# LINEAR PROGRAMMING

MATTHEW GREENBERG

## 1. Separation

**Theorem 1.1.** *Let $S$ be a closed, convex subset of $\mathbb{R}^n$, let $x \in \mathbb{R}^n \setminus S$, and let $y^* \in S$ be the point closest to $x$. Then*

$$(x - y^*) \cdot (y - y^*) \leq 0$$

*for all $y \in S$.*

*Proof.* Let $y \in S$. The result clearly holds if $y = y^*$, so assume that $y \neq y^*$. Suppose, for the purpose of getting a contradition, that

$$(x - y^*) \cdot (y - y^*) > 0.$$

Let $L$ be the line through $y^*$ and $y$ and let $z$ be the point on $L$ closest to $x$:

$$z = (1-t)y^* + ty, \quad \text{where} \quad t = \frac{(x - y^*) \cdot (y - y^*)}{\|y - y^*\|^2} > 0.$$

By the law of cosines,

$$\begin{aligned} \|x - y\|^2 &= \|x - y^*\|^2 + \|y^* - y\|^2 - 2(x - y^*)^T (y - y^*) \\ &= \|x - y^*\|^2 + \|y^* - y\|^2 - 2t\|y - y^*\|^2 \\ &= \|x - y^*\|^2 + (1 - 2t)\|y^* - y\|^2. \end{aligned}$$

Therefore,

$$1 - 2t = \frac{\|x - y\|^2 - \|x - y^*\|^2}{\|y^* - y\|} > 0$$

as $\|x - y^*\| < \|x - y\|$, and it follows that

$$0 < t < \frac{1}{2}.$$

Therefore, $z \in (y^*, y) \subseteq S$. By the Pythagorean theorem,

$$\|x - z\|^2 < \|x - z\|^2 + \|z - y^*\|^2 = \|x - y^*\|^2,$$

contradicting the minimality of $\|x - y^*\|$. $\square$

*Date*: May 4, 2022.

**Corollary 1.2.** *Let $S$ be a closed, convex subset of $\mathbb{R}^n$ and let $x \in \mathbb{R}^n \setminus S$. Then there is a vector $a \in \mathbb{R}^n$ such that*

$$a \cdot x > a \cdot y \quad \textit{for all} \quad y \in S.$$

*Proof.* Let $y^* \in S$ be the point of $S$ closest to $x$ and take $a = x - y^*$. □

*Exercise:* Let $S$ be a convex subset of $\mathbb{R}^n$. Prove that $\operatorname{int} \bar{S} = \operatorname{int} S$. In particular, $S = \mathbb{R}^n$ if and only if $\bar{S} = \mathbb{R}^n$.

**Corollary 1.3.** *Let $S$ be a convex subset of $\mathbb{R}^n$ with $S \neq \mathbb{R}^n$. Then $S$ is contained in some half-space.*

**Corollary 1.4.** *Let $S$ be a nonempty, closed, convex cone in $\mathbb{R}^n$ with $S \neq \mathbb{R}^n$. Then there is a nonzero vector $a \in \mathbb{R}^n$ such that $a \cdot y \leq 0$ for all $y \in S$.*

*Proof.* Let $x \in \mathbb{R}^n \setminus S$. Then, by the preceding theorem, there is a vector $a \in \mathbb{R}^n$, such that $a \cdot x > a \cdot y$ for all $y \in S$. Since $S$ is assumed nonempty, $a \neq 0$. Since $S$ is a closed, convex cone, $0 \in S$. Therefore, $a \cdot x > 0$. Suppose $y \in S$ and $a \cdot y > 0$. Then

$$t := \frac{a \cdot x}{a \cdot y} > 0,$$

so $ty \in C$ as $C$ is a cone. But then

$$a \cdot x > a \cdot (ty) = t(a \cdot y) = a \cdot x,$$

a contradiction. □

**Theorem 1.5.** *Let $S$ be an $n$-dimensional, convex subset of $\mathbb{R}^n$. Let $y \in \mathbb{R}^n$, and let*

$$N_y := \{v \in \mathbb{R}^n : y + \varepsilon v \in S \textit{ for some } \varepsilon > 0\}$$

*be the cone of feasible directions at $y$. Then $N_y = \mathbb{R}^n$ if and only if $y \in \operatorname{int} S$.*

*Proof.* Suppose $N_y = \mathbb{R}^n$. Then there is an $\varepsilon > 0$ such that $y \pm \varepsilon e_i \in S$ for all $i = 1, \ldots, n$. The convex hull of these points is contained in $S$ and contains $y$ in its interior, putting $y$ in $\operatorname{int} S$.

Conversely, suppose $N_y \neq \mathbb{R}^n$. By Corollary 1.4, there is a vector nonzero $a \in \mathbb{R}^n$ such that $a \cdot v \leq 0$ for all $v \in N_y$. In particular, $a \notin N_y$. Therefore, $y + \varepsilon a \notin S$ for any $\varepsilon > 0$. Therefore, $y \notin \operatorname{int} S$. □

**Theorem 1.6.** *Let $S$ be a convex subset of $\mathbb{R}^n$ and let $x \in \mathbb{R}^n$. Then $x \in \operatorname{bd} S$ if and only if there is a closed, convex cone $N_x \subsetneq \mathbb{R}^n$ such that $S \subseteq x + N_x$.*

**Theorem 1.7.** *Let $S$ be an $n$-dimensional, convex subset of $\mathbb{R}^n$ and let $x \in \operatorname{bd} S$. Then there is a nonzero vector $a \in \mathbb{R}^n$ such that $a \cdot y \leq a \cdot x$ for all $y \in S$.*

*Proof.* We can assume, without loss of generality, that $S$ is closed. Since $x \in \operatorname{bd} S$, $x \notin \operatorname{int} S$. Therefore, $N_y \neq \mathbb{R}^n$ and, consequently, $N_y^* \neq \mathbb{R}^n$. Let $a \in -N_y^*$, so that $a \cdot v \leq 0$ for all $v \in N_y$. Let $y \in S$. Then $y - x \in N_y$, so $a \cdot (y - x) \leq 0$. Rearranging gives $a \cdot y \leq a \cdot x$. □

**Corollary 1.8.** *Let $S$ be an $n$-dimensional, closed, convex subset of $\mathbb{R}^n$ and let $x \in \operatorname{bd} S$. Then $x$ is contained in a face of $S$ of dimension $< n$.*

Let $S$ be a convex subset of $\mathbb{R}^n$. A hyperplane $H$ in $\mathbb{R}^n$ is called a *supporting hyperplane* of $S$ if

(1) $S$ is contained in one of the closed half-spaces associated to $H$,
(2) $\bar{S} \cap H \neq \emptyset$.

**Theorem 1.9.** *Let $\mathcal{H}_S$ be the set of supporting hyperplanes of $S$. For each $H \in \mathcal{H}_S$, choose $a_H \in \mathbb{R}^n$ and $b_H \in \mathbb{R}$ such that*

$$H = \{x \in \mathbb{R}^n : a_H \cdot x = b_H\}$$

*and*

$$S \subseteq \{x \in \mathbb{R}^n : a_H \cdot x \leq b_H\} =: H^-.$$

*Then*

$$\bar{S} = \bigcap_{H \in \mathcal{H}_S} H^-.$$

*Proof.* Obviously, $\bar{S} \subset \bigcap_{H \in \mathcal{H}_S} H^-$. Suppose $x \notin \bar{S}$. Let $y^*$ be the point in $\bar{S}$ closest to $x$. Then

$$(x - y^*) \cdot (y - y^*) \leq 0$$

for all $y \in \bar{S}$. Let $a = x - y^*$, let $b = a \cdot y^*$, and let

$$H = \{z \in \mathbb{R}^n : a \cdot z = b\}.$$

Then $\bar{S} \subseteq H^-$ and $y^* \in \bar{S} \cap H$, so $H \in \mathcal{H}_S$. Since $y^* \neq x$,

$$a \cdot (x - y^*) = \|x - y^*\|^2 > 0.$$

Therefore, $a \cdot x > b$ and $x \notin H^-$. Therefore, $x \notin \bigcap_{H \in \mathcal{H}_S} H^-$. □

A *face* of $S$ is the intersection of $S$ with one of its supporting hyperplanes.

## 2. General position

**Theorem 2.1.** *The following are equivalent for points $x_0, \dots, x_k \in \mathbb{R}^n$:*

(1) *$x_0, \dots, x_k$ are not contained in any $(k-1)$-dimensional, affine subspace of $\mathbb{R}^n$.*
(2) *$x_1 - x_0, \dots, x_k - x_0$ are linearly independent.*

*Proof.*

(1) implies (2): Suppose $x_1 - x_0, \dots, x_k - x_0$ are linearly dependent. Then they lie in a $(k-1)$-dimensional subspace $U$ of $\mathbb{R}^n$ and $Z = U + x_0$ is a $(k-1)$-dimensional, affine subspace of $\mathbb{R}^n$ containing $x_0, \dots, x_k$.

(2) implies (1): Suppose $x_0, \dots, x_k \in Z$, where $Z$ is a $(k-1)$-dimensional, affine subspace of $\mathbb{R}^n$. Then $U := Z - x_0$ is a $(k-1)$-dimensional subspace of $\mathbb{R}^n$ and, therefore, the vectors $x_1 - x_0, \dots, x_k - x_0 \in U$ are linearly dependent. □

**Theorem 2.2.** *Let $x_0, \dots, x_n \in \mathbb{R}^n$ and let*

$$\Delta(x_0, \dots, x_n) = \{t_0 x_0 + \cdots t_n x_n : t \in \Delta^n\}.$$

*Then $x_0, \dots, x_n$ are in general position if and only if*

$$\operatorname{int} \Delta(x_0, \dots, x_n) = \{t_0 x_0 + \cdots t_n x_n : t \in \operatorname{int} \Delta^n\}.$$

*Proof.*

$$\begin{aligned} \operatorname{int} \Delta(x_0, \dots, x_n) &= x_0 + \operatorname{int} \Delta(0, x_1 - x_0, \dots, x_n - x_0) \\ &= x_0 + \operatorname{int} \left\{ \sum_{i=1}^n t_i (x_i - x_0) : t_i \geq 0,\ \sum_{i=1}^n t_i \leq 1 \right\} \end{aligned}$$

Suppose $x_0, \dots, x_n$ are in general position. Then $x_1 - x_0, \dots, x_n - x_0$ are a basis of $\mathbb{R}^n$ and there is a unique linear automorphism $f : \mathbb{R}^n \to \mathbb{R}^n$ such that $f(e_i) = x_i - x_0$, where $e_1, \dots, e_n$ is the standard basis of $\mathbb{R}^n$. Applying $f$ to the identity

$$\operatorname{int} \left\{ t \in \mathbb{R}^n : t_i \geq 0,\ \sum_{i=1}^n t_i \leq 1 \right\} = \left\{ t \in \mathbb{R}^n : t_i > 0,\ \sum_{i=1}^n t_i < 1 \right\}$$

gives

$$\operatorname{int}\left\{\sum_{i=1}^{n} t_i(x_i - x_0) : t_i \geq 0,\ \sum_{i=1}^{n} t_i \leq 1\right\}$$
$$= \left\{\sum_{i=1}^{n} t_i(x_i - x_0) : t_i > 0,\ \sum_{i=1}^{n} t_i < 1\right\}.$$

Therefore,

$$x_0 + \operatorname{int}\left\{\sum_{i=1}^{n} t_i(x_i - x_0) : t_i \geq 0,\ \sum_{i=1}^{n} t_i \leq 1\right\}$$
$$= x_0 + \left\{\sum_{i=1}^{n} t_i(x_i - x_0) : t_i > 0,\ \sum_{i=1}^{n} t_i < 1\right\}$$
$$= \left\{\sum_{i=0}^{n} t_i x_i : t \in \operatorname{int} \Delta^n\right\}.$$

Conversely, suppose $x_0, \ldots, x_n$ are not in general position. Then they lie in an $(n-1)$-dimensional, affine subspace $Z$ of $\mathbb{R}^n$. As $\Delta(x_0, \ldots, x_n)$ is containe in $Z$ and $Z$ has empty interior, so does $\Delta(x_0, \ldots, x_n)$. On the other hand, $\{t_0 x_0 + \cdots t_n x_n : t \in \operatorname{int} \Delta^n\}$ is obviously nonempty. $\square$

## 3. Affinity

An *affine combination* of points $x_1, \ldots, x_m \in \mathbb{R}^n$ is a point of the form $\sum t_i x_i$, where $\sum t_i = 1$.

A subset $A$ of $\mathbb{R}^n$ is a *affine subspace of* $\mathbb{R}^n$ if it is closed under convex combination.

**Lemma 3.1.** *Let $A \subset \mathbb{R}^n$. Then $A$ is affine if and only if for every pair $x, y$ of distinct points of $A$, the line through $x$ and $y$ is contained in $A$.*

**Lemma 3.2.** *The following are equivalent for a subset $A$ of $\mathbb{R}^n$.*

1. *$A$ is an affine subspace of $\mathbb{R}^n$.*
2. *$A - a$ is a linear subspace of $\mathbb{R}^n$ for any $a \in A$.*
3. *$A - a$ is a linear subspace of $\mathbb{R}^n$ for some $a \in A$.*

*Proof.* (1) implies (2): Let $A$ is an affine subspace of $\mathbb{R}^n$ and let $a \in A$. Let $x, y \in A - a$. Then

$$x + y + a = (x + a) + (y + a) - a$$

is an affine combination of $x + a$, $y + a$, and $a$, all elements of $A$. Therefore, $x + y + a \in A$ and $x + y \in A - a$. Now let $x \in A - a$ and let $t \in \mathbb{R}$. Then

$$tx + a = t(x + a) + (1 - t)a$$

is an affine combination of $x+a$ and $a$, both in $A$. Therefore, $tx+a \in A$ and $tx \in A - a$. Thus, $A - a$ is a linear subspace of $\mathbb{R}^n$.

(2) implies (3): Obvious.

(3) implies (1): Let $A \subset \mathbb{R}^n$, let $a \in A$, and suppose that $A - a$ is a linear subspace of $\mathbb{R}^n$. Suppose $\sum t_i a_i$ is an affine combination of $a_1, \ldots, a_n \in A$. Then

$$\begin{aligned} \sum t_i a_i &= \sum t_i(a_i - a) + \sum t_i a \\ &= \sum t_i(a_i - a) + a \\ &\in (A - a) + a \\ &= A. \end{aligned}$$

Thus, $A$ is an affine subspace of $\mathbb{R}^n$. $\square$

Let $A$ be an affine subspace of $\mathbb{R}^n$. Call a vector $x \in \mathbb{R}^n$ a *translation of* $A$ if $A + x \subseteq A$. Let $T(A)$ be the set of all translations of $A$.

**Lemma 3.3.** *Let $A$ be an affine subspace of $\mathbb{R}^n$. Then:*

1. *For any $a \in A$, $T(A) = A - a$.*
2. *$T(A)$ is a linear subspace of $\mathbb{R}^n$.*

*Proof.* It is clear that $T(A)$ is closed under addition. To see that it's closed under scalar multiplication, let $x \in T(A)$ and let $t \in \mathbb{R}$. If $a \in A$, then $x + a \in A$ and

$$a + tx = (1 - t)a + t(x + a),$$

being a convex combination of elements of $A$, also belongs to $A$.

Let $x \in A - a$ and let $b \in A$. Then

$$b + x = b + (x + a) - a,$$

being a convex combination of elements of elements of $A$, belongs to $A$. Therefore, $A - a \subseteq T(A)$. Conversely, suppose $x \in T(A)$. Then $a + x \in A$, so $x \in A - a$. Thus, $T(A) \subseteq A - a$. $\square$

**Lemma 3.4.** *Let $x, y \in \mathbb{R}^n$ and let $V$ and $W$ be a linear subspaces of $\mathbb{R}^n$. Then $x + V = y + W$ if and only if $V = W$ and $y - x \in V$. In* *particular, the linear subspace $W$ is uniquely determined by the affine subspace $x + W$.*

*Proof.* Since $0 \in W$, $y \in x + W$ and $y - x \in W$. Let $v \in V$. Then $x + v = y + w$ for some $w \in W$ and $v = (y - x) + w \in W$. Therefore, $V \subseteq W$. The reverse containment holds by symmetry, so $V = W$. □

The *dimension* $\dim A$ of an affine subset $A$ of $\mathbb{R}^n$ is defined by

$$\dim A := \dim T(A).$$

The *affine hull* of a set $S \subseteq \mathbb{R}^n$, denoted $\operatorname{aff} S$, is the set of affine combinations of finite subsets of $S$.

**Lemma 3.5.** *Let $S \subseteq \mathbb{R}^n$. Then $\operatorname{aff} S$ is an affine subspace of $\mathbb{R}^n$.*

## 4. Convexity

The *standard $n$-simplex* is

$$\Delta^n = \left\{ (t_0, \dots, t_n) \in [0,1]^{n+1} : \sum t_i = 1 \right\} \subseteq \mathbb{R}^{n+1}.$$

A *convex combination* of points $x_1, \dots, x_m \in \mathbb{R}^n$ is a point of the form $\sum t_i x_i$, where $t \in \Delta^{m-1}$. A set $S \subset \mathbb{R}^n$ is *convex* if it is closed under convex combination.

A convex combination of points $x, y \in \mathbb{R}^n$ is a point of the form $(1 - t)x + ty$, where $t \in [0, 1]$.

Define the *closed interval* $[x, y]$ to be the set of all convex combinations of $x$ and $y$:

$$[x, y] = \{(1 - t)x + ty : t \in [0, 1]\}.$$

**Lemma 4.1.** *The following are equivalent for $S \subset \mathbb{R}^n$:*

(1) *$S$ is convex.*
(2) *$[x, y] \subseteq S$ for all $x, y \in S$.*

*Proof.* (1) implies (2): Obvious.

(2) implies (1): Suppose (2) holds. Let $x_0, \dots x_m \in S$ and let $t \in \Delta^m$. We prove that $y := \sum t_i x_i \in S$ by induction on $m$. The case $m = 1$ follows directly from our hypothesis. Now suppose that the result holds for $m - 1$. Let

$$u_i = \frac{t_i}{1 - t_m}, \qquad i < m.$$

Then $u \in \Delta^{m-1}$ as

$$t_0 + \cdots + t_{m-1} = 1 - t_m.$$

By the inductive hypothesis,

$$\sum_{i<m} u_i x_i \in S.$$

Therefore,

$$y = (1 - t_m)z + t_m x_m \in [z, x_m] \subseteq S. \qquad \square$$

A set $S \subseteq \mathbb{R}^n$ is *convex* if $[x, y] \subseteq S$ for all $x, y \in S$.

**Lemma 4.2.** *Let $A$ be an affine subset of $\mathbb{R}^n$. Then $A$ is convex.*

*Proof.* Convex combinations are affine combinations. $\square$

Let $S$ be a convex subset of $\mathbb{R}^n$. The *dimension* of $S$, denoted $\dim S$, is defined by

$$\dim S = \dim \operatorname{aff} S.$$

Let $X$ be a subset of $\mathbb{R}^n$. The *convex hull* of $X$, denoted $\operatorname{conv} X$, is the set of convex combinations of finite subsets of $X$. It is easy to see that $X$ is convex if and only if $X = \operatorname{conv} X$.

**Theorem 4.3** (Carathéodory)**.** *Let $X \subset \mathbb{R}^n$. Then every point in* $\operatorname{conv} X$ *is a combination of $n + 1$ elements of $X$.*

*Proof.* Suppose not. Then there is a point $x \in \mathbb{R}^n$ and a $p > n+1$ such that

(1) $x$ is a convex combination of $p$ points $x_1, \ldots, x_p \in X$, and
(2) $x$ cannot be expressed as a convex combination of fewer than $p$ points of $X$.

Let $t \in \Delta^{p-1}$ be such that

$$x = \sum_{i=1}^{p} t_i x_i. \tag{†}$$

By (2), we must have $t_i > 0$ for all $i$.

Identify $\mathbb{R}^n$ with $\mathbb{R}^{n\times 1}$ and let

$$A = \begin{pmatrix} 1 & \cdots & 1 \\ x_1 & \cdots & x_p \end{pmatrix} \in \mathbb{R}^{(n+1)\times p}.$$

Since $p > n+1$, the nullspace $N(A)$ is nonzero. Let $s \in N(A)$ be such that $s_i > 0$ for some $i$. Then

$$\sum_{i=1}^{p} s_i = 0 \qquad \text{and} \qquad \sum_{i=1}^{p} s_i x_i = 0.$$

Let $i_0$ be an index such that

$$\frac{t_{i_0}}{s_{i_0}} = \min\left\{\frac{t_i}{s_i} : s_i > 0\right\}$$

and solve for $x_{i_0}$ in terms of $x_i$, $i \neq i_0$:

$$x_{i_0} = -\sum_{i \neq i_0} \frac{s_i}{s_{i_0}} x_i$$

Substituting into (†) gives

$$(\ddagger) \qquad x = \sum_{i \neq i_0} r_i x_i, \qquad r_i := t_i - \frac{t_{i_0}}{s_{i_0}} s_i.$$

If $s_i \leq 0$, then $r_i > 0$ as $t_i$, $t_{i_0}$, and $s_{i_0}$ are all positive. If $s_i > 0$, then $r_i \geq 0$ by definition of $i_0$.

Since $r_{i_0} = 0$ and $s_1 + \cdots + s_p = 0$,

$$\sum_{i \neq i_0} r_i = \sum_{i=1}^{p} r_i = \sum_{i=1}^{p} t_i - \frac{t_{i_0}}{s_{i_0}} \sum_{i=1}^{p} s_i = 1 - 0 = 1.$$

Thus, (‡) expresses $x$ as a convex combination of the $p$ vectors $x_i$, $i \neq i_0$, contradicting (2). □

**Corollary 4.4.** *Let $K \subset \mathbb{R}^n$ be a compact set. Then* $\operatorname{conv} K$ *is compact.*

*Proof.* Define

$$f : \Delta^n \times K^{n+1} \longrightarrow \mathbb{R}^n$$

by

$$f(t, (x_0, \ldots, x_n)) = \sum_{i=0}^{n} t_i x_i.$$

Then $f$ is continuous and, by Carathéodory's Theorem, has image $\operatorname{conv} K$. The result follows. □

A function $f : \mathbb{R}^n \to \mathbb{R}$ is *convex* if

$$f((1-t)x + ty) \leq (1-t)f(x) + tf(y)$$

for all $x, y \in \mathbb{R}^n$ and all $t \in [0, 1]$.

**Lemma 4.5.** *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function. Then*

$$S := \{x \in \mathbb{R}^n : f(x) \leq M\}$$

*is convex.*

*Proof.* Let $x, y \in S$ and let $t \in [0, 1]$. Then

$$f((1-t)x + ty) \leq (1-t)f(x) + tf(y) \leq (1-t)M + tM = M,$$

implying $(1-t)x + ty \in S$. □

**Theorem 4.6.** *Let $S \subseteq \mathbb{R}^n$ be a convex set and let $f : \mathbb{R}^n \to \mathbb{R}$ be strictly convex. Suppose that $x^* \in S$ and $x^{**} \in S$ satisfy*

$$f(x^*) = \inf\{f(x) : x \in S\} = f(x^{**}).$$

*Then $x^* = x^{**}$.*

*Proof.* By the convexity of $S$, $y := \frac{1}{2}(x^* + x^{**}) \in S$. If $x^* \neq x^{**}$, then

$$f(y) < \frac{1}{2}(f(x^*) + f(x^{**})) = \inf\{f(x) : x \in S\}$$

by the strict convexity of $f$. This contradicts the minimality of $f(x^*)$. □

**Theorem 4.7.** *Let $S \subseteq \mathbb{R}^n$ be convex and let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function. Suppose $x^* \in S$ is a local minimum of $f$ on $S$, i.e., there is an $\varepsilon > 0$ such that*

$$f(x^*) \leq f(x) \quad \text{for all} \quad x \in S, \ \|x - x^*\| < \varepsilon.$$

*Then $x^*$ is a global mimimum of $f$ on $S$, i.e.,*

$$f(x^*) \leq f(x) \quad \text{for all} \quad x \in S.$$

*Proof.* Suppose $x^{**} \in S$ satisfies $f(x^{**}) < f(x^*)$. Let

$$t = \frac{\varepsilon}{2\|x^* - x^{**}\|}$$

and let

$$y = (1-t)x^* + tx^{**}.$$

Then $y \in S$ and

$$\|y - x^*\| = \frac{\varepsilon}{2} < \varepsilon.$$

Moreover,

$$f(y) \leq (1-t)f(x^{**}) + tf(x^*) < f(x^*),$$

contradicting the local minimality of $x^*$. □

## 5. Fourier-Motzkin Elimination

**Lemma 5.1.** *Let $A \in \mathbb{R}^{n\times n}$ be an invertible matrix linear isomorphism and let $P \subseteq \mathbb{R}^n$ be a polyhedron. Then $AP$ is a polyhedron.*

*Proof.* It suffices to consider the case where $P$ is a half space, which is easily dealt with:

$$A\{x \in \mathbb{R}^n : ax \leq b\} = \{y \in \mathbb{R}^n : (aA^{-1})y \leq b\}$$

□

**Theorem 5.2** (Fourier-Motzkin Elimination)**.** *Let $\pi : \mathbb{R}^n \to \mathbb{R}^{n-1}$ be projection onto the first $n-1$ coordinates:*

$$f(x_1, \ldots, x_n) = (x_1, \ldots, x_{n-1})$$

*Let $P \subseteq \mathbb{R}^n$ be a polyhedron. Then $f(P)$ is a polyhedron.*

*Proof.* Suppose

$$P = \{x \in \mathbb{R}^n : Ax \leq b\},$$

where $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^{m\times 1}$. Let $I_-$, $I_0$, and $I_+$ be the sets of indices $i$ for which $a_{in} < 0$, $a_{in} = 0$, and $a_{in} > 0$, respectively.

A point $(x_1, \ldots, x_{n-1})$ belongs to $f(P)$ if and only if

$$(\dagger) \qquad \sum_{j=1}^{n-1} a_{ij}x_j \leq b_i \quad \text{for all} \quad i \in I_0$$

and there is an $x_n \in \mathbb{R}$ such that

$$(\ddagger) \qquad \sum_{j=1}^{n} a_{ij}x_j \leq b_i \quad \text{for all} \quad i \in I_- \cup I_+.$$

System (‡) is equivalent to

$$x_n \geq \frac{b_i}{a_{in}} - \sum_{j=1}^{n-1} \frac{a_{ij}}{a_{in}}x_j \qquad \text{for all } i \in I_-,$$

$$x_n \leq \frac{b_i}{a_{in}} - \sum_{j=1}^{n-1} \frac{a_{ij}}{a_{in}}x_j \qquad \text{for all } i \in I_+.$$

Therefore, and there is an $x_n \in \mathbb{R}^n$ satisfying (‡) if and only if

$$\frac{b_p}{a_{pn}} - \sum_{j=1}^{n-1} \frac{a_{pj}}{a_{pn}}x_j \leq \frac{b_q}{a_{qn}} - \sum_{j=1}^{n-1} \frac{a_{qj}}{a_{qn}}x_j \quad \text{for all} \quad p \in I_-,\ q \in I_+.$$

$$\sum_{j=1}^{n-1}\left(\frac{a_{qj}}{a_{qn}}-\frac{a_{pj}}{a_{pn}}\right)x_j \le \frac{b_q}{a_{qn}}-\frac{b_p}{a_{pn}}, \quad p\in I_-,\ q\in I_+.$$

□

This theorem is false with "polyhedron" replaced by "closed, convex set":

$$C := \{(x,y)\in\mathbb{R}^2 : x\ge e^y\}$$

is closed and convex, but

$$\{x\in\mathbb{R} : (x,y)\in C \text{ for some } y\in\mathbb{R}\} = (0,\infty]$$

is not closed.

**Theorem 5.3.** *Let $P\subseteq\mathbb{R}^n$ be a polyhedron and let $A\in\mathbb{R}^{m\times n}$. Then*

$$Q := \left\{\begin{pmatrix}x\\y\end{pmatrix}\in\mathbb{R}^{n+m} : x\in P \text{ and } y = Ax\right\}$$

*is a polyhedron.*

*Proof.* Suppose

$$P = \{x\in\mathbb{R}^n : Bx\le c\}.$$

Then

$$Q = \left\{\begin{pmatrix}x\\y\end{pmatrix}\in\mathbb{R}^{n+m} : \begin{pmatrix}B & 0\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}\le c \text{ and } \begin{pmatrix}A & -I\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix} = 0\right\}$$

is evidently a polytope. □

**Theorem 5.4.** *Let $P\subseteq\mathbb{R}^m$ and $Q\subseteq\mathbb{R}^n$ be polyhedra. Then $P\times Q\subseteq\mathbb{R}^{m+n}$ is a polyhedron.*

*Proof.* Suppose

$$P = \{x\in\mathbb{R}^m : Ax\le b\} \quad \text{and} \quad Q = \{y\in\mathbb{R}^n : Cx\le d\}.$$

Then

$$P\times Q = \left\{\begin{pmatrix}x\\y\end{pmatrix}\in\mathbb{R}^{m+n} : \begin{pmatrix}A & 0\\ 0 & C\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}\le\begin{pmatrix}b\\d\end{pmatrix}\right\}$$

is evidently a polyhedron. □

Let $P,Q\subseteq\mathbb{R}^n$ be polytopes. Define

$$P+Q = \{x+y : x\in P,\ y\in Q\}.$$

**Theorem 5.5.** *$P+Q$ is a polytope.*

*Proof.* $P+Q$ is the image of $P\times Q$ under the linear map $(x,y)\mapsto x+y$. □

**Corollary 5.6.** *Let $P \subseteq \mathbb{R}^n$ be a polyhedron and let $A \in \mathbb{R}^{m\times n}$. Then $AP \subseteq \mathbb{R}^m$ is a polytope.*

## 6. Cones

A *conic combination* of vectors $x_1, \ldots, x_m \in \mathbb{R}^n$ is a vector of the form $\sum t_i x_i$ where $t_i \geq 0$ for all $i$. A subset $C$ of $\mathbb{R}^n$ is a *cone* if it is closed under conic combination.

The *conic hull* of a set $S \subseteq \mathbb{R}^n$, denoted $\operatorname{cone} S$, is the set of convex combinations of finite subsets of $S$.

**Lemma 6.1.** *A subset $C$ of $\mathbb{R}^n$ is a cone if and only if $C = \operatorname{cone} C$.*

**Lemma 6.2.** *Let $C$ be a cone in $\mathbb{R}^n$. Then $C$ is convex.*

*Proof.* Convex combinations are conic combinations. □

Let $\mathcal{F} \subset \mathbb{R}^{n\times 1}$ and let $x^* \in \mathcal{F}$. Let

$$C = \{c \in \mathbb{R}^{1\times n} : cx^* = \inf\{cx : x \in \mathcal{F}\}\}.$$

The $C$ is a cone.

A cone $C \subseteq \mathbb{R}^n$ is *finitely generated* if is the conic hull of a finite set of points.

**Lemma 6.3.** *Let $x_1, \ldots, x_m \in \mathbb{R}^n$ be linearly independent. Then $\operatorname{cone}\{x_1, \ldots, x_m\}$ is closed.*

*Proof.* Let $x_{m+1}, \ldots, x_n \in \mathbb{R}^n$ be such that $x_1, \ldots, x_m, x_{m+1}, \ldots, x_n$ is a basis of $\mathbb{R}^n$. Let $e_1, \ldots, e_n$ be the standard basis of $\mathbb{R}^n$ and define a linear automorphism $f : \mathbb{R}^n \to \mathbb{R}^n$ by $f(e_i) = x_i$. In particular, $f$ is closed. Let

$$C = \{t \in \mathbb{R}^n : t_1, \ldots, t_m \geq 0,\ t_{m+1}, \ldots, t_n = 0\}.$$

Then $C$ is closed and $\operatorname{cone}\{x_1, \ldots, x_m\} = f(C)$. Since $f$ is closed, so is $\operatorname{cone}\{x_1, \ldots, x_n\}$. □

**Theorem 6.4.** *Let $C$ be a finitely generated cone in $\mathbb{R}^n$. Then $C$ is closed.*

*Proof.* Let $y \in C$. Let $X$ be a subset of $\{x_1, \ldots, x_m\}$ of minimal size such that $y \in \operatorname{cone} X$. We may assume, without loss of generality, that $X = \{x_1, \ldots, x_\ell\}$ and write

$$(\dagger) \qquad y = \sum_{i=1}^{\ell} t_i x_i \quad \text{with} \quad t_i > 0.$$

Suppose $X$ is linearly dependent. Then there are scalars $s_i$, at least one of which is positive, such that

$$\sum_{i=1}^{\ell} s_i x_i = 0.$$

Without loss of generality, we may assume that

$$\frac{s_1}{t_1} \leq \cdots \leq \frac{s_\ell}{t_\ell}.$$

It follows that $s_\ell > 0$. Therefore,

$$t_\ell \frac{s_i}{s_\ell} \leq t_i, \qquad i < \ell. \tag{‡}$$

Solve for $x_\ell$ in terms of $x_1, \ldots, x_{\ell-1}$:

$$x_\ell = -\sum_{i=1}^{\ell-1} \frac{s_i}{s_\ell} x_i$$

Substituting into (†), we get

$$y = \sum_{i=1}^{\ell-1} \left( t_i - t_\ell \frac{s_i}{s_\ell} \right) x_i.$$

By (‡), this is a conic combination, contradicting the minimality of $\ell$. Therefore, $X$ must be linearly independent as claimed.

Let $\mathcal{X}$ be the set of linearly independent subsets of $\{x_1, \ldots, x_m\}$. It follows from the above that

$$\operatorname{cone}\{x_1, \ldots, x_m\} = \bigcup_{X \in \mathcal{X}} \operatorname{cone} X.$$

By the preceding lemma, each set $\operatorname{cone} X$ is closed. Since $\mathcal{X}$ is finite, $\operatorname{cone}\{x_1, \ldots, x_m\}$ is closed, too. $\square$

## 7. Remedial linear algebra

**Theorem 7.1.** *Let $W$ be a subspace of $\mathbb{R}^n$. Then $W$ has an orthonormal basis.*

**Theorem 7.2.** *Let $W$ be a subspace of $\mathbb{R}^n$ and let $v \in \mathbb{R}^n$. Then there are unique vectors $v_W \in W$ and $v_{W^\perp} \in W^\perp$ such that $v = v_W + v_{W^\perp}$.*

*Proof.* Let $u_1, \ldots, u_r$ be an orthonormal basis of $W$ and set

$$v_W = \sum_{j=1}^{r} (u_j^T v) u_j, \quad v_{W^\perp} = v - v_W$$

Then $v = v_W + v_{W^\perp}$ and it's easy to check that $v_{W^\perp} \in W^\perp$.

To prove uniqueness, suppose $v_W, v'_W \in W$ and $v_{W^\perp}, v'_{W^\perp} \in W^\perp$ satisfy

$$v_W + v_{W^\perp} = v'_W + v'_{W^\perp}.$$

Then

$$v_W - v'_W = v'_{W^\perp} - v_{W^\perp},$$

putting $v_W - v_{W'}$ in $W \cap W^\perp = \{0\}$. Therefore, $v_W = v'_W$, and it follows that $v_{W^\perp} = v'_{W^\perp}$, proving uniqueness. ☐

**Corollary 7.3.** *Let $W$ be a subspace of $\mathbb{R}^n$. Then:*

(1) $\dim W + \dim W^\perp = n$,
(2) $(W^\perp)^\perp = W$.

**Corollary 7.4.**

**Theorem 7.5** (Rank-Nullity Theorem)**.** *Let $A \in \mathbb{R}^{m\times n}$. Then*

$$\dim C(A) + \dim N(A) = n.$$

*Proof.* Let $U \in \mathbb{R}^{n\times q}$ be a basis matrix of $N(A)$ and let $W \in \mathbb{R}^{m\times r}$ be a basis matrix of $C(A)$. Then, by the definition of the column space, there is a matrix $V \in \mathbb{R}^{n\times r}$ such that $W = AV$. The linear independence of the coluns of $V$ follows immediately from that of the columns of $W$.

I claim that the $\begin{bmatrix} U & V \end{bmatrix} \in \mathbb{R}^{m\times(q+r)}$ is a basis matrix of $\mathbb{R}^n$. To see that its columns span $\mathbb{R}^n$, let $x \in \mathbb{R}^n$. Then $Ax \in C(A)$ so there is a $z \in \mathbb{R}^r$ such that $Ax = Wz$. It follows that $x - Vy \in N(A)$ as

$$A(x - Vy) = Ax - AVy = Ax - Wy = Ax - Ax = 0.$$

Therefore, $x - Vz = Uy$ for some $y \in \mathbb{R}^q$. Rearranging, we have

$$x = Uy + Vz = \begin{bmatrix} U & V \end{bmatrix} \begin{bmatrix} y \\ z \end{bmatrix},$$

putting $x$ in the column space of $\begin{bmatrix} U & V \end{bmatrix}$. Thus, the columns of $\begin{bmatrix} U & V \end{bmatrix}$ span $\mathbb{R}^n$.

To see that the columns of $\begin{bmatrix} U & V \end{bmatrix}$ are linearly independent, suppose

$$Uy + Vz = 0.$$

Then

$$0 = A0 = A(Uy + Vz) = AUy + AVz = 0 + Wz = Wz.$$

Since the columns of $W$ are linearly independent, $z = 0$. Consequently, $Uy = 0$ and, by the linear independence of the columns of $U$, $y = 0$. Thus, the columns of $\begin{bmatrix} U & V \end{bmatrix}$ are linearly independent. ☐

**Corollary 7.6.** *Let $A \in \mathbb{R}^{m\times n}$. Then*

$$C(A) = N(A^T)^\perp.$$

*Proof.* The inclusion $C(A) \subseteq N(A^T)^\perp$ is obvious: if $b = Ax \in C(A)$ and $y \in N(A^T)$, then

$$y^T b = y^T(Ax) = (y^T A)x = 0x = 0,$$

implying $b \in N(A^T)^\perp$.

Conversely, suppose $y \in C(A)^\perp$. Then

$$(A^T y)^T x = y^T(Ax) = 0$$

for all $x \in \mathbb{R}^n$, implying $A^T y = 0$. Therefore, $y \in N(A^T)$ and we've shown that $C(A)^\perp \subseteq N(A^T)$. Taking orthogonal complements, we get

$$N(A^T)^\perp \subseteq (C(A)^\perp)^\perp = C(A). \qquad \square$$

**Corollary 7.7.** *Let $A \in \mathbb{R}^{m\times n}$. Then*

$$\dim C(A) = \dim C(A^T).$$

*Proof.*

$$\begin{aligned}\dim C(A) &= \dim N(A^T)^\perp \\ &= m - \dim N(A^T) \\ &= \dim C(A^T)\end{aligned} \qquad \square$$

**Theorem 7.8.** *Let $A \in \mathbb{R}^{m\times n}$ and let $b \in \mathbb{R}^m$. Then exactly one of the following conclusions hold:*

(1) *$Ax = b$ has a solution.*
(2) *There is a $y \in \mathbb{R}^m$ such that $y^T A = 0$ and $y^T b \neq 0$.*

*Proof.* Suppose (1) doesn't hold. Then $b \notin C(A)$. As $C(A) = N(A^T)^\perp$, there is a $y \in N(A^T)$ such that $y^T b \neq 0$. Since $y \in N(A^T)$, we have $y^T A = 0$. Therefore, (2) holds. $\square$

## 8. Convex sets

**Theorem 8.1.** *Let $C \subseteq \mathbb{R}^n$ be a nonempty, closed, convex set and let $b \in \mathbb{R}^n \setminus C$. Then there is a unique $c^* \in C$ such that*

$$\|b - c^*\| = \inf_{c\in C} \|b - c\|.$$

*Proof.* Let $c_0 \in C$ and let $r = \|b - c_0\|$. Then

$$C_r := C \cap \{y \in \mathbb{R}^n : \|b - y\| \leq r\}$$

is a nonempty, compact set not containing $b$. Therefore, there is a $c^* \in C_r$ such that

$$\|b - c^*\| = \inf_{c \in C_r} \|b - c\|.$$

If $c \in C \setminus C_r$, then

$$\|b - c^*\| \leq r < \|b - c\|.$$

Therefore,

$$\|b - c^*\| = \inf_{c \in C} \|b - c\|.$$

It remains to prove the unicity of $c^*$. Suppose that $c^{**} \in C$ satisfies

$$\|b - c^{**}\| = \|b - c^*\|.$$

Then

$$c^{***} := \frac{1}{2}(c^* + c^{**}) \in C$$

by convebity. By the Pythagorean Theorem,

$$\|b - c^*\|^2 = \|b - c^{**}\|^2 + \|c^{**} - c^{***}\|^2$$

Since $\|b - c^{**}\| = \|b - c^*\|$, it follows that $c^{***} = c^{**}$ and, by definition of $c^{***}$, that $c^{**} = c^*$. □

**Theorem 8.2** (Separating Hyperplane Theorem)**.** *Let $C \subset \mathbb{R}^n$ be a nonempty, closed, convex set and let $b \in \mathbb{R}^n \setminus C$. Then there is a vector $y \in \mathbb{R}^n$ such that*

$$y^T c < y^T b$$

*for all $c \in C$.*

**Theorem 8.3** (Farkas's Lemma)**.** *Let $A \in \mathbb{R}^{m \times n}$, let $b \in \mathbb{R}^m$, and suppose that $Ax = b$ has no solution with $x \geq 0$. Then there is a $y \in \mathbb{R}^m$ such that $y^T A \geq 0$ and $y^T b < 0$.*

*Proof.* Let

$$C = \{Ax : x \geq 0\}.$$

Then $C$ is a nonempty, closed (why?), convex set not containing $b$. Therefore, there is a $y \in \mathbb{R}^m$ such that $y^T b < y^T c$ for all $c \in C$. Since $0 \in C$, we have $y^T b < 0$. For every $c \in C$ and every $t > 0$ we have $tc \in C$, implying

$$\frac{1}{t} y^T b < \frac{1}{t} y^T (tc) = y^T c.$$

Letting $t \to \infty$ gives $y^T c \geq 0$. Taking $c$ to be the columns of $A$ gives $y^T A \geq 0$. □

## 9. Duality

Let

$$A = \begin{bmatrix} a_1 \\ \vdots \\ a_m \end{bmatrix} \in \mathbb{R}^{m\times n},$$

let $b \in \mathbb{R}^{m\times 1}$, and let $c \in \mathbb{R}^{1\times n}$. Define

$$\mathcal{F} = \{x \in \mathbb{R}^{n\times 1} : Ax \geq b\},$$
$$\mathcal{G} = \{y \in \mathbb{R}^{1\times m} : y \geq 0 \text{ and } yA = c\}.$$

**Theorem 9.1** (Weak Duality)**.** *Let $x \in \mathcal{F}$ and let $y \in \mathcal{G}$. Then*

$$yb \leq cx.$$

*Proof.* We have:

$$\begin{aligned} yb &\leq yAx && (\text{as } b \leq Ax \text{ and } y \geq 0) \\ &= cx && (\text{as } yA = c) \end{aligned}$$

□

**Theorem 9.2** (Strong Duality)**.** *Suppose there is a point $x^* \in \mathcal{F}$ such that*

$$cx^* = \inf_{x\in\mathcal{F}} cx.$$

*Then $c$ belongs to the cone spanned by the constraints $a_i$ active at $x^*$:*

$$c \in \operatorname{cone}\{a_i : a_i x^* = b_i\}$$

*In other words, there is a point $y^* \in \mathcal{G}$ such that*

$$y_i^*(b_i - a_i x^*) = 0 \quad \text{for all } i.$$

*Moreover,*

$$cx^* = y^*b = \sup\{yb : y \in \mathcal{G}\}.$$

*Proof.* Let $I = \{i : a_x^* = b_i\}$. By Farkas's Lemma, to show that $c \in \operatorname{cone}\{a_i : i \in I\}$, it suffces to show that

$$a_i v \geq 0 \text{ for all } i \in I \quad \text{implies} \quad cv \geq 0.$$

So suppose $a_i v \geq 0$ for all $i \in I$. If $i \in I$ and $\varepsilon > 0$, then

$$a_i(x^* + \varepsilon v) = a_i x^* + \varepsilon a_i v \geq b_i$$

while if $i \notin I$ and $\varepsilon > 0$ is sufficiently small, then

$$a_i(x^* + \varepsilon v) = a_i x^* + \varepsilon a_i v > b_i.$$

Therefore, for sufficiently small $\varepsilon > 0$,

$$x^* + \varepsilon v \in \mathcal{F}$$

and, by the optimaliy of $x^*$,

$$cx^* \le c(x^* + \varepsilon v).$$

It follows that $cv \ge 0$.

Since $c \in \operatorname{cone}\{a_i : i \in I\}$, there is a vector $y^* \ge 0$ such that $y_i^* = 0$ for all $i \notin I$ and

$$c = yA = \sum_{i \in I} y_i a_i.$$

In particular, $y^* \in \mathcal{G}$. Also,

$$cx^* = yAx^* = \sum_{i \in I} y_i^* a_i x^* = \sum_{i \in I} y_i^* b_i = \sum_i y_i^* b_i = y^* b.$$

The identity

$$y^* b = \sup\{yb : y \in \mathcal{G}\}$$

now follows from weak duality. □

## 10. Polyhedra

A *polyhedron* is the solution set of a system of finitely many linear inequalities of the form $a_i x \le b_i$.

A *polytope* is the convex hull of a finite set of points.

**Theorem 10.1.** *A polytope is a polyhedron.*

*Proof.* Suppose $P$ is the convex hull of $x_1, \ldots, x_n \in \mathbb{R}^m$. Define $f : \mathbb{R}^n \to \mathbb{R}^m$ by

$$f(t_1, \ldots, t_n) = \sum_i t_i x_i.$$

Then $f$ is a linear map and $P = f(\Delta^{n-1})$. Therefore, $P$ is a polyhedron. □

Let $C \subset \mathbb{R}^n$ be a closed and convex.

A *face* of $C$ is a set of the form

$$F = \{x \in C : \ell(x) = b\},$$

where $\ell$ is a linear functional on $\mathbb{R}^n$, $b \in \mathbb{R}$, and $\ell(x) \le b$ for all $x \in C$. Note that faces of $C$ are, themselvess, closed and convex.

If $F \ne \emptyset$ and $F \ne C$, then $F$ is called a *proper face* of $C$.

A point $v \in C$ is called an *extreme point* of $C$ If $x, y \in C$ and $v = \frac{1}{2}(x+y)$ imply $v = x = y$.

**Theorem 10.2.** *Let $F = \{x \in C : \ell(x) \leq b\}$ be a face of $C$ and let $v$ be an extreme point of $F$. Then $v$ is an extreme point of $C$.*

*Proof.* Suppose $v = \frac{1}{2}(x+y)$ with $x, y \in C$. As $\ell(v) = b$,

$$b = \ell(v) = \ell(\tfrac{1}{2}(x+y)) = \tfrac{1}{2}(\ell(x) + \ell(y)).$$

As $\ell(x) \leq b$ and $\ell(y) \leq b$, we must have

$$b = \ell(x) = \ell(y),$$

implying $x, y \in F$. Since $v$ is an extreme point of $F$, $v = x = y$. It follows that $v$ is an extreme point of $C$. $\square$

10.1. **Extreme points of polyhedra.** Let

$$A = \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} \in \mathbb{R}^{m \times n}$$

and let $P = \{x \in \mathbb{R}^{n \times 1} : Ax \leq b\}$.

For $x \in P$, let $I_x$ be the *active set* of $x$:

$$I_x = \{i : a_i x = b_i\}$$

**Theorem 10.3.** *Let $v \in P$. Then $v$ is an extreme point of $P$ if and only if*

$$\operatorname{span}\{a_i : i \in I_x\} = \mathbb{R}^{1 \times n}.$$

## 11. Recession Cones

For subset $X$ of $\mathbb{R}^n$ and a point $x \in X$, define the *recession cone to $X$ at $x$* by

$$R_{X,x} = \{v \in \mathbb{R}^n : x + tv \in X \text{ for all } t \geq 0\}.$$

**Lemma 11.1.** *Let $C$ be convex and let $x \in C$. Then $R_{C,x}$ is closed under addition.*

*Proof.* Let $v, w \in R_{C,x}$. Then $x + 2v \in C$, $x + 2w \in C$, and

$$x + (v+w) = \frac{1}{2}(x+2v) + \frac{1}{2}(x+2w) \in [x+2v, x+2w] \subseteq C. \quad \square$$

**Theorem 11.2.** *Let $C$ be a nonempty, closed, convex subset of $\mathbb{R}^n$. Then $R_{C,x} = R_{C,y}$ for all $x, y \in C$.*

*Proof.* Let $x, y \in C$. If suffices to show that if $v \in R_{C,x}$ then $y+v \in C$. Let $v \in R_{C,x}$ and define

$$t_n = \frac{\|v\|}{\|x+nv-y\|}, \qquad z_n = (1-t_n)y + t_n(x+nv).$$

If $n$ is suffciently large, then $0 \leq t_n \leq 1$ and $z_n \in [y, x+nv] \subseteq C$ by the convexity of $C$. It's routine to check that $z_n \to y+v$ as $n \to \infty$. Since $C$ is closed, $y+v \in C$. $\square$

The preceding theorem justifies abbreviating $R_{C,x}$ to $R_C$ for $C$ closed and convex.

## 12. Basic solutions

Let $A \in \mathbb{R}^{m\times n}$ and let $b \in \mathbb{R}^m$. A vector $x^* \in \mathbb{R}^n$ is a *basic solution of* $Ax = b$ if $Ax^* = b$ and there are sets of indices

$$M \subseteq \{1, \dots, m\} \quad \text{and} \quad N \subseteq \{1, \dots, n\}$$

such that $x_i^* = 0$ for all $i \in N$ and

$$\operatorname{rank} \begin{pmatrix} A[M, :] \\ I[N, :] \end{pmatrix} = n.$$

Note that $Ax^* = b$ and $x_i^* = 0$ for all $i \in N$ imply

$$\begin{pmatrix} A[M, :] \\ I[N, :] \end{pmatrix} x^* = \begin{pmatrix} b[M] \\ 0 \end{pmatrix}.$$

**Lemma 12.1.** *Let $x^* \in \mathbb{R}^n$ and let $B = \{i : x_i^* \neq 0\}$. Then $x^*$ is a basic solution of $Ax = b$ if and only if $A[:, B]z = b$ has unique solution $z = x^*[B]$.*

*Proof.* First, observe that if $\{1, \dots, n\} = B \sqcup N$ then

$$\left\{x \in \mathbb{R}^n : \begin{pmatrix} A \\ I[N, :] \end{pmatrix} x = \begin{pmatrix} b \\ 0 \end{pmatrix}\right\} = \{x \in \mathbb{R}^n : Ax = b,\ x[N] = 0\}$$

and the mapping

$$\{x \in \mathbb{R}^n : Ax = b,\ x[N] = 0\} \longrightarrow \{z \in \mathbb{R}^k : A[:, B]z = b\}$$

defined by $x \mapsto x[B]$ is a bijection.

Now let $x^* \in \mathbb{R}^n$ and let $B = \{i : x_i^* \neq 0\}$. Suppose $z = x^*[B]$ is the unique solution of $A[:, B]z = b$. Then

$$\begin{pmatrix} A \\ I[N, :] \end{pmatrix} x = \begin{pmatrix} b \\ 0 \end{pmatrix}.$$

has unique solution $x = x^*$. In particular,

$$\operatorname{rank}\begin{pmatrix} A \\ I[N,:] \end{pmatrix} = n.$$

Therefore, $x^*$ is a basic solution of $Ax = b$.

Conversely, suppose that $x^*$ is a basic solution of $Ax = b$. Then $Ax^* = b$ and there are sets of indices

$$M \subseteq \{1,\ldots,m\} \quad \text{and} \quad N \subseteq \{1,\ldots,n\}$$

such that $x_i^* = 0$ for all $i \in N$ and

$$\operatorname{rank}\begin{pmatrix} A[M,:] \\ I[N,:] \end{pmatrix} = n.$$

We have $B \subseteq \{1,\ldots,n\} \setminus N$.

□

**Theorem 12.2.** *Suppose* $\operatorname{rank} A = m$. *Then $x^*$ is a basic solution of $Ax = b$ if and only if $Ax^* = b$ and there is a set $B \subseteq \{1,\ldots,n\}$ of size $m$ such that* $\operatorname{rank} A[:,B] = m$ *and $x_i^* = 0$ for all $i \notin B$.*

*Proof.* Suppose that $Ax^* = b$ and that there is a set $B \subseteq \{1,\ldots,n\}$ of size $m$ such that $\operatorname{rank} A[:,B] = m$ and $x_i^* = 0$ for all $i \notin B$. Take $M = \{1,\ldots,m\}$ and $N = \{1,\ldots,n\} \setminus B$. By design, $x_i^* = 0$ for all $i \in N$. To show that

$$\operatorname{rank}\begin{pmatrix} A[M,:] \\ I[N,:] \end{pmatrix} = n,$$

is suffices to show that $x^*$ is the unique solution of

$$(\dagger) \qquad \begin{pmatrix} A \\ I[N,:] \end{pmatrix} x = \begin{pmatrix} b \\ 0 \end{pmatrix}.$$

If $x$ is any solution of this equation, then $Ax = b$ and $x[N] = 0$. Therefore,

$$b = Ax = A[:,B]x[B].$$

Since $\operatorname{rank} A[:,B]$ is assumed equal to $m$, the $m \times m$ matrix $A[:,B]$ is invertible and

$$x[B] = A[:,B]^{-1}b.$$

Thus, $(\dagger)$ has a unique solution. Therefore, $x^*$ is a basic solution of $Ax = b$.

Conversely, suppose that $x^*$ is a basic solution of $Ax = b$. Then $Ax^* = b$ and there are sets of indices

$$M \subseteq \{1,\ldots,m\} \quad \text{and} \quad N \subseteq \{1,\ldots,n\}$$

such that $x_i^* = 0$ for all $i \in N$ and

$$\operatorname{rank}\begin{pmatrix} A[M,:] \\ I[N,:] \end{pmatrix} = n.$$

We may assume, without loss of generality, that $|M| + |N| = n$. Let $B = \{1, \ldots, n\} \setminus N$. Then

$$A[:,B]z = b$$

has a unique solution: $z = x^*[B]$ is clearly such a solution, and if $z$ is any such solution, then the vector $x \in \mathbb{R}^n$ characterized by $x[B] = z$ and $x[N] = 0$ a solution of $Ax = b$. Since $x^*$ is the unique solution of $Ax = b$, it follows that $z = x^*[B]$.

$\square$