Linear Programming: Basics

Lecturer: Peter Shor

Linear Programming deals with the problem of optimizing a linear *objective function* subject to linear equality and inequality *constraints* on the *decision variables*. Linear programming has many practical applications (in transportation, production planning, ...). It is also the building block for combinatorial optimization. One aspect of linear programming which is often forgotten is the fact that it is also a useful proof technique. In this first chapter, we describe some linear programming *formulations* for some classical problems. We also show that linear programs can be expressed in a variety of equivalent ways.

1 Formulations

1.1 The Diet Problem

In the diet model, a list of available foods is given together with the nutrient content and the cost per unit weight of each food. A certain amount of each nutrient is required per day. For example, here is the data corresponding to a civilization with just two types of grains (G1 and G2) and three types of nutrients (starch, proteins, vitamins):

	Starch	Proteins	Vitamins	Cost ($/kg)
G1	5	4	2	0.6
G2	7	2	1	0.35

Nutrient content and cost per kg of food.

The requirement per day of starch, proteins and vitamins is 8, 15 and 3 respectively. The problem is to find how much of each food to consume per day so as to get the required amount per day of each nutrient at minimal cost.

When trying to formulate a problem as a linear program, the first step is to decide which *decision variables* to use. These variables represent the unknowns in the problem. In the diet problem, a very natural choice of decision variables is:

- x_1: number of units of grain G1 to be consumed per day,
- x_2: number of units of grain G2 to be consumed per day.

The next step is to write down the *objective function*. The objective function is the function to be minimized or maximized. In this case, the objective is to minimize the total cost per day which is given by $z = 0.6x_1 + 0.35x_2$ (the value of the objective function is often denoted by z).

Finally, we need to describe the different *constraints* that need to be satisfied by x_1 and x_2. First of all, x_1 and x_2 must certainly satisfy $x_1 \geq 0$ and $x_2 \geq 0$. Only nonnegative amounts of food can be eaten! These constraints are referred to as *nonnegativity constraints*. Nonnegativity constraints appear in most linear programs. Moreover, not all possible values for x_1 and x_2 give rise to a diet with the required amounts of nutrients per day. The amount of starch in x_1 units of G1 and x_2 units of G2 is $5x_1 + 7x_2$ and this amount must be at least 8, the daily requirement of starch. Therefore, x_1 and x_2 must satisfy $5x_1 + 7x_2 \geq 8$. Similarly, the requirements on the amount of proteins and vitamins imply the constraints $4x_1 + 2x_2 \geq 15$ and $2x_1 + x_2 \geq 3$.

This diet problem can therefore be formulated by the following linear program:

$$\begin{aligned} \text{Minimize} \quad & z = 0.6x_1 + 0.35x_2 \\ \text{subject to:} \quad & \\ & 5x_1 + 7x_2 \geq 8 \\ & 4x_1 + 2x_2 \geq 15 \\ & 2x_1 + x_2 \geq 3 \\ & x_1 \geq 0, x_2 \geq 0. \end{aligned}$$

Some more terminology. A *solution* $x = (x_1, x_2)$ is said to be *feasible* with respect to the above linear program if it satisfies all the above constraints. The set of feasible solutions is called the *feasible space* or *feasible region*. A feasible solution is *optimal* if its objective function value is equal to the smallest value z can take over the feasible region.

1.2 The Transportation Problem

Suppose a company manufacturing widgets has two factories located at cities F1 and F2 and three retail centers located at C1, C2 and C3. The monthly demand at the retail centers are (in thousands of widgets) 8, 5 and 2 respectively while the monthly supply at the factories are 6 and 9 respectively. Notice that the total supply equals the total demand. We are also given the cost of transportation of 1 widget between any factory and any retail center.

	C1	C2	C3
F1	5	5	3
F2	6	4	1

Cost of transportation (in 0.01\$/widget).

In the *transportation problem*, the goal is to determine the quantity to be transported from each factory to each retail center so as to meet the demand at minimum total shipping cost.

In order to formulate this problem as a linear program, we first choose the decision variables. Let x_{ij} ($i = 1, 2$ and $j = 1, 2, 3$) be the number of widgets (in thousands) transported from factory Fi to city Cj. Given these x_{ij}'s, we can express the total shipping cost, i.e. the objective function to be minimized, by

$$5x_{11} + 5x_{12} + 3x_{13} + 6x_{21} + 4x_{22} + x_{23}.$$

We now need to write down the constraints. First, we have the nonnegativity constraints saying that $x_{ij} \geq 0$ for $i = 1, 2$ and $j = 1, 2, 3$. Moreover, we have that the demand at each retail center must be met. This gives rise to the following constraints:

$$x_{11} + x_{21} = 8,$$

$$x_{12} + x_{22} = 5,$$

$$x_{13} + x_{23} = 2.$$

Finally, each factory cannot ship more than its supply, resulting in the following constraints:

$$x_{11} + x_{12} + x_{13} \leq 6,$$

$$x_{21} + x_{22} + x_{23} \leq 9.$$

These inequalities can be replaced by equalities since the total supply is equal to the total demand. A linear programming formulation of this transportation problem is therefore given by:

$$\begin{array}{rl} \text{Minimize} & 5x_{11} + 5x_{12} + 3x_{13} + 6x_{21} + 4x_{22} + x_{23} \\ \text{subject to:} & \\ & x_{11} + x_{21} = 8 \\ & x_{12} + x_{22} = 5 \\ & x_{13} + x_{23} = 2 \\ & x_{11} + x_{12} + x_{13} = 6 \\ & x_{21} + x_{22} + x_{23} = 9 \\ & x_{11} \geq 0, x_{21} \geq 0, x_{31} \geq 0, \\ & x_{12} \geq 0, x_{22} \geq 0, x_{32} \geq 0. \end{array}$$

Among these 5 equality constraints, one is *redundant*, i.e. it is implied by the other constraints or, equivalently, it can be removed without modifying the feasible space. For example, by adding the first 3 equalities and substracting the fourth equality we obtain the last equality. Similarly, by adding the last 2 equalities and substracting the first two equalities we obtain the third one.

2 Representations of Linear Programs

A linear program can take many different forms. First, we have a minimization or a maximization problem depending on whether the objective function is to be minimized or maximized. The constraints can either be inequalities ($\leq$ or $\geq$) or equalities. Some variables might be unrestricted in sign (i.e. they can take positive or negative values; this is denoted by $\gtreqless 0$) while others might be restricted to be nonnegative. A general linear program in the decision variables $x_1, \ldots, x_n$ is therefore of the following form:

$$\begin{array}{rl} \text{Maximize or Minimize} & z = c_0 + c_1 x_1 + \ldots + c_n x_n \\ \text{subject to:} & \\ & a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n \begin{array}{c} \leq \\ \geq \\ = \end{array} b_i \qquad i = 1, \ldots, m \\ & x_j \left\{ \begin{array}{l} \geq 0 \\ \gtreqless 0 \end{array} \right. \qquad j = 1, \ldots, n. \end{array}$$

The problem data in this linear program consists of c_j ($j = 0, \ldots, n$), b_i ($i = 1, \ldots, m$) and a_{ij} ($i = 1, \ldots, m$, $j = 1, \ldots, n$). c_j is referred to as the objective function coefficient of x_j or, more simply, the *cost coefficient* of x_j. b_i is known as the *right-hand-side* (RHS) of equation i. Notice that the constant term c_0 can be omitted without affecting the set of optimal solutions.

A linear program is said to be in *standard form* if

- it is a maximization program,
- there are only equalities (no inequalities) and
- all variables are restricted to be nonnegative.

In matrix form, a linear program in standard form can be written as:

$$\begin{aligned} \text{Max} \quad & z = c^T x \\ \text{subject to:} & \\ & Ax = b \\ & x \geq 0. \end{aligned}$$

where

$$c = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}, b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}, x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

are column vectors, c^T denote the transpose of the vector c, and $A = [a_{ij}]$ is the $m \times n$ matrix whose $i, j-$element is a_{ij}.

Any linear program can in fact be transformed into an equivalent linear program in standard form. Indeed,

- If the objective function is to minimize $z = c_1 x_1 + \ldots + c_n x_n$ then we can simply maximize $z' = -z = -c_1 x_1 - \ldots - c_n x_n$.
- If we have an inequality constraint $a_{i1} x_1 + \ldots + a_{in} x_n \leq b_i$ then we can transform it into an equality constraint by adding a *slack* variable, say s, restricted to be nonnegative: $a_{i1} x_1 + \ldots + a_{in} x_n + s = b_i$ and $s \geq 0$.
- Similarly, if we have an inequality constraint $a_{i1} x_1 + \ldots + a_{in} x_n \geq b_i$ then we can transform it into an equality constraint by adding a *surplus* variable, say s, restricted to be nonnegative: $a_{i1} x_1 + \ldots + a_{in} x_n - s = b_i$ and $s \geq 0$.
- If x_j is unrestricted in sign then we can introduce two new decision variables x_j^+ and x_j^- restricted to be nonnegative and replace every occurrence of x_j by $x_j^+ - x_j^-$.

For example, the linear program

$$\begin{aligned} \text{Minimize} \quad & z = 2x_1 - x_2 \\ \text{subject to:} & \\ & x_1 + x_2 \geq 2 \\ & 3x_1 + 2x_2 \leq 4 \\ & x_1 + 2x_2 = 3 \\ & x_1 \gtrless 0, x_2 \geq 0. \end{aligned}$$

is equivalent to the linear program

$$\begin{aligned} \text{Maximize} \quad & z' = -2x_1^+ + 2x_1^- + x_2 \\ \text{subject to:} \quad & \\ & x_1^+ - x_1^- + x_2 - x_3 = 2 \\ & 3x_1^+ - 3x_1^- + 2x_2 + x_4 = 4 \\ & x_1^+ - x_1^- + 2x_2 = 3 \\ & x_1^+ \geq 0, x_1^- \geq 0, x_2 \geq 0, x_3 \geq 0, x_4 \geq 0. \end{aligned}$$

with decision variables $x_1^+, x_1^-, x_2, x_3, x_4$. Notice that we have introduced different slack or surplus variables into different constraints.

In some cases, another form of linear program is used. A linear program is in *canonical form* if it is of the form:

$$\begin{aligned} \text{Max} \quad & z = c^T x \\ \text{subject to:} \quad & \\ & Ax \leq b \\ & x \geq 0. \end{aligned}$$

A linear program in canonical form can be replaced by a linear program in standard form by just replacing $Ax \leq b$ by $Ax + Is = b$, $s \geq 0$ where s is a vector of slack variables and I is the $m \times m$ identity matrix. Similarly, a linear program in standard form can be replaced by a linear program in canonical form by replacing $Ax = b$ by $A'x \leq b'$ where $A' = \begin{bmatrix} A \\ -A \end{bmatrix}$ and $b' = \begin{pmatrix} b \\ -b \end{pmatrix}$.

3 Basic Feasible Solutions

Although there is a continuum of feasible solutions in a linear program (unless the LP is infeasible or the feasible region contains only a single point), linear programs can also be shown to behave in a discrete fashion, in the sense one can focus on a discrete set of feasible solutions. These special feasible solutions are called *basic feasible solutions* and it is guaranteed that there is always an optimum solution that will be a bfs (if the LP is feasible and bounded). This is discussed in this section.

Consider an LP in standard form:

$$\begin{aligned} \text{Max} \quad & z = c^T x \\ \text{subject to:} \quad & \\ & Ax = b \\ & x \geq 0. \end{aligned}$$

We have seen how to transform any LP in this form.

First, we focus on the system of linear equalities $Ax = b$. Clearly if $\{x : Ax = b\} = \emptyset$ then our feasible region is empty and our LP is infeasible. Remark, however, that the converse is not true; namely, the LP might be infeasible although $Ax = b$ has a solution (with some negative components). So, we can assume that $Ax = b$ has a solution, and we can remove redundant inequalities. Therefore, we can assume that $rank(A) = m$ where A is $m \times n$; this says that A is of full row-rank. Since the column-rank and the row-rank of a matrix coincide, we know that we can find subsets of m columns which are linearly independent. We'll say that $B \subseteq \{1, \cdots, n\}$ with $|B| = m$ forms a *basis* if the corresponding $m \times m$ matrix A_B consisting of all columns of A indexed by indices in B is invertible (i.e. of rank m). There are at most $\binom{n}{m}$ bases. Once we consider a basis B, it is convenient to let the set of (non-basic) indices be denoted by $N = \{1, \cdots, n\} \setminus B$. In block format, we can write

$$A = [A_B | A_N],$$

and if we partition the (decision) variables x also into $x = (x_B, x_N)$, we get that the system $Ax = b$ can be written as $A_B x_B + A_N x_N = b$.

Remember that if we take any invertible $m \times m$ matrix P then the set of solutions of $Ax = b$ or of $PAx = Pb$ are the same (since we can premultiply the second system by P^{-1} to get the first system). So, if we have a basis B and premultiply $Ax = b$ by A_B^{-1} we get an equivalent system:

$$x_B + A_B^{-1} A_N x_N = A_B^{-1} b.$$

This says that if we set the values of x_N, we can deduce a unique set of values for x_B satisfying our system of equalities, namely

$$x_B = A_B^{-1} b - A_B^{-1} A_N x_N.$$

In particular, there is a particularly simple solution if one sets $x_N = 0$; in this case we have $x = (x_B, x_N) = (A_B^{-1} b, 0)$.

But so far, we have not considered the nonnegativity constraints $x \geq 0$. We say that x is a *basic feasible solution* (bfs) if there exists a basis B such that $x = (x_B, x_N) = (A_B^{-1} b, 0) \geq 0$. The importance of basic feasible solutions is given by the following theorem.

Theorem 1. *Consider any linear program in standard form with $rank(A) = m$ where A is $m \times n$. If the LP is feasible and bounded then there exists an optimum solution that is a basic feasible solution.*

We won't prove this although this follows from linear algebra. Here is a brief sketch of a proof (with missing details and explanations). Let x^* be any optimum solution to the linear program. Let $J = \{j : x_j^* > 0\}$. We consider two cases, either $rank(A_J) = |J| \leq m$ or $rank(A_J) < |J|$. In the first case, one argues that x^* is already a basic feasible solution corresponding to a basis B, where B is obtained from J by adding $m - |J|$ columns so that $rank(A_B) = |B|$ (which is always possible since $rank(A) = m$). In the second case, we take a solution $y \neq 0$ to $A_J y_J = 0$, $y_j = 0$ for $j \notin J$ and $c^T y = 0$ (by rank considerations, one can argue that such y exists). We then consider $x^* + \lambda y$ and choose λ so that this vector is still nonnegative but at least one index in J has now zero value, and we repeat this argument until we are in case 1.

MIT OpenCourseWare
https://ocw.mit.edu

18.200 Principles of Discrete Applied Mathematics
Spring 2024

For information about citing these materials or our Terms of Use, visit: https://ocw.mit.edu/terms.