1

Introduction

Broadly speaking, optimization is the problem of minimizing or maximizing a function subject to a number of constraints. Optimization problems are ubiquitous. Every chief executive officer (CEO) is faced with the problem of maximizing profit given limited resources. In general, this is too general a problem to be solved exactly; however, many aspects of decision making can be successfully tackled using optimization techniques. This includes, for instance, production, inventory, and machine-scheduling problems. Indeed, the overwhelming majority of Fortune 500 companies make use of optimization techniques. However, optimization problems are not limited to the corporate world. Every time you use your GPS, it solves an optimization problem, namely how to minimize the travel time between two different locations. Your hometown may wish to minimize the number of trucks it requires to pick up garbage by finding the most efficient route for each truck. City planners may need to decide where to build new fire stations in order to efficiently serve their citizens. Other examples include: how to construct a portfolio that maximizes its expected return while limiting volatility; how to build a resilient tele-communication network as cheaply as possible; how to schedule flights in a cost-effective way while meeting the demand for passengers; or how to schedule final exams using as few classrooms as possible.

Suppose that you are a consultant hired by the CEO of the WaterTech company to solve an optimization problem. Say for simplicity that it is a maximization problem. You will follow a two-step process:

(1) find a *formulation* of the optimization problem,

(2) use a suitable *algorithm* to solve the formulation.

A formulation is a mathematical representation of the optimization problem. The various parameters that the WaterTech CEO wishes to determine are represented as *variables* (unknowns) in your formulations. The *objective function* will represent the quantity that needs to be maximized. Finally, every constraint to the problem is expressed as a *mathematical constraint*.

Now given a mathematical formulation of an appropriate form, you need to develop (or use an existing) algorithm to solve the formulation. By an algorithm, we mean in this case a finite procedure (something that can be coded as a computer program) that will take as input the formulation, and return an assignment of values to each of the variables such that all constraints are satisfied, and, subject to these conditions, maximizes the objective function. The values assigned to the variables indicate the optimal choices for the parameters that the CEO of WaterTech wishes to determine.

This two-step process is summarized in Figure 1.1. In this chapter, we will focus our attention on the first step, namely how to formulate optimization problems.

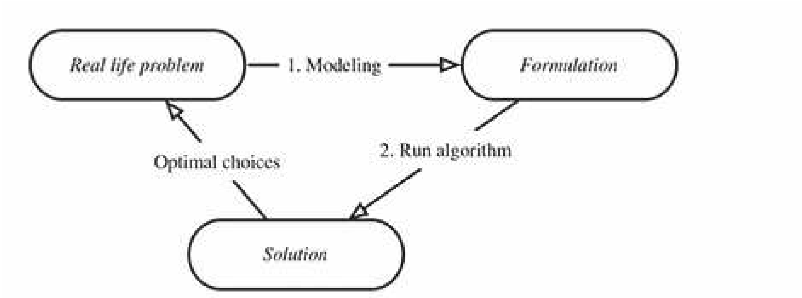


Figure 1.1 Modeling and solving optimization problems.

1.1 A first example

To clarify these ideas, let us consider a simple example. Suppose WaterTech manufactures four products, requiring time on two machines and two types (skilled and unskilled) of labor. The amount of machine time and labor (in hours) needed to produce a unit of each product and the sales prices in dollars per unit of each product are given in the following table:

Product	Machine 1	Machine 2	Skilled labor	Unskilled labor	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, 700 hours are available on machine 1 and 500 hours on machine 2. Each month, WaterTech can purchase up to 600 hours of skilled labor at \$8 per hour and up to 650 hours of unskilled labor at \$6 per hour. The company wants to determine how much of each product it should produce each month and how much labor to purchase in order to maximize its profit (i.e. revenue from sales minus labor costs).

1.1.1 The formulation

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

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

Objective function. Deciding on a production plan now amounts to finding values for variables x_1,

. . . , x_4, y_s and y_u. Once the values for these variables have been found, WaterTech's profit is easily expressed by the following function:

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

and the company wants to maximize this quantity.

Constraints. We manufacture x_1 units of product 1 and each unit of product 1 requires 11 hours on machine 1. Hence, product 1 will use $11x_1$ hours on machine 1. Similarly, for machine 1, product 2 will use $7x_2$ hours, product 3 will use $6x_3$ hours, and product 4 will use $5x_4$ hours. Hence, the total amount of time needed on machine 1 is given by

$$11x_1 + 7x_2 + 6x_3 + 5x_4,$$

and this must not exceed the available 700 hours of time on that machine. Thus

$$11x_1 + 7x_2 + 6x_3 + 5x_4 \leq 700. \tag{1.1}$$

In a similar way, we derive a constraint for machine 2

$$4x_1 + 6x_2 + 5x_3 + 4x_4 \leq 500. \tag{1.2}$$

Analogously, once we decide how much of each product should be produced, we know how much skilled and unskilled labor is needed. Naturally, we need to make sure that enough hours of each type of labor are purchased. The following two constraints enforce this:

$$8x_1 + 5x_2 + 5x_3 + 6x_4 \leq y_s, \tag{1.3}$$

$$7x_1 + 8x_2 + 7x_3 + 4x_4 \leq y_u. \tag{1.4}$$

Finally, we need to add constraints that force each of the variables to take on only nonnegative values as well as constraints that limit the number of hours of skilled and unskilled labor purchased. Combining the objective function with (1.1)–(1.4) gives the following formulation:

$$
\begin{array}{ll}
\max & 300x_1 + 260x_2 + 220x_3 + 180x_4 - 8y_s - 6y_u \\
\text{subject to} & \\
& 11x_1 + 7x_2 + 6x_3 + 5x_4 \leq 700 \\
& 4x_1 + 6x_2 + 5x_3 + 4x_4 \leq 500 \\
& 8x_1 + 5x_2 + 5x_3 + 6x_4 \leq y_s \\
& 7x_1 + 8x_2 + 7x_3 + 4x_4 \leq y_u \\
& y_s \leq 600 \\
& y_u \leq 650 \\
& x_1, x_2, x_3, x_4, y_u, y_s \geq 0.
\end{array} \tag{1.5}
$$

1.1.2 Correctness

Is the formulation given by (1.5) correct, i.e. does this formulation capture exactly the WaterTech problem? We will argue that it does and outline a procedure to verify whether a given formulation is correct. Let us introduce a bit of terminology. By the *word description* of the optimization problem, we mean a description of the optimization problem in plain English. This is the description that the CEO of WaterTech would give you. By a *formulation*, we mean the mathematical formulation, as in (1.5). A *solution* to the formulation is an assignment of values to each of the variables of the formulation. A solution is *feasible* if it has the property that all the constraints are satisfied. An *optimal solution* to the formulation is a feasible solution that maximizes the objective function (or minimizes it if the optimization problem is a minimization problem). Similarly, we define a solution to the word description of the optimization problem to be a choice for the unknowns, and a feasible solution to be such a choice that satisfies all the constraints.

To construct a formulation for an optimization problem, there are many approaches. Not all of them may apply to a given problem. Conceptually, an easy approach is to make sure that there is a mapping between feasible solutions of the word description and feasible solutions of the formulation and vice versa (between feasible solutions of the formulation and feasible solutions of the word description). For instance, a feasible solution for WaterTech is to produce 10 units of product 1, 50 of product 2, 0 units of product 3, and 20 of product 4, and buy 600 hours of both skilled and unskilled labor. This corresponds to the following feasible solution of the formulation:

$$x_1 = 10, x_2 = 50, x_3 = 0, x_4 = 20, y_s = 600, y_u = 600. \tag{1.6}$$

Conversely, given the feasible solution (1.6), we can construct a feasible solution for the word description of the WaterTech problem. Note that this works for every feasible solution. When constructing a formulation using this approach, you need to make sure that through the map that you defined:

(1) every feasible solution of the word description gives a feasible solution of the mathematical formulation, and

(2) every feasible solution of the mathematical formulation gives a feasible solution of the word description.

If (2) does not hold, feasible solutions to the formulation may violate constraints of the word description, and if (1) does not hold, then the formulation is more restrictive than the word description. A common mistake is to violate (2) by forgetting some constraint when writing down the formulation. For instance, we may forget to write down the condition for WaterTech that y_s is nonnegative. In this case, when solving the formulation using an algorithm, we may end up with a negative value for y_s, i.e. we buy a negative amount of skilled labor or equivalently we sell skilled labor; the latter is not allowed in our word description. Thus far, we have only discussed feasible solutions. Clearly, we also need to verify that the objective function in the word description and the formulation are the same. This is clearly the case for the WaterTech formulation, and is usually straightforward to verify.

We used an algorithm to find an optimal solution to (1.5) and obtained

$$x_1 = 16 + \frac{2}{3},\ x_2 = 50,\ x_3 = 0,\ x_4 = 33 + \frac{1}{3},\ y_s = 583 + \frac{1}{3},\ y_u = 650, \tag{1.7}$$

achieving a total profit of \$15, 433 + $\frac{1}{3}$. Thus, the optimal strategy for WaterTech is to manufacture 16 + $\frac{2}{3}$ units of product 1, 50 units of product 2, 0 units of product 3, 33 + $\frac{1}{3}$ units of product 4, and to buy 583 hours of skilled labor and 650 units of unskilled labor.

Since constructing the formulation is only our first step (see Figure 1.1) and we need to use an algorithm to find an optimal solution to the formulation, we will strive to get, among all possible formulations, one that is as simple as possible. In the remainder of the chapter, we will introduce three types of formulation:

- linear programs (Section 1.2),
- integer programs (Section 1.3), and
- nonlinear programs (Section 1.6).

There are efficient techniques to solve linear programs and we will see some of these in Chapters 2 and 7. Integer programs and nonlinear programs can be hard to solve however. Thus, we will always attempt to formulate our problem as a linear program. Unfortunately, this may not always be possible, and sometimes the only valid formulation is an integer program or a nonlinear program.

1.2 Linear programs

A function $f : \mathbb{R}^n \to \mathbb{R}$ is an *affine* function if $f(x) = a^\top x + \beta$, where x and a are vectors with n entries and β is a real number. If $\beta = 0$, then f is a *linear function*. Thus, every linear function is affine, but the converse is not true.

Example 1 Suppose $x = (x_1, x_2, x_3, x_4)^\top$. Then:

1. $f(x) = x_1 - x_3 + x_4$ is a linear function,
2. $f(x) = 2x_1 - x_3 + x_4 - 6$ is an affine function, but not a linear function, and
3. $f(x) = 3x_1 + x_2 - 6x_3 x_4$ is not an affine function (because of the product $x_3 x_4$).

A *linear constraint* is any constraint that is of one of the following forms (after moving all variables to the left-hand side and all constants to the right-hand side):

$$f(x) \leq \beta \quad \text{or} \quad f(x) \geq \beta \quad \text{or} \quad f(x) = \beta,$$

where $f(x)$ is a linear function, and β is a real number. A *linear program* (LP) is the problem of maximizing or minimizing an affine function subject to a finite number of linear constraints. We will abbreviate the term linear program by LP, throughout this book.

Example 2

(a) The following is an LP:

$$\begin{array}{ll} \min & x_1 - 2x_2 + x_4 \\ \text{subject to} & \\ & \begin{aligned} x_1 - x_3 &\leq 3 \\ x_2 + x_4 &\geq 2 \\ x_1 + x_2 &= 4 \end{aligned} \\ & x_1, x_2, x_3, x_4 \geq 0. \end{array}$$

(b) The following is not an LP, as $x_2 + x_3 < 3$ is not a linear constraint:[1]

$$\begin{array}{ll} \max & x_1 \\ \text{subject to} & \\ & \begin{aligned} x_2 + x_3 &< 3 \\ x_1 - x_4 &\geq 1. \end{aligned} \end{array}$$

(c) The following is not an LP, as the objective function is not affine:

$$\begin{array}{ll} \min & \frac{1}{x_1} \\ \text{subject to} & \\ & x_1 \geq 3. \end{array}$$

(d) The following is not an LP, as there are an infinite number of constraints:

$$\begin{array}{ll} \max & x_1 \\ \text{subject to} & \\ & x_1 + \alpha x_2 \leq 5 \qquad \text{(for all real numbers } \alpha \in [2, 3]). \end{array}$$

Observe also that (1.5) is an example of an LP. In that example, the constraint $8x_1 + 5x_2 + 5x_3 + 6x_4 \leq y_s$ can be rewritten as $8x_1 + 5x_2 + 5x_3 + 6x_4 - y_s \leq 0$.

1.2.1 Multiperiod models

In this section, we present another example of a type of optimization problem that can be formulated as an LP. In the problem discussed in Section 1.1, we were asked to make a one-time decision on a production plan. Often times, the decision-making process has a temporal component; time is split into so-called *periods* and we have to make certain decisions at the beginning or end of each of them. Each of these decisions will naturally determine the final outcome at the end of all periods. We introduce this area with an example. KWOil is a local supplier of heating oil. The company has been around for many years, and knows its home turf. In particular, KWOil has developed a dependable model to forecast future demand for oil. For each of the following four months, the company expects the following amounts of demand for heating oil.

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

At the beginning of each of the four months, KWOil may purchase heating oil from a regional supplier at the current market rate. The following table shows the projected price per litre at the beginning of each of these months:

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

KWOil has a small storage tank on its facility. The tank can hold up to 4000 litres of oil, and currently (at the beginning of month 1) contains 2000 litres. The company wants to know how much oil it should purchase at the beginning of each of the four months such that it satisfies the projected customer demand at the minimum possible total cost. Note, oil that is delivered at the beginning of each month can be delivered directly to the customer, it does not need to be first put into storage; only oil that is left over at the end of the month goes into storage. We wish to find an LP formulation for this problem. Thus, we need to determine the variables, the objective function, and the constraints.

Variables. KWOil needs to decide how much oil to purchase at the beginning of each of the four months. We therefore introduce variables p_i for $i \in \{1, 2, 3, 4\}$ denoting the number of litres of oil purchased at the beginning of month i for $i \in \{1, 2, 3, 4\}$. We also introduce variables t_i for each $i \in \{1, 2, 3, 4\}$ to denote the number of litres of heating oil in the company's tank at the beginning of month i (we already know that $t_1 = 2000$ – we can substitute this value later as we finish constructing our mathematical formulation). Thus, while every unknown of the word description always needs to be represented as a variable in the formulation, it is sometimes useful, or necessary, to introduce additional variables to keep track of various parameters.

Objective function. Given the variables defined above, it is straightforward to write down the cost incurred by KWOil. The objective function of KWOil's problem is

$$\min \quad 0.75p_1 + 0.72p_2 + 0.92p_3 + 0.90p_4. \tag{1.8}$$

Constraints. In each month i, the company needs to have enough heating oil available to satisfy customer demand. The amount of available oil at the beginning of month 1, for example, is comprised of two parts: the p_1 litres of oil purchased in month 1, and the t_1 litres contained in the tank. The sum of these two quantities needs to cover the demand in month 1, and the excess is stored in the local tank. Hence, we obtain the following constraint:

$$p_1 + t_1 = 5000 + t_2. \tag{1.9}$$

We obtain similar constraints for months 2 and 3

$$p_2 + t_2 = 8000 + t_3, \tag{1.10}$$

$$p_3 + t_3 = 9000 + t_4. \tag{1.11}$$

Finally, in order to satisfy the demand in month 4, we need to satisfy the following constraint:

$$p_4 + t_4 \geq 6000. \tag{1.12}$$

Notice that each of the variables t_i for $i \in \{2, 3, 4\}$ appears in two of the constraints (1.9) and (1.12). The constraints are therefore linked by the variables t_i. Such linkage is a typical feature in multiperiod models. Constraints (1.9) and (1.12) are sometimes called *balance constraints* as they *balance* demand and inventory between periods.

We now obtain the entire LP for the KWOil problem by combining (1.8)–(1.12), and by adding upper bounds and initialization constraints for the tank contents, as well as non-negativity constraints

$$
\begin{array}{lll}
\min & 0.75p_1 + 0.72p_2 + 0.92p_3 + 0.90p_4 & \\
\text{subject to} & & \\
& p_1 + t_1 = 5000 + t_2 & \\
& p_2 + t_2 = 8000 + t_3 & \\
& p_3 + t_3 = 9000 + t_4 & \\
& p_4 + t_4 \geq 6000 & \\
& t_1 \quad = 2000 & \\
& t_i \quad \leq 4000 & (i = 2, 3, 4) \\
& t_i, p_i \quad \geq 0 & (i = 1, 2, 3, 4).
\end{array}
$$

Solving this LP yields

$$p_1 = 3000, p_2 = 12000, p_3 = 5000, p_4 = 6000, t_1 = 2000, t_2 = 0, t_3 = 4000, t_4 = 0,$$

corresponding to a total purchasing cost of \$20 890. Not surprisingly, this solution suggests to take advantage of the low oil prices in month 2, while no oil should be stored in month 3 when prices are higher.

Exercises

1 Consider the following table indicating the nutritional value of different food types:

Foods	Price ($) per serving	Calories per serving	Fat (g) per serving	Protein (g) per serving	Carbohydrate (g) per serving
Raw carrots	0.14	23	0.1	0.6	6
Baked potatoes	0.12	171	0.2	3.7	30
Wheat bread	0.2	65	0	2.2	13
Cheddar cheese	0.75	112	9.3	7	0
Peanut butter	0.15	188	16	7.7	2

You need to decide how many servings of each food to buy each day so that you minimize the total cost of buying your food while satisfying the following daily nutritional requirements:

- calories must be at least 2000,
- fat must be at least 50g,
- protein must be at least 100g,
- carbohydrates must be at least 250g.

Write an LP that will help you decide how many servings of each of the aforementioned foods are needed to meet all the nutritional requirement, while minimizing the total cost of the food (you may buy fractional numbers of servings).

2 MUCOW (Milk Undertakings, C and O, Waterloo) owns a herd of Holstein cows and a herd of Jersey cows. For each herd, the total number of litres produced each day, and milk-fat content (as a *percentage*) are as follows:

	Litres produced	Percent milk-fat
Holstein	500	3.7
Jersey	250	4.9

The fat is split off and blended again to create various products. For each product, the volume, required milk-fat percentage, and profit are as follows. In particular, the milk-fat percentage must exactly equal what is specified.

	Skimmed milk	2%	Whole milk	Table cream	Whipping cream
Volume (litres)	2	2	2	0.6	0.3
Milk-fat requirement	0%	2%	4%	15%	45%
Profit ($)	0.1	0.15	0.2	0.5	1.2

(a) Formulate as an LP the problem of deciding how many items of each type to produce, so that the total profit is maximized. Don't worry about fractional numbers of items. Write your formulation in matrix notation.

(b) MUCOW is told of a regulation change: 'skimmed milk' can now contain anything up to 0.1% milk-fat, but no more. How does the formulation of the problem change? Note the resulting formulation should also be an LP.

3 The director of the CO-Tech startup needs to decide what salaries to offer its employees for the coming year. In order to keep the employees satisfied, she needs to satisfy the following constraints:

- Tom wants at least \$20 000 or he will quit;
- Peter, Nina, and Samir each want to be paid at least \$5000 more than Tom;
- Gary wants his salary to be at least as high as the combined salary of Tom and Peter;
- Linda wants her salary to be \$200 more than Gary;
- the combined salary of Nina and Samir should be at least twice the combined salary of Tom and Peter;
- Bob's salary is at least as high as that of Peter and at least as high as that of Samir;
- the combined salary of Bob and Peter should be at least \$60 000;
- Linda should not make more money than the combined salary of Bob and Tom.

(a) Write an LP that will determine salaries for the employees of CO-tech that satisfy each of these constraints while minimizing the total salary expenses.

(b) Write an LP that will determine salaries for the employees of CO-tech that satisfy each of these constraints while minimizing the salary of the highest paid employee.

(c) What is the relation between the solutions for (a) and (b)?

4 You wish to build a house and you have divided the process into a number of tasks, namely:

B. excavation and building the foundation,

F. raising the wooden frame,

E. electrical wiring,

P. indoor plumbing,

D. dry walls and flooring,

L. landscaping.

You estimate the following duration for each of the tasks (in weeks):

Task	B	F	E	P	D	L
Duration	3	2	3	4	1	2

Some of the tasks can only be started when some other tasks are completed. For instance, you can only build the frame once the foundation has been completed, i.e. F can start only after B is completed. All the precedence constraints are summarized as follows:

- F can start only after B is completed,
- L can start only after B is completed,
- E can start only after F is completed,
- P can start only after F is completed,
- D can start only after E is completed,
- D can start only after P is completed.

The goal is to schedule the starting time of each task such that the entire project is completed as soon as possible.

As an example, here is a feasible schedule with a completion time of ten weeks.

Tasks	B	F	E	P	D	L
Starting time	0	3	6	5	9	6
End time	3	5	9	9	10	8

Formulate this problem as an LP. Explain your formulation. Note, that there is no limit on the number of tasks that can be done in parallel.

HINT: Introduce variables to indicate the times that the tasks start.

5 The CRUD chemical plant produces as part of its production process a noxious compound called chemical X. Chemical X is highly toxic and needs to be disposed of properly. Fortunately, CRUD is linked by a pipe system to the FRESHAIR recycling plant that can safely reprocess chemical X. On any give day, the CRUD plant will produce the following amount of Chemical X (the plant operates between 9am and 3pm only):

Time	9–10 am	10–11 am	11am–12pm	12–1 pm	1–2 pm	2–3 pm
Chemical X (in litres)	300	240	600	200	300	900

Because of environmental regulation, at no point in time is the CRUD plant allowed to keep more than 1000 litres on site and no chemical X is allowed to be kept overnight. At the top of every hour, an arbitrary amount of chemical X can be sent to the FRESHAIR recycling plant. The cost of recycling chemical X is different for every hour:

Time	10am	11am	12pm	1pm	2pm	3pm
Price ($ per litre)	30	40	35	45	38	50

You need to decide how much chemical to send from the CRUD plant to the FRESHAIR recycling plant at the top of each hour, so that you can minimize the total recycling cost but also meet all the environmental constraints. Formulate this problem as an LP.

6 We are given an m by n matrix A and a vector b in $\mathbb{R}^m$, for which the system $Ax = b$ has no solution. Here is an example:

$$2x_1 - x_2 = -1$$

$$x_1 + x_2 = 1$$

$$x_1 + 3x_2 = 4$$

$$-2x_1 + 4x_2 = 3.$$

We are interested in finding a vector $x \in \mathbb{R}^n$ that "comes close" to solving the system. Namely, we want to find an $x \in \mathbb{R}^n$ whose *deviation* is minimum, and where the deviation of x is defined to be

$$\sum_{i=1}^{n} |b_i - \sum_{j=1}^{n} a_{ij}x_j|.$$

(For the example system above, the vector $x = (1, 1)^\top$ has deviation $2 + 1 + 0 + 1 = 4$.)

(a) Show that a solution to the optimization problem

$$\text{minimize} \sum_{i=1}^{m} y_i$$

$$\text{subject to}$$

$$|\sum_{j=1}^{n} a_{ij}x_j - b_i| \leq y_i \qquad (i = 1, \ldots, m)$$

will give a vector x of minimum deviation.

(b) The problem of part (a) is not an LP. (Why?) Show that it can be formulated as an LP.

(c) Suppose that we had instead defined the deviation of x as

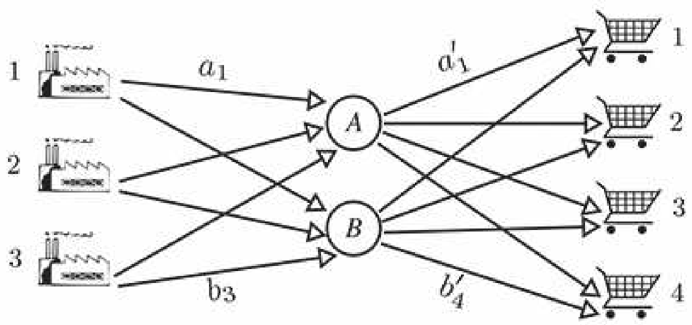
$$\max_{1 \leq i \leq m} |b_i - \sum_{j=1}^{n} a_{ij}x_j|.$$

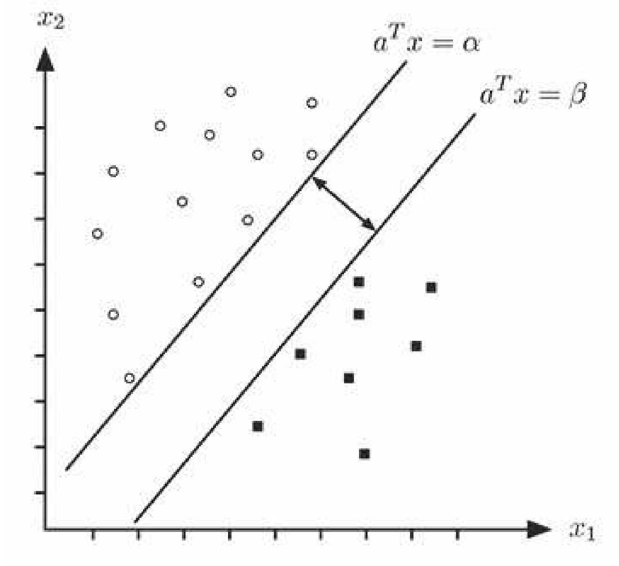
(According to this definition, in the example above $x = (1, 1)^\top$ would have deviation max(2, 1, 0,

1) = 2.) With this new definition, write the problem of finding a vector of minimum deviation as an optimization problem, and show that this problem can also be formulated as an LP.

7 Consider the following set up: we have factories 1 through m and stores 1 through n. Each factory i produces u_i units of a commodity and each store j requires ℓ_j units of that commodity. Note, each factory produces the same commodity, and each store requires the same commodity. The goal is to transfer the commodity from the factories to the stores. All the commodities going from the factories to the stores are first sent to one of two central storage hubs A and B. The cost of transporting one unit of commodity from factory i to hub A (resp. B) is given by a_i (resp. b_i). The cost of transporting one unit of commodity from hub A (resp. B) to store j is given by a'_j (resp. b'_j).[2] In the figure on top of the next page, we illustrate the case of three factories and four stores. The problem is to decide how much to send from each factory to each hub and how much to send from each hub to each store so that each store receives the amount of commodity it requires, no factory sends out more commodity than it produces, and such that the total transportation cost is minimized. Formulate this problem as an LP (we may assume that the number of units of commodity sent may be fractional).

8 We are given a matrix $A \in \mathbb{R}^{m\times n}$ and a matrix $B \in \mathbb{R}^{p\times n}$ so that the rows of A denote observations for healthy human tissue and the rows of B denote observations for unhealthy human tissue. We would like to find $a \in \mathbb{R}^n$ and $\alpha, \beta \in \mathbb{R}$ such that all rows of A are in the set $\{x \in \mathbb{R}^n : a^\top x \leq \alpha\}$, all rows of B are in the set $\{x \in \mathbb{R}^n : a^\top x \geq \beta\}$, and such that the distance between the sets $\{x \in \mathbb{R}^n : a^\top x = \alpha\}$ and $\{x \in \mathbb{R}^n : a^\top x = \beta\}$ is maximized. The following figure illustrates the situation for $n = 2$; circles correspond to rows in A, and squares to rows in B.





Formulate the problem of computing a, α, and β achieving the above-mentioned goals as an LP.

1.3 Integer programs

An *integer program* is obtained by taking a linear program and adding the condition that a nonempty subset of the variables be required to take integer values. When all variables are required to take integer values, the integer program is called a *pure integer program* otherwise it is called a *mixed integer program*. We will abbreviate the term integer program by IP, throughout this book.

Example 3 The following is a mixed IP, where variables x_1 and x_3 are required to take integer values:

$$
\begin{array}{lrcr}
\max & x_1 + x_2 + 2x_4 & & \\
\text{subject to} & & & \\
& x_1 + x_2 & \leq & 1 \\
& -x_2 - x_3 & \geq & -1 \\
& x_1 + x_3 & = & 1 \\
& \multicolumn{3}{l}{x_1, x_2, x_3 \geq 0 \text{ and } x_1, x_3 \text{ integer.}}
\end{array}
$$

In Section 1.1, we introduced the WaterTech production problem. We gave an LP formulation (1.5) and a solution to that formulation in (1.7). This solution told us to manufacture, 16+ $\frac{2}{3}$ units of product 1. Depending on the nature of product 1, it may not make sense to produce a fractional number of units of this product. Thus, we may want to add the condition that each of x_1, x_2, x_3, x_4 is an integer. The resulting program would be an IP. In this example, we could try to ignore the integer condition, and round down the solution, hoping to get a reasonably good approximation to the optimal solution.