The Mathematics of Decision Making I

Joshua Maglione

October 14, 2024

Contents

1 Introduction

The mathematics of decision making is very closely tied to the field of mathematical optimization. One of the primary ways mathematics is used to help guide decisions is by maximizing (or minimizing) specific outcomes subject to a list of constraints. Mathematical optimization provides the formal tools to model and solve such problems.

There are many kinds of mathematical optimization. There are two basic types depending on whether the variables to optimize or discrete or continuous. A few types of optimization are[1]

[1] "Program" is not a computer program but comes from the United States military's use of the word for training and logistics schedules.

- Linear Programming,
- Integer Programming,
- Stochastic programming,
- Combinatorial optimization,
- Dynamic programming.

Unsurprisingly there are many real-world applications; to list a few we have network optimization, pricing strategy, scheduling, supervised machine learning training, supply chain optimization, and transportation problems.

In this module, we will introduce the fundamentals of **linear programming**, also called *linear optimization* and *operations research*, such as the simplex method, polyhedral geometry, and the notion of duality. Depending on the time, we may also delve into **integer programming**.

1.1 History

Mathematical optimization has quite an interesting history. In the 17th century, combinatorial optimization problems were solved using game theory, combinatorics, and ad hoc methods. In the 19th century, transportation problems involving post and rail were studied and solved. And in the 20th century with the two World Wars and rise of the assembly line, operations research took off developing the mathematics for all kinds of optimization problems.

One of the most influential figures in mathematical optimization, and linear programming in particular, is George Dantzig. He was the recipient of the President's National Medal of Science in 1975 [3] and was credited for

> *inventing linear programming and discovering methods that led to wide-scale scientific and technical applications to important problems in logistics, scheduling, and network optimization, and to the use of computers in making efficient use of the mathematical theory.*

The proof of the simplex method, name coined by Motskin, was developed by Dantzig in the late 1940s [2]. I find it interesting that the "inductive proof of the simplex method" was published by the Mathematics Division of the RAND Corporation in 1960 (by Dantzig) and was made classified [1]. Now, of course, it is no longer classified.

After explaining the Simplex Method to John von Neumann at the Institute of Advanced Study in Princeton during 1948, von Neumann immediately conjectured the notion of duality because of his recent foray into game theory.

1.2 Four examples

We describe four example problems that touch on the tools we will develop in this module. For now, these problems are meant to introduce basic concepts and vocabulary.

1.2.1 A diet problem

Erin is planning her breakfast and wants to make oats with milk. (These numbers of simplified and not accurate to real life.)

	Milk (100ml)	Oats (100g)
fat	2g	3g
carbohydrates	1g	3g
protein	4g	3g

Erin wants the meal to provide at least 18g of fat, at least 12g of carbohydrates, and at least 24g of protein. If milk costs 20 cents per 100ml and oats 25 cents per 100g, what mixture minimizes the cost of the desired meal?

We could express this more mathematically. For example, let x and y be variables such that $x = 1$ means 100ml of milk and $y = 1$ means 100g of oats. Calculating the grams of fat relative to x and y is

$$2x + 3y.$$

For carbohydrates it is $x + 3y$, and for protein it is $4x + 3y$. Because we want *at least* 18g of fat, we express this via

$$2x + 3y \geqslant 18.$$

We can set up similar inequalities for the other two:

$$\begin{aligned} 2x + 3y &\geqslant 18, \\ x + 3y &\geqslant 12, \\ 4x + 3y &\geqslant 24. \end{aligned}$$

Since we cannot have negative amounts of milk or oats, we have $x \geqslant 0$ and $y \geqslant 0$. Since we want to minimize costs, we want to minimize

$$C = 0.2x + 0.25y.$$

Putting all of this together, we have the following optimization problem.

Determine values for x and y that minimize

$$C = 0.2x + 0.25y$$

subject to the constraints: $x \geqslant 0$, $y \geqslant 0$, and

$$\begin{aligned} 2x + 3y &\geqslant 18, \\ x + 3y &\geqslant 12, \\ 4x + 3y &\geqslant 24. \end{aligned}$$

1.2.2 A transportation problem

Javier has two production sites: one in Sligo and another in Kilkenny. There are three distributing warehouses in Dublin, Galway, and Cork. The Sligo site can supply 120 products per week, whereas the site in Kilkenny can supply 140 per week. The warehouses in Dublin, Galway, and Cork need 100, 60, and 80 products per week respectively to meet demand. The shipping costs are giving in the following table.

	Dublin	Galway	Cork
Sligo	5	7	9
Kilkenny	6	7	10

How many products should Javier ship from each production site to minimize total shipping costs while still meeting demand?

We need many variables, so let's define a variable for each shipment—for example, from Kilkenny to Dublin. Write them as

$$x_{kd}, x_{kg}, x_{kc}, x_{sd}, x_{sg}, x_{sc}.$$

Since Kilkenny and Sligo can only produce 140 and 120 products, respectively, we have

$$\begin{aligned} x_{kd} + x_{kg} + x_{kc} &\leqslant 140, \\ x_{sd} + x_{sg} + x_{sc} &\leqslant 120. \end{aligned}$$

We need to meet demands, so we have

$$\begin{aligned} x_{kd} + x_{sd} &\geqslant 100, \\ x_{kg} + x_{sg} &\geqslant 60, \\ x_{kc} + x_{sc} &\geqslant 80. \end{aligned}$$

Lastly, we want to minimize cost, so we want to minimize

$$C = 6x_{kd} + 7x_{kg} + 10x_{kc} + 5x_{sd} + 7x_{sg} + 9x_{sc}.$$

Altogether we have the following linear program.

Minimize

$$C = 6x_{kd} + 7x_{kg} + 10x_{kc} + 5x_{sd} + 7x_{sg} + 9x_{sc}$$

subject to the constraints: $x_{ij} \geqslant 0$ for all i and j and

$$\begin{aligned} x_{kd} + x_{kg} + x_{kc} &\leqslant 140, \\ x_{sd} + x_{sg} + x_{sc} &\leqslant 120, \\ x_{kd} + x_{sd} &\geqslant 100, \\ x_{kg} + x_{sg} &\geqslant 60, \\ x_{kc} + x_{sc} &\geqslant 80. \end{aligned}$$

1.2.3 The travelling salesperson problem

Kofi need to deliver n products in n different cities starting in Paris. He wants to do this by visiting each city exactly one time and then returning back to Paris at the end. Which path minimizes the distance traveled?

This problem is perhaps the most famous combinatorial optimization problem and is the core problem of many other more complex problems. We will not do much more with this, but note that different "distance functions" can allow for all kinds of slow-downs and speed-ups.

1.2.4 A financial problem

Julia runs an investment and must invest exactly €100,000 in two types of securities: bond A paying a dividend of 7% and stock B paying a dividend of 9%. Due to her incredible experience, she knows that

- no more than €40,000 can be invested in stock B and
- the amount invested in bond A must be at least twice that in stock B.

How much should Julia invest in each security to maximize her return?

See if you can get the following set up.

Maximize

$$z = 0.07A + 0.09B$$

subject to the constraints: $A \geqslant 0$, $B \geqslant 0$, and

$$\begin{aligned} A + B &= 100000, \\ B &\leqslant 40000, \\ A &\geqslant 2B. \end{aligned}$$

2 General linear programming

Linear programs are the basis of what we consider throughout this module. In the example problems above, we sometimes wanted to maximize and sometimes we wanted to minimize. Although these are technically different, we can treat them as the same. Suppose f is some function we want to maximize. Then

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

So maximizing f is the same as minimizing $-f$. Thus, we can use the two interchangeably—as long as we correctly compensate!

General linear program

Determine values for $x_1, x_2, \ldots, x_n$ that maximize

$$z = c_1x_1 + c_2x_2 + \cdots c_nx_n$$

subject to the constraints:

$$\begin{array}{ccccccccc}
a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & \square & b_1, \\
a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & \square & b_2, \\
\vdots & & \vdots & & \vdots & & \vdots & \square & \vdots \\
a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & \square & b_m,
\end{array}$$

where each of the $\square$ can be replaced with one of $\{=, \leqslant, \geqslant\}$.

Definition 2.1. A ***linear program (LP) problem*** is a problem of the form above. The function z is called the ***objective function***, and the m (in-)equalities are called the ***constraints***.

A key feature of LPs is that the objective function as well as each of the constraint (in-)equalities are *linear* in the $x_1, x_2, \ldots, x_n$.

2.1 Standard form

Can we play around with the constants a_{ij} and b_k to get all of the (in-)equalities into the same "shape"? For example,

$$4x_1 - 5x_2 - x_3 \geqslant 1$$

is equivalent to

$$-4x_1 + 5x_2 + x_3 \leqslant -1.$$

Thus, if we have an inequality, we can force it to use just $\leqslant$. Moreover, if we have an equality, we can use two inequalities to obtain the same solutions:

$$4x_1 - 5x_2 - x_3 = 1 \qquad \text{is equivalent to} \qquad \begin{cases} 4x_1 - 5x_2 - x_3 \geqslant 1 \text{ and} \\ 4x_1 - 5x_2 - x_3 \leqslant 1 \end{cases}$$

So we can transform equalities to inequalities, but what about the other way around? We will look at this soon.

In some examples, variables only took on non-negative values. This actually has an advantage of constraining the possible values of the variables, and it is something we will come back to later on. But what about situations were variables are allowed to have negative values? Suppose x_i can be negative. We can introduce two new variables, say, x_i^+ and x_i^-, and we can rewrite x_i as follows:

$$x_i = x_i^+ - x_i^-.$$

In this way, x_i can be negative while both x_i^+ and x_i^- are non-negative. Thus, we can replace all instances of x_i with $x_i^+ - x_i^-$, so that all variables take non-negative values.

Now we can define the standard form for an LP.

Linear program standard form

Determine values for $x_1, x_2, \ldots, x_n$ that maximize

$$z = c_1x_1 + c_2x_2 + \cdots c_nx_n$$

subject to the constraints: for all $i \in \{1, \ldots, n\}$, $x_i \geqslant 0$ and

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & \leqslant & b_1, \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & \leqslant & b_2, \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & \leqslant & b_m. \end{array}$$

Example 2.2. The following LP is not in standard form.

Determine values for x and y that minimize

$$z = 3x + 2y$$

subject to the constraints: $x \geqslant 0$, $y \geqslant 0$, and

$$\begin{aligned} 2x + y &\leqslant 4 \\ 3x - 2y &\leqslant 6. \end{aligned}$$

We can put it into standard form as follows.

Determine values for x and y that maximize

$$z = -3x - 2y$$

subject to the constraints: $x \geqslant 0$, $y \geqslant 0$, and

$$\begin{aligned} 2x + y &\leqslant 4 \\ 3x - 2y &\leqslant 6. \end{aligned}$$

Example 2.3. Put the following LP into standard form.

Determine values for x and y that minimize

$$z = -4x + y$$

subject to the constraints:

$$\begin{aligned} x - 3y &= 2, \\ x + y &\leqslant 6. \end{aligned}$$

2.2 Canonical form

The canonical form is slightly different to that of the standard form of an LP.

Linear program canonical form

Determine values for $x_1, x_2, \ldots, x_s$ that maximize

$$z = c_1 x_1 + c_2 x_2 + \cdots c_s x_s$$

subject to the constraints: for all $i \in \{1, \ldots, s\}$, $x_i \geqslant 0$ and

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1s}x_s &= b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2s}x_s &= b_2, \\ \vdots \quad\quad \vdots \quad\quad\quad \vdots \quad\quad & \vdots \\ a_{r1}x_1 + a_{r2}x_2 + \cdots + a_{rs}x_s &= b_r. \end{aligned}$$

Proposition 2.4. *Every LP in standard form can be brought into canonical form. In other words, every LP has an associated LP in canonical form.*

Week 1

Proof. Since we have already convinced ourselves that every LP can be brought into standard form, it suffices to show that we can convert every LP in standard form into canonical form.

The only difference between the two forms are in the constraints; namely, we need to convert an inequality of the form

$$a_1 x_1 + \cdots + a_n x_n \leqslant b \tag{2.1}$$

to an equality. To accomplish this, we introduce *slack* variables—these are just variables with a pretentious title. They simply exist to "pick up the slack". The "slack" is just the difference of the right hand side and the left hand side of (2.1).

Let s be a (slack) variable. Then (2.1) is equivalent to

$$\begin{aligned} s &\geqslant 0, \\ a_1 x_1 + \cdots + a_n x_n + s &= b. \end{aligned}$$

Hence, we can introduce a new variable for each inequality and obtain an LP in canonical form. $\square$

Example 2.5. A tailor is producing jumpers and trousers. They first need to cut the fabric and then sew it together. It takes 2 hours to cut the fabric for either a pair of trousers or a jumper. It takes 5 hours to sew a pair of trousers and 3 hours for a jumper. Scissors can be used for 8 hours per day, wheres the sewing machine can be used for 15 hours per day. If a pair of trousers is sold for €120 and a jumper for €100, how many of each should be made to maximize revenue? (Let's ignore demand.)

Write an LP in canonical form for this scenario.

Maximize

$$z = 100J + 120T,$$

subject to $J, T, s_1, s_2 \geqslant 0$ and

$$\begin{aligned} 2J + 2T + s_1 &= 8, \\ 3J + 5T + s_2 &= 15. \end{aligned}$$

2.3 Matrix notation

Instead of writing out all of the constraints and all the terms of the objective function, we can compactly describe the same data using matrices. Define

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}, \quad c = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}.$$

We use the relations $\leqslant$ and $\geqslant$ like we do with $=$ when applied to vectors, that is, they are determined coordinate wise. For example

$$\begin{bmatrix} 1 \\ 4 \end{bmatrix} \leqslant \begin{bmatrix} 2 \\ 5 \end{bmatrix}, \qquad \begin{bmatrix} 2 \\ 3 \end{bmatrix} \not\leqslant \begin{bmatrix} 5 \\ 1 \end{bmatrix}.$$

In symbols, $x \leqslant y$ if and only if $x_i \leqslant y_i$ for all i.

LP standard form (matrices)

For $A \in \mathrm{Mat}_{m\times n}(\mathbb{R})$, $b \in \mathbb{R}^m$, and $c \in \mathbb{R}^n$, maximize

$$z = c^\top x$$

subject to $x \geqslant 0$ and

$$Ax \leqslant b.$$

Notation 2.6. The letter n is the number of variables in the objective function, and m is the number of inequalities separate from $x \geqslant 0$.

If it is not already clear how to convert all the previous example above into the matrix form, try to work this out yourself.

Definition 2.7. A vector $x \in \mathbb{R}^n$ satisfying all the constraints of an LP (in standard form) is a ***feasible solution***.

Example 2.8. Recall Example 2.5 the "sewing problem". The following vectors are all feasible solutions:

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 4 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ \pi \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

The following vectors are not feasible solutions:

$$\begin{bmatrix} -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 5 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 4 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \end{bmatrix}.$$

Definition 2.9. A feasible solution that maximizes the objective function of an LP is an ***optimal solution***.

Now we describe the corresponding matrix form for the canonical form of an LP. It is built *from* the standard form of an LP. We write $I_n \in \mathrm{Mat}_n(\mathbb{R})$ for the identity matrix. For matrices $A, B \in \mathrm{Mat}_{m\times n}(\mathbb{R})$, we set

$$[A \mid B] = \begin{bmatrix} a_{11} & \cdots & a_{1n} & b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_{m1} & \cdots & b_{mn} \end{bmatrix} \in \mathrm{Mat}_{m\times 2n}(\mathbb{R}).$$

LP canonical form (matrices)

For $A \in \mathrm{Mat}_{m\times(n+s)}(\mathbb{R})$, $b \in \mathbb{R}^m$, and $c \in \mathbb{R}^{n+s}$, maximize

$$z = c^\top x,$$

subject to $x \geqslant 0$ and

$$Ax = b,$$

where $c^\top = (c_1, \dots, c_n, 0, \dots, 0)$ and $x^\top = (x_1, \dots, x_n, \dots, x_{n+s})$.

One can take an LP in standard form and construct one in canonical form with the (main) constraint equation:

$$[A \mid I_m]\, x = b.$$

Exercise 1. Show that a feasible solution for an LP in standard form induces a feasible solution in canonical form. Is the converse true?

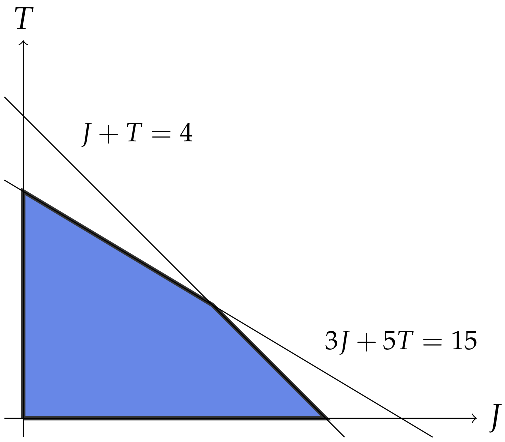
2.4 Geometry of the feasibly set

Now we begin our analysis of the set of feasible solutions to an LP. We begin by looking at the features of its geometry.

Example 2.10. Let's consider the standard form of the LP in Example 2.5. In particular, the feasible solutions are constrained by $J, T \geqslant 0$ and

$$\begin{aligned} 2J + 2T &\leqslant 8, \\ 3J + 5T &\leqslant 15. \end{aligned}$$

We can plot the region in $\mathbb{R}^2$ as follows:



Let's consider one of our constraint inequalities:

$$a_{i1}x_1 + \cdots + a_{in}x_n \leqslant b_i.$$

This can be compactly written as $a^\top x \leqslant b_i$ for $a \in \mathbb{R}^n$ and $b_i \in \mathbb{R}$. The *equation*

$$a^\top x = b_i$$

defines a ***hyperplane*** in $\mathbb{R}^n$: the vector a describes the "slope" and the scalar b_i describes how far the hyperplanes shifts away from the origin. The hyperplane $a^\top x = b_i$ is the boundary of the set of solutions to $a^\top x \leqslant b_i$. In $\mathbb{R}^2$, hyperplanes are lines, and in $\mathbb{R}^3$ they are planes.

Hyperplanes H in $\mathbb{R}^n$ partition $\mathbb{R}^n$ into three sets: the points "below" H, the points "above" H, and the points on H. The set of points below H define a ***half-space***, and similarly for the set of points above H. More precisely, if

$$H = \{x \in \mathbb{R}^n \mid a^\top x = b\},$$

then both

$$\begin{aligned} H^+ &= \{x \in \mathbb{R}^n \mid a^\top x > b\}, \\ H^- &= \{x \in \mathbb{R}^n \mid a^\top x < b\} \end{aligned}$$

are half-spaces of $\mathbb{R}^n$. The ***closed half-spaces*** are

$$\begin{aligned} \overline{H^+} &= \{x \in \mathbb{R}^n \mid a^\top x \geqslant b\} = H^+ \cup H, \\ \overline{H^-} &= \{x \in \mathbb{R}^n \mid a^\top x \leqslant b\} = H^- \cup H. \end{aligned}$$

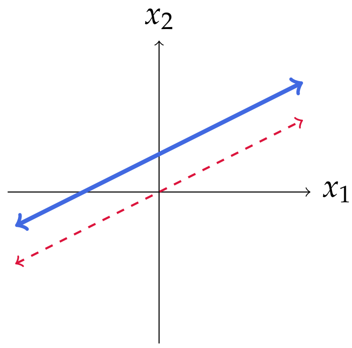


Figure 2.1: The line given by $a = (-1, 2)$, $b = 1$ in blue, and the line with $a = (-1, 2)$, $b = 0$ is in red.

Example 2.11. The sets

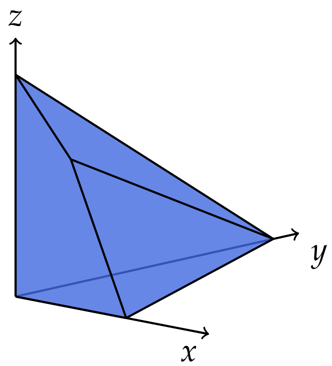
$$\begin{aligned} X &= \{x \in \mathbb{R}^4 \mid -x_1 - 4x_2 + \sin(1)x_3 \leqslant \pi\} \\ Y &= \{y \in \mathbb{R}^4 \mid 7y_1 + 7y_2 + 7y_3 + 7y_4 = 2\}, \\ Z &= \{z \in \mathbb{R}^4 \mid -8z_1 + 4z_2 - 2z_3 + z_4 > 0\} \end{aligned}$$

respectively define a closed half-space, hyperplane, and half-space in $\mathbb{R}^4$.

What does this have to do with LPs? Our constraint inequalities define closed half-spaces, and if we want to look at the set of feasible solutions, such points must satisfy all inequalities. Geometrically, the feasible solutions are contained in the intersection of all of the close half-spaces, and every point in this intersection must therefore be a feasible solution since it satisfies all of the constraints. All of this implies that the set of feasible solutions is a finite intersection of closed half-spaces.

Example 2.12. The following are the constraints for some LP and the corresponding set of feasible solutions.

$$\begin{aligned} x, y, z &\geqslant 0, \\ 5x + 3y + 5z &\leqslant 15, \\ 10x + 4y + 5z &\leqslant 20. \end{aligned}$$



But wait, there's more! The objective function in an LP in standard form is linear:

$$z = c^\top x.$$

We rephrase the LP in the following way.

Hyperplanes all the way down

Find the largest $k \in \mathbb{R}$ such that

$$c^{\top} x = k$$

subject to $x \geqslant 0$ and

$$Ax \leqslant b.$$

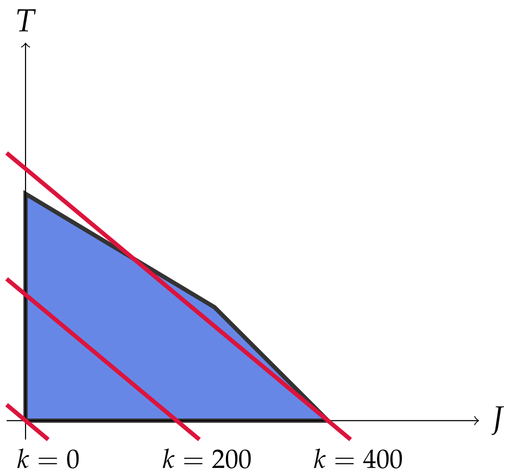
Example 2.13. Let's bring in the objective function from the sewing problem in Example 2.5. We want to maximize

$$z = 100J + 120T.$$

Instead, let's plot a number of hyperplanes (i.e. lines) of the form

$$100J + 120T = k$$

for different values of k. We use the feasibility region plotted in Example 2.10.



We can see the optimal solution to the sewing problem.

Example 2.14. Consider the following LP.

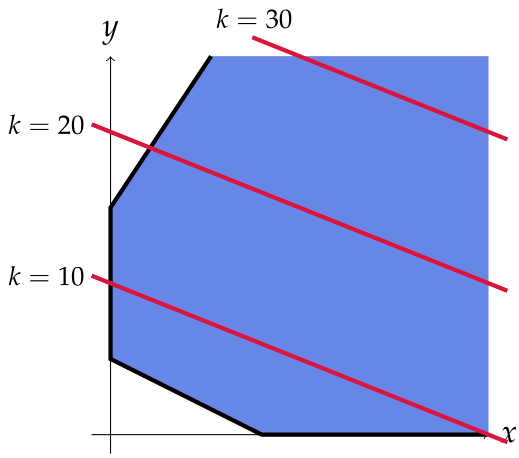
Maximize

$$z = 2x + 5y$$

subject to $x, y \geqslant 0$ and

$$\begin{aligned} -3x + 2y &\leqslant 6, \\ -x - 2y &\leqslant -2 \end{aligned}$$

Does the LP have an optimal solution? Let's plot the feasible solutions and a few hyperplanes of the form $2x + 5y = k$.



Note that the feasible solutions are not *bounded*—more on this later. No matter how large a k we get, we can always find feasible solutions that yield a larger k. Hence, there is no optimal solution.

Example 2.15. We take the LP from Example 2.14 and change the objective function slightly. Does it have an optimal solution?

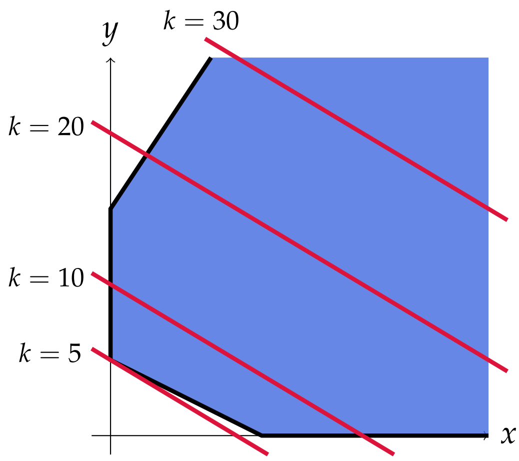
Minimize

$$z = 3x + 5y$$

subject to $x, y \geqslant 0$ and

$$\begin{aligned} -3x + 2y &\leqslant 6, \\ -x - 2y &\leqslant -2 \end{aligned}$$

Again, we'll just plot it.



There is an optimal solution. It occurs at $(0, 1)$ where $z = 5$.

Definition 2.16. A set $S \subseteq \mathbb{R}^n$ is ***bounded*** if there exists $r > 0$ such that for all $u, v \in S$ the Euclidean distance $d(u, v) \leqslant r$.

Note the order of quantifiers in Definition 2.16! Informally speaking, a set is bounded if we can wrap it in a ball (of finite radius). Are feasible solutions always bounded? Unbounded?

2.5 Convexity

Let's look at some of the feasible solutions we have plotted so far.

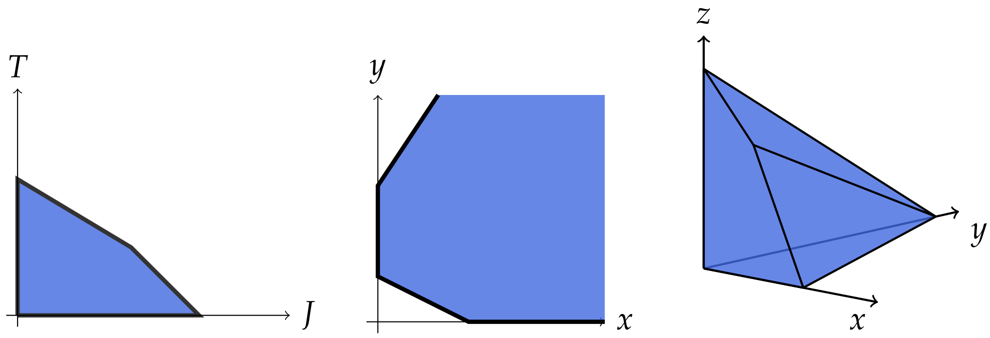


Figure 2.2: Three sets of feasible solutions

These regions have the property that if one takes two points in that region, say u and v, then all of the points on the line segment between u and v are also in that region. This is not true of all sets in $\mathbb{R}^n$; can you draw an example?

The formula for the line segment between points u and v is

$$L_{u,v}(t) = vt + (1-t)u$$

where $t \in [0,1]$. Note that at the endpoints, we have

$$L_{u,v}(0) = u, \qquad\qquad L_{u,v}(1) = v.$$

In the middle, we have points like $L_{u,v}(1/2) = (u+v)/2$ and $L_{u,v}(1/5) = (4u+v)/5$.

Definition 2.17. A set $S \subseteq \mathbb{R}^n$ is ***convex*** if it contains all points on all line segments between every pair of points in S. In symbols, this means that for all $u, v \in S$,

$$\{L_{u,v}(t) \mid t \in [0,1]\} \subseteq S.$$

Proposition 2.18. *The feasible set of solutions of an LP forms a convex set.*

Week 2

Try to prove Proposition 2.18 yourself. Consider first proving that closed half-spaces are convex.

Let's consider two distinct feasible solutions u and v to an LP. We have two cases: either they take the same value in the objective function or they have different values. Assume the first, that is, suppose

$$c^\top u = c^\top v.$$

What can we say about the values of the points on the line segment $L_{u,v}$? Let $w = vt + (1-t)u$ for some $t \in [0,1]$. Then

$$\begin{aligned} c^\top w &= c^\top (tv + (1-t)u) \\ &= c^\top tv + c^\top (1-t)u \\ &= tc^\top v + c^\top u - tc^\top u \\ &= c^\top u. \end{aligned}$$

Hence, all points on the $L_{u,v}$ have the same value under the objective function.

Let's consider the second case, that is, the values are distinct. Assume that

$$c^\top u < c^\top v,$$

and suppose $w = tv + (1-t)u$ for some $t \in [0,1]$. Some of the same analysis applies in this case: namely,

$$\begin{aligned} c^\top w &= tc^\top v + c^\top u - tc^\top u \\ &= c^\top u + t(c^\top v - c^\top u). \end{aligned}$$

Therefore, for all $t \in [0,1]$, we have

$$c^\top u \leqslant c^\top w \leqslant c^\top v.$$

Hence, the endpoint of $L_{u,v}$ have the extreme values.

We summarize all of this in the following proposition.

Proposition 2.19. *Let S be the set of feasible solutions to an LP. If $L \subseteq S$ is a line segment, then one of the following holds.*

1. *The objective function is constant on L.*

2. *The endpoints of L are the extreme points under the objective function.*

2.6 Convex polyhedra

We briefly leave the world of linear optimization and discuss some polyhedral geometry.

Definition 2.20. A ***convex polyhedron*** is a finite intersection of closed half-spaces in $\mathbb{R}^n$.

Examples include regular polygons in $\mathbb{R}^2$, infinite cone in $\mathbb{R}^2$, and the platonic solids in $\mathbb{R}^3$. As we have discussed before, the set of feasible solutions of an LP are, therefore, convex polyhedra. Some non-examples include balls in every dimension and any non-convex set.

Definition 2.21. A point $x \in \mathbb{R}^n$ is a ***convex combination*** of points $t_1, \ldots, t_r \in \mathbb{R}^n$ if there exist $\lambda_1, \ldots, \lambda_r \in [0,1]$, with $\lambda_1 + \cdots + \lambda_r = 1$, such that

$$x = \lambda_1 t_1 + \cdots + \lambda_r t_r.$$

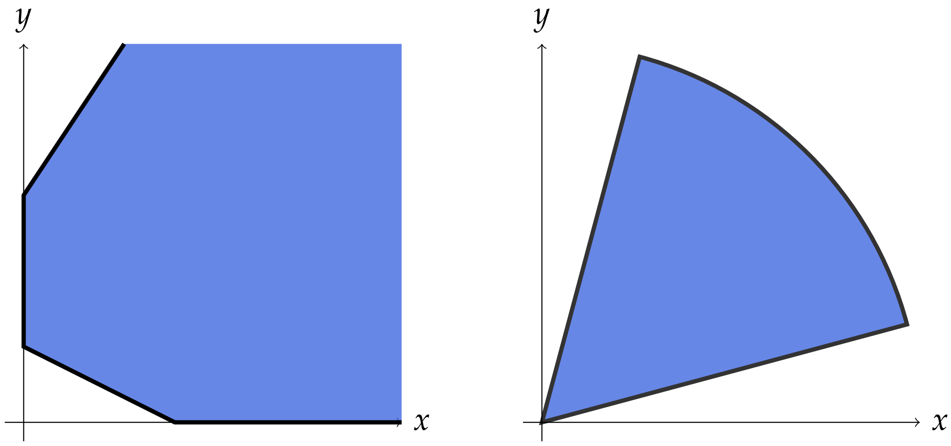


Figure 2.3: Two convex sets

Example 2.22. The point $(2,1,0)^\top$ is a convex combination of

$$\begin{bmatrix} 4 \\ 4 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ -4 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ -4 \end{bmatrix}.$$

Take $(\lambda_1, \lambda_2, \lambda_3) = (1/2, 1/4, 1/4)$.

The reason for the name "convex combination" is that the set of points that are convex combinations of a set of points is convex.

Definition 2.23. A point x in a convex set $S \subseteq \mathbb{R}^n$ is ***extreme*** if for every line segment in S, the point x is not in the interior.

Example 2.24. In Figure 2.3, one of the convex sets has infinitely many extreme points, and the other has exactly three. Which is which?

Proposition 2.25. *Let $S \subseteq \mathbb{R}^n$ be convex. A point $x \in S$ is extreme if and only if x is not a convex combination of other points in S.*

We won't prove Proposition 2.25.

2.7 Extreme point theorem

We state and prove some of the fundamental theorems in the theory of linear optimization.

Theorem 2.26 (Extreme points of an LP)**.** *Let S be the set of feasible solutions to an LP.*

(1) *If S is non-empty and bounded, then an optimal solution exists and occurs as an extreme point of S.*

(2) *If S is non-empty, unbounded, and contains an optimal solution, then the optimal solution occurs as an extreme point of S.*

(3) *If an optimal solution does not exist, then either S is empty or unbounded.*

Proof sketch. A subset of $\mathbb{R}^n$ is compact if and only if it is closed and bounded (Heine–Borel Theorem). Continuous real-valued functions on compact sets have a global maximum (fact from metric spaces or topology). The seat of feasible solutions form a convex polyhedron. By Proposition 2.25, an optimal solution is an extreme point. □

Example 2.27. Show that the following LP has infinitely many optimal solutions for $f(x,y) = 4x + 4y$. Show that for $f(x,y) = 4x + y$ there is a unique optimal solution.

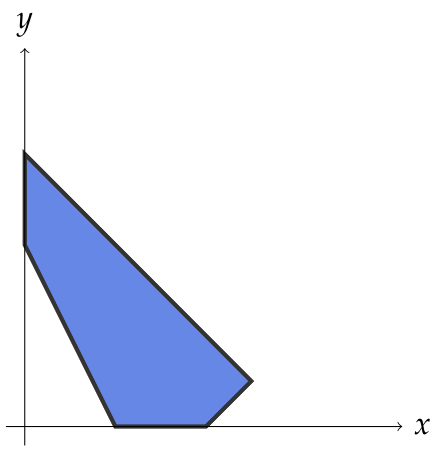
Maximize:

$$z = f(x,y),$$

subject to the constraints: $x, y \geqslant 0$ and

$$\begin{aligned} -2x - y &\leqslant -2, \\ x - y &\leqslant 2, \\ x + y &\leqslant 3. \end{aligned}$$

The plot of this LP looks like the following.



The extreme points are given by

$$\{(1,0),\ (2,0),\ (5/2,1/2),\ (0,2),\ (0,3)\}.$$

With $f(x,y) = 4x + 4y$, the values are $4, 8, 12, 8, 12$, respectively. Therefore, all points on the line segment between $(0,3)$ and $(5/2,1/2)$ are optimal solutions. If, instead, $f(x,y) = 4x + y$, then the values are $4, 8, 10.5, 2, 3$. Hence, we have a unique optimal solution at $(5/2,1/2)$.

Note that we had infinitely many optimal solutions when the line determined by $f(x,y) = 0$ was parallel to one of our constraints—the converse is not true in general: try $f(x,y) = 4x + 2y$, which is parallel to the first constraint.

Example 2.27 is simple because we can visualize fairly easily all of the extreme points. In fact, all extreme points for an LP with two variables occur at intersections of lines, which are easy to handle. Without some additional tools, it is not an easy task to find extreme points of higher-dimensional LPs. The next two theorems provide a roadmap to find these extreme points.

Theorem 2.28. *Suppose we have an LP in canonical form with constraints $x \geqslant 0$ and $Ax = b$ for some $A = \begin{bmatrix} a_1 & a_2 & \cdots & a_{n+s} \end{bmatrix} \in \mathrm{Mat}_{m\times n+s}(\mathbb{R})$. Assume that*

- *the first m columns of A, i.e. $\{a_1, \dots, a_m\}$, are linearly independent, and*
- *for some $x_1', \dots, x_m' \geqslant 0$, we have $x_1'a_1 + \cdots + x_m'a_m = b$.*

Then the following is an extreme point of the set of feasible solutions:

$$(x_1', \dots, x_m', 0, \dots, 0).$$

Proof. By assumptions, we know that $x = (x_1', \dots, x_m', 0 \dots, 0)$ is feasible. We need to show it is extreme. Assume x is not extreme, so there exists feasible points $u, v \in \mathbb{R}^m$ and $t \in (0, 1)$ such that

$$x = tv + (1-t)u.$$

This implies that for all $i \in m+1, \dots, n+s$ and all $j \in \{1, \dots, m\}$ we have

$$\begin{aligned} tv_i + (1-t)u_i &= 0, \\ tv_j + (1-t)u_j &= x_j'. \end{aligned}$$

Since $t \in (0, 1)$ and $u_i, v_i \geqslant 0$, it follows that $u_i = 0 = v_i$ for $i \in \{m+1, \dots, n+s\}$. As u is feasible, $Au = b$. Since $u_{m+1} = \cdots = u_{n+s} = 0$, we have

$$u_1a_1 + \cdots + u_ma_m = b.$$

By our assumptions, $u_j = x_j'$ for all $j \in \{1, \dots, m\}$, which is a contradiction. Hence, x is extreme. □

Theorem 2.29. *Suppose we have an LP in canonical form. If x is an extreme point of the set of feasible solutions, then the columns of A corresponding to positive coordinates of x form a set of linearly independent vectors of $\mathbb{R}^m$.*

Week 3

Proof. Reorganize the variables so that the first k coordinates of x are positive and all others zero. Thus,

$$x_1'a_1 + \cdots + x_k'a_k = b.$$

Suppose that $\{a_1, \dots, a_k\}$ is linearly dependent, so there exist scalars such that

$$\lambda_1a_1 + \cdots + \lambda_ka_k = 0,$$

where not all $\lambda_1, \dots, \lambda_k$ are zero. Therefore, we have two feasible solutions:

$$u = (x'_1 - \lambda_1, \dots, x'_k - \lambda_k, 0 \dots, 0), \qquad v = (x'_1 + \lambda_1, \dots, x'_k + \lambda_k, 0 \dots, 0).$$

Moreover $x = L_{u,v}(1/2)$, contradicting the fact that it is extreme. Hence, the set $\{a_1, \dots, a_k\}$ is linearly independent. $\square$

So from Theorem 2.29, the columns of A corresponding to positive entries of an extreme point x (contained in the set of feasible solutions to an LP in canonical form) are linearly independent. Since they exist in $\mathbb{R}^m$, and we cannot have more than m linearly independent vectors in $\mathbb{R}^m$, we have the following corollary to Theorem 2.29.

Corollary 2.30. *At most m entries of an extreme point can be positive. The rest are zero.*

Given an LP in canonical form with constraint matrix $A \in \mathrm{Mat}_{m\times s}(\mathbb{R})$ where $s \geqslant m$, we can select subsets of m columns of A that are linearly independent to find extreme points. Let's first give a name to these points.

Definition 2.31. A ***basic solution*** to $Ax = b$ is a vector x with exactly m nonzero entries. The variables associated to the zero entries of x are called ***non-basic variables***, and the others are called ***basic variables***.

Example 2.32. A basic solution to the equation

$$\begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & -1 & -1 & 0 & -1 & -1 \\ 1 & 2 & 2 & 1 & 1 & 1 \end{bmatrix} x = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

is obtained by finding three linearly independent columns. For example, the first, fourth, and fifth columns:

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & -1 \\ 1 & 1 & 1 \end{bmatrix} x' = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

yield the basic solution $x = (b_1 + b_2, 0, 0, b_3 - b_1, -b_2, 0)$.

Note that basic solutions need not be feasible solutions, that is, they may contain negative entries. A feasible solution that is also a basic solution is called a ***basic feasible solution***.

Exercise 2. Consider an LP in canonical form with

$$A = \begin{bmatrix} 2 & 3 & 1 & 0 & 0 \\ -1 & 1 & 0 & 2 & 1 \\ 0 & 6 & 1 & 0 & 3 \end{bmatrix}, \qquad b = \begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix}.$$

Which of the following points are basic solutions?

$$u = \begin{bmatrix} 0 \\ 2 \\ -5 \\ 0 \\ -1 \end{bmatrix}, \qquad v = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \qquad w = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 1 \\ 0 \end{bmatrix}$$

Are any of them basic feasible solutions?

Theorem 2.33 (Extremely basic feasible solutions)**.** *Every basic feasible solution of an LP in canonical form is an extreme point of the set of feasible solutions. The converse is also true.*

Assuming the constraint matrix $A \in \mathrm{Mat}_{m\times s}(\mathbb{R})$, with $s \geqslant m$, then we know an *upper bound* on the number of basic feasible solutions. It is

$$\binom{s}{m} = \frac{s!}{m!(s-m)!}.$$

We have been dealing primarily with canonical form. What can we do about standard form?

Suppose $x' \in \mathbb{R}^s$ is an extreme point of the set of feasible solutions in canonical form. Then by *truncating* x' to $x \in \mathbb{R}^m$ we obtain an extreme point of the set of feasible solutions in standard form. Thus, we go from SF to CF via adding slack variables, and from CF to SF by truncating those slack variables.

Try these problems out yourself.

Exercise 3. Consider an LP in CF with

$$A = \begin{bmatrix} 3 & 0 & 1 & 1 & 0 \\ 2 & 1 & 0 & 0 & 0 \\ 4 & 0 & 3 & 0 & 1 \end{bmatrix}, \qquad b = \begin{bmatrix} 5 \\ 3 \\ 6 \end{bmatrix}.$$

Which of the points

$$x_1 = \begin{bmatrix} 0 \\ 3 \\ 0 \\ 5 \\ 6 \end{bmatrix}, \quad x_2 = \begin{bmatrix} 0 \\ 3 \\ 5 \\ 0 \\ -9 \end{bmatrix}, \quad x_3 = \begin{bmatrix} 1 \\ 1 \\ 1/2 \\ 3/2 \\ 1/2 \end{bmatrix}, \quad x_4 = \begin{bmatrix} 1/2 \\ 1 \\ 1 \\ 0 \\ 2 \end{bmatrix}, \quad x_5 = \begin{bmatrix} 3/2 \\ 0 \\ 0 \\ 1/2 \\ 0 \end{bmatrix}$$

is

(i) a basic solution,

(ii) a basic feasible solution,

(iii) an extreme point of the set of feasible solutions,

(iv) a feasible solution.

Exercise 4. Consider the following LP.

Maximize

$$z = 4x + 2y + 7z$$

subject to the constraints $x, y, z \geqslant 0$ and

$$2x - y + 4z \leqslant 18,$$
$$4x + 2x + 5z \leqslant 10.$$

1. Put this into canonical form.
2. For each extreme point of the LP in canonical form, identify the basic variables.
3. Write down all of the extreme points for both the standard form and canonical form.
4. Which of the extreme points are optimal solutions?

3 The simplex method

By Section 2.7, optimal solutions to LPs are extreme points of a convex polyhedron. Although only finitely many, running through all of these points can be expensive. The key result of the simplex method is that we do not need to consider *all* extreme points. Instead, we start at an extreme point, and then move to a "neighboring" extreme point if it further maximizes our objective function.

3.1 Build up

Definition 3.1. Two distinct extreme points of an LP in CF are ***adjacent*** if as basic feasible solutions they have all but one basic variable in common.

Example 3.2. The pair of extreme points

$$(0, 0, 8, 15)^\top \qquad (3, 0, 2, 0)^\top$$

are adjacent, but the following pairs of points are not:

$$(0, 0, 8, 15)^\top \qquad (3/2, 5/2, 0, 0)^\top.$$

Week 4

Let's consider the sewing problem from Example 2.5 again. In canonical form, we have the following LP.

Maximize

$$z = 100J + 120T$$

subject to the constraints $J, T, s_1, s_2 \geqslant 0$ and

$$\begin{aligned} 2J + 2T + s_1 \qquad &= 8, \\ 3J + 5Y \qquad + s_2 &= 15. \end{aligned}$$

There is an obvious recipe to cook up a basic feasible solution: set all non-slack variables to 0 and "solve" for the slack variables. In our case, the point is

$(0,0,8,15)^\top$, and the basic variables are s_1 and s_2. Really, this should be thought of taking the third and fourth columns and solving.

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} = \begin{bmatrix} 8 \\ 15 \end{bmatrix}.$$

We can move to an adjacent extreme point by swapping out a basic variable. We could replace s_1 with J for example. Thus we would take the first and last columns of A to solve

$$\begin{bmatrix} 2 & 0 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} J \\ s_2 \end{bmatrix} = \begin{bmatrix} 8 \\ 15 \end{bmatrix}.$$

This yields the solution $(4,0,0,3)^\top$. We could have instead exchanged s_2 with T yielding

$$\begin{bmatrix} 1 & 2 \\ 0 & 5 \end{bmatrix} \begin{bmatrix} s_1 \\ T \end{bmatrix} = \begin{bmatrix} 8 \\ 15 \end{bmatrix}$$

and a solution of $(0,3,2,0)^\top$. For now, these choices are arbitrary, but this is the game that is played with the simplex method.

We will initialize a tableau to guide us through extreme points.

	J	T	s_1	s_2	z	
s_1	2	2	1	0	0	8
s_2	3	5	0	1	0	15
	-100	-120	0	0	1	0

Here, we have the two basic variables in the first (leftmost) column. The top row has all of the variables plus the objective function. The second row corresponds to the first constraint equation: $2J + 2T + s_1 = 8$, and the third row corresponds to the second constraint equation. The last row is the ***objective row***, and it corresponds to $-100J - 120T + z = 0$. The final column are the values of the equations at the particular extreme point determined by the basic variables in the first column. For example, the current extreme point is $(0,0,8,15)$, and evaluating this at $2J + 2T + s_1$, at $3J + 5T + s_2$, and $-100J - 120T + z$ yields 8, 15, and 0, respectively.

3.2 Basic Simplex Method

The Simplex Method can be thought of as a game, where we hop around extreme points until we are at an optimal solution. Once we hit an optimal solution, we are done. (We've won!) Otherwise we continue playing. We will address how we can tell whether or not we are on an optimal solution later. For now, we discuss how to play.

We move from one basic feasible solution to another adjacent one. We need to determine a pair of variables to exchange. The ***entering variable*** is the new basic variable; the ***departing variable*** is the outgoing variable. The fundamental questions we are to answer right now is how we decide which variable is which. To determine the entering variable, we apply a simple check: the entering variable corresponds to the column with the most negative (i.e. negative and largest in absolute value) coefficient in the objective row. The departing variable is a bit more complicated and will be more easily seen in an example.

Now we continue with our example. Note that the variable T has the most negative coefficient; therefore it is our entering variable. By how much can we increase T? Let's consider s_1 and s_2 has functions of the other variables:

$$\begin{aligned} s_1 &= 8 - 2J - 2T, \\ s_2 &= 15 - 3J - 5T. \end{aligned}$$

Since we are at $(0, 0, 8, 15)$ and we will have T as the new basic variable, we have that J is non-basic. Hence, $J = 0$, so we can simplify the above equations:

$$\begin{aligned} s_1 &= 8 - 2T, \\ s_2 &= 15 - 5T. \end{aligned}$$

Because $s_1, s_2 \geqslant 0$, we have two implied inequalities:

$$\begin{aligned} 4 &\geqslant T, \\ 3 &\geqslant T. \end{aligned} \tag{3.1}$$

Hence, the most restrictive constraint is that $3 \geqslant T$. In other words, if T were larger than 3, then $s_2 < 0$ which is not allowed. However, if $T = 3$, then $s_2 = 0$ and, hence, is non-basic. Thus, s_2 is the departing variable.

The values 4 and 3 in Equation (3.1) were obtained from $8/2$ and $15/5$, respectively. They are called ***θ-ratios***. We decide which variable should become the departing variable by taking the variable that corresponds to the smallest positive θ-ratio. Note that one can see these ratios in the initial tableau by looking at the final column "divided by" the T-column. Here is the tableau once again.

	J	T	s_1	s_2	z	
s_1	2	2	1	0	0	8
s_2	3	5	0	1	0	15
	-100	-120	0	0	1	0

The column corresponding to the entering variable is called the ***entering column*** or the ***pivotal column***. The row corresponding to the departing variable is called the ***departing row*** or the ***pivotal row***. The entry in both the pivotal row and column is called the ***pivot***. Hence, the 5 is the current pivot.

To form a new table, we update the column of basic variables and apply Gaussian elimination. We want that the entry in both the s_1 row and column to be 1 and likewise for T.

	J	T	s_1	s_2	z	
s_1	2	2	1	0	0	8
T	$3/5$	1	0	$1/5$	0	3
	-100	-120	0	0	1	0

Now we apply Gaussian elimination to clear the entries above and below the pivot to obtain our new tableau:

	J	T	s_1	s_2	z	
s_1	$4/5$	0	1	$-2/5$	0	2
T	$3/5$	1	0	$1/5$	0	3
	-28	0	0	24	1	360

Note we have moved from $(0,0,8,15)$ to $(0,3,2,0)$, and in the standard form of the same LP, we have moved from $(0,0)$ to $(0,3)$. Because we have negative entires in the objective row, we continue playing.

Our new entering variable is J since it corresponds to the most (and only) negative coefficient in the objective row. The corresponding θ-ratios are

$$s_1 : \frac{2}{4/5} = 5/2,$$
$$T : \frac{3}{3/5} = 5.$$

Since s_1 corresponds to the smallest positive ratio, it is the new departing variable. We update our tableau accordingly.

	J	T	s_1	s_2	z	
J	1	0	$5/4$	$-1/2$	0	$5/2$
T	0	1	$-3/4$	$1/2$	0	$3/2$
	0	0	35	10	1	430

Since there are no negative entries in the objective row, we are done and are therefore at an optimal solution. The extreme point we are on is $(5/2, 3/2, 0, 0)$, and the corresponding value under the objective function is 430. To address a question previously raised: we know when we are at an optimal solution if none of the entries in the objective row are negative.

I purposefully skipped over two potential issues. Notice that with θ-ratios, we consider only those that are positive.

1. What do we do with the negative θ-ratios?

2. What is we get a divide by zero error?

We'll consider both of these cases separately and with our running example—though slightly tweaked to fit these cases.

Suppose we have the following tableau.

	J	T	s_1	s_2	z	
s_1	$4/5$	0	1	$-2/5$	0	2
T	$-3/5$	1	0	$1/5$	0	3
	-28	0	0	24	1	360

Our entering variable is again J, and the θ-ratio corresponding to T is negative. Let's consider what this implies about the system. Treating T as a function of the rest, we have

$$T = 3 - \frac{1}{5}s_2 + \frac{3}{5}J = 3 + \frac{3}{5}J.$$

Since $T \geqslant 0$, the equality together with this inequality imply that

$$-5 \leqslant T,$$

but this is not new information since we are already assuming that $T \geqslant 0$. Hence, it can be ignored. And this is precisely the conclusion when we encounter θ-ratios: they do not impose any new constraints on our system and, therefore, can be ignored.

Similarly, we assume now that we have the following tableau.

	J	T	s_1	s_2	z	
s_1	$4/5$	0	1	$-2/5$	0	2
T	0	1	0	$1/5$	0	3
	-28	0	0	24	1	360

We cannot form the θ-ratio for T since that would involve dividing by 0. Retreating to the equation of T in terms of the rest yields

$$T = 3 - \frac{1}{5}s_2 = 3.$$

This again is not new information and can be deduced from the tableau. Hence, this (lack of a) θ-ratio can be ignored.

We arrive at two conclusions concerning θ-ratios:

Conclusion 1: Only the positive θ-ratios impose new constraints. All other can be ignored.

Conclusion 2: If in the pivotal column all entries above the objective row are non-positive, then the corresponding LP has no (finite) optimal solution.

Week 5

Example 3.3. Let's use this method to solve the following LP.

Maximize

$$z = x_1 + 3x_2 + 5x_3$$

subject to $x_1, x_2, x_3 \geqslant 0$ and

$$\begin{aligned} 2x_1 - 5x_2 + x_3 &\leqslant 3, \\ x_1 + 4x_2 \phantom{{}+x_3} &\leqslant 5. \end{aligned}$$

First we will convert this to canonical form by adding in two slack variables s_1 and s_2, and we will start with the extreme point $(0,0,0,3,5)$. Thus, our initial tableau is

	x_1	x_2	x_3	s_1	s_2	z	
s_1	2	-5	1	1	0	0	3
s_2	1	4	0	0	1	0	5
	-1	-3	-5	0	0	1	0

Our first entering variable is x_3 because -5 is the most negative entry in the objective row. There is only one θ-ratio that is positive, and it corresponds to s_1. Hence s_1 is our first departing variable. Thus, the new tableau is

	x_1	x_2	x_3	s_1	s_2	z	
x_3	2	-5	1	1	0	0	3
s_2	1	4	0	0	1	0	5
	9	-28	0	5	0	1	15

Now we see that the next entering variable is x_2. The two θ-ratios are $-3/5$ and $5/4$. Since we ignore negative θ-ratios, the next departing variable is s_2. Hence the new tableau is

	x_1	x_2	x_3	s_1	s_2	z	
x_3	$13/4$	0	1	1	$5/4$	0	$37/4$
x_2	$1/4$	1	0	0	$1/4$	0	$5/4$
	16	0	0	5	7	1	50

Therefore the optimal solution to the LP in canonical form is $(0,5/4,37/4,0,0)$, and hence, the optimal solution to the original LP is $(0,5/4,37/4)$.

3.3 Artificial variables

Two problems could arise that would prevent us from using what we discussed. Suppose we have an LP in canonical form with constraint matrix $A \in \mathrm{Mat}_{m\times s}(\mathbb{R})$ and $b \in \mathbb{R}^m$.

1. What if $b \not\geq 0$?
2. What if the last m columns of A are not an identity matrix?

In both cases, we do not get an initial basic feasible solution by $(0,b) \in \mathbb{R}^{m+s}$.

We modify the Simplex Method from Section 3.2 to have two phases, called the Two-Phase Method. We must augment our LP to account for these problems.

We saw in Section 2.1 that we can put all LPs into standard form, but now we want to make sure that we do so in such a way that the b vector is non-negative. Suppose we have the following *general* LP.

Maximize

$$z = c^\top x$$

subject to $x \geqslant 0$ and

$$A_1 x \geqslant b_1, \qquad A_2 x \leqslant b_2, \qquad A_3 x = b_3,$$

where $A_i \in \mathrm{Mat}_{r_i \times n}(\mathbb{R})$ and $r_1 + r_2 + r_3 = m$ and $b_i \geqslant 0$.

Normally, we would negate and add slack; however instead of negating, we will just include the slack. We add r_2 "positive" slack variables and r_1 "negative" slack variables. Thus we have, for $t = (t_1, \ldots, t_{r_1})$ and $s = (s_1, \ldots, s_{r_2})$,

$$A_1 x - I_{r_1} t = b_1, \qquad A_2 x + I_{r_2} s = b_2, \qquad A_3 x = b_3.$$

Example 3.4. Rewrite the LP so that it is in canonical form and the b-vector is non-negative:

Maximize

$$z = 2x_1 + 5x_2$$

subject to $x \geqslant 0$ and

$$\begin{aligned} 2x_1 + 3x_2 &\leqslant 6, \\ -2x_1 + x_2 &\leqslant -2, \\ x_1 - 6x_2 &= -2. \end{aligned}$$

We can rewrite it as follows.

Maximize

$$z = 2x_1 + 5x_2$$

subject to $x \geqslant 0$ and

$$\begin{aligned} 2x_1 + 3x_2 + x_3 &= 6, \\ 2x_1 - x_2 - x_4 &= 2, \\ -x_1 + 6x_2 &= 2. \end{aligned}$$

Although the b-vector is non-negative in Example 3.4, we still do not have an initial basic feasible solution. Now we must take an arbitrary LP in canonical form, without the condition that the last m columns are an identity matrix, and add ***artificial variables***. We simply add a new variable to each constraint equation—we could be more conservative with these artificial variables, but we don't care. We can rewrite the LP from Example 3.4 as follows:

Maximize

$$z = 2x_1 + 5x_2$$

subject to $x, y \geqslant 0$ and

$$\begin{aligned} 2x_1 + 3x_2 + x_3 + y_1 &= 6, \\ 2x_1 - x_2 - x_4 + y_2 &= 2, \\ -x_1 + 6x_2 + y_3 &= 2. \end{aligned}$$

Clearly $x = 0$ and $y = b$ yields an initial basic feasible solution now. However, we want $y = 0$, so we apply a two-phase approach, which we describe next.

3.4 Two-Phase Method

Suppose we have an LP in canonical form with a non-negative b-vector and each constraint equation has an artificial variable. We will apply the Simplex Method to a different LP in order to produce a basic feasible solution to the original problem (without artificial variables). This is called Phase 1. The second phase begins when we have a basic feasible solution to the original problem; then we can apply the Simplex Method as before. This is Phase 2.

Suppose the following is our LP in canonical form with artificial variables y:

Maximize

$$z = c^\top x$$

subject to $x, y \geqslant 0$

$$Ax + Iy = b.$$

We want $y = 0$, so in order to ensure this, we want to minimize

$$z' = y_1 + y_2 + \cdots + y_m.$$

Since $y \geqslant 0$, $z' = 0$ implies that $y_1 = \cdots = y_m = 0$. But instead of minimizing an objective function, we can maximize its negative. Hence, we want to maximize

$$z = -y_1 - y_2 - \cdots - y_m.$$

To make the artificial variables like slack variables from before, we need them out of the objective function. Since each artificial variable occurs in exactly one constraint equation, we can solve for the artificial variables. Instead of writing out the formulas, we illustrate by example.

Example 3.5. Use artificial variables to convert this LP in canonical form to a canonical form with artificial variables to use with the 2-Phase Method.

Maximize

$$z = 2x_1 + 5x_2$$

subject to $x \geqslant 0$ and

$$\begin{aligned} 2x_1 + 3x_2 + x_3 &= 6, \\ 2x_1 - x_2 - x_4 &= 2, \\ -x_1 + 6x_2 &= 2. \end{aligned}$$

We define artificial variables y_1, y_2, y_3 and define our constraints to be

$$\begin{aligned} 2x_1 + 3x_2 + x_3 + y_1 &= 6, \\ 2x_1 - x_2 - x_4 + y_2 &= 2, \\ -x_1 + 6x_2 + y_3 &= 2. \end{aligned}$$

Now we want to maximize

$$z = -y_1 - y_2 - y_3.$$

Using the constraint equations, we can remove the y_i from the objective function:

$$\begin{aligned} z &= -y_1 - y_2 - y_3 \\ &= -(6 - 2x_1 - 3x_2 - x_3) - (2 - 2x_1 + x_2 + x_4) - (2 + x_1 - 6x_2) \\ &= 3x_1 + 8x_2 + x_3 - x_4 - 10. \end{aligned}$$

Thus, the LP we want is

Maximize

$$z = 3x_1 + 8x_2 + x_3 - x_4 - 10$$

subject to $x, y \geqslant 0$ and

$$\begin{aligned} 2x_1 + 3x_2 + x_3 + y_1 &= 6, \\ 2x_1 - x_2 - x_4 + y_2 &= 2, \\ -x_1 + 6x_2 + y_3 &= 2. \end{aligned}$$

We can simply apply the Simplex Method to the LP we get after this process. This is Phase 1, and there are two possible outcomes that mark the end of it:

1. all artificial variables are non-basic variables,

2. it is not possible to get all artificial variables non-basic.

We will discuss only the first case right now.

The final tableau from Phase 1 will be the initial tableau in Phase 2 with the following modifications.

1. Delete all columns labeled with artificial variables.

2. Calculate a new objective row as follows. Start with the original objective function (not the one we got after modifying). Use the basic variables to clear their corresponding entries in the objective row.

Phase 2 then continues by applying the Simplex Method.

To summarize before looking at examples, in less than ideal cases, we add artificial variables and use the Simplex Method on an appropriate (but different) LP. This will hopefully yield a basic feasible solution of our original problem, which we can then feed into the Simplex Method again from before to find an optimal solution.

References

[1] G. B. Dantzig. Inductive proof of the simplex method. https://apps.dtic.mil/sti/tr/pdf/AD0224306.pdf, 1960.

[2] G. B. Dantzig. Reminiscences about the origins of linear programming. *Operations Research Letters*, 1(2):43–48, 1982.

[3] National Science Foundation. The President's National Medal of Science: Recipient Details. https://www.nsf.gov/od/nms/recip_details.jsp?recip_id=95. Accessed: 2024-09-08.