2. LINEAR PROGRAMMING

2.1 The primal and dual problems

Constrained optimization problems where both the objective function and the the functional constraints are linear in $\boldsymbol{x}$ are known as **linear programming** problems; that is, when $f(\boldsymbol{x}) = \boldsymbol{c}^\top \boldsymbol{x}$ and $g(\boldsymbol{x}) = A\boldsymbol{x}$, where $\boldsymbol{c} \in \mathbb{R}^n$ and A is an $m \times n$ matrix. The problem can be one of either minimization or maximization, the functional constraints may be either inequalities or equalities and the set X is typically the non-negative orthant $X = \{\boldsymbol{x} : \boldsymbol{x} \geqslant 0\}$. To be specific, we will investigate the following general pair of problems:

$$\begin{array}{llll} \text{P:} & \text{maximize } \boldsymbol{c}^\top \boldsymbol{x} & \text{subject to} & A\boldsymbol{x} \leqslant \boldsymbol{b},\ \boldsymbol{x} \geqslant 0, \\ \text{D:} & \text{minimize } \boldsymbol{b}^\top \boldsymbol{\lambda} & \text{subject to} & A^\top \boldsymbol{\lambda} \geqslant \boldsymbol{c},\ \boldsymbol{\lambda} \geqslant 0; \end{array}$$

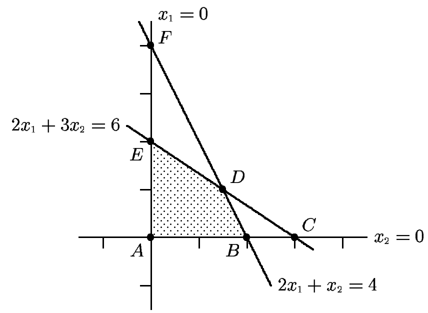
P will be referred to as the primal problem and D is its dual. The solutions of these two problems are intimately related in a way that we will explore in this chapter.

To make the situation concrete, we will carry through detailed consideration of the following particular example of these problems:

$$\begin{array}{lll} \text{P:} & \text{maximize} & 3x_1 + 2x_2 \\ & \text{subject to} & 2x_1 + x_2 \leqslant 4 \\ & & 2x_1 + 3x_2 \leqslant 6 \\ & & x_1,\ x_2 \geqslant 0 \end{array} \qquad \begin{array}{lll} \text{D:} & \text{minimize} & 4\lambda_1 + 6\lambda_2 \\ & \text{subject to} & 2\lambda_1 + 2\lambda_2 \geqslant 3 \\ & & \lambda_1 + 3\lambda_2 \geqslant 2 \\ & & \lambda_1,\ \lambda_2 \geqslant 0 \end{array}$$

We begin by plotting the feasible region for the problem P; that is the set of values x_1, x_2 which are feasible for the problem and we observe that it is the shaded region in the diagram bounded by the four lines $x_1 = 0$, $x_2 = 0$, $2x_1 + x_2 = 4$ and $2x_1 + 3x_2 = 6$.

We notice that the feasible region is a convex set, which is always true for a linear programming problem, and since it is bounded by lines (in higher dimensions the feasible region is bounded by hyperplanes) it is an example of what is known as a polyhedral set. We further observe that the contours of the objective function $f = 3x_1 + 2x_2 = \text{constant} = \alpha$, say, are straight lines so the linear programming problem is equivalent to finding the largest



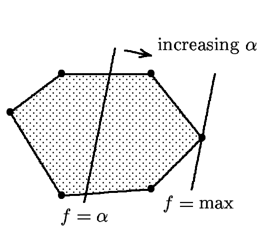
value of α for which these lines intersect the feasible set. It is immediately clear that, since the feasible region is bounded by lines, as we increase α the last time (assuming that there is a last time so that there is a finite maximum for the problem) the contour intersects the feasible set it will pass through at least one ‘corner’ or vertex of the feasible set. Thus if there is a finite optimal solution to a linear programming problem then there is a solution that occurs at a vertex; this statement is formalized in the Theorem below which will be proved later. The vertices are known as extreme points of the feasible set.

Definition A point $\boldsymbol{x} \in X$ is an **extreme point** of a convex set X if $\boldsymbol{x} = \theta\boldsymbol{y} + (1-\theta)\boldsymbol{z}$, for $\boldsymbol{y}, \boldsymbol{z} \in X$ and $0 < \theta < 1$, implies that $\boldsymbol{x} = \boldsymbol{y} = \boldsymbol{z}$.

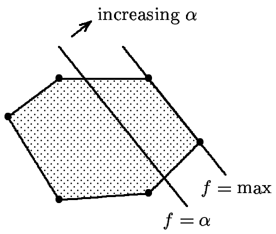
Theorem 2.1 *If a linear programming problem has a finite optimum then there is an optimal solution which is an extreme point of the feasible set.*

The diagram below shows the two situations, firstly where, as α increases, the contour of the objective function last intersects the feasible set in just one point, which is necessarily an extreme point, and secondly where the slope of the objective function is parallel to an edge of the feasible set (or face in higher dimensions); in this latter case there will be (infinitely) many solutions including the two extreme points which are the endpoints of the edge.

The upshot of this is that to solve a linear programming problem it is sufficient to search throught the extreme points of the feasible set to find the best one. There are only a finite number of extreme points, but the number increases very quickly with the size of the problem so that direct enumeration is not a viable approach to find a solution in



One optimal solution



Many optimal solutions

general.

Before we move on to explore the simplex algorithm, which is a highly effective algorithm that searches through the extreme points in an efficient way, we will confirm that D is indeed the dual of the problem P. If we write P in equality form, by adding slack variables $\boldsymbol{z} \geqslant 0$, we have the problem as:

$$\text{maximize} \quad \boldsymbol{c}^\top \boldsymbol{x} \quad \text{subject to} \quad A\boldsymbol{x} + \boldsymbol{z} = \boldsymbol{b},\ \boldsymbol{x} \geqslant 0,\ \boldsymbol{z} \geqslant 0;$$

its Lagrangian is

$$L(\boldsymbol{x}, \boldsymbol{z}, \lambda) = \boldsymbol{c}^\top \boldsymbol{x} + \boldsymbol{\lambda}^\top (\boldsymbol{b} - A\boldsymbol{x} - \boldsymbol{z}).$$

Maximizing in $\boldsymbol{z} \geqslant 0$, we see that for a finite maximum we require $\boldsymbol{\lambda} \geqslant 0$ and then at the maximum we will have $\boldsymbol{\lambda}^\top \boldsymbol{z} = 0$, which is complementary slackness. The remaining terms in L may be written as

$$L = \left(\boldsymbol{c} - A^\top \boldsymbol{\lambda}\right)^\top \boldsymbol{x} + \boldsymbol{b}^\top \boldsymbol{\lambda};$$

maximizing over $\boldsymbol{x} \geqslant 0$, for a finite maximum we require $A^\top \boldsymbol{\lambda} \geqslant \boldsymbol{c}$ and then we will have the further complementary slackness condition $\left(\boldsymbol{c} - A^\top \boldsymbol{\lambda}\right)^\top \boldsymbol{x} = 0$ at the maximum, which leaves

$$\max_{\boldsymbol{x} \geqslant 0, \boldsymbol{z} \geqslant 0} L(\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{\lambda}) = \boldsymbol{b}^\top \boldsymbol{\lambda} = h(\boldsymbol{\lambda}), \quad \text{the dual objective function.}$$

Hence the dual problem is to minimize $\boldsymbol{b}^\top \boldsymbol{\lambda}$ subject to the conditions that ensure a finite minimum, viz. $A^\top \boldsymbol{\lambda} \geqslant \boldsymbol{c}$ and $\boldsymbol{\lambda} \geqslant 0$. One can ask what is the dual of the dual problem? As an exercise, you should verify that here the dual of the dual is the primal problem (which is always the case for linear programming problems).

2.2 Basic solutions

First, we write the problem introduced in the last section in equation form:

$$\begin{aligned} \text{P:} \quad \text{maximize} \quad & f = 3x_1 + 2x_2 \\ \text{subject to} \quad & 2x_1 + x_2 + z_1 = 4 \\ & 2x_1 + 3x_2 + z_2 = 6 \\ & x_1,\ x_2 \geqslant 0,\ z_1,\ z_2 \geqslant 0. \end{aligned}$$

Now we consider the values taken on by the variables at each of the six points A, B, C, D, E and F in the diagram of the feasible set for the problem.

	x_1	x_2	z_1	z_2	f
A	0	0	4	6	0
B	2	0	0	2	6
C	3	0	-2	0	9
D	$\frac{3}{2}$	1	0	0	$\frac{13}{2}$
E	0	2	2	0	4
F	0	4	0	-6	8

The first thing to notice is that in each case two of the variables are zero with the other two being non-zero; this is clear from the picture since each point is determined by the intersection of two of the lines $x_1 = 0$, $x_2 = 0$, $2x_1 + x_2 = 4$ and $2x_1 + 3x_2 = 6$, but these last two are also $z_1 = 0$ and $z_2 = 0$. Another way of thinking about these points is that we consider the two constraints

$$\begin{aligned} 2x_1 + x_2 + z_1 &= 4 \\ 2x_1 + 3x_2 + z_2 &= 6 \end{aligned}$$

and then in turn we take two of the variables, set them to zero and solve for the remaining two; these are known as the basic solutions of these two linear equations.

To accord with later practice we will relabel the slack variables z_1 and z_2 as x_3 and x_4 so that the problem becomes

$$\begin{aligned} \text{P:} \quad \text{maximize} \quad & f = 3x_1 + 2x_2 \\ \text{subject to} \quad & 2x_1 + x_2 + x_3 = 4 \\ & 2x_1 + 3x_2 + x_4 = 6 \\ & x_1,\ x_2,\ x_3,\ x_4 \geqslant 0, \end{aligned}$$

so that it is in what is known as **standard form**:

$$\text{maximize} \quad \boldsymbol{c}^\top \boldsymbol{x} \quad \text{subject to} \quad A\boldsymbol{x} = \boldsymbol{b},\ \boldsymbol{x} \geqslant 0.$$

For this example, the matrix

$$A = \begin{pmatrix} 2 & 1 & 1 & 0 \\ 2 & 3 & 0 & 1 \end{pmatrix}.$$

In general, we will assume that A has m rows and n columns, $n > m$, and that the rank of the matrix is m; we will further assume that any set of m columns of A is linearly independent.

Definition A **basic** solution, $\boldsymbol{x}$, of the system of equations $A\boldsymbol{x} = \boldsymbol{b}$ is a solution for which at least $n - m$ entries of $\boldsymbol{x}$ are zero.

Since any m columns of A form a non-singular matrix by our assumption above, any basic solution may be obtained by setting $n-m$ components of $\boldsymbol{x}$ equal to zero and solving uniquely for the remaining m variables. The $n-m$ variables set equal to zero are known as the **non-basic variables** of the basic solution, the remaining m variables are the **basic variables.**

Definition A **basic feasible solution**, $\boldsymbol{x}$, of the linear programming problem (in standard form) is a basic solution of the equations $A\boldsymbol{x} = \boldsymbol{b}$ for which $\boldsymbol{x} \geqslant 0$.

If a basic feasible solution has one, or more, basic variables equal to zero it is said to be **degenerate**. The problems that we will deal with will be assumed to be non-degenerate in that they have no degenerate basic feasible solutions; we will discuss the relevance of degeneracy to the progress of the simplex algorithm later on.

In the particular example we are considering the six points $A - F$ are all basic solutions but only A, B, D and E are basic feasible solutions. The importance of these ideas comes from the following equivalence, which we will prove later but which is clear in the case of our example.

Theorem 2.2 *A point $\boldsymbol{x}$ is a basic feasible solution of $A\boldsymbol{x} = \boldsymbol{b}$ if and only if it is an extreme point of the feasible set $\{\boldsymbol{x} : A\boldsymbol{x} = \boldsymbol{b},\ \boldsymbol{x} \geqslant 0\}$ for the linear programming problem.*

This result, which will be proved later, gives us an algebraic characterization of extreme points of the feasible set and when combined with our previous observation in Theorem 2.1 it gives the result which is the basis of the simplex algorithm.

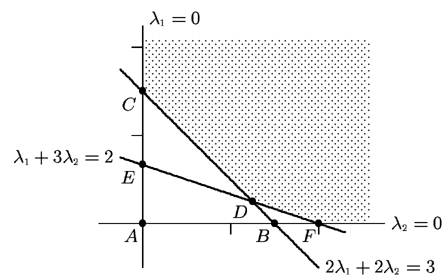
Theorem 2.3 *If the linear programming problem has a finite optimum then there is an optimal solution which is a basic feasible solution.*

2.3 Relationship between the primal and dual

Add slack variables to the dual problem so that it becomes a problem with constraints in equation form:

$$\begin{aligned} \text{D:} \quad \text{minimize} \quad & 4\lambda_1 + 6\lambda_2 \\ \text{subject to} \quad & 2\lambda_1 + 2\lambda_2 - v_1 \phantom{{}-v_2} = 3 \\ & \lambda_1 + 3\lambda_2 \phantom{{}-v_1} - v_2 = 2 \\ & \lambda_1,\ \lambda_2 \geqslant 0,\ v_1,\ v_2 \geqslant 0. \end{aligned}$$

We can plot the constraints and the feasible set for the dual problem in the λ_1-λ_2 plane, and we will observe that the basic solutions for the dual problem can be paired with those for the primal problem.



The basic solutions for D are obtained as for the primal problem by taking the intersections in pairs of the constraint $2\lambda_1 + 2\lambda_2 = 3$, $\lambda_1 + 3\lambda_2 = 2$, $\lambda_1 = 0$ and $\lambda_2 = 0$; they are paired up with the basic solutions to the primal by seeing that complementary slackness holds, that is $\boldsymbol{\lambda}^\top \boldsymbol{z} = 0$ and $\boldsymbol{v}^\top \boldsymbol{x} = \left(\boldsymbol{c} - A^\top \boldsymbol{\lambda}\right)^\top \boldsymbol{x} = 0$, or

$$\lambda_1 x_3 = 0 = \lambda_2 x_4 \quad \text{and} \quad v_1 x_1 = 0 = v_2 x_2,$$

and the fact that the values of the respective objective functions f and h for the primal and dual are the same at the paired basic solutions. These relationships hold for all pairs of primal and dual linear programming problems. As we mentioned above, there are four basic feasible solutions for the primal problem, A, B, D, and E and you should observe that there are three basic feasible solutions for the dual, C, D and F; there is only one basic solution which is feasible for both the primal and dual problem, viz. D, which is the optimal solution to both problems. You should further note that the value of the dual objective function at any basic feasible solution to the dual is greater than or equal to the value of the primal objective function at any basic feasible solution to the primal (which we know from weak duality) and furthermore, at the common optimal solution D the values of the two objective functions are the same ('the two problems have the same value').

P:

	x_1	x_2	x_3	x_4	f
A	0	0	4	6	0
B	2	0	0	2	6
C	3	0	-2	0	9
D	$\frac{3}{2}$	1	0	0	$\frac{13}{2}$
E	0	2	2	0	4
F	0	4	0	-6	8

D:

	v_1	v_2	λ_1	λ_2	h
A	-3	-2	0	0	0
B	0	$-\frac{1}{2}$	$\frac{3}{2}$	0	6
C	0	$\frac{5}{2}$	0	$\frac{3}{2}$	9
D	0	0	$\frac{5}{4}$	$\frac{1}{4}$	$\frac{13}{2}$
E	$-\frac{5}{3}$	0	0	$\frac{2}{3}$	4
F	1	0	2	0	8

These observations are special cases of the general relationship between the two problems:

$$\begin{array}{llll} \text{P:} & \text{maximize } \boldsymbol{c}^\top \boldsymbol{x} & \text{subject to} & A\boldsymbol{x} \leqslant \boldsymbol{b},\ \boldsymbol{x} \geqslant 0, \\ \text{D:} & \text{minimize } \boldsymbol{b}^\top \boldsymbol{\lambda} & \text{subject to} & A^\top \boldsymbol{\lambda} \geqslant \boldsymbol{c},\ \boldsymbol{\lambda} \geqslant 0. \end{array}$$

We have encountered weak duality in the context of general constrained problems and here it says that if $\boldsymbol{x}$ is a feasible solution to P and $\boldsymbol{\lambda}$ is a feasible solution to D then $\boldsymbol{c}^\top \boldsymbol{x} \leqslant \boldsymbol{b}^\top \boldsymbol{\lambda}$. The most important relationship between the two problems characterizes the necessary and sufficient conditons for optimality for a point to be an optimal solution of P.

Theorem 2.4 Necessary and sufficient conditions for optimality *A vector $\boldsymbol{x}$ is an optimal solution to the problem P if and only if there exists a vector $\boldsymbol{\lambda}$ such that the pair $\boldsymbol{x}$ and $\boldsymbol{\lambda}$ satisfy:*

$$\begin{array}{lll} \text{(i)} & A\boldsymbol{x} \leqslant \boldsymbol{b},\ \boldsymbol{x} \geqslant 0 & \text{(primal feasibility)} \\ \text{(ii)} & A^\top \boldsymbol{\lambda} \geqslant \boldsymbol{c},\ \lambda \geqslant 0 & \text{(dual feasibility)} \\ \text{(iii)} & \boldsymbol{\lambda}^\top (\boldsymbol{b} - A\boldsymbol{x}) = 0 = \boldsymbol{x}^\top \left(A^\top \boldsymbol{\lambda} - \boldsymbol{c}\right); & \text{(complementary slackness)} \end{array}$$

furthermore, if condition (iii) *holds, then $\boldsymbol{c}^\top \boldsymbol{x} = \boldsymbol{b}^\top \boldsymbol{\lambda}$.*

Proof. We will only establish the sufficiency here (since it is essentially the Lagrangian Sufficiency Theorem it drops out as easily as that result); the necessity is harder and we will discuss it further later on. Suppose that $\boldsymbol{x}$ is primal feasible and $\boldsymbol{\lambda}$ is dual feasible, then

$$\boldsymbol{c}^\top \boldsymbol{x} \leqslant \boldsymbol{c}^\top \boldsymbol{x} + \underbrace{\boldsymbol{\lambda}^\top (\boldsymbol{b} - A\boldsymbol{x})}_{\geqslant 0} = \underbrace{\boldsymbol{x}^\top \left(\boldsymbol{c} - A^\top \boldsymbol{\lambda}\right)}_{\leqslant 0} + \boldsymbol{b}^\top \boldsymbol{\lambda} \leqslant \boldsymbol{b}^\top \boldsymbol{\lambda};$$

this statement is weak duality, with the addition of complementary slackness the two inequalities are replaced by equalities and the sufficiency is established. □

Remark

Although the primal problem was formulated here with inequality constraints and sign constraints on the primal variables any formulation of a primal linear programming problem (equality constraints or mixed inequality-equality constraints, with or without sign constraints on the variables) leads to the same three necessary and sufficient conditions: primal feasibility, dual feasibility and complementary slackness. The form that these conditions take will depend on the formulation of the primal in each case.