Chapter 2
Convex sets

A polyhedron $P \subseteq \mathbb{R}^n$ is a set of the form $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ for some $A \in \mathbb{R}^{m \times n}$ and some $b \in \mathbb{R}^m$. The set of feasible solutions of a linear program $\text{max} \{c^T x : Ax \leq b\}$ is a polyhedron. Polyhedra are convex sets. Convex sets are the main objects of study of this chapter.

Linear, affine, conic and convex hulls

Let $X \subseteq \mathbb{R}^n$ be a set of $n$-dimensional vectors. The linear hull, affine hull, conic hull and convex hull of $X$ are defined as follows.

Fig. 6 The convex hull of 7 points in $\mathbb{R}^2$. 
\begin{align*}
\text{lin.hull}(X) &= \{ \lambda_1 x_1 + \cdots + \lambda_t x_t \mid t \geq 0, x_1, \ldots, x_t \in X, \lambda_1, \ldots, \lambda_t \in \mathbb{R} \} \quad (7) \\
\text{affine.hull}(X) &= \{ \lambda_1 x_1 + \cdots + \lambda_t x_t \mid t \geq 1, \\
x_1, \ldots, x_t \in X, \sum_{i=1}^{t} \lambda_i = 1, \lambda_1, \ldots, \lambda_t \in \mathbb{R} \} \quad (8) \\
\text{cone}(X) &= \{ \lambda_1 x_1 + \cdots + \lambda_t x_t \mid t \geq 0, \\
x_1, \ldots, x_t \in X, \lambda_1, \ldots, \lambda_t \in \mathbb{R}_{\geq 0} \} \quad (9) \\
\text{conv}(X) &= \{ \lambda_1 x_1 + \cdots + \lambda_t x_t \mid t \geq 1, \\
x_1, \ldots, x_t \in X, \sum_{i=1}^{t} \lambda_i = 1, \lambda_1, \ldots, \lambda_t \in \mathbb{R}_{\geq 0} \} \quad (10)
\end{align*}

Fig. 7 Two points with their convex hull on the left and their affine hull on the right.

Fig. 8 Two points with their conic hull
Proposition 2.1. Let $X \subseteq \mathbb{R}^n$ and $x_0 \in X$. One has

$$\text{affine.hull}(X) = x_0 + \text{lin.hull}(X - x_0),$$

where for $u \in \mathbb{R}^n$ and $V \subseteq \mathbb{R}^n$, $u + V$ denotes the set $u + V = \{u + v \mid v \in V\}$.

Proof. Let $x_0$ be a point in $X$. Let $x \in \text{affine.hull}(X)$, i.e., there exists a natural number $t \geq 1$ and $\lambda_1, \ldots, \lambda_t \in \mathbb{R}$, with $x = \lambda_1 x_1 + \cdots + \lambda_t x_t$ and $\sum_{i=1}^t \lambda_i = 1$. Now

$$x = x_0 - x_0 + \lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_t x_t$$
$$= x_0 - \lambda_1 x_0 - \cdots - \lambda_t x_0 + \lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_t x_t$$
$$= x_0 + \lambda_1 (x_1 - x_0) + \cdots + \lambda_t (x_t - x_0)$$

This shows that $x \in x_0 + \text{lin.hull}(X - x_0)$. If on the other hand $x \in x_0 + \text{lin.hull}(X - x_0)$, then there exist $\lambda_1, \ldots, \lambda_t \in \mathbb{R}$ with $x = x_0 + \lambda_1 (x_1 - x_0) + \cdots + \lambda_t (x_t - x_0)$. With $\lambda_0 = 1 - \sum_{i=1}^t \lambda_i$ one has $\sum_{i=0}^t \lambda_i = 1$ and

$$x = x_0 + \lambda_1 (x_1 - x_0) + \cdots + \lambda_t (x_t - x_0)$$
$$= \lambda_0 x_0 + \cdots + \lambda_t x_t$$

and thus that $x \in \text{affine.hull}(X)$. \qed

Definition 2.1. The convex hull of two distinct points $u \neq v \in \mathbb{R}^n$ is called a line segment and is denoted by $\overline{uv}$.

Definition 2.2. A set $K \subseteq \mathbb{R}^n$ is convex if for each $u \neq v$, the line-segment $\overline{uv}$ is contained in $K$, $\overline{uv} \subseteq K$.

Fig. 9 The set on the left is convex, the set on the right is non-convex.

In other words, a set $K \subseteq \mathbb{R}^n$ is convex, if for each $u, v \in K$ and $\lambda \in [0, 1]$ the point $\lambda u + (1 - \lambda) v$ is also contained in $K$.

Theorem 2.1. Let $X \subseteq \mathbb{R}^n$ be a set of points. The convex hull, $\text{conv}(X)$, of $X$ is convex.
Proof. Let \( u \) and \( v \) be points in \( \text{conv}(X) \). This means that there exist a natural number \( t \geq 1 \), real numbers \( \alpha_i, \beta_i \geq 0 \), and points \( x_i \in X \), \( i = 1, \ldots, t \) with \( \sum_{i=1}^{t} \alpha_i = 1 \) and \( u = \sum_{i=1}^{t} \alpha_i x_i \) and \( \sum_{i=1}^{t} \beta_i = 1 \) and \( v = \sum_{i=1}^{t} \beta_i x_i \). For \( \lambda \in [0, 1] \) one has \( \lambda \alpha_i + (1 - \lambda) \beta_i \geq 0 \) for \( i = 1, \ldots, t \) and \( \sum_{i=1}^{t} \lambda \alpha_i + (1 - \lambda) \beta_i = 1 \). This shows that for \( \lambda \in [0, 1] \) one has
\[
\lambda u + (1 - \lambda) v = \sum_{i=1}^{t} \left( \lambda \alpha_i + (1 - \lambda) \beta_i \right) x_i \in \text{conv}(X),
\]
and therefore that \( \text{conv}(X) \) is convex.

Theorem 2.2. Let \( X \subseteq \mathbb{R}^n \) be a set of points. Each convex set \( K \) containing \( X \) also contains \( \text{conv}(X) \).

Proof. Let \( K \) be a convex set containing \( X \), and let \( x_1, \ldots, x_t \in X \) and \( \lambda_i \in \mathbb{R} \) with \( \lambda_i > 0 \), \( i = 1, \ldots, t \) and \( \sum_{i=1}^{t} \lambda_i = 1 \). We need to show that \( u = \sum_{i=1}^{t} \lambda_i x_i \) is contained in \( K \). This is true for \( t = 2 \) by the definition of convex sets. We argue by induction. Suppose that \( t \geq 3 \). If one of the \( \lambda_i \) is equal to 0, then one can represent \( u \) as a convex combination of \( t - 1 \) points in \( X \) and, by induction, \( u \in K \). Since \( t \geq 3 \), each \( \lambda_i > 0 \) and \( \sum_{i=1}^{t} \lambda_i = 1 \) one has \( 0 < \lambda_i < 1 \) for \( i = 1, \ldots, t \) and thus we can write
\[
u = \lambda_1 x_1 + (1 - \lambda_1) \sum_{i=2}^{t} \frac{\lambda_i}{1 - \lambda_1} x_i.
\]
One has \( \lambda_i / (1 - \lambda_1) > 0 \) and
\[
\sum_{i=2}^{t} \frac{\lambda_i}{1 - \lambda_1} = 1,
\]
which means that the point \( \sum_{i=2}^{t} \frac{\lambda_i}{1 - \lambda_1} x_i \) is in \( K \) by induction. Again, by the definition of convex sets, we conclude that \( u \) lies in \( K \).

Theorem 2.2 implies that \( \text{conv}(X) \) is the intersection of all convex sets containing \( X \), i.e.,
\[
\text{conv}(X) = \bigcap_{K \subseteq X \text{ convex}} K.
\]

Definition 2.3. A set \( C \subseteq \mathbb{R}^n \) is a cone, if it is convex and for each \( c \in C \) and each \( \lambda \in \mathbb{R}_{\geq 0} \) one has \( \lambda \cdot c \in C \).

Similarly to Theorem 2.1 and Theorem 2.2 one proves the following.

Theorem 2.3. For any \( X \subseteq \mathbb{R}^n \), the set \( \text{cone}(X) \) is a cone.

Theorem 2.4. Let \( X \subseteq \mathbb{R}^n \) be a set of points. Each cone containing \( X \) also contains \( \text{cone}(X) \).

These theorems imply that \( \text{cone}(X) \) is the intersection of all cones containing \( X \), i.e.,
\[
\text{cone}(X) = \bigcap_{C \subseteq X \text{ cone}} C.
\]
Radon's lemma and Carathéodory's theorem

**Theorem 2.5 (Radon's lemma).** Let $A \subseteq \mathbb{R}^n$ be a set of $n+2$ points. There exist disjoint subsets $A_1, A_2 \subseteq A$ with $\text{conv}(A_1) \cap \text{conv}(A_2) \neq \emptyset$.

*Proof.* Let $A = \{a_1, \ldots, a_{n+2}\}$. We embed these points into $\mathbb{R}^{n+1}$ by appending a 1 in the $n+1$-st component, i.e., we construct $A' = \{(a_1, 1), \ldots, (a_{n+2}, 1)\} \subseteq \mathbb{R}^{n+1}$.

The set $A'$ consists of $n+2$ vectors in $\mathbb{R}^{n+1}$. Those vectors are linearly dependent. Let

\[ 0 = \sum_{i=1}^{n+2} \lambda_i (a_i, 1) \]  

be a nontrivial linear representation of 0, i.e., not all $\lambda_i$ are 0. Furthermore, let $P = \{i : \lambda_i \geq 0, i = 1, \ldots, n+2\}$ and $N = \{i : \lambda_i < 0, i = 1, \ldots, n+2\}$. We claim that $\text{conv}(\{a_i : i \in P\}) \cap \text{conv}(\{a_i : i \in N\}) \neq \emptyset$.

It follows from (11) and the fact that the $n+1$-st component of the vectors is 1 that $\sum_{i \in P} \lambda_i = -\sum_{i \in N} \lambda_i = s > 0$. It follows also from (11) that

\[ \sum_{i \in P} \lambda_i a_i = \sum_{i \in N} -\lambda_i a_i. \]

The point $u = \sum_{i \in P} (\lambda_i/s) a_i = \sum_{i \in N} (-\lambda_i/s) a_i$ is contained in $\text{conv}(\{a_i : i \in P\}) \cap \text{conv}(\{a_i : i \in N\})$, implying the claim. $\square$

**Theorem 2.6 (Carathéodory's theorem).** Let $X \subseteq \mathbb{R}^n$, then for each $x \in \text{cone}(X)$ there exists a set $\bar{X} \subseteq X$ of cardinality at most $n$ such that $x \in \text{cone}(\bar{X})$. The vectors in $\bar{X}$ are linearly independent.

*Proof.* Let $x \in \text{cone}(X)$, then there exist $t \in \mathbb{N}_+\times, x_i \in X$ and $\lambda_i \geq 0, i = 1, \ldots, t$, with $x = \sum_{i=1}^t \lambda_i x_i$. Suppose that $t \in \mathbb{N}_+$ is minimal such that $x$ can be represented as above. We claim that $t \leq n$. If $t \geq n+1$, then the $x_i$ are linearly dependent. This means that there are $\mu_i \in \mathbb{R}$, not all equal to 0 with

\[ \sum_{i=1}^t \mu_i x_i = 0. \]  

By multiplying each $\mu_i$ in (12) with $-1$ if necessary, we can assume that at least one of the $\mu_i$ is strictly larger than 0. One has for each $\epsilon \in \mathbb{R}$

\[ x = \sum_{i=1}^t (\lambda_i - \epsilon \cdot \mu_i) x_i. \]  

(13)
Corollary 2.1 (Carathéodory’s theorem for convex hulls). Let \( X \subseteq \mathbb{R}^n \), then for each \( x \in \text{conv}(X) \) there exists a set \( X \subseteq \mathbb{R}^n \) of cardinality at most \( n + 1 \) such that \( x \in \text{conv}(X) \).

### Separation theorem and Farkas’ lemma

We recall a basic fact from analysis, see, e.g. [1, Theorem 4.4.1].

**Theorem 2.7.** Let \( X \subseteq \mathbb{R}^m \) be compact and \( f : X \to \mathbb{R} \) be continuous. Then \( f \) is bounded and there exist points \( x_1, x_2 \in X \) with \( f(x_1) = \sup \{ f(x) : x \in X \} \) and \( f(x_2) = \inf \{ f(x) : x \in X \} \).

**Theorem 2.8.** Let \( K \subseteq \mathbb{R}^n \) be a closed convex set and \( x^* \in \mathbb{R}^n \setminus K \), then there exists an inequality \( a^T x \geq \beta \) such that \( a^T y \geq \beta \) holds for all \( y \in K \) and \( a^T x^* < \beta \).

**Proof.** Since the mapping \( f(x) = \| x^* - x \| \) is continuous and since for any \( k \in K \), \( K \cap \{ x \in K : \| x^* - x \| \leq \| x^* - k \| \} \) is compact, there exists a point \( k^* \in K \) with minimal distance to \( x^* \). Consider the midpoint \( m = 1/2(k^* + x^*) \) on the line-segment \( k^*x^* \) and the hyperplane \( a^T x = \beta \) with \( \beta = a^T m \) and \( a = (k^* - x^*) \). Clearly, \( a^T x^* = \beta - 1/2 \| k^* - x^* \|^2 \) and \( a^T k^* = \beta + 1/2 \| k^* - x^* \|^2 \). Suppose that there exists a \( k' \in K \) with \( a^T k' \leq \beta \). The points \( \lambda k^* + (1 - \lambda)k' \), \( \lambda \in [0, 1] \) are in \( K \) by the convexity of \( K \), thus we can also assume that \( k' \) lies on the hyperplane, i.e., \( a^T k' = \beta \). This means that there exist a vector \( x' \) which is orthogonal to \( a \) and \( k' = m + x' \). The distance squared of a point \( \lambda k^* + (1 - \lambda)k' \) with \( \lambda \in [0, 1] \) to \( m \) is, by Pythagoras equal to

\[
\lambda^2 \| a \|^2 + (1 - \lambda)^2 \| x' \|^2.
\]

As a function of \( \lambda \), this is increasing at at \( \lambda = 1 \). Thus there exists a point on the line-segment \( \lambda x^* + (1 - \lambda)k \) which is closer to \( m \) than \( k^* \). This point is also closer to \( x^* \) than \( k^* \), which is a contradiction. Therefore \( a^T k > \beta \) for each \( k \in K \). 


Theorem 2.9 (Farkas’ lemma). Let $A \in \mathbb{R}^{m \times n}$ be a matrix and $b \in \mathbb{R}^m$ be a vector. The system $Ax = b$, $x \geq 0$ has a solution if and only if for all $\lambda \in \mathbb{R}^m$ with $\lambda^T A \geq 0$ one has $\lambda^T b \geq 0$.

Proof. Suppose that $x^* \in \mathbb{R}^n$ satisfies $Ax^* = b$ and let $\lambda \in \mathbb{R}^m$ with $\lambda^T A \geq 0$. Then $\lambda^T b = \lambda^T A x^* \geq 0$, since $\lambda^T A \geq 0$ and $x^* \geq 0$.

Now suppose that $Ax = b$, $x \geq 0$ does not have a solution. Then, with $X \subseteq \mathbb{R}^n$ being the set of column vectors of $A$, $b$ is not in $\text{cone}(X)$. The set $\text{cone}(X)$ is convex and closed, see exercise 5. Theorem 2.8 implies that there is an inequality $\lambda^T y \geq \beta$ for each $y \in \text{cone}(X)$ and $\lambda^T b < \beta$. Since for each $a \in X$ and each $\mu \geq 0$ one has $\mu \cdot a \in \text{cone}(X)$ and thus $\lambda^T (\mu \cdot a) > \beta$, it follows that $\lambda^T a \geq 0$ for each $a \in X$. Furthermore, since $0 \in \text{cone}(X)$ it follows that $0 \geq \beta$ and thus that $\lambda^T b < 0$.

Exercises

1) Let $\{C_i\}_{i \in I}$ be a family of convex subsets of $\mathbb{R}^n$. Show that the intersection $\bigcap_{i \in I} C_i$ is convex.
2) Show that the set of feasible solutions of a linear program is convex.
3) Prove Carathéodory’s Theorem for convex hulls, Corollary 2.1.
4) Let $A \in \mathbb{R}^{n \times n}$ be a non-singular matrix and let $a_1, \ldots, a_n \in \mathbb{R}^n$ be the columns of $A$. Show that $\text{cone}(\{a_1, \ldots, a_n\})$ is the polyhedron $P = \{y \in \mathbb{R}^n : A^{-1} y \geq 0\}$.

Show that $\text{cone}(\{a_1, \ldots, a_k\})$ for $k \leq n$ is the set $P_k = \{y \in \mathbb{R}^n : a_i^{-1} x \geq 0, i = 1, \ldots, k, a_k^{-1} x = 0, i = k + 1, \ldots, n\}$, where $a_i^{-1}$ denotes the $i$-th row of $A^{-1}$.
5) Prove that for a finite set $X \subseteq \mathbb{R}^n$ the conic hull $\text{cone}(X)$ is closed and convex. Hint: Use Carathéodory’s theorem and exercise 4.
6) Find a countably infinite set $X \subseteq \mathbb{R}^2$ such that $\text{cone}(X)$ is not closed. Are there any cones that are open?
7) Prove Theorem 2.3.
8) Prove Theorem 2.4.
9) Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a linear map.

a) Show that $f(K) = \{f(x) : x \in K\}$ is convex if $K$ is convex. Is the reverse also true?

b) For $X \subseteq \mathbb{R}^n$ arbitrary, prove that $\text{conv}(f(X)) = f(\text{conv}(X))$.
10) Using Theorem 2.9 prove the following variant of Farkas’ lemma: Let $A \in \mathbb{R}^{m \times n}$ be a matrix and $b \in \mathbb{R}^m$ be a vector. The system $Ax \leq b$, $x \in \mathbb{R}^n$ has a solution if and only if for all $\lambda \in \mathbb{R}^m_{\geq 0}$ with $\lambda^T A = 0$ one has $\lambda^T b \geq 0$.
11) Provide an example of a convex and closed set $K \subseteq \mathbb{R}^2$ and a linear objective function $c^T x$ such that $\min \{c^T x : x \in K\} > -\infty$ but there does not exist an $x^* \in K$ with $c^T x^* \leq c^T x$ for all $x \in K$.
12) Consider the vectors.
Let $A = \{x_1, \ldots, x_5\}$. Find two disjoint subsets $A_1, A_2 \subseteq A$ such that

$$\text{conv}(A_1) \cap \text{conv}(A_2) \neq \emptyset.$$ 

**Hint:** Recall the proof of Radon’s lemma

13) Consider the vectors

$$x_1 = \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix}, x_2 = \begin{pmatrix} 1 \\ 2 \\ 5 \end{pmatrix}, x_3 = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}, x_4 = \begin{pmatrix} 2 \\ 4 \\ 3 \end{pmatrix}, x_5 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

The vector

$$v = x_1 + 3x_2 + 2x_3 + x_4 + 3x_5 = \begin{pmatrix} 15 \\ 14 \\ 25 \end{pmatrix}$$

is a conic combination of the $x_i$.

Write $v$ as a conic combination using only three vectors of the $x_i$.

**Hint:** Recall the proof of Carathéodory’s theorem

**References**