MA423 - Fundamentals of Operations Research
Lecture 3: The Simplex Method

Katerina Papadaki[1]

Academic year 2017/18

[1]London School of Economics and Political Science. Houghton Street London WC2A 2AE (k.p.papadaki@lse.ac.uk)

Contents

Chapter 1

Preliminaries: definitions and geometrical insights

In the last two lectures we looked at formulations of linear programs, linear programs in standard form and duality. In this lecture we will introduce the simplex method an algorithm for solving linear programs. But before we do that we need to look at some geometric concepts regarding linear programming.

1.1 Basic solutions and extreme points

Consider the LP problem

$$\begin{array}{rrcl} \text{maximise} & 2x_1 + 8x_2 & & \\ \text{subject to} & 2x_1 + x_2 & \leq & 10 \\ & x_1 + 2x_2 & \leq & 10 \\ & x_1 + x_2 & \leq & 6 \\ & x_1 + 3x_2 & \leq & 12 \\ & 3x_1 - x_2 & \geq & 0 \\ & x_1 + 4x_2 & \geq & 4 \\ & x_1, x_2 & \geq & 0 \end{array} \tag{1.1}$$

depicted graphically in Figure 1.1.

The optimal solution of the LP (1.1) is the point A of coordinates $x_1 = 1.2$, $x_2 = 3.6$. It can be observed from the diagram that this point is a corner point of the polygon that defines the feasible region. As we shall see, the fact that the optimum is achieved by some "corner point" of the feasible region is true in general, even for problems with more than two variables. Firstly, we need to define

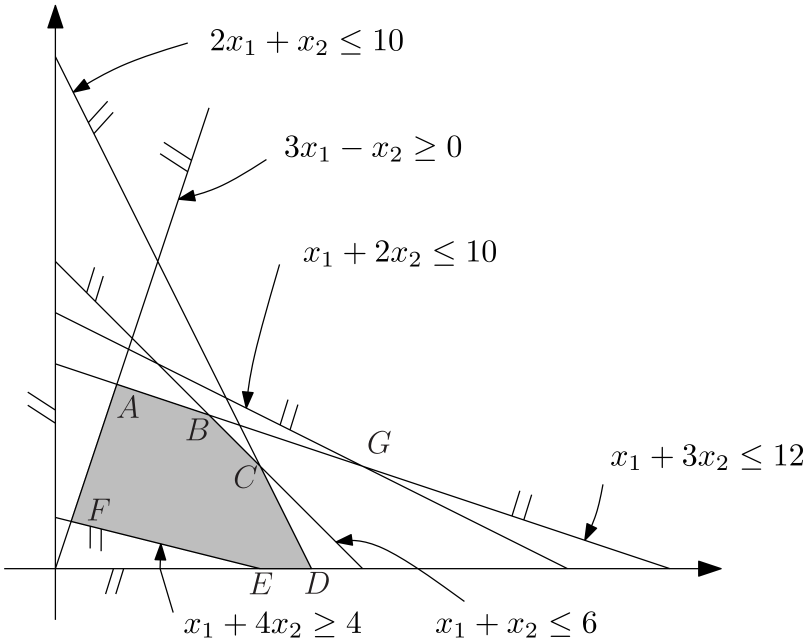

Figure 1.1: The feasible region is represented by the area shaded gray.

what we mean by "corner points", which in the language of Linear Programming are called *extreme points*.

In two dimensions a corner point is one at the intersection of two non-parallel inequalities. In order to formalise the concept to the case of more variables, we need to recall the following concept from linear algebra.

Independent constraints: Given a system of n linear inequalities in n variables, we say that the n inequalities are *(linearly) independent* if there exists a unique solution satisfying all n of them at equality.

For example, the inequalities

$$\begin{array}{ccccc} x_1 & +x_2 & & \leq & 2 \\ x_1 & & +x_3 & = & 2 \\ & x_2 & +x_3 & \geq & 2 \end{array}$$

are independent because the only solution satisfying all three of them at equality is the point $(x_1^*, x_2^*, x_3^*) = (1, 1, 1)$, as one can verify.

On the other hand, the inequalities

$$\begin{array}{rrrcl} x_1 & +x_2 & & \leq & 2 \\ x_1 & & +x_3 & = & 2 \\ & -x_2 & +x_3 & \geq & 0 \end{array}$$

are not independent because there are multiple solutions satisfying all three of them at equality, for example the points $(1, 1, 1)$ and $(0, 2, 2)$.

Note that in 2 dimensions two equations are independent if they do not define parallel lines. If they do define parallel lines and if the two parallel lines are distinct then there are no solutions satisfying both equations, while if the two lines coincide then there are infinitely many solutions satisfying both equations.

To simplify matters, in all our examples and in all exercises we will always work under the assumption that any n equalities we deal with are independent, unless otherwise specified.

> Consider a system of linear constraints in n variables.
>
> **Basic point:** A *basic point* for the system is a point satisfying at equality n independent constraints from the system.
>
> **Defining constraints (active constraints):** Given a basic point $\bar{x}$ for the system, a set of *defining constraints for* $\bar{x}$ (also called *active constraints for* $\bar{x}$) is any choice of n independent constraints satisfied at equality by $\bar{x}$.

Note that there might be more than one set of defining constraints for each basic point as we will see later.

Example. Consider the LP in equation (1.1), whose feasible region is represented in Figure 1.1. The basic point labelled B is defined by the intersection of constraints 3 and 4, so it is the unique solution of the system:

$$\begin{aligned} x_1 + x_2 &= 6 \\ x_1 + 3x_2 &= 12 \end{aligned}$$

Basic point D is defined by constraint 1 and the nonnegativity of x_2, that is:

$$\begin{aligned} 2x_1 + x_2 &= 10 \\ x_2 &= 0 \end{aligned}$$

Note that, in the definition, we do not insist that a basic point is feasible for the whole system of constraints. For example, basic points B and D are feasible. Basic point G defined by

$$\begin{aligned} 2x_1 + x_2 &= 10 \\ x_1 + 3x_2 &= 12 \end{aligned}$$

is infeasible.

Extreme point: A point that is both basic and feasible for a given system of linear constrains.

For example, points A to F in Figure 1.1 are extreme points while points G and the origin are basic but they are not extreme.

1.1.1 Is a point an extreme point?

For a system of linear constraints in n-variables $x_1, \ldots, x_n$, to determine whether a point x^* is extreme:

- Substitute the values of $x_1^*, x_2^*, \ldots, x_n^*$ into the constraints, and verify if the point is feasible.
- Determine whether n (or more) of them are satisfied as equalities,
- Among the constraints satisfied at equality by x^*, determine if there are n of them that are independent.

For example, consider as usual the system of constraints (1.1), and let $x^* = (4, 2)$.

Substituting the values $x_1 = 4$, $x_2 = 2$ into the constraints, we obtain

$$\begin{array}{rlcll} 2 \cdot 4 + 2 & = 10 & = & 10 & \surd \\ 4 + 2 \cdot 2 & = 8 & \leq & 10 & \\ 4 + 2 & = 6 & = & 6 & \surd \\ 4 + 3 \cdot 2 & = 10 & \leq & 12 & \\ 3 \cdot 4 - 2 & = 10 & \geq & 0 & \\ 4 + 4 \cdot 2 & = 12 & \geq & 4 & \\ 4, 2 & & \geq & 0 & \end{array}$$

The point is basic as constraint 1 and constraint 3 are satisfied as equalities and they are independent. The point (4, 2) is point C in the diagram in Figure 1.1. Point C is an extreme point as these calculations show that it is feasible as well as basic.

Note: In general, a basic point might have more than one set of defining constraints. For example, point $\bar{x} = (1, 1, 1)$ is basic for the following system

$$\begin{array}{rrrcl} x_1 & +x_2 & & \leq & 2 \\ x_1 & & +x_3 & = & 2 \\ & x_2 & +x_3 & \geq & 2 \\ x_1 & -x_2 & +x_3 & \geq & 1 \end{array}$$

but each subset consisting of three constraints is defining for $\bar{x}$, since $\bar{x}$ satisfies at equality all four constraints and every three of them are independent.

1.1.2 Optimality and extreme points

The following is one of the most fundamental and useful facts in Linear Programming.

> **Extreme optimal solutions.** Whenever an LP admits an optimal solution, there exists some optimal solution that is an extreme point of the feasible region.

This fact tells us that, when solving an LP, one only needs to search among the extreme points, and pick the one with best objective value.

Note: a linear programming problem may have non-extreme points that are also optimal, but according to the statement above there is always at least an extreme point on the optimal contour. For example, consider the LP in (1.2) depicted in Figure (1.2).

$$\begin{array}{rrcl} \text{maximise} & 2x_1 + 2x_2 & & \\ \text{subject to} & x_1 + x_2 & \leq & 4 \\ & x_2 & \leq & 3 \\ & x_1, x_2 & \geq & 0 \end{array} \tag{1.2}$$

(1.2) has two extreme points that are optimal, namely the points $(1, 3)$ and $(4, 0)$, but there are also infinitely many optimal solutions that are not extreme, namely all points in the line segment between $(1, 3)$ and $(4, 0)$.

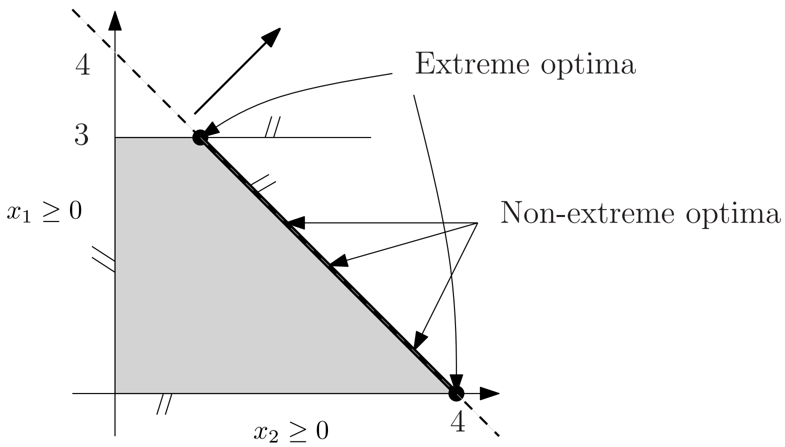

Figure 1.2: Problem (1.2) has two extreme optimal solutions and infinitely many non-extreme optimal solutions.

1.2 Effective Constraints and basic variables

Constraints of the form $x_i \geq 0$ are called *nonnegativity constraints* and all other constraints are called *resource constraints*.

We have seen that an extreme point of an LP in n variables is defined by n independent constraints satisfied at equality at the point. The constraints that define the extreme point can either belong to the resource constraints or to the nonnegativity constraints. This leads to the following definitions.

Consider an extreme point $\bar{x}$ of a system of linear constraints, and a set S of constraints defining $\bar{x}$.
Effective constraints. The resource constraints in S which are defining for $\bar{x}$ are called the *effective constraints* at $\bar{x}$ with respect to S. The remaining resource constraints are *ineffective*.

Basic variables. If a nonnegativity constraint, say constraint $x_j \geq 0$, is defining at $\bar{x}$ with respect to S, then we say that x_j is a *non-basic variable* with respect to S. The other variables are *basic variables* at that point with respect to S.

Observe that non-basic variables always have value 0. Thus, the only variables that may take positive values in the extreme point are the basic ones. However, it is possible for a basic variable to take 0 value (see section 1.4).

Given that an extreme point is defined by n independent constraints, the number of nonbasic variables plus the number of effective constraints must be n. Considering that the number of basic variables is n minus the number of nonbasic variables, it follows that:

> At any extreme point, the number of effective constraints equals the number of basic variables.

For example, consider the linear program

$$\begin{array}{rrcl}
\text{maximise} & 3x_1 + 5x_2 + 2x_3 & & \\
\text{subject to} & 2x_1 + 2x_2 + x_3 & \leq & 4 \\
& 3x_1 - x_2 + 2x_3 & \leq & 5 \\
& x_1 - x_2 - x_3 & \geq & 1 \\
& x_1 \qquad\qquad & \geq & 0 \\
& x_2 \qquad & \geq & 0 \\
& x_3 & \geq & 0
\end{array} \tag{1.3}$$

For the LP in (1.3), point x^* of coordinates $x_1 = 1.5$, $x_2 = 0.5$, $x_3 = 0$ has one set of three defining constraints 1, 3 and 6. Thus this is basic point and since it is also feasible this is an extreme point. Further, the effective constraints are constraints 1 and 3, variable x_3 is non-basic. Constraint 2 is ineffective and variables are x_1 and x_2 are basic. Note that the number of basic variables is equal to the number of effective constraints.

1.3 Proving Optimality Revisited

In the last lecture we created the dual problem in order to prove optimality of a feasible solution x^*. We found the dual values (positive multipliers) for all the constraints to derive an inequality with the same coefficients as the objective.

However, if x^* is a basic point and we know its n defining constraints then we only need to calculate coefficients (dual values) for the defining constraints to prove optimality. Let's see how this is done geometrically.

For example, consider the two inequalities

$$\begin{array}{rcl}
2x_1 + x_2 & \leq & 10 \\
x_1 + 3x_2 & \leq & 12
\end{array} \tag{1.4}$$

which intersect at $x_1 = 3.6$, $x_2 = 2.8$. Deriving a new constraint with arbitrarily chosen weights of 2 for the first constraint and 1 for the second gives

$$\begin{array}{rrcl}
2 \cdot (& 2x_1 + x_2 & \leq & 10) \\
1 \cdot (& x_1 + 3x_2 & \leq & 12) \\
\hline
& 5x_1 + 5x_2 & \leq & 32
\end{array}$$

The point $x_1 = 3.6$, $x_2 = 2.8$ satisfies $5x_1 + 5x_2 \leq 32$ as an equality. This relationship is illustrated in Figure 1.3.

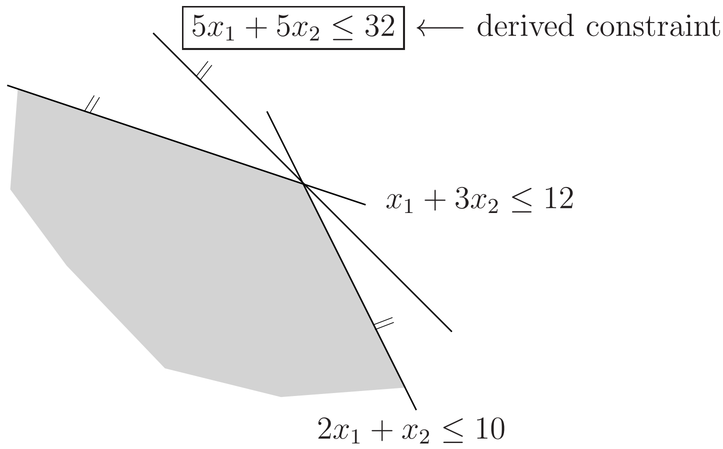

Figure 1.3: Derived constraint.

Now, consider an LP where we want to maximize $5x_1 + 5x_2$ with several constraints including the ones in (1.4) but point $(3.6, 2.8)$ satisfies only the two constraints from (1.4) at equality. So, constraints (1.4) are its defining constraints. A derived inequality such as $5x_1 + 5x_2 \leq 32$ is satisfied by all feasible solutions (since it is a non-negative linear combination of the constraints). Since the objective $5x_1 + 5x_2$ is always less than or equal to 32 and point $(3.6, 2.8)$ achieves the value of 32, this must be an optimal point.

We now have the ingredients for a test to determine whether or not a given point is optimal. Ask the question: is it possible to represent the optimal contour of the objective function as the non-negative combination of a set of defining constraints at that the point? If the answer is yes, then the point is optimal. If the answer is no and there is no other set of defining constraints then the point is not optimal. If there are other sets of defining constraints then we need to try the same thing with the other sets of defining constraints (see section 1.4).

Thus, we only need the defining constraints to prove optimality.

1.4 A degenerate example.

Consider the LP

$$\begin{array}{lrcl} \max & x_1 - x_2 & & \\ & x_1 + x_2 & \leq & 1 \\ & x_1 & \leq & 1 \\ & x_2 & \geq & 0 \end{array}$$

The problem is represented in the Figure 1.4

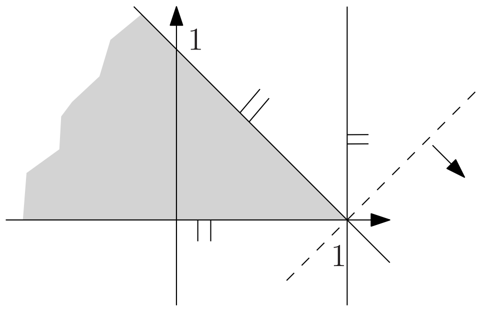

Figure 1.4: Representation of the LP (1.5). The jagged edge of the shaded area indicates that the feasible region continues indefinitely.

It is immediately clear from the figure that the optimal solution is the point $x^* = (1, 0)$, with objective value equal to 1. This problem is "degenerate" in the sense that x^* satisfies as equality three constraints, whereas the problem has only two variables. Since, there can only be two independent equalities that pass from the same point (in two dimensions), the third equality can be derived from the other two. Subtracting $x_1 + x_2 = 1$ from $x_1 = 1$ term by term gives us $x_2 = 0$. This is what makes this example to be degenerate.

Therefore there are three choices of defining constraints for x^*, namely: constraints 1 and 2; constraints 1 and 3; constraints 2 and 3. However, note that not every choice of defining constraints for x^* provides an optimality proof. Indeed, consider the three cases.

- Constraints 1 and 2. In this case, to prove optimality we need to find multipliers y_1, y_2 such that

$$\begin{array}{rcl} y_1 \cdot (x_1 + x_2 & \leq & 1) \\ y_2 \cdot (\ x_1 \quad\quad & \leq & 1) \\ \hline x_1 - x_2 & \leq & 1 \end{array}$$

  Clearly the only solution is $y_1 = -1$, $y_2 = 2$. Since y_2 is negative, this choice of defining constraints does not prove optimality of x^*. Note that with this choice of defining constraints the effective constraints are 1 and 2 and there are no non-basic variables. The basic variables are x_1 and x_2 but basic variable x_2 takes value 0.

- Constraints 1 and 3. In this case, to prove optimality we need to find multipliers y_1, y_3 such that

$$\begin{array}{rrcl} y_1 \cdot (x_1 + x_2 & \leq & 1) \\ y_3 \cdot (\quad -x_2 & \leq & 0) \\ \hline x_1 - x_2 & \leq & 1 \end{array}$$

Clearly the only solution is $y_1 = 1$, $y_3 = 2$. Since y_1 and y_3 are non-negative, this choice of defining constraints proves optimality of x^*. Note that with this choice of defining constraints the effective constraint is 1 the non-basic variable is x_2. The basic variable is x_1 and it takes a positive value.

- Constraints 2 and 3. In this case, to prove optimality we need to find multipliers y_2, y_3 such that

$$\begin{array}{rrcl} y_2 \cdot (\; x_1 \qquad & \leq & 1) \\ y_3 \cdot (\quad -x_2 & \leq & 0) \\ \hline x_1 - x_2 & \leq & 1 \end{array}$$

Clearly the only solution is $y_2 = 1$, $y_3 = 1$. Since y_2 and y_3 are non-negative, this choice of defining constraints proves optimality of x^*. Note that with this choice of defining constraints the effective constraint is 2 and the non-basic variable is x_2. The basic variable is x_1 and it takes a positive value.

Figure 1.5 illustrates the three LPs obtained by only considering each of the three possible choices of the defining constraints. As can be seen from the figure, for the LP comprised of only constraints 1 and 2 the point $x^* = (0, 1)$ is not optimal, which explains why this choice of defining constraints does not prove optimality. For the LP comprised of only constraints 1 and 3, instead, the point $x^* = (1, 0)$ is optimal, and similarly for the LP comprised of only constraints 2 and 3.

Constraint 3, nonnegativity constraint $x_2 \geq 0$ needs to be in the defining set of constraints for point x^* otherwise we cannot prove optimality and we end up with a basic variable x_2 taking 0 value.

1.5 Multiple optimal solutions

It may happen that an objective function reaches its optimum value at more than one extreme point. Consider the LP

$$\begin{array}{rrcl} \text{maximise} & 3x_1 + 3x_2 & & \\ \text{subject to} & x_1 + x_2 & \leq & 5 \\ & 2x_1 - x_2 & \leq & 5 \\ & x_2 & \leq & 3 \\ & x_1, x_2 & \geq & 0 \end{array}$$

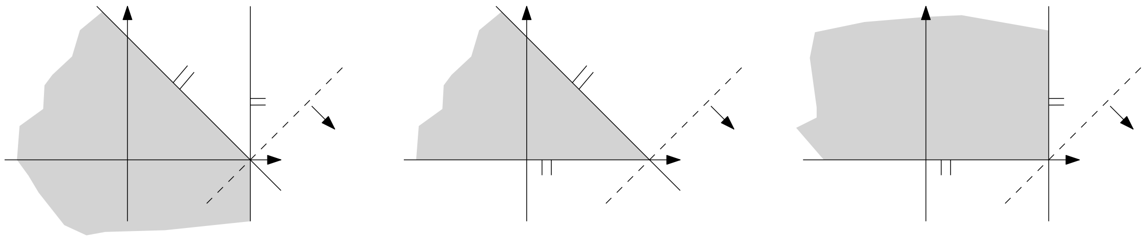

Figure 1.5: From left to right, the above diagrams represent the LPs comprised only of, respectively, constraints 1 and 2, constraints 1 and 3, constraints 2 and 3.

which is illustrated in the diagram in Figure 1.6. The optimal extreme solutions are points P^*, P^{**}, and all points in the line segment between P^* and P^{**} are also optimal, but not extreme. Any computer code will find only one of these points. The question of interest is: if we have the solution at one point can we tell whether or not another optimal point exists?

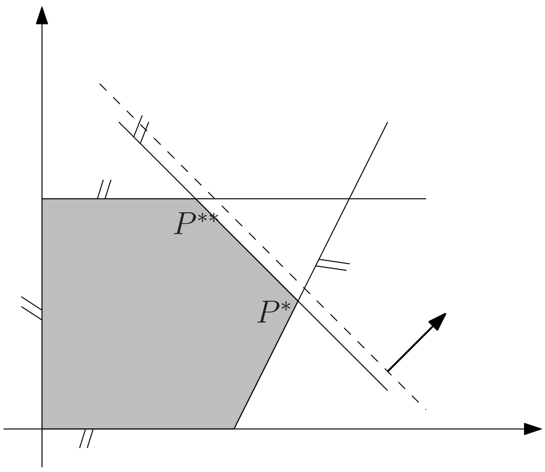

Figure 1.6: A problem with multiple optimal solutions: all points in the line segment between P^* and P^{**} are optimal.

At point P^* the point is defined by constraints 1 and 2. Solving the equations to find the intersection point gives point P^* as $(\frac{10}{3}, \frac{5}{3})$, and a value of the objective function of 15. To determine the value of the dual variables to prove optimality of the extreme point P^*, we need to solve the system

$$\begin{aligned} y_1 + 2y_2 &= 3 \\ y_1 - y_2 &= 3 \end{aligned}$$

which gives $y_1 = 3$, $y_2 = 0$. The dual values of the constraints not defining point P^* are 0. Note that one of the dual value of defining constraint 2 is zero. This is because constraint 1 is parallel to the objective contours and thus it is the only one needed to generate the objective function coefficients.

At point P^{**} the point is defined by constraints 1 and 3. Solving the equations to find the intersection point gives point P^{**} as $(2, 3)$, and a value of the objective function of 15. To determine the value of the dual variables to prove optimality of the extreme point P^{**}, we need to solve the system

$$\begin{aligned} y_1 \quad &= 3 \\ y_1 + y_3 &= 3 \end{aligned}$$

which gives $y_1 = 3$, $y_3 = 0$. The dual values of the constraints not defining point P^{**} are 0. Again note that only constraint 1, which is parallel to the objective, has positive dual value.

Note that, at either point, the dual value of one of the constraints is 0. Indeed, the only constraint with a positive dual value is constrain 1, which defines the line where both P^* and P^{**} lie. In general, in order for the problem to have multiple optimal solutions, the following condition must be met.

Necessary condition for multiple optimal solutions
Given an LP problem, let x^* be an optimal extreme solution. If the problem has multiple optimal solutions then some constraint defining x^* must have dual value 0.

In general, the above condition is necessary but not sufficient for multiple optimal solutions to exist. This means that there are examples where some of the constraints defining an extreme optimal solutions have zero dual value, yet there is only one optimum. The following example illustrates this situation.

$$\begin{aligned} \text{maximise} \quad & 3x_1 + 3x_2 \\ \text{subject to} \quad x_1 + x_2 &\leq 5 \\ 2x_1 - x_2 &\leq 5 \\ x_2 &\leq 3 \\ -x_1 + 2x_2 &\leq 0 \\ x_1, x_2 &\geq 0 \end{aligned}$$

From Figure 1.7, it appears that the unique optimal solution is the point Q, which lies at the intersection of constraints 1, 2, and 4. To prove optimality of point Q, we need to choose two independent constraints among 1, 2 and 4 as defining constraints.

If we select constraints 1 and 2 as the defining ones, the corresponding dual values are $y_1 = 3$, $y_2 = 0$ (the dual values of all remaining constraints being zero). This proves that point Q is optimal, because $y_1, y_2 \geq 0$, but it would not allow us to conclude that Q is the unique optimum, since constraint 2 defines Q but it has a dual value of 0.

Note that, if instead we selected constraints 2 and 4 as defining constraints, the corresponding dual values would be $y_2 = 3$, $y_4 = 3$ (the dual values of all remaining constraints being zero). This proves that the point Q is optimal, because $y_2, y_3 \geq 0$, but also that there is no other optimal solution, because $y_2, y_3 \neq 0$.

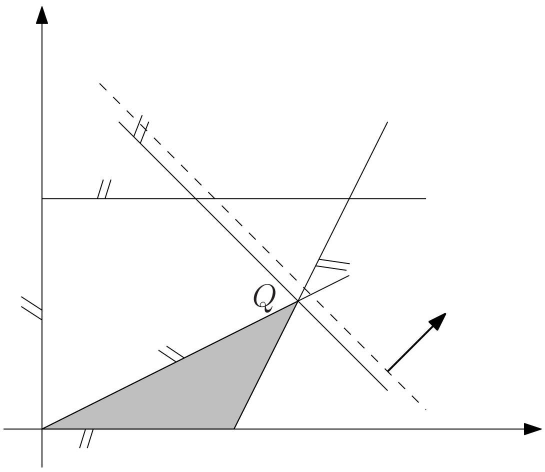

Figure 1.7: A problem with a unique optimum but a dual solution with dual value 0 on some constraint defining the optimal solution.

Note that when we selected constraints 1 and 2, we ignored constraint 4. Without constraint 4 Q is still an optimal solution but it is not a unique optimum. However when we selected constraints 2 and 4, we ignored constraint 1, and without constraint 1 Q is not only an optimal solution but it is the unique optimum.

Chapter 2

The simplex method

Linear programming emerged after the end of world war II as a practical, powerful tool in a wide array of applications. This was made possible by the convergence of two events: the advent of computers, and the development of the first effective method for solving LP problems, the Simplex Method. Devised by George Dantzig in the 1940's, the method remains today at the core of all commercial Linear Programming solvers. Solvers such as IBM's Cplex, Gurobi, or FICO's Xpress, to name a few, all implement some version of the method.

2.1 The Simplex Method: example

Consider the following LP problem

$$
\begin{array}{rrcrcr}
\max & 3x_1 & + & 2x_2 & & \\
s.t. & x_1 & + & x_2 & \leq & 6 \\
 & x_1 & + & 2x_2 & \leq & 10 \\
 & x_1 & - & x_2 & \leq & 4 \\
 & & & x_1, x_2 & \geq & 0
\end{array}
\tag{2.1}
$$

Let us first transform the problem in standard equality form:

$$
\begin{array}{rrcrcrcrcrcr}
\max & 3x_1 & + & 2x_2 & & & & & & & & \\
s.t. & x_1 & + & x_2 & + & x_3 & & & & & = & 6 \\
 & x_1 & + & 2x_2 & & & + & x_4 & & & = & 10 \\
 & x_1 & - & x_2 & & & & & + & x_5 & = & 4 \\
 & & & & & & & & x_1, x_2, x_3, x_4, x_5 & & \geq & 0
\end{array}
\tag{2.2}
$$

Note that there is an obvious feasible solution to start from, namely the point $x^* = (0, 0, 6, 10, 4)$ with objective function of 0. The solution x^* is extreme in $\mathbb{R}^5$ since all three resource constraints and the two non-negativity constraints for the variables x_1, x_2 in the objective are defining for x^* (they are independent and satisfied at equality). The equivalent solution $(0, 0) \in \mathbb{R}^2$ in the original problem (2.1) is also extreme.

Note that in (2.2) the non-basic variables x_1, x_2 are in the objective and the basic variables x_3, x_4, x_5 are written is terms of the non-basic ones, where the coefficient of the basic variables in the resource constraints is $+1$. This way the basic solution corresponding to (2.2) is given by setting to zero the non-basic variables, which gives the basic variables the value of the right hand side parameters. This layout of the LP is called a *dictionary*. And the obvious solution x that arises by setting the variables in the objective to 0 is called the *dictionary solution*.

If the original problem is in standard form with n variables and m resource constraints then in standard equality form it will have $n + m$ variables and m equality resource constraints. The dictionary solution x satisfies m resource constraints and n nonnegativity constraints at equality; let us assume that these are independent then the dictionary solution x is a basic point. We define a *basis* to be the set of the basic variables of the dictionary solution. Thus any variables not in the basis have to be set to 0 (the non-basic variables). If the problem is not degenerate then there is a one to one correspondence between a basis and a dictionary/ dictionary solution. The basis that corresponds to (2.2) is $B = \{3, 4, 5\}$. When the solution that corresponds to a basis or a dictionary is feasible then we call it a *feasible basis* and a *feasible dictionary*.

First iteration Going back to the example (2.2), can we find a better solution?

Note that, if we increase x_2 by $t \geq 0$ and leave $x_1 = 0$, the objective value increases by $2t$.

However, to insure that the resource constraints are satisfied, the remaining components must become

$$\begin{aligned} x_3(t) &= 6 - t \\ x_4(t) &= 10 - 2t \\ x_5(t) &= 4 + t \end{aligned}$$

We can only increase t as long as the above three variables take nonnegative values. What is the maximum value of t we can choose? In order to have $x_3(t) \geq 0$, we need $t \leq 6$. For $x_4(t) \geq 0$, we need $t \leq 5$. Note instead that $x_5(t) \geq 0$ for all nonnegative values of t. Thus the largest t we can choose is $t = 5$.

The new solution we obtain is therefore the point $(0, 5, 1, 0, 9)$. Note that there are still three effective constraints and two non-basic variables (namely, x_1, x_4). The solution is extreme. The value of the solution is $2 \cdot 5 = 10$. Here we say that x_2 is the entering variable, i.e. it enters the basis and from non-basic it becomes basic; and x_4 is the leaving variable, i.e. it leaves the basis and from basic it becomes non-basic. The new basis is $B = \{2, 3, 5\}$.

Second iteration Can we find a yet better solution? What we did before, was to increase the value of a nonbasic variable with positive coefficient in the objective function.

In (2.2) the objective and the basic variables were expressed in terms of the non-basic variables. So setting the non-basic variables to zero leaves the basic variables to take the value of the right hand side, where if positive then the solution is feasible. So when we increased the value of x_2, the basic variables x_3, x_4, x_5 were adjusted in value to maintain feasibility. But their change in value did not affect the objective since they do not appear in the objective. This guarantees that by increasing a non-basic variable with positive objective coefficient the new solution value we find will have better objective value.

So we would like only the nonbasic variables x_1, x_4 to appear in the objective function. To accomplish this, we resort to the following trick. Let us introduce a new variable z, and set

$$z = 3x_1 + 2x_2.$$

Now, the original problem can be stated as

$$\begin{array}{rrcrcrcrcrcl}
\max z & & & & & & & & & & & \\
z = & 3x_1 & + & 2x_2 & & & & & & & & \\
 & x_1 & + & x_2 & + & x_3 & & & & & = & 6 \\
 & x_1 & + & 2x_2 & & & + & x_4 & & & = & 10 \\
 & x_1 & - & x_2 & & & & & + & x_5 & = & 4 \\
 & & & & & & & & & x_1, x_2, x_3, x_4, x_5 & \geq & 0
\end{array}$$

Note that the second constraint of the original problem expresses x_4 in terms of the original non-basic variables x_1, x_2 and thus we can solve for the new basic variable x_2 in terms of the new nonbasic variables x_1, x_4, that is

$$x_2 = 5 - \frac{1}{2}x_1 - \frac{1}{2}x_4.$$

Substituting x_2 with $5 - \frac{1}{2}x_1 - \frac{1}{2}x_4$ in the equation defining z, we obtain

$$z = 10 + 2x_1 - x_4.$$

Therefore, if we increase the value of x_1 from 0 to t, while leaving the value of x_4 at 0, the next solution will have value $10 + 2t$, which is greater than the current one. We need to compute the values of the basic variables in order to insure that the next solution satisfies the resource constraints. In order to do this, we need to know how x_2, x_3, x_5 change as x_1 increases. Thus, we need to solve the basic variables in terms of the nonbasic one. This can be done by substituting x_2 with $5 - \frac{1}{2}x_1 - \frac{1}{2}x_4$ in all the constraints of the problem.

We obtain the following problem, equivalent to the previous one.

$$
\begin{array}{rlllllllll}
\max z & & & & & & & & & \\
z = 10+ & 2x_1 & & & & - & x_4 & & & \\
& \frac{1}{2}x_1 & & & + \; x_3 & - & \frac{1}{2}x_4 & & = & 1 \\
& \frac{1}{2}x_1 & + & x_2 & & + & \frac{1}{2}x_4 & & = & 5 \\
& \frac{3}{2}x_1 & & & & + & \frac{1}{2}x_4 & + \; x_5 & = & 9 \\
& & & & & & x_1, x_2, x_3, x_4, x_5 & & \geq & 0
\end{array}
\tag{2.3}
$$

Note that in (2.3) the basic variables x_2, x_3, x_4 are written with respect to the non-basic variables and the thus values of the basic variables are given by the right hand side values. This is the dictionary that corresponds to the dictionary solution $(0, 5, 1, 0, 9)$ and to the basis $B = \{2, 3, 5\}$.

If we increase x_1 by $t \geq 0$ and leave $x_4 = 0$, the objective value increases by $2t$.

The remaining components must become

$$
\begin{aligned}
x_2(t) &= 5 - \frac{1}{2}t \\
x_3(t) &= 1 - \frac{1}{2}t \\
x_5(t) &= 9 - \frac{3}{2}t
\end{aligned}
$$

The variables x_2, x_3, x_5 remain nonnegative as long as t satisfies, respectively, $t \leq 10$, $t \leq 2$, and $t \leq 6$. Thus the largest t for which the solution is nonnegative is $t = 2$. The new solution is $(2, 4, 0, 0, 6)$, with value 14. The variables x_1, x_2, x_5 are basic, while x_3, x_4 are nonbasic. In this iteration x_1 is the entering basic variable, while the variable x_3 is the leaving basic variable. The new basis is $B = \{1, 2, 5\}$.

Third iteration As before, we want to express z and the basic variables x_1, x_2, x_5 in terms of the nonbasic variables x_3, x_4. We can use the constraint $\frac{1}{2}x_1 + x_3 - \frac{1}{2}x_4 = 1$ to express x_1 in terms of x_3 and x_4, namely

$$x_1 = 2 - 2x_3 + x_4.$$

Substituting x_1 with $2 - 2x_3 + x_4$ in all constraints of the previous LP, we get

$$
\begin{array}{rllllllll}
\max z = 14 & & & - \; 4x_3 & + & x_4 & & & \\
x_1 & & & + \; 2x_3 & - & x_4 & & = & 2 \\
& + & x_2 & - \; x_3 & + & x_4 & & = & 4 \\
& & & - \; 3x_3 & + & 2x_4 & + \; x_5 & = & 6 \\
& & & & & x_1, x_2, x_3, x_4, x_5 & & \geq & 0
\end{array}
$$

The above dictionary corresponds to the dictionary solution $(2, 4, 0, 0, 6)$ and basis $B = \{1, 2, 5\}$.

If we increase x_4 by $t \geq 0$ and leave $x_3 = 0$, the objective value increases by t.

The remaining components must become

$$\begin{aligned} x_1(t) &= 2 + t \\ x_2(t) &= 4 - t \\ x_5(t) &= 6 - 2t \end{aligned}$$

The largest value of t for which the basic variables are nonnegative is $t = 3$. The new basic solution is $(5, 1, 0, 3, 0)$, with value 17 and the new basis is $B = \{1, 2, 4\}$.

Fourth iteration Substituting x_4 with $3+\frac{3}{2}x_3-\frac{1}{2}x_5$ in all constraints, we obtain

$$\begin{array}{rcccccccccl} \max z = 17 & & & & - & \frac{5}{2}x_3 & & & - & \frac{1}{2}x_5 & \\ & x_1 & & & + & \frac{1}{2}x_3 & & & + & \frac{1}{2}x_5 & = 5 \\ & & + & x_2 & + & \frac{1}{2}x_3 & & & - & \frac{1}{2}x_5 & = 1 \\ & & & & - & \frac{3}{2}x_3 & + & x_4 & + & \frac{1}{2}x_5 & = 3 \\ & & & & & & & & & x_1, x_2, x_3, x_4, x_5 & \geq 0 \end{array}$$

The above dictionary corresponds to the dictionary solution $(5, 1, 0, 3, 0)$ and the basis $B = \{1, 2, 4\}$.

Can there be a better solution? Since the coefficient of x_3 and x_5 in the objective function are negative, it is impossible to achieve a value better than 17. Therefore our current solution is optimal.

We have visited the following solutions:

$$(0, 0, 6, 10, 4) \to (0, 5, 1, 0, 9) \to (2, 4, 0, 0, 6) \to (5, 1, 0, 3, 0).$$

The components relative to the original variables x_1 and x_2 can be plotted in a diagram, shown in Figure 2.1.

2.2 The simplex method

In order to be applied, the simplex method requires the problem to be converted initially into standard equality form. Note that LP solvers accept LP problems in any form, so the user needs not be concerned with entering the problem in the appropriate form.

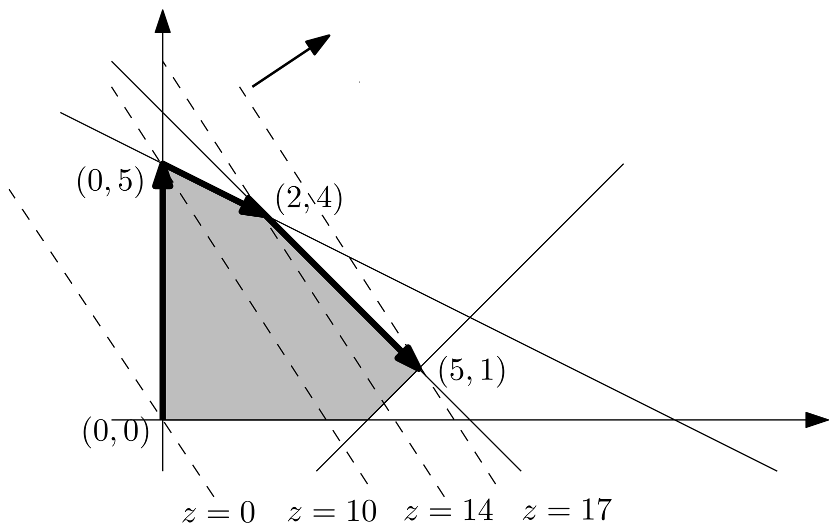

Figure 2.1: Sequence of points visited by the simplex.

Thus, we assume that we are dealing with an LP problem in the form

$$\begin{aligned} \max cx & \\ Ax = & \ b \\ x \geq & \ 0 \end{aligned} \tag{2.4}$$

where A is an $m \times (n+m)$ matrix, c is an $(n+m)$-dimensional row vector, b is an m-dimensional column vector, and x is a column vector of $n+m$ variables. We will also assume that the m resource constraints are independent. Since all resource constraints are equalities and they are independent, every basic solution will have m effective constraints and thus m basic variables and n nonbasic variables.

It will be convenient to write (2.4) in the following equivalent form:

$$\begin{aligned} \max z & \\ z = cx & \\ Ax = & \ b \\ x \geq & \ 0 \end{aligned} \tag{2.5}$$

The Simplex Method needs to be given an initial extreme solution to start form, say $\bar{x}$, and the corresponding set of basic variables (in the example in the previous section the initial extreme solution was the point $(0, 0, 6, 10, 4)$).

Let us assume that the basic variables are $x_{B[1]}, x_{B[2]} \ldots, x_{B[m]}$, belonging to basic B and let us denote by N the set of indices of the n nonbasic variables. In the previous example, for the extreme point $(0, 0, 6, 10, 4)$, $x_{B[1]} = x_3$, $x_{B[2]} = x_4$, $x_{B[3]} = x_5$ corresponding to basis $B = \{3, 4, 5\}$, and $N = \{1, 2\}$.

We solve the variables z, $x_{B[1]}, x_{B[2]} \ldots, x_{B[m]}$ in terms of the nonbasic variables, so that the problem is in dictionary form:

$$\begin{array}{lllcll} \max & z & & & & \\ & z = & \bar{z} & + \sum_{j \in N} \bar{c}_j x_j & & \\ & & x_{B[i]} & + \sum_{j \in N} \bar{a}_{ij} x_j & = & \bar{b}_i, \quad i = 1, \ldots, m \\ & & x & & \geq & 0. \end{array} \tag{2.6}$$

The extreme solution or dictionary solution corresponding to the above dictionary $\bar{x}$ is defined by

$$\begin{array}{rcl} \bar{x}_{B[i]} & = & \bar{b}_i; \quad i = 1, \ldots, m \\ \bar{x}_j & = & 0; \quad j \in N \end{array}$$

(so in particular $\bar{b}_1, \ldots, \bar{b}_m \geq 0$), and has objective value $\bar{z}$. The values $\bar{c}_j$ of the dictionary that corresponds to basis B are called *reduced costs* corresponding to basis B for the non-basic variables in N. The reduced costs corresponding to basis B for the basic variables are set to 0. Reduced costs have a nice intuitive interpretation: it is the marginal increase in the objective value when one of the non-basic variables (now set to 0) is increased.

There are two possible cases:

Case 1. There exists a nonbasic variable x_k such that $\bar{c}_k > 0$ (reduced cost is positive).

Ideally, we would like to find a new feasible basis whose objective value is better (or not worse) than $\bar{z}$. Note that, since $\bar{c}_k > 0$, if we increase the value of x_k from 0 to some number $t \geq 0$, while leaving at 0 the value of all other nonbasic variables, then the value of the new solution in the objective function will increase by $\bar{c}_j t$, thus improving the value of the objective function. However, in order to do so, we need to adjust the values of the basic variables to maintain feasibility.

More formally, for $t \geq 0$, the basic variables need to be defined as

$$x_{B[i]}(t) = \bar{b}_i - t\bar{a}_{ik}, \qquad i = 1, \ldots, m; \tag{2.7}$$

By construction, the new solution satisfies the resource constraints. We would like to increase the value of t as much as possible. The only thing preventing t from increasing indefinitely is the nonnegativity of the variables. Thus we can have two sub-cases

a) For some value of $t = t^*$, the value of some basic variable, say x_ℓ, goes to 0, while keeping the values of all other variables nonnegative. In this

case, the new extreme solution is obtained by replacing t^* into (2.7). The variable x_k becomes basic, while the variable x_ℓ become nonbasic, and we repeat.
(In the first iteration in the example in the previous section, $x_k = x_2$, $x_\ell = x_4$, $t^* = 5$, and the new solution is $(0, t^*, 6 - t^*, 0, 4 + t^*)$.)

b) The values of the variables remain nonnegative for any positive value of t.
In this case we can increase the value of t as much as we want, thus obtaining feasible solutions of arbitrarily large value. It follows that the problem is unbounded.

Case 2. $\bar{c}_j \leq 0$ for all indices $j \in N$.
This means that all coefficient in the objective function of the LP (2.6) are ≤ 0. So no solution can have value greater than $\bar{z}$. Since $\bar{x}$ has value $\bar{z}$, it must be optimal. Thus the algorithm stops.

The Simplex Method

Start from an extreme solution, with basic variables $x_{B[1]}, \ldots, x_{B[m]}$.

1. Write the LP in dictionary form (2.6) by expressing the variables $z, x_{B[1]}, \ldots, x_{B[m]}$ in terms of the nonbasic variables.
2. If $\bar{c}_j \leq 0$ for all $j \in N$, then the current solution is optimal, STOP.
3. Otherwise, pick a nonbasic variable x_k such that $\bar{c}_k > 0$.
   Compute the largest value t^* of t such that the solution (2.7) is feasible.

   3a. If $t^* < +\infty$, then some basic variable x_ℓ takes value 0. Compute the new extreme solution, and replace x_ℓ with x_k as basic variable. Return to 1.

   3b. If $t^* = +\infty$, then the problem is unbounded. STOP.

2.3 Issues that arise when using the simplex method

In this section we discuss problems that one might face when running the simplex method. For example, how do we pick an entering basic variable? What happens when the problem is degenerate? What if the problem is unbounded? What happens if there are multiple optima?

1. **Tie in the nonbasic entering variables** We did not specify how to choose which nonbasic variable to increase. In example (2.1), in the first iteration of the simplex algorithm we chose non-basic variable x_2 to become a basic variable, which had coefficient 2, but we could have also chosen non-basic variable x_1, which had a higher coefficient of value 3.

   If we chose the non-basic variable with the highest coefficient this will have the highest marginal increase in the objective value. However, this does not mean that it will result to the highest total increase in the objective since we might not be able to increase that nonbasic variable by a large amount whereas another basic variable with a smaller objective coefficient might result in a higher total increase in the objective.

2. **Degeneracy** When a problem is degenerate this might result in a basic variable taking the value of zero (a degenerate variable). This could be problematic because of the following. Suppose that one of these degenerate basic variables retains its value of zero until it is chosen at a subsequent iteration to be a leaving basic variable. But the corresponding entering basic variable cannot be increased in value without the leaving basic variable becoming negative. In this case the basis will change but the value of the objective will not change because we were unable to increase the value of the entering basic variable. If z remains the same in an iteration there is a chance that the simplex method may go around in a loop repeating the same sequence of solutions periodically without increasing the objective value.

   Fortunately, though a perpetual loop is possible in theory, it rarely occurs in practice. Further, one could always change the choice of the leaving basic variable to get out of the loop. Finally, special rules have been constructed for breaking ties of leaving basic variables so that such loops are always avoided.

   For an example of this please see exercises 3 and 4 and corresponding solutions.

3. **Multiple Optima** Some LPs have multiple optimal extreme solutions. When this happens the optimal dictionary will have a nonbasic variable with 0 coefficient. To find the other extreme optimal solutions you can increase the value of this nonbasic variable. See exercise 5 for an example.