C&O 250

Introduction to Optimization

Catherine Zhou

April 19, 2022

Contents

Preface

Module 1

Formulations

1.1 Overview

What is optimization? Abstractly, we will focus on **abstract optimization problem (P)**:

Given a set $A \subseteq \mathbb{R}^n$ and a function $f : A \to \mathbb{R}$

Goal find $x \in A$ that minimizes/maximizes f

(the above problem is very hard to solve and may not even be well-defined).

There are three special cases of P in this course:

- Linear Programming (LP): A is implicitly given by *linear* constraints, and f is a *linear* function.
- Integer Programming (IP): we want the max or min over the *integer* points in A.
- Nonlinear Programming (NLP): A is given by *non-linear* constraints, and f is a *non-linear* function.

1.1.1 Typical Workflow

Practical problem : text description of practical problem

Mathematical model : we will develop this model for the problem, capturing problem in mathematics. LP, IP, NLP appears here.

Practical implementation : we feed the model and data into a solver

This process interates.

Example 1.1.1. *WaterTech*

WaterTech produces 4 products, $\mathcal{P} = \{1, 2, 3, 4\}$:

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

Some restrictions:

- *WaterTech has 700h on machine 1 and 500h on machine 2 available.*
- *It can purchase 600h of skilled labour at \$8 per hour and at most 650h of unskilled labour at \$6 per hour.*

Question: How much of each product should WaterTech produce in order to maximize profit? We can formulate this as a mathematical program!

1.1.2 Ingredients of a Math Model

Decision Variables : capture unknown information

Constraints : describe which assignments to variables are feasible

Objective function : a function of the variables that we would like to maximize/minimize

WaterTech Model

Variables : It needs to decide how many units of each product to produce $\implies$ introduce x_i for number of units of product i to produce. For convenience, we also have y_s, y_u: number of hours of skilled/unskilled labour to purchase.

Constraints : What makes an assignment to $\{x_i\}_{i\in\mathcal{P}}, y_s, y_u$ a feasible? - Restricted available time on machine 1 and machine 2,

$$11x_1 + 7x_2 + 6x_3 + 5x_4 \leq 700$$

$$4x_1 + 6x_2 + 5x_3 + 4x_4 \leq 500$$

and the amount of time skilled or unskilled labour can work.

$$8x_1 + 5x_2 + 5x_3 + 6x_4 \leq y_s$$

$$7x_1 + 8x_2 + 7x_3 + 4x_4 \leq y_u$$

and $y_s \leq 600$, $y_u \leq 650$.

Objective Function : the revenue from sales:

$$300x_1 + 260x_2 + 220x_3 + 180x_4$$

the cost of labour:

$$8y_s + 6y_u$$

So we want to maximize the objective function:

$$300x_1 + 260x_2 + 220x_3 + 180x_4 - (8y_s + 6y_u)$$

Solution. To find $\max\{300x_1 + 260x_2 + 220x_3 + 180x_4 - (8y_s + 6y_u)\}$ with the above constraints, and using CPLEX, we find

$$\begin{aligned} x &= (16 + \frac{2}{3}, 50, 0, 33 + \frac{1}{3})^T \\ y_s &= 583 + \frac{1}{3} \\ y_u &= 650 \end{aligned}$$

with profit $\$15433 + \frac{1}{3}$. ■

1.1.3 Correctness of Model

How do we know if our model is correct or not? We have the word description of problem and the formulation, and we find a solution to the formulation, which is an assignment to all of its variables. This is **feasible** if all the constraints are satisfied, and **optimal** if no better feasible solution exists.

Note: a solution to the *word description* is an assignment to the unkonwns.

One way of showing correctness: define a **mapping** between feasible solutions to the word description, and feasible solutions to the model, and vise versa.

Feasible solution to WaterTech problem

The solution to the word description is given by

1. Producing 10 units of product 1, 50 units of product 2, 0 units of product 3, and 20 units of product 2, and by
2. purchasing 600 units of skilled and unskilled labour

which is equivalent to

$$x_1 = 10, x_2 = 50, x_3 = 0, x_4 = 20, y_s = y_u = 600$$

and feasible for the mathematical program we wrote.

Your mapping should **preserve cost**. In the above example, the profit from the solution to the word description should correspond to the objective value of its image (under map), and vice versa. **You need to check this!**

1.2 LP Models

In this course, we consider optimization problems of this form:

$$\min\{f(x) : g_i(x) \leq b_i, (1 \leq i \leq m), x \in \mathbb{R}^n\}$$

where

- $n, m \in \mathbb{N}$
- $b_1, \dots, b_m \in \mathbb{R}$
- $f, g_1, \dots, g_n$ are functions from $\mathbb{R}^n$ to $\mathbb{R}$

Problems like the above are very hard to solve in general, so we only focus on the special case - **all functions are affine**.

1.2.1 Modelling: Linear Problems

Definition 1: Affine function

A function $f : \mathbb{R}^n \to \mathbb{R}$ is affine if $f(x) = a^T x + \beta$ for $a \in \mathbb{R}^n$, $\beta \in \mathbb{R}$. It is linear if, in addition, $\beta = 0$.

Example of Affine functions

1. $f(x) = 2x_1 + 3x_2 - x_3 + 7$ is affine, but not linear
2. $f(x) = -3x_1 + 5x_3$ is linear
3. $f(x) = 5x - 3\cos(x) + \sqrt{x}$ is not affine nor linear.

Definition 2: Linear program

The optimization problem

$$\min\{f(x) : g_i(x) \leq b_i,\ \forall\ 1 \leq i \leq m, x \in \mathbb{R}^n\}$$

is called a linear program if f is affine, and $g_1, \dots, g_m$ is **finite** number of linear functions.

Notes:

- Instead of set notation, we often write LPs more verbosely
- Often give non-negativity constraints separately
- May use $\max$ instead of $\min$
- Sometimes replace subject to by s.t.
- We often write $x \geq \mathbb{0}$ as a short form for all variabels are non-negative

- This is not an LP:

$$\begin{aligned} \max \quad & -1/x_1 - x_3 \\ \text{s.t.} \quad & 2x_1 + x_2 < 3 \\ & x_1 + \alpha x_2 = 2 \ \forall \ \alpha \in \mathbb{R} \end{aligned}$$

  for the following reasons:

  1. Dividing by variables is not allowed
  2. Cannot have strict inequalities
  3. Must have finite number of constaints

Example 1.2.1. *LP Model*

$$\begin{aligned} \min \quad & x_1 - 2x_2 + x_4 \\ \text{s.t.} \quad & x_1 - x_3 \leq 3 \\ & x_2 + x_4 \geq 2 \\ & x_1 + x_2 = 4 \\ & x_1, x_2, x_3, x_4 \geq 0 \end{aligned}$$

1.2.2 Multiperiod Methods

A main feature of the WaterTech model is that the decisions about production levels have to be made once and for all. In practice, we often have to make a **series** of decisions that influence each other.

One such example is **multiperiod models**:

- time is split into **periods**
- we have to make a decision **in each period**
- all decisions influence the final outcome

Example 1.2.2. ***KW Oil***

KW Oil is the local supplier of heating oil. It needs to decide on how much oil to purchase in order to satisfy demand of its customers. Years of experience give the following demand forecast for the next 4 months:

Month	1	2	3	4
Demand (l)	5000	8000	9000	6000

The projected price of oil fluctuates from month to month:

Month	1	2	3	4
Price ($\$/l$)	0.75	0.72	0.92	0.90

Question: *when should we purchase how much oil when the goal is to minimize overall total cost?*

Additional complication: *The company has a storage tank that*

- *has a capacity of 4000 litres of oil*
- *currently (beginning of month 1) contains 2000 litres of oil*

Assumption: *Oil is delivered at the beginning of the month, and consumption occurs in the middle of the month.*

We first need to decide how many litres of oil to purchase in each month $i \implies$ variable p_i for $i \in [1, 4]$, and how much oil is stored in the tank at the beginning of month $i \implies$ variable t_i for $i \in [1, 4]$.

Objective function:

Minimize cost of oil procurement

$$\min \quad 0.75p_1 + 0.72p_2 + 0.92p_3 + 0.90p_4$$

Constaints: when do

$$t_1, \ldots, t_4, p_1, \ldots, p_4$$

corresponds to a feasible purchasing scheme?

By assumption, oil is purchased at the beginning of month, and is consumed afterwards. Therefore, we need

$$p_i + t_i \geq \{\text{demand in month } i\} \implies p_i + t_i = \{\text{demand in month } i\} + t_{i+1}$$

We have the following equations:

$$\begin{aligned} p_1 + 2000 &= 5000 + t_2 \\ p_2 + t_2 &= 8000 + t_3 \\ p_3 + t_3 &= 9000 + t_4 \\ p_4 + t_4 &\geq 6000 \end{aligned}$$

The entire LP is

$$\begin{aligned} \min \quad & 0.75p_1 + 0.72p_2 + 0.92p_3 + 0.90p_4 \\ \text{s.t.} \quad & \end{aligned}$$

$$\begin{aligned} p_1 + 2000 &= 5000 + t_2 \\ p_2 + t_2 &= 8000 + t_3 \\ p_3 + t_3 &= 9000 + t_4 \\ p_4 + t_4 &\geq 6000 \\ t_1 &= 2000 \\ t_i &\leq 4000 \;\forall\; i \in [2,4] \\ t_i, p_i &\geq 0 \;\forall\; i \in [1,4] \end{aligned}$$

Solution. We get $p = (3000, 12000, 5000, 6000)^T$ and $t = (2000, 0, 4000, 0)^T$. ■

We can always add additional add-on features to the example:

- storage comes at a cost, \$1.5 per litre/month - add $\sum_{i=1}^{4} 0.15t_i$ to objective
- minimize the maximum # of litres of oil purchased over all months
  - we will need a new variable M for maximum # of litres purchased
  - we will have to add constaints $p_i \leq M$ for all $i \in [1,4]$
  - We need to replace the objective function with $\min \quad M$ such that

$$\begin{aligned} \min \quad & M \\ \text{s.t.} \quad & \\ & p_1 + 2000 && = 5000 + t_2 \\ & p_2 + t_2 && = 8000 + t_3 \\ & p_3 + t_3 && = 9000 + t_4 \\ & p_4 + t_4 && \geq 6000 \\ & t_1 && = 2000 \\ & t_i && \leq 4000 \;\forall\; i \in [2,4] \\ & p_i && \leq M \;\forall\; i \in [1,4] \\ & t_i, p_i && \geq 0 \;\forall\; i \in [1,4] \end{aligned}$$

Correctness:

Why is this a correct model?

Suppose that $M, p_1, \ldots, p_4, t_1, \ldots, t_4$ is an optimal solution to the LP, clearly $M \geq \max_i \; p_i$. Since M, p, t is optimal we must have $M = \max_i \; p_i$. Why?

Otherwise, we could decrease M by a little bit, without violating the feasibility. This would contradict optimality because we would get a new feasible solution with a smaller objective function.

1.3 IP Models

Recap the solution to the WaterTech problem. However, fractional solutions are often not desirable! Can we force solutions to take on only integer values?

Yes! An **integer program** is a linear program with added integrality constaints for some/all of the variables. We call an IP **mixed** if there are integer and fractional variabels, and **pure** otherwise.

The difference between LPs and IPs is **subtle**, yet, LPs are easy to solve, IPs do not!

Can we solve IPs?

- Every problem instance has a **size** which we normally denote by n. Think: $n \sim$ number of variables/constaints of IP.
- The **running time** of an algorithm is then the number of steps that an algorithm takes.
- It is stated as a **function** of n: $f(n)$ measures the largest number of steps an algorithm takes on an instance of size n.

An algorithm is **efficient** if its running time $f(n)$ is a polynomial of n. LPs can be solved efficiently. IPs are very unlikely to have efficient algorithms!
It is very important to find an efficient algorithm of a problem.

1.3.1 IP Models: Knapsack

Example 1.3.1. ***KitchTech Shipping***

A company wishes to ship crates from Toronto to Kitchener. Each crate type has a weight and a value:

Type	1	2	3	4	5	6
weight (lbs)	30	20	30	90	30	70
value ($)	60	70	40	70	20	90

The total weight of crates shipped must not exceed 10,000 lbs. The goal is to ***maximizes*** *the total value of shipped goods.*

Variables x_i for the number of crates of type i to pack

Constraints total weight of creates picked must not exceed 10,000 lbs

$$30x_1 + 20x_2 + 3x_3 + 90x_4 + 30x_5 + 70x_6 \leq 10,000$$

Objective function maximize the total value:

$$\max \quad 60x_1 + 70x_2 + 40x_3 + 70x_4 + 20x_5 + 90x_6$$

We have the IP Model:

$$\begin{aligned} \max \quad & 60x_1 + 70x_2 + 40x_3 + 70x_4 + 20x_5 + 90x_6 \\ \text{s.t.} \quad & 30x_1 + 20x_2 + 3x_3 + 90x_4 + 30x_5 + 70x_6 \leq 10,000 \\ & x_i \geq 0 \ (i \in [1,6]) \\ & x_i \in \mathbb{Z} \ (i \in [1,6]) \end{aligned}$$

Example 1.3.2. ***KitchTech: Additional Conditions***

Suppose that we must not send more than 10 crates of the same type, and we can only send crates of type 3, if we send at least 1 crate of type 4. Note that we can send at least 10 crates of type 3 by the previous constaints!

The new IP model becomes:

$$\begin{array}{rl} \max & 60x_1 + 70x_2 + 40x_3 + 70x_4 + 20x_5 + 90x_6 \\ \text{s.t.} & 30x_1 + 20x_2 + 3x_3 + 90x_4 + 30x_5 + 70x_6 \leq 10,000 \\ & x_3 \leq 10x_4 \\ & 0 \leq x_i \leq 10 \ (i \in [1,6]) \\ & x_i \in \mathbb{Z} \ (i \in [1,6]) \end{array}$$

Example 1.3.3. ***KitchTech: 1 more tricky case***

Suppose that we must

- *take a total of at least 4 crates of type 1 or 2, or*
- *take at least 4 crates of type 5 or 6*

We will create a new variable y such that

- $y = 1 \implies x_1 + x_2 \geq 4$
- $y = 0 \implies x_5 + x_6 \geq 4$

and y has to take value 0 or 1.

The new IP model becomes:

$$\begin{array}{rl} \max & 60x_1 + 70x_2 + 40x_3 + 70x_4 + 20x_5 + 90x_6 \\ \text{s.t.} & 30x_1 + 20x_2 + 3x_3 + 90x_4 + 30x_5 + 70x_6 \leq 10,000 \\ & x_3 \leq 10x_4 \\ & x_1 + x_2 \geq 4y \\ & x_5 + x_6 \geq 4(1-y) \\ & 0 \leq y \leq 1 \\ & 0 \leq x_i \leq 10 \ (i \in [1,6]) \\ & x_i \in \mathbb{Z} \ (i \in [1,6]) \\ & y \in \mathbb{Z} \end{array}$$

In this example, y is called a binary variable. These are very useful for modeling **logical constraints** of the form [Condition (A or B) and C] → D.

1.3.2 IP Models: Scheduling

Example 1.3.4. ***CoffeeShop***

The neighbourhood coffee shop only opens on workdays. The daily demand for workers is

Mon	*Tues*	*Wed*	*Thurs*	*Fri*
3	*5*	*9*	*2*	*7*

Each worker works for 4 consecutive days and has one day off. The goal is to hire the samllest number of workers so that the demand can be met!

Variables : What do we need to decide on?

variable x_d for every $d \in \{M, T, W, Th, F\}$, the number of people to hire with starting day d.

Objective function : What do we want to minimize?

the total number of people hired:

$$\min \quad x_M + x_T + x_W + x_{Th} + x_F$$

Constraints : We need to ensure that enough people work on each of the days:

Question: given a solution $(x_M, x_T, x_W, x_{Th}, x_F)$, how many people work on Monday?

All but those start on Tuesdays (because they rest on Monday), i.e. $x_M + x_W + x_{Th} + x_F$.

The entire LP is

$$
\begin{aligned}
\min \quad & x_M + x_T + x_W + x_{Th} + x_F \\
\text{s.t.} \quad & x_M + x_W + x_{Th} + x_F \geq 3 \\
& x_M + x_T + x_{Th} + x_F \geq 5 \\
& x_M + x_T + x_W + x_F \geq 9 \\
& x_M + x_T + x_W + x_{Th} \geq 2 \\
& x_T + x_W + x_{Th} + x_F \geq 7 \\
& x \geq \mathbb{0} \\
& x \in \mathbb{Z}
\end{aligned}
$$

Example 1.3.5. ***Quiz***

We are given an integer program with integer values $x_1, \ldots, x_6$. Let

$$\mathcal{S} := \{127, 289, 1310, 2754\}$$

We want to add constaints and/or variables to the IP that enforce that the $x_1 + \cdots + x_6 \in \mathcal{S}$. How?

Solution. We can add binary variables y_n where $n \in \mathcal{S}$. Then exactly 1 of these variables to take the value 1 in a feasible solution. If $y_n = 1$, for some $n \in \mathcal{S}$, then $\sum_{i=0}^{6} x_i = n$.

The constraint is:

$$
\begin{aligned}
& \sum_{n \in \mathcal{S}} y_n = 1 \\
& \sum_{i=1}^{6} x_i = \sum_{i \in \mathcal{S}} i y_i \\
& 0 \leq y_i \leq 1 \\
& y_i \in \mathbb{Z} \ (\forall\ i \in \mathcal{S})
\end{aligned}
$$

■

1.4 Optimization on Graphs

Familiar problem - starting at location s, we wish to travel to t, what is the best (shortest) route?

Goal: Write the problem of finding the shortest route between s and t as an integer program!

Rephrasing the problem in the language of graphy theory helps.

A graph G consists of

Vertices $u, w, \ldots \in V$ (drawn as filled circles)

Edges $uw, wz, \ldots \in E$ (drawn as lines connecting circles)

Two vertices u and v are **adjacent** if $uv \in E$. Vertices u and v are the **endpoints** of edge $uv \in E$, and edge $e \in E$ is **incident** to $u \in V$ if u is an endpoint of e.

A s,t-path in $G = (V, E)$ is a sequence

$$v_1v_2, v_2v_3, v_3v_4, \ldots, v_{k-2}v_{k-1}, v_{k-1}v_k$$

where

- $v_i \in V$ and $v_i v_{i+1} \in E$ for all i
- $v_1 = s$, $v_k = t$, and $v_i \neq v_j$ for all $i \neq j$. Without this, it is called s,t-walk.

Graphs are useful to compactly model amny real-world entities.

Example 1.4.1. ***Map of Water Town***

We can think of the street map as a graph, G.

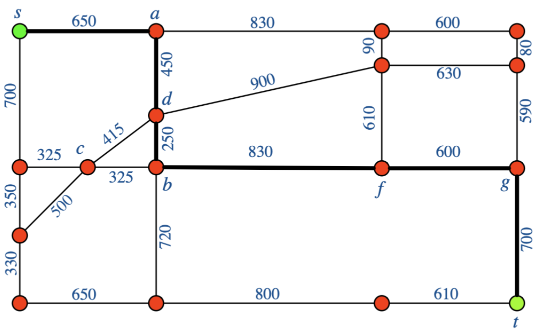


Figure 1.1: Map of Water Town

Vertices: road intersections
Edges: Road segments connecting adjacent intersections.

Each edge $e \in E$ is labelled by its length $c_e \geq 0$. We are looking for a path connecting s and t of smallest total length.

Solution. The shortest path to the Water Town problem is $P = sa, ad, db, bf, fg, gt$ with

$$\begin{aligned} c(P) &= c_{sa} + c_{ad} + c_{db} + c_{bf} + c_{fg} + c_{gt} \\ &= 650 + 490 + 250 + 830 + 600 + 700 = 3520 \end{aligned}$$

■

The length of a path $P = v_1v_2, \ldots, v_{k-1}v_k$ is the **sum of the lengths** of the edges on P:

$$c(P) := \sum (c_e : e \in P)$$

1.4.1 Matching Problem

Example 1.4.2. ***WaterTech - Job Assignment***

WaterTech has a collection of important jobs: $J = \{1', 2', 3', 4'\}$ that it needs to handle urgently. It also has 4 employees: $E = \{1, 2, 3, 4\}$ that need to handle these jobs. Employees have different skill-sets and may take diffent amounts of time to execute a job.

Employees	Jobs			
	1'	2'	3'	4'
1	-	5	-	7
2	8	-	2	-
3	-	1	-	-
4	8	-	3	-

Note: some workers are not able to handle certain jobs.

***Goal**: Assign each worker to exactly one task so that the **total execution time** is smallest!*

Solution. We will rephrase this in the language of graphs.

We create a graph with **one vertex** for each employee and job.
Add an edge ij for $i \in Em$ and $j \in J$ if employee i can handle job J.
Let the **cost** c_{ij} of edge ij be the amount of time needed by i to complete j.

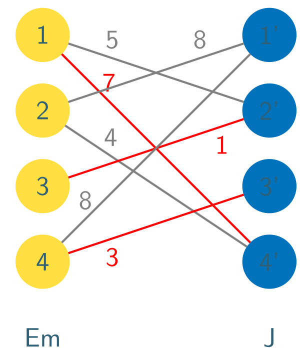


Figure 1.2: WaterTech Job Assignment Graph

■

Definition 3: Matching

A collection $M \subseteq E$ is a matching if no two edges ij, $i'j' \in M$ $(ij \neq i'j')$ share an endpoint; i.e. $\{i,j\} \cap \{i',j'\} = \emptyset$

The cost of matching M is the sum of costs of its edges:

$$c(M) = \sum (c_e : e \in M)$$

Definition 4: Perfect Matching

A mathcing M is perfect if every vertex v in the graph is incident to an edge in M.

Note: Perfect matchings correspond to feasible assignments of workers to jobs!

Solution. Continued from above, we can see that in Figure 1.1, $M = \{14', 21', 32', 43'\}$ is a perfect matching, thus one solution to the problem would be

$$1 \to 4', 2 \to 1', 3 \to 2', 4 \to 3'$$

whose execution time equals $c(M) = 19$.

Restatement of original question: find a perfect matching M in our graph of smallest cost. ■

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

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

Theorem 1: Perfect Matching

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

The IP will have a **binary variable** x_e for every edge $e \in E$. The idea is

$$x_e = 1 \leftrightarrow e \in M$$

Constraints: $\forall\ v \in V$, we need

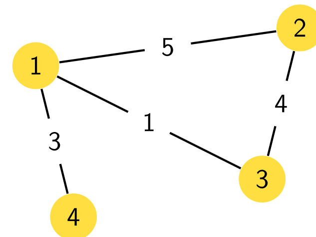
$$\sum (x_e : e \in \delta(v)) = 1$$

Objective:

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

An IP for Perfect Matching

We have the graph to the right, and want to find a perfect matching with minimum cost.



Solution.

$$\begin{aligned} \min \quad & \sum(c_e x_e : e \in E) & & \min \quad (5,1,3,4)x \\ \text{s.t.} \quad & \sum(x_e : e \in \delta(v)) = 1 \ (\forall\ v \in V) & \implies & \text{s.t.} \quad \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} x = \mathbb{1} \\ & x \geq \mathbb{0},\ x \in \mathbb{Z} & & \quad x \geq \mathbb{0},\ x \in \mathbb{Z} \end{aligned} \tag{1}$$

where $x = \{x_e : x \in \delta(v)\}$ for all $v \in V$. (1) gives the vector of $\sum(x_e : e \in \delta(v))$ for any $v \in V$. ■

1.5 Shortest Paths

Given : Graph $G = (V, E)$, length $c_e \geq 0$ for all $e \in E$, $s, t \in V$

Find : Minimum-length s, t-path P

Useful observation: Let $C \subseteq E$ be a set of edges whose removal **disconnects** s and t $\rightarrow$ Every s, t-path must have at least one edge in C.

Definition 5

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

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

Definition 6: Cuts

$\delta(S)$ is an s, t-cut if $s \in S$ and $t \notin S$.

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. Suppose 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:

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

Then by assumption, $t \notin R$ since S doesn't contain a 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 abour 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 a s,t-path. □

An IP for Shortest Paths

Variables : We have one binary variable x_e for each edge $e \in E$. We want

$$x_e = \begin{cases} 1 & e \in P \\ 0 & \text{otherwise} \end{cases}$$

Constraints : We have one constraint for each s,t-cut $\delta(U)$, forcing P to have an edge from $\delta(S)$.

$$\sum(x_e : e \in \delta(U)) \geq 1$$

for all s,t-cuts $\delta(U)$.

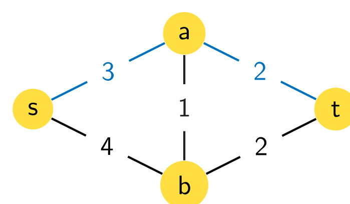
Objective :

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

.

By Remark 1.5.1 and Remark 1.5.2, the s,t-path P will contain at least one edge from every s,t-cut, i.e. for any $\delta(U)$, P must contain at least one edge from it. This makes the constraint. And to optimize the set, we try to find the path with least cost.

We have the graph to the right, and want to find a perfect matching with minimum cost.



Solution.

$$\begin{aligned} \min \quad & \sum(c_e x_e : e \in E) \\ \text{s.t.} \quad & \sum(x_e : e \in \delta(U)) \geq 1 \ (U \subseteq V, s \in U, t \notin U) \\ & x_e \geq 0,\ x_e \in \mathbb{Z} \ (e \in E) \\ & \quad \Downarrow \\ \min \quad & (3,4,1,2,2)x \\ \text{s.t.} \quad & \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_{sa} \\ x_{sb} \\ x_{ab} \\ x_{at} \\ x_{bt} \end{pmatrix} \geq \mathbb{1} \\ & x \geq \mathbb{0},\ x \in \mathbb{Z} \end{aligned}$$

■

For an optimal solution, $x_e \leq 1$ for all $e \in E$, since if $x_e > 1$, making $x_e = 1$ would be cheaper and maintains feasibility!

For a binary solution x, define

$$S_x = \{e \in E : x_e = 1\}$$

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 a shortest s,t-path.

1.6 Nonlinear Models

A **nonlinear program (NLP)** is of the form

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & g_1(x) \leq 0 \\ & g_2(x) \leq 0 \\ & \cdots \\ & g_m(x) \leq 0 \end{aligned}$$

where

- $x \in \mathbb{R}^n$
- $f : \mathbb{R}^n \to \mathbb{R}$
- $g_i : \mathbb{R}^n \to \mathbb{R}$

Example 1.6.1. ***Finding Close Points in LP***

We are given an LP(P), and an infeasible point $\bar{x}$. The goal is to find a point $x \in P$ that is as close as possible to $\bar{x}$, i.e. find a point $x \in P$ that minimizes the Euclidean distance to $\bar{x}$.

$$\|x - \bar{x}\|_2 = \sqrt{\sum_{i=1}^{n} (x_i - \bar{x}_i)^2}$$

Note that $\|p\|_2$ is called the L^2-norm of p.

We have

$$\begin{aligned} \min \quad & \|x - \bar{x}\|_2 \\ \text{s.t.} \quad & x \in P, \ P = \{x : Ax \leq b\} \end{aligned}$$

Example 1.6.2. ***Binary IP via NLP***

Suppose we are given a binary IP (i.e. an integer program all of those variables are binary). Recall that (binary) IPs are generally hard to solve. Now, we can write any binary IP as an NLP.

Binary IP:

$$\begin{aligned} \max \quad & c^T x \\ \text{s.t.} \quad & Ax \leq b \\ & x \geq \mathbb{0} \\ & x_j \in \{0, 1\} \ (j \in \{1, \ldots, n\}) \end{aligned}$$

NLP:

$$\begin{aligned} \max \quad & c^T x \\ \text{s.t.} \quad & Ax \leq b \\ & x \geq \mathbb{0} \\ & x_j(1 - x_j) = 0 \ (j \in [n]) \end{aligned}$$

$$
\begin{aligned}
\max \quad & c^T x \\
\text{s.t.} \quad & Ax \leq b \\
& x \geq \mathbb{0} \\
& \sin(\pi x_j) = 0 \ (j \in [n])
\end{aligned}
$$

Example 1.6.3. ***Fermat's Last Theorem***

There are no integers $x, y, z \geq 1$ and $n \geq 3$ such that

$$x^n + y^n = z^n$$

NLP for Fermat's Last Theorem

$$
\begin{aligned}
\min \quad & (x_1^{x_4} + x_2^{x_4} - x_3^{x_4})^2 \\
& + (\sin \pi x_1)^2 + (\sin \pi x_2)^2 + (\sin \pi x_3)^2 + (\sin \pi x_4)^2 \\
\text{s.t.} \quad & x_i \geq 1 \ (i = 1 \cdots 3) \\
& x_4 \geq 3
\end{aligned}
$$

The NLP is trivially feasible, and the value of any feasible solution is non-negative as its objective is the **sum of squares**.

In fact, the value of a solution (x_1, x_2, x_3, x_4) is 0 iff

- $x_1^{x_4} + x_2^{x_4} = x_3^{x^4}$
- $\sin \pi x_i = 0$ for all $i = 1, \ldots, 3$

Remark 4

Fermat's Last Theorem is true iff the NLP has optimal value greater than 0.

Note: It is well known that there is an infinite sequence of feasible solutions whose objective value converges to 0. Proving Fermat's Last Theorem suffices to show that the value 0 cannot be attained.

Module 2

Linear Programs

2.1 Possible Outcomes

When we solve an optimization problem, the input will be a LP/IP/NLP program, and the algorithm (software) outputs the solution.

Definition 7: Feasible Solution

All assignment of values to each of the variables is a feasible solution if all the constraints are satisfied.

Definition

An optimization problem is **feasible** if it has at least one feasible solution. It is **infeasible** otherwise.

Optimal solution

- 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 can have several optimal solutions.

unbounded

- A maximization problem is **unbounded** if for every value M, there exists an 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.

There are three possible outcomes for an optimization problem:

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

But, there can be other outcomes!

Example 2.1.1. *Consider*

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

This is feasible since one solution could be $x = 0$, and it is not unbounded since 1 is an upper bound. However, this model has no optimal solution.

This is because the model is not a linear program, it contains strict inequality.

Theorem 2: Foundamental Theorem of Linear Programming

For any linear prorgam, **exactly** one of the following holds:

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

What it means by solving a LP:

- It has an optimal solution: return an optimal solution $\bar{x}$ and proof that $\bar{x}$ is optimal
- It is infeasible: return a proof that LP is infeasible
- It is unbounded: return a proof that LP is unbounded

2.2 Certificates

How can we prove that a solution is infeasible?

2.2.1 Infeasibilty of LP Model

Example 2.2.1. *The following LP is infeasible:*

$$\begin{aligned} \max \quad & (3,4,-1,2)^T x \\ \text{s.t.} \quad & \begin{pmatrix} 3 & -2 & -6 & 7 \\ 2 & -1 & -2 & 4 \end{pmatrix} x = \begin{pmatrix} 6 \\ 2 \end{pmatrix} \\ & x \geq \mathbb{0} \end{aligned}$$

Proof. One way of proving is to construct a system of equations and show that the system has no solutions:

$$\begin{cases} \begin{pmatrix} -3 & 2 & 6 & -7 \end{pmatrix} x = 6 \\ \begin{pmatrix} 4 & -2 & -4 & 8 \end{pmatrix} x = 4 \end{cases}$$

After we do $-1 \times (1) + 2 \times (2)$, we have

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

Suppose there exists $\bar{x} \geq \mathbb{0}$ satisfying (1), (2). Then $\bar{x}$ satisfies the last equation we produce:

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

leads to a contradiction.

Another way of proving this is using matrix formulations. Suppose for a contradiction there is a solution $\bar{x}$ to $x \geq \mathbb{0}$ and

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

We construct a new equation:

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

$$\begin{pmatrix} 1 & 0 & 2 & 1 \end{pmatrix} x = -2 \qquad (y^T A x = y^T b)$$

Since $\bar{x}$ satisfies the last equation which means that

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

This is a contradiction. □

Theorem 3: Farkas' Lemma

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

$$y^T A \geq \mathbb{0}^T \qquad y^T b < 0$$

2.2.2 Optimality

We cannot try all possible feasible solutions to find the optimal solution.

Example 2.2.2. *We have*

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

We claim that $\bar{x}$ with $\bar{x} = \begin{pmatrix} 0 & 0 & 4 & 5 \end{pmatrix}$ is feasible solution of value 4 (easy to prove), and 4 is an upper bound.

Proof. Let x' be an aribitrary feasible solution, then

$$z(x') = \underbrace{\begin{pmatrix} -1 & -4 & 0 & 0 \end{pmatrix}}_{\leq \mathbb{0}} \underbrace{x'}_{\geq \mathbb{0}} + 4 \leq 4$$

□

2.2.3 Unboundedness

Example 2.2.3. *We have*

$$\begin{array}{ll} \max & z := \begin{pmatrix} -1 & 0 & 0 & 1 \end{pmatrix} x \\ \text{s.t.} & \begin{pmatrix} -1 & -1 & 1 & 0 \\ -2 & 1 & 0 & 1 \end{pmatrix} x = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \\ & x \geq \mathbb{0} \end{array}$$

How can we prove that this problem is unbounded?

The idea is to construct a family of feasible solutions $x(t)$ for all $t \geq 0$ and show that as t goes to infinity, the value of the objective function goes to infinity.

Proof. We solve the matrix equation

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

and get

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

Claim 1: $x(t)$ is feasible for all $t \geq 0$.

Since for all $t \geq 0$ as $\bar{x}, r \geq \mathbb{0}$,

$$x(t) = \bar{x} + tr \geq \mathbb{0} \to Ax(t) = A[\bar{x} + tr] = \underbrace{A\bar{x}}_{b} + t \underbrace{Ar}_{\mathbb{0}} = b$$

Claim 2: $z \to \infty$ when $t \to \infty$.

Let $c^T = \begin{pmatrix} -1 & 0 & 0 & 1 \end{pmatrix}$,

$$z = c^T x(T) = c^T[\bar{x} + tr] = c^T\bar{x} + t\underbrace{c^t r}_{=1>0}$$

□

Remark 5

The linear program

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

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

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

2.3 Standard Equality Forms

Definition 8: SEF

A LP is in Standard Equality Form (SEF) if

- it is a maximization problem
- for every variable x_j, we have the constraint $x_j \geq 0$ and
- all other constraints are equality constraints

Remark 6

For the following LP:

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

there is no constraint $x_2 \geq 0$, we say x_2 is free. Though $x_2 \geq 0$ is implied by the constraints, x_2 is still free since $x_2 \geq 0$ is not given explicitly.

We will develop an algorithm called the Simplex that can solve any LP **as long as it is in Standard Equality Form (SEF)**.

Idea:

1. Find an "equivalent" LP in SEF
2. Solve the "equivalent" LP using Simplex
3. Use the solution of "equivalent" LP to get the solution of the original LP

Definition 9: Equivalent

LP (P) and (Q) are equivalent if

- (P) infeasible $\implies$ (Q) infeasible
- (P) unbounded $\implies$ (Q) unbounded
- can construct optimal solution of (P) from optimal solution of (Q)
- can construct optimal solution of (Q) from optimal solution of (P)

Theorem 4

Every LP is equivalent to an LP in SEF.

How do we change minimum problem to maximum problem?

Take the oppsite sign of the objective function and find its maximum.

How do we replace an inequality with an equality?

Suppose an LP has the constraint

$$x_1 - x_2 + x_4 \leq 7$$

We can replace it by

$$x_1 - x_2 + x_4 + s = 7 \qquad s \geq 0$$

Suppose an LP has the constraint

$$x_1 - x_2 + x_4 \geq 7$$

We can replace it by

$$x_1 - x_2 + x_4 - s = 7 \qquad s \geq 0$$

What if we have a free variable?

Example 2.3.1.

$$\begin{aligned} \max \quad & z = (1,2,3)(x_1, x_2, x_3)^T \\ \text{s.t.} \quad & \begin{pmatrix} 1 & 5 & 3 \\ 2 & -1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 5 \\ 4 \end{pmatrix} \\ & x_1, x_2 \geq 0, \ x_3 \textit{ is free} \end{aligned}$$

The idea is that any number is the difference between two **non-negative** numbers.

Solution. Set $x_3 := a - b$ where $a, b \geq 0$.

$$\begin{aligned} z &= (1,2,3)(x_1, x_2, x_3)^T \\ &= x_1 + 2x_2 + 3x_3 \\ &= x_1 + 2x_2 + 3(a-b) \\ &= x_1 + 2x_2 + 3a - 3b \\ &= (1,2,3,-3)(x_1, x_2, a, b)^T \end{aligned}$$

and

$$\begin{pmatrix} 1 & 5 & 3 \\ 2 & -1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 & 5 & 3 & -3 \\ 2 & -1 & 2 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ a \\ b \end{pmatrix} = \begin{pmatrix} 5 \\ 4 \end{pmatrix}$$

■

2.4 Simplex - A First Attempt

A naive way to solve an LP:

Step 1 Find a feasible solution x

Step 2 If x is optimal, STOP

Step 3 If LP is unbounded, STOP

Step 4 Find a "better" feasible solution

Some questions we have with this method:

- How do we find a feasible solution?
- How do we find a "better" solution?

- Will this ever terminate?

Example 2.4.1. *We want to solve*

$$\begin{aligned} \max \quad & (4,3,0,0)x + 7 \\ \text{s.t.} \quad & \begin{pmatrix} 3 & 2 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{pmatrix} x = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \\ & x_1, x_2, x_3, x_4 \geq 0 \end{aligned}$$

We have a feasible solution $x = (0,0,2,1)^T$ and the objective function has value 7. Can we find a feasible solution larger than 7?

The idea is to increase x_1 as much as possible, but keep x_2 unchanged.

Solution. Let $x_1 = t$, $x_2 = 0$, the equality constraints and the non-negativity constraints need to be satisfied. We get

$$\begin{pmatrix} x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} - t \begin{pmatrix} 3 \\ 1 \end{pmatrix}$$

that equality constraints hold for any choice of t.

$$\begin{aligned} x_3 = 2 - 3t \geq 0 &\implies t \leq \frac{2}{3} \\ x_4 = 1 - t \geq 0 &\implies t \leq 1 \end{aligned}$$

The largest possible t is $\min\{1, \frac{2}{3}\} = \frac{2}{3}$, the new solution is then

$$x = (t, 0, 2-3t, 1-t)^T = \left(\frac{2}{3}, 0, 0, \frac{1}{3}\right)^T$$

■

Is this new solution optimal? NO! Can we use the same trick to get a better solution? NO! To make it work, the LP needs to be in "canonical" form.

Revised Strategy

Step 1 Find a feasible solution x

Step 2 Rewrite LP so that it is in canonical form

Step 3 If x is optimal, STOP

Step 4 If LP is unbounded, STOP

Step 5 Find a "better" feasible solution

We need to define canonical and prove that we can always rewrite LP in canonical form.

2.5 Basis

Notation: let B be a subset of column indices, then A_B is a column sub-matrix of A indexed by set B. A_j denotes the column j of A.

Definition 10: Basis

Let B be a subset of column indices, B is a basis if

1. A_B is a square matrix
2. A_B is non-sigular (columns are independent)

Does every matrix have a basis? NO!

Theorem 5

Max number of independent columns = Max number of independent rows

Remark 7

Let A be a matrix with independent rows, then B is a basis iff B is a maximal set of independent columns of A.

Definition 11: Basic solution

x is a basic solution for basis B if

1. $Ax = b$
2. $x_j = 0$ whenever $j \notin B$

Example 2.5.1. *For the following equation*

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

$x = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}$ *is a basic solution for* $B = \{1, 2, 3\}$ *since*

1. $Ax = b$
2. $x_4 = x_5 = 0$

Find a Basic Solution

How to find 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

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

and solving the equation gives us

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

so the basic solution is then $(4, 0, 0, 2)^T$. ■

Remark 8

Consider $Ax = b$ and a basis B of A, 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 non-singular - A_B^{-1} exists. Hence, $x_B = A_B^{-1} b$. □

Definition 12: Basic Solution of LP

Consider $Ax = b$ with independent rows, vector x is a basic solution if it is a basic solution for *some* basis B.

Example 2.5.2. *Consider the following equation:*

$$\underbrace{\begin{pmatrix} 3 & 2 & 1 & 4 & 1 \\ -1 & 1 & 0 & 2 & 1 \end{pmatrix}}_{A} x = \underbrace{\begin{pmatrix} 6 \\ 3 \end{pmatrix}}_{b}$$

is $x = (0, 1, 0, 1, 0)^T$ *basic?*

Proof. No. By contradiction, suppose x is basic for basis B.

- $x_2 = 1 \neq 0 \implies 2 \in B$
- $x_4 = 1 \neq 0 \implies 4 \in B$

Thus,

$$A_{\{2,4\}} = \begin{pmatrix} 2 & 4 \\ 1 & 2 \end{pmatrix}$$

is a column matrix of A_B. But the columns of $A_{\{2,4\}}$ are dependent, so A_B is singular and B is not a basis - contradiction. □

Remark 9

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

Consider the problem in SEF:

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

If the rows of A are dependent, then either

- there is no solution to $Ax = b$, (P) is infeasible
- a constriant of $Ax = b$ can be removed *without changing the solutions*

Remark 10

We may assume, when trying to solve (P), that rows of A are independent.

Definition 13: Basic Feasible Solution

A basic solution x of $Ax = b$ is feasible if $x \geq 0$, i.e. if it is feasible for (P). A basic solution is feasible if it is non-negative.

2.6 Canonical Forms

Consider the problem in SEF:

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

Definition 14: Canonical Form

Let B be a basis of A, then (P) is in **canonical form** for B if
P1 $A_B = I$, and
P2 $c_j = 0$ for all $j \in B$.

Idea: for any basis B we can rewrite (P) so that it is in canonical form for a basis B and such that the resulting LP behaves the same as (P). More foramlly, we will show the following:

Remark 11

For any basis B, there exists (P') in canonical form of B such that
1. (P) and (P') have the same feasible region, and
2. **feasible solutions** have the same objective value for (P) and (P').

Rewrite LP in Canonical Form

We have the LP model

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

How do we rewrite (P) in canonical form for basis $B = \{2, 3\}$?

Solution. We have the following steps:

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

$$\begin{aligned} & \begin{pmatrix} 1 & 0 & 1 & -1 \\ 0 & 1 & 1 & 2 \end{pmatrix} x = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \\ \leftrightarrow \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}^{-1} & \begin{pmatrix} 1 & 0 & 1 & -1 \\ 0 & 1 & 1 & 2 \end{pmatrix} x = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ 2 \end{pmatrix} \\ \leftrightarrow & \begin{pmatrix} -1 & 1 & 0 & 3 \\ 1 & 0 & 1 & -1 \end{pmatrix} x = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \end{aligned}$$

since

$$Ax = b \leftrightarrow \underbrace{A_B^{-1} A}_{A'}\, x = \underbrace{A_B^{-1} b}_{b'}$$

P2 Replace $c^T x$ by $\bar{c}^T x + \bar{z}$ 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, and
- adding the result constraints to the objective function

Step 2 choose y_1, y_2 to get $\bar{c}_B = 0$

We have

$$\begin{aligned} &0 = -(y_1, y_2)\begin{pmatrix} 1 & 0 & 1 & -1 \\ 0 & 1 & 1 & 2 \end{pmatrix} x + (y_1, y_2)\begin{pmatrix} 1 \\ 2 \end{pmatrix} \\ &z = (0,0,2,4)x \\ \implies &z = \left[(0,0,2,4) - (y_1, y_2)\begin{pmatrix} 1 & 0 & 1 & -1 \\ 0 & 1 & 1 & 2 \end{pmatrix}\right] x + (y_1, y_2)\begin{pmatrix} 1 \\ 2 \end{pmatrix} \end{aligned}$$

Remark 12

For any choice of y_1, y_2 and any feasible solution x, objective value of x for **old** objective function $=$ objective value of x for **new** objective function.

$$z = \underbrace{\left[(0,0,2,4) - (y_1, y_2)\begin{pmatrix} 1 & 0 & 1 & -1 \\ 0 & 1 & 1 & 2 \end{pmatrix}\right]}_{\bar{c}^T} x + \underbrace{(y_1, y_2)\begin{pmatrix} 1 \\ 2 \end{pmatrix}}_{\bar{z}}$$

$$\begin{aligned} &(0,0) = \bar{c}_B^T = (0,2) - (y_1, y_2)\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \\ \leftrightarrow \quad &(y_1, y_2)\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} = (0,2) \\ \leftrightarrow \quad &\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}^{-1}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \end{pmatrix} \\ \leftrightarrow \quad &\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}^{-1}\begin{pmatrix} 0 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix} \end{aligned}$$

Hence, we choose $(y_1, y_2) = (2, 0)$ and

$$z = (-2, 0, 0, 6)x + 2$$

In general, we have

$$\begin{aligned} &0 = -y^T Ax + y^T b \\ &z = c^T x \\ &z = [c^T - y^T A]x + y^T b \end{aligned}$$

■

Consider

$$z = \underbrace{[c^T - y^T A]}_{\bar{c}^T} x + \underbrace{y^T b}_{\bar{z}}$$

How do we choose y such that $\bar{c}_B = 0$ for a basis B?

$$\begin{aligned} &\mathbb{0}^T = \bar{c}_B^T = c_B^T - y^T A_B \\ \leftrightarrow \quad &y^T A_B = c_B^T \\ \leftrightarrow \quad &A_B^T y = c_B \\ \leftrightarrow \quad &y = (A_B^T)^{-1} c_B = A_B^{-T} c_B \end{aligned}$$

Remark 13

For any non-singular matrix M,
$$(M^T)^{-1} = (M^{-1})^T =: M^{-T}$$

Theorem 6

Consider A with basis B,
(P)

$$\begin{aligned} \max \quad & c^T x \\ \text{s.t.} \quad & Ax = b \\ & x \geq \mathbb{0} \end{aligned}$$

(P')

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

where $y = A_B^{-T}$, then

1. (P') is in canonical form for basis B, i.e. $\bar{c}_B = 0$ and $A'_B = I$
2. (P) and (P') have the same feasible region
3. **feasible solutions** have the same objective value for (P) and (P').

2.7 Formalizing the Simplex

Example 2.7.1. *Consider*

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

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

- *A_B is square and non-singular $\to$ B is a basis*
- *$A_B = I$ and $c_B = \mathbb{0}$ $\to$ LP is in canonical form for B*
- *$\bar{x} = (2, 0, 0, 5)^T$ is a basic solution*
- *$\bar{x} \geq \mathbb{0} \to \bar{x}$ is feasible, i.e. B is feasible.*

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.

We pick $k = 2$, set $x_2 = t \geq 0$, keep $x_3 = 0$. We want to choose basic variables such that $Ax = b$ holds.

We find
$$\underbrace{\begin{pmatrix} x_1 \\ x_4 \end{pmatrix}}_{x_B} = \underbrace{\begin{pmatrix} 2 \\ 5 \end{pmatrix}} - t \underbrace{\begin{pmatrix} 1 \\ 1 \end{pmatrix}}_{A_2}$$

and choose t as large as possible and basic variables must remain non-negative.

$$x_1 = 2 - t \geq 0 \implies t \leq 2$$
$$x_4 = 5 - t \geq 0 \implies x \leq 5$$

Thus, the largest possible $t = \min\{2, 5\} = 2$, and the new feasible solution is $x = (0, 2, 0, 3)^T$ with objective value $2 > 0$.

The new feasible solution is a basic solution on basis $B = \{2, 4\}$.

Old $\{1, 4\}$ is a feasible basis

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

New $\{2, 4\}$ is a feasible basis

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

Remark 14

We only need to know how to go from the OLD basis to a NEW basis!

In the above example, 2 entered the basis and 1 left the basis. Why?

We picked $k = 2 \notin B$, so that 2 enters the basis. We choose $t = 2$ instead of 5 makes $x_1 = 0$ and 1 leaves the basis.

If we now pick $k = 3 \notin B$ and set $x_3 = t$, 3 then enters the basis. We have

$$\begin{pmatrix} x_2 \\ x_4 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \end{pmatrix} - t \begin{pmatrix} 2 \\ -1 \end{pmatrix}$$

and get $t = \min\{1, -\} = 1$ thus $x_2 = 0$, making 2 leaving the basis.

The NEW basis is $B = \{3, 4\}$, and $x = (0, 0, 1, 4)^T$ is a basic solution.

$$\begin{array}{ll} \max & (-1.5,-0.5,0,0)x + 3 \\ \text{s.t.} & \begin{pmatrix} 0.5 & 0.5 & 1 & 0 \\ -0.5 & 0.5 & 0 & 1 \end{pmatrix} x = \begin{pmatrix} 1 \\ 4 \end{pmatrix} \\ & x \geq \mathbb{0} \end{array}$$

Claim: $(0, 0, 1, 4)^T$ has value 3, it is optimal since 3 is an upper bound.

Proof. Let x be a feasible solution, then

$$(-1.5, -0.5, 0, 0)x + 3 \leq 3$$

□

Example 2.7.2.

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

with $\{1, 4, 5\}$ *as a feasible basis.*

Solution. Pick $k = 3 \notin B$ and let $x_3 = t$, 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 $t = \min\{1, -.-\} = 1$, thus $x_1 = 0 \implies 1$ leaves the basis.

The NEW basis is then $B = \{3, 4, 5\}$.

We then choose $k = 2 \notin B$ and set $x_2 = t$, 2 enters the basis, and then

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

leads to an unbounded solution. ■

Claim: the LP is unbounded.

Proof.

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

where $\bar{x}, r$ are certificates of unboundedness.

- $x(t)$ is feasible for all $t \geq 0$
- $z \to \infty$ as $t \to \infty$

□

2.7.1 The Simplex Algorithm

LP model:

$$\begin{aligned} \max \quad & c^T x \\ \text{s.t.} \quad & Ax = b \\ & x \geq \mathbb{0} \end{aligned}$$

Algorithm 1: Simplex

Input : a feasible basis B
Output: an optimal solution OR it detects that LP is unbounded

while *true* **do**
  Rewrite in canonical form for the basis B
  Get $\bar{x}$ as a basic solution
  /* Find a better basis B or get required outcome */
  if $c_N \leq \mathbb{0}$ **then**
    STOP, the basic solution $\bar{x}$ is optimal
    return $\bar{x}$
  Pick $k \notin B$ such that $c_k > 0$ and set $x_k = t$
  Pick $x_B = b - tA_k$
  if $A_k \leq \mathbb{0}$ **then**
    STOP, the LP is unbounded
    return *Unboundedness outcome*
  Choose $t = \min\{\frac{b_i}{A_{ik}} :$ for all i such that $A_{ik} > 0\}$
  Let x_r be a basic variable forced to 0
  The new basis is obtained by having k enter and r leave

Remark 15

Simplex tells the truth

- If it claims that the LP is unbounded, it is unbounded
- If it claims the solution is optimal, it is optimal

Example 2.7.3. *We have the following model:*

$$\begin{aligned} \max \quad & (5,0,0,0,-3)\mathbf{x} + 12 \\ \text{s.t.} \quad & \begin{pmatrix} -1 & 1 & 0 & 0 & 1 \\ 2 & 0 & 1 & 0 & -1 \\ 3 & 0 & 0 & 1 & -1 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 4 \\ 2 \\ 6 \end{pmatrix} \\ & \mathbf{x} \geq \mathbb{0} \end{aligned}$$

One feasible solution we find is $(1,5,0,3,0)^T$ *with objective function value* 17.

Solution. We apply *Simplex* on our example above:

Transform the LP into canonical form for basis $\{1,2,4\}$:

$$\begin{aligned} \max \quad & \left(0,0,-\frac{5}{2},0,-\frac{1}{2}\right)\mathbf{x} + 17 \\ \text{s.t.} \quad & \begin{pmatrix} 1 & 0 & 1/2 & 0 & -1/2 \\ 0 & 1 & 1/2 & 0 & 1/2 \\ 0 & 0 & -3/2 & 1 & 1/2 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 1 \\ 5 \\ 3 \end{pmatrix} \\ & \mathbf{x} \geq \mathbb{0} \end{aligned}$$

Note that the objective function vector is $\leq \mathbb{0}$, so the objective function value is ≤ 17. Since our basic solution achieves this value, it is optimal. The algorithm then terminates.

■

Example 2.7.4. *We have another LP in SEF:*

$$\begin{aligned} \max \quad & (-1,3,0,0,1)\mathbf{x} \\ \text{s.t.} \quad & \begin{pmatrix} -2 & 4 & 1 & 0 & 1 \\ 3 & 7 & 0 & 1 & 1 \end{pmatrix}\mathbf{x} = \begin{pmatrix} 1 \\ 3 \end{pmatrix} \\ & \mathbf{x} \geq \mathbb{0} \end{aligned}$$

The LP is in canonical form for basis $\{3,4\}$ *with basic solution* $(0,0,1,3,0)$ *and objective function value* 0.

Solution. Running *Simplex* for one iteration, we get bfs $(0,0,0,2,1)$ for basis $\{4,5\}$. LP in canonical form for $\{4,5\}$ is

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

Only 1st coordinate of objective vector > 0, so set $x_1 := t$, $x_2, x_3 = 0$, and find x_4, x_5. We need to solve

$$A(t,0,0,x_4,x_5)^T = (2,1)$$

$$\begin{aligned} &\implies \begin{pmatrix} -1 \\ -2 \end{pmatrix} t + \begin{pmatrix} x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \\ &\implies x_4 = 2 + t, \quad x_5 = 1 + 2t \\ &\mathbf{x} \geq \mathbb{0} \implies t \geq 0, \quad 2 + t \geq 0, \quad 1 + 2t \geq 0 \end{aligned}$$

These are satisfied $\forall\ t \geq 0$. So, by taking t as large as we like, we see that the LP has arbitrarily large objective function value, i.e. the LP is unbounded.

■

Remark 16

Whenever the column A_k of the constraint matrix A is non-positive ($A_k \leq 0$), where k is the entering variable, the LP is unbounded.

Continued from previous example

Solution. The certificate of unboundedness:

We get the feasible solutions

$$\begin{aligned} f(t) &= (t,0,0,2+t,1+2t) \\ &= \underbrace{(0,0,0,2,1)}_{e} + t\underbrace{(1,0,0,1,2)}_{d} \end{aligned}$$

We notice that

1. e is feasible (previous bfs).
2. $d \geq \mathbb{0}$, $Ad = 0$, $c^T d > \mathbb{0}$.

(e,d) are a certificate of unboundedness. ■

Is the *Simplex* a correct algorithm? NOT AS STATED! IT MAY NOT TERMINATE.

Potential problem: infinite loop (cycling)

$$B_1 \rightsquigarrow B_2 \rightsquigarrow B_3 \rightsquigarrow \cdots \rightsquigarrow B_{k-1} \rightsquigarrow B_k = B_1$$

When there is an optimal solution, the algorithm may cycle through several bfs with the same objective function value.

However, with the **Bland Rule**, the algorithm terminates.

Theorem 7

If we use the **Bland's Rule**, then the Simplex algorithm always terminates.

Definition 15: Bland's Rule

- If we have a choice for the element **entering** the basis, pick the smallest one
- If we have a choice for the element **leaving** the basis, pick the smallest one

2.7.2 Finding a Feasible Solution

Example 2.7.5. ***Finding a feasible solution***

We have the following LP:

$$\begin{aligned} \max \quad & (2,-1,2)\mathbf{x} \\ \text{s.t.} \quad & \begin{pmatrix} -1 & -2 & 1 \\ 1 & 1 & 1 \end{pmatrix} \mathbf{x} = \begin{pmatrix} -1 \\ 3 \end{pmatrix} \\ & \mathbf{x} \geq \mathbb{0} \end{aligned} \tag{P1}$$

Is this feasible? If so, find a bfs.

Solution. We follow the following steps to find the feasible solution:

Step 1: Check if the equality constaints are feasible. We may do this by std. linear algebra (compute the RREF). If infeasible, the LP is infeasible, STOP.

Step 2: If $Ax = b$ is feasible (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: We will "bootstrap" to find a feasible solution (≥ 0). We introduce two new auxilary variables (in general, as many as the number of rows) - x_4, x_5 for our example.

We ask $x_4, x_5 \geq 0$, and we form a new LP that is guaranteed to be feasible.

But now b has a negative coordinate. If b has a negative coordinate, multiply the corresponding equations by -1 on both sides: new constraints:

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

Then argument A with the identity:

$$\begin{pmatrix} 1 & 2 & -1 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \quad x \geq \mathbb{0},\ x \in \mathbb{R}^5 \tag{P2}$$

These are feasible! $(0,0,0,1,3)$ is one, and is basic for $\{4,5\}$.

Step 4: we use the *Simplex* algorithm to try to find a feasible solution to the original LP (P1): run the algorithm on

$$\begin{aligned} \max \quad & (0,0,0,-1,-1)\mathbf{x} \\ \text{s.t.} \quad & \begin{pmatrix} 1 & 2 & -1 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 1 \\ 3 \end{pmatrix} \\ & \mathbf{x} \geq \mathbb{0} \end{aligned} \tag{P3}$$

Note that $\max(-x_4 - x_5) = -\min(x_4 + x_5)$.

If the max is 0, we get a feasible solution to (P1). Also, if (P1) is feasible, we can augment it with $x_4 = x_5 = 0$, to get a feasible solution for (P2), with value 0. Since the objective function vector is $\leq \mathbb{0}$, the LP (P2) is bounded, and since it is feasible, *Simplex* terminates with an optimal solution.
This tells us if (P1) is feasible or not: **if optimal value $< \mathbb{0} \implies$ (P1) is infeasible, otherwise (P1) is feasible.**

Solving (P3) using Simplex, we get:

- optimal basis $= \{1, 3\}$
- cooresponding basic feasible solution $= (2, 0, 1, 0, 0)^T$
- optimum $= 0$

This gives us a feasible solution for (P2): $(2, 0, 1)^T$ basic feasible solution for basis $\{1, 3\}$. Now we can run Simplex on (P2) with this solution. Doing this, we get optimal solution $(0, 4, 7)^T$, corresponding to basis $\{2, 3\}$ with optimum 10.

■

This is called the two-phase simplex method.

Remark 17

If the optimal solution of the derived program does not have an optimal value equal to 0, then the original program is infeasible and does not have a solution.

Two-phase Simplex Method

This method gives us:

- Since (P3) is feasible by construction and the optimal value is $\leq \mathbb{0}$, *Simplex* returns an optimal solution.
- If the optimal value is $< \mathbb{0}$, then (P2) is infeasible.
- If the optimal value $= 0$, we run Phase 2 of Simplex with the bfs given by Phase 1 on (P3).
- Simplex either tells us (P2) is unbounded, or it tells us it has an optimal solution. It gives us certificates in either case.

We can derive a certificate of Infeasibilty for (P2) if Phase 1 Simplex has optimal value $< \mathbb{0}$, by using the certificate of optimality Simplex gives us.

Recall: Fundamental Theorem of Linear Programming.

Theorem 8

Given

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

Exactly one of the following holds for the LP:

- it is feasible
- it is unbounded
- it has an optimal solution that is basic.

Remark 18

A finite number of basis implies a finite number of basic solutions.

Example 2.7.6. *For the following NLP:*

$$\begin{aligned} \min \quad & \frac{1}{x} \\ \text{s.t.} \quad & x \geq \mathbb{0} \end{aligned}$$

*This has an optimal value $\mathbb{0}$, but no $x > \mathbb{0}$ has $\frac{1}{x} = \mathbb{0}$. No feasible solution achieves the optimal value. This **does not** happen for LPs.*

2.8 Halfspaces and Convexity

It'll be conveient to work with an LP in **SIF** (Standard Inequality Form)

$$\begin{aligned} \max \quad & c^T \mathbf{x} \\ \text{s.t.} \quad & A\mathbf{x} \le b \\ & \mathbf{x} \ge \mathbb{0} \\ & c, \mathbf{x} \in \mathbb{R}^n,\ b \in \mathbb{R}^m,\ A \in \mathbb{R}^{m \times n} \end{aligned}$$

We can replace $A\mathbf{x} = b$ in SEF by

$$A\mathbf{x} \le b \iff -A\mathbf{x} \le -b$$

to get an LP in SIF. The constraints are

$$\begin{aligned} & A\mathbf{x} \le b \\ & -A\mathbf{x} \le -b \end{aligned}$$

Consider one row of the constriants:

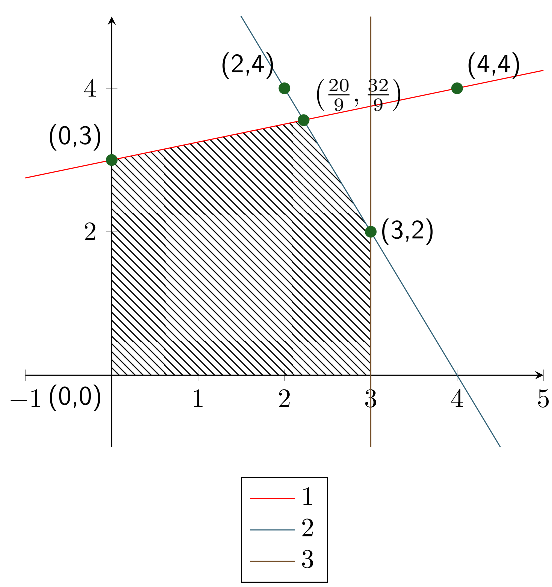
$$a_i^T \cdot \mathbf{x} \le b_i \quad \text{for some } i,\ 1 \le i \le m$$

What set of points H satisfy this inequality?

$$H := \{\mathbf{x} \in \mathbb{R}^n : a_i^T \mathbf{x} \le b_i\}$$

Example 2.8.1. ***LP in SIF***

$$\begin{aligned} \max \quad & (1, -1)\mathbf{x} \\ \text{s.t.} \quad & -x_1 + 4x_2 \le 12 && (1) \\ & 2x_1 + x_2 \le 8 && (2) \\ & x_1 \le 3 && (3) \\ & x_1, x_2 \ge \mathbb{0} && (4,\ 5) \end{aligned}$$

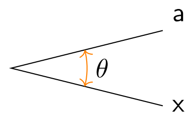


Each constraint defines a half-plane, and the set of feasible solutions, **("feasible regions")** is the intersection of these half-planes. The situation is similar in higher dimensions.

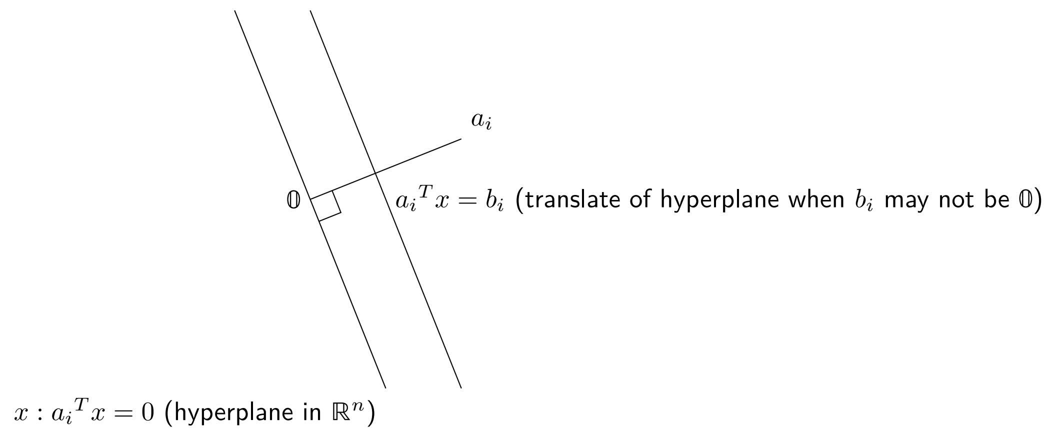
Definition 16: Feasible Region

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

The set of points $H_0 = \{x \in \mathbb{R}^n : {a_i}^T x = b_i\}$ defines a hyperplane, (when it is non-empty/non-trivial). Recap: $a^T x = \|a\| \cdot \|x\| \cdot \cos\theta$



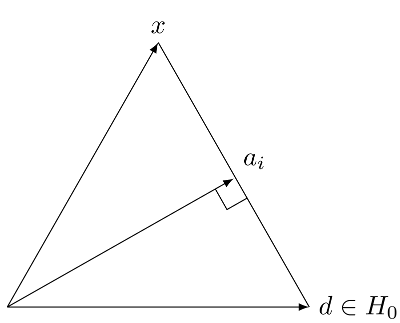
Consider ${a_i}^T x = b_1, a_i \neq \mathbb{0}$, $x \in \mathbb{R}^n$ that satisfy this are

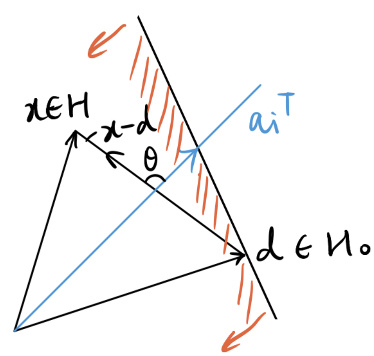


Consider $d \in H_0$:

$$
\begin{aligned}
{a_i}^T x &= b_i \\
{a_i}^T d &= b_i \\
{a_i}^T (x - d) = \mathbb{0} &\iff x - d \perp a_i
\end{aligned}
$$

$H = \{x : {a_i}^T x \leq b_i\}$ is the set of points on one side of H_0, a **"half-space"**.





$$\begin{aligned}
&{a_i}^T(x-d) = \underbrace{{a_i}^T x}_{\le b_i} - \underbrace{{a_i}^T d} = b_i \le \mathbb{0} \\
&\leftrightarrow \underbrace{\|a_i\|}_{>\mathbb{0}} \cdot \underbrace{\|x-d\|}_{>\mathbb{0}} \cdot \cos\theta \le \mathbb{0} \\
&\leftrightarrow \cos\theta \le \mathbb{0} \\
&\leftrightarrow 90^\circ \le \theta \le 270^\circ \qquad (\theta \in [\tfrac{\pi}{2}, \tfrac{3\pi}{2}])
\end{aligned}$$

The set of $x \in \mathbb{R}^n$ that satisfy $Ax \le b$ thus equals the intersection of square number of half-spaces. This is called **polyhedron** (called polytope if it is bounded).

Definition 17: Polyhedron

$P \subset \mathbb{R}^n$ is a polyhedron if there exists a matrix A and a vector b such that

$$P = \{x : Ax \le b\}$$

So the feasible region of an LP (in SIF) is a polyhedron (which is nice properties).

Definition 18: Geometry of Polyhedra

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

1. $\{x : a^T x = \beta\}$ is a hyperplane.
2. $\{x : a^T x \le \beta\}$ is a halfspace.

A hyperplane is the set of solutions to a single linear equation, while the halfspace is the set of solutions to a single linear inequality.

Remark 19

A polyhedron is the intersection of a finite set of halfspaces.

Example 2.8.2. *Suppose vector $a \neq \mathbb{0}, \beta = 0$, then the hyperplane is $H = \{x : a^T x = \beta\}$. The halfspace $F = \{x : a^T x \le \beta\}$.*

Remark 20

1. H is the set of vectors **orthogonal** to a.
2. F is the set of vectors on side of H not containing a.

Definition 19: Translate

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

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

Remark 21

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

$$H := \{x : a^T x = \beta\}, \qquad H_0 := \{x : a^T x = 0\}$$

It follows that H is a translate of H_0.
Let

$$F := \{x : a^T x \le \beta\} \qquad F_0 := \{x : a^T x \le 0\}$$

It follows that F is a translate of F_0.

Theorem 9: Dimension of Hyperplane

The dimension of a hyperplane in $\mathbb{R}^n$ is $n-1$.

Proof. Let $a \in \mathbb{R}^n, a \neq \mathbb{0}$, and let $\beta \in \mathbb{R}$. Define

$$H = \{x : a^T x = \beta\}, \qquad H_0 = \{x : a^T x = 0\}$$

We define the dimension of H to be the dimension of H_0. H_0 is a vector space and its dimension can be computed as

$$\dim(H_0) = n - \operatorname{rank}(a) = n - 1$$

□

Note: a polyhedron has no "dents" and no "holes".

Definition 20

Let $x^{(1)}, x^{(2)} \in \mathbb{R}^n$. The line through $x^{(1)}$ and $x^{(2)}$ is defined as

$$L = \{x = \lambda x^{(1)} + (1-\lambda)x^{(2)} : \lambda \in \mathbb{R}\}$$

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 21: Convexity

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

Remark 22

Polyhedra are convex.

Proof. Suppose a polyhedron P is specified by inequalities $Ax \leq b$. Suppose $a, a' \in P$, $\lambda \in [0,1]$, we have

$$\left.\begin{array}{l} Aa \leq b \implies \lambda Aa \leq \lambda b \\ Aa' \leq b \implies (1-\lambda)Aa' \leq (1-\lambda)b \end{array}\right\} \text{as } \lambda \geq 0, 1-\lambda \geq 0$$
$$A(\underbrace{\lambda a + (1-\lambda)a'}_{\lambda a + (1-\lambda)a' \in P}) \leq b$$

□

The feasible region of an LP is **always a convex set**!

2.9 Extreme Points

Definition 22: Properly Contained

Point $x \in \mathbb{R}^n$ is properly contained in the line segment L if
- $x \in L$ and
- x is distinct from the endpoints of L.

The objective function of an LP is linear:

Let $z(x) := c^T x$, consider two **feasible** solutions, a, a', and a convex combination $\underbrace{\lambda a + (1-\lambda)a}_{d}$, $\lambda \in [0,1]$.

$$
\begin{aligned}
& z(d) = c^T(\lambda a + (1-\lambda)a) = \lambda(c^T a) + (1-\lambda)(c^T a) \le \lambda(c^T a') + (1-\lambda)(c^T a') && \text{(if } c^T a' > c^T a\text{)} \\
\implies & z(d) = \max\{c^T a,\ c^T a'\}
\end{aligned}
$$

So the objective function value is bounded by that at one of the "extreme endpoints" of the same polyhedron (if the LP is bounded).

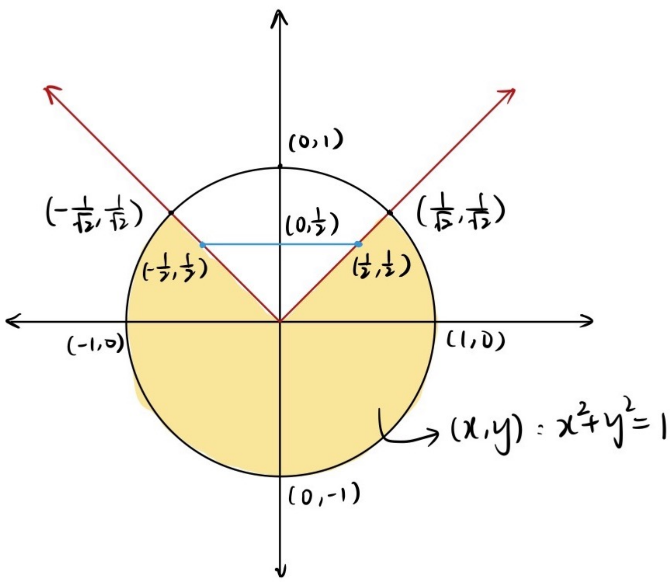
Definition 23: Extreme Point

Given a convex set S, $x \in S$ is called an extreme point if it is not a **non-trivial** convex combination of two distinct points in S, i.e. x is an extreme point iff there is no $\lambda \in (0,1)$ and no two points $a_1, a_1 \in S$, $a_1 \neq a_2$, such that $x = \lambda a_1 + (1-\lambda)a_2$.
Another way to state this definition is that x is NOT an extreme point if there exists a line segment $L \subseteq S$ where L properly contains x.

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

Example 2.9.1. *We have the following examples:*

(i). $T := \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \le 1, y \le |x|\}$. *Is this a convex set?*

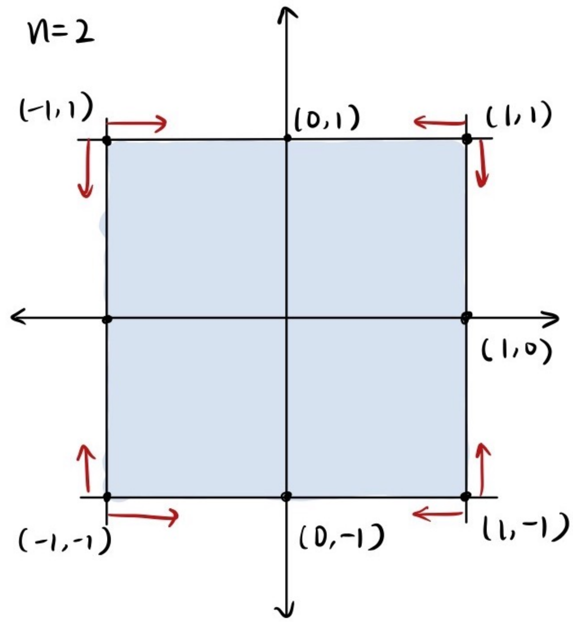


Solution. T is not convex as

$$(0, \frac{1}{2}) = \frac{1}{2}\left((-\frac{1}{2}, \frac{1}{2}) + (\frac{1}{2}, \frac{1}{2})\right)$$

but $(0, \frac{1}{2}) \notin T$. ∎

(ii). Consider $C := \{x \in \mathbb{R}^n : |x_i| \le 1,\ \forall\, i \in \{1, \dots, n\}\}$. *Is* C *convex? If so, what are its extreme points?*



Solution. Yes C is convex!

Here $(1,1),(1,-1),(-1,-1),(-1,1)$ are all the extreme points. ■

We claim that the extreme points of C are: $\{x \in \mathbb{R}^n : x_i \in \{1,-1\},\ \forall\, i = 1,\dots,n\}$.

Proof. We will show that $x \in C$ is NOT an extreme point iff $\exists\, i \in \{1,\dots,n\}$ such that $x_i \in (-1,1)$.

$\Rightarrow$ Suppose $x \in C$ is not an extreme point. There is a $\lambda \in (0,1),\ a,b \in C,\ a \neq b,\ x = \lambda a + (1-\lambda)b$. Suppose a,b differ in jth coordinate: $a_j \neq b_j$, WLOG, let $a_j < b_j$.

$$x_j = \lambda a_j + (1-\lambda)b_j \quad\implies\quad -1 \le a_j < x_j < b_j \le 1$$

Since $\lambda > 0,\ a_j < b_j$ and $1-\lambda > 0$, we have

$$x_j < \lambda b_j + (1-\lambda)b_j = b_j \le 1 \implies x_j < 1$$
$$x_j > \lambda a_j + (1-\lambda)a_j = a_j \ge -1 \implies x_j > -1$$

Thus, $x_j \in (-1,1)$. This proves the forward direction.

$\Leftarrow$ Suppose x has a coordinate $x_j \in (-1,1)$, that is not an extreme point. We'll show that x is a non-trivial convex combination of two points $a,b \in C,\ a \neq b$.

We find λ:

$$x_i = \lambda(-1) + (1-\lambda)(1) = 1 - 2\lambda \implies \lambda = \frac{1-x_i}{2}$$

Since

$$x_i < 1, \lambda > \frac{1-1}{2} = 0$$
$$-x_1 \ge -1, \lambda < \frac{1-(-1)}{2} = 1 \implies \lambda \in (0,1)$$

Define $a,b \in C$: $a_j = b_j = x_j$ for all $j \neq i$, and $x = \lambda a + (1-\lambda)b$. $a_j = -1, b_j = 1 \implies a \neq b$. So x is not an extreme point.

□

Theorem 10

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

Remark 23

The optimum of an LP in SEF, when it exists, is achieved by a basic feasible solution.

The *Simplex* algorithm iterates through basic feasible solution, i.e. extreme points of the feasible region, improving the objective function value, until it finds an optimum (or concludes that the LP is unbounded).

We can find extreme points of LPs in SEF by listing all possible bases and finding the corresponding basic feasible solution. Here is a related method for a general polyhedron:

Suppose $P := \{x \in \mathbb{R}^n : \underbrace{Ax \le b}_{\substack{\text{if derived} \\ \text{form an LP in SIF} \\ \text{it includes the } x \ge 0 \text{ constraints}}} \}$ is a polyhedron.

Definition 24: Tight

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

Example 2.9.2. *Consider Example 2.8.1,*

(i). $(3,0)$*: tight inequalities are* $x_1 \le 3,\ x_2 \ge 0$

(ii). $(\frac{20}{9}, \frac{32}{9})$*:* $2x_1 + x_2 \le 8,\ -x_1 + 4x_2 \le 12$ *are not tight.*

(iii). $(1,1)$*: there are no tight inequalities.*

(iv). $(3,1)$*:* $x_1 \le 3$ *is the only tight constraint.*

Theorem 11

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

1. If $\text{rank}(\bar{A}) = n$, then $\bar{x}$ is an extreme point.
2. If $\text{rank}(\bar{A}) < n$, then $\bar{x}$ is NOT an extreme point.

Note: this means, in particular that there are n tight constriants.

We can use the following remark to prove this theorem:

Remark 24

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

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

then $a = b = c$.

Proof. If we fix c and $a \ne b$, ie. $a > b$, then

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

contradicts the fact that $a = a$ and $a \not< a$. Similarly, if we fix b and $a \ne c$, $a > c$, then

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

□

Proof for the first bullet point of previous theorem:

Proof. Suppose $\bar{x}$ is not an extreme point, then there exists a line segment L_s connecting x_1 and x_2 such that

$$\bar{x} = \lambda x_1 + (1-\lambda)x_2 \to \bar{A}\bar{x} = \bar{b} = \lambda \bar{A}x_1 + (1-\lambda)\bar{A}x_2$$

with some $0 < \lambda < 1$. Since $\bar{A}x_1 \leq \bar{b}$ and $\bar{A}x_2 \leq \bar{b}$ by assumption, with the previous remark, we have $\bar{b} = \bar{A}x_1 = \bar{A}x_2$. If $\operatorname{rank}(A) = n$, then A is non-singular and invertible, implies that there is only one unique solution to $\bar{A}\bar{x} = \bar{b}$. Hence, $\bar{x} = \bar{x_1} = \bar{x_2}$ meaning that $\bar{x}$ is an extreme point, contradicting the assumption. Then $\operatorname{rank}(A) \neq n$.

By its contrapositive, the original statement is then true. $\square$

Module 3

Duality

3.1 Duality through Examples

3.1.1 Shortest Paths

Given a graph $G = (V, E)$, a non-negative length c_e for each edge $e \in E$, and a pair of verticies s and t in V. Our goal is to compute an s,t-path P of smallest total length.

Finding an Intuitive Lower Bound

We will first consider the cardinality special case of the shortest path problem. We consider shortest path instances where

- each edge $e \in E$ has length 1, and
- we are therefore looking for an s,t-path with the smallest number of edges.

Recall:

- If P is an s,t-path and $\delta(U)$ is an s,t-cut, then P contains an edge of $\delta(U)$.
- If $S \subseteq E$ contains an edge from every s,t-cut, then S contains an s,t-path.

Note that $\delta(U_i) \cap \delta(U_j) = \emptyset$ if $i \neq j$ and an s,t-path must contain an edge from $\delta(U_i)$ for all i. If h_i is not in any of the $\delta(U_i)$, then h_i is not on any shortest s,t-path, since an s,t-path that contains h_i must also contain an edge from **each** of the s,t-cuts $\delta(U_i)$.

Back to the General Case. In general instances, we assign a **non-negative width** y_U to every s,t-cut $\delta(U)$.

Definition 25: Width Assignment

A width assignment $\{y_U : \delta(U)\ s,t\text{-cut}\}$ is feasible if, for every edge $e \in E$, the total width of all cuts containing e is no more than c_e.
Using math: y is feasible if for all e

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

Remark 25

If y is a feasible width assignment, then any s,t-path must have length at least $\sum(y_U : \delta(U)\ s,t\text{-cut})$.

Proof. Consider an s, t-path P, it follows that

$$\begin{aligned} c(P) &= \sum(c_e : e \in P) \\ &\geq \sum(\sum(y_u : e \in \delta U) : e \in P) \\ &\geq \sum(y_U : \delta U \ s, t\text{-cut}) \end{aligned}$$

where the last inequality follows from the feasibility of y.

$\square$

3.2 Weak Duality

Example 3.2.1.

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

We want to find a **lower-bound** on the optimal value (objective value). Suppose x is feasible, then x satisfies

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

for $y_1, y_2, y_3 \geq 0$. So, if x is feasible for the LP, it also satisfies

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

for any $y_1, y_2, y_3 \geq 0$, e.g. for $y = (0, 2, 1)^T$, we obtain $(1,3)x \geq 44$. Therefore,

$$\begin{aligned} z(x) &= (2,3)x \\ &\geq (2,3)x + 44 - (1,3)x = 44 + (1,0)x \end{aligned}$$

Since $x \geq \mathbb{0}$, it follows that $z(x) \geq 44$ for every feasible solution x. The optimal value of the LP is in the interval $[44, 49]$ since we have one feasible solution $x = (5, 13)^T$ with objective value 49.

Can we find a better **lower bound** on $z(x)$ for feasible x?

From above, we obtain

$$z(x) \geq (y_1, y_2, y_3) \begin{pmatrix} 20 \\ 18 \\ 8 \end{pmatrix} + \left((2,3) - (y_1, y_2, y_3) \begin{pmatrix} 2 & 1 \\ 1 & 1 \\ -1 & 1 \end{pmatrix} \right) x \tag{3.1}$$

We want the second term to be non-negative. Since $x \geq \mathbb{0}$, this amounts to choose $y \geq \mathbb{0}$ such that

$$(y_1, y_2, y_3) \begin{pmatrix} 2 & 1 \\ 1 & 1 \\ -1 & 1 \end{pmatrix} \leq (2,3)$$

which yields

$$z(x) \geq (y_1, y_2, y_3) \begin{pmatrix} 20 \\ 18 \\ 8 \end{pmatrix}$$

This makes a Linear Program:

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

Solving it gives

$$\bar{y_1} = 0, \bar{y_2} = \frac{5}{2}, \bar{y_3} = \frac{1}{2}$$

and the objective value is 49. There is no feasible solution x to the original LP which has objective value smaller than 49.

Suppose now we are given the LP

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & Ax \geq b \\ & x \geq \mathbb{0} \end{aligned}$$

Any **feasible** solution x must satisfy

$$y^T Ax \geq y^T b$$

for $y \geq \mathbb{0}$, and hence also

$$0 \geq y^T b - y^T Ax$$

If we also know that $A^T y \leq c$ then $x \geq \mathbb{0}$ implies that $z(x) \geq y^T b$. The best lower-bound on $z(x)$ can be found by the following LP:

$$\begin{aligned} \max \quad & b^T y \\ \text{s.t.} \quad & A^T y \leq c \\ & y \geq \mathbb{0} \end{aligned}$$

Definition 26: Dual & Primal

The linear program (D) is called the dual of primal LP (P).

$$\begin{aligned} \max \quad & b^T y & \quad (D) \\ \text{s.t.} \quad & A^T y \leq c & \\ & y \geq \mathbb{0} & \end{aligned} \qquad\qquad \begin{aligned} \min \quad & c^T x & \quad (P) \\ \text{s.t.} \quad & Ax \geq b & \\ & x \geq \mathbb{0} & \end{aligned}$$

Theorem 12: Weak Duality

If $\bar{x}$ is feasible for (P) and $\bar{y}$ is feasible for (D), then $b^T \bar{y} \leq c^T \bar{x}$.

Proof.

$$\begin{aligned} b^T \bar{y} &= \bar{y}^T b \\ &\leq \bar{y}^T (A\bar{x}) && \text{(as } \bar{y} \geq \mathbb{0} \text{ and } b \leq A\bar{x}\text{)} \\ &= (A^T \bar{y})^T \bar{x} \\ &\leq c^T \bar{x} && \text{(as } \bar{x} \geq \mathbb{0} \text{ and } A^T \bar{y} \leq c\text{)} \end{aligned}$$

□

3.2.1 Lowerbounding the Length of s,t-Paths

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

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

Note that the optimal value of the shortest path IP is, at most, the length of a shortest s,t-path.

Also, **dropping the integrality restriction** cannot increase the optimal value (since IP is the special case of LP). The resulting LP is called linear programming relaxation of the IP.

See assignment 6 question 2 for linear relaxation.

Remark 26

The dual of (P) has optimal value no larger than that of (P)!

We can rewrite the shortest-path LP as

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & Ax \geq \mathbb{1} \\ & x \geq \mathbb{0} \end{aligned} \tag{P}$$

where

(i) A has a column for every edge and a row for every s,t-cut $\delta(U)$.

(ii) $A[U, e] = 1$ if $e \in \delta U$ and 0 otherwise.

Its dual is of the form:

$$\begin{aligned} \max \quad & \mathbb{1}^T y \\ \text{s.t.} \quad & A^T y \leq c \\ & y \geq \mathbb{0} \end{aligned} \tag{D}$$

Note that the dual has a constraint for every edge $e \in E$. The left-hand side of this constraint is $\sum(y_U : e \in \delta U)$ and the right-hand size is c_e.

Remark 27

Feasible solutions to (D) correspond precisely to feasible width assignments. Weak Duality implies that $\sum y_U$ is, at most, the length of a shortest s,t-path.

3.3 Shortest Path Algorithm

Shortest Path LP

$$\begin{aligned} \min \quad & \sum(x_e : e \in E) \\ \text{s.t.} \quad & \sum(x_e : e \in \delta(S)) \geq 1,\ (\delta(S) \text{ is } s,t\text{-cut}) \\ & x \geq \mathbb{0} \end{aligned}$$

Shortest Path Dual

$$\begin{aligned} \max \quad & \sum(y_S : \delta(S)\ s,t\text{-cut }) \\ \text{s.t.} \quad & \sum(y_S : e \in \delta(S)) \leq c_e\ (e \in E) \\ & y \geq \mathbb{0} \end{aligned}$$

So far, we know that edges of a graph $G = (V, E)$ are unordered pairs of vertices. Now, we'll introduce arcs - ordered pairs of vertices. We denote an arc from u to v as $\vec{uv}$, and draw it as an arrow from u to v.

Definition 27: Directed Path

A directed path is then a **sequence of arcs**

$$v_1\vec{v}_2, v_2\vec{v}_3, \cdots, v_{k-1}v_k$$

where $v_i\vec{v_{i+1}}$ is an arc in the given graph, and $v_i \neq v_j$ for all $i \neq j$.

Definition 28: Slack

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

$$\text{slack}_y(e) = c_e - \sum(y_U : \delta(U) \ s,t\text{-cut},\ e \in \delta(U))$$

Shortest Paths: Building Duals Incrementally

We start with the trivial dual $y = \mathbb{0}$.

The simplest s,t-cut is $\delta(\{s\})$.
$\longrightarrow$ Increase $y_{\{s\}}$ as much as we can while still maintaining feasibility
$\longrightarrow$ $y_{\{s\}} = 1$

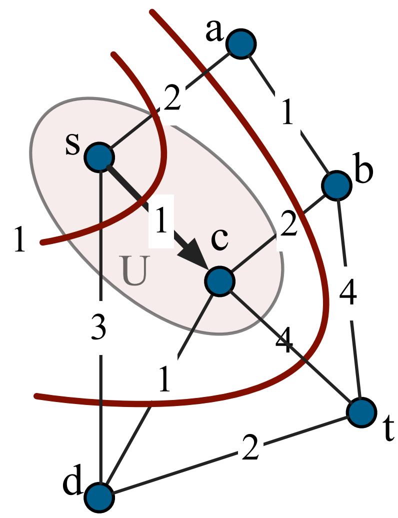
Note: This decreases the slack of sc to 0!
$\longrightarrow$ Replace sc by $\overrightarrow{sc}$

Next we look at all vertices that are reachable from s via directed paths:

$$U = \{s, c\}$$

and consider increasing y_U.

Q: By how much can we increase y_U?



$$\max \quad \sum (y_S \ : \ \delta(S) \ s,t\text{-cut})$$

$$\text{s.t.} \quad \sum (y_S \ : \ e \in \delta(S)) \le c_e \quad (e \in E)$$

$$y \geq \mathbb{0}$$

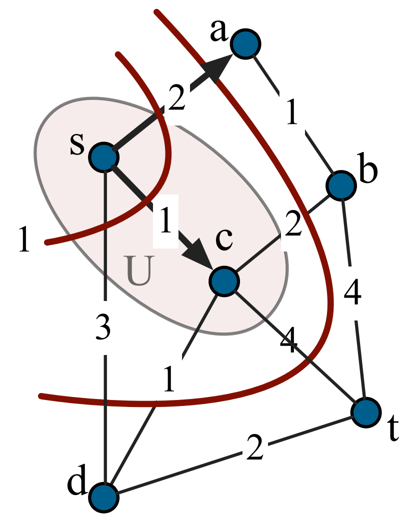
Shortest Paths: Building Duals Incrementally

Q: By how much can we increase y_U?

The maximum increase possible for $y_{\{s,c\}}$ is determined by the slack of edges in $\delta(\{s,c\})$!

$$\begin{aligned}
\mathsf{slack}_y(sa) &= 2-1=1\\
\mathsf{slack}_y(cb) &= 2\\
\mathsf{slack}_y(ct) &= 4\\
\mathsf{slack}_y(cd) &= 1\\
\mathsf{slack}_y(sd) &= 3-1=2
\end{aligned}$$

Edges cd and sa minimize slack. If we pick one arbitrarily, sa for example, we can then set $y_U = \mathsf{slack}_y(sa) = 1$ and convert sa into arc $\overrightarrow{sa}$.



$$\begin{aligned}
\max \quad & \sum (y_S \,:\, \delta(S)\ s,t\text{-cut})\\
\text{s.t.} \quad & \sum (y_S \,:\, e \in \delta(S)) \le c_e \\
& \qquad\qquad (e \in E)\\
& y \ge \mathbb{0}
\end{aligned}$$

Shortest Paths: Building Duals Incrementally

Q: Which vertices are reachable from s via directed paths?

$$U = \{s, a, c\}$$

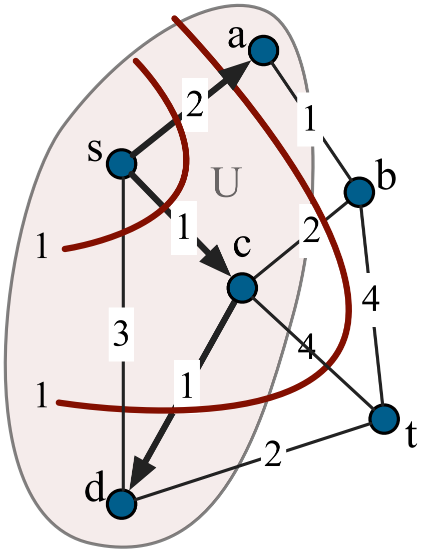
Natural idea: Increase $y_{\{s,a,c\}}$ by as much as we can. How much is this?
$\longrightarrow$ the slack of cd is 0, and hence

$$y_{\{s,a,c\}} = 0$$

Also: we can change cd into $\overrightarrow{cd}$ and let

$$U = \{s, a, c, d\}$$

be the reachable vertices from s.



$$\begin{aligned} \max \quad & \sum (y_S \,:\, \delta(S) \; s,t\text{-cut}) \\ \text{s.t.} \quad & \sum (y_S \,:\, e \in \delta(S)) \le c_e \\ & \qquad (e \in E) \\ & y \ge \mathbb{0} \end{aligned}$$

Shortest Paths: Building Duals Incrementally

The vertices reachable from s by directed paths are in

$$U = \{s, a, c, d\}$$

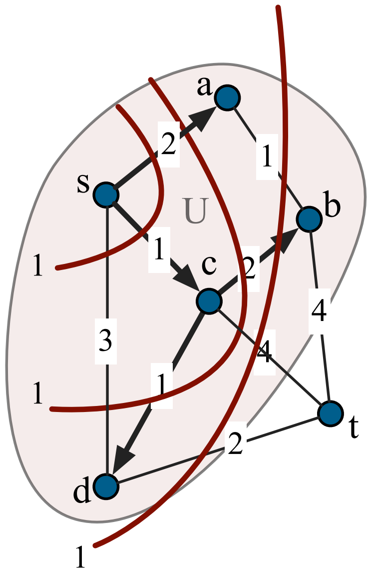
Let us compute the slack of edges in $\delta(U)$.

$$\begin{aligned}
\text{slack}_y(ab) &= 1 \\
\text{slack}_y(cb) &= 2 - 1 = 1 \\
\text{slack}_y(ct) &= 4 - 1 = 3 \\
\text{slack}_y(dt) &= 2
\end{aligned}$$

We let $y_{\{s,a,c,d\}} = 1$, add the equality arc $\overrightarrow{cb}$, and update the set

$$U = \{s, a, b, c, d\}$$

of vertices reachable from s.



$$\begin{aligned}
\max \quad & \sum (y_S : \delta(S) \ s,t\text{-cut}) \\
\text{s.t.} \quad & \sum (y_S : e \in \delta(S)) \le c_e \\
& \qquad (e \in E) \\
& y \ge \mathbb{0}
\end{aligned}$$

Shortest Paths: Building Duals Incrementally

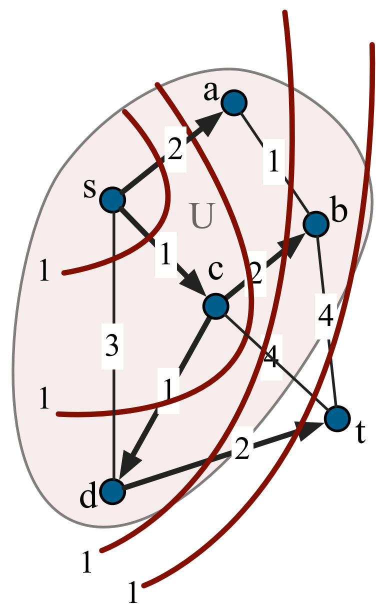
The vertices reachable from s by directed paths are now in

$$U = \{s, a, b, c, d\}$$

Let us compute the slack of edges in $\delta(U)$:

$$\begin{aligned}\text{slack}_y(bt) &= 4 \\ \text{slack}_y(ct) &= 4 - 2 = 2 \\ \text{slack}_y(dt) &= 2 - 1 = 1\end{aligned}$$

We let $y_{\{s,a,b,c,d\}} = 1$ and add the equality arc $\overrightarrow{dt}$.



$$\begin{aligned}\max \quad & \sum (y_S : \delta(S)\ s,t\text{-cut}) \\ \text{s.t.} \quad & \sum (y_S : e \in \delta(S)) \le c_e \quad (e \in E) \\ & y \ge \mathbb{0}\end{aligned}$$

Shortest Paths: Building Duals Incrementally

Note: We now have a directed s,t-path in our graph:

$$P = \overrightarrow{sc}, \overrightarrow{cd}, \overrightarrow{dt},$$

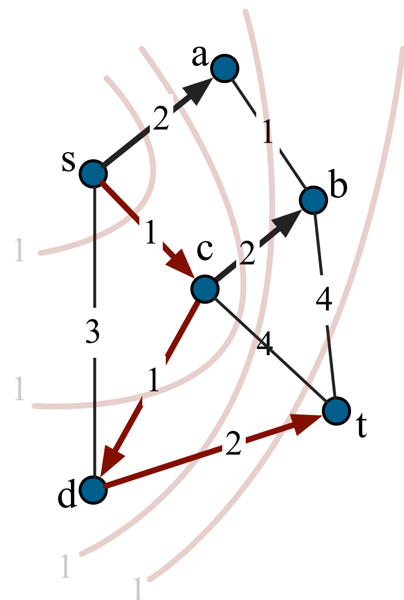
Its length is 4 and its value if 4!

We also have a feasible dual solution:

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

and $y_U = 0$ otherwise.

Therefore, we know that path P is a shortest path!



$$\begin{aligned} \max \quad & \sum(y_S \,:\, \delta(S)\ s,t\text{-cut}) \\ \text{s.t.} \quad & \sum(y_S \,:\, e \in \delta(S)) \le c_e \\ & \qquad\qquad (e \in E) \\ & y \ge \mathbb{0} \end{aligned}$$

Shortest Path Algorithm

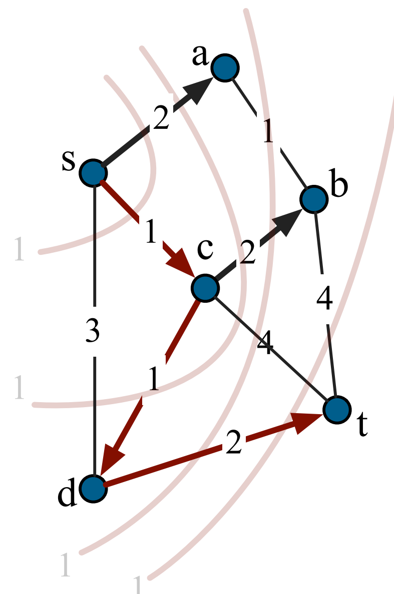
To compute the shortest Path for the instance on the right, we used the following algorithm:

Algorithm 3.2 Shortest path.

Input: Graph $G = (V, E)$, costs $c_e \geq 0$ for all $e \in E$, $s, t \in V$ where $s \neq t$.

Output: A shortest st-path P

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



3.4 Correctness

Recall that the slack of an edge $uv \in E$ for a feasible dual solution y is

$$c_{uv} - \sum(y_U : uv \in \delta(U))$$

We call an edge $uv \in E$ an equality edge if its slack is 0. We also call a cut $\delta(U)$ active for a dual solution y if $y_U > 0$.

Theorem 13

Let y be a feasible dual solution, and P is an s,t-path. P is a shortest path if

- all edges on P are equality edges, and
- every active cut $\delta(U)$ has exactly one edge of P.

To show that the shortest-path algorithm is correct, it suffices to show that

Theorem 14

The Shortest Path Algorithm maintains throughout its execution if

1. y is a feasible dual
2. arcs are equality arcs (i.e. have 0 slack)
3. no active cut $\delta(U)$ has an entering arc: an arc wu with $w \notin U$ and $u \in U$.
4. for every $u \in U$ there is directed s,t-path, and
5. arcs have both ends in U.

Suppose the invariants hold when the algorithm terminates, then

- $t \in U$ and (4) implies that there is a directed s,t-path.
- y is feasible by (1),
- arcs on P are equality arcs by (2).

We want to show that $\delta(U)$ is active $\rightarrow P$ has exactly one edge in $\delta(U)$.

For **contradiction**, suppose $\delta(U)$ active and P has more than one edge in $\delta(U)$. Let e and e' be the first two edges on P that leave $\delta(U)$. Then, there must also be an arc f on P that enters U - since e and e' are both arc leaving U. This contradicts (3).

We now want to prove Theorem 14.

Proof. It is trivial that (1) to (5) holds after Step 1 (initialization). Suppose (1) - (5) hold before Step 3 (find the smallest slak in $\delta(U)$), we will show that they also hold after Step 6 (change edge ab to $\vec{ab}$).

Note that only y_U for the current U changes in step 3-6. y_U arises only on the left-hand sides of constraints for edges in $\delta(U)$. The smallest slack - $c_{uv} - \sum(y_U : uv \in \delta(U))$ - of any of these constraints is precisely the increase in y_U.

$$c_{uv} = \sum(y_U : uv \in \delta(U)) + \text{slack}(uv) = \underbrace{\sum(y'_U : uv \in \delta(U))}_{\text{increased width assignment}}$$

Therefore, y remains feasible! (1) holds.

Also, the constraint of the newly created arc holds with equality after the increase $\rightarrow$ (2) continues to hold and constraints for arcs have slack 0.

The only new active cut created is $\delta(U)$, and then all old arcs have both ends in U. One new arc has tail in U and head outside $U \rightarrow$ (3) holds after Step 6.

The only new arc added is ab and b is added to U at the end of the loop, both (4) and (5) hold.

□

We have already seen that the shortest path algorithm

1. always produces an s, t-path P, and
2. produces a feasible dual solution y

Moreover, the length of P equals the objective value of y, and hence, P must be a shortest s, t-path. Implicitly, we therefore conclude that the shortest path LP always has an optimal integer solution.

Module 4

Duality Theory

Recall the shortest path dual:

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

If (P) is a shortest path LP, then we can rewrite (D) as

$$\begin{aligned} \max \quad & \sum (y_U : s \in U, t \notin U) \\ \text{s.t.} \quad & \sum (y_U : e \in \delta(U)) \leq c_e, \ e \in E \\ & y \geq \mathbb{0} \end{aligned}$$

Using the Weak Duality Theorem, it is equivalent that y is feasible widths and P is an s,t-path $\rightarrow \mathbb{1}^T y \leq c(P)$.

4.1 Weak Duality

In the primal-dual pair

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

- each **non-negative variable** x_e in (P) corresponds to an $\leq$-constraint in (D)
- each $\geq$-constraint in (P) corresponds to a **non-negative variable** y_U in (D).

How can we find the dual LP for every primal LP?

As before,

$$\text{primal variables} \equiv \text{dual constraints}$$
$$\text{primal constraints} \equiv \text{dual variables}$$

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

$(\mathbb{P}_{\max})$			$(\mathbb{P}_{\min})$
	$\leq$ constraint	≥ 0 variable	
max $c^\top x$	$=$ constraint	free variable	min $b^\top y$
subject to	$\geq$ constraint	≤ 0 variable	subject to
Ax **?** b	≥ 0 variable	$\geq$ constraint	$A^\top y$ **?** c
x **?** $\mathbb{0}$	free variable	$=$ constraint	y **?** $\mathbb{0}$
	≤ 0 variable	$\leq$ constraint	

Example 4.1.1.

$$\begin{array}{rl} \max & (1,0,2)x \\ \text{s.t.} & \begin{pmatrix} 3 & -1 & 0 \\ 1 & 0 & 1 \end{pmatrix} x \begin{array}{c} \leq \\ = \end{array} \begin{pmatrix} 3 \\ 4 \end{pmatrix} \\ & x_1, x_2 \geq 0,\ x_3 \textit{ free} \end{array} \tag{P}$$

has its dual LP

$$\begin{array}{rl} \min & (3,4)y \\ \text{s.t.} & \begin{pmatrix} 3 & 1 \\ -1 & 0 \\ 0 & 1 \end{pmatrix} y \begin{array}{c} \geq \\ \geq \\ = \end{array} \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} \\ & y_1 \geq 0,\ y_2 \textit{ free} \end{array} \tag{D}$$

Example 4.1.2.

$$\begin{array}{rl} \min & d^T y \\ \text{s.t.} & W^T y \geq e \\ & y \geq \mathbb{0} \end{array} \tag{P}$$

has its dual LP

$$\begin{array}{rl} \max & e^T x \\ \text{s.t.} & Wx \leq d \\ & x \geq \mathbb{0} \end{array} \tag{D}$$

Example 4.1.3.

$$\begin{array}{rl} \max & (12,26,20)x \\ \text{s.t.} & \begin{pmatrix} 1 & 2 & 1 \\ 4 & 6 & 5 \\ 2 & -1 & -3 \end{pmatrix} x \begin{array}{c} \geq \\ \leq \\ = \end{array} \begin{pmatrix} -2 \\ 2 \\ 13 \end{pmatrix} \\ & x_1 \geq 0,\ x_2 \textit{ free },x_3 \geq 0 \end{array} \tag{P}$$

has its dual LP

$$\begin{array}{rl} \min & (-2,2,13)y \\ \text{s.t.} & \begin{pmatrix} 1 & 4 & 2 \\ 2 & 6 & -1 \\ 1 & 5 & -3 \end{pmatrix} y \begin{array}{c} \geq \\ = \\ \geq \end{array} \begin{pmatrix} 12 \\ 26 \\ 20 \end{pmatrix} \\ & y_1 \leq 0,\ y_2 \geq 0,\ y_3 \textit{ free} \end{array} \tag{D}$$

Theorem 15: Weak Duality Theorem

Let (P_{max}) and (P_{min}) represent the above. If $\bar{x}$ and $\bar{y}$ are feasible for the two LPs, then

$$c^T \bar{x} \leq b^T \bar{Y}$$

If $c^T \bar{x} = b^T \bar{Y}$, then $\bar{x}$ is optimal for (P_{max}) and $\bar{y}$ is optimal for (P_{min}).

We can rewrite the general primal LP and the dual using slack variables.

$$\begin{array}{llcll}
\max & c^T x & \qquad & \min & b^T y \\
\text{s.t.} & Ax + x = b & & \text{s.t.} & A^T y + w = c \\
& s_i \geq 0 \ (i \in R_1) & & & w_j \leq 0 \ (j \in C_1) \\
& s_i \leq 0 \ (i \in R_2) & & & w_j \geq 0 \ (j \in C_2) \\
& s_i = 0 \ (i \in R_3) & & & w_j = 0 \ (j \in C_3) \\
& x_j \geq 0 \ (j \in C_1) & & & y_i \geq 0 \ (i \in R_1) \\
& x_j \leq 0 \ (j \in C_2) & & & y_i \leq 0 \ (i \in R_2) \\
& x_j \text{ free } (j \in C_3) & & & y_i \text{ free } (i \in R_3)
\end{array}$$

Suppose $\bar{x}$ and $\bar{y}$ are feasible for the original primal and dual LPs, let $\bar{s} = b - A\bar{x}$ and $\bar{w} = c - A^T\bar{y}$. It follows that

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

We can show that $\bar{w}^T\bar{x} \leq 0$ and $\bar{y}^T\bar{s} \geq 0 \Rightarrow \bar{y}^T b \geq c^T\bar{x}$.

Since for all $j \in C_1$, $w_j \leq 0$ and $x_j \geq 0$, for all $j \in C_2$, $w_j \geq 0$ and $x_j \leq 0$, and for all $j \in C_3$, $w_j = 0$,

$$\bar{w}^T\bar{x} = \underbrace{\sum_{j \in C_1} \bar{w}_j\bar{x}_j}_{\leq 0} + \underbrace{\sum_{j \in C_2} \bar{w}_j\bar{x}_j}_{\leq 0} + \underbrace{\sum_{j \in C_3} \bar{w}_j\bar{x}_j}_{=0} \leq 0$$

Similarly, for all $i \in R_1$, $s_i \geq 0$ and $y_i \geq 0$, for all $i \in R_2$, $s_i \leq 0$ and $y_i \leq 0$, and for all $i \in R_3$, $s_i = 0$,

$$\bar{y}^T\bar{s} = \underbrace{\sum_{i \in R_1} \bar{s}_i\bar{y}_i}_{\geq 0} + \underbrace{\sum_{i \in R_2} \bar{s}_i\bar{y}_i}_{\geq 0} + \underbrace{\sum_{i \in R_3} \bar{s}_i\bar{y}_i}_{=0} \geq 0$$

The formal proof of Theorem 15:

Proof. There are three cases:

1. (P_{max}) is unbounded $\rightarrow$ (P_{min}) is infeasible.

   Suppose, for its contrapositive, that $\bar{y}$ is feasible for (P_{min}). By Weak Duality, $c^T\bar{x} \leq b^T\bar{y}$ for all $\bar{x}$ feasible for (P_{max}), and hence the latter is bounded.

2. (P_{min}) is unbounded $\rightarrow$ (P_{max}) is infeasible.

   Similar to 1.

3. (P_{max}) and (P_{min}) are feasible $\rightarrow$ both must have optimal solutions.

   By Weak duality, both are bounded, and by Foundamental Theorem of LP, both must have optimal solution!

□

4.2 Strong Duality

Can we always find feasible solutions $\bar{x}$ and $\bar{y}$ to a primal-dual pair, such that $c^T\bar{x} = b^T\bar{y}$?

Theorem 16: Strong Duality Theorem

If (P_{max}) has an optimal solution $\bar{x}$, then (P_{min}) has an optimal solution $\bar{y}$ such that $c^T\bar{x} = b^T\bar{y}$.

We can prove Strong Duality Theorem in the special case when (P)=(P_{max}) is in SEF.

$$\begin{aligned} \max \quad & c^T x & \quad (P) \qquad \qquad \min \quad & b^T y & \quad (D) \\ \text{s.t.} \quad & Ax = b & \text{s.t.} \quad & A^T y \geq c & \\ & x \geq \mathbb{0} & & & \end{aligned}$$

Assume (P) has an optimal solution, 2-Phase Simplex terminates with an optimal basis B.

We can rewrite (P) for basis B:

$$\begin{aligned} \max \quad & z = \bar{y}^T b + \bar{c}^T x \\ \text{s.t.} \quad & x_B + A_B^{-1} A_N x_N = A_B^{-1} b \\ & x \geq \mathbb{0} \end{aligned} \tag{P'}$$

where $\bar{y} = A_B^{-T} c_B$ and $\bar{c}^T = c^T - \bar{y}^T A$. Thus, $\bar{x}_N = \mathbb{0}$ and $\bar{x}_B = A_B^{-1} b$. Recall that P and P' are equivalent, $\bar{x}$ has the same objective value in P and P'.

$$\begin{aligned} c^T \bar{x} &= \bar{y}^T b + \bar{c}^T \bar{x} \\ &= \bar{y}^T b + \bar{c}_N^T \bar{x}_N \\ &= b^T \bar{y} \end{aligned}$$

and we can show that $\bar{y}$ is dual feasible.

B is an optimal basis $\rightarrow \bar{c} \leq \mathbb{0}$, $c^T - \bar{y}^T A \leq \mathbb{0}$. Equivalently, $A^T \bar{y} \geq c$, meaning $\bar{y}$ is dual feasible.

Note: (P) is feasible and (D) is feasible $\implies$ (P) cannot be unbounded. By Foundamental Theorem of LP, (P) has an optimal solution.

Subtly different version via previous results:

Theorem 17: Strong Duality Theorem - Feasibility Version

Let (P) and (D) be primal-dual pair of LPs, if both are feasible, then both have optimal solutions of the same objective value.

(D)\ (P)	optimal solution	unbounded	infeasible
optimal solution	possible (1)	impossible (2)	impossible (3)
unbounded	impossible (4)	impossible (5)	possible (6)
infeasible	impossible (7)	possible (8)	possible (9)

- (1), (6), (8) many examples exists
- (2) follows directly from Weak Duality as follows:

  Suppose, for contradiction, that (D) has an optimal solution $\bar{y}$, $c^T \bar{x} \leq b^T \bar{y}$ for all feasible primal solutions $\bar{x}$ by Weak Duality, then (P) is bounded. Similar arguments apply to (4) and (5).
- (3), (7) follow directly from Strong Duality.

4.3 Geometric Optimality

We know that the feasible region of an LP is a polyhedron, and basic soltions corresponds to the extreme points of this polyhedron. When is an extreme point optimal?

We can rewrite (P) using **slack variables** s:

$$\begin{aligned} \max \quad & c^T x \\ \text{s.t.} \quad & Ax + s = b \\ & s \geq \mathbb{0} \end{aligned} \tag{P'}$$

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

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

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

and Strong Duality tells us that

$$\begin{aligned} \bar{x}, \bar{y} \text{ both optimal } &\Leftrightarrow c^T \bar{x} = \bar{y}^T b \\ &\Leftrightarrow \bar{y}^T \bar{s} = 0 \end{aligned} \tag{*}$$

By feasibility, $\bar{x} \geq \mathbb{0}$ and $\bar{s} \geq \mathbb{0}$, hence (*) holds if and only if $\bar{y}_i = 0$ or $\bar{s}_i = 0$ for every $1 \leq i \leq m$.

Theorem 18: Complementary Slackness - Special Case

Let $\bar{x}$ and $\bar{y}$ be feasible for (P) and (D),

$$\begin{aligned} \max \quad & c^T x \qquad (P) \\ \text{s.t.} \quad & Ax \leq b \end{aligned} \qquad\qquad \begin{aligned} \min \quad & b^T y \qquad (D) \\ \text{s.t.} \quad & A^T y = c \\ & y \geq \mathbb{0} \end{aligned}$$

Then $\bar{x}$ and $\bar{y}$ are optimal if and only if
- $\bar{y}_i = 0$, or
- the ith constraint of (P) is **tight** for $\bar{x}$

for every row index i.

Theorem 19: Complementary Slackness

Feasible solutions $\bar{x}$ and $\bar{y}$ for (P) and (D) are optimal if and only if $\bar{y}_i = 0$ or the ith primal constraint is tight for $\bar{x}$ for all row indices i.

Example 4.3.1. *Consider the following LP:*

$$\begin{aligned} \max \quad & (5, 3, 5)x \\ \text{s.t.} \quad & \begin{pmatrix} 1 & 2 & -1 \\ 3 & 1 & 2 \\ -1 & 1 & 1 \end{pmatrix} x \leq \begin{pmatrix} 2 \\ 4 \\ -1 \end{pmatrix} \end{aligned} \tag{P}$$

Its dual is

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

We have $\bar{x} = (1, -1, 1)^T$ and $\bar{y} = (0, 2, 1)^T$. It is easy to check if $\bar{x}$ and $\bar{y}$ are feasible.

- $\bar{y}_1 = 0$ *or* $(1, 2, -1)\bar{x} = 2$
- $\bar{y}_2 = 0$ *or* $(3, 1, 2)\bar{x} = 4$
- $\bar{y}_3 = 0$ *or* $(-1, 1, 1)\bar{x} = -1$

$\Rightarrow$ *$\bar{x}$ and $\bar{y}$ are optimal!*

$\bar{x}$ and $\bar{y}$ satisfy the **complementary slackness conditions** if

for all variables x_j of (P_{max}):

- $\bar{x}_j = 0$ or
- jth constraint of (P_{min}) is satisfied with equality for $\bar{y}$

for all variables y_i of (P_{min}):

- $\bar{y}_i = 0$ or
- ith constraint of (P_{max}) is satisfied with equality for $\bar{x}$

The two or's above are inclusive!

Theorem 20: Complementary Slackness Theorem

Let (P) and (D) be an arbitrary primal-dual pair of LPs, and let $\bar{x}$ and $\bar{y}$ be feasible solutions. Then these solutions are optimal if and only if CS conditions hold.

4.3.1 Cones of Vectors

Definition 29: Cone of Vectors

Let $a^{(1)}, \dots, a^{(k)}$ be vectors in $\mathbb{R}^n$. 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 rows of tight constraints.

Theorem 21

Let $\bar{x}$ be a feasible solution to

$$\max\{c^T x : Ax \leq b\}$$

Then $\bar{x}$ is optimal if and only if c is in the cone of tight constraints for $\bar{x}$.

Example 4.3.2. *Consider the LP*

$$\max\left\{\left(\frac{3}{2}, \frac{1}{2}\right) x : x \in P\right\} \tag{*}$$

where

$$P = \left\{x \in \mathbb{R}^2 : \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{pmatrix} x \leq \begin{pmatrix} 2 \\ 3 \\ 2 \end{pmatrix}\right\}$$

Tight constriants at $\bar{x} = (2,1)^T$:

$$(1,0)\bar{x} = 2 \tag{1}$$
$$(1,1)\bar{x} = 3 \tag{2}$$

Note that $c = (3/2, 1/2)^T$ is in the cone of tight constraints as

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

Proving the if direction of the above theorem amouns to

- finding a feasible solution $\bar{y}$ to the dual of (*) and
- showing that $\bar{x}$ and $\bar{y}$ satisfy the CS conditions

Proof. Suppose $\bar{x}$ is a solution to (P), and let $J(\bar{x})$ be the indicies of **tight constraints** for $\bar{x}$, ie.

$$\mathrm{Row}_i(A)\bar{x} = b_i$$

for $i \in J(\bar{x})$ and

$$\mathrm{Row}_i(A)\bar{x} < b_i$$

for $i \notin J(\bar{x})$.

Suppose c is in the cone of tight constraints at $\bar{x}$, and thus

$$c = \sum_{i \in J(\bar{x})} \lambda_i \mathrm{Row}_i(A)^T = A^T \bar{y}$$

for some $\lambda \geq \mathbb{0}$, where we define

$$\bar{y}_i = \begin{cases} \lambda_i & i \in J(\bar{x}) \\ 0 & \text{otherwise} \end{cases}$$

Also note that $\bar{y}_i > 0$ only if $\mathrm{Row}_i(A)\bar{x} = b_i \Rightarrow$ CS conditions (*) hold!

Hence, by CS theorem, $(\bar{x}, \bar{y})$ is then optimal. $\square$

Module 5

Integer Programs

5.1 IP vs. LP (Convex Hulls)

LINEAR PROGRAMMING	INTEGER PROGRAMMING
Can solve very large instances	Some small instances cannot be solved
Algorithms exist that are guaranteed to be fast	No fast algorithm exists
Short certificate of infeasibility (Farka's Lemma)	Does not always exist
Short certificate of optimality (Strong Duality)	Does not always exist
The only possible outcomes are infeasible, unbounded, or optimal	Can have other outcomes

Example 5.1.1. *Consider the following IP:*

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

It is feasible, bounded, and has no optimal solution.

Proof. Suppose, for a contradiction, there exists optimal x_1, x_2, let

$$x_1' = 2x_1 + x_2 \qquad x_2' = x_1 + 2x_2$$

Claim: x_1', x_2' are feasible. Since $x_1' = 2x_1 + 2x_1 \geq 1$ and $x_2' = x_1 + 2x_2 \geq 1$,

$$\begin{aligned} & x_1' \overset{?}{\leq} \sqrt{2}x_2' \\ \Leftrightarrow\ & 2x_1 + 2x_2 \overset{?}{\leq} \sqrt{2}(x_1 + 2x_2) = \sqrt{x_1} + 2\sqrt{2}x_2 \\ \Leftrightarrow\ & x_1(2 - \sqrt{2}) \overset{?}{\leq} (2\sqrt{2} - 2)x_2 \\ \Leftrightarrow\ & x_1 \overset{?}{\leq} \frac{2\sqrt{2} - 2}{2 - \sqrt{2}}x_2 = 2\sqrt{2}x_2 \end{aligned}$$

Claim: $x_1' - \sqrt{2}x_2' > x_1 - \sqrt{2}x_2$

$$(2x_1 + 2x_2) - \sqrt{2}(x_1 + 2x_2) \overset{?}{>} x_1 - \sqrt{2}x_2$$

Simplifying, we obtain

$$\sqrt{2}x_2 \overset{?}{>} x_1$$

- $\geq$ since x_1, x_2 are feasible for (P)
- $>$ otherwise $\sqrt{2} = \frac{x_1}{x_2}$ but $\sqrt{2}$ is not rational number

□

Remark 28

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

Definition 30: Convex Hull

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

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

Theorem 22: Meyer's 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.

Remark 29

The condition that all entries of A and b are rational numbers cannot be excluded from the hypothesis.

Let A, b be rational,

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

The convex hull of all feasible solutions of (IP) is a polyhedron $\{x : A'x \leq b\}$:

$$\max\{c^T x : A'x \leq b', x \in \mathbb{Z}\} \tag{LP}$$

Theorem 23

- (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)

Conceptual way of solving (IP):

Step 1 Compute A', b'

Step 2 Use Simplex to find an extreme optimal solution to (LP)

Note that this is NOT a practical way to solve an LP, since we do not know how to compute A', b', and A', b' can be MUCH MORE complicated than A, b.

5.2 Cutting Planes

Definition 31: Cutting Plane

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

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

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

Example 5.2.1. *Consider the IP:*

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

Using Simplex, we can find that $\bar{x} = \left(\frac{8}{3}, \frac{4}{3}\right)$ is optimal, but they are not integers.

A cutting plane for this IP is

$$x_1 + 3x_2 \leq 6 \tag{*}$$

After adding (*) to our relaxation, we get

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

Using Simplex, we get $x' = (3,1)^T$ is optimal, and this is the optimal solution for IP.

Algorithm 2: Cutting Plane Scheme

Input : (IP)$=\max\{c^T x : Ax \leq b, x \in \mathbb{Z}\}$

```
repeat
    Let (P) denote max{c^T x : Ax ≤ b} (integer program relaxation)
    if (P) is infeasible then
        return (IP) is also infeasible
    x̄ ← optimal solution to (P)
    if x̄ is integral then
        return x̄ is also optimal for (IP)
    Finding a cutting plane a^T x ≤ β for x̄
    Add a constraint a^T x ≤ β to the system Ax ≤ b
until
```

We use Simplex to find the cutting plane.

Solve the relaxation and get the LP in a canonical form for B:

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

where

$$N = \{j : j \notin B\}$$
$$\bar{x} \text{ basic } (\bar{x}_N = \mathbb{0}, \bar{x}_B = b)$$
$$r(i) \text{ index of } i^{th} \text{ basic variable}$$

Suppose $\bar{x}$ is not integer, then b_i is fractional for some value i. We know that every feasible solution to the LP relaxation satisfies

$$x_{r(i)} + \sum_{j \in N} A_{ij} x_j = b_i \Rightarrow \underbrace{x_{r(i)} + \sum_{j \in N} \lfloor A_{ij} \rfloor x_j}_{\text{integer for all } x \text{ integer}} \leq b_i$$

Hence, every feasible solution to IP satisfies

$$x_{r(i)} + \sum_{j \in N} \lfloor A_{ij} \rfloor x_j \leq \lfloor b_i \rfloor \tag{*}$$

However, $\bar{x}$ does not satisfy this as

$$\underbrace{x_{r(i)}}_{b_i} + \sum_{j \in N} \lfloor A_{ij} \rfloor \underbrace{x_j}_{=0} = b_i > \lfloor b_i \rfloor$$

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

Module 6

Nonlinear Programs

6.1 Convexity

Definition 32: NLP

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

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

where
$f \;\; \mathbb{R}^n \to \mathbb{R}$, and
$g_i \;\; \mathbb{R}^n \to \mathbb{R}$ for $i = 1, \ldots, k$

Remark 30

There aren't any restrictions regarding the type of functions.

This is a very general model, but NLP can be very hard to solve.

Remark 31

We may assume $f(x)$ is a linear function, ie. $f(x) = c^T x$

We can rewrite (P) as

$$\begin{aligned} \min \quad & \lambda \\ \text{s.t.} \quad & \lambda \geq f(x) \\ & g_i(x) \leq 0 \; (i = 1, \ldots, k) \end{aligned} \tag{Q}$$

The optimal solution to (Q) will have $\lambda = f(x)$.

Example 6.1.1.

$$\begin{aligned} \max \quad & x_1 + x_2 \\ \text{s.t.} \quad & 2x_1 - x_2 \geq 3 \\ & x_1 - x_2 = 4 \\ & x_1, x_2 \geq 0 \end{aligned}$$

$$\begin{aligned} \min \quad & -x_1 - x_2 \\ \text{s.t.} \quad & -2x_1 + x_2 + 3 \leq 0 \\ & x_1 - x_2 - 4 \leq 0 \\ & -x_1 + x_2 + 4 \leq 0 \\ & -x_1 \leq 0 \\ & -x_2 \leq 0 \end{aligned}$$

Nonlinear programs an also generalize integer programs.

Example 6.1.2. *We have the* $0, 1$ *IP:*

$$\begin{aligned} \max \quad & c^T x \\ \text{s.t.} \quad & Ax \le b \\ & x_j \in \{0,1\} \ (j = 1, \ldots, n) \end{aligned}$$

The idea is

$$x_j \in \{0,1\} \Leftrightarrow x_j(1 - x_j) = 0$$

and we have the quadratic NLP:

$$\begin{aligned} \min \quad & -c^T x \\ \text{s.t.} \quad & Ax \le b \\ & x_j(1 - x_j) \le 0 \ (j = 1, \ldots, n) \\ & -x_j(1 - x_j) \le 0 \ (j = 1, \ldots, n) \end{aligned}$$

Note that $0, 1$ *IPs are hard to solve, thus, quadratic NLP are also hard to solve.*

Example 6.1.3. *We have the pure IP:*

$$\begin{aligned} \max \quad & c^T x \\ \text{s.t.} \quad & Ax \le b \\ & x_j \in \mathbb{Z} \ (j = 1, \ldots, n) \end{aligned}$$

The idea is

$$x_j \in \mathbb{Z} \Leftrightarrow \sin(\pi x) = 0$$

and we have the NLP:

$$\begin{aligned} \min \quad & -c^T x \\ \text{s.t.} \quad & Ax \le b \\ & \sin(\pi x) = 0 \ (j = 1, \ldots, n) \end{aligned}$$

IPs are hard to solve, so NLPs are hard to solve.

Definition 33: Local Optimum

Consider

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

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

$$\forall\ x' \in S,\ \|x' - x\| \le \delta$$

and we have $f(x) \le f(x')$.

Remark 32

Consider

$$\min\{c^T x : x \in S\} \tag{P}$$

If S is a convex and x is a local optimum, then x is optimal.

Proof. Suppose $\exists\ x' \in S$ with $c^T x' < c^T x$, let $y = \lambda x' + (1 - \lambda)x$ for $\lambda > 0$ small. Since S is a convex, $y \in S$, as λ

small $\|y - x\| \leq \delta$,

$$\begin{aligned} c^T y &= c^T(\lambda x' + (1-\lambda)x) \\ &= \underbrace{\lambda}_{\geq 0} \underbrace{c^T x'}_{< c^T x} + \underbrace{(1-\lambda)}_{\geq 0} c^T x \\ &< \lambda c^T x + (1-\lambda) c^T x \\ &= c^T x \end{aligned}$$

This is a contradiction.

□

We want to study the cases where feasible region of (P) is **convex**.

Definition 34: Convex

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 6.1.4. *We claim that $f(x) = x^2$ is convex. Pick $a, b \in \mathbb{R}$ and pick λ where $0 \leq \lambda \leq 1$. We check that*

$$\lambda(1-\lambda)2ab - [\lambda(1-\lambda)(a^2+b^2)] = -\lambda(1-\lambda)(a-b)^2 < 0$$

since $\lambda, (1-\lambda) > 0$ and $(a-b)^2 \geq 0$. Hence,

$$[\lambda a + (1-\lambda)b]^2 \leq \lambda a^2 + (1-\lambda) b^2$$

Remark 33: Convex Set

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 $a, b \in S$, and λ where $0 \leq \lambda \leq 1$. Let $x = \lambda a + (1-\lambda)b$, our goal si to show that $x \in S$, that $g(x) \leq \beta$.

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

□

Remark 34

Suppose

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & g_i(x) \leq 0 \ (i = 1, \ldots, k) \end{aligned} \tag{P}$$

If all functions g_i are convex, then the feasible region of (P) is convex.

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

Definition 35: Epigraph

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

$$\text{epi}(f) = \left\{ \begin{pmatrix} y \\ x \end{pmatrix} : y \geq f(x), x \in \mathbb{R}^n \right\} \subseteq \mathbb{R}^{n+1}$$

Remark 35

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

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

6.2 The KKT Theorem

How can we prove a feasible solution $\bar{x}$ is optimal to the NLP?

Step 1 Find a relaxation of the NLP.

Step 2 Prove $\bar{x}$ is optimal for the relaxation.

Step 3 Deduce that $\bar{x}$ is optimal for the NLP.

Example 6.2.1. ***Claim:*** *$\bar{x} = (1,1)^T$ is an optimal solution to*

$$\begin{aligned} \min \quad & -x_1 - x_2 \\ \text{s.t.} \quad & -x_2 + x_1^2 \leq 0 && (1) \\ & -x_1 + x_2^2 \leq 0 && (2) \\ & -x_1 + \frac{1}{2} \leq 0 && (3) \end{aligned}$$

Proof. Tight constraints for $\bar{x}$ are (a) and (b). **Goal:** show that the objective function is in the cone of tight constraints.

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} \overset{?}{\in} \text{cone}\left\{ \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \end{pmatrix} \right\} \Leftarrow \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 1 \times \begin{pmatrix} 2 \\ -1 \end{pmatrix} + 1 \times \begin{pmatrix} -1 \\ 2 \end{pmatrix}$$

Original NLP:

$$\begin{aligned} \min \quad & -x_1 - x_2 \\ \text{s.t.} \quad & -x_2 + x_1^2 \leq 0 && (1) \\ & -x_1 + x_2^2 \leq 0 && (2) \\ & -x_2 + \frac{1}{2} \leq 0 && (3) \end{aligned}$$

Relaxation (we'll show why this is the relaxation later):

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

It is clear that $\bar{x} = (1,1)^T$ is an optimal solution to the relaxation. $\implies$ $\bar{x}$ is an optimal solution to the original NLP. $\square$

We use **subgradients** in general to solve this kind of problem.

Definition 36: Subgradient

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

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

Example 6.2.2. *Consider the NLP in 6.2.1, $f : \mathbb{R}^2 \to \mathbb{R}$ where $f(x) = -x_1 + x_2^2$ and $\bar{x} = (1,1)^T$. We claim that $(-1,2)^T$ is a subgradient of f at $\bar{x}$.*

$$h(x) = f(\bar{x}) + s^T(x - \bar{x}) = 0 + (-1,2)(x - (1,1)^T) = -(x_1 - 1) + 2(x_2 - 1) = -x_1 + 2x_2 - 1$$

Check: $h(x) \leq f(x)$ *for all* $x \in \mathbb{R}^n$.

$$-x_1 + 2x_2 - 1 \overset{?}{\leq} -x_1 + x_2^2 \Leftrightarrow x_2^2 - 2x_2 + 1 \overset{?}{\geq} 0$$

which is the case as $x_2^2 - 2x_2 + 1 = (x_2 - 1)^2 \geq 0$.

Definition 37: Supporting set

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

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

Remark 36

let $g : \mathbb{R}^n \to \mathbb{R}$ be convex and let $\bar{x}$ where $g(\bar{x}) = 0$. Let s be a subgradient of g at $\bar{x}$. Let $C = \{x : g(x) \leq 0\}$, $F = \{x : h(x) := g(\bar{x}) + s^T(x - \bar{x}) \leq 0\}$. 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 a affine function, and
- $h(\bar{x}) = g(\bar{x}) = 0$; thus, $\bar{x}$ is on the boundary of F.

Proof. **Claim:** $C \subseteq F$. Let $x \in C$ and thus $g(x) \leq 0$. By definition of a subgradient, we know that $h(x) \leq g(x)$. It follows that $h(x) \leq g(x) \leq 0$. Hence, $x \in F$.

Claim: $h(\bar{x}) = 0$. $h(\bar{x}) = g(\bar{x}) = 0$. □

Example 6.2.3. *Continue from 6.2.2. Let* $g(x) = x_2^2 - x_1$, $\bar{x} = (1,1)^T$, *and* $s = (-1,2)^T$ *is a subgradient at* $\bar{x}$.

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

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

We can use this to construct relaxations of NLPs. Given constriant $g_i(x) \leq 0$, if we replace the nonlinear contraint by the linear constraint $h(x) = g_i(\bar{x}) + s^T(x - \bar{x}) \leq 0$, we get a relaxation.

Theorem 24

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & g_i(x) \leq 0 \quad (i = 1, \ldots, k) \end{aligned}$$

- $g_1, \ldots, g_k$ 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 $-c \in \text{cone}\{s^{(i)} : i \in I\}$, then $\bar{x}$ is optimal.

Proof.

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & g_i(x) \leq 0 \quad (i \in I) \end{aligned}$$

We proved that the set of solutions to $g_i(x) \leq 0$ is contained in the set of solutions to $g_i(\bar{x}) + s^{(i)}(x - \bar{x}) \leq 0$, which can be rewritten as $s^{(i)}x \leq s^{(i)}\bar{x} - g_i(\bar{x})$.

We then have a relaxation

$$\begin{array}{ll} \max & -c^T x \\ \text{s.t.} & s^{(i)} x \le s^{(i)} \bar{x} - g_i(\bar{x}) \quad (i \in I) \end{array}$$

Then, $\bar{x}$ is optimal for the relaxation if $-c \in \text{cone}\{s^{(i)} : i \in I\}$. This means that $\bar{x}$ is also optimal for the NLP. □

Example 6.2.4. *Consider the NLP in 6.2.1, we know that $\bar{x} = (1,1)^T$ is feasible, $I = \{1,2\}$ (where $g_i(\bar{x}) = 0$). $(2,-1)^T$ is subgradient for g_i at $\bar{x}$. $(-1,2)^T$ is subgradient for g_2 at $\bar{x}$.*

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

Theorem 25

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 $j = 1, \ldots, n$, then the gradient $\nabla f(\bar{x})$ is obtained by evaluating for $\bar{x}$,

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

Example 6.2.5. *Computing the gradient of the convex function $f(x) = -x_2 + x_1^2$ at $\bar{x} = (-1,-1)^T$, we have*

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

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

Definition 38: Slater Point

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

$$\begin{array}{ll} \min & c^T x \\ \text{s.t.} & g_i(x) \le 0 \quad (i = 1, \ldots, k) \end{array}$$

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

In 6.2.1, $\bar{x} = \left(\frac{3}{4}, \frac{3}{4}\right)^T$ is a slater point.

Theorem 26: The Karush-Kuhn-Tucker (KKT) Theorem

Consider the following NLP:

$$\begin{array}{ll} \min & c^T x \\ \text{s.t.} & g_i(x) \le 0 \quad (i = 1, \ldots, k) \end{array}$$

Suppose that
1. $g_1, \ldots, g_k$ are all convex
2. there exists a slater point
3. $\bar{x}$ is a feasible solution
4. I is the set of indicies i for which $g_i(\bar{x}) = 0$, and
5. $\forall\ i \in I$, there exists a gradient $\nabla g_i(\bar{x})$ of g_i at $\bar{x}$.

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