# 4 Duality Theory

Recall from Section 1 that the dual to an LP in standard form

$$(\mathcal{P}) \qquad \begin{array}{ll} \text{maximize} & c^T x \\ \text{subject to} & Ax \leq b,\ 0 \leq x \end{array}$$

is the LP

$$(\mathcal{D}) \qquad \begin{array}{ll} \text{minimize} & b^T y \\ \text{subject to} & A^T y \geq c,\ 0 \leq y. \end{array}$$

Since the problem $\mathcal{D}$ is a linear program, it too has a dual. The *duality* terminology suggests that the problems $\mathcal{P}$ and $\mathcal{D}$ come as a pair implying that the dual to $\mathcal{D}$ should be $\mathcal{P}$. This is indeed the case as we now show:

$$\begin{array}{ll} \text{minimize} & b^T y \\ \text{subject to} & A^T y \geq c, \\ & 0 \leq y \end{array} \quad = \quad \begin{array}{ll} -\text{maximize} & (-b)^T y \\ \text{subject to} & (-A^T) y \leq (-c), \\ & 0 \leq y. \end{array}$$

The problem on the right is in standard form so we can take its dual to get the LP

$$\begin{array}{ll} \text{minimize} & (-c)^T x \\ \text{subject to} & (-A^T)^T x \geq (-b),\ 0 \leq x \end{array} \quad = \quad \begin{array}{ll} \text{maximize} & c^T x \\ \text{subject to} & Ax \leq b,\ 0 \leq x\ . \end{array}$$

The primal-dual pair of LPs $\mathcal{P} - \mathcal{D}$ are related via the Weak Duality Theorem.

**Theorem 4.1 (Weak Duality Theorem)** *If $x \in \mathsf{R}^n$ is feasible for $\mathcal{P}$ and $y \in \mathsf{R}^m$ is feasible for $\mathcal{D}$, then*

$$c^T x \leq y^T A x \leq b^T y.$$

*Thus, if $\mathcal{P}$ is unbounded, then $\mathcal{D}$ is necessarily infeasible, and if $\mathcal{D}$ is unbounded, then $\mathcal{P}$ is necessarily infeasible. Moreover, if $c^T \bar{x} = b^T \bar{y}$ with $\bar{x}$ feasible for $\mathcal{P}$ and $\bar{y}$ feasible for $\mathcal{D}$, then $\bar{x}$ must solve $\mathcal{P}$ and $\bar{y}$ must solve $\mathcal{D}$.*

We now use The Weak Duality Theorem in conjunction with The Fundamental Theorem of Linear Programming to prove the *Strong Duality Theorem*. The key ingredient in this proof is the general form for simplex tableaus derived at the end of Section 2 in (2.5).

**Theorem 4.2 (The Strong Duality Theorem)** *If either $\mathcal{P}$ or $\mathcal{D}$ has a finite optimal value, then so does the other, the optimal values coincide, and optimal solutions to both $\mathcal{P}$ and $\mathcal{D}$ exist.*

Remark: This result states that the finiteness of the optimal value implies the existence of a solution. This is not always the case for nonlinear optimization problems. Indeed, consider the problem

$$\min_{x \in \mathsf{R}} e^x.$$

This problem has a finite optimal value, namely zero; however, this value is not attained by any point $x \in \mathsf{R}$. That is, it has a finite optimal value, but a solution does not exist. The existence of solutions when the optimal value is finite is one of the many special properties of linear programs.

PROOF: Since the dual of the dual is the primal, we may as well assume that the primal has a finite optimal value. In this case, the Fundamental Theorem of Linear Programming says that an optimal basic feasible solution exists. By our formula for the general form of simplex tableaus (2.5), we know that there exists a nonsingular record matrix $R \in \mathsf{R}^{n\times n}$ and a vector $y \in \mathsf{R}^m$ such that the optimal tableau has the form

$$\begin{bmatrix} R & 0 \\ -y^T & 1 \end{bmatrix} \begin{bmatrix} A & I & b \\ c^T & 0 & 0 \end{bmatrix} = \begin{bmatrix} RA & R & Rb \\ c^T - y^T A & -y^T & -y^T b \end{bmatrix}.$$

Since this is an optimal tableau we know that

$$c - A^T y \leq 0, \qquad\qquad -y^T \leq 0$$

with $y^T b$ equal to optimal value in the primal problem. But then $A^T y \geq c$ and $0 \leq y$ so that $y$ is feasible for the dual problem $\mathcal{D}$. In addition, the Weak Duality Theorem implies that

$$\begin{array}{rclll} b^T y & = & \text{maximize} & c^T x & \leq \; b^T \widehat{y} \\ & & \text{subject to} & Ax \leq b,\ 0 \leq x & \end{array}$$

for every vector $\widehat{y}$ that is feasible for $\mathcal{D}$. Therefore, $y$ solves $\mathcal{D}$!!!! ■

This is an amazing fact! Our method for solving the primal problem $\mathcal{P}$, the simplex algorithm, simultaneously solves the dual problem $\mathcal{D}$! This fact will be of enormous practical value when we study sensitivity analysis.

## 4.1 Complementary Slackness

The Strong Duality Theorem tells us that optimality is equivalent to equality in the Weak Duality Theorem. That is, $x$ solves $\mathcal{P}$ and $y$ solves $\mathcal{D}$ if and only if $(x, y)$ is a $\mathcal{P} - \mathcal{D}$ feasible pair and

$$c^T x = y^T A x = b^T y.$$

We now carefully examine the consequences of this equivalence. Note that the equation $c^T x = y^T A x$ implies that

$$0 = x^T (A^T y - c) = \sum_{j=1}^{n} x_j \Big(\sum_{i=1}^{m} a_{ij} y_i - c_j\Big). \tag{4.1}$$

In addition, feasibility implies that

$$0 \leq x_j \qquad \text{and} \qquad 0 \leq \sum_{i=1}^{m} a_{ij} y_i - c_j \qquad \text{for } j = 1, \ldots, n,$$

and so

$$x_j(\sum_{i=1}^{m} a_{ij}y_i - c_j) \geq 0 \qquad \text{for} \qquad j = 1, \ldots, n.$$

Hence, the only way (4.1) can hold is if

$$x_j(\sum_{i=1}^{m} a_{ij}y_i - c_j) = 0 \qquad \text{for} \qquad j = 1, \ldots, n.$$

or equivalently,

$$(4.2) \qquad x_j = 0 \qquad \text{or} \qquad \sum_{i=1}^{m} a_{ij}y_i = c_j \qquad \text{or both for } j = 1, \ldots, n.$$

Similarly, (**??**) implies that

$$0 = y^T(b - Ax) = \sum_{i=1}^{m} y_i(b_i - \sum_{j=1}^{n} a_{ij}x_j).$$

Again, feasibility implies that

$$0 \leq y_i \qquad \text{and} \qquad 0 \leq b_i - \sum_{j=1}^{n} a_{ij}x_j \qquad \text{for } i = 1, \ldots, m.$$

Thus, we must have

$$y_i(b_i - \sum_{j=1}^{n} a_{ij}x_j) = 0 \qquad \text{for} \qquad j = 1, \ldots, n,$$

or equivalently,

$$(4.3) \qquad y_i = 0 \qquad \text{or} \qquad \sum_{j=1}^{n} a_{ij}x_j = b_i \qquad \text{or both for } i = 1, \ldots, m.$$

The two observations (4.2) and (4.3) combine to yield the following theorem.

**Theorem 4.8 (The Complementary Slackness Theorem)** *The vector $x \in \mathsf{R}^n$ solves $\mathcal{P}$ and the vector $y \in \mathsf{R}^m$ solves $\mathcal{D}$ if and only if $x$ is feasible for $\mathcal{P}$ and $y$ is feasible for $\mathcal{D}$ and*

*(i) either $0 = x_j$ or $\sum_{i=1}^{m} a_{ij}y_i = c_j$ or both for $j = 1, \ldots, n$, and*

*(ii) either $0 = y_i$ or $\sum_{j=1}^{n} a_{ij}x_j = b_i$ or both for $i = 1, \ldots, m$.*

PROOF: If $x$ solves $\mathcal{P}$ and $y$ solves $\mathcal{D}$, then by the Strong Duality Theorem we have equality in the Weak Duality Theorem. But we have just observed that this implies (4.2) and (4.3) which are equivalent to (i) and (ii) above.

Conversely, if (i) and (ii) are satisfied, then we get equality in the Weak Duality Theorem. Therefore, by Theorem 4.2, $x$ solves $\mathcal{P}$ and $y$ solves $\mathcal{D}$. ■

The Complementary Slackness Theorem can be used to develop a test of optimality for a putative solution to $\mathcal{P}$ (or $\mathcal{D}$). We state this test as a corollary.

**Corollary 4.1** *The vector $x \in \mathsf{R}^n$ solves $\mathcal{P}$ if and only if $x$ is feasible for $\mathcal{P}$ and there exists a vector $y \in \mathsf{R}^m$ feasible for $\mathcal{D}$ and such that*

*(i) for each $i \in \{1, 2, \ldots, m\}$, if $\sum_{j=1}^{n} a_{ij} x_j < b_i$, then $y_i = 0$, and*

*(ii) for each $j \in \{1, 2, \ldots, n\}$, if $0 < x_j$, then $\sum_{i=1}^{m} a_{ij} y_i = c_j$.*

PROOF: (i) and (ii) implies equality in the Weak Duality Theorem. The primal feasibility of $x$ and the dual feasibility of $y$ combined with Theorem 4.1 yield the result. ■

We now show how to apply this Corollary to test whether or not a given point solves an LP. Recall that all of the nonbasic variables in an optimal BFS take the value zero, and, if the BFS is nondegenerate, then all of the basic variables are nonzero. That is, $m$ of the variables in the optimal BFS are nonzero since every BFS has $m$ basic variables. Consequently, among the $n$ original decision variables and the $m$ slack variables, $m$ variables are nonzero at a nondegenerate optimal BFS. That is, among the constraints

$$\begin{array}{ll} 0 \le x_j & j = 1, \ldots, n, \\ 0 \le x_{n+i} = c_i - \displaystyle\sum_{i \in N} a_{ij} x_j & i = 1, \ldots, m \end{array}$$

$m$ of them are strict inequalities. If we now look back at Corollary 4.1, we see that every nondegenerate optimal basic feasible solution yields a total of $m$ equations that an optimal dual solution $y$ must satisfy. That is, Corollary 4.1 tells us that the $m$ optimal dual variables $y_i$ satisfy $m$ equations. Therefore, we can write an $m \times m$ system of equations to solve for $y$. We illustrate this by applying Corollary 4.1 to the following LP

$$\begin{array}{lrcrcrcrcrcl} \text{maximize} & 7x_1 & + & 6x_2 & + & 5x_3 & - & 2x_4 & + & 3x_5 & & \\ \text{subject to} & x_1 & + & 3x_2 & + & 5x_3 & - & 2x_4 & + & 2x_5 & \le & 4 \\ & 4x_1 & + & 2x_2 & - & 2x_3 & + & x_4 & + & x_5 & \le & 3 \\ & 2x_1 & + & 4x_2 & + & 4x_3 & - & 2x_4 & + & 5x_5 & \le & 5 \\ & 3x_1 & + & x_2 & + & 2x_3 & - & x_4 & - & 2x_5 & \le & 1 \\ & & & & & \multicolumn{5}{c}{0 \le x_1, x_2, x_3, x_4, x_5.} & & \end{array} \tag{4.10}$$

Does the point

$$x^T = (x_1, x_2, x_3, x_4, x_5) = (0, \frac{4}{3}, \frac{2}{3}, \frac{5}{3}, 0)$$

solves this LP? Following Corollary 4.1, if $x$ is optimal, then there exists a vector $y \in \mathsf{R}^4$ feasible for the dual LP to (4.10) and which satisfies the conditions given in items (i) and (ii) of the corollary. Plugging $x$ into the constraints for (4.10) we see that equality is attained in each of the constraints except the third:

$$\begin{array}{rcrcrcrcrcl}
(0) & + & 3\left(\frac{4}{3}\right) & + & 5\left(\frac{2}{3}\right) & - & 2\left(\frac{5}{3}\right) & + & 2(0) & = & 4 \\
4(0) & + & 2\left(\frac{4}{3}\right) & - & 2\left(\frac{2}{3}\right) & + & \left(\frac{5}{3}\right) & + & (0) & = & 3 \\
2(0) & + & 4\left(\frac{4}{3}\right) & + & 4\left(\frac{2}{3}\right) & - & 2\left(\frac{5}{3}\right) & + & 5(0) & < & 5 \\
3(0) & + & \left(\frac{4}{3}\right) & + & 2\left(\frac{2}{3}\right) & - & \left(\frac{5}{3}\right) & - & 2(0) & = & 1.
\end{array}$$

By item (i) of Corollary 4.1, we see that the vector $y \in \mathsf{R}^4$ that we seek must have

$$y_3 = 0. \tag{4.11}$$

Since $x_2 > 0$, $x_3 > 0$, and $x_4 > 0$, item (ii) of Corollary 4.1 implies that the vector $y$ we are looking for must also satisfy the equations

$$\begin{array}{rcrcrcrcr}
3y_1 & + & 2y_2 & + & 4y_3 & + & y_4 & = & 6 \\
5y_1 & - & 2y_2 & + & 4y_3 & + & 2y_4 & = & 5 \\
-2y_1 & + & y_2 & - & 2y_3 & - & y_4 & = & -2.
\end{array} \tag{4.12}$$

Putting (4.11) and (4.12) together, we see that $y$ must satisfy

$$\begin{bmatrix} 3 & 2 & 4 & 1 \\ 5 & -2 & 4 & 2 \\ -2 & 1 & -2 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} 6 \\ 5 \\ -2 \\ 0 \end{pmatrix},$$

where the first three rows come from (4.12) and the last row comes from (4.11). We reduce the associated augmented system as follows:

$$
\begin{array}{rrrr|r l}
3 & 2 & 4 & 1 & 6 & \\
5 & -2 & 4 & 2 & 5 & \\
-2 & 1 & -2 & -1 & -2 & \\
0 & 0 & 1 & 0 & 0 & \\
\hline
3 & 2 & 0 & 1 & 6 & \quad r_1 - 4r_4 \\
5 & -2 & 0 & 2 & 5 & \quad r_2 - 4r_4 \\
-2 & 1 & 0 & -1 & -2 & \quad r_3 + 2r_4 \\
0 & 0 & 1 & 0 & 0 & \\
\hline
1 & 3 & 0 & 0 & 4 & \quad r_1 + r_3 \\
1 & 0 & 0 & 0 & 1 & \quad r_2 + 2r_3 \\
-2 & 1 & 0 & -1 & -2 & \\
0 & 0 & 1 & 0 & 0 & \\
\hline
0 & 3 & 0 & 0 & 3 & \quad r_1 - r_2 \\
1 & 0 & 0 & 0 & 1 & \\
0 & 1 & 0 & -1 & 0 & \quad r_3 + 2r_2 \\
0 & 0 & 1 & 0 & 0 & \\
\hline
1 & 0 & 0 & 0 & 1 & \quad r_2 \\
0 & 1 & 0 & 0 & 1 & \quad \frac{1}{3}r_1 \\
0 & 0 & 1 & 0 & 0 & \quad r_4 \\
0 & 0 & 0 & 1 & 1 & \quad -r_3 + \frac{1}{3}r_1
\end{array}
$$

This gives $y^T = (1, 1, 0, 1)$ as the only possible vector $y$ that can satisfy the requirements of (i) and (ii) in Corollary 4.1. To satisfy these requirements, we need only check that $y$ is feasible for the dual LP to (4.10):

$$
\begin{array}{lrcrcrcrcr}
\text{minimize} & 4y_1 & + & 3y_2 & + & 5y_3 & + & y_4 & & \\
\text{subject to} & y_1 & + & 4y_2 & + & 2y_3 & + & 3y_4 & \geq & 7 \\
 & 3y_1 & + & 2y_2 & + & 4y_3 & + & y_4 & \geq & 6 \\
 & 5y_1 & - & 2y_2 & + & 4y_3 & + & 2y_4 & \geq & 5 \\
 & -2y_1 & + & y_2 & - & 2y_3 & - & y_4 & \geq & -2 \\
 & 2y_1 & + & y_2 & + & 5y_3 & - & 2y_4 & \geq & 3 \\
 & & & \multicolumn{5}{l}{0 \leq y_1, y_2, y_3, y_4.} & &
\end{array}
$$

Clearly, $0 \leq y$ and by construction the 2nd, 3rd, and 4th of the linear inequality constraints are satisfied with equality. Thus, it only remains to check the first and fifth inequalities:

$$
\begin{array}{rcrcrcrcl}
(1) & + & 4(1) & + & 2(0) & + & 3(1) & = & 8 \geq 7 \\
2(1) & + & (1) & + & 5(0) & - & 2(1) & = & 1 \not\geq 3.
\end{array}
$$

Therefore, $y$ is not dual feasible. But as observed, this is the only possible vector $y$ satisfying (i) and (ii) of Corollary (4.1), hence $x^T = (0, \frac{4}{3}, \frac{2}{3}, \frac{5}{3}, 0)$ cannot be a solution to the LP (4.10).

## 4.2 General Duality Theory

Thus far we have discussed duality theory as it pertains to LPs in standard form. Of course, one can always transform any LP into one in standard form and then apply the duality theory. However, from the perspective of applications, this is cumbersome since it obscures the meaning of the dual variables. It is very useful to be able to compute the dual of an LP without first converting to standard form. In this section we show how this can easily be done. For this, we still make use of a standard form, but now we choose one that is much more flexible:

$$\begin{array}{lll} \mathcal{P} & \max & \sum_{j=1}^n c_j x_j \\ & \text{subject to} & \sum_{j=1}^n a_{ij} x_j \le b_i \qquad i \in I \\ & & \sum_{j=1}^n a_{ij} x_j = b_i \qquad i \in E \\ & & 0 \le x_j \qquad\qquad\quad j \in R \quad . \end{array}$$

Here the index sets $I$, $E$, and $R$ are such that

$$I \cap E = \emptyset,\ I \cup E = \{1, 2, \dots, m\},\ \text{and } R \subset \{1, 2, \dots, n\}.$$

We use the following primal-dual correspondences to compute the dual of an LP.

| In the Dual | In the Primal |
|---|---|
| Restricted Variables | Inequality Constraints |
| Free Variables | Equality Constraints |
| Inequality Constraints | Restricted Variables |
| Equality Constraints | Free Variables |

Using these rules we obtain the dual to $\mathcal{P}$.

$$\begin{array}{lll} \mathcal{D} & \min & \sum_{i=1}^m b_i y_i \\ & \text{subject to} & \sum_{i=1}^m a_{ij} y_i \ge c_j \qquad j \in R \\ & & \sum_{i=1}^m a_{ij} y_i = c_j \qquad j \in F \\ & & 0 \le y_i \qquad\qquad\quad i \in I \quad , \end{array}$$

where $F = \{1, 2, \dots, n\} \setminus R$.

For example, the LP

$$\begin{array}{lll} \text{maximize} & x_1 - 2x_2 + 3x_3 & \\ \text{subject to} & 5x_1 + x_2 - 2x_3 & \le 8 \\ & -x_1 + 5x_2 + 8x_3 & = 10 \\ & x_1 \le 10,\ 0 \le x_3 & \end{array}$$

has dual

$$\begin{array}{lll} \text{minimize} & 8y_1 + 10y_2 + 10y_3 & \\ \text{subject to} & 5y_1 - y_2 + y_3 & = 1 \\ & y_1 + 5y_2 & = -2 \\ & -2y_1 + 8y_2 & \ge 3 \\ & 0 \le y_1,\ 0 \le y_3 & . \end{array}$$

The primal-dual pair $\mathcal{P}$ and $\mathcal{D}$ above are related by the following weak duality theorem.

**Theorem 4.9** *[General Weak Duality Theorem]*
*Let $A \in \mathsf{R}^{m\times n}$, $b \in \mathsf{R}^m$, and $c \in \mathsf{R}^n$. If $x \in \mathsf{R}^n$ is feasible for $\mathcal{P}$ and $y \in \mathsf{R}^m$ is feasible for $\mathcal{D}$, then*
$$c^T x \le y^T A x \le b^T y.$$
*Moreover, the following statements hold.*

*(i) If $\mathcal{P}$ is unbounded, then $\mathcal{D}$ is infeasible.*

*(ii) If $\mathcal{D}$ is unbounded, then $\mathcal{P}$ is infeasible.*

*(iii) If $\bar{x}$ is feasible for $\mathcal{P}$ and $\bar{y}$ is feasibe for $\mathcal{D}$ with $c^T\bar{x} = b^T\bar{y}$, then $\bar{x}$ is and optimal solution to $\mathcal{P}$ and $\bar{y}$ is an optimal solution to $\mathcal{D}$.*

Proof: Suppose $x \in \mathsf{R}^n$ is feasible for $\mathcal{P}$ and $y \in \mathsf{R}^m$ is feasible for $\mathcal{D}$. Then

$$\begin{aligned}
c^T x &= \sum_{j\in R} c_j x_j + \sum_{j\in F} c_j x_j \\
&\le \sum_{j\in R}(\sum_{i=1}^m a_{ij}y_i)x_j + \sum_{j\in F}(\sum_{i=1}^m a_{ij}y_i)x_j \\
&\qquad\qquad \text{(Since } c_j \le \textstyle\sum_{i=1}^n a_{ij}y_i \text{ and } x_j \ge 0 \text{ for } j \in R \\
&\qquad\qquad\quad \text{and } c_j = \textstyle\sum_{i=1}^n a_{ij}y_i \text{ for } j \in F.) \\
&= \sum_{i=1}^m \sum_{j=1}^n a_{ij} y_i x_j \\
&= y^T A x \\
&= \sum_{i\in I}(\sum_{j=1}^n a_{ij}x_j)y_i + \sum_{i\in E}(\sum_{j=1}^n a_{ij}x_j)y_i \\
&\le \sum_{i\in I} b_i y_i + \sum_{i\in E} b_i y_i \\
&\qquad\qquad \text{(Since } \textstyle\sum_{j=1}^n a_{ij}x_j \le b_i \text{ and } 0 \le y_i \text{ for } i \in I \\
&\qquad\qquad\quad \text{and } \textstyle\sum_{j=1}^n a_{ij}x_j = b_i \text{ for } i \in E. \\
&= \sum_{i=1}^m b_i y_i \\
&= b^T y\ .
\end{aligned}$$

■

## 4.3 The Dual Simplex Algorithm

Consider the linear program

$$
\begin{array}{lll}
\mathcal{P} & \text{maximize} & -4x_1 - 2x_2 - x_3 \\
 & \text{subject to} & -x_1 - x_2 + 2x_3 \leq -3 \\
 & & -4x_1 - 2x_2 + x_3 \leq -4 \\
 & & x_1 + x_2 - 4x_3 \leq 2 \\
 & & 0 \leq x_1, x_2, x_3 \ .
\end{array}
$$

The dual to this LP is

$$
\begin{array}{llr}
\mathcal{D} & \text{minimize} & -3y_1 - 4y_2 + 2y_3 \\
 & \text{subject to} & -y_1 - 4y_2 + y_3 \geq -4 \\
 & & -y_1 - 2y_2 + y_3 \geq -2 \\
 & & 2y_1 + y_2 - 4y_3 \geq -1 \\
 & & 0 \leq y_1, y_2, y_3 \ .
\end{array}
$$

The problem $\mathcal{P}$ does not have feasible origin, and so it appears that one must apply phase I of the two phase simplex algorithm to obtain an initial basic feasible solution. On the other hand, the dual problem $\mathcal{D}$ does have feasible origin, so why not just apply the simplex algorithm to $\mathcal{D}$ and avoid phase I altogether? This is exactly what we will do. However, we do it in a way that may at first seem odd. We *reverse* the usual simplex procedure by choosing a pivot row first, and then choosing the pivot column. The initial tableau for the problem $\mathcal{P}$ is

$$
\left|\begin{array}{rrrrrr|r}
-1 & -1 & 2 & 1 & 0 & 0 & -3 \\
-4 & -2 & 1 & 0 & 1 & 0 & -4 \\
1 & 1 & -4 & 0 & 0 & 1 & 2 \\
\hline
-4 & -2 & -1 & 0 & 0 & 0 & 0
\end{array}\right|
$$

A striking and important feature of this tableau is that every entry in the cost row is nonpositive! This is exactly what we are trying to achieve by our pivots in the simplex algorithm. This is a consequence of the fact that the dual problem $\mathcal{D}$ has feasible origin. Any tableau having this property we will call *dual feasible*. Unfortunately, the tableau is not feasible since some of the right hand sides are negative. Henceforth, we will say that such a tableau is not *primal feasible*. That is, instead of saying that a tableau (or dictionary) is feasible or infeasible in the usual sense, we will now say that the tableau is *primal feasible*, respectively, *primal infeasible*.

Observe that if a tableau is *both* primal and dual feasible, then it must be optimal, i.e. the basic feasible solution that it identifies is an optimal solution. We now describe an implementation of the simplex algorithm, called the *dual simplex algorithm*, that can be applied to tableaus that are dual feasible but not primal feasible. Essentially it is the simplex algorithm applied to the dual problem but using the tableau structure associated with the primal problem. The goal is to use simplex pivots to attain primal feasibility while maintaining dual feasibility.

Consider the tableau above. The right hand side coefficients are $-3$, $-4$, and 2. These correspond to the cost coefficients of the dual objective. Not that this tableau also identifies a basic feasible solution for the dual problem by setting the dual variable equal to the negative of the cost row coefficients associated with the slack variables:

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} .$$

The dual variables are currently nonbasic and so their values are zero. Next note that by increasing the value of either $y_1$ or $y_2$ we decrease the value of the dual objective since the coefficients of these variables are $-3$ and $-4$. In the simplex algorithm terminology, we can pivot on either the first or second row to decrease the value of the dual objective. Let's choose the first row as our pivot row. How do we choose the pivot column? Similar to the primal simplex algorithm, we choose the pivot column to maintain dual feasibility. For this we again must compute ratios, but this time it is the ratios of the negative entries in the pivot row with the corresponding cost row entries:

ratios for the first two columns are 4 and 2

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| -1 | $\boxed{-1}$ | 2 | 1 | 0 | 0 | -3 | ← pivot row |
| -4 | -2 | 1 | 0 | 1 | 0 | -4 | |
| 1 | 1 | -4 | 0 | 0 | 1 | 2 | |
| -4 | -2 | -1 | 0 | 0 | 0 | 0 | |

The smallest ratio is 2 so the pivot column is column 2 in the tableau, and the pivot element is therefore the (1,2) entry of the tableau. Note that this process of choosing the pivot is the reverse of how the pivot is chosen in the primal simplex algorithm. In the dual simplex algorithm we fist choose a pivot row, then compute ratios to determine the pivot column which identifies the pivot. We now successive apply this process to the above tableau until optimality is achieved.

$$
\begin{array}{|rrrrrr|r|l}
-1 & \boxed{-1} & 2 & 1 & 0 & 0 & -3 & \leftarrow \text{pivot row} \\
-4 & -2 & 1 & 0 & 1 & 0 & -4 & \\
1 & 1 & -4 & 0 & 0 & 1 & 2 & \\
\hline
-4 & -2 & -1 & 0 & 0 & 0 & 0 & \\
\hline
 & & & & & & & \\
1 & 1 & -2 & -1 & 0 & 0 & 3 & \\
-2 & 0 & -3 & -2 & 1 & 0 & 2 & \\
0 & 0 & \boxed{-2} & 1 & 0 & 1 & -1 & \leftarrow \text{pivot row} \\
\hline
-2 & 0 & -5 & -2 & 0 & 0 & 6 & \\
\hline
1 & 1 & 0 & -2 & 0 & -1 & 4 & \\
-2 & 0 & 0 & -7/2 & 1 & -3/2 & 7/2 & \\
0 & 0 & 1 & -1/2 & 0 & -1/2 & 1/2 & \\
\hline
-2 & 0 & 0 & -9/2 & 0 & -5/2 & 17/2 & \text{optimal}
\end{array}
$$

Therefore, the optimal solutions to $\mathcal{P}$ and $\mathcal{D}$ are

$$
\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 4 \\ 1/2 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 9/2 \\ 0 \\ 5/2 \end{pmatrix} ,
$$

respectively, with optimal value $-17/2$.

Next consider the LP

$$
\begin{array}{lll}
\mathcal{P} & \text{maximize} & -4x_1 - 2x_2 - x_3 \\
 & \text{subject to} & -x_1 - x_2 + 2x_3 \leq -3 \\
 & & -4x_1 - 2x_2 + x_3 \leq -4 \\
 & & x_1 + x_2 - x_3 \leq 2 \\
 & & 0 \leq x_1, x_2, x_3 \ .
\end{array}
$$

This LP differs from the previous LP only in the $x_3$ coefficient of the third linear inequality. Let's apply the dual simplex algorithm to this LP.

$$
\begin{array}{|rrrrrr|r|l}
-1 & \boxed{-1} & 2 & 1 & 0 & 0 & -3 & \leftarrow \text{pivot row} \\
-4 & -2 & 1 & 0 & 1 & 0 & -4 & \\
1 & 1 & -1 & 0 & 0 & 1 & 2 & \\
\hline
-4 & -2 & -1 & 0 & 0 & 0 & 0 & \\
\hline
 & & & & & & & \\
1 & 1 & -2 & -1 & 0 & 0 & 3 & \\
-2 & 0 & -3 & -2 & 1 & 0 & 2 & \\
0 & 0 & 1 & 1 & 0 & 1 & -1 & \leftarrow \text{pivot row} \\
\hline
-2 & 0 & -5 & -2 & 0 & 0 & 6 &
\end{array}
$$

The first dual simplex pivot is given above. Repeating this process again, we see that there is only one candidate for the pivot row in our dual simplex pivoting strategy. What do we do now? It seems as though we are stuck since there are no negative entries in the third row with which to compute ratios to determine the pivot column. What does this mean? Recall that we chose the pivot row because the negative entry in the right hand side implies that we can decrease the value of the dual objective by bring the dual variable $y_3$ into the dual basis. The ratios are computed to preserve dual feasibility. In this problem, the fact that there are no negative entries in the pivot row implies that we can increase the value of $y_3$ as much as we want without violating dual feasibility. That is, the dual problem is unbounded below, and so by the weak duality theorem the primal problem must be infeasible!

We will make extensive use of the dual simplex algorithm in our discussion of sensitivity analysis in linear programming.

# 5 LP Geometry

We now briefly turn to a discussion of LP geometry extending the geometric ideas developed in Section 1 for 2 dimensional LPs to $n$ dimensions. In this regard, the key geometric idea is the notion of a hyperplane.

**Definition 5.1** *A hyperplane in* $\mathsf{R}^n$ *is any set of the form*

$$H(a, \beta) = \{x : a^T x = \beta\}$$

*where* $a \in \mathsf{R}^n$, $\beta \in \mathsf{R}$, *and* $a \neq 0$.

We have the following important fact whose proof we leave as an exercise for the reader.

**Fact 5.2** $H \subset \mathsf{R}^n$ *is a hyperplane if and only if the set*

$$H - x_0 = \{x - x_0 : x \in H\}$$

*where* $x_0 \in H$ *is a subspace of* $\mathsf{R}^n$ *of dimension* $(n-1)$.

Every hyperplane $H(a, \beta)$ generates two closed half spaces:

$$H_+(a, \beta) = \{x \in \mathsf{R}^n : a^T x \geq \beta\}$$

and

$$H_-(a, \beta) = \{x \in \mathsf{R}^n : a^T x \leq \beta\}.$$

Note that the constraint region for a linear program is the intersection of finitely many closed half spaces: setting

$$H_j = \{x : e_j^T x \geq 0\} \qquad \text{for } j = 1, \ldots, n$$

and

$$H_{n+i} = \{x : \sum_{j=1}^{n} a_{ij} x_j \leq b_i\} \qquad \text{for } i = 1, \ldots, m$$

we have

$$\{x : Ax \leq b, 0 \leq x\} = \bigcap_{i=1}^{n+m} H_i.$$

Any set that can be represented in this way is called a *convex polyhedron.*

**Definition 5.3** *Any subset of* $\mathsf{R}^n$ *that can be represented as the intersection of finitely many closed half spaces is called a convex polyhedron.*

Therefore, a linear programming is simply the problem of either maximizing or minimizing a linear function over a convex polyhedron. We now develop some of the underlying geometry of convex polyhedra.

**Fact 5.4** *Given any two points in* $\mathsf{R}^n$*, say* $x$ *and* $y$*, the line segment connecting them is given by*

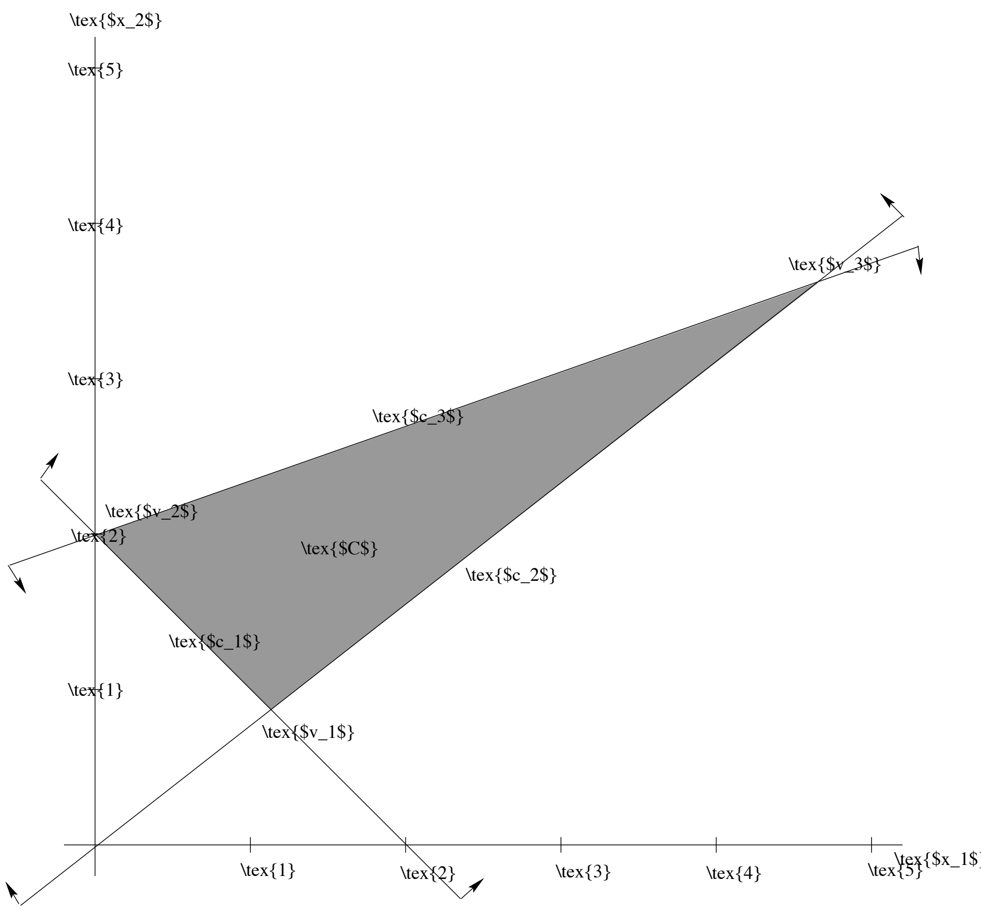
$$[x, y] = \{(1-\lambda)x + \lambda y : 0 \leq \lambda \leq 1\}.$$

**Definition 5.5** *A subset* $C \in \mathsf{R}^n$ *is said to be convex if* $[x, y] \subset C$ *whenever* $x, y \in C$*.*

**Fact 5.6** *A convex polyhedron is a convex set.*

We now consider the notion of vertex, or corner point, for convex polyhedra in $\mathsf{R}^2$. For this, consider the polyhedron $C \subset \mathsf{R}^2$ defined by the constraints

$$\begin{aligned} c_1 &: -x_1 - x_2 \leq -2 \\ c_2 &: 3x_1 - 4x_2 \leq 0 \\ c_3 &: -x_1 + 3x_2 \leq 6. \end{aligned} \tag{5.6}$$




The vertices are $v_1 = \left(\frac{8}{7}, \frac{6}{7}\right)$, $v_2 = (0, 2)$, and $v_3 = \left(\frac{24}{5}, \frac{18}{5}\right)$. One of our goals in this section is to discover an intrinsic geometric property of these vertices that generalizes to $n$ dimensions and simultaneously captures our intuitive notion of what a vertex is. For this we examine our notion of convexity which is based on line segments. Is there a way to use line segments to make precise our notion of vertex?

Consider any of the vertices in the polyhedron $C$ defined by (5.7). Note that any line segment in $C$ that contains one of these vertices must have the vertex as one of its end points. Vertices are the only points that have this property. In addition, this property easily generalizes to convex polyhedra in $\mathsf{R}^n$. This is the rigorous mathematical formulation for our notion of vertex that we seek. It is simple, has intuitive appeal, and yields the correct objects in dimensions 2 and 3.

**Definition 5.7** *Let $C$ be a convex polyhedron. We say that $x \in C$ is a vertex of $C$ if whenever $x \in [u, v]$ for some $u, v \in C$, it must be the case that either $x = u$ or $x = v$.*

This definition says that a point is a vertex if and only if whenever that point is a member of a line segment contained in the polyhedron, then it must be one of the end points of the line segment. In particular, this implies that vertices must lie in the boundary of the set and the set must somehow make a corner at that point. Our next result gives an important and useful characterization of the vertices of convex polyhedra.

**Theorem 5.8 (Fundamental Representation Theorem for Vertices)** *A point $x$ in the convex polyhedron $C = \{x \in \mathsf{R}^s \,|\, Tx \leq g\}$, where $T = (t_{ij})_{s \times n}$ and $g \in \mathsf{R}^s$, is a vertex of this polyhedron if and only if there exist an index set $\mathcal{I} \subset \{1, \ldots, s\}$ with such that $x$ is the unique solution to the system of equations*

$$\sum_{j=1}^{n} t_{ij} x_j = g_i \quad i \in \mathcal{I}. \tag{5.7}$$

*Moreover, if $x$ is a vertex, then one can take $|\mathcal{I}| = n$ in (5.7), where $|\mathcal{I}|$ denotes the number of elements in $\mathcal{I}$.*

Proof: We first prove that if there exist an index set $\mathcal{I} \subset \{1, \ldots, s\}$ such that $x = \bar{x}$ is the unique solution to the system of equations (5.7), then $\bar{x}$ is a vertex of the polyhedron $C$. We do this by proving the contraposition, that is, we assume that $\bar{x} \in C$ is not a vertex and show that it cannot be the unique solution to any system of the form (5.7) with $\mathcal{I} \subset \{1, 2, \ldots, s\}$.

If $\bar{x}$ is not a vertex of $C$, then there exist $u, v \in C$ and $0 < \lambda < 1$ such that $\bar{x} = (1 - \lambda)u + \lambda v$. Let $\mathsf{A}(x)$ denote the set of *active indices* at $x$:

$$\mathsf{A}(x) = \left\{ i \;\middle|\; \sum_{j=1}^{n} t_{ij} x_j = g_i \right\}.$$

For every $i \in \mathsf{A}(\bar{x})$

$$\sum_{j=1}^{n} t_{ij} \bar{x}_j = g_i, \; \sum_{j=1}^{n} t_{ij} u_j \leq g_i, \text{ and } \sum_{j=1}^{n} t_{ij} v_j \leq g_i. \tag{5.8}$$

Therefore,

$$\begin{aligned} 0 &= g_i - \sum_{j=1}^n t_{ij}\bar{x}_j \\ &= (1-\lambda)g_i + \lambda g_i - \sum_{j=1}^n t_{ij}((1-\lambda)u + \lambda v) \\ &= (1-\lambda)\left[g_i - \sum_{j=1}^n t_{ij}u_j\right] \ + \ \lambda\left[g_i - \sum_{j=1}^n t_{ij}v_j\right] \\ &\geq 0 \ . \end{aligned}$$

Hence,

$$0 = (1-\lambda)\left[g_i - \sum_{j=1}^n t_{ij}u_j\right] \ + \ \lambda\left[g_i - \sum_{j=1}^n t_{ij}v_j\right]$$

which implies that

$$g_i = \sum_{j=1}^n t_{ij}u_j \text{ and } g_i = \sum_{j=1}^n t_{ij}v_j$$

since both $\left[g_i - \sum_{j=1}^n t_{ij}u_j\right]$ and $\left[g_i - \sum_{j=1}^n t_{ij}v_j\right]$ are non-negative. That is, $\mathsf{A}(\bar{x}) \subset \mathsf{A}(u) \cap \mathsf{A}(v)$. Now if $\mathcal{I} \subset \{1, 2, \ldots, s\}$ is such that (5.7) holds at $x = \bar{x}$, then $\mathcal{I} \subset \mathsf{A}(\bar{x})$. But then (5.7) must also hold for $x = u$ and $x = v$ since $\mathsf{A}(\bar{x}) \subset \mathsf{A}(u) \cap \mathsf{A}(v)$. Therefore, $\bar{x}$ is not a unique solution to (5.7) for any choice of $\mathcal{I} \subset \{1, 2, \ldots, s\}$.

Let $\bar{x} \in C$. We now show that if $\bar{x}$ is a vertex of $C$, then there exist an index set $\mathcal{I} \subset \{1, \ldots, s\}$ such that $x = \bar{x}$ is the unique solution to the system of equations (5.7). Again we establish this by contraposition, that is, we assume that if $\bar{x} \in C$ is such that, for every index set $\mathcal{I} \subset \{1, 2, \ldots, s\}$ for which $x = \bar{x}$ satisfies the system (5.7) there exists $w \in \mathsf{R}^n$ with $w \neq \bar{x}$ such that (5.7) holds with $x = w$, then $\bar{x}$ cannot be a vertex of $C$ . To this end take $\mathcal{I} = \mathsf{A}(\bar{x})$ and let $w \in \mathsf{R}^n$ with $w \neq \bar{x}$ be such that (5.7) holds with $x = w$ and $\mathcal{I} = \mathsf{A}(\bar{x})$, and set $u = w - \bar{x}$. Since $\bar{x} \in C$, we know that

$$\sum_{j=1}^n t_{ij}\bar{x}_j < g_i \quad \forall \ i \in \{1, 2, \ldots, s\} \setminus \mathsf{A}(\bar{x}) \ .$$

Hence, by continuity, there exists $\tau \in (0, 1]$ such that

$$\sum_{j=1}^n t_{ij}(\bar{x}_j + tu_j) < g_i \quad \forall \ i \in \{1, 2, \ldots, s\} \setminus \mathsf{A}(\bar{x}) \text{ and } |t| \leq \bar{\tau}. \tag{5.9}$$

Also note that

$$\sum_{j=1}^n t_{ij}(\bar{x}_j \pm \tau u_j) = (\sum_{j=1}^n t_{ij}\bar{x}_j) \pm \tau(\sum_{j=1}^n t_{ij}w_j \ - \ \sum_{j=1}^n t_{ij}\bar{x}_j) = g_i \pm \tau(g_i - g_i) = g_i \ \forall \ i \in \mathsf{A}(\bar{x}).$$

Combining these equivalences with (5.9) we find that $\bar{x}+\tau u$ and $\bar{x}-\tau u$ are both in $C$. Since $x = \frac{1}{2}(x+\tau u) + \frac{1}{2}(x-\tau u)$ and $\tau u \neq 0$, $\bar{x}$ cannot be a vertex of $C$.

It remains to prove the final statement of the theorem. Let $\bar{x}$ be a vertex of $C$ and let $\mathcal{I} \subset \{1, 2, \ldots, s\}$ be such that $\bar{x}$ is the unique solution to the system (5.7). First note that since the system (5.7) is consistent and its solution unique, we must have $|\mathcal{I}| \geq n$; otherwise, there are infinitely many solutions since the system has a non-trivial null space when $n > |\mathcal{I}|$. So we may as well assume that $|\mathcal{I}| > n$. Let $\mathcal{J} \subset \mathcal{I}$ be such that the vectors $t_{i\cdot} = (t_{i1}, t_{i2}, \ldots, t_{in})^T$, $i \in \mathcal{J}$ is a maximally linearly independent subset of the set of vectors $t_{i\cdot} = (t_{i1}, t_{i2}, \ldots, t_{in})^T$, $i \in \mathcal{I}$. That is, the vectors $t_{i\cdot}$ $i \in \mathcal{J}$ form a basis for the subspace spanned by the vectors $t_{i\cdot}$, $i \in \mathcal{I}$. Clearly, $|\mathcal{J}| \leq n$ since these vectors reside in $\mathsf{R}^n$ and are linearly independent. Moreover, each of the vectors $t_{r\cdot}$ for $r \in \mathcal{I} \setminus \mathcal{J}$ can be written as a linear combination of the vectors $t_{i\cdot}$ for $i \in \mathcal{J}$;

$$t_{r\cdot} = \sum_{i \in J} \lambda_{ri} t_{i\cdot}, \ r \in \mathcal{I} \setminus \mathcal{J}.$$

Therefore,

$$g_r = t_{r\cdot}^T \bar{x} = \sum_{i \in \mathcal{J}} \lambda_{ri} t_{i\cdot}^T \bar{x} = \sum_{i \in \mathcal{J}} \lambda_{ri} g_i, \quad r \in \mathcal{I} \setminus \mathcal{J},$$

which implies that any solution to the system

$$t_{i\cdot}^T x = g_i, \ i \in \mathcal{J} \tag{5.10}$$

is necessarily a solution to the larger system (5.7). But then the smaller system (5.10) must have $\bar{x}$ as its unique solution; otherwise, the system (5.7) has more than one solution. Finally, since the set of solutions to (5.10) is unique and $|\mathcal{J}| \leq n$, we must in fact have $|\mathcal{J}| = n$ which completes the proof. ■

We now apply this result to obtain a characterization of the vertices for the constraint region of an LP in standard form.

**Corollary 5.1** *A point $x$ in the convex polyhedron described by the system of inequalities*

$$Ax \leq b \qquad \textit{and} \qquad 0 \leq x,$$

*where $A = (a_{ij})_{m \times n}$, is a vertex of this polyhedron if and only if there exist index sets $\mathcal{I} \subset \{1, \ldots, m\}$ and $\mathcal{J} \in \{1, \ldots, n\}$ with $|\mathcal{I}| + |\mathcal{J}| = n$ such that $x$ is the unique solution to the system of equations*

$$\begin{aligned} \sum_{j=1}^{n} a_{ij} x_j &= b_i \qquad i \in \mathcal{I}, \qquad \textit{and} \\ x_j &= 0 \qquad j \in \mathcal{J}. \end{aligned} \tag{5.9}$$

PROOF: Take

$$T = \begin{bmatrix} A \\ -I \end{bmatrix} \quad \text{and} \quad g \begin{bmatrix} b \\ 0 \end{bmatrix}$$

in the previous theorem. ■

Recall that the symbols $|\mathcal{I}|$ and $|\mathcal{J}|$ denote the number of elements in the sets $\mathcal{I}$ and $\mathcal{J}$, respectively. The constraint hyperplanes associated with these indices are necessarily a subset of the set of *active* hyperplanes at the solution to (5.9).

Theorem 5.1 is an elementary yet powerful result in the study of convex polyhedra. We make strong use of it in our study of the geometric properties of the simplex algorithm. As a first observation, recall from Math 308 that the coefficient matrix for the system (5.9) is necessarily non-singular if this $n \times n$ system has a unique solution. How do we interpret this system geometrically, and why does Theorem 5.1 make intuitive sense?

To answer these questions, let us return to the convex polyhedron $C$ defined by (5.7). In this case, the dimension $n$ is 2. Observe that each vertex is located at the intersection of precisely two of the bounding constraint lines. Thus, each vertex can be represented as the unique solution to a $2 \times 2$ system of equations of the form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1 \\ a_{21}x_1 + a_{22}x_2 &= b_2, \end{aligned}$$

where the coefficient matrix

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

is non-singular. For the set $C$ above, we have the following:

(a) The vertex $v_1 = (\frac{8}{7}, \frac{6}{7})$ is given as the solution to the system

$$\begin{aligned} -x_1 - x_2 &= -2 \\ 3x_1 - 4x_2 &= 0, \end{aligned}$$

(b) The vertex $v_2 = (0, 2)$ is given as the solution to the system

$$\begin{aligned} -x_1 - x_2 &= -2 \\ -x_1 + 3x_2 &= 6, \end{aligned}$$

and

(c) The vertex $v_3 = \left(\frac{24}{5}, \frac{18}{5}\right)$ is given as the solution to the system

$$\begin{aligned} 3x_1 - 4x_2 &= 0 \\ -x_1 + 3x_2 &= 6. \end{aligned}$$

Theorem 5.1 indicates that any subsystem of the form (5.9) for which the associated coefficient matrix is non-singular, has as its solution a vertex of the polyhedron

$$Ax \leq b, \qquad 0 \leq x \tag{5.10}$$

if this solution is in the polyhedron. We now connect these ideas to the operation of the simplex algorithm.

The system (5.10) describes the constraint region for an LP in standard form. It can be expressed componentwise by

$$\begin{array}{rcll} \sum_{j=1}^{n} a_{ij}x_j & \leq & b_i & \qquad i = 1, \ldots, m \\ 0 & \leq & x_j & \qquad j = 1, \ldots, n. \end{array}$$

The associated slack variables are defined by the equations

$$x_{n+i} = b_i - \sum_{j=1}^{n} a_{ij}x_j \qquad i = 1, \ldots, m. \tag{5.11}$$

Let $\bar{x} = (\bar{x}_1, \ldots, \bar{x}_{n+m})$ be any solution to the system (5.11) and set $\widehat{x} = (\bar{x}_1, \ldots, \bar{x}_n)$ ($\widehat{x}$ gives values for the decision variables associated with the underlying LP). Note that if for some $j \in \mathcal{J} \subset \{1, \ldots, n\}$ we have $\bar{x}_j = 0$, then the hyperplane

$$H_j = \{x \in \mathsf{R}^n : e_j^T x = 0\}$$

is *active* at $\widehat{x}$, i.e., $\widehat{x} \in H_j$. Similarly, if for some $i \in \mathcal{I} \subset \{1, 2, \ldots, m\}$ we have $\bar{x}_{n+i} = 0$, then the hyperplane

$$H_{n+i} = \{x \in \mathsf{R}^n : \sum_{j=1}^{n} a_{ij}x_j = b_i\}$$

is *active* at $\widehat{x}$, i.e., $\widehat{x} \in H_{n+i}$. Next suppose that $\bar{x}$ is a basic feasible solution for the LP

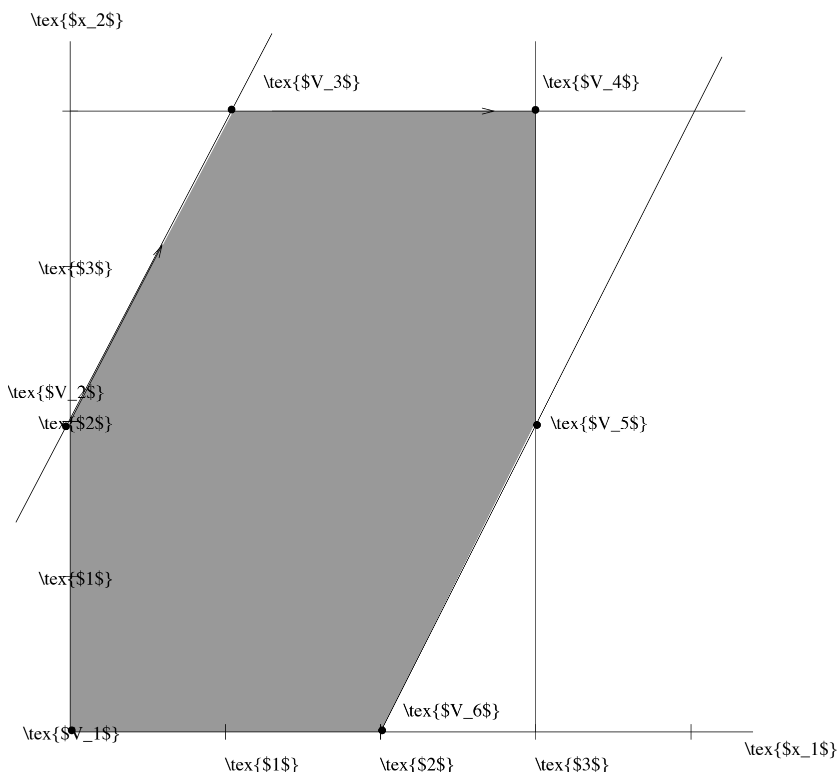
$$(\mathcal{P}) \qquad \begin{array}{l} \max c^T x \\ \text{subject to } Ax \leq b, 0 \leq x. \end{array}$$

Then it must be the case that $n$ of the components $\bar{x}_k$, $k \in \{1, 2, \ldots, n+m\}$ are assigned to the value zero since every dictionary has $m$ basic and $n$ non-basic variables. That is, every basic feasible solution is in the polyhedron defined by (5.10) and is the unique solution to a system of the form (5.9). But then, by Theorem 5.1, basic feasible solutions correspond precisely to the vertices of the polyhedron defining the constraint region for the LP $\mathcal{P}$!! This amazing geometric fact implies that the simplex algorithm proceeds by moving from vertex to adjacent vertex of the polyhedron given by (5.10). This is the essential underlying geometry of the simplex algorithm for linear programming!

By way of illustration, let us observe this behavior for the LP

$$
\begin{array}{lll}
\text{maximize} & 3x_1 + 4x_2 & \\
\text{subject to} & -2x_1 + x_2 & \leq 2 \\
& 2x_1 - x_2 & \leq 4 \\
& 0 \leq x_1 \leq 3, & 0 \leq x_2 \leq 4.
\end{array}
\tag{5.12}
$$

The constraint region for this LP is graphed on the next page.




The simplex algorithm yields the following pivots:

$$
\begin{array}{rrrrrr|r l}
-2 & \boxed{1} & 1 & 0 & 0 & 0 & 2 & \text{vertex} \\
2 & -1 & 0 & 1 & 0 & 0 & 4 & V_1 = (0,0) \\
1 & 0 & 0 & 0 & 1 & 0 & 3 & \\
0 & 1 & 0 & 0 & 0 & 1 & 4 & \\
\hline
3 & 4 & 0 & 0 & 0 & 0 & 0 & \\
\hline\hline
-2 & 1 & 1 & 0 & 0 & 0 & 2 & \text{vertex} \\
0 & 0 & 1 & 1 & 0 & 0 & 6 & V_2 = (0,2) \\
1 & 0 & 0 & 0 & 1 & 0 & 3 & \\
\boxed{2} & 0 & -1 & 0 & 0 & 1 & 2 & \\
\hline
11 & 0 & -4 & 0 & 0 & 0 & -8 & \\
\hline\hline
0 & 1 & 0 & 0 & 0 & 1 & 4 & \text{vertex} \\
0 & 0 & 1 & 1 & 0 & 0 & 6 & V_3 = (1,4) \\
0 & 0 & \boxed{\frac{1}{2}} & 0 & 1 & -\frac{1}{2} & 2 & \\
1 & 0 & -\frac{1}{2} & 0 & 0 & \frac{1}{2} & 1 & \\
\hline
0 & 0 & \frac{3}{2} & 0 & 0 & \frac{-11}{2} & -19 & \\
\hline\hline
0 & 1 & 0 & 0 & 0 & 1 & 4 & \text{vertex} \\
0 & 0 & 0 & 1 & -2 & 1 & 2 & V_4 = (3,4) \\
0 & 0 & 1 & 0 & 2 & -1 & 4 & \\
1 & 0 & 0 & 0 & 1 & 0 & 3 & \\
\hline
0 & 0 & 0 & 0 & -3 & -4 & -25 &
\end{array}
$$

## The Geometry of Degeneracy

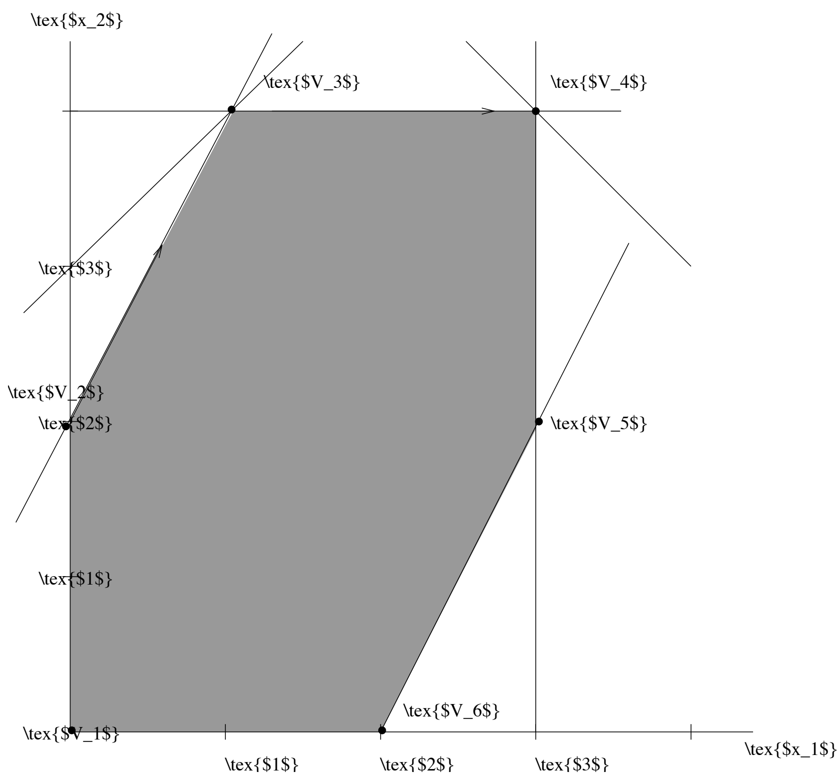
We now give a geometric interpretation of degeneracy in linear programming. Recall that a basic feasible solution, or vertex, is said to be degenerate if one or more of the basic variables is assigned the value zero. In the notation of (5.11) this implies that more than $n$ of the hyperplanes $H_k$, $k = 1, 2, \ldots, n+m$ are active at this vertex. By way of illustration, suppose we add the constraints

$$-x_1 + x_2 \leq 3$$

and

$$x_1 + x_2 \leq 7$$

to the system of constraints in the LP (5.12). The picture of the constraint region now looks as follows:




Notice that there are redundant constraints at both of the vertices $V_3$ and $V_4$. Therefore, as we pivot we should observe that the tableaus associated with these vertices are degenerate.

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| -2 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 2 | vertex |
| 2 | -1 | 0 | 1 | 0 | 0 | 0 | 0 | 4 | $V_1 = (0, 0)$ |
| -1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 3 | |
| 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 7 | |
| 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 3 | |
| 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 4 | |
| 3 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| -2 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 2 | vertex |
| 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 6 | $V_2 = (0, 2)$ |
| 1 | 0 | -1 | 0 | 1 | 0 | 0 | 0 | 1 | |
| 3 | 0 | -1 | 0 | 0 | 1 | 0 | 0 | 5 | |
| 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 3 | |
| 2 | 0 | -1 | 0 | 0 | 0 | 0 | 1 | 2 | |
| 11 | 0 | -4 | 0 | 0 | 0 | 0 | 0 | -8 | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 1 | -1 | 0 | 2 | 0 | 0 | 0 | 4 | vertex |
| 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 6 | $V_3 = (1, 4)$ |
| 1 | 0 | -1 | 0 | 1 | 0 | 0 | 0 | 1 | |
| 0 | 0 | 2 | 0 | -3 | 1 | 0 | 0 | 2 | |
| 0 | 0 | 1 | 0 | -1 | 0 | 1 | 0 | 2 | |
| 0 | 0 | 1 | 0 | -2 | 0 | 0 | 1 | 0 | degenerate |
| 0 | 0 | 7 | 0 | -11 | 0 | 0 | 0 | -19 | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 4 | vertex |
| 0 | 0 | 0 | 1 | 2 | 0 | 0 | 1 | 6 | $V_3 = (1, 4)$ |
| 1 | 0 | 0 | 0 | -1 | 0 | 0 | 1 | 1 | |
| 0 | 0 | 0 | 0 | 1 | 1 | 0 | -2 | 2 | |
| 0 | 0 | 0 | 0 | 1 | 0 | 1 | -1 | 2 | |
| 0 | 0 | 1 | 0 | -2 | 0 | 0 | 1 | 0 | degenerate |
| 0 | 0 | 0 | 0 | 3 | 0 | 0 | -7 | -19 | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 4 | vertex |
| 0 | 0 | 0 | 1 | 0 | -2 | 0 | 5 | 2 | $V_4 = (3, 4)$ |
| 1 | 0 | 0 | 0 | 0 | 1 | 0 | -1 | 3 | |
| 0 | 0 | 0 | 0 | 1 | 1 | 0 | -2 | 2 | optimal |
| 0 | 0 | 0 | 0 | 0 | -1 | 1 | 1 | 0 | degenerate |
| 0 | 0 | 1 | 0 | 0 | 2 | 0 | -3 | 4 | |
| 0 | 0 | 0 | 0 | 0 | -3 | 0 | -1 | -25 | |

In this way we see that a degenerate pivot arises when we represent the same vertex as the intersection point of a different subset of $n$ active hyperplanes. Cycling implies that we are cycling between different representations of the same vertex. In the example given above, the third pivot is a degenerate pivot. In the third tableau, we represent the vertex

$V_3 = (1, 4)$ as the intersection point of the hyperplanes

$$\begin{aligned} -2x_1 + x_2 &= 2 \qquad\qquad \text{(since } x_3 = 0\text{)} \\ \text{and} \qquad -x_1 + x_2 &= 3. \qquad\qquad \text{(since } x_5 = 0\text{)} \end{aligned}$$

The third pivot brings us to the 4th tableau where the vertex $V_3 = (1, 4)$ is now represented as the intersection of the hyperplanes

$$\begin{aligned} -x_1 + x_2 &= 3 \qquad\qquad \text{(since } x_5 = 0\text{)} \\ \text{and} \qquad x_2 &= 4 \qquad\qquad \text{(since } x_8 = 0\text{).} \end{aligned}$$

Observe that the final tableau is both optimal and degenerate. Just for the fun of it let's try pivoting on the only negative entry in the 5th row of this tableau (we choose the 5th row since this is the row that exhibits the degeneracy). Pivoting we obtain the following tableau.

$$\begin{array}{cccccccc|c} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 1 & 0 & 0 & -2 & 3 & 2 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & -1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 2 & -1 & 4 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & -2 & -3 & -25 \end{array}$$

Observe that this tableau is also optimal, but it provides us with a different set of optimal dual variables. In general, a degenerate optimal tableau implies that the dual problem has infinitely many optimal solutions.

FACT: If an LP has an optimal tableau that is degenerate, then the dual LP has infinitely many optimal solutions.

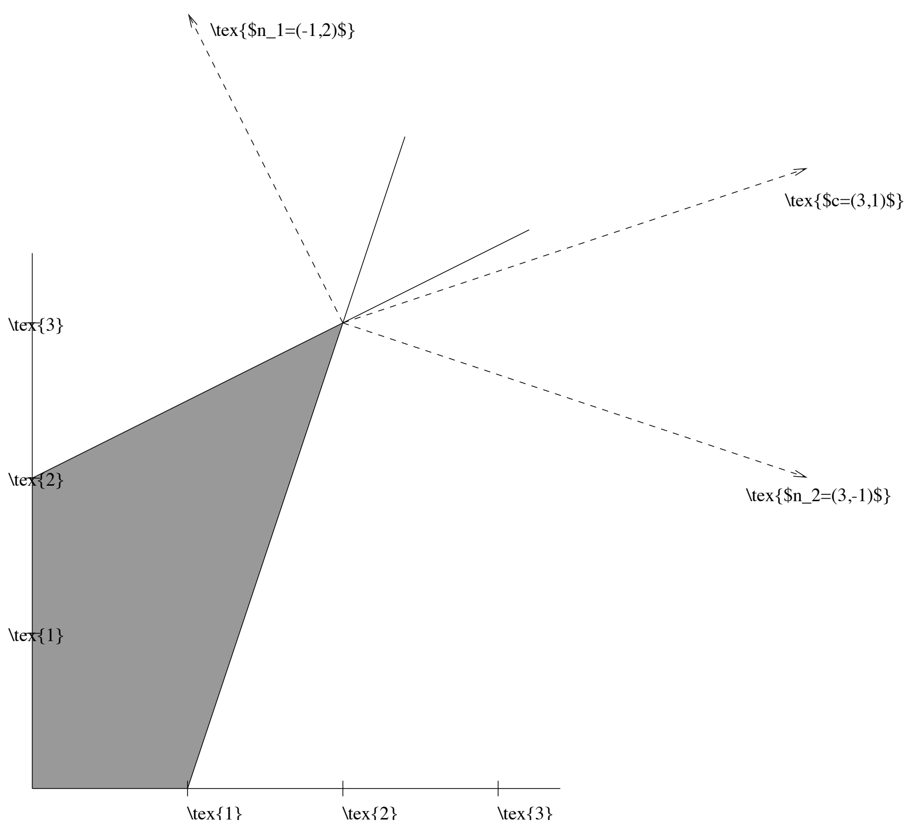
We will arrive at an understanding of why this fact is true after we examine the geometry of duality.

## The Geometry of Duality

Consider the linear program

$$(5.13) \qquad \begin{array}{lll} \text{maximize} & 3x_1 + x_2 & \\ \text{subject to} & -x_1 + 2x_2 & \leq 4 \\ & 3x_1 - x_2 & \leq 3 \\ & 0 \leq x_1, & x_2. \end{array}$$

This LP is solved graphically below.




The solution is $x = (2, 3)$. In the picture, the vector $n_1 = (-1, 2)$ is the normal to the hyperplane

$$-x_1 + 2x_2 = 4,$$

the vector $n_2 = (3, -1)$ is the normal to the hyperplane

$$3x_1 - x_2 = 3,$$

and the vector $c = (3, 1)$ is the objective normal. Geometrically, the vector $c$ lies *between* the vectors $n_1$ and $n_2$. That is to say, the vector $c$ can be represented as a non-negative linear combination of $n_1$ and $n_2$: there exist $y_1 \geq 0$ and $y_2 \geq 0$ such that

$$c = y_1 n_1 + y_2 n_2,$$

or equivalently,

$$\begin{aligned} \begin{pmatrix} 3 \\ 1 \end{pmatrix} &= y_1 \begin{pmatrix} -1 \\ 2 \end{pmatrix} + y_2 \begin{pmatrix} 3 \\ -1 \end{pmatrix} \\ &= \begin{bmatrix} -1 & 3 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}. \end{aligned}$$

Solving for $(y_1, y_2)$ we have

$$\begin{array}{rr|r} -1 & 3 & 3 \\ 2 & -1 & 1 \\ \hline 1 & -3 & -3 \\ 0 & 5 & 7 \\ \hline 1 & -3 & -3 \\ 0 & 1 & \frac{7}{5} \\ \hline 1 & 0 & \frac{6}{5} \\ 0 & 1 & \frac{7}{5} \end{array}$$

or $y_1 = \frac{6}{5}$, $y_2 = \frac{7}{5}$. I claim that the vector $y = (\frac{6}{5}, \frac{7}{5})$ is the optimal solution to the dual! Indeed, this result follows from the complementary slackness theorem and gives another way to recover the solution to the dual from the solution to the primal, or equivalently, to check whether a point that is feasible for the primal is optimal for the primal.

**Theorem 5.14 (Geometric Duality Theorem)** *Consider the LP*

$$(\mathcal{P}) \qquad \begin{array}{ll} \textit{maximize} & c^T x \\ \textit{subject to} & Ax \leq b, 0 \leq x. \end{array}$$

*where $A$ is an $m \times n$ matrix. Given a vector $\bar{x}$ that is feasible for $\mathcal{P}$, define*

$$\mathcal{Z}(\bar{x}) = \{j \in \{1, 2, \ldots, n\} : \bar{x}_j = 0\} \text{ and } \mathcal{E}(\bar{x}) = \{i \in \{1, \ldots, m\} : \sum_{j=1}^{n} a_{ij}\bar{x}_j = b_i\}.$$

*The indices $\mathcal{Z}(\bar{x})$ and $\mathcal{E}(\bar{x})$ are the* active *indices at $\bar{x}$ and correspond to the active hyperplanes at $\bar{x}$. Then $\bar{x}$ solves $\mathcal{P}$ if and only if there exist non-negative scalars $r_j$, $j \in \mathcal{Z}(\bar{x})$ and $y_i$, $i \in \mathcal{E}(\bar{x})$ such that*

$$c = -\sum_{j \in \mathcal{Z}(\bar{x})} r_j e_j + \sum_{i \in \mathcal{E}(\bar{x})} y_i a_{i\bullet} \tag{5.14}$$

*where for each* $i = 1, \ldots, m$, $a_{i\bullet} = (a_{i1}, a_{i2}, \ldots, a_{in})^T$ *is the* $i$*th column of the matrix* $A^T$*, and, for each* $j = 1, \ldots, n$, $e_j$ *is the* $j$*th unit coordinate vector. Moreover, the vector* $\bar{y} \in \mathsf{R}^m$ *given by*

$$\bar{y}_i = \begin{cases} y_i & \textit{for } i \in \mathcal{E}(\bar{x}) \\ 0 & \textit{otherwise} \end{cases} , \tag{5.15}$$

*solves the dual problem*

$$(\mathcal{D}) \qquad \begin{array}{ll} \textit{maximize} & b^T x \\ \textit{subject to} & A^T y \geq c, 0 \leq y. \end{array}$$

Proof: Let us first suppose that $\bar{x}$ solves $\mathcal{P}$. Then there is a $\bar{y} \in \mathsf{R}^n$ solving the dual $\mathcal{D}$ with $c^T\bar{x} = \bar{y}^T A\bar{x} = b^T\bar{y}$ by the Strong Duality Theorem. We need only show that there exist $r_j$, $j \in \mathcal{Z}(\bar{x})$ such that (5.14) and (5.15) hold. The Complementary Slackness Theorem implies that

$$\bar{y}_i = 0 \text{ for } i \in \{1, 2, \ldots, m\} \setminus \mathcal{E}(\bar{x}) \tag{5.16}$$

and

$$\sum_{i=1}^{m} \bar{y}_i a_{ij} = c_j \text{ for } j \in \{1, \ldots, n\} \setminus \mathcal{Z}(\bar{x}). \tag{5.17}$$

Note that (5.16) implies that $\bar{y}$ satisfies (5.15). Define $r = A^T\bar{y} - c$. Since $\bar{y}$ is dual feasible we have both $r \geq 0$ and $\bar{y} \geq 0$. Moreover, by (5.17), $r_j = 0$ for $j \in \{1, \ldots, n\} \setminus \mathcal{Z}(\bar{x})$, while

$$r_j = \sum_{i=1}^{n} \bar{y}_i a_{ij} - c_j \geq 0 \text{ for } j \in \mathcal{Z}(\bar{x}),$$

or equivalently,

$$c_j = -r_j + \sum_{i=1}^{m} \bar{y}_i a_{ij} \text{ for } j \in \mathcal{Z}(\bar{x}). \tag{5.18}$$

Combining (5.18) with (5.17) and (5.16) gives

$$c = -\sum_{j \in \mathcal{Z}(\bar{x})} r_j e_j \; + \; A^T\bar{y} = -\sum_{j \in \mathcal{Z}(\bar{x})} r_j e_j \; + \sum_{i \in \mathcal{E}(\bar{x})} \bar{y}_i a_{i\bullet},$$

so that (5.14) and (5.15) are satisfied with $\bar{y}$ solving $\mathcal{D}$.

Next suppose that $\bar{x}$ is feasible for $\mathcal{P}$ with $r_j$, $j \in \mathcal{Z}(\bar{x})$ and $\bar{y}_i$, $i \in \mathcal{E}(\bar{x})$ non-negative and satisfying (5.14). We must show that $\bar{x}$ solves $\mathcal{P}$ and $\bar{y}$ solves $\mathcal{D}$. Let $\bar{y} \in \mathsf{R}^m$ be such that its components are given by the $\bar{y}_i$'s for $i \in \mathcal{E}(\bar{x})$ and by (5.15) otherwise. Then the non-negativity of the $r_j$'s in (5.14) imply that

$$A^T\bar{y} = \sum_{i\in\mathcal{E}(\bar{x})} \bar{y}_i a_{i\bullet} \geq -\sum_{j\in\mathcal{Z}(\bar{x})} r_j e_j \; + \sum_{i\in\mathcal{E}(\bar{x})} \bar{y}_i a_{i\bullet} = c,$$

so that $\bar{y}$ is feasible for $\mathcal{D}$. Moreover,

$$c^T\bar{x} = -\sum_{j\in\mathcal{Z}(\bar{x})} r_j e_j^T \bar{x} \; + \sum_{i\in\mathcal{E}(\bar{x})} \bar{y}_i a_{i\bullet}^T \bar{x} = \bar{y}^T A\bar{x} = \bar{y}^T b,$$

where the final equality follows from the definition of the vector $\bar{y}$ and the index set $\mathcal{E}(\bar{x})$. Hence, by the Weak Duality Theorem $\bar{x}$ solves $\mathcal{P}$ and $\bar{y}$ solves $\mathcal{D}$ as required. ■

REMARK: As is apparent from the proof, the Geometric Duality Theorem is nearly equivalent to the complementary Slackness Theorem even though it provides a superficially different test for optimality.

We now illustrate how to apply this result with an example. Consider the LP

$$\begin{array}{llrrrrl}
 & \text{maximize} & x_1 & +x_2 & -x_3 & +2x_4 & \\
 & \text{subject to} & x_1 & +3x_2 & -2x_3 & +4x_4 & \leq -3 \\
(5.15) & & & 4x_2 & -2x_3 & +3x_4 & \leq 1 \\
 & & & -x_2 & +x_3 & -x_4 & \leq 2 \\
 & -x_1 & & -x_2 & +2x_3 & -x_5 & \leq 4 \\
 & 0 \leq & & x_1, & x_2, & x_3, & x_4 \qquad .
\end{array}$$

Does the vector $\bar{x} = (1, 0, 2, 0)^T$ solve this LP? If it does, then according to Theorem 5.14 we must be able to construct the solution to the dual of (5.15) by representing the objective vector $c = (1, 1, -1, 2)^T$ as a non-negative linear combination of the outer normals to the active hyperplanes at $\bar{x}$. Since the active hyperplanes are

$$\begin{array}{rcrcrcrcr}
x_1 & + & 3x_2 & - & 2x_3 & + & 4x_4 & = & -3 \\
 & - & x_2 & + & x_3 & - & x_4 & = & 2 \\
 & - & x_2 & & & & & = & 0 \\
 & & & & & - & x_4 & = & 0.
\end{array}$$

This means that $y_2 = y_4 = 0$ and $y_1$ and $y_3$ are obtained by solving

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 3 & -1 & -1 & 0 \\ -2 & 1 & 0 & 0 \\ 4 & -1 & 0 & -1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_3 \\ r_2 \\ r_4 \end{bmatrix} = \begin{pmatrix} 1 \\ 1 \\ -1 \\ 2 \end{pmatrix}.$$

Row reducing, we get

$$
\begin{array}{cccc|c}
1 & 0 & 0 & 0 & 1 \\
3 & -1 & -1 & 0 & 1 \\
-2 & 1 & 0 & 0 & -1 \\
4 & -1 & 0 & -1 & 2 \\
\hline
1 & 0 & 0 & 0 & 1 \\
0 & 1 & 1 & 0 & 2 \\
0 & 1 & 0 & 0 & 1 \\
0 & 1 & 0 & 1 & 2
\end{array}.
$$

Therefore, $y_1 = 1$ and $y_3 = 1$. We now check to see if the vector $\bar{y} = (1, 0, 1, 0)$ does indeed solve the dual to (5.15);

$$
\text{(5.16)} \qquad
\begin{array}{lrcrcrcrcr}
\text{minimize} & -3y_1 & + & y_2 & + & 2y_3 & + & 4y_4 & & \\
\text{subject to} & y_1 & & & & & - & y4 & \geq & 1 \\
 & 3y_1 & + & 4y_2 & - & y_3 & - & y_4 & \geq & 1 \\
 & -2y_1 & - & 2y_2 & + & y_3 & + & 2y_4 & \geq & -1 \\
 & 4y_1 & + & 3y_2 & - & y_3 & - & y_4 & \geq & 2 \\
 & \multicolumn{9}{l}{\quad 0 \leq y_1, y_2, y_3, y_4.}
\end{array}
$$

Clearly, $\bar{y}$ is feasible for (5.16). In addition,

$$
b^T \bar{y} = -1 = c^T \bar{x}.
$$

Therefore, $\bar{y}$ solves (5.16) and $\bar{x}$ solves (5.15) by the Weak Duality Theorem.

# 6 Sensitivity Analysis

In this section we study general questions involving the sensitivity of the solution to an LP under changes to its input data. As it turns out LP solutions can be extremely sensitive to such changes and this has very important practical consequences for the use of LP technology in applications. Let us now look at a simple example to illustrate this fact.

Consider the scenario where we be believe the federal reserve board is set to decrease the prime rate at its meeting the following morning. If this happens then bond yields will go up. In this environment, you have calculated that for every dollar that you invest today in bonds will give a return of a half percent tomorrow so, as a bond trader, you decide to invest in lots of bonds today. But to do this you will need to borrow money on margin. For the 24 hours that you intend to borrow the money you will need to place a reserve with the exchange that is un-invested, and then you can borrow up to 100 times this reserve. Regardless of how much you borrow, the exchange requires that you pay them back 10% of your reserve tomorrow. To add an extra margin of safety you will limit the sum of your reserve and one hundreth of what you borrow to be less than 200,000 dollars. Model the problem of determining how much money should be put on reserve and how much money should be borrowed to maximize your return on this 24 hour bond investment.

To model this problem, let $R$ denote your reserve in \$10,000 units and let $B$ denote the amount you borrow in the same units. Due to the way you must pay for the loan (i.e. it depends on the reserve, not what you borrow), your goal is to

$$\text{maximize } 0.005B - 0.1R\ .$$

Your borrowing constraint is

$$B \leq 100R\ ,$$

and your safety constraint is

$$\frac{B}{100} + R \leq 20\ .$$

The full LP model is

$$\begin{array}{ll} \text{maximize} & 0.005B - 0.1R \\ \text{subject to} & B - 100R \leq 0 \\ & 0.01B + R \leq 20 \\ & 0 \leq B,\ R\ . \end{array}$$

We conjecture that the solution occurs at the intersection of the two nontrivial constraint lines. We check this by applying the geometric duality theorem, i.e., we solve the system

$$\begin{pmatrix} 0.005 \\ -0.1 \end{pmatrix} = y_1 \begin{pmatrix} 1 \\ -100 \end{pmatrix} + y_2 \begin{pmatrix} 0.01 \\ 1 \end{pmatrix}$$

which gives $(y_1, y_2) = (0.003,\ 0.2)$. Since the solution is non-negative, the solution does occur at the intersection of the two nontrivial constraint lines giving $(B, R) = (1000,\ 10)$

with dual solution $(y_1, y_2) = (0.003,\ 0.2)$ and optimal value 4, or equivalently a profit of \$40,000 on a \$100,000 investment (the cost of the reserve).

But suppose that somehow your projections are wrong, and the Fed left rates alone and bond yields dropped by half a percent rather than increase by half a percent. In this scenario you would have lost $\$60,000$ on the \$100,000 investment. That is, the difference between a rise of the interest rate by half a percent to a drop in the interest rate by half a percent is one hundred thousand dollars. Clearly, this is a very risky investment opportunity. In this environment the downside risks must be fully understood before an investment is made. Doing this kind of analysis is called *sensitivity analysis*. We will look at some techniques for sensitivity analysis in this section. All of our discussion will be motivated by examples.

In practice, performing sensitivity analysis on solutions to LPs is absolutely essential. One should never report a solution to an LP without the accompanying sensitivity analysis. This is because all of the numbers defining the LP are almost always subject to error. The errors may be modeling errors, statistical errors, or data entry errors. Such errors can lead to catastrophically bad optimal solutions to the LP. Sensitivity analysis techniques provide tools for detecting and avoiding bad solutions.

## 6.1 Break-even Prices and Reduced Costs

The first type of sensitivity problem we consider concerns variations or *perturbations* to the objective coefficients. For this we consider the following LP problem.

**SILICON CHIP CORPORATION**

A Silicon Valley firm specializes in making four types of silicon chips for personal computers. Each chip must go through four stages of processing before completion. First the basic silicon wafers are manufactured, second the wafers are laser etched with a micro circuit, next the circuit is laminated onto the chip, and finally the chip is tested and packaged for shipping. The production manager desires to maximize profits during the next month. During the next 30 days she has enough raw material to produce 4000 silicon wafers. Moreover, she has 600 hours of etching time, 900 hours of lamination time, and 700 hours of testing time. Taking into account depreciated capital investment, maintenance costs, and the cost of labor, each raw silicon wafer is worth \$1, each hour of etching time costs \$40, each hour of lamination time costs \$60, and each hour of inspection time costs \$10. The production manager has formulated her problem as a profit maximization linear program with the following initial tableau:

| | $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ | $x_6$ | $x_7$ | $x_8$ | $b$ |
|---|---|---|---|---|---|---|---|---|---|
| raw wafers | 100 | 100 | 100 | 100 | 1 | 0 | 0 | 0 | 4000 |
| etching | 10 | 10 | 20 | 20 | 0 | 1 | 0 | 0 | 600 |
| lamination | 20 | 20 | 30 | 20 | 0 | 0 | 1 | 0 | 900 |
| testing | 20 | 10 | 30 | 30 | 0 | 0 | 0 | 1 | 700 |
| | 2000 | 3000 | 5000 | 4000 | 0 | 0 | 0 | 0 | 0 |

where $x_1$, $x_2$, $x_3$, $x_4$ represent the number of 100 chip batches of the four types of chips and the objective row coefficients for these variables correspond to dollars profit per 100 chip batch. After solving by the Simplex Algorithm, the final tableau is:

| $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ | $x_6$ | $x_7$ | $x_8$ | $b$ |
|---|---|---|---|---|---|---|---|---|
| 0.5 | 1 | 0 | 0 | .015 | 0 | 0 | $-.05$ | 25 |
| $-5$ | 0 | 0 | 0 | $-.05$ | 1 | 0 | $-.5$ | 50 |
| 0 | 0 | 1 | 0 | $-.02$ | 0 | .1 | 0 | 10 |
| 0.5 | 0 | 0 | 1 | .015 | 0 | $-.1$ | .05 | 5 |
| $-1500$ | 0 | 0 | 0 | $-5$ | 0 | $-100$ | $-50$ | $-145,000$ |

Thus the optimal production schedule is $(x_1, x_2, x_3, x_4) = (0,\ 25,\ 10,\ 5)$. In this solution we see that type 1 chip is not efficient to produce.

The first problem we address is to determine the sale price at which it is efficient to produce type 1 chip. That is, what sale price $p$ for which it is not efficient to produce type 1 chip below this sale price, but it is efficient to produce above this sale price? This is called the *breakeven sale price* of type 1 chip. As a first step let us compute the current sale price of type 1 chip. From the objective row we see that each 100 type 1 chip batch has a profit of \$2000. The cost of production of each 100 unit batch of type 1 chip is given by

$$\text{chip cost} + \text{etching cost} + \text{lamination cost} + \text{inspection cost},$$

where

$$\begin{aligned}
\text{chip cost} &= \text{no. chips} \times \text{cost per chip} = 100 \times 1 = 100 \\
\text{etching cost} &= \text{no. hours} \times \text{cost per hour} = 10 \times 40 = 400 \\
\text{lamination cost} &= \text{no. hours} \times \text{cost per hour} = 20 \times 60 = 1200 \\
\text{inspection cost} &= \text{no. hours} \times \text{cost per hour} = 20 \times 10 = 200\ .
\end{aligned}$$

Hence the costs per batch of 100 type 1 chips is \$1900. Therefore, the sale price of each batch of 100 type 1 chips is \$2000 + \$1900 = \$3900, or equivalently, \$39 per chip.

Since we do not produce type 1 chip in our optimal production mix, the breakeven sale price must be greater than \$39 per chip. Let $\theta$ denote the amount by which we need to increase the current sale price of type 1 chip so that it enters the optimal production mix. With this change to the sale price of type 1 chip the initial tableau for the LP becomes

| | $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ | $x_6$ | $x_7$ | $x_8$ | $b$ |
|---|---|---|---|---|---|---|---|---|---|
| raw wafers | 100 | 100 | 100 | 100 | 1 | 0 | 0 | 0 | 4000 |
| etching | 10 | 10 | 20 | 20 | 0 | 1 | 0 | 0 | 600 |
| lamination | 20 | 20 | 30 | 20 | 0 | 0 | 1 | 0 | 900 |
| testing | 20 | 10 | 30 | 30 | 0 | 0 | 0 | 1 | 700 |
| | $2000+\theta$ | 3000 | 5000 | 4000 | 0 | 0 | 0 | 0 | 0 |

.

Next let us suppose that we repeat on this tableau all of the pivots that led to the previously optimal tableau given above. What will the new tableau look like? That is, how does $\theta$ appear in this new tableau? This question is easily answered by recalling our general observations on simplex pivoting as left multiplication of an augmented matrix by a sequence of Gaussian elimination matrices.

Recall that given a problem in standard form,

$$\begin{array}{lll} \mathcal{P} & \text{maximize} & c^T x \\ & \text{subject to} & Ax \le b,\ 0 \le x\ , \end{array}$$

the initial tableau is an augmented matrix whose block form is given by

$$\begin{bmatrix} A & I & b \\ c^T & 0 & 0 \end{bmatrix} . \tag{6.1}$$

Pivoting to an optimal tableau corresponds to left multiplication by a matrix of the form

$$G = \begin{bmatrix} R & 0 \\ -y^T & 1 \end{bmatrix}$$

where the nonsingular matrix $R$ is called the *record matrix* and where the block form of $G$ is conformal with that of the initial tableau. Hence the optimal tableau has the form

$$\begin{bmatrix} R & 0 \\ -y^T & 1 \end{bmatrix} \begin{bmatrix} A & I & b \\ c^T & 0 & 0 \end{bmatrix} = \begin{bmatrix} RA & R & Rb \\ (c - A^T y)^T & -y^T & -b^T y \end{bmatrix} , \tag{6.2}$$

where $0 \le y,\ A^T y \ge c$, and the optimal value is $b^T y$. Now changing the value of one (or more) of the objective coefficients $c$ corresponds to replacing $c$ by a vector of the form $c + \Delta c$. The corresponding new initial tableau is

$$\begin{bmatrix} A & I & b \\ (c + \Delta c)^T & 0 & 0 \end{bmatrix} . \tag{6.3}$$

Performing the same simplex pivots on this tableau as before simply corresponds to left multiplication by the matrix $G$ given above. This yields the simplex tableau

$$\begin{bmatrix} RA & R & Rb \\ (c + \Delta c - A^T y)^T & -y^T & -b^T y \end{bmatrix} = \begin{bmatrix} RA & R & Rb \\ \Delta c^T + (c - A^T y)^T & -y^T & -b^T y \end{bmatrix} . \tag{6.4}$$

That is, we just add $\Delta c$ to the objective row in the old optimal tableau. Observe that this matrix may or may not be a simplex tableau since some of the basic variable cost row coefficients in $c - A^T y$ (which are zero) may be non-zero in $\Delta c + (c - A^T y)$. To completely determine the effect of the perturbation $\Delta c$ one must first use Gaussian elimination to return the basic variabl coefficients in $\Delta c + (c - A^T y)$ to zero. After returning (6.4) to a simplex tableau, the resulting tableau is optimal if it is dual feasible, that is, if all of the objective row coefficients are non-positive. These non-positivity conditions place restrictions on how large the entries of $\Delta c$ can be before one must pivot to obtain the new optimal tableau.

Let us apply these observations to the Silicon Chip Corp. problem and the question of determining the breakeven sale price of type 1 chip. In this case the expression $\Delta c$ takes the form $\Delta c = \theta e_1$, where $e_1$ is the first unit coordinate vector. Plugging this into (6.4) gives

$$(c - A^T y) + \Delta c = \begin{pmatrix} -1500 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} \theta \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \theta - 1500 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Therefore, the perturbed tableau (6.4) remains optimal if and only if $\theta \leq 1500$. That is, as soon as $\theta$ increases beyond 1500, type 1 chip enters the optimal production mix, and for $\theta = 1500$ we obtain multiple optimal solutions where type 1 chip may be in the optimal production mix if we so choose. The number 1500 appearing in the optimal objective row is called the *reduced cost* for type 1 chip. In general, the negative of the objective row coefficient for decision variables in the optimal tableau are the reduced costs of these variables. The reduced cost of a decision variable is the precise amount by which one must increase its objective row coefficient in order for it to be included in the optimal solution. Therefore, for nonbasic variables one can compute breakeven sale prices by simply reading off the reduced costs from the optimal tableau. In the case of the type 1 chip in the Silicon Chip Corp. problem above, this gives a breakeven sale price of

$$\begin{aligned} \text{breakeven price} &= \text{current price} + \text{reduced cost} \\ &= \$39 + \$15 = \$54 \ . \end{aligned}$$

With the long winded derivation of breakeven prices behind us, let us now consider a more intuitive and simpler explanation. One way to determine the breakeven sale price, is to determine by how much our profit is reduced if we produce one batch of these chips. Recall that the objective row coefficients in the optimal tableau correspond to the following expression for the objective variable $z$:

$$z = 145000 - 1500x_1 - 5x_5 - 100x_7 - 50x_8.$$

Hence, if we make one batch of type 1 chip, we reduce our optimal value by \$1500. Thus, to recoup this loss we must charge \$1500 more for these chips yielding a breakeven sale price of $\$39 + \$15 = \$54$ per chip.

## 6.2 Range Analysis for Objective Coefficients

Range analysis is a tool for understanding the effects of both objective coefficient variations as well as resource availability variations. In this section we examine objective coefficient variations. In the previous section we studied the idea of break-even sale prices. These prices are associated with activities that do not play a role in the currently optimal production schedule. In computing a breakeven price one needs to determine the change in the associated objective coefficient that make it efficient to introduce this activity into the optimal production mix, or equivalently, to determine the smallest change in the objective coefficient of this currently nonbasic decision variable that requires one to bring it into the basis in order to maintain optimality. A related question that can be asked of any of the objective coefficients is *what is the range of variation of a given objective coefficient that preserves the current basis as optimal*? The answer to this question is an interval, possibly unbounded, on the real line within which a given objective coefficient can vary but these variations do not effect the currently optimal basis.

For example, consider the objective coefficient on type 1 chip analyzed in the previous section. The range on this objective coefficient is $(-\infty, 3500]$ since within this range one need not change the basis to preserve optimality. Note that for any nonbasic decision variable $x_i$ the range at optimality is given by $(-\infty, c_i + r_i]$ where $r_i$ is the reduced cost of this decision variable in the optimal tableau.

How does one compute the range of a basic decision variable? That is, if an activity is currently optimal, what is the range of variations in its objective coefficient within which the optimal basis does not change. The answer to this question is again easily derived by considering the effect of an arbitrary perturbation to the objective vector $c$. For this we again consider the perturbation $\Delta c$ to $c$ and the associated initial tableau (6.3). This tableau yields the perturbed optimal tableau (6.4). When computing the range for the objective coefficient of a optimal basic variable $x_i$ one sets $\Delta c = \theta e_i$.

For example, in the Silicon Chip Corp. problem the decision variable $x_3$ associated with type 3 chips is in the optimal basis. For what range of variations in $c_3 = 5000$ does the current optimal basis $\{x_2, x_3, x_4, x_6\}$ remain optimal? Setting $\Delta c = \theta e_3$ in (6.4) we get the perturbed tableau

$$\begin{array}{cccccccc|c}
x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & x_7 & x_8 & b \\
0.5 & 1 & 0 & 0 & .015 & 0 & 0 & -.05 & 25 \\
-5 & 0 & 0 & 0 & -.05 & 1 & 0 & -.5 & 50 \\
0 & 0 & 1 & 0 & -.02 & 0 & .1 & 0 & 10 \\
0.5 & 0 & 0 & 1 & .015 & 0 & -.1 & .05 & 5 \\
\hline
-1500 & 0 & \theta & 0 & -5 & 0 & -100 & -50 & -145,000
\end{array} .$$

This augmented matrix is no longer a simplex tableau since the objective row coefficient of one of the basic variables, namely $x_3$, is not zero. To convert this to a proper simplex tableau we must eliminate $\theta$ from the objective row entry under $x_3$. Multiplying the third row by $-\theta$ and adding to the objective row gives the tableau

$$\begin{array}{cccccccc|c}
x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & x_7 & x_8 & b \\
0.5 & 1 & 0 & 0 & .015 & 0 & 0 & -.05 & 25 \\
-5 & 0 & 0 & 0 & -.05 & 1 & 0 & -.5 & 50 \\
0 & 0 & 1 & 0 & -.02 & 0 & .1 & 0 & 10 \\
0.5 & 0 & 0 & 1 & .015 & 0 & -.1 & .05 & 5 \\
\hline
-1500 & 0 & 0 & 0 & -5 + 0.02\theta & 0 & -100 - 0.1\theta & -50 & -145,000 - 10\theta
\end{array} .$$

For this tableau to remain optimal it must be both primal and dual feasible. Obviously primal feasibility is not an issue, but dual feasibility is due to the presence of $\theta$ in the objective row. For dual feasibility to be preserved the entries in the objective row must remain nonpositive; otherwise, a primal simplex pivot must be taken which will alter the currently optimal basis. That is, to preserve the current basis as optimal, we must have

$$\begin{array}{rcr} -5 + 0.02\theta \le 0, & \text{or equivalently,} & \theta \le 250 \\ -100 - 0.1\theta \le 0, & \text{or equivalently,} & -1000 \le \theta \end{array} .$$

Thus, the range of $\theta$ that preserves the current basis as optimal is

$$-1000 \le \theta \le 250,$$

and the corresponding range for $c_3$ that preserves the current basis as optimal is

$$4000 \le c_3 \le 5250.$$

Similarly, we can consider the range of the objective coefficient for type 4 chips. For this we simply multiply the fourth row by $-\theta$ and add it to the objective row to get the new objective row

$$-1500 - 0.5\theta \quad 0 \quad 0 \quad 0 \quad -5 - 0.015\theta \quad 0 \quad -100 + 0.1\theta \quad -50 - 0.05\theta \quad | \quad -145,000 - 5\theta \quad .$$

Again, to preserve dual feasibility we must have

$$\begin{array}{rcr} -1500 - 0.5\theta \le 0, & \text{or equivalently,} & -3000 \le \theta \\ -5 - 0.015\theta \le 0, & \text{or equivalently,} & -333.\bar{3} \le \theta \\ -100 + 0.1\theta \le 0, & \text{or equivalently,} & \theta \le 1000 \\ -50 - 0.05\theta \le 0, & \text{or equivalently,} & -1000 \le \theta \end{array} .$$

Thus, the range of $\theta$ that preserves the current basis as optimal is

$$-333.\bar{3} \le \theta \le 1000,$$

and the corresponding range for $c_4$ that preserves the current basis as optimal is

$$3666.\bar{6} \le c_4 \le 5000.$$

## 6.3 Resource Variations, Marginal Values, and Range Analysis

We now consider questions concerning the effect of resource variations on the optimal solution. We begin with a concrete instance of such a problem in the case of the Silicon Chip Corp. problem above.

Suppose we wish to purchase more silicon wafers this month. Before doing so, we need to answer three obvious questions.

a) How many should we purchase?

b) What is the most that we should pay for them?

c) After the purchase, what is the new optimal production schedule?

The technique we develop for answering these questions is similar to the technique used to determine objective coefficient ranges.. We begin by introducing a variable $\theta$ for the number of silicon wafers that will be purchased, and then determine how this variable appears in the tableau after using the same simplex pivots encoded in the matrix $G$ given above. In this case the new initial tableau looks like

| | $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ | $x_6$ | $x_7$ | $x_8$ | $b$ |
|---|---|---|---|---|---|---|---|---|---|
| raw wafers | 100 | 100 | 100 | 100 | 1 | 0 | 0 | 0 | $4000+\theta$ |
| etching | 10 | 10 | 20 | 20 | 0 | 1 | 0 | 0 | 600 |
| lamination | 20 | 20 | 30 | 20 | 0 | 0 | 1 | 0 | 900 |
| testing | 20 | 10 | 30 | 30 | 0 | 0 | 0 | 1 | 700 |
| | 2000 | 3000 | 5000 | 4000 | 0 | 0 | 0 | 0 | 0 |

.

In the general case, we are discussing perturbations, or variations, to the resource vector $b$ in the LP $\mathcal{P}$ in standard form given above. If we let $\Delta b$ denote this variation, then the associated initial tableau is

$$\begin{bmatrix} A & I & b+\Delta b \\ c^T & 0 & 0 \end{bmatrix} .$$

Again, multiplying on the left by the matrix $G$ gives

$$\begin{bmatrix} RA & R & Rb+R\Delta b \\ (c-A^Ty)^T & -y^T & -y^Tb-y^T\Delta b \end{bmatrix} .$$

This time the terms $R\Delta b$ and $y^T\Delta b$ encode the complete change to the optimal tableau by introducing the perturbation $\Delta b$. Clearly this new tableau is dual feasible and so it remains optimal as long as it remains primal feasible. That is, the new tableau is optimal as long as $0 \le Rb + R\Delta b$, or equivalently,

$$-Rb \le R\Delta b \; . \tag{6.5}$$

These inequalities place restrictions on the values $\Delta b$ may take and still preserve the optimality of the tableau. If (6.5) holds, then the new optimal value is $y^Tb + y^T\Delta b$. That is, the rate of change in the optimal value is given by the vector $y$, the solution to the dual LP.

In the case of the Silicon Chip Corp. problem where we are interested in varying the number of silicon wafers available, we have $\Delta b = \theta e_1$ and the matrix $R$ and vector $y$ are given by

$$R = \begin{bmatrix} .015 & 0 & 0 & -.05 \\ -.05 & 1 & 0 & -.5 \\ -.02 & 0 & .1 & 0 \\ .015 & 0 & -.1 & .05 \end{bmatrix} \quad \text{and} \quad y = \begin{pmatrix} 5 \\ 0 \\ 100 \\ 50 \end{pmatrix} .$$

Therefore, the inequality (6.5) takes the form

$$-\begin{pmatrix} 25 \\ 50 \\ 10 \\ 5 \end{pmatrix} \le \theta \begin{pmatrix} 0.015 \\ -0.05 \\ -0.02 \\ 0.015 \end{pmatrix},$$

or equivalently,

$$\begin{array}{rcrlcrr} -25 & \le & .015\theta & \text{implies} & \theta & \ge & -5000/3 \\ -50 & \le & -.05\theta & \text{implies} & \theta & \le & 1000 \\ -10 & \le & -.02\theta & \text{implies} & \theta & \le & 500 \\ -5 & \le & .015\theta & \text{implies} & \theta & \ge & -1000/3\ , \end{array}$$

which reduces to the simple inequality

$$-\frac{1000}{3} \le \theta \le 500.$$

This is the interval on which we may vary $\theta$ and not change the optimal basis. This interval is called the *range of the raw chip resource in the optimal solution.* If the variation $\theta$ stays within this interval, then the optimal solution is given by

$$\begin{pmatrix} x_2 \\ x_6 \\ x_3 \\ x_4 \end{pmatrix} = Rb + R\Delta b = \begin{pmatrix} 25 + .015\theta \\ 50 - .05\theta \\ 10 - .02\theta \\ 5 + .015\theta \end{pmatrix}$$

with optimal value

$$y^T b + y^T \Delta b = 145000 + 5\theta.$$

Observe from the expression for the optimal value that the profit increases by \$5 for every new silicon wafer that we get (up to 500 wafers). That is, if we pay less than \$5 over current costs for new wafers, then our profit increases. The dual value 5 is called the *shadow price*, or *marginal value*, for the raw silicon wafer resource. It represents the increased value of this resource due to the production process. It tells us the rate at which the optimal value increases due to increases in this resource. For example, we know that we currently pay \$1 per wafer. If another vendor offers wafers to us for \$2.50 per wafer, then we should buy them since our unit increase in profit with this purchase price is $\$5 - \$1.5 = \$3.5$ since \$2.5 is \$1.5 greater than the \$1 we now pay. So in answer to the questions we started out this discussion with, it seems that we should purchase 500 raw wafers at a purchase price of no more than $\$5 + \$1 = \$6$ dollars per wafer. With this purchase the new optimal production schedule is

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 25 + .015\theta \\ 10 - .02\theta \\ 5 + .015\theta \end{pmatrix}_{\theta=500} = \begin{pmatrix} 0 \\ 32.5 \\ 0 \\ 12.5 \end{pmatrix}.$$

Is this all of the wafers we should purchase? The answer to this question is not immediately obvious. This is because all we know about the range of values $-\frac{1000}{3} \le \theta \le 500$ is that if we move $\theta$ beyond these range boundaries, then the optimal basis will change. In particular, moving $\theta$ above 500 will introduce a negative entry in the third row of the simplex tableau. But the tableau will remain dual feasible. Hence to determine then new optimal solution for $\theta$ above 500 we must perform a dual simplex pivot in the third row. We can formally perform this pivot with the variable $\theta$ staying in the tableau:

$$
\begin{array}{rrrrrrrr|r}
x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & x_7 & x_8 & b \\
0.5 & 1 & 0 & 0 & .015 & 0 & 0 & -.05 & 25 + .015\theta \\
-5 & 0 & 0 & 0 & -.05 & 1 & 0 & -.5 & 50 - .05\theta \\
0 & 0 & 1 & 0 & -.02 & 0 & .1 & 0 & 10 - .02\theta \\
0.5 & 0 & 0 & 1 & .015 & 0 & -.1 & .05 & 5 + .015\theta \\
\hline
-1500 & 0 & 0 & 0 & -5 & 0 & -100 & -50 & -145,000 - 5\theta \\
\hline
\\
0.5 & 1 & .75 & 0 & 0 & 0 & .075 & -.05 & 32.5 \\
-5 & 0 & -2.5 & 0 & 0 & 1 & -.25 & -.5 & 25 \\
0 & 0 & -50 & 0 & 1 & 0 & -5 & 0 & -500 + \theta \\
0.5 & 0 & .75 & 1 & 0 & 0 & -.025 & .05 & 12.5 \\
\hline
-1500 & 0 & -250 & 0 & 0 & 0 & -125 & -50 & -147500
\end{array}
.
$$

Observe that for $\theta > 500$ we have pivoted to an optimal tableau with the slack for raw silicon wafers basic. Hence we cannot use any more wafers and their shadow price has fallen to zero. The new optimal solution is

$$
\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 32.5 \\ 0 \\ 12.5 \end{pmatrix},
$$

and this solution persists regardless of how many more raw silicon wafers we get. The final conclusion is that we should only buy 500 raw wafers at a price less than 6 per wafer.

Let us briefly review the range analysis for the right hand side coefficients. In the range analysis for a right hand side coefficient the goal is to determine the range of variation in a particular right hand side coefficient $b_i$ within which the optimal basis does not change. In this regard it is very similar to objective coefficient range analysis, but in this case we add $\theta$ times the column associated with the slack variable $s_i$ (or $x_{n+i}$) to the right hand side coefficients in the optimal tableau and then determine the variations in $\theta$ that preserve the primal feasibility of this tableau.

In the discussion above we computed the range for $b_1$, or the raw wafer resource. Let us now do a range analysis on $b_2$ the etching time resource. Note that this resource is slack in the optimal tableau since it is in the basis. Regardless, the new right hand side resulting from the perturbation $\Delta b = \theta e_2$ to $b$ is

$$Rb + R\Delta b = \begin{pmatrix} 25 \\ 50+\theta \\ 10 \\ 5 \end{pmatrix} .$$

To preserve primal feasibility we only require $0 \leq 50+\theta$, or equivalently, $-50 \leq \theta$. Therefore, the range for $b_2$ is

$$[550, \ +\infty) .$$

The upper bound of $+\infty$ make sense since etching time is already slack so any more etching time won't make any difference. The lower bound of 550 implies that if etching time drops to 550 or less, then this constraint will be binding in the optimal tableau.

Similarly, we can compute the range for the lamination and testing time resources. For example, to compute the range of the lamination time resource we simply add $\theta$ times the $x_7$ column to the optimal right hand side to get

$$Rb + R\Delta b = \begin{pmatrix} 25 \\ 50 \\ 10+0.1\theta \\ 5-0.1\theta \end{pmatrix} .$$

To preserve primal feasibility we must have

$$\begin{array}{rcr} 0 \leq 10+0.1\theta, & \text{or equivalently,} & -100 \leq \theta \\ 0 \leq 5-0.1\theta, & \text{or equivalently,} & \theta \leq 50 \end{array} .$$

Therefore, the range on $b_3$ is

$$800 \leq b_3 \leq 950 .$$

## 6.4 Pricing Out New Products

Next we consider the problem of adding a new product to our product line. In the context of the Silicon Chip Corp. problem, we consider a new chip that requires ten hours each of etching, lamination, and testing time per 100 chip batch. If it can be sold for \$ 33.10 per chip, we would like to know the answer to the following two questions:

(a) Is it efficient to produce?

(b) If it efficient to produce, what is the new production schedule?

We analyze this problem in the same way that we analyzed the two previous problems. That is, we first determine how this new chip effects the original initial tableau, and then see how the original pivoting process effects the new initial tableau by multiplying this new tableau on the left by the matrix $G$.

In the context of a new product, the original initial tableau is altered by the addition of a new column:

$$\begin{bmatrix} a_{\text{new}} & A & I & b \\ c_{\text{new}} & c^T & 0 & 0 \end{bmatrix} .$$

Again, multiplying on the left by the matrix $G$ gives

$$\begin{bmatrix} Ra_{\text{new}} & RA & R & Rb \\ c_{\text{new}} - a_{\text{new}}^T y & (c - A^T y)^T & -y^T & -y^T b \end{bmatrix} . \tag{6.6}$$

The expression $(c_{\text{new}} - a_{\text{new}}^T y)$ determines whether this new tableau is optimal or not. The act of forming this expression is called *pricing out* the new product. If this number is non-positive, then the new product does not price out, and we do not produce it since in this case the new tableau is optimal with the new product nonbasic. If, on the other hand, $(c_{\text{new}} - a_{\text{new}}^T y) > 0$, then we say that the new product does price out and it should be introduced into the optimal production mix. The new optimal production mix is found by applying the standard primal simplex algorithm to the tableau (6.6) since this tableau is primal feasible but not dual feasible.

Let us return to the Silicon Chip Corp. problem and the new chip under consideration. In this case we have

$$a_{\text{new}} = \begin{pmatrix} 100 \\ 10 \\ 10 \\ 10 \end{pmatrix} .$$

We also need to compute $c_{\text{new}}$. The stated sale price or revenue for each 100 chip batch of the new chip is \$3310. We need to subtract from this number the cost of producing each 100 chip batch. Recall from the Silicon Chip Corp. problem statement that each raw silicon wafer is worth \$1, each hour of etching time costs \$40, each hour of lamination time costs \$60, and each hour of inspection time costs \$10. Therefore, the cost of producing each 100 chip batch of these new chips is

$$\begin{array}{rl} 100 & \text{(cost of the raw wafers)} \\ +10 \times 40 & \text{(cost of etching time)} \\ +10 \times 60 & \text{(cost of lamination time)} \\ +10 \times 10 & \text{(cost of testing time)} \\ \hline 1200 & \text{(total cost)} . \end{array}$$

Hence the profit on each 100 chip batch of these new chips is \$3310 − \$1200 = \$2110, or \$21.10 per chip, and so $c_{\text{new}} = 2110$. Pricing out the new chip gives

$$c_{\text{new}} - a_{\text{new}}^T y = 2110 - \begin{pmatrix} 100 \\ 10 \\ 10 \\ 10 \end{pmatrix}^T \begin{pmatrix} 5 \\ 0 \\ 100 \\ 50 \end{pmatrix} = 2110 - 2000 = 110 ,$$

which is positive, and so this chip prices out. The new column in the tableau associated with this chip is

$$\begin{pmatrix} Ra_{\text{new}} \\ c_{\text{new}} - a_{\text{new}}^T y \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 1 \\ 110 \end{pmatrix} .$$

Pivoting on the new tableau yields

$$\begin{array}{rrrrrrrrr|r}
x_{\text{new}} & x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & x_7 & x_8 & b \\
1 & 0.5 & 1 & 0 & 0 & .015 & 0 & 0 & -.05 & 25 \\
0 & -5 & 0 & 0 & 0 & -.05 & 1 & 0 & -.5 & 50 \\
-1 & 0 & 0 & 1 & 0 & -.02 & 0 & .1 & 0 & 10 \\
1 & 0.5 & 0 & 0 & 1 & .015 & 0 & -.1 & .05 & 5 \\
\hline
110 & -1500 & 0 & 0 & 0 & -5 & 0 & -100 & -50 & -145,000 \\
\hline
0 & 0 & 1 & 0 & -1 & 0 & 0 & .1 & -.1 & 20 \\
0 & -5 & 0 & 0 & 0 & -.05 & 1 & 0 & -.5 & 50 \\
0 & .5 & 0 & 1 & 1 & -.005 & 0 & 0 & .05 & 15 \\
1 & 0.5 & 0 & 0 & 1 & .015 & 0 & -.1 & .05 & 5 \\
\hline
0 & -1555 & 0 & 0 & -110 & -6.65 & 0 & -88.9 & -55.5 & -145550
\end{array} .$$

The new optimal solution is

$$\begin{pmatrix} x_{\text{new}} \\ x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 5 \\ 0 \\ 20 \\ 15 \\ 0 \end{pmatrix} .$$

## 6.5 Fundamental Theorem on Sensitivity Analysis

The purpose of this section is to state and prove a theorem that captures much of the flavor of the results on sensitivity analysis that we have seen in this section. While there are many possible results one might choose to present, the theorem we give is a stepping stone to the more advanced theory of *Lagrangian duality*. This result focuses on variations in resource availability. We presented this result in Section 1 of these notes on 2-dimensional LPs.

Let $A \in \mathsf{R}^{m\times n}$, $b \in \mathsf{R}^m$, and $c \in \mathsf{R}^n$. This data defines an LP in standard form by

$$\begin{array}{lll} \mathcal{P} & \text{maximize} & c^T x \\ & \text{subject to} & Ax \le b,\ 0 \le x \ . \end{array}$$

We associate $\mathcal{P}$ the *optimal value function* $V : \mathsf{R}^m \mapsto \mathsf{R} \cup \{\pm\infty\}$ defined by

$$\begin{array}{rll} V(u) = & \text{maximize} & c^T x \\ & \text{subject to} & Ax \le b + u,\ 0 \le x \end{array}$$

for all $u \in \mathsf{R}^m$. Let

$$\mathcal{F}(u) = \{x \in \mathsf{R}^n \mid Ax \le b + u,\ 0 \le x\}$$

denote the feasible region for the LP associated with value $V(u)$. If $\mathcal{F}(u) = \emptyset$ for some $u \in \mathsf{R}^m$, we define $V(u) = -\infty$.

**Theorem 6.1 (Fundamental Theorem on Sensitivity Analysis)** *If $\mathcal{P}$ is primal nondegenerate, i.e. the optimal value is finite and no basic variable in any optimal tableau takes the value zero, then the dual solution $y^*$ is unique and there is an $\epsilon > 0$ such that*

$$V(u) = b^T y^* + u^T y^* \quad \text{whenever } |u_i| \le \epsilon,\ i = 1, \dots, m\ .$$

*Thus, in particular, the optimal value function $V$ is differentiable at $u = 0$ with $\nabla V(0) = y^*$.*

Proof: Let

$$\begin{bmatrix} RA & R & Rb \\ (c - A^T y^*)^T & -(y^*)^T & -b^T y^* \end{bmatrix}$$

be any optimal tableau for $\mathcal{P}$. Primal nondegeneracy implies that every component of the vector $Rb$ is strictly positive. If there is another dual optimal solution $\tilde{y}$ associated with another tableau, then we can pivot to it using simplex pivots. All of these simplex pivots must be degenerate since the optimal value cannot change. But degenerate pivots can only be performed if the tableau is degenerate, i.e. there is an index $i$ such that $(Rb)_i = 0$. But then the basic variable associated with $(Rb)_i$ must take the value zero contradicting the hypothesis that $Rb$ is a strictly positive vector. Hence the only possible optimal tableau is the one given. The only other way to have multiple dual solutions is if there is an unbounded ray of optimal solutions emanating from the optimal solution identified by the unique optimal tableau. For this to occur, there must be a row in the optimal tableau such that any positive multiple of that row can be added to the objective row without changing the optimal value. Again, this can only occur if some $(Rb)_i$ is zero leading to the same contradiction. Therefore, primal nondegeneracy implies the uniqueness of the dual solution $y^*$.

Next let $0 < \delta < \min\{(Rb)_i \,|\, i = 1, \dots, m\}$ . Due to the continuity of the mapping $u \to Ru$, there is an $\epsilon > 0$ such that $|(Ru)_i| \le \delta\ i = 1, \dots m$ whenever $|u_j| \le \epsilon\ j = 1, \dots, n$. Hence, if we perturb $b$ by $u$, then

$$R(b + u) = Rb + Ru \ge Rb - \delta \mathsf{e} > 0$$

whenever $|u_j| \le \epsilon\ j = 1, \dots, n$, where $\mathsf{e}$ is the vector of all ones. Therefore, if we perturb $b$ by $u$ in the optimal tableau with $|u_j| \le \epsilon\ j = 1, \dots, n$ , we get the tableau

$$\begin{bmatrix} RA & R & Rb + Ru \\ (c - A^T y^*)^T & -(y^*)^T & -b^T y^* - b^T u \end{bmatrix}$$

which is still both primal and dual feasible, hence optimal with optimal value $V(u) = b^T y^* + b^T u$ proving the theorem. ■