EXTREME POINTS AND BASIC SOLUTIONS:

In Linear Programming, the feasible region in $\mathbb{R}^n$ is defined by $P := \{x \in \mathbb{R}^n | Ax = b, x \geq 0\}$. The set $P$, as we have seen, is a convex subset of $\mathbb{R}^n$. It is called a **convex polytope**. The term **convex polyhedron** refers to convex polytope which is **bounded**. Polytopes in two dimensions are often called **polygons**. Recall that the vertices of a convex polytope are what we called extreme points of that set.

Recall that extreme points of a convex set are those which **cannot** be represented as a proper convex combination of two other (distinct) points of the convex set. It may, or may not be the case that a convex set **has** any extreme points as shown by the example in $\mathbb{R}^2$ of the strip $S := \{(x, y) \in \mathbb{R}^2 | 0 \leq x \leq 1, y \in \mathbb{R}\}$. On the other hand, the square defined by the inequalities $|x| \leq 1, |y| \leq 1$ has exactly four extreme points, while the unit disk described by the inequality $x^2 + y^2 \leq 1$ has infinitely many. These examples raise the question of finding conditions under which a convex set has extreme points. The answer in general vector spaces is answered by one of the “big theorems” called the Krein-Milman Theorem. However, as we will see presently, our study of the linear programming problem actually answers this question for convex polytopes without needing to call on that major result.

The algebraic characterization of the vertices of the feasible polytope confirms the observation that we made by following the steps of the Simplex Algorithm in our introductory example. Some of the techniques used in proving the preceeding theorem come into play in making this characterization as we will now discover.

**Theorem 1.1**
The set of extreme points, $E$, of the feasible region $P$ is exactly the set, $B$ of all basic feasible solutions of the linear programming problem.

**Proof:** We wish to show that $E = B$ so, as usual, we break the proof into two parts.

(a) $B \subset E$.

Suppose that $x^{(b)} \in B$. Then, for the index set $J(x^{(b)}) \subset \{1, 2, \ldots, n\}$ is defined by $j \in J(x^{(b)})$ if and only if $x_j^{(b)} > 0$. Now suppose that $x^{(b)} \not\in E$. Then there exist two distinct feasible points $y, z \in P$ and a $\lambda \in (0, 1)$ for which $x^{(b)} = (1 - \lambda) y + \lambda z$. Observe that for any integer $k \not\in J$, it must be true that $(1 - \lambda) y_k + \lambda z_k = 0$. Since $0 < \lambda < 1$ and $y, z \geq 0$, this implies that, for all such indices $k$, $y_k = z_k = 0$. 

Now $x^{(b)}$ is basic, so that the columns of $A$ corresponding to the non-zero components form a linearly independent set in $\mathbb{R}^n$. Since the only non-zero components of $y$ and $z$ have the same indices as the non-zero components of $x^{(b)}$, we have span$\{a^{(j)}\}, j \in J(x^{(b)})$ contains $x^{(b)}, y$ and $z$. Moreover, since $y$ and $z$ are feasible, we have $Ay = b$ and $Az = b$ so that

$$b = (1 - \lambda) b + \lambda b = (1 - \lambda) Ay + \lambda Az.$$  

Now the system $Ax = b$ is uniquely solvable on the set $\{x \in \mathbb{R}^n | x_i = 0, i \notin J(x^{(b)})\}$, so that we must have $Ax^{(b)} = Ay = Az$ and hence $x^{(b)}$ cannot be written as a proper convex combination of two other distinct points of $P$, which means that $x^{(b)}$ is an extreme point of $P$.

(b) $E \subset B$

If $x^{(e)}$ is an extreme point of $P$, let us assume that the columns associated with non-zero components correspond to columns $a^{(j_i)}, i = 1, \ldots, k$. Then, as before, we have

$$x^{(e)}_{j_1} a^{(j_1)} + x^{(e)}_{j_2} a^{(j_2)} + \ldots + x^{(e)}_{j_k} a^{(j_k)} = b,$$

where the $x^{(e)}_{j_i} > 0, i = 1, \ldots, k$. To show that $x^{(e)}$ is a basic feasible solution, we must show that the set of vectors $\{a^{(j_1)}, a^{(j_2)}, \ldots, a^{(j_k)}\}$ is a linearly independent set. Let us suppose that this is not the case. Then there are constants $\alpha_{j_1}, \ldots, \alpha_{j_k}$, not all zero, such that

$$\alpha_{j_1} a^{(j_1)} + \alpha_{j_2} a^{(j_2)} + \alpha_{j_k} a^{(j_k)} = 0.$$

We set $\alpha = \text{col} (\alpha_{\ell})$ where $\alpha_{\ell} = 0$ if $\ell \neq j_i$ for some $i$, and $\alpha_{\ell} = \alpha_{j_i}$ if $\ell = j_i$ for some $i$. Then, since all the non-zero components of $x^{(e)}$ are strictly positive, it is possible to find an $\epsilon > 0$ such that

$$x^{(e)} + \epsilon \alpha \geq 0, \text{ and } x^{(e)} - \epsilon \alpha \geq 0.$$  

Note that each satisfies $Ax = b$ since $A(\epsilon \alpha) = \epsilon A(\alpha) = 0$ by the dependence relation.
We then have
\[ x^{(e)} = \frac{1}{2} (x^{(e)} + \epsilon \alpha) + \frac{1}{2} (x^{(e)} - \epsilon \alpha), \]
which expresses \( x^{(e)} \) as a proper convex combination of two distinct vectors in \( P \).
But this is impossible since \( x^{(e)} \) is an extreme point of \( P \). It follows that the set \( \{a^{(j_1)}, a^{(j_2)}, \ldots, a^{(j_k)}\} \) is a linearly independent set and so \( x^{(e)} \) is a basic feasible solution. \( \square \)

There are several results that are useful to state that follow immediately from this result.

**Corollary 1.2** If the convex set \( P \) is non-empty, then it has an extreme point.

Indeed, this fact follows from the Fundamental Theorem of LP and this last equivalence theorem.

**Corollary 1.3** If there is a finite optimal solution for the problem, then there is a finite optimal solution which is an extreme point.

**Corollary 1.4** The constraint set \( P \) contains at most a finite number of extreme points.

**Corollary 1.5** If the set \( P \) is bounded, then \( P \) consists of points that are finite linear combinations of a finite number of points of \( P \).

Finally, we have a direct proof of the following result:

**Theorem 1.6** If \( P \) is bounded, then a linear cost function \( \langle c, x \rangle \) achieves its minimum on \( P \) at an extreme point of \( P \).

**Proof:** Let \( x_1, x_2, \ldots, x_k \) be the extreme points of \( P \). Then any point \( x \in P \) can be written as a convex combination
\[ x = \sum_{i=1}^{k} \lambda_i x_i, \quad \text{where} \quad \lambda_i \geq 0, \ i = 1, 2, \ldots, k, \quad \text{and} \quad \sum_{i=1}^{k} \lambda_i = 1. \]
Then

$$\langle \mathbf{c}, \mathbf{x} \rangle = \sum_{i=1}^{k} \lambda_i \langle \mathbf{c}, \mathbf{x}_i \rangle.$$  

Now let $\iota := \min \{ \langle \mathbf{c}, \mathbf{x}_i \rangle \mid i = 1, 2, \ldots, k \}$. Then from the representation of $\mathbf{x}$ in terms of the extreme points we have

$$\langle \mathbf{c}, \mathbf{x} \rangle = \sum_{i=1}^{k} \lambda_i \langle \mathbf{c}, \mathbf{x}_i \rangle \geq (\lambda_1 + \lambda_2 + \ldots + \lambda_k) \iota = \iota,$$

and so the minimum of the cost functional is $\iota$ which is a value taken on at one of the extreme points. \qed