1 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 1.1 *A hyperplane in $\mathbb{R}^n$ is any set of the form*

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

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

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

Fact 1.2 *$H \subset \mathbb{R}^n$ 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 $\mathbb{R}^n$ of dimension $(n-1)$.

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

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

and

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

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 1.3 *Any subset of $\mathbb{R}^n$ 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 1.4 *Given any two points in* $\mathbb{R}^n$, *say* x *and* y, *the line segment connecting them is given by*

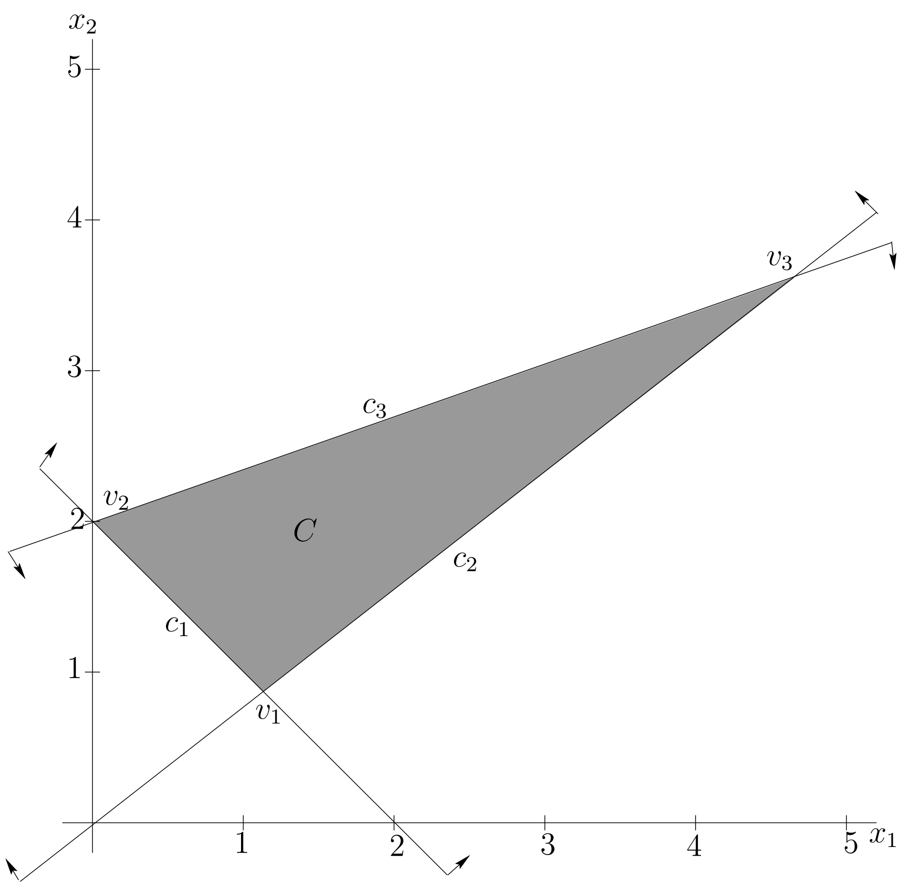
$$[x, y] = \{(1 - \lambda)x + \lambda y : 0 \leq \lambda \leq 1\}.$$

Definition 1.5 *A subset* $C \in \mathbb{R}^n$ *is said to be convex if* $[x, y] \subset C$ *whenever* $x, y \in C$.

Fact 1.6 *A convex polyhedron is a convex set.*

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

$$\begin{aligned} c_1 &: -x_1 - x_2 \leq -2 \\ c_2 &: 3x_1 - 4x_2 \leq 0 \\ c_3 &: -x_1 + 3x_2 \leq 6. \end{aligned} \tag{1.6}$$



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 (1.7). 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 $\mathbb{R}^n$. 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 1.7 *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 1.8 (Fundamental Representation Theorem for Vertices) *A point x in the convex polyhedron $C = \{x \in \mathbb{R}^s \,|\, Tx \le g\}$, where $T = (t_{ij})_{s\times n}$ and $g \in \mathbb{R}^s$, 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*

$$\sum_{j=1}^{n} t_{ij}x_j = g_i \quad i \in \mathcal{I}. \tag{1.7}$$

Moreover, if x is a vertex, then one can take $|\mathcal{I}| = n$ in (1.7), 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 (1.7), 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 (1.7) 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 $\mathbb{A}(x)$ denote the set of *active indices* at x:

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

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

$$\sum_{j=1}^{n} t_{ij}\bar{x}_j = g_i, \; \sum_{j=1}^{n} t_{ij}u_j \le g_i, \; \text{and} \; \sum_{j=1}^{n} t_{ij}v_j \le g_i. \tag{1.8}$$

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, $\mathbb{A}(\bar{x}) \subset \mathbb{A}(u) \cap \mathbb{A}(v)$. Now if $\mathcal{I} \subset \{1, 2, \ldots, s\}$ is such that (1.7) holds at $x = \bar{x}$, then $\mathcal{I} \subset \mathbb{A}(\bar{x})$. But then (1.7) must also hold for $x = u$ and $x = v$ since $\mathbb{A}(\bar{x}) \subset \mathbb{A}(u) \cap \mathbb{A}(v)$. Therefore, $\bar{x}$ is not a unique solution to (1.7) 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 (1.7). 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 (1.7) there exists $w \in \mathbb{R}^n$ with $w \neq \bar{x}$ such that (1.7) holds with $x = w$, then $\bar{x}$ cannot be a vertex of C . To this end take $\mathcal{I} = \mathbb{A}(\bar{x})$ and let $w \in \mathbb{R}^n$ with $w \neq \bar{x}$ be such that (1.7) holds with $x = w$ and $\mathcal{I} = \mathbb{A}(\bar{x})$, and set $u = w - \bar{x}$. Since $\bar{x} \in C$, we know that

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

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

$$\sum_{j=1}^n t_{ij}(\bar{x}_j + tu_j) < g_i \quad \forall\ i \in \{1, 2, \ldots, s\} \setminus \mathbb{A}(\bar{x}) \text{ and } |t| \leq \bar{\tau}. \tag{1.9}$$

Also note that

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

Combining these equivalences with (1.9) 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 (1.7). First note that since the system (1.7) 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 $\mathbb{R}^n$ 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

$$t_{i\cdot}^T x = g_i, \ i \in \mathcal{J} \tag{1.10}$$

is necessarily a solution to the larger system (1.7). But then the smaller system (1.10) must have $\bar{x}$ as its unique solution; otherwise, the system (1.7) has more than one solution. Finally, since the set of solutions to (1.10) 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 1.1 *A point x in the convex polyhedron described by the system of inequalities*

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

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

$$\begin{aligned} \sum_{j=1}^{n} a_{ij} x_j &= b_i \qquad i \in \mathcal{I}, \qquad \text{and} \\ x_j &= 0 \qquad j \in \mathcal{J}. \end{aligned} \tag{1.9}$$

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 (1.9).

Theorem 1.1 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 (1.9) 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 1.1 make intuitive sense?

To answer these questions, let us return to the convex polyhedron C defined by (1.7). 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 2×2 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 1.1 indicates that any subsystem of the form (1.9) for which the associated coefficient matrix is non-singular, has as its solution a vertex of the polyhedron

$$Ax \leq b, \qquad 0 \leq x \tag{1.10}$$

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

The system (1.10) 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

$$x_{n+i} = b_i - \sum_{j=1}^{n} a_{ij}x_j \qquad i = 1, \ldots, m. \tag{1.11}$$

Let $\bar{x} = (\bar{x}_1, \ldots, \bar{x}_{n+m})$ be any solution to the system (1.11) 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

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

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

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

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

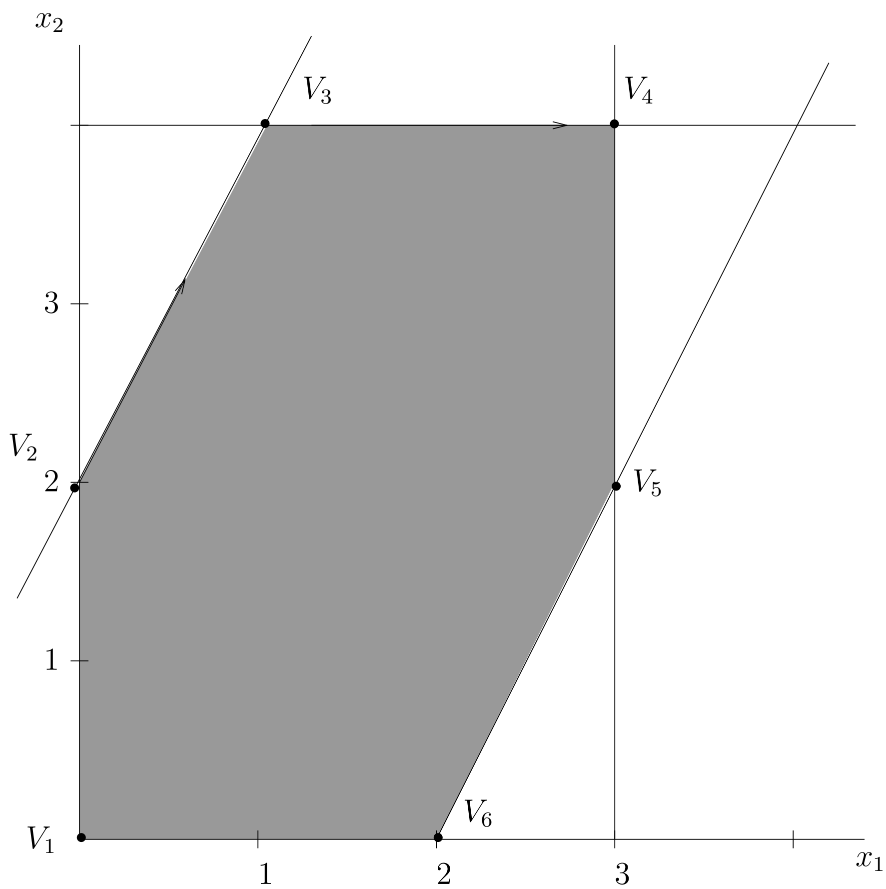
$$(\mathcal{P}) \qquad \begin{array}{l} \max c^T x \\ \text{subject to } Ax \leq b, 0 \leq x. \end{array}$$

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 (1.10) and is the unique solution to a system of the form (1.9). But then, by Theorem 1.1, 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 (1.10). 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

$$\text{(1.12)} \qquad \begin{array}{lll} \text{maximize} & 3x_1 + 4x_2 & \\ \text{subject to} & -2x_1 + x_2 & \leq 2 \\ & 2x_1 - x_2 & \leq 4 \\ & 0 \leq x_1 \leq 3, & 0 \leq x_2 \leq 4. \end{array}$$

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



The simplex algorithm yields the following pivots:

-2	$\boxed{1}$	1	0	0	0	2	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	
3	4	0	0	0	0	0	

-2	1	1	0	0	0	2	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	
11	0	-4	0	0	0	-8	

0	1	0	0	0	1	4	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	
0	0	$\frac{3}{2}$	0	0	$\frac{-11}{2}$	-19	

0	1	0	0	0	1	4	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	
0	0	0	0	-3	-4	-25	

The Geometry of Degeneracy

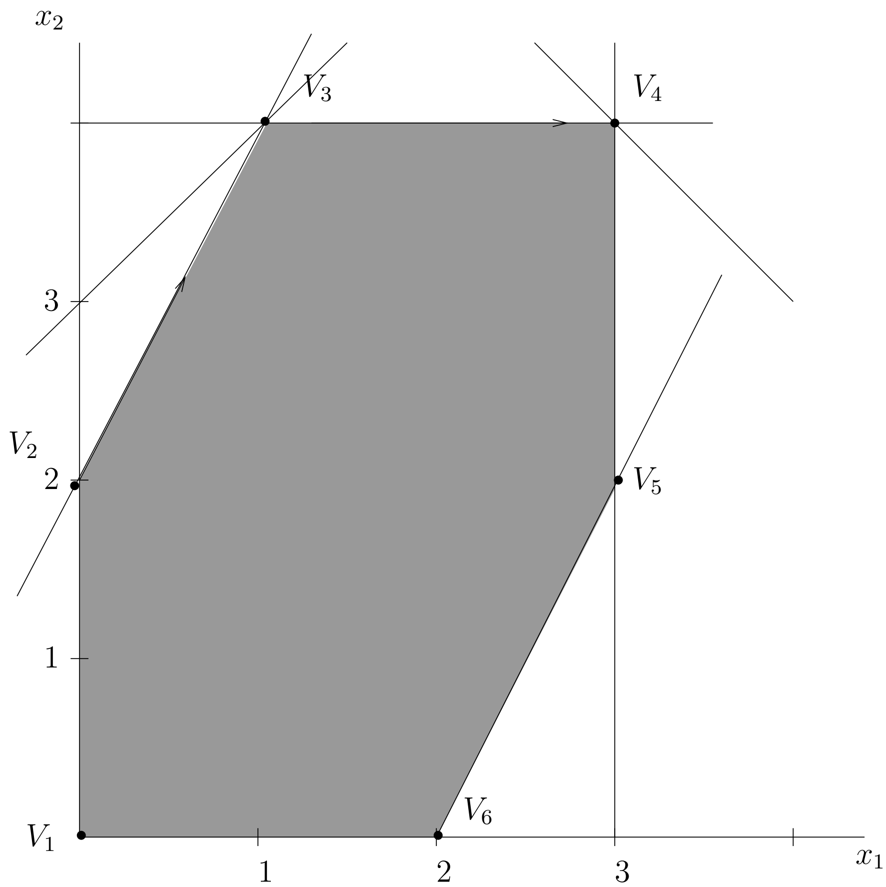
We now give a geometric interpretation of degeneracy in linear programming. Recall that a basic feasible solution, or vertex, is said to be degenerate if one or more of the basic variables is assigned the value zero. In the notation of (1.11) this implies that more than n of the hyperplanes H_k, $k = 1, 2, \ldots, n+m$ 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 (1.12). 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

$$-2x_1 + x_2 = 2 \qquad \text{(since } x_3 = 0\text{)}$$

and

$$-x_1 + x_2 = 3. \qquad \text{(since } x_5 = 0\text{)}$$

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

$$-x_1 + x_2 = 3 \qquad \text{(since } x_5 = 0\text{)}$$

and

$$x_2 = 4 \qquad \text{(since } x_8 = 0\text{).}$$

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.

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
0	0	0	0	0	0	-2	-3	-25

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.

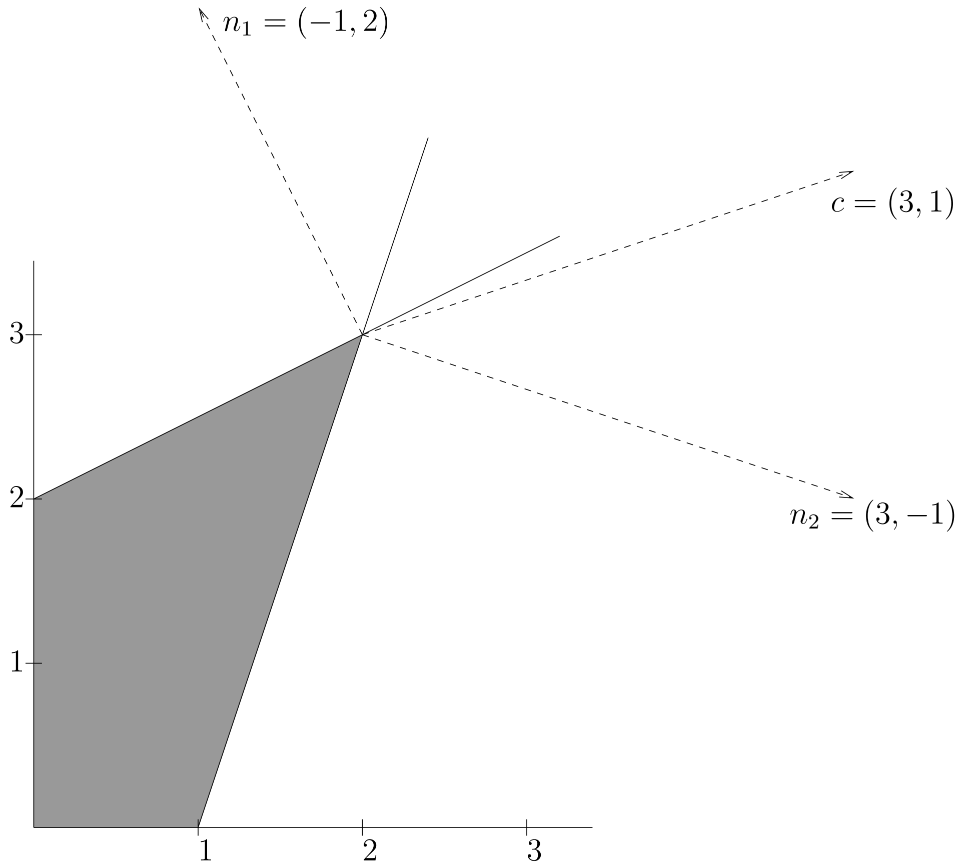
We will arrive at an understanding of why this fact is true after we examine the geometry of duality.

The Geometry of Duality

Consider the linear program

$$(1.13) \qquad \begin{array}{lll} \text{maximize} & 3x_1 + x_2 & \\ \text{subject to} & -x_1 + 2x_2 & \leq 4 \\ & 3x_1 - x_2 & \leq 3 \\ & 0 \leq x_1, & x_2. \end{array}$$

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

-1	3	3
2	-1	1
1	-3	-3
0	5	7
1	-3	-3
0	1	$\frac{7}{5}$
1	0	$\frac{6}{5}$
0	1	$\frac{7}{5}$

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 1.14 (Geometric Duality Theorem) *Consider the LP*

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

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

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

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*

$$c = -\sum_{j \in \mathcal{Z}(\bar{x})} r_j e_j + \sum_{i \in \mathcal{E}(\bar{x})} y_i a_{i\bullet} \tag{1.14}$$

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

$$\bar{y}_i = \begin{cases} y_i & \textit{for } i \in \mathcal{E}(\bar{x}) \\ 0 & \textit{otherwise} \end{cases}, \tag{1.15}$$

solves the dual problem

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

PROOF: Let us first suppose that $\bar{x}$ solves $\mathcal{P}$. Then there is a $\bar{y} \in \mathbb{R}^n$ 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 (1.14) and (1.15) hold. The Complementary Slackness Theorem implies that

$$\bar{y}_i = 0 \text{ for } i \in \{1, 2, \ldots, m\} \setminus \mathcal{E}(\bar{x}) \tag{1.16}$$

and

$$\sum_{i=1}^{m} \bar{y}_i a_{ij} = c_j \text{ for } j \in \{1, \ldots, n\} \setminus \mathcal{Z}(\bar{x}). \tag{1.17}$$

Note that (1.16) implies that $\bar{y}$ satisfies (1.15). 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 (1.17), $r_j = 0$ 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,

$$c_j = -r_j + \sum_{i=1}^{m} \bar{y}_i a_{ij} \text{ for } j \in \mathcal{Z}(\bar{x}). \tag{1.18}$$

Combining (1.18) with (1.17) and (1.16) 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 (1.14) and (1.15) 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 (1.14). We must show that $\bar{x}$ solves $\mathcal{P}$ and $\bar{y}$ solves $\mathcal{D}$. Let $\bar{y} \in \mathbb{R}^m$ be such that its components are given by the $\bar{y}_i$'s for $i \in \mathcal{E}(\bar{x})$ and by (1.15) otherwise. Then the non-negativity of the r_j's in (1.14) 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

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

Does the vector $\bar{x} = (1, 0, 2, 0)^T$ solve this LP? If it does, then according to Theorem 1.14 we must be able to construct the solution to the dual of (1.15) 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

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

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 (1.15);

$$
(1.16)\qquad
\begin{array}{lrcrcrcrcr}
\text{minimize} & -3y_1 & + & y_2 & + & 2y_3 & + & 4y_4 & & \\
\text{subject to} & y_1 & & & & & - & y4 & \ge & 1 \\
 & 3y_1 & + & 4y_2 & - & y_3 & - & y_4 & \ge & 1 \\
 & -2y_1 & - & 2y_2 & + & y_3 & + & 2y_4 & \ge & -1 \\
 & 4y_1 & + & 3y_2 & - & y_3 & - & y_4 & \ge & 2 \\
 & & & & & & & & & \\
\end{array}
$$
$$0 \le y_1, y_2, y_3, y_4.$$

Clearly, $\bar{y}$ is feasible for (1.16). In addition,

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

Therefore, $\bar{y}$ solves (1.16) and $\bar{x}$ solves (1.15) by the Weak Duality Theorem.