

Midterm 1 was written on this day, as a result no classes were held.

---

2019-10-22

---

Given any LP problem, we know how to convert it into an equivalent LP problem in SEF:

(P)

$$\max z := \boldsymbol{c}^\top \boldsymbol{x}$$

subject to

$$\begin{aligned} A\boldsymbol{x} &= \boldsymbol{b} \\ \boldsymbol{x} &\geqslant 0 \end{aligned}$$

where $A \in \mathbb{R}^{m \times n}$ has $\operatorname{rank}(A) = m$.

Given an LP in SEF, with a given basic feasible solution, we know how to solve it.

## 3.5 Finding Feasible Solutions

Given an LP in SEF with $\operatorname{rank}(A) = m$, how do we find a feasible solution or prove that none exists.

We will construct an *auxiliary LP problem*.

We can always make sure $\boldsymbol{b} \geqslant \mathbf{0}$. (If any $b_i < 0$, multiply both sides of that equation by $(-1)$) Introduce artificial variables $x_{n+1}, \ldots, x_{n+m}$

**DEFINITION 3.5.1.** Given (P): $\max \left\{ \boldsymbol{c}^\top \boldsymbol{x},\ A\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geqslant \mathbf{0} \right\}$, we define the *auxiliary linear program* of (P) as:
$(P_{aux})$

$$\min w := x_{n+1} + \cdots + x_{n+m}$$

subject to

$$\left[\ A \mid I\ \right] \underbrace{\begin{bmatrix} x_1 \\ \vdots \\ x_n \\ x_{n+1} \\ \vdots \\ x_{n+m} \end{bmatrix}}_{\boldsymbol{x}} = \boldsymbol{b}$$

$$\boldsymbol{x} \geqslant \mathbf{0}$$

where $\boldsymbol{b} \geqslant \mathbf{0}$, and $I$ is the $m \times m$ identity matrix.
We call the variables $x_{n+1}, \ldots, x_{n+m}$ *auxiliary variables*.

For every feasible solution of $(P_{aux})$, $w \geqslant 0$.

Therefore, $(P_{aux})$ is not unbounded.

If the optimal value of $(P_{aux})$ is zero, let $(\hat{x}_1, \ldots, \hat{x}_{n+m})^\top$ be the basic feasible solution of $(P_{aux})$. Then, $(\hat{x}_1, \ldots, \hat{x}_n)^\top$ is a basic feasible solution of (P).

It is basic since $\{A_j : \hat{x_j} > 0\}$ is linearly independent where $J$ is the column indices $j$ of $A$ for which $\hat{x}_j \neq 0$.

If $|\{j : \hat{x_j} > 0\}| = m$, this index set is a basis of $A$ which determines $(\hat{x}_1, \ldots, \hat{x}_n)^\top$.

If $|\{j : \hat{x_j} > 0\}| \leqslant m - 1$, we can extend this index set to be a basis of $A$, since $\operatorname{rank}(A) = m$.

If the optimal value of $(P_{aux})$ is positive, then (P) is infeasible. We state this as a theorem.

**THEOREM 3.5.2.** *Let $\bar{\boldsymbol{x}} = (\bar{x}_1, \ldots, \bar{x}_{n+m})^\top$ be an optimal solution to $(P_{aux})$.*
*(1) if $w = 0$, then $(\bar{x}_1, \ldots, \bar{x}_n)^\top$ is a solution to (P).*
*(2) if $w > 0$, then (P) is infeasible.*

*Proof.* Let $\bar{\boldsymbol{x}} = (\bar{x}_1, \ldots, \bar{x}_{n+m})^\top$ be an optimal solution to $(P_{aux})$.

(1) Assume $w = 0$, then $\bar{x}_{n+1} = \cdots = \bar{x}_{n+m} = 0$. Thus $(\bar{x}_1, \ldots \bar{x}_n)^\top$ is a feasible solution of (P).

(2) Assume $w > 0$. Suppose for a contradiction that there exists a feasible solution $(\bar{x}_1, \ldots \bar{x}_n)^\top$ to (P). Then, $(\bar{x}_1, \ldots \bar{x}_n, \underbrace{0, \ldots, 0}_{m \text{ terms}})$ is a feasible solution to $(P_{aux})$ with optimal objective value $0$ which is a contradiction to the fact that $\bar{\boldsymbol{x}}$ is optimal. $\square$

**Algorithm 2:** Two Phase Method

**Input :** $A, \boldsymbol{b}, \boldsymbol{c}$ data for LP in SEF such that full row rank and $\boldsymbol{b} \geqslant \mathbf{0}$.

1 Construct $(P_{aux})$ put into SEF, $B := \{n+1, n+2, \ldots, n+m\}$
2 Put $(P_{aux})$ into the canonical form determined by $B$.
3 Solve $(P_{aux})$ starting with basis $B$ by Simplex Method.
4 If the optimal value of $(P_{aux})$ is zero, then we have a basic feasible solution of (P). Solve (P) using Simplex Method. This is known as Phase II.
5 If the optimal objective value of $(P_{aux})$ is not zero, then (P) is infeasible (a certificate of infeasibility is given by the last $\bar{\boldsymbol{y}}$ computed).

As seen above, the original LP can either have an optimal solution or be infeasible when performing the Two Phase Method.

### 3.5.1 The Two Phase Simplex Algorithm—An Optimal Example

**EXAMPLE 3.5.3** (Two Phase—Optimal)**.** (P)

$$\max z := \begin{bmatrix} 1 & 2 & -1 \end{bmatrix} \boldsymbol{x}$$

subject to

$$\begin{bmatrix} 1 & -2 & -3 \\ -1 & 1 & 1 \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} -3 \\ 1 \end{bmatrix}$$

$$\boldsymbol{x} \geqslant \mathbf{0}$$

Since $b_1 < 0$ we write

$$\begin{bmatrix} -1 & 2 & 3 \\ -1 & 1 & 1 \end{bmatrix} x = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

we do this because we will not have a feasible solution if $\boldsymbol{b} < \mathbf{0}$.
Introduce artificial variables: $x_4, x_5$
<u>Phase I</u>
$(P_{aux})$

$$\max -w := \begin{bmatrix} 0 & 0 & 0 & -1 & -1 \end{bmatrix} \boldsymbol{x}$$

subject to

$$\begin{bmatrix} -1 & 2 & 3 & 1 & 0 \\ -1 & 1 & 1 & 0 & 1 \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

$$\boldsymbol{x} := (x_1, x_2, x_3, x_4, x_5)^\top \geqslant \mathbf{0}$$

Turn $(P_{aux})$ into canonical form for $B := \{4, 5\}$

$$\max -w := \begin{bmatrix} -2 & 3 & 4 & 0 & 0 \end{bmatrix} \boldsymbol{x}$$

subject to

$$\begin{bmatrix} -1 & 2 & 3 & 1 & 0 \\ -1 & 1 & 1 & 0 & 1 \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

$$\boldsymbol{x} \geqslant \mathbf{0}$$

We solve the LP via Simplex Algorithm and obtain the following LP corresponding to the optimal basis of $B = \{1, 2\}$

$$\max -w := \begin{bmatrix} 0 & 0 & 0 & -1 & -1 \end{bmatrix} \boldsymbol{x}$$

subject to

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

$$\boldsymbol{x} \geqslant \mathbf{0}$$

End of Phase I.
Phase II
We rewrite the original LP in canonical form corresponding to basis $B = \{1, 2\}$ to obtain

$$\max z := \begin{bmatrix} 0 & 0 & -6 \end{bmatrix} \boldsymbol{x} + 5$$

subject to

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$\boldsymbol{x} := (x_1, x_2, x_3)^\top \geqslant \mathbf{0}$$

We obtain the optimal basic feasible solution of $(P_{aux})$ via the Simplex Algorithm

$$(\hat{x_1}, \hat{x_2}, \hat{x_3}, \hat{x_4}, \hat{x_5}) := (1, 2, 0, 0, 0)^\top$$

Thus, the corresponding basic feasible solution of (P) is

$$(\hat{x_1}, \hat{x_2}, \hat{x_3}) = (1, 2, 0)^\top$$

with an optimal objective value of $z := 5$.

### 3.5.2 The Two Phase Simplex Algorithm—An Infeasible Example

**EXAMPLE 3.5.4** (Two Phase—Infeasible)**.** (P)

$$\max z := \begin{bmatrix} 3 & 2 & 4 \end{bmatrix} \boldsymbol{x}$$

subject to

$$\underbrace{\begin{bmatrix} 5 & 1 & 1 \\ -1 & 1 & 2 \end{bmatrix}}_{A} \boldsymbol{x} = \begin{bmatrix} 1 \\ 5 \end{bmatrix}$$

$$\boldsymbol{x} \geqslant \mathbf{0}$$

$(P_{aux})$

$$\max -w := \underbrace{\begin{bmatrix} 0 & 0 & 0 & -1 & -1 \end{bmatrix}}_{\tilde{c_B}} \boldsymbol{x}$$

subject to

$$\underbrace{\begin{bmatrix} 5 & 1 & 1 & 1 & 0 \\ -1 & 1 & 2 & 0 & 1 \end{bmatrix}}_{\tilde{A}} \boldsymbol{x} = \begin{bmatrix} 1 \\ 5 \end{bmatrix}$$

$$\boldsymbol{x} := (x_1, x_2, x_3, x_4, x_5)^\top \geqslant \mathbf{0}$$

Turn $(P_{aux})$ into canonical form for $B := \{4, 5\}$ (by adding the constraints up to the original objective function).

$$\max -w := \begin{bmatrix} 4 & 2 & 3 & 0 & 0 \end{bmatrix} \boldsymbol{x} - 4$$

subject to

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

Starting with the basis $B = \{4, 5\}$, solve $(P_{aux})$ with the Simplex Algorithm to get:

$$\max -w = \begin{bmatrix} -11 & -1 & 0 & -3 & 0 \end{bmatrix} \boldsymbol{x} - 3$$

subject to

$$\begin{bmatrix} 5 & 1 & 1 & 1 & 0 \\ -11 & -1 & 0 & -2 & 1 \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$
$$\boldsymbol{x} \geqslant \boldsymbol{0}$$

The optimal value of $(P_{aux})$ is not zero. Therefore, (P) is infeasible. The basis in the last iteration was $B = \{3, 5\}$.

$$\boldsymbol{y}^\top = \tilde{\boldsymbol{c}}_{\boldsymbol{B}}^\top \tilde{A}_B^{-1} \iff \boldsymbol{y}^\top = \underbrace{\begin{bmatrix} 0 & -1 \end{bmatrix}}_{\text{SEF of } (P_{aux})} \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}^{-1}$$

$\bar{\boldsymbol{y}} = (2, -1)^\top$ is a certificate of infeasibility of (P).
Compute $\bar{\boldsymbol{y}}^\top A = \begin{bmatrix} 11 & 1 & 0 \end{bmatrix} \geqslant \boldsymbol{0}^\top$ and $\bar{\boldsymbol{y}}^\top \boldsymbol{b} = -3 = \boldsymbol{c}^\top \boldsymbol{x}$ where $\bar{\boldsymbol{x}} = (0, 0, 0, 1, 3)^\top$.
Thus, $\bar{\boldsymbol{y}}$ is a certificate of optimality for $(P_{aux})$.

**THEOREM 3.5.5** (Fundamental Theorem of LP (SEF)). *Let (P) be an LP problem in SEF, where $A \in \mathbb{R}^{m \times n}$ has $\operatorname{rank}(A) = m$.*
*(1) if (P) does not have an optimal solution, then (P) is either infeasible or unbounded.*
*(2) if (P) has a feasible solution, then (P) has a basic feasible solution.*
*(3) if (P) has an optimal solution, then (P) has an optimal basic feasible solution.*

---

2019-10-24

---

**THEOREM 3.5.6** (Fundamental Theorem of LP). *Let (P) be an LP problem. Then exactly one of the following holds:*

- *(P) is infeasible*
- *(P) is unbounded*
- *(P) has an optimal solution*

**REMARK 3.5.7.** Fundamental Theorem of LP and Fundamental Theorem of LP (SEF) are not the same, they are two completely different theorems!

## 3.6 Geometry

### 3.6.1 Feasible Region of LPs and Polyhedra

**DEFINITION 3.6.1.** Let $\boldsymbol{a} \in \mathbb{R}^n \setminus \boldsymbol{0}$, $\beta \in \mathbb{R}$.
$H := \left\{ \boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{a}^\top \boldsymbol{x} = \beta \right\}$ is a *hyperplane*.
$F := \left\{ \boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{a}^\top \boldsymbol{x} \leqslant \beta \right\}$ is a *half-space*.

Solution sets of linear equations are intersections of hyperplanes.

**DEFINITION 3.6.2.** Let $A \in \mathbb{R}^{m \times n}$, $\boldsymbol{b} \in \mathbb{R}^m$. $P := \{\boldsymbol{x} \in \mathbb{R}^n : A\boldsymbol{x} \leqslant \boldsymbol{b}\}$ is a *polyhedron*.

**REMARK 3.6.3.** The set of solutions to any one of the inequalities of $A\boldsymbol{x} \leqslant \boldsymbol{b}$ is a half-space.

**THEOREM 3.6.4.** *The feasible region of an LP is a polyhedron or equivalently the intersection of a finite number of half-spaces.*

*Proof.* Let $\boldsymbol{a} \in \mathbb{R}^n, \boldsymbol{x} \in \mathbb{R}^n, \beta \in \mathbb{R}$.

Given an inequality of the form $\boldsymbol{a}^\top \boldsymbol{x} \geqslant \beta$, we can rewrite it as $-\boldsymbol{a}^\top \boldsymbol{x} \leqslant -\beta$.

Given an equation of the form $\boldsymbol{a}^\top \boldsymbol{x} = \beta$ we can rewrite it as $\boldsymbol{a}^\top \boldsymbol{x} \geqslant \beta$ and $-\boldsymbol{a}^\top \boldsymbol{x} \leqslant -\beta$.

Thus, any set of linear constraints can be rewritten as $A\boldsymbol{x} \leqslant \boldsymbol{b}$ for some $A \in \mathbb{R}^{m \times n}$ and $\boldsymbol{b} \in \mathbb{R}^m$, where $\boldsymbol{a}^\top$ can correspond to each row of $A$, and $\beta$ can correspond to each row of the column vector $\boldsymbol{b}$. $\square$

Solutions sets of $A\boldsymbol{x} = \boldsymbol{b}$ are either $\emptyset$, a single point, a line, or in general, an intersection of a hyperplane.

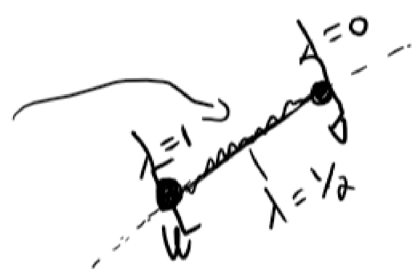
Note that already in $\mathbb{R}^2$ there are already equivalent polyhedra. The mathematical modelling power of LPs are significantly more than that of linear systems of equations.

### 3.6.2 Convexity

**DEFINITION 3.6.5.** The *line segment* joining two points, $x^{(1)}$ and $x^{(2)}$ is

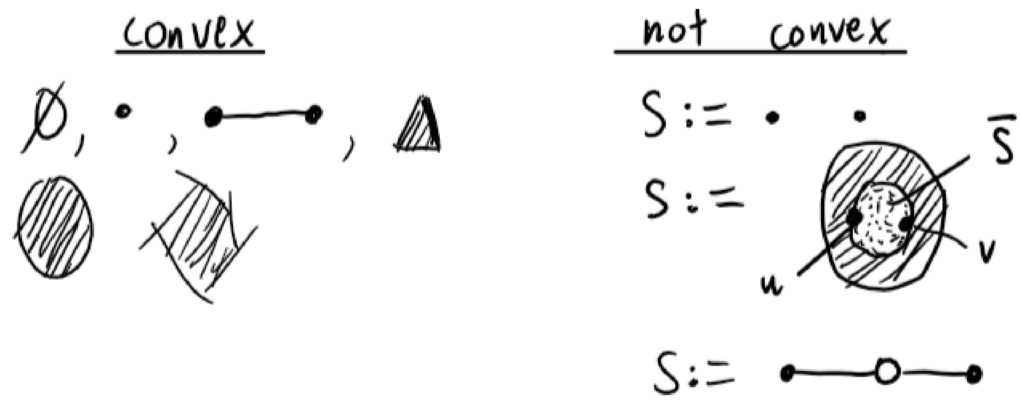
$$\left\{\lambda x^{(1)} + (1-\lambda)x^{(2)} : \lambda \in [0,1]\right\}$$

Graphically, the line segment can be seen as:




**DEFINITION 3.6.6.** A subset $S \subseteq \mathbb{R}^n$ is *convex* if for every pair of points $x^{(1)}, x^{(2)} \in S$, the line segment with ends $x^{(1)}$ and $x^{(2)}$ is included in $S$. That is,

$$\left\{\lambda x^{(1)} + (1-\lambda)x^{(2)} : \lambda \in [0,1]\right\} \subseteq S$$




**THEOREM 3.6.7.** *Half-spaces are convex.*

*Proof.* Let $H \subseteq \mathbb{R}^n$ be a half-space. Then $\boldsymbol{a} \in \mathbb{R}^n \setminus \mathbf{0}$ and $\beta \in \mathbb{R}^n$ such that

$$H = \left\{ \boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{a}^\top \boldsymbol{x} \leq \beta \right\}$$

Let $\bar{x} = \lambda x^{(1)} + (1-\lambda)x^{(2)}$ where $\lambda \in [0,1]$.

$$\boldsymbol{a}^\top \bar{x} = \boldsymbol{a}^\top \left[ \lambda x^{(1)} + (1-\lambda)x^{(2)} \right] = \underbrace{\lambda}_{\geqslant 0} \underbrace{\boldsymbol{a}^\top x^{(1)}}_{\leqslant \beta} + \underbrace{(1-\lambda)}_{\geqslant 0} \underbrace{\boldsymbol{a}^\top x^{(2)}}_{\leqslant \beta} \leqslant \lambda\beta + (1-\lambda)\beta = \beta$$

Thus, $H$ is convex since $\bar{x} \in H$. $\square$

**THEOREM 3.6.8.** *The intersection of any collection of convex sets is convex. That is, a convex set $C_j$ for all $j \in J$, the intersection*

$$C := \bigcap_{j \in J} C_j$$

*is convex.*

*Proof.* Let $u, v$ be two points in $C$. Let $w$ lie on the line segment between $u$ and $v$. Then, $w \in C_j$ since $C_j$ is convex for each $j \in J$. Thus, $w \in C$. $\square$

**REMARK 3.6.9.** $J$ can be infinite. That is, the intersection of infinitely many convex sets is convex, which can be formally proved by strong induction.

**THEOREM 3.6.10.** *Polyhedra are convex.*

**DEFINITION 3.6.11.** We say that a point $x$ is *properly contained* in a line segment if it is in the line segment and not an endpoint.

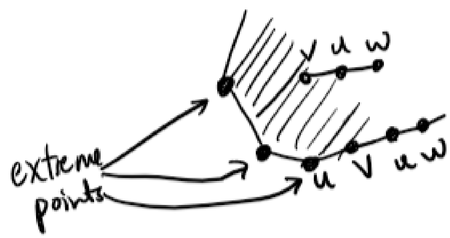
### 3.6.3 Extreme Points

**DEFINITION 3.6.12.** Let $S \subseteq \mathbb{R}^n$ be a convex set. Let $\bar{x} \in \mathbb{R}^n$. $\bar{x}$ is an *extreme point* of $S$, if $\bar{x} \in S$ and no line segment that properly contains $\bar{x}$ is included in $S$.
Equivalently, $\bar{x}$ is an *extreme point* of $S$, if $\bar{x} \in S$ and no two distinct $x^{(1)}, x^{(2)} \in S$ exist satisfying

$$\bar{x} = \lambda x^{(1)} + (1-\lambda)x^{(2)}$$

for some $\lambda \in (0,1)$.
Equivalently, $\bar{x}$ is an *extreme point* of $S$, if $\bar{x} \in S$ and no two distinct $x^{(1)}, x^{(2)} \in S$ exist satisfying

$$\bar{x} = \frac{1}{2}x^{(1)} + \frac{1}{2}x^{(2)}$$




2019-10-30

Recall the notions: hyperplane, half-space, polyhedron, feasible regions of LPs, convex sets, extreme points of convex sets.

**THEOREM 3.6.13.** *Let $S \subseteq \mathbb{R}^n$ be a convex set and $\bar{x} \in S$. $\bar{x}$ is an extreme point of $S$ if and only if $S \setminus \{\bar{x}\}$ is convex.*

*Proof.* Assume $S \subseteq \mathbb{R}^n$ is convex and $\bar{x} \in S$.

$\Rightarrow$ Suppose $\bar{x}$ is an extreme point of $S$. Pick two points $x^{(1)}, x^{(2)} \in S \setminus \{\bar{x}\}$ and $\lambda \in [0,1]$ and set $\bar{x} = \lambda x^{(1)} + (1-\lambda)x^{(2)}$. To show that $S \setminus \{\bar{x}\}$ is convex, we have to verify that $x \in S \setminus \{\bar{x}\}$. Now the set $S$ is convex, so $x \in S$. It remains to show that $x \neq \bar{x}$.

Case 1: $\lambda = 0$. Then $x = x^{(1)} \neq \bar{x}$.

Case 2: $\lambda = 1$. Then $x = x^{(2)} \neq \bar{x}$.

Case 3: $\lambda \in (0,1)$. We know $x$ must be different from $\bar{x}$, otherwise we would contradict the fact that $\bar{x}$ is an extreme point of $S$.

$\Leftarrow$ We prove the contrapositive. Assume that $\bar{x}$ is not an extreme point of $S$. Then there exists two points $x^{(1)}, x^{(2)} \in S$ with $x^{(1)} \neq x^{(2)}$ and some $\lambda \in (0,1)$ such that $\bar{x} = \lambda x^{(1)} + (1-\lambda)x^{(2)}$. Note that $x^{(1)}, x^{(2)} \in S \setminus \{\bar{x}\}$ since $x^{(1)} \neq x^{(2)}$. Thus $S \setminus \{\bar{x}\}$ is not convex. $\square$

**DEFINITION 3.6.14.** Let $A \in \mathbb{R}^{m \times n}$, $\boldsymbol{b} \in \mathbb{R}^m$. Consider the polyhedron $P \subseteq \mathbb{R}^n$ $P := \{\boldsymbol{x} \in \mathbb{R}^n : A\boldsymbol{x} \leqslant \boldsymbol{b}\}$. We say that a constraint $\boldsymbol{\alpha}^\top \boldsymbol{x} \leqslant \beta$ of $A\boldsymbol{x} \leqslant \boldsymbol{b}$ is *tight* for $\bar{\boldsymbol{x}}$ if $\boldsymbol{\alpha}^\top \bar{\boldsymbol{x}} = \beta$. We denote the set of all inequalities of $A\boldsymbol{x} \leqslant \boldsymbol{b}$ that are tight at $\bar{\boldsymbol{x}}$ by $A^= \bar{\boldsymbol{x}} = b^=$.

**THEOREM 3.6.15.** *Let $P = \{\boldsymbol{x} \in \mathbb{R}^n : A\boldsymbol{x} \leqslant \boldsymbol{b}\}$ be a polyhedron and let $\bar{\boldsymbol{x}} \in P$. Let $A^{=}\boldsymbol{x} = \boldsymbol{b}^{=}$ be the set of tight constraints for $\bar{\boldsymbol{x}}$. Then $\operatorname{rank}(A^{=}) = n$ if and only if $\bar{\boldsymbol{x}}$ is an extreme point of $P$.*

*Proof.* $\implies$ [$\operatorname{rank}(A^{=}) = n \implies \bar{\boldsymbol{x}}$ is an extreme point of $P$]

Suppose $\operatorname{rank}(A^{=}) = n$. Suppose for a contradiction that $\bar{\boldsymbol{x}}$ is not an extreme point. Then there exists $\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)} \in P$, where $\boldsymbol{x}^{(1)} \neq \boldsymbol{x}^{(2)}$ and $\lambda \in (0, 1)$ such that

$$\bar{\boldsymbol{x}} = \lambda \boldsymbol{x}^{(1)} + (1 - \lambda)\boldsymbol{x}^{(2)}$$

Thus,

$$\begin{aligned}
\boldsymbol{b}^{=} &= A^{=}\bar{\boldsymbol{x}} \\
&= A^{=}[\lambda \boldsymbol{x}^{(1)} + (1-\lambda)\boldsymbol{x}^{(2)}] \\
&= \underbrace{\lambda}_{>0} \underbrace{A^{=}\boldsymbol{x}^{(1)}}_{\leqslant \boldsymbol{b}^{=}} + \underbrace{(1-\lambda)}_{>0} \underbrace{A^{=}\boldsymbol{x}^{(2)}}_{\leqslant \boldsymbol{b}^{=}} \\
&\leq \lambda \boldsymbol{b}^{=} + (1-\lambda)\boldsymbol{b}^{=} \\
&= \boldsymbol{b}^{=}
\end{aligned}$$

Thus, we must have that everything in the inequality chain starting and ending with $\boldsymbol{b}^{=}$ is equal. Thus, $A^{=}\boldsymbol{x}^{(1)} = A^{=}\boldsymbol{x}^{(2)} = \boldsymbol{b}^{=}$. $\operatorname{rank}(A^{=}) = n$ implies there is a unique solution to $A^{=}\bar{\boldsymbol{x}} = \boldsymbol{b}^{=}$, so we have $\bar{\boldsymbol{x}} = \boldsymbol{x}^{(1)} = \boldsymbol{x}^{(2)}$, a contradiction that $\bar{\boldsymbol{x}}$ is an extreme point.

$\impliedby$ [$\operatorname{rank}(A^{=}) = n \impliedby \bar{\boldsymbol{x}}$ is an extreme point of $P$]

We will prove the contrapositive of this. That is, we will be prove $\operatorname{rank}(A^{=}) \neq n \implies \bar{\boldsymbol{x}}$ is not an extreme point of $P$.

Suppose that $\operatorname{rank}(A^{=}) \neq n$, that is $\operatorname{rank}(A^{=}) < n$, which means that the columns of $A^{=}$ are linearly dependent. Thus, $\exists \boldsymbol{d}$ such that $A^{=}\boldsymbol{d} = \boldsymbol{0}$. Let $\varepsilon > 0$ be arbitrarily small and define

$$\boldsymbol{x}^{(1)} := \bar{\boldsymbol{x}} + \varepsilon \boldsymbol{d}$$

$$\boldsymbol{x}^{(2)} := \bar{\boldsymbol{x}} - \varepsilon \boldsymbol{d}$$

Hence, $\bar{\boldsymbol{x}} = \frac{1}{2}\boldsymbol{x}^{(1)} + \frac{1}{2}\boldsymbol{x}^{(2)}$, where $\boldsymbol{x}^{(1)}$ and $\boldsymbol{x}^{(2)}$ are distinct. Thus, $\bar{\boldsymbol{x}}$ is in the line segment between $\boldsymbol{x}^{(1)}$ and $\boldsymbol{x}^{(2)}$.

We need to show that $\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)} \in P$ for $\varepsilon > 0$ arbitrarily small. We have

$$\begin{aligned}
A^{=}\boldsymbol{x}^{(1)} &= A^{=}(\bar{\boldsymbol{x}} + \varepsilon \boldsymbol{d}) \\
&= A^{=} \underbrace{\bar{\boldsymbol{x}}}_{=\boldsymbol{b}^{=}} + \varepsilon \underbrace{A^{=}\boldsymbol{d}}_{=\boldsymbol{0}} \\
&= \boldsymbol{b}^{=}
\end{aligned}$$

Similarly, $A^{=}\boldsymbol{x}^{(2)} = \boldsymbol{b}^{=}$. Let $\boldsymbol{a}^{\top}\boldsymbol{x} \leqslant \beta$ be any of the inequalities of $A\boldsymbol{x} \leqslant \boldsymbol{b}$ that is not in $A^{=}\boldsymbol{x} \leqslant \boldsymbol{b}^{=}$. It follows for $\varepsilon > 0$ arbitrarily small that:

$$\begin{aligned}
\boldsymbol{a}^{\top}\boldsymbol{x}^{(1)} &= \boldsymbol{a}^{\top}(\bar{\boldsymbol{x}} + \varepsilon \boldsymbol{d}) \\
&= \underbrace{\boldsymbol{a}^{\top}\bar{\boldsymbol{x}}}_{\leqslant \beta} + \varepsilon \boldsymbol{a}^{\top}\boldsymbol{d} \\
&\leqslant \beta
\end{aligned}$$

hence $\boldsymbol{x}^{(1)} \in P$. Similarly, $\boldsymbol{x}^{(2)} \in P$. Thus, $\bar{\boldsymbol{x}}$ is properly contained in $P$ and hence is not an extreme point. $\square$

**Example**

$$F := \left\{ \boldsymbol{x} \in \mathbb{R}^2 : \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ -1 & 0 \\ 0 & -1 \end{bmatrix} \boldsymbol{x} \leq \begin{bmatrix} 4 \\ 2 \\ 0 \\ 0 \end{bmatrix}, \boldsymbol{x} \geqslant \mathbf{0} \right\}$$

(i) $\boldsymbol{x}^{(1)} := \begin{bmatrix} 2 \\ 0 \end{bmatrix}$, $A^{=} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$

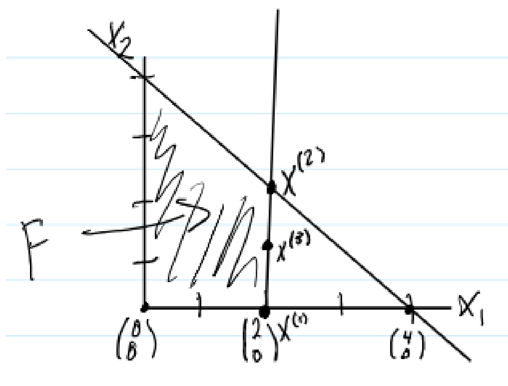
$\operatorname{rank}(A^{=}) = 2 = n$, therefore $\boldsymbol{x}^{(1)}$ is an extreme point of $F$.

(ii) $\boldsymbol{x}^{(2)} := \begin{bmatrix} 2 \\ 2 \end{bmatrix}$, $A^{=} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$

$\operatorname{rank}(A^{=}) = 2 = n$, therefore $\boldsymbol{x}^{(2)}$ is an extreme point of $F$.

(iii) $\boldsymbol{x}^{(3)} := \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, $A^{=} = \begin{bmatrix} 1 & 0 \end{bmatrix}$

$\operatorname{rank}(A^{=}) = 1 < 2 = n$, therefore $\boldsymbol{x}^{(3)}$ is not an extreme point of $F$.




**THEOREM 3.6.16.** *Let $A \in \mathbb{R}^{m \times n}$ with $\operatorname{rank}(A) = m$. Let $P = \{\boldsymbol{x} \in \mathbb{R}^n : A\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geqslant \mathbf{0}\}$, and let $\bar{\boldsymbol{x}} \in P$. $\bar{\boldsymbol{x}}$ is an extreme point of $P$ if and only if $\bar{\boldsymbol{x}}$ is a basic feasible solution of $A\boldsymbol{x} = \boldsymbol{b}$.*

$$F := \left\{ \boldsymbol{x} \in \mathbb{R}^2 : \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ -1 & 0 \\ 0 & -1 \end{bmatrix} \boldsymbol{x} \leq \begin{bmatrix} 4 \\ 2 \\ 0 \\ 0 \end{bmatrix}, \boldsymbol{x} \geqslant \mathbf{0} \right\}$$

$$P := \left\{ \boldsymbol{x} \in \mathbb{R}^4 : \begin{bmatrix} 1 & 1 & -1 & 0 \\ 1 & 0 & 0 & -1 \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}, \boldsymbol{x} \geqslant \mathbf{0} \right\}$$

Note that for every feasible solution $\begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \end{bmatrix} \in F$, $\begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ 4 - \bar{x}_1 - \bar{x}_3 \\ 2 - \bar{x}_1 \end{bmatrix} \in P$.

Conversely, for every $\begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \hat{x}_3 \\ \hat{x}_4 \end{bmatrix} \in P$, $\begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} \in F$

Consider the basis $B := \{3, 4\}$ of $A$. The corresponding basic feasible solution is $\bar{\boldsymbol{x}} = (0, 0, 4, 2)^\top$. Thus, $\bar{\boldsymbol{x}}$ is an extreme point of $P$.

### 3.6.4 Geometric Interpretation of Simplex Algorithm

(P)

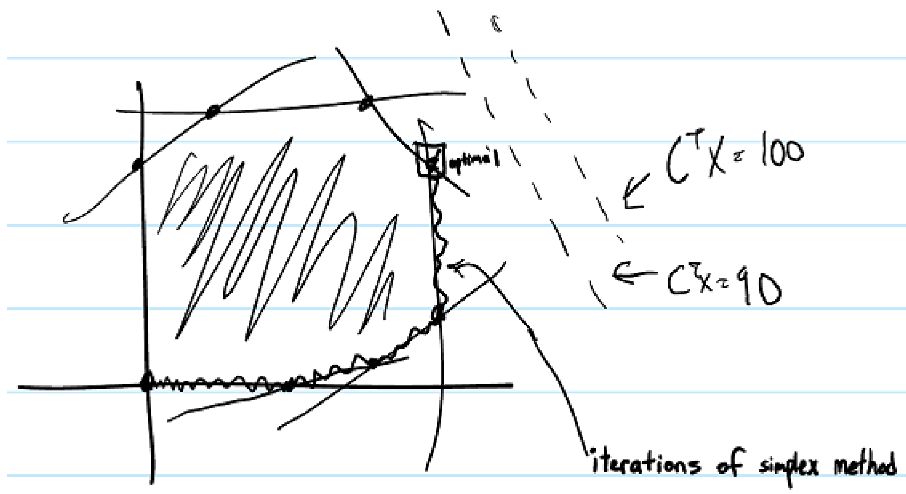
$$\max z := \boldsymbol{c}^\top x$$

subject to

$$A\boldsymbol{x} \leqslant \boldsymbol{b}$$

$$\boldsymbol{x} \geqslant \boldsymbol{0}$$

Suppose $n = 1$ and $m = 6$ with $\boldsymbol{b} \geqslant \boldsymbol{0}$.



# Chapter 4

# Duality Theory

(P)

$$\max\{\boldsymbol{c}^\top \boldsymbol{x} : A\boldsymbol{x} = \boldsymbol{b},\, \boldsymbol{x} \geqslant \boldsymbol{0}\}$$

Recall the notation of optimality certificate $\bar{\boldsymbol{y}} \in \mathbb{R}^m$ such that $A^\top \bar{\boldsymbol{y}} \geqslant \boldsymbol{c}$. We noted that for every feasible $\boldsymbol{x}$ in (P), $A\boldsymbol{x} = \boldsymbol{b} \implies \bar{\boldsymbol{y}}^\top A\boldsymbol{x} = \bar{\boldsymbol{y}}^\top \boldsymbol{b}$

Since $A^\top \bar{\boldsymbol{y}} \geqslant \boldsymbol{c}$ and $\boldsymbol{x} \geqslant \boldsymbol{0}$, we have $\boldsymbol{c}^\top \boldsymbol{x} \leqslant \bar{\boldsymbol{y}}^\top A\boldsymbol{x} = \bar{\boldsymbol{y}}^\top \boldsymbol{b}$. So as long as $\boldsymbol{y} \in \mathbb{R}^m$ with $A^\top \boldsymbol{y} \geqslant \boldsymbol{c}$, we can get an upper bound of $\boldsymbol{b}^\top \boldsymbol{y}$ on the optimal objective value of (P).

We want to minimize $\boldsymbol{b}^\top \boldsymbol{y}$ subject to $A^\top \boldsymbol{y} \geqslant \boldsymbol{c}$

**DEFINITION 4.0.1.** Given (P)

$$\max\{\boldsymbol{c}^\top \boldsymbol{x} : A\boldsymbol{x} = \boldsymbol{b},\, \boldsymbol{x} \geqslant qslant\boldsymbol{0}\}$$

we define (D)

$$\min\{\boldsymbol{b}^\top \boldsymbol{y} : A^\top \boldsymbol{y} \geqslant qslant\boldsymbol{c},\, \boldsymbol{y} \geqslant qslant\boldsymbol{0}\}$$

to be the *dual* of (P).

**EXAMPLE 4.0.2** (Dual)**.** Find the dual of $(P_1)$: $\max\{\boldsymbol{c}^\top \boldsymbol{x} : A\boldsymbol{x} \leqslant \boldsymbol{b},\, \boldsymbol{x} \geqslant \boldsymbol{0}\}$.
**Solution.**
Convert to SEF by introducing slack variables: $\boldsymbol{s} = (s_1, \ldots, s_n)^\top$.
$(P_2)$

$$\max \boldsymbol{c}^\top \begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{s} \end{bmatrix}$$

subject to

$$\left[\begin{array}{c|c} A & I \end{array}\right] \begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{s} \end{bmatrix} = \boldsymbol{b}$$

$$(\boldsymbol{x}, \boldsymbol{s})^\top \geqslant \boldsymbol{0}$$

$(D_2)$

$$\min \boldsymbol{b}^\top \boldsymbol{y}$$

$$\begin{bmatrix} A^T \\ I \end{bmatrix} \geqslant \begin{bmatrix} \boldsymbol{c} \\ \boldsymbol{0} \end{bmatrix}$$

$$\boldsymbol{y} \geqslant \boldsymbol{0}$$

Thus, the dual of $(P_1)$ is $(D_2)$: $\min\{\boldsymbol{b}^\top \boldsymbol{y} : A^\top \boldsymbol{y} \geqslant \boldsymbol{c},\, \boldsymbol{y} \geqslant \boldsymbol{0}\}$.

2019-10-31

Last lecture, we defined the dual of LPs in SIF and found the dual of $(P_1)$.

$$(P_1):\ \max\{\boldsymbol{c}^\top \boldsymbol{x} : A\boldsymbol{x} \leqslant \boldsymbol{b},\ \boldsymbol{x} \geqslant \boldsymbol{0}\}$$

$$(D_1):\ \min\{\boldsymbol{b}^\top \boldsymbol{y} : A^\top \boldsymbol{y} \geqslant \boldsymbol{c},\ \boldsymbol{y} \geqslant \boldsymbol{0}\}$$

**EXAMPLE 4.0.3** (Directly Writing the Dual of an LP)**.** Suppose $A \in \mathbb{R}^{3\times 4}$.
(P)

$$\max 10x_1 + 20x_2 + 30x_3 + 40x_4$$

subject to

$$\begin{bmatrix} \boldsymbol{a_1}^\top \\ \boldsymbol{a_2}^\top \\ \boldsymbol{a_3}^\top \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \begin{matrix} \leq \\ = \\ \geqslant \end{matrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

$$x_1 \geqslant 0, x_2 \geqslant 0, x_3 \leqslant 0, x_4 \text{ free}$$

where $\boldsymbol{a_1}, \boldsymbol{a_2}, \boldsymbol{a_3} \in \mathbb{R}^4$.
(D)

$$\min \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

subject to

$$\begin{bmatrix} \boldsymbol{a_1} & \boldsymbol{a_2} & \boldsymbol{a_3} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \begin{matrix} \geqslant \\ \geqslant \\ \leq \\ = \end{matrix} \begin{bmatrix} 10 \\ 20 \\ 30 \\ 40 \end{bmatrix}$$

$$y_1 \geqslant 0, y_2 \text{ free}, y_3 \leqslant 0$$

Dual of the dual is the original problem, the primal.
Since:

1. constraint 1 in (P) is $\leqslant$, then $y_1 \geqslant 0$
2. constraint 2 in (P) is $=$, then $y_2$ free
3. constraint 3 in (P) is $\geqslant$, then $y_3 \leqslant 0$
4. $x_1, x_2 \geqslant 0$, then constraint 1, 2 in (D) is $\geqslant$
5. $x_3 \leqslant 0$, then constraint 3 in (D) is $\leq$
6. $x_4$ free, then constraint 4 in (D) is $=$

## 4.1 Weak Duality

**THEOREM 4.1.1** (Weak Duality - Special Form)**.** *Consider (P)*

$$\max\{\boldsymbol{c}^\top \boldsymbol{x} : A\boldsymbol{x} \leqslant \boldsymbol{b},\ \boldsymbol{x} \geqslant qslant\boldsymbol{0}\}$$

*and (P)'s dual (D)*

$$\min\{\boldsymbol{b}^\top \boldsymbol{y} : A^\top \boldsymbol{y} \geqslant qslant\boldsymbol{c},\ \boldsymbol{y} \geqslant qslant\boldsymbol{0}\}$$

*Let $\bar{\boldsymbol{x}}$ be a feasible solution for (P) and $\bar{\boldsymbol{y}}$ be a feasible solution for (D). Then*
*(1) $\boldsymbol{c}^\top \bar{\boldsymbol{x}} \leqslant \boldsymbol{b}^\top \bar{\boldsymbol{y}}$*
*(2) if $\boldsymbol{c}^\top \bar{\boldsymbol{x}} = \boldsymbol{b}^\top \bar{\boldsymbol{y}}$, then $\bar{\boldsymbol{x}}$ is an optimal solution for (P).*

*Proof.* Let $\bar{\boldsymbol{x}}$ be a feasible solution of (P) and let $\bar{\boldsymbol{y}}$ be a feasible solution of (D). Then

$$\begin{aligned} \boldsymbol{c}^\top \bar{\boldsymbol{x}} &\leqslant (\bar{\boldsymbol{y}}^\top A)\bar{\boldsymbol{x}} \\ &= \bar{\boldsymbol{y}}^\top (A\bar{\boldsymbol{x}}) \\ &\leqslant \bar{\boldsymbol{y}}^\top \boldsymbol{b} \end{aligned}$$

If $\boldsymbol{c}^\top \bar{\boldsymbol{x}} = \boldsymbol{b}^\top \bar{\boldsymbol{y}}$ it follows that $\bar{\boldsymbol{x}}$ is optimal for (P). □

**THEOREM 4.1.2.** *Let (P) and (D) be a pair of primal-dual LPs. Then*
*(1) if (P) is unbounded, then (D) is infeasible*
*(2) if (D) is unbounded, then (P) is infeasible*
*(3) if (P) and (D) are both feasible, then they both have optimal solutions*

*Proof.* (1) We prove the contrapositive, that is we prove: [(D) feasible $\implies$ (P) not unbounded]

Suppose that (D) is feasible. By Weak Duality theorem, we know that $\boldsymbol{b}^\top \boldsymbol{y}$ is an upper bound on (P). Thus, (P) is not unbounded.

(2) We prove the contrapositive, that is we prove: [(P) feasible $\implies$ (D) not unbounded]

Suppose that (P) is feasible. By Weak Duality theorem, we know that $\boldsymbol{c}^\top \boldsymbol{x}$ is a lower bound on (D). Thus, (D) is not unbounded.

(3) Assume (P) and (D) are both feasible. By (1) and (2) we know that (P) and (D) are both not unbounded. By Fundamental Theorem of Linear Programming, we know that exactly one of the following holds from (I)-(III):

(I) (P) and (D) are infeasible

(II) (P) and (D) are unbounded

(III) (P) and (D) both have optimal solutions

Clearly, we know that (I) and (II) both do not hold in this part of the proof, thus (P) and (D) both have optimal solutions. □

## 4.2 Strong Duality

**THEOREM 4.2.1** (Strong Duality)**.** *Let (P) and (D) be a pair of primal-dual LPs. Then*
*(1) If there exists an optimal solution $\bar{\boldsymbol{x}}$ of (P), then there exists an optimal solution $\bar{\boldsymbol{y}}$ of (D).*
*(2) The value of $\bar{\boldsymbol{x}}$ in (P) equals the value of $\bar{\boldsymbol{y}}$ in (D).*

## 4.3 A Geometric Characterization of Optimality

### 4.3.1 Complementary Slackness

Recall our proof of Weak Duality. Then for LPs in SEF: $\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}}$ are feasible in (P) and (D) respectively. $\bar{\boldsymbol{x}}$ is optimal in (P), $\bar{\boldsymbol{y}}$ is optimal in (D) if and only if

$$\boldsymbol{c}^\top \bar{\boldsymbol{x}} = (A^\top \bar{\boldsymbol{y}})^\top \bar{\boldsymbol{x}} = \bar{\boldsymbol{y}}^\top (A\bar{\boldsymbol{x}}) = \bar{\boldsymbol{y}}^\top \boldsymbol{b}$$

The first equality came from $(A^\top \boldsymbol{y})^\top = \boldsymbol{c}^\top$, and the last equality came from $A\bar{\boldsymbol{x}} = \boldsymbol{b}$ (check Weak Duality Theorem - Special Form). That is, if and only if

$$\bar{\boldsymbol{x}}^\top (A^\top \bar{\boldsymbol{y}} - \boldsymbol{c}) = 0$$

and

$$\bar{\boldsymbol{y}}^\top(\boldsymbol{b} - A\bar{\boldsymbol{x}}) = 0$$

That is, if and only if $\forall j \in \{1, \dots, n\}$ either $x_j = 0$ or $(A^\top \bar{\boldsymbol{y}} - \boldsymbol{c})_j = 0$ possibly both, and $\forall i \in \{1, \dots, m\}$ either $y_i = 0$ or $(\boldsymbol{b} - A\bar{\boldsymbol{x}})_i = 0$ possibly both. We call these the Complementary Slackness Conditions (CS).

**THEOREM 4.3.1** (Complementary Slackness)**.** *Let (P) and (D) be an arbitrary primal-dual pair. Let $\bar{\boldsymbol{x}}$ be a feasible solution to (P) and let $\bar{\boldsymbol{y}}$ be a feasible solution to (D). Then, $\bar{\boldsymbol{x}}$ is an optimal solution to (P) and $\bar{\boldsymbol{y}}$ is an optimal solution to (D) if and only if the complementary slackness conditions hold.*

**EXAMPLE 4.3.2** (Complementary Slackness)**.** (P)

$$\max \begin{bmatrix} -2 & -1 & 0 \end{bmatrix} \boldsymbol{x}$$

subject to

$$\begin{bmatrix} 1 & 3 & 2 \\ -1 & 4 & 2 \end{bmatrix} \boldsymbol{x} \begin{matrix} \geqslant \\ \leq \end{matrix} \begin{bmatrix} 5 \\ 7 \end{bmatrix}$$

$$x_1 \leqslant 0,\ x_2 \geqslant 0,\ x_3 \text{ free}$$

(1) Write the dual (D) of (P)
(2) Write the complementary slackness (CS) conditions for (P) and (D)
(3) Use weak duality to prove that $\bar{\boldsymbol{x}}$ is optimal for (P) and $\bar{\boldsymbol{y}}$ is optimal for (D)
(4) Use CS to prove that $\bar{\boldsymbol{x}}$ is optimal for (P) and $\bar{\boldsymbol{y}}$ is optimal for (D)
and

$$\bar{\boldsymbol{x}} = (-1, 0, 3)^\top \qquad \bar{\boldsymbol{y}} = (-1, 1)^\top$$

**Solution.**
(1)
(D)

$$\min \begin{bmatrix} 5 & 7 \end{bmatrix}$$

subject to

$$\begin{bmatrix} 1 & -1 \\ 3 & 4 \\ 2 & 2 \end{bmatrix} \boldsymbol{y} \begin{matrix} \geqslant \\ \leq \\ = \end{matrix} \begin{bmatrix} -2 \\ -1 \\ 0 \end{bmatrix}$$

$$y_1 \leqslant 0,\ y_2 \geqslant 0,\ y_3 \text{ free}$$

(2)
- $x_1 = 0$ OR $\boxed{y_1 - y_2 = -2}$
- $\boxed{x_2 = 0}$ OR $3y_1 + 4y_2 = -1$
- $y_1 = 0$ OR $\boxed{x_1 + 3x_2 + 2x_3 = 5}$
- $y_2 = 0$ OR $\boxed{-x_1 + 4x_2 + 2x_3 = 7}$

(3) Verify that $\bar{\boldsymbol{x}}$ and $\bar{\boldsymbol{y}}$ are feasible for (P) and (D), and check $\boldsymbol{c}^\top \bar{\boldsymbol{x}} = \boldsymbol{b}^\top \bar{\boldsymbol{y}}$.
(4) By Complementary Slackness Theorem, this is trivially true as seen boxed above.

### 4.3.2 Geometry

(P)

$$\max 2x_1 + x_2$$

subject to

$$\begin{bmatrix} 1 & 1 \\ -1 & 0 \\ 0 & -1 \end{bmatrix} \boldsymbol{x} \leq \begin{bmatrix} 4 \\ 0 \\ 0 \end{bmatrix}$$

(D)

$$\min 4y_1$$

subject to

$$\begin{bmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix} \boldsymbol{y} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$y_1,\, y_2,\, y_3 \geqslant 0$$

Let $\bar{\boldsymbol{x}}$ be a feasible solution of (P) and $J(\bar{\boldsymbol{x}})$ denote the indices of the tight constraints at $\bar{\boldsymbol{x}}$.

Then, $\bar{\boldsymbol{x}}$ is optimal in (P) if and only if $c$ is a non-negative linear combination of tights rows ($A^=$).

Suppose $\bar{\boldsymbol{x}} = (4,0)^\top$. Then $A^= = \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix}$. That is, $\bar{\boldsymbol{x}}$ is tight at $\mathrm{row}_1(A)$ and $\mathrm{row}_3(A)$.

$J(\bar{\boldsymbol{x}}) := \{1,3\}$, so $\bar{\boldsymbol{x}}$ is optimal if and only if $\exists\, \bar{y}_1,\, \bar{y}_2$ such that

$$\boldsymbol{c}^\top = \bar{y}_1 \left[\mathrm{row}_1(A)\right] + \bar{y}_2 \left[\mathrm{row}_3(A)\right]$$

---

2019-11-05

---

**DEFINITION 4.3.3.** Let $a^{(1)}, \dots, a^{(k)} \in \mathbb{R}^n$. We define the *cone* generated by $a^{(1)}, \dots, a^{(k)}$ to be the set

$$C = \text{cone}\left\{a^{(1)}, \dots, a^{(k)} \in \mathbb{R}^n\right\} = \left\{\sum_{i=1}^{k} \lambda_i a^{(i)} : \lambda_i \geqslant 0,\, \forall i \in \{1, \dots k\}\right\}$$

We define the cone generated by an empty set to be $\{\mathbf{0}\}$.

**DEFINITION 4.3.4.** Let $P := \{\boldsymbol{x} : A\boldsymbol{x} \leqslant \boldsymbol{b}\}$ and let $\bar{\boldsymbol{x}} \in P$. Let $J(\bar{\boldsymbol{x}})$ be the row indices of $A$ corresponding to the tight constraints of $A\boldsymbol{x} \leqslant \boldsymbol{b}$, that is, $i \in J(\bar{\boldsymbol{x}})$ if and only if $\mathrm{row}_i(A)\bar{\boldsymbol{x}} = b_i$. We define the *cone of tight constraints* for $\bar{\boldsymbol{x}}$ to be the cone $C$ generated by the rows of $A$ corresponding to the tight constraints, that is

$$C = \text{cone}\left\{\mathrm{row}_i(A)^\top : i \in J(\bar{\boldsymbol{x}})\right\}$$

**THEOREM 4.3.5.** *Let $\bar{\boldsymbol{x}}$ be a feasible solution to $\max\left\{\boldsymbol{c}^\top \boldsymbol{x} : A\boldsymbol{x} \leqslant \boldsymbol{b}\right\}$. Then $\bar{\boldsymbol{x}}$ is optimal if and only if $c$ is in the cone of tight constraints for $\bar{\boldsymbol{x}}$.*

# Chapter 5

# Duality Through Examples

2019-11-12

## 5.1 The Shortest Path Problem

Given a graph $G = (V, E)$, two distinct distinguished nodes $s, t \in V$, $c_e \geqslant 0 \,\forall e \in E$, we want to find a shortest path from $s$ to $t$

**DEFINITION 5.1.1.** Let $G = (V, E)$ be a graph, and let $U \subseteq V$. We let $\delta(U)$ denote the set of edges that have exactly one endpoint in $U$. That is,

$$\delta(U) = \{uv \in E : u \in U,\, v \in U\}$$

**DEFINITION 5.1.2.** Let $G = (V, E)$ be a graph, and let $U \subseteq V$. Suppose $G$ has a distinct pair of vertices $s$ and $t$. An *$s$-$t$ cut* is the set of edges of the form $\delta(U)$ where $s \in U$, $t \notin U$.

**DEFINITION 5.1.3.** Let $G = (V, E)$ be a graph. Suppose $G$ has a distinct pair of vertices $s$ and $t$. An *$s$-$t$ path* $\mathcal{P}$ in $G$ is the following sequence of edges of $G$

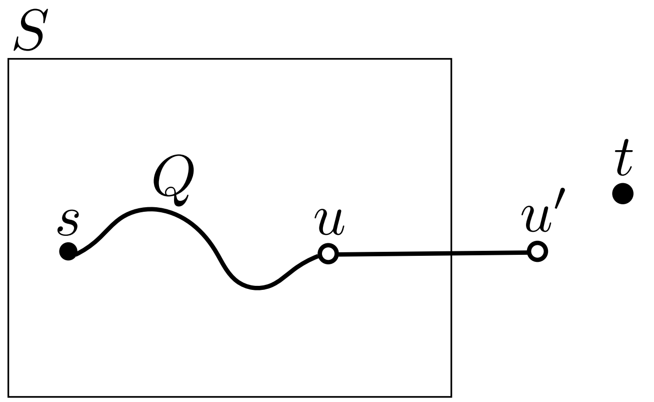
$$\{v_1, \ldots, v_k\}$$

where $v_1 = s$, $v_k = t$, and $v_i \neq v_j$ for $i \neq j$ (as $s$ and $t$ are distinct).

**THEOREM 5.1.4.** *An $s$-$t$ path $\mathcal{P}$ intersects every $s$-$t$ cut.*

*Proof.* Let $G = (V, E)$ be a graph with a distinct pair of vertices $s$ and $t$, let $\mathcal{P}$ be an $s$-$t$ path of $G$, let $U \subseteq V$, and let $\delta(U)$ be an arbitrary $s$-$t$ cut of $G$. Follow the path $\mathcal{P}$ starting from $s$ to $u$ where $u$ is the last vertex of $\mathcal{P}$ in $U$, and denote $u'$ the vertex that follows $u$ in $\mathcal{P}$. Note that $u$ exists since $s \in U$, $t \notin U$. Then by definition $uu'$ is an edge that is in $\delta(U)$. $\square$

**THEOREM 5.1.5.** *Let $S \subseteq E$ be a set of edges that contains at least one edge from every $s$-$t$ cut. Then there exists an $s$-$t$ path $\mathcal{P}$ that is contained in the edges of $S$.*




*Proof.* Let $S \subseteq E$ be a set of edges that contains at least one edge from every $s$-$t$ cut. Let $U$ be the set of vertices that contain $s$ as well as all vertices $u$ for which there exists a path from $s$ to $u$ only using edges in $S$. We need to show that there exists an $s$-$t$ path that is contained in the set of edges of $S$, i.e. $t \in U$. Suppose for a contradiction that $t \notin U$. We know $\delta(U)$ is an $s$-$t$ cut by definition. By our hypothesis, there exists an edge $uu' \in S$, where $u \in U$, $u' \notin U$. By construction, there exists and $s$-$u$ path $Q$ that is contained in $S$. Then the path obtained from $Q$ by adding an edge $uu'$ is an $s$-$u'$ path contained in $S$. By definition of $U$, we have that $u' \in U$, contradiction. □

Suppose $G$ contains an $s$-$t$ path. Let $S \subseteq E$ be a set of edges that contains at least one edge from every $s$-$t$ cut. Then such $S$ contains an $s$-$t$ path.

$$x_e := \begin{cases} 1, & \text{if edge } e \text{ is in the shortest } s\text{-}t \text{ path} \\ 0, & \text{otherwise} \end{cases}$$

(IP)

$$\min \sum_{e \in E} c_e x_e$$

subject to

$$\begin{array}{rcll} \sum_{e \in \delta(U)} x_e & \geqslant & 1 & (U \subseteq V,\, s \in U,\, t \notin U) \\ x_e & \geqslant & 0 & (e \in E) \\ x_e & \in & \mathbb{Z} & (e \in E) \end{array}$$

If $c_e > 0$ $(e \in E)$ then optimal solutions of this IP correspond to shortest $s$-$t$ paths. If some $c_e = 0$ $(e \in E)$, then some optimal solutions will correspond to sets like set $S$ above which contains a shortest $s$-$t$ path.

Let (P) denote the LP relaxation of (IP) (replace $x_e\{0, 1\}$ with $0 \leqslant x_e \leqslant 1$ $(e \in E)$). Write down the dual of (P): (D)

$$\max \sum (y_U : \delta(U) \text{ is an } s\text{-}t \text{ cut})$$

subject to

$$\begin{array}{rcll} \sum (y_U : \delta(U) \text{ is an } s\text{-}t \text{ cut containing } e) & \leqslant & c_e & (e \in E) \\ y_U & \geqslant & 0 & (U \subseteq V,\, s \in U,\, t \notin U) \end{array}$$

A consequence of Complementary Slackness for a shortest path problem is: Let $\mathcal{P}$ be an $s$-$t$ path (as set of edges) and let $\bar{y}$ be a feasible solution of (D). Suppose

- every edge in $\mathcal{P}$ corresponds to an equality edge $\Sigma(\bar{y}_U : \delta(U)$ is an $st$-cut containing $e\ ) = c_e$
- for every active cut (i.e. $st$-cut $\delta(U)$ such that $\bar{y}_U > 0$, $\mathcal{P}$ must contain at least one edge from that $st$-cut.

Then $\mathcal{P}$ is a shortest $s$-$t$ path.

**REMARK 5.1.6.** $\bar{y}_U := 0$ for all $s$-$t$ cuts $\delta(U)$ gives a feasible solution of (D)

## 5.2 An Algorithm

Given $y$, $\text{slack}_y(e)$ is

$$c_e - \Sigma(y_U : \delta(U) \text{ is a } s\text{-}t \text{ cut containing } e)$$

That is, $\text{slack}_y(e)$ is the length of $e$ minus the total width of all $s$-$t$ cuts using $e$.

**Algorithm 3:** Shortest path

**Input :** Graph $G = (V, E)$, costs $c_e \geqslant 0$ for all $e \in E$, $s, t \in V$, where $s \neq t$.
**Output:** A shortest $s$-$t$ path $\mathcal{P}$.

$y_w := 0$ for all $s$-$t$ cuts $\delta(W)$. Set $U := \{s\}$
**while** $t \notin U$ **do**
  Let $ab$ be an edge in $\delta(U)$ of smallest slack for $y$ where $a \in U$, $b \notin U$
  $y_U := \text{slack}_y(ab)$
  $U := U \cup \{b\}$
  change edge $ab$ into an arc $\overrightarrow{ab}$
**end**
**return** A directed $s$-$t$ path $\mathcal{P}$.

This is similar to Dijkstra's shortest path algorithm, but out algorithm also generates an optimality certificate.

2019-11-14

## 5.3 Minimum Cost Perfect Matching Problem in Bipartite Graphs

Given a biparite graph $G = (V, E)$ with biparition $[W, J]$ where $W \cap J = \varnothing$ and $W \cup J = V$ with cost $c_e \in \mathbb{R}$ for each $e \in E$, we want to find a perfect matching of minimum cost.

What are necessary conditions for having a perfect matching in $G$?

- $|W| = |J| = \frac{|V|}{2}$
- For every $S \subseteq W$, the *neighbour set of* $S$,

$$\mathcal{N}(S) := \{v \in V : uv \in E, u \in S, v \notin S\}$$

  has as many vertices as $S$.

**THEOREM 5.3.1** (Hall's Theorem)**.** *Let $G$ be a bipartite graph with biparition $[W, J]$. Suppose $|W| = |S|$, then $G$ has a perfect matching if and only if*

$$|S| \leqslant |\mathcal{N}(S)|$$

*for each $S \in W$.*

Therefore, if $G$ does not have a perfect matching, then $G$ has a *deficient set*; a subset $S \subseteq W$ such that

$$|\mathcal{N}(S)| < |S|$$

Moreover, there is an efficient algorithm to find such deficient sets.

(IP)

$$\min \sum_{e \in E} c_e x_e$$

subject to

$$\begin{array}{rcll} \sum_{e \in \delta(V)} x_e & = & 1 & \forall v \in V \\ \boldsymbol{x} & \geqslant & \mathbf{0} & \\ \boldsymbol{x} & \in & \mathbb{Z}^E & \end{array}$$

Recall the meaning of $x_e$:

$$x_e := \begin{cases} 1 & \text{if edge } e \text{ is in the minimum cost perfect matching} \\ 0 & \text{otherwise} \end{cases}$$

(D)

$$\max \sum_{v \in V} y_v$$

subject to

$$y_u + y_v \leqslant c_{uv} \qquad \forall uv \in E$$

$$\bar{y} := \min_{uv \in E} \frac{\{c_{uv}\}}{2}$$

is a feasible solution for (D).

Complementary Slackness Theorem implies: a feasible solution $\bar{y}$ of (D) is optimal if and only if there exists a feasible solution $\bar{x}$ of (P) such that

$$\bar{\boldsymbol{x}} > 0 \implies \bar{y}_u + \bar{y}_v = c_{uv}$$

Let $H = (V, E(\bar{y}))$,

$$E(\bar{y}) = \{uv \in E : \bar{y}_u + \bar{y}_v = c_{uv}\}$$

if we can find a perfect matching in $H$, then that perfect matching is an optimal solution of (IP).

If $H$ does not have a perfect matching, then $H$ has a deficient set $S \subseteq W$.

Modify our dual solution:

$$\bar{y}_v := \begin{cases} \bar{y}_v + \varepsilon & v \in S \\ \bar{y}_v - \varepsilon & v \in \mathcal{N}_H(S) \\ \bar{y}_v & \text{otherwise} \end{cases}$$

Choose

$$\varepsilon := \min\{c_{uv} - \bar{y}_u + \bar{y}_v : uv \in E, u \in S, v \notin \mathcal{N}_H(S)\}$$

If no such $\varepsilon$ exists, ($\mathcal{N}_H(S) = \mathcal{N}(S)$), then $S$ is a definite set in $G$; moreover, (IP) and (P) are infeasible and (D) is unbounded.

**Algorithm 4:** Minimum Cost Perfect Matching in Bipartite Graphs

**Input** : Biparite graph $G = (V, E)$ with biparition $[W, J]$ such that $|W| = |J|$, $c_e \in \mathbb{R}$, for each $e \in E$

**Output:** A minimum cost perfect matching, or a deficient set $S$.

1 $\bar{y}_v := 1/2 \min_{e \in E}\{c_e\}$ for all $v \in V$

2 Construct $H = (V, E(\bar{y}))$

3 Find a perfect matching in $H$. If yes, STOP; we have a minimum cost perfect matching. Otherwise, find a deficient set $S$ in $H$.

4 If $\nexists\, uv \in E$ such that $u \in S$, $v \notin \mathcal{N}_H(S)$, then STOP; $G$ has no perfect matching, $S$ is a deficient set in $G$.

5 $\varepsilon := \min\{c_{uv} - \bar{y}_u + \bar{y}_v : uv \in E, u \in S, v \notin \mathcal{N}_H(S)\}$

6 $\bar{y}_v := \begin{cases} \bar{y}_v + \varepsilon & v \in S \\ \bar{y}_v - \varepsilon & v \in \mathcal{N}_H(S) \\ \bar{y}_v & \text{otherwise} \end{cases}$

7 Go to step 2.

Note that in every iteration, $\varepsilon > 0$ and the objective value goes up by

$$\varepsilon(|S| - |\mathcal{N}_H(S)|) > 0$$

Note that when the algorithm stops with a minimum cost perfect matching, the current $\bar{y}$ is a certificate of optimality.

## 2020-11-19

Left class early due to questionable lecturer.

## 2019-11-21

**DEFINITION 5.3.2.** Given a set $S \subseteq \mathbb{R}^n$, the *convex hull* of $S$ is the smallest convex set which contains $S$. Equivalently,

$$\operatorname{conv}(S) = \bigcap C$$

where $C \subseteq S$ is convex.

Let $S_1 \supseteq S_2$, $c \in \mathbb{R}^n$ and consider

$$(P_2)\ \max\{\boldsymbol{c}^\top \boldsymbol{x} : \boldsymbol{x} \in S_2\}$$

$$(P_1)\ \max\{\boldsymbol{c}^\top \boldsymbol{x} : \boldsymbol{x} \in S_1\}$$

If we have an optimal solution $\bar{\boldsymbol{x}}$ of $(P_1)$ and $\bar{\boldsymbol{x}} \in S_2$, then $\bar{\boldsymbol{x}}$ is an optimal solution of $(P_2)$. Regardless of whether $\bar{\boldsymbol{x}} \in S_2$, $\boldsymbol{c}^\top \bar{\boldsymbol{x}}$ is an upper bound on the optimal value of $(P_2)$.

$$(IP)\ \max \boldsymbol{c}^\top \boldsymbol{x}$$

subject to

$$\begin{array}{rcl} A\boldsymbol{x} & = & \boldsymbol{b} \\ \boldsymbol{x} & \geqslant & \boldsymbol{0} \\ \boldsymbol{x} & \in & \mathbb{Z}^n \end{array}$$

$$(LP)\ \max \boldsymbol{c}^\top \boldsymbol{x}$$

subject to

$$\begin{array}{rcl} A\boldsymbol{x} & = & \boldsymbol{b} \\ \boldsymbol{x} & \geqslant & \mathbf{0} \end{array}$$

Find an optimal solution $\bar{\boldsymbol{x}}$ of $(LP)$. If $\bar{\boldsymbol{x}} \in \mathbb{Z}^n$, then $\bar{\boldsymbol{x}}$ is optimal in $(IP)$. Otherwise, find a cut ($\bar{\boldsymbol{x}} \in \mathbb{R}^n \setminus \mathbb{Z}^n$)

$$\boldsymbol{a}^\top \boldsymbol{x} \geqslant \alpha$$

such that

(i) $\boldsymbol{a}^\top \boldsymbol{x} \leqslant \alpha$

(ii) $\boldsymbol{a}^\top \bar{\boldsymbol{x}} > \alpha$

for all $\boldsymbol{x}$ feasible in $(IP)$. The inequality "cuts" the current optimal solution $\bar{\boldsymbol{x}}$ of $(LP)$.

**EXAMPLE 5.3.3** (Cutting Plane Algorithm).

$$(IP)\max x_2$$

subject to

$$\begin{array}{rcl} 3x_1 + 2x_2 & \leqslant & 6 \\ -3x_1 + 2x_2 & \leqslant & 0 \\ \boldsymbol{x} = (x_1, x_2)^\top & \geqslant & \mathbf{0} \\ \boldsymbol{x} & \in & \mathbb{Z}^2 \end{array}$$

Introduce slack variables $x_3, x_4 \in \mathbb{Z}_{\geqslant 0}$. Then solve the LP relaxation:

$$\begin{array}{ccccccccc} z & = & & -\frac{1}{4}x_3 & - & \frac{1}{4}x_4 & + & \frac{3}{2} \\ x_1 & & + & \frac{1}{6}x_3 & - & x_4 & = & 1 \\ x+2 & & + & \frac{1}{4}x_3 & + & \frac{1}{4}x_3 & = & \frac{3}{2} \end{array}$$

For every feasible solution of $(IP)$,

$$\begin{aligned} x_3 + \frac{1}{4}x_3 + \frac{1}{4}x_4 &= \frac{3}{2} \\ \implies x_2 + \left\lfloor \frac{1}{4} \right\rfloor x_3 + \left\lfloor \frac{1}{4} \right\rfloor x_4 &\leqslant \frac{3}{2} \end{aligned}$$

Since there are no integers in $(1, 3/2)$, every feasible solution of the $(IP)$

$$x_2 \leqslant \left\lfloor \frac{3}{2} \right\rfloor = 1$$

we call this a *cut* since:

(i) we proved above

(ii) $\bar{\boldsymbol{x}}_2 = 3/2 > 1$

Add the constraints $x_2 + x_5 = 1$, with $x_5 \geqslant 0$ to the LP relaxation and solve.

$$\begin{array}{ccccccccccc} z & = & & & & & - & x_5 & + & 1 \\ x_1 & & & & - & \frac{1}{3}x_4 & + & \frac{2}{5}x_5 & = & \frac{2}{3} \\ & x_2 & & & & & + & x_5 & = & 1 \\ & & x_3 & + & & x_4 & - & 4x_5 & = & 2 \end{array}$$

Add the constraints $x_1 + x_4 + x_6 = 0$ with $x_6 \geqslant 0$ to the LP relaxation and solve.

$$\begin{array}{ccccccccccc} z & = & & & - & x_5 & & & + & 1 \\ x_1 & & & & + & x_5 & - & \frac{1}{2}x_6 = & = & 1 \\ & x_2 & & & + & x_5 & & & = & 1 \\ & & x_3 & & - & 5x_5 & + & \frac{3}{2}x_6 & = & 1 \end{array}$$

$\boldsymbol{x}^* = (1, 1)^\top$ with objective value $z = 1$ is optimal in $(LP)$.

Another idea for solving IPs is **Branch-and-Bound** (related to *Divide-and-Conquer*). Seperate the problem at hand into exaustive and mutually exclusive subproblems (*Branching*). For each subproblem, solve its relaxation and get an upper bound on the optimal objective value of the subproblem. If the upper bound is less than the objective value of the current best integer solution, fathom this branch.

# Chapter 6

# Non-linear Optimization

2019-11-26

All linear programs are non-linear programs.

## 6.1 Convexity

**DEFINITION 6.1.1.** Let $f : \mathbb{R}^n \to \mathbb{R}$. Then $\bar{\boldsymbol{x}} \in \mathbb{R}^n$ is a *local minimizer* if for some $\varepsilon \in \mathbb{R}_{>0}$ we have that $f(\boldsymbol{x}) \geqslant f(\bar{\boldsymbol{x}})$ where

$$||\boldsymbol{x} - \bar{\boldsymbol{x}}|| < \varepsilon$$

for all $\boldsymbol{x} \in \mathbb{R}^n$.

**DEFINITION 6.1.2.** Let $f : \mathbb{R}^n \to \mathbb{R}$. Then $\bar{\boldsymbol{x}} \in \mathbb{R}^n$ is a *global minimizer* if for each $\boldsymbol{x} \in \mathbb{R}^n$,

$$f(\boldsymbol{x}) \geqslant f(\bar{\boldsymbol{x}})$$

### 6.1.1 Convex Functions and Epigraphs

**DEFINITION 6.1.3.** Let $C \subseteq \mathbb{R}^n$ be a convex set. Let $f : C \to \mathbb{R}$. $f$ is *convex* if for each $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ and $\lambda \in [0, 1]$

$$f(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \leqslant \lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y})$$

**DEFINITION 6.1.4.** Let $f : \mathbb{R}^n \to \mathbb{R}$. The *epigraph* of $f$ is the following set:

$$\operatorname{epi}(f) := \left\{ \begin{pmatrix} \mu \\ \boldsymbol{x} \end{pmatrix} \in \mathbb{R} \times \mathbb{R}^n : f(\boldsymbol{x}) \leqslant \mu \right\}$$

**THEOREM 6.1.5.** *Let $f : \mathbb{R}^n \to \mathbb{R}$ be convex. $f$ is convex if and only if epi$(f)$ is convex.*

*Proof.* ($\Longrightarrow$) Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is convex. Let $\begin{pmatrix} \mu_1 \\ \boldsymbol{x} \end{pmatrix}, \begin{pmatrix} \mu_2 \\ \boldsymbol{y} \end{pmatrix} \in \text{epi}(f)$ and $\lambda \in [0,1]$. We have

$$f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \leqslant \lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y}) \leqslant \lambda \mu_1 + (1-\lambda)\mu_2$$

which implies

$$\begin{pmatrix} \lambda \mu_1 + (1-\lambda)\mu_2 \\ \lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y} \end{pmatrix} \in \text{epi}(f)$$

($\Longleftarrow$) Suppose epi$(f)$ is convex. Let $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ and $\lambda \in [0,1]$. Then, $\begin{pmatrix} f(\boldsymbol{x}) \\ \boldsymbol{x} \end{pmatrix}, \begin{pmatrix} f(\boldsymbol{y}) \\ \boldsymbol{y} \end{pmatrix} \in \text{epi}(f)$. Hence,

$$\lambda \begin{pmatrix} f(\boldsymbol{x}) \\ \boldsymbol{x} \end{pmatrix} + (1-\lambda) \begin{pmatrix} f(\boldsymbol{y}) \\ \boldsymbol{y} \end{pmatrix} \in \text{epi}(f)$$

Hence, by definition of epi$(f)$,

$$f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \leqslant \lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y})$$

Thus, $f$ is a convex function. □

### 6.1.2 Level Sets and Feasible Region

**DEFINITION 6.1.6.** Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function and $\beta \in \mathbb{R}$. The *level set* of $f$ is:

$$S_\beta = \{\boldsymbol{x} \in \mathbb{R}^n : f(\boldsymbol{x}) \leqslant \beta\}$$

**THEOREM 6.1.7.** *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex and $\beta \in \mathbb{R}$. Then the level set is a convex set.*

*Proof.* Since $f$ is convex, we know that

$$\begin{aligned} f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) &\leqslant \lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y}) \\ &\leqslant \lambda \beta + (1-\lambda)\beta \\ &= \beta \end{aligned}$$

Thus, $\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y} \in S_\beta$. Thus, $S_\beta$ is a convex set. □

---

2019-11-28

---

## 6.2 Relaxing Convex NLPs

### 6.2.1 Subgradients

**DEFINITION 6.2.1.** Let $g : \mathbb{R}^n \to \mathbb{R}$ be convex, and let $\bar{\boldsymbol{x}} \in \mathbb{R}^n$. We say that $\boldsymbol{s} \in \mathbb{R}$ is a *subgradient* at $\bar{\boldsymbol{x}}$ if for every $\boldsymbol{x} \in \mathbb{R}^n$ the following inequality holds:

$$g(\bar{\boldsymbol{x}}) + \boldsymbol{s}^\top (\boldsymbol{x} - \bar{\boldsymbol{x}}) \leqslant g(\boldsymbol{x})$$

$\{\boldsymbol{x} \in \mathbb{R}^n : g(\boldsymbol{x}) \leqslant 0\} \subseteq \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{s}^\top \boldsymbol{x} \leqslant \boldsymbol{s}^\top \bar{\boldsymbol{x}} - g(\bar{\boldsymbol{x}})\}$ Consider the NLP

$$(P_C) \min \boldsymbol{c}^\top \boldsymbol{x}$$

subject to

$$g_i(x) \leqslant 0$$

$g_i$ convex for all $i \in 1, \ldots, m$.

Let $\bar{\boldsymbol{x}}$ be feasible in $(P_C)$: $g_i(\boldsymbol{x}) \leqslant 0$ for all $i \in \{1, \ldots, m\}$.

Consider the LP problem

$$(LP_C) \min \boldsymbol{c}^\top \boldsymbol{x}$$

subject to

$$(\boldsymbol{s}^{(i)})^\top \boldsymbol{x} \leqslant (\boldsymbol{s}^{(i)})^\top \bar{\boldsymbol{x}} - g_i(\bar{\boldsymbol{x}})$$

$$s^{(i)} \in \mathbb{R}^n \text{ is a subgradient of } g_i \text{ at } \bar{\boldsymbol{x}}$$

For $(LP_C)$, we know that $\bar{\boldsymbol{x}}$ is optimal if and only if

$$-c \in \{\text{cone of tight constraints at } \bar{\boldsymbol{x}}\}$$

If this condition holds, then $\bar{\boldsymbol{x}}$ is optimal in $(LP_C)$ as well as $(P_C)$.

**THEOREM 6.2.2.** *Let $\bar{\boldsymbol{x}} \in \mathbb{R}^n$ be a feasible solution of $(P_C)$. Let $\boldsymbol{s}^{(i)} \in \mathbb{R}^n$ denote the subgradients of $g_i$ at $\bar{\boldsymbol{x}}$ for all $i \in J(\bar{\boldsymbol{x}})$, where $J(\bar{\boldsymbol{x}}) := \{i : g_i(\bar{\boldsymbol{x}}) = 0\}$. Then, if*

$$-c \in \textit{cone}\{s^{(i)} : i \in J(\bar{\boldsymbol{x}})\}$$

*$\bar{\boldsymbol{x}}$ is optimal in $(P_C)$.*

The convex function $f$ is differentiable at $\bar{\boldsymbol{x}}$, but not differentiable at $\hat{\boldsymbol{x}}$.

Let $f : \mathbb{R}^n \to \mathbb{R}$, let $\bar{\boldsymbol{x}} \in \mathbb{R}^n$, then $f$ is differentiable at $\bar{\boldsymbol{x}}$ if there exists $\boldsymbol{s} \in \mathbb{R}^n$ such that

$$\lim_{h \to 0} \frac{f(\bar{\boldsymbol{x}} + h) - f(\bar{\boldsymbol{x}}) - \boldsymbol{s}^\top h}{||h||_2} = 0$$

If $f$ is differentiable at $\bar{\boldsymbol{x}}$, then such an $\boldsymbol{s}$ is a subgradient of $f$ at $\bar{\boldsymbol{x}}$ (and it is unique). We denote such $\boldsymbol{s}$ by

$$\nabla f(\bar{\boldsymbol{x}})$$

gradient of $f$ at $\bar{\boldsymbol{x}}$. If $f$ is continuously differentiable at $\bar{\boldsymbol{x}}$, then

$$\nabla f(\bar{\boldsymbol{x}}) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix}$$

For example, consider $f : \mathbb{R}^2 \to \mathbb{R}$, $f(\boldsymbol{x}) = \frac{1}{4}{x_1}^2 + x_2^2$. What is the subgradient of $f$ at $\bar{\boldsymbol{x}} := (2, 1)^\top$?

$$\nabla f(\boldsymbol{x}) = (\frac{1}{2}x_1, 2x_2)^\top$$

The subgradient at $\bar{\boldsymbol{x}}$ is $\nabla f(\bar{\boldsymbol{x}}) = (1, 2)^\top$.

Consider the constraint $g(\boldsymbol{x}) = \frac{1}{4}x_1^2 + x_2^2 - 2 \leqslant 0$.

Consider a convex NLP:

$$\min f(\boldsymbol{x})$$

subject to

$$g_1(x) \leqslant 0$$
$$\vdots$$
$$g_m(x) \leqslant 0$$

Note that we can relax any constraint $g_i(\boldsymbol{x}) \leqslant 0$ of NLP by replacing it with

$$\boldsymbol{s}^\top \boldsymbol{x} \leqslant \boldsymbol{s}^\top \bar{\boldsymbol{x}} - g_i(\bar{\boldsymbol{x}})$$

where $\bar{\boldsymbol{x}}$ is a subgradient of $g_i$ at $\bar{\boldsymbol{x}}$. Note that since $f$ is convex, if it differentiable, then

$$f(\boldsymbol{x}) \geqslant f(\bar{\boldsymbol{x}}) + \nabla f(\bar{\boldsymbol{x}})(\boldsymbol{x} - \bar{\boldsymbol{x}})$$

In dealing with the convex NLP, we can use the LP problem

$$\min \nabla f(\bar{\boldsymbol{x}})^\top \boldsymbol{x} + f(\bar{\boldsymbol{x}}) - \nabla f(\bar{\boldsymbol{x}})^\top \bar{\boldsymbol{x}}$$

subject to

$$\nabla g_1(\bar{\boldsymbol{x}})^\top \leqslant g_1(\bar{\boldsymbol{x}})^\top - g_1(\bar{\boldsymbol{x}})$$
$$\vdots$$
$$\nabla g_m(\bar{\boldsymbol{x}})^\top \leqslant g_m(\bar{\boldsymbol{x}})^\top - g_m(\bar{\boldsymbol{x}})$$

$\hat{\boldsymbol{x}}$ is called a *Slater point* for the NLP if

$$g_i(\hat{\boldsymbol{x}}) < 0 \qquad \forall i \in \{1, \ldots, m\}$$

**THEOREM 6.2.3** (Karush-Kuhn-Tucker)**.** *Consider a convex NLP that has a Slater point. Let $\bar{\boldsymbol{x}} \in \mathbb{R}^n$ be a feasible solution and assume that $f, g_1, g_2, \ldots, g_m$ are differentiable at $\bar{\boldsymbol{x}}$. Then $\bar{\boldsymbol{x}}$ is an optimal solution of NLP if and only if*

$$-\nabla f(\bar{\boldsymbol{x}}) \in \mathit{cone}\{\nabla g_i(\bar{\boldsymbol{x}}) : i \in J(\bar{\boldsymbol{x}})\}$$

**EXAMPLE 6.2.4.** (NLP)

$$\min f(\boldsymbol{x}) := -x_1 + x_2$$
$$g_1(\boldsymbol{x}) := (x_1 - 1)^2 + (x_2 - 1)^2 - 1 \leqslant 0$$
$$g_2(\boldsymbol{x}) := x_1^2 + x_2^2 - 1 \leqslant 0$$

- Is $\tilde{\boldsymbol{x}} := (0, 1)^\top$ optimal?
- Is $\bar{\boldsymbol{x}} := (1, 0)^\top$ optimal?

**Solution.** We can see that $\hat{\boldsymbol{x}} := (1/2, 1/2)^\top$ is a Slater point.

$$-\nabla f(\bar{\boldsymbol{x}}) = (1, -1)^\top$$

$$g_1(\bar{\boldsymbol{x}}) = (0, -2)^\top, \ g_2(\bar{\boldsymbol{x}}) = (2, 0)^\top$$

Thus,

$$-\nabla f(\bar{\boldsymbol{x}}) = \frac{1}{2} g_1(\bar{\boldsymbol{x}}) + \frac{1}{2} g_2(\bar{\boldsymbol{x}})$$
$$\implies -\nabla f(\bar{\boldsymbol{x}}) \in \text{cone}\{\nabla g_1(\bar{\boldsymbol{x}}), g_2(\bar{\boldsymbol{x}})\}$$

Hence, $\bar{\boldsymbol{x}}$ is optimal in NLP.
<u>Exercise</u>: Show that $\tilde{\boldsymbol{x}}$ is not optimal.

## 2019-12-03

Graphical representation of last example was done.

**DEFINITION 6.2.5.** Given an NLP, let $F \subseteq \mathbb{R}^n$ denote its feasible region. $\bar{\boldsymbol{x}} \in F$ is called a *local minimizer* if $\exists\, \varepsilon > 0$ such that

$$f(\bar{\boldsymbol{x}}) < f(\boldsymbol{x})$$

for all $\boldsymbol{x} \in F \cap \{\boldsymbol{x} : ||\boldsymbol{x} - \bar{\boldsymbol{x}}||_2 < \varepsilon\}$

**DEFINITION 6.2.6.** Given an NLP, let $F \subseteq \mathbb{R}^n$ denote its feasible region. $\bar{\boldsymbol{x}} \in F$ is called a *global minimizer* if

$$f(\bar{\boldsymbol{x}}) < f(\boldsymbol{x})$$

for all $\boldsymbol{x} \in F$.

**THEOREM 6.2.7.** *For convex NLPs, every local minimizer is a global minimizer.*

*Proof.* Let $F$ denote the feasible region of the NLP, and let $\bar{\boldsymbol{x}} \in F$ be a local minimizer. Suppose for a contradiction that $\bar{\boldsymbol{x}}$ is not a global minimizer. Then there exists a $\hat{\boldsymbol{x}} \in F$ such that

$$f(\hat{\boldsymbol{x}}) < f(\bar{\boldsymbol{x}})$$

Since $f$ is convex, for all $\lambda \in [0, 1]$,

$$[\lambda\bar{\boldsymbol{x}} + (1 - \lambda)\hat{\boldsymbol{x}} \in F]$$

So,

$$f(\lambda\bar{\boldsymbol{x}} + (1 - \lambda)\hat{\boldsymbol{x}}) \leqslant \lambda f(\bar{\boldsymbol{x}}) + (1 - \lambda)f(\hat{\boldsymbol{x}}) < f(\bar{\boldsymbol{x}})$$

contradiction. $\square$

Interesting Future (Optimization) Courses

- CO 351: Network Flow Theory (Assignment, Shortest Path, Perfect Matching)
- CO 367: Non-linear Optimization
- CO 353: Computational Discrete Optimization (Transportation, Branch and Bound)
- CO 370: Deterministic OR Models (First Weeks of CO 250)
- CO 372: Portfolio Optimization (2019-09-19)
- CO 342: Graph Theory
- CO 454: Scheduling
- CO 456: Introduction to Game Theory (Dual Variables in Shortest Path, Allocation of Resources)
- *Advanced Courses* ($80\%+^*$):

  CO 450, 452, 463, 466, 471, 673