8 Algebra of Linear Programming

8.1 Basic feasible solutions and extreme points

To prove the equivalence of basic feasible solutions and extreme points which was stated in Theorem 2.1 we need to recall some results on the solution of simultaneous linear equations.

1. If A is a $m \times m$ invertible matrix, $b \in \mathbb{R}^m$, then $Ax = b$ has a unique solution, $x = A^{-1}b$.

2. If A is a $m \times n$ with $n > m$ and no row of A is a linear combination of other rows (so there are no redundant equations or inconsistent equations, i.e., $\operatorname{rank}(A) = m$) then there are (a) infinitely many solutions to $Ax = b$, and (b) infinitely many solutions to $Ay = 0$.

We will also find it convenient to divide the variables into two sets, and to split the matrix A accordingly. For example, given

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

we can partition the variables into two disjoint subsets (basic and non-basic) $B = \{1, 2\}$ and $N = \{3\}$ and rewrite the equation

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} a_{13} \\ a_{23} \end{pmatrix} x_3 = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

or

$$A_B x_B + A_N x_N = b,$$

where $x_B = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ contains the variables in B and $x_N = \begin{pmatrix} x_3 \end{pmatrix}$ contains the variables in N, and A_B has the columns of A corresponding to variables in B (columns 1 and 2) and A_N has columns corresponding to variables from N (column 3).

You should convince yourself that the two forms of the linear equations are equivalent and that the same trick would work for general $m \times n$ matrices A and partition of variables into two sets. If $A = (a_1, \ldots, a_n)$, where a_i is the ith column, then

$$Ax = \sum_{i \in B} x_i a_i + \sum_{i \in N} x_i a_i = A_B x_B + A_N x_N = b.$$

We usually choose B to have m elements (a basis) and N to have $n - m$ elements. Then setting $x_N = 0$, we solve $A_B x_B = b$ where A_B is an $m \times m$ matrix to find the basic solution $x_B = A_B^{-1}b$, $x_N = 0$. We also make the following assumptions

Non-degeneracy assumptions for A.
(a) The matrix A has linearly independent rows (i.e., rank$(A) = m$), and
(b) Any $m \times m$ matrix formed from m columns of A is non-singular (i.e., any set of m columns from A is linearly independent).

Non-degeneracy of basic solutions.
(c) All basic solutions to $x_B = A_B^{-1}b$, $x_N = 0$ have exactly m non-zero coordinates, i.e., $x_i \neq 0$ for $i \in B$.

Remarks.
(a) is a reasonable assumption for small problems such as we can do by hand, though it may be difficult to check if a large set of equations is inconsistent or contains redundancy. Assumptions (b) and (c) are just made to simplify the theory and there is no trouble in solving LPs when A does not satisfy these. For (b) there may be no solution corresponding to some choice of basis (e.g., the equations obtained by adding slack variables to $x_1 \leq 2$, $x_2 \leq 3$), but the simplex algorithm still works perfectly well. Or there may be many (e.g., $x_1 + x_2 + z_1 = 1$, $2x_1 + 2x_2 + z_2 = 2$) and here there are solutions such as $x_1 = x_2 = 1/2$, $z_1 = z_2 = 0$ which are not extreme points. The answer here is to alter the definition of basic solution to exclude those obtained when A_B is not invertible. You will see on the examples sheet that the failure of (c) to hold causes no trouble.
The proof which follows can be adapted to work if A does not satisfy the non-degeneracy assumptions as stated above.

Lemma 8.1 *If $x \in X_b$ and $x = \lambda y + (1-\lambda)z$ for $y, z \in X_b$, $y \neq z$, $0 < \lambda < 1$, (i.e., x is not extreme) then $x_i = 0$ implies $y_i = z_i = 0$.*

Proof. $x_i = \lambda y_i + (1-\lambda)z_i$. Suppose $x_i = 0$. Then since $\lambda, 1-\lambda, y_i, z_i$ are all ≥ 0, we must have $y_i = z_i = 0$. ■
This result says that if x is on some face or edge of the simplex X_b then y and z must be in the same face or edge.

Lemma 8.2 *If x is a basic feasible solution then x is an extreme point of X_b.*

Proof. If x is a b.f.s. then $\exists$ a basis B and non-basis N such that the non-basic components of x satisfy $x_N = 0$. Suppose x is not extreme. Then $\exists$ y and z such that x lies on the line segment between y and z and hence $y_N = z_N = 0$. Furthermore, y and z are feasible, so $A_B y_B = A_B z_B = b$ and this gives $A_B(y_B - z_B) = 0$. But as A_B is non-singular (by assumption (b)) the unique solution to $A_B v_B = 0$ is $v_B = 0$. But this implies $y_B = z_B$, and hence $y = z$. But this is a contradiction. So x is extreme.
■

Lemma 8.3 *If x is an extreme point of X_b then x is a basic feasible solution.*

Proof. Since $x \in X_b$ we know it is feasible. We just need to show it is basic. Again, we argue by contradiction.

Suppose x is not a b.f.s. Then the number of non-zero coordinates, say p, is greater than m. (If the number is equal to m the solution is basic, and we can't have less than m by the non-degeneracy assumption.) Intuitively, x is not at a vertex, but is in some face, or edge specified by setting $n-p$ variables to zero.

Let $P = \{i : x_i > 0\}$ and $Q = \{i : x_i = 0\}$. Since x is feasible, $Ax = A_P x_P + A_Q x_Q = b \Longrightarrow A_P x_P = b$. But this is m equations in $p > m$ variables, so $\exists$ non-zero y_P s.t. $A_P y_P = 0$. We put $y_Q = 0$ and $y = \begin{pmatrix} y_P \\ y_Q \end{pmatrix}$.

Consider $x + \epsilon y$ and $x - \epsilon y$ for small ϵ. For $\epsilon > 0$ small enough these two points are both feasible since $A(x \pm \epsilon y) = Ax \pm \epsilon Ay = b$, and $x \pm \epsilon y \geq 0$ (for small ϵ since $y_i > 0$ implies $x_i > 0$). Hence

$$x = \tfrac{1}{2}(x + \epsilon y) + \tfrac{1}{2}(x - \epsilon y),$$

and so x is not extreme. ∎

Corollary 8.4 *x is an extreme point of X_b if and only if x is a basic feasible solution of $Ax = b$. (This is a restatement of Theorem 2.1.)*

We now need to prove that finite optimums occur at extreme points. The proof of this result is very like the proof of Lemma 8.3. The intuition is that if x is optimal and lies on a face (but not a vertex) then the whole face is optimal. So we construct a line segment through x which lies in the face (as in Lemma 8.3) and all solutions on this line segment are optimal. We extend the line segment till we come to an edge (or a face of one less dimension) and now have an optimal solution with one more zero coordinate. Continue until you end at an optimal vector.

Lemma 8.5 *Suppose minimize $c^\top x$ s.t. $x \in X_b$ has a finite optimal solution. Then there exists an extreme point which is optimal. (This gives Theorem 2.2.)*

Proof. Suppose x is an optimal solution with the smallest number of non-zero components, but is not extreme. Then $\exists\ y, z$ such that x lies on the line segment between y and z. Since x is optimal, $c^\top x \leq c^\top y$ and $c^\top x \leq c^\top z$. But $c^\top x = \lambda c^\top y + (1-\lambda) c^\top z$ and hence $c^\top x = c^\top y = c^\top z$. Thus y and z are also optimal.

Now consider the line segment $x(\epsilon) = x + \epsilon(y - z)$ for small ϵ. Note that $c^\top x(\epsilon) = c^\top x$. So all points on the line are optimal. Furthermore, $x_i = 0 \Longrightarrow x(\epsilon)_i = 0$ for all ϵ and $x_i > 0 \Longrightarrow x(\epsilon)_i > 0$ for small ϵ. It is clear that by choosing ϵ large enough and of the right sign we can force at least one more component of $x(\epsilon)$ to become zero.

Now since x is not extreme, the number of non-zero components is $p > m$. But we have now constructed $x(\epsilon)$ which is optimal, feasible and has $p-1$ non-zero components. This is a contradiction. So an optimal solution with the smallest possible number of non-zero components must be extreme. ■

8.2 Algebra of the simplex method

Let us take problem P in the form

$$\text{P: } \max c^\top x \quad \text{s.t. } Ax = b, \ x \geq 0.$$

Given a choice of basis B we can rewrite the problem as above

$$\max \{c_B^\top x_B + c_N^\top x_N\} \quad \text{s.t. } A_B x_B + A_N x_N = b, \ x_B, x_N \geq 0.$$

At this stage we are just rewriting the equations in terms of the two sets of variables x_B and x_N. The equations hold for any feasible x. Now A_B is invertible by our non-degeneracy assumptions. Thus we can write

$$x_B = A_B^{-1} b - A_B^{-1} A_N x_N, \tag{1}$$

and

$$\begin{aligned} f &= c_B^\top x_B + c_N^\top x_N \\ &= c_B^\top (A_B^{-1} b - A_B^{-1} A_N x_N) + c_N^\top x_N \\ &= c_B^\top A_B^{-1} b + (c_N^\top - c_B^\top A_B^{-1} A_N) x_N \end{aligned} \tag{2}$$

Equation (1) gives the constraints in a form where x_B appears with an identity matrix and (2) gives the objective function in terms of the non-basic variables. Thus the tableau corresponding to basis B will contain (after appropriate rearrangement of rows and columns)

basic	non-basic	
I	$A_B^{-1} A_N$	$A_B^{-1} b$
0	$c_N^\top - c_B^\top A_B^{-1} A_N$	$-c_B^\top A_B^{-1} b$

Thus, given a basis (choice of variables) we can work out the appropriate tableau by inverting A_B. Note that for many choices of B we will find that $A_B^{-1} b$ has negative components and so the basic solution $x_B = A_B^{-1} b$ is not feasible; we need the simplex algorithm to tell which Bs to look at.

In programming a computer to implement the simplex algorithm you need only remember what your current basis is, since the whole tableau can then be computed from A_B^{-1}. The **revised simplex algorithm** works this way and employs various tricks to compute the inverse of the new A_B from the old A_B^{-1} (using the fact that only one variable enters and one leaves). This can be very much more efficient.

Now we know that the simplex algorithm terminates at an optimal feasible basis when the coefficients of the objective row are all ≤ 0. In other words, there is a basis B for which

$$c_N^\top - c_B^\top A_B^{-1} A_N \leq 0.$$

Recall the dual problem is

$$\text{D: } \min \lambda^\top b \quad \text{s.t. } A^\top \lambda \geq c.$$

Let us write $\lambda = (A_B^{-1})^\top c_B$. Then we have $A_B^\top \lambda = c_B$ and $A_N^\top \lambda \geq c_N$, and hence λ is feasible. Furthermore, $x_B = A_B^{-1} b$, $x_N = 0$ is a basic solution for P and complementary slackness is satisfied since

$$c_B - A_B^\top \lambda = 0 \Longrightarrow (c_B - A_B^\top \lambda)^\top x_B = 0,$$
$$x_N = 0 \Longrightarrow (c_N - A_N^\top \lambda)^\top x_N = 0.$$

Consequently, $x_B = A_B^{-1} b$, $x_N = 0$ and $\lambda = (A^{-1})^\top c_B$ are respectively optimal for the primal and dual. We also have that with these solutions

$$f = c_B^\top x_B = c_B^\top A_B^{-1} b = \lambda^\top b.$$

So we have a proof of Theorem 2.5, that the primal and dual have the same objective value (if we accept that the simplex algorithm terminates at an optimum with ≤ 0 objective row) for the case of LP problems.

Remark

We have shown that in general the objective row of the final (optimal) tableau will contain $c_N^\top - \lambda^\top A_N$, where λ are dual variables. This is consistent with our observation that the final tableau contains $-\lambda$ when we start with a primal $Ax \leq b$ and add slack variables z. In this case the c_i corresponding to the slack variables are zero and columns of A_N corresponding to original variables are the columns of an identity matrix. So the part of the objective row beneath the original slack variables will contain $-\lambda^\top$, and λ are the dual variables corresponding to the primal constraints. The rest of the objective row, beneath the original x, contains $c_B^\top - \lambda^\top A_B$, i.e., the slack variables for the dual problem. This is what we observed in our earlier example.