Chapter 1

MATHEMATICAL BACKGROUND

This chapter discusses the mathematics that is necessary for the development of the theory of linear programming. We are particularly interested in the solutions of a system of linear equations.

In this notes, matrices are denoted by capital letters $A, B, C, \cdots$. Vectors are denoted by bold face smaller letters $\mathbf{a}, \mathbf{b}, \mathbf{c}, \cdots, \mathbf{x}, \mathbf{y}, \mathbf{z}$. Vectors are column vectors. The transpose of a matrix A is A^T. $\mathbf{0}$ is the zero vector and $\mathbf{1}$ is a column vector of all 1's.

1.1 Linear Equations and Linear Programming

In *linear programming* (LP), we usually want to optimize the *objective variable* z which is a linear function of the *decision variables* x_i $(i = 1, 2, \cdots, n)$ under some linear constraints on x_i. Symbolically, we have

$$\begin{array}{ll} \text{optimize} & z = c_1x_1 + c_2x_2 + \cdots + c_nx_n \\ \text{subject to} & \begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n \leq b_1, \\ \quad\vdots \qquad\qquad \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \leq b_m, \end{cases} \\ \text{where} & x_i \geq 0, \quad i = 1, 2, \cdots, n. \end{array}$$

In matrix form, we have

$$\begin{array}{ll} \text{optimize} & z = \mathbf{c}^T\mathbf{x} \\ \text{subject to} & \begin{cases} A\mathbf{x} \leq \mathbf{b} \\ \mathbf{x} \geq \mathbf{0}\ . \end{cases} \end{array}$$

This is called the *canonical form* of an LP problem. Since inequality constraints are difficult to handle, in actual computation, we usually consider LP problems of the form

$$\begin{array}{ll} \max & z = \mathbf{c}^T\mathbf{x} \\ \text{subject to} & \begin{cases} A\mathbf{x} = \mathbf{b} \\ \mathbf{x} \geq \mathbf{0}\ . \end{cases} \end{array}$$

This is called the *standard form* of the LP problem. Notice that the constraints in the canonical form are transformed to a system of linear equations. The equation $z = \mathbf{c}^T\mathbf{x}$ is called the *objective* *function* and the set of all $\mathbf{x} \in \mathbb{R}^n$ that satisfy this equation is a *hyperplane* in $\mathbb{R}^n$. The rest of this chapter is devoted to the study of the solutions of systems of linear equations and the properties of hyperplanes.

1.2 Systems of Linear Equations and Their Solutions

Consider a system of m simultaneous linear equations in n unknown variables $x_1, \cdots, x_n$:

$$\begin{cases} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1, \\ \vdots & & & & & & & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m. \end{cases}$$

In matrix form, we have

$$A\mathbf{x} = \mathbf{b}. \tag{1.1}$$

To solve this system of equations is to find the values of $x_1, x_2, \cdots, x_n$ that satisfy the equation. The corresponding vector $\mathbf{x} = [x_1, x_2, \cdots, x_n]^T$ will be called a *solution* to (1.1).

To find the solutions of (1.1), we construct the *augmented matrix* A_b of A that is defined by

$$A_b = [A, \mathbf{b}] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{bmatrix} = \left[\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n, \mathbf{b}\right],$$

where $\mathbf{a}_i$ is the i-th column of A. For the solution of (1.1), there are two cases to consider.

(a) $\text{rank}(A) < \text{rank}(A_b)$.

Then $\mathbf{b}$ and the columns of A are linearly independent. Hence there are no x_i such that

$$\sum_{i=1}^{n} x_i \mathbf{a}_i = \mathbf{b}.$$

In particular, the system $A\mathbf{x} = \mathbf{b}$ has no solutions. In that case, we call the system *inconsistent*. Notice that here we have $\text{rank}(\mathbf{b}) = 1$ and $\text{rank}\ (A_b) = \text{rank}\ (A) + 1$.

(b) $\text{rank}(A) = \text{rank}(A_b) = k$.

Then every column of A_b, in particular the vector $\mathbf{b}$, can be expressed as a linear combination of k linearly independent columns of A, i.e. there exist $x_{i_1}, x_{i_2}, \cdots x_{i_k}$ not all zero such that

$$\sum_{j=1}^{k} x_{i_j} \mathbf{a}_{i_j} = \mathbf{b}.$$

Thus at least one solution exists in this case. We remark that if $m = n = \text{rank}(A)$, then the solution is also unique and $\mathbf{x} = A^{-1}\mathbf{b}$. However, in LP, we usually have $\text{rank}(A) = m < n$ and $A\mathbf{x} = \mathbf{b}$ usually has more than one solution.

1.3 Properties of Solutions of Systems of Linear Equations

Let us suppose that $A\mathbf{x} = \mathbf{b}$ has more than one solution, say $\mathbf{x}_1$ and $\mathbf{x}_2$ with $\mathbf{x}_1 \neq \mathbf{x}_2$. Then for any $\lambda \in [0, 1]$,

$$A[\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2] = \lambda A\mathbf{x} + (1-\lambda)A\mathbf{x}_2 = \lambda\mathbf{b} + (1-\lambda)\mathbf{b} = \mathbf{b}.$$

Thus $\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2$ is also a solution. Hence we have proved that if $A\mathbf{x} = \mathbf{b}$ has more than one solution, then it has infinite many solutions.

To characterize these solutions, let us first suppose that A is an m-by-n matrix with $\text{rank}(A) = m < n$. Then $A\mathbf{x} = \mathbf{b}$ can be written as

$$B\mathbf{x}_B + R\mathbf{x}_\beta = \mathbf{b} \ ,$$

where

$$B = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mm} \end{bmatrix} = [\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_m],$$

$$R = \begin{bmatrix} a_{1m+1} & a_{1m+2} & \cdots & a_{1n} \\ a_{2m+1} & a_{2m+2} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{mn+1} & a_{mn+2} & \cdots & a_{mn} \end{bmatrix} = [\mathbf{a}_{m+1}, \mathbf{a}_{m+2}, \cdots, \mathbf{a}_n],$$

$$\mathbf{x}_B = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} \quad \text{and} \quad \mathbf{x}_\beta = \begin{bmatrix} x_{m+1} \\ x_{m+2} \\ \vdots \\ x_n \end{bmatrix} .$$

Since the ordering of the variables x_i are irrelevant, we can assume that the variables have been reordered so that B is nonsingular. Or equivalently, we can consider the nonsingular B as being formed by suitably picking m columns of A. Then we have

$$\mathbf{x}_B = B^{-1}(\mathbf{b} - R\mathbf{x}_\beta) \ . \tag{1.2}$$

Hence given any $\mathbf{x}_\beta$ we can solve uniquely for $\mathbf{x}_B$ in terms of $\mathbf{x}_\beta$. Thus to find the solution to (1.1), we can assign arbitrary values to the $(n-m)$ variables in $\mathbf{x}_\beta$ and determine from (1.2) the values for the remaining m variables in $\mathbf{x}_B$. Then $\mathbf{x} = \begin{bmatrix} \mathbf{x}_B \\ \mathbf{x}_\beta \end{bmatrix}$ is a solution to $A\mathbf{x} = \mathbf{b}$.

1.4 Homogeneous Systems of Linear Equations

A homogeneous system of linear equations is one that is of the form $A\mathbf{x} = \mathbf{0}$. Its solutions form an $(n-m)$ dimensional subspace of $\mathbb{R}^n$. In fact, the solution space is just the kernel of the linear mapping represented by the matrix A. Since A is a mapping from $\mathbb{R}^n$ to $\mathbb{R}^m$, we see that

$$\begin{aligned} \text{dimension of kernel of } A &= \text{dimension of } \mathbb{R}^n - \text{dimension of range of } A \\ &= n - \text{rank}(A) = n - m. \end{aligned}$$

Note that if $\mathbf{x}_1$ is a solution of $A\mathbf{x} = \mathbf{b}$ and $\mathbf{x}_0 \neq \mathbf{0}$ is a solution of $A\mathbf{x} = \mathbf{0}$, then $\mathbf{x}_1 + \mathbf{x}_0$ is a solution of $A\mathbf{x} = \mathbf{b}$. In fact,

$$A(\mathbf{x}_1 + \mathbf{x}_0) = A\mathbf{x}_1 + A\mathbf{x}_0 = \mathbf{b} + \mathbf{0} = \mathbf{b} \ .$$

Clearly $\mathbf{x}_1 \neq \mathbf{x}_1 + \mathbf{x}_0$. Hence we have found two distinct solutions to $A\mathbf{x} = \mathbf{b}$ and by the results in §3, we see that $A\mathbf{x} = \mathbf{b}$ has infinite many solutions. Thus we have proved that if $n > m$, and if $A\mathbf{x} = \mathbf{b}$ is not inconsistent, then $A\mathbf{x} = \mathbf{b}$ has infinite many solutions.

We remark that if $\mathbf{b} \neq \mathbf{0}$, the solutions of $A\mathbf{x} = \mathbf{b}$ do not form a subspace of $\mathbb{R}^n$, but is a space translated away from the origin. Such a space is called an *affine space.*

1.5 Basic Solutions

Definition 1.1. Consider $A\mathbf{x} = \mathbf{b}$ with $\text{rank}(A) = m < n$. Let B be a matrix formed by choosing m columns out of the n columns of A. If B is a non-singular matrix and if all the $(n-m)$ variables not associated with these columns are set equal to zero, then the solution to the resulting system of equations is called a *basic solution*. We call the m variables x_i associated with these m columns the *basic variables*, the other variables are called the *non-basic variables*.

Using (1.2), we see that $\mathbf{x}_B$ are the basic variables and $\mathbf{x}_\beta$ are the non-basic variables. To get a basic solution, we set $\mathbf{x}_\beta = \mathbf{0}$. Then $\mathbf{x}_B = B^{-1}\mathbf{b}$ and $\mathbf{x} = \begin{bmatrix} B^{-1}\mathbf{b} \\ \mathbf{0} \end{bmatrix}$ is a basic solution.

Definition 1.2. A basic solution to $A\mathbf{x} = \mathbf{b}$ is *degenerate* if one or more of the m basic variables vanish.

Example 1.1. Consider the system $A\mathbf{x} = \mathbf{b}$ where

$$A = \begin{bmatrix} 1 & 2 & 1 & 0 \\ 0 & 1 & 2 & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

If we take $B = [\mathbf{a}1 \ \mathbf{a}_2] = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$. In this case $\mathbf{x}_B = [-1, 1]^T$ and $\mathbf{x} = [-1, 1, 0, 0]^T$ is the corresponding basic solution.

We can also take $B = [\mathbf{a}_1 \ \mathbf{a}_4] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. Then $\mathbf{x}_B = [1, 1]^T$ and $\mathbf{x} = [1, 0, 0, 1]^T$ is the corresponding basic solution.

For $B = [\mathbf{a}_2 \ \mathbf{a}_3] = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ we have $\mathbf{x}_B = [1/3, 1/3]^T$. Hence $\mathbf{x} = [0, 1/3, 1/3, 0]^T$ is the corresponding basic solution.

Note that all these basic solutions are non degenerate.

Thus if any one of the elements of $\mathbf{x}_B$ is zero, the basic solution is degenerate.

Although the number of solutions to $A\mathbf{x} = \mathbf{b}$ are in general infinite, we will see later that the optimal solutions in LP problems are basic solutions. Therefore, we will like to know how many basic solutions are there. This is equivalent to asking how many such nonsingular matrices B can possibly be formed from the columns of A. It is obvious that the number of such matrices is bounded by $C_m^n = \dfrac{n!}{m!(n-m)!}$.

Theorem 1.1. *A necessary and sufficient condition for the existence and non-degeneracy of all possible basic solutions of $A\mathbf{x} = \mathbf{b}$ is the linearly independence of every set of m columns of the augmented matrix $A_b = [A, \mathbf{b}]$.*

Proof. ($\Rightarrow$) Suppose that all basic solutions exist and are not degenerate. Then for any set of m columns of A, say

$$\mathbf{a}_{i_1}, \mathbf{a}_{i_2}, \cdots, \mathbf{a}_{i_m}, \qquad 1 \le i_j \le n, 1 \le j \le m,$$

there exists x_{B_j}, $j = 1, 2, \cdots, m$ such that

$$\sum_{j=1}^{m} x_{B_j} \mathbf{a}_{i_j} = \mathbf{b}.$$

Since the solution is non degenerate, all $x_{B_j} \neq 0$. We first claim that $\{\mathbf{a}_{i_j}\}_{j=1}^m$ is linearly independent. For if not, then there exist c_j, $j = 1, \cdots, m$ not all zeros, such that

$$\sum_{j=1}^{m} c_j \mathbf{a}_{i_j} = \mathbf{0}.$$

Let us suppose that $c_r \neq 0$. Then

$$\sum_{j=1}^{m} \tilde{x}_{B_j} \mathbf{a}_{i_j} = \sum_{j=1}^{m} (x_{B_j} - \frac{x_{B_r}}{c_r} c_j) \mathbf{a}_{i_j} = \mathbf{b} \ .$$

Since $\tilde{x}_{B_r} = 0$, we have found a degenerate solution to $A\mathbf{x} = \mathbf{b}$, a contradiction. Thus $\{\mathbf{a}_{i_j}\}_{j=1}^{m}$ is linearly independent. Using the fact that $x_{B_j} \neq 0$ for all j, we can, by replacement method, show that any $m-1$ vectors of $\{\mathbf{a}_{i_j}\}_{j=1}^{m}$ together with the vector $\mathbf{b}$ are also linearly independent. In fact, for all $x_{B_r} \neq 0$, we have

$$-\sum_{\substack{j=1 \\ j \neq r}}^{m} \frac{x_{B_j}}{x_{B_r}} \mathbf{a}_{i_j} - \frac{1}{x_{B_r}} \mathbf{b} = \mathbf{a}_{i_r}.$$

Using similar argument as above, we conclude that the set

$$\{\mathbf{a}_{i_1}, \cdots \mathbf{a}_{i_{r-1}}, \mathbf{a}_{i_{r+1}}, \cdots \mathbf{a}_{i_m}, \mathbf{b}\}$$

is also linearly independent.
($\Leftarrow$) Let $\{\mathbf{a}_{i_1}, \mathbf{a}_{i_2}, \cdots, \mathbf{a}_{i_m}\}$ be an arbitrary set of m columns of A. By assumption, it is linearly independent. Since $\mathbf{a}_{i_j}$ are m-vectors, we see that the set $\{\mathbf{a}_{i_1}, \mathbf{a}_{i_2}, \cdots, \mathbf{a}_{i_m}, \mathbf{b}\}$ is linearly dependent. Hence there exist x_{B_j}, $j = 1, \cdots, m$ such that

$$\mathbf{b} = \sum_{j=1}^{m} x_{B_j} \mathbf{a}_{i_j}.$$

Thus basic solution exists for such choice of m columns of A. Next we claim that $x_{B_j} \neq 0$ for all $j = 1, 2, \cdots, m$. Suppose that $x_{B_r} = 0$. Then

$$\mathbf{b} - \sum_{\substack{j=1 \\ j \neq r}}^{m} x_{B_j} \mathbf{a}_{i_j} = \mathbf{0} \ .$$

That means that the set

$$\{\mathbf{a}_{i_1}, \cdots \mathbf{a}_{i_{r-1}}, \mathbf{a}_{i_{r+1}}, \cdots \mathbf{a}_{i_m}, \mathbf{b}\}$$

is linearly dependent, hence a contradiction to our assumption. Thus $x_{B_j} \neq 0$ for all $j = 1, \cdots, m$. □

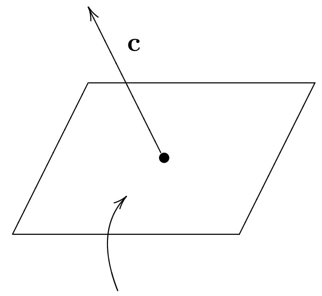
By the same arguments used in the proof of the above theorem, we have the following corollary.

Corollary 1.1. *Given a basic solution to* $A\mathbf{x} = \mathbf{b}$ *with basic variables* $x_{i_1}, \cdots, x_{i_m}$*, a necessary and sufficient condition for the solution to be non-degenerate is the linearly independence of* $\mathbf{b}$ *with every* $m-1$ *columns of* $\{\mathbf{a}_{i_1}, \cdots, \mathbf{a}_{i_m}\}$.

1.6 Hyperplanes

Definition 1.3. A *hyperplane* in $\mathbb{R}^n$ is defined to be the set of all points in $\{\mathbf{x} \mid \mathbf{c}^T\mathbf{x} = z\}$ where $\mathbf{c}$ is a fixed nonzero vector in $\mathbb{R}^n$ and $z \in \mathbb{R}$.

In $\mathbb{R}^2$, a hyperplane is given by the equation $c_1x_1 + c_2x_2 = z$, which is a straight line. In $\mathbb{R}^3$, hyperplanes are just planes.



Given a hyperplane $\mathbf{c}^T\mathbf{x} = z$, it is clear that it passes through the origin if and only if $z = 0$. In that case, $\mathbf{c}$ is orthogonal to every vector $\mathbf{x}$ of the hyperplane, and the hyperplane forms a vector subspace of $\mathbb{R}^n$ with dimension $n-1$.

If $z \neq 0$, then for any two vectors $\mathbf{x}_1 \neq \mathbf{x}_2$ in the hyperplane, we have

$$\mathbf{c}^T(\mathbf{x}_1 - \mathbf{x}_2) = \mathbf{c}^T\mathbf{x}_1 - \mathbf{c}^T\mathbf{x}_2 = z - z = 0.$$

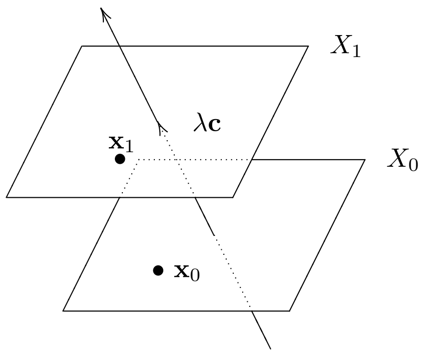
Thus $\mathbf{c}$ is orthogonal to $\mathbf{x}_1 - \mathbf{x}_2$ which is a vector lying on the hyperplane, i.e. $\mathbf{c}$ is also perpendicular to the hyperplane. We note that in this case, the hyperplane is an affine space.

Definition 1.4. Given the hyperplane $\mathbf{c}^T\mathbf{x} = z$, the vector $\mathbf{c}$ is called the *normal* of the hyperplane. Two hyperplanes are said to be *parallel* if their normals are parallel vectors.

Example 1.1. Let $\mathbf{x}_0$ be arbitrarily chosen from a hyperplane $X_0 = \{\mathbf{x} \mid \mathbf{c}^T\mathbf{x} = z_0\}$. Let $\lambda > 0$ be fixed. Then the point $\mathbf{x}_1 = \mathbf{x}_0 + \lambda\mathbf{c}$ satisfies

$$\mathbf{c}^T\mathbf{x}_1 = \mathbf{c}^T\mathbf{x}_0 + \lambda|\mathbf{c}|^2 = z_0 + \lambda|\mathbf{c}|^2 > z_0.$$

Here $|\cdot|$ denotes the Euclidean norm in $\mathbb{R}^n$. Let $z_1 \equiv z_0 + \lambda|\mathbf{c}|^2$ and define the hyperplane $X_1 = \{\mathbf{x} \mid \mathbf{c}^T\mathbf{x} = z_1\}$. We see that the hyperplanes X_0 and X_1 are parallel and X_1 is lying in the direction of $\mathbf{c}$ from X_0. The distance between the hyperplanes is λ.



1.7 Convex Sets

Definition 1.5. A set C is said to be *convex* if for all $\mathbf{x}_1$ and $\mathbf{x}_2$ in C and $\lambda \in [0,1]$, we have $\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2 \in C$.

Geometrically, that means that given any two points in C, then all points on the line segment joining the given two points should also be in C.

Definition 1.6. A point $\mathbf{x}$ is an *extreme point* of a convex set C if there exist no two distinct points $\mathbf{x}_1$ and $\mathbf{x}_2 \in C$ such that $\mathbf{x} = \lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2$ for some $\lambda \in (0,1)$.

Geometrically, extreme points are just the corner points of C.

Example 1.2. $\mathbb{R}^n$ is convex. Let W be a subspace of $\mathbb{R}^n$ and $\mathbf{x}_1, \mathbf{x}_2 \in W$. Thus any linear combination of $\mathbf{x}_1$ and $\mathbf{x}_2$ is also in W, in particular the linear combination $\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2 \in W$ for $\lambda \in [0,1]$. This shows that W is convex.

Example 1.3. The n dimensional open ball centered at $\mathbf{x}_0$ with radius r is defined as

$$B_r(\mathbf{x}_0) = \{\mathbf{x} \mid |\mathbf{x} - \mathbf{x}_0| < r\}.$$

The n dimensional closed ball centered at $\mathbf{x}_0$ with radius r is defined as

$$\overline{B_r(\mathbf{x}_0)} = \{\mathbf{x} \mid |\mathbf{x} - \mathbf{x}_0| < r\}.$$

Both the open ball and the closed ball are convex. We prove it for the open ball. Let $\mathbf{x}_1, \mathbf{x}_2 \in B(\mathbf{x}_0)$ and $\lambda \in [0,1]$. Then

$$\begin{aligned} |(\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2) - \mathbf{x}_0| &= |\lambda(\mathbf{x}_1 - \mathbf{x}_0) + (1-\lambda)(\mathbf{x}_2 - \mathbf{x}_0)| \\ &\leq \lambda|\mathbf{x}_1 - \mathbf{x}_0| + (1-\lambda)|\mathbf{x}_2 - \mathbf{x}_0| \\ &\leq \lambda r + (1-\lambda)r = r. \end{aligned}$$

Let S be a subset of $\mathbb{R}^n$. A point $\mathbf{x}$ is a *boundary point* of S if every open ball centered at $\mathbf{x}$ contains both a point in S and a point in $\mathbb{R}^n - S$. Note that a boundary point can either be in S or not in S. The set of all boundary points of S, denoted by ∂S, is the *boundary* of S. A set S is *closed* if $\partial S \subset S$. A set S is open if its complement $\mathbb{R}^n - S$ is closed. Note that a set that is not closed is *not* necessarily open; and a set that is not open is *not necessarily closed*. There are sets that are neither open nor closed. The *closure* of a set S is the set $\overline{S} = S \cup \partial S$. The *interior* of a set S is the set $S^\circ = S - \partial S$. A set S is closed if and only if $S = \overline{S}$. A set S is open if and only if $S = S^\circ$.

Example 1.4. $\mathbb{R}^n$ is both open and closed. The empty set $\emptyset$ is both open and closed.

Example 1.5. The hyperplane $P = \{\mathbf{x} \mid \mathbf{c}^T\mathbf{x} = z\}$ is closed in $\mathbb{R}^n$. In fact we will show that $P \subseteq \partial P$. Without loss of generality we may assume $|\mathbf{c}| = 1$. Let $\mathbf{x} \in P$ and $B_r(\mathbf{x})$ is an open ball centered at $\mathbf{x}$ with radius r. Since $\mathbf{x} \in B_r(\mathbf{x})$ it remains to show that $B_r(\mathbf{x})$ contains a point not in P. Let

$$\mathbf{y} = \mathbf{x} + \frac{r}{2}\mathbf{c}$$

then

$$\mathbf{c}^T\mathbf{y} = \mathbf{c}^T\mathbf{x} + \frac{r}{2}\mathbf{c}^t\mathbf{c} = z + \frac{r}{2} > z.$$

Hence $\mathbf{y} \notin P$. But $|\mathbf{y} - \mathbf{x}| = \frac{r}{2}$ therefore $\mathbf{y} \in B_r(\mathbf{x}$.

Example 1.6. The half spaces

$$X_1 = \{\mathbf{x} \mid \mathbf{c}^T\mathbf{x} \leq z\} \quad \text{and} \quad X_2 = \{\mathbf{x} \mid \mathbf{c}^T\mathbf{x} \geq z\}$$

are closed in $\mathbb{R}^n$. In fact we have $\partial X_1 = \partial X_2 =$ the hyperplane $P = \{\mathbf{x} \mid \mathbf{c}^T\mathbf{x} = z\}$. We will show that $\partial X_1 = P$, the proof for $\partial X_2 = P$ is similar.

($P \subset \partial X_1$) Let $\mathbf{x} \in P$. For any $r > 0$, let

$$\mathbf{y}_1 = \mathbf{x} + \frac{r}{2|\mathbf{c}|}\mathbf{c}, \quad \mathbf{y}_2 = \mathbf{x} - \frac{r}{2|\mathbf{c}|}\mathbf{c}.$$

We see that $|\mathbf{x} - \mathbf{y}_1| = \frac{r}{2} = |\mathbf{x} - \mathbf{y}_2|$ so both $\mathbf{y}_1, \mathbf{y}_2 \in B_r(\mathbf{x})$. Moreover

$$\mathbf{c}^T\mathbf{y}_1 = \mathbf{c}^T\mathbf{x} + \frac{r}{2|\mathbf{c}|}\mathbf{c}^T\mathbf{c} = r + \frac{r}{2} > r$$

and therefore $\mathbf{y}_1 \notin X_1$. On the other hand

$$\mathbf{c}^T\mathbf{y}_2 = \mathbf{c}^T\mathbf{x} - \frac{r}{2|\mathbf{c}|}\mathbf{c}^T\mathbf{c} = r - \frac{r}{2} < r$$

so $\mathbf{y}_2 \in X_1$. This shows $\mathbf{x} \in \partial X_1$.

($\partial X_1 \subset P$) Suppose $\mathbf{x} \notin P$. If $\mathbf{c}^T\mathbf{x} = z_1 < z$ then let $r = \frac{z-z_1}{2} > 0$. The open ball $B_r(\mathbf{x})$ lies entirely in X_1. So $B_r(\mathbf{x})$ contains no point outside of X_1, hence $\mathbf{x} \notin \partial X_1$. If $\mathbf{c}^T\mathbf{x} = z_1 > z$ then let $r = \frac{z_1-z}{2} > 0$. The open ball $B_r(\mathbf{x})$ lies entirely outside of X_1. So $B_r(\mathbf{x})$ contains no point of X_1, hence $\mathbf{x} \notin \partial X_1$. In either case $\mathbf{c} \notin \partial X_1$.

Lemma 1.1. *(a) All hyperplanes are convex.*

(b) The closed half-space $\{\mathbf{x} \mid \mathbf{c}^T\mathbf{x} \le z\}$ and $\{\mathbf{x} \mid \mathbf{c}^T\mathbf{x} \ge z\}$ are convex.

(c) The open half-space $\{\mathbf{x} \mid \mathbf{c}^T\mathbf{x} < z\}$ and $\{\mathbf{x} \mid \mathbf{c}^T\mathbf{x} > z\}$ are convex.

(d) Any intersection of convex sets is still convex.

(e) The set of all feasible solutions *to a linear programming problem is a convex set.*

Proof. (a) Let $X = \{\mathbf{x} \mid \mathbf{c}^T\mathbf{x} = z\}$ be our hyperplane. For all $\mathbf{x}_1, \mathbf{x}_2 \in X$ and $\lambda \in [0,1]$, we have

$$\mathbf{c}^T[\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2] = \lambda\mathbf{c}^T\mathbf{x}_1 + (1-\lambda)\mathbf{c}^T\mathbf{x}_2 = \lambda z + (1-\lambda)z = z.$$

Thus $\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2 \in X$. Hence X is convex.

(b) and (c) can be proved similarly by replacing the equality signs in (a) by the corresponding inequality signs.

(d) Let $C = \bigcap_{\alpha\in I} C_\alpha$, where C_α are convex for all α in the index set I. Then for all $\mathbf{x}_1, \mathbf{x}_2 \in C$, we have $\mathbf{x}_1, \mathbf{x}_2 \in C_\alpha$ for all $\alpha \in I$. Hence for all $\lambda \in [0,1]$,

$$\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2 \in C_\alpha$$

for all $\alpha \in I$. Thus $\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2 \in C$, and C is convex.

(e) For any LP problem, the constraints can be written as $\mathbf{a}_i\mathbf{x} \le b_i$ or $\mathbf{a}_i\mathbf{x} = b_i$ etc. The set of points that satisfy any one of these constraints is thus a half space or a hyperplane. By (a), (b) and (c), they are convex. By (d), the intersection of all these sets, which is defined to be the set of feasible solutions, is a convex set. □

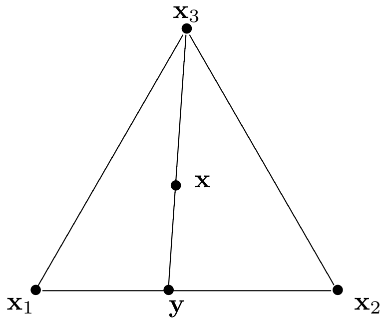
Definition 1.7. Let $\{x_1, \cdots x_k\}$ be a set of given points. Let

$$\mathbf{x} = \sum_{i=1}^{k} M_i\mathbf{x}_i$$

where $M_i \ge 0$ for all i and $\sum_{i=1}^{k} M_i = 1$. Then $\mathbf{x}$ is called a *convex combination* of the points $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_k$.

Example 1.7. Consider the triangle on the plane with vertices $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$. Then any point $\mathbf{x}$ in the triangle is a convex combinations of $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$. In fact, let $\mathbf{y}$ be the extension of the line segment

$\mathbf{x}\mathbf{x}_3$ and such that $\mathbf{y}$ lies on the line segment $\mathbf{x}_1\mathbf{x}_2$.



Then

$$\mathbf{y} = p\mathbf{x}_1 + q\mathbf{x}_2,$$

where

$$p = \frac{|\mathbf{x}_2\mathbf{y}|}{|\mathbf{x}_2\mathbf{x}_1|} \quad \text{and} \quad q = \frac{|\mathbf{x}_1\mathbf{y}|}{|\mathbf{x}_2\mathbf{x}_1|}.$$

Since

$$\mathbf{x} = \ell\mathbf{x}_3 + k\mathbf{y},$$

where

$$\ell = \frac{|\mathbf{x}\mathbf{y}|}{|\mathbf{y}\mathbf{x}_3|} \quad \text{and} \quad k = \frac{|\mathbf{x}\mathbf{x}_3|}{|\mathbf{y}\mathbf{x}_3|},$$

we have

$$\mathbf{x} = kp\mathbf{x}_1 + kq\mathbf{x}_2 + \ell\mathbf{x}_3.$$

Clearly $kp + kq + \ell = 1$.

Definition 1.8. The *convex hull* $\langle A\rangle$ of a set A is defined to be the intersection of all convex sets that contain A.

Lemma 1.2. *(a) If $A \neq \phi$, then $\langle A\rangle \neq \phi$.*

(b) If $A \subseteq B$, then $\langle A\rangle \subseteq \langle B\rangle$.

(c) $\langle A\rangle$ is the smallest convex set that contains A.

(d) If A is convex, then $\langle A\rangle = A$.

Proof. Part (a) follows from the facts that $A \subseteq \mathbb{R}^n$ and $\mathbb{R}^n$ is convex. For (b), we note that any convex set that contains B should also contain A. Finally, (c) and (d) are trivial. □

Theorem 1.2. *The convex hull of a finite number of points $\mathbf{x}_1$, $\mathbf{x}_2, \cdots, \mathbf{x}_k$ is the set of all convex combinations of $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_k$.*

Proof. Let

$$A = \left\{ \mathbf{x} \mid \mathbf{x} = \sum_{i=1}^{k} M_i\mathbf{x}_i, M_i \geq 0, \sum_{i=1}^{k} M_i = 1 \right\} ,$$

be the set of all convex combinations of $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_k$. Let $B = \langle\{\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_k\}\rangle$ be the convex hull of $\{\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_k\}$. We claim that $A = B$.

We first prove that A is convex. For all $\mathbf{y}_1, \mathbf{y}_2 \in A$, we have

$$\mathbf{y}_1 = \sum_{i=1}^{k} M_i\mathbf{x}_i \quad \text{and} \quad \mathbf{y}_2 = \sum_{i=1}^{k} N_i\mathbf{x}_i,$$

where

$$\sum_{i=1}^{k} M_i = 1, \quad \sum_{i=1}^{k} N_i = 1$$

and $M_i, N_i \geq 0$ for all i. Hence for all $\lambda \in [0,1]$,

$$\begin{aligned} \mathbf{y} &= \lambda \mathbf{y}_1 + (1-\lambda)\mathbf{y}_2 \\ &= \lambda \sum_{i=1}^{k} M_i \mathbf{x}_i + (1-\lambda) \sum_{i=1}^{k} N_i \mathbf{x}_i \\ &= \sum_{i=1}^{k} \left(\lambda M_i + (1-\lambda) N_i\right) \mathbf{x}_i = \sum_{i=1}^{k} P_i \mathbf{x}_i \end{aligned}$$

where $P_i = \lambda M_i + (1-\lambda)N_i \geq 0$ and

$$\begin{aligned} \sum_{i=1}^{k} P_i &= \sum_{i=1}^{k} (\lambda M_i + (1-\lambda)N_i) \\ &= \lambda \sum_{i=1}^{k} M_i + (1-\lambda) \sum_{i=1}^{k} N_i \\ &= \lambda \cdot 1 + (1-\lambda) \cdot 1 = 1. \end{aligned}$$

Thus $\mathbf{y}$ is a convex combination of $\mathbf{x}_1, \cdots, \mathbf{x}_k$, hence is in A. Therefore A is convex.

Next we claim that $B \subseteq A$. For all $i = 1, \cdots k$, since

$$\mathbf{x}_i = \sum_{j=1}^{k} \delta_{ij} \mathbf{x}_j,$$

where δ_{ij} is the Kronecker delta[1], we see that $\mathbf{x}_i \in A$. Thus A is a convex set containing $\{\mathbf{x}_1, \cdots, \mathbf{x}_k\}$. By Lemma 1.2 (b) and (d), $B \subseteq A$.

Finally, we claim that $A \subseteq B$. We prove this by mathematical induction on k. If $k = 1$, then $B = B_1 = \langle\{\mathbf{x}_1\}\rangle$. For all $\mathbf{x} \in A_1$, $\mathbf{x} = 1 \cdot \mathbf{x}_1 = \mathbf{x}_1$. Thus $A_1 = \{\mathbf{x}_1\} \subseteq B_1$. Hence by Lemma 1.2 (b) and (d), $A_1 = \langle A \rangle \subseteq B_1$. Now assume that the claim holds for $k-1$, i.e. $A_{k-1} \subseteq B_{k-1}$, where

$$B_{k-1} = \langle\{\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_{k-1}\}\rangle,$$

and

$$A_{k-1} = \left\{ \mathbf{x} \mid \mathbf{x} = \sum_{i=1}^{k-1} M_i \mathbf{x}_i, M_i \geq 0, i = 1, 2, \cdots, k-1\ , \sum_{i=1}^{k-1} M_i = 1 \right\}.$$

We claim that $A_k \subseteq B_k = \langle\{\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_k\}\rangle$. Let $\mathbf{x} \in A_k$, then

$$\mathbf{x} = \sum_{i=1}^{k} M_i \mathbf{x}_i$$

where $M_i \geq 0$ for all i and $\sum_{i=1}^{k} M_i = 1$. If $\mathbf{x} = \mathbf{x}_k$, then $\mathbf{x} \in B_k$. Hence we may assume that $\mathbf{x} \neq \mathbf{x}_k$, or equivalently, $M_k \neq 1$. Since $\sum_{i=1}^{k} M_i = 1$, we have

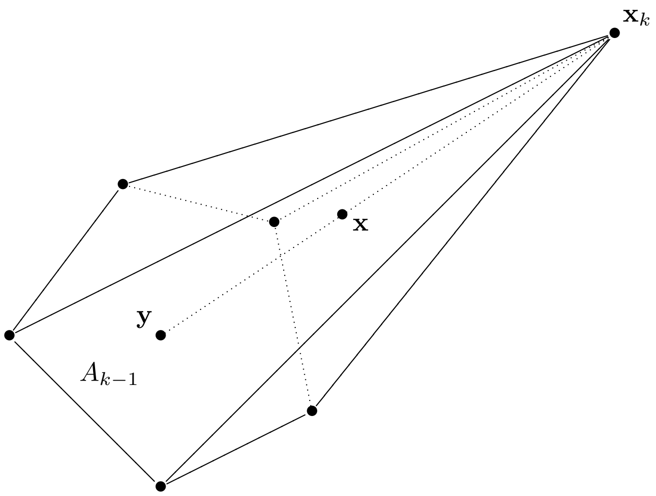
$$\sum_{i=1}^{k-1} \frac{M_i}{1 - M_k} = \frac{1}{1 - M_k} \sum_{i=1}^{k-1} M_i = 1.$$

[1]The Kronecker delta is defined by $\delta_{ii} = 1$ and $\delta_{ij} = 0$ for $i \neq j$.

Consider the vector $\mathbf{y} \in A_{k-1}$ given by

$$\mathbf{y} = \sum_{i=1}^{k-1} \frac{M_i}{1 - M_k} \mathbf{x}_i,$$

we have $\dfrac{M_i}{1 - M_k} \geq 0$ for all $i = 1, 2, \cdots, k-1$ and $\sum\limits_{i=1}^{k-1} \dfrac{M_i}{1 - M_k} = 1$. You can think of $\mathbf{y}$ as the projection of $\mathbf{x}$ onto the convex set A_{k-1}:



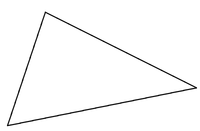
Thus by the induction hypothesis $\mathbf{y} \in B_{k-1}$. Since $\{\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_{k-1}\} \subseteq \{\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_k\}$, $B_{k-1} \subseteq B_k$. Therefore, $\mathbf{y} \in B_k$. Since B_k is convex by definition and $\mathbf{y}$, $\mathbf{x}_k \in B_k$, we see that

$$\mathbf{x} = \sum_{i=1}^{k} M_i \mathbf{x}_i = (1 - M_k)\mathbf{y} + M_k \mathbf{x}_k \in B_k.$$

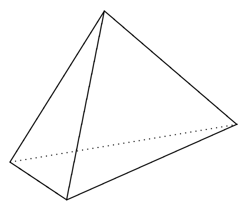
Thus $A_k \subseteq B_k$ for all k. □

Definition 1.9. Let $A = \{\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n\}$ be a set of points. The convex hull of A is called the *convex polyhedron* spanned by A. The convex hull of any set of $n+1$ points in $\mathbb{R}^n$ which do not lie on a hyperplane is called a *simplex.*

Thus in $\mathbb{R}^2$, the triangle formed by any three non-collinear points is a simplex while in $\mathbb{R}^3$, the tetrahedron formed by any four points that are not on a plane is a simplex.



A simplex in $\mathbb{R}^2$



A simplex in $\mathbb{R}^3$

1.8 Supporting Hyperplanes

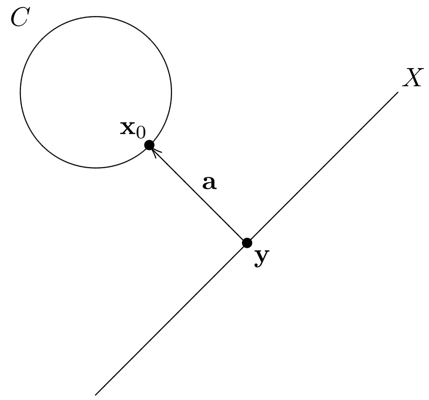
Theorem 1.3. *Let C be a closed convex set and $\mathbf{y}$ be a point not in C. Then there is a hyperplane $X = \{\mathbf{x} \mid \mathbf{a}^T\mathbf{x} = z\}$ that contains $\mathbf{y}$ and such that $C \subseteq \{\mathbf{x} \mid \mathbf{a}^T\mathbf{x} < z\}$ or $C \subseteq \{\mathbf{x} \mid \mathbf{a}^T\mathbf{x} > z\}$.*

Proof. Let

$$\delta = \inf_{\mathbf{x}\in C} |\mathbf{x} - \mathbf{y}| > 0\ .$$

Then there is an $\mathbf{x}_0$ on the boundary of C such that $|\mathbf{x}_0 - \mathbf{y}| = \delta$. This follows because the continuous function $f(\mathbf{x}) = |\mathbf{x} - \mathbf{y}|$ achieves its minimum over any closed and bounded set and it is clear that it suffices to consider $\mathbf{x}$ in the intersection of the closure of C and the sphere of radius 2δ centered at $\mathbf{y}$.

Define $\mathbf{a} = \mathbf{x}_0 - \mathbf{y}$ and $z = \mathbf{a}^T\mathbf{y}$. Consider the hyperplane $X = \{\mathbf{x} \mid \mathbf{a}^T\mathbf{x} = z\}$ which is illustrated by the following diagram



Clearly $\mathbf{y} \in X$. We claim that $C \subseteq \{\mathbf{x} \mid \mathbf{a}^T\mathbf{x} > z\}$. Let $\mathbf{x} \in C$. For any $\lambda \in (0, 1)$, since $\mathbf{x}_0 \in C$,

$$\mathbf{x}_0 + \lambda(\mathbf{x} - \mathbf{x}_0) = \lambda\mathbf{x} + (1 - \lambda)\mathbf{x}_0 \in C.$$

Thus by definition of $\mathbf{a}$,

$$|\mathbf{x}_0 + \lambda(\mathbf{x} - \mathbf{x}_0) - \mathbf{y}|^2 \geq \delta \geq |\mathbf{x}_0 - \mathbf{y}|^2\ ,$$

i.e.

$$|\lambda(\mathbf{x} - \mathbf{x}_0) + \mathbf{a}|^2 \geq |\mathbf{a}|^2.$$

Expanding it, we have

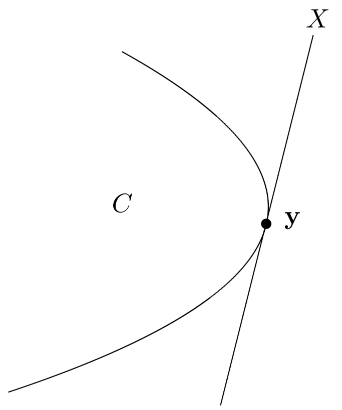
$$2\lambda\mathbf{a}^T(\mathbf{x} - \mathbf{x}_0) + \lambda^2|\mathbf{x} - \mathbf{x}_0|^2 \geq 0\ .$$

Letting $\lambda \to 0^+$, we obtain $\mathbf{a}^T(\mathbf{x} - \mathbf{x}_0) \geq 0$. Hence

$$\mathbf{a}^T\mathbf{x} \geq \mathbf{a}^T\mathbf{x}_0 = \mathbf{a}^T(\mathbf{y} + \mathbf{a}) = \mathbf{a}^T\mathbf{y} + |\mathbf{a}|^2 = \mathbf{a}^T\mathbf{y} + \delta^2 = z + \delta^2 > z\ .$$

Thus $C \subseteq \{\mathbf{x} \mid \mathbf{a}^T\mathbf{x} > z = \mathbf{a}^T\mathbf{y}\}$. □

Definition 1.10. Let $\mathbf{y}$ be a boundary point of a convex set C. Then $X = \{\mathbf{x}|\mathbf{a}^T\mathbf{x} = z\}$ is called a *supporting hyperplane* of C at $\mathbf{y}$ if $\mathbf{y} \in X$ and either $C \subseteq \{\mathbf{x}|\mathbf{a}^T\mathbf{x} \geq z\}$ or $C \subseteq \{\mathbf{x}|\mathbf{a}^T\mathbf{x} \leq z\}$.



Theorem 1.4. *If* $\mathbf{y}$ *is a boundary point of a closed convex set* C*, then there is at least one supporting hyperplane at* $\mathbf{y}$.

Proof. Since $\mathbf{y}$ is a boundary point of C, for any positive integer k, let $\mathbf{y}_k \in B_{\frac{1}{k}}(\mathbf{y})$ and $\mathbf{y}_k \notin C$. Then $\{\mathbf{y}_k\}$ be a sequence of vectors not in C but converges to $\mathbf{y}$. Let $\{\mathbf{a}_k\}$ be the sequence of corresponding normal vectors constructed according to Theorem 1.3, normalized so that $|\mathbf{a}_k| = 1$ and such that $C \subseteq \{\mathbf{x} \mid \mathbf{a}_k^T\mathbf{x} > \mathbf{a}_k^T\mathbf{y}_k\}$. Since $\{\mathbf{a}_k\}$ is a bounded sequence, it has a convergent subsequence $\{\mathbf{a}_{k_j}\}$ with limit $\mathbf{a}$. Let $z = \mathbf{a}^T\mathbf{y}$. We claim that $X = \{\mathbf{x} \mid \mathbf{a}^T\mathbf{x} = z\}$ is a supporting hyperplane at $\mathbf{y}$. Clearly $\mathbf{y} \in X$. For any $\mathbf{x} \in C$, we have

$$\mathbf{a}^T\mathbf{x} = \lim_{j\to\infty} \mathbf{a}_{k_j}^T\mathbf{x} \geq \lim_{j\to\infty} \mathbf{a}_{k_j}^T\mathbf{y}_{k_j} = \mathbf{a}^T\mathbf{y} = z.$$

Thus $C \subseteq \{\mathbf{x} \mid \mathbf{a}^T\mathbf{x} \geq z\}$. □

Note that in general a supporting hyperplane is not unique.

Definition 1.11. A set C is said to be *bounded from below* if

$$\inf\{x_j \mid x_j \text{ is the } j\text{-th component of } \mathbf{x} \text{ in } C\} > -\infty \quad \text{for all } j.$$

Thus

$$\mathbb{R}_+^n \equiv \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x}^T = [x_1, \cdots, x_n], x_j \geq 0, j = 1, \cdots, n\} \tag{1.3}$$

is a set bounded from below. However, the upper half plane $\{(x_1, x_2) \mid x_2 \geq 0\}$ is not bounded from below. Clearly, any subset of $\mathbb{R}^{n+}$ is also bounded from below.

Theorem 1.5. *Let* C *be a closed convex set which is bounded from below. Then every supporting hyperplane of* C *contains an extreme point of* C.

Proof. Let $\mathbf{x}_0$ be a boundary point of C and $X = \{\mathbf{x} \mid \mathbf{c}^T\mathbf{x} = z\}$ be a supporting hyperplane at $\mathbf{x}_0$. Define $T = X \cap C$. We will show that T contains an extreme point of C. Since $\mathbf{x}_0 \in X$ and $\mathbf{x}_0 \in C$, $\mathbf{x}_0 \in T$ and hence T is nonempty.

We first claim that any extreme point of T is also an extreme point of C, or equivalently, if $\mathbf{t}$ is not an extreme point of C, then $\mathbf{t}$ cannot be an extreme point of T. If $\mathbf{t} \notin T$, then we are done. Let us assume that $\mathbf{t} \in T$, then $\mathbf{t} \in C$. By Theorem 1.4, we may assume that $C \subseteq \{\mathbf{x} \mid \mathbf{c}^T\mathbf{x} \geq z\}$. Since $\mathbf{t}$ is not an extreme point in C, there exist two distinct points $\mathbf{x}_1$ and $\mathbf{x}_2 \in C$ such that $\mathbf{t} = \lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2$ for some $\lambda \in (0, 1)$. It then suffices to show that $\mathbf{x}_1, \mathbf{x}_2 \in T$. Since

$$\mathbf{c}^T\mathbf{t} = z = \lambda\mathbf{c}^T\mathbf{x}_1 + (1-\lambda)\mathbf{c}^T\mathbf{x}_2$$

and $C \subseteq \{\mathbf{x} \mid \mathbf{c}^T\mathbf{x} \geq z\}$, we see that $\mathbf{c}^T\mathbf{x}_1 = z = \mathbf{c}^T\mathbf{x}_2$. Hence $\mathbf{x}_1, \mathbf{x}_2 \in T$. Thus $\mathbf{t}$ is not an extreme point in T.

We next claim that there is an extreme point in T. We first find a $\mathbf{t}^1 = [t_1^1, t_2^1, \cdots, t_n^1]^T$ in T such that

$$t_1^1 = \inf\{t_1 \mid \mathbf{t} = [t_1, \cdots, t_n]^T \in T\} \ .$$

Since $T \subseteq C$ and C is bounded from below, the infimum t_1^1 is well-defined. If $\mathbf{t}^1$ is unique, then we prove below that it is our extreme point. If $\mathbf{t}^1$ is not unique, we continue to find a $\mathbf{t}^2 = (t_1^2, t_2^2, \cdots, t_n^2)^T \in T$ such that

$$\begin{aligned} t_1^2 &= \inf\{t_1 | \mathbf{t} = [t_1, t_2, \cdots, t_n]^T \in T\} \\ t_2^2 &= \inf\{t_2 | \mathbf{t} = [t_1, t_2, \cdots, t_n]^T \in T \text{ with } t_1 = t_1^2\} \ . \end{aligned}$$

Since $C \subseteq \mathbb{R}^n$, this process must stop at most after n steps.

We claim that when the process stops, the point found is an extreme point. Let $\mathbf{t}^j$ be that point. Then for all $k = 1, \cdots, j$,

$$t_k^j = \inf\{t_k \mid \mathbf{t} = [t_1, t_2, \cdots, t_n]^T \in T \text{ with } t_i = t_i^j \text{ for } i = 1, \cdots k-1\} \ .$$

Suppose that $\mathbf{t}^j$ is not an extreme point of T, then there exist two distinct points $\mathbf{y}^1, \mathbf{y}^2 \in T$ such that

$$\mathbf{t}^j = \lambda \mathbf{y}^1 + (1-\lambda)\mathbf{y}^2$$

for some $\lambda \in (0, 1)$. Thus

$$t_i^j = \lambda y_i^1 + (1-\lambda) y_i^2, \quad \text{for all} i = 1, 2, \cdots, n \ .$$

For all $i = 1, \cdots, j$, since t_i^j is the infimum, $t_i^j \leq y_i^1$ and $t_i^j \leq y_i^2$. It follows that $t_i^j = y_i^1 = y_i^2$, $i = 1, \cdots, j$, a contradiction to the uniqueness of $\mathbf{t}^j$. □