Polyhedral Theory

Andrew Eberhard
# Contents

1 Summary of basic concepts from Convexity ........................................ 1
   1.1 Convexity Preserving Operations .............................................. 4
       1.1.1 Intersection ..................................................................... 4
       1.1.2 Cross Products .................................................................. 5
       1.1.3 Affine Images and Preimages .......................................... 5

2 Supports, Separation and Farka’s Lemma ............................................. 9
   2.1 Supports and Separation ............................................................. 9
       2.1.1 Theorems of the Alternative - Farkas’ Lemma ................... 13

3 Polyhedral Theory ........................................................................... 17
   3.1 Dimensionality ......................................................................... 17
   3.2 Polyhedra and Dimension ......................................................... 20
   3.3 Faces and Facets and Dimensions ............................................. 22
   3.4 The Finite Basis Theorems (unbounded polyhedral sets) ........... 28
Introduction

These note were produced for the MINLP reading group that ran during 2013 at Melbourne University. They are an attempt by the author to reconcile concepts he as familiar within convex analysis with those was asked to present from some integer programming literature.
Chapter 1

Summary of basic concepts from Convexity

There are two ways of defining polyhedra and polytopes. Some texts begin with the definition of a polytope as a convex combination of a finite number of points. Other begin with the definition of polyhera via the intersection of finitely many half spaces. In every approach some familiarity with basic notions from convex analysis are of a help in order to unify these approaches.

Roughly speaking convexity refers to a geometric concept of shape of an object. A set in a vector space is convex if it contains all lines that can be drawn between two points within this set. This corresponds to a rotund shape.

**Definition 1** Let $C \subseteq X$, where $X$ is a vector space. A set $C$ is convex if for all $x_1, x_2 \in C$ and $\lambda \in [0, 1]$ we have

$$\lambda x_1 + (1 - \lambda) x_2 \in C.$$ 

**Remark 2** If $C$ is convex we have for any finite subset $\{x_i \mid i = 1, \ldots, m\} \subseteq C$ that
\[ \sum_i \lambda_i x_i \in C \text{ for any } \lambda_i \geq 0 \text{ and } \sum_i \lambda_i = 1. \] This may be done by induction using 1. above. Suppose \( m = 3 \) then we have
\[ \lambda_1 x_1 + \lambda_2 x_2 + \lambda_3 x_3 = \left( \frac{\lambda_1 + \lambda_2}{\lambda_1 + \lambda_2 + \lambda_3} \right) \left( \frac{\lambda_1}{\lambda_1 + \lambda_2} \right) x_1 + \left( \frac{\lambda_2}{\lambda_1 + \lambda_2} \right) x_2 + \left( \frac{\lambda_3}{\lambda_1 + \lambda_2 + \lambda_3} \right) x_3 \in C \]
because \( \lambda_1 + \lambda_2 + \lambda_3 = 1 \). It may be shown that if \( C = \{ x_1, \ldots, x_m \} \subseteq \mathbb{R}^n \) for \( m > n \) then we only need to use sums from \( i = 1, \ldots, m = n + 1 \) to represent a point \( x \in C \) (this is called Carathéodory’s theorem).

**Definition 3** The convex hull of a subset \( C \subseteq \mathbb{R}^n \) is the smallest convex set containing \( C \) and is denoted by \( \text{co} \ C \).

**Proposition 4** If \( C \subseteq \mathbb{R}^n \) then
\[
\text{conv} \ C = \left\{ \sum_i \lambda_i x_i \mid \sum_i \lambda_i = 1, \lambda_i \geq 0 \text{ and } x_i \in C \right\}.
\] (1.1)

**Proof.** Let \( \hat{C} \) denote the right hand side (1.1). Clearly \( \hat{C} \) is convex and \( C \subseteq \hat{C} \) (just take \( \lambda_1 = 1 \) and \( x_1 = x \in C \)). As \( \text{conv} \ C \) is the smallest convex set containing \( C \) and \( \hat{C} \) is an example of such a set we have \( \text{conv} \ C \subseteq \hat{C} \). Now suppose there exists \( y \in \hat{C} \) with \( y \notin \text{conv} \ C \). Then \( y = \sum_i \lambda_i x_i \) with \( x_i \in C \subseteq \text{conv} \ C \). Then by the convexity of \( \text{co} \ C \) we must have \( y = \sum_i \lambda_i x_i \in \text{conv} \ C \), a contradiction. 

We may now define what is meant by a convex function on a real variable.

**Definition 5** A convex cone is a convex set \( C \) in \( X \) with the additional property that \( x \in C \) whenever \( x \in C \) and \( \lambda \geq 0 \).

Given any convex set \( C \) we may generate the smallest convex cone containing \( C \) as
\[
\text{cone} \ C = \{ \lambda x \mid x \in C \text{ and } \lambda \geq 0 \}.
\]

If \( C \) is not convex then using the characterization of the smallest convex set containing \( C \) in (1.1) one may see that the smallest convex cone containing \( C \) is given by
\[
\text{cone} \ C = \left\{ \sum_i \lambda_i x_i \mid \lambda_i \geq 0 \text{ and } x_i \in C \right\}.
\]

**Example 6** The simplest example of a convex cone is the positive orthant in \( \mathbb{R}^n \)
\[
\mathbb{R}_+^n := \{ x := (x_1, \ldots, x_n) \in \mathbb{R}^n \mid x_i \geq 0 \text{ for all } i = 1, \ldots, n \}.
\]
(Check that this is indeed a convex cone).

**Example 7** A half-space in \( X \) is given by
\[
H_{y,\alpha} := \{ x \in X \mid f(x) \geq \alpha \}
\]
where \( f(x) \) is a linear mapping on the space \( X \). When \( X = \mathbb{R}^n \) then one could take \( f(x) = x \cdot y \) (the usual dot product of vectors). Show \( H_{y,\alpha} \) is a convex set.
**Solution:** Take \( x, z \in H_{y, \alpha} \) and \( \lambda \in [0, 1] \) then

\[
f(\lambda x + (1 - \lambda) z) = \lambda f(x) + (1 - \lambda) f(z) \geq \lambda \alpha + (1 - \lambda) \alpha = \alpha
\]

and so \( \lambda x + (1 - \lambda) z \in H_{y, \alpha} \). Having verified convexity we now show that \( H_{y, 0} \) is a cone. This follows immediately since \( \lambda \geq 0 \) and \( x \in H_{y, 0} \) implies

\[
f(\lambda x) = f(\lambda x + (1 - \lambda)0) = \lambda f(x) + (1 - \lambda) f(0) = \lambda f(x) \geq \lambda 0 = 0
\]

and so \( \lambda x \in H_{y, 0} \) for all \( \lambda \geq 0 \).

**Example 8** We denote the vector space of all symmetric \( n \times n \) matrices by

\[
S(n) := \{ X \in \mathbb{R}^{n \times n} \mid X = X^T \}
\]

where the superscript \( T \) denotes the transpose operation. A natural cone in the space is the cone of positive semidefinite matrices given by

\[
P(n) := \{ X \in S(n) \mid x^T X x \geq 0 \text{ for all } x \in \mathbb{R}^n \}
\]

Even in two dimensions this set is complex. For

\[
X = \begin{pmatrix} x & y \\ y & z \end{pmatrix} \in P(n)
\]

we need \( x \geq 0, z \geq 0 \) and \( \det X = xz - y^2 \geq 0 \) or \( xz \geq y^2 \).

![The boundary of the positive semidefinite cone \( P(2) \) plotted in \( \mathbb{R}^3 \).](image)

**Remark 9** The symmetric matrices: Denote by \( S(n) \) the \( n \times n \) real, symmetric matrices and by \( P(n) \) the set of all positive semi–definite, real symmetric matrices i.e.

\[
P(n) = \{ Q \in S(n) \mid x^T Q x \geq 0 \text{ for all } x \in \mathbb{R}^n \}.
\]
CHAPTER 1. SUMMARY OF BASIC CONCEPTS FROM CONVEXITY

One can show that all \( Q \) in \( S(n) \) are finitely decomposable and so \( S(n) \) is a finite dimensional space but its dimension is not \( n \) but actually \( \frac{1}{2}n(n+1) \). To make it a linear space define the trace of a matrix to be the sum of the diagonal elements i.e. 
\[
\text{tr} Q = \sum_{j=1}^n q_{jj}
\]
for \( Q = [q_{ij}] \in S(n) \). We now show \( S(n) \) is a vector space with an inner product \( \langle A, B \rangle = \text{tr} B^T A (= \text{tr} BA) \). Define for \( A, B \in S(n) \) the sum \( A + B = \{a_{ij} + b_{ij}\} \) (i.e. component wise sum). Multiplication by a scalar is similarly component wise i.e. \( \alpha A = \{\alpha a_{ij}\} \). Clearly this defines a vector space (essentially isomorphic to a subspace of \( \mathbb{R}^n \)). Then given a fixed \( z z^T \in \mathcal{P}(n) \) then we may form a hold space via
\[
H_{z, \alpha} := \{X \in \mathcal{P}(n) \mid \langle X, zz^T \rangle = z^T X z \geq \alpha\}.
\]

1.1 Convexity Preserving Operations

One can construct complicated convex sets from simpler ones by using various convexity preserving operations.

1.1.1 Intersection.

**Lemma 10** Convexity is preserved under arbitrary intersection. That is if \( \{C_\alpha\}_{\alpha \in \Lambda} \) is a family of convex sets in a vector space \( X \) then so is
\[
C := \bigcap_{\alpha \in \Lambda} C_\alpha.
\]

If all \( C_\alpha \) are cones then so is \( C \).

**Proof.** Let \( x_1, x_2 \in C \) and \( \lambda \in [0, 1] \) then \( x_1, x_2 \in C \) for all \( \alpha \in \Lambda \) and as \( C_\alpha \) is convex we have \( \lambda x_1 + (1 - \lambda) x_2 \in C_\alpha \) for all \( \alpha \in \Lambda \). But then
\[
\lambda x_1 + (1 - \lambda) x_2 \in \bigcap_{\alpha \in \Lambda} C_\alpha = C.
\]

Now suppose all \( C_\alpha \) are cones, \( x \in C \) and \( \lambda \geq 0 \). As \( x \in C_\alpha \) we have \( \lambda x \in C_\alpha \) for all \( \alpha \in \Lambda \) and so \( \lambda x \in \bigcap_{\alpha \in \Lambda} C_\alpha = C \). \( \blacksquare \)

**Example 11** The positive semidefinite cone \( \mathcal{P}(n) \) in \( S(n) \) can be written as
\[
\mathcal{P}(n) = \bigcap_{x \in \mathbb{R}^n} \{X \in S(n) \mid x^T X x \geq 0\}.
\]

The simpler set \( \{X \in S(n) \mid x^T X x \geq 0\} \) is clearly convex since it is actually an half-space for each fixed \( x \) as \( X \mapsto x^T X x \) is a linear mapping i.e. for all \( X_1, X_2 \in S(n) \) and \( \alpha_1, \alpha_2 \in \mathbb{R} \) we have
\[
x^T (\alpha_1 X_1 + \alpha_2 X_2) x = \alpha_1 (x^T X_1 x) + \alpha_2 (x^T X_2 x).
\]

As \( \mathcal{P}(n) \) is an intersection of convex cones and so is also a convex cone.
1.1. CONVEXITY PRESERVING OPERATIONS

This kind of argument may also be used to show that for an arbitrary convex set $C$ we have

$$C = \bigcap \{H \mid C \subseteq H, \text{ and } H \text{ is a half-space}\}.$$  \hfill (1.2)

Indeed the right hand side of (1.2) can be shown to be just $\text{conv } C$ even if $C$ is not convex.

1.1.2 Cross Products

Lemma 12 Given a family of convex sets $\{C_{\alpha}\}_{\alpha \in \Lambda}$ the cross product

$$C = \bigotimes_{\alpha \in \Lambda} C_{\alpha} := \{(c_{\alpha})_{\alpha \in \Lambda} \mid c_{\alpha} \in C_{\alpha}\}$$

is convex. If all $C_{\alpha}$ are cones then so is $C$.

Proof. If $(c_{\alpha}^{1})_{\alpha \in \Lambda}$ and $(c_{\alpha}^{2})_{\alpha \in \Lambda}$ are in $C$ then for any $\lambda \in [0,1]$ and each $\alpha \in \Lambda$ we have

$$\lambda c_{\alpha}^{1} + (1 - \lambda)c_{\alpha}^{2} \in C_{\alpha}.$$

That is

$$\left(\lambda c_{\alpha}^{1} + (1 - \lambda)c_{\alpha}^{2}\right)_{\alpha \in \Lambda} \in \bigotimes_{\alpha \in \Lambda} C_{\alpha} = C.$$

If all $C_{\alpha}$ are cones $(c_{\alpha})_{\alpha \in \Lambda} \in C$ and $\lambda \geq 0$ we have $\lambda(c_{\alpha})_{\alpha \in \Lambda} = (\lambda c_{\alpha})_{\alpha \in \Lambda} \in C$. ■

When $\Lambda$ is finite index set $\Lambda = \{\alpha_{1}, \ldots, \alpha_{k}\}$ we denote

$$\bigotimes_{\alpha \in \Lambda} C_{\alpha} = C_{\alpha_{1}} \times C_{\alpha_{2}} \times \cdots \times C_{\alpha_{k}}.$$

1.1.3 Affine Images and Preimages

Let $h : X \to Y$ be an linear function i.e. for all $\alpha_{1}, \alpha_{2} \in \mathbb{R}$ and $x_{1}, x_{2} \in X$ we have

$$h(\alpha_{1}x_{1} + \alpha_{2}x_{2}) = \alpha_{1}h(x_{1}) + \alpha_{2}h(x_{2}).$$

We say $f : X \to Y$ is an affine mapping if it is a translation of a linear mapping i.e.

$$f(x) = h(x) + \alpha \quad \text{for some } \alpha \in Y.$$

Then the image of a convex set $C$ in $X$ is a convex set $f(C)$ in $Y$ where

$$f(C) := \{f(x) \mid x \in C\}.$$

Also the preimage of a convex set $S$ in $Y$ is a convex set $f^{-1}(S)$ in $X$ where

$$f^{-1}(S) := \{x \in X \mid f(x) \in S\}.$$

Exercise 13 Prove that $f(C)$ and $f^{-1}(S)$ are convex for convex sets $C$ in $X$ and $S$ in $Y$. (Discuss in class)
The following operations can all be shown to preserve convexity.

- **Translation**: \( y_1 + (1 - \lambda) y_2 = \lambda y_1 + (1 - \lambda) y_2 \)
- **Scalar multiplication**: \( f(\lambda x) = \lambda f(x) \)
- **The polyhedron** (as described as an intersection of a finite number of half spaces)
- **The sum of two sets**

**Example 14** The following operations can all be shown to preserve convexity.

- **Scalar multiplication**: \( \alpha C = \{ \alpha x \mid x \in C \} \), for \( \alpha \in \mathbb{R} \)

  *Why?* This is the image of \( C \) under the affine mapping \( f(x) = \alpha x \).

- **Translation**: \( C + y := \{ x + y \mid x \in C \} \), where \( y \in \mathbb{R}^n \)

  *Why?* This is the image of \( C \) under the affine transformation \( f(x) = x + y \).

- **The sum of two sets** \( S_1 \) and \( S_2 \) in \( \mathbb{R}^n \) given by

  \[ S_1 + S_2 := \{ x_1 + x_2 \mid x_1 \in S_1 \text{ and } x_2 \in S_2 \} \]

  *Why?* This set may be viewed as the image of the convex set \( S_1 \times S_2 := \{ (x_1, x_2) \mid x_1 \in S_1 \text{ and } x_2 \in S_2 \} \) under the affine mapping

  \[ f(x_1, x_2) = x_1 + x_2 \]

- **The polyhedron** (as described as an intersection of a finite number of half spaces)

  \[ C := \{ x \mid Ax \leq b, Cx = d \} \]

  where \( A \) is an \( n \times m \) matrix \( b \in \mathbb{R}^m \) and \( C \) is an \( n \times l \) matrix and \( d \in \mathbb{R}^l \).

  *Why?* This may be shown to convex by noting that it is representable as \( f^{-1}(S) \) where \( S = \mathbb{R}^m_+ \times \{ 0 \} \) is a convex set (the cross product of two convex sets) and

  \[ f(x) := (b - Ax, d - Cx) = (b, d) - (A, C)x \]

is an affine mapping. That is, \( x \in f^{-1}(\mathbb{R}^m_+ \times \{ 0 \}) \) or \( f(x) \in \mathbb{R}^m_+ \times \{ 0 \} \) which means \( b - Ax \in \mathbb{R}^m_+ \) and \( d - Cx = 0 \).

**Proof.** Consider first proving the convexity of \( f(C) \). We take \( y_1, y_2 \in f(C) \) and \( \lambda \in [0, 1] \). Then we note that \( y_i = f(x_i) \) for some \( x_i \in C \) so

\[
\lambda y_1 + (1 - \lambda) y_2 = \lambda f(x_1) + (1 - \lambda) f(x_2)
= \lambda h(x_1) + (1 - \lambda) h(x_2) + \lambda \alpha + (1 - \lambda) \alpha
= h(\lambda x_1 + (1 - \lambda) x_2) + \alpha = f(\lambda x_1 + (1 - \lambda) x_2) \in f(C).
\]

Now consider proving the convexity of \( f^{-1}(C) \). Take \( x_1, x_2 \in f^{-1}(C) \) and \( \lambda \in [0, 1] \). Then we have \( f(x_i) \in C \) for each \( i \) and so

\[
f(\lambda x_1 + (1 - \lambda) x_2) = h(\lambda x_1 + (1 - \lambda) x_2) + \alpha
= \lambda h(x_1) + (1 - \lambda) h(x_2) + \lambda \alpha + (1 - \lambda) \alpha
= \lambda f(x_1) + (1 - \lambda) f(x_2) \in C \quad \text{(by the convexity of } C)\]

Thus it follows that

\[
\lambda x_1 + (1 - \lambda) x_2 \in f^{-1}(C).\]
• Suppose $C \subseteq X \times Y$ is a convex set then under the linear projection mapping $P(x, y) = x$ the set

$$P(C) := \{(x \mid \exists y \text{ such that } (x, y) \in C\}$$

is convex (put possibly not closed).

**Why?** The projection $P$ is linear because for any $A, B \in \mathbb{R}$ we have

$$P(A(x_1, y_1) + B(x_2, y_2)) = P(Ax_1 + Bx_2, Ay_1 + By_2) = Ax_1 + Bx_2 = AP(x_1, y_1) + BP(x_2, y_2).$$

It may not be closed as seen in the following example: let

$$C = \{(x, y) \mid y \geq 1/x\} \text{ and } P(x, y) = y.$$

Then

$$P(C) = (0, +\infty) \text{ is an open convex set.}$$
CHAPTER 1. SUMMARY OF BASIC CONCEPTS FROM CONVEXITY
Chapter 2

Supports, Separation and Farka’s Lemma

2.1 Supports and Separation

In the following we denote the inner product by $\langle \cdot, \cdot \rangle$. Suppose $C$ is a closed, convex set and $x \notin C$. The problem of finding the closest point in $C$ to $x$ is well defined for a convex set. This problem has many important application and may be formulated as:

$$\text{Find } y \in \arg \min \{ \|x - y\| \mid y \in C\}$$

where the argmin refers to the elements that are arguments of the minimized function value. When $X$ is an inner product space the following classical analysis may be performed.

$$\|x - (y + \alpha z)\|^2 = \|x - y\|^2 - \alpha \langle z, x - y \rangle + \alpha^2 \|z\|^2 \geq \|x - y\|^2$$

$y$ is the closest point to $x$ in the convex set $C$. The distance form $x$ to $y + \alpha z$ is larger than that to $y$. Thus $\|x - (y + \alpha z)\| \geq \|x - y\|$ and so
implying
\[ \alpha \|z^2\| \geq \langle z, x - y \rangle. \]
As this is true for all sufficiently small \( \alpha \geq 0 \) we obtain
\[ 0 \geq \langle z, x - y \rangle. \tag{2.1} \]
That is the angle between the vector \( x - y \) and \( \alpha z \) is obtuse. Placing \( w := x - y \) we find that or any arbitrary point \( h := y + \alpha z \in C \) we have
\[ \langle w, h \rangle = \langle x - y, y + \alpha z \rangle \\
= \alpha \langle x - y, z \rangle + \langle x - y, y \rangle \\
\leq \langle x - y, y \rangle = \langle w, y \rangle. \]
Put this another way we have
\[ y \in \arg \max \{ \langle w, h \rangle \mid h \in C \}. \]
That is \( y \) is a maximizing element of the following problem
\[ \sup \{ \langle w, h \rangle \mid h \in C \}. \]
This prompts us to define the support function to \( C \) to be
\[ v \mapsto S(C, v) := \sup \{ \langle v, h \rangle \mid h \in C \} \tag{2.2} \]
for which \( S(C, w) = \langle y, w \rangle \) if we take \( w = x - y \) (with \( y \) the closest point to \( x \) in \( C \)). Note that for \( \lambda > 0 \)
\[ S(\lambda C, v) = \sup \{ \langle v, \lambda h \rangle \mid h \in C \} = \lambda \sup \{ \langle v, h \rangle \mid h \in C \} = \lambda S(C, v) . \]
Finally, when \( x \notin C \) and \( C \) is closed we have \( \|x - y\|^2 > 0 \) and so
\[ \langle x, w \rangle - \langle y, w \rangle = \langle x - y, w \rangle = \langle x - y, x - y \rangle = \|x - y\|^2 > 0 \]
or
\[ \langle x, w \rangle > \langle y, w \rangle = S(C, w) \geq \langle h, w \rangle \quad \text{for all } h \in C. \tag{2.3} \]
When such a \( w \) can be found we say that we have separated \( x \) from \( C \). It can be shown that for two convex sets \( C \) and \( D \) with \( \text{int } D \neq \emptyset \) and
\[ C \cap \text{int } D = \emptyset \]
then there exists \( w \) such that
\[ \langle c, w \rangle \geq \langle d, w \rangle \quad \text{for all } c \in C \text{ and } d \in D. \]
A similar result even holds in infinite dimensions (in real Banach spaces) and is usually referred to as strong separation of convex sets.
The two sets $C$ and $D$ are separated by a hyperplane.

The inner product is very important in the study of convexity since it can be used to generate a half-space via the linear mapping $x \mapsto \langle x, y \rangle$ i.e.

$$H := \{ x \in X \mid \langle x, y \rangle \geq \alpha \} \quad \text{and} \quad H' := \{ x \in X \mid \langle x, y \rangle \leq \alpha \}$$

are a half space for each $y$.

We denote the smallest closed convex set containing $C$ by $\text{conv} C$. Half spaces are always closed and convex. We have noted earlier in example 7 that half spaces are convex. To see that they are closed take a sequence $x_n \to x$ then when $x_n \in H'$ (say) for all $n$ we have

$$\langle x, y \rangle = \langle x - x_n, y \rangle + \langle x_n, y \rangle \leq \| y \| \| x - x_n \| + \alpha \to_{n \to \infty} \alpha.$$ 

Consequently $x \in H'$ establishing closedness.

**Lemma 15** Suppose $C \subseteq X$, an inner product space, then $\text{conv} C = \bigcap \{ H \mid C \subseteq H, \text{ and } H \text{ is a half-space} \}$.

**Proof.** Clearly $C \subseteq \bigcap \{ H \mid C \subseteq H, \text{ and } H \text{ is a half-space} \}$ and the larger set is closed and convex. Thus $\text{conv} C \subseteq \bigcap \{ H \mid C \subseteq H, \text{ and } H \text{ is a half-space} \}$ as $\text{conv} C$ is the smallest convex set containing $C$. Now suppose $\text{conv} C$ is a proper subset of $\bigcap \{ H \mid C \subseteq H, \text{ and } H \text{ is a half-space} \}$ then exists $x \notin \text{conv} C$ with $x \in \bigcap \{ H \mid C \subseteq H, \text{ and } H \text{ is a half-space} \}$. Then using the half spaced defined by $w = y - x$, where $y$ is the closest point to $x$ in $\text{conv} C$, we find on using (2.3) that

$$\text{conv} C \subseteq H := \{ z \in X \mid \langle z, w \rangle \leq \alpha := \langle y, w \rangle \} \quad \text{but } x \notin H.$$ 

But this cannot be so as we have assumed $x \in \bigcap \{ H \mid C \subseteq H, \text{ and } H \text{ is a half-space} \}$. Thus no such $x$ can exist and $\text{conv} C \supseteq \bigcap \{ H \mid C \subseteq H, \text{ and } H \text{ is a half-space} \}$. $\blacksquare$

We say that $y$ is supported by the hyperplane $H := \{ x \in X \mid \langle x, w \rangle = S(C, w) \}$ when $y \in C \cap H$.  

![Diagram](https://example.com/diagram.png)
A convex set is uniquely defined by its supporting hyperplanes.

A better way to write this relation is to use support functional. For any \( y \in X \) we need to find an \( \alpha \) so that \( C \subseteq \{ x \in X \mid \langle x, y \rangle \leq \alpha \} \). The smallest such value would be

\[
\alpha = S(C, y) := \sup_{x' \in C} \langle x', y \rangle \geq \langle x, y \rangle \quad \text{for all } x \in C.
\]

So clearly:

**Corollary 16** Suppose \( C \subseteq X \), an inner product space, then \( \text{conv} C = \{ x \in X \mid \langle x, y \rangle \leq S(C, y) , \text{ for all } y \in X \} \).

**Definition 17** When we only have a finite number of points in \( C \) then \( \text{conv} C \) is a polyhedral set. A convex polyhedron set is simply the intersection of a finite number of half spaces and so can be represented as:

\[
P := \{ x \in X \mid \langle x, y_i \rangle \leq \alpha_i , \ i = 1, \ldots , m \} = \{ x \in X \mid Ax \leq b \}
\]

where \( A \) contains \( y_i \) as rows and \( b \) contains \( \alpha_i \) as coordinates. A polytope is some time defined as a bounded polyhedron.

If one starts with the definition of a polytopes as a convex hull of a finite number of points then one needs to show it is bounded using the finite basis theorem which we will shortly derive.

**Example 18** (Integer programming and Chvatal–Gomory cuts). Given a sets

\[
C = \{ x \mid Ax \leq b, \ x \in \mathbb{Z}^n \}
\]

where \( \mathbb{Z} \) denotes the integral points \( \mathbb{Z} = \{ \ldots , -1, 0, 1, 2, 3, \ldots \} \). How can we find \( \text{conv} C \)? This can be done (in principle) by the addition of valid inequalities (i.e. those that separate points not in \( \text{conv} C \) from the set \( \text{conv} C \)). Consider for example let \( C \) be...
defined by the inequalities
\[
7x_1 - 2x_2 \leq 14 \\
x_2 \leq 3 \\
2x_1 - 2x_2 \leq 3 \\
x \geq 0, \ x \in \mathbb{Z}^2.
\]

Try first adding up the constraints with the weights \((\frac{2}{7}, \frac{37}{63}, 0)\) to obtain a valid constraint
\[
2x_1 + \frac{1}{63} x_2 \leq \frac{121}{21}.
\]

next note that the we may reduce the left hand side to the nearest integral value and obtain a weaker constraint (as \(x \geq 0\)). This gives
\[
2x_1 + 0x_2 \leq \frac{121}{21} = 5 + \frac{16}{21}.
\]

As the left hand side should take an integral value we may reduce the right hand side to an integral value without removing any integral points i.e.
\[
2x_1 \leq 5
\]

We could even continue this process by writing
\[
x_1 \leq \frac{5}{2} \implies x_1 \leq 2.
\]

Then the convex set \(B\) defined by
\[
7x_1 - 2x_2 \leq 14 \\
x_2 \leq 3 \\
2x_1 - 2x_2 \leq 3 \\
x_1 \leq 2 \\
x \geq 0
\]
satisfies
\[
B \supseteq \text{conv}C.
\]

It is a surprising but true fact that all valid inequalities can be generated in this manor. That is a finite (but possibly) large set of inequalities generated in the way defines \(\text{conv}C\).

### 2.1.1 Theorems of the Alternative - Farkas’ Lemma

This result is extremely useful and in some ways more flexible than linear programming duality. Farkas’ Lemma is about containment of a polyhedral set in a half space but can be formulated as an alternative. More precisely two logical propositions \(P\) and \(Q\) are said to form an alternative iff one and only one is true:
\[
P \iff \neg Q \quad \text{[or]} \quad Q \iff \neg P
\]

Farkas’ Lemma is really just about containment of convex cones.
Lemma 19 Let \( b, a_1, \ldots, a_m \) be given data in \( \mathbb{R}^n \). Then the convex cone
\[
C := \text{conv cone} \{a_1, \ldots, a_m\} = \left\{ \sum_{i=1}^{m} \lambda_i a_i \mid \lambda_i \geq 0, i = 1, \ldots, m \right\}
\]
is closed.

Proof. If all \( \{a_1, \ldots, a_m\} \) are linearly independent and \( a^k = \sum_{i=1}^{m} \lambda^k_i a_i \to y \) then as \( y \) can expressed as \( y = \sum_{i=1}^{m} \lambda_i a_i \) and \( y - a^k = \sum_{i=1}^{m} (\lambda^k_i - \lambda_i) a_i \to 0 \) by linear independence \( \lambda^k_i - \lambda_i \to 0 \). Consequently \( \lambda_i \geq 0 \) and \( y \in C \).

Now suppose \( \sum_{i=1}^{m} \beta_i a_i = 0 \) has a nonzero solution and we may suppose there is a \( j \) with \( \beta_j < 0 \) (take \( -\beta \) if not, as \( \sum_{i=1}^{m} (-\beta_i) a_i = 0 \)). Then note that if \( x \in C \) then there exists \( \lambda_i \geq 0 \) such that
\[
x = \sum_{i=1}^{m} \lambda_i a_i = \sum_{i=1}^{m} (\lambda_i + t^* (x) \beta_i) a_i \quad \text{as} \quad t^* (x) \sum_{i=1}^{m} \beta_i a_i = 0
\]
\[
= \sum_{i \neq j} \lambda'_i a_i \quad \text{where} \quad \lambda'_i := \lambda_i + t^* (x) \beta_i.
\]

When \( \beta_i \geq 0 \) it is clear that \( \lambda'_i \geq 0 \) and if \( \beta_i < 0 \) in order that \( \lambda'_i \geq 0 \) we require
\[
t^* (x) \leq \frac{\lambda_i}{\beta_i} \quad \text{for all} \quad i \quad \text{for which} \quad \beta_i < 0.
\]

If we choose \( j (x) \in \arg \min_{\beta_i<0} -\frac{\lambda_i}{\beta_i} \) and \( t^* (x) := -\frac{\lambda_{j(x)}}{\beta_{j(x)}} \) then \( \lambda'_i \geq 0 \) for all \( i \) and \( \lambda'_{j(x)} = 0 \).

In effect we have removed one element \( y_{j(x)} \) from this sum representing \( x \). If the resultant collection \( \{y_i\}_{i \neq j(x)} \) is still linearly dependent we can repeat this process and remove another. Eventually we arrive at a linearly independent set that allows us to represent \( x \). Thus to each \( x \) we may associate a subset of linearly independent \( \{a_1, \ldots, a_k\} \) that represent \( x \). Thus we arrive at a finite number of possible decompositions of \( C \) into a union of cones \( C_k \) for \( k = 1, \ldots, p \) where each \( C_i \) is generated by a combination of \( k \) linearly independent vectors i.e. \( C = \bigcup_{k=1}^{p} C_k \). As all these convex \( C_k \) (as shown in the first part) are closed then so is \( C \). □

Theorem 20 Let \( b, a_1, \ldots, a_m \) be given data in \( \mathbb{R}^n \). Then the set
\[
\{ v \in \mathbb{R}^n \mid v^T A \leq 0 \} = \{ v \in \mathbb{R}^n \mid \langle a_i, v \rangle \leq 0 \quad \text{for} \quad i = 1, \ldots, m \}
\]
is contained in the set
\[
\{ x \in \mathbb{R}^n \mid \langle b, v \rangle \leq 0 \}
\]
iff
\[
b \in \text{conv cone} \{a_1, \ldots, a_m\}.
\]

Proof. One direction is immediate. If \( b \in C := \text{conv cone} \{a_1, \ldots, a_m\} \) and
\[
z \in \{ v \in \mathbb{R}^n \mid \langle a_i, v \rangle \leq 0 \quad \text{for} \quad i = 1, \ldots, m \}
\]
then
\[
\langle b, z \rangle = \sum_{i=1}^{m} \lambda_i a_i, z = \sum_{i=1}^{m} \lambda_i \langle a_i, z \rangle \leq 0.
\]
2.1. SUPPORTS AND SEPARATION

Now if \( b \notin C \) then we may separate \( b \) from this closed convex cone \( C \). Thus there is a \( d \) such that \( \langle d, a \rangle \leq \alpha \) for all \( a \in C \) and with \( \langle b, d \rangle > \alpha \) where \( \alpha := S(C, d) \). Note now that as \( 0 \in C \) we have \( \alpha \geq 0 \). As \( \lambda a \in C \) for any \( y \in C \) we have \( \alpha / \lambda := \frac{1}{\lambda}S(C, d) = S(\lambda C, d) = S(C, d) = \alpha \) for all \( \lambda > 0 \) and so \( \alpha = 0 \). Consequently \( \langle b, d \rangle > 0 \) and so \( d \notin \{ v \in \mathbb{R}^n \mid \langle b, v \rangle \leq 0 \} \) but \( \langle d, a \rangle \leq 0 \) for all \( a \) (because \( \langle d, a \rangle \leq 0 \) for all \( a \in C \)) or

\[
d \in \{ v \in \mathbb{R}^n \mid \langle a_i, v \rangle \leq 0 \text{ for } i = 1, \ldots, m \}
\]
counter to the asserted containment. ■

Another way to look at the containment \( b \in \text{conv cone } \{a_1, \ldots, a_m\} \) is to say that the system of equations

\[
b = \sum_{i=1}^{m} a_i x_i = Ax, \quad x_i \geq 0, \quad i = 1, \ldots, m \tag{2.4}
\]

has a solution.

**Lemma 21 (Frakas’ lemma)** Let \( b, a_1, \ldots, a_m \) be given data in \( \mathbb{R}^n \). Then exactly one of the following statements are true:

1. The system of equations (2.4) has a solution:

2. The system of inequalities:

\[
v^T b = \langle b, v \rangle > 0, \quad \langle a_i, v \rangle \leq 0 \text{ for } i = 1, \ldots, m \quad \text{(ie. } v^T A \leq 0) \tag{2.5}
\]

has a solution \( v \).

**Proof.** Observe that of (2.4) has a solution then \( b \in \text{conv cone } \{a_1, \ldots, a_m\} \) and hence but the previous theorem (2.5) has no solution. When \( b \notin \text{conv cone } \{a_1, \ldots, a_m\} \) the previous proof establishes that (2.5) has a solution. ■

There are many variants of this theorem: (denote \( \mathbb{R}_m^m := \{ x \in \mathbb{R}^m \mid x \geq 0 \} \))

1. Either \( \{ x \in \mathbb{R}_m^m \mid Ax \leq b \} \neq \emptyset \) of \( \{ v \in \mathbb{R}_m^n \mid v^T A \geq 0, \ v^T b < 0 \} \neq \emptyset \).

2. Either \( \{ x \in \mathbb{R}^m \mid Ax \leq b \} \neq \emptyset \) of \( \{ v \in \mathbb{R}_m^n \mid v^T A = 0, \ v^T b < 0 \} \neq \emptyset \).

Let’s see why the first holds. Extend \( Ax \leq b \) to

\[
By = [A \ I] \begin{bmatrix} x \\ s \end{bmatrix} = b, \quad y \in \mathbb{R}_m^{m+n}.
\]

Then by Lemma 21 we don’t have a solution to this system then this is one to the system

\[
v^T b > 0 \quad \text{and} \quad \begin{bmatrix} A^T \\ I \end{bmatrix} v = \begin{bmatrix} A^T v \\ v \end{bmatrix} \leq 0
\]

or \( -v \geq 0 \) and \( (-v)^T A \geq 0 \) with \( (-v)^T b < 0 \).
Proposition 22 (LP duality) If both the primal LP
\[
\max \{ c^T x \mid Ax \leq b \}
\]
and the dual LP
\[
\min \{ b^T \lambda \mid A^T \lambda = c, \lambda \geq 0 \}
\]
then the two optimal values coincide.

**Proof.** Let \( x^* \) and \( \lambda^* \) be optimal for their respective problems then
\[
\text{primal optimal} = c^T x^* = (A^T \lambda^*)^T x^* = \lambda^T Ax^* \leq \lambda^T b = \text{dual optimal}.
\]
For the converse we will invoke Farkas’ lemma. Let \( v^* \) be the optimal value of the primal then the following linear system is infeasible for all \( \varepsilon > 0 \):
\[
c^T x \geq v^* + \varepsilon \quad \text{(or} \quad -c^T x \leq -v^* - \varepsilon
Ax \leq b
\]
where the first inequality is essential to the infeasibility of the system. By Farkas’ lemma there must be a solution to
\[
\lambda \geq 0, \quad \begin{bmatrix} -c & A^T \end{bmatrix} \begin{bmatrix} \lambda_0 \\ \lambda \end{bmatrix} = 0, \quad \begin{bmatrix} -v^* + \varepsilon & b^T \end{bmatrix} \begin{bmatrix} \lambda_0 \\ \lambda \end{bmatrix} < 0
\]
or \( \lambda \geq 0 \) with
\[-\lambda_0 c + A^T \lambda = 0 \quad \text{and} \quad (-v^* + \varepsilon) \lambda_0 + b^T \lambda < 0.
\]
Now \( \lambda_0 > 0 \) (as the first inequality was essential to the infeasibility of the system) and so
\[
A^T \left( \frac{\lambda}{\lambda_0} \right) = c \quad \text{and} \quad b^T \left( \frac{\lambda}{\lambda_0} \right) < v^* - \varepsilon.
\]
Letting \( \varepsilon \downarrow 0 \) we get a solution of the dual with a value less than that of the primal. ■
Chapter 3

Polyhedral Theory

In integer programming we would like a linear inequality description of sets like \( \{x \in \mathbb{Z}^n_+ \mid Ax \leq b\} \). This corresponds to the convex hull of a set discrete set of points and hence of a polyhedron. It mat not be bounded and this leads to some greater complexity in its description.

3.1 Dimensionality

Definition 23 Suppose \( C \subseteq \mathbb{R}^n \).

1. The conic hull of \( C \) is smallest convex conic set containing \( C \).

2. A set \( H \) is an affine subset if it contains all if affine combinations i.e. \( \{x^1, x^2, \ldots, x^k\} \subseteq H \) implies \( \sum_i \lambda_i x_i \in H \) for \( \sum_i \lambda_i = 1 \).

3. The affine hull aff \( C \) is is the smallest affine set containing \( C \).

4. The dimension of a convex set \( C \) is

\[
d := \dim \text{[aff } C - x] \text{ for some } x \in C.
\]

[That is it’s dimension is that of the smallest affine subspace that contains it.]

5. A set of points \( \{x^1, x^2, \ldots, x^k\} \subseteq \mathbb{R}^n \) are affinely independent iff the unique solution of \( \sum_{i=1}^{k} \alpha_i x_i = 0, \sum_{i=1}^{k} \alpha_i = 0 \) is \( \alpha_i = 0, i = 1, \ldots, k \).

6. A set of points \( \{x^1, x^2, \ldots, x^k\} \subseteq \mathbb{R}^n \) are linearly independent iff the unique solution of \( \sum_{i=1}^{k} \alpha_i x_i = 0 \) is \( \alpha_i = 0, i = 1, \ldots, k \).

Proposition 24 Suppose \( C \subseteq \mathbb{R}^n \) then

1. \( \text{aff } C := \{\sum_i \lambda_i x_i \mid \sum_i \lambda_i = 1, \text{ for } \{x_i\}_{i \in F} \subseteq C, F \text{ a finite set}\} \) and

2. \( \text{cone } C := \{\sum_i \lambda_i x_i \mid \lambda_i \geq 0, \text{ for } \{x_i\}_{i \in F} \subseteq C, F \text{ a finite set}\} \).
Proof. Denote

\[ D := \left\{ \sum_i \lambda_i x_i \mid \sum_i \lambda_i = 1, \text{ for } \{x_i\}_{i \in F} \subseteq C, \ F \text{ a finite set} \right\} \]

which is clearly an affine set by construction. Also \( D \supseteq C \) by construction so \( D \supseteq \text{aff } C \) being the smallest such set. Since \( C \subseteq \text{aff } C \) we may take \( \{x_i\}_{i \in F} \subseteq C \) for a \( F \) a finite set and hence \( \{x_i\}_{i \in F} \subseteq \text{aff } C \). Thus \( \sum_i \lambda_i x_i \in \text{aff } C \) when \( \sum_i \lambda_i = 1 \) by the definition of affineness of \( \text{aff } C \) and hence \( D \subseteq \text{aff } C \). The other follows in a similar way. ■

Recall that the rank \( A \) corresponds the number of linearly independent column (or rows as these numbers always coincide). Also we have the following well known result from linear algebra.

1. The dimension of the range of \( A \) is its rank ie.

\[ \dim \text{Range } A = \text{rank } A. \]

2. The kernel of \( A \) is the subspace

\[ \ker A := \{ x \in \mathbb{R}^n \mid Ax = 0 \}, \]

and the nullity if its dimension

\[ \text{nullity } A := \dim \ker A \]

3. We have

\[ n = \text{rank } A + \text{nullity } A. \]

Proposition 25 The following are equivalent:

1. \( \{x^1, x^2, \ldots, x^k\} \subseteq \mathbb{R}^n \) are affinely independent:

2. \( \{x^2 - x^1, \ldots, x^k - x^1\} \) are linearly independent:

3. \( k - 1 = \dim \text{conv } \{x^1, x^2, \ldots, x^k\} \)

Proof. 1. \( \Rightarrow \) 2. Consider

\[ \sum_{i=2}^k \alpha_i (x^i - x^1) = 0 \] (3.1)

Suppose we do not have linear independence.

Case 1: Assume \( \left( \sum_{i=2}^k \alpha_i \right) \neq 0 \) then we have (3.1) equivalent to

\[ \sum_{i=2}^k \frac{\alpha_i}{\sum_{i=2}^k \alpha_i} x^i - x^1 = 0 \]

where

\[ \sum_{i=2}^k \frac{\alpha_i}{\sum_{i=2}^k \alpha_i} - 1 = 0. \]
By affine independence we should have \( \left( \sum_{i=2}^{k} \alpha_i \right) = 0 \) for all \( i \), a contradiction.

Case 2: We have \( \left( \sum_{i=2}^{k} \alpha_i \right) = 0 \). Then (3.1) equivalent to

\[
\sum_{i=2}^{k} \alpha_i x^i = 0 \quad \text{and} \quad \left( \sum_{i=2}^{k} \alpha_i \right) = 0.
\]

By affine independence of \( \{ x^2, \ldots, x^k \} \) we must have \( \alpha_i = 0 \) for \( i = 2, \ldots, k \) as desired.

[2. \Rightarrow 1.] Suppose \( \sum_{i=1}^{k} \alpha_i x^i = 0, \sum_{i=1}^{k} \alpha_i = 0 \)
then \( \alpha_1 = - \sum_{i=2}^{k} \alpha_i \) and so

\[
\sum_{i=1}^{k} \alpha_i x^i = 0 \quad \Rightarrow \quad \sum_{i=1}^{k} \alpha_i x^i - \left( \sum_{i=2}^{k} \alpha_i \right) x^1 = \sum_{i=2}^{k} \alpha_i (x^i - x^1) = 0.
\]

By linear independence we have \( \alpha_i = 0 \) for \( i = 2, \ldots, k \) and so \( \alpha_1 = - \sum_{i=2}^{k} \alpha_i = 0 \) as well.

[2. \Rightarrow 3.] The \( \dim \text{conv} \{ x^1, x^2, \ldots, x^k \} \) is the same as the dimension \( \dim \text{aff} C - x \) where \( x \in \text{aff} C \). Indeed

\[
\dim \text{conv} \{ x^1, x^2, \ldots, x^k \} \geq \dim \{ x^2 - x^1, \ldots, x^k - x^1 \} = k - 1,
\]

by linear independence in 2. Now \( x \in \text{aff} C \) iff \( x = \sum_{i=1}^{k} \alpha_i x^i \) for \( \sum_{i=1}^{k} \alpha_i = 1 \). Thus

\[
x = \sum_{i=2}^{k} \alpha_i x^i - \left( \sum_{i=2}^{k} \alpha_i - 1 \right) x^1
= \sum_{i=2}^{k} \alpha_i (x^i - x^1) + x^1
\]

so

\[
x - x^1 = \sum_{i=2}^{k} \alpha_i (x^i - x^1).
\]

By the linear independence in 2. we have \( \dim \text{aff} C - x = k - 1 \).

[3. \Rightarrow 2.] We have \( k - 1 = \dim \text{aff} C - x_1 \leq \dim \text{span} \{ x^2 - x^1, \ldots, x^k - x^1 \} \)

implying linear independence. \( \blacksquare \)

We will call \( k \) the affine dimension. We also need the next rank result.

**Proposition 26** If \( \{ x \in \mathbb{R}^n \mid Ax = b \} \neq \emptyset \), the maximum number of affinely independent solutions of \( Ax = b \) is \( n + 1 - \text{rank } A \).
CHAPTER 3. POLYHEDRAL THEORY

**Proof.** We have \( C = \{ x \in \mathbb{R}^n \mid Ax = b \} \) a convex set and for any \( x^1 \in C \) we have \( Ax^1 = b \) so
\[
C - x^1 = \{ x \in \mathbb{R}^n \mid Ax = 0 \}.
\]
The the smallest subspace containing \( C - x^1 \) is clearly \( \{ x \in \mathbb{R}^n \mid Ax = 0 \} = \ker A \) so
\[
\dim \text{aff} \ C - x^1 = \dim \ker A = \text{nullity} \ A.
\]
Now by Proposition 25 we have
\[
k = \text{nullity} \ A + 1
\]
where \( k \) equals the maximum number of affinely independent solutions of \( Ax = b \) (the affine dimension). But
\[
n = \text{rank} \ A + \text{nullity} \ A
\]
gives \( n + 1 - \text{rank} \ A = \text{nullity} \ A + 1 = k \).

3.2 Polyhedra and Dimension

**Definition 27** A polyhedron \( P \subseteq \mathbb{R}^n \) is the convex set generated by the intersection of a finite number of hyperplanes i.e.
\[
P = \{ x \in \mathbb{R}^n \mid Ax \leq b \}
\]
where \( (A,b) \) is an \( n \times (m+1) \) matrix. A polyhedron is called rational if there exists representation of \( P \) with \( (A,b) \) containing rational coefficients.

Clear the representation of \( P \) via \( (A,b) \) need not be unique.

**Proposition 28** The polyhedron \( \{ x \in \mathbb{R}^n \mid Ax \leq 0 \} \) is a cone.

**Proof.** An exercise ■

**Definition 29** We say a polyhedron is of dimension \( k \) iff it contains \( k + 1 \) affinely independent points. Thus a polyhedron of full dimension would have \( \dim P = n \).

Denote \( I := \{1,2,\ldots,m\} \) and
\[
I^= := \{ i \in I \mid a^i x = b_i, \text{ for all } x \in P \} \quad \text{and} \quad I^\leq := \{ i \in I \mid a^i x < b_i \text{ for some } x \in P \}.
\]
Let \( (A^=, b^=) \) be the corresponding of \( (A,b) \) corresponding to \( I^= \) and \( (A^\leq, b^\leq) \) be the corresponding rows of \( (A,b) \) then
\[
P = \{ x \in \mathbb{R}^n \mid A^= x = b^=, A^\leq x \leq b^\leq \}.
\]

**Definition 30** A point \( x \in P \) is call an inner point (or relative interior point) of \( P \) if \( a^i x < b_i \) for all \( i \in M^\leq \). An interior point of \( P \) satisfies \( a^i x < b_i \) for all \( i \in I \).

Inner points can be viewed as interior points ”relative to aff \( P \).” A polyhedron does not necessarily have an interior point (and won’t if \( I^= \neq \emptyset \)).
3.2. POLYHEDRA AND DIMENSION

Proposition 31 Every polyhedron $P$ has an inner point.

Proof. If $I^<=\emptyset$ then every point of $P$ is an inner point. Otherwise take $x^i \in P$ with $a^i x^i < b_i$ and place

$$\bar{x} := \frac{1}{|I^<=|} \sum_{i \in I^<=} x^i \in P$$

by convexity. Then

$$a^p \bar{x} = \frac{1}{|I^<=|} \sum_{i \in I^<=} a^p x^i \leq \frac{1}{|I^<=|} \left\{ \sum_{i \in I^<= \backslash \{p\}} b^p + a^p x^p \right\} < b^p$$

for all $p \in I^<=$. ■

The dimension of $P$ is related to the rank of $A^-$. 

Proposition 32 If $P \subseteq \mathbb{R}^n$ then $\dim P + \operatorname{rank} (A^-, b^-) = n$.

Proof. The dimension of $P$ is one less than the number of affinely independent points it contains. By Proposition 26 we have the maximum number of affinely independent solutions of $A^- x = b^-$ is $n + 1 - \operatorname{rank} A^-$. Thus we only need show that $\dim P = \dim \{ x \in \mathbb{R}^n \mid A^- x = b^- \}$. Now $\dim P = \dim [\text{aff } P - x]$ for any $x \in P$ and

$$P - x = \{ x \in \mathbb{R}^n \mid A^- x = 0, A^- x \leq 0 \} \subseteq \{ x \in \mathbb{R}^n \mid A^- x = 0 \}$$

which is a subspace and clearly the smallest such subspace. Thus $\text{aff } P - x = \{ x \in \mathbb{R}^n \mid A^- x = 0 \}$ and so

$$\dim P = \dim [\text{aff } P - x] = \dim \{ x \in \mathbb{R}^n \mid A^- x = b^- \}.$$ 

Finally note that $\{ x \in \mathbb{R}^n \mid A^- x = b^- \} \neq \emptyset$ iff $\operatorname{rank} (A^-, b^-) = \operatorname{rank} A^-$. ■

When $\operatorname{rank} (A^-, b^-) = 0$ we have $I^- = \emptyset$.

Corollary 33 A polyhedron $P$ is of full dimension iff it has an interior point.

We finish this section with Carathéodory’s theorem. This asserts that in finite dimensions we only require at most $n + 1$ points in any of these sums.

Theorem 34 (Carathéodory) For any subset $C \subseteq \mathbb{R}^n$ with $\dim \text{aff } C = n$, the convex hull $\operatorname{conv} C$ is the set of convex combinations using at most $n + 1$ points in $C$.

Proof. Let $y \in \operatorname{conv} C$ and suppose we need $r > n + 1$ points to express $y$ i.e.

$$y = \sum_{i=1}^{r} \lambda_i x^i, \lambda_i > 0 \text{ and } \sum_{i=1}^{r} \lambda_i = 1. \quad (3.2)$$

As $\dim \text{aff } C = n$ there are at most $n + 1$ affinely independent points in $C$ and so (renumber if needed)

$$x^1 = \sum_{i=2}^{r} \theta_i x^i \text{ and } \sum_{i=2}^{r} \theta_i = 1. \quad (3.3)$$
Using (3.2) and (3.3)
\[ y = \sum_{i=1}^{r} \lambda_i x^i + \beta \left( x^1 - \sum_{i=2}^{r} \theta_i x^i \right) \]
\[ = (\lambda_1 + \beta) x^1 + \sum_{i=2}^{r} (\lambda_i - \beta \theta_i) x^i \]
for all real \( \beta \). Now all \( \lambda_i > 0 \) and at least one \( \theta_i > 0 \) (as \( \sum_{i=2}^{r} \theta_i = 1 \)) there is a positive \( \beta > 0 \) such that \( (\lambda_i - \beta \theta_i) \geq 0 \) with at least one \( \lambda_k - \beta \theta_k = 0 \) (note that as \( \beta > 0 \) we have \( \lambda_1 + \beta > 0 \)). As
\[ (\lambda_1 + \beta) + \sum_{i=2}^{r} (\lambda_i - \beta \theta_i) = \beta - \beta \sum_{i=2}^{r} \theta_i + \sum_{i=1}^{r} \lambda = 1 \]
we have \( y \) an affine combination of less that \( r - 1 \) points. Now proceed iteratively to reduce \( r \) till \( r = n + 1 \) (the minimal set as there are \( n + 1 \) affinely independent points in \( C \)).

### 3.3 Faces and Facets and Dimensions

We need to deal with unbounded polyhedral sets.

**Definition 35** Let \( C \subseteq \mathbb{R}^n \) be a convex set.

1. The recession cone of \( C \):
   \[ \text{rec} \, C := \{ x \in \mathbb{R}^n \mid C + x \subseteq C \} \, . \]

2. The linearity space of \( C \):
   \[ \text{lin} \, C := \text{rec} \, C \cap (-\text{rec} \, C) \, . \]

3. A ray \( r \) of \( C \) satisfies
   \[ x + \lambda r \in C \quad \text{for all} \, x \in C \, \text{and} \, \lambda \geq 0 \, . \]

4. A ray is extremal iff there do not exists rays \( r_1, r_2 \in C \) such that \( r = \lambda r_1 + (1 - \lambda) r_2 \) for some \( \lambda \in (0, 1) \).

Clearly a ray \( r \in \text{rec} \, P \). For a polyhedron \( P = \{ x \in \mathbb{R}^n \mid Ax \leq b \} \) we have
\[ \text{rec} \, P = \{ x \in \mathbb{R}^n \mid Ax \leq 0 \} \quad \text{and} \]
\[ \text{lin} \, P = \{ x \in \mathbb{R}^n \mid Ax = 0 \} \, . \]

A polyhedral cone \( P = \{ x \in \mathbb{R}^n \mid Ax \leq 0 \} \) is pointed if \( \text{lin} \, P = \{ 0 \} \).

It is our alternate objective to derive a representation of a polyhedral set \( P \) as
\[ P = \text{conv} \{ x_1, \ldots, x_q \} + \text{conv} \{ r_{q+1}, \ldots, r_{q+k} \} \, \quad (3.4) \]
where \( \{ x_1, \ldots, x_q \} \subseteq P \) and \( \{ r_{q+1}, \ldots, r_{q+k} \} \) are extremal rays of \( P \). To do so we need to describe \( P \) via its faces and facets.
Definition 36 Let $P$ be a polyhedral set.

1. A Face $F \subseteq P$ of $P$ is a subset supported by a hyperplane i.e.

$$F = P \cap H$$

where $H := \{ x \in \mathbb{R}^n \mid c^T x = c_0 \}$.

2. A facet is a face of dimension $\dim F = \dim P - 1$.

3. A face $F$ is proper if $F \subsetneq P$.

One can partially order faces via inclusion and the minimal faces ordered by inclusion play a role in the choice of $\{x_1, \ldots, x_q\}$ in (3.4) while the maximal faces correspond to facets.

Proposition 37 If

$$P = \{ x \in \mathbb{R}^n \mid A^a x = b^a, A^\leq x \leq b^\leq \}$$

is polyhedral and $F$ a face. Then there exists $I_F^\geq \supseteq I^\geq$ and $I_F^\leq := I \setminus I_F^\geq$ such that

$$F = \{ x \in \mathbb{R}^n \mid A_F^\geq x = b_F^\geq, A_F^\leq x \leq b_F^\leq \}$$

where $(A_F^\geq, b_F^\geq)$ are the rows of $(A, b)$ corresponding to the indices $I_F^\geq$ and likewise for $(A_F^\leq, b_F^\leq)$. The number of faces are finite.

Proof. Let $c = \sum_{i \in I^\geq} a^i$ and $c_0 = \sum_{i \in I^\leq} b_i$. Then $F \subseteq \{ x \in \mathbb{R}^n \mid c^T x = c_0 \}$ as these are sums of equality constraints. Also $F \subseteq \{ x \in \mathbb{R}^n \mid Ax \leq b \} \cap \{ x \in \mathbb{R}^n \mid c^T x = c_0 \}$. Then $F$ is contained in the set of optimal solutions to the LP $v^* := \max \{ c^T x \mid Ax \leq b \}$ with $c_0$ its optimal value. If $\lambda^*$ is an optimal solution of the dual $v^* = \min \{ b^T \lambda \mid A^T \lambda = c, \lambda \geq 0 \}$. Let $I^* := \{ i \in I \mid \lambda^*_i > 0 \}$ (the active constraints in the primal). Now consider the polyhedron

$$F^* := \{ x \in \mathbb{R}^n \mid a^i x = b_i \text{ for } i \in I^* \text{ and } a^i x \leq b_i \text{ for } i \in I \setminus I^* \}.$$ 

We show $F = F^*$. If $x \in F^*$ then

$$c^T x = \lambda^* A x = \sum_{i \in I^*} \lambda^*_i a^i x = \sum_{i \in I^*} \lambda^*_i b_i = c_0 \text{ (primal optimal).}$$

But if $x \in P \setminus F^*$, then $a^k x < b_k$ for some $k \in I^*$, so $\lambda^*_k > 0$ and

$$c^T x = \sum_{i \in I^*} \lambda^*_i a^i x < \sum_{i \in I^*} \lambda^*_i b_i = c_0$$

hence $F = F^*$ and $F$ is a polyhedron. Since $F \subseteq P$ the set $I_F^\geq$ has the required property (and $I_F^\leq := I \setminus I^*$). As $I$ is finite so are the number of faces. ■

Proposition 38 If $F$ is a facet of $P$, there is exists some inequality in $a^k x \leq b_k$ for $k \in I^\geq$ representing $F$ i.e.

$$F = P \cap \{ x \in \mathbb{R}^n \mid a^k x = b_k \}.$$
Proof. Since \( \dim F = \dim P - 1 \) we have from Proposition 32 that
\[
\text{rank} (A^*_F, b^*_F) = \text{rank} (A, b) + 1
\]
from which it follows that \( F \) requires an additional constraint from \( I^\circ \setminus I^*_F \). ■

Remark 39 It should be clear that one requires one of the inequalities representing \( P \) to describe any facet \( F \).

The description of \( P \) is much easier when \( \text{lin} P = \emptyset \). In this case we have the existence of extremal points. These are points \( x \in P \) for which there does not exist \( x_1, x_2 \in P \) and \( \lambda \in (0, 1) \) for which \( x = \lambda x_1 + (1 - \lambda) x_2 \). When \( \text{lin} P \neq \emptyset \) then we do not have extremal points.

Proposition 40 If \( P = \{ x \in \mathbb{R}^n \mid Ax \leq b \} \neq \emptyset \) and \( \text{rank} A = n - k \), \( P \) has faces of dimension \( k \) and no proper faces of lower dimension.

Proof. For any face \( F \neq \emptyset \) of \( P \), \( \text{rank} (A^*_F, b^*_F) \leq n - k \). By Proposition 32 we have \( \dim F \geq k \). Now let \( F \) be a face of minimal dimension. If \( \dim F = 0 \) there is nothing to prove, so assume \( \dim F > 0 \).

Let \( \bar{x} \) be an inner point of \( F \). Since \( \dim F > 0 \), there is some other point \( y \) in \( F \). Joining \( \bar{x} \) and \( y \) we get \( z(\lambda) := \bar{x} + \lambda (y - \bar{x}) \) where \( \lambda \in \mathbb{R} \). Suppose that this line intersects \( a^i x = b_i \) for some \( i \in I^*_F \). Set \( \lambda^* = \min \{|\lambda^i| \mid i \in I^*_F, z(\lambda^i) \text{ lies in } a^i x = b_i \} = \lambda^{i^*} \)
to be the smallest \( \lambda \) for which one equality becomes an equality. Now \( \lambda^* \neq 0 \) since \( \bar{x} \) is an inner point. Then
\[
F_{\lambda^*} := \{ A^*_F x = b^*_F, a^{i^*} x = b_i^* \} \neq \emptyset
\]
is a small face, a contradiction.

Thus the line does not intersect any \( a^i x = b_i \) for any \( i \in I^*_F \). Hence
\[
A \bar{x} + \lambda A (y - \bar{x}) \leq b \quad \text{for all } \lambda \in \mathbb{R}.
\]

Dividing by \( \lambda > 0 \) gives
\[
\lim_{\lambda \to +\infty} \frac{1}{\lambda} A \bar{x} + A (y - \bar{x}) \leq \lim_{\lambda \to +\infty} \frac{1}{\lambda} b
\]
or \( A (y - \bar{x}) \leq 0 \).

Dividing by \( \lambda < 0 \) gives
\[
\lim_{\lambda \to -\infty} \frac{1}{\lambda} A \bar{x} + A (y - \bar{x}) \geq \lim_{\lambda \to 1^{-}} \frac{1}{\lambda} b
\]
or \( A (y - \bar{x}) \geq 0 \).

Hence \( A (y - \bar{x}) = 0 \) for all \( y \in F \) or \( F = \{ y \in P \mid Ay = A \bar{x} \} \). Since \( \text{rank} A = n - k \) Proposition 32 gives \( \dim (F) = k \). ■

Definition 41 An \( x \in P \) is an extremal point iff
\[
x = \frac{1}{2} x_1 + \frac{1}{2} x_2, \quad x_1, x_2 \in P \quad \implies \quad x_1 = x_2 = x.
\]
Proposition 42 If \( x \) is a extremal point of \( P \) iff it is a zero dimensional face.

Proof. Suppose \( x \) is zero dimensional face. Then \( \text{rank} (A^\top_F) = n \) (by Proposition 32). Let \((\bar{A}, \bar{b})\) be a submatrix of \((A, b)\) with \(\bar{A}\) nonsingular, so \( x = \bar{A}^{-1}\bar{b} \). If \( x = \frac{1}{2}x_1 + \frac{1}{2}x_2 \) with \( x_1, x_2 \in P \), then
\[
\bar{b} = \bar{A}x = \frac{1}{2}\bar{A}x_1 + \frac{1}{2}\bar{A}x_2 \leq \bar{b},
\]
so
\[
\frac{1}{2}(\bar{b} - \bar{A}x_1) + \frac{1}{2}(\bar{b} - \bar{A}x_2) = 0
\]
and since \(\bar{A}x_i \leq \bar{b}\) (or \((\bar{b} - \bar{A}x_i) \geq 0\) we have \( \bar{b} = \bar{A}x_1 = \bar{A}x_2 \) or \( \bar{A}^{-1}\bar{b} = x_1 = x_2 \). So \( x \) is extremal.
Conversely if \( x \in P \) is not in a zero dimensional face of \( P \) then \( x \in F \) with \( \text{rank} (A^\top_F) < n \). But then there exists \( y \neq 0 \) such that \( A^\top_F y = 0 \) and then for sufficiently small \( \varepsilon \) we have \( x^1 = x + \varepsilon y \in F \) and \( x^1 = x - \varepsilon y \in F \) giving
\[
x = \frac{1}{2}x^1 + \frac{1}{2}x^1,
\]
and \( x \) is not an extremal point. \( \blacksquare \)

Corollary 43 If \( P = \{ x \in \mathbb{R}^n \mid Ax \leq b \} \) is not empty and \( \text{rank} A = n \), then \( F \) is a minimal face iff it is an extremal point.

Proof. As \( F \) is minimal it has dimension \( \text{dim} F = n - n = 0 \) and so is a point. \( \blacksquare \)

Remark 44 From this analysis we see that notion of a basic feasible point arises as a description of an extremal point.

If \( A \) is and \( n \times m \) (\( n \leq m \)) of full rank then there are a number of submatrices \( B \) such that \( B \) is square and invertible. We call the columns of \( B \) are called a basis and often
we identify $B$ via the notation $A_B$ where $B$ now refers to the indices associated with the columns of $A$ that form a basis. Then $\bar{x}_B := A_B^{-1}b$ provide the nonzero values of the basic solution $x_B$ associated with $B$. We place $x_i = 0$ for all $i \notin B$. When $x_B \geq 0$ we say it is basic feasible.

**Proposition 45 (Basic Feasible Solutions)** A vector $\bar{x}$ is a basic feasible solution of the system $Ax \leq b$ iff $\bar{x}$ is an extremal point of the polyhedron $P := \{x \in \mathbb{R}^n \mid Ax \leq b\}$.

We say a constraint is redundant iff it’s removal from the description of $P$ does not change $P$.

**Proposition 46** If $P = \{x \in \mathbb{R}^n \mid Ax \leq b\}$ and $q \in I^\leq$ then $F = \{x \in P \mid a^q x = b_q\}$ is a facet of $P$ (ie. $\dim F = \dim P - 1$) iff $a^q x \leq b_q$ is not a redundant constraint.

\[\iff \text{only.}\] To show $\dim P = \dim F + 1$ we need to show that $\rank([A^=, a^q]) = \rank A^= + 1$. That is $a^q$ is not a linear combination of the columns of $A^=$. Suppose this is not so then $a^q = \sum_{i \in I^=} \lambda_i a^i$. Let $\bar{b} := \sum_{i \in I^=} \lambda_i b_i$ then $q \in I^\leq$ implies

\[a^q x = \sum_{i \in I^=} \lambda_i a^i x = \sum_{i \in I^=} \lambda_i b_i = \bar{b}\]

for all $x \in P$ and for $x \in P \setminus F$ we have $a^q x = \bar{b} < b_q$. Then we have $\{x \in \mathbb{R}^n \mid a^q x \leq b_q\} \supseteq \{x \in \mathbb{R}^n \mid a^q x = \bar{b}\} \supseteq P$ and so $a^q x \leq b_q$ is a redundant constraint. Thus $a^q$ is linearly independent of $A^=$ and $\rank([A^=, a^q]) = \rank A^= + 1$. Hence

\[\dim F = n - \rank([A^=, a^q]) = n - (\rank A^= + 1) = \dim P - 1.\]

The final one in this cycle of results is!

**Proposition 47** Every inequality $a^q x \leq b_q$ for $q \in I^\leq$ that represents a face of $P$ of dimension less than $\dim P - 1$ is redundant to the description of $P$.

**Proof.** Suppose $a^q x \leq b_q$ represents a face of $P$ of dimension $\dim P - k$ with $k > 1$ and the constraint is not redundant. Then there exists $x \in \mathbb{R}^n$ such that $A^= x = b^=, a^i x \leq b_i$ for all $i \in I^\setminus\{q\}$ and $a^q x > b_q$. Let $\bar{x}$ be an inner point of $P$. Draw a line from $x$ to $\bar{x}$ then on the line there exists $z \in F$ such that $A^= z = b^=, a^i z \leq b_i$ for all $i \in I^\setminus\{q\}$ and $a^q z = b_q$. Hence the equality set of $F$ is $(A^=, b^=)$ and $(a^q, b_q)$ which is of rank $n - \rank P + 1$ implying

\[\dim F = n - (n - \rank P + 1)\]

\[= \rank P - 1\]

a contradiction. ■
Example 48 Consider $P$ defined by the system

$$
\begin{align*}
&x_1 + x_2 + x_3 \leq 1 \\
&-x_1 - x_2 - x_3 \leq -1 \\
&x_1 + x_3 \leq 1 \\
&-x_1 \leq 0 \\
&-x_2 \leq 0 \\
&x_3 \leq 2 \\
&x_1 + x_2 + 2x_3 \leq 2 
\end{align*}
$$

1. The three points $(1,0,0)$, $(0,1,0)$, $(0,0,1)$ satisfy the inequalities and are affinely independent. So $\dim P \geq 2$. Because all points of $P$ satisfy $x_1 + x_2 + x_3 = 1$ (conjunction of first two constraints) we have $\text{rank } (A^=, b^=) \geq 1$ so $\dim P \leq 3 - 1 = 2$. Thus $\dim P = 2$.

2. Note

$$F := \{ x \in P \mid -x_1 + x_2 + x_3 = 1 \} = \{(0,0,1)\}$$

is a face of $P$ as it is supported by a hyperplane (not in the description of $P$). It is indeed a face of zero dimension and so an extremal point.

3. Now consider

$$F_2 := \{ x \in P \mid 2x_1 - 7x_2 + 2x_3 = 2 \}$$

The two points $(1,0,0)$ and $(0,0,1)$ satisfy this system so are in $F_2$. As these two points are affinely independent we have $\dim F_2 \geq 1$. As $(0,1,0)$ is not in $F_2$ we have $F \nsubseteq P$ and so $1 \leq \dim F_2 < \dim P = 2$ so $\dim F_2 = 1$. Thus $F_2$ is a facet. It must be representable by a system of constraints (augmented by one non-redundant constraint) taken from $I^\leq$. Indeed

$$F_2 = \{ x \in P \mid x_1 + x_3 = 1 \}.$$ 

This also implies $x_2 = 0$.

4. Note that

$$\max \{ x_3 \mid x \in P \} = 1 < 2$$

so the constraint $x_3 \leq 2$ is redundant.

5. Note that for $F_2 = \{ x \in P \mid x_1 + x_3 = 1 \}$ we have

$$(A^=_{F_2}, b^=_{F_2}) = \begin{pmatrix}
1 & 1 & 1 & 1 \\
-1 & -1 & -1 & -1 \\
1 & 0 & 0 & 1 \\
0 & -1 & 0 & 0 
\end{pmatrix}$$

[Note that $x_1 + x_2 + 2x_3 < 2$ on part of $F_2$, try $(1,0,0)$]. This is matrix of rank 2 so $\dim F = n - \dim (A^=, b^=) = 2 + 1 - 2 = 1$. 

6. Now consider the constraint \( x_1 + x_2 + 2x_3 \leq 2 \), which clearly supports \( P \) (try the point \((0,0,1)\)). Let 

\[
F_3 := \{ x \in P \mid x_1 + x_2 + 2x_3 = 2 \} \\
= \{ x \in \mathbb{R}^n \mid x_1 + x_2 + x_3 = 1, x_1 + x_3 = 1, -x_1 = 0, -x_2 \leq 0, x_1 + x_2 + 2x_3 = 2 \}
\]

Hence the rank of \((A_F^\pi, b_F^\pi)\) is 3. Thus 

\[
\dim F_3 = n + 1 - \text{rank } (A_F^\pi, b_F^\pi) \\
= 3 - 3 = 0.
\]

Thus \( \dim F_3 < \dim P - 1 = 2 - 1 = 1 \) and so is redundant.

7. The minimal description of \( P \) is 

\[
\begin{align*}
& x_1 + x_2 + x_3 \leq 1 \\
& -x_1 - x_2 - x_3 \leq -1 \\
& x_1 + x_3 \leq 1 \\
& -x_1 \leq 0 \\
& -x_2 \leq 0
\end{align*}
\]

### 3.4 The Finite Basis Theorems (unbounded polyhedral sets)

We defined extremal rays before.

**Proposition 49** If \( P := \{ x \in \mathbb{R}^n \mid Ax \leq b \} \neq \emptyset \), \( r \) is an extremal ray of \( P \) iff \( \{ \lambda r \mid \lambda \geq 0 \} \) is a one-dimensional face of \( P^0 := \{ x \in \mathbb{R}^n \mid Ax \leq 0 \} \).

\( \iff \) only. Let \( A_r^\pi := \{ a^i \mid i \in I^\pi, a^i r = 0 \} \). If \( \{ \lambda r \mid \lambda \geq 0 \} \) is a one dimensional face of \( P^0 \) then \( \text{rank } (A_r^\pi) = n - 1 \) and \( A_r^\pi y = 0 \) has on degree of freedom i.e. \( y = \lambda r \), for \( \lambda \in \mathbb{R} \) i.e. \( r \) is a ray. ■

It should be clear that a polyhedron has a finite number of extremal points and rays. It can have zero extremal points if \( \dim P < n \). When they do exists it is commonly known that they are optimal solutions to linear programming problems. In fact for every extremal point there is a uniquely supporting hyperplane that defines a linear objective that yield that extremal point as an optimal solution. Moreover all finitely obtained objective values are attained by an extremal point of \( P \). A similar situation exists for extremal rays.

**Proposition 50** If \( P \neq \emptyset \) and \( \text{rank } A = n \) and \( \max \{ c^T x \mid x \in P \} \) is unbounded, then \( P \) has an extremal ray \( r^* \) with \( c^T r^* > 0 \).

**Proof.** Now \( \max \{ c^T x \mid x \in P \} \) unbounded mean \( \{ u \in \mathbb{R}_+^n \mid A^T u = c \} = \emptyset \) so by Farkas Lemma there exists \( r \) such that \( Ar \leq 0 \) and \( cr > 0 \). Now consider the linear program 

\[
\max \{ cr \mid Ar \leq 0, cr \leq 1 \} = 1
\]
must have an optimal extrema point solution \( r^* \in P^0 = \{ r \in \mathbb{R}^n \mid Ar \leq 0 \} \) such that

\[
\dim F_{r^*} = n - \text{rank} ([A_{r^*}^T, c], [0, 1]) = 0
\]

and so \( A_{r^*}^T = n - 1 \) and \( cr^* > 0 \) which is thus an extremal ray. \( \blacksquare \)

We can now prove Minkowski’s theorem. This version of the theorem assumes \( \text{lin } P = \emptyset \) or rank \( A = n \). When this is not the case we lack extremal points and the results less useful (from a practical point of view).

**Theorem 51 (Minkowski’s Theorem)** If \( P \neq \emptyset \) and rank \( A = n \), then

\[
P = \left\{ x \in \mathbb{R}^n \mid x = \sum_{k \in K} \lambda_k x^k + \sum_{j \in J} \mu_j r^j, \sum_{k \in K} \lambda_k = 1, \lambda_k \geq 0, k \in K \text{ and } \mu_j \geq 0, j \in J \right\},
\]

where \( \{x^k\}_{k \in K} \) is the set of extremal points of \( P \) and \( \{r^j\}_{j \in J} \) are the set of extremal rays of \( P \).

**Proof.** Let

\[
Q := \left\{ x \in \mathbb{R}^n \mid x = \sum_{k \in K} \lambda_k x^k + \sum_{j \in J} \mu_j r^j, \sum_{k \in K} \lambda_k = 1, \lambda_k \geq 0, k \in K \text{ and } \mu_j \geq 0, j \in J \right\}.
\]

Since \( x^k \in P \) for \( k \in K \) then \( \sum_{k \in K} \lambda_k x^k \in P \), for \( \sum_{k \in K} \lambda_k = 1, \lambda_k \geq 0 \). As \( \{r^j\}_{j \in J} \) are the set of extremal rays of \( P \) we have \( \sum_{j \in J} \mu_j r^j \in \text{rec } P \) for \( \mu_j \geq 0, j \in J \) hence

\[
\sum_{k \in K} \lambda_k x^k + \sum_{j \in J} \mu_j r^j \in P + \text{rec } P \subseteq P.
\]

Suppose \( Q \subseteq P \) so there exists \( y \in P \setminus Q \). Then the system

\[
y = \sum_{k \in K} \lambda_k x^k + \sum_{j \in J} \mu_j r^j, \sum_{k \in K} \lambda_k = 1, \lambda_k \geq 0, k \in K \text{ and } \mu_j \geq 0, j \in J
\]

has no solution in \((\lambda, \mu)\). By Farkas Lemma there exists \((c, c_0)\) such that \( c^T x^k \leq c_0 \) for \( k \in K \), \( c^T r^j \leq 0 \) for \( j \in J \) and \( c^T y > c_0 \).

Consider the LP

\[
\max \{ c^T x \mid x \in P \}.
\]

If it has a finite value it is obtained by an extremal point \( x^k \). However \( y \in P \) and

\[
c^T y > c_0 \geq c^T x^k, \text{ for } \forall k \in K
\]

a contradiction to optimality. On the other hand if the value is unbounded then there is an extremal ray \( r^j \) with \( c^T r^j > 0 \). But again \( c^T r^j \leq 0 \) for all \( j \in J \) is a contradiction. \( \blacksquare \)

**Example 52** For the example given we have \((1, 0, 0)\) and extremal point and one can easily check \((1, 0, -1)\) and \((0, 0, -1)\) are extremal rays. Our constraint set contains

\[
-x_1 \leq 0, \quad -x_2 \leq 0, \quad x_3 \leq 2
\]
which has rank 3. Hence rank \( A = 3 \). We know \((0, 0, 1)\) is an extremal point. Now note that in \( P^0 \) we have the constraints 

\[
x_1 + x_2 + x_3 = 0, x_1 + x_3 = 0, x_3 = 0
\]

which are satisfied by \( r^1 = (1, 0, -1) \). Since \[
\begin{pmatrix}
1 & 1 & 1 \\
1 & 0 & 1 \\
0 & 1 & 0
\end{pmatrix}
\]
has rank 2 then \( r^1 \) is an extremal ray. We have

\[
P = \left\{ x \in \mathbb{R}^n \mid x = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \lambda^1 + \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \mu^1 + \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \mu^2, \lambda^1 = 1, \mu^i \geq 0 \right\}.
\]

When we have rank \( A < n \) then the approach of [2] section A4 is better. This we describe briefly as the last topic.

The problem with the previous analysis when rank \( A < n \) is that all minimal faces are rays on the linearity space which are not too interesting from the basis theorem viewpoint. Thus we need to define the concept of a minimal proper face. Consider for now only the cone 

\[
C := \{ x \in \mathbb{R}^n \mid Ax \leq 0 \}.
\]

For \( I \subseteq \{1, \ldots, n\} \) denote 

\[
C(I) := \{ x \in C \mid a^i x = 0, i \in I \}.
\]

Thus \( C(\emptyset) = C \). Next denote 

\[
\mathcal{F} := \{ I \subseteq \{1, \ldots, n\} \mid \text{there exists } x \in C(I) \text{ with } a^i \bar{x} < 0, \text{ for all } i \notin I \}.
\]

Thus \( C(I^\infty) = C \) and if \( I^\infty \subseteq I \) we have \( C(I) \) a proper face of \( C \).

**Definition 53** We say \( I \) is maximal with respect to containment if \( I \not\subseteq J \) implies \( J \notin \mathcal{F} \). In this event we have \( C(I) \) a minimal proper face.

**Lemma 54** If \( \bar{x} \) is an inner point of \( C \) and \( z^1 \) is an element of \( C \) not in \( \text{lin} C \), then there is a \( \bar{\lambda} > 0 \) and an index set \( I \subseteq \{1, \ldots, n\} \) such that \( x - \bar{\lambda} z^1 \in C(I) \) and \( \dim C(I) < \dim C \).

**Proof.** Since \( a^i \bar{x} < 0 \) and for all \( i \in I^\subseteq \) and \( a^i z^1 < 0 \) for some \( i \in I^\subseteq \) there exists a sufficiently small \( \bar{\lambda} > 0 \) such that 

\[
a^i (\bar{x} - \bar{\lambda} z^1) = a^i \bar{x} - \bar{\lambda} a^i z^1 = 0, \text{ for } i \in I,
\]

\[
a^i (\bar{x} - \bar{\lambda} z^1) = a^i \bar{x} - \bar{\lambda} a^i z^1 < 0, \text{ for } i \notin I.
\]

[We take the smallest \( \lambda \) for which inequalities are driven to zero.] Thus \( \bar{x} - \bar{\lambda} z^1 \in C(I) \) (by definition). By construction \( I^\subseteq I \) and \( C(I) \) is a proper face implying \( \dim C(I) < \dim C \). 

The following justifies the definition of a minimal proper face. Recall that for \( \text{lin} C = \{ x \in \mathbb{R}^n \mid a^i x = 0, i \in \{1, \ldots, n\} \} = \{ x \in \mathbb{R}^n \mid Ax = 0 \} \).
Lemma 55 If $I \in \mathcal{F}$ is maximal then $\dim C(I) = \dim (\text{lin} \, C) + 1$.

Proof. If $I$ is maximal then for all $J \supseteq I$, that $J \notin \mathcal{F}$. Thus there does not exist a $\bar{x}$ such that $a^t \bar{x} = 0$ for all $i \in J$ and $a^t \bar{x} < 0$ for $i \notin J$ implying $C(J) = \text{lin} \, C$. To assume $C(J) \neq \text{lin} \, C$ we would have some $K \subseteq \{1, \ldots, n\} \setminus J$ such that $a^K \bar{x} < 0$, for $k \in K$ and $a^t \bar{x} = 0$ for all $i \in K^c := \{1, \ldots, n\} \setminus K$ implying $K^c \supseteq J \supseteq I$ is bigger with respect to containment.

By the definition of $\mathcal{F}$ if $k \in \{1, \ldots, n\} \setminus I$ we have an $\bar{x}$ such that $a^K \bar{x} < 0$ and $a^t \bar{x} = 0$ for $i \in I$. Thus

$$\begin{align*}
\text{rank} \left[ a^k, a^l \right]_{i \in I} &= \text{rank} \left[ a^l \right]_{i \in I} + 1 \\
\text{and so } \dim C(I \cup \{k\}) &= \dim \text{lin} \, C = \dim C(I) - 1.
\end{align*}$$

[It takes one extra constraint to define $\text{lin} \, C$ than $C(I)$].

We see from this that all faces of higher dimension than the minimal proper faces contain the subspace $\text{lin} \, C$. We first show the basis theorem for the cone

$$C := \{ x \in \mathbb{R}^n \mid Ax \leq 0 \}$$

and then extend to $P := \{ x \in \mathbb{R}^n \mid Ax \leq b \}$.

Theorem 56 (Minkowski’s theorem for conex) Let $C := \{ x \in \mathbb{R}^n \mid Ax \leq 0 \}$. Then there exists $x^1, \ldots, x^r \in C$ such that $C = \text{cone} \{ x^1, \ldots, x^r \}$ where each $x^i$ is either an element of the $\text{lin} \, C$ or an element of a minimal proper face of $C$.

Proof. Select one point $z^h \in C(I)$ such that $z^h$ is an inner point for each minimal proper face $C(I)$ of $C$. There is a finite number $\{z^1, \ldots, z^s\}$. The linearly subspace $\text{lin} \, C$ is spanned by a finite number of points $\{y^1, \ldots, y^r\}$. Then clearly

$$D := \text{cone} \{ z^1, \ldots, z^s, -y^1, \ldots, -y^r, y^1, \ldots, y^r \} \subseteq C.$$ 

We show the reverse containment by induction on the $\dim C$. Assume $\dim C = 0$ then $C = \{0\}$ and the result trivially follows as $\text{lin} \, C = \{0\}$. Assume the result for $\dim C \leq n$ and prove it for $\dim C = n + 1$. We take $\bar{x} \in C$ and show $x \in D$. If $x \in \text{lin} \, C$ we are finished so from now on assume $\bar{x} \notin \text{lin} \, C$.

If $\bar{x} \in C(I)$ for some proper face containing minimal proper faces then $\dim C(I) < \dim C$ and $\text{lin} \, C = \text{lin} \, C(I)$. Also any minimal proper face of $C(I)$ is also one of $C$. Therefore by the induction hypothesis we have some subset of $\{z^1, \ldots, z^s\}$ along with $\text{lin} \, C(I) = \{-y^1, \ldots, -y^r, y^1, \ldots, y^r\}$ with which we may represent $\bar{x}$ as a conic combination.

The last case is when $\bar{x} \notin C(I)$ for all proper face. Thus $\bar{x}$ must be an inner point. Apply Lemma 55 to get $\lambda > 0$ such that $\bar{x} - \lambda z^1 \in C(I)$ for $\dim C(I) \leq n$ (as $z^1$ is not in $\text{lin} \, C$). By the induction hypothesis we have subset of $\{z^1, \ldots, z^s\}$ along with $\text{lin} \, C(I) = \{-y^1, \ldots, -y^r, y^1, \ldots, y^r\}$ with which we may represent $\bar{x}$ as a conic combination.
combination i.e.

\[
\bar{x} - \bar{\lambda} z^1 = \sum_{i=1}^{s} \alpha_i z^i + \sum_{i=1}^{r} \beta_i y^i - \sum_{i=1}^{r} \gamma_i y^i
\]

or \( \bar{x} = \bar{\lambda} z^1 + \sum_{i=1}^{s} \alpha_i z^i + \sum_{i=1}^{r} \beta_i y^i - \sum_{i=1}^{r} \gamma_i y^i \)

\( \in \text{cone} \{ z^1, \ldots, z^s, -y^1, \ldots, -y^r, y^1, \ldots, y^r \} \).

\[\text{Theorem 57 (Finite Basis Theorem)}\]

If \( P := \{ x \in \mathbb{R}^n \mid Ax \leq b \} \neq \emptyset \) then there exists \( x^1, \ldots, x^r \in P \) such that

\[ P = \text{conv} \{ x^1, \ldots, x^q \} + \text{cone} \{ x^{q+1}, \ldots, x^r \} \]

and \( \text{lin} \ C = \text{cone} \{ x^{q+1}, \ldots, x^r \} \).

\[\text{Proof.}\]

Define

\[ C = \{ (x, x_0) \mid Ax - x_0 b \leq 0, x_0 \geq 0 \} . \]

Then by Minkowski’s theorem

\[ C = \text{cone} \{ (x^1, x_0^1), \ldots, (x^r, x_0^r) \} . \]

As we are taking a conic combination we may rescale each \( (x^i, x_0^i) \) with \( x_0^i \neq 0 \) so that \( x_0^i = 1 \) and otherwise \( x_0^i = 0 \). Reordering so that \( \{(x^1, 1), \ldots, (x^q, 1)\} \) (had \( x_0^i \neq 0 \)) and the remaining set \( \{(x^{q+1}, 0), \ldots, (x^r, 0)\} \) we have \((\bar{x}, 1) \in C \cap \{(x, 1) \mid x \in \mathbb{R}^n \} = \{(x, 1) \mid Ax - b \leq 0 \} \) given by

\[
(\bar{x}, 1) = \sum_{i=1}^{q} \bar{\lambda}_i (x^i, 1) + \sum_{i=q+1}^{r} \bar{\lambda}_i (x^i, 0)
\]

or \( \bar{x} = \sum_{i=1}^{q} \bar{\lambda}_i x^i + \sum_{i=q+1}^{r} \bar{\lambda}_i x^i \) where \( \sum_{i=1}^{q} \bar{\lambda}_i = 1 \) and \( \bar{\lambda}_i \geq 0 \).

\[\text{Example 58} \]

Consider

\[ C = \{ (x_1, x_2) \mid x_2 - x_1 \leq 0, -x_1 - x_2 \leq 0 \} . \]

Then \( \mathcal{F} = \{ \{1\}, \{2\}, \emptyset \} \)

Pick a point

\[ z^1 = (10, 10) \in C (\{1\}) := \{ (x_1, x_2) \in C \mid x_2 - x_1 = 0 \} \quad \text{and} \quad z^2 = (10, -10) \in C (\{2\}) := \{ (x_1, x_2) \in C \mid -x_1 - x_2 = 0 \} . \]

Note that \( \text{lin} C = \{ 0 \} \) and so for \( \bar{x} = (5, 1) \) we find

\[
a^1 \bar{x} - \lambda a^1 z^1 = -4 + \lambda 0 < 0 \]

and \( a^2 \bar{x} - \lambda a^2 z^1 = -6 + \lambda 20 \leq 0 \)
3.4. THE FINITE BASIS THEOREMS (UNBOUNDED POLYHEDRAL SETS)

\[ \bar{\lambda} = \frac{3}{10}. \] Then

\[ \bar{x} - \lambda z^1 = (5, 1) - \frac{3}{10} (10, 10) = (2, -2) \in C(\{2\}). \]

Thus

\[ \bar{x} - \lambda z^1 = (2, 2) = \frac{1}{5} (10, -10) \]

or

\[ \bar{x} = (5, 1) = \frac{3}{10} (10, 10) + \frac{1}{5} (10, -10) = \frac{3}{10} z^1 + \frac{1}{5} z^2. \]
Bibliography

