OPTIMISATION 3: NOTES ON THE SIMPLEX ALGORITHM

Abstract. These notes give a summary of the essential ideas and results. It is not a complete account; see **Winston** Chapters 4, 5 and 6. The conventions and notation may be slightly different from those used in Opt 2. We introduce MAPLE's `simplex` package.

1. Fundamental Ideas

A linear program (abbreviated LP; also called "linear programming problem") is a maximisation or minimisation problem where the objective function and constraints are all linear.

A LP can be expressed in many different forms (see Problem Sheet 1, Q. 3), and different authors use words such as standard, basic, canonical in different senses. By and large, we shall follow **Winslow**'s practice and terminology.

Notation (vector notation). We use the column notation

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

to specify a vector in $\mathbb{R}^n$; in other words, x is viewed as an $n \times 1$ matrix. Then

$$x^T = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} .$$

The most common types of LP that we shall encounter are the so-called *normal* (max or min) problems defined below.

Definition 1.1. The LP

$$\text{maximise } z = c^T x \quad \text{subject to} \quad Ax \le b \text{ and } x \ge 0\,, \tag{1.1}$$

is called a *normal max problem.*

The LP

$$\text{minimise } z = c^T x = c^T x \quad \text{subject to} \quad Ax \ge b \text{ and } x \ge 0\,, \tag{1.2}$$

is called a *normal min problem.* Here, $c \in \mathbb{R}^n$, the $m \times n$ matrix A, and $b \in \mathbb{R}^m$ are given, whilst $x \in \mathbb{R}^n$ is the vector of unknowns.

We will usually write "$\max z = c^T x$" instead of "maximise $z := c^T x$". and "$\min z = c^T x$" instead of "minimise $z := c^T x$". Some conventions and terminology:

- x is the vector of *decision variables.* z is the objective function and c is the vector of objective coefficients.
- The inequality operators $\ge$ and $\le$ apply entrywise, i.e. $x \ge 0$ means
$$\forall\, 1 \le i \le n, \quad x_i \ge 0\,.$$
- The conditions $x \ge 0$ and $Ax \le b$ or $Ax \ge b$ are called the *constraints.*

There are many other kinds of LP involving any mixture of $\leq$, $\geq$ and $=$ type constraints. In order to develop a general theory and some solution algorithms, we need a notation that encompasses all these different forms of LP. A LP is in *standard form* when it is expressed as follows: all variables are constrained to be ≥ 0; all the other constraints, called "structural constraints", are linear equations. Every LP can be put into standard form by using slack (and excess) variables and other techniques. In other words, we can always express the LP as

$$\max z = c^T x \quad \text{subject to} \quad Ax = b \text{ and } x \geq 0\,. \tag{1.3}$$

Example 1.1. *Here is a typical LP: maximise $z = 2x_1 + 5x_2$ subject to*

$$\begin{aligned} 2x_1 - 3x_2 &\geq -6 \\ 7x_1 - 2x_2 &\leq 14 \\ x_1 + x_2 &\leq 5 \\ x_1, x_2 &\geq 0. \end{aligned}$$

We introduce the excess variable x_3 and the slack variables x_4 and x_5 defined by

$$\begin{aligned} -2x_1 + 3x_2 + x_3 &= 6 \\ 7x_1 - 2x_2 + x_4 &= 14 \\ x_1 + x_2 + x_5 &= 5 \end{aligned}$$

So we have expressed the LP in the form (1.3) with $n = 5$, $m = 3$,

$$A = \begin{pmatrix} -2 & 3 & 1 & 0 & 0 \\ 7 & -2 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} 6 \\ 14 \\ 5 \end{pmatrix} \quad \text{and} \quad c = \begin{pmatrix} 2 \\ 5 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Winston *would write the vector x in the form*

$$x = \begin{pmatrix} x_1 \\ x_2 \\ e_1 \\ s_1 \\ s_2 \end{pmatrix}$$

where e_1 is the excess and the s_i are the slack variables.

Suppose that we have converted the LP to the standard form (1.3), where A is of rank m. A solution x of $Ax = b$ is said to be *feasible* if it is in $\mathbb{R}^n_+$. Typically, $m < n$, and so the linear system $Ax = b$ will not determine all of the x_n. We can supplement the system with additional equations by setting $n - m$ of the x_i to zero; a *basic solution* is any vector x satisfying $Ax = b$ obtained in this way. The $n - m$ entries of the variables x_i that are set to zero are called the *nonbasic variables* whereas the remaining m are called the *basic variables.* The set of the basic variables is called a *basis.* Different choices of basis generally yield different basic solutions, and there can be up to

$$\binom{n}{m} = \frac{n!}{(n-m)!m!}$$

basic solutions, although these need not all be feasible. We shall see that, if a LP has an optimal solution, then there is a basic solution that is optimal. There may be more than one optimal solution (in which case there are nonbasic optimal solutions), or there may be no feasible solutions at all (*infeasible* LP), or there may be feasible solutions but no optimal solution (*unbounded* LP).

The *value* of a LP is the optimal value of the objective function.

Example 1.2. *Consider Example 1.1. Here* $m = 3$ *and* $n = 5$*, and so there may be as many as* 10 *basic solutions. We shall compute only three of them, by choosing two nonbasic variables with indices in* $\{3, 4, 5\}$.

(1) *Set* $x_3 = 0$ *and* $x_4 = 0$. *Then* $x_1 = 54/17$, $x_2 = 70/17$ *and so* $x_5 = -39/17$*; this solution is not feasible.*
(2) *Set* $x_3 = 0$ *and* $x_5 = 0$. *Then* $x_1 = 9/5$, $x_2 = 16/5$ *and so* $x_4 = 39/5$*; this solution is feasible and yields the objective value* $98/5$.
(3) *Set* $x_4 = 0$ *and* $x_5 = 0$. *Then* $x_1 = 8/3$, $x_2 = 7/3$ *and so* $x_3 = 13/3$*; this solution is also feasible and yields the objective value* 17.

It turns out that the second choice yields the feasible solution with the greatest value of the objective function; it is an optimal solution and the value of the LP is $98/5$.

2. Geometrical Viewpoint

Definition 2.1. A set S is said to be *convex* if, for every $x, y \in S$ and every $\lambda \in (0, 1)$, we have

$$\lambda x + (1 - \lambda)y \in S\,.$$

The set of vectors satisfying any given linear constraint is obviously convex, and the intersection of two convex sets is also a convex set. It follows that the set of all feasible vectors for an LP— called the *feasible region* of the LP— is a convex region in $\mathbb{R}^n_+$ where n is the number of variables. The problem is to find a point of the feasible region where the objective function is maximum or minimum. The fundamental region, say S, may be bounded (meaning $\exists\, C$ such that $\forall\, x \in S$, $\|x\| \le C$) or unbounded (meaning $\forall\, C\ \exists x \in S$ such that $\|x\| > L$). If the feasible region is bounded and nonempty, then the LP has an optimal solution. If the feasible region is unbounded, then the LP may be bounded or unbounded.

Definition 2.2. Let S be a convex set. We say that a point $x \in S$ is a *vertex* (or *extreme point*) of S if there exists no pair $(u, v) \in S \times S$ such that $u \neq v$ and

$$x = \frac{u + v}{2}\,.$$

Geometrically, each constraint defines a hyperplane such that any feasible vector must lie on the same side of the hyperplane as the origin. So the feasible region is a region bounded by hyperplanes (the geometrical expression of the constraints); hence it is a convex polyhedron (a geometrical figure whose surface consists of a number of flat portions connected by edges and vertices) which may extend to infinity.

The objective function has a gradient independent of x and so it cannot have a local extremum inside the feasible region. If it has a max or min value, it occurs at the boundary of the feasible region.

Theorem 2.1. *If the LP has an optimal solution, then there is an optimal solution that is a vertex.*

Theorem 2.2. *A feasible solution is a vertex if and only if it is a feasible basic solution.*

Thus if the LP has an optimal solution, then there is a basic solution that is optimal. The upshot is that the optimal solution should be sought in the set of the basic solutions. This set is finite, but large, and grows factorially with $n-m$; hence it is not practical to merely compute *every* basic solution and choose the one that yields the largest value of the objective function. The *simplex method* is an iterative algorithm for searching in the set of basic solutions, so as to lead to an optimal basic solution efficiently if one exists.

3. Tableau Notation

Given a LP in standard form, there are various ways of writing a basic solution as a *tableau.* We shall follow **Winston**.

Example 3.1. *Consider the LP*

$$\begin{aligned} \max z &= 2x_1 + 5x_2 \\ -2x_1 + 3x_2 + x_3 &= 6 \\ 7x_1 - 2x_2 + x_4 &= 14 \\ x_1 + x_2 + x_5 &= 5 \end{aligned}$$

We obtain a basic feasible solution by setting $x_1 = x_2 = 0$*, so that* $x_3 = 6$*,* $x_4 = 14$ *and* $x_5 = 5$*. We write this in tableau form as*

z	x_1	x_2	x_3	x_4	x_5	*rhs*	BV
1	-2	-5	0	0	0	0	$z=0$
0	-2	3	1	0	0	6	$x_3=6$
0	7	-2	0	1	0	14	$x_4=14$
0	1	1	0	0	1	5	$x_5=5$

Winston's form of the simplex tableau.

Here, BV stands for basic variable*; the column specifies a basic solution such that any variable not appearing in the column is implicitly assumed to have the value* 0*.*

The tableau is said to be in *canonical form* if, for every row, the entry corresponding to the basic variable is unity, and all the other entries in the column vanish. We can always go from a basic form to a canonical form by performing elementary row operations.

4. Phase I of the Simplex Method

Phase II of the simplex method starts from a basic feasible solution and *either* finds an optimal solution *or* shows that the LP is infeasible or unbounded.

Phase I is concerned with finding a first basic feasible solution; this can be difficult, if m is large. When a basic feasible vector is found, Phase II can begin. There are various systematic ways of doing Phase I. Here is one way (**Winston** p. 178):

- Introduce an artificial variable for each constraint which is *not* of the form $\ldots \leq b_i$ with $b_i \geq 0$;
- let z_a be the sum of the artificial variables;
- temporarily forget the actual objective function and minimise z_a;

– if the minimum value of z_a is > 0 then the LP is infeasible;
– if the minimum value is 0, then the artificial variables will all be 0, delete them from the standard form to get a feasible basic solution, then move to Phase II.

Another method is the so-called "big-M" method (see **Winston**, pp. 172).

5. Phase II of the Simplex Method

Phase II begins with a basic feasible solution, and a corresponding tableau in canonical form. The idea is to find a new basic feasible solution that will differ from the current one in at most two entries; more precisely, one variable will be dropped from the current set of basic variables (the *leaving variable*), and one of the current nonbasic variable will take its place (the *entering variable*). The following example illustrates the strategy for choosing the entering and leaving variables. The iteration ends by reducing the new tableau to canonical form.

PROBLEM: maximise $z = 2x_1 + 5x_2$ subject to

$$\begin{aligned} 2x_1 - 3x_2 &\geq -6 \\ 7x_1 - 2x_2 &\leq 14 \\ x_1 + x_2 &\leq 5 \end{aligned}$$

STANDARD FORM:

$$\begin{aligned} \max z &= 2x_1 + 5x_2 \\ -2x_1 + 3x_2 + x_3 &= 6 \\ 7x_1 - 2x_2 + x_4 &= 14 \\ x_1 + x_2 + x_5 &= 5 \end{aligned}$$

z	x_1	x_2	x_3	x_4	x_5	rhs	BV	ratio
1	-2	-5	0	0	0	0	$z = 0$	—
0	-2	**3**	1	0	0	6	$x_3 = 6$	2
0	7	-2	0	1	0	14	$x_4 = 14$	—
0	1	1	0	0	1	5	$x_5 = 5$	5

First canonical form. The pivot is in bold.

The basic solution $x^T = (0, 0, 6, 14, 5)$ is feasible. Row 0 says

$$z - 2x_1 - 5x_2 = 0\,.$$

The strategy for choosing the entering variable is to look for the variable with respect to which the objective function would increase most rapidly; this is clearly x_2. Another way of putting it is: choose the variable corresponding to the most negative entry in Row 0. We wish to make x_2 a new basic variable; so we need to discard one of the current basic variable. Which should be the leaving variable? For a tableau in canonical form, there can only be one basic variable per row, so the problem of choosing the leaving variable is equivalent to that of choosing the row in which x_2 becomes the basic variable; this is called the *pivot row*. The last three rows of the tableau say

$$\begin{aligned} &\text{Row 1}: && -2x_1 + 3x_2 + x_3 && = 6 \\ &\text{Row 2}: && 7x_1 - 2x_2 + x_4 && = 14 \\ &\text{Row 3}: && x_1 + x_2 + x_5 && = 5 \end{aligned}$$

Ideally, we would like to have x_2 as large as possible whilst maintaining feasibility. Consider each possible choice:

(1) If we chose Row 1, then the new basis would be $\{x_2, x_4, x_5\}$ and the new basic vector would be $x^T = (0, 2, 0, 18, 3)$; this is feasible.
(2) If we chose Row 2, we would find $x_2 = -7$; this in infeasible.
(3) If we chose Row 3, then the new basis would be $\{x_2, x_3, x_4\}$ and the new basic vector would be $x^T = (0, 5, 0, 24, -9)$; this is infeasible.

So we are led to choosing Row 1 as the pivot row.

There is an easier, equivalent way of determining the pivot row:

> **The ratio test:** For each row (other than Row 0) with a positive entry in the column of the entering variable, compute the ratio of the entry in that column to that of the entry in the rhs column. The row with the smallest ratio is said to be the *winner of the ratio test*; take it as the pivot row.

The ratio test "works" because choosing the winner of the ratio test is the only way to ensure that the new basic vector is feasible.

Returning to our particular example, we make the entry in the second row, third column the *pivot* for the elementary row operations required to obtain the new canonical form:

z	x_1	x_2	x_3	x_4	x_5	rhs	BV	ratio
1	$-16/3$	0	$5/3$	0	0	10	$z = 10$	—
0	$-2/3$	1	$1/3$	0	0	2	$x_2 = 2$	—
0	$17/3$	0	$2/3$	1	0	18	$x_4 = 18$	54/17
0	**5/3**	0	$-1/3$	0	1	3	$x_5 = 3$	9/5

Second canonical form. The pivot is in bold.

After this first iteration of the Phase II of the Simplex method; we have

$$z - \frac{16}{3}x_1 + \frac{5}{3}x_3 = 10 \,.$$

We are now ready for the second iteration, and repeat the same steps: we choose x_1 as the entering variable, and x_5 as the leaving variable. By performing elementary row operations, we obtain the third canonical form

z	x_1	x_2	x_3	x_4	x_5	rhs	BV	ratio
1	0	0	$3/5$	0	$16/5$	98/5	$z = 98/5$	
0	1	0	$-1/5$	0	$3/5$	9/5	$x_1 = 9/5$	
0	0	1	$1/5$	0	$2/5$	16/5	$x_2 = 16/5$	
0	0	0	$9/5$	1	$-17/5$	39/5	$x_4 = 39/5$	

Third canonical form.

At this point, all the entries in the first row are positive; hence we have found the optimal solution.

6. A minimisation problem

The Simplex method can also be used to solve minimisation problems. For example, consider the

PROBLEM: minimise $w = 48y_1 + 20y_2 + 8y_3$ subject to

$$8y_1 + 4y_2 + 2y_3 \geq 60$$
$$6y_1 + 2y_2 + 1.5y_3 \geq 30$$
$$2y_1 + 1.5y_2 + 0.5y_3 \geq 20$$

This is equivalent to the problem of maximising $z = -w$ subject to the same constraints, so we have the

STANDARD FORM:

$$\max z = -48y_1 - 20y_2 - 8y_3$$
$$-8y_1 - 4y_2 - 2y_3 + y_4 = -60$$
$$y_5 = -30 + 6y_1 + 2y_2 + 1.5y_3$$
$$y_6 = -20 + 2y_1 + 1.5y_2 + 0.5y_3$$

Choosing y_1, y_2 and y_6 as basic variables, we obtain the

z	y_1	y_2	y_3	y_4	y_5	y_6	rhs	BV
1	48	20	8	0	0	0	0	$z = -300$
0	-8	-4	-2	1	0	0	-60	$y_2 = 15$
0	-6	-2	-1.5	0	1	0	-30	$y_1 = 0$
0	-2	-1.5	-0.5	0	0	1	-20	$y_6 = 2.5$

Basic form.

The basic solution $(0, 15, 0, 0, 0, 2.5)$ is feasible. The canonical form is

z	y_1	y_2	y_3	y_4	y_5	y_6	rhs	BV	ratio
1	0	0	-4	3	4	0	-300	$z = -300$	—
0	0	1	0	-0.75	1	0	15	$y_2 = 15$	∞
0	1	0	**0.25**	0.25	-0.5	0	0	$y_1 = 0$	0
0	0	0	0	-0.625	0.5	1	2.5	$y_6 = 2.5$	∞

First canonical form.

The entering variable is y_3 and the leaving variable is y_1.

z	y_1	y_2	y_3	y_4	y_5	y_6	rhs	BV	ratio
1	4	0	0	7	-4	0	-300	$z = -300$	—
0	0	1	0	-0.75	1	0	15	$y_2 = 15$	15
0	1	0	1	1	-2	0	0	$y_3 = 0$	—
0	0	0	0	-0.625	**0.5**	1	2.5	$y_6 = 2.5$	5

Second canonical form.

The entering variable is y_5, the leaving variable is y_6.

z	y_1	y_2	y_3	y_4	y_5	y_6	rhs	BV	ratio
1	4	0	0	2	0	8	-280	$z = -280$	
0	0	1	0	0.5	0	-2	10	$y_2 = 10$	
0	1	0	1	-1.5	0	4	10	$y_3 = 10$	
0	0	0	0	-1.25	1	2	5	$y_5 = 5$	

Third canonical form.

At this point, the first row contains only positive entries, and so the optimal solution is $(0, 10, 10, 0, 5, 0)$.

7. The optimal tableau in terms of the LP data

We consider the LP in the satndard form

$$\max z = c^T x \quad \text{subject to} \quad Ax = b \ \text{ and } \ x \geq 0\,, \tag{7.1}$$

where A is $m \times n$. Suppose that there is an optimal solution. Set

$$BV := \{x_{j_1}, \ldots, x_{j_m}\} \quad \text{and} \quad NBV := \{x_{j_{m+1}}, \ldots, x_{j_n}\}$$

for the sets of the basic variables and the non-basic variables respectively in the optimal tableau. We shall express the optimal tableau in terms of BV and the data A, b, c of the LP. This will be useful when we study duality and sensitivity later.

Let $\mathbf{a}_j$ denote the jth column of A, and define the matrices

$$B := \begin{pmatrix} \mathbf{a}_{j_1} & \cdots & \mathbf{a}_{j_m} \end{pmatrix} \quad \text{and} \quad N := \begin{pmatrix} \mathbf{a}_{j_{m+1}} & \cdots & \mathbf{a}_{j_n} \end{pmatrix}$$

as well as the vectors

$$x_{BV} = \begin{pmatrix} x_{j_1} \\ \vdots \\ x_{j_m} \end{pmatrix}, \quad c_{BV} = \begin{pmatrix} c_{j_1} \\ \vdots \\ c_{j_m} \end{pmatrix}, \quad x_{NBV} = \begin{pmatrix} x_{j_{m+1}} \\ \vdots \\ x_{j_n} \end{pmatrix} \quad \text{and} \quad c_{NBV} = \begin{pmatrix} c_{j_{m+1}} \\ \vdots \\ c_{j_n} \end{pmatrix}.$$

Then the LP (7.1) can be written as: maximise

$$z = c_{BV}^T x_{BV} + c_{NBV}^T x_{NBV} \tag{7.2}$$

subject to

$$Bx_{BV} + Nx_{NBV} = b\,, \quad x_{BV}\,, x_{NBV} \geq 0\,. \tag{7.3}$$

Multiply the last equation by B^{-1} to obtain

$$x_{BV} + B^{-1}Nx_{NBV} = B^{-1}b\,. \tag{7.4}$$

Then, substitute the result in Equation (7.2); this yields

$$z + \left(c_{BV}^T B^{-1}N - c_{NBV}^T\right) x_{NBV} = c_{BV}^T B^{-1} b\,. \tag{7.5}$$

When written out in tableau form, Equations (7.5) and (7.4) yield a tableau in canonical form that uses the optimal basis; hence this tableau is the optimal tableau. Equation (7.5) gives Row 0 of the optimal tableau; Equation (7.4) gives the remaining rows. Thus, we have shown that the optimal tableau can be expressed as

z	x_{j_1}	x_{j_2}	$\cdots$	x_{j_m}	$x_{j_{m+1}}$	$\cdots$	x_{j_n}	rhs	BV
1	0	0	$\cdots$	0	$\bar{c}_{j_{m+1}}$	$\cdots$	$\bar{c}_{j_n}$	$c_{BV}^T B^{-1} b$	z
0	1	0	$\cdots$	0	$(B^{-1}\mathbf{a}_{j_{m+1}})_1$	$\cdots$	$(B^{-1}\mathbf{a}_{j_n})_1$	$(B^{-1}b)_1$	x_{j_1}
$\vdots$	$\ddots$	$\ddots$	$\ddots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$
0	0	0	$\cdots$	1	$(B^{-1}\mathbf{a}_{j_{m+1}})_m$	$\cdots$	$(B^{-1}\mathbf{a}_{j_n})_m$	$(B^{-1}b)_m$	x_{j_m}

Optimal tableau in terms of the LP data.

where

$$\bar{c}_j = c_{BV}^T \left(B^{-1}\mathbf{a}_j\right) - c_j\,. \tag{7.6}$$

8. Duality

We can associate to any given LP another LP called its *dual*; we then refer to the original, i.e. the given LP as the *primal*. We proceed to explain how the dual of a given LP is defined. We begin with normal max and normal min problems.

Definition 8.1. The *dual* of the normal max LP (1.1) is the normal min LP

$$\min w = b^T y \quad \text{subject to} \quad A^T y \geq c \ \text{ and } \ y \geq 0\,. \tag{8.1}$$

More generally, given a max LP (not necessarily normal), we can define its dual by first reformulating it as a normal max problem; see **Winston**, §6.5. It is not immediately clear that the dual thus obtained is independent of the particular reformulation, but every reformulation of the given LP as a normal problem yield dual problems that are essentially equivalent.

In the remainder of this section, we shall restrict our attention to the case where the primal problem is the normal max LP (1.1). **Thus the dual problem will be the LP** (8.1).

Lemma 8.1 (Weak duality)**.** *For every feasible vector x of the primal problem, and every feasible vector y of the dual problem,*

$$z(x) \leq w(y)\,.$$

Proof. Let $x \in \mathbb{R}^n_+$ be a feasible vector for the primal problem, and $y \in \mathbb{R}^m_+$ be a feasible vector for the dual problem. Then

$$\begin{aligned} z(x) = c^T x = x^T c \leq x^T \left(A^T y\right) = \left(x^T \left(A^T y\right)\right)^T = y^T \left(Ax\right) \\ \leq y^T b = b^T y = w(y)\,. \end{aligned}$$

since $x^T \geq 0$, $y^T \geq 0$, $A^T y \geq c$ and $Ax \leq b$. □

Corollary 8.2. *Let x be primal-feasible and y be dual-feasible. If*

$$z(x) = w(y)$$

then x is primal-optimal and y is dual-optimal.

Proof. Assume that x is not primal-optimal. Then there exists a primal-feasible x^* such that

$$z(x^*) > z(x) = w(y)\,.$$

This contradicts weak duality. The same argument goes for y. □

Corollary 8.3. *If the primal is unbounded, then the dual is infeasible.*

Proof. Assume the contrary. Then we can find a dual-feasible vector y and, by weak duality,

$$z(x) \leq w(y)$$

for every primal-feasible x. This contradicts the assumption that the primal is unbounded. □

Corollary 8.4. *If the dual is unbounded, then the primal is infeasible.*

The concept of duality is closely related to the concept of sentivity to changes in the data of the primal problem. The *sensitivity* of a LP is defined as follows. Write the constraints in the form (linear combination of variables $\Diamond\, b_i$) for $i = 1, \dots, k$, where $\Diamond$ stands for one of $=, \leq, \geq$. Regard the value V of the LP as a function of the b_i. Then the sensitivity to the i-th constraint is $\partial V / \partial b_i$.

To investigate sensitivity, it is useful to express the optimal tableau in terms of the data of the problem. In §7, we worked out a formula for the case where the LP is in standard form. The matrix A and the vectors x and c appearing there are not the same as those appearing in the normal max problem (1.1), so we first need to "translate" the result of §7 in a form appropriate to the normal max problem.

To this end, we introduce slack variables s_i, $1 \leq i \leq m$, to put the normal max problem in the standard form

$$\max z = \tilde{c}^T \tilde{x} \quad \text{subject to} \quad \tilde{A}\tilde{x} = b\,, \;\; \tilde{x} \geq 0\,. \tag{8.2}$$

In this expression,

$$\tilde{x} := \begin{pmatrix} x_1 \\ \vdots \\ x_n \\ s_1 \\ \vdots \\ s_m \end{pmatrix}, \quad \tilde{c} := \begin{pmatrix} c_1 \\ \vdots \\ c_n \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad \text{and} \quad \tilde{A} = \begin{pmatrix} A & \mathbf{e}_1 & \cdots & \mathbf{e}_m \end{pmatrix},$$

where $\mathbf{e}_j$ is the jth column of the $m \times m$ identity matrix. We denote by

$$BV = \{\tilde{x}_{j_1}, \dots, \tilde{x}_{j_m}\} \quad \text{and} \quad NBV = \left\{\tilde{x}_{j_{m+1}}, \dots, \tilde{x}_{j_{m+n}}\right\}$$

the set of the optimal basic and nonbasic variables respectively. We also set

$$B := \begin{pmatrix} \tilde{\mathbf{a}}_{j_1} & \dots & \tilde{\mathbf{a}}_{j_m} \end{pmatrix}, \;\; N := \begin{pmatrix} \tilde{\mathbf{a}}_{j_{m+1}} & \dots & \tilde{\mathbf{a}}_{j_{m+n}} \end{pmatrix},$$

$$\tilde{c}_{BV} := \begin{pmatrix} \tilde{c}_{j_1} \\ \vdots \\ \tilde{c}_{j_m} \end{pmatrix} \quad \text{and} \quad \tilde{c}_{NBV} := \begin{pmatrix} \tilde{c}_{j_{m+1}} \\ \vdots \\ \tilde{c}_{j_{m+n}} \end{pmatrix}.$$

Then, if we repeat the calculation of §7, we find

z	$\tilde{x}_{j_1}$	$\tilde{x}_{j_2}$	$\cdots$	$\tilde{x}_{j_m}$	$\tilde{x}_{j_{m+1}}$	$\cdots$	$\tilde{x}_{j_{m+n}}$	rhs	BV
1	0	0	$\cdots$	0	$\bar{c}_{j_{m+1}}$	$\cdots$	$\bar{c}_{j_{m+n}}$	$\tilde{c}_{BV}^T B^{-1} b$	z
0	1	0	$\cdots$	0	$(B^{-1}\tilde{\mathbf{a}}_{j_{m+1}})_1$	$\cdots$	$(B^{-1}\tilde{\mathbf{a}}_{j_{m+n}})_1$	$(B^{-1}b)_1$	$\tilde{x}_{j_1}$
$\vdots$	$\ddots$	$\ddots$	$\ddots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$
0	0	0	$\cdots$	1	$(B^{-1}\tilde{\mathbf{a}}_{j_{m+1}})_m$	$\cdots$	$(B^{-1}\tilde{\mathbf{a}}_{j_{m+n}})_m$	$(B^{-1}b)_m$	$\tilde{x}_{j_m}$

Optimal tableau in terms of the data for the normal max problem (1.1).

where

$$\bar{c}_{j_k} = \tilde{c}_{BV}^T \left(B^{-1}\tilde{\mathbf{a}}_{j_k}\right) - \tilde{c}_{j_k}\,, \quad m+1 \leq k \leq m+n\,. \tag{8.3}$$

Proposition 8.5. *We have the following formulae for the entries in Row 0 of the optimal tableau for the normal max problem:*

$$\begin{aligned} \text{entry in the } x_j \text{ column} &= \tilde{c}_{BV}^T \left(B^{-1}\mathbf{a}_j\right) - c_j \\ \text{entry in the } s_j \text{ column} &= \tilde{c}_{BV}^T \left(B^{-1}\mathbf{e}_j\right) \\ \text{entry in the rhs column} &= \tilde{c}_{BV}^T \left(B^{-1}b\right) \end{aligned}$$

Proof. We begin with the formula for the decision variables. Consider x_j where $1 \leq j \leq n$ and let $1 \leq k \leq m+n$ be the index such that $\mathbf{a}_j = \tilde{\mathbf{a}}_{j_k}$. Suppose that $x_j \in BV$. Then $1 \leq k \leq m$; in other words, $\mathbf{a}_j$ is the kth column of B and, by definition of the inverse of a matrix,

$$B^{-1}\mathbf{a}_j = \mathbf{e}_k\,.$$

Hence

$$\tilde{c}_{BV}^T \left(B^{-1}\mathbf{a}_j\right) - c_j = \tilde{c}_{BV}^T \mathbf{e}_k - c_j = \tilde{c}_{j_k} - c_j = c_j - c_j = 0\,.$$

This shows that the formula is correct if $x_j \in BV$. Next, suppose that $x_j \in NBV$. Then $m+1 \leq k \leq m+n$ and Equation (8.3) shows the validity of the formula.

Turning now to the formula for the slack variables, consider s_j where $1 \leq j \leq m$ and let $1 \leq k \leq m+n$ be the index such that $\mathbf{e}_j = \tilde{\mathbf{a}}_{j_k}$. Suppose that $s_j \in BV$. Then $1 \leq k \leq m$ and $\mathbf{e}_j$ is the kth column of B; so

$$B^{-1}\mathbf{e}_j = \mathbf{e}_k$$

and

$$\tilde{c}_{BV}^T \left(B^{-1}\mathbf{e}_j\right) = \tilde{c}_{BV}^T \mathbf{e}_k = \tilde{c}_{j_k} = 0\,.$$

So, if $s_j \in BV$, the formula is correct. Suppose now that $s_j \in NBV$ Then $m+1 \leq k \leq m+n$ and Equation (8.3) shows, again, the validity of the formula. □

Theorem 8.6 (The Duality Theorem)**.** *If the primal problem has an optimal solution, then the vector $y \in \mathbb{R}^m$ defined by*

$$y^T = \tilde{c}_{BV}^T B^{-1}$$

is an optimal solution for the dual problem.

Proof. Suppose that the primal has an optimal solution. Then the primal value is the entry of rhs in Row 0 of the optimal tableau, as given in Proposition 8.5. So if we can show that y is dual-feasible, then the theorem will follow immediately from Lemma 8.2.

Proposition 8.5 shows that the entry of s_j in Row 0 of the optimal primal tableau is precisely y_j, whilst the entry of x_j is the jth entry of $y^T A - c^T$. By the optimality of the tableau, the entries in Row 0 are non-negative. Hence

$$y^T \geq 0\,.$$

and

$$y^T A - c^T \geq 0\,.$$

This is the same as

$$A^T y \geq c \quad \text{and} \quad y \geq 0\,.$$

So y is indeed dual-feasible. □

Corollary 8.7. *We have*

$$y_j = \frac{\partial V}{\partial b_j}\,, \quad 1 \le j \le m\,.$$

In other words, the dual solution is a measure of the sensitivity of the primal problem to changes in the right-hand of the constraints.

Corollary 8.8 (Complementary Slackness Property)**.** *Set*

$$e := A^T y - c \in \mathbb{R}^n\,.$$

We have

$$\forall\, 1 \le j \le m\,, \quad s_j\, y_j = 0 \quad \text{and} \quad \forall\, 1 \le j \le n\,, \quad e_j\, x_j = 0\,.$$

Proof. If $s_j \in BV$, then the entry of s_j in Row 0 of the optimal tableau vanishes; so, by Proposition 8.5, $y_j = 0$. On the other hand, if $s_j \in NBV$, then $s_j = 0$. In the same way, if $x_j \in BV$, then the entry of s_j in Row 0 of the optimal tableau vanishes; so, by Proposition 8.5, the jth entry of the vector $y^T A - c$ vanishes. This is just the same as $e_j = 0$. On the other hand, if $x_j \in NBV$, then $x_j = 0$. □

Example 8.1 (**Winston**'s Dakota problem)**.** *Consider the normal max LP: maximise $z = 60x_1 + 30x_2 + 20x_3$ subject to*

$$\begin{aligned} 8x_1 + 6x_2 + x_3 &\le 48 \\ 4x_1 + 2x_2 + 1.5x_3 &\le 20 \\ 2x_1 + 1.5x_2 + 0.5x_3 &\le 8 \end{aligned}$$

Hence,

$$\tilde{A} = \begin{pmatrix} 8 & 6 & 1 & 1 & 0 & 0 \\ 4 & 2 & 1.5 & 0 & 1 & 0 \\ 2 & 1.5 & 0.5 & 0 & 0 & 1 \end{pmatrix}\,, \quad b = \begin{pmatrix} 48 \\ 20 \\ 8 \end{pmatrix} \quad \text{and} \quad \tilde{c} = \begin{pmatrix} 60 \\ 30 \\ 20 \\ 0 \\ 0 \\ 0 \end{pmatrix}\,.$$

The simplex method yields the final canonical form

z	x_1	x_2	x_3	s_1	s_2	s_3	*rhs*	BV	*ratio*
1	0	5	0	0	10	10	280	$z = 280$	
0	0	0	0	1	2	-8	24	$s_1 = 24$	
0	0	-2	1	0	2	-4	8	$x_3 = 8$	
0	1	1.25	0	0	-0.5	1	2	$x_1 = 2$	

Final canonical form for the primal problem.

We have

$$B = \begin{pmatrix} 1 & 1 & 8 \\ 0 & 1.5 & 4 \\ 0 & 0.5 & 2 \end{pmatrix} \quad \text{and} \quad \tilde{c}_{BV} = \begin{pmatrix} 0 \\ 20 \\ 60 \end{pmatrix}\,.$$

Therefore

$$\tilde{c}_{BV}^T B^{-1} = \begin{pmatrix} 0 & 10 & 10 \end{pmatrix}\,.$$

The dual problem is that considered in §6; this illustrates the theorem's validity.

9. The dual of a general LP

So far in our discussion of duality, we have considered only normal max problems. Our purpose in this section is to explain how one can define the dual of an arbitrary LP. We shall generalise the definition in a way that makes duality a *symmetric* relation; in other words, if LP* denotes the dual of LP, then

$$(\mathrm{LP}^*)^* = \mathrm{LP}\,. \tag{9.1}$$

With this in mind, it is immediately clear how to define the dual of a normal min problem.

Definition 9.1. The dual of the normal min LP:

$$\min w = c^T y \quad \text{subject to} \quad Ay \geq b\,, \;\; y \geq 0\,,$$

is the normal max LP:

$$\max z = b^T x \quad \text{subject to} \quad A^T x \leq c\,, \;\; x \geq 0\,.$$

Next, consider a LP that is not in normal form, i.e. the LP is such that

(1) Some decision variables are *unrestricted in sign* (urs).
(2) Some of the constraints are $=$ constraints.
(3) Some of the constraints are $\geq$ constraints.

The following steps will convert the LP into normal max form:

(1) Suppose x_i is urs. Introduce two new non-negative variables x_i' and x_i'' and make the substitution
$$x_i := x_i' - x_i''\,.$$
This eliminates x_i from the problem.
(2) Replace any $=$ constraint of the form
$$\text{lhs} = \text{rhs}$$
by the equivalent pair of $\leq$ constraints
$$\text{lhs} \leq \text{rhs} \quad \text{and} \quad -\text{lhs} \leq -\text{rhs}\,.$$
(3) Replace any $\geq$ constraint of the form
$$\text{lhs} \geq \text{rhs}$$
by the equivalent $\leq$ constraint
$$-\text{lhs} \leq -\text{rhs}\,.$$

Example 9.1. *Consider the LP:* $\max z = \nu$ *subject to*

$$\begin{aligned} \nu &\leq x_2 - x_3 \\ \nu &\leq -x_1 + x_3 \\ \nu &\leq x_1 - x_2 \\ x_1 + x_2 + x_3 &= 1 \\ x_1,\, x_2,\, x_3 &\geq 0,\; \nu \text{ urs} \end{aligned}$$

Set $\nu = x_4 - x_5$. *Then the LP can be put in the normal form:* $\max z = x_4 - x_5$ *subject to*

$$\begin{aligned} -x_2 + x_3 + x_4 - x_5 &\le 0 \\ x_1 - x_3 + x_4 - x_5 &\le 0 \\ -x_1 + x_2 + x_4 - x_5 &\le 0 \\ x_1 + x_2 + x_3 &\le 1 \\ -x_1 - x_2 - x_3 &\le -1 \\ x &\ge 0 \end{aligned}$$

or, equivalently,

$$\max z = c^T x \quad \textit{subject to} \quad Ax \le b,\ x \ge 0,$$

where

$$A := \begin{pmatrix} 0 & -1 & 1 & 1 & -1 \\ 1 & 0 & -1 & 1 & -1 \\ -1 & 1 & 0 & 1 & -1 \\ 1 & 1 & 1 & 0 & 0 \\ -1 & -1 & -1 & 0 & 0 \end{pmatrix}, \quad b := \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ -1 \end{pmatrix} \quad \textit{and} \quad c := \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ -1 \end{pmatrix}.$$

The normal form of the dual is therefore

$$\min w = b^T y \quad \textit{subject to} \quad A^T y \ge c,\ y \ge 0\,.$$

That is: $\min w = y_4 - y_5$ *subject to*

$$\begin{aligned} y_2 - y_3 + y_4 - y_5 &\ge 0 \\ -y_1 + y_3 + y_4 - y_5 &\ge 0 \\ y_1 - y_2 + y_4 - y_5 &\ge 0 \\ y_1 + y_2 + y_3 &\ge 1 \\ -y_1 - y_2 - y_3 &\ge -1 \\ y &\ge 0 \end{aligned}$$

Bearing in mind our earlier method of converting an LP to an equivalent one in normal form, we see that the dual can also be expressed in the non-normal form: $\min w$ *subject to*

$$\begin{aligned} w &\ge y_3 - y_2 \\ w &\ge y_1 - y_3 \\ w &\ge y_2 - y_1 \\ y_1 + y_2 + y_3 &= 1 \\ y_1,\, y_2,\, y_3 &\ge 0,\ w \text{ urs} \end{aligned}$$

10. The dual simplex method

For a normal max problem, two properties of the tableaux produced by the simplex method are:

- At every step, the entries of the rhs column are all non-negative; this expresses the fact that the current vector is feasible.

- Only in the final tableau are the entries in the first row non-negative; this expresses the optimality property of the final vector.

The proof of Theorem 8.6 shows that the positivity of the entries in Row 0 is also equivalent to the dual-feasibility of the vector defined by

$$y^T = \tilde{c}_{BV}^T B^{-1} .$$

The dual simplex method turns the simplex strategy around; it produces a sequence of tableaux such that

- At every step, the entries in Row 0 are all non-negative (dual-feasibility is maintained).
- Only in the final tableau are the entries in the rhs column non-negative; this expresses the (primal) feasibility of the final vector.

This variant of the simplex algorithm is particularly useful when, after solving a given LP, new constraints are added. A description is given by **Winston**, §6.11.

11. MAPLE's `simplex` package

MAPLE has a useful package that implements the simplex method.

```
> with(simplex);
Warning, the protected names maximize and minimize have been redefined
and unprotected
[basis, convexhull, cterm, define_zero, display,
dual, feasible, maximize,
minimize, pivot, pivoteqn, pivotvar, ratio, setup, standardize]
```

We use **Winston**'s Dakota problem, i.e. Example 8.1, to illustrate how it works.

First, we enter the objective function and the constraints:

```
> z := 60*x1 + 30*x2+20*x3;
                    z := 60 x1 + 30 x2 + 20 x3
> constraints := {8*x1+6*x2+x3<=48,4*x1+2*x2+1.5*x3<=20,
                                        2*x1+1.5*x2+0.5*x3<=8};
constraints := {8 x1 + 6 x2 + x3 <= 48, 4 x1 + 2 x2 + 1.5 x3 <= 20,
                                        2 x1 + 1.5 x2 + 0.5 x3 <= 8}
```

It's a good idea to check that the problem is feasible:

```
> feasible(constraints,NONNEGATIVE);
                               true
```

The `setup` functions puts the problem in standard form:

```
> C1 := setup(constraints,NONNEGATIVE);
C1 := {_SL1 = 48 - 8 x1 - 6 x2 - x3, _SL2 = 20 - 4 x1 - 2 x2 - 1.5 x3,
                                    _SL3 = 8 - 2 x1 - 1.5 x2 - 0.5 x3}
```

The functions `maximize` and `minimize` simply return the solution. By using the other functions, it is possible to carry out the individual steps of the simplex algorithm. We have

```
> basis(C1);display(C1);z;
                      {_SL1, _SL2, _SL3}
                 [8    6    1  1  0  0]    [ 48]
                 [                    ]    [   ]
                 [4    2  1.5  0  1  0] =  [20.]
```

```
                [                    ]    [   ]
                [2  1.5  0.5  0  0  1]    [ 8.]
                    60 x1 + 30 x2 + 20 x3
```

This is in canonical form. It is not optimal, since the expression for z in terms of the non-basic variables contains positive coefficients. We look for the entering variable:

```
> pivotvar(z);
                                x1
```

Then we look for the leaving variable:

```
> eqn := pivoteqn(C1,x1);
           eqn := [_SL3 = 8 - 2 x1 - 1.5 x2 - 0.5 x3]
```

We obtain the new canonical form by performing elementary row operations:

```
> C2 := pivot(C1,x1,eqn);
C2 := {_SL2 = 4. + 2.000000000 _SL3 + 1.000000000 x2 - 0.500000000 x3,
                    _SL1 = 16. + 4.000000000 _SL3 + 1.000000000 x3,
    x1 = -0.5000000000 _SL3 + 4. - 0.7500000000 x2 - 0.2500000000 x3}
```

We express z in terms of the non-basic variables:

```
> solve(eqn[1],x1);
  -0.5000000000 _SL3 + 4. - 0.7500000000 x2 - 0.2500000000 x3
> z := subs(x1=%,z);
 z := -30.00000000 _SL3 + 240. - 15.00000000 x2 + 5.00000000 x3
```

We then display the current canonical form, the current basis, and the expression for z:

```
> basis(C2);display(C2);z;
                          {x1, _SL1, _SL2}
  [0  -1.000000000   0.500000000  0  1  -2.000000000]    [ 4.]
  [                                                 ]    [   ]
  [0             0  -1.000000000  1  0  -4.000000000] = [16.]
  [                                                 ]    [   ]
  [1  0.7500000000  0.2500000000  0  0  0.5000000000]    [ 4.]
   -30.00000000 _SL3 + 240. - 15.00000000 x2 + 5.00000000 x3
```

This is not optimal; indeed

```
> pivotvar(z);
                                x3
```

So x_3 is the new entering variable. We proceed as before to find the next canonical form:

```
> eqn := pivoteqn(C2,x3);
eqn := [_SL2 = 4. + 2.000000000 _SL3 + 1.000000000 x2
                                          - 0.500000000 x3]
> C3 := pivot(C2,x3,eqn);
C3 := {
x1 = -1.500000000 _SL3 + 2.000000000 - 1.250000000 x2
                                     + 0.5000000000 _SL2,
                   x3 = -2. _SL2 + 8. + 4. _SL3 + 2. x2,
    _SL1 = 24.00000000 + 8.000000000 _SL3 - 2.000000000 _SL2
                                     + 2.000000000 x2}
```

```
> solve(eqn[1],x3);
                -2. _SL2 + 8. + 4. _SL3 + 2. x2
> z := subs(x3=%,z);
z := -10.00000000 _SL3 + 280.0000000 - 5.00000000 x2
                                     - 10.00000000 _SL2
> basis(C3);display(C3);z;
                         {x1, x3, _SL1}
[0  -2.000000000  0  1    2.000000000  -8.000000000]   [24.00000000]
[                                                  ]   [           ]
[1   1.250000000  0  0  -0.5000000000   1.500000000] = [2.000000000]
[                                                  ]   [           ]
[0           -2.  1  0             2.           -4.]   [         8.]
-10.00000000 _SL3 + 280.0000000 - 5.00000000 x2 - 10.00000000 _SL2
```

The coefficients of the non-basic variables in the expression for z are all negative, so we have found an optimal solution. Indeed,

```
> pivotvar(z);
                              FAIL
```

and the algorithm terminates.

The MAPLE worksheet may be found at

http://www.maths.bris.ac.uk/~mayt/MATH32500/simplex.mw