C&O 250

A Gentle Introduction to Optimization

Yuan Si
y3si@uwaterloo.ca
Combinatorics & Optimization
University of Waterloo

7/14/2024

Contents

Preface

- This textbook is an edited and restructured version of the C&O250 textbook, all rights belong to the Department of Combinatorics and Optimization, University of Waterloo. This textbook may not be used commercially or for profit without permission from the department and me.

- Since this textbook is completely made by me, there might be some mistakes, please bear with me. The textbook may not fully cover the content in the future C&O250 course, so please combine this with the official textbook.

- I would like to express my heartfelt gratitude to my professors, Kanstantsin Pashkovich and Vijay Bhattiprolu, for inspiring my deep interest in the field of Combinatorics and Optimization. Their guidance is invaluable.

- I am also deeply grateful to my friends, Andy Bukovcan and Shamar Phillips, whose irreplaceable assistance has greatly supported my study in CO.

Prerequisite

0.1 Matrix product

Definition: Matrix Multiplication

Let $A \in M_{m\times n}(\mathbb{F})$ and $B \in M_{n\times p}(\mathbb{F})$. We define the **matrix product** $AB = C$ to be the matrix $C \in M_{m\times p}(\mathbb{F})$, constructed as follows:

$$C = AB = A\begin{bmatrix}\vec{b}_1 & \vec{b}_2 & \cdots & \vec{b}_p\end{bmatrix} = \begin{bmatrix}A\vec{b}_1 & A\vec{b}_2 & \cdots & A\vec{b}_p\end{bmatrix}.$$

That is, the j^{th} column of C, $\vec{c}_j$, is obtained by multiplying the matrix A by the j^{th} column of the matrix B:

$$\vec{c}_j = A\vec{b}_j, \quad \text{for all } j = 1, \ldots, p.$$

Example

Given $A = \begin{bmatrix}1 & 2\\ 3 & 5\\ 8 & 7\end{bmatrix}, B = \begin{bmatrix}-1 & 3\\ 2 & -4\end{bmatrix}$, calculate the products of AB, BA if possible.

Solution:

- AB: since the # of columns of A is equal to the # of rows of B, this product is defined. Then, we have:

$$\begin{aligned}
AB &= \begin{bmatrix}1 & 2\\ 3 & 5\\ 8 & 7\end{bmatrix}\begin{bmatrix}-1 & 3\\ 2 & -4\end{bmatrix}\\
&= \begin{bmatrix}\begin{bmatrix}1 & 2\\ 3 & 5\\ 8 & 7\end{bmatrix}\begin{bmatrix}-1\\ 2\end{bmatrix} & \begin{bmatrix}1 & 2\\ 3 & 5\\ 8 & 7\end{bmatrix}\begin{bmatrix}3\\ -4\end{bmatrix}\end{bmatrix}\\
&= \begin{bmatrix}(1)(-1)+(2)(2) & (1)(3)+(2)(-4)\\ (3)(-1)+(5)(2) & (3)(3)+(5)(-4)\\ (8)(-1)+(7)(2) & (8)(3)+(7)(-4)\end{bmatrix}\\
&= \begin{bmatrix}3 & -5\\ 7 & -11\\ 6 & -4\end{bmatrix}
\end{aligned}$$

- The product BA is undefined, since B has 2 columns, A has 3 rows, $2 \neq 3$.

0.2 Inverse of a Matrix

Definition: Inverse of a Matrix

If an $n \times n$ matrix A is invertible, we refer to the matrix B such that $AB = I_n$ as the **inverse** of A. We denote the inverse of A by A^{-1}. The inverse of A satisfies:

$$AA^{-1} = A^{-1}A = I_n$$

Example

Determine whether the matrix $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ is invertible. If invertible, find inverse.

Solution: We will solve the super-augmented matrix:

$$[A|I_2] = \left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \end{array}\right]$$

By EROs, we can obtain:

$$\left[\begin{array}{cc|cc} 1 & 0 & -2 & 1 \\ 0 & 1 & 3/2 & -1/2 \end{array}\right]$$

Thus, we conclude that the inverse of A is:

$$A^{-1} = \begin{bmatrix} -2 & 1 \\ 3/2 & -1/2 \end{bmatrix}$$

Let's verify the calculations: $AA^{-1} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} -2 & 1 \\ 3/2 & -1/2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2$

There's also a quick way to calculate the inverse of a matrix in $\mathbb{F}^{2\times 2}$:

Let $B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, we have:

- If $ad - bc = 0$, it is not invertible.
- If $ad - bc \neq 0$, it is invertible, and the inverse is given by:

$$B^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -c \\ -b & a \end{bmatrix}$$

0.3 Linear Dependence/Independence

Definition: Linear Dependence

- **Linear Dependence**:

  we say that the vectors $\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_k \in \mathbb{F}^n$ are **linearly dependent** if there exist scalars $c_1, c_2, \cdots, c_k \in \mathbb{F}$, not all zero, such that

  $$c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_k\vec{v}_k = \mathbb{0}$$

  ($\mathbb{0}$ means $\vec{0}$, can fit any dimentions.)

- **Linear Independence**:

  we say that the vectors $\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_k \in \mathbb{F}^n$ are **linearly independent** if there do not exist scalars $c_1, c_2, \cdots, c_k \in \mathbb{F}$, not all zero, such that

  $$c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_k\vec{v}_k = \mathbb{0}$$

  Equivalently we say that $\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_k \in \mathbb{F}^n$ are **linearly indepeent** if the only solution to the equation:

  $$c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_k\vec{v}_k = \mathbb{0}$$

  is the **trivial solution** $c_1 = c_2 = \cdots = c_k = 0$.

- **Edge case**: the empty set $\emptyset$ is considered to be linearly independent.

Example of linearly indepeent

Determine whether the set $\mathbb{V} = \left\{ \binom{1}{0}, \binom{0}{1} \right\}$ is linearly indepeent.
Solution: let $c_1, c_2 \in \mathbb{F}$, such that $c_1\binom{1}{0} + c_2\binom{0}{1} = \mathbb{0}$. That is, $\binom{c_1}{c_2} = \mathbb{0}$, which implies $c_1 = c_2 = 0$, as the unique solution to that equation.
Therefore, the set $\mathbb{V}$ is linearly independent.

Example of linearly depeent

Determine whether the set $\mathbb{T} = \left\{ \binom{1}{0}, \binom{2}{0} \right\}$ is linearly indepeent.
Solution: let $c_1, c_2 \in \mathbb{F}$, such that $c_1\binom{1}{0} + c_2\binom{2}{0} = \mathbb{0}$. That is, $\binom{c_1+2c_2}{0} = \mathbb{0}$, we can indeed find some c_1, c_2, not both zero, satisfying this equation. For example, $c_1 = 2, c_2 = -1$. Therefore, the set $\mathbb{T}$ is linearly independent.

0.4 Matrix inequality

Definition: Matrix inequality

The matrix inequality is a short version of a set of inequalities:

$$Ax \underbrace{\leq}_{or \geq, =} b, A \in \mathbb{R}^{m\times n}, x \in \mathbb{R}^n, b \in \mathbb{R}^m$$

means

$$\begin{aligned} A_{row_1}x &\leq b_{row_1} \\ A_{row_2}x &\leq b_{row_2} \\ &\vdots \\ A_{row_m}x &\leq b_{row_m} \end{aligned}$$

Example

If we say:

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} x \leq \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}$$

it actually means:

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &\leq 4 \\ 4x_1 + 5x_2 + 6x_3 &\leq 5 \\ 7x_1 + 8x_2 + 9x_3 &\leq 6 \end{aligned}$$

0.5 Partial Derivative

Definition: Partial Derivative

Just like ordinary derivatives, the **paritial derivative** is defined as a limit. Let U be an open subset of $\mathbb{R}^n$ and $f : U \to \mathbb{R}$ a function. The partial derivative of f at point $\alpha = (a_1, a_2, \cdots, a_n) \in U$ with respect to the i-th variable x_i is defined as:

$$
\begin{aligned}
\frac{\partial f(\alpha)}{\partial x_i} &= \lim_{g \to 0} \frac{f(a_1, \cdots, a_{i-1}, a_i + h, a_{i+1}, \cdots, a_n) - f(a_1, \cdots, a_i, \cdots, a_n)}{h} \\
&= \lim_{h \to 0} \frac{f(\alpha + he_i) - f(\alpha)}{h}
\end{aligned}
$$

Where e_i is the unit vector of i-th variable x_i.

Remark

The method to calculate the partial derivative (with respect to x_i) is to treat all the variables other than x_i as constant.

Example

- $z = f(x, y) = x^2 + xy + y^2$, then

$$
\begin{aligned}
\frac{\partial z}{\partial x} &= 2x + y + 0 = 2x + y \\
\frac{\partial z}{\partial y} &= 0 + x + 2y = x + 2y
\end{aligned}
$$

- $z = f(a, b, c) = a^2 + b^3 + c^4 + ab + 2bc + 3ca + 4abc$, then

$$
\begin{aligned}
\frac{\partial z}{\partial a} &= 2a + 0 + 0 + b + 0 + 3c + 4bc = 2a + b + 3c + 4bc \\
\frac{\partial z}{\partial b} &= 0 + 3b^2 + 0 + a + 2c + 0 + 4ac = 3b^2 + a + 2c + 4ac \\
\frac{\partial z}{\partial c} &= 0 + 0 + 4c^3 + 0 + 2b + 3a + 4ab = 4c^3 + 2b + 3a + 4ab
\end{aligned}
$$

- $z = f(x_1, x_2) = x_1^2 - x_2$, then

$$
\begin{aligned}
\frac{\partial z}{\partial x_1} &= 2x_1 + 0 = 2x_1 \\
\frac{\partial z}{\partial x_2} &= 0 - 1 = -1
\end{aligned}
$$

Now you are fully capable of exploring the wonderful C&O. Hope all goes well with your studies. See you in the next chapter.

Chapter 1

Formulations

1.1 Overview

What is optimization? It is the action of making the best or most effective use of a situation or resource. In this course, broadly speaking, optimization is the problem of minimizing or maximizing a function subject to several constraints. Abstractly, we focus on **abstract optimization problem**, (P):
Given a **variable** set $X \subseteq \mathbb{R}^n$, **constraints**, and an **objective function** $f : X \to \mathbb{R}$.
Goal is to find some $x \in X$ that **minimizes** or **maximizes** f. (May not exist all time)
In this course, we will focus on three special cases of P:

- **Linear Program(LP)**.
- **Integer Program (IP)**.
- **Nonlinear Program (NLP)**.

Before giving the definitions of the programs above, we introduce:

Definition: Affine function

A function $f : \mathbb{R}^n \to \mathbb{R}$ is **affine** if $f(x) = \alpha^\top x + \beta$ for $\alpha \in \mathbb{R}^n, \beta \in \mathbb{R}$.
Additionally, f is **linear** if $\beta = 0$.

Examples

1. $f(x) = x_1 + 3x_2$ is affine, and also linear.
2. $f(x) = -x_1 + x_2 + 1$ is affine, but not linear.
3. $f(x) = 2x_1 + 4x_2 + sin(x_1)$ is not affine nor linear.

Definition: Affine constraint

A constraint:

$$g(x) \leq b$$

is linear if g is affine. (Can also be $\geq, =$ signs)

In the following sections, we will introduce LP, IP and NLP.

1.2 LP models

Definition: LP

The optimization problem

$$\underbrace{\max}_{or\ \min}\{f(x) : g_i(x) \underbrace{\leq}_{or \geq, =} b_i, \forall i = 1, 2, \cdots, m, x \in \mathbb{R}^m\}$$

is called a **linear program** (LP) if $f, g_1, g_2, \cdots, g_m$ are all **affine** functions.

Examples

- **Linear Program (LP)**:

$$\begin{aligned} \max\ & x_1 + x_2 \\ s.t.\ & x_1 + 3x_2 \leq 4 \\ & x_1 - x_2 \leq 2 \\ & x \geq \mathbb{0} \text{ (This stands for all } x_i \geq 0). \end{aligned}$$

- **Not linear Programm**: (I underbraced the place that makes it not an LP)

$$\max\ x_1 + \underbrace{\frac{1}{x_2}}_{!!!}$$

$$s.t.\ x_1 + 3x_2 \underbrace{<}_{!!!} 4$$

$$\underbrace{x_1x_2}_{!!!} \leq 2$$

Remark

Sometimes we can find that by doing some simplification, we can make a constraint from non-linear into linear. For example, the P is given by:

$$\begin{aligned} \max\ & x_1 + x_2 + x_3 \\ s.t.\ & x_1 + 3x_2 + \frac{1}{x_3} \leq 4 + \frac{1}{x_3} \\ & x_1 - x_2 \leq 2 \end{aligned}$$

By subtracting $\frac{1}{x_3}$ on both sides of the first constraint, this constraint can be linear. But **attention** here, even though we may be able to do some simplifications like this, P is still **not** an LP.

Before we move into Integer programming (IP), let's apply the LP we just learned first!

1.2.1 The formulation of LP

A formulation is a mathematical representation of the optimization problem, modeling the real-life problem into formulations, which is the fundamental of optimization. In this chapter, we will focus on how to formulate optimization problems.
More clearly, to formulate the problems in this chapter, we need to determine the **variables**, the **objective function**, and the **constraints**. Here is an example of formulating a real-life problem into formulations.

Real life problem Examples

Suppose a factory manufactures four products, requiring time on two machines and two types (skilled and unskilled) of labor. The amount of machine time and labor (in hours) needed to produce a unit of each product and the sales prices in dollars per unit of each product are given in the following table:

Product	Machine 1	Machine 2	Skilled labor	Unskilled labor	Price
1	11	4	8	7	300
2	7	6	5	8	260
3	6	5	5	7	220
4	5	4	6	4	180

There are also some constraints:

- Each month, 700 hours are available on machine 1, 500 hours on machine 2.
- Each month, the factory can purchase up to 600 hours of skilled labor at \$8 per hour and up to 650 hours of unskilled labor at \$6 per hour.

The factory wants to determine how much of each product it should produce each month and how much labor to purchase to maximize its profit (i.e. revenue from sales minus labor costs).

We wish to find a formulation for this problem, that is, we need to determine the variables, the objective function, and the constraints.

- **Variables**: The factory has to decide how much of each product to manufacture; we capture this by introducing a variable x_i for each $i = 1, 2, 3, 4$ for the number of units of product i to manufacture. As part of the planning process, the company must also decide on the number of hours of skilled and unskilled labor that it wants to purchase. We therefore introduce variables $y_{skilled}$ and $y_{unskilled}$ for the number of purchased hours of skilled and unskilled labor, respectively.

- **Objective function**: Deciding on a production plan now amounts to finding values for variables $x_1, \cdots, x_4, y_{skilled}$ and $y_{unskilled}$. Once the values for these variables have been found, the factory's profit is easily expressed by the following function:

$$\underbrace{300x_1 + 260x_2 + 220x_3 + 180x_4}_{\text{Profit from sales}} - \underbrace{(8y_{skilled} + 6y_{unskilled})}_{\text{Labor costs}}$$

  and the factory wants to maximize this quantity.

- Constraints:
  - The total amount of time needed on machine 1 can't exceed 700 hours:
$$11x_1 + 7x_2 + 6x_3 + 5x_4 \leq 700$$
  - The total amount of time needed on machine 2 can't exceed 500 hours:
$$4x_1 + 6x_2 + 5x_3 + 4x_4 \leq 500$$
  - Enough hours of skilled labor are purchased:
$$8x_1 + 5x_2 + 5x_3 + 6x_4 \leq y_{skilled}$$
  - Enough hours of unskilled labor are purchased:
$$7x_1 + 8x_2 + 7x_3 + 4x_4 \leq y_{unskilled}$$
  - The factory can purchase up to 600 hours of skilled labor:
$$y_{skilled} \leq 600$$
  - The factory can purchase up to 650 hours of unskilled labor:
$$y_{unskilled} \leq 650$$
  - All variables are non-negative by common sense:
$$x_1, x_2, x_3, x_4, y_{skilled}, y_{unskilled} \geq 0$$
- **Final formulation**:

$$\begin{aligned}
\max\ & 300x_1 + 260x_2 + 220x_3 + 180x_4 - 8y_{skilled} - 6y_{unskilled} \\
s.t.\ & 11x_1 + 7x_2 + 6x_3 + 5x_4 \leq 700 \\
& 4x_1 + 6x_2 + 5x_3 + 4x_4 \leq 500 \\
& 8x_1 + 5x_2 + 5x_3 + 6x_4 \leq y_{skilled} \\
& 7x_1 + 8x_2 + 7x_3 + 4x_4 \leq y_{unskilled} \\
& y_{skilled} \leq 600 \\
& y_{unskilled} \leq 650 \\
& x_1, x_2, x_3, x_4, y_{skilled}, y_{unskilled} \geq 0
\end{aligned}$$

We will learn how to solve this LP soon, but now, just focus on how to formulate the real-life problem!

1.3 IP models

Definition: IP

An **Integer Program** (IP) is a linear program with added integrality constraints for some/all (at least 1) of the variables.
We call an IP **mixed** if there are integral and fractional variables and **pure** otherwise.

Examples

- **Integer Program (IP)**:

$$\begin{aligned} \max \ & x_1 + x_2 \\ s.t. \ & x_1 + 3x_2 \leq 4 \\ & x_1 - x_2 \leq 2 \\ & x_1 \in \mathbb{Z}. \end{aligned}$$

- **Not Integer Programm**:

$$\begin{aligned} \max \ & x_1 + 5x_2 \\ s.t. \ & x_1 + 3x_2 \leq 4 \\ & x_1 - 5x_2 \leq 2 \end{aligned}$$

Remark 1

As the definition, an Integer Program (IP) is a linear program with added integrality constraints for some/all (at least 1) of the variables. It is easy to think IPis a special case of LP. But **attention** here, IP is **not** a special case of LP since there doesn't exist any affine function that constrains the variable as an integer.

Remark 2

Sometimes we have to let a variable x, be zero or one, to express some status. For example, $x = 0$ means the machine is not working, and $x = 1$ means the machine is working. We can use the (Binary) IP in this case, where the constraints include:

$$\begin{aligned} x &\geq 0 \\ x &\leq 1 \\ x &\in \mathbb{Z} \end{aligned}$$

which means x is zero or one. But to prevent the conflict with the definition, try to avoid writing: $x(x-1) = 0$ or $x \in \{0, 1\}$ or $sin(\pi x) = 0$ etc. But it depends, please verify with your instructors whether you can use these expressions when writing an IP.

1.3.1 Language of Graph theory

A familiar problem: starting at location s, we wish to travel to t, what is the best (shortest) route? Before, we usually found it by staring at it or using a computer to forcibly enumerate. But we indeed have some more elegant ways to do this kind of question. The goal of this section is to be able to formulate these kinds of questions into an Integer Program (IP).

Definition: Graph

A graph G consists of

- **Vertices**: $u, w, \cdots \in V(G)$. (drawn as filled circles)
- **Edges**: $uw, wz, \cdots \in E(G)$. (drawn as connecting circles)
- **Weight**: Mostly it is the length of the edge, or the difficulty of passing.
- **Adjacent**: Two vertices u and v are adjacent if $uv \in E(G)$.
- **Endpoints**: Vertices u and v are the endpoints of edge $uv \in E(G)$.
- **Incident**: Edge $e \in E(G)$ is incident to $u \in V(G)$ if u is and endpoint of e.
- ***s,t-path***: A *s,t-path* is a sequence

$$sv_1, v_1v_2, v_2v_3, \cdots, v_{k-1}v_k, v_kt$$

  where

  - $v_i \in V(G)$ and $sv_1, v_kt, v_iv_i + 1 \in E(G), \forall i$
  - $\forall i \neq j, v_i \neq v_j$ Without this, it is called ***s,t-walk***

- **Matching**: A collection $M \subseteq E(G)$, no two edges in M shares an endpoint.
- **Perfect Matching**: $V(M) = V(G)$, given that M is a matching of G.

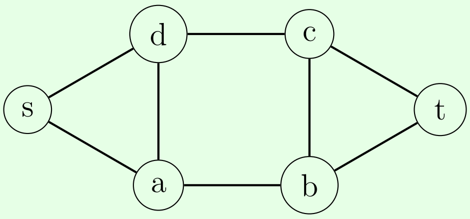
Remark

Edge's compositions' order doesn't matter, for example, uw has the exactly same meaning as wu since we are only focused on the undirected graph in this course.

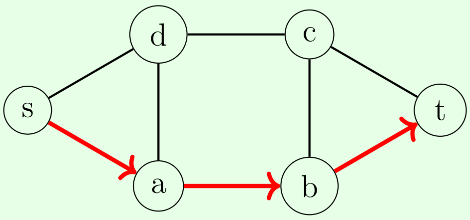
It is very easy to lose concentration if we don't draw the graph, for most of the situations, drawing a graph to help to understand is encouraged. On the next page, we provide a drawing to help understand these concepts.

Examples

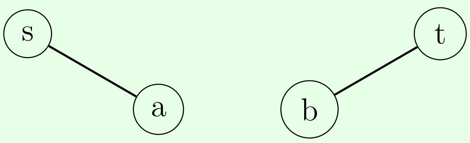
Consider this graph, we have:



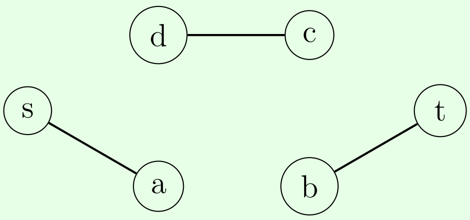
- **Vertices**: s, a, b, c, d, t.
- **Edges**: $sd, dc, ct, ad, bc, sa, ab, bt$.
- **Adjacent**: $(s, d), (d, c), (c, t), (a, d), (b, c), (s, a), (a, b), (b, t)$ are adjacent.
- **Endpoints**: d, a are the endpoints of edge da.
- **Incident**: Edge dc is incident to d and c.
- ***s,t-path***: $s \to a \to b \to t$ is one $s, t - path$.



- ***s,t-walk***: $s \to a \to d \to a \to b \to t$ is one $s, t - walk$.
- **Matching**: Here, $M = \{sa, bt\}$ is a matching, but not perfect.



- **Perfect Matching**: Here, $M = \{sa, dc, bt\}$ is a perfect matching.



1.3.2 Perfect Matching Optimization

Perfect Matching Theorem

Use $\delta(v)$ to denote the set of edges incident to v, i.e.

$$\delta(v) = \{e \in E : e = vu \text{ for some } u \in V\}$$

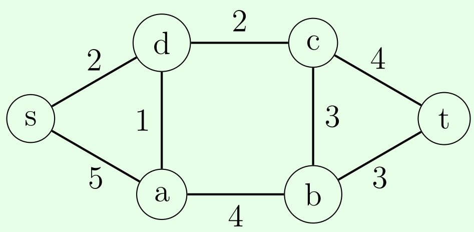
Given $G = (V, E), M \subseteq E$ is a perfect matching if and only if $M \cap \delta(v)$ contains a single edge for all $v \in V$.

By this theorem, we can formulate the perfect matching optimization problem using the Binary Integer Program. The IP will have a binary variable x_e for every edge $e \in E$, and $x_e = 1$ if $e \in M$, otherwise $x_e = 0$. Also, here, we represent the weight of edge e as c_e. Then, we can write our IP as:

$$\begin{aligned} \min \quad & \sum(c_e x_e : e \in E) \\ s.t. \quad & \sum(x_e : e \in \delta(v)) = 1(\forall v \in V) \\ & x \geq \mathbb{0}, x \in \mathbb{Z} \end{aligned}$$

Note here, that we don't need to state that every $x_e \leq 1$, since it is a minimized question, if $x_e > 1$, it will cause a greater objective value, which won't be picked up. Let's apply!

Examples



Objective function:

$$\sum(c_e x_e : e \in E) = 2x_{sd} + 2x_{dc} + 4x_{ct} + x_{da} + 3x_{cb} + 5x_{sa} + 4x_{ab} + 3x_{bt}$$

Constraints:

- $v = s, \delta(s) = \{sd, sa\} \Rightarrow x_{sd} + x_{sa} = 1$
- $v = d, \delta(d) = \{ds, da, dc\} \Rightarrow x_{sd} + x_{da} + x_{dc} = 1$
- $v = c, \delta(c) = \{cd, cb, ct\} \Rightarrow x_{cd} + x_{cb} + x_{ct} = 1$
- $v = t, \delta(t) = \{tc, tb\} \Rightarrow x_{tc} + x_{tb} = 1$
- $v = a, \delta(a) = \{as, ad, ab\} \Rightarrow x_{as} + x_{ad} + x_{ab} = 1$
- $v = b, \delta(b) = \{ba, bc, bt\} \Rightarrow x_{ba} + x_{bc} + x_{bt} = 1$
- $x \geq \mathbb{0}, x \in \mathbb{Z}$

1.3.3 The Shortest Paths Optimization

In this section, we wish to formulate the shortest path optimization problem. Now, let $C \subseteq E$ be a set of edges whose removal **disconnects** s and t. This implies that every $s, t-path$ must have at least one edge in C.

Definition: cut

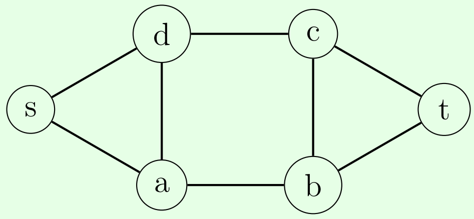
For $S \subseteq V$, we let $\delta(S)$ be the set of edges with **exactly one endpoint** in S.

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

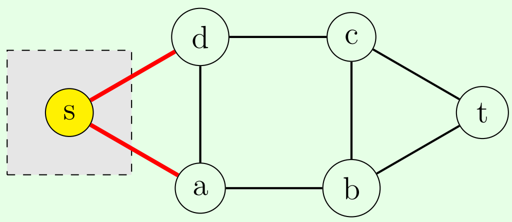
We call $\delta(S)$ an $s, t-cut$ and $t \notin S$

Examples

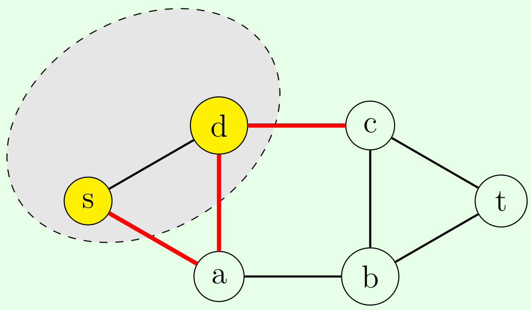
Consider this graph, we have:



- $\delta(s) = sd, sa$



- $\delta(s, d) = sa, da, dc$



Remark 1

If P is an $s, t-path$ and $\delta(S)$ is an $s, t-cut$, then P must have an edge from $\delta(S)$.

Remark 2

If $S \subseteq E$ contains at least one edge from every $s, t - cut$, then S contains an $s, t - path$.

Proof. Assume, by contradiction, S has an edge from every $s, t - cut$, but S has no $s, t - path$. Let R be the set of vertices reachable from s in S:

$$R = \{u \in V : S \text{ has an } s, u - path\}$$

Then, by assumption, $t \notin R$ since S doesn't contain an $s, t - path$. However, $\delta(R)$ is an $s, t - cut$ since $s \in R, t \notin R$. Then, $\exists e = (v_1, v_2) \in S$, such that $e \in \delta(R)$ where $v_1 \in R, v_2 \notin R$. This contradicts our assumption about R since if v_2 is connected to v_1, v_2 should be in R as well.
Hence, $\delta(R) \cap S = \emptyset$ contradicts our assumption. Therefore, S contains an $s, t - path$. $\square$

Now, we can try to formulate the shortest path optimization problem with a Binary Integer Program, here, we still use c_e to express the weight of edge e.

- **Variables**: We have one binary variable x_e for each edge $e \in E$. When $e \in P$, $x_e = 1$, otherwise $x_e = 0$.
- **Constraints**: We have one constraint for each $s, t - cut$ $\delta(U)$, forcing P to have an edge from $\delta(S)$. That is,

  $$\sum (x_e : e \in \delta(U)) \geq 1, (\forall U \subseteq V, s \in U, t \notin U)$$

  Also, we wish all $e \in E, x_e \geq 0, x_e \in \mathbb{Z}$.

  Note here, that we don't need to state that every $x_e \leq 1$, since it is a minimized question, if $x_e > 1$, it will cause a greater objective value, which won't be picked up.
- **Objective function**:

  $$\sum (c_e x_e : e \in E)$$

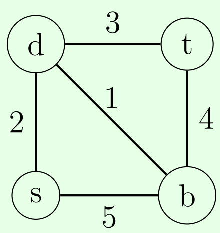
Remark 3

If x is an optimal solution for the above IP and $c_e > 0$ for all $e \in E$, then S_x contains the edges of the shortest $s, t - path$.

Now, let's apply!

Examples

Consider the graph below, we wish to find the shortest $s, t - path$.

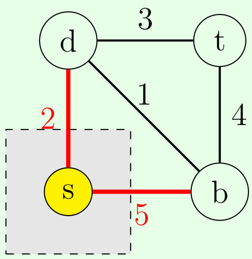


Examples

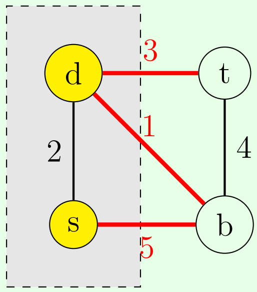
- **Objective function**:

$$\sum(c_e x_e : e \in E) = 5x_{sb} + 4x_{bt} + 3x_{td} + 2x_{ds} + x_{bd}$$

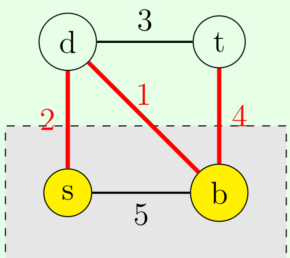
- **Constraints**:
  - $U = \{s\}, \delta(U) = \{sb, sd\} \Rightarrow x_{sd} + x_{sb} \geq 1.$



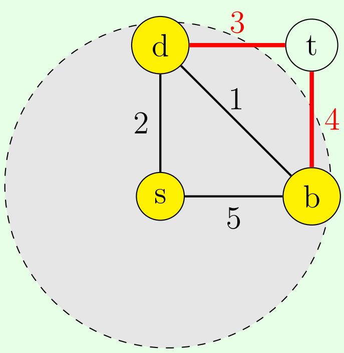
  - $U = \{s, d\}, \delta(U) = \{sb, db, dt\} \Rightarrow x_{sb} + x_{db} + x_{dt} \geq 1.$



  - $U = \{s, b\}, \delta(U) = \{sd, bd, bt\} \Rightarrow x_{sd} + x_{bd} + x_{bt} \geq 1.$



  - $U = \{s, b, d\}, \delta(U) = \{dt, bt\} \Rightarrow x_{dt} + x_{bt} \geq 1.$



  - $x \geq \mathbb{0}, x \in \mathbb{Z}.$

1.4 NLP models

Definition: NLP

The optimization problem

$$\underbrace{\max}_{or\min}\{f(x) : g_i(x) \underbrace{\leq}_{or \geq, =} b_i, \forall i = 1, 2, \cdots, m, x \in \mathbb{R}^m\}$$

is called a **Nonlinear program** (NLP) if one of $f, g_1, g_2, \cdots, g_m$ is not **affine** functions.

Examples

- **Nonlinear program**:

$$\begin{aligned} \max\ & x_1 + x_2 \\ s.t.\ & x_1 + 3x_2 \leq 4 \\ & x_1 x_2 \leq 2 \end{aligned}$$

- **Nonlinear program**:

$$\begin{aligned} \max\ & x_1 + x_2 \\ s.t.\ & x_1 + 3x_2 \leq 4 \\ & \frac{x_1}{x_2} \leq 2 \end{aligned}$$

- **Nonlinear program**:

$$\begin{aligned} \max\ & x_1 + x_2 \\ s.t.\ & x_1 + 3x_2 \leq 4 \\ & 5sin(x_1) \leq 4 \end{aligned}$$

Remark

Sometimes we can find some constraints are always valid, for example: $(x_1)^2 \geq 0$ or $sin(x_1) \leq 10$, but as for definition, we can't delete them when determining the type of the optimization problems. (But you can surely delete them when you run the algorithm for the optimal value)

Due to the complexity of NLP, not too many applications for NLP models are covered in this course, we just need to recognize the different types of optimization problems.
This is the end of Chapter 1, in this chapter, we introduced the optimization, the three programs, and their corresponding applications. In Chapter 2, we will focus on the outcomes of these programs, and on how to solve these programs by using the Simplex algorithm. I hope your study in C&O250 goes well, see you in the next chapter!

Chapter 2

Linear Programs

2.1 Possible Outcomes

In this course, when solving an optimization problem, the input will be an LP/IP/NLP program, and the algorithm (software) outputs the solution.

Definition

All assignment of values to each of the variables is a **feasible solution** if all the constraints are satisfied.
An optimization problem is **feasible** if it has at least one feasible solution. It is **infeasible** otherwise.

Definition

- For a **maximization** problem, an **optimal** solution is a feasible solution that **maximizes** the objective function.
- For a **minimization** problem, an **optimal** solution is a feasible solution that **minimizes** the objective function.
- An optimization problem may have several optimal solutions, and may also have no optimal solutions.

Definition

- A maximization problem is **unbounded** if for every value M, there exists a feasible solution with objective value **greater** than M.
- A minimization problem is **unbounded** if for every value M, there exists a feasible solution with objective value **smaller** than M.

Theorem: Fundamental Theorem of Linear Programming

There are **exactly three** possible outcomes for an optimization problem for LP:

- It has an **optimal** solution.
- It is **infeasible**.
- It is **unbounded**.

What does it mean to solve an LP using some algorithms (software)?

- **Optimal**: return an optimal solution $\bar{x}$ and a proof (certificate) that $\bar{x}$ is optimal.
- **Infeasible**: return a proof (certificate) that LP is infeasible.
- **Unbounded**: return a proof (certificate) that LP is unbounded.

In the next section, we will explain what exactly the certificate is for each outcome.

Remark

Why did I say the three outcomes are only for LP? Are there any other outcomes for other types of optimization problems?
The answer is **Yes**, here is an example, consider:

$$\begin{aligned} \max \ & x \\ s.t. \ & x < 1 \end{aligned}$$

- This program is none of LP, IP or NLP.
- This program is feasible, since $x = 0$ is one feasible solution.
- This program is not unbounded since 1 is an upper bound.
- This model has no optimal solution, consider $0 < \bar{x} < 1$ is a feasible solution, with objective value $\bar{x}$. Then, we can find that $\bar{x} + \frac{1-\bar{x}}{2}$ is also feasible, with a greater objective value. Therefore, due to the definition, this program doesn't have an optimal solution.

As we listed above, we can't find any outcomes corresponding to this problem indeed.

2.2 Certificates of Outcomes

2.2.1 Certificate of Optimality

First, due to the definition, an **optimal** solution is a feasible solution that maximizes/minimizes the objective function, and we can't find any other feasible solution that achieves a better value. But as for elegance, we can't try all the possible feasible solutions to verify, and sometimes we are not able to do that.
Here, a brief certificate of optimality is provided (we will figure this more later):
Given an LP:

$$\begin{aligned} \max\ & z(x) := c^\top x + m \\ s.t.\ & Ax = b \\ & x \geq \mathbb{0} \end{aligned}$$

We wish to make some change to the objective function z, say the new function is $\bar{z}$:

$$z(x) = \bar{z}(x) := \underbrace{\bar{c}^\top}_{\leq \mathbb{0}} \underbrace{x}_{\geq \mathbb{0}} + n$$

Then, if we can find some $\bar{x}$ such that $\bar{c}^\top \bar{x} = 0$, then $\bar{x}$ is optimal, since $\forall x$:

$$\bar{z}(\bar{x}) = \bar{c}^\top \bar{x} + n = n \geq \underbrace{\bar{c}^\top}_{\leq \mathbb{0}} \underbrace{x}_{\geq \mathbb{0}} + n = \bar{z}(x)$$

Sounds crazy and the gap is huge! Let's give an example for better understanding.

Example

We have:

$$\begin{aligned} \max\ & z(x) := \begin{pmatrix} -2 & -1 & 1 & 0 \end{pmatrix} x + 1 \\ s.t.\ & \begin{pmatrix} -1 & 3 & 1 & 0 \\ -2 & 6 & 0 & 1 \end{pmatrix} x = \begin{pmatrix} 4 \\ 5 \end{pmatrix} \\ & x \geq \mathbb{0} \end{aligned}$$

we can find that

$$z(x) = \begin{pmatrix} -1 & -4 & 0 & 0 \end{pmatrix} x + \underbrace{\begin{pmatrix} -1 & 3 & 1 & 0 \end{pmatrix} x}_{=4,\text{ by first constraint}} + 1 = \underbrace{\begin{pmatrix} -1 & -4 & 0 & 0 \end{pmatrix}}_{\leq \mathbb{0}} \underbrace{x}_{\geq \mathbb{0}} + 5 := \bar{z}(x)$$

We claim that $\bar{x} = \begin{pmatrix} 0 & 0 & 4 & 5 \end{pmatrix}^\top$ is a feasible solution of value 5 (easy to prove), and:

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

Thus, we conclude that $\bar{x}$ is an optimal solution, and this LP's outcome is **optimal**.

On the next page, we will introduce the **Certificate of Optimality**.

Start with the example:

Example

We have:

$$\begin{aligned} \max \ z(x) &:= (-1,3,-5,2,1)x - 3 \\ s.t. \ \begin{pmatrix} 1 & -2 & 1 & 0 & 2 \\ 0 & 1 & -1 & 1 & 3 \end{pmatrix} x &= \begin{pmatrix} 2 \\ 4 \end{pmatrix}, x \geq \mathbb{0} \end{aligned}$$

We wish to construct a new $\bar{z}(x) := \bar{c}^\top x + n$, such that $\bar{c}^\top \leq \mathbb{0}$. Now, by multiplying the first constraint of the LP by -1, the second by 2, and by adding the two constraints together, we obtain:

$$\underbrace{(-1,2)}_{y^\top} \begin{pmatrix} 1 & -2 & 1 & 0 & 2 \\ 0 & 1 & -1 & 1 & 3 \end{pmatrix} x = \underbrace{(-1,2)}_{y^\top} \begin{pmatrix} 2 \\ 4 \end{pmatrix}$$

and after simplifying, we have

$$(-1,4,-3,2,4)x = 6 \Leftrightarrow -(-1,4,-3,2,4)x + 6 = 0$$

Let $\bar{z}(x) := z(x) + \underbrace{-(-1,4,-3,2,4)x + 6}_{0} = \underbrace{(0,-1,-2,0,-3)}_{\bar{c}^\top \leq \mathbb{0}} x + 3$. We can find a feasible solution $\bar{x} = (2,0,0,4,0)^\top$ such that $\bar{c}^\top \bar{x} = 0$.
Then, we can conclude that $\bar{x}$ is the optimal solution.

Definition

In our example, the scalars vector $y^\top = (-1,2)$ is the **Certificate of Optimality**.

But still, I believe it is a bit confusing:

- How can we find the scalars vector $y^\top$ as the certificate of optimality?
- How can we find $\bar{z}(x)$? Is there exists a method to help find it?
- How can we find the optimal solution $\bar{x}$? Is there exists a method to help find it?

The answer is **Yes**, we do have some methods for it, and we will introduce them soon!

2.2.2 Certificate of Infeasibility

One way is trying to convert $Ax = b$ into RREF, if the RREF of $Ax = b$ is unsolvable, then this LP is infeasible as well. But we missed the constraint $x \geq 0$ here, we may have some cases in which $Ax = b$ is solvable without constraints $x \geq \mathbb{0}$, and the LP is infeasible.

Example

$$x_1 + x_2 = -1, x \geq \mathbb{0}$$

We only need to give a certificate of infeasibility when $Ax = b$ is solvable, but the LP is infeasible. For the case $Ax = b$ unsolvable, just state how $[A|b]$ is inconsistent.

Theorem: Ferka's Lemma

There is no solution to $Ax = b, x \geq \mathbb{0}$ if there exists a y such that:

$$y^\top A \geq \mathbb{0}, y^\top b < 0$$

Example

Consider the LP:

$$\begin{aligned} \max\ & (3, 4, -1, 2)^\top x \\ s.t.\ & \underbrace{\begin{pmatrix} 3 & -2 & -6 & 7 \\ 2 & -1 & -2 & 4 \end{pmatrix}}_{A} x = \underbrace{\begin{pmatrix} 6 \\ 2 \end{pmatrix}}_{b} \\ & x \geq \mathbb{0} \end{aligned}$$

Here, we can construct a $y^\top := \begin{pmatrix} -1 & 2 \end{pmatrix}$, and then we have:

$$\underbrace{\begin{pmatrix} -1 & 2 \end{pmatrix}}_{y^\top} \underbrace{\begin{pmatrix} 3 & -2 & -6 & 7 \\ 2 & -1 & -2 & 4 \end{pmatrix}}_{A} x = \underbrace{\begin{pmatrix} -1 & 2 \end{pmatrix}}_{y^\top} \underbrace{\begin{pmatrix} 6 \\ 2 \end{pmatrix}}_{b}$$

$$\underbrace{\begin{pmatrix} 1 & 0 & 2 & 1 \end{pmatrix}}_{\geq \mathbb{0}} \underbrace{x}_{\geq \mathbb{0}} = \underbrace{-2}_{<0}$$

Suppose there's a solution $\bar{x}$, then we have

$$\underbrace{\begin{pmatrix} 1 & 0 & 2 & 1 \end{pmatrix}}_{\geq \mathbb{0}} \underbrace{\bar{x}}_{\geq \mathbb{0}} = \underbrace{-2}_{<0}$$

Which makes no sense, since the LHS is non-negative and the RHS is negative. Therefore, there's no solution to this LP, which means this LP is **infeasible**. Moreover, here, the $y^\top$ we constructed is the **certificate of infeasibility**

- How did we find the $y^\top$ for the certificate? Is there exists a method to help find it?

The answer is **Yes**! We do have some methods for it, and we will introduce them soon!

2.2.3 Certificate of Unboundedness

In this section, we will only focus on the certificate of unboundedness of maximization optimization problems since the minimization one is similar.
Recall the definition of unboundedness, a maximization problem is **unbounded** if for every value M, there exists a feasible solution with objective value **greater** than M.

Remark

Our idea to prove the LP is unbounded is to construct a family of feasible solutions $x(t), \forall t \geq 0$, and show that as t goes to infinity, the value of the objective function goes to infinity, that is, the LP

$$\max\{c^\top x : Ax = b, x \geq \mathbb{0}\}$$

is unbounded if we can find $\bar{x}$ and r such that

$$\bar{x} \geq \mathbb{0}, r \geq \mathbb{0}, \ A\bar{x} = b, Ar = 0, \ c^\top r > 0$$

so that we can rewrite the LP as:

$$\begin{aligned} \max \ & c^\top \bar{x} + \underbrace{t}_{\geq 0} \underbrace{c^\top r}_{>0} \\ s.t. \ & \underbrace{A\bar{x}}_{b} + \underbrace{Ar}_{0} = b \end{aligned}$$

Now, as t goes to infinity, the objective value will go to infinity as well, which shows that this LP is unbounded. Moreover, here, the pair $(\bar{x}, r)$ is the **Certificate of Unboundedness**.

Example

We have the LP

$$\begin{aligned} \max \ & \underbrace{\begin{pmatrix} -1 & 0 & 0 & 1 \end{pmatrix}}_{c^\top} x \\ s.t. \ & \underbrace{\begin{pmatrix} -1 & -1 & 1 & 0 \\ -2 & 1 & 0 & 1 \end{pmatrix}}_{A} x = \underbrace{\begin{pmatrix} 2 \\ 1 \end{pmatrix}}_{b}, x \geq \mathbb{0} \end{aligned}$$

By solving the matrix equation $Ax = b$ using some EROs, we can obtain:

$$x(t) := \underbrace{\begin{pmatrix} 0 & 0 & 2 & 1 \end{pmatrix}^\top}_{\bar{x}} + t \underbrace{\begin{pmatrix} 1 & 0 & 1 & 2 \end{pmatrix}^\top}_{r}$$

And note here t must forcibly be non-negative since if it is negative, $x_1 = t < 0$ which contradicts $x \geq \mathbb{0}$. Now, we check:

- $x(t)$ is feasible for all $t \geq \mathbb{0}$: $Ax(t) = A(\bar{x} + tr) = A\bar{x} + tAr = b$ indeed, and $x(t) = \begin{pmatrix} t & 0 & 2+t & 1+2t \end{pmatrix} \geq \mathbb{0}$ as $t \geq 0$. (More on next page)

Example

- $c^\top x \to \infty$ when $t \to \infty$:

$$c^\top x(t) = c^\top \bar{x} + t \underbrace{c^\top r}_{1} = \underbrace{c^\top \bar{x}}_{\text{fixed}} + \underbrace{t}_{\infty} \to \infty$$

We now have a clearer realization of the certificate of each type of outcome, now, we can start to fill the gap left before!

2.3 Standard Equality Forms

Definition

An LP is in **Standard Equality Form (SEF)** if

- It is a maximization problem.
- $x \geq \mathbb{0}$.
- All constraints are equality constraints except the $x \geq \mathbb{0}$.

Example

- This is an LP in SEF:

$$\begin{aligned} \max\ & x_1 + x_2 + 17 \\ s.t.\ & x_1 - x_2 = 2 \\ & x_1, x_2 \geq 0. \end{aligned}$$

- This LP is not in SEF since $x_2 \geq 0$ is missed:

$$\begin{aligned} \max\ & x_1 + 11x_2 + 9 \\ s.t.\ & x_1 - 2x_2 = 3 \\ & x_1 \geq 0. \end{aligned}$$

Remark

For this LP

$$\begin{aligned} \max\ & x_1 + x_2 + 17 \\ s.t.\ & x_1 - x_2 = 0 \\ & x_1 \geq 0. \end{aligned}$$

Though $x_2 \geq 0$ can be implied by the constraint, it is not given **explicitytly**, so this LP is not in SEF. Moreover, we call the x_i **free** if there's no constraint $x_i \geq 0$.

Definition

We say two LPs (P) and (Q) are **equivalent** if

- (P) is infeasible implies (Q) is infeasible.
- (P) is unbounded implies (Q) is unbounded.
- Can construct an optimal solution of (P) from the optimal solution of (Q).
- Can construct an optimal solution of (Q) from the optimal solution of (P).

Remark

Every LP is equivalent to an LP in SEF.

Here are some methods to convert an LP into SEF.

- **How do we change a minimum problem into a maximum problem?**

  Answer: Take the opposite sign of the objective function and find its maximum.

  Example

  Consider this minimum problem's objective function:

  $$\min\ (x_1 + x_2 - 3x_3 + 9) \rightarrow -\max\ (-x_1 - x_2 + 3x_3 - 9)$$

- **How do we replace an inequality with an equality?**

  Answer: Add/minus a positive variable to fill the difference.

  Example

  Consider these constraints

  $$\begin{array}{rcl} x_1 + x_2 - 4x_3 \leq 7 & & x_1 + x_2 - 4x_3 + s = 7 \\ 2x_1 - x_2 + x_3 \geq 10 & \Rightarrow & 2x_1 - x_2 + x_3 - t = 10 \\ x_1, x_2, x_3 \geq 0 & & x_1, x_2, x_3, s, t \geq 0 \end{array}$$

- **What if we have free variables?**

  Answer: The idea is, any number can be expressed by the difference of two non-negative numbers. So if x_i is free, we introduce $x_i^+, x_i^- \geq 0$, such that $x_i = x_i^+ - x_i^-$.

  Example

  Consider these constraints

  $$\begin{aligned} x_1 + x_2 - 4x_3 &= 7 \\ x_1, x_2 &\geq 0 \\ \Rightarrow x_1 + x_2 - 4(x_3^+ - x_3^-) &= 7 \\ x_1, x_2, x_3^+, x_3^- &\geq 0 \end{aligned}$$

2.4 Basis

Notation let B be a subset of column indices, then A_B is a columns sub-matrix of A indexed by set B. A_i denotes the columns i of A.

Definition

Let B be a subset of column indices, B is a basis if A_B is invertible (non-sigular).

Remark

Some properties of basis:

- Max number of independent columns = Max number of independent rows.
- B is a basis if and only if B is a maximal set of independent columns of A.
- Not every matrix has basis, consider $A = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$, then all A_B are singular.

Example

Given $A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 5 & 3 \end{pmatrix}$, the possible subset of column indices are: $\{1,2\}, \{1,3\}, \{2,3\}$. (We wish A_B be square, as a prerequisite to be invertible).

- $B = \{1,2\}$, in this case $A_B = \begin{pmatrix} 1 & 2 \\ 1 & 5 \end{pmatrix}$, and using determinant we have: $1 \times 5 - 2 \times 1 \neq 0$, therefore, $B = \{1,2\}$ **is** a basis for A.
- $B = \{1,3\}$, in this case $A_B = \begin{pmatrix} 1 & 3 \\ 1 & 3 \end{pmatrix}$, and using determinant we have: $1 \times 3 - 3 \times 1 = 0$, therefore, $B = \{1,3\}$ is a **not** a basis for A.
- $B = \{2,3\}$, in this case $A_B = \begin{pmatrix} 2 & 3 \\ 5 & 3 \end{pmatrix}$, and using determinant we have: $2 \times 3 - 3 \times 5 \neq 0$, therefore, $B = \{2,3\}$ **is** a basis for A.

Definition

- x is a **basic solution (bs)** for basis B if
  - $Ax = b$.
  - $x_i = 0$ for all $i \notin B$.
- x is a **basic feasible solution (bfs)** for basis B if
  - x is a basic solution for basis B.
  - $x \geq 0$

  And in this case, we say B is a **feasible basis**.

Example

How to find a basic solution for

$$\underbrace{\begin{pmatrix} 1 & 0 & 1 & -1 \\ 0 & 1 & 1 & 1 \end{pmatrix}}_{A} x = \underbrace{\begin{pmatrix} 2 \\ 2 \end{pmatrix}}_{b}$$

When $B = \{1, 4\}$?
Solution: We have:

$$\begin{aligned}
\begin{pmatrix} 2 \\ 2 \end{pmatrix} &= \begin{pmatrix} 1 & 0 & 1 & -1 \\ 0 & 1 & 1 & 1 \end{pmatrix} x \\
&= x_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \underbrace{x_2}_{0} \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \underbrace{x_3}_{0} \begin{pmatrix} 1 \\ 1 \end{pmatrix} + x_4 \begin{pmatrix} -1 \\ 1 \end{pmatrix} \\
&= \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_4 \end{pmatrix}
\end{aligned}$$

Since A_B is invertible, multiplying A_B^{-1} on both sides gives:

$$\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_4 \end{pmatrix}$$
$$\begin{pmatrix} 4 \\ 2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_4 \end{pmatrix}$$

So the basic solution $x = (4, 0, 0, 2)^\top$, since $x \geq \mathbb{0}$, it is also a basic feasible solution.

Remark

A basic solution can be the basic solution for more than one basis.

Remark

If x is a basic solution for basis B, $x_i = 0$ for all $i \notin B$, what about x_j, for $j \in B$, do all $x_j \neq 0$?
The answer is **No**, x_j can be zero sometimes.

Example

Consider $Ax = b$,

$$A := \begin{pmatrix} 4 & 1 & -1 & 0 \\ 5 & 0 & 8 & 0 \\ 7 & 0 & -4 & 1 \end{pmatrix}, b := \begin{pmatrix} 3 \\ 13 \\ 3 \end{pmatrix}, B_1 = \{1, 2, 3\}, B_2 = \{1, 3, 4\}$$

Then, we can easily show that A_{B_1} and A_{B_2} are both invertible, and $x = (1, 0, 1, 0)^\top$ is a basic feasible solution for both B_1, B_2, with $x_2 = 0, 2 \in B_1$. This example proves our remarks above once.

Remark

Consider $Ax = b$ and a basis B of A, then there exists a **unique basic solution** x for B. Columns of A_B and elements of x_B are ordered by B.

Proof.

$$\begin{aligned} b = Ax &= \sum_j A_j x_j \\ &= \sum_{j \in B} A_j x_j + \sum_{j \notin B} A_j \underbrace{x_j}_{0} \\ &= \sum_{j \in B} A_j x_j = A_B x_B. \end{aligned}$$

Since B is a basis, it implies A_B is invertible, and hence, $x_B = A_B^{-1} b$. ☐

Definition

For all indexes of $A, \notin B$, we say the set of these elements N.

2.5 Canonical Forms

Consider the problem (P) in SEF:

$$\max\{c^\top x + m : Ax = b, x \geq \mathbb{0}\}$$

Definition

Let B be a basis of A, then (P) is in **canonical form** for B if

- $A_B = I$.
- $c_j = 0$ for all $j \in B$

Also, we say B is an optimal basis if and only if $c^\top \leq \mathbb{0}$.

Example

This is an LP in Canonical Form for $B = \{1, 2\}$:

$$\begin{aligned} \max \quad & \underbrace{(0, 0, 2, 4)}_{c^\top} x \\ s.t. \quad & \underbrace{\begin{pmatrix} 1 & 0 & 1 & -1 \\ 0 & 1 & 1 & 2 \end{pmatrix}}_{A} x = \underbrace{\begin{pmatrix} 1 \\ 2 \end{pmatrix}}_{b}, x \geq \mathbb{0} \end{aligned}$$

We can verify:

- $A_B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2$.
- $c_1 = c_2 = 0$.

Remark

We can always rewrite an LP in Canonical Form on any basis B.

Here are the steps with an example: consider, we have the LP model:

$$\max \underbrace{(4,5,2,3,-2)}_{c^\top} x + 10$$
$$s.t. \underbrace{\begin{pmatrix} 1 & 2 & 0 & 0 & 3 \\ 1 & 1 & 1 & 0 & 8 \\ 1 & 5 & 0 & 1 & 2 \end{pmatrix}}_{A} x = \underbrace{\begin{pmatrix} 2 \\ 6 \\ 3 \end{pmatrix}}_{b}$$
$$x \geq \mathbb{0}$$

Now, we bring this LP into canonical form for the basis $\{1,2,4\}$, the steps are below:

P1. Replace $Ax = b$ by $A'x = b'$ with $A'_B = I$

$$\begin{aligned}
&\begin{pmatrix} 1 & 2 & 0 & 0 & 3 \\ 1 & 1 & 1 & 0 & 8 \\ 1 & 5 & 0 & 1 & 2 \end{pmatrix} x = \begin{pmatrix} 2 \\ 6 \\ 3 \end{pmatrix} \\
\leftrightarrow &\begin{pmatrix} 1 & 2 & 0 \\ 1 & 1 & 0 \\ 1 & 5 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 2 & 0 & 0 & 3 \\ 1 & 1 & 1 & 0 & 8 \\ 1 & 5 & 0 & 1 & 2 \end{pmatrix} x = \begin{pmatrix} 1 & 2 & 0 \\ 1 & 1 & 0 \\ 1 & 5 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 2 \\ 6 \\ 3 \end{pmatrix} \\
\leftrightarrow &\begin{pmatrix} -1 & 2 & 0 \\ 1 & -1 & 0 \\ -4 & 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 0 & 0 & 3 \\ 1 & 1 & 1 & 0 & 8 \\ 1 & 5 & 0 & 1 & 2 \end{pmatrix} x = \begin{pmatrix} -1 & 2 & 0 \\ 1 & -1 & 0 \\ -4 & 3 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 6 \\ 3 \end{pmatrix} \\
\leftrightarrow &\begin{pmatrix} 1 & 0 & 2 & 0 & 13 \\ 0 & 1 & -1 & 0 & -5 \\ 0 & 0 & 3 & 1 & 14 \end{pmatrix} x = \begin{pmatrix} 10 \\ -4 \\ 13 \end{pmatrix}
\end{aligned}$$

P2. Replace $c^\top x$ by $\bar{c}^\top x + \bar{z}$ with $\bar{c}_B = \mathbb{0}$ ($\bar{z}$ is a constant)

Step 1. construct a new objective function by

- multiplying constraint 1 by y_1
- multiplying constraint 2 by y_2
- multiplying constraint 3 by y_3, and
- adding the result constraints to the objective function.

Step 2. choose y_1, y_2, y_3 to get $\bar{c}_B = \mathbb{0}$, (which means $\bar{c}_B^\top = \mathbb{0}$ too).

We have:

$$
\begin{aligned}
0 &= -(y_1, y_2, y_3)\begin{pmatrix}1 & 2 & 0 & 0 & 3\\ 1 & 1 & 1 & 0 & 8\\ 1 & 5 & 0 & 1 & 2\end{pmatrix} x + (y_1, y_2, y_3)\begin{pmatrix}2\\6\\3\end{pmatrix}\\
z &= (4,5,2,3,-2)x + 10\\
\to z &= \underbrace{\left[(4,5,2,3,-2) - (y_1, y_2, y_3)\begin{pmatrix}1 & 2 & 0 & 0 & 3\\ 1 & 1 & 1 & 0 & 8\\ 1 & 5 & 0 & 1 & 2\end{pmatrix}\right]}_{\bar{c}^\top} x + \underbrace{(y_1, y_2, y_3)\begin{pmatrix}2\\6\\3\end{pmatrix} + 10}_{\bar{z}}\\
\mathbb{0} = \bar{c}_B^\top &= (4,5,3) - (y_1, y_2, y_3)\begin{pmatrix}1 & 2 & 0\\ 1 & 1 & 0\\ 1 & 5 & 1\end{pmatrix}\\
\leftrightarrow &(y_1, y_2, y_3)\begin{pmatrix}1 & 2 & 0\\ 1 & 1 & 0\\ 1 & 5 & 1\end{pmatrix} = (4,5,3)\\
\leftrightarrow &\begin{pmatrix}y_1\\y_2\\y_3\end{pmatrix} = (4,5,3)\begin{pmatrix}1 & 2 & 0\\ 1 & 1 & 0\\ 1 & 5 & 1\end{pmatrix}^{-1} = \begin{pmatrix}-11\\12\\3\end{pmatrix}
\end{aligned}
$$

Hence, we choose $(y_1, y_2, y_3) = (-11, 12, 3)$ and

$$
\begin{aligned}
z &= \left[(4,5,2,3,-2) - (y_1, y_2, y_3)\begin{pmatrix}1 & 2 & 0 & 0 & 3\\ 1 & 1 & 1 & 0 & 8\\ 1 & 5 & 0 & 1 & 2\end{pmatrix}\right] x + (y_1, y_2, y_3)\begin{pmatrix}2\\6\\3\end{pmatrix} + 10\\
&= (0,0,-10,0,-71)x + 69
\end{aligned}
$$

Therefore, our final Canonical Form is:

$$
\begin{aligned}
&\max (0,0,-10,0,-71)x + 69\\
s.t. &\begin{pmatrix}1 & 0 & 2 & 0 & 13\\ 0 & 1 & -1 & 0 & -5\\ 0 & 0 & 3 & 1 & 14\end{pmatrix} x = \begin{pmatrix}10\\-4\\13\end{pmatrix}, x \geq \mathbb{0}
\end{aligned}
$$

Wow, this is complex, however, it does have a formula! (Showing on next page)

Remark

For any invertible matrix M, we have:

$$(M^\top)^{-1} = (M^{-1})^\top =: M^{-\top}$$

Remark

Consider converting this LP (P) into canonical form $(P')$with basis B:
(P)

$$\begin{aligned} \max\ & c^\top x + m \\ s.t.\ & Ax = b \\ & x \geq \mathbb{0} \end{aligned}$$

(P')

$$\begin{aligned} \max\ & \underbrace{[c^\top - y^\top A]}_{\bar{c}^\top} x + y^\top b + m \\ s.t.\ & \underbrace{A_B^{-1} A}_{A'} x = \underbrace{A_B^{-1} b}_{b'} \\ & x \geq \mathbb{0} \end{aligned}$$

where $y = A_B^{-\top} c_B$.

Example

Consider, we have the LP model:

$$\begin{aligned} \max\ & \underbrace{(4,5,2,3,-2)}_{c^\top} x + \underbrace{10}_{m} \\ s.t.\ & \underbrace{\begin{pmatrix} 1 & 2 & 0 & 0 & 3 \\ 1 & 1 & 1 & 0 & 8 \\ 1 & 5 & 0 & 1 & 2 \end{pmatrix}}_{A} x = \underbrace{\begin{pmatrix} 2 \\ 6 \\ 3 \end{pmatrix}}_{b}, x \geq \mathbb{0} \end{aligned}$$

Now, we bring this LP into canonical form for the basis $\{1,2,4\}$ using the remark:

$$y = A_B^{-\top} c_B = \begin{pmatrix} 1 & 2 & 0 \\ 1 & 1 & 0 \\ 1 & 5 & 1 \end{pmatrix}^{-\top} \begin{pmatrix} 4 \\ 5 \\ 3 \end{pmatrix} = \begin{pmatrix} -1 & 2 & 0 \\ 1 & -1 & 0 \\ -4 & 3 & 1 \end{pmatrix}^\top \begin{pmatrix} 4 \\ 5 \\ 3 \end{pmatrix} = \begin{pmatrix} -11 \\ 12 \\ 3 \end{pmatrix}$$

$$y^\top = \begin{pmatrix} -11 & 12 & 3 \end{pmatrix},\ y^\top A = \begin{pmatrix} 4 & 5 & 12 & 3 & 69 \end{pmatrix}, y^\top b + m = 69$$

$$A_B^{-1} A = \begin{pmatrix} 1 & 0 & 2 & 0 & 13 \\ 0 & 1 & -1 & 0 & -5 \\ 0 & 0 & 3 & 1 & 14 \end{pmatrix}, A_B^{-1} b = \begin{pmatrix} 10 \\ -4 \\ 13 \end{pmatrix}.$$

Therefore, our final Canonical Form is:

$$\begin{aligned} & \max(0,0,-10,0,-71)x + 69 \\ s.t.\ & \begin{pmatrix} 1 & 0 & 2 & 0 & 13 \\ 0 & 1 & -1 & 0 & -5 \\ 0 & 0 & 3 & 1 & 14 \end{pmatrix} x = \begin{pmatrix} 10 \\ -4 \\ 13 \end{pmatrix}, x \geq \mathbb{0} \end{aligned}$$

Another version of the converting formula is: (just deleted the use of y and $y^\top$)

Remark

Consider converting this LP (P) into canonical form $(P')$with basis B:
(P)

$$\begin{aligned} \max\ & c^\top x + m \\ s.t.\ & Ax = b \\ & x \geq \mathbb{0} \end{aligned}$$

(P')

$$\begin{aligned} \max\ & [c^\top - c_B^\top A_B^{-1} A]x + c_B^\top A_B^{-1} b + m \\ s.t.\ & A_B^{-1} A x = A_B^{-1} b \\ & x \geq \mathbb{0} \end{aligned}$$

So what you have to calculate is:

$$\begin{tikzcd} & \boxed{E_1 := A_B^{-1}} \arrow[dl] \arrow[dr] & \\ \boxed{E_2 := E_1 A} \arrow[d] & & \boxed{E_3 := E_1 b} \arrow[d] \\ \boxed{E_4 := c_B^\top E_2} & & \boxed{E_5 := c_B^\top E_3} \end{tikzcd}$$

And the (P') now can be written as:

$$\begin{aligned} \max\ & (c^\top - E_4)x + E_5 + m \\ s.t.\ & E_2 x = E_3 \\ & x \geq \mathbb{0} \end{aligned}$$

Remark

Recall that for the LP in canonical form for

$$B : \max\{c^\top x + m : Ax = b, x \geq \mathbb{0}\}$$

We say B is an optimal basis if and only if $c^\top \leq \mathbb{0}$.
Question: Provide an LP of the form $\max\{c^\top x : Ax = b, x \geq \mathbb{0}\}$ and basis B, where B is not optimal basis, and the basic solution $\bar{x}$ for B is an optimal solution.
Answer: Seems really counter-intuitive, but it does exist. Consider:

$$\max\{(1,0)x : \begin{pmatrix} 1 & 1 \end{pmatrix} x = 0, x \geq \mathbb{0}\}, B = \{2\}$$

This LP is in canonical form for B, with $c^\top \not\leq \mathbb{0}$, but the corresponding basic feasible solution $\bar{x} = (0,0)$ is an optimal solution. (Since it is the only feasible solution for this LP).

2.6 Simplex-Optimality

Consider the LP:

$$\begin{aligned} \max\ & (0,1,3,0)x \\ s.t.\ & \begin{pmatrix} 1 & 1 & 2 & 0 \\ 0 & 1 & 1 & 1 \end{pmatrix} x = \begin{pmatrix} 2 \\ 5 \end{pmatrix}, x \geq \mathbb{0} \end{aligned}$$

and $B = \{1,4\}$, then

- A_B is invertible $\to$ B is a basis.
- $A_B = I_2$ and $c_B = \mathbb{0} \to$ LP is in canonical form for B.
- $\bar{x} = (2,0,0,5)^\top$ is a basic feasible solution for basis B.

So how do we find a better solution?
The idea is to pick $k \notin B$ such that $c_k > 0$, set $x_k = t \geq 0$ as large as possible and keep all other non-basic variables at 0. So here, we pick $k = 2$, set $x_2 = t \geq 0$, keep $x_3 = 0$. We want to choose basic variables such that $Ax = b$ still holds.

$$\begin{aligned} & \begin{pmatrix} 1 & 1 & 2 & 0 \\ 0 & 1 & 1 & 1 \end{pmatrix} x = \begin{pmatrix} 2 \\ 5 \end{pmatrix} \\ \Rightarrow & \begin{pmatrix} 1 & 1 & 2 & 0 \\ 0 & 1 & 1 & 1 \end{pmatrix} (x_1, t, 0, x_4)^\top = \begin{pmatrix} 2 \\ 5 \end{pmatrix} \\ \Rightarrow & \begin{matrix} x_1 = 2 - t \geq 0 \to t \leq 2 \\ x_4 = 5 - t \geq 0 \to t \leq 5 \end{matrix} \end{aligned}$$

Thus, the largest possible $t = \min\{2,5\} = 2$, and the new feasible solution is $x = (2-t, t, 0, 5-t)^\top = (0,2,0,3)^\top$. The new feasible solution is a basic feasible solution for basis $B = \{2,4\}$. Now, convert this LP into canonical form with new basis $B = \{2,4\}$: (P')

$$\begin{aligned} \max\ & (-1,0,1,0)x + 2 \\ s.t.\ & \begin{pmatrix} 1 & 1 & 2 & 0 \\ -1 & 0 & -1 & 1 \end{pmatrix} x = \begin{pmatrix} 2 \\ 3 \end{pmatrix}, x \geq \mathbb{0} \end{aligned}$$

Continue picking $k \notin B$ such that $c_k > 0$, set $x_k = t \geq 0$ as large as possible and keep all other non-basic variables at 0. So here, we pick $k = 3$, set $x_3 = t \geq 0$, keep $x_1 = 0$. We want to choose basic variables such that $Ax = b$ still holds.

$$\begin{aligned} & \begin{pmatrix} 1 & 1 & 2 & 0 \\ -1 & 0 & -1 & 1 \end{pmatrix} x = \begin{pmatrix} 2 \\ 3 \end{pmatrix} \\ \Rightarrow & \begin{pmatrix} 1 & 1 & 2 & 0 \\ -1 & 0 & -1 & 1 \end{pmatrix} (0, x_2, t, x_4)^\top = \begin{pmatrix} 2 \\ 3 \end{pmatrix} \\ \Rightarrow & \begin{matrix} x_2 = 2 - 2t \geq 0 \to t \leq 1 \\ x_4 = 3 + t \geq 0 \to t \leq \infty \text{ no upper bound} \end{matrix} \end{aligned}$$

Thus, the largest possible $t = \min\{1, \infty\} = 1$, and the new feasible solutino is $x = (0,0,1,4)^\top$. The new feasible solution is a basic feasible solution for basis $B = \{3,4\}$. Now, convert this LP into canonical form with new basis $B = \{3,4\}$:

$(P")$

$$\max \underbrace{(-3/2, -1/2, 0, 0)}_{\bar{c}^\top} x + 3$$
$$s.t. \begin{pmatrix} 1/2 & 1/2 & 1 & 0 \\ -1/2 & 1/2 & 0 & 1 \end{pmatrix} x = \begin{pmatrix} 1 \\ 4 \end{pmatrix}, x \geq \mathbb{0}$$

Since we have : $\bar{c}^\top \leq \mathbb{0}$, and we can indeed find a feasible solution $\bar{x} = (0, 0, 1, 4)^\top, \bar{c}^\top \bar{x} = 0$, we say this LP is optimal wth optimal solution $\bar{x}$, with optimal basis $\bar{B} = \{3, 4\}$.

Definition

- The algorithm we applied above is called **Simplex Algorithm**.
- During the example above, we know that the essence of this algorithm is to let an element leave the basis while letting an element enter the basis. In our example's first step, 1 leaves and 2 enters the basis, so the new basis is $\{2, 4\}$.

Now, we can find the **Certificate of Optimality**!

Definition

Given an LP:

$$\begin{aligned} \max\ & z(x) := c^\top x + m \\ s.t.\ & Ax = b \\ & x \geq \mathbb{0} \end{aligned}$$

After **Simplex**, we can find the optimal basis $\bar{B}$. The **Certificate of Optimality** is then the scalars vector defined as:

$$y^\top := c_{\bar{B}}^\top A_{\bar{B}}^{-1}$$

Remark

Wait, is the Simplex above really correct? During Simplex, we said the idea is to let one $x_i, i \in B$ that has been forcibly zero leave, and let another element in, but what if we have more than one choice of entering or leaving? This is an important question, if we ignore this and choose the c_k randomly, we might fall into an infinity loop! To solve this, we introduce **Bland's Rule**.

Theorem: Bland's Rule (Smallest-subscript rule)

If we have a choice for element entering/leaving, pick the one with the smallest index.

Remark

Still not the end! What if this LP is infeasible or unbounded? Seems like the current version algorithm can't handle it, we will complete it in a few sections later.

2.7 Simplex-Unboundedness

Consider the LP:

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

with $B = \{1,4,5\}$ as a feasible basis. Follow the Simplex algorithm we introduced:
Solution: Pick $k = 3 \notin B$ and let $x_3 = t, x_2 = 0$, then:

$$\begin{pmatrix} x_1 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} - t \begin{pmatrix} 1 \\ -3 \\ -2 \end{pmatrix}$$

with $t = \min\{1, \infty, \infty\} = 1$, thus $x_1 = 0 \rightarrow 1$ leaves the basis, 3 enter the basis. $\rightarrow B = \{3,4,5\}$, our LP in canonical form with this basis is:

$$\begin{aligned} \max\ & (-3,2,0,0,0)x + 3 \\ s.t.\ & \begin{pmatrix} 1 & -2 & 1 & 0 & 0 \\ 3 & -1 & 0 & 1 & 0 \\ 2 & 0 & 0 & 0 & 1 \end{pmatrix} x = \begin{pmatrix} 1 \\ 4 \\ 4 \end{pmatrix}, x \geq \mathbb{0} \end{aligned}$$

Pick $k = 2 \notin B$ and set $x_2 = t, x_1 = 0$, then:

$$\begin{pmatrix} x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 1 \\ 4 \\ 4 \end{pmatrix} - t \begin{pmatrix} -2 \\ -1 \\ 0 \end{pmatrix}$$

With $t = \min\{\infty, \infty, \infty\}$. Wait, can we choose the minimum number from three infinity? The answer is **No**, and here, back to the feasible solution we constructed:

$$x = \begin{pmatrix} 0 \\ t \\ 1+2t \\ 4+t \\ 4 \end{pmatrix} = \underbrace{\begin{pmatrix} 0 \\ 0 \\ 1 \\ 4 \\ 4 \end{pmatrix}}_{\bar{x}} + t \underbrace{\begin{pmatrix} 0 \\ 1 \\ 2 \\ 1 \\ 0 \end{pmatrix}}_{r}$$

Recall the certificate of unboundedness:

Definition

The LP

$$\max\{c^\top x : Ax = b, x \geq \mathbb{0}\}$$

is unbounded if we can find $\bar{x}$ and r such that

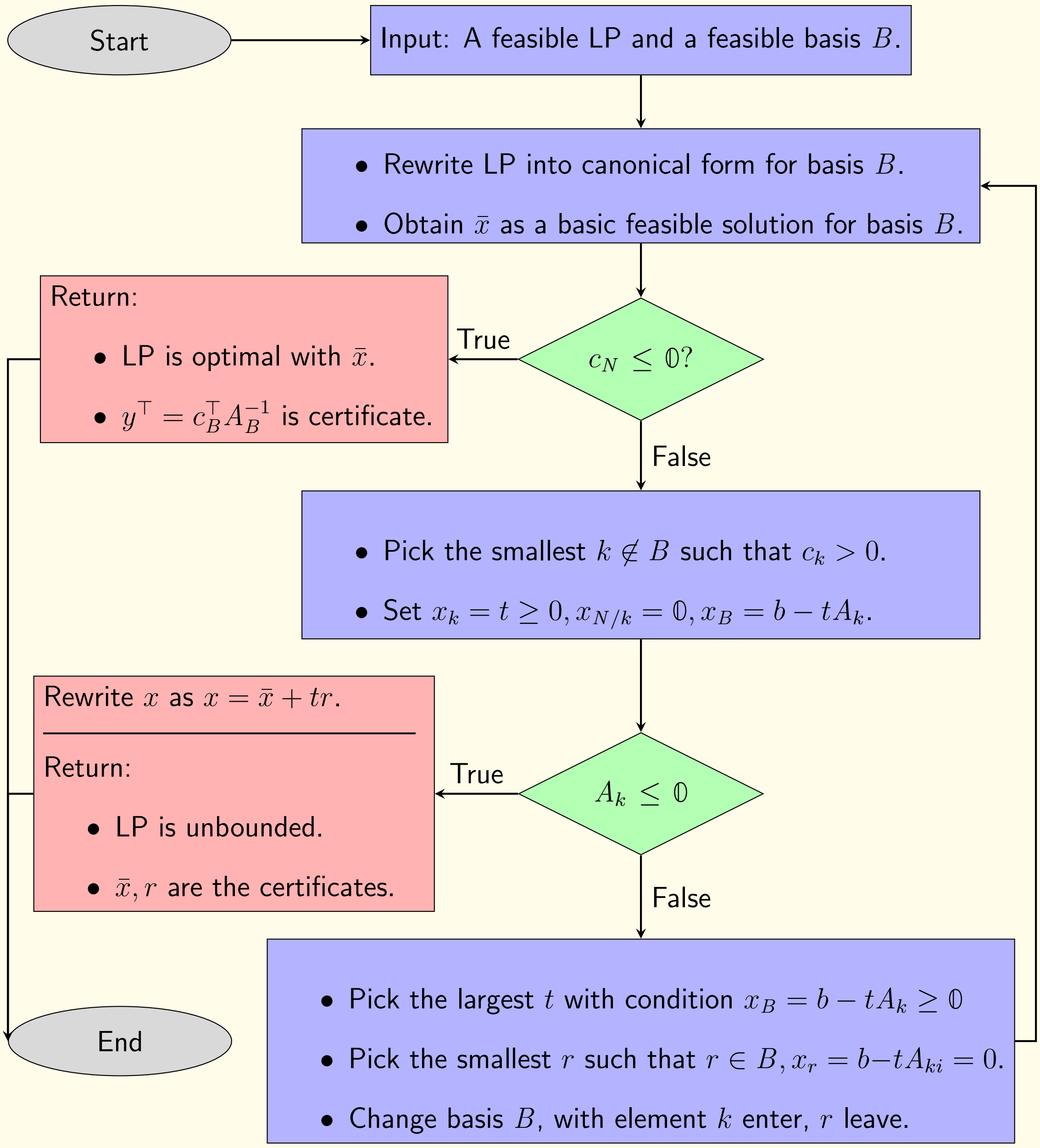
$$\bar{x} \geq \mathbb{0}, r \geq \mathbb{0},\ A\bar{x} = b, Ar = 0,\ c^\top r > 0$$

And the $\bar{x}, r$ we constructed can satisfy these conditions well! Therefore, we can conclude that this LP is unbounded. So in general, what is the key point we found the LP unbounded? Back to the procedure, $A_k \geq \mathbb{0}$ is essential for our construction.

2.8 Simplex Algorithm Procedure

Till now, we've solved the optimality and unboundedness, what about the Infeasibility? Unfortunately, only Simplex Algorithm can't detect all situations of Infeasibility with a certificate, but, we will introduce a method named **Two-phase Simplex** to handle that, before that, we may assume every LP that has been applied to Simplex Algorithm is feasible. In this section, we will describe the complete procedure of Simplex Algorithm:

Definition



Remark

Recall that N is the set of all indexes of A that not $\in B$.

$$A = \begin{pmatrix} a & b & c & d & e \end{pmatrix}, B = \{1\}, N = \{2, 3, 4, 5\}$$

2.9 Two-phase Simplex (Certificate of Infeasibility)

In this section, we will extend our Simplex algorithm, and after this section, we can:

- Determine whether an LP is feasible or not.
- Find the basic feasible solution for a feasible LP.
- Give the certificate of infeasibility of an infeasible LP.

Still, we start with an example:

Example

Consider the LP (P1):

$$\begin{aligned} \max \ & (2,-1,2)x \\ s.t. \ & \begin{pmatrix} -1 & -2 & 1 \\ 1 & 1 & 1 \end{pmatrix} x = \begin{pmatrix} -1 \\ 3 \end{pmatrix}, x \geq \mathbb{0} \end{aligned}$$

Is the (P1) feasible? If so, find a basic feasible solution.
Solution: We follow the following steps to find the feasible solution:

- Since b has a negative coordinate, we multiply the corresponding equations by minus 1 on both sides and then we can obtain the new constraints:

$$\begin{pmatrix} 1 & 2 & -1 \\ 1 & 1 & 1 \end{pmatrix} x = \begin{pmatrix} 1 \\ 3 \end{pmatrix}, x \geq \mathbb{0}$$

- Since we have two equational constraints, we set up two auxiliary variables, x_4, x_5. We form a new LP (P2) that is guaranteed to be feasible:

$$\begin{aligned} \max \ & (0,0,0,-1,-1)x \\ s.t. \ & \begin{pmatrix} 1 & 2 & -1 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 \end{pmatrix} x = \begin{pmatrix} 1 \\ 3 \end{pmatrix}, x \geq \mathbb{0}, x \in \mathbb{R}^5 \end{aligned}$$

  since $(0,0,0,1,3)$ is a basic feasible solution for basis $\{4,5\}$.

- Recall that we required the LP to be feasible when applying Simplex, so now, we run Simpex on (P2), which gives:
  - (P2) is optimal with basis $\{1,3\}$, basic feasible solution $(2,0,1,0,0)^\top$.
  - The optimal value is 0.
- Note that $x_4 = x_5 = 0$ in the basic feasible solution we computed above, deleting x_4, x_5 gives a vector: $\bar{x} = (2,0,1)^\top$, which is a basic feasible solution for (P1) with basis $\{1,3\}$.
- Now we can conclude (P1) is feasible, with basic feasible solution $\bar{x} = (2,0,1)^\top$ with basis $\{1,3\}$.

The method we used above is called **Two-phase simplex**.

In this page, we will provide a procedure of two-phase simplex:

Definition

- **Step 1**: Check if the equality constaints are feasible. We may do this by converting the $Ax = b$ into RREF. If it is consistent, then LP is infeasible, we stop the algorithm, and no certificate of infeasibility is needed.
- **Step 2**: If $Ax = b$ is consistent, there may be redundant constriants. Remove those so that A has full row rank (all rows are linearly independent). This will ensure that we can find a basis (of columns).
- **Step 3**: Use row operations to make $b \geq \mathbb{0}$ (like multiplying minus one on both sides of the equation)
- **Step 4**: We introduce some # of new auxiliary variables (in general, as many as the number of rows). And then augment A with the identity matrix I.
- **Step 5**: We use the Simplex algorithm to try to find a feasible solution to the original LP. If the max is 0, we get a feasible solution. Also, if the LP is feasible, we can augment it by letting all the augmented variables be 0, to get a feasible solution for the auxiliary problem, with value 0. Since the objective function vector is $\leq \mathbb{0}$, the auxiliary problem is bounded, and since it is feasible, Simplex terminates with an optimal solution. This tells us if the original LP is feasible or not: **If optimal value $< 0 \Rightarrow$ (P1) is infeasible, otherwise (P1) is feasible.**

Remark

- The auxiliary LP is always feasible and optimal since it is bounded by 0.
- By running Simplex on the auxiliary LP, if the optimal value $< \mathbb{0}$, then the original LP is infeasible, and the certificate of infeasibility can be derived from the optimal basis of the auxiliary LP. Detailedly, to obtain the certificate $y^\top$:
  - Get $\bar{A}_B^{-\top}\bar{c}_B$, where $\bar{A}, \bar{c}$ are the constraints matrix and objective of original auxiliary LP.
  - Flip the row's sign in $\bar{A}_B^{-\top}\bar{c}_B$ if we fliped such row's sign to let $b \geq \mathbb{0}$.

Finally! We filled the gap we left before! In the later content, we will extend the SEF into the form $Ax \leq b$, and combine the graphs, geometry etc.
In the next few pages, we will provide a complete procedure to determine the type of outcome of an LP, with a certificate. We will also provide some examples of using the Two-phase Simplex Algorithm.

Definition

Start

Input: An LP.

Rewrite this LP into SEF, $\max\{c^\top x : Ax = b, x \geq \mathbb{0}\}$

$[A|b]$ consistent?

False → Return: LP is infeasible.

True

- Make sure the rows of $[A|b]$ are linearly independent, one way is to turn it into RREF: $[A^{'}|b^{'}]$.
- Let $Row_i = -Row_i$ if $b_i^{'} < 0$, define $S = \{i : b_i^{'} < 0\}$.

The new LP is: $\max\{c^\top, A^{"}x = b^{"}, x \geq \mathbb{0}\}$, with $b \geq \mathbb{0}$, $A^{"} \in \mathbb{R}^{m\times n}$.

- Define the auxiliary LP:

$$\max \ (\underbrace{0, 0, \cdots, 0}_{\#n}, \underbrace{-1, -1, \cdots, -1}_{\#m})x$$
$$s.t. \ [A^{"}|I_m]x = b^{"}, x \geq \mathbb{0}, x \in \mathbb{R}^{m+n}$$

- Run **Simplex** on auxiliary LP with basis $\{n+1, n+2, \cdots, n+m\}$:
  - The auxilary LP is optimal with basis $\bar{B}$.
  - Denote the auxiliary LP's optimal value as $\bar{z}$.

$\bar{z} = 0?$

False:

The original LP is infeasible.

- $\bar{y} = (A^{"}_{\bar{B}})^{-\top}(\underbrace{0, 0, \cdots, 0}_{\#n}, \underbrace{-1, -1, \cdots, -1}_{\#m})^\top_{\bar{B}}$
- Obtain y by flipping the sign of $\bar{y}_i$, where $i \in S$.

$y^\top$ is the certificate of infeasibility.

True:

The original LP is feasible.

Run **Simplex** on LP with $\bar{B}$.

Two possible outcomes

This LP is **infeasible** with certificate $y^\top$.

This LP is **Optimal** with certificate $y^\top$.

This LP is **Unboundede** with certificate $\bar{x}, r$.

Example

Determine the outcome of the following LP:

$$\max\ (-1,-4,-4,-2)x$$
$$s.t.\ \underbrace{\begin{pmatrix}2 & 2 & 3 & 3\\ 2 & 0 & 3 & 2\\ 1 & -1 & 1 & 0\end{pmatrix}}_{A} x = \underbrace{\begin{pmatrix}4\\ 2\\ -1\end{pmatrix}}_{b}, x \geq \mathbb{0}$$

Solution:

Step 1: Because this LP is already in SEF, no operation is needed here.

Step 2: Since $[A|b]$ is consistent, we can go to the next step.

Step 3: $b_3 = -1$ is negative, so we multiply the last row by -1.

Step 4: Define the auxiliary LP:

$$\max\ (0,0,0,0,-1,-1,-1)x$$
$$s.t.\ \underbrace{\begin{pmatrix}2 & 2 & 3 & 3 & 1 & 0 & 0\\ 2 & 0 & 3 & 2 & 0 & 1 & 0\\ -1 & 1 & -1 & 0 & 0 & 0 & 1\end{pmatrix}}_{[A|I_3]} x = \underbrace{\begin{pmatrix}4\\ 2\\ 1\end{pmatrix}}_{b"}, x \geq \mathbb{0}, x \in \mathbb{R}^7$$

Step 5: $\bar{x} = (0,0,0,0,4,2,1)^\top$ is a basic feasible solution for basis $\{5,6,7\}$, we now run Simplex on the auxiliary LP above with basis $\{5,6,7\}$.

Step 6: The procedure of Simplex. (We skip the steps here since we're familiar with that, but in Assignments, and Exams, you still have to write it)

Step 7: The last step of Simplex gives us the LP with optimal basis $\bar{B} = \{2,4,7\}$:

$$\max\ (-1/2,0,-1/4,0,-3/2,-1/4,0)x - 1/2$$
$$s.t.\ \begin{pmatrix}-1/2 & 1 & -3/4 & 0 & 1/2 & -3/4 & 0\\ 1 & 0 & 3/2 & 1 & 0 & 1/2 & 0\\ -1/2 & 0 & -1/4 & 0 & -1/2 & 3/4 & 1\end{pmatrix} x = \begin{pmatrix}1/2\\ 1\\ 1/2\end{pmatrix}, x \geq \mathbb{0}, x \in \mathbb{R}^7$$

But the optimal value is $-\frac{1}{2} < 0$, which means the original LP is infeasible.

Step 8: $$\bar{y} = (A_{\bar{B}})^{-\top}(0,0,0,0,-1,-1,-1)^\top_{\bar{B}} = \begin{pmatrix}2 & 3 & 0\\ 0 & 2 & 0\\ 1 & 0 & 1\end{pmatrix}^{-\top} \begin{pmatrix}0\\ 0\\ -1\end{pmatrix} = \begin{pmatrix}1/2\\ -3/4\\ -1\end{pmatrix}.$$

Step 9: Remind that we flipped the third row in Step 3, so we have to flip back: $y = (1/2,-3/4,1)^\top$, then we can get our certificate $y^\top = (1/2,-3/4,1)$.

Step 10: In conclusion, this LP is infeasible, with certificate $y^\top = (1/2,-3/4,1)$.

(* Indeed, we have $y^\top A = (1/2,0,1/4,0) \geq \mathbb{0}$ and $y^\top b = -1/2 < 0$ as expected.)

2.10 Standard inequality Form

In all the sections above, we're deliberately using SEF, and sometimes even if it is an inequality, we convert it into an equation. But actually, working with LP in **Standard Inequality Form** is very convenient too.

Definition

The LP with the form

$$\begin{aligned} \max\ & c^\top x + m \\ s.t.\ & Ax \le b \\ & x \ge \mathbb{0} \end{aligned}$$

is in **Standard Inequality Form (SIF)**.

Remark

We can replace $Ax = b$ in SEF by:

$$Ax \le b, -Ax \le -b$$

Example

Convert the following LP in SEF into SIF:

$$\begin{aligned} \max\ & (1, -1, -9)x \\ s.t.\ & \begin{pmatrix} 1 & 2 & 8 \\ 9 & 10 & -1 \end{pmatrix} x = \begin{pmatrix} 1 \\ -2 \end{pmatrix}, x \ge \mathbb{0} \end{aligned}$$

Solution:

$$\begin{aligned} \max\ & (1, -1, -9)x \\ s.t.\ & \begin{pmatrix} 1 & 2 & 8 \\ -1 & -2 & -8 \\ 9 & 10 & -1 \\ -9 & -10 & 1 \end{pmatrix} x \le \begin{pmatrix} 1 \\ -1 \\ -2 \\ 2 \end{pmatrix}, x \ge \mathbb{0} \end{aligned}$$

Moreover, you can even put $x \ge \mathbb{0}$ in:

$$\begin{aligned} \max\ & (1, -1, -9)x \\ s.t.\ & \begin{pmatrix} 1 & 2 & 8 \\ -1 & -2 & -8 \\ 9 & 10 & -1 \\ -9 & -10 & 1 \\ -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} x \le \begin{pmatrix} 1 \\ -1 \\ -2 \\ 2 \\ 0 \\ 0 \\ 0 \end{pmatrix} \end{aligned}$$

Definition

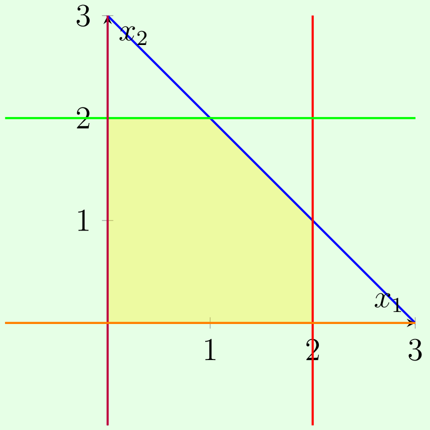
For an **optimal** problem, the **feasible region** is the set of all feasible solutions.

Let's combine this with geometry:

Example

The feasible region of the following LP is:

$$\begin{aligned} &\max\ (1,-1)x \\ &s.t.\ \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} x \leq \begin{pmatrix} 3 \\ 2 \\ 2 \end{pmatrix}, x \geq \mathbb{0} \end{aligned}$$



Note that each constraint (including $x \geq \mathbb{0}$) defines a half-plane, and the feasible region is the intersection of these half-planes. The situation is similar in higher dimensions, but due to the complexity of higher dimensions, we only show the example of $\mathbb{R}^2$.

Definition

- The intersection of **finate** half-planes is also called **polyhedron**.
- A polyhedron is **polytope** if it is bounded.
- The **hyperplane** is the set of solutions to a single linear equation.
- The **half-space** is the set of solutions to a single linear inequality.

Mathematically:

- $\{x : Ax \leq b\}$ is a **polyhedron**.
- $\{x : A_i x = b_i\}$ is a **hyperplane**.
- $\{x : A_i x \leq b_i\}$ is a **half-space**.

Definition

Let $S, S^{'} \subseteq \mathbb{R}^n$, then $S^{'}$ is a **translate** of S if there exists $p \in \mathbb{R}^n$ and

$$S^{'} = \{s + p : s \in S\}$$

Example

Let $\alpha \neq \mathbb{0}$ be a vector and β a real number, and:

- let:
$$H := \{x : \alpha^\top x = \beta\}, H_0 := \{x : \alpha^\top x = 0\}$$
Then, H is a translate of H_0, and H_0 is a translate of H as well.
- let:
$$F := \{x : \alpha^\top x \leq \beta\}, F_0 := \{x : \alpha^\top x \leq 0\}$$
Then, F is a translate of F_0, and F_0 is a translate of F as well.

Definition

Let $x_1, x_2 \in \mathbb{R}^n$, then the **line** through x_1, x_2 is defined as:

$$L := \{x = \lambda x_1 + (1 - \lambda)x_2 : \lambda \in \mathbb{R}\}$$

and the **line segment** between x_1 and x_2 is:

$$S := \{x = \lambda x_1 + (1 - \lambda)x_2 : \lambda \in \mathbb{R}, 0 \leq \lambda \leq 1\}$$

Definition

Given two points $x, y \in \mathbb{R}^n, x \neq y$, the line segment joining x and y in the set $\{\underbrace{\lambda x + (1 - \lambda)y}_{\text{convex combination of } x,y} : 0 \leq \lambda \leq 1\}$, we saya set $S \subset \mathbb{R}^n$ is **convex** if for every pair of points $x, y \in \mathbb{S}, x \neq y$, the line segment joining x and y also is contained in S.

Remark

If a set $S \subseteq \mathbb{R}^n$ is a polyhedron, then it is convex.

Proof. Suppose a polyhedron $S \subseteq \mathbb{R}^n$ is specified by inequalities $Ax \leq b$. Suppose, $a, a^{'} \in S, 0 \leq \lambda \leq 1$, we have:

- Since $\lambda \geq 0$, $Aa \leq b$, we have $\lambda Aa \leq \lambda b$.
- Since $1 - \lambda \geq 0$, $Aa^{'} \leq b$, we have $(1 - \lambda)Aa^{'} \leq (1 - \lambda)b$.

Adding up the two inequalities above, we have:

$$A(\underbrace{\lambda a + (1 - \lambda)a^{'}}_{\in S}) \leq b$$

Therefore, any polyhedra is convex, and we can also deduce that the feasible region of an LP is convex. □

Remark

A circle: $C := \{x : x_1^2 + x_2^2 \leq 1\}$ is convex.

Proof. Let $a, b \in C, a \neq b$, we wish the convex combination of a and b are also in C. That is, we wish to prove:

$$(\lambda a_1 + (1-\lambda)a_2)^2 + (\lambda b_1 + (1-\lambda)b_2)^2 \leq 1, 0 \leq \lambda \leq 1$$

From the LHS:

$$\begin{aligned}
&(\lambda a_1 + (1-\lambda)a_2)^2 + (\lambda b_1 + (1-\lambda)b_2)^2 \\
=&\lambda^2(a_1^2 + a_2^2) + 2\lambda(1-\lambda)(a_1a_1 + b_1b_2) + (1-\lambda)^2(b_1^2 + b_2^2) \\
\leq&\lambda^2 + 2\lambda(1-\lambda)(\frac{(a_1^2 + a_2^2) + (b_1^2 + b_2^2)}{2}) + (1-\lambda)^2 \\
\leq&\lambda^2 + 2\lambda(1-\lambda) + (1-\lambda)^2 = 1
\end{aligned}$$

Note that we used the inequalities: $mn \leq \frac{m^2+n^2}{2}, a_1^2 + a_2^2 \leq 1, b_1^2 + b_2^2 \leq 1$.
Therefore, a circle is indeed convex. □

2.11 Extreme Points

Definition

Point $x \in \mathbb{R}^n$ is **properly contained** in the line segment L if

- $x \in L$ and x is **not** the endpoints of L.

Definition

Point $x \in \mathbb{R}^n$ is an **extreme point** for S if and only if there's no $0 < \lambda < 1$ and no two points $a_1, a_2 \in S, a_1 \neq a_2$ such that $x = \lambda a_1 + (1-\lambda)a_2$.

Remark

By the definitions above, we can also deduce:
Point $x \in \mathbb{R}^n$ is **NOT** an **extreme point** for S if there exists a line segment $L \subseteq S$ where L properly contains x.

Remark

A convex set may have **infinite** number of extreme points.

Proof. A circle $C := \{x : x_1^2 + x_2^2 \leq 1\}$ can be an example of our remark. Assume, there exists a non-extreme point x on the boundary of C $(x_1^2 + x_2^2 = 1)$, then, there exists a line segment $L \subseteq C$ where L properly contains x. But by definition, a circle is curved everywhere, which means there's even no segment in a circle. This leads to a contradiction to our assumption. Therefore, every point in the boundary of a circle is an extreme point. And since there are infinite points in the boundary of the circle, there are an infinite number of extreme points. □

Definition

Suppose $d \in P$, we say a constraint $a_i^\top \leq b_i$ is **tight** for d_i if $a_i^\top d = b_i$. The set of all tight constraints is denoted $\bar{A}x \leq \bar{b}$.

Example

Consider $Ax = b$: $\begin{pmatrix} 1 & 2 \\ 4 & 5 \end{pmatrix} x \leq \begin{pmatrix} 3 \\ 10 \end{pmatrix}$, then:

- The first constraint is tight for $d = (1, 1)$.
- The second constraint is tight for $d = (5/2, 0)$.

Theorem

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

- If $\operatorname{rank}(\bar{A}) = n$, then $\bar{x}$ is an extreme point.
- If $\operatorname{rank}(\bar{A}) < n$, then $\bar{x}$ is **Not** an extreme point.

We can use the following remark to prove this theorem.

Remark

Let $a, b, c \in \mathbb{R}$, and $0 < \lambda < 1$, then if

$$a = \lambda b + (1 - \lambda)c, b \leq a, c \leq a$$

the $a = b = c$.

- Proof of remark:

  Proof. Fix c and $a \neq b$, i.e. $a > b$, then:

  $$a = \lambda b + (1 - \lambda)c < \lambda a + (1 - \lambda)c \leq \lambda a + (1 - \lambda)a = a$$

  contradicts the fact $a = a$. Similarly, if we fix b and $a \neq c$, i.e. $a > c$, then:

  $$a = \lambda b + (1 - \lambda)c < \lambda b + (1 - \lambda)a \leq \lambda a + (1 - \lambda)a = a$$

  Therefore, we must have $a = b = c$. □

- Proof of theorem:

  Proof. Suppose $\bar{x}$ is not an extreme point, then there exists a line segment L_S connecting x_1, x_2, such that $\bar{x} = \lambda x_1 + (1 - \lambda)x_2 \rightarrow \bar{A}\bar{x} = \bar{b} = \lambda \bar{A}x_1 + (1 - \lambda)\bar{A}x_2$ with some $0 < \lambda < 1$. Since $\bar{A}x_1 \leq \bar{b}$ and $\bar{A}x_2 \leq \bar{b}$, we have $\bar{b} = \bar{A}x_1 = \bar{A}x_2$. If $\operatorname{rank}(A) = n$, then A is invertible, which means the solution of $\bar{A}x = \bar{b}$ is unique, hence, $\bar{x} = \bar{x_1} = \bar{x_2}$ means that $\bar{x}$ is an extreme point, contradicting the assumption. Then $\operatorname{rank}(A) = n$.

  By its contrapositive, the original statement is then true. □

Remark

Consider the LP: $\max\{c^\top x : Lx = d, x \geq \mathbb{0}\}$, where $c \in \mathbb{R}^n, L \in \mathbb{R}^{m\times n}, d \in \mathbb{R}^m$, and L has linearly independent rows. Then,

- Any extreme point of the feasible region of the LP is a basic feasible solution.

Proof. Let $\bar{x}$ be an arbitrary extreme point in this LP.

- Assume for contradiction, it is not a basic feasible solution. Say $B = \{i : \bar{x}_i > 0\}$, with $|B| = k$.
- Then L_B must not have linearly independent columns, since if it has independent columns, then $\bar{x}$ would be basic, which contradicts our assumption.
- Let $d' \neq \mathbb{0} \in \mathbb{R}^k$, such that $L_C d' = 0$. And define $d \in \mathbb{R}^n$ with $d_i = d'_i$ for $i \in C$, and $d_i = 0$ otherwise.
- Then, for some small number $\epsilon > 0$, we have distinct points $\bar{x} \pm \epsilon d$ both in P, since $L_C d' = 0$ and thus $Cd = 0$, and with $\bar{x} \pm \epsilon d \geq 0$ for some small ϵ since $d_i = 0$ whenever $\bar{x}_i = 0$.
- But since $\bar{x}$ is an extreme point, there are no such two distinct points, and this leads to a contradiction.
- Therefore, the $\bar{x}$ must be basic, and it is a feasible basic solution since it satisfies every constraint well.

□

Here is an example of applying the remark:

Example

Find all extreme points of the polyhedron

$$P := \left\{ x \in \mathbb{R}^4 : \begin{pmatrix} 1 & -1 & 3 & 1 \\ -1 & 3 & 0 & 1 \end{pmatrix} x = \begin{pmatrix} 3 \\ 1 \end{pmatrix}, x \geq \mathbb{0} \right\}$$

Solution: Firstly, the matrix $\begin{pmatrix} 1 & -1 & 3 & 1 \\ -1 & 3 & 0 & 1 \end{pmatrix}$ has linearly independent rows obviously, and this is the condition we can use the remark's conclusion in this question. And since any extreme point of the feasible region of the LP is a basic feasible solution, we just simply find every basic feasible solution of this LP and check if they are extreme points for polyhedra R by using theorem:

- Let $P = \{x \in \mathbb{R}^n, Ax \leq b\}$ be a polyhedron and let $\bar{x} \in P$.
  - If $\text{rank}(\bar{A}) = n$, then $\bar{x}$ is an extreme point.
  - If $\text{rank}(\bar{A}) < n$, then $\bar{x}$ is **Not** an extreme point.

Example

And also, we provide the standard inequality form for R: (say R')

$$R' = \left\{ x \in \mathbb{R}^4 : \begin{pmatrix} 1 & -1 & 3 & 1 \\ -1 & 3 & 0 & 1 \\ -1 & 1 & -3 & -1 \\ 1 & -3 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} x \leq \begin{pmatrix} 3 \\ 1 \\ -3 \\ -1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\}$$

- Basis $B = \{1, 2\}$, (it is a valid basis), the corresponding basic solution is (5, 2, 0, 0), and it is feasible. By checking this in R' using theorem 2.20, we have:

$$\bar{A} = \begin{pmatrix} 1 & -1 & 3 & 1 \\ -1 & 3 & 0 & 1 \\ -1 & 1 & -3 & -1 \\ 1 & -3 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}, \bar{b} = \begin{pmatrix} 3 \\ 1 \\ -3 \\ -1 \\ 0 \\ 0 \end{pmatrix}$$

  And since $rank(\bar{A}) = 4 = n$, we can say this is an extreme point.

- Basis $B = \{1, 3\}$, (it is a valid basis), the corresponding basic solution is (-1, 0, 4/3, 0), which is not feasible since it is $\not\geq 0$.

- Basis $B = \{1, 4\}$, (it is a valid basis), the corresponding basic solution is (1, 0, 0, 2), and it is feasible. By checking this in R' using theorem 2.20, we have:

$$\bar{A} = \begin{pmatrix} 1 & -1 & 3 & 1 \\ -1 & 3 & 0 & 1 \\ -1 & 1 & -3 & -1 \\ 1 & -3 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \bar{b} = \begin{pmatrix} 3 \\ 1 \\ -3 \\ -1 \\ 0 \\ 0 \end{pmatrix}$$

  And since $rank(\bar{A}) = 4 = n$, we can say this is an extreme point.

- By using the same method, we can find the rest of the extreme points:(0, 1/3, 10/9, 0), (0, 0, 2/3, 1).

Therefore, the all extreme points are:

$$\begin{aligned} &(5, 2, 0, 0) \\ &(1, 0, 0, 2) \\ &(0, 1/3, 10/9, 0) \\ &(0, 0, 2/3, 1) \end{aligned}$$

We have a theorem that builds the relation between the extreme point and the basic feasible solution: (no proof provided here)

Theorem

Suppose we have an LP in SEF, then the extreme points of its feasible region are exactly the basic feasible solution of the LP.

By using this theorem, after finding the basic feasible solutions, we don't have to check if it is an extreme point anymore since they are one-to-one correspondence.

Till now, we learned the full procedure to determine the outcomes with certificates for an LP and also learned the extreme points, which gives us a better understanding of LP. Later, our key point is Duality, which can make it even easier to get the optimal solution and value. We will also look back at the Shortest path problems and perfect matching problems!

Chapter 3

Duality

3.1 Weak Duality

Start with an example:

$$\begin{aligned} \min \;\; & (2,3)x \\ s.t. \;\; & \begin{pmatrix} 2 & 1 \\ 1 & 1 \\ -1 & 1 \end{pmatrix} x \geq \begin{pmatrix} 20 \\ 18 \\ 8 \end{pmatrix}, x \geq \mathbb{0} \end{aligned}$$

We wish to find a lower bound to the objective value. Suppose x is a feasible solution, then x satisfies:

$$\begin{aligned} y_1 \cdot (2,1)x &\geq y_1 \cdot 20 \\ y_2 \cdot (1,1)x &\geq y_2 \cdot 18 \\ y_3 \cdot (-1,1)x &\geq y_3 \cdot 8 \end{aligned}$$

for $y_1, y_2, y_3 \geq 0$. Adding up these inequalities gives:

$$\begin{aligned} (2y_1 + y_2 - y_3, y_1 + y_2 + y_3)x &\geq 20y_1 + 18y_2 + 8y_3 \\ \Rightarrow \begin{pmatrix} y_1 & y_2 & y_3 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 1 \\ -1 & 1 \end{pmatrix} x &\geq \begin{pmatrix} y_1 & y_2 & y_3 \end{pmatrix} \begin{pmatrix} 20 \\ 18 \\ 8 \end{pmatrix} \end{aligned}$$

Then, we have our objective function:

$$\begin{aligned} (2,3)x &\geq (2,3)x + \underbrace{\begin{pmatrix} y_1 & y_2 & y_3 \end{pmatrix} \begin{pmatrix} 20 \\ 18 \\ 8 \end{pmatrix} - \begin{pmatrix} y_1 & y_2 & y_3 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 1 \\ -1 & 1 \end{pmatrix} x}_{\leq 0} \\ &= \underbrace{\begin{pmatrix} y_1 & y_2 & y_3 \end{pmatrix} \begin{pmatrix} 20 \\ 18 \\ 8 \end{pmatrix}}_{a} + \left(\underbrace{(2,3) - \begin{pmatrix} y_1 & y_2 & y_3 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 1 \\ -1 & 1 \end{pmatrix}}_{b} \right) \underbrace{x}_{\geq \mathbb{0}} \end{aligned}$$

If $b \leq 0$, then the lower bound is a. The optimal value of our LP is then the greatest lower bound. i.e. maximization of a.

Let us reorder our idea:

- We wish to maximize a, with constraints $b \leq 0, y \geq \mathbb{0}$

This gives us a new LP!

$$\max \quad \begin{pmatrix} y_1 & y_2 & y_3 \end{pmatrix} \begin{pmatrix} 20 \\ 18 \\ 8 \end{pmatrix}$$
$$s.t. \ (2,3) - \begin{pmatrix} y_1 & y_2 & y_3 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 1 \\ -1 & 1 \end{pmatrix} \geq 0, y \geq \mathbb{0}$$

Still not very familiar, rewrite it by transposing, we obtain:

$$\max \quad \begin{pmatrix} 20, 18, 8 \end{pmatrix} y$$
$$s.t. \quad \begin{pmatrix} 2 & 1 & -1 \\ 1 & 1 & 1 \end{pmatrix} y \leq (2,3), y \geq \mathbb{0}$$

This LP is called the **Dual** of the original LP. Let's go further:

- Use the method to determine the outcomes on the Dual LP gives: the optimal value is 49, with optimal solution $\bar{y} = (0, 5/2, 1/2)^\top$.
- Then 49 is also the optimal value of the original LP.

This method supports us to solve the LP by solving another LP.

Definition

The LP (D) is called **dual** of the **primal** (original) LP (P):

- (P): $\min\{c^\top x : Ax \geq b, x \geq \mathbb{0}\}$.
- (D): $\max\{b^\top y : A^\top y \leq c, y \geq \mathbb{0}\}$.

Weak Duality

Let (P) be a minimization problem, and (D) be a maximization problem.

- If $\bar{x}$ is feasible for (P), $\bar{y}$ is feasible for (D), then $b^\top \bar{y} \leq c^\top \bar{x}$.
- If $\bar{x}$ is optimal solution for (P), $\bar{y}$ is optimal solution for (D), then $b^\top \bar{y} = c^\top \bar{x}$.

Proof.

$$\begin{aligned} b^\top \bar{y} &= \bar{y}^\top b \\ &\leq \bar{y}^\top (A\bar{x}) && \text{equal if } \bar{x} \text{ optimal} \\ &= (A^\top \bar{y})^\top \bar{x} \\ &\leq c^\top \bar{x} && \text{equal if } \bar{y} \text{ optimal} \end{aligned}$$

□

That proves both points of our theorem, we will extend this theorem later.

In the primal-dual pair:

- (P): $\min\{c^\top x : Ax \geq b, x \geq \mathbb{0}\}$.
- (D): $\max\{b^\top y : A^\top y \leq c, y \geq \mathbb{0}\}$.

We can find some interesting corresponding:

- Each non-negative variable x_e in (P) corresponds to a $\leq$-constraint in (D).
- Each $\geq$-constraints in (P) corresponds to a non-negative variable y_e in (D).
- Primal variables $\equiv$ Dual constraints.
- Primal constraints $\equiv$ Dual variabless.

The following table shows how constraints and variables in primal and dual LPs correspond:

<table>
<tr><td colspan="2">(P_{max})</td><td colspan="2">(P_{min})</td></tr>
<tr><td rowspan="3">constraint</td><td>$\leq$</td><td rowspan="3">variable</td><td>≥ 0</td></tr>
<tr><td>$=$</td><td>free</td></tr>
<tr><td>$\geq$</td><td>≤ 0</td></tr>
<tr><td rowspan="3">variable</td><td>≥ 0</td><td rowspan="3">constraint</td><td>$\geq$</td></tr>
<tr><td>free</td><td>$=$</td></tr>
<tr><td>≤ 0</td><td>$\leq$</td></tr>
<tr><td colspan="2">$\max\ c^\top x$
$s.t.\ Ax \underline{?} b$
$x \underline{?} \mathbb{0}$</td><td colspan="2">$\min\ b^\top y$
$s.t.\ A^\top y \underline{?} c$
$y \underline{?} \mathbb{0}$</td></tr>
</table>

Example

(P):

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

(D):

$$\begin{aligned} \min\ & (3,4)y \\ s.t.\ & \begin{pmatrix} 3 & 1 \\ -1 & 0 \\ 0 & 1 \end{pmatrix} y \begin{matrix} \geq \\ \geq \\ = \end{matrix} \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} \\ & y_1 \geq 0, y_2\ free \end{aligned}$$

Weak Duality Theorem

Let (P_{max}) and (P_{min}) be primal-dual pair. Say $\bar{x}$ is feasible for (P_{max}), $\bar{y}$ is feasible for (P_{min}), then we have:

$$c^\top \bar{x} \leq b^\top \bar{y}$$

If they are equal, then $\bar{x}$ is optimal for (P_{max}) and $\bar{y}$ is optimal for (P_{min}).

Remark

1. If the primal LP is unbounded, then its dual is infeasible.
2. If the primal LP and its dual are feasible, then they both have optimal solutions.

Proof. 1. Suppose, the primal LP is a maximization problem. For its contradictive, that $\bar{y}$ is feasible for the dual. By Weak Duality, $c^\top \bar{x} \leq b^\top \bar{y}$ for all $\bar{x}$ that feasible for the primal LP, this means the objective value of the primal LP has an upper bound, which leads to a contradiction. Therefore, the dual is not feasible. The situation in which the primal LP is a minimization problem is similar.

2. By Weak Duality, both of them are bounded, therefore both have optimal solutions. □

3.2 Strong Duality

Strong Duality Theorem

Let (P_{max}) and (P_{min}) be primal-dual pair.

- If one has an optimal solution, then the other has an optimal solution too.

Say $\bar{x}$ is optimal for (P_{max}), $\bar{y}$ is optimal for (P_{min}), then we have:

$$c^\top \bar{x} = b^\top \bar{y}$$

The table of possible combinations of primal-dual pairs of LPs:

(D) \ (P)	optimal	unbounded	infeasible
optimal	✓	×	×
unbounded	×	×	✓
infeasible	×	✓	✓

Example

(P) and (D) both infeasible:

- (P): $\min \left\{ x_1 + 2x_2 : \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} x = \binom{1}{3}, x \ free \right\}$.
- (D): $\max \left\{ y_1 + 3y_2 : \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} y = \binom{1}{2}, y \ free \right\}$.

Example

(P) infeasible, (D) unbounded:

- (P): min $\left\{x_1 + x_2 : \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} x = \binom{1}{3}, x\ free\right\}$.
- (D): max $\left\{y_1 + 3y_2 : \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} y = \binom{1}{1}, y\ free\right\}$.

Remark

Consider the LPs:

- (P): $\max\{c^\top x : Ax \leq b, x\ free\}$.
- $(P^{'})$: $\max\{c^\top x : Ax \leq b, x \geq \mathbb{0}\}$.

Suppose that (P) has an optimal solution, $(P^{'})$ is infeasible. Then the dual of $(P^{'})$ is unbounded.

Proof. Let (D) be the dual of (P), $(D^{'})$ be the dual of $(P^{'})$, we have:

- (D): $\max\{b^\top y : A^\top y = c, y \geq \mathbb{0}\}$.
- $(D^{'})$: $\max\{b^\top y : A^\top y \geq c, y \geq \mathbb{0}\}$.
- Since (P) is optimal, by Strong Duality Theorem, (D) is optimal (at $\bar{y}$).
- Since $(P^{'})$ is unbounded, by the table of possible primal-dual pairs, $(D^{'})$ is infeasible or unbounded.

Now, because (D) is optimal (feasible) at $\bar{y}$, we have $A^\top \bar{y} = c, \bar{y} \geq \mathbb{0}$, we can found that $\bar{y}$ is also a feasible solution to $(D^{'})$. This proves $(D^{'})$ is not infeasible. Hence, $(D^{'})$ is unbounded. $\square$

In the later few pages, we will revisit some problems we've already introduced.

3.3 Shortest Path problem revisit

Given the shortest path instance $G = (V, E)$ with $s, t \in V, c_e \geq 0$ for all $e \in E$, then the shortest-path IP is:

$$\begin{aligned} \min \ & \sum (c_e x_e : e \in E) \\ s.t. \ & \sum (x_e : e \in \delta(U), U \subseteq V, s \in U, t \notin U) \geq 1 \\ & x \geq \mathbb{0}, x \in \mathbb{Z} \end{aligned}$$

Definition

By dropping the **integrality** restriction of this IP, the result LP is called **relaxation** of the original IP.

Remark

Since we obtain the relaxation LP by dropping the integrality restriction of the original IP, the range of objective value of the original IP is the subset of relaxation LP's. So if we can find the optimal value of the relaxation LP, then the optimal value can be a lower bound of the original IP.

We can write the shortest-path LP relaxation as:

$$\begin{aligned} \min \ & c^\top x \\ s.t. \ & Ax \geq \mathbb{1} \\ & x \geq \mathbb{0} \end{aligned}$$

where

- A has a column for every edge and a row for every s,t-cut $\delta(U)$.
- $A[U, e] = 1$ if $e \in \delta(U)$, and 0 otherwise.

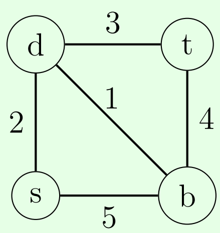
We can obtain its dual:

$$\begin{aligned} \max \ & \mathbb{1}^\top y \\ s.t. \ & A^\top y \leq c \\ & y \geq \mathbb{0} \end{aligned}$$

Let's go further on this dual, can we find that the new constraints have some special meaning in the graph?

Examples

Consider the graph below, we wish to find the shortest $s, t - path$.



Examples

We can write an IP for this problem:

- **Objective function**:

$$\sum(c_e x_e : e \in E) = 5x_{sb} + 4x_{bt} + 3x_{td} + 2x_{ds} + x_{bd}$$

- **Constraints**:
  - $U = \{s\}, \delta(U) = \{sb, sd\} \Rightarrow x_{sd} + x_{sb} \geq 1$.
  - $U = \{s, d\}, \delta(U) = \{sb, db, dt\} \Rightarrow x_{sb} + x_{db} + x_{dt} \geq 1$.
  - $U = \{s, b\}, \delta(U) = \{sd, bd, bt\} \Rightarrow x_{sd} + x_{bd} + x_{bt} \geq 1$.
  - $U = \{s, b, d\}, \delta(U) = \{dt, bt\} \Rightarrow x_{dt} + x_{bt} \geq 1$.
  - $x \geq \mathbb{0}, x \in \mathbb{Z}$.

IP:

$$\begin{aligned} \min\ & (5,4,3,2,1)x \\ s.t.\ & \\ & \begin{array}{r c} & \begin{matrix} sb & bt & td & ds & bd \end{matrix} \\ \begin{matrix} \delta(\{s\}) \rightarrow \\ \delta(\{s,d\}) \rightarrow \\ \delta(\{s,b\}) \rightarrow \\ \delta(\{s,b,d\}) \rightarrow \end{matrix} & \begin{pmatrix} 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 \end{pmatrix} \end{array} x \geq \mathbb{1}, x \geq \mathbb{0}, x \in \mathbb{Z} \end{aligned}$$

Our (relaxation) dual is:

$$\begin{aligned} \max\ & (1,1,1,1)y \\ s.t.\ & \\ & \begin{array}{r c} & \begin{matrix} \underbrace{\delta(\{s\})}_{\downarrow} & \uparrow\ \overbrace{\delta(\{s,d\})} & \underbrace{\delta(\{s,b\})}_{\downarrow} & \uparrow\ \overbrace{\delta(\{s,b,d\})} \end{matrix} \\ \begin{matrix} sb \rightarrow \\ bt \rightarrow \\ td \rightarrow \\ ds \rightarrow \\ bd \rightarrow \end{matrix} & \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix} \end{array} y \leq \begin{pmatrix} 5 \\ 4 \\ 3 \\ 2 \\ 1 \end{pmatrix}, y \geq \mathbb{0} \end{aligned}$$

We can then conclude the constraints:

$$\sum(y_U : \underbrace{\delta(U)}_{s,t\text{-cut contains } e}) \leq c_e, e \in E$$

In the rest of this section, we will present an algorithm to solve the shortest path problem based on our duality, before that, we need to know some definitions:

Definition

We call the ordered pairs of vertices **arcs**.

Example

We denote an arc from u to v as $\vec{uv}$, which is different from $\vec{vu}$.

Definition

A **directed path** in G is a **sequence of arcs**

$$\vec{v_1v_2}, \vec{v_2v_3}, \cdots, \vec{v_{k-1}v_k}$$

where all $v_iv_{i+1} \in E(G)$, and $v_i \neq v_j$ if $i \neq j$.

Definition

Let y be a feasible dual solution. The **slack** of an edge $e \in E$ is defined as:

$$slack_y(e) = c_e - \sum(y_U : \underbrace{\delta(U)}_{s,t\text{-cut contains } e})$$

Now, we introduce the algorithm to find the shortest path:

Remark

Input: Graph $G = (V, E)$, costs $c_e \geq 0$ for all $e \in E. s, t \in V, s \neq t$.
Output: The shortest s, t-path P.

1: Start with $y = \mathbb{0}$, then set $U := \{s\}$.

2: **while** $t \notin U$, **do**:

- Let ab be an edge in $\delta(U)$, with the smallest slack for y where $a \in U, b \notin U$.
- Set $y_U := slack_y(ab)$.
- Reset $U := \underbrace{U}_{\text{original } U} \cup \{b\}$.
- Change edge ab into an arc $\vec{ab}$.

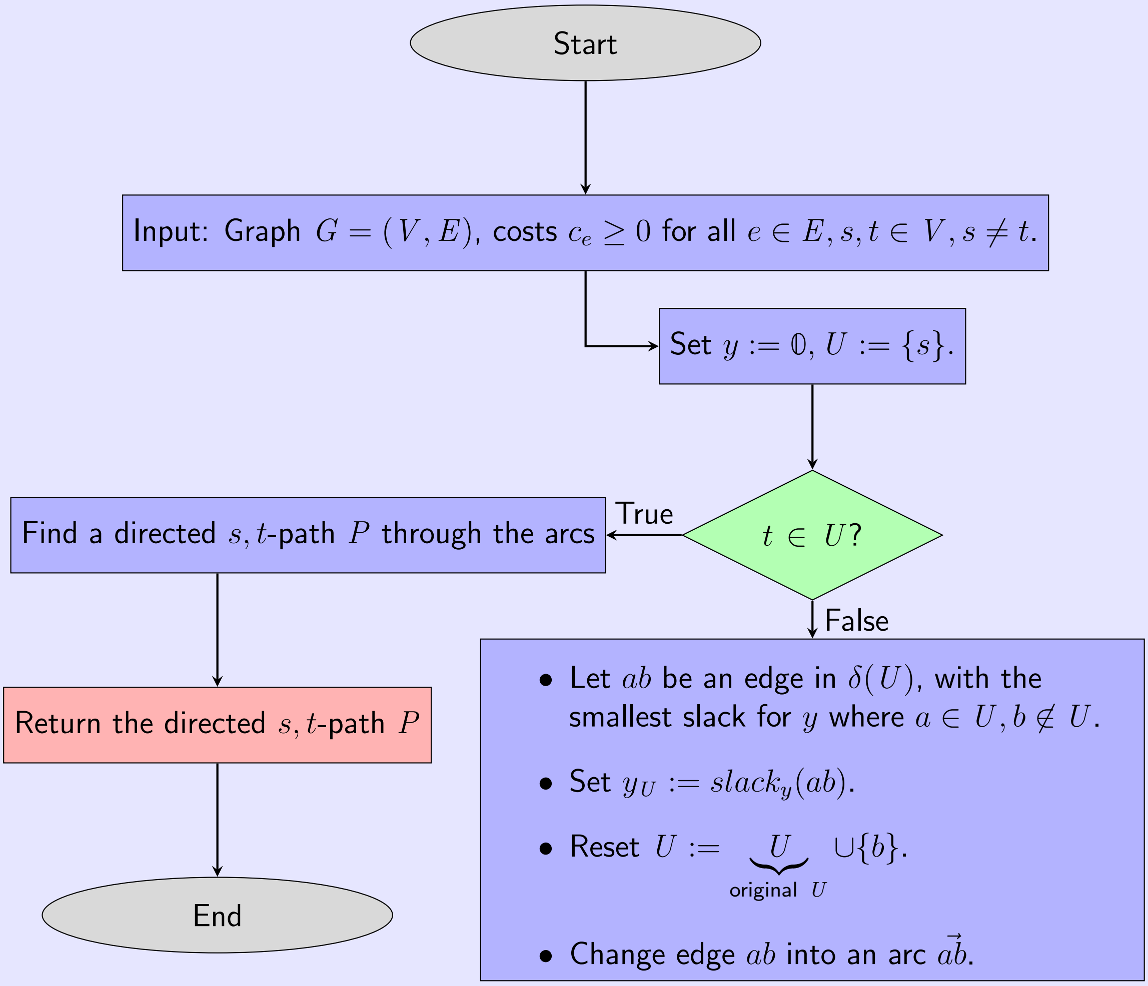
3: **End while**.

4: Find a directed s, t-path P through the arcs.

5: **return** a directed s, t-path P.

* Note that we won't cover the method of finding a directed s, t-path P through the arcs. As before, a procedure is provided. (On the next page)

Remark

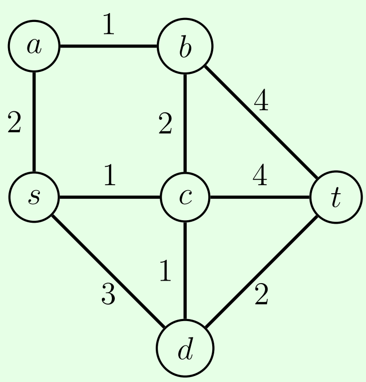


Remark

If we have a choice, just choose randomly.

Example

We will apply the algorithm on the following graph:



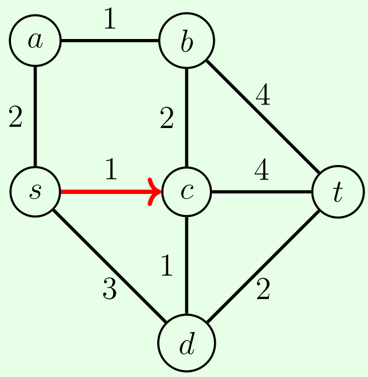
* The solution is on the next page.

Example

Solution:

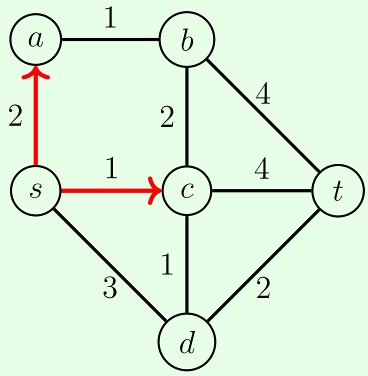
- Start with $U = \{s\}$, $\delta(U) = sa, sc, sd$.
  - $slack_y(sa) = 2 - 0 = 2$.
  - $slack_y(sc) = 1 - 0 = 1$.
  - $slack_y(sa) = 3 - 0 = 3$.

  We can find $slack_y(sc) = 1$ is the smallest one.
- Set $y_U := slack_y(sc) = 1$.
- Reset $U := \{s\} \cup \{c\} = \{s, c\}$.
- Change edge sc into an arc $\vec{sc}$.



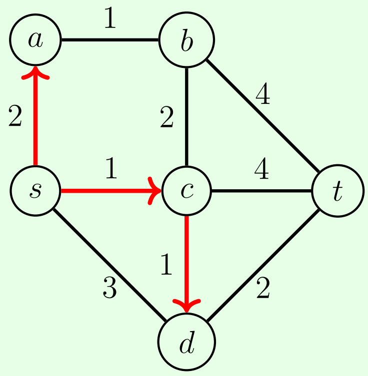
- Start with $U = \{s, c\}$, $\delta(U) = sa, cb, ct, cd, sd$.
  - $slack_y(sa) = 2 - 1 = 1$.
  - $slack_y(cb) = 2 - 0 = 2$.
  - $slack_y(ct) = 4 - 0 = 4$.
  - $slack_y(cd) = 1 - 0 = 1$.
  - $slack_y(sd) = 3 - 1 = 2$.

  We can find $slack_y(sa) = 0$ is the smallest one.
- Set $y_U := slack_y(sa) = 1$.
- Reset $U := \{s, c\} \cup \{a\} = \{s, c, a\}$.
- Change edge sa into an arc $\vec{sa}$.



- Start with $U = \{s, c, a\}$, $\delta(U) = ab, cb, ct, cd, sd$.
  - $slack_y(ab) = 1 - 0 = 1$.
  - $slack_y(cb) = 2 - 1 = 1$.
  - $slack_y(ct) = 4 - 1 = 3$.
  - $slack_y(cd) = 1 - 1 = 0$.
  - $slack_y(sd) = 3 - 1 = 2$.

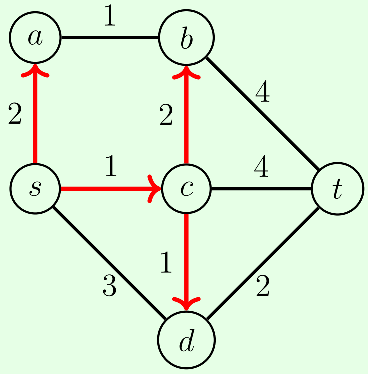
  We can find $slack_y(sa) = 1$ is the smallest one.
- Set $y_U := slack_y(cd) = 0$.
- Reset $U := \{s, c, a\} \cup \{d\} = \{s, c, a, d\}$.
- Change edge cd into an arc $\vec{cd}$.



Example

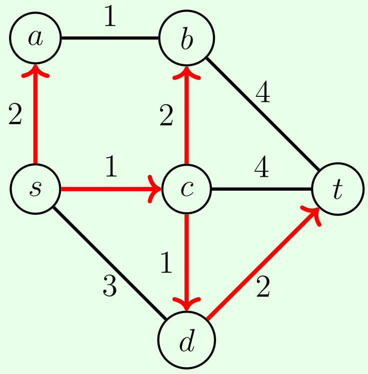
- Start with $U = \{s, c, a, d\}$, $\delta(U) = ab, cb, ct, dt$.
  - $slack_y(ab) = 1 - 0 = 1$.
  - $slack_y(cb) = 2 - 1 = 1$.
  - $slack_y(ct) = 4 - 1 = 3$.
  - $slack_y(dt) = 2 - 0 = 2$.

  We can find $slack_y(cb) = 1$ is the smallest one.
- Set $y_U := slack_y(cb) = 1$.
- Reset $U := \{s, c, a, d\} \cup \{b\} = \{s, c, a, d, b\}$.
- Change edge cb into an arc $\vec{cb}$.



- Start with $U = \{s, c, a, d, b\}$, $\delta(U) = bt, ct, dt$.
  - $slack_y(bt) = 4 - 0 = 4$.
  - $slack_y(ct) = 4 - 2 = 2$.
  - $slack_y(dt) = 2 - 1 = 1$.

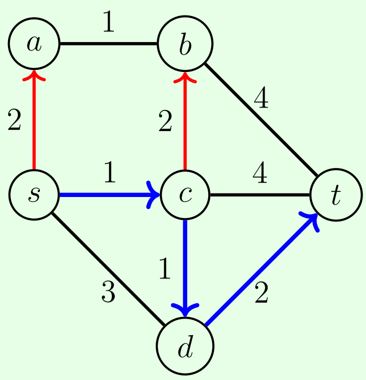
  We can find $slack_y(dt) = 1$ is the smallest one.
- Set $y_U := slack_y(dt) = 1$.
- Reset $U := \{s, c, a, d\} \cup \{t\} = \{s, c, a, d, b, t\}$.
- Change edge dt into an arc $\vec{dt}$.



Now $t \in U = \{s, c, a, d, b, t\}$, we can find an s, t-path easily through the arcs:

$$P = \vec{dc}, \vec{cd}, \vec{dt}$$

with objective value 4.



We also have our feasible dual solution:

$$y_{\{s\}} = y_{\{s,c\}} = y_{\{s,c,a,d\}} = y_{\{s,c,a,b,d\}} = 1$$

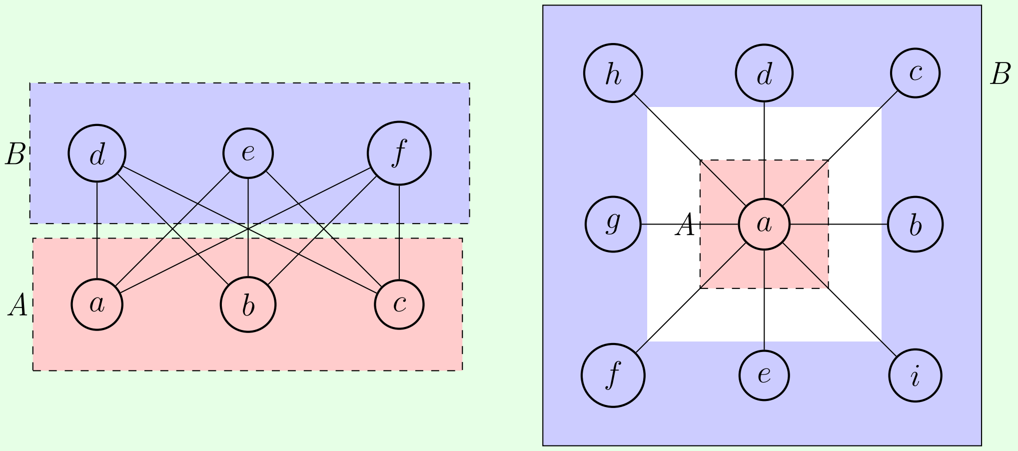
and $y_U = 0$ otherwise. Therefore, we find P as the **shortest path**!

3.4 Perfect Matching problem revisit

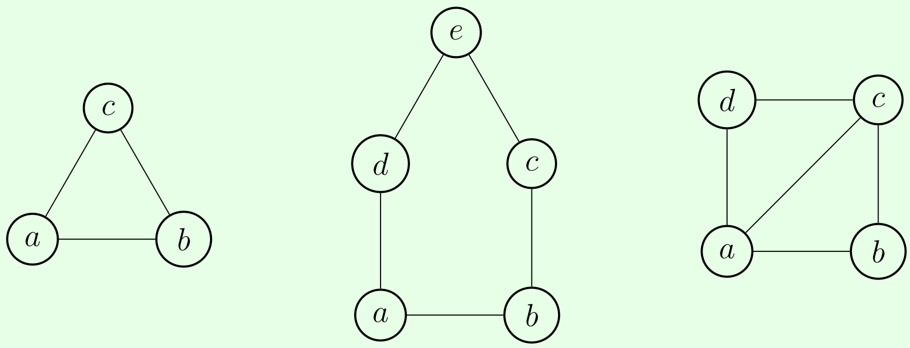
Definition

A graph is **bipartite** if and only if there is a partition (A, B) of the vertices such that all edges join a vertex in A with a vertex in B.

Examples of bipartite graph



Examples of non-bipartite graph



Remark

- A **cycle** is a path with the same start and end point.
- A graph is bipartite if and only if it doesn't contain a cycle with odd # of edges.

* One taking this course doesn't need to know the knowledge of the remark, putting it here just aims for a better determination.
On the next page, we will introduce the algorithm to find the minimum-cost perfect matching. Before that, a few definitions are needed.

Definition

- Let S be a set, then $|S|$ is the # of elements in S.
- Consider a graph $G = (V, E)$, let $S \subseteq V$ be a set of vertices. Then, we denote the set of neighbors of S in graph G by

$$N_G(S) := \{r \in V/S : sr \in E, s \in S\}$$

- Let $S \subseteq V$ be a set of vertices, and with its neighbors set $N_G(S)$. If $|S| > |N_G(S)|$, then we call such S a **deficient set**.

Theorem

Let $G = (V, E)$ be a bipartite graph with bipartition U, W, where $|U| = |W|$. Then there **exists** a perfect matching M in G if and only if G has **NO** deficient set $S \subseteq U$.

Here is the algorithm to find the minimum-cost perfect matching for bipartite graphs.

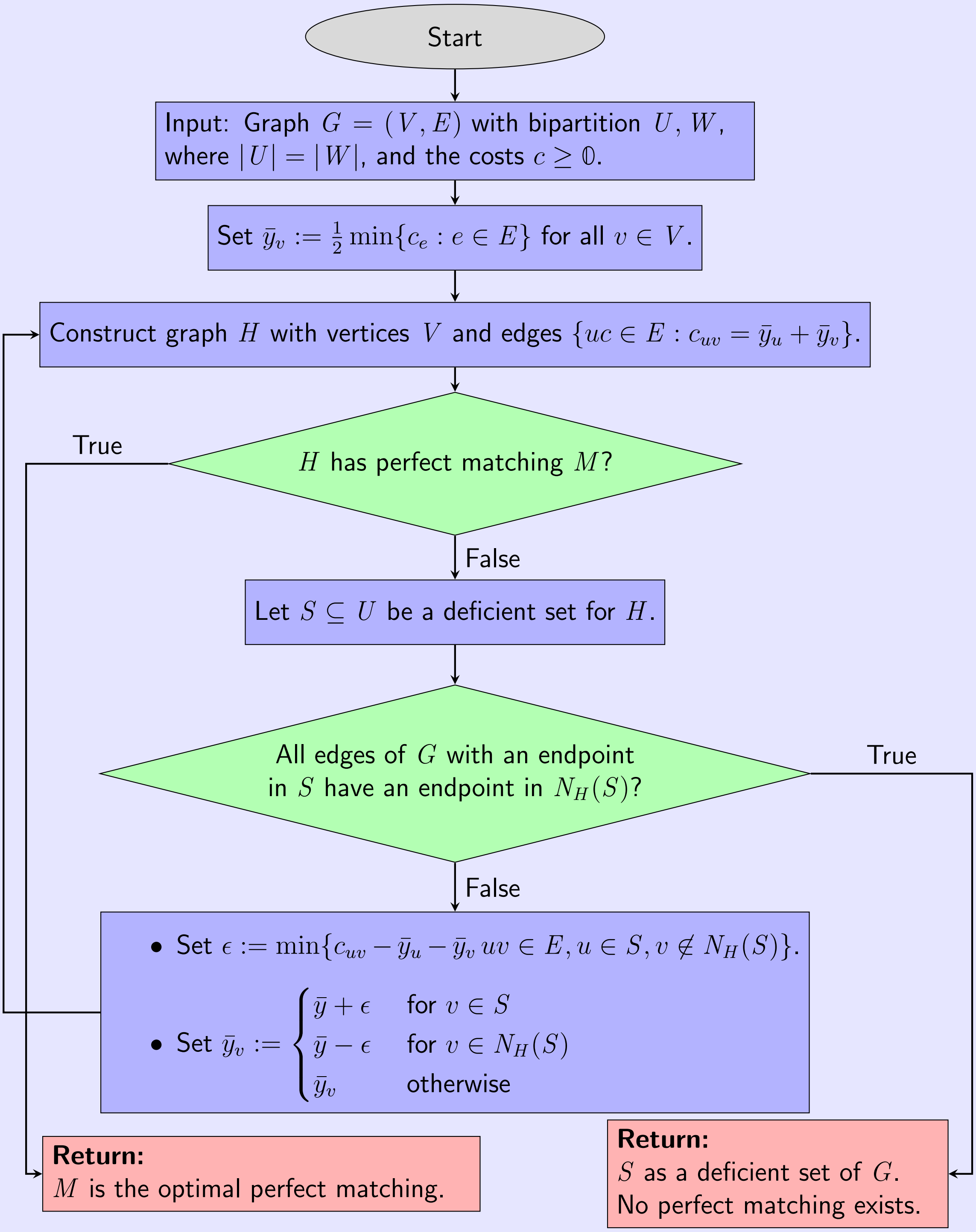
Remark

Input: $G = (V, E)$ with bipartition U, W, where $|U| = |W|$, and the costs $c \geq \mathbb{0}$.
Output: A minimum cost perfect matching M or a deficient set S.

1. Set $\bar{y}_v := \frac{1}{2}\min\{c_e : e \in E\}$ for all $v \in V$.
2. **loop**
   - Construct graph H with vertices V and edges $\{uc \in E : c_{uv} = \bar{y}_u + \bar{y}_v\}$.
   - **if** H has perfect matching M, **then**
     return M as a minimum-cost perfect matching of G.
   - Let $S \subseteq U$ be a deficient set for H.
   - **if** all edges of G with an endpoint in S have an endpoint in $N_H(S)$, **then**
     return S as a deficient set of G, no prefect marching exists.
   - Set $\epsilon := \min\{c_{uv} - \bar{y}_u - \bar{y}_v \; uv \in E, u \in S, v \notin N_H(S)\}$.
   - Set $\bar{y}_v := \begin{cases} \bar{y} + \epsilon & \text{for } v \in S \\ \bar{y} - \epsilon & \text{for } v \in N_H(S) \\ \bar{y}_v & \text{otherwise} \end{cases}$

As before, a procedure and an example are provided: (see next page)

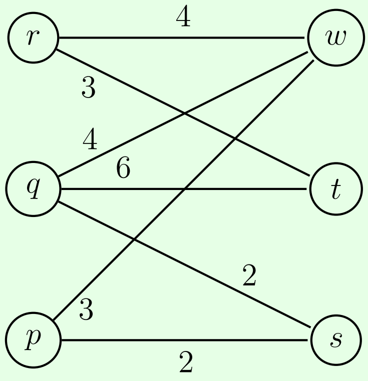
Remark



Here's an example of the application of this algorithm:

Example

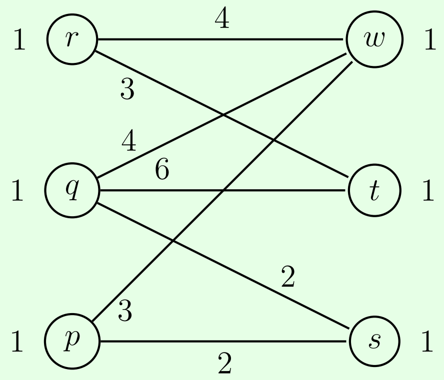
Find the minimum-cost perfect matching on the following graph G, if there's no perfect matching, provide a deficient set of G.



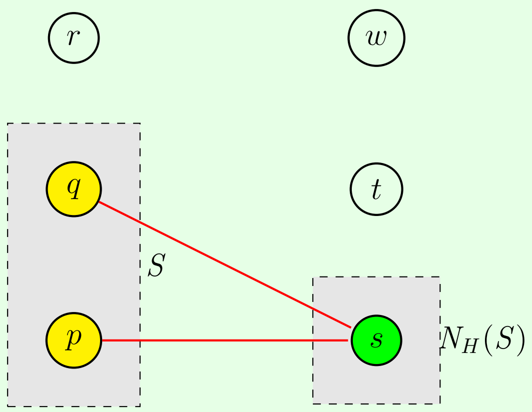
Solution:
Since the minimum cost is 2, and followed by the guide, we initially set potentials:

$$\bar{y}_r = \bar{y}_q = \bar{y}_p = \bar{y}_w = \bar{y}_t = \bar{y}_s = 1$$

Now the minimum cost perfect matching problem can be described as follows:



Then, we indicate the graph H, the deficient set S and $N_H(S)$:

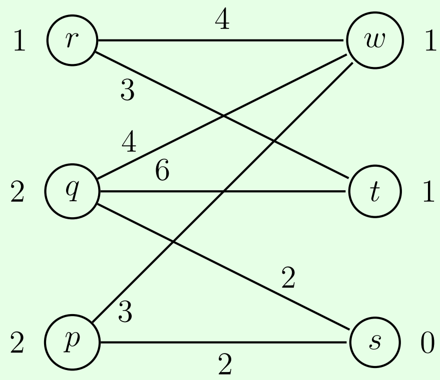


Example

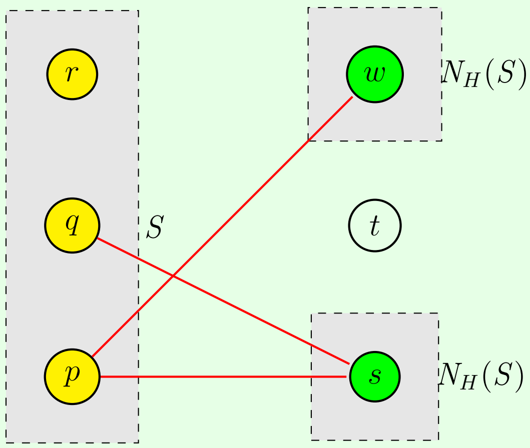
The edges of G with one endpoint in S and the other endpoint not in $N_H(S)$ are qw, qt, pw, hence

$$\begin{aligned}\epsilon &= \min\{c_{qw} - \bar{y}_q - \bar{y}_w, c_{qt} - \bar{y}_q - \bar{t}_w, c_{pw} - \bar{y}_p - \bar{y}_w\} \\ &= \min\{4-1-1, 6-1-1, 3-1-1\} \\ &= 1\end{aligned}$$

Therefore, we can redraw the problem as follows:



Then, we indicate the graph H, the deficient set S and $N_H(S)$:

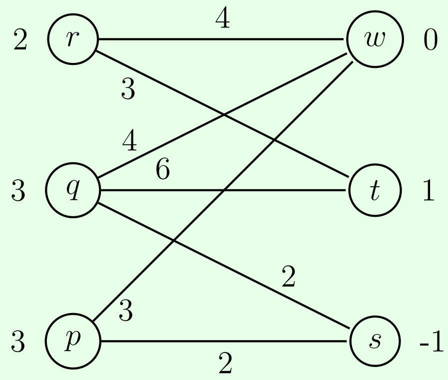


The edges of G with one endpoint in S and the other endpoint not in $N_H(S)$ are rt, qt, hence

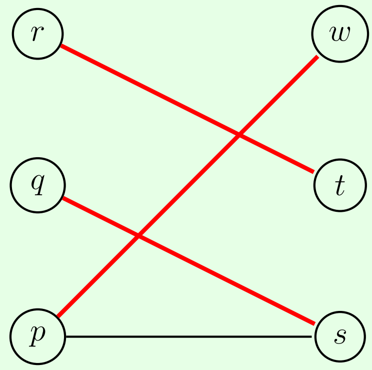
$$\begin{aligned}\epsilon &= \min\{c_{rt} - \bar{y}_r - \bar{y}_t, c_{qt} - \bar{y}_q - \bar{t}_w\} \\ &= \min\{3-1-1, 6-1-1\} \\ &= 1\end{aligned}$$

Example

Therefore, we can redraw the problem as follows:



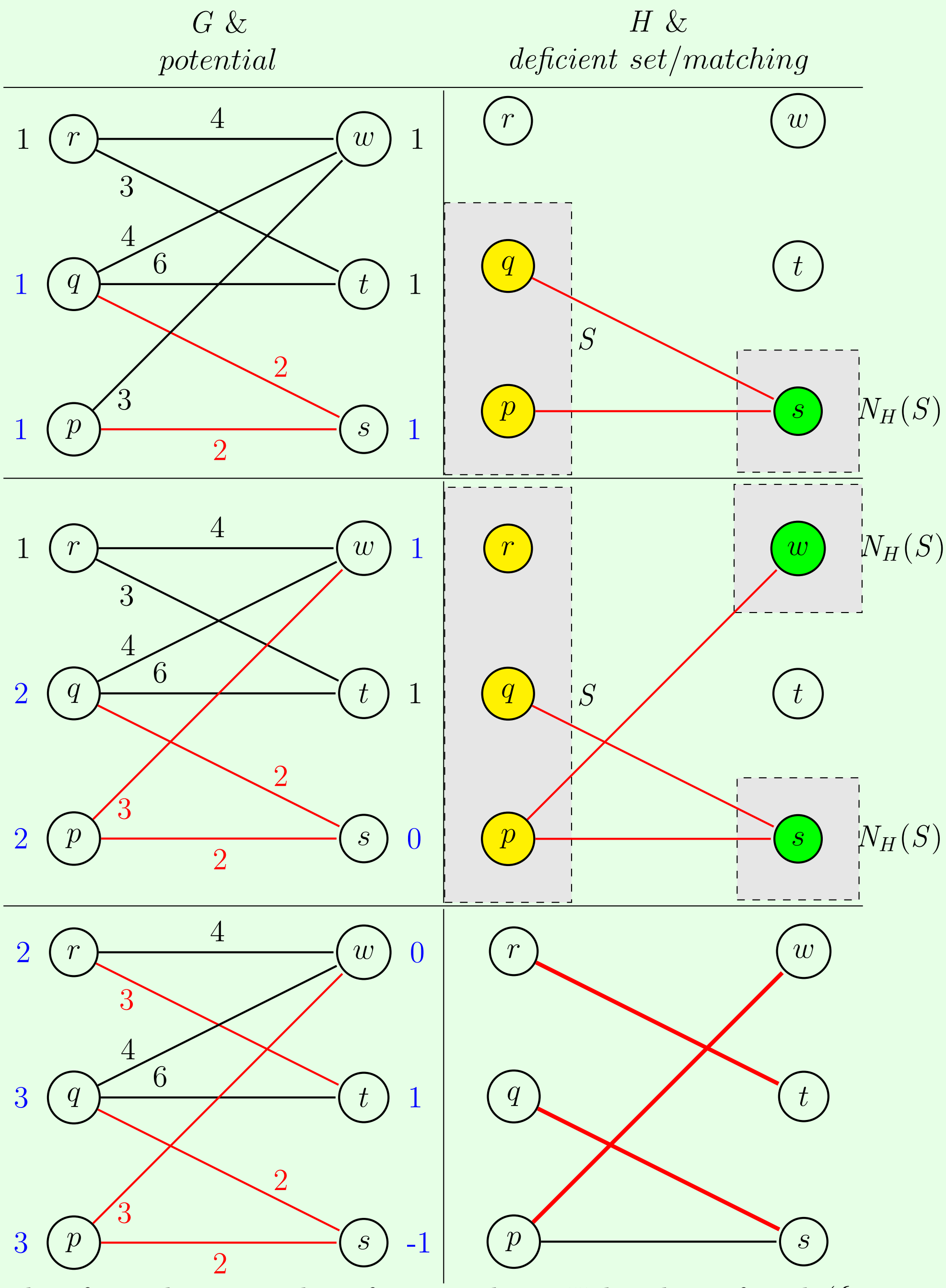
Then, we indicate the graph H, the deficient set S and $N_H(S)$:



This graph has no deficient set, but it has a perfect matching $M = \{rt, qs, pq\}$. This matching is a minimum cost perfect matching of G, with the optimal cost equal to 8, the loop in the Algorithm ends, the Algorithm ends.

As for a more convenient demonstration, we provide a procedure on the next page:

Example



Therefore, the optimal perfect matching H has been found ($\{rt, qs, pw\}$), with optimal value $3 + 2 + 3 = 8$.

3.5 Geometric Optimality

3.5.1 Conplementary Slackness

We know that the feasible region of an LP is a polyhedron, and basic solutions correspond to the extreme points of this polyhedron. **When is an extreme point optimal?**
We can rewrite (P) using **slack variables** s:
(P'):

$$\max\ c^\top x$$
$$s.t.\ Ax + s = b, x, s \geq \mathbb{0}$$

Note that

- (x, s) is feasible for $(P') \to x$ is feasible for (P).
- x is feasible for $(P) \to (x, b - Ax)$ is feasible for (P').

Suppose that $\bar{x}$ is feasible for (P), and $\bar{y}$ is feasible for (D), the dual of (P). Then, $(\bar{x}, \underbrace{b - A\bar{x}}_{\bar{s}})$ is feasible for (P'). Recall that

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

and by **Strong Duality**:

$$\bar{x}, \bar{y} \text{ both optimal} \Leftrightarrow c^\top \bar{x} = \bar{y}^\top b$$
$$\Leftrightarrow \bar{y}^\top \bar{s} = 0$$

By feasibility, $\bar{x}, \bar{s} \geq \mathbb{0}$, hence $\bar{y}^\top \bar{s} = 0$ holds if and only if $\bar{y}_i = 0$ or $\bar{s}_i = 0$ for $1 \leq i \leq m$. Therefore, we can conclude two theorems:

Complementary Slackness Theorem - Special Case

Let $\bar{x}$ and $\bar{y}$ be feasible for $(P) : \max\{c^\top x : Ax \leq b, x\ free\}$ and its dual $(D) : \min\{b^\top y : A^\top y = c, y \geq \mathbb{0}\}$.
Then $\bar{x}$ and $\bar{y}$ are optimal if and only if

- $\bar{y}_i = 0$, or
- the i-th constraint of (P) is **tight** for $\bar{x}$. $(\bar{s}_i = 0)$

Complementary Slackness Theorem

Feasible solution $\bar{x}$ and $\bar{y}$ for (P) and its dual (D) are optimal if and only if:

- $\bar{x}_i = 0$ or the i-th primal constraint is tight for $\bar{y}$ for all row indices i.
- $\bar{y}_i = 0$ or the i-th primal constraint is tight for $\bar{x}$ for all row indices i.

Here's an example:

Example

Consider the LP: (P):

$$\begin{aligned}
&\max\ (5,3,5)x \\
&s.t.\ \begin{pmatrix} 1 & 2 & -1 \\ 3 & 1 & 2 \\ -1 & 1 & 1 \end{pmatrix} x \geq \begin{pmatrix} 2 \\ 4 \\ -1 \end{pmatrix}, x\ free
\end{aligned}$$

Its dual (D) is:

$$\begin{aligned}
&\min\ (2,4,-1)y \\
&s.t.\ \begin{pmatrix} 1 & 3 & -1 \\ 2 & 1 & 1 \\ -1 & 2 & 1 \end{pmatrix} y = \begin{pmatrix} 5 \\ 3 \\ 5 \end{pmatrix}, y \geq \mathbb{0}
\end{aligned}$$

We have $\bar{x} = (1,-1,1)^\top, \bar{y} = (0,2,1)^\top$, and it is easy to check if $\bar{x}$ and $\bar{y}$ are feasible. Now, we check the optimality.

- $\bar{x}_1 = 0$ or $\underbrace{(1,3,-1)\bar{y} = 5}_{\checkmark}$.
- $\bar{x}_2 = 0$ or $\underbrace{(2,1,1)\bar{y} = 3}_{\checkmark}$.
- $\bar{x}_3 = 0$ or $\underbrace{(-1,2,1)\bar{y} = 5}_{\checkmark}$.
- $\underbrace{\bar{y}_1 = 0}_{\checkmark}$ or $(1,2,-1)\bar{x} = 2$.
- $\bar{y}_2 = 0$ or $\underbrace{(3,1,2)\bar{x} = 4}_{\checkmark}$.
- $\bar{y}_3 = 0$ or $\underbrace{(-1,1,1)\bar{x} = -1}_{\checkmark}$.

Since at least one of the conditions is satisfied, we can conclude that $\bar{x}$ and $\bar{y}$ are optimal solutions.

3.5.2 Cones of Vectors

Definition

Let $a_1, a_2, \cdots, a_k$ be vectors in $\mathbb{R}$. Then the **cone** generated by these vectors is given by:

$$C = \{\lambda_1 a_1 + \lambda_2 a_2 + \cdots + \lambda_k a_k : \lambda \geq \mathbb{0}\}$$

Cone of tight constraints is the cone generated by **transpose of rows of tight constraints**.

Example

Consider the constraints: $\begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{pmatrix} x \leq \begin{pmatrix} 2 & 3 & 2 \end{pmatrix}$, the tight constraints at $\bar{x} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ are: $(1, 0)x \leq 2$ and $(1, 1)x \leq 3$.
Then, the cone of tight constraints for $\bar{x}$ is:

$$C(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}) = \left\{\lambda_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix}\right\}$$

Definition

Let $\bar{x}$ be a feasible solution to $\max\{c^\top x : Ax \leq b\}$.
Then $\bar{x}$ is **optimal** if and only if c is in the **cone of tight constraints** for $\bar{x}$.

Example

Consider the LP:

$$\begin{aligned} \max \ & (3/2, 1/2)x \\ s.t. \ & \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{pmatrix} x \leq \begin{pmatrix} 2 \\ 3 \\ 2 \end{pmatrix} \end{aligned}$$

and feasible solution $\bar{x} = (2, 1)^\top$, we've found cone of tight constraints for $\bar{x}$ is $C(\binom{1}{0}, \binom{1}{1}) = \{\lambda_1 \binom{1}{0} + \lambda_2 \binom{1}{1}\}$.
Note that $c = (3/2, 1/2)^\top$ is in the cone of tight constraints as:

$$\begin{pmatrix} 3/2 \\ 1/2 \end{pmatrix} = 1 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 1/2 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Nice, the chapters above are all content for Linear Programs, in the next chapter, we will focus on another type of optimization problem: Integer Program.

Chapter 4

Integer Programs

- **LP**:
  - Can solve very large instances.
  - Algorithms exist that are guaranteed to be fast.
  - Short certificates of optimality (Strong Duality).
  - The only possible outcomes are infeasible, unbounded or optimal.
- **IP**:
  - Some small instances can't be solved.
  - No fast algorithm exists.
  - Doesn't always exist certificate of infeasibility, or optimality.
  - May have an outcome that none of infeasible, unbounded and optimal.

Example

Consider the following IP:

$$\begin{aligned} \max\ & x_1 - \sqrt{2}x_2 \\ s.t.\ & x_1 \leq \sqrt{2}x_2 \\ & x_1, x_2 \geq 1, x_1, x_2 \in \mathbb{Z} \end{aligned}$$

- $x_1 - \sqrt{2}x_2 \leq 0$, so this IP is not unbounded.
- $x_1 = x_2 = 1$ is a feasible solution, so this IP is not infeasible.

Therefore, we wish to prove this IP is not optimal as well. Assume, for a contradiction, $(\bar{x}_1, \bar{x}_2)$ is optimal solution, and let $x_1^{'} = 2\bar{x}_1 + 2\bar{x}_2, x_2^{'} = \bar{x}_1 + 2\bar{x}_2$. $(x_1^{'}, x_2^{'})$ is feasible since: $x_1^{'} = 2\bar{x}_1 + 2\bar{x}_2 \geq 4 > 1, x_2^{'} = \bar{x}_1 + 2\bar{x}_2 \geq 3 > 1$, and:

$$\frac{x_1^{'}}{x_2^{'}} = \frac{2\bar{x}_1 + 2\bar{x}_2}{\bar{x}_1 + 2\bar{x}_2} = 1 + \frac{\bar{x}_1}{\bar{x}_1 + 2\bar{x}_2} = 1 + \frac{1}{1 + \frac{2\bar{x}_2}{\bar{x}_2}} \leq 1 + \frac{1}{1 + \sqrt{2}} = \sqrt{2}$$

But we have $0 \geq (\bar{x}_1 - \sqrt{2}\bar{x}_2) - (x_1^{'} - x_2^{'}) = (\sqrt{2} - 1)(\bar{x}_1 - \sqrt{2}\bar{x}_2) \leq 0$, this means $\bar{x}_1 - \sqrt{2}\bar{x}_2 = 0$, which implies at least one of $\bar{x}_1, \bar{x}_2$ are not integer, this leads to a contradiction. Therefore, this IP is not optimal.

4.1 Convex Hulls

Remark

There will **NOT** a practical procedure to solve IPs, but it will suggest a strategy.

Definition

Let C be a subset of $\mathbb{R}^n$, the **convex hull** of C is the **smallest convex set** that constraints C.

Remark

Given $C \subset \mathbb{R}^n$, there is a unique smallest convex set containing C

Theorem

Consider $P = \{x : Ax \leq b\}$ where A, b are rational. Then, the convex hull of all integer points in P is a polyhedron.

Let A, b be rational, and the (IP) is:

$$\max\{c^\top x : Ax \leq b, x \in \mathbb{Z}\}$$

The convex hull of all feasible solutions of (IP) is a polyhedron: $\{x : A^{'}x \leq b^{'}\}$.
Then the LP with convex hull as constraints is:

$$\max\{c^\top x : A^{'}x \leq b^{'}\}$$

Remark

- (IP) is infeasible if and only if (LP) is infeasible.
- (IP) is unbounded if and only if (LP) is unbounded.
- An optimal solution to (IP) is an optimal solution to (LP).
- An **extreme** optimal solution to (LP) is an optimal solution to (IP).

So the "way" of solving an IP is:

- Compute $A^{'}, b^{'}$.
- Use Simpex to find an extreme optimal solution to LP.

Easier said than done, it is very hard to find such $A^{'}, b^{'}$, and they might be **much more** complicated than A, b.
In the next section, we introduce a method called **Cutting Planes** to get closer to $A^{'}, b^{'}$.

4.2 Cutting Planes

Definition

Suppose a constraint $\alpha^\top x \leq \beta$ that

- is satisfied for all feasible solutions to the IP, and
- is not satisfied for $\bar{x}$

We then call this constraint a **cutting plane** for $\bar{x}$.

- $\lfloor x \rfloor$ means **floor** x, which gives the greatest integer less than or equal to x.
- $\lceil x \rceil$ means **ceil** x, which gives the smallest integer greater than or equal to y.

Example

- $\lfloor 5.3 \rfloor = 5, \lfloor 3 \rfloor = 3, \lfloor 0 \rfloor = 0, \lfloor -1.5 \rfloor = -2, \lfloor -3 \rfloor = -3.$
- $\lceil 5.3 \rceil = 6, \lceil 5 \rceil = 5, \lceil 0 \rceil = 0, \lceil -1.2 \rceil = -1, \lceil -3 \rceil = -3.$

Remark

Do **Not** think floor/ceil is round to 0!

Example

Consider the IP:

$$\begin{aligned} \max\ & (2,5)x \\ s.t.\ & \begin{pmatrix} 1 & 4 \\ 1 & 1 \end{pmatrix} x \leq \begin{pmatrix} 8 \\ 4 \end{pmatrix}, x \geq \mathbb{0}, x \in \mathbb{Z}. \end{aligned}$$

Run Simplex on the IP without integrality constraint, we can find that $\bar{x} = (8/3, 4/3)$ is optimal, but note that they are not integers.
A cutting plane for $\bar{x}$ is $x_1 + 3x_2 \leq 6$. (We will learn how to find the cutting plan later). Then, adding this into our IP:

$$\begin{aligned} \max\ & (2,5)x \\ s.t.\ & \begin{pmatrix} 1 & 4 \\ 1 & 1 \\ 1 & 3 \end{pmatrix} x \leq \begin{pmatrix} 8 \\ 4 \\ 6 \end{pmatrix}, x \geq \mathbb{0}, x \in \mathbb{Z}. \end{aligned}$$

Run Simplex on the new IP without integrality constraints, we can find that $x' = (3,1)^\top$ is optimal, and since it is integral, x' is also the optimal solution for IP.

We will use Simplex to help find the cutting plane:

- Solve the relaxation LP (by dropping the integrality constraint) and get the LP into the canonical form for basis B:

$$\begin{aligned} &\max\ \bar{c}^\top x + \bar{z} \\ &s.t.\ x_B + A_N x_N = b, x \geq \mathbb{0} \end{aligned}$$

  where $N = \{j : j \notin B\}, \bar{x}_N = \mathbb{0}, \bar{x}_B = b$.

- If $\bar{x} \in \mathbb{Z}$, then $\bar{x}$ is optimal for the original IP, algorithm end.
- If $\bar{x} \notin \mathbb{Z}$, denote the i^{th} basic variable $r(i)$, and say $\bar{x}_i$ be fractional, then b_i is fractional too. We know that every feasible solution to the IP satisfies:

$$\begin{aligned} b_i &= x_{r(i)} + \sum_{j \in N} A_{ij} x_j \\ \Rightarrow \underbrace{b_i}_{\notin \mathbb{Z}} &\geq \underbrace{x_{r(i)} + \sum_{j \in N} \lfloor A_{ij} \rfloor x_j}_{\in \mathbb{Z}} \\ \Rightarrow \lfloor b_i \rfloor &\geq x_{r(i)} + \sum_{j \in N} \lfloor A_{ij} \rfloor x_j \qquad (*) \end{aligned}$$

- $(*)$ holds for all $x \in \mathbb{Z}$, but invalid for $\bar{x}$, since:

$$\underbrace{\bar{x}_{r(i)}}_{b_i} + \sum_{j \in N} \lfloor A_{ij} \rfloor \underbrace{\bar{x}_j}_{0} = \underbrace{b_i}_{fractional} > \lfloor b_i \rfloor$$

  Therefore, by definition, $(*)$ is the cutting plane for $\bar{x}$.

  Remark

  We floored down to get the cutting plane above, but we can also ceil up!

$$\begin{aligned} b_i &= x_{r(i)} + \sum_{j \in N} A_{ij} x_j \\ \Rightarrow \underbrace{b_i}_{\notin \mathbb{Z}} &\leq \underbrace{x_{r(i)} + \sum_{j \in N} \lceil A_{ij} \rceil x_j}_{\in \mathbb{Z}} \\ \Rightarrow \lceil b_i \rceil &\leq x_{r(i)} + \sum_{j \in N} \lceil A_{ij} \rceil x_j \qquad (**) \end{aligned}$$

  $(**)$ holds for all $x \in \mathbb{Z}$, but invalid for $\bar{x}$, since:

$$\underbrace{\bar{x}_{r(i)}}_{b_i} + \sum_{j \in N} \lceil A_{ij} \rceil \underbrace{\bar{x}_j}_{0} = \underbrace{b_i}_{fractional} < \lceil b_i \rceil$$

  Therefore, by definition, $(**)$ is also a cutting plane for $\bar{x}$.

Example

Consider the IP:

$$\begin{aligned} &\max \ (2,5)x \\ &s.t. \ \begin{pmatrix} 1 & 4 \\ 1 & 1 \end{pmatrix} x \le \begin{pmatrix} 8 \\ 4 \end{pmatrix}, x \ge \mathbb{0}, x \in \mathbb{Z}. \end{aligned}$$

Transform it into SEF and drop the integrality constraints:

$$\begin{aligned} &\max \ (2,5,0,0)x \\ &s.t. \ \begin{pmatrix} 1 & 4 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{pmatrix} x = \begin{pmatrix} 8 \\ 4 \end{pmatrix}, x \ge \mathbb{0}. \end{aligned}$$

Note here we have $x_1 + 4x_2 + x_3 = 8, x_1 + x_2 + x_4 = 4$. Run Simplex on the LP gives:

$$\begin{aligned} &\max \ (0,0,-1,-1)x + 12 \\ &s.t. \ \begin{pmatrix} 1 & 0 & -1/3 & 4/3 \\ 0 & 1 & 1/3 & -1/3 \end{pmatrix} x = \begin{pmatrix} 8/3 \\ 4/3 \end{pmatrix}, x \ge \mathbb{0} \end{aligned}$$

with optimal solution $(0,0,8/3,4/3)$, which leads to the solution $(8/3,4/3)$ of our original IP, which is not valid, so we have to find a cutting plane:

- Pick the first constraint (we can also pick the second one), and use the floor-up method, it gives:

  $$\begin{aligned} &x_1 - 1/3x_3 + 4/3x_4 = 8/3 \\ \Rightarrow &x_1 + \lfloor -1/3 \rfloor x_3 + \lfloor 4/3 \rfloor x_4 \le \lfloor 8/3 \rfloor \\ \Rightarrow &x_1 - x_3 + x_4 \le 2 \end{aligned}$$

  Substitute $x3, x_4$ in terms of x_1, x_2 gives:

  $$x_1 + 3x_2 \le 6$$

- Use the ceil-up method on the first constraint gives:

-

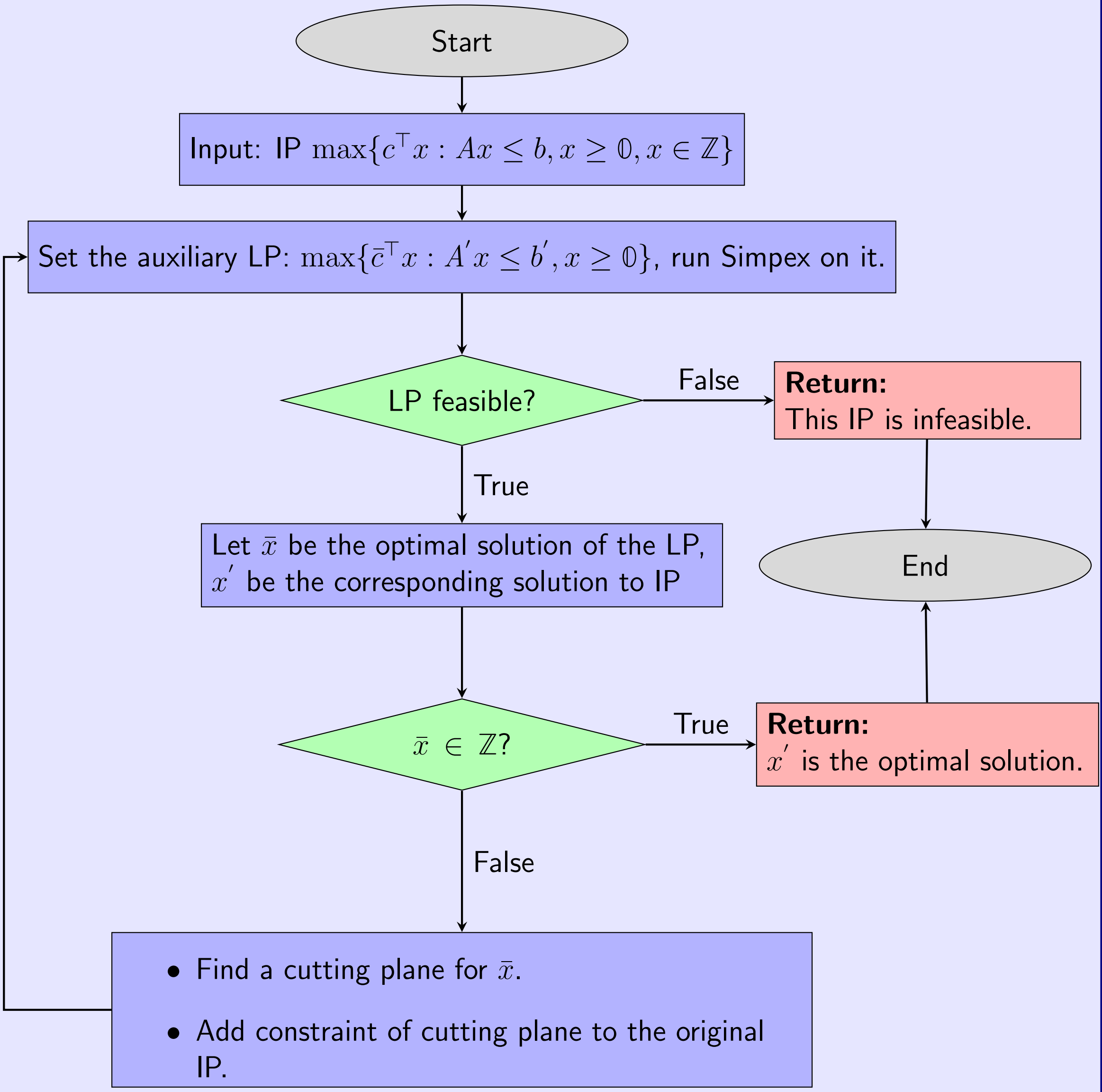
  $$\begin{aligned} &x_1 - 1/3x_3 + 4/3x_4 = 8/3 \\ \Rightarrow &x_1 + \lceil -1/3 \rceil x_3 + \lceil 4/3 \rceil x_4 \ge \lceil 8/3 \rceil \\ \Rightarrow &x_1 + 2x_4 \ge 3 \end{aligned}$$

  Substitute x_4 in terms of x_1, x_2 gives:

  $$x_1 + 2x_2 \le 5$$

Our procedure for using the cutting plane to get the optimal solution of an IP is:

Remark



Remark

In normal cases, A and b, c have only rational entries ($\in \mathbb{R}$), we have:

- LP is unbounded & IP is feasible $\rightarrow$ IP is unbounded.
- LP has an optimal solution & IP is feasible $\rightarrow$ IP has an optimal solution.

So even if IP is unbounded, we can still use this procedure, and then we can obtain a certificate of unboundedness, $\bar{x}, d$. If $\bar{x}$ is integral, we can then conclude that the IP is unbounded, if $\bar{x}$ is not integral, we continue finding more cutting planes.

That's the content of the Integer Program! In the next and last chapter, we will learn some interesting properties of Nonlinear Programs.

Chapter 5

Nonlinear Programs

5.1 Convexity

Definition

A **nonlinear program (NLP)** is a problem of the form:

$$\underbrace{\min}_{or\ \max} \{f(x) : g_i(x) \leq 0, i = 1, 2, \cdots, k\}$$

where $f : \mathbb{R}^n \to \mathbb{R}, g_i : \mathbb{R}^n \to \mathbb{R}, i = 1, 2, \cdots, k$.

In this section, we will only focus on the case $f(x)$ is linear.

Definition

Consider, the NLP:

$$\min\{f(x) : x \in S\}$$

$x \in S$ is a **local optimum** if there exists $\delta > 0$ such that

$$\forall x^{'} \in \mathbb{S}, -\delta \leq x^{'} - x \leq \delta$$

and we have $f(x) \leq f(x^{'})$.

Remark

Consider : $\min\{c^\top x : x \in S\}$. If S is a **convex** and x is a **local optimum**, then x is optimal.

Proof. Suppose $\exists x^{'} \in S$ with $c^\top x^{'} \leq c^\top x$, let $y = \lambda x^{'} + (1-\lambda)x$ for $\lambda > 0$ small. Since S is a convex, $y \in S$, as λ small $-\delta \leq y - x \leq \delta$. Then, we have:

$$\begin{aligned} c^\top y &= c^\top(\lambda x^{'} + (1-\lambda)x) \\ &= \underbrace{\lambda}_{\geq 0} \underbrace{c^\top x^{'}}_{< c^\top x} + \underbrace{(1-\lambda)}_{\geq 0} x \\ &< \lambda c^\top x + (1-\lambda) c^\top x = c^\top x \end{aligned}$$

This leads to a contradiction. $\square$

Definition

Function $f : \mathbb{R}^n \to \mathbb{R}$ is **convex** if for all $a, b \in \mathbb{R}^n$:

$$f(\lambda a + (1-\lambda)b) \leq \lambda f(a) + (1-\lambda)f(b)$$

for all $0 \leq \lambda \leq 1$.

Example

Prove $f(x) = x^2$ is convex.
Solution: Pick $a, b \in \mathbb{R}$ and pick λ such that $0 \leq \lambda \leq 1$, then

$$\begin{aligned}
&f(\lambda a + (1-\lambda)b) - \lambda f(a) - (1-\lambda)f(b) \\
=&[\lambda a + (1-\lambda)b]^2 - \lambda a^2 - (1-\lambda)b^2 \\
=&2(1-\lambda)2b - \lambda(1-\lambda)(a^2+b^2) \\
=&\underbrace{-\lambda}_{\leq 0}\underbrace{(1-\lambda)}_{\geq 0}\underbrace{(a-b)^2}_{\geq 0} \leq 0.
\end{aligned}$$

Remark

Let $g : \mathbb{R}^n \to \mathbb{R}$ be a convex function and $\beta \in \mathbb{R}$, it follows that $S = \{x \in \mathbb{R}^n : g(x) \leq \beta\}$ is a **convex set**.

Proof. Pick $a, b \in \mathbb{S}$, and λ where $0 \leq \lambda \leq 1$. Let $x = \lambda a + (1-\lambda)b$, our goal is to show that if $x \in S$, then $g(x) \leq \beta$.

$$\begin{aligned}
g(x) &= g(\lambda a + (1-\lambda b)) \\
&\leq \underbrace{\lambda}_{\geq 0}\underbrace{g(a)}_{\leq \beta} + \underbrace{(1-\lambda)}_{\geq 0}\underbrace{g(b)}_{\leq \beta} \\
&\leq \lambda\beta + (1-\lambda)\beta \\
&= \beta
\end{aligned}$$

□

Remark

Suppose NLP:

$$\min\{c^\top x : g_i(x) \leq 0, i = 1, 2, \cdots, k\}$$

If all functions g_i are convex, then the feasible region of NLP is convex.

Proof. Let $S_i = \{x : g_i(x) \leq 0\}$, by previous result, S_i is convex. The feasible region of the NLP is $\bigcap_{i=1}^{k} S_i$. Since the intersection of convex sets is still convex, the result follows. □

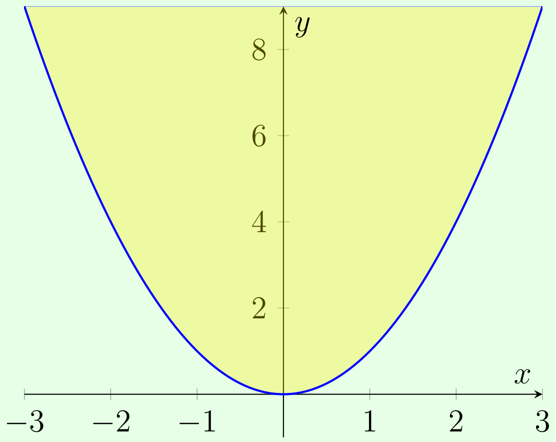
Definition

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function. The **epigraph** of f is given by:

$$epi(f) = \left\{ \begin{pmatrix} y \\ x \end{pmatrix} : y \geq f(x), x \in \mathbb{R}^n \right\}$$

Example

Consider $f(x) = x^2$, then the draw of $epi(f)$ is:



Remark

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function, it follows that:

$$f \text{ is convex} \Leftrightarrow epi(f) \text{ is convex.}$$

5.2 The KKT Theorem

How can we prove a feasible solution $\bar{x}$ is optimal for an NLP?
We can still use the method to handle Integer Program, that is:

- Find a relaxation of the NLP.
- Prove $\bar{x}$ is optimal for the relaxation.
- Deduce that $\bar{x}$ is optimal for the NLP.

In this section, we will go a bit further on this method.

Example

Prove the $\bar{x} = (1,1)^\top$ is an optimal solution to:

$$\begin{aligned}
\min \quad & -x_1 - x_2 \\
s.t. \quad & -x_2 + x_1^2 \leq 0 \\
& -x_1 + x_2^2 \leq 0 \\
& -x_2 + 1/2 \leq 0
\end{aligned}$$

We can find a relaxation LP (We will show why this is the relaxation later)

$$\begin{aligned}
\min \quad & -x_1 - x_2 \\
s.t. \quad & 2x_1 - x_2 \leq 1 \\
& -x_1 + 2x_2 \leq 1
\end{aligned}$$

$\bar{x}$ is an optimal solution for the relaxation, therefore, $\bar{x}$ is also an optimal solution for the original NLP.

Definition

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function and $\bar{x} \in \mathbb{R}^n$, then, $s \in \mathbb{R}^n$ is a **subgradient** of f at $\bar{x}$ if

$$h(x) := f(\bar{x}) + s^\top (x - \bar{x}) \leq f(x), \forall x \in \mathbb{R}^n$$

Example

In our previous example, we have $f(x) = -x_1 + x_2^2$ and $\bar{x} = (1,1)^\top$, we claim that $(-1,2)^\top$ is a subgradient of f at $\bar{x}$:

$$h(x) := f(\bar{x}) + s^\top (x - \bar{x}) = -x_1 + 2x_2 - 1$$

and check $h(x) \leq f(x)$:

$$h(x) - f(x) = -x_1 + 2x_2 - 1 + x_1 - x_2^2 = -(x_2 - 1)^2 \leq 0$$

Now we have a new question, how to get the subgradient s? We will cover that soon.

Definition

Let $C \in \mathbb{R}^n$ be a convex set and let $\bar{x} \in C$. The halfspace $F = \{x : s^\top x \leq \beta\}$ is **supporting** C at $\bar{x}$ if:

- $C \subseteq F$ and
- $s^\top \bar{x} = \beta$. That is, $\bar{x}$ is on the boundary of F.

Remark

Let $g : \mathbb{R}^n \to \mathbb{R}$ be a convex and let $\bar{x}$ where $g(\bar{x}) = 0$. Let s be a subgradient of g at $\bar{x}$, let $C = \{x : g(x) \leq 0\}, F = \{x : h(x) := g(\bar{x}) + s^\top(x - \bar{x}) \leq 0\}$. Then F is a supporting halfspace of C at $\bar{x}$.

- C is convex, as g is a convex function.
- F is a halfspace, as $h(x)$ is an affine function.
- $h(\bar{x}) = g(\bar{x}) = 0$, thus, $\bar{x}$ is on the boundary of F.

Example

Continue our NLP example: Let $g(x) = x_2^2 - x_1, \bar{x} = (1,1)^\top$, and $s = (-1,2)^\top$ as a subgradient at $\bar{x}$, we have:

$$h(x) = 0 + (-1,2)\left[\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \end{pmatrix}\right] = -x_1 + 2x_2 - 1$$

and

$$F = \{x : -x_1 + 2x_2 \leq 1\}$$

We can use this to construct relaxations of NLPs. Given constraint $g_i(x) \leq 0$. If we replace the nonlinear constraint with the linear constraint $h(x) := g_i(\bar{x}) + s^\top(x - \bar{x})$, we then get a relaxation.

KKT Theorem - Subgradient

Let I denote the set of indices i for which $g_i(\bar{x}) = 0$, consider the NLP:

$$\begin{aligned} &\min\ c^\top x \\ &s.t.\ g_i(x) \leq 0, i = 1, 2, \cdots, k \end{aligned}$$

where

- $g_1, g_2, \cdots, g_k$ are all convex.
- $\bar{x}$ is a feasible solution.
- $\forall i \in I, g_i(\bar{x}) = 0$.
- $\forall i \in I, s^{(i)}$ is a subgradient for g_i at $\bar{x}$.

If $(-c) \in cone\{s^{(i)} : i \in I\}$, then $\bar{x}$ is **optimal**.

Example

In our previous NLP example, we have:

$$-\begin{pmatrix} -1 \\ -1 \end{pmatrix} \in cone\left\{\begin{pmatrix} 2 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \end{pmatrix}\right\} \Rightarrow \bar{x} \text{ optimal}$$

Remark

- Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function and let $\bar{x} \in \mathbb{R}^n$. If the **gradient** $\nabla f(\bar{x})$ of f exists at $\bar{x}$, then it is a subgradient.
- If the partial derivative $\frac{\partial f(x)}{\partial x_j}$ exists for f at $\bar{x}$ for all $j = 1, 2, \cdots, n$, then the gradient $\nabla f(\bar{x})$ is obtained by evaluating for $\bar{x}$,

$$\left(\frac{\partial f(x)}{\partial x_1}, \frac{\partial f(x)}{\partial x_2} \cdots, \frac{\partial f(x)}{\partial x_n}\right)^\top$$

Example

In our previous NLP example, we have:

$$\left(\frac{\partial f(x)}{\partial x_1}, \frac{\partial f(x)}{\partial x_2}\right)^\top = (2x_1, -1)^\top$$

For $\bar{x}$, we get $\nabla f(\bar{x}) = (2, -1)^\top$. Since $(2, -1)^\top$ is the gradient of f at $\bar{x}$, it is a subgradient as well.

Definition

A feasible solution to $\bar{x}$ is a **Slater point** of

$$\begin{aligned} &\min\ c^\top x \\ &\ s.t.\ g_i(x) \leq 0, i = 1, 2, \cdots, k \end{aligned}$$

if $g_i(\bar{x}) < 0$ for all $i = 1, 2, \cdots, k$.

Example

In our previous NLP example, $\bar{x} = (3/4, 3/4)^\top$ is a slater point.

KKT Theorem

Let I denote the set of indices i for which $g_i(\bar{x}) = 0$, consider the following NLP:

$$\begin{aligned} &\min\ c^\top x \\ &\ s.t.\ g_i(x) \leq 0, i = 1, 2, \cdots, k \end{aligned}$$

Suppose that $\bar{x}$ is a feasible solution, and

- $g_1, g_2, \cdots, g_k$ are all convex.
- There exists a slater point.
- $\forall i \in I$, there exists a gradient $\nabla g_i(\bar{x})$ of g_i at $\bar{x}$.

Then, $\bar{x}$ is optimal $\Leftrightarrow (-c) \in cone\{\nabla g_i(\bar{x}) : i \in I\}$

Now, looking back to our example, here is the whole solution:

Example

Prove the $\bar{x} = (1,1)^\top$ is an optimal solution to:

$$\begin{aligned} \min \quad & -x_1 - x_2 \\ s.t. \quad & -x_2 + x_1^2 \le 0 \\ & -x_1 + x_2^2 \le 0 \\ & -x_2 + 1/2 \le 0 \end{aligned}$$

Solution: Let $g_1(x) = -x_2 + x_1^2, g_2(x) = -x_1 + x_2^2, g_3(x) = -x_1 + 1/2$. Then, $g_1(\bar{x}) = -1 + 1 = 0, g_2(\bar{x}) = -1 + 1 = 0, g_3 = -1 + 1/2 < 0, I = \{1,2\}$. Since we have:

- g_1, g_2, g_3 convex.
- There exists a slater point $x^{'} = (3/4, 3/4)^\top$.
- $\forall i \in I$, there exists a gradient $\nabla g_i(\bar{x})$ of g_i at $\bar{x}$.
  - When $i = 1$:

    $$\begin{aligned} \frac{\partial g_1(x)}{\partial x_1} &= \frac{\partial(-x_2 + x_1^2)}{\partial x_1} = 2x_1 \\ \frac{\partial g_1(x)}{\partial x_2} &= \frac{\partial(-x_2 + x_1^2)}{\partial x_2} = -1 \end{aligned}$$

    Here, embed $\bar{x}$, we have $2\bar{x}_1 = 2$ exist, and $\nabla g_1(\bar{x}) = \binom{2}{-1}$.
  - When $i = 2$:

    $$\begin{aligned} \frac{\partial g_2(x)}{\partial x_1} &= \frac{\partial(-x_1 + x_2^2)}{\partial x_1} = -1 \\ \frac{\partial g_2(x)}{\partial x_2} &= \frac{\partial(-x_1 + x_2^2)}{\partial x_2} = 2x_2 \end{aligned}$$

    Here, embed $\bar{x}$, we have $2\bar{x}_2 = 2$ exist, and $\nabla g_2(\bar{x}) = \binom{-1}{2}$.
- $cone\{\nabla g_1(\bar{x}), \nabla g_2(\bar{x})\} = cone\left\{\binom{2}{-1}, \binom{-1}{2}\right\}$
- $-c = \binom{1}{1} = \underbrace{1}_{\lambda_1 \ge 0} \times \binom{2}{-1} + \underbrace{1}_{\lambda_2 \ge 0} \times \binom{-1}{2}$.

Therefore, by KKT Theorem, we can conclude $\bar{x} = (1,1)^\top$ is optimal of the NLP.

On the next page, we will extend the KKT Theorem, which supports us to solve the case when the objective function is not linear.

Recall the conditions of using KKT:

Remark

Let I denote the set of indices i for which $g_i(\bar{x}) = 0$, consider the following NLP:

$$\begin{aligned} &\min\ c^\top x \\ &s.t.\ g_i(x) \le 0, i = 1, 2, \cdots, k \end{aligned}$$

Suppose that $\bar{x}$ is a feasible solution, and

- $g_1, g_2, \cdots, g_k$ are all convex.
- There exists a slater point.
- $\forall i \in I$, there exists a gradient $\nabla g_i(\bar{x})$ of g_i at $\bar{x}$.

Then, $\bar{x}$ is optimal $\Leftrightarrow$ $(-c) \in cone\{\nabla g_i(\bar{x}) : i \in I\}$. We can extend this to:

Definition

$\bar{x}$ is optimal $\Leftrightarrow$ $(-\nabla f(\bar{x})) \in cone\{\nabla g_i(\bar{x}) : i \in I\}$, where f is the objective function.

Remark

For the case f is linear, we indeed have $(-\nabla f(x)) = (-c)$. (easy to prove)

Finally, we finished all the content of C&O250!, we took a broad introduction to the field of optimization, and we discussed the applications like shortest path problem, perfect matching problem etc. We also learned some solution techniques, mathematical models for real-life applications, algorithms, aspects of computational complexity, geometry, and linear duality. I hope you can gain something and wish you all the best in your future study journey. Thank you all for your reading!

References

- Catherine Zhou (2022), "C&O250, introduction to Optimization". [Online] https://student.cs.uwaterloo.ca/~z374zhou/co250-note.pdf. [Accessed: 18/07/2024]
- J. KÖNEMANN, J. KÖNEMANN, L. TUNÇEL (2014), "A Gentle Introduction to Optimization". Cambridge University Press. [Physical book].
- © Faculty of Mathematics, University of Waterloo (2022), "MATH 136: Linear Algebra 1 for Honours Mathematics". [Physical book].