18.310 lecture notes September 2, 2013

Linear programming

Lecturer: Michel Goemans

1 Basics

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.

1.1 Formulations

1.1.1 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:

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

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.

1.1.2 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.

1.2 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 $\gtrless 0$) 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:

$$\begin{array}{rl} \text{Maximize or Minimize} & z = c_0 + c_1 x_1 + \ldots + c_n x_n \\ \text{subject to:} & \end{array}$$

$$a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n \begin{array}{c} \leq \\ \geq \\ = \end{array} b_i \qquad\qquad i = 1, \ldots, m$$

$$x_j \left\{ \begin{array}{l} \geq 0 \\ \gtrless 0 \end{array} \right. \qquad\qquad\qquad\qquad j = 1, \ldots, n.$$

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

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

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 $b' = \begin{pmatrix} b \\ -b \end{pmatrix}$.

2 The Simplex Method

In 1947, George B. Dantzig developed a technique to solve linear programs — this technique is referred to as the *simplex method.*

2.1 Brief Review of Some Linear Algebra

Two systems of equations $Ax = b$ and $\bar{A}x = \bar{b}$ are said to be equivalent if $\{x : Ax = b\} = \{x : \bar{A}x = \bar{b}\}$. Let E_i denote equation i of the system $Ax = b$, i.e. $a_{i1}x_1 + \ldots + a_{in}x_n = b_i$. Given a system $Ax = b$, an *elementary row operation* consists in replacing E_i either by αE_i where α is a *nonzero* scalar or by $E_i + \beta E_k$ for some $k \neq i$. Clearly, if $\bar{A}x = \bar{b}$ is obtained from $Ax = b$ by an elementary row operation then the two systems are equivalent. (Exercise: prove this.) Notice also that an elementary row operation is reversible.

Let a_{rs} be a nonzero element of A. A pivot on a_{rs} consists of performing the following sequence of elementary row operations:

- replacing E_r by $\bar{E}_r = \frac{1}{a_{rs}} E_r$,
- for $i = 1, \ldots, m$, $i \neq r$, replacing E_i by $\bar{E}_i = E_i - a_{is}\bar{E}_r = E_i - \frac{a_{is}}{a_{rs}} E_r$.

After pivoting on a_{rs}, all coefficients in column s are equal to 0 except the one in row r which is now equal to 1. Since a pivot consists of elementary row operations, the resulting system $\bar{A}x = \bar{b}$ is equivalent to the original system.

Elementary row operations and pivots can also be defined in terms of matrices. Let P be an $m \times m$ invertible (i.e. P^{-1} exists[1]) matrix. Then $\{x : Ax = b\} = \{x : PAx = Pb\}$. The two types of elementary row operations correspond to the matrices (the coefficients not represented are equal to 0):

$$P = \left(\begin{array}{ccccccc} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \alpha & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{array}\right)\begin{array}{l} \\ \\ \\ \leftarrow i \\ \\ \\ \\ \end{array} \text{ and } P = \left(\begin{array}{ccccccc} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & \beta & & \\ & & & \ddots & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{array}\right)\begin{array}{l} \\ \\ \leftarrow i \\ \\ \leftarrow k \\ \\ \\ \end{array}.$$

Pivoting on a_{rs} corresponds to premultiplying $Ax = b$ by

$$P = \left(\begin{array}{ccccccc} 1 & & & -a_{1s}/a_{rs} & & & \\ & \ddots & & & & & \\ & & 1 & -a_{r-1,s}/a_{rs} & & & \\ & & & 1/a_{rs} & & & \\ & & & -a_{r+1,s}/a_{rs} & 1 & & \\ & & & & & \ddots & \\ & & & -a_{ms}/a_{rs} & & & 1 \end{array}\right)\begin{array}{l} \\ \\ \\ \leftarrow r. \\ \\ \\ \\ \end{array}$$

2.2 The Simplex Method on an Example

For simplicity, we shall assume that we have a linear program of (what seems to be) a rather special form (we shall see later on how to obtain such a form):

- the linear program is in standard form,
- $b \geq 0$,
- there exists a collection B of m variables called a *basis* such that
  - the submatrix A_B of A consisting of the columns of A corresponding to the variables in B is the $m \times m$ identity matrix and
  - the cost coefficients corresponding to the variables in B are all equal to 0.

For example, the following linear program has this required form:

[1] This is equivalent to saying that $\det P \neq 0$ or also that the system $Px = 0$ has $x = 0$ as unique solution

$$
\begin{array}{llllllllll}
\text{Max } z = 10 + & 20\, x_1 & + & 16\, x_2 & + & 12\, x_3 & & & & \\
\quad\text{subject to} & & & & & & & & & \\
 & x_1 & & & & & + x_4 & & & = \; 4 \\
 & 2\, x_1 & + & x_2 & + & x_3 & & + x_5 & & = \; 10 \\
 & 2\, x_1 & + & 2x_2 & + & x_3 & & & + x_6 & = \; 16 \\
 & \multicolumn{9}{l}{x_1, x_2, x_3, x_4, x_5, x_6 \geq 0.}
\end{array}
$$

In this example, $B = \{x_4, x_5, x_6\}$. The variables in B are called *basic* variables while the other variables are called *nonbasic*. The set of nonbasic variables is denoted by N. In the example, $N = \{x_1, x_2, x_3\}$.

The advantage of having $A_B = I$ is that we can quickly infer the values of the basic variables given the values of the nonbasic variables. For example, if we let $x_1 = 1, x_2 = 2, x_3 = 3$, we obtain

$$x_4 = 4 - x_1 = 3,$$

$$x_5 = 10 - 2x_1 - x_2 - x_3 = 3,$$

$$x_6 = 16 - 2x_1 - 2x_2 - x_3 = 7.$$

Also, we don't need to know the values of the basic variables to evaluate the cost of the solution. In this case, we have $z = 10 + 20x_1 + 16x_2 + 12x_3 = 98$. Notice that there is no guarantee that the so-constructed solution be feasible. For example, if we set $x_1 = 5, x_2 = 2, x_3 = 1$, we have that $x_4 = 4 - x_1 = -1$ does not satisfy the nonnegativity constraint $x_4 \geq 0$.

There is an assignment of values to the nonbasic variables that needs special consideration. By just letting all nonbasic variables to be equal to 0, we see that the values of the basic variables are just given by the right-hand-sides of the constraints and the cost of the resulting solution is just the constant term in the objective function. In our example, letting $x_1 = x_2 = x_3 = 0$, we obtain $x_4 = 4, x_5 = 10, x_6 = 16$ and $z = 10$. Such a solution is called a *basic feasible solution* or *bfs*. The feasibility of this solution comes from the fact that $b \geq 0$. Later, we shall see that, when solving a linear program, we can restrict our attention to basic feasible solutions. The simplex method is an iterative method that generates a sequence of basic feasible solutions (corresponding to different bases) and eventually stops when it has found an optimal basic feasible solution.

Instead of always writing explicitely these linear programs, we adopt what is known as the *tableau format*. First, in order to have the objective function play a similar role as the other constraints, we consider z to be a variable and the objective function as a constraint. Putting all variables on the same side of the equality sign, we obtain:

$$-z + 20x_1 + 16x_2 + 12x_3 = -10.$$

We also get rid of the variable names in the constraints to obtain the tableau format:

$-z$	x_1	x_2	x_3	x_4	x_5	x_6	
1	20	16	12				-10
	1	0	0	1			4
	2	1	1		1		10
	2	2	1			1	16

Our bfs is currently $x_1 = 0, x_2 = 0, x_3 = 0, x_4 = 4, x_5 = 10, x_6 = 16$ and $z = 10$. Since the cost coefficient c_1 of x_1 is positive (namely, it is equal to 20), we notice that we can increase z by increasing x_1 and keeping x_2 and x_3 at the value 0. But in order to maintain feasibility, we must have that $x_4 = 4 - x_1 \geq 0$, $x_5 = 10 - 2x_1 \geq 0$, $x_6 = 16 - 2x_1 \geq 0$. This implies that $x_1 \leq 4$. Letting $x_1 = 4, x_2 = 0, x_3 = 0$, we obtain $x_4 = 0, x_5 = 2, x_6 = 8$ and $z = 90$. This solution is also a bfs and corresponds to the basis $B = \{x_1, x_5, x_6\}$. We say that x_1 has entered the basis and, as a result, x_4 has left the basis. We would like to emphasize that there is a *unique* basic solution associated with any basis. This (not necessarily feasible) solution is obtained by setting the nonbasic variables to zero and deducing the values of the basic variables from the m constraints.

Now we would like that our tableau reflects this change by showing the dependence of the new basic variables as a function of the nonbasic variables. This can be accomplished by *pivoting* on the element a_{11}. Why a_{11}? Well, we need to pivot on an element of column 1 because x_1 is entering the basis. Moreover, the choice of the row to pivot on is dictated by the variable which leaves the basis. In this case, x_4 is leaving the basis and the only 1 in column 4 is in row 1. After pivoting on a_{11}, we obtain the following tableau:

$-z$	x_1	x_2	x_3	x_4	x_5	x_6	
1		16	12	-20			-90
	1	0	0	1			4
		1	1	-2	1		2
		2	1	-2		1	8

Notice that while pivoting we also modified the objective function row as if it was just like another constraint. We have now a linear program which is equivalent to the original one from which we can easily extract a (basic) feasible solution of value 90. Still z can be improved by increasing x_s for $s = 2$ or 3 since these variables have a positive cost coefficient[2] $\bar{c}_s$. Let us choose the one with the greatest $\bar{c}_s$; in our case x_2 will enter the basis. The maximum value that x_2 can take while x_3 and x_4 remain at the value 0 is dictated by the constraints $x_1 = 4 \geq 0$, $x_5 = 2 - x_2 \geq 0$ and $x_6 = 8 - 2x_2 \geq 0$. The tightest of these inequalities being $x_5 = 2 - x_2 \geq 0$, we have that x_5 will leave the basis. Therefore, pivoting on $\bar{a}_{22}$, we obtain the tableau:

$-z$	x_1	x_2	x_3	x_4	x_5	x_6	
1			-4	12	-16		-122
	1		0	1	0		4
		1	1	-2	1		2
			-1	2	-2	1	4

The current basis is $B = \{x_1, x_2, x_6\}$ and its value is 122. Since $12 > 0$, we can improve the current basic feasible solution by having x_4 enter the basis. Instead of writing explicitely the constraints on x_4 to compute the level at which x_4 can enter the basis, we perform the *min ratio test.* If x_s is the variable that is entering the basis, we compute

$$\min_{i:\bar{a}_{is}>0} \{\bar{b}_i / \bar{a}_{is}\}.$$

The argument of the minimum gives the variable that is exiting the basis. In our example, we obtain $2 = \min\{4/1, 4/2\}$ and therefore variable x_6 which is the basic variable corresponding to row 3 leaves the basis. Moreover, in order to get the updated tableau, we need to pivot on $\bar{a}_{34}$. Doing so, we obtain:

[2] By simplicity, we always denote the data corresponding to the current tableau by $\bar{c}$, $\bar{A}$, and $\bar{b}$.

$-z$	x_1	x_2	x_3	x_4	x_5	x_6	
1			2		-4	-6	-146
	1		1/2		1	-1/2	2
		1	0		-1	1	6
			-1/2	1	-1	1/2	2

Our current basic feasible solution is $x_1 = 2, x_2 = 6, x_3 = 0, x_4 = 2, x_5 = 0, x_6 = 0$ with value $z = 146$. By the way, why is this solution feasible? In other words, how do we know that the right-hand-sides (RHS) of the constraints are guaranteed to be nonnegative? Well, this follows from the min ratio test and the pivot operation. Indeed, when pivoting on $\bar{a}_{rs}$, we know that

- $\bar{a}_{rs} > 0$,
- $\frac{\bar{b}_r}{\bar{a}_{rs}} \leq \frac{\bar{b}_i}{\bar{a}_{is}}$ if $\bar{a}_{is} > 0$.

After pivoting the new RHS satisfy

- $\bar{b}_r = \frac{\bar{b}_r}{\bar{a}_{rs}} \geq 0$,
- $\bar{b}_i = \bar{b}_i - \frac{\bar{a}_{is}}{\bar{a}_{rs}} \geq \bar{b}_i \geq 0$ if $\bar{a}_{is} \leq 0$ and
- $\bar{b}_i = \bar{b}_i - \frac{\bar{a}_{is}}{\bar{a}_{rs}} = \bar{a}_{is} \left(\frac{\bar{b}_i}{\bar{a}_{is}} - \frac{\bar{b}_r}{\bar{a}_{rs}} \right) \geq 0$ if $\bar{a}_{is} > 0$.

We can also justify why the solution keeps improving. Indeed, when we pivot on $\bar{a}_{rs} > 0$, the constant term $\bar{c}_0$ in the objective function becomes $\bar{c}_0 + \bar{b}_r * \bar{c}_s / \bar{a}_{rs}$. If $\bar{b}_r > 0$, we have a strict improvement in the objective function value since by our choice of entering variable $\bar{c}_s > 0$. We shall deal with the case $\bar{b}_r = 0$ later on.

The bfs corresponding to $B = \{1, 2, 4\}$ is not optimal since there is still a positive cost coefficient. We see that x_3 can enter the basis and, since there is just one positive element in row 3, we have that x_1 leaves the basis. We thus pivot on $\bar{a}_{13}$ and obtain:

$-z$	x_1	x_2	x_3	x_4	x_5	x_6	
1	-4				-8	-4	-154
	2		1		2	-1	4
	0	1			-1	1	6
	1			1	0	0	4

The current basis is $\{x_3, x_2, x_4\}$ and the associated bfs is $x_1 = 0, x_2 = 6, x_3 = 4, x_4 = 4, x_5 = 0, x_6 = 0$ with value $z = 154$. This bfs is *optimal* since the objective function reads $z = 154 - 4x_1 - 8x_5 - 4x_6$ and therefore cannot be more than 154 due to the nonnegativity constraints.

Through a sequence of pivots, the simplex method thus goes from one linear program to another equivalent linear program which is trivial to solve. Remember the crucial observation that a pivot operation does not alter the feasible region.

In the above example, we have not encountered several situations that may typically occur. First, in the min ratio test, several terms might produce the minimum. In that case, we can arbitrarily select one of them. For example, suppose the current tableau is:

$-z$	x_1	x_2	x_3	x_4	x_5	x_6	
1		16	12	-20			-90
	1	0	0	1			4
		1	1	-2	1		**2**
		2	1	-2		1	**4**

and that x_2 is entering the basis. The min ratio test gives $2 = \min\{2/1, 4/2\}$ and, thus, either x_5 or x_6 can leave the basis. If we decide to have x_5 leave the basis, we pivot on $\bar{a}_{22}$; otherwise, we pivot on $\bar{a}_{32}$. Notice that, in any case, the pivot operation creates a zero coefficient among the RHS. For example, pivoting on $\bar{a}_{22}$, we obtain:

$-z$	x_1	x_2	x_3	x_4	x_5	x_6	
1			-4	12	-16		-122
	1		0	1	0		4
		1	1	-2	1		2
			-1	2	-2	1	**0**

A bfs with $\bar{b}_i = 0$ for some i is called *degenerate*. A linear program is *nondegenerate* if *no* bfs is degenerate. Pivoting now on $\bar{a}_{34}$ we obtain:

$-z$	x_1	x_2	x_3	x_4	x_5	x_6	
1			2		-4	-6	-122
	1		1/2		1	-1/2	4
		1	0		-1	1	2
			-1/2	1	-1	1/2	0

This pivot is degenerate. A pivot on $\bar{a}_{rs}$ is called *degenerate* if $\bar{b}_r = 0$. Notice that a degenerate pivot alters neither the $\bar{b}_i$'s nor $\bar{c}_0$. In the example, the bfs is $(4, 2, 0, 0, 0, 0)$ in both tableaus. We thus observe that several bases can correspond to the same basic feasible solution.

Another situation that may occur is when x_s is entering the basis, but $\bar{a}_{is} \leq 0$ for $i = 1, \ldots, m$. In this case, there is no term in the min ratio test. This means that, while keeping the other nonbasic variables at their zero level, x_s can take an arbitrarily large value without violating feasibility. Since $\bar{c}_s > 0$, this implies that z can be made arbitrarily large. In this case, the linear program is said to be *unbounded* or *unbounded from above* if we want to emphasize the fact that we are dealing with a maximization problem. For example, consider the following tableau:

$-z$	x_1	x_2	x_3	x_4	x_5	x_6	
1		16	12	20			-90
	1	0	0	-1			4
		1	1	0	1		2
		2	1	-2		1	8

If x_4 enters the basis, we have that $x_1 = 4 + x_4$, $x_5 = 2$ and $x_6 = 8 + 2x_4$ and, as a result, for *any* nonnegative value of x_4, the solution $(4 + x_4, 0, 0, x_4, 2, 8 + 2x_4)$ is feasible and its objective function value is $90 + 20x_4$. There is thus no *finite optimum.*

2.3 Detailed Description of Phase II

In this section, we summarize the different steps of the simplex method we have described in the previous section. In fact, what we have described so far constitutes Phase II of the simplex method. Phase I deals with the problem of putting the linear program in the required form. This will be described in a later section.

Phase II of the simplex method

1. Suppose the initial or current tableau is

$$\begin{array}{c|ccccc|c}
-z & x_1 & \dots & x_s & \dots & x_n & \\
1 & \bar{c}_1 & \dots & \bar{c}_s & \dots & \bar{c}_n & -\bar{c}_0 \\
\hline
 & \bar{a}_{11} & \dots & \bar{a}_{1s} & \dots & \bar{a}_{1n} & \bar{b}_1 \geq 0 \\
 & \vdots & & \vdots & & \vdots & \vdots \\
 & \bar{a}_{r1} & \dots & \bar{a}_{rs} & \dots & \bar{a}_{rn} & \bar{b}_r \geq 0 \\
 & \vdots & & \vdots & & \vdots & \vdots \\
 & \bar{a}_{m1} & \dots & \bar{a}_{ms} & \dots & \bar{a}_{mn} & \bar{b}_m \geq 0
\end{array}$$

   and the variables can be partitioned into $B = \{x_{j_1}, \dots, x_{j_m}\}$ and N with

   - $\bar{c}_{j_i} = 0$ for $i = 1, \dots, m$ and
   - $$\bar{a}_{kj_i} = \begin{cases} 0 & k \neq i \\ 1 & k = i. \end{cases}$$

   The current basic feasible solution is given by $x_{j_i} = \bar{b}_i$ for $i = 1, \dots, m$ and $x_j = 0$ otherwise. The objective function value of this solution is $\bar{c}_0$.

2. If $\bar{c}_j \leq 0$ for all $j = 1, \dots, n$ then the current basic feasible solution is optimal. STOP.

3. Find a column s for which $\bar{c}_s > 0$. x_s is the variable entering the basis.

4. Check for unboundedness. If $\bar{a}_{is} \leq 0$ for $i = 1, \dots, m$ then the linear program is unbounded. STOP.

5. Min ratio test. Find row r such that

$$\frac{\bar{b}_r}{\bar{a}_{rs}} = \min_{i:\bar{a}_{is}>0} \frac{\bar{b}_i}{\bar{a}_{is}}.$$

6. Pivot on $\bar{a}_{rs}$. I.e. replace the current tableau by:

$$
\begin{array}{c|cccccc|c}
-z & \dots & x_s & \dots & x_j & \dots & & \\
1 & \dots & 0 & \dots & \bar{c}_j - \dfrac{\bar{a}_{rj}\bar{c}_s}{\bar{a}_{rs}} & \dots & & -\bar{c}_0 - \dfrac{\bar{b}_r\bar{c}_s}{\bar{a}_{rs}} \\
\hline
 & & \vdots & & \vdots & & & \vdots \\
\text{row } r & \dots & 1 & \dots & \dfrac{\bar{a}_{rj}}{\bar{a}_{rs}} & \dots & & \dfrac{\bar{b}_r}{\bar{a}_{rs}} \\
 & & \vdots & & \vdots & & & \vdots \\
\text{row } i & \dots & 0 & \dots & \bar{a}_{ij} - \dfrac{\bar{a}_{rj}\bar{a}_{is}}{\bar{a}_{rs}} & \dots & & \bar{b}_i - \dfrac{\bar{b}_r\bar{a}_{is}}{\bar{a}_{rs}} \\
 & & \vdots & & \vdots & & & \vdots
\end{array}
$$

Replace x_{j_r} by x_s in B.

7. Go to step 2.

2.4 Convergence of the Simplex Method

As we have seen, the simplex method is an iterative method that generates a sequence of basic feasible solutions. But, do we have any guarantee that this process eventually terminates? The answer is yes if the linear program is nondegenerate.

Theorem 2.1. The simplex method solves a nondegenerate linear program in finitely many iterations.

Proof. For nondegenerate linear programs, we have a strict improvement (namely of value $\frac{\bar{b}_r\bar{c}_s}{\bar{a}_{rs}} > 0$) in the objective function value at each iteration. This means that, in the sequence of bfs produced by the simplex method, each bfs can appear at most once. Therefore, for nondegenerate linear programs, the number of iterations is certainly upper bounded by the number of bfs. This latter number is finite (for example, it is upper bounded by $\binom{n}{m}$) since any bfs corresponds to m variables being basic[3]. □

However, when the linear program is degenerate, we might have degenerate pivots which give no strict improvement in the objective function. As a result, a subsequence of bases might repeat implying the nontermination of the method. This phenomenon is called *cycling*.

2.4.1 An Example of Cycling

The following is an example that will cycle if unfortunate choices of entering and leaving variables are made (the pivot element is within a box).

[3] Not all choices of basic variables give rise to *feasible* solutions.

$-z$	x_1	x_2	x_3	x_4	x_5	x_6	
1	4	1.92	-16	-0.96			0
	-12.5	-2	12.5	1	1		0
	$\boxed{1}$	0.24	-2	-0.24		1	0

$-z$	x_1	x_2	x_3	x_4	x_5	x_6	
1		0.96	-8	0		-4	0
		$\boxed{1}$	-12.5	-2	1	12.5	0
	1	0.24	-2	-0.24		1	0

$-z$	x_1	x_2	x_3	x_4	x_5	x_6	
1			4	1.92	-0.96	-16	0
		1	-12.5	-2	1	12.5	0
	1		$\boxed{1}$	0.24	-0.24	-2	0

$-z$	x_1	x_2	x_3	x_4	x_5	x_6	
1	-4			0.96	0	-8	0
	12.5	1		$\boxed{1}$	-2	-12.5	0
	1		1	0.24	-0.24	-2	0

$-z$	x_1	x_2	x_3	x_4	x_5	x_6	
1	-16	-0.96			1.92	4	0
	12.5	1		1	-2	-12.5	0
	-2	-0.24	1		0.24	$\boxed{1}$	0

$-z$	x_1	x_2	x_3	x_4	x_5	x_6	
1	-8	0	-4		0.96		0
	-12.5	-2	12.5	1	$\boxed{1}$		0
	-2	-0.24	1		0.24	1	0

$-z$	x_1	x_2	x_3	x_4	x_5	x_6	
1	4	1.92	-16	-0.96			0
	-12.5	-2	12.5	1	1		0
	$\boxed{1}$	0.24	-2	-0.24		1	0

2.4.2 Bland's Anticycling Rule

The simplex method, as described in the previous section, is ambiguous. First, if we have several variables with a positive $\bar{c}_s$ (cfr. Step 3) we have not specified which will enter the basis. Moreover, there might be several variables attaining the minimum in the minimum ratio test (Step 5). If so, we need to specify which of these variables will leave the basis. A *pivoting rule* consists of an *entering variable rule* and a *leaving variable rule* that unambiguously decide what will be the entering and leaving variables.

The most classical entering variable rule is:

Largest coefficient entering variable rule: Select the variable x_s with the largest $\bar{c}_s > 0$. In case of ties, select the one with the smallest subscript s.

The corresponding leaving variable rule is:

Largest coefficient leaving variable rule: Among all rows attaining the minimum in the minimum ratio test, select the one with the largest pivot $\bar{a}_{rs}$. In case of ties, select the one with the smallest subscript r.

The example of subsection 2.4.1 shows that the use of the largest coefficient entering and leaving variable rules does not prevent cycling. There are two rules that avoid cycling: the lexicographic rule and Bland's rule (after R. Bland who discovered it in 1976). We'll just describe the latter one, which is conceptually the simplest.

Bland's anticycling pivoting rule: Among all variables x_s with positive $\bar{c}_s$, select the one with the smallest subscript s. Among the eligible (according to the minimum ratio test) leaving variables x_l, select the one with the smallest subscript l.

Theorem 2.2. The simplex method with Bland's anticycling pivoting rule terminates after a finite number of iterations.

Proof. The proof is by contradiction. If the method does not stop after a finite number of iterations then there is a cycle of tableaus that repeats. If we delete from the tableau that initiates this cycle the rows and columns not containing pivots during the cycle, the resulting tableau has a cycle with the same pivots. For this tableau, all right-hand-sides are zero throughout the cycle since all pivots are degenerate.

Let t be the largest subscript of the variables remaining. Consider the tableau T_1 in the cycle with x_t leaving. Let $B = \{x_{j_1}, \ldots, x_{j_m}\}$ be the corresponding basis (say $j_r = t$), x_s be the associated entering variable and, a^1_{ij} and c^1_j the constraint and cost coefficients. On the other hand, consider the tableau T_2 with x_t entering and denotes by a^2_{ij} and c^2_j the corresponding constraint and cost coefficients.

Let x be the (infeasible) solution obtained by letting the nonbasic variables in T_1 be zero except for $x_s = -1$. Since all RHS are zero, we deduce that $x_{j_i} = a_{is}$ for $i = 1, \ldots, m$. Since T_2 is obtained from T_1 by elementary row operations, x must have the same objective function value in T_1 and T_2. This means that

$$c^1_0 - c^1_s = c^2_0 - c^2_s + \sum_{i=1}^{m} a^1_{is} c^2_{j_i}.$$

Since we have no improvement in objective function in the cycle, we have $c^1_0 = c^2_0$. Moreover, $c^1_s > 0$ and, by Bland's rule, $c^2_s \leq 0$ since otherwise x_t would not be the entering variable in T_2. Hence,

$$\sum_{i=1}^{m} a^1_{is} c^2_{j_i} < 0$$

implying that there exists k with $a^1_{ks} c^2_{j_k} < 0$. Notice that $k \neq r$, i.e. $j_k < t$, since the pivot element in T_1, a^1_{rs}, must be positive and $c^2_t > 0$. However, in T_2, all cost coefficients c^2_j except c^2_t are nonnegative; otherwise x_j would have been selected as entering variable. Thus $c^2_{j_k} < 0$ and $a^1_{ks} > 0$. This is a contradiction because Bland's rule should have selected x_{j_k} rather than x_t in T_1 as leaving variable. □

2.5 Phase I of the Simplex Method

In this section, we show how to transform a linear program into the form presented in Section 2.2. For that purpose, we show how to find a basis of the linear program which leads to a basic feasible solution. Sometimes, of course, we may inherit a bfs as part of the problem formulation. For example, we might have constraints of the form $Ax \leq b$ with $b \geq 0$ in which case the slack variables constitute a bfs. Otherwise, we use the two-phase simplex method to be described in this section.

Consider a linear program in standard form with $b \geq 0$ (this latter restriction is without loss of generality since we may multiply some constraints by -1). In phase I, instead of solving

$$(P) \qquad \begin{array}{rl} \text{Max} & z = c_0 + c^T x \\ \text{subject to:} & \\ & Ax = b \\ & x \geq 0 \end{array}$$

we add some artificial variables $\{x_i^a : i = 1, \ldots, m\}$ and consider the linear program:

$$\begin{array}{rl} \text{Min} & w = \sum_{i=1}^{m} x_i^a \\ \text{subject to:} & \\ & Ax + Ix^a = b \\ & x \geq 0, x^a \geq 0. \end{array}$$

This program is not in the form required by the simplex method but can easily be transformed to it. Changing the $\min w$ by $\max w' = -w$ and expressing the objective function in terms of the initial variables, we obtain:

$$(Q) \qquad \begin{array}{rl} \text{Max} & w' = -e^T b + (e^T A)x \\ \text{subject to:} & \\ & Ax + Ix^a = b \\ & x \geq 0, x^a \geq 0 \end{array}$$

where e is a vector of 1's. We have artificially created a bfs, namely $x = 0$ and $x^a = b$. We now use the simplex method as described in the previous section. There are three possible outcomes.

1. w' is reduced to zero and no artificial variables remain in the basis, i.e. we are left with a basis consisting only of original variables. In this case, we simply delete the columns corresponding to the artificial variables, replace the objective function by the objective function of (P) after having expressed it in terms of the nonbasic variables and use Phase II of the simplex method as described in Section 2.3.

2. $w' < 0$ at optimality. This means that the original LP (P) is infeasible. Indeed, if x is feasible in (P) then $(x, x^a = 0)$ is feasible in (Q) with value $w' = 0$.

3. w' is reduced to zero but some artificial variables remain in the basis. These artificial variables must be at zero level since, for this solution, $-w' = \sum_{i=1}^{m} x_i^a = 0$. Suppose that the ith variable ofthe basis is artificial. We may pivot on any nonzero (not necessarily positive) element $\bar{a}_{ij}$ of row i corresponding to a non-artificial variable x_j. Since $\bar{b}_i = 0$, no change in the solution or in w' will result. We say that we are *driving the artificial variables out of the basis*. By repeating this for all artificial variables in the basis, we obtain a basis consisting only of original variables. We have thus reduced this case to case 1.

   There is still one detail that needs consideration. We might be unsuccessful in driving one artificial variable out the basis if $\bar{a}_{ij} = 0$ for $j = 1, \ldots, n$. However, this means that we have arrived at a zero row in the original matrix by performing elementary row operations, implying that the constraint is redundant. We can delete this constraint and continue in phase II with a basis of lower dimension.

Example

Consider the following example already expressed in tableau form.

$-z$	x_1	x_2	x_3	x_4	
1	20	16	12	5	0
	1	0	1	2	4
	0	1	2	3	2
	0	1	0	2	2

We observe that we don't need to add three artificial variables since we can use x_1 as first basic variable. In phase I, we solve the linear program:

w	x_1	x_2	x_3	x_4	x_1^a	x_2^a	
1		2	2	5			4
	1	0	1	2			4
		$\boxed{1}$	2	3	1		2
		1	0	2		1	2

The objective function is to minimize $x_1^a + x_2^a$ and, as a result, the objective function coefficients of the nonbasic variables as well as $-\bar{c}_0$ are obtained by taking the negative of the sum of all rows corresponding to artificial variables. Pivoting on $\bar{a}_{22}$, we obtain:

w	x_1	x_2	x_3	x_4	x_1^a	x_2^a	
1			-2	-1	-2		0
	1		1	2	0		4
		1	2	3	1		2
			-2	$\boxed{-1}$	-1	1	0

This tableau is optimal and, since $w = 0$, the original linear program is feasible. To obtain a bfs, we need to drive x_1^a out of the basis. This can be done by pivoting on say $\bar{a}_{34}$. Doing so, we get:

w	x_1	x_2	x_3	x_4	x_1^a	x_2^a	
1			0		-1	-1	0
	1		-3		-2	2	4
		1	-4		-2	3	2
			2	1	1	-1	0

Expressing z as a function of $\{x_1, x_2, x_4\}$, we have transformed our original LP into:

$-z$	x_1	x_2	x_3	x_4	
1			126		-112
	1		-3		4
		1	-4		2
			2	1	0

This can be solved by phase II of the simplex method.

3 Linear Programming in Matrix Form

In this chapter, we show that the entries of the current tableau are uniquely determined by the collection of decision variables that form the basis and we give matrix expressions for these entries.

Consider a feasible linear program in standard form:

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

where A has full row rank. Consider now any intermediate tableau of phase II of the simplex method and let B denote the corresponding collection of basic variables. If D (resp. d) is an $m \times n$ matrix (resp. an n-vector), let D_B (resp. d_B) denote the restriction of D (resp. d) to the columns (resp. rows) corresponding to B. We define analogously D_N and d_N for the collection N of nonbasic variables. For example, $Ax = b$ can be rewritten as $A_B x_B + A_N x_N = b$. After possible regrouping of the basic variables, the current tableau looks as follows:

	x_B	x_N	
$-z$	0	$\bar{c}_N^T$	$-\bar{c}_0$
	$\bar{A}_B = I$	$\bar{A}_N$	$\bar{b}$.

Since the current tableau has been obtained from the original tableau by a sequence of elementary row operations, we conclude that there exists an invertible matrix P (see Section 2.1) such that:

$$PA_B = \bar{A}_B = I$$

$$PA_N = \bar{A}_N$$

and

$$Pb = \bar{b}.$$

This implies that $P = A_B^{-1}$ and therefore:

$$\bar{A}_N = A_B^{-1} A_N$$

and

$$\bar{b} = A_B^{-1} b.$$

Moreover, since the objective functions of the original and current tableaus are equivalent (i.e. $c_B^T x_B + c_N^T x_N = \bar{c}_0 + \bar{c}_B^T x_B + \bar{c}_N^T x_N = \bar{c}_0 + \bar{c}_N^T x_N$) and $x_B = \bar{b} - \bar{A}_N x_N$, we derive that:

$$\bar{c}_N^T = c_N^T - c_B^T \bar{A}_N = c_N^T - c_B^T A_B^{-1} A_N$$

and

$$\bar{c}_0 = c_B^T \bar{b} = c_B^T A_B^{-1} b.$$

This can also be written as:

$$\bar{c}^T = c^T - c_B^T A_B^{-1} A.$$

As we'll see in the next chapter, it is convenient to define an m-vector y by $y^T = c_B^T A_B^{-1}$. In summary, the current tableau can be expressed in terms of the original data as:

	x_B	x_N	
$-z$	0	$c_N^T - y^T A_N$	$-y^T b$
	I	$A_B^{-1} A_N$	$A_B^{-1} b.$

The simplex method could be described using this matrix form. For example, this optimality criterion becomes $c_N^T - y^T A_N \leq 0$ or, equivalently, $c^T - y^T A \leq 0$, i.e. $A^T y \geq c$ where $y^T = c_B^T A_B^{-1}$.

4 Duality

Duality is the most important and useful structural property of linear programs. We start by illustrating the notion on an example.

Consider the linear program:

$$\begin{aligned} \text{Max} \quad & z = 5x_1 + 4x_2 \\ \text{subject to:} \quad & \\ & x_1 \leq 4 && (1) \\ & x_1 + 2x_2 \leq 10 && (2) \\ & 3x_1 + 2x_2 \leq 16 && (3) \\ & x_1, x_2 \geq 0. \end{aligned}$$

We shall refer to this linear program as the *primal*. By exhibiting any feasible esolution, say $x_1 = 4$ and $x_2 = 2$, one derives a *lower bound* (since we are maximizing) on the optimum value z^* of the linear program; in this case, we have $z^* \geq 28$. How could we derive *upper bounds* on z^*? Multiplying inequality (3) by 2, we derive that $6x_1 + 4x_2 \leq 32$ for any feasible (x_1, x_2). Since $x_1 \geq 0$, this in turn implies that $z = 5x_1 + 4x_2 \leq 6x_1 + 4x_2 \leq 32$ for any feasible solution and, thus, $z^* \leq 32$. One can even combine several inequalities to get upper bounds. Adding up all three inequalities, we get $5x_1 + 4x_2 \leq 30$, implying that $z^* \leq 30$. In general, one would multiply inequality (1) by some nonnegative scalar y_1, inequality (2) by some nonnegative y_2 and inequality (3) by some nonnegative y_3, and add them together, deriving that

$$(y_1 + y_2 + 3y_3)x_1 + (2y_2 + 2y_3)x_2 \leq 4y_1 + 10y_2 + 16y_3.$$

To derive an upper bound on z^*, one would then impose that the coefficients of the x_i's in this implied inequality dominate the corresponding cost coefficients: $y_1+y_2+3y_3 \geq 5$ and $2y_2+2y_3 \geq 4$. To derive the *best* upper bound (i.e. smallest) this way, one is thus led to solve the following so-called *dual* linear program:

$$\begin{aligned} \text{Min} \quad & w = 4y_1 + 10y_2 + 16y_3 \\ \text{subject to:} \quad & \\ & y_1 + y_2 + 3y_3 \geq 5 \\ & 2y_2 + 2y_3 \geq 4 \\ & y_1 \geq 0, y_2 \geq 0, y_3 \geq 0. \end{aligned}$$

Observe how the dual linear program is constructed from the primal: one is a maximization problem, the other a minimization; the cost coefficients of one are the RHS of the other and vice versa; the constraint matrix is just transposed (see below for more precise and formal rules). The optimum solution to this linear program is $y_1 = 0$, $y_2 = 0.5$ and $y_3 = 1.5$, giving an upper bound of 29 on z^*. What we shall show in this chapter is that this upper bound is in fact *equal* to the optimum value of the primal. Here, $x_1 = 3$ and $x_2 = 3.5$ is a feasible solution to the primal of value 29 as well. Because of our upper bound of 29, this solution must be optimal, and thus duality is a way to prove optimality.

4.1 Duality for Linear Programs in canonical form

Given a linear program (P) in canonical form

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

we define its *dual* linear program (D) as

$$(D) \qquad \begin{aligned} \text{Min} \quad & w = b^T y \\ \text{subject to:} \quad & \\ & A^T y \geq c \\ & y \geq 0. \end{aligned}$$

(P) is called the *primal* linear program. Notice there is a dual variable associated with each primal constraint, and a dual constraint associated with each primal variable. In fact, the primal and dual are indistinguishable in the following sense:

Proposition 4.1. The dual of the dual is the primal.

Proof. To construct the dual of the dual, we first need to put (D) in canonical form:

$$(D') \qquad \begin{array}{rl} \text{Max} & w' = -w = -b^T y \\ \text{subject to:} & \\ & -A^T y \leq -c \\ & y \geq 0. \end{array}$$

Therefore the dual (DD') of D is:

$$(DD') \qquad \begin{array}{rl} \text{Min} & z' = -c^T x \\ \text{subject to:} & \\ & -Ax \geq -b \\ & x \geq 0. \end{array}$$

Transforming this linear program into canonical form, we obtain (P). □

Theorem 4.2 (Weak Duality)**.** If x is feasible in (P) with value z and y is feasible in (D) with value w then $z \leq w$.

Proof.

$$z = c^T x \overset{x \geq 0}{\leq} (A^T y)^T x = y^T A x \overset{y \geq 0}{\leq} y^T b = b^T y = w.$$

□

Any *dual feasible solution* (i.e. feasible in (D)) gives an upper bound on the optimal value z^* of the primal (P) and vice versa (i.e. any primal feasible solution gives a lower bound on the optimal value w^* of the dual (D)). In order to take care of infeasible linear programs, we adopt the convention that the maximum value of any function over an empty set is defined to be $-\infty$ while the minimum value of any function over an empty set is $+\infty$. Therefore, we have the following corollary:

Corollary 4.3 (Weak Duality)**.** $z^* \leq w^*$.

What is more surprising is the fact that this inequality is in most cases an equality.

Theorem 4.4 (Strong Duality)**.** If z^* is finite then so is w^* and $z^* = w^*$.

Proof. The proof uses the simplex method. In order to solve (P) with the simplex method, we reformulate it in standard form:

$$(P) \qquad \begin{array}{rl} \text{Max} & z = c^T x \\ \text{subject to:} & \\ & Ax + Is = b \\ & x \geq 0, s \geq 0. \end{array}$$

Let $\tilde{A} = (A \; I)$, $\tilde{x} = \begin{pmatrix} x \\ s \end{pmatrix}$ and $\tilde{c} = \begin{pmatrix} c \\ 0 \end{pmatrix}$. Let B be the optimal basis obtained by the simplex method. The optimality conditions imply that

$$\tilde{A}^T y \geq \tilde{c}$$

where

$$y^T = (\tilde{c}_B)^T \tilde{A}_B^{-1}.$$

Replacing $\tilde{A}$ by $(A\ I)$ and $\tilde{c}$ by $\begin{pmatrix} c \\ 0 \end{pmatrix}$, we obtain:

$$A^T y \geq c$$

and

$$y \geq 0.$$

This implies that y is a dual feasible solution. Moreover, the value of y is precisely $w = y^T b = (\tilde{c}_B)^T \tilde{A}_B^{-1} b = (\tilde{c}_B)^T \tilde{x}_B = z^*$. Therefore, by weak duality, we have $z^* = w^*$. □

Since the dual of the dual is the primal, we have that if either the primal or the dual is feasible and bounded then so are both of them and their values are equal. From weak duality, we know that if (P) is unbounded (i.e. $z^* = +\infty$) then (D) is infeasible ($w^* = +\infty$). Similarly, if (D) is unbounded (i.e. $w^* = -\infty$) then (P) is infeasible ($z^* = -\infty$). However, the converse to these statements are not true: There exist dual pairs of linear programs for which both the primal and the dual are infeasible. Here is a summary of the possible alternatives:

Primal Dual	z^* finite	unbounded ($z^* = \infty$)	infeasible ($z^* = -\infty$)
w^* finite	$z^* = w^*$	impossible	impossible
unbounded ($w^* = -\infty$)	impossible	impossible	possible
infeasible ($w^* = +\infty$)	impossible	possible	possible

4.2 The dual of a linear program in general form

In order to find the dual of any linear program (P), we can first transform it into a linear program in canonical form (see Section 1.2), then write its dual and possibly simplify it by transforming it into some equivalent form.

For example, considering the linear program

$$(P) \qquad \begin{array}{lll} \text{Max} & z = c^T x & \\ \text{subject to:} & & \\ & \sum_j a_{ij} x_j \leq b_i & i \in I_1 \\ & \sum_j a_{ij} x_j \geq b_i & i \in I_2 \\ & \sum_j a_{ij} x_j = b_i & i \in I_3 \\ & x_j \geq 0 & j = 1, \ldots, n, \end{array}$$

we can first transform it into

$$
(P')\qquad
\begin{array}{lll}
\text{Max} & z = c^T x & \\
\text{subject to:} & & \\
& \sum_j a_{ij} x_j \leq b_i & i \in I_1 \\
& -\sum_j a_{ij} x_j \leq -b_i & i \in I_2 \\
& \sum_j a_{ij} x_j \leq b_i & i \in I_3 \\
& -\sum_j a_{ij} x_j \leq -b_i & i \in I_3 \\
& x_j \geq 0 & j = 1, \dots, n.
\end{array}
$$

Assigning the vectors y^1, y^2, y^3 and y^4 of dual variables to the first, second, third and fourth set of constraints respectively, we obtain the dual:

$$
(D')\qquad
\begin{array}{lll}
\text{Min} & w = \sum_{i \in I_1} b_i y_i^1 - \sum_{i \in I_2} b_i y_i^2 + \sum_{i \in I_3} b_i y_i^3 - \sum_{i \in I_3} b_i y_i^4 & \\
\text{subject to:} & & \\
& \sum_{i \in I_1} a_{ij} y_i^1 - \sum_{i \in I_2} a_{ij} y_i^2 + \sum_{i \in I_3} a_{ij} y_i^3 - \sum_{i \in I_3} a_{ij} y_i^4 \geq c_j & j = 1, \dots, n \\
& y^1, y^2, y^3, y^4 \geq 0. &
\end{array}
$$

This dual can be written in a simplified form by letting

$$
\begin{cases}
y_i = y_i^1 & i \in I_1 \\
y_i = -y_i^2 & i \in I_2 \\
y_i = y_i^3 - y_i^4 & i \in I_3.
\end{cases}
$$

In terms of y_i, we obtain (verify it!) the following equivalent dual linear program

$$
(D)\qquad
\begin{array}{lll}
\text{Min} & w = \sum_{i \in I} b_i y_i & \\
\text{subject to:} & & \\
& \sum_{i \in I} a_{ij} y_i \geq c_j & j = 1, \dots, n \\
& y_i \geq 0 & i \in I_1 \\
& y_i \leq 0 & i \in I_2 \\
& y_i \gtreqless 0 & i \in I_3,
\end{array}
$$

where $I = I_1 \cup I_2 \cup I_3$.

We could have avoided all these steps by just noticing that, if the primal program is a maximization program, then inequalities with a $\leq$ sign in the primal correspond to nonnegative dual variables, inequalities with a $\geq$ sign correspond to nonpositive dual variables, and equalities correspond to unrestricted in sign dual variables.

By performing similar transformations for the restrictions on the primal variables, we obtain the following set of rules for constructing the dual linear program of any linear program:

Primal	$\longleftrightarrow$	**Dual**
Max	$\longleftrightarrow$	Min
$\sum_j a_{ij}x_j \leq b_i$	$\longleftrightarrow$	$y_i \geq 0$
$\sum_j a_{ij}x_j \geq b_i$	$\longleftrightarrow$	$y_i \leq 0$
$\sum_j a_{ij}x_j = b_i$	$\longleftrightarrow$	$y_i \gtrless 0$
$x_j \geq 0$	$\longleftrightarrow$	$\sum_i a_{ij}y_i \geq c_j$
$x_j \leq 0$	$\longleftrightarrow$	$\sum_i a_{ij}y_i \leq c_j$
$x_j \gtrless 0$	$\longleftrightarrow$	$\sum_i a_{ij}y_i = c_j.$

If the primal linear program is in fact a minimization program then we simply use the above rules from right to left. This follows from the fact that the dual of the dual is the primal.

4.3 Complementary slackness

Consider a pair of dual linear programs

$$(P)\qquad \begin{array}{rl} \text{Max} & z = c^Tx \\ \text{subject to:} & \\ & Ax \leq b \\ & x \geq 0 \end{array}$$

and

$$(D)\qquad \begin{array}{rl} \text{Min} & w = b^Ty \\ \text{subject to:} & \\ & A^Ty \geq c \\ & y \geq 0. \end{array}$$

Strong duality allows to give a simple test for optimality.

Theorem 4.5 (Complementary Slackness)**.** *If x is feasible in (P) and y is feasible in (D) then x is optimal in (P) and y is optimal in (D) iff $y^T(b - Ax) = 0$ and $x^T(A^Ty - c)$.*

The latter statement can also be written as
either $y_i = 0$ or $(Ax)_i = b_i$ (or both) and
either $x_j = 0$ or $(A^Ty)_j = c_j$ (or both).

Proof. By strong duality we know that x is optimal in (P) and y is optimal in (D) iff $c^Tx = b^Ty$. Moreover, (cfr. Theorem 4.2) we always have that:

$$c^Tx \leq y^TAx \leq y^Tb = b^Ty.$$

Therefore, $c^Tx = b^Ty$ is equivalent to $c^Tx = y^TAx$ and $y^TAx = y^Tb$. Rearranging these expressions, we obtain $x^T(A^Ty - c) = 0$ and $y^T(b - Ax) = 0$. □

Corollary 4.6. Let x be feasible in (P). Then x is optimal iff there exists y such that

$$\begin{array}{lll} A^T y \left\{ \begin{array}{c} \geq \\ = \end{array} \right. c_j & \text{if} & \left\{ \begin{array}{l} x_j = 0 \\ x_j > 0 \end{array} \right. \\ y_i \left\{ \begin{array}{c} \geq \\ = \end{array} \right. 0 & \text{if} & \left\{ \begin{array}{l} (Ax)_i = b_i \\ (Ax)_i < b_i. \end{array} \right. \end{array}$$

As a result, the optimality of a given primal feasible solution can be tested by checking the feasibility of a system of linear inequalities and equalities.

As should be by now familiar, we can write similar conditions for linear programs in other forms. For example,

Theorem 4.7. Let x be feasible in

$$\begin{array}{llr} & \text{Max} & z = c^T x \\ & \text{subject to:} & \\ (P) & & Ax = b \\ & & x \geq 0 \end{array}$$

and y feasible in

$$\begin{array}{llr} & \text{Min} & w = b^T y \\ & \text{subject to:} & \\ (D) & & A^T y \geq c. \end{array}$$

Then x is optimal in (P) and y is optimal in (D) iff $x^T(A^T y - c) = 0$.

4.4 The separating hyperplane theorem

In this section, we use duality to obtain a necessary and sufficient condition for feasibility of a system of linear inequalities and equalities.

Theorem 4.8 (The Separating Hyperplane Theorem)**.** *$Ax = b$, $x \geq 0$ has no solution iff $\exists y \in \mathbb{R}^m$: $A^T y \geq 0$ and $b^T y < 0$.*

The geometric interpretation behind the separating hyperplane theorem is as follows: Let $a_1, \ldots, a_n \in \mathbb{R}^m$ be the columns of A. Then b does not belong to the cone $K = \{\sum_{i=1}^n a_i x_i : x_i \geq 0$ for $i = 1, \ldots, n\}$ generated by the a_i's iff there exists an hyperplane $\{x : x^T y = 0\}$ (defined by its normal y) such that K is entirely on one side of the hyperplane (i.e. $a_i^T y \geq 0$ for $i = 1, \ldots, n$) while b is on the other side ($b^T y < 0$).

Proof. Consider the pair of dual linear programs

$$\begin{array}{llr} & \text{Max} & z = 0^T x \\ & \text{subject to:} & \\ (P) & & Ax = b \\ & & x \geq 0 \end{array}$$

and

$$\begin{array}{llr} & \text{Min} & w = b^T y \\ & \text{subject to:} & \\ (D) & & A^T y \geq 0. \end{array}$$

Notice that (D) is certainly feasible since $y = 0$ is a feasible solution. As a result, duality implies that (P) is infeasible iff (D) is unbounded. However, since λy is dual feasible for any $\lambda \geq 0$ and any dual feasible solution y, the unboundedness of (D) is equivalent to the existence of y such that $A^T y \geq 0, y \geq 0$ and $b^T y < 0$. □

Other forms of the separating hyperplane theorem include:

Theorem 4.9. *$Ax \leq b$ has no solution iff $\exists y \geq 0 : A^T y = 0$ and $b^T y < 0$.*

5 Zero-Sum Matrix Games

In a matrix game, there are two players, say player I and player II. Player I has m different *pure strategies* to choose from while player II has n different pure strategies. If player I selects strategy i and player II selects strategy j then this results in player I gaining a_{ij} units and player II losing a_{ij} units. So, if a_{ij} is positive, player II pays a_{ij} units to player I while if a_{ij} is negative then player I pays $-a_{ij}$ units to player II. Since the amounts gained by one player equal the amounts paid by the other, this game is called a *zero-sum* game. The matrix $A = [a_{ij}]$ is known to both players and is called the *payoff matrix*. In a sequence of games, player I (resp. player II) may decide to randomize his choice of pure strategies by selecting strategy i (resp. j) with some probability y_i (resp. x_j). The vector y (resp. x) satisfies

$$\sum_{i=1}^{m} y_i = 1 \qquad\qquad (resp. \sum_{j=1}^{n} x_j = 1),$$

$y_i \geq 0$ (resp. $x_j \geq 0$) and defines a *mixed strategy*.

If player I adopts the mixed strategy y then his expected gain g_j if player II selects strategy j is given by:

$$g_j = \sum_i a_{ij} y_i = (y^T A)_j = y^T A e_j.$$

By using y, player I assures himself a guaranteed gain of

$$g = \min_j g_j = \min_j (y^T A)_j.$$

Similarly, if player II adopts the mixed strategy x then his expected loss l_i if player I selects strategy i is given by:

$$l_i = \sum_j a_{ij} x_j = (Ax)_i = e_i^T Ax$$

and his guaranteed loss[4] is

$$l = \max_i l_i = \max_i (Ax)_i.$$

If player I uses the mixed strategy y and player II uses the mixed strategy x then the expected gain of player I is $h = \sum_{i,j} y_i a_{ij} x_j = y^T Ax$.

Theorem 5.1. If y and x are mixed strategies respectively for players I and II then $g \leq l$.

Proof. We have that

$$h = y^T Ax = \sum_i y_i (Ax)_i \leq l \sum_i y_i = l$$

and

$$h = y^T Ax = \sum_j (y^T A)_j x_j \geq g \sum_j x_j = g$$

proving the result. □

Player I will try to select y so as to maximize his guaranteed gain g while player II will select x so as to minimize l. From the above result, we know that the optimal guaranteed gain g^* of player I is at most the optimal guaranteed loss l^* of player II.

The main result in zero-sum matrix games is the following result obtained by Von Neumann and called the minimax theorem.

Theorem 5.2 (The Minimax Theorem)**.** There exist mixed strategies x^* and y^* such that $g^* = l^*$.

Proof. In order to prove this result, we formulate the objectives of both players as linear programs. Player II's objective is to minimize l. This can be expressed by:

$$(P) \qquad \begin{array}{rl} \text{Min} & l \\ \text{subject to:} & \\ & Ax \leq le \\ & e^T x = 1 \\ & x \geq 0, l \gtreqless 0 \end{array}$$

where e is a vector of all 1's. Indeed, for any optimal solution x^*, l^* to (P), we know that $l^* = \max_i (Ax^*)_i$ since otherwise l^* could be decreased without violating feasibility.

Similarly, player I's objective can be expressed by:

$$(D) \qquad \begin{array}{rl} \text{Max} & g \\ \text{subject to:} & \\ & A^T y \geq ge \\ & e^T y = 1 \\ & y \geq 0, g \gtreqless 0 \end{array}$$

Again, any optimal solution to the above program will satisfy $g^* = \min_j (A^T y^*)_j$.

The result follows by noticing that (P) and (D) constitute a pair of dual linear programs (verify it!) and, therefore, by strong duality we know that $g^* = l^*$. □

[4] Here guaranteed means that he'll loose *at most* l.

The above Theorem can be rewritten as follows (This explains why it is called the minimax theorem):

Corollary 5.3.

$$\max_{e^Ty=1,y\geq 0}\ \min_{e^Tx=1,x\geq 0} y^TAx = \min_{e^Tx=1,x\geq 0}\ \max_{e^Ty=1,y\geq 0} y^TAx.$$

Indeed

$$\min_{e^Tx=1,x\geq 0} y^TAx = \min_j (y^TA)_j = g$$

and

$$\max_{e^Ty=1,y\geq 0} y^TAx = \max_i (Ax)_i = l.$$

Example

Consider the game with payoff matrix

$$A = \begin{pmatrix} 1 & -3 \\ -2 & 4 \end{pmatrix}.$$

Solving the linear program (P), we obtain the following optimal mixed strategies for both players (do it by yourself!):

$$x^* = \begin{pmatrix} 7/10 \\ 3/10 \end{pmatrix} \qquad \text{and} \qquad y^* = \begin{pmatrix} 6/10 \\ 4/10 \end{pmatrix},$$

for which $g^* = l^* = -2/10$.

A matrix game is said to be *symmetric* if $A = -A^T$. Any symmetric game is *fair*, i.e. $g^* = l^* = 0$.