CO 250 Spring 2022 Week 1 (Monday May 2)

Class outline.

1. Starting example
2. What is optimization?
3. What are linear programs?
4. Linear programming formulations
5. Integer programming formulations

1 Starting example

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

Variables: Let x_1 be the amount of apples we stock in kg. Let x_2 be the amount of bananas we stock in kg.

What are we looking for?

Objective function: maximize $2x_1 + 1.5x_2$. (This is what we are optimizing)

Constraints: (What limits our variable values?)

1. Weight constraints: $x_1 + x_2 \leq 75$
2. Space constraints $0.08x_1 + 0.05x_2 \leq 4$
3. Non-nongetivity constriaints: $x_1, x_2 \geq 0$

Full linear program:

$$\begin{array}{rl} \max & 2x_1 + 1.5x_2 \\ \text{subject to} & x_1 + x_2 \leq 75 \\ & 0.08x_1 + 0.05x_2 \leq 4 \\ & x_1, x_2 \geq 0 \end{array}$$

Possible solutions:
$x_1 = 30, x_2 = 20$ satisfy all constraints, this is a "feasible solution"
Total sales value is $2 \cdot 30 + 1.5 \cdot 20 = 90$, objective value is 90.

Another solution: $x_1 = 50, x_2 = 0$, feasible, objective value is 100.

Another solution: $x_1 = 8\frac{1}{3}, x_2 = 66\frac{2}{3}$, feasible, objective value is $116\frac{2}{3}$.
We claim without proof, that this is the best possible, optimal solution.

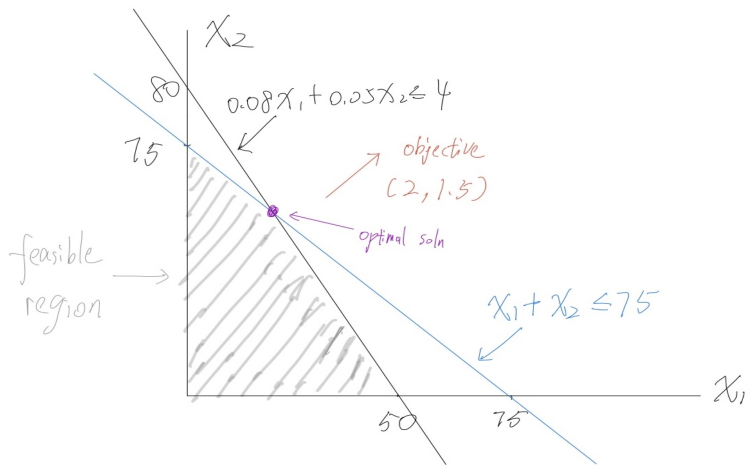


Figure 1: linear program example

2 What is optimization?

Abstractly, optimization problems have this form: We are given $A \subseteq \mathbb{R}^n$ and a function $f : A \to \mathbb{R}$. The goal is to find $x \in A$ that miminizes or maximizes $f(x)$.

Rough idea: we have decisions to make They are possibilities of decisions that are viable. Among them, find the decisions that are best for us.

Optimization problems are applicable and widely used, but often very difficult.

Three components of an optimization problem (constrained optimization):

1. Decision variables
2. Objective function: $\max f(x)$ or $\min f(x)$
3. Constraints: generally $g_i(x) \leq b_i,\ b_i \in \mathbb{R}$.

3 What are linear programs (LPs)?

Linear programs (LPs) is a class of optimization problems that can be efficiently solved. Using variables $x_1, \cdots, x_n$, a function is <u>affine</u> if it has the form $a_1x_1 + \cdots a_nx_n + b$ for some constants $a_1, \cdots a_n, b$. It is called <u>linear</u> if in addition, $b = 0$. In vector form, if $x = (x_1, \cdots, x_n)^T$ and $a = (a_1, \cdots, a_n)^T$, then $a^Tx + b$ is affine, and a^Tx is linear.

A linear program has objective function $\max f(x)$ or $\min f(x)$ where $f(x)$ is affine, and constraints of the form $g(x) \leq b, g(x) = b, g(x) \geq b$ for some linear function g and constant b.

Example: previous LP can be written as following:

$$\begin{aligned} \max \quad & (2, 1.5)(x_1, x_2)^T \\ \text{subject to} \quad & \begin{pmatrix} 1 & 1 \\ 0.08 & 0.05 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \leq \begin{pmatrix} 75 \\ 4 \end{pmatrix} \\ & (x_1, x_2) \geq 0 \end{aligned}$$

Let $c = \begin{pmatrix} 2 \\ 1.5 \end{pmatrix}$, $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, $A = \begin{pmatrix} 1 & 1 \\ 0.08 & 0.05 \end{pmatrix}$ and $b = \begin{pmatrix} 75 \\ 4 \end{pmatrix}$, we can rewrite it in the matrix form:

$$\max\{c^Tx : Ax \leq b,\ x \geq 0\}$$

which is a standard form for a linear programing problem.

Note: Constraints <u>CANNOT</u> be strict inequalities in an LP, i.e., constraints like $3x_1 + 3x_2 < 5$ are illegal!

4 LP formulations

Example 1. A company makes 4 types of products, each requiring time on two different machines and two types of labour. The amount of machine time and labour needed to produce one unit of each product along with its sale price are summarized in the following table.

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

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

Variables: Let x_1, x_2, x_3, x_4 be the number of units of products 1,2,3,4 that we make.
Let y_s, y_u be the hours of skilled, unskilled labour that we hire.

Objective function: $\max 300x_1 + 260x_2 + 220x_3 + 180x_4 - 8y_s - 6y_u$ (Money from selling products subtract labour costs)

Amount of time on machine 1 is at most 700, so $11x_1 + 7x_2 + 6x_3 + 5x_4 \leq 700$

Amount of time on machine 2 is at most 500, so $4x_1 + 6x_2 + 5x_3 + 4x_4 \leq 500$

Amount of time on skilled labour is at most 600, so $y_s \leq 600$.

Amount of time on unskilled labour is at most 650, so $y_u \leq 650$.

Use of skilled labour is at most y_s: $8x_1 + 5x_2 + 5x_3 + 6x_4 \leq y_s$.

Use of unskilled labour is at most y_u: $7x_1 + 8x_2 + 7x_3 + 4x_4 \leq y_u$.

Non-negativity constraints: $x_1, x_2, x_3, x_4, y_s, y_u \geq 0$

Optimal solution; $x_1 = 16\frac{2}{3}, x_2 = 50, x_3 = 0, x_4 = 33\frac{1}{3}$, $y_s = 583\frac{1}{3}, y_u = 650$.

Problems: not integers. Need integer programs (IPs): LP except variables need to be integral. IPs are not efficient to solve.

5 Linear programming formulations

Example 2. A certain company provides heating oil for the local community. They have historical data that helps them predict the demands for heating oil in the next four months.

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

At the beginning of each month, they can purchase oil from the supplier at the current market rate. The projected rates are given in this table.

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

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

Variables : At the beginning of each month, the company purchase oil, so denote x_i as the amount of oil in litres that we buy in month $i, i = 1, 2, 3, 4$.

Also at the beginning of each month, there are oil left in the storage tank, so let y_i be the amount of oil in the storage tank at the beginning of month $i, i = 1, 2, 3, 4$

The company wants to minimize the total money spent. So the objective function is $\min 0.75x_1 + 0.72x_2 + 0.92x_3 + 0.90x_4$

The constraints are:

- Non-nongetivity $x_i, y_i \geq 0, i = 1, ..., 4$.
- The limit on the storage tank is 4000L: $y_i \leq 4000, i = 1, 2, 3, 4$.
- Month 1, $y_1 = 2000$.
- For each month, we get oil from what we buy and also the storage tank, so in total we get $x_i + y_i$ for month i we sent the demanded oil and what is left is in the storage tank for the next month. So

- Month 1: $x_1 + y_1 = 5000 + y_2$.
- Month 2: $x_2 + y_2 = 8000 + y_3$.
- Month 3: $x_3 + y_3 = 9000 + y_4$.
- Month 4: $x_4 + y_4 \geq 6000$. (could be =)

Now we can write down the full LP formulation:

$$\begin{aligned} \min \quad & 0.75x_1 + 0.72x_2 + 0.92x_3 + 0.90x_4 \\ & y_1 = 2000 \\ & y_i \leq 4000, i = 1, 2, 3, 4 \\ & x_1 + y_1 = 5000 + y_2 \\ & x_2 + y_2 = 8000 + y_3 \\ & x_3 + y_3 = 9000 + y_4 \\ & x_4 + y_4 \geq 6000 \\ & x_i, y_i \geq 0, i = 1, ..., 4 \end{aligned}$$

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

Let $M = \max\{0.75x_1, 0.72x_2, 0.92x_3, 0.90x_4\}$, we want to minimize M. Minimize the maximum ,i.e.,

$$\min \max\{0.75x_1, 0.72x_2, 0.92x_3, 0.90x_4\}$$

How do we convert this min max problem into LPs?

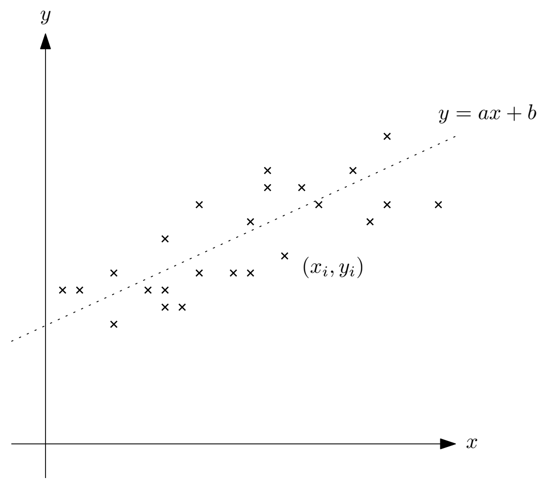
We keep the constraints from the previous example and add new constraints.
We define M to be an upper bound, i.e., $M \geq 0.75x_1, M \geq 0.72x_2, M \geq 0.92x_3, M \geq 0.90x_4$

The objective function is $\min M$. (mininize the upper bound)

We claim that in the optima solution, $M = \max\{0.75x_1, 0.72x_2, 0.92x_3, 0.90x_4\}$, why?

If not, then M must be greater than all 4 numbers. So we can reduce M by just a little so that the 4 constraints hold, but we have a better solution (lower objective value). So it could not have been optimal.

Example 3. We are given a set of data points $\{(x_i, y_i) \ : \ i = 1, \ldots, n\}$ on the plane, perhaps from an experiment. Plotting them on a graph suggests that the x values and the y values may be linearly related. Find a line $y = ax + b$ that "best fits" this set of data points.



By "best fit" here, we mean minimizing the total vertical distance from each point to the line.

Given point (x_i, y_i), its vertical distance to the line $y = ax + b$ is $|y_i - ax_i - b|$.

Variables: a, b in $y = ax + b$.

The objective function is $\min \sum_{i=1}^{n} |y_i - ax_i - b|$. (minimize sum of all errors)

The objective fuction is not affine nor linear!

Define variables e_i for each $i = 1, .., n$, we want $e_i = |y_i - ax_i - b|$.
Then $e_i = \max\{y_i - ax_i - b, -y_i + ax_i + b\}$.
Set constraints $e_i \geq y_i - ax_i - b$, $e_i \geq -y_i + ax_i + b$.

The objective function: $\min \sum_{i=1}^{n} e_i$. (The variables e_i's are independent with each other, so at the optimal solution, it is guaranteed that $e_i = |y_i - ax_i - b|$

Full LP:

$$\begin{array}{ll} \min & \sum_{i=1}^{n} e_i \\ \text{s.t.} & e_i \geq y_i - ax_i - b, \quad i = 1, ..., n \\ & e_i \geq -y_i + ax_i + b, \quad i = 1, ..., n. \end{array}$$

Exercise:

1. Modify this to find the best fit parabola. (Is it linear?)

2. Suppose instead we are maximizing the absolute value $\max\{|x| : -6 \le x \le 5\}$, is it possible to convert it to a linear program?

6 Integer programming formulations

An integer program (IP) is a linear program except we restrict variables to take on only integer values. IPs are generally difficult to solve, and there are no known efficient algorithms for them. But IPs are more powerful in formulating problems.

Example 1. Consider the job application process where a company has 3 positions available, and there are 4 applicants for these jobs. For each applicant and position, the company assigns a number indicating how well the applicant is suited for the position. The goal is to hire a different applicant for each position to maximize the total suitability.

		Candidates			
		1	2	3	4
Positions	1	3	5	2	4
	2	3	1	4	3
	3	1	4	2	3

Want to know: For each position, who gets that position?

Variables: x_{ij} for each position and candidate j. In total, there are 12 variables.

Represent the solution by $x_{ij} = \begin{cases} 1 & \text{if position } i \text{ is given to candidate } j \\ 0 & \text{otherwise} \end{cases}$

Those variables are binary variables.

The objective function: $\max \sum_{i=1}^{3} \sum_{j=1}^{4} c_{ij} x_{ij}$ (sums all c_{ij} where $x_{ij} = 1$.)

Constraints:

1. $x_{ij} \in \{0, 1\}$, $\forall\ i = 1, 2, 3,\ j = 1, 2, 3, 4$. (Integer programs only) Instead, we can have $x_{ij} \in \mathbb{Z}, 0 \leq x_{ij} \leq 1$
2. Each position can be filled by at most 1 candidates: $x_{i1} + x_{i2} + x_{i3} + x_{i4} \leq 1, i = 1, 2, 3$
3. Each candidate can hold at most 1 position: $x_{1j} + x_{2j} + x_{3j} \leq 1, j = 1, 2, 3, 4$.

CO 250 Winter 2022 Week 2

Class outline.

1. Integer programming formulations
2. Optimization on graphs

7 Integer programming formulations

IPs can be used to model logical statements.

Example 2: Knapsack Problem. There are 4 types of items that you can put into your backpack. You can take any integer number of units of any item. However, you can only carry a maximum of 40 pounds. Each unit of item you take is also worth a certain amount of money. The goal is to maximize the total value of the items you carry.

Item	A	B	C	D
Weight (lbs)	1	7	3	2
Value ($)	10	50	20	15

Variables: Let x_A, x_B, x_C, x_D represent the number of units of items A, B, C, D we take.

Objective function: $\max 10x_A + 50x_B + 20x_C + 15x_D$ (maximize total value).
Constraints:

1. Non-nonegativity $x_A, x_B, x_C, x_D \geq 0$.
2. Weight cannot exceed 40 lbs: $x_A + 7x_B + 3x_c + 2x_D \leq 40$.
3. Integral variables: $x_A, x_B, x_C, x_D \in \mathbb{Z}$

Full Integer program formulation:

$$\begin{array}{rl} \max & 10x_A + 50x_B + 20x_C + 15x_D \\ \text{s.t.} & x_A + 7x_B + 3x_C + 2x_D \leq 40 \\ & x_i \geq 0, x_i \in \mathbb{Z}, \text{ for all } i \in \{A, B, C, D\} \end{array}$$

Variation 1. Suppose we are allowed to take A only if we take at least one unit of B.
Want: If $x_B = 0$, then x_A must be zero. If $x_B \geq 1$, then x_A could be any value.
Consider the constraint $x_A \leq 40x_B$. When $x_B = 0$, we have $x_A \leq 0$, so this forces $x_A = 0$. When $x_B \geq 1$, we have $x_A \leq 40$ or 40 or more. But we cannot take more than 40 units of A to begin with, so this does not restrict values of x_A.

Variation 2. Suppose we want at least one of the following conditions to hold.

(a) We carry at least 5 units of items A and/or B; or

(b) We carry at least 7 units of items C and/or D.

We define a new binary variable y.
We want

$$y = \begin{cases} 1 \implies & \text{(a) is satisfied.} \\ 0 \implies & \text{(b) is satisfied.} \end{cases}$$

Consider $x_A + x_B \geq 5y$. When $y = 1$, this gives $x_A + x_B \geq 5$, so (b) is satisfied. When $y = 0$, this gives $x_A + x_B \geq 0$, which does not restrict x_A, x_B, so this constrait is redundant. (Note: (a) could still be satisfied when $y = 0$.

Consider $x_C + x_D \geq 7(1 - y)$. When $y = 0$, this give $x_C + x_D \geq 7$, so (b) is satisfied. When $y = 1$, this gives $x_C + x_D \geq 0$, which is redundant.

Exercise: Is it possible to formulate this as "exclusive or" (exactly one condition holds) ?
Consider the following

$$\begin{cases} x_A + x_B & \geq & 5y \\ x_A + x_B & \leq & 4 + 36y \end{cases}$$

where $y = \{0, 1\}$.
So when $y = 0$, we have $x_A + x_B \geq 0$ which is redundant, and $x_A + x_B \leq 4$ which means (a) fails. When $y = 1$, we have $x_A + x_B \geq 5$ which means (a) holds and $x_A + x_B \leq 40$ which is redundant.

Similarly, for x_C and x_D, we have the following:

$$\begin{cases} x_C + x_D & \geq & 7(1 - y) \\ x_C + x_D & \leq & 6 + 34(1 - y) \end{cases}$$

Variation 3. Suppose that the value of item A is \$10 for the first 5 units, but any more units beyond that has value \$5.

We divide the variables into two parts x_{A_1} and x_{A_2} where x_{A_1} represent the unit of item purchased at \$10, and x_{A_2} represent the unit of item purchased at \$5.
In the objective function, we replace $10x_A$ with $10x_{A_1} + 5x_{A_2}$ and replace x_A with $x_{A_1} + x_{A_2}$ in the knapsack constraint.
We add constraint $x_{A_1} \leq 5$.
In addition we have the following constraints: If $x_{A_1} \leq 4$, then $x_{A_2} = 0$. If $x_{A_1} \geq 5$, then x_{A_2} is free. why ??
We introduce a binary variable $y = \{0, 1\}$ with the constraints .

$$\begin{cases} x_{A_1} & \leq & 4 + y \\ x_{A_2} & \leq & 40y \end{cases}$$

and

$$x_{A_1} \geq 5y$$

8 Optimization in graphs

What is a graph?

A graph $G = (V, E)$ consists of a set of objects V called vertices and a collection of unordered pair of vertices E called edges.
Example: Define $G = (V, E)$ by $V = \{1, 2, 3, 4\}$, $E = \{12, 23, 34, 41, 24\}$
A drawing of G:

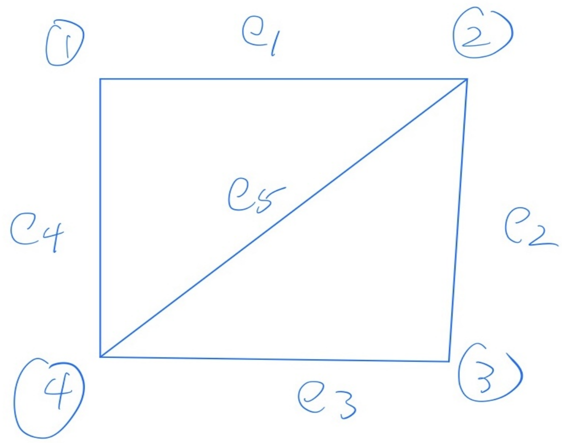


Figure 2: Graph G

Incidence relation.

For an edge $e = uv$, we say e is incident with u and v.

Example above: e_1 is incident with vertices 1 and 2.

Vertex 2 is incident with edges e_1, e_5, e_2.

We use $\delta(v)$ to represent the set of all edges incident with v.

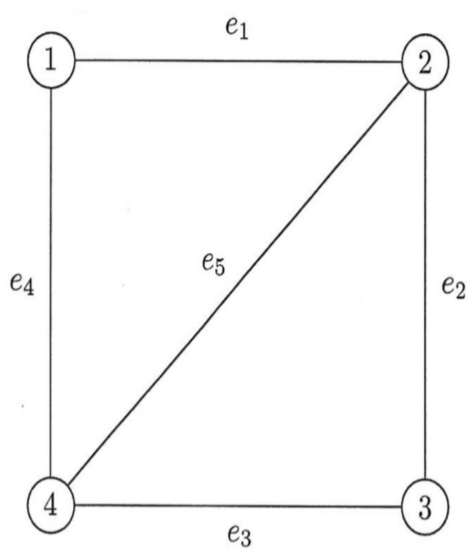
$$\delta(2) = \{e_1, e_5, e_2\}, \quad \delta(3) = \{e_2, e_3\}.$$

Class outline.

1. Matchings
2. Shortest paths

Incidence matrix.

The incidence matrix of $G = (V, E)$ is a $|V| \times |E|$ matrix where the rows are indexed by the vertices, the columns are indexed by the edges, and the entry corresponding to row V and column e is 1 if e is incident with v; is 0 otherwise.

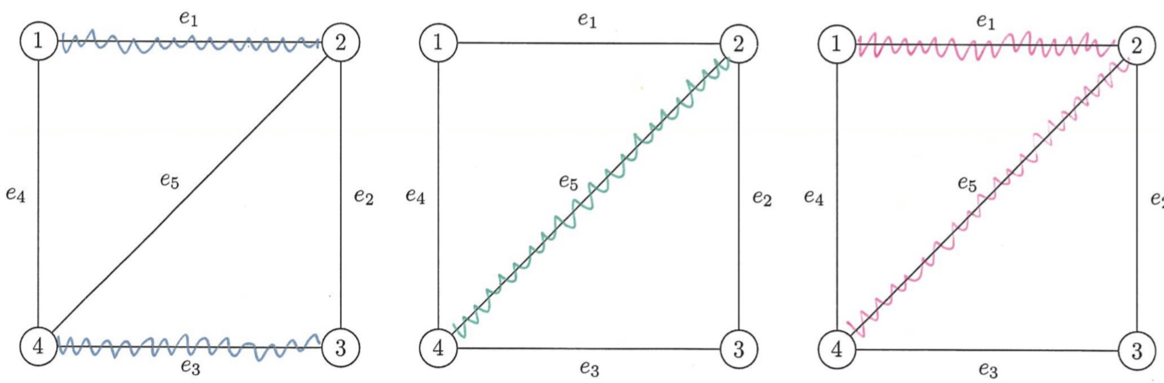


$$\begin{array}{c|c} & \begin{matrix} e_1 & e_2 & e_3 & e_4 & e_5 \end{matrix} \\ \hline \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \end{matrix} & \begin{pmatrix} 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \end{pmatrix} \end{array}$$

Note: Each column has exactly 2 1's.

9 Matchings

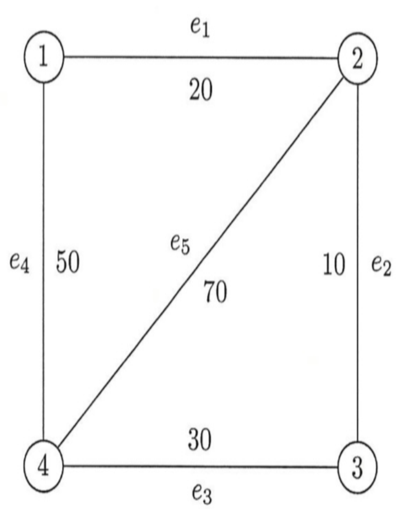
Given a graph $G = (V, E)$, a matching is a set of edges M where each vertex is incident with at most one edge in M. In other words, $|M \cap \delta(v)| \leq 1$ for all $v \in V$.



$M = \{e_1, e_3\}$ matching $M = \{e_5\}$ matching Not a matching, Vertex 2 is incident with 2 edges.

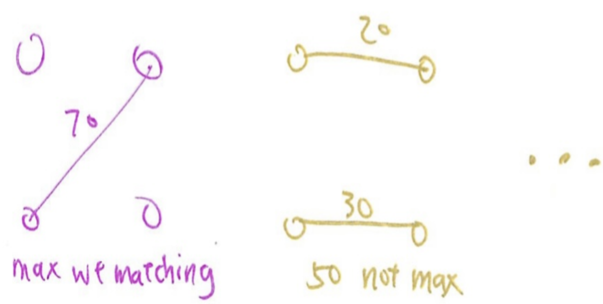
Maximum weight matching problem.

Given: graph $G = (V, E)$, and weight w_e for each $e \in E$.
Goal: find a matching in G where total edge weight is maximized.
Examples:



Given: Graph $G=(V,E)$, and weight w_e for each $e \in E$.

Goal: Find a matching in G whose total edge weight is maximized.



IP formulation.

1. variables: need to decide which edge are in the matching, thus define x_e for each $e \in E$, we want this to represent

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

2. objective function: $\max \sum_{e \in E} w_e x_e$ maximize total weight of edges in the matching. In the above example, $\max\{20x_{e_1} + 10x_{e_2} + 30x_{e_3} + 50x_{e_4} + 70x_{e_5}\}$, $w = (20, 10, 30, 50, 70)^T$. The objective function is $\max w^T x$

3. inequality constraints: For each vertex v, we want at most one edge in $\delta(v)$ to be in the matching, so we have $\sum_{e \in \delta(v)} x_e \leq 1$, for each $v \in V$. In the example:

$$
\begin{aligned}
x_{e_1} + x_{e_4} &\leq 1 \quad \text{(vertex 1)} \\
x_{e_1} + x_{e_2} + x_{e_5} &\leq 1 \quad \text{(vertex 2)} \\
x_{e_2} + x_{e_3} &\leq 1 \quad \text{(vertex 3)} \\
x_{e_3} + x_{e_4} + x_{e_5} &\leq 1 \quad \text{(vertex 4)}
\end{aligned}
$$

4. matrix form of the constraints

$$
\begin{pmatrix} 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \end{pmatrix}
\begin{pmatrix} x_{e_1} \\ x_{e_2} \\ x_{e_3} \\ x_{e_4} \\ x_{e_5} \end{pmatrix}
\leq
\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}
$$

Full IP for general maximal weight matching problem:

Let A be the incidence matrix of G, let $\mathbb{1}$ be a vector of all 1's. Then

$$
\begin{aligned}
\max \quad & w^T x \\
\text{s.t} \quad & Ax \leq \mathbb{1} \\
& x_e \in \{0, 1\} \quad \forall e \in E.
\end{aligned}
$$

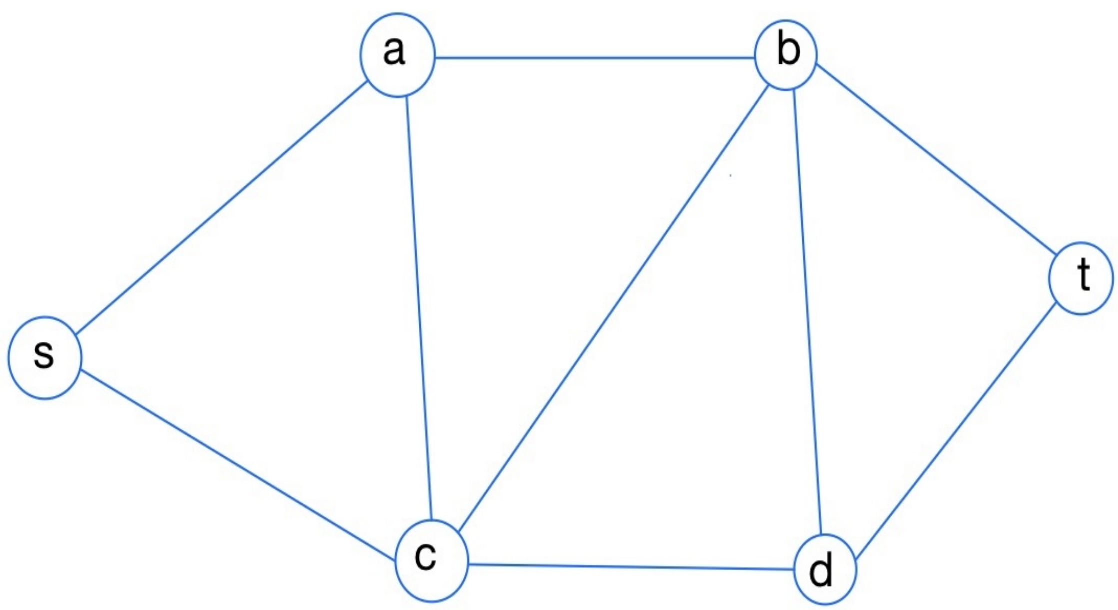
10 Shortest paths

What is a path?

Definitions:

A path is a sequence of edges $v_1v_2, v_2v_3, \cdots, v_{k-1}v_k$ such that $v_1, \cdots, v_k$ are distinct. It is an s,t - path if $v_1 = s$ and $v_k = t$.

Examples:

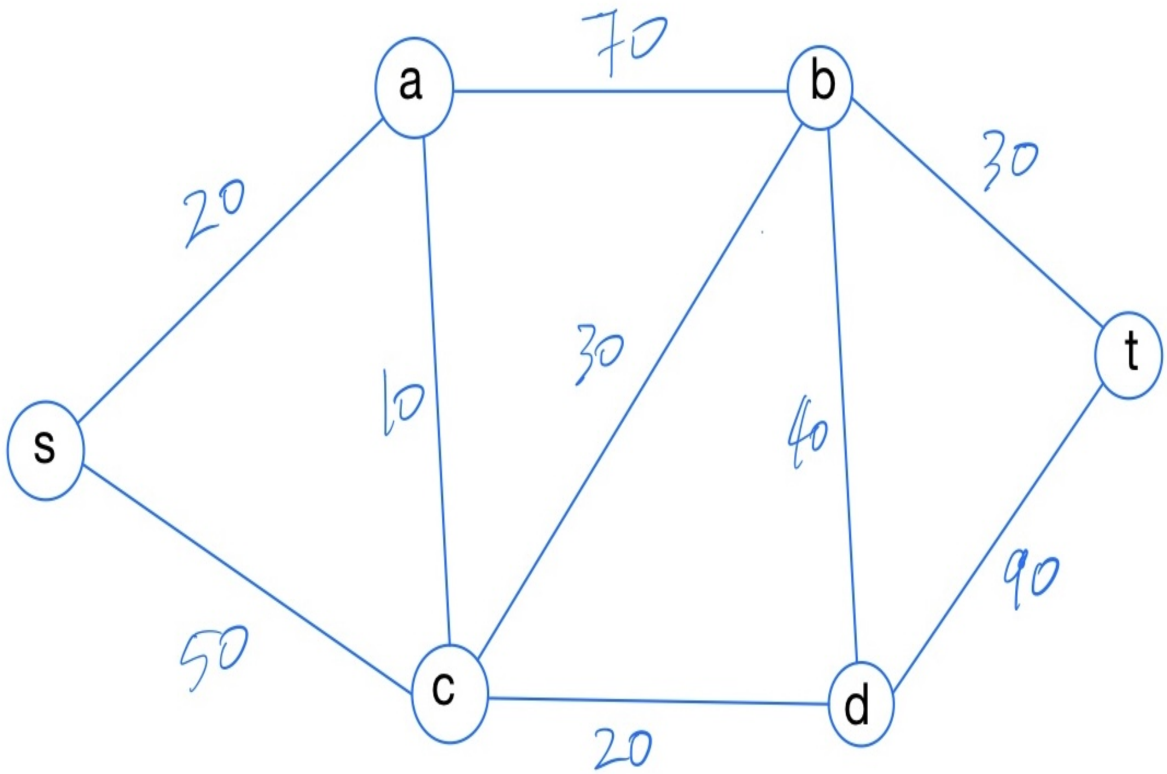


1. $\{sa, ab, bc, cd\}$ is a path but not an s,t-path.
2. $\{sa, ab, bc, cd\}$ is an s,t-path.
3. $\{sc, dt\}$ is not a path.
4. $\{sc, cb, ba, ac, cd, dt\}$ is not a path. (Vertices are not distinct, contains a cycle)

Shortest path problem.

Given: graph $G = (V, E)$, positive edge weights w_e for each edge e, two vertices s, t.

Goal: Find an s,t-path with minimum total edge weight.
Examples:



- $\{sa, ab, bt\}$ with total edge weight 120.
- $\{sa, ac, cb, bt\}$ with total edge weight 90. Is this the shortest s,t-path?

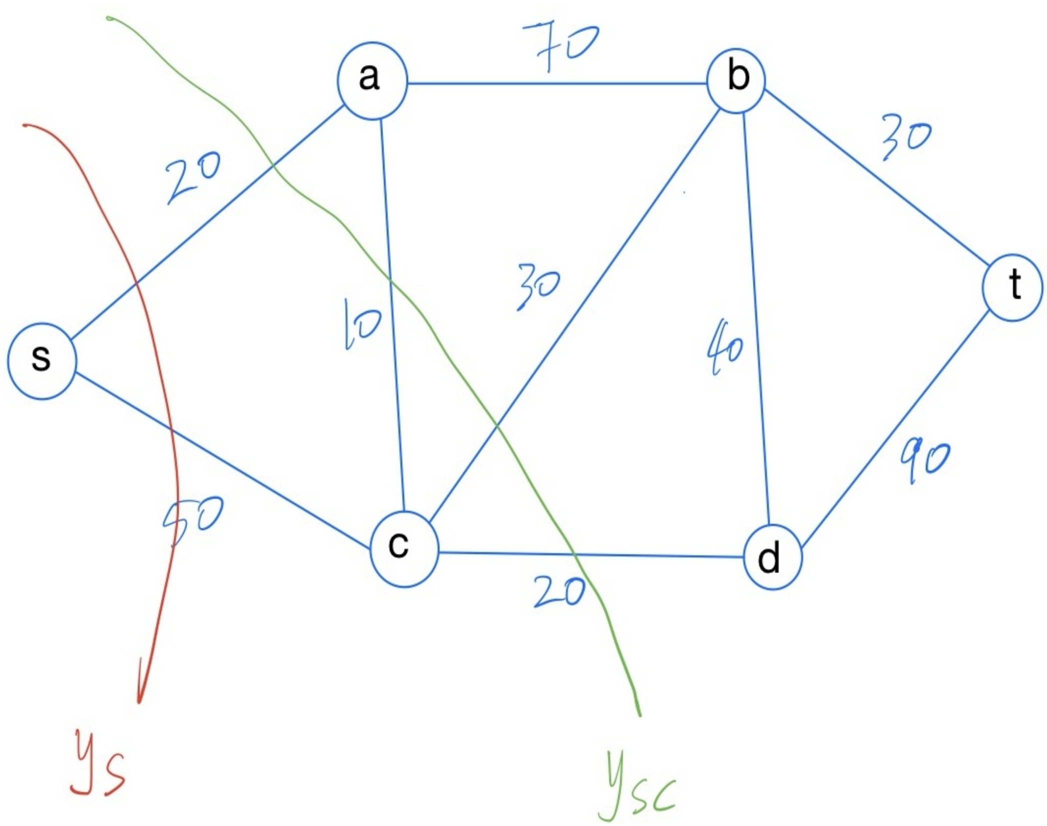
IP formulation of the example:

1. variables: Define x_e for each $e \in E$ to represent
$$x_e = \begin{cases} 1 & e \text{ is in the path} \\ 0 & \text{otherwise} \end{cases}$$
2. objective function: $\min \sum_{e \in E} w_e x_e$ or $\min w^T x$.
3. constraints: how to add constraints??

We need to add constraints to force optimal solutions to form s,t - paths. The key idea is using "cuts"

Why we need cuts? Imagine vertices as cities, a cut is like a river that separates city s from city t. The edges on the river are bridges. If we want to reach city t from city s, we must cross the river (cut) that separate s from t, i.e., we must cross at least 1 bridge on the river.

Examples following the example of paths with a few rivers (cuts) added:



- constraint corresponding to cut y_s: $x_{sc} + x_{sa} \geq 1$.
- constraint corresponding to cut y_{sc}: $x_{sa} + x_{ac} + x_{bc} + x_{cd} \geq 1$

Cuts vs paths.

Formal definition of cuts and s,t cuts:

Examples of cuts:

Formalizing the idea of river and bridge:

given any s,t-cut $\delta(w)$, any s,t-path must cross at least one edge in $\delta(w)$.

Examples with all s.t-cuts added:

Possible General IP formulation for shortest path problem (Constraints corresponding to all s.t-cuts.):
Problem: a feasible solution may not be an s,t-path. For example: a solution of all ones.
Resolution: an optimal solution must be an s,t- path.
Why??
The idea:

1. If a set of edges intersects every s,t-cut, then this set of edges must contain an s,t-path. (i.e., it contains an s,t-path plus possibly some redundant edges)

2. If we have a solution which contains redundant edges ($x_e = 1$ for redundant edge e), since $w_e > 0$, we can set $x_e = 0$ to get a better solution! Therefore, an optimal solution of the general IP formulation must be an s.t-path.

3. By the same argument, we don't need $x_e \leq 1$, any higher values of x_e would not be optimal.

Full shortest path IP simplied:

CO 250 Winter 2022 Week 3

Class outline.

1. shortest paths

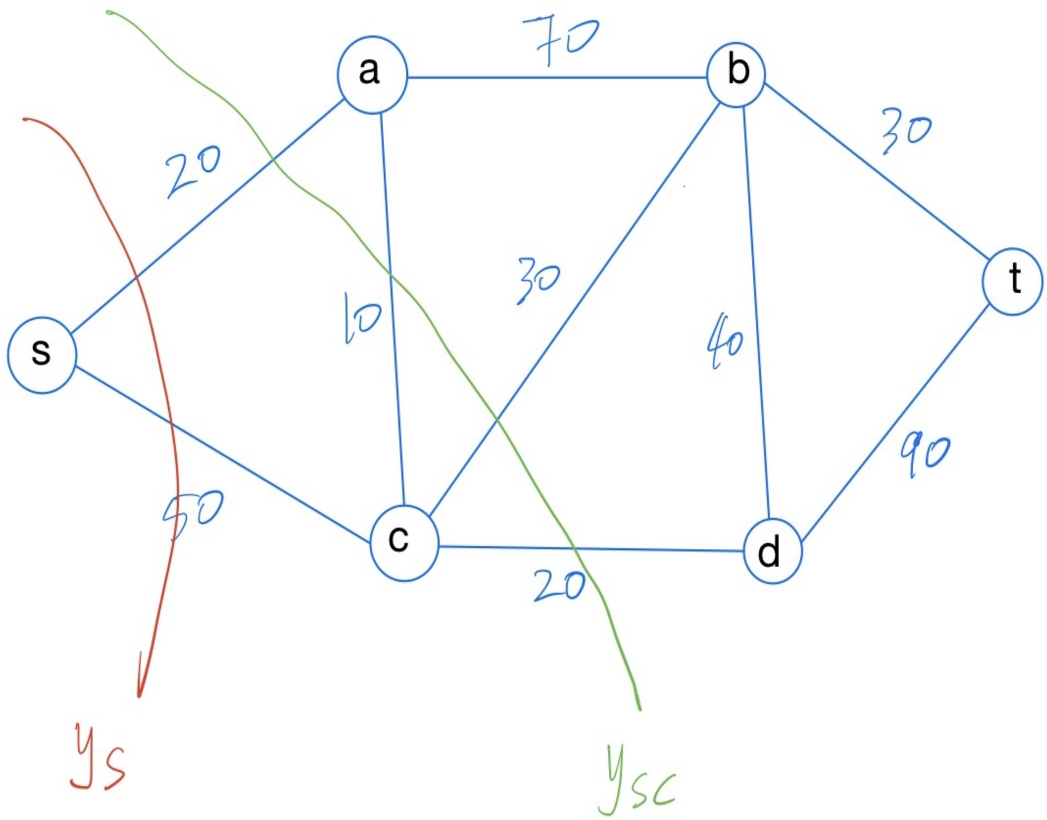
2. Nonlinear programing (NLP)

3. Solving linear programs

4. Possible outcomes

11 Shortest path problem.

Cuts vs paths



Cuts vs paths.
Formal definition of cuts and s,t cuts:

Definition 1. *Given $W \subseteq V$, the cut induced by W is the set of all edges with exactly one endpoint in W. Notation: $\delta(W)$. $\delta(W) = \{uv \in E : u \in W, v \notin W\}$. This $\delta(W)$ is an s,t-path if $s \in W$ and $t \notin W$.*

Examples of cuts:

$$W_1 = \{s, a, c\}, \quad \delta(W_1) = \{ab, cb, cd\}$$

$$W_2 = \{s, a, b\}, \quad \delta(W_2) = \{sc, ac, bc, bd, bt\}$$

$$W_3 = \{s, a, d\}, \quad \delta(W_2) = \{ab, ac, sc, cd, bd, dt\}$$

Formalizing the idea of river and bridge:
Claim: given any s,t-cut $\delta(w)$, any s,t-path must cross at least one edge in $\delta(w)$.

Proof. Suppose there exists an s,t-cut $\delta(w)$ such that $\delta \cap L = \emptyset$. Assume $\delta(w)$ separate the set of vertices V of G into two disjoint parts w and $V \backslash w$ such that $s \in w$ and $t \in V \backslash w$. Since $L = \{v_1v_2, ..., v_{k-1}v_k\}$ $(v_1 = s, v_k = t)$ is a path from s to t with $v_1 \in w$ and $v_k \in V \backslash w$, there must exist an edge $v_i v_{i+1}$ with $v_i \in w, v_{i+1} \in V \backslash w$, however, this edge $v_i v_{i+1}$ must be contained in the s,t-cut $\delta(w)$, a contradiction. $\square$

Examples with all s.t-cuts added:
Possible General IP formulation for shortest path problem (Constraints corresponding to all s.t-cuts.):

$$\sum_{e \in \delta(W)} X_e \geq 1 \quad \text{for all s,t-cuts } \delta(W)$$

Problem: a feasible solution may not be an s,t-path. For example: a solution of all ones.
Resolution: an optimal solution must be an s,t- path.
Why??
The idea:

1. If a set of edges intersects every s,t-cut, then this set of edges must contain an s,t-path. (i.e., it contains an s,t-path plus possibly some redundant edges)

   Proof: Let C be a set of edges which intersect every s,t-cut. We start with $W_1 = s$ and the corresponding cut $\delta(W_1)$. We let $v_1 = s$ and $L_1 = \emptyset$. Then we choose one edge $e_1 \in \delta(W_1) \cap C$. ($\delta(W_1) \cap C$ is non-empty since C intersects every s,t-cut). We add e_1 to L_1 to get $L_2 = L_1 \cap \{e_1\}$. Also one of the two ends of e_1, say v_2, must not lie in W_1. We add v_2 to W_1 to obtain W_2 and the corresponding cut $\delta(W_2)$. We repeat this process, after at most n steps (where n is the number of total vertices), we will have $t \in W_i$ for some $i \leq n$. We claim there is a path from s to t in L_k which is an s,t-path. We can prove this by induction, at step k, the induction hypothesis assumes there exists a path from v_1 to every vertex in W_k in L_k. Now we add an edge e_k from the cut $\delta(W_k)$ and corresponding vertex v_{k+1}, and v_{k+1} is connected to exactly one vertex in W_k by this edge e_k, therefore there must exist a path from v_1 to every vertex in W_{k+1} in L_{k+1}, hence the induction hypothesis holds. In the end we will get an s,t-path from C in at most n steps.

2. If we have a solution which contains redundant edges ($x_e = 1$ for redundant edge e), since $w_e > 0$, we can set $x_e = 0$ to get a better solution! Therefore, an optimal solution of the general IP formulation must be an s.t-path.

3. By the same argument, we don't need $x_e \leq 1$, any higher values of x_e would not be optimal.

Full shortest path IP simplied:

$$\begin{aligned} \min \quad & w^T x \\ \text{s.t.} \quad & \sum_{e \in \delta(W)} X_e \geq 1 \quad \forall \text{ s,t-cuts } \delta(W) \\ & X_e \geq 0, X_e \in \mathbb{Z}, \forall e \in E \end{aligned}$$

12 Nonlinear Programming

General form of nonlinear programming:
A short NLP example:
NLP vs IP vs LP

1. LP is a special NLP.
2. For IPs (unless otherwise stated, we mean linear IPs.), we can simply enforce the integer constraints by using $\sin(\pi x) = 0$, which implies $\pi x = \pi k, k \in \mathbb{Z}$, so $x \in \mathbb{Z}$. For binary variables $x \in \{0, 1\}$, we can add $x^2 - x \leq 1, x - x^2 \leq 1$. Any IPs can be solved as an NLPs.
3. NLPs are harder to solve.

A general nonlinear programming problem has the following form:

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

where $f(x) : \mathbb{R}^n \to \mathbb{R}$ and $g_i(x) : \mathbb{R}^n \to \mathbb{R}$ for each i. Example:

Among all the pts x that satisfy $Ax \leq b$, find one that is closest to target point y. we minimize the distance

$$||x - y||^2 = (x - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2$$

An NLP formulation:

$$\begin{aligned} \min \quad & ||x - y|| \\ \text{s.t.} \quad & Ax \leq b \end{aligned}$$

NLP vs LP vs IP

1. An LP is an NLP.
2. Any IP is a special NLP. reason: since we can always impose a constraint $\sin(\pi x) = 0$, which implies $\pi x = k\pi$, $k = 1, 2, \ldots$. Hence $x \in \mathbb{Z}$
3. NLPs are harder to solve than IPs

13 Solving linear programs

13.1 Infeasible LPs

Examples:

$$\begin{aligned}
\max \quad & (3, 1, -7, 4)(x_1, x_2, x_3, x_4)^T \\
\text{s.t.} \quad & \begin{pmatrix} 5 & -4 & -3 & 1 \\ -2 & -1 & 5 & -3 \\ 1 & 3 & -1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} -3 \\ 2 \\ -1 \end{pmatrix} \\
& x \geq \mathbb{0}
\end{aligned}$$

Consider $y = (2, 3, 4)^T$, rewrite the linear system in matrix form $Ax = b$ and multiply with y to get $y^T Ax = y^T b$. By computation, we have $y^T A = (-8, -1, -5, -1)^T$ and $y^T b = 4$. Since $x \geq 0$, we have $y^T Ax \leq 0$. However, $y^T b = 4$, this is impossible, so this linear program is infeasible.

Proposition 1. *The system $Ax = b$, $x \geq 0$ is infeasible if there exists a vector y such that $y^T A \geq 0$ and $y^T b < 0$.*

Proof. Suppose the system is feasible with x being a feasible solution. By assumption, $Ax = b$. Multiply y^T to get $y^T Ax = y^T b$. By assumption, $y^T A \geq 0$ and $x \geq 0$, so $y^T Ax \geq 0$. But $y^T b < 0$, a contradiction. So this system must be infeasible. $\square$

Is the converse also true? Yes, a certificate always exists for an infeasible system, according to "Farkas Lemma".

13.2 Unbounded LPs

Definition 2. *A maximization LP is unbounded if there exists a series of feasible solutions $x(t)$ such that the objective value approaches ∞ as $t \to \infty$.*

Examples:

$$\begin{aligned}
\max \quad & (-1,2,3,4)(x_1,x_2,x_3,x_4)^T \\
\text{s.t.} \quad & \begin{pmatrix} 3 & 0 & 2 & -5 \\ -2 & 3 & -4 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 4 \\ 1 \end{pmatrix} \\
& x \geq \mathbb{0}
\end{aligned}$$

We first rewrite the LP in matrix form, $\max c^T x : Ax = b, x \geq 0$.
We consider the vectors $\bar{x} = (3,1,0,1)^T$, which is a feasible solution such that $A\bar{x} = b$ and $d = (0,4,5,2)^T$. We can check that d is in the nullspace of A, i.e., $Ad = 0$. We can let $x(t) = \bar{x} + td$, then $Ax(t) = b$, and $x(t) \geq 0$ for any $t > 0$. Also $c^T d = 31$. The objective value is $c^T(x(t)) = 3 + 31t \to \infty$ as $t \to \infty$. Therefore it is an unbounded LP.
The certificate consists of both $\bar{x}$ and d. They are not unique

Proposition 2. *The LP* $\max\{c^T x : Ax = b, x \geq 0\}$ *is unbounded if there exists a feasible solution* $\bar{x}$ *and a vector* d *such that* $Ad = 0, d \geq 0, c^t d > 0$.

Proof. Exercise □

The converse is also true (duality theory).

13.3 LPs with optimal solution

Example:

$$\begin{aligned}
\max \quad & -2x_2 - 3x_3 + 7 \\
\text{s.t.} \quad & \begin{pmatrix} 1 & 3 & -5 & 0 \\ 0 & -1 & 2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 6 \\ 9 \end{pmatrix} \\
& x \geq \mathbb{0}
\end{aligned}$$

We see that $\bar{x} = (6,0,0,9)^T$ is feasible, with objective value 7.
For any feasible solution $\bar{x}$, its objective value is $-2x_1 - 3x_3 + 7 \leq 7$ since $x_2, x_3 \geq 0$. So 7 is an upper bound for this maximization problem and it is attained by $\bar{x} = (6,0,0,9)^T$, we know $\bar{x}$ must be an optimal solution.
Why this works: two observation: (1) $c \leq 0$. (2) $\bar{x}$ is feasible where $c^T\bar{x} = 0$. The certificate consist of both c and $\bar{x}$.

13.4 Fundamental Theorem of Linear Programming

Theorem 1. *For any LP, exactly one of these three holds:*

1. *It is infeasible.*
2. *It is unbounded.*
3. *It has an optimal solution.*

CO 250 Winter 2022 Week 4

Class outline.

1. Foundamental Theorem of LP
2. Equivalent forms of LP
3. Simplex method

13.5 Fundamental Theorem of Linear Programming

Theorem 2. *For any LP, exactly one of these three holds:*

1. *It is infeasible (We can find a certificate).*
2. *It is unbounded. (We can find a certificate.*
3. *It has an optimal solution (We can find a certificate.*

Example of non-LP where one of these 3 conditions hold:

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

This problem is feasible and bounded, but there is no optimal solution, the optimal can not be attained, our solution can be arbitrarily close to 1 but never equal to 1.

14 Equivalent LPs

Definition 3. *Two LPs (P) and (P') are <u>equivalent</u> if*

1. *(P) is infeasible if and only if (P') is infeasible.*
2. *(P) is unbounded if and only if (P') is unbounded.*
3. *Given any optimal solution to (P), one can construact an optimal solution to (P'), and vice versa.*

14.1 Methods to convert LP to equivalent forms

1. minimization $\Longleftrightarrow$ maximization

$$\min -f(x) = -\max f(x)$$

   So when we minimize $f(x)$ and it is unbounded (below), we can see it is unbounded above when we maximize $-f(x)$.

   $\min 2x - 2y$ is equivalent to $\max -2x + 2y$

2. Inequalities to equalities with nonnegative constraints. Suppose we have $Ax \leq b$, we can add nonnegative "slack" variabes z such that $Ax+z = b, z \geq 0$

3. Free variables: Suppose $Ax = b$, x are free variables, we can introduce new variables x_1 and x_2 such that $x_1 \geq 0, x_2 \geq 0$ and let $x = x_1 - x_2$. So $Ax = b$ becomes $A(x_1 - x_2) = b$ and $x_1 \geq 0, x_2 \geq 0$.

   Example $x + y = 5$, we can turn it in to

$$\begin{aligned} x_1 - x_2 + y_1 - y_2 &= 5 \\ x_1, x_2, y_1, y_2 \geq 0 & \end{aligned}$$

14.2 Convert an LP to standard equality form

Any LP can be written as the following standard equality form (SEF)

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

where $\bar{z}$ denotes some constants.

Example:

(a) convert $\min c^T x : Ax = b$ into SEF:
Add variables $x = x_1 - x_2$ with $x_1 \geq 0, x_2 \geq 0$

(b) convert $\min c^T x : Ax \leq b$ into SEF:
Add slack variables z first, so we have

$$\min c^T x : Ax + z = b, z \geq 0$$

Then add $x = x_1 - x_2, x_1 \geq 0, x_2 \geq 0$, so we have

$$\min c^T (x_1 - x_2) : A(x_1 - x_2) + z = b, z \geq 0, x_1 \geq 0, x_2 \geq 0$$

Example: Find an equivalent LP in SEF

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

- $\min(-1,2,-3)x \to \max(1,-2,3)$.
- $x_1 + 5x_2 + 3x_3 \leq 5 \to x_1 + 5x_2 + 3x_3 + x_4 = 5, x_4 \geq 0$. ($x_4$) is a new variable.
- $2x_1 - x_2 + 2x_3 \geq 4 \to 2x_1 - x_2 + 2x_3 - x_5 = 4, x_4 \geq 0$. ($x_5$ is a new variable)
- x_3 is free, so add new variables x_3^+, x_3^-, replace x_3 with $x_3^+ - x_3^-, x_3^+ \geq 0, x_3^- \geq 0$

Let $x = (x_1, x_2, x_3^+, x_3^-, x_4, x_5)^T$. Then

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

An optimal solution to the LP in SEF is $(\frac{11}{4}, 0, \frac{3}{4}, 0, 0, 3)^T$ with optimal value 5.

This corresponds to the optimal solution of the LP in the original form: $(\frac{11}{4}, 0, \frac{3}{4})^T$ with optimal value -5.

15 A taste of simplex

Example:

$$\begin{array}{rl} \max & (3,-2,0,0,0)x \\ \text{s.t.} & \begin{pmatrix} 4 & -1 & 1 & 0 & 0 \\ 3 & -3 & 0 & 1 & 0 \\ -2 & 2 & 0 & 0 & 1 \end{pmatrix} x = \begin{pmatrix} 8 \\ 9 \\ 1 \end{pmatrix} \\ & x \geq \mathbb{0} \end{array}$$

A feasible solution of this problem is $\bar{x} = (0,0,8,9,1)^T$.
The objective value is 0.
Question: Is $\bar{x}$ an optimal solution?

We want to find a direction to increase the objective value while maintain feasibility.
How to find such a direction?
By observation, we see the coefficient of x_1 in the objective function is 3, therefore if we can increase the value of x_1 while fixing the value of x_2, we will increase the objective value.
So let $x_1 = t, x_2 = 0$, then we have $x_3 = -4t, x_4 = -3t, x_5 = 2t$. ($t \geq 0$ since we want to increase the objective value)
So we have found a direction $d = (1, 0, -4, -3, 2)$. (it is a vector in the nullspace of matrix A.)
Now we also want $x + td \geq 0$ (feasibility). So we need $x_3 = 8 - 4t \geq 0$, $x_4 = 9 - 3t \geq 0$, $x_5 = 1 + 2t \geq 0$. which gives us

$$t \leq 2, t \leq 3, t \geq -\frac{1}{2}$$

So $t = 2$ is the maximum t possible. (we use a greedy approach)
A new feasible solution $x^* = \bar{x} + 2d = (2, 0, 0, 3, 5)^T$.
Observation:

1. Identity matrix for certain columns (column 1,4,5)

2. Coefficient 0 in the objective function for the variables corresponding to the identity matrix. (We call these variables basic variables or basis)

16 Simplex method: Bases and canonical form

16.1 Bases

In order to find an identity matrix in a matrix, we need linearly independent rows. If we have linearly dependent rows, then we have redundant constraints, we can remove the redundant constraints. Without of genericity, we assume our coefficient matrix has full row rank.

Definition 4 (Basis)**.** *Let A be an $m \times n$ matrix ($m \leq n$) with full row rank. The columns are indexed from 1 to n. Denote A_j to be column j of A. Let $J \subseteq \{1, ..., n\}$ be a set of column indices. Denote A_J to be the submatrix of all columns A_j where $j \in J$.*
We say $B \subseteq \{1, ..., n\}$ is a basis if A_B is invertible. For a basis B, we denote $N = \{1, ..., n\} \backslash B$ to be indices of non-basic variables
Let $Ax = b$ where $A = [A_B, A_N]$ and A_B is invertible, then a basic solution is

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

Definition 5 (Basic feasible solution)**.** *A basic solution $\bar{x}$ is a* <u>*basic feasible solution*</u> *if $\bar{x} \geq 0$.*

Examples:
Given

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

1. Given basis $B = \{3, 4, 5\}$, the corresponding $x_B = (8, 9, 1)^T$ which is a basic feasible solution. For another basis $B = \{1, 4, 5\}$, the corresponding $x_B = (2, 3, 5)$ which is also a basic feasible solution.

2. $B = \{1, 2, 3\}$ is not a basis since $A_1 + A_2 - 3A_3 = 0$ (A_B is not invertible).

3. Exercise $B = \{1, 4, 5\}$ is basis, x_B is a basic solution but no a basic feasible solution.

CO 250 Winter 2022 Week 5

Class outline.

1. Canonical forms

2. Simplex method

17 Canonical form

An LP in SEF $\max\{c^T x + \bar{z} : Ax = b, x \geq 0\}$ is in canonical form with respect to a basis B if $A_B = I$ and $c_B = 0$.
Examples. Convert this LP into a canonical form for some basis.

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

First we choose $B = \{2, 3\}$, then $A_B = \begin{pmatrix} 1 & -2 \\ 0 & -1 \end{pmatrix}$ and $A_B^{-1} = \begin{pmatrix} 1 & -2 \\ 0 & -1 \end{pmatrix}$, so B is a basis.
What about $B = \{1, 5\}$? It is not basis!
we multiply $Ax = b$ by A_B^{-1}.

$$\begin{pmatrix} 1 & -2 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 2 & 1 & -2 & 5 & -4 \\ -1 & 0 & -1 & -3 & 2 \end{pmatrix} x = \begin{pmatrix} 1 & -2 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 3 \\ -2 \end{pmatrix}$$

which is

$$\begin{pmatrix} 4 & 1 & 0 & 11 & -8 \\ 1 & 0 & 1 & 3 & -2 \end{pmatrix} x = \begin{pmatrix} 7 \\ 2 \end{pmatrix}$$

Objective function is $c_B = (-1, 4)^T$, we want the objective function to be $(0, 0)^T$. We want to reduce the coefficient of x_2 and x_3 in the objective function to zero using the constraint equalities. Since the coefficient of x_2 in the objective is -1, we multiply the first constraint by 1, since the coefficient of x_3 in the objective is 4, we multiply the second constraint by -4. Then we add them to the objective to cancel the coefficient of x_2 and x_3 in the objective function.
So we multiply $(1, -4)$ on the left:

$$(1, -4)\begin{pmatrix} 4 & 1 & 0 & 11 & -8 \\ 1 & 0 & 1 & 3 & -2 \end{pmatrix} x = (1, -4)\begin{pmatrix} 7 \\ 2 \end{pmatrix}$$

which is

$$\begin{aligned} (0, 1, -4, -1, 0)x &= -1 \\ (0, 1, -4, -1, 0)x + 1 &= 0 \end{aligned}$$

We add this expression to the objective (since it equals to 0) to get

$$\begin{aligned} &(3, -1, 4, 0, -1)x + 2 + (0, 1, -4, -1, 0)x + 1 \\ &= (3, 0, 0, -1, -1)x + 3 \end{aligned}$$

So the canonical form corresponding to the basis $B = \{2, 3\}$ is

$$\begin{array}{ll} \max & (3, 0, 0, -1, -1)x + 3 \\ \text{s.t.} & \begin{pmatrix} 4 & 1 & 0 & 11 & -8 \\ 1 & 0 & 1 & 3 & -2 \end{pmatrix} x = \begin{pmatrix} 7 \\ 2 \end{pmatrix} \\ & x \geq \mathbb{0} \end{array}$$

Is this optimal? Not yet, since the coefficient of x_1 is 3 which is positive, we need all the coeeficients to be nonpositive.
The corresponding basic solution is also a basic feasible solution.
What about $B = \{2, 5\}$? It is basis but x_B is not a basic feasible solution

17.1 Canonical form in symbols

Consider the LP in SEF $\{\max c^T x + \bar{z} : Ax = b, x \geq 0\}$, let B be a basis and partition A as $A = [A_B, A_N]$, we can multiply A_B^{-1} on both sides of $Ax = b$

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

We can then multiply the new constraints by $-c_B^T$ on the left, which is

$$
\begin{aligned}
-c_B^T x_B - c_B^T A_B^{-1} A_N x_N &= -c_B^T A_B^{-1} b \\
-c_B^T x_B - c_B^T A_B^{-1} A_N x_N + c_B^T A_B^{-1} b &= 0
\end{aligned}
$$

We then add 0 to the the objective function $c^T x + \bar{z}$

$$
c^T x + \bar{z} + 0
$$

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

If we define $y^T = c_B^T A_B^{-1}$ meaning $y = A_B^{-T} c_B$. The objective becomes

$$
(c_N^T - y^T A_N) x_N + (\bar{z} + y^T b)
$$

Hence the new LP in canonical form is

$$
\begin{array}{ll}
\max & (c_N^T - y^T A_N) x_N + (\bar{z} + y^T b) \\
\text{s.t.} & x_B + A_B^{-1} A_N x_N = A_B^{-1} b \\
& x \geq \mathbb{0}
\end{array}
$$

Vector y will be "certificates" of optimality or infeasibility in the future.

Example:

$$
\begin{array}{ll}
\max\limits_{x} & (3, -2, 0, 0, 0) x \\
\text{s.t.} &
\end{array}
$$

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

This is canonical form for basi $B = \{3, 4, 5\}$ with BFS $(0, 0, 8, 9, 1)^T$.
We used one simplex iteration to derive a better solution $(2, 0, 0, 3, 5)^T$ with a basis $(1, 4, 5)$. We now find the canonical form using the formula we just derived.

$$
A_B = \begin{pmatrix} 4 & 0 & 0 \\ 3 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix}, \quad A_B^{-1} = \begin{pmatrix} \frac{1}{4} & 0 & 0 \\ -\frac{3}{4} & 1 & 0 \\ \frac{1}{2} & 0 & 1 \end{pmatrix}
$$

We compute

$$y = A_B^{-T} c_B = \begin{pmatrix} \frac{1}{4} & -\frac{3}{4} & \frac{1}{2} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{3}{4} \\ 0 \\ 0 \end{pmatrix}$$

and

$$c_N^T - y^T A_N = (-2, 0) - (3/4, 0, 0)^T \begin{pmatrix} -1 & 1 \\ -3 & 0 \\ 2 & 0 \end{pmatrix} = (-\frac{5}{4}, -\frac{3}{4})$$

and

$$y^T b = 6$$

and

$$A_B^{-1} A_N = \begin{pmatrix} \frac{1}{4} & 0 & 0 \\ -\frac{3}{4} & 1 & 0 \\ \frac{1}{2} & 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & 1 \\ -3 & 0 \\ 2 & 0 \end{pmatrix} = \begin{pmatrix} -1/4 & 1/4 \\ -9/4 & -3/4 \\ 3/2 & 1/2 \end{pmatrix}$$

and

$$A_B^{-1} b = \begin{pmatrix} 2 \\ 3 \\ 5 \end{pmatrix}$$

So the caninical form for basis $B = \{1, 4, 5\}$ is

$$\begin{aligned} \max_x \quad & (0, -5/4, -3/4, 0, 0)x + 6 \\ \text{s.t.} \quad & \begin{pmatrix} 1 & -1/4 & 1/4 & 0 & 0 \\ 0 & -9/4 & -3/4 & 1 & 0 \\ 0 & 3/2 & 1/2 & 0 & 1 \end{pmatrix} x = \begin{pmatrix} 2 \\ 3 \\ 5 \end{pmatrix} \\ & x \geq \mathbb{0} \end{aligned}$$

18 Simplex

Main idea: Go from BFS to BFS, attempt to increase the objective value along the way.

1. Start with a feasible basis B.

   (It is nontrivial to find a basic feasible solution, we will discuss this later in "2-phase simplex"

2. Rewrite the LP in canonical form w.r.t B

3. Pick an entering variable

   To increase the objective value , pick a non-basic variable x_k where $c_k > 0$ in the objective function. If there are multiple choices of k, we can use Bland's rule which is to pick x_k with smallest index k.

   If there are no $c_k > 0$, then our BFS is optimal. Stop

4. Pick a leaving variable.

   To maintain feasibility (non-negativity), we need to first pick some index i such that $A_{ik} > 0$. Also we want the ration $\dfrac{b_i}{A_{ik}}$ to be minumum among all such i. So we pick

   $$j = \arg\min_i \{\frac{b_i}{A_{ik}} : A_{ik} > 0\}$$

   We then pick x_l corresponding to the jth row A_j as the leaving variable. i.e., we pick index $l \in B$ such that $A_{jl} = 1$.

   If there are multiple choice of l, we pick the smallest one (Bland's rule).

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

5. Repeat step 2.

Example

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

1. The initial basis $B = \{4, 5, 6\}$, with BFS $x = (0, 0, 0, 5, 3, 5)^T$.

2. This is in canonical form. We pick $c_k > 0$ with smallest k from the nonbasic variables. So we pick $c_2 = 2 > 0$ so the entering variable x_2.

   Now we need to pick a leaving variable, so we look at column 2 which is $A_2 = (-1, 1, 1)^T$, the positive elements in A_2 are $A_{22} = 1$ and $A_{32} = 1$. we look for the index j with the minimum ratio ($\min\{\frac{b_2}{A_{22}}, \frac{b_3}{A_{32}}\} = \min\{3, 5\}$), so index $j = 2$ achieves the minimum . The basic variable corresponds to the jth row is x_5. So x_5 is the leaving variable.

   Hence the new basis is $\{2, 4, 6\}$

After some elementary row operations (We use A_{22} to do Gaussian eliminations), the canonical form for the new basis $B = \{2, 4, 6\}$ is

$$
\begin{array}{ll}
\max & (4, 0, 1, 0, -2, 0)x + 6 \\
\text{s.t.} & \\
& \begin{pmatrix} -1 & 0 & 1 & 1 & 1 & 0 \\ -2 & 1 & 0 & 0 & 1 & 0 \\ 2 & 0 & -2 & 0 & -1 & 1 \end{pmatrix} x = \begin{pmatrix} 8 \\ 3 \\ 2 \end{pmatrix} \\
& x \geq \mathbb{0}
\end{array}
$$

3. We then again pick $c_k > 0$ with smallest k from the non-basic variables which is $c_1 = 4 > 0$. So x_1 is the entering variable.

   We then need to pick a positive pivot element from $A_1 = (-1, -2, 2)^T$. We compute the ratio $\min\{\cdot, \cdot, \frac{2}{2}\}$, so A_{31} is chosen as the pivot element, the corresponding basic variable is x_6, so x_6 is the leaving variable.

   The new basis is therefore $\{1, 2, 4\}$

   We then use $A_{31} = 2$ as the pivot element to do Gaussian eliminations to reduce the LP into canonical form for the new basis $\{1, 2, 4\}$ which is

$$
\begin{array}{ll}
\max & (0, 0, 5, 0, 0, -2)x + 10 \\
\text{s.t.} & \\
& \begin{pmatrix} 0 & 0 & 0 & 1 & 1/2 & 1/2 \\ 0 & 1 & -2 & 0 & 0 & 1 \\ 1 & 0 & -1 & 0 & -1/2 & 1/2 \end{pmatrix} x = \begin{pmatrix} 9 \\ 5 \\ 1 \end{pmatrix} \\
& x \geq \mathbb{0}
\end{array}
$$

4. We then pick up x_3 as the entering variable ($c_3 = 5 > 0$). However, $A_3 = (0, -2, -1)^T$ is non-positive. This means the LP is unbounded. Why?

   We let $x_3 = t, x_5 = 0, x_6 = 0$, then from the constraints we have

$$x_1 = 1 + t, x_2 = 5 + 2t, x_4 = 9$$

   and the objective is $5t + 10$. So as $t \to +\infty$, we have $x_1 > 0, x_2 > 0, x_3 > 0, x_4 = 9, x_5 = x_6 = 0$, but the objective goes to infinity.

   In fact the basic feasible solution $\bar{x} = (1, 5, 0, 9, 0, 0)^T$ and $d = (1, 2, 1, 0, 0, 0)^T$ is a certificate of unboundedness. (check)

CO 250 Winter 2022 Week 6

Class outline.

1. Termination of simplex
2. Two-phase Simplex method

19 Termination of simplex

Does simplex end in finitely many steps? If at each iteration we can increase the objective value. Then we will never visit the same BFS twice, therefore if objective values strictly increases at each iteration, then simplex terminiate in finitely many steps.
However, if the objective value does not increase, then simplex method can run in cycles, never end.

19.1 Degenerate iterations

A degenerate iteration is an iteration where objective value does not change, or BFS does not change. (The degenerate variables are zero)
Example

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

Basis $B = \{3,4\}$, BFS $(0,0,0,2)^T$. The degenerate basic variable is x_3
Entering variable x_1.
Leaving variable x_3.
New basis $B = \{1,4\}$. BFS $(0,0,0,2)$.
No change in the objective value or BFS. This is a degenerate iteration.

19.2 Cycling

With unfortunate choices of entering/leaving variables, we may get a series of degenerate iterations that repeat a basis already used. This may cycle through indefinitely.
Example:

$$\begin{array}{ll} \max & 10x_1 - 57x_2 - 9x_3 - 24x_4 \\ \text{s.t.} & \frac{1}{2}x_1 - \frac{11}{2}x_2 - \frac{5}{2}x_3 + 9x_4 + s_1 = 0 \\ & \frac{1}{2}x_1 - \frac{3}{2}x_2 - \frac{1}{2}x_3 + x_4 + s_2 = 0 \\ & x_1 + x_2 + x_3 + x_4 + s_3 = 1 \\ & x_1, x_2, x_3, x_4, s_1, s_2, s_3 \geq 0 \end{array}$$

The above LP as a series of basis which form a cycle (exercise)

$$(s_1,s_2,s_3) \to (x_1,s_2,s_3) \to (x_1,x_2,s_3) \to (x_2,x_3,s_3) \to \\ (x_3,x_4,s_3) \to (x_4,s_1,s_3) \to (s_1,s_2,s_3)$$

19.3 prevent cycling

Theorem 3. *Bland's rule prevents cycling, guaranteeing that simplex terminates in finitely many steps. (Bland's rule does not prevent degenerate iteration)*

Sketch of Proof (additional reading, not required for the class): Suppose $B_1, ..., B_m, B_1$ is a cycle, We first define flickle variables as those variables entering or leaving some basis in the cycle. If x_j is a flickle variable, then $x_j = 0$ in each BFS during the cycle.
We let x_t to be the largest "flickle" variable. Assume $x_t \in B_i$ but $x_t \notin B_{i+1}$. Assume s is the entering variable from B_i to B_{i+1}. Thus $s < t$ (x_s and x_t are both flickle variables).
Let

$$\begin{aligned} x_s &= q \\ x_i &= 0, i \notin B_i,\ i \neq s \\ x_k &= b_{l_k} - A_{l_k,s}q, \quad \forall k \in B_i \end{aligned}$$

be a Not Necessary Feasible solution corresponding to the dictionary of B_i. The objective value from the dictionary of B_i is

$$v = c_s q.$$

Assume x_t reenters at B_f, then we can plug the above solution into the SEF corresponding to B_f, and the objective value should stay the same which is

$$v = c_s^f q + \sum_{k \in B_i} c_k^f (b_{l_k} - A_{l_k,s}q)$$

By equating the two objective values, we have

$$c_s q = c_s^f q + \sum_{k \in B_i} c_k^f (b_{l_k} - A_{l_k,s}q)$$

which is

$$(c_s - c_s^f + \sum_{k \in B_i} c_k^f A_{l_k,s})q = \sum_{k \in B_i} c_k^f b_{l_k}$$

This holds true for any q, hence

$$c_s - c_s^f + \sum_{k \in B_i} c_k^f A_{l_k,s} = 0$$

We know $c_s > 0$ since s is the entering variable at B_i. Also we know $c_s^f \leq 0$ (if $c_s^f > 0$, then x_s enters instead of x_t at B_f by Bland's rule), therefore

$$\sum_{k \in B_i} c_k^f A_{l_k,s} < 0$$

Hence there exists some $x_r \in B_i$ such that $c^f_r A_{l_r,s} < 0$.
We know $c^f_t > 0$ (since x_t enters at B_f). Also we know $A_{l_t,s} > 0$ (since x_t leaves at B_i), so $c^f_t A_{l_t,s} > 0$.
Therefore we conclude $r \neq t$.
Also since $c^f_r A_{l_r,s} < 0$, we know $c^f_r \neq 0$, so $x_r \notin B_f$. But $x_r \in B_i$, so x_r is a flickle variable. This means $r < t$ since t is the largest flickle variable.
Hence $c^f_r < 0$ (otherwise x_r enters at B_f instead of x_t by Bland's rule), hence $A_{l_r,s} > 0$.
We know $b_{lr} = 0$ since x_r is a flickle variable and $x_r \in B_i$. And $r < t$, hence x_r leaves B_i instead x_t by Bland's rule, a contradiction. Therefore cycle does not exist.

19.4 Two phase simplex

At the start of simplex algorithm, we need a basic feasible solution, what if there is no obvious BFS?
Example:

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

$\{3,4,5\}$ is a basis, but its basic solution $(0,0,-1,-4,4)^T$ is not feasible.
How to find a BFS? We ignore the objective, we form an auxiliary problem.

1. Form an <u>auxiliary LP</u> with an obvious BFS.
2. solve this auxillary problem using simplex method, the result of simplex will determine if the original LP is feasible (has a BFS) or not.

How to form an auxiliary problem with an obvious BFS.
Step 1: For constraints wholse RHS is negatie, multiply by -1.

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

Step 2: Add new auxiliary variables, one for each constraint. They form a basis that is feasible.
Add x_6, x_7, x_8

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

Basis is $\{6,7,8\}$ with BFS $(0,0,0,0,0,1,4,4)^T$ Key observation:
The original LP has a feasible solution if and only if the auxillary LP has a feasible solution where all auxiliary variables are 0.
Step 3: Add new objective function to maximize the sum of the negatives of all auxiliary variables.

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

Note: the auxiliary LP is feasible and bounded by 0. Hence it must have an optimal solution.

Proposition 3. *An LP has a feasible solution or BFS if and only if its auxiliary LP (maximizing the negative sum of auxiliary variables) has optimal value 0.*

Two-phase simplex summary

1. Run simplex on the auxiliary LP.
   - Optimal value less than 0: Original LP is infeasible.
   - Optimal value is 0: The original LP is feasible and we have an BFS for the original LP.
2. Run simplex for the original LP starting with the BFS.
   - The optimal solution (optimal BFS) of the original LP is obtained.
   - The original LP is unbounded.

The auxiliary LP

$$\begin{array}{ll} & \text{Auxiliary LP} \\ \max & (0,0,0,0,0,-1,-1,-1)^T x \\ \text{s.t.} & \begin{pmatrix} 1 & -1 & -1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 2 & 0 & -1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \end{pmatrix} x = \begin{pmatrix} 1 \\ 4 \\ 4 \end{pmatrix} \\ & x \geq 0 \end{array}$$

has a basis $6,7,8$ with an obvious BFS $(0,0,0,0,0,1,4,4)^T$. Muptiply constraints by $(1,1,1)$ to get $(2,2,-1,-1,1,1,1,1)x-9=0$. The new objective is $(2,2,-1,-1,1,0,0,0)$

9. We run simplex on the auxiliary LP and obtain

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

with Basis $\{1,3,5\}$ and BFS $(4,0,3,0,4,0,0,0)^T$ and optimal value 0.
Therefore we concluded that the original LP is feasible $(4,0,3,0,4)^T$ is a BFS with basis $\{1,3,5\}$.
Phase II:
Solve original LP with a starting basis $\{1,3,5\}$. Run simplex. Results:

$$\begin{array}{ll} \max & (0,0,-\frac{8}{3},-\frac{1}{3},0)^T x - 4 \\ \text{s.t.} & \begin{pmatrix} 1 & 0 & -\frac{2}{3} & -\frac{1}{3} & 0 \\ 0 & 1 & \frac{1}{3} & -\frac{1}{3} & 0 \\ 0 & 0 & -\frac{1}{3} & \frac{1}{3} & 1 \end{pmatrix} x = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} \\ & x \geq 0 \end{array}$$

Optimal solution (BFS) is $(2,1,0,0,3)^T$, with optimal value -4.
Example:

LP

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

Auxiliary LP

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

Auxiliary LP in canonical form

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

Run simplex on auxiliary LP, results:

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

Basis $\{1,6\}$. BFS is $\{\frac{5}{2},0,0,0,0,\frac{1}{2}\}$.So the optimal value is $-\frac{1}{2<0}$. So the original LP is infeasible.

Certificate of infeasibility Recall: a certificate of infeasibility is a vector y where $y^T A \geq 0$ and $y^T b < 0$.
The certificate $y = A_B^{-T} c_B$ where $B = \{1,6\}$.

$$A_B = \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix} \quad c_B = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$$

Therefore

$$y = \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix}^{-T} \begin{pmatrix} 0 \\ -1 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ -1 \end{pmatrix}$$

Multiply y^T to the left of the (original) constraints:

$$(0, \frac{7}{2}, \frac{1}{2}, 1)^T x = -\frac{1}{2}$$

Exercise: show that $y = A_B^{-T} c_B$ is a certificate of infeasibility for the original LP is B is the optimal basis for the Auxiliary problem with optimal value less than 0.

19.5 Tableau method

Tableau method is suitable for doing simplex by hand, we use a matrix or tableau to represent all data in the LP. We then do row reductions on that tableau/matrix as we did in linear algebra course.
Example:

$$\begin{array}{ll} \max & z = (3,-2,0,0,0)^T x + 0 \\ \text{s.t.} & \begin{pmatrix} 4 & -1 & 1 & 0 & 0 \\ 3 & -3 & 0 & 1 & 0 \\ -2 & 2 & 0 & 0 & 1 \end{pmatrix} x = \begin{pmatrix} 8 \\ 9 \\ 1 \end{pmatrix} \end{array}$$

From the objective function we have $z - (3,-2,0,0,0)^T x = 0$.
The tableau is the following:

$$\left(\begin{array}{c|ccccc|c} 1 & -3 & 2 & 0 & 0 & 0 & 0 \\ \hline 0 & 4 & -1 & 1 & 0 & 0 & 8 \\ 0 & 3 & -3 & 0 & 1 & 0 & 9 \\ 0 & -2 & 2 & 0 & 0 & 1 & 1 \end{array}\right)$$

The current basis is $\{3,4,5\}$

1. Entering variable: Instead of looking for positive coefficient, we now look for negative coefficient in the first row of the tableau. The coefficient of x_1 is -3. So entering variable is x_1.

2. leaving variable: we look for the minimum ratio in the column corresponding to x_1. Since $\min\{\frac{8}{4}, \frac{9}{3}\} = \frac{8}{4}$. So 4 is the pivot element, the corresponding basic variable is x_3, so x_3 leaves.

3. The new basic is $\{1, 4, 5\}$. Update the tableau by elementary row operations to get the canonical form:

$$\left(\begin{array}{c|cccccc|c} 1 & 0 & \frac{5}{4} & \frac{3}{4} & 0 & 0 & 6 \\ \hline 0 & 1 & -\frac{1}{4} & \frac{1}{4} & 0 & 0 & 2 \\ 0 & 0 & -\frac{9}{4} & -\frac{3}{4} & 1 & 0 & 3 \\ 0 & 0 & \frac{3}{2} & \frac{1}{2} & 0 & 1 & 5 \end{array}\right)$$

   So every element in the first row of the tableau is nonnegative. The tableau is optimal, with optimal BFS being $(2, 0, 0, 3, 5)^T$. The optimal value is 6.

Advantages and disdvantages of Tableau method:

- Advantage: Tableau method is fast for computation by hand.
- Disadvantage: when doing tableau method on a computer, the rounding errors would accumulate. To avoid accumulation of rounding errors, each time when we get a new basis B, we can recalculate A_B^{-1} and $y = A_B^{-1} c_B$ using the original data to convert the LP into canonical form corresponding to basis B (instead of using Tableau from the previous basis).

20 Geometry of of simplex and linear program

Definition 6 (Informal definition of extreme point)**.** *Roughly speaking, extreme point of an LP is a point in the "vertex" of the feasible region*

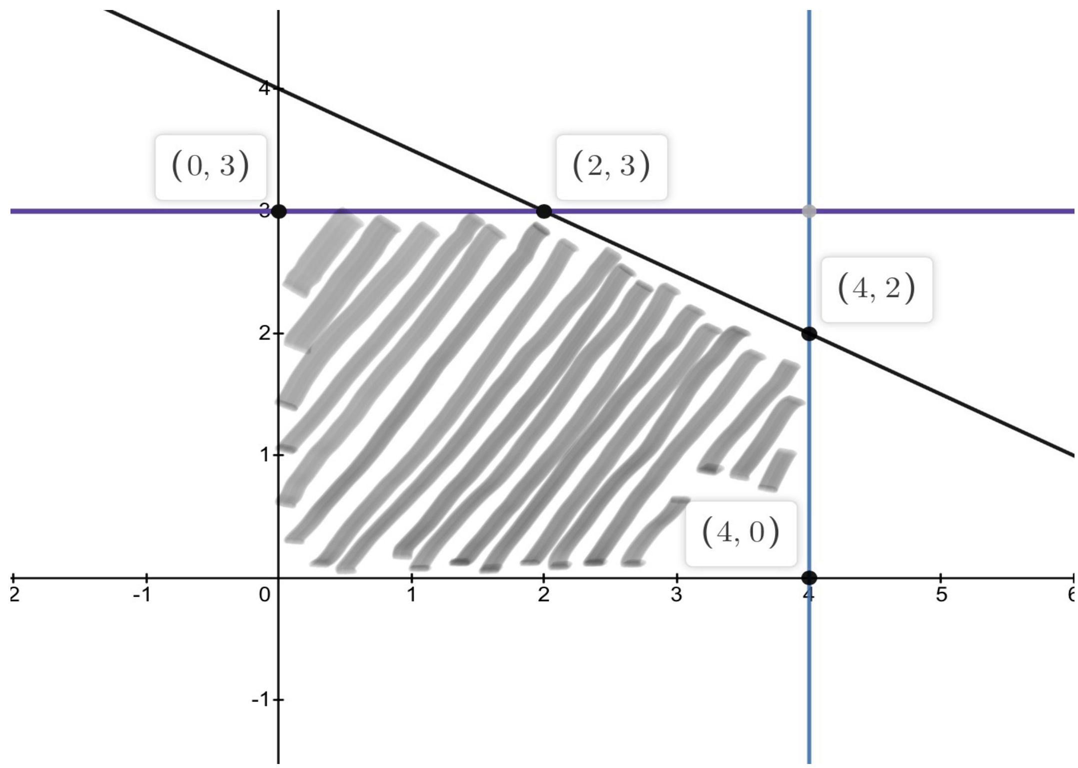
Key concept: If there is an optimal solution to an LP, then there is an optimal solution that is an extreme point.
Example:

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

We add auxiliary variables to get SEF for the above LP

$$\begin{array}{ll} \max & (1,1,0,0,0)x \\ \text{s.t.} & \begin{pmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 2 & 0 & 0 & 1 \end{pmatrix} x = \begin{pmatrix} 4 \\ 3 \\ 8 \end{pmatrix} \\ & x \geq 0 \end{array} \tag{2}$$

The feasible region in $\mathbb{R}^2$



There are 5 extreme points in the feasible region. Each extreme point corresponds to two equality constraints. Among the 5 inequalities in (1), we obtain an extreme point by setting two inequalities to equalities, this is equivalent to setting two non-basic variable in SEF to zero.
The key observation: setting a variable to 0 in SEF (2) is equivalent to setting an inequality to equality in (1)

1. For $(0,0)$, we set $x_1 = 0, x_2 = 0$ in SEF, so the basis in SEF is $\{3,4,5\}$
2. For $(0,3)$, we set $x_1 = 0$ and $x_4 = 0$ in SEF, so the basis in SEF is $\{2,3,5\}$.
3. For $(2,3)$, we set $x_4 = 0, x_5 = 0$, so the basis in SEF is $\{1,2,3\}$
4. For $(4,2)$, we set $x_3 = 0, x_5 = 0$, so the basis is $\{1,2,4\}$
5. For $(4,0)$, we set $x_2 = 0, x_3 = 0$, so the basis is $\{1,4,5\}$

Observation: extreme point of the feasible region of LP (1) corresponds to basic feasible solutions in LP (2).

Definition 7. *An inequality that is satisfied with equality by $\bar{x}$ is a tight constrait.*

Remark 1 (degeneracy)**.** *In the above example, we have 2 tight constraints in $\mathbb{R}^2$ for each extreme point (corresponding to 2 non-basic variables being set to 0), but it is possible to have more than 2 (or in general n in $\mathbb{R}^n$) tight constraints for an extreme point. In this case, some basic variables have to be equal to 0, the corresponding BFS is degenerate (multiple bases have the same BFS). See figure 3*

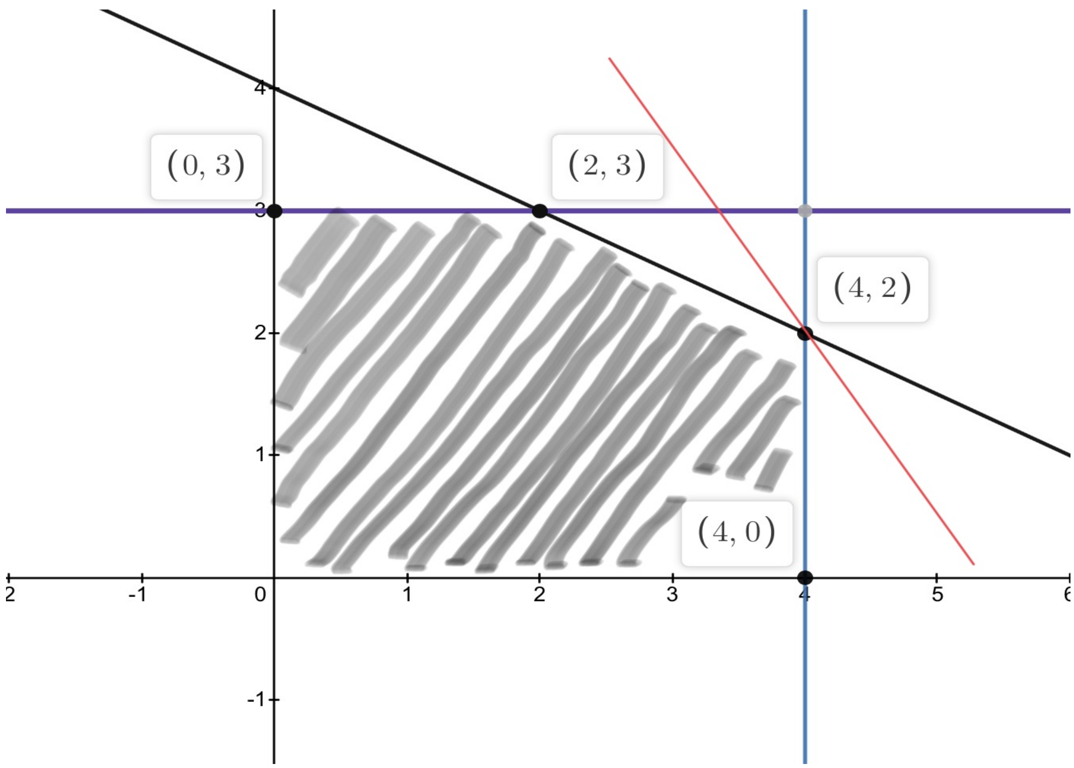


Figure 3: $(4, 2)$ is generate if there is a third constraint passing through it

.

Geometry of simplex

Key concept: simplex method moves from basis to basis. Geometrically, simplex is moving from one extreme point to a neighbouring extreme point, in the direction of the objective function.

CO 250 Winter 2022 Week 7

Class outline.

1. Geometry of Linear program and simplex
2. Duality theory

21 Geometry of of simplex and linear program

Definition 8 (Informal definition of extreme point)**.** *Roughly speaking, extreme point of an LP is a point in the "vertex" of the feasible region*

Geometry of simplex

Key concept:

- Simplex method moves from basis to basis. Geometrically, simplex is moving from one extreme point to a neighbouring extreme point, in the direction of the objective function.
- If there is an optimal solution to an LP, then there is an optimal solution that is an extreme point.

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

We add auxiliary variables to get SEF for the above LP

$$\begin{aligned} \max \quad & (1,1,0,0,0)x \\ \text{s.t.} \quad & \begin{pmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 2 & 0 & 0 & 1 \end{pmatrix} x = \begin{pmatrix} 4 \\ 3 \\ 8 \end{pmatrix} \\ & x \geq 0 \end{aligned} \tag{4}$$

We start at extreme point $x_1 = 0, x_2 = 0$, which corresponds to BFS $(0,0,4,3,8)$. The corresponding basis $B = \{3,4,5\}$.
The corresponding tableau:

$$\begin{array}{ccccc|c} 1 & 1 & 0 & 0 & 0 & 0 \\ \hline 1 & 0 & 1 & 0 & 0 & 4 \\ 0 & 1 & 0 & 1 & 0 & 3 \\ 1 & 2 & 0 & 0 & 1 & 8 \end{array}$$

So x_1 enters and x_3 leaves. The new basis is $\{1,4,5\}$, with extreme point $x_1 = 4, x_2 = 0$ and BFS $(4,0,0,3,4)^T$. The tableau for the canonical form is:

$$\begin{array}{ccccc|c} 0 & 1 & -1 & 0 & 0 & -4 \\ \hline 1 & 0 & 1 & 0 & 0 & 4 \\ 0 & 1 & 0 & 1 & 0 & 3 \\ 0 & 2 & -1 & 0 & 1 & 4 \end{array}$$

So x_2 enters and x_5 leaves, the new basis is $\{1,2,4\}$, with extreme point $x_1 = 4, x_2 = 2$ and BFS $(4,2,0,1,0)^T$. The tableau for the canonical form is:

$$\begin{array}{ccccc|c}
0 & 0 & -1/2 & 0 & -1/2 & -6 \\
\hline
1 & 0 & 1 & 0 & 0 & 4 \\
0 & 0 & 1/2 & 1 & -1/2 & 1 \\
0 & 1 & -1/2 & 0 & 1/2 & 2
\end{array}$$

The tableau is optimal with optimal value being $-(-6) = 6$.
Observation: In each iteration, we swap one tight constraint for another, we move to a "neighbouring" extreme point.
Degenerate iteration: Swapping tight constraints for the same BFS.
Suppose we add a new constraint $x_1 + x_2 = 0$, so three constraints passing through $(0,0)$. We start at extreme point $x_1 = 0, x_2 = 0$ with basis $B = \{3,4,5,6\}$. The corresponding tableau is

$$\begin{array}{cccccc|c}
1 & 1 & 0 & 0 & 0 & 0 & 0 \\
\hline
1 & 0 & 1 & 0 & 0 & 0 & 4 \\
0 & 1 & 0 & 1 & 0 & 0 & 3 \\
1 & 2 & 0 & 0 & 1 & 0 & 8 \\
1 & 1 & 0 & 0 & 0 & 1 & 0
\end{array}$$

After one step of simplex iteration, x_1 enters, x_5 leaves, the tight constraints are $x_1 + x_2 = 0$ and $x_2 = 0$, so we swapped constraint $x_1 = 0$ with constraint $x_1 + x_2 = 0$.

22 Geometry of Linear Program and Simplex: theory

22.1 polyhedron

We consider constraints of the form $Ax \leq b$ where $x \in \mathbb{R}^n$.

- An equation of the form $a^T x = \beta$ is a hyperplane in $\mathbb{R}^n$.
- An inequality of the form $a^T x \leq \beta$ is a halfspace in $\mathbb{R}^n$
- $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ is a polyhedron.

Note: A polyhedron is the intersection of halfspaces.

22.2 Convexity

A set of points in $\mathbb{R}^n$ is convex if the line segment joining any two points in the set is also contained in the set.
Pictures of convex set and non-convex sets

Definition 9. *A set $C \subseteq \mathbb{R}^n$ is* <u>*convex*</u> *if for any $x, y \in C$, we have $\lambda x + (1-\lambda)y \in C$ for all $\lambda \in [0,1]$. Such a vector is a* <u>*convex combination*</u> *of x and y.*

Proposition 4. *A hyperplane is convex*

Proposition 5. *A halfspace is convex*

Proposition 6. *The intersection of convex sets is convex*

Proposition 7. *A polyhedron is convex*

22.3 Extreme point

Definition 10. *Let C be a convex set. Then $\bar{x} \in C$ is an* <u>*extreme point*</u> *if $\bar{x}$ is not a* ***proper*** *convex combination of any two other points in C. In other words, $\bar{x}$ is not* ***properly*** *contained in any line segment of C. (there do not exist any $x, y \in C$ and $\lambda \in (0,1)$ where $\bar{x} = \lambda x + (1-\lambda)y$.)*

Remark 2. *$x \in C$ is not an extreme point of C if and only if*

$$x = \lambda x + (1-\lambda)y$$

for distinct $x, y \in C$ and λ with $0 < \lambda < 1$.

22.4 Charaterization of extreme points of polyhedron

Theorem 4. *Let $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ be a polyhedron, and let $\bar{x} \in P$. Let $A^{=}x = b^{=}$ be the set of all tight constraints for $\bar{x}$. Then $\bar{x}$ is an extreme point of P if and only if rank$(A^{=}) = n$.*

Proof. Suppose that rank$(A^{=}) = n$. We will show that x is an extreme point. Suppose for a contradiction this is not the case, Then there exist $x^{(1)}, x^{(2)} \in P$, where $x^{(1)} \neq x^{(2)}$ and λ where $0 < \lambda < 1$ for which $x = \lambda x^{(1)} + (1-\lambda)x^{(2)}$. So $A^{=}x^{(1)} \leq b^{=}$ and $A^{=}x^{(2)} \leq b^{=}$. Thus

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

Therefore we have $b^{=} \leq b^{=}$ which means every inequality should hold as equality. Hence $A^{=}x^{(1)} = b$ and $A^{=}x^{(2)} = b$. Because rank $(A^{=}) = n$, the linear system $A^{=}x = b$ has a unique solution, thus $x^{(1)} = x^{(2)}$, a contradiction.
To prove the other direction, let $\bar{x}$ be an extreme point of P, we want to show rank $(A^{=}) = n$. Suppose to the contrary rank $(A^{=}) < n$, then there exists some vector d such that $A^{=}d = 0$. We now let $x^{(1)} = \bar{x} - \epsilon d$ and $x^{(2)} = \bar{x} + \epsilon d$, it is easy to check $\bar{x} = \frac{1}{2}x^{(1)} + \frac{1}{2}x^{(1)}$.

It is easy to check that $A^{=}x^{(1)} \le b^{=}$ and $A^{=}x^{(2)} \le b^{=}$.
Now partition A as $A = \begin{bmatrix} A^{=} \\ A^{\neq} \end{bmatrix}$ where $A^{\neq}$ corresponds to the constraints which are not tight, i.e., $A^{\neq}\bar{x} < b^{\neq}$
So $A^{\neq}x^{(1)} = A^{\neq}\bar{x} + \epsilon A^{\neq}d$. Since $A^{\neq}\bar{x} < b^{\neq}$, we know $A^{\neq}x^{(1)} \le b^{\neq}$ for sufficiently small ϵ, similarly $A^{\neq}x^{(2)} \le b^{\neq}$ for sufficiently small ϵ too. Therefore $Ax^{(1)} \le b$ and $Ax^{(2)} \le b$ for sufficiently small ϵ, so $x^{(1)} \in P$ and $x^{(2)} \in P$, $\bar{x}$ is properly contained in some line segment of P, which means $\bar{x}$ is not an extreme point, a contradiction. □

Proposition 8. *Let $P = \{x \in \mathbb{R}^n : Ax = b\}$ and A has linearly independent rows, then $\bar{x}$ is a BFS of P if and only if $\bar{x}$ is an extreme point of P.*

Proof. Exercise (use Theorem 4). □

23 Duality Theory of LP

Recall the LP in SEF

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

Our goal is to derive an certificate of optimality for this LP.
The idea is to derive an upper bound for $c^T x$ such that the upper bound is tight (It is attained by a feasible solution).
Example

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

Choose $B = \{1, 2\}$ as a basis, the canonical form is

$$\begin{array}{ll} \max & (0, 0, -8/3)x + 1 \\ \text{s.t.} & \begin{pmatrix} 1 & 0 & -2/3 \\ 0 & 1 & -5/3 \end{pmatrix} x = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \\ & x \ge 0 \end{array}$$

So basis $B = \{1, 2\}$ is optimal. The optimal solution being $(2, 1, 0)^T$ with optimal value 1.

To certify the optimality, we let $y = (8/3, -1/3)^T$ and multiply the constraint equations by y^T, we get

$$(8/3, -1/3)\begin{pmatrix} 1 & -1 & 1 \\ 2 & 1 & -3 \end{pmatrix} x = (8/3, -1/3)\begin{pmatrix} 1 \\ 5 \end{pmatrix}$$

which gives

$$(2, -3, 11/3)x = 1$$

However we know $2x_1 - 3x_2 + \frac{11}{3}x_3 \geq 2x_1 - 3x_2 + x_3$ since x_1, x_2, x_3 are all non-negative. Hence $2x_1 - 3x_2 + x_3 \leq 1$ for any feasible x, so 1 is the upper bound of our LP, and this upper bound is attained in by solution $(2, 1, 0)^T$ so we can certify that $(2, 1, 0)^T$ is indeed optimal.
In general, consider the an LP in SEF

$$(P) \qquad \begin{array}{cc} \max & c^T x \\ & Ax = b \\ & x \geq 0 \end{array}$$

If we can find some y such that

$$y^T A \geq c^T.$$

Since $x \geq 0$, by multiplying x, we have

$$y^T Ax \geq c^T x.$$

Since $Ax = b$, we have

$$y^T b \geq c^T x.$$

So $b^T y$ is an upper bound for the optimal value of (P).
In general, to obtain an upper bound, we try to find y such that $A^T y \geq c$ and compute $b^T y$. To make the upper bound as tight as possible, we solve the following LP,

$$(D) \qquad \begin{array}{cc} \min & b^T y \\ \text{s.t.} & A^T y \geq c \\ & y \text{ is free} \end{array}$$

This LP (D) is called the dual linear program.

CO 250 Winter 2022 Week 8

Class outline.

1. Weak Duality
2. Strong duality

24 Weak Duality

In general, consider the an LP in SEF

$$
(P) \qquad \begin{array}{rl} p^* = \max & c^T x \\ & Ax = b \\ & x \geq 0 \end{array}
$$

We call this primal problem

$$
(D) \qquad \begin{array}{rl} d^* = \min & b^T y \\ \text{s.t.} & A^T y \geq c \\ & y \text{ is free} \end{array}
$$

This LP (D) is called the dual linear program.
Together (P) and (D) is called a primal-dual pair.

Theorem 5. *Assume both (P) and (D) have an optimal solution, then $p^* \leq d^*$*

Proof. Let y be any feasible solution of (D) and x be any feasible solution of (D). Then

$$
\begin{aligned}
b^T y &= (Ax)^T y \\
&= x^T A^T y \\
&\geq x^T c \text{ (since } x \geq 0 \text{ and } A^T y \geq c) \\
&= c^T x
\end{aligned}
$$

□

1. Weak duality says any feasible solution of the primal problem (P) gives a lower bound of the optimal value of the dual problem (D), and any feasible solution of the dual problem (D) gives an upper bound of the primal problem (P).

2. If the primal problem (P) is unbounded (above), then the dual problem (D) must be infeasible. Otherwise every feasible solution of the dual problem (D) will give an upper bound for the primal.

3. Similarly if the dual problem (D) is unbounded (below), then the primal problem (P) must be infeasible. Otherwise every feasible solution of the primal problem (P) will yield a lower bound for the dual.

4. If the primal is infeasible, then we say $p^* = -\infty$. If the dual is infeasible, we say $d^* = +\infty$. In summary, in all cases, we always have $p^* \leq d^*$.

Example
Consider the primal LP

$$\begin{array}{ll} \max & (1,1)x \\ \text{s.t.} & x_1 - x_2 = 1 \\ x \geq \mathbb{0} & \end{array}$$

This LP is unbounded, i.e., $x_1 = 10000, x_2 = 9999$
The dual problem is given as

$$\begin{array}{ll} \min & y \\ & y \geq 1 \\ & -y \geq 1 \end{array}$$

So in the dual we have $y \geq 1$ and $y \leq -1$ which is infeasible.

24.1 Different formulation of primal and dual

If the primal problem is given as

$$(P) \quad \begin{array}{ll} p^* = \max & c^T x \\ & Ax \leq b \\ & x \geq 0 \end{array}$$

Then the dual problem is given as

$$(D) \quad \begin{array}{ll} d^* = \min & b^T y \\ \text{s.t.} & A^T y \geq c \\ & y \geq 0 \end{array}$$

This can be derived by adding the auxiliary variable to make the primal into SEF and write down the dual. (Exercise)

24.2 dual of dual is primal

Consider the primal-dual pair,

$$(P) \quad \begin{array}{ll} p^* = \max & c^T x \\ & Ax = b \\ & x \geq 0 \end{array}, \quad (D) \quad \begin{array}{ll} d^* = \min & b^T y \\ \text{s.t.} & A^T y \geq c \\ & y \text{ free} \end{array}$$

We add slack variables and convert a free variable to two nonnegative variable to the dual to get

$$\begin{array}{ll} d^* = \min & b^T y \\ \text{s.t.} & A^T(y^+ - y^-) - z = c \\ & y^+, y^-, z \geq 0 \end{array}$$

Then

$$\begin{aligned} -d^* = \max \quad & -b^T y \\ \text{s.t.} \quad & (A^T, -A^T, -I) \begin{pmatrix} y^+ \\ y^- \\ z \end{pmatrix} = c \\ & y^+, y^-, z \geq 0 \end{aligned}$$

The rest is just exercise

25 Strong duality

Theorem 6 (Strong duality)**.** *If either primal (P) or dual (D) are feasible, then* $p^* = d^*$*.*

Proof. Since the dual of dual is primal, we can just assume (D) is feasible (the case of (P) being feasible can be proved similarly.
Then the proof of this theorem has two parts:

1. If the primal (P) is feasible, then $p^* = d^*$ (both has optimal solution).
2. If the primal (P) is infeasible, then $p^* = d^* = -\infty$ (dual (D) is unbounded)

We will prove part (2) first, in fact, you already proved it in your assignment (A6-2 part(c)). We run simplex, since (P) is infeasible, from Phase I of simplex, we get $y = A_B^{-T} c_B$ which is a certificate of infeasibility, so $A^T y \leq 0, b^T y > 0$, this shows that the dual problem must be unbounded if the dual problem is feasible.

To prove part (1), we assume the primal (P) is feasible, since the dual is also feasible, and the dual gives an upper bound, the primal (P) must have an optimal solution, and we run simplex, we know simplex method will terminate with an optimal BFS $\bar{x}$ and optimal basis B.
We let $\bar{y} = A_B^{-T} c_B$, then $c^T - \bar{y}^T A \leq \mathbb{0}^{\mathbb{T}}$ since B is the optimal basis (the coefficients of the objective function in the optimal canonical form must be nonpositive). This shows $\bar{y}$ is a feasible solution of the dual problem (D).
To show both optimal value are equal, we notice

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

Since B is a basis and $\bar{x}$ is a BFS, we know $(c^T - \bar{y}^T A)_i = 0$ for any $i \in B$ and $\bar{x}_j = 0$ for any $j \in N$, so $(c^T - \bar{y}^T A)\bar{x} = 0$, therefore $c^T \bar{x} = b^T \bar{y}$, and by weak duality, we have $p^* = d^*$. $\square$

We summarize strong duality and weak duality

primal\ dual	O	U	I
O	✓	×	×
U	×	×	✓
I	×	✓	✓

25.1 Example of both and dual are infeasible

The following example shows a primal-dual pair that both are infeasible

$$\begin{array}{lll} (P) & \max & x_1 - x_2 \\ & \text{s.t.} & \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \leq \begin{pmatrix} 1 \\ -1 \end{pmatrix} \\ & & x_1, x_2 \geq 0 \end{array}$$

$$\begin{array}{lll} (D) & \min & y_1 - y_2 \\ & \text{s.t.} & \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \geq \begin{pmatrix} 1 \\ -1 \end{pmatrix} \\ & & y_1, y_2 \geq 0 \end{array}$$

26 Geometry of the duality and complementary slackness

We consider the following example

$$\begin{array}{llc} (P) & \max & (0,1)x \\ & \text{s.t.} & -x_1 + x_2 \leq 0 \\ & & -x_1 - x_2 \leq -1 \\ & & x_1, x_2 \geq 0 \end{array}$$

$$\begin{array}{llc} (D) & \min & (0,-1)y \\ & \text{s.t.} & -y_1 - y_2 \geq 0 \\ & & y_1 - y_2 \geq 1 \\ & & y_1, y_2 \geq 0 \end{array}$$

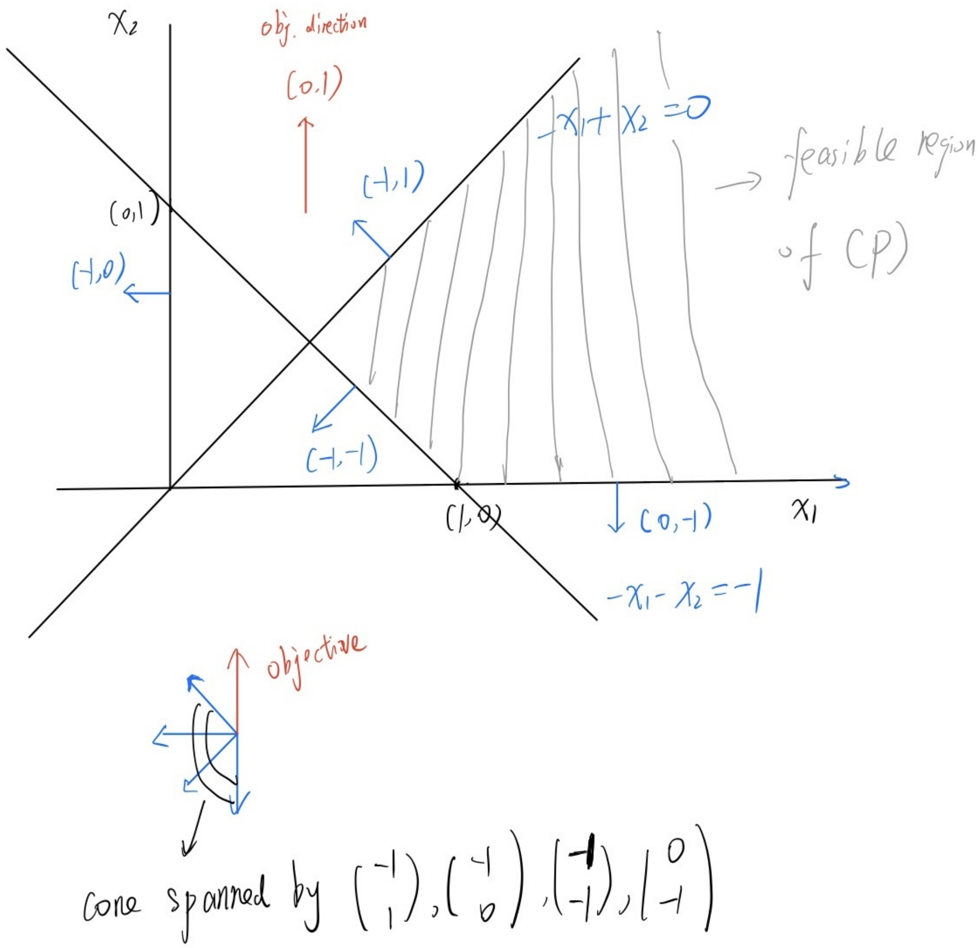
The primal problem is unbounded, that is because $(0,1)^T$ is not in the cone spanned by the normal vectors of constraints $(-1,1)^T, (-1,-1)^T, (-1,0)^T, (0,-1)^T$. The dual problem requires

$$y_1 \begin{pmatrix} -1 \\ 1 \end{pmatrix} + y_2 \begin{pmatrix} -1 \\ -1 \end{pmatrix} \geq \begin{pmatrix} 0 \\ 1 \end{pmatrix} \text{ for some } y_1 \geq 0, y_2 \geq 0$$

Equivalently,

$$y_1 \begin{pmatrix} -1 \\ 1 \end{pmatrix} + y_2 \begin{pmatrix} -1 \\ -1 \end{pmatrix} + y_3 \begin{pmatrix} -1 \\ 0 \end{pmatrix} + y_4 \begin{pmatrix} 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \text{ for some } y_1 \geq 0, y_2 \geq 0, y_3 \geq 0, y_4 \geq 0.$$

This is saying $(0,1)^T$ must lie in the cone spanned by the normal vectors of constraints $(-1,1)^T, (-1,-1)^T, (-1,0)^T, (0,-1)^T$. Hence unbouneedness of primal problem implies the infeasibility of dual problem.



26.1 Dual of other forms

Summary chart Primal

Primal: Max $c^T x$	Dual: Min $b^T y$
$\leq$ constraints	$\geq$ variables
$=$ constraints	free variables
$\geq$ constraints	$\leq$ variables
$\geq$ variables	$\geq$ constraints
free variables	$=$ constraints
$\leq$ variables	$\leq$ constraints

CO 250 Winter 2022 Week 9

Class outline.

1. Farkas Lemma
2. Geometry of optimality and strong duality
3. Complementary slackness

26.2 Dual of other forms

Summary chart Primal

Primal: Max $c^T x$	Dual: Min $b^T y$
$\leq$ constraints	$\geq$ variables
$=$ constraints	free variables
$\geq$ constraints	$\leq$ variables
$\geq$ variables	$\geq$ constraints
free variables	$=$ constraints
$\leq$ variables	$\leq$ constraints

Example:

$$\begin{array}{ll} \max & (3,2,4,1)x \\ \text{s.t.} & \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \\ 2 & 2 & 2 & 2 \end{pmatrix} \begin{array}{c} \geq \\ \leq \\ = \end{array} \begin{pmatrix} 3 \\ 4 \\ 5 \end{pmatrix} \\ & x_1, x_2 \geq 0, x_3 \leq 0, x_4 \text{ free} \end{array}$$

Dual:

$$\begin{array}{ll} \min & (3,4,5)y \\ \text{s.t.} & \begin{pmatrix} 1 & 4 & 2 \\ 2 & 3 & 2 \\ 3 & 2 & 2 \\ 4 & 1 & 2 \end{pmatrix} \begin{array}{c} \geq \\ \geq \\ \leq \\ = \end{array} \begin{pmatrix} 3 \\ 2 \\ 4 \\ 1 \end{pmatrix} \\ & y_1 \leq 0, y_2 \geq 0, y_3 \text{ free} \end{array}$$

27 Farkas Lemma

Theorem 7. *Exactly one of the following two conditions hold:*

- *There exists $x \geq 0$ such that $Ax = c$.*
- *There exists some vector d such that $d^T A \leq 0$ and $d^T c > 0$*

Geometric interpretation: Hyperplane seperation theorem.

Definition 11. *Let $v^1, v^2, \cdots, v^k \in \mathbf{R}^n$. The cone generated by $v^1, \cdots, v^k$ is*

$$\mathcal{C} = \{\lambda_1 v^1 + \cdots + \lambda_k x^k : \lambda_i \geq 0 \, \forall i\}$$

27.1 Geometry of optimality

Theorem 8. *Consider the following LP:*

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

Assume $c \neq \mathbb{0}^T$. Let $\bar{x}$ be a feasible solution, then $\bar{x}$ is optimal if and only if c is in the cone of normal vectors of tight constraints for $\bar{x}$.

Proof. Let $\bar{x}$ be a feasible solution, and let the tight constraints of $\bar{x}$ be $A^=$, so we can partition A as $A = \begin{pmatrix} A^= \\ A^< \end{pmatrix}$ such that $A^= \bar{x} = b^=$ and $A^< \bar{x} < b^<$. Suppose c is not in the cone of normal vectors of the tight constraints, i.e., there does not exists $y \geq 0$ such that $(A^=)^T y = c$. Then by Farkas Lemma, we know there exists d such that $d^T (A^=)^T \leq 0$ and $d^T c > 0$, which is equivalent to $(A^=)d \leq 0$ and $c^T d > 0$. So we let $x^* = \bar{x} + \epsilon d$ for some small $\epsilon > 0$, then x^* is feasible.
However $c^T x^* = c^T \bar{x} + \epsilon c^T d$ which is greater than $c^T \bar{x}$, contradiction to $\bar{x}$ being optimal solution.
For the other direction, let $\bar{x}$ be feasible and let c be in the cone of the normal vectors of tight constraints, then we have

$$c^T = y^T A^=$$

where $y \geq 0$.
Let $\bar{y} = (y, 0, \cdots, 0)$, then obviously $\bar{y}$ is feasible for the dual. Now we have $b^T \bar{y} = y^T (b^=) + \mathbb{0}^T (b^<) = y^T b^= = y^T (A^=) \bar{x} = c^T \bar{x}$. By weak duality $\bar{x}$ is optimal for the primal. □

An illustration of Theorem 8 is in Figure 4.

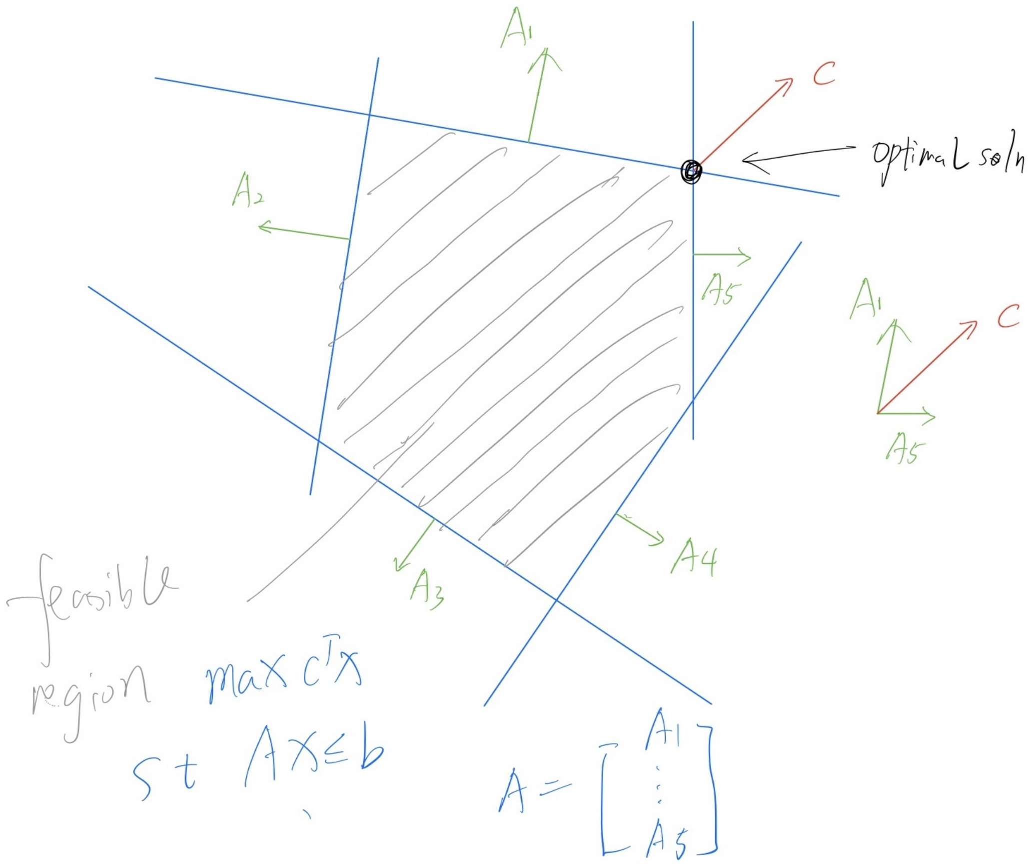


Figure 4: At the optimal vertex, c is in the cone generated by A_1 and A_5

28 Strong duality and complementary slackness

Remark 3 (Strong duality by Farkas Lemma)**.** *We have shown that if $\bar{x}$ is an optimal solution, then c must be in the cone of normal vectors of the tight constraints. Now we can easily construct a dual solution such that strong duality holds: Let $c = (A^=)^T y^=$ for some $y^= \geq 0$, we can then construct a feasible dual solution y by adding zeros to $y^=$. So let $y = \begin{pmatrix} y^= \\ 0 \end{pmatrix}$, and we have $c = A^T y$. So y is feasible for the dual problem:*

$$(D) \qquad \begin{array}{ll} \max & b^T y \\ s.t. & A^T y = c \\ & y \geq 0. \end{array}$$

Then $c^T \bar{x} = (A^T y)^T \bar{x} = y^T A \bar{x} = (y^=)^T (A^= \bar{x}) + 0^T (A^< \bar{x}) = (y^=)^T b^= + 0^T b^< = y^T b$, strong duality holds.

28.1 Complementary Slackness

Definition 12 (Complementary slackness conditions). *Given any primal-dual pair (P), (D), complementary slackness conditions are:*

- $x_i = 0$ *or the corresponding dual constraint is tight*
- $y_j = 0$ *or the corresponding primal constraint is tight.*

Theorem 9 (Complementary Slackness theorem). *Let $\bar{x}, \bar{y}$ be two feasible solutions to (P) and (D) repectively. Then $\bar{x}, \bar{y}$ are optimal for (P), (D) respectively if and only if $\bar{x}, \bar{y}$ satisfy all complementary slackness conditions.*

Proof. Omitted, exercise. $\square$

Example

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

$$
(D) \quad \begin{array}{ll} \min & (2,1,4)y \\ \text{s.t.} & \begin{pmatrix} 1 & 2 & 0 \\ -1 & 1 & 1 \end{pmatrix} y \begin{array}{c} \geq \\ = \end{array} \begin{pmatrix} 1 \\ -2 \end{pmatrix} \\ & y_1 \text{ free, } y_2 \leq 0, y_3 \geq 0. \end{array}
$$

CS conditions are

- $x_1 = 0$ or $y_1 + 2y_2 = 0$
- $y_2 = 0$ or $2x_1 + x_2 = 1$
- $y_3 = 0$ or $x_2 = 4$

Let $\bar{x} = (1,-1)^T$ be a feasible solution to the primal, to verify if it is optimal, we need to compute y feasible for the dual (D) and satisfy the CS conditions

- From the first CS condition we have $y_1 + 2y_2 = 0$ since $\bar{x}_1 \neq 0$.
- Also $2\bar{x}_1 + \bar{x}_2 = 1$, so the second CS condition is satisfied.
- From the third CS condition, we have $y_3 = 0$ since $\bar{x}_2 \neq 4$.

Also we have $-y_1 + y_2 + y_3 = -2$ from the feasibility for (D).
Therefore we calculate $y_1 = -\frac{1}{3}, y_2 = \frac{5}{3}, y_3 = 0$. We can check $y_1 + 2y_2 = 3 \geq 1$,so $y = (-\frac{1}{3}, \frac{5}{3}, 0)$ is feasible for (D), therefore $\bar{x}$ is optimal for the primal and y is optimal for the dual.

Class outline.

1. Simplex from the dual point of view
2. Shortest path algorithm

29 Simplex from the primal point of view

Consider the following LP:

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

and its dual

$$(D) \quad \begin{array}{ll} \max & b^T y \\ \text{s.t.} & A^T y = c \\ & y \ge 0. \end{array}$$

From a primal point of view, simple method try to improve the objective value while staying in the feasible region (polyhedron). Once the objective c is in the cone of the tight constraints, we know we are at the optimal point (vertex).

30 Simplex from the dual point of view

From the dual point of view, we consider the primal in the SEF.
Assume primal problem is

$$(P) \quad \begin{array}{ll} \max & c^T x \\ \text{s.t.} & Ax = b \\ & x \ge 0 \end{array}$$

and its dual

$$(D) \quad \begin{array}{ll} \min & b^T y \\ \text{s.t.} & A^T y \ge c \\ & y \text{ free} \end{array}$$

From each basis B in simplex iterations, we can compute $y = A_B^{-T} c_B$ which gives us a "solution" to the dual problem which may not be necessarily feasible. In an simplex iteration, we perform the minimum ratio test, which is to maintain feasibility of the primal problem. From the dual point of view, this is to make sure the objective b stays in the cone of tight constraints of the dual problem , i.e., $b = A^= x^=$ for some $x^= \ge 0$ (we assume $A = [A^=, A^<]$). In simplex iterations, we always pick the positive coefficient from the objective function as the entering variable. From the dual point of view, this is to pick a constraint that is violated by $y = A_B^{-T} c_B$ in the dual and we will make this constraint tight in the new basis, i.e., we are locally improving the feasibility of the dual solution y while keeping the objective function b in the cone of tight constraints. Once the dual solution y is feasible, we have optimality for the dual (also for the primal) since the objective function b always stays in the cone of tight constraints.

In the dual problem, when we move from one vertex to a nearby vertex, there are three rules to follow:

1. The new constraint needs to be violated by the current vertex (We need to pick the positive coefficient c_j from the objective function in the canonical form).

2. The tight inequality which is left needs to be consistent in the new vertex (not being violated). In simplex, this is saying we need to choose the positive A_{kj} from column j).

3. The objective direction needs to stay in the cone of normal vectors of tight constraints. (We perform minimum ratio test in simplex to make sure the basic solution is nonnegative.)

An illustration of simplex iterations from the dual point of view is in Figure 5

1. We start at vertex 1, where b is in the cone of tight constraints, so it is primal feasible.

2. From vertex 1, we have 3 neighbors: 2, 3 and 4. We don't move to vertex 3 because the new tight constraint at vertex 3 is not violated by vertex 1. We don't move to vertex 4 because the tight constraint at vertex 1 is violated by vertex 4, so we move to vertex 2, and the objective b stays in the cone of tight constraints.

3. From vertex 2, we have three choices: 4, 5 and 6. (In simplex, we don't immediately come back to the previous vertex.) We don't move to 4 because the new tight constraint at vertex 4 is not violated by vertex 2. We don't move to vertex 6 because b is not in the cone of tight constraints at vertex 6. So we move to vertex 5. So we have dual feasibility and since b is always in the cone of tight constraints, we arrive at an optimal solution. One can visually confirm vertex 5 is indeed the optimal solution for the objective function b (we are minimizing instead of maximizing, so 5 is far most in the opposite direction of b).

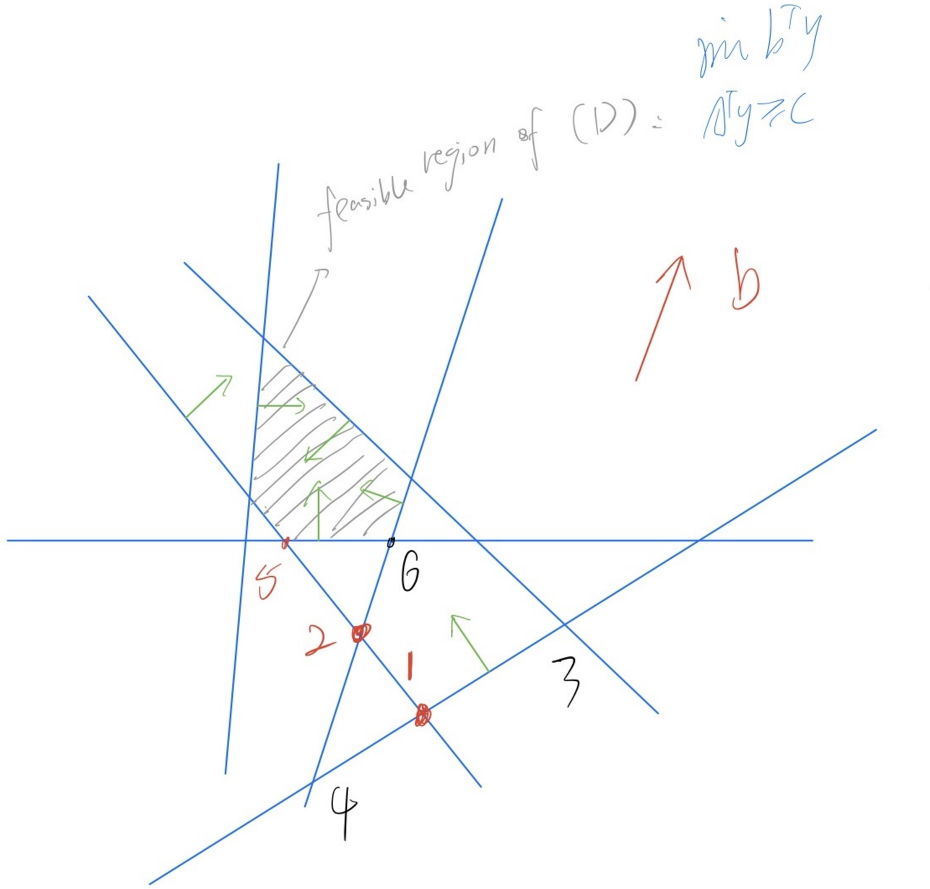


Figure 5: Simplex iterations: $1 \to 2 \to 5$

Summary

1. From a primal point of view, simplex method maintains primal feasibility and aim to include the objective function c in the cone of tight constraints. Once c is the cone of tight constraints, the solution is optimal.

2. From the dual point of view, simplex method keeps the objective b inside the cone of tight constraints and aim to have dual feasibility. Once a dual feasible solution is obtained, the solution is optimal.

30.1 Exercise

Consider the following primal-dual pair linear program

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

Show that

1. The feasibility of the primal problem is equivalent to the objective b being in the cone of tight constraints of the dual.
2. The objective c in the cone of tight constraints in the primal is equivalent to the feasibility of the dual.

31 Shortest path algorithm

Recall the shortest path problem: Given a graph $G = (V, E)$ with two distinct specified vertices vertices s, t and positive edge lengths c_e, find a shortest $s, t-$path in G.
Recall the IP formulations of finding the shortest path problem:

$$\begin{aligned} \min \quad & \sum_{e \in E} c_e x_e \\ \text{s.t.} \quad & \sum_{e \in \delta(S)} x_e \geq 1, \quad \forall \text{ s,t-cuts } \delta(S) \\ & x_e \geq 0, x_e \in \mathbb{Z}, \forall e \in E \end{aligned}$$

We know that for any s,t-cut $\delta(S)$ and any s,t-path, the path must have at least one edge in $\delta(S)$. We also know that an edge set that intersects all s,t-cuts must contain an s,t-path (with possible redundant edges). So given any feasible solution of the IP formulation, we can get an s,t-path by removing the redundant edges, therefore if we find the minimum solution of the IP formulation, it must be an shortest s,t-path.

Integer programming problems are generally hard to solve, However, the shortest path problem has a special structure which we can use. We will relax this IP problem to an LP problem, and solve the LP problem. The relaxed LP problem has special primal-dual structure which we can use to design efficient algorithms to solve it. It also has an optimal solution of integers, therefore when we solve the LP problem, we solved the original IP problem already.

We start with a simple example in Figure 6:

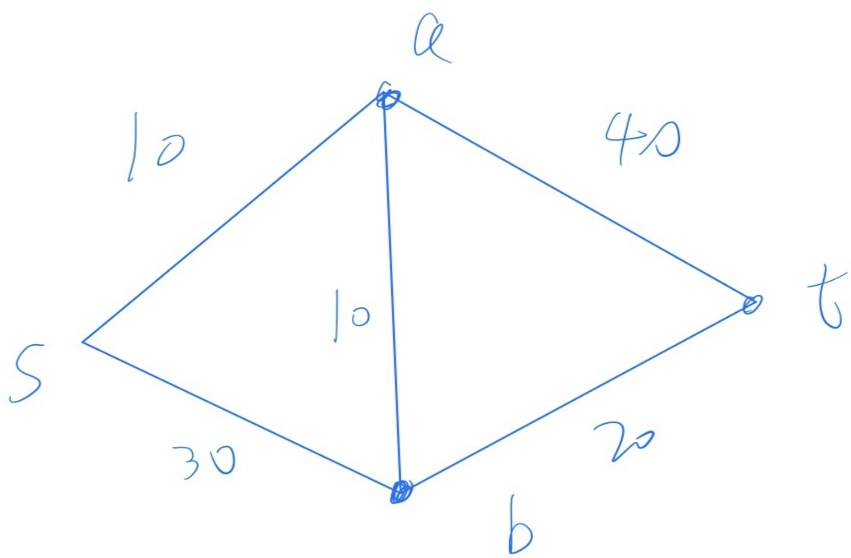


Figure 6: shortest path from s to t

$$
(P) \quad \begin{array}{ll}
\min & 10x_{sa} + 30x_{sb} + 10x_{ab} + 40x_{at} + 20x_{bt} \\
\text{s.t.} & x_{sa} + x_{sb} \geq 1 \\
& x_{sb} + x_{ab} + x_{at} \geq 1 \\
& x_{sa} + x_{ab} + x_{bt} \geq 1 \\
& x_{at} + x_{bt} \geq 1 \\
& x \geq 0
\end{array}
$$

There are 4 constraints, each constraint corresponds to an s,t-cut.
The dual LP is the following:

$$
(D) \quad \begin{array}{ll}
\max & y_s + y_{s,a} + y_{s,b} + y_{s,a,b} \\
\text{s.t.} & y_s + y_{s,b} \leq 10 \\
& y_s + y_{sa} \leq 30 \\
& y_{s,a} + y_{s,b} \leq 10 \\
& y_{s,a} + y_{s,a,b} \leq 40 \\
& y_{s,b} + y_{s,a,b} \leq 20 \\
& y \geq 0
\end{array}
$$

The general primal-dual LP for the shortest path problem is

$$
(P_{s,t}) \qquad \begin{aligned} \min \quad & \sum_{e\in E} c_e x_e \\ \text{s.t.} \quad & \sum_{e\in\delta(S)} x_e \geq 1, \quad \forall \text{ s,t-cuts } \delta(S) \\ & x_e \geq 0, \forall e \in E \end{aligned}
$$

$$
(D_{s,t}) \qquad \begin{aligned} \max \quad & \mathbb{1}^T y \\ \text{s.t.} \quad & \sum_S (y_S : \delta(S) \text{ is an s,t-cut containing } e) \leq c_e, \forall e \in E \\ & y \geq 0, \end{aligned}
$$

Remark 4 (Interpreting the dual:)**.** *Every s,t-path must cross all s,t-cuts, so the length of each s,t-path is at least the sum of all y-values. This is also a consequence of weak duality.*

31.1 Complementary Slack conditions

At optimality, the CS conditions for the shortest path LP problem is

1. For each s,t-cut $\delta(S)$, either $y_S = 0$ or $\sum(x_e : e \in \delta(S)) = 1$. (Primal CS condition)

2. For each $e \in E$, either $x_e = 0$ or $\sum_S(y_S : \delta(S)$ is an s,t-cut containing $e) = c_e$. In other words, if $x_e \neq 0$ in the primal, then the sum of all the values of cuts crossing e must be equal to its weight c_e. (Dual CS condition)

Examples: An illustration is in Figure 7.

An algorithm for solving shortest path problem:

1. Initialization: $W = \{s\}$ is (set to store vertices), $C = \emptyset$ (set to store cuts), $T = \emptyset$ (set to store edges), set $x_e = 0$ for all the edge e. (Iteration 0)

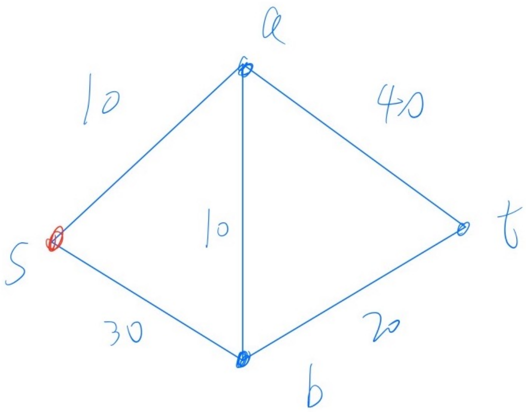
2. For i from 1 to $n-1$ do the following:

   (a) Compute the slacks $f(e)$ of all the edges e in the cut $\delta(\bar{V})$, i.e., ($f(e) = c_e - \sum_S y_S : \delta(S)$ is an cut in C containing e.)

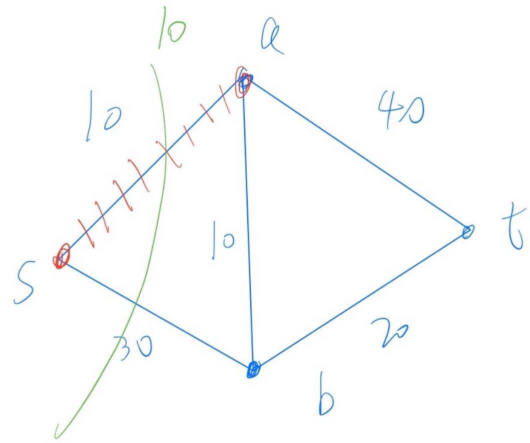
   (b) Find an edge $uv \in \delta(W)$ with minimum slacks among all the edges e in $\delta(W)$. ($u \in W, v \notin W$.)

   (c) Set $y_W = f(uv)$

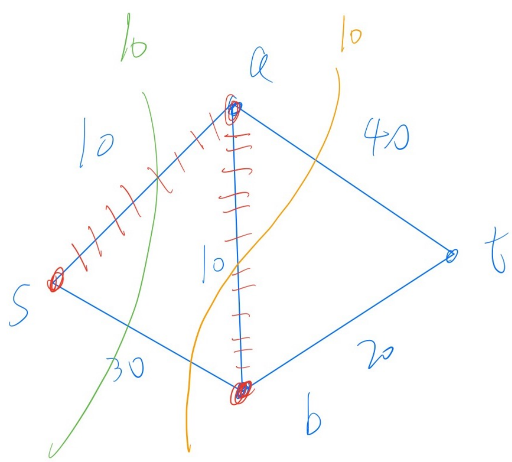
   (d) Add cut $\delta(W)$ to C.



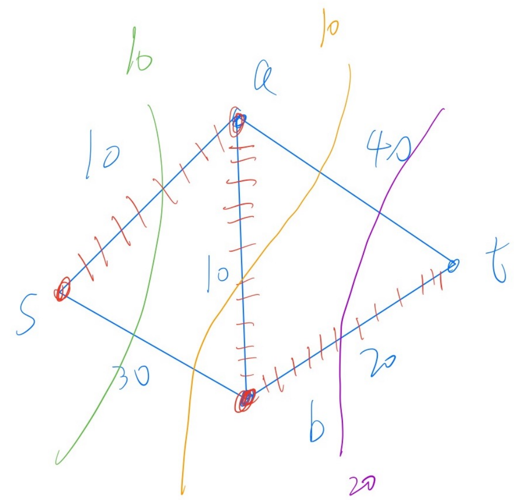
(a) iteration 0



(b) iteration 1



(c) iteration 2



(d) iteration 3

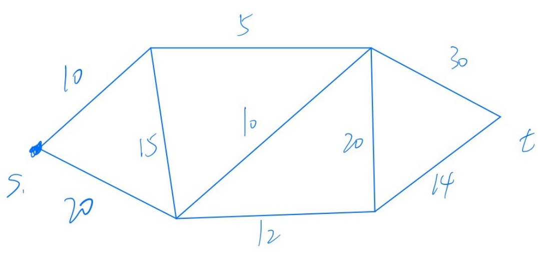
Figure 7

(e) Add v to W.

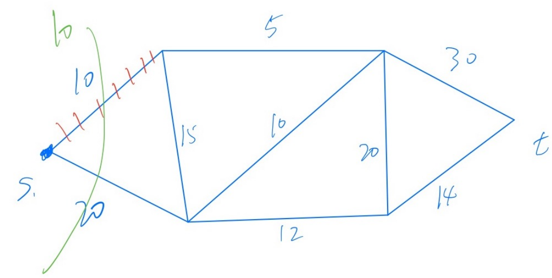
(f) Add uv to T

(g) There exists a unique path from s to v_i. (this is indeed the shortest path from s to v_i.)
(One can find the path by backtracking from v_i to s.)

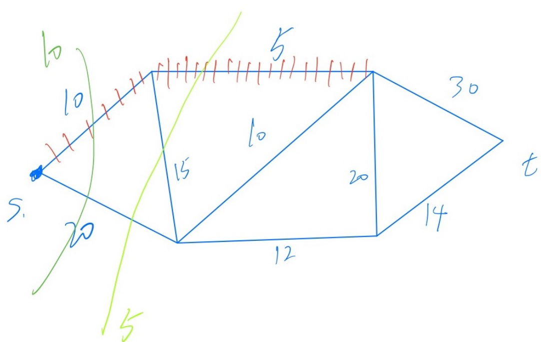
(h) Stop when $v_i = t$



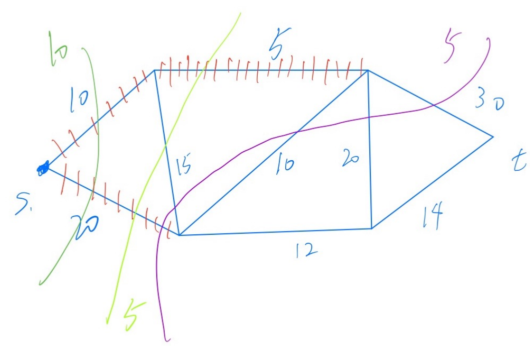
(a) iteration 0



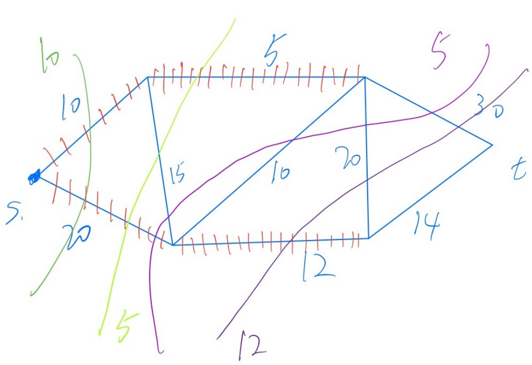
(b) iteration 1



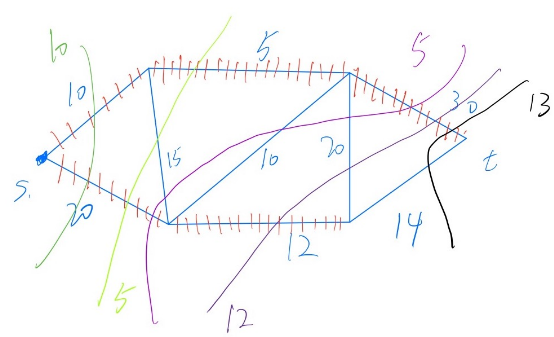
(c) iteration 2



(d) iteration 3



(e) iteration 4



(f) iteration 5

Figure 8

Another example in Figure 8

31.2 Correctness of the algorithm

At the end of step i, the edge set T and vertex set W forms a tree, and we have

1. The unique path from s to v_i gives us a feasible solution to the primal LP P_{s,v_i}. (An s, v_i-path)

2. The values of y gives a feasible solution to the dual LP D_{s,v_i}. (Assign of values to the cuts)

3. For each $e \in T$, we have $\sum_S (y_S : \delta(S)$ is an cut in C containing $e) = c_e$ due to the algorithm. (we choose the edge uv with the minimum slack and set $y_W = f(uv)$) (Dual CS condition is satisfied)

4. For each cut $\delta(S) \in C$, it only cuts the s, v_i-path exactly once. It must cut the s, v_i-path by at least once, because it is an s, v_i-cut. It can not cut it more than once, because W and T forms a tree. (Primal CS condition is satisfied)

So we have primal and dual feasibility plus CS conditions, the path we find must be optimal, i.e., it is an shortest path from s to v_i.

CO 250 Winter 2022 Week 11

Class outline.

1. LP relaxation of integer programming

LP problems are easy to solve, IP problems are hard to solve. No known fast algorithms exist for solving general IPs.

Fundamental Theorem for LP does apply to Integer programming, there exists IP which is bounded, feasible, but the optimal solution is not obtained when the entries are irrational.

The idea to solve IPs, we relax IP to LPs, If the extreme point of the relaxed LPs are integers, then we solve the LPs and get an optimal solution at an extreme point, it must integral, therefore it must be the optimal solution of the IPs.

We consider the following IP example:

$$\begin{array}{ll} \max & (1,1)x \\ \text{s.t.} & \begin{pmatrix} 2 & 4 \\ 2 & 0 \\ 1 & -4 \\ -2 & 1 \\ -3 & -2 \end{pmatrix} x \leq \begin{pmatrix} 11 \\ 5 \\ 4.5 \\ 1.5 \\ 0.5 \end{pmatrix} \\ & x \in \mathbb{Z} \end{array}$$

Let P be the polyhedron defined by the linear constraints. If the run Simplex method for the LP relaxation, we obtain the optimal solution $(\frac{5}{2}, \frac{3}{2})^T$ which is an extreme point of P. But $(\frac{5}{2}, \frac{3}{2})^T$ is fractional not integral!, which is problematic. Let is consider all the integral oints in P. Define $S = P \cap \mathbb{Z}^2$. Then

$$S = \{(0,0), (1,0), (2,0), (0,1), (1,1), (2,1), (1,2)\}$$

We can then find a polyhedron Q such that $S \subseteq Q$ and all extreme points of Q are in S.
In this example, Q can be computed as

$$\begin{aligned} x_1 &\geq 0 \\ x_1 &\geq 0 \\ x_1 + x_2 &\leq 3 \\ -x_1 + x_2 &\leq 1 \\ x_1 &\leq 2 \end{aligned}$$

We then run Simplex to maximize $(1,1)x$ over Q, we will obtain the optimal solution $(2,1)^T$

31.3 Convex hull

There are two equivalent definition of convex hull.

Definition 13 (Convex hull)**.**

- *The convex hull conv* $(x^{(1)}, \cdots, x^{(k)})$ *of a set of points* $x^{(1), \cdots, x^{(k)}}$ *is the smallest convex set that contains all those points.*
- *Equivalently, the convex hull is*

  $$conv\ (x^{(1)}, \cdots, x^{(k)} = \{\lambda_1 x^{(1)} + \cdots + \lambda_k x^{(k)} : \lambda_1 + \cdots + \lambda_k = 1, \lambda_1, \cdots, \lambda_k \geq 0\}$$

Example of convex hull of some points:

31.4 Convex hull of IP

Theorem 10 (Fundamental theorem of IP). *Let $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ be a polyhedron where A, b have rational entries. Let S be the set of integer points in P. Then the convex hull of S is a polyhedron of the form $\{x \in \mathbb{R}^n : \bar{A}x \leq \bar{b}\}$ where $\bar{A}, \bar{b}$ have rational entries.*

We can now relate an IP with its LP relaxations using the convex hull of the feasible solutions of the IP problem.

Theorem 11. *Let (IP) be*

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

Let $S = \{x \in \mathbb{Z}^n : Ax \leq b\}$ and conv $(S) = \{x \in \mathbb{R}^n : \bar{A}x \leq \bar{b}\}$. Let (LP) be

$$\begin{array}{ll} \max & c^T x \\ s.t. & \bar{A}x \leq \bar{b} \\ & x \in \mathbb{R}^n \end{array}$$

Then

1. *(IP) is infeasible if and only if (LP) is infeasible.*
2. *(IP) is unbounded if and only if (LP) is unbounded.*
3. *Every optimal solution of (IP) is an optimal solution of (LP).*
4. *Every optimal solution of (LP) which is an extreme point is also an optimal solution of (IP).*

31.5 Cutting planes

Example:

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

The optimal solution for the LP is $\bar{x} = (\frac{8}{3}, \frac{4}{3})^T$.

Definition 14 (Cutting plane). *A half space that $\alpha^T x \leq \beta$ such that*

- *The optimal solution $\bar{x}$ to the LP relaxation does not satisfy the cutting plane inequality.*

- *Every feasible solution to the IP satisfies the cutting plane inequality.*

Verifying the linear constraints

$$(1,3)x \leq 6$$

is a cutting plane:
We add 2 times the first row and 1 times the second row to get

$$3x_1 + 9x_2 = 20$$

which holds for all feasible solutions. We divide this by 3 to get

$$x_1 + 3x_2 \leq \frac{20}{3}.$$

The right hand side is not an integer, so we can take the floor function of it and it will still hold for any feasible integer solution of the (IP), thus

$$x_1 + 3x_2 \leq \lfloor\frac{20}{3}\rfloor = 6$$

$(\frac{8}{3}, \frac{4}{3})$ violates this constraints, and all feasible integer solution satisfy this constraints.

31.6 Cutting plane and simplex algorithm

We can obtain this cutting plane by simplex method.
We run simplex on the LP relaxation and the optimal basis is $\{1,2\}$ and the corresponding canonical form is

$$\begin{aligned} \max \quad & (0,0,-1,-1)x + 12 \\ \text{s.t.} \quad & \begin{pmatrix} 1 & 0 & -\frac{1}{3} & \frac{4}{3} \\ 0 & 1 & \frac{1}{3} & -\frac{1}{3} \end{pmatrix} x = \begin{pmatrix} \frac{8}{3} \\ \frac{4}{3} \end{pmatrix} \\ & x \geq 0 \end{aligned}$$

Note that the RHS of the first row and second row of the constraint matrix A are not integers, therefore the BFS are not integers. To cut the non-integer optimal solution, we use the first row (one can also use the second row) which is

$$x_1 - \frac{1}{3}x_3 + \frac{4}{3}x_4 = \frac{8}{3}$$

Since $x_3 \geq 0, x_4 \geq 0$, we can take the floor function of the coefficients of x_3 and x_4, so

$$x_1 + \lfloor -\frac{1}{3} \rfloor x_3 + \lfloor \frac{4}{3} \rfloor x_4 \leq \frac{8}{3}$$

which is

$$x_1 - x_3 + x_4 \leq \frac{8}{3}$$

Now we take floor function of $\frac{8}{3}$ to cut off the optimal solution which is not integral from the LP relaxation, thus

$$x_1 - x_3 + x_4 \leq \lfloor \frac{8}{3} \rfloor = 2$$

Note if we substitute the slack variables x_3, x_4 for the original variables, i.e., we plug in

$$\begin{aligned} x_3 &= 8 - x_1 - 4x_2 \\ x_4 &= 4 - x_1 - x_2, \end{aligned}$$

then we get

$$x_1 - (8 - x_1 - 4x_2) + 4 - x_1 - x_2 \leq 2$$

which is

$$x_1 + 3x_2 \leq 6$$

If we add the new cutting plane to the LP relaxation,

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

The optimal solution for the above LP is $(3,1)^T$. We can check the tight constraints for $(3,1)$ are $(1,1)x \leq 4$ and $(1,3)x \leq 6$. The objective $(2,5)^T$ is indeed in the cone generated by $(1,1)^T$ and $(1,3)^T$. So the optimal solution of the original (IP) problem is $(3,1)^T$.

31.7 The cutting plane algorithm

1. **Start loop**
2. Solve (LP): $\max\{c^T x : Ax = b, x \geq 0\}$
3. if (LP) is infeasible then (IP) is infeasible.
4. Let $\bar{x}$ be an optimal solution on (LP)
5. If $\bar{x}$ is integral then
6.   Stop $\bar{x}$ is an optimal solution to (IP)

7. end if
8. Find a cutting plane $\alpha^T x \leq \beta$ for $\bar{x}$ (where α, β are integers) (by taking the floor function of a row with rational RHS from the canonical form of the optimal basis in simplex).
9. Add slack variable z to constraints such that $\alpha^T x + z = \beta$
10. Add constraint $\alpha^T x + z = \beta$ to the system $Ax = b$.
11. **End loop**

Example Consider the following example

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

First we run simplex on the LP relaxation, the optimal basis is $\{1\}$ with an optimal BFS $\bar{x} = \frac{3}{2}$.
We take the constraint $x_1 + x_2 + \frac{1}{2}x_3 = \frac{3}{2}$, and take the floor function of the coefficient to get

$$x_1 + x_2 \leq 1$$

We add a slack variable x_4 to get $x_1 + x_2 + x_4 = 1$.
We solve the following LP

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

The optimal basis is $\{1,3\}$ with optimal BFS being $(1,0,1,0)^T$ which is an optimal solution to the (IP)

Remark 5. *In general if $a^T x = \beta$ is a constraint in the (IP) $\max c^T x : Ax = b, x \geq 0, x \in \mathbb{Z}$. Then $\lfloor a \rfloor x \leq \lfloor \beta \rfloor$ is a valid cutting plane of the feasible region of the (IP). So we can assume that A and b are integers.*

Theorem 12. *Assume A has rational entries, if (LP) is unbounded, then (IP) is unbounded or infeasible.*

Proof. Suppose that (IP) has a feasible solution $\bar{x}$. Since (LP) is unbounded, there exists d such that $Ad = 0$, $d \geq 0$ and $c^T d > 0$. Since A has rational entries, we know d must have rational entries, so we can multiply d by some integer $\alpha > 0$ such that αd have integer entries. So $\bar{x}$ and αd is a certificate of unboundness of the (IP). □

Note if A has some non-rational entries, anything can happen.
Example:

$$\begin{aligned} \max \quad & (0,1)x \\ \\ \text{s.t.} \quad & (1, -\sqrt{2})x = 0 \\ & x \geq 0, \text{ integer} \end{aligned}$$

(LP) is unbounded with $x = (0,0)^T$ and $d = (\sqrt{2}, 1)$, Howver (IP) has only one feasible solution $(0,0)$.

CO 250 Winter 2022 Week 12

Class outline.

1. Nonlinear programming (Convex)

31.8 Convex functions

Definition 15. *A function $f : \mathbb{R}^n \to \mathbb{R}$ is convex if for every pair of points $x^{(1)}$ and $x^{(2)} \in \mathbb{R}^n$ and $\lambda_1, \lambda_2 \geq 0$, $\lambda_1 + \lambda_2 = 1$ we have*

$$f(\lambda_1 x^{(1)} + \lambda_2 x^{(2)}) \leq \lambda_1 f(x^{(1)}) + \lambda_2 f(x^{(2)})$$

Example of convex functions
Prove $f(x) = x^2$ is a convex function.
Is $f(x) = |x|$ convex? Yes
Is $f(x) = x^3$ convex? No

Definition 16. *level set Let $f : \mathbb{R}^n \to \mathbb{R}$, the set*

$$\{x \in \mathbb{R}^n : f(x) \leq \beta\}$$

is called a level set of the function f.

Example: let $f(x) = x^2$, then the level set

$$S = \{x \in \mathbb{R}^n : x^2 \leq 1\} = \{[-1, 1]\}$$

Theorem 13. *If f is a convex function, then the level set*

$$S = \{x \in \mathbb{R}^n : f(x) \leq \beta\}$$

is a convex set.

Proof. Exercise $\square$

Remark 6. *The inverse of the previous theorem is not true. For example $f(x) = x^3$ is not convex but level set is always convex.*

31.9 Epigraph

Definition 17. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function. Then the set*

$$epi(f) = \{(\alpha, x)^T \in \mathbb{R} \times \mathbb{R}^n : f(x) \leq \alpha\}$$

is called the epigraph of f.

Epigraph is the region above the graph $f(x)$.

Proposition 9. *f is convex if and only if $epi(f)$ is convex.*

Proof. Exercise $\square$

A convex function f on $\mathbb{R}^n$ is continuous. ($f < +\infty$).

31.10 Subgradient of convex function

Definition 18. *We say a vector $g \in \mathbb{R}^n$ is a subgradient of $f : \mathbb{R}^n \to \mathbb{R}$ at $x \in \mathbb{R}^n$ if for all $z \in \mathbb{R}^n$,*

$$f(z) \geq f(x) + g^T(z - x)$$

The set of subgradients of f at the point x is called the subdifferential of f at x, and is denoted $\partial f(x)$.

Geometrically, a vector g is a subgradient of f at x if the affine function (of z) $f(x) + g^T(z - x)$ is a global underestimator of f. Geometrically, g is a subgradient of f at x if $(g, -1)$ supports epi$f$ at $(x, f(x))$,

Remark 7. *If f is differential, then the subgradient coincides with the gradient.*

Example: consider the function $f(x) = |x|$,
The set of subgradients at $x = 0$ is the interval $[-1, 1]$
Example:
Consider the function $f(x, y) = x^2 + y^2$ which is differentiable. The gradient of f is $(2x, 2y)^T$ which is also the unique subgradient for any x.

One can check for example let $\bar{x} = (1,1)^T$, so that $g = \nabla f(1,1) = (2,2)^T$. Then

$$\begin{aligned} f(z_1, z_2) &\geq f(1,1) + (2,2)(z_1 - 1, z_2 - 1)^T \\ z_1^2 + z_2^2 &\geq 2 + 2(z_1 - 1) + 2(z - 2 - 1) \\ (z_1 - 1)^2 + (z_2 - 1)^2 &\geq 0 \end{aligned}$$

So $g = (2,2)^T$ is indeed a subgradient of f at $(1,1)^T$.

Lemma 1. *The set of all subgradient at x (subdifferential) is a convex set.*

Proof. Exercise. □

31.11 Supporting halfspace

Consider a convex set $C \subseteq \mathbb{R}^n$ and let $x \in C$. We say that the halfspace $F = \{x \in \mathbb{R}^n : g^T x \leq b\}(g \in \mathbb{R}^n$ and $\beta \in \mathbb{R})$ is a supporting halfspace of C at x if the following conditions hold:

1. $C \subseteq F$.
2. $g^T x = b$, i.e. x is on the hyperplane that defines the boundary of F.

Proposition 10. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function, let $\bar{x} \in \mathbb{R}^n$ such that $f(\bar{x}) = \beta$, and let $g \in \mathbb{R}^n$ be a subgradient of f at $\bar{x}$. Denote by $C = \{x \in \mathbb{R}^n : f(x) \leq \beta\}$ and $F = \{x \in \mathbb{R}^n : g^T(x - \bar{x}) \leq 0\}$. Then F is a supporting halfspace of C at $\bar{x}$.*

Proof. We need to verify conditions (1),(2) of supporting halfspaces hold.
To verify (1), let $z \in C$, then $f(z) \leq \beta$. Since g is a subgradient of f at $\bar{x}$, we have $f(\bar{x}) + g^T(z - \bar{x}) \leq f(z)$. It follows $f(\bar{x}) + g^T(z - \bar{x}) \leq \beta$, i.e., $g^T(z - \bar{x}) \leq 0$ and $z \in F$.
To verify (2), it is easy to see $g^T(\bar{x} - \bar{x}) = 0$.

□

The converse is also true with some mild assumptions on F.

Proposition 11. *If $0 \notin \partial f(\bar{x})$ and $F = \{x \in \mathbb{R}^n : g^T(x - \bar{x}) \leq 0\}$ is a supporting halfspace of $C = \{x \in \mathbb{R}^n : f(x) \leq f(\bar{x})\}$ at $\bar{x}$. Then λg is a subgradient of f at $\bar{x}$ for some $\lambda > 0$.*

Proof. Omitted. (covered in more advanced nonlinear optimization course) □

Example:

$$f(x_1, x_2) = |x_1| + 2|x_2|$$

Let $\bar{x} = \{0, \frac{1}{2}\}$ and $f(\bar{x}) = 1$. Consider the level set $C = \{(x_1, x_2) : |x_1| + 2|x_2| \leq 1\}$.
Let $g_1 = (1, 2)^T$ and $g_2 = (-1, 2)^T$, one can check $F_1 = \{x : g_1^T(x - \bar{x}) \leq 0\}$ and $F_2 = \{x : g_2^T(x - \bar{x}) \leq 0\}$ are indeed supporting halfspaces of C.

Also $|x_1| + 2|x_2| \geq g_1^T(x - \bar{x}) + f(\bar{x}) = x_1 + 2x_2$. So g_1 is a subgradient of f at $\bar{x}$

Similarly $|x_1| + 2|x_2| \geq g_2^T(x - \bar{x}) + f(\bar{x}) = -x_1 + 2x_2$. So g_2 is also a subgradient of f at $\bar{x}$.

The subdifferential $\partial f(\bar{x}) = \{\lambda_1 g_1 + \lambda_2 g_2 : \lambda_1 + \lambda_2 \leq 1\}$

31.12 NLP with linear objective

Consider the NLP with linear objective

$$\begin{array}{cc} \min & f(x) \\ \text{s.t.} & g_1(x) \leq 0 \\ & \vdots \\ & g_m(x) \leq 0 \end{array}$$

which is equivalent to

$$\begin{array}{cc} \min & x_{n+1} \\ \text{s.t.} & f(x) - x_{n+1} \leq 0 \\ & g_1(x) \leq 0 \\ & \vdots \\ & g_m(x) \leq 0 \end{array}$$

Corollary 1. *Consider an NLP of the form given in the following:*

$$\begin{array}{cc} \min & c^T x \\ s.t. & g_1(x) \leq 0 \\ & \vdots \\ & g_m(x) \leq 0. \end{array} \tag{5}$$

Let $\bar{x}$ be a feasible solution and suppose that constraint $g_i(\bar{x}) \leq 0$ is tight for some $i \in J(\bar{x})$. Suppose g_i is a convex function that has a subgradient s_i at $\bar{x}$. Then the NLP obtained by replacing constraint $g_i(x) \leq 0$ by the supporting halfspace $s^T(x - \bar{x}) \leq 0$ is a relaxation of the original NLP.

Let $\bar{x}$ be a feasible solution, then the relaxed problem is

$$\begin{array}{ll} \max & -c^T x \\ \text{s.t.} & s_i^T(x-\bar{x}) \leq 0 \text{ for all } i \in J(\bar{x}) \end{array}$$

where $J(\bar{x})$ is the index set of all the tight constraints g_i and s_i is a subgradient of g_i at $\bar{x}$.

Proposition 12. *Consider the NLP in (5) and assume that $g_1, ..., g_m$ are convex functions. Let $\bar{x}$ be a feasible solution and suppose that for all $i \in J(\bar{x})$ we have a subgradient s_i at $\bar{x}$. If $-c \in$ cone $\{s_i : i \in J(\bar{x})\}$, then $\bar{x}$ is an optimal solution of Problem (5).*

Thus, we have sufficient condition for optimality.

CO 250 Winter 2022 Week 13

Class outline.

1. Convex NLP, Slater point.
2. KKT conditions

31.13 Convex NLP

Consider the NLP

$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & g_1(x) \leq 0 \\ & \vdots \\ & g_m(x) \leq 0 \end{array}$$

If all functions $f : \mathbb{R}^n \to \mathbb{R}, g_1, ..., g_n : \mathbb{R}^n \to \mathbb{R}$ are convex, then the NLP is called Convex NLP Equivalently, a convex NLP can be written in the following form:

$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & h_i(x) = 0, i \in \{1, ..., m_1\} \\ & g_j(x) \leq 0, j \in \{m_1, ..., m\} \end{array}$$

where $h_i(x)$ are affine functions.

A convex NLP can be written in the following equivalent form

$$\begin{array}{rc} \min & x_{n+1} \\ \text{s.t.} & f(x) - x_{n+1} \leq 0 \\ & g_1(x) \leq 0 \\ & \vdots \\ & g_m(x) \leq 0 \end{array}$$

Definition 19 (set of indices of tight constraints)**.** *Given a feasible solution $\bar{x}$, we define $J(\bar{x}) = \{i : g_i(\bar{x}) = 0\}$ to be the set of indices of the tight constraints.*

Corollary 2. *Consider an NLP of the form given in the following:*

$$\begin{array}{rc} \min & c^T x \\ s.t. & g_1(x) \leq 0 \\ & \vdots \\ & g_m(x) \leq 0. \end{array} \tag{6}$$

Let $\bar{x}$ be a feasible solution and suppose that constraint $g_i(\bar{x}) \leq 0$ is tight for some $i \in J(\bar{x})$. Suppose g_i is a convex function that has a subgradient s_i at $\bar{x}$. Then the NLP obtained by replacing constraint $g_i(x) \leq 0$ by the supporting halfspace $s^T(x - \bar{x}) \leq 0$ is a relaxation of the original NLP.

Let $\bar{x}$ be a feasible solution, then the relaxed problem is

$$\begin{array}{rc} \max & -c^T x \\ \text{s.t.} & s_i^T(x - \bar{x}) \leq 0 \text{ for all } i \in J(\bar{x}) \end{array}$$

where $J(\bar{x})$ is the index set of all the tight constraints g_i and s_i is a subgradient of g_i at $\bar{x}$.

Proposition 13 (Sufficient condition for optimality)**.** *Consider the NLP in* (5) *and assume that $g_1, \ldots, g_m$ are convex functions. Let $\bar{x}$ be a feasible solution and suppose that for all $i \in J(\bar{x})$ we have a subgradient s_i at $\bar{x}$. If $-c \in$ cone $\{s_i : i \in J(\bar{x})\}$, then $\bar{x}$ is an optimal solution of Problem* (5)*.*

Thus, we have sufficient condition for optimality.
Examples:
Given an NLP

$$\begin{array}{rc} \min & -x_1 - x_2 \\ \text{s.t.} & x_2^2 - x_1 \leq 0 \\ & x_1^2 - x_2 \leq 0 \\ & -x_1 + \frac{1}{2} \leq 0 \end{array}$$

Prove that $\bar{x} = (1, 1)^T$ is an optimal solution to the LP.

First this NLP is a convex NLP: the objective function is convex, all the constraints are convex since the sum of convex functions are also convex.
Now we check the indices of tight constraints for $\bar{x}$. In fact the indices of tight constraints are $\{1, 2\}$.

The gradient of $g_1(x)$ is $\nabla(-x_1 + x_2^2) = (-1, 2x_2)^T = (-1, 2)^T$ (evaluated at $(1, 1)^T$).

The gradient of $g_2(x)$ is $\nabla(x_1^2 - x_2) = (2x_1, -1)^T$. We evaluate it at $(1, 1)^T$ which is $(2, -1)^T$.

The cone of the tight constraints is

$$\text{cone}\{\begin{pmatrix} 2 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \end{pmatrix}\}$$

One can check

$$-c = \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \end{pmatrix} + \begin{pmatrix} -1 \\ 2 \end{pmatrix}$$

So $-c$ is in the cone of the tight constraints.

31.14 Necessary condition for optimality, Slater point and KKT

The converse of Proposition 13 is not true in general.
Consider the following convex NLP

$$\begin{array}{ll} \min & c^T x \\ & \text{s.t. } x_1^2 + x_2^2 \leq 0 \end{array}$$

The only feasible solution to this convex NLP is $(0, 0)$. Therefore $(0, 0)^T$ is optimal for any c. Howeer The gradient of $x_1^2 + x_2^2 - 1$ at $(0, 0)^T$ is $(0, 0)^T$. Any nonzero c is not in cone of $(0, 0)^T$.

Definition 20 (Slater point)**.** *Consider a convex NLP*

$$\begin{array}{ll} \min & f(x) \\ s.t. & g_1(x) \leq 0 \\ & \vdots \\ & g_m(x) \leq 0 \end{array}$$

A Slater point $\bar{x}$ is a feasible point $\bar{x}$ such that $\bar{x}$ satisfies all the non-affine inequalities ***strictly****, i.e., if $g_i(x)$ is a non-affine function, then $g_i(\bar{x}) < 0$.*

Theorem 14 (Karush-Kuhn-Tucker or KKT condition for optimality)**.** *Consider a convex NLP*

$$\begin{array}{ll} \min & f(x) \\ s.t. & g_1(x) \leq 0 \\ & \vdots \\ & g_m(x) \leq 0 \end{array}$$

that has a Slater point x^*. Let $\bar{x} \in \mathbb{R}^n$ be a feasible solution and assume that $f, g_1, g_2, ..., g_m$ are differentiable at $\bar{x}$. Then $\bar{x}$ is an optimal solution of the convex NLP if and only if

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

Equivalently the KKT condition can be written in the following formulation:

$$\begin{array}{l} \nabla f(\bar{x}) + \sum_{i=1}^m \lambda_i \nabla g_i(\bar{x}) = 0 \\ \lambda_i g_i(\bar{x}) = 0, \quad i = 1, .., m \\ g_i(\bar{x}) \leq 0, \quad i = 1, .., m \\ \lambda_i \geq 0, \quad i = 1, .., m. \end{array}$$

In the nondifferential setting, the KKT condition can be written in terms of the subdifferential at $\bar{x}$.

$$\begin{array}{l} 0 \in \partial f(\bar{x}) + \sum_{i=1}^m \lambda_i \partial g_i(\bar{x}) \\ \lambda_i g_i(\bar{x}) = 0, \quad i = 1, .., m \\ g_i(\bar{x}) \leq 0, \quad i = 1, .., m \\ \lambda_i \geq 0, \quad i = 1, .., m. \end{array}$$

Examples:

$$\begin{array}{ll} \min & (-1, 1)x \\ \text{s.t.} & x^2 + y^2 \leq 1 \\ & -2x - y \leq 0 \end{array}$$

Slater point: One can check that $(\frac{1}{2}, 0)$ is indeed a Slater point of the problem. KKT conditions for optimality:

$$\begin{array}{c} \lambda_1(2x, 2y)^T + \lambda_2(-2, -1)^T + (-1, -1)^T = 0 \\ \lambda_1(x^2 + y^2 - 1) = 0 \\ \lambda_2(-2x - y) = 0 \\ x^2 + y^2 \leq 1, -2x - y \leq 0 \\ \lambda_1 \geq 0, \lambda_2 \geq 0 \end{array}$$

32 Review

1. LP formulation, Integer Programming formulation, Some tricks: binary variables, very large constant, inequalities.

2. Simplex method, Bland rule, Simplex in symbolic format, canonical form.

3. Two-phase simplex, auxiliary LP (throw away the objective and add new objective,

4. Dual LP, weak duality, strong duality, CS conditions

5. Using CS condition to determine optimality

6. Optimal condition for LP: cone of tight constraints. Justify an LP and its dual LP have infinitely many optimal solution, unique optimal solution.

7. Shortest path algorithm, CS condition for shortest path. Using the cuts to justify the optimal primal solution.

8. Integer programming, definition of cutting plane.

9. Nonlinear Optimization, Convex NLP, Slater point, KKT optimality condition.