Supplementary notes on extreme points

Math 354-04, Spring 2006

This is a supplement to Sections 1.3 and 1.4 of the textbook, with a more detailed treatment of convex polyhedra and extreme points. We also establish the Extreme Point Theorem, which we state with a minor correction.

1. Lines in $\mathbf{R}^n$

Visualizing $\boldsymbol{R}^n$ for $n > 3$ is a challenge, to say the least. Nevertheless, effective and simple geometric ideas can be brought into the study of LPP's by thinking only about points and lines in $\boldsymbol{R}^n$, no matter how large n may be. Starting with any two distinct points (column vectors) $\boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{R}^n$, we can parametrize the line joining them as

$$\text{(A)} \qquad \boldsymbol{x} = \boldsymbol{x}(t) = \boldsymbol{u} + t(\boldsymbol{v} - \boldsymbol{u}) = (1-t)\boldsymbol{u} + t\boldsymbol{v}, \quad -\infty < t < \infty.$$

Each real number t corresponds to a unique point on the line, and conversely. From this parametrization we may think of the line as a "t-axis." The point $t = 0$ is $\boldsymbol{x} = \boldsymbol{u}$. The point $t = 1$ is $\boldsymbol{x} = \boldsymbol{v}$. The point $t = 1/2$ is the midpoint between $\boldsymbol{u}$ and $\boldsymbol{v}$. The line **segment** between $\boldsymbol{u}$ and $\boldsymbol{v}$ corresponds to the values $0 \le t \le 1$. Observe that the coefficients of $\boldsymbol{u}$ and $\boldsymbol{v}$ add up to 1 at **every** point of the whole line, and the line **segment** is precisely that portion of the line where the coefficients of $\boldsymbol{u}$ and $\boldsymbol{v}$ are both nonnegative.

2. Linear Functions on $\mathbf{R}^n$

A **linear function on $\boldsymbol{R}^n$** is a real-valued function f defined on $\boldsymbol{R}^n$ by a formula[1]

$$f(\boldsymbol{x}) = f(x_1, \ldots, x_n) = a_1 x_1 + \cdots + a_n x_n,$$

in which $a_1, \ldots, a_n$ are constants. More compactly:

$$\text{(B)} \qquad f(\boldsymbol{x}) = \boldsymbol{a}^T \boldsymbol{x}, \quad \text{where } \boldsymbol{a}^T = [a_1 \cdots a_n].$$

Linear functions occur in LPP's in two ways: (a) the objective function z is itself a linear function; and (b) every constraint is of the form $f(\boldsymbol{x}) \le b$, $f(\boldsymbol{x}) \ge b$ or $f(\boldsymbol{x}) = b$ for some linear function f and some constant b. Even the constraints $x_i \ge 0$ are of this type; for any fixed i, the function $f(\boldsymbol{x}) = x_i$ is definitely a linear function.

[1] Terminology disagreement: Some mathematicians allow a term $+b$ in a "linear" function. They would call our f a linear **homogeneous** function, to indicate that there is no $+b$ term on the right side. For these notes it is important, in Theorem 1 in particular, that there be no $+b$ term.

Theorem 1. *Any linear function f on $\boldsymbol{R}^n$ has the property:*

$$f(c\boldsymbol{x} + c'\boldsymbol{x}') = cf(\boldsymbol{x}) + c'f(\boldsymbol{x}') \quad \text{for all } c, c' \in \boldsymbol{R} \text{ and all } \boldsymbol{x}, \boldsymbol{x}' \in \boldsymbol{R}^n.$$

In the language of linear algebra, f is a linear transformation from $\boldsymbol{R}^n$ to $\boldsymbol{R}$.

Proof. With the notation of Eq. (B), this follows from basic rules of matrix algebra:

$$\begin{aligned} f(c\boldsymbol{x} + c'\boldsymbol{x}') &= \boldsymbol{a}^T(c\boldsymbol{x} + c'\boldsymbol{x}') = \boldsymbol{a}^T(c\boldsymbol{x}) + \boldsymbol{a}^T(c'\boldsymbol{x}') \\ &= c(\boldsymbol{a}^T\boldsymbol{x}) + c'(\boldsymbol{a}^T\boldsymbol{x}') = cf(\boldsymbol{x}) + c'f(\boldsymbol{x}'). \qquad \square \end{aligned}$$

Here is a comforting consequence, which is important for these notes: any linear function f, when the independent variable is "restricted" to a single line in $\boldsymbol{R}^n$, has the algebraic form familiar from beginning algebra.

Theorem 2. *Let ℓ be a line in $\boldsymbol{R}^n$. Let $\boldsymbol{u}$ and $\boldsymbol{v}$ be distinct points on ℓ and parametrize ℓ as in (A), with parameter t. Let f be any linear function on $\boldsymbol{R}^n$. Then on the line ℓ, f takes the following form when expressed in terms of the parameter t:*

(C) $$f(\boldsymbol{x}) = f(\boldsymbol{x}(t)) = Ct + D \quad \text{for certain constants } C \text{ and } D.$$

Proof. Let $C = f(\boldsymbol{v} - \boldsymbol{u})$ and $D = f(\boldsymbol{u})$. Then using Theorem 1,

$$f(\boldsymbol{x}) = f(\boldsymbol{x}(t)) = f[\boldsymbol{u} + t(\boldsymbol{v} - \boldsymbol{u})] = f(\boldsymbol{u}) + tf(\boldsymbol{v} - \boldsymbol{u}) = Ct + D. \quad \square$$

Corollary 3. *Let ℓ be a line in $\boldsymbol{R}^n$ and let f be a linear function on $\boldsymbol{R}^n$. Let U be a closed interval or closed ray in ℓ. Then one of the following holds:*

(1) *On U, f is constant.*
(2) *At points of U, f takes on values tending to ∞.*
(3) *On U, f assumes a maximum value at a unique point $\boldsymbol{u} \in U$, and $\boldsymbol{u}$ is an endpoint of U.*

Proof. Write $f(\boldsymbol{x}) = Ct + D$ as in Eq. (C). Case 1: Suppose that $C = 0$. Then f is constant on ℓ, hence f is constant on U, and (1) holds. Case 2: Suppose that $C > 0$. Since U is a closed interval or a closed ray, the points of U correspond to the values of the parameter t satisfying $a \leq t \leq b$, or $t \leq b$, or $t \geq a$, for some fixed real numbers a and/or b. If U is defined by $a \leq t \leq b$ or by $t \leq b$, then since $C > 0$, f assumes a maximum value only at the endpoint $t = b$. In these cases (3) holds. If U is defined by $t \geq a$, then (2) holds, again since $C > 0$. Case 3: $C < 0$. This is similar to Case 2; details are left to the reader. $\square$

3. Convex Polyhedra

Definition. *A* **convex polyhedron** *in* $\boldsymbol{R}^n$ *is the set of all solutions of a finite system of linear inequalities and equations, each of the form*

$$f(\boldsymbol{x}) \le b, \;\; f(\boldsymbol{x}) = b \text{ or } f(\boldsymbol{x}) \ge b$$

for some linear function f *and real number* b. (*Note: the plural of "polyhedron" is "polyhedra."*)

Since $f(\boldsymbol{x}) \ge b$ is equivalent to $-f(\boldsymbol{x}) \le -b$, and since $f(\boldsymbol{x}) = b$ is equivalent to the pair $f(\boldsymbol{x}) \le b$ and $f(\boldsymbol{x}) \ge b$, we can and shall assume for the purposes of these notes that all the conditions defining S are inequalities of the form $f(\boldsymbol{x}) \le b$. We don't have to do this, but it simplifies the discussion a bit.

Thus, a convex polyhedron in $\boldsymbol{R}^n$ is the simultaneous solution set of a set of finitely many inequalities

$$f_i(\boldsymbol{x}) \le b_i, \qquad i = 1, \ldots, N,$$

for some integer N, linear functions $f_1, \ldots, f_N$, and real numbers $b_1, \ldots, b_n$.

As examples: in $\boldsymbol{R}^2$, any convex polygon is a convex polyhedron. In $\boldsymbol{R}^3$, cubes, rectangular solids, parallelepipeds, tetrahedra (triangular pyramids), regular dodecahedra and regular icosahedra are convex polyhedra. And

THE FEASIBLE REGION OF ANY LPP IS A CONVEX POLYHEDRON IN $\boldsymbol{R}^n$, WHERE n IS THE NUMBER OF DECISION VARIABLES IN THE PROBLEM.

The terminology "convex polyhedron" is also consistent with the notion of "convex set" discussed in Section 1.3; any convex polyhedron S is a convex set, meaning that for any two points $\boldsymbol{u}, \boldsymbol{v} \in S$, the entire line segment joining $\boldsymbol{u}$ and $\boldsymbol{v}$ lies in S.

4. Extreme Points

Definition. *Let* S *be a convex polyhedron. An* **extreme point of** S *is a point* $\boldsymbol{e} \in S$ *such that* **no line segment entirely within** S **has** $\boldsymbol{e}$ **in its interior**.

Equivalently, a point $\boldsymbol{e} \in S$ is **not** an extreme point of S if and only if there exist distinct points $\boldsymbol{u}, \boldsymbol{v} \in S$, both different from $\boldsymbol{e}$, such that $\boldsymbol{e}$ lies on the line segment joining $\boldsymbol{u}$ and $\boldsymbol{v}$.

Examples: the extreme points of a convex polygon, a cube, a pyramid, or a parallelepiped are just its vertices.[2]

We know (Theorem 1.11) that for a LPP in super-canonical form, the extreme points of the feasible region are the basic feasible solutions, and these are finite in number.[3] In particular:

Theorem 4. *The feasible region of a LPP in canonical form possesses only finitely many extreme points.*

5. Finding Extreme Points in a Convex Polyhedron

The dirty work of these notes comes in this section. If S is a convex polyhedron and $\boldsymbol{x} \in S$, then $\boldsymbol{x}$ satisfies all the (linear) inequalities defining S. Each of these inequalities is of the form

$$f_i(\boldsymbol{x}) \leq b_i$$

where f_i is some linear function and b_i is a real number. For any such inequality, let us say that a point $\boldsymbol{x}_0 \in S$ **is on the edge** of that inequality if $f_i(\boldsymbol{x}_0) = b_i$. Thus if $f_i(\boldsymbol{x}_0) < b_i$, then $\boldsymbol{x}_0$ is **not** on the edge of the above inequality. A typical point $\boldsymbol{x} \in S$ satisfies all the defining inequalities, and $\boldsymbol{x}$ could be on the edge of some of those inequalities but not others. In fact it is rather special to be on the edge of one or more of the inequalities. Let's define

$$\begin{aligned} e(\boldsymbol{x}) &= \text{the } \textbf{edginess} \text{ of } \boldsymbol{x} \\ &= \text{the number of defining inequalities of } S \text{ for which } \boldsymbol{x} \text{ is on the edge.} \end{aligned}$$

Example: let S be defined by $.2x + .6y \leq 3000$, $.8x + .4y \leq 5000$, $x \geq 0$, $y \geq 0$. Then $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ is in S and has edginess 2 (it satisfies $x = 0$ and $y = 0$), while $\begin{bmatrix} 3000 \\ 4000 \end{bmatrix}$

[2] Perhaps surprisingly, there are convex polyhedra that have no extreme points. The x-axis in $\boldsymbol{R}^2$, defined by $y = 0$, is such an example. However, unless a convex polyhedron contains an entire line, it must have extreme points, cf. Theorem 6 below. In solving LPP's, we will stay away from such bizarre polyhedra by considering only those LPP's that are in standard or canonical form. The feasible region of such an LPP contains no lines, because of the constraints $x_1 \geq 0, \ldots, x_n \geq 0$.

[3] And any canonical form LPP is equivalent to a super-canonical form LPP or is infeasible. If the rows of A are not linearly independent, but the LPP is not infeasible, pick some equation from $A\boldsymbol{x} = \boldsymbol{b}$ which is a linear combination of the other equations and drop it. Continue doing this till super-canonical form is reached.

is in S and has edginess 1 (it satisfies $.2x + .6y = 3000$), and $\begin{bmatrix}1000\\2000\end{bmatrix}$ is in S with edginess 0. If you sketch S, you will see that the four extreme points have edginess 2, and the points on the boundary of S are just those with positive edginess.

We also will say that given a convex polyhedron S and points $\boldsymbol{x}, \boldsymbol{x}' \in S$, $\boldsymbol{x}'$ is **edgier**[4] than $\boldsymbol{x}$ if and only if $e(\boldsymbol{x}') > e(\boldsymbol{x})$.

The next lemma brings us to the critical idea for our proof of the Extreme Point Theorem. It implies that the edgiest points of a convex polyhedron are extreme points. But it says even more.

Lemma 5. *Suppose that S is a convex polyhedron containing no (complete) line. Let $\boldsymbol{x}_0 \in S$. If $\boldsymbol{x}_0$ is not an extreme point of S, then there is $\boldsymbol{X} \in S$ such that $\boldsymbol{X}$ is edgier than $\boldsymbol{x}_0$. In fact, for any line ℓ through $\boldsymbol{x}_0$ such that $\boldsymbol{x}_0$ is not an endpoint of $\ell \cap S$, any endpoint of $\ell \cap S$ is edgier than $\boldsymbol{x}_0$.*

Proof. Since $\boldsymbol{x}$ is not an extreme point of S, there is a line segment in S with $\boldsymbol{x}$ in its interior. Let ℓ be the line of which this segment is a part. Then $\boldsymbol{x}_0$ is not an endpoint of $\ell \cap S$. Moreover $\ell \cap S$ is not all of ℓ since S contains no line. For the rest of the proof these last two sentences are the only properties of ℓ and $\boldsymbol{x}_0$ that we shall use.

Pick any two points $\boldsymbol{u} \neq \boldsymbol{v}$ on ℓ and as before, parametrize ℓ as $\boldsymbol{x} = (1-t)\boldsymbol{u} + t\boldsymbol{v}$, $t \in \boldsymbol{R}$. Consider any one of the inequalities $f_i(\boldsymbol{x}) \leq b_i$ $(i = 1, \ldots, N)$ defining S. On ℓ, we know that this inequality takes the following form

$$C_i t + D_i \leq b_i$$

Such an inequality is

true for all t in the closed ray $t \leq (b_i - D_i)/C_i$,	if $C_i > 0$;
true for all t in the closed ray $t \geq (b_i - D_i)/C_i$,	if $C_i < 0$;
true for all t,	if $C_i = 0$ and $D_i \leq b_i$;
false for all t,	if $C_i = 0$ and $D_i > b_i$.

As $\ell \cap S$ is defined by finitely many such inequalities, and S does not contain all of ℓ, the only possibilities are that ℓ is either a closed ray ($t \geq a$ or $t \leq a$) or a closed interval ($a \leq t \leq b$), with $\boldsymbol{x}_0$ in the interior in either case.

Consider any defining inequality $f_i(\boldsymbol{x}) \leq b_i$ for S that $\boldsymbol{x}_0$ is on the edge of. That is, $f_i(\boldsymbol{x}_0) = b_i$. Then on $\ell \cap S$, f_i achieves its maximum b_i at $\boldsymbol{x}_0$, which is not

[4] We should probably say that $\boldsymbol{x}'$ is S-edgier than $\boldsymbol{x}$, because the definition clearly depends on S. However, for our purposes below, there will only be one S under consideration, so there should be no danger of confusion if we just say "edgier".

an endpoint. By Corollary 3, f_i must be constant on ℓ. So all points of $\ell \cap S$ are also on the edge of this inequality. Therefore in order to find a point $\boldsymbol{X}$ such that $e(\boldsymbol{X}) > e(\boldsymbol{x}_0)$, which is our goal, it remains only to find one point $\boldsymbol{X}$ on $\ell \cap S$ and a defining inequality for S with $\boldsymbol{X}$ on its edge but $\boldsymbol{x}_0$ not on its edge.

Where do we find such a point? Any **endpoint** $\boldsymbol{X}$ of $\ell \cap S$ will do: let us show that $\boldsymbol{X}$ is edgier than $\boldsymbol{x}_0$. The defining inequalities $f_i(\boldsymbol{x}) \leq b_i$ are all true for all $\boldsymbol{x} \in \ell$ between $\boldsymbol{X}$ and $\boldsymbol{x}_0$, but since $\boldsymbol{X}$ is an endpoint of $\ell \cap S$, at least one of the defining inequalities for S, say

$$(D) \qquad f_j(\boldsymbol{x}) \leq b_j$$

for a certain j, is false on that part of ℓ just beyond $\boldsymbol{X}$. Therefore $\boldsymbol{x}_0$ **is not** on the edge of inequality (D). (If it were on the edge, we have just seen that this inequality would be satisfied by **all** points of ℓ.) However, $f_j(\boldsymbol{X}) = b_j$, since the inequality $f_j(\boldsymbol{x}) \leq b_j$ has the form $Ct + D \leq b_j$ for some constants C, D, b_j, and is true on one side of $\boldsymbol{X}$ but not the other. Therefore $e(\boldsymbol{X}) \geq e(\boldsymbol{x}_0) + 1$. □

Theorem 6. *A nonempty convex polyhedron containing no line possesses at least one extreme point.*

Proof. Say that there are N inequalities defining the convex polyhedron S. Then obviously for any $\boldsymbol{x} \in S$, $e(\boldsymbol{x})$ is one of the numbers $0, 1, \ldots, N$. Hence we may choose $\boldsymbol{x}_0 \in S$ such that $e(\boldsymbol{x}_0)$ is as large as possible. By Lemma 5, $\boldsymbol{x}_0$ is an extreme point of S. □

6. Extreme Points of the Feasible Region, and the Objective Function

We have been considering a convex polyhedron S. For the study of a LPP the relevant example of S is the feasible region of the LPP. Now let's bring the objective function z of our LPP into the picture. By returning to the situation of Lemma 5 and Theorem 6, we can get a stronger conclusion. Given any $\boldsymbol{x}_0 \in S$, we will be able to find an extreme point $\boldsymbol{e}$ of S with the additional property:

such that $z(\boldsymbol{e}) \geq z(\boldsymbol{x}_0) \ldots$ unless z is unbounded on S.

Therefore (unless z is unbounded on the feasible region) a feasible solution x to a LPP only has to be as good as every extreme point in order to be an optimal solution. This observation, once proved, will be conceptually useful because there are only finitely many extreme points.

Lemma 7. *Suppose that S is a convex polyhedron containing no complete line. Let z be any linear function on $\boldsymbol{R}^n$. Let $\boldsymbol{x}_0 \in S$. Then one of the following holds:*

(1) *There is an extreme point $\boldsymbol{e}$ of S such that $z(\boldsymbol{e}) \geq z(\boldsymbol{x}_0)$.*
(2) *z is unbounded on S.*

Proof. Assume that z is not unbounded on S. We have to find an extreme point $\boldsymbol{e}$ of S as in (1). If $\boldsymbol{x}_0$ is an extreme point of S, then we can take $\boldsymbol{e} = \boldsymbol{x}_0$. So we can assume for the rest of the proof that $\boldsymbol{x}_0$ is not an extreme point. Then Lemma 5 applies, but let's actually re-run and improve the proof of Lemma 5. There's a line ℓ such that $\ell \cap S$ is either a ray or closed interval in ℓ, and $\boldsymbol{x}_0$ is in the interior of $\ell \cap S$. Moreover any endpoint $\boldsymbol{X}$ of $\ell \cap S$ is edgier than $\boldsymbol{x}_0$.

Consider z on $\ell \cap S$ and apply Corollary 3 again. By assumption z is not unbounded on S, so alternative (2) of Corollary 3 cannot hold. Therefore on $\ell \cap S$, z assumes a maximum at some endpoint $\boldsymbol{X}$ of $\ell \cap S$ (even if z is constant). Then $\boldsymbol{X} \in S$,

$$z(\boldsymbol{X}) \geq z(\boldsymbol{x}_0) \text{ and } \boldsymbol{X} \text{ is edgier than } \boldsymbol{x}_0.$$

Now replace $\boldsymbol{x}_0$ by $\boldsymbol{X}$ and repeat the argument. Either $\boldsymbol{X}$ is an extreme point of S (in which case (1) holds and we're done) or we get a point $\boldsymbol{Y}$ such that $z(\boldsymbol{Y}) \geq z(\boldsymbol{X})$ and $\boldsymbol{Y}$ is edgier than $\boldsymbol{X}$. And so on. But this cannot go on indefinitely, since $e(\boldsymbol{x}) \leq N$ for all $\boldsymbol{x} \in S$. So in no more than N steps, we reach some extreme point $\boldsymbol{W}$ of S such that $z(\boldsymbol{W}) \geq z(\boldsymbol{x}_0)$. □

7. The Extreme Point Theorem

Extreme Point Theorem[5]**.** *Consider any LPP in standard or canonical form, with feasible region S and objective function $z = z(\boldsymbol{x})$. Then*

(1) *The problem is infeasible ($S = \emptyset$), or*
(2) *The problem is unbounded (z is unbounded on S), or*
(3) *The problem has an optimal solution at some extreme point of S.*

Proof. If $S = \emptyset$ or z is unbounded on S, there's nothing to prove. So assume that $S \neq \emptyset$, and that z is bounded on S. Since the LPP is in standard or canonical form, every decision variable is constrained to be nonnegative, so S contains no line. (Any

[5] This is slightly different from Theorem 1.7, which is incorrect because of the nuisance possibility that for a general LPP, S might contain complete lines. It is no more than a nuisance, in the sense that every LPP can be restated in equivalent fashion so that S contains no complete lines.

line, as in Eq. (1), has some point with a negative coordinate, as can be found by taking t extremely large or extremely negative). So we are set up to use Lemma 7.

By Theorem 4, S has finitely many extreme points. Choose an extreme point $\boldsymbol{e}$ of S for which $z(\boldsymbol{e})$ is largest (compared to extreme points only). Claim: $\boldsymbol{e}$ is an optimal solution of the LPP. Proof: Let $\boldsymbol{y} \in S$. By Lemma 7, there is an extreme point $\boldsymbol{e}_0$ of S such that $z(\boldsymbol{e}_0) \geq z(\boldsymbol{y})$. By our choice of $\boldsymbol{e}$, $z(\boldsymbol{e}) \geq z(\boldsymbol{e}_0)$. Combining these, $z(\boldsymbol{e}) \geq z(\boldsymbol{y})$. Since $\boldsymbol{y} \in S$ was arbitrary, $\boldsymbol{e}$ is therefore optimal. □

A completely different proof of the Extreme Point Theorem can be given by a simple analysis of the two-phase method, but this involves a complication: one must prove that cycling can be avoided, e.g. by using Bland's Rule.

8. Bounded Convex Polyhedra

When a LPP has a feasible region that is **bounded**, the Extreme Point Theorem can also be understood from a somewhat different point of view.

Definition. *Let S be a subset of $\boldsymbol{R}^n$. Then S is* **bounded** *if and only if there exists a number B such that $|x_i| \leq B$ for all $\boldsymbol{x} = [x_1 \cdots x_n]^T \in S$.*

Be careful here: this definition does not directly address the boundedness or unboundedness of the objective function of a LPP on the feasible region S. It's just a matter of whether S contains points with arbitrarily large (or negatively large) coordinates. Notice: if S is bounded, then it cannot contain an entire line, or even a ray.

Definition. *Let $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m \in \boldsymbol{R}^n$. A* **convex combination** *of $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m$ is a linear combination*

$$\boldsymbol{x} = c_1\boldsymbol{x}_1 + \cdots + c_m\boldsymbol{x}_m$$

$$\text{such that } c_1 + \cdots + c_m = 1 \text{ and } c_i \geq 0 \text{ for all } i = 1, \ldots, m.$$

Examples: **The set of all convex combinations of two distinct points is the line segment joining them.** If $\boldsymbol{u}$ and $\boldsymbol{v}$ are distinct elements of $\boldsymbol{R}^n$, these combinations are $c\boldsymbol{u} + d\boldsymbol{v}$ with $c, d \geq 0$ and $c + d = 1$, that is, $(1 - d)\boldsymbol{u} + d\boldsymbol{v}$ with $0 \leq d \leq 1$.

The set of all convex combinations of three noncollinear points in $\boldsymbol{R}^n$ is the triangle (interior and boundary) formed by those points.

The following lemma helps to understand this; any point $\boldsymbol{y}$ on the perimeter of the triangle is a convex combination of two of the vertices. Then any point in the interior of the triangle is on some line segment joining some two points on the perimeter. So $\boldsymbol{y}$ is a convex combination of two convex combinations of the vertices of the triangle.

Lemma 8. *Suppose that $\boldsymbol{y}^1, \dots, \boldsymbol{y}^r$ are all convex combinations of $\boldsymbol{x}_1, \dots, \boldsymbol{x}_m$. If $\boldsymbol{w}$ is any convex combination of $\boldsymbol{y}^1, \dots, \boldsymbol{y}^r$, then $\boldsymbol{w}$ is also a convex combination of $\boldsymbol{x}_1, \dots, \boldsymbol{x}_m$.*

Proof. The assumptions mean that there are scalars c_{ij}, $i = 1, \dots, m$, $j = 1, \dots, r$, and d_j, $j = 1, \dots, r$, such that

$$\boldsymbol{y}^j = c_{1j}\boldsymbol{x}_1 + \cdots + c_{mj}\boldsymbol{x}_m \text{ for each } j = 1, \dots, r, \text{ and } \boldsymbol{w} = d_1\boldsymbol{y}^1 + \cdots d_r\boldsymbol{y}^r,$$
$$\text{all } c_{ij} \geq 0,\ d_j \geq 0,\ d_1 + \cdots + d_r = 1,\ c_{1j} + \cdots + c_{mj} = 1.$$

Then $\boldsymbol{w} = \sum_{j=1}^r d_j \boldsymbol{y}^j = \sum_{j=1}^r d_j \sum_{i=1}^m c_{ij}\boldsymbol{x}_i = \sum_{i=1}^m b_i \boldsymbol{x}_i$, where $b_i = \sum_{j=1}^r d_j c_{ij}$ for each $i = 1, \dots, m$. As all $c_{ij} \geq 0$ and all $d_j \geq 0$, clearly $b_i \geq 0$ for all i. Furthermore

$$\sum_{i=1}^r b_i = \sum_{i=1}^m \sum_{j=1}^r d_j c_{ij} = \sum_{j=1}^r d_j \left(\sum_{i=1}^m c_{ij} \right) = \sum_{j=1}^r d_j = 1.$$

Therefore $\boldsymbol{w}$ is a convex combination of $\boldsymbol{x}_1, \dots, \boldsymbol{x}_m$. □

Corollary 9. *Given any subset $V \subseteq \boldsymbol{R}^n$, the set of all convex combinations of (finite subsets of) elements of V is a convex set. (It is called the* **convex hull** *of V.)*

Proof. Suppose that $\boldsymbol{u}, \boldsymbol{v}$ are convex combinations of elements of S. By adding terms with 0 coefficients, we can consider them to be convex combinations of the same (finite number of) elements of S. The line segment joining $\boldsymbol{u}$ and $\boldsymbol{v}$ consists of all the points $\boldsymbol{x} = (1-t)\boldsymbol{u} + t\boldsymbol{v}$, $0 \leq t \leq 1$, and these are all convex combinations of $\boldsymbol{u}$ and $\boldsymbol{v}$. They are therefore convex combinations of elements of S by Lemma 8. □

Theorem 10. *Let S be convex polyhedron in $\boldsymbol{R}^n$, and assume that S is* **bounded**. *Let E be the set of all extreme points of S. Then every element of S is a convex combination of elements of E. (Restatement: S is the convex hull of its extreme points.)*

(Remark: The same assertion can be proved for any bounded and closed convex set S. However, for a convex polyhedron S, the set E is finite, while for a general convex set S, such as a disk, E might be infinite.)

Proof. As before, S is defined by finitely many inequalities $f_i(\boldsymbol{x}) \leq b_i$, $i = 1, \ldots, r$. Again for each $\boldsymbol{x}_0 \in S$, let $e(\boldsymbol{x})$ be the number of these inequalities such that $f_i(\boldsymbol{x}_0) = b_i$, i.e., the number of inequalities with $\boldsymbol{x}_0$ on the edge.

We suppose that the theorem is false, and try for a contradiction. Since the theorem is false there is $\boldsymbol{x} \in S$ such that $\boldsymbol{x}$ is not a convex combination of elements of E. Among all such $\boldsymbol{x}$, we choose one such that $e(\boldsymbol{x})$ is as large as possible. (Note: the possible values of $e(\boldsymbol{x})$ are the integers from 0 to the number of defining inequalities of S.)

Now $\boldsymbol{x}$ cannot be an extreme point of S ($\boldsymbol{x} = 1 \cdot \boldsymbol{x}$ is a convex combination). So there is a line ℓ such that $\boldsymbol{x}$ is in the interior of $\ell \cap S$. Since S is bounded, $\ell \cap S$ must be a closed interval in ℓ. Let its endpoints be $\boldsymbol{u}$ and $\boldsymbol{v}$. By Lemma 5, $e(\boldsymbol{u}) > e(\boldsymbol{x})$ and $e(\boldsymbol{v}) > e(\boldsymbol{x})$. Therefore by our maximal choice of $e(\boldsymbol{x})$, the endpoints $\boldsymbol{u}$ and $\boldsymbol{v}$ must be convex combinations of elements of E. However, $\boldsymbol{x} = (1-t)\boldsymbol{u} + t\boldsymbol{v}$ for some $0 \leq t \leq 1$, so $\boldsymbol{x}$ is a convex combination of $\boldsymbol{u}$ and $\boldsymbol{v}$. Combining these facts with Lemma 8, we deduce that $\boldsymbol{x}$ is a convex combination of elements of E. This contradicts our choice of $\boldsymbol{x}$. $\square$

As a result, there is a pretty proof of the Extreme Point Theorem in the special case that the feasible region is bounded. Corollary 11, to repeat, is a special case of the Extreme Point Theorem.

Corollary 11. *Consider any LPP whose feasible region S is nonempty and bounded. Then some extreme point of S is an optimal solution.*

Proof. We already know that S has finitely many extreme points $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_r$, say. Let z be the objective function as usual. As before, we pick an extreme point (for simplicity assume it's $\boldsymbol{e}_1$) such that $z(\boldsymbol{e}_1) \geq z(\boldsymbol{e}_i)$ for all $i = 1, \ldots, r$, and prove that $\boldsymbol{e}_1$ is optimal.

Take any $\boldsymbol{y} \in S$. By Theorem 10, there are scalars $c_1, \ldots, c_r$ such that

$$\begin{aligned} &\boldsymbol{y} = c_1\boldsymbol{e}_1 + \cdots + c_r\boldsymbol{e}_r, \\ &c_i \geq 0 \quad \text{for all } i = 1, \ldots, r, \text{ and} \\ &c_1 + \cdots + c_r = 1. \end{aligned}$$

$$\begin{aligned} \text{Then } z(\boldsymbol{y}) &= z(c_1\boldsymbol{e}_1 + \cdots + c_r\boldsymbol{e}_r) \\ &= c_1 z(\boldsymbol{e}_1) + c_2 z(\boldsymbol{e}_2) + \cdots + c_r z(\boldsymbol{e}_r) \quad \text{by Theorem 1} \\ &\leq c_1 z(\boldsymbol{e}_1) + c_2 z(\boldsymbol{e}_1) + \cdots + c_r z(\boldsymbol{e}_1) \quad \text{since } c_i \geq 0 \text{ for all } i \\ &= (c_1 + \cdots + c_r) z(\boldsymbol{e}_1) = 1 \cdot z(\boldsymbol{e}_1) = z(\boldsymbol{e}_1), \text{ which proves optimality.} \quad \square \end{aligned}$$