Definition: A set $C \subseteq \mathbb{R}^n$ is called convex if for any $x, y \in C$ and $\lambda \in [0, 1]$, the point $\lambda x + (1 - \lambda)y$ belongs to $C$.

- The above definition is equivalent to saying that for any $x, y \in C$, the line segment $[x, y]$ is also in $C$. 

convex sets

nonconvex sets
Examples of Convex Sets

- **Lines:** A line in $\mathbb{R}^n$ is a set of the form
  \[ L = \{ z + td : t \in \mathbb{R} \}, \]
  where $z, d \in \mathbb{R}^n$ and $d \neq 0$.

- $[x, y], (x, y)$ for $x, y \in \mathbb{R}^n (x \neq y)$.

- $\emptyset, \mathbb{R}^n$.

- A hyperplane is a set of the form
  \[ H = \{ x \in \mathbb{R}^n : a^T x = b \} \quad (a \in \mathbb{R}^n \setminus \{0\}, b \in \mathbb{R}) \]

  The associated half-space is the set
  \[ H^- = \{ x \in \mathbb{R}^n : a^T x \leq b \} \]

Both hyperplanes and half-spaces are convex sets.
Convexity of Balls

Lemma. Let $c \in \mathbb{R}^n$ and $r > 0$. Then the open ball

$$B(c, r) = \{x \in \mathbb{R}^n : \|x - c\| < r\}$$

and the closed ball

$$B[c, r] = \{x \in \mathbb{R}^n : \|x - c\| \leq r\}$$

are convex.

Note that the norm is an arbitrary norm defined over $\mathbb{R}^n$.

Proof. In class
$l_1$, $l_2$ and $l_\infty$ balls

Figure: $l_1$, $l_2$ and $l_\infty$ balls in $\mathbb{R}^2$
Convexity of Ellipsoids

An **ellipsoid** is a set of the form

$$E = \{x \in \mathbb{R}^n : x^T Q x + 2b^T x + c \leq 0\},$$

where $Q \in \mathbb{R}^{n \times n}$ is positive semidefinite, $b \in \mathbb{R}^n$ and $c \in \mathbb{R}$.

**Lemma:** $E$ is convex.

**Proof.**

- Write $E$ as $E = \{x \in \mathbb{R}^n : f(x) \leq 0\}$ where $f(x) \equiv x^T Q x + 2b^T x + c$. 

  ▶ Take $x, y \in E$ and $\lambda \in [0, 1]$. Then $f(x) \leq 0$, $f(y) \leq 0$.

  ▶ The vector $z = \lambda x + (1-\lambda)y$ satisfies

    $$z^T Q z = \lambda^2 x^T Q x + (1-\lambda)^2 y^T Q y + 2\lambda(1-\lambda)x^T Q y.$$ 

    ▶ $x^T Q y \leq \|Q^{1/2} x\| \cdot \|Q^{1/2} y\| = \sqrt{x^T Q x} \cdot \sqrt{y^T Q y}$.

    ▶ $z^T Q z \leq \lambda x^T Q x + (1-\lambda)y^T Q y + 2\lambda(1-\lambda)\sqrt{x^T Q x} \cdot \sqrt{y^T Q y}$. 

    ▶ $f(z) = z^T Q z + 2b^T z + c \leq \lambda f(x) + (1-\lambda)f(y) \leq 0$, 

where $\lambda f(x) + (1-\lambda)f(y) \leq 0$. 

This shows that $E$ is convex.
Convexity of Ellipsoids

An ellipsoid is a set of the form

$$E = \{ x \in \mathbb{R}^n : x^T Q x + 2 b^T x + c \leq 0 \},$$

where $Q \in \mathbb{R}^{n \times n}$ is positive semidefinite, $b \in \mathbb{R}^n$ and $c \in \mathbb{R}$.

**Lemma:** $E$ is convex.

**Proof.**

- Write $E$ as $E = \{ x \in \mathbb{R}^n : f(x) \leq 0 \}$ where $f(x) \equiv x^T Q x + 2 b^T x + c$.
- Take $x, y \in E$ and $\lambda \in [0, 1]$. Then $f(x) \leq 0$, $f(y) \leq 0$. 
Convexity of Ellipsoids

An ellipsoid is a set of the form

\[ E = \{ x \in \mathbb{R}^n : x^T Q x + 2b^T x + c \leq 0 \}, \]

where \( Q \in \mathbb{R}^{n \times n} \) is positive semidefinite, \( b \in \mathbb{R}^n \) and \( c \in \mathbb{R} \).

**Lemma:** \( E \) is convex.

**Proof.**

- Write \( E \) as \( E = \{ x \in \mathbb{R}^n : f(x) \leq 0 \} \) where \( f(x) \equiv x^T Q x + 2b^T x + c \).
- Take \( x, y \in E \) and \( \lambda \in [0, 1] \). Then \( f(x) \leq 0, f(y) \leq 0 \).
- The vector \( z = \lambda x + (1 - \lambda) y \) satisfies
  \[ z^T Q z = \lambda^2 x^T Q x + (1 - \lambda)^2 y^T Q y + 2\lambda(1 - \lambda)x^T Q y. \]
Convexity of Ellipsoids

An **ellipsoid** is a set of the form

$$E = \{ x \in \mathbb{R}^n : x^T Q x + 2b^T x + c \leq 0 \},$$

where $Q \in \mathbb{R}^{n \times n}$ is positive semidefinite, $b \in \mathbb{R}^n$ and $c \in \mathbb{R}$.

**Lemma:** $E$ is convex.

**Proof.**

- Write $E$ as $E = \{ x \in \mathbb{R}^n : f(x) \leq 0 \}$ where $f(x) \equiv x^T Q x + 2b^T x + c$.
- Take $x, y \in E$ and $\lambda \in [0, 1]$. Then $f(x) \leq 0, f(y) \leq 0$.
- The vector $z = \lambda x + (1 - \lambda)y$ satisfies
  $$z^T Q z = \lambda^2 x^T Q x + (1 - \lambda)^2 y^T Q y + 2\lambda(1 - \lambda)x^T Q y.$$
- $x^T Q y \leq \|Q^{1/2}x\| \cdot \|Q^{1/2}y\| = \sqrt{x^T Q x} \sqrt{y^T Q y} \leq \frac{1}{2}(x^T Q x + y^T Q y)$
Convexity of Ellipsoids

An ellipsoid is a set of the form

\[ E = \{ x \in \mathbb{R}^n : x^T Q x + 2 b^T x + c \leq 0 \}, \]

where \( Q \in \mathbb{R}^{n \times n} \) is positive semidefinite, \( b \in \mathbb{R}^n \) and \( c \in \mathbb{R} \).

**Lemma:** \( E \) is convex.

**Proof.**
- Write \( E \) as \( E = \{ x \in \mathbb{R}^n : f(x) \leq 0 \} \) where \( f(x) \equiv x^T Q x + 2 b^T x + c \).
- Take \( x, y \in E \) and \( \lambda \in [0, 1] \). Then \( f(x) \leq 0, f(y) \leq 0 \).
- The vector \( z = \lambda x + (1 - \lambda) y \) satisfies
  \[ z^T Q z = \lambda^2 x^T Q x + (1 - \lambda)^2 y^T Q y + 2\lambda(1 - \lambda) x^T Q y. \]
- \( x^T Q y \leq \|Q^{1/2} x\| \cdot \|Q^{1/2} y\| = \sqrt{x^T Q x} \sqrt{y^T Q y} \leq \frac{1}{2} (x^T Q x + y^T Q y) \)
- \( z^T Q z \leq \lambda x^T Q x + (1 - \lambda) y^T Q y \)
Convexity of Ellipsoids

An 

ellipsoid

is a set of the form

\[ E = \{ x \in \mathbb{R}^n : x^T Q x + 2 b^T x + c \leq 0 \}, \]

where \( Q \in \mathbb{R}^{n \times n} \) is positive semidefinite, \( b \in \mathbb{R}^n \) and \( c \in \mathbb{R} \).

**Lemma:** \( E \) is convex.

**Proof.**

- Write \( E \) as \( E = \{ x \in \mathbb{R}^n : f(x) \leq 0 \} \) where \( f(x) \equiv x^T Q x + 2 b^T x + c \).
- Take \( x, y \in E \) and \( \lambda \in [0, 1] \). Then \( f(x) \leq 0, f(y) \leq 0 \).
- The vector \( z = \lambda x + (1 - \lambda) y \) satisfies
  \[ z^T Q z = \lambda^2 x^T Q x + (1 - \lambda)^2 y^T Q y + 2 \lambda (1 - \lambda) x^T Q y. \]
- \( x^T Q y \leq \| Q^{1/2} x \| \cdot \| Q^{1/2} y \| = \sqrt{x^T Q x} \sqrt{y^T Q y} \leq \frac{1}{2} (x^T Q x + y^T Q y) \)
- \( z^T Q z \leq \lambda x^T Q x + (1 - \lambda) y^T Q y \)

\[
  f(z) = z^T Q z + 2b^T z + c \\
  \leq \lambda x^T Q x + (1 - \lambda) y^T Q y + 2 b^T x + 2(1 - \lambda) b^T y + \lambda c + (1 - \lambda) c \\
  = \lambda f(x) + (1 - \lambda) f(y) \leq 0,
\]
Algebraic Operations Preserving Convexity

Lemma. Let $C_i \subseteq \mathbb{R}^n$ be a convex set for any $i \in I$ where $I$ is an index set (possibly infinite). Then the set $\bigcap_{i \in I} C_i$ is convex.

Proof. In class

Example: Consider the set

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

where $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. $P$ is called a convex polyhedron and it is indeed convex. Why?
Algebraic Operations Preserving Convexity

preservation under addition, cartesian product, forward and inverse linear mappings

Theorem.

1. Let $C_1, C_2, \ldots, C_k \subseteq \mathbb{R}^n$ be convex sets and let $\mu_1, \mu_2, \ldots, \mu_k \in \mathbb{R}$. Then the set $\mu_1 C_1 + \mu_2 C_2 + \ldots + \mu_k C_k$ is convex.

2. Let $C_i \subseteq \mathbb{R}^{k_i}, i = 1, \ldots, m$ be convex sets. Then the cartesian product $C_1 \times C_2 \times \cdots \times C_m = \{(x_1, x_2, \ldots, x_m) : x_i \in C_i, i = 1, 2, \ldots, m\}$ is convex.

3. Let $M \subseteq \mathbb{R}^n$ be a convex set and let $A \in \mathbb{R}^{m \times n}$. Then the set $A(M) = \{Ax : x \in M\}$ is convex.

4. Let $D \subseteq \mathbb{R}^m$ be convex and let $A \in \mathbb{R}^{m \times n}$. Then the set $A^{-1}(D) = \{x \in \mathbb{R}^n : Ax \in D\}$ is convex.
Convex Combinations

Given $m$ points $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m \in \mathbb{R}^n$, a convex combination of these $m$ points is a vector of the form $\lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2 + \cdots + \lambda_m \mathbf{x}_m$, where $\lambda_1, \lambda_2, \ldots, \lambda_m$ are nonnegative numbers satisfying $\lambda_1 + \lambda_2 + \ldots + \lambda_m = 1$.

- A convex set is defined by the property that any convex combination of two points from the set is also in the set.
- We will now show that a convex combination of any number of points from a convex set is in the set.
Convex Combinations

**Theorem.** Let $C \subseteq \mathbb{R}^n$ be a convex set and let $x_1, x_2, \ldots, x_m \in C$. Then for any $\lambda \in \Delta_m$, the relation $\sum_{i=1}^{m} \lambda_i x_i \in C$ holds.

**Proof by induction on $m$.**

- For $m = 1$ the result is obvious.
Convex Combinations

**Theorem.** Let $C \subseteq \mathbb{R}^n$ be a convex set and let $x_1, x_2, \ldots, x_m \in C$. Then for any $\lambda \in \Delta_m$, the relation $\sum_{i=1}^m \lambda_i x_i \in C$ holds.

**Proof by induction on $m$.**

- For $m = 1$ the result is obvious.
- The induction hypothesis is that for any $m$ vectors $x_1, x_2, \ldots, x_m \in C$ and any $\lambda \in \Delta_m$, the vector $\sum_{i=1}^m \lambda_i x_i$ belongs to $C$. We will now prove the theorem for $m + 1$ vectors.
Convex Combinations

Theorem. Let $C \subseteq \mathbb{R}^n$ be a convex set and let $x_1, x_2, \ldots, x_m \in C$. Then for any $\lambda \in \Delta_m$, the relation $\sum_{i=1}^{m} \lambda_i x_i \in C$ holds.

Proof by induction on $m$.

- For $m = 1$ the result is obvious.
- The induction hypothesis is that for any $m$ vectors $x_1, x_2, \ldots, x_m \in C$ and any $\lambda \in \Delta_m$, the vector $\sum_{i=1}^{m} \lambda_i x_i$ belongs to $C$. We will now prove the theorem for $m + 1$ vectors.
- Suppose that $x_1, x_2, \ldots, x_{m+1} \in C$ and that $\lambda \in \Delta_{m+1}$. We will show that $z \equiv \sum_{i=1}^{m+1} \lambda_i x_i \in C$.
- If $\lambda_{m+1} = 1$, then $z = x_{m+1} \in C$ and the result obviously follows.
**Theorem.** Let $C \subseteq \mathbb{R}^n$ be a convex set and let $x_1, x_2, \ldots, x_m \in C$. Then for any $\lambda \in \Delta_m$, the relation $\sum_{i=1}^m \lambda_i x_i \in C$ holds.

**Proof by induction on $m$.**

- For $m = 1$ the result is obvious.
- The induction hypothesis is that for any $m$ vectors $x_1, x_2, \ldots, x_m \in C$ and any $\lambda \in \Delta_m$, the vector $\sum_{i=1}^m \lambda_i x_i$ belongs to $C$. We will now prove the theorem for $m + 1$ vectors.
- Suppose that $x_1, x_2, \ldots, x_m+1 \in C$ and that $\lambda \in \Delta_{m+1}$. We will show that $z \equiv \sum_{i=1}^{m+1} \lambda_i x_i \in C$.
- If $\lambda_{m+1} = 1$, then $z = x_{m+1} \in C$ and the result obviously follows.
- If $\lambda_{m+1} < 1$ then
  
  $$z = \sum_{i=1}^m \lambda_i x_i + \lambda_{m+1} x_{m+1} = (1 - \lambda_{m+1}) \sum_{i=1}^m \frac{\lambda_i}{1 - \lambda_{m+1}} x_i + \lambda_{m+1} x_{m+1}.$$
Convex Combinations

**Theorem.** Let \( C \subseteq \mathbb{R}^n \) be a convex set and let \( x_1, x_2, \ldots, x_m \in C \). Then for any \( \lambda \in \Delta_m \), the relation \( \sum_{i=1}^{m} \lambda_i x_i \in C \) holds.

**Proof by induction on** \( m \).

- For \( m = 1 \) the result is obvious.
- The induction hypothesis is that for any \( m \) vectors \( x_1, x_2, \ldots, x_m \in C \) and any \( \lambda \in \Delta_m \), the vector \( \sum_{i=1}^{m} \lambda_i x_i \) belongs to \( C \). We will now prove the theorem for \( m + 1 \) vectors.
- Suppose that \( x_1, x_2, \ldots, x_{m+1} \in C \) and that \( \lambda \in \Delta_{m+1} \). We will show that \( z \equiv \sum_{i=1}^{m+1} \lambda_i x_i \in C \).
- If \( \lambda_{m+1} = 1 \), then \( z = x_{m+1} \in C \) and the result obviously follows.
- If \( \lambda_{m+1} < 1 \) then

\[
  z = \sum_{i=1}^{m} \lambda_i x_i + \lambda_{m+1} x_{m+1} = (1 - \lambda_{m+1}) \sum_{i=1}^{m} \frac{\lambda_i}{1 - \lambda_{m+1}} x_i + \lambda_{m+1} x_{m+1}.
\]

- \( v \in C \) and hence \( z = (1 - \lambda_{m+1}) v + \lambda_{m+1} x_{m+1} \in C \).
The Convex Hull

Definition. Let $S \subseteq \mathbb{R}^n$. The convex hull of $S$, denoted by $\text{conv}(S)$, is the set comprising all the convex combinations of vectors from $S$:

$$\text{conv}(S) \equiv \left\{ \sum_{i=1}^{k} \lambda_i \mathbf{x}_i : \mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k \in S, \lambda \in \Delta_k \right\}.$$
The Convex Hull

The convex hull \( \text{conv}(S) \) is “smallest” convex set containing \( S \).

**Lemma.** Let \( S \subseteq \mathbb{R}^n \). If \( S \subseteq T \) for some convex set \( T \), then \( \text{conv}(S) \subseteq T \).

**Proof.**

- Suppose that indeed \( S \subseteq T \) for some convex set \( T \).
The Convex Hull

The convex hull conv(S) is “smallest” convex set containing S.

**Lemma.** Let $S \subseteq \mathbb{R}^n$. If $S \subseteq T$ for some convex set $T$, then conv($S$) $\subseteq T$.

**Proof.**

- Suppose that indeed $S \subseteq T$ for some convex set $T$.
- To prove that conv($S$) $\subseteq T$, take $z \in$ conv($S$).

$\square$
The Convex Hull

The convex hull $\text{conv}(S)$ is “smallest” convex set containing $S$.

**Lemma.** Let $S \subseteq \mathbb{R}^n$. If $S \subseteq T$ for some convex set $T$, then $\text{conv}(S) \subseteq T$.

**Proof.**

- Suppose that indeed $S \subseteq T$ for some convex set $T$.
- To prove that $\text{conv}(S) \subseteq T$, take $z \in \text{conv}(S)$.
- There exist $x_1, x_2, \ldots, x_k \in S \subseteq T$ (where $k$ is a positive integer), and $\lambda \in \Delta_k$ such that $z = \sum_{i=1}^{k} \lambda_i x_i$. 


The Convex Hull

The convex hull $\text{conv}(S)$ is “smallest” convex set containing $S$.

Lemma. Let $S \subseteq \mathbb{R}^n$. If $S \subseteq T$ for some convex set $T$, then $\text{conv}(S) \subseteq T$.

Proof.

- Suppose that indeed $S \subseteq T$ for some convex set $T$.
- To prove that $\text{conv}(S) \subseteq T$, take $z \in \text{conv}(S)$.
- There exist $x_1, x_2, \ldots, x_k \in S \subseteq T$ (where $k$ is a positive integer), and $\lambda \in \Delta_k$ such that $z = \sum_{i=1}^{k} \lambda_i x_i$.
- Since $x_1, x_2, \ldots, x_k \in T$, it follows that $z \in T$, showing the desired result.
Carathéodory theorem

Theorem. Let $S \subseteq \mathbb{R}^n$ and let $x \in \text{conv}(S)$. Then there exist $x_1, x_2, \ldots, x_{n+1} \in S$ such that $x \in \text{conv}\left(\{x_1, x_2, \ldots, x_{n+1}\}\right)$, that is, there exist $\lambda \in \Delta_{n+1}$ such that

$$x = \sum_{i=1}^{n+1} \lambda_i x_i.$$ 

Proof. Let $x \in \text{conv}(S)$. Then $\exists x_1, x_2, \ldots, x_k \in S$ and $\lambda \in \Delta_k$ s.t.

$$x = \sum_{i=1}^{k} \lambda_i x_i.$$

We can assume that $\lambda_i > 0$ for all $i = 1, 2, \ldots, k$.

If $k \leq n+1$, the result is proven.

Otherwise, if $k \geq n+2$, then the vectors $x_2 - x_1, x_3 - x_1, \ldots, x_k - x_1$, being more than $n$ vectors in $\mathbb{R}^n$, are necessarily linearly dependent $\Rightarrow \exists \mu_2, \mu_3, \ldots, \mu_k$ not all zeros s.t.

$$\sum_{i=2}^{k} \mu_i (x_i - x_1) = 0.$$
Carathéodory theorem

Theorem. Let $S \subseteq \mathbb{R}^n$ and let $x \in \text{conv}(S)$. Then there exist $x_1, x_2, \ldots, x_{n+1} \in S$ such that $x \in \text{conv} \{x_1, x_2, \ldots, x_{n+1}\}$, that is, there exist $\lambda \in \Delta_{n+1}$ such that

$$x = \sum_{i=1}^{n+1} \lambda_i x_i.$$  

Proof.

Let $x \in \text{conv}(S)$. Then $\exists x_1, x_2, \ldots, x_k \in S$ and $\lambda \in \Delta_k$ s.t.

$$x = \sum_{i=1}^{k} \lambda_i x_i.$$
Carathéodory theorem

Theorem. Let $S \subseteq \mathbb{R}^n$ and let $x \in \text{conv}(S)$. Then there exist $x_1, x_2, \ldots, x_{n+1} \in S$ such that $x \in \text{conv}(\{x_1, x_2, \ldots, x_{n+1}\})$, that is, there exist $\lambda \in \Delta_{n+1}$ such that

$$ x = \sum_{i=1}^{n+1} \lambda_i x_i. $$

Proof.

- Let $x \in \text{conv}(S)$. Then $\exists x_1, x_2, \ldots, x_k \in S$ and $\lambda \in \Delta_k$ s.t.

$$ x = \sum_{i=1}^{k} \lambda_i x_i. $$

- We can assume that $\lambda_i > 0$ for all $i = 1, 2, \ldots, k$. 

Theorem. Let $S \subseteq \mathbb{R}^n$ and let $x \in \text{conv}(S)$. Then there exist $x_1, x_2, \ldots, x_{n+1} \in S$ such that $x \in \text{conv}\left(\{x_1, x_2, \ldots, x_{n+1}\}\right)$, that is, there exist $\lambda \in \Delta_{n+1}$ such that

$$x = \sum_{i=1}^{n+1} \lambda_i x_i.$$ 

Proof.

- Let $x \in \text{conv}(S)$. Then $\exists x_1, x_2, \ldots, x_k \in S$ and $\lambda \in \Delta_k$ s.t.

$$x = \sum_{i=1}^{k} \lambda_i x_i.$$ 

- We can assume that $\lambda_i > 0$ for all $i = 1, 2, \ldots, k$.

- If $k \leq n + 1$, the result is proven.
Carathéodory theorem

**Theorem.** Let \( S \subseteq \mathbb{R}^n \) and let \( \mathbf{x} \in \text{conv}(S) \). Then there exist \( \mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{n+1} \in S \) such that \( \mathbf{x} \in \text{conv} (\{ \mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{n+1} \}) \), that is, there exist \( \lambda \in \Delta_{n+1} \) such that

\[
\mathbf{x} = \sum_{i=1}^{n+1} \lambda_i \mathbf{x}_i.
\]

**Proof.**

- Let \( \mathbf{x} \in \text{conv}(S) \). Then \( \exists \mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k \in S \) and \( \lambda \in \Delta_k \) s.t.

\[
\mathbf{x} = \sum_{i=1}^{k} \lambda_i \mathbf{x}_i.
\]

- We can assume that \( \lambda_i > 0 \) for all \( i = 1, 2, \ldots, k \).
- If \( k \leq n + 1 \), the result is proven.
- Otherwise, if \( k \geq n + 2 \), then the vectors \( \mathbf{x}_2 - \mathbf{x}_1, \mathbf{x}_3 - \mathbf{x}_1, \ldots, \mathbf{x}_k - \mathbf{x}_1 \), being more than \( n \) vectors in \( \mathbb{R}^n \), are necessarily linearly dependent \( \Rightarrow \exists \mu_2, \mu_3, \ldots, \mu_k \) not all zeros s.t.

\[
\sum_{i=2}^{k} \mu_i (\mathbf{x}_i - \mathbf{x}_1) = 0.
\]
Proof of Carathéodory Theorem Contd.

- Defining \( \mu_1 = - \sum_{i=2}^{k} \mu_i \), we obtain that

\[
\sum_{i=1}^{k} \mu_i x_i = 0,
\]

Not all of the coefficients \( \mu_1, \mu_2, ..., \mu_k \) are zeros and \( \sum_{i=1}^{k} \mu_i = 0 \).

There exists an index \( \mu_i < 0 \). Let \( \alpha \in \mathbb{R}^+ \). Then

\[
x = \sum_{i=1}^{k} \lambda_i x_i + \alpha \sum_{i=1}^{k} \mu_i x_i = \sum_{i=1}^{k} \left( \lambda_i + \alpha \mu_i \right) x_i.
\]

(1)

We have \( \sum_{i=1}^{k} \left( \lambda_i + \alpha \mu_i \right) = 1 \), so (1) is a convex combination representation iff \( \lambda_i + \alpha \mu_i \geq 0 \) for all \( i = 1, ..., k \).

(2)

Since \( \lambda_i > 0 \) for all \( i \), it follows that (2) is satisfied for all \( \alpha \in [0, \varepsilon] \) where \( \varepsilon = \min_{i: \mu_i < 0} \{-\lambda_i \mu_i\} \).
Proof of Carathéodory Theorem Contd.

- Defining \( \mu_1 = - \sum_{i=2}^{k} \mu_i \), we obtain that

\[
\sum_{i=1}^{k} \mu_i x_i = 0,
\]

- Not all of the coefficients \( \mu_1, \mu_2, \ldots, \mu_k \) are zeros and \( \sum_{i=1}^{k} \mu_i = 0 \).
Proof of Carathéodory Theorem Contd.

- Defining $\mu_1 = -\sum_{i=2}^{k} \mu_i$, we obtain that
  \[
  \sum_{i=1}^{k} \mu_i x_i = 0,
  \]

- Not all of the coefficients $\mu_1, \mu_2, \ldots, \mu_k$ are zeros and $\sum_{i=1}^{k} \mu_i = 0$.
- There exists an index $i$ for which $\mu_i < 0$. Let $\alpha \in \mathbb{R}_+$. Then
  \[
  x = \sum_{i=1}^{k} \lambda_i x_i = \sum_{i=1}^{k} \lambda_i x_i + \alpha \sum_{i=1}^{k} \mu_i x_i = \sum_{i=1}^{k} \left( \lambda_i + \alpha \mu_i \right) x_i.
  \]
Proof of Carathéodory Theorem Contd.

- Defining \( \mu_1 = -\sum_{i=2}^{k} \mu_i \), we obtain that

\[
\sum_{i=1}^{k} \mu_i x_i = 0,
\]

- Not all of the coefficients \( \mu_1, \mu_2, \ldots, \mu_k \) are zeros and \( \sum_{i=1}^{k} \mu_i = 0 \).
- There exists an index \( i \) for which \( \mu_i < 0 \). Let \( \alpha \in \mathbb{R}_+ \). Then

\[
x = \sum_{i=1}^{k} \lambda_i x_i = \sum_{i=1}^{k} \lambda_i x_i + \alpha \sum_{i=1}^{k} \mu_i x_i = \sum_{i=1}^{k} (\lambda_i + \alpha \mu_i) x_i. \tag{1}
\]

- We have \( \sum_{i=1}^{k} (\lambda_i + \alpha \mu_i) = 1 \), so (1) is a convex combination representation iff

\[
\lambda_i + \alpha \mu_i \geq 0 \text{ for all } i = 1, \ldots, k. \tag{2}
\]
Proof of Carathéodory Theorem Contd.

- Defining $\mu_1 = -\sum_{i=2}^{k} \mu_i$, we obtain that

$$\sum_{i=1}^{k} \mu_i x_i = 0,$$

- Not all of the coefficients $\mu_1, \mu_2, \ldots, \mu_k$ are zeros and $\sum_{i=1}^{k} \mu_i = 0$.
- There exists an index $i$ for which $\mu_i < 0$. Let $\alpha \in \mathbb{R}^+$. Then

$$x = \sum_{i=1}^{k} \lambda_i x_i = \sum_{i=1}^{k} \lambda_i x_i + \alpha \sum_{i=1}^{k} \mu_i x_i = \sum_{i=1}^{k} (\lambda_i + \alpha \mu_i) x_i. \quad (1)$$

- We have $\sum_{i=1}^{k} (\lambda_i + \alpha \mu_i) = 1$, so (1) is a convex combination representation iff

$$\lambda_i + \alpha \mu_i \geq 0 \text{ for all } i = 1, \ldots, k. \quad (2)$$

- Since $\lambda_i > 0$ for all $i$, it follows that (2) is satisfied for all $\alpha \in [0, \varepsilon]$ where

$$\varepsilon = \min_{i: \mu_i < 0} \left\{-\frac{\lambda_i}{\mu_i}\right\}.$$
Proof of Carathéodory Theorem Contd.

- If we substitute \( \alpha = \varepsilon \), then (2) still holds, but \( \lambda_j + \varepsilon \mu_j = 0 \) for

\[
j \in \arg\min_{i: \mu_i < 0} \left\{ -\frac{\mu_i}{\lambda_i} \right\}.
\]

This means that we found a representation of \( x \) as a convex combination of \( k - 1 \) (or less) vectors.

This process can be carried on until a representation of \( x \) as a convex combination of no more than \( n + 1 \) vectors is derived.
Proof of Carathéodory Theorem Contd.

- If we substitute $\alpha = \varepsilon$, then (2) still holds, but $\lambda_j + \varepsilon \mu_j = 0$ for $j \in \text{argmin}_{i: \mu_i \leq 0} \left\{ -\frac{\mu_i}{\lambda_i} \right\}$.

- This means that we found a representation of $x$ as a convex combination of $k - 1$ (or less) vectors.
Proof of Carathéodory Theorem Contd.

- If we substitute $\alpha = \varepsilon$, then (2) still holds, but $\lambda_j + \varepsilon \mu_j = 0$ for $j \in \arg\min_{i: \mu_i < 0} \left\{ -\frac{\mu_i}{\lambda_i} \right\}$.

- This means that we found a representation of $x$ as a convex combination of $k - 1$ (or less) vectors.

- This process can be carried on until a representation of $x$ as a convex combination of no more than $n + 1$ vectors is derived.
Example

For \( n = 2 \), consider the four vectors

\[
x_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad x_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad x_3 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad x_4 = \begin{pmatrix} 2 \\ 2 \end{pmatrix},
\]

and let \( x \in \text{conv}(\{x_1, x_2, x_3, x_4\}) \) be given by

\[
x = \frac{1}{8}x_1 + \frac{1}{4}x_2 + \frac{1}{2}x_3 + \frac{1}{8}x_4 = \begin{pmatrix} \frac{13}{8} \\ \frac{11}{8} \end{pmatrix}.
\]

Find a representation of \( x \) as a convex combination of no more than 3 vectors.

In class
Convex Cones

- A set $S$ is called a cone if it satisfies the following property: for any $x \in S$ and $\lambda \geq 0$, the inclusion $\lambda x \in S$ is satisfied.
- The following lemma shows that there is a very simple and elegant characterization of convex cones.

**Lemma.** A set $S$ is a convex cone if and only if the following properties hold:

A. $x, y \in S \Rightarrow x + y \in S$.
B. $x \in S, \lambda \geq 0 \Rightarrow \lambda x \in S$.

Simple exercise
Examples of Convex Cones

- The convex polytope
  \[ C = \{ x \in \mathbb{R}^n : Ax \leq 0 \}, \]
  where \( A \in \mathbb{R}^{m \times n} \).

- Lorentz Cone
  The Lorenz cone, or *ice cream cone* is given by
  \[ L^n = \left\{ \begin{pmatrix} x \\ t \end{pmatrix} \in \mathbb{R}^{n+1} : \|x\| \leq t, x \in \mathbb{R}^n, t \in \mathbb{R} \right\}. \]

- nonnegative polynomials. set consisting of all possible coefficients of polynomials of degree \( n - 1 \) which are nonnegative over \( \mathbb{R} \):
  \[ K^n = \{ x \in \mathbb{R}^n : x_1 t^{n-1} + x_2 t^{n-2} + \ldots + x_{n-1} t + x_n \geq 0 \forall t \in \mathbb{R} \} \]
The Conic Hull

Definition. Given $m$ points $x_1, x_2, \ldots, x_m \in \mathbb{R}^n$, a conic combination of these $m$ points is a vector of the form $\lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_m x_m$, where $\lambda \in \mathbb{R}^m_+$. The definition of the conic hull is now quite natural.

Definition. Let $S \subseteq \mathbb{R}^n$. Then the conic hull of $S$, denoted by $\text{cone}(S)$ is the set comprising all the conic combinations of vectors from $S$:

$$\text{cone}(S) \equiv \{ k \sum_{i=1}^{k} \lambda_i x_i : x_1, x_2, \ldots, x_k \in S, \lambda \in \mathbb{R}^k_+ \}.$$ 

Similarly to the convex hull, the conic hull of a set $S$ is the smallest cone containing $S$.

Lemma. Let $S \subseteq \mathbb{R}^n$. If $S \subseteq T$ for some convex cone $T$, then $\text{cone}(S) \subseteq T$. 

Amir Beck "Introduction to Nonlinear Optimization" Lecture Slides - Convex Sets
The Conic Hull

**Definition.** Given $m$ points $x_1, x_2, \ldots, x_m \in \mathbb{R}^n$, a conic combination of these $m$ points is a vector of the form $\lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_m x_m$, where $\lambda \in \mathbb{R}_m^+$. 

The definition of the *conic hull* is now quite natural.

**Definition.** Let $S \subseteq \mathbb{R}^n$. Then the conic hull of $S$, denoted by $\text{cone}(S)$ is the set comprising all the conic combinations of vectors from $S$:

$$\text{cone}(S) \equiv \left\{ \sum_{i=1}^{k} \lambda_i x_i : x_1, x_2, \ldots, x_k \in S, \lambda \in \mathbb{R}_+^k \right\}.$$ 

Similarly to the convex hull, the conic hull of a set $S$ is the smallest cone containing $S$.

**Lemma.** Let $S \subseteq \mathbb{R}^n$. If $S \subseteq T$ for some convex cone $T$, then $\text{cone}(S) \subseteq T$. 

Amir Beck “Introduction to Nonlinear Optimization” Lecture Slides - Convex Sets 18 / 32
Representation Theorem for Conic Hulls

a similar result to Carathéodory theorem

**Conic Representation Theorem.** Let $S \subseteq \mathbb{R}^n$ and let $x \in \text{cone}(S)$. Then there exist $k$ linearly independent vector $x_1, x_2, \ldots, x_k \in S$ such that $x \in \text{cone} (\{x_1, x_2, \ldots, x_k\})$, that is, there exist $\lambda \in \mathbb{R}^k_+$ such that

$$x = \sum_{i=1}^{k} \lambda_i x_i.$$ 

In particular, $k \leq n$.

Proof very similar to the proof of Carathéodory theorem. See page 107 of the book for the proof.
Basic Feasible Solutions

▶ Consider the convex polyhedron.

\[ P = \{ x \in \mathbb{R}^n : Ax = b, x \geq 0 \}, \quad (A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m) \]

▶ the rows of \( A \) are assumed to be linearly independent.

▶ The above is a standard formulation of the constraints of a linear programming problem.

Definition. \( \bar{x} \) is a basic feasible solution (abbreviated bfs) of \( P \) if the columns of \( A \) corresponding to the indices of the positive values of \( \bar{x} \) are linearly independent.
Basic Feasible Solutions

Consider the convex polyhedron.

\[ P = \{ x \in \mathbb{R}^n : Ax = b, x \geq 0 \}, \quad (A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m) \]

the rows of \( A \) are assumed to be linearly independent.

The above is a standard formulation of the constraints of a linear programming problem.

Definition. \( \bar{x} \) is a basic feasible solution (abbreviated bfs) of \( P \) if the columns of \( A \) corresponding to the indices of the positive values of \( \bar{x} \) are linearly independent.

Example. Consider the linear system:

\[
\begin{align*}
    x_1 + x_2 + x_3 &= 6 \\
    x_2 + x_4 &= 3 \\
    x_1, x_2, x_3, x_4 &\geq 0.
\end{align*}
\]

Find all the basic feasible solutions. In class
Existence of bfs's

**Theorem.** Let \( P = \{ x \in \mathbb{R}^n : Ax = b, x \geq 0 \} \), where \( A \in \mathbb{R}^{m \times n} \) and \( b \in \mathbb{R}^m \). If \( P \neq \emptyset \), then it contains at least one bfs.
Existence of bfs’s

**Theorem.** Let \( P = \{ x \in \mathbb{R}^n : Ax = b, x \geq 0 \} \), where \( A \in \mathbb{R}^{m \times n} \) and \( b \in \mathbb{R}^m \). If \( P \neq \emptyset \), then it contains at least one bfs.

**Proof.**

- \( P \neq \emptyset \Rightarrow b \in \text{cone}(\{a_1, a_2, \ldots, a_n\}) \) where \( a_i \) denotes the \( i \)-th column of \( A \).
Existence of bfs’s

Theorem. Let \( P = \{ x \in \mathbb{R}^n : Ax = b, x \geq 0 \} \), where \( A \in \mathbb{R}^{m \times n} \) and \( b \in \mathbb{R}^m \).

If \( P \neq \emptyset \), then it contains at least one bfs.

Proof.

- \( P \neq \emptyset \Rightarrow b \in \text{cone}(\{a_1, a_2, \ldots, a_n\}) \) where \( a_i \) denotes the \( i \)-th column of \( A \).
- By the conic representation theorem, there exist indices \( i_1 < i_2 < \ldots < i_k \) and \( k \) numbers \( y_{i_1}, y_{i_2}, \ldots, y_{i_k} \geq 0 \) such that \( b = \sum_{j=1}^{k} y_{i_j} a_{i_j} \) and \( a_{i_1}, a_{i_2}, \ldots, a_{i_k} \) are linearly independent.
Existence of bfs's

Theorem. Let $P = \{ x \in \mathbb{R}^n : Ax = b, x \geq 0 \}$, where $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. If $P \neq \emptyset$, then it contains at least one bfs.

Proof.

- $P \neq \emptyset \Rightarrow b \in \text{cone}(\{a_1, a_2, \ldots, a_n\})$ where $a_i$ denotes the $i$-th column of $A$.
- By the conic representation theorem, there exist indices $i_1 < i_2 < \ldots < i_k$ and $k$ numbers $y_{i_1}, y_{i_2}, \ldots, y_{i_k} \geq 0$ such that $b = \sum_{j=1}^{k} y_{i_j} a_{i_j}$ and $a_{i_1}, a_{i_2}, \ldots, a_{i_k}$ are linearly independent.
- Denote $\bar{x} = \sum_{j=1}^{k} y_{i_j} e_{i_j}$. Then obviously $\bar{x} \geq 0$ and in addition

$$A\bar{x} = \sum_{j=1}^{k} y_{i_j} Ae_{i_j} = \sum_{j=1}^{k} y_{i_j} a_{i_j} = b.$$
Existence of bfs’s

**Theorem.** Let $P = \{x \in \mathbb{R}^n : Ax = b, x \geq 0\}$, where $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. If $P \neq \emptyset$, then it contains at least one bfs.

**Proof.**

- $P \neq \emptyset \Rightarrow b \in \text{cone}(\{a_1, a_2, \ldots, a_n\})$ where $a_i$ denotes the $i$-th column of $A$.
- By the conic representation theorem, there exist indices $i_1 < i_2 < \ldots < i_k$ and $k$ numbers $y_{i_1}, y_{i_2}, \ldots, y_{i_k} \geq 0$ such that $b = \sum_{j=1}^{k} y_{i_j} a_{i_j}$ and $a_{i_1}, a_{i_2}, \ldots, a_{i_k}$ are linearly independent.
- Denote $\bar{x} = \sum_{j=1}^{k} y_{i_j} e_{i_j}$. Then obviously $\bar{x} \geq 0$ and in addition

$$A\bar{x} = \sum_{j=1}^{k} y_{i_j} Ae_{i_j} = \sum_{j=1}^{k} y_{i_j} a_{i_j} = b.$$ 

- Therefore, $\bar{x}$ is contained in $P$ and the columns of $A$ corresponding to the indices of the positive components of $\bar{x}$ are linearly independent, meaning that $P$ contains a bfs.
Topological Properties of Convex Sets

Theorem. Let \( C \subseteq \mathbb{R}^n \) be a convex set. Then \( \text{cl}(C) \) is a convex set.

Proof.

- Let \( x, y \in \text{cl}(C) \) and let \( \lambda \in [0, 1] \).
Topological Properties of Convex Sets

**Theorem.** Let $C \subseteq \mathbb{R}^n$ be a convex set. Then $\text{cl}(C)$ is a convex set.

**Proof.**

- Let $x, y \in \text{cl}(C)$ and let $\lambda \in [0, 1]$.
- There exist sequences $\{x_k\}_{k \geq 0} \subseteq C$ and $\{y_k\}_{k \geq 0} \subseteq C$ for which $x_k \to x$ and $y_k \to y$ as $k \to \infty$. 
Topological Properties of Convex Sets

**Theorem.** Let $C \subseteq \mathbb{R}^n$ be a convex set. Then $\text{cl}(C)$ is a convex set.

**Proof.**
- Let $x, y \in \text{cl}(C)$ and let $\lambda \in [0, 1]$.
- There exist sequences $\{x_k\}_{k \geq 0} \subseteq C$ and $\{y_k\}_{k \geq 0} \subseteq C$ for which $x_k \to x$ and $y_k \to y$ as $k \to \infty$.
- $(\ast)$ $\lambda x_k + (1 - \lambda)y_k \in C$ for any $k \geq 0$. 
Theorem. Let $C \subseteq \mathbb{R}^n$ be a convex set. Then $\text{cl}(C)$ is a convex set.

Proof.

- Let $x, y \in \text{cl}(C)$ and let $\lambda \in [0, 1]$.
- There exist sequences $\{x_k\}_{k \geq 0} \subseteq C$ and $\{y_k\}_{k \geq 0} \subseteq C$ for which $x_k \to x$ and $y_k \to y$ as $k \to \infty$.
- $(\ast) \quad \lambda x_k + (1 - \lambda)y_k \in C$ for any $k \geq 0$.
- $(\ast\ast) \quad \lambda x_k + (1 - \lambda)y_k \to \lambda x + (1 - \lambda)y$. 
Theorem. Let $C \subseteq \mathbb{R}^n$ be a convex set. Then $\text{cl}(C)$ is a convex set.

Proof.

- Let $x, y \in \text{cl}(C)$ and let $\lambda \in [0, 1]$.
- There exist sequences $\{x_k\}_{k \geq 0} \subseteq C$ and $\{y_k\}_{k \geq 0} \subseteq C$ for which $x_k \to x$ and $y_k \to y$ as $k \to \infty$.
- $\lambda x_k + (1 - \lambda) y_k \in C$ for any $k \geq 0$.
- $\lambda x_k + (1 - \lambda) y_k \to \lambda x + (1 - \lambda) y$.
- $(*) + (**)$ implies $\lambda x + (1 - \lambda) y \in \text{cl}(C)$. 
The Line Segment Principle

Theorem. Let $C$ be a convex set and assume that $\text{int}(C) \neq \emptyset$. Suppose that $x \in \text{int}(C)$ and $y \in \text{cl}(C)$. Then $(1 - \lambda)x + \lambda y \in \text{int}(C)$ for any $\lambda \in [0, 1)$.

Proof. ▶ There exists $\varepsilon > 0$ such that $B(x, \varepsilon) \subseteq C$. ▶ Let $z = (1 - \lambda)x + \lambda y$. We will show that $B(z, (1 - \lambda)\varepsilon) \subseteq C$. ▶ Let $w \in B(z, (1 - \lambda)\varepsilon)$. Since $y \in \text{cl}(C)$, $\exists w_1 \in C$ s.t. $\|w_1 - y\| < (1 - \lambda)\varepsilon - \|w - z\|\lambda$. (3) ▶ Set $w_2 = (1 - \lambda)(w - \lambda w_1)$. Then $\|w_2 - x\| = \|w - \lambda w_1 + \lambda x - \lambda x\| \leq (1 - \lambda)(\|w - z\| + \lambda\|w_1 - y\|) < \varepsilon$, ▶ Hence, since $B(x, \varepsilon) \subseteq C$, it follows that $w_2 \in C$. Finally, since $w = \lambda w_1 + (1 - \lambda)w_2$ with $w_1, w_2 \in C$, we have that $w \in C$.
The Line Segment Principle

**Theorem.** Let $C$ be a convex set and assume that $\text{int}(C) \neq \emptyset$. Suppose that $x \in \text{int}(C)$ and $y \in \text{cl}(C)$. Then $(1 - \lambda)x + \lambda y \in \text{int}(C)$ for any $\lambda \in [0, 1)$.

**Proof.**
- There exists $\varepsilon > 0$ such that $B(x, \varepsilon) \subseteq C$. 


The Line Segment Principle

**Theorem.** Let \( C \) be a convex set and assume that \( \text{int}(C) \neq \emptyset \). Suppose that \( x \in \text{int}(C) \) and \( y \in \text{cl}(C) \). Then \( (1 - \lambda)x + \lambda y \in \text{int}(C) \) for any \( \lambda \in [0, 1) \).

**Proof.**
- There exists \( \varepsilon > 0 \) such that \( B(x, \varepsilon) \subseteq C \).
- Let \( z = (1 - \lambda)x + \lambda y \). We will show that \( B(z, (1 - \lambda)\varepsilon) \subseteq C \).
The Line Segment Principle

**Theorem.** Let $C$ be a convex set and assume that $\text{int}(C) \neq \emptyset$. Suppose that $x \in \text{int}(C)$ and $y \in \text{cl}(C)$. Then $(1 - \lambda)x + \lambda y \in \text{int}(C)$ for any $\lambda \in [0, 1)$.

**Proof.**

- There exists $\varepsilon > 0$ such that $B(x, \varepsilon) \subseteq C$.
- Let $z = (1 - \lambda)x + \lambda y$. We will show that $B(z, (1 - \lambda)\varepsilon) \subseteq C$.
- Let $w \in B(z, (1 - \lambda)\varepsilon)$. Since $y \in \text{cl}(C)$, $\exists w_1 \in C$ s.t.

$$\|w_1 - y\| < \frac{(1 - \lambda)\varepsilon - \|w - z\|}{\lambda}.$$  \hspace{1cm} (3)
The Line Segment Principle

**Theorem.** Let $C$ be a convex set and assume that $\text{int}(C) \neq \emptyset$. Suppose that $x \in \text{int}(C)$ and $y \in \text{cl}(C)$. Then $(1 - \lambda)x + \lambda y \in \text{int}(C)$ for any $\lambda \in [0, 1)$.

**Proof.**

- There exists $\varepsilon > 0$ such that $B(x, \varepsilon) \subseteq C$.
- Let $z = (1 - \lambda)x + \lambda y$. We will show that $B(z, (1 - \lambda)\varepsilon) \subseteq C$.
- Let $w \in B(z, (1 - \lambda)\varepsilon)$. Since $y \in \text{cl}(C)$, there exists $w_1 \in C$ s.t.

  $$\|w_1 - y\| < \frac{(1 - \lambda)\varepsilon - \|w - z\|}{\lambda}. \quad (3)$$

- Set $w_2 = \frac{1}{1 - \lambda}(w - \lambda w_1)$. Then

  $$\|w_2 - x\| = \|\frac{w - \lambda w_1}{1 - \lambda} - x\| = \frac{1}{1 - \lambda} \|(w - z) + \lambda(y - w_1)\| \leq \frac{1}{1 - \lambda} (\|w - z\| + \lambda\|w_1 - y\|) \overset{(3)}{<} \varepsilon,$$
The Line Segment Principle

**Theorem.** Let $C$ be a convex set and assume that $\text{int}(C) \neq \emptyset$. Suppose that $x \in \text{int}(C)$ and $y \in \text{cl}(C)$. Then $(1 - \lambda)x + \lambda y \in \text{int}(C)$ for any $\lambda \in [0, 1)$.

**Proof.**
- There exists $\varepsilon > 0$ such that $B(x, \varepsilon) \subseteq C$.
- Let $z = (1 - \lambda)x + \lambda y$. We will show that $B(z, (1 - \lambda)\varepsilon) \subseteq C$.
- Let $w \in B(z, (1 - \lambda)\varepsilon)$. Since $y \in \text{cl}(C)$, $\exists w_1 \in C$ s.t. $\|w_1 - y\| < \frac{(1 - \lambda)\varepsilon - \|w - z\|}{\lambda}$.

\[\|w_1 - y\| < \frac{(1 - \lambda)\varepsilon - \|w - z\|}{\lambda}.\] (3)

- Set $w_2 = \frac{1}{1 - \lambda}(w - \lambda w_1)$. Then

\[\|w_2 - x\| = \left\|\frac{w - \lambda w_1}{1 - \lambda} - x\right\| = \frac{1}{1 - \lambda} \|(w - z) + \lambda(y - w_1)\|
\leq \frac{1}{1 - \lambda} (\|w - z\| + \lambda\|w_1 - y\|) \quad \text{(3)} < \varepsilon,

- Hence, since $B(x, \varepsilon) \subseteq C$, it follows that $w_2 \in C$. Finally, since $w = \lambda w_1 + (1 - \lambda)w_2$ with $w_1, w_2 \in C$, we have that $w \in C$.  

Amir Beck  
“Introduction to Nonlinear Optimization” Lecture Slides - Convex Sets
Convexity of the Interior

**Theorem.** Let $C \subseteq \mathbb{R}^n$ be a convex set. Then $\text{int}(C)$ is convex.

**Proof.**

- If $\text{int}(C) = \emptyset$, then the theorem is obviously true.
- Otherwise, let $x_1, x_2 \in \text{int}(C)$, and let $\lambda \in (0, 1)$.
- By the LSP, $\lambda x_1 + (1 - \lambda) x_2 \in \text{int}(C)$, establishing the convexity of $\text{int}(C)$. 
Lemma. Let C be a convex set with a nonempty interior. Then
1. \( \text{cl}(\text{int}(C)) = \text{cl}(C) \).
2. \( \text{int}(\text{cl}(C)) = \text{int}(C) \).

Proof of 1. Obviously, \( \text{cl}(\text{int}(C)) \subseteq \text{cl}(C) \) holds.
To prove that opposite, let \( x \in \text{cl}(C) \), \( y \in \text{int}(C) \).
Then \( x_k = \frac{1}{k} y + \left(1 - \frac{1}{k}\right) x \in \text{int}(C) \) for any \( k \geq 1 \).
Since \( x \) is the limit (as \( k \to \infty \)) of the sequence \( \{x_k\} \), it follows that \( x \in \text{cl}(\text{int}(C)) \).
For the proof of 2, see pages 109, 110 of the book for the proof of Lemma 6.30(b).
Combination of Closure and Interior

**Lemma.** Let $C$ be a convex set with a nonempty interior. Then

1. $\text{cl}(\text{int}(C)) = \text{cl}(C)$.
2. $\text{int}(\text{cl}(C)) = \text{int}(C)$.

**Proof of 1.**
- Obviously, $\text{cl}(\text{int}(C)) \subseteq \text{cl}(C)$ holds.
Combination of Closure and Interior

**Lemma.** Let $C$ be a convex set with a nonempty interior. Then

1. $\text{cl}(\text{int}(C)) = \text{cl}(C)$.
2. $\text{int}(\text{cl}(C)) = \text{int}(C)$.

**Proof of 1.**

- Obviously, $\text{cl}(\text{int}(C)) \subseteq \text{cl}(C)$ holds.
- To prove that opposite, let $x \in \text{cl}(C), y \in \text{int}(C)$. 


Combination of Closure and Interior

Lemma. Let $C$ be a convex set with a nonempty interior. Then

1. $\text{cl}(\text{int}(C)) = \text{cl}(C)$.
2. $\text{int}(\text{cl}(C)) = \text{int}(C)$.

Proof of 1.

- Obviously, $\text{cl}(\text{int}(C)) \subseteq \text{cl}(C)$ holds.
- To prove that opposite, let $x \in \text{cl}(C)$, $y \in \text{int}(C)$.
- Then $x_k = \frac{1}{k}y + (1 - \frac{1}{k})x \in \text{int}(C)$ for any $k \geq 1$. 
Combination of Closure and Interior

**Lemma.** Let $C$ be a convex set with a nonempty interior. Then

1. $\text{cl}(\text{int}(C)) = \text{cl}(C)$.
2. $\text{int}(\text{cl}(C)) = \text{int}(C)$.

**Proof of 1.**

- Obviously, $\text{cl}(\text{int}(C)) \subseteq \text{cl}(C)$ holds.
- To prove that opposite, let $x \in \text{cl}(C)$, $y \in \text{int}(C)$.
- Then $x_k = \frac{1}{k}y + (1 - \frac{1}{k})x \in \text{int}(C)$ for any $k \geq 1$.
- Since $x$ is the limit (as $k \to \infty$) of the sequence $\{x_k\}_{k \geq 1} \subseteq \text{int}(C)$, it follows that $x \in \text{cl}(\text{int}(C))$. 

For the proof of 2, see pages 109,110 of the book for the proof of Lemma 6.30(b).
Combination of Closure and Interior

Lemma. Let $C$ be a convex set with a nonempty interior. Then

1. $\text{cl}(\text{int}(C)) = \text{cl}(C)$.
2. $\text{int}(\text{cl}(C)) = \text{int}(C)$.

Proof of 1.

- Obviously, $\text{cl}(\text{int}(C)) \subseteq \text{cl}(C)$ holds.
- To prove that opposite, let $x \in \text{cl}(C), y \in \text{int}(C)$.
- Then $x_k = \frac{1}{k}y + \left(1 - \frac{1}{k}\right)x \in \text{int}(C)$ for any $k \geq 1$.
- Since $x$ is the limit (as $k \to \infty$) of the sequence $\{x_k\}_{k \geq 1} \subseteq \text{int}(C)$, it follows that $x \in \text{cl}(\text{int}(C))$.

For the proof of 2, see pages 109,110 of the book for the proof of Lemma 6.30(b).
Compactness of the Convex Hull of Convex Sets

**Theorem.** Let $S \subseteq \mathbb{R}^n$ be a compact set. Then $\text{conv}(S)$ is compact.

**Proof.**
Compactness of the Convex Hull of Convex Sets

**Theorem.** Let $S \subseteq \mathbb{R}^n$ be a compact set. Then $\text{conv}(S)$ is compact.

**Proof.**

1. $\exists M > 0$ such that $\|x\| \leq M$ for any $x \in S$. 

...
Compactness of the Convex Hull of Convex Sets

**Theorem.** Let \( S \subseteq \mathbb{R}^n \) be a compact set. Then \( \text{conv}(S) \) is compact.

**Proof.**

- \( \exists M > 0 \) such that \( \|x\| \leq M \) for any \( x \in S \).
- Let \( y \in \text{conv}(S) \). Then there exist \( x_1, x_2, \ldots, x_{n+1} \in S \) and \( \lambda \in \Delta_{n+1} \) for which \( y = \sum_{i=1}^{n+1} \lambda_i x_i \) and therefore

\[
\|y\| = \left\| \sum_{i=1}^{n+1} \lambda_i x_i \right\| \leq \sum_{i=1}^{n+1} \lambda_i \|x_i\| \leq M \sum_{i=1}^{n+1} \lambda_i = M,
\]

establishing the boundedness of \( \text{conv}(S) \).
Compactness of the Convex Hull of Convex Sets

Theorem. Let $S \subseteq \mathbb{R}^n$ be a compact set. Then $\text{conv}(S)$ is compact.

Proof.

- $\exists M > 0$ such that $\|x\| \leq M$ for any $x \in S$.
- Let $y \in \text{conv}(S)$. Then there exist $x_1, x_2, \ldots, x_{n+1} \in S$ and $\lambda \in \Delta_{n+1}$ for which $y = \sum_{i=1}^{n+1} \lambda_i x_i$, and therefore

$$
\|y\| = \left\| \sum_{i=1}^{n+1} \lambda_i x_i \right\| \leq \sum_{i=1}^{n+1} \lambda_i \|x_i\| \leq M \sum_{i=1}^{n+1} \lambda_i = M,
$$

establishing the boundedness of $\text{conv}(S)$.

- To prove the closedness of $\text{conv}(S)$, let $\{y_k\}_{k \geq 1} \subseteq \text{conv}(S)$ be a sequence converging to $y \in \mathbb{R}^n$. 

Compactness of the Convex Hull of Convex Sets

**Theorem.** Let \( S \subseteq \mathbb{R}^n \) be a compact set. Then \( \text{conv}(S) \) is compact.

**Proof.**

- **\( \exists M > 0 \) such that \( \|x\| \leq M \) for any \( x \in S \).**
- **Let \( y \in \text{conv}(S) \).** Then there exist \( x_1, x_2, \ldots, x_{n+1} \in S \) and \( \lambda \in \Delta_{n+1} \) for which \( y = \sum_{i=1}^{n+1} \lambda_i x_i \) and therefore

\[
\|y\| = \left\| \sum_{i=1}^{n+1} \lambda_i x_i \right\| \leq \sum_{i=1}^{n+1} \lambda_i \|x_i\| \leq M \sum_{i=1}^{n+1} \lambda_i = M,
\]

establishing the boundedness of \( \text{conv}(S) \).

- **To prove the closedness of \( \text{conv}(S) \), let \( \{y_k\}_{k \geq 1} \subseteq \text{conv}(S) \) be a sequence converging to \( y \in \mathbb{R}^n \).**
- **There exist \( x_1^k, x_2^k, \ldots, x_{n+1}^k \in S \) and \( \lambda^k \in \Delta_{n+1} \) such that**

\[
y_k = \sum_{i=1}^{n+1} \lambda_i^k x_i^k. \tag{4}
\]
Proof Contd.

- By the compactness of $S$ and $\Delta_{n+1}$, it follows that 
  \[
  \{(\lambda^k, x^k_1, x^k_2, \ldots, x^k_{n+1})\}_{k \geq 1}
  \]
  has a convergent subsequence

  \[
  \{(\lambda^{k_j}, x^{k_j}_1, x^{k_j}_2, \ldots, x^{k_j}_{n+1})\}_{j \geq 1}
  \]
  whose limit will be denoted by

  \[
  (\lambda, x_1, x_2, \ldots, x_{n+1})
  \]

  with $\lambda \in \Delta_{n+1}, x_1, x_2, \ldots, x_{n+1} \in S$.
Proof Contd.

By the compactness of $S$ and $\Delta_{n+1}$, it follows that 
\[
\{(\lambda^k, x_1^k, x_2^k, \ldots, x_{n+1}^k)\}_{k \geq 1}
\]
has a convergent subsequence 
\[
\{(\lambda^{k_j}, x_1^{k_j}, x_2^{k_j}, \ldots, x_{n+1}^{k_j})\}_{j \geq 1}
\]
whose limit will be denoted by 
\[
(\lambda, x_1, x_2, \ldots, x_{n+1})
\]
with $\lambda \in \Delta_{n+1}$, $x_1, x_2, \ldots, x_{n+1} \in S$.

Taking the limit $j \to \infty$ in 
\[
y_{k_j} = \sum_{i=1}^{n+1} \lambda^{k_j}_i x^{k_j}_i,
\]
we obtain that 
\[
y = \sum_{i=1}^{n+1} \lambda_i x_i \in \text{conv}(S)
\]
as required.
By the compactness of $S$ and $\Delta_{n+1}$, it follows that
\[ \{(\lambda^k, x_1^k, x_2^k, \ldots, x_{n+1}^k)\}_{k \geq 1} \]
has a convergent subsequence
\[ \{(\lambda^{k_j}, x_1^{k_j}, x_2^{k_j}, \ldots, x_{n+1}^{k_j})\}_{j \geq 1} \]
whose limit will be denoted by
\[ (\lambda, x_1, x_2, \ldots, x_{n+1}) \]
with $\lambda \in \Delta_{n+1}$, $x_1, x_2, \ldots, x_{n+1} \in S$.

Taking the limit $j \to \infty$ in
\[ y_{k_j} = \sum_{i=1}^{n+1} \lambda_i^{k_j} x_i^{k_j}, \]
we obtain that $y = \sum_{i=1}^{n+1} \lambda_i x_i \in \text{conv}(S)$ as required.

Example: $S = \{(0,0)^T\} \cup \{(x,y)^T : xy \geq 1\}$
Closedness of the Conic Hull of a Finite Set

Theorem. Let $a_1, a_2, \ldots, a_k \in \mathbb{R}^n$. Then $\text{cone}(\{a_1, a_2, \ldots, a_k\})$ is closed.

Proof.
Closedness of the Conic Hull of a Finite Set

Theorem. Let \( a_1, a_2, \ldots, a_k \in \mathbb{R}^n \). Then \( \text{cone}(\{a_1, a_2, \ldots, a_k\}) \) is closed.

Proof.

By the conic representation theorem, each element of \( \text{cone}(\{a_1, a_2, \ldots, a_k\}) \) can be represented as a conic combination of a linearly independent subset of \( \{a_1, a_2, \ldots, a_k\} \).
Closedness of the Conic Hull of a Finite Set

Theorem. Let \( a_1, a_2, \ldots, a_k \in \mathbb{R}^n \). Then cone(\( \{a_1, a_2, \ldots, a_k\} \)) is closed.

Proof.

\begin{itemize}
  \item By the conic representation theorem, each element of cone(\( \{a_1, a_2, \ldots, a_k\} \)) can be represented as a conic combination of a linearly independent subset of \( \{a_1, a_2, \ldots, a_k\} \).
  \item Therefore, if \( S_1, S_2, \ldots, S_N \) are all the subsets of \( \{a_1, a_2, \ldots, a_k\} \) comprising linearly independent vectors, then
  \[\text{cone}(\{a_1, a_2, \ldots, a_k\}) = \bigcup_{i=1}^{N} \text{cone}(S_i).\]
\end{itemize}
Closedness of the Conic Hull of a Finite Set

**Theorem.** Let \( a_1, a_2, \ldots, a_k \in \mathbb{R}^n \). Then \( \text{cone}(\{a_1, a_2, \ldots, a_k\}) \) is closed.

**Proof.**

- By the conic representation theorem, each element of \( \text{cone}(\{a_1, a_2, \ldots, a_k\}) \) can be represented as a conic combination of a linearly independent subset of \( \{a_1, a_2, \ldots, a_k\} \).
- Therefore, if \( S_1, S_2, \ldots, S_N \) are all the subsets of \( \{a_1, a_2, \ldots, a_k\} \) comprising linearly independent vectors, then

\[
\text{cone}(\{a_1, a_2, \ldots, a_k\}) = \bigcup_{i=1}^{N} \text{cone}(S_i).
\]

- It is enough to show that \( \text{cone}(S_i) \) is closed for any \( i \in \{1, 2, \ldots, N\} \). Indeed, let \( i \in \{1, 2, \ldots, N\} \). Then

\[
S_i = \{b_1, b_2, \ldots, b_m\},
\]

where \( b_1, b_2, \ldots, b_m \) are linearly independent.
Closedness of the Conic Hull of a Finite Set

Theorem. Let $a_1, a_2, \ldots, a_k \in \mathbb{R}^n$. Then $\text{cone}(\{a_1, a_2, \ldots, a_k\})$ is closed.

Proof.

- By the conic representation theorem, each element of $\text{cone}(\{a_1, a_2, \ldots, a_k\})$ can be represented as a conic combination of a linearly independent subset of $\{a_1, a_2, \ldots, a_k\}$.
- Therefore, if $S_1, S_2, \ldots, S_N$ are all the subsets of $\{a_1, a_2, \ldots, a_k\}$ comprising linearly independent vectors, then
  \[
  \text{cone}(\{a_1, a_2, \ldots, a_k\}) = \bigcup_{i=1}^{N} \text{cone}(S_i).
  \]

- It is enough to show that $\text{cone}(S_i)$ is closed for any $i \in \{1, 2, \ldots, N\}$. Indeed, let $i \in \{1, 2, \ldots, N\}$. Then
  \[
  S_i = \{b_1, b_2, \ldots, b_m\},
  \]
  where $b_1, b_2, \ldots, b_m$ are linearly independent.
- $\text{cone}(S_i) = \{B y : y \in \mathbb{R}^m_+\}$, where $B$ is the matrix whose columns are $b_1, b_2, \ldots, b_m$. 
Proof Contd.

- Suppose that $x_k \in \text{cone}(S_i)$ for all $k \geq 1$ and that $x_k \to \bar{x}$.

\[ y_k = (B^T B)^{-1} B^T x_k. \]

Taking the limit as $k \to \infty$ in the last equation, we obtain that $y_k \to \bar{y}$ where

\[ \bar{y} = (B^T B)^{-1} B^T \bar{x}. \]

Thus, taking the limit in (5), we conclude that $\bar{x} = B \bar{y}$ with $\bar{y} \in \mathbb{R}^m^+$, and hence $\bar{x} \in \text{cone}(S_i)$. 

Amir Beck “Introduction to Nonlinear Optimization” Lecture Slides - Convex Sets
Proof Contd.

- Suppose that $x_k \in \text{cone}(S_i)$ for all $k \geq 1$ and that $x_k \to \bar{x}$.
- $\exists y_k \in \mathbb{R}^m_+$ such that

$$x_k = By_k.$$  \hspace{1cm} (5)

- $\bar{y} \in \mathbb{R}^m_+$.
- Thus, taking the limit in (5), we conclude that $\bar{x} = B\bar{y}$ with $\bar{y} \in \mathbb{R}^m_+$, and hence $\bar{x} \in \text{cone}(S_i)$. 

Proof Contd.

- Suppose that $x_k \in \text{cone}(S_i)$ for all $k \geq 1$ and that $x_k \to \bar{x}$.
- $\exists y_k \in \mathbb{R}^m_+$ such that $x_k = B y_k$. \hfill (5)

\[ y_k = (B^T B)^{-1} B^T x_k. \]

- Taking the limit as $k \to \infty$ in the last equation, we obtain that $y_k \to \bar{y}$ where $\bar{y} = (B^T B)^{-1} B^T \bar{x}$. 

$\bar{y} \in \mathbb{R}^m_+$. 

Thus, taking the limit in (5), we conclude that $\bar{x} = B \bar{y}$ with $\bar{y} \in \mathbb{R}^m_+$, and hence $\bar{x} \in \text{cone}(S_i)$. 

Amir Beck “Introduction to Nonlinear Optimization” Lecture Slides - Convex Sets 29 / 32
Proof Contd.

- Suppose that \( x_k \in \text{cone}(S_i) \) for all \( k \geq 1 \) and that \( x_k \to \bar{x} \).
- \( \exists y_k \in \mathbb{R}^m_+ \) such that
  \[
  x_k = B y_k. 
  \]  
  \( \quad \)\( (5) \)

- \[
  y_k = (B^T B)^{-1} B^T x_k. 
  \]

- Taking the limit as \( k \to \infty \) in the last equation, we obtain that \( y_k \to \bar{y} \) where
  \[
  \bar{y} = (B^T B)^{-1} B^T \bar{x}. 
  \]
- \( \bar{y} \in \mathbb{R}^m_+ \).
Proof Contd.

- Suppose that $x_k \in \text{cone}(S_i)$ for all $k \geq 1$ and that $x_k \to \bar{x}$.

- $\exists y_k \in \mathbb{R}_+^m$ such that

\[ x_k = B y_k. \]  \hfill (5)

- $y_k = (B^T B)^{-1} B^T x_k$.

- Taking the limit as $k \to \infty$ in the last equation, we obtain that $y_k \to \bar{y}$ where

\[ \bar{y} = (B^T B)^{-1} B^T \bar{x}. \]

- $\bar{y} \in \mathbb{R}_+^m$.

- Thus, taking the limit in (5), we conclude that $\bar{x} = B \bar{y}$ with $\bar{y} \in \mathbb{R}_+^m$, and hence $\bar{x} \in \text{cone}(S_i)$. 

-Amir Beck “Introduction to Nonlinear Optimization” Lecture Slides - Convex Sets
Extreme Points

Definition. Let $S \subseteq \mathbb{R}^n$ be a convex set. A point $x \in S$ is called an extreme point of $S$ if there do not exist $x_1, x_2 \in S (x_1 \neq x_2)$ and $\lambda \in (0, 1)$, such that $x = \lambda x_1 + (1 - \lambda)x_2$.

- The set of extreme point is denoted by $\text{ext}(S)$.
- For example, the set of extreme points of a convex polytope consists of all its vertices.
Equivalence Between bfs’s and Extreme Points

**Theorem.** Let $P = \{x \in \mathbb{R}^n : Ax = b, x \geq 0\}$, where $A \in \mathbb{R}^{m \times n}$ has linearly independent rows and $b \in \mathbb{R}^m$. The $\bar{x}$ is a basic feasible solution of $P$ if and only if it is an extreme point of $P$.

Theorem 6.34 in the book.
Krein-Milman Theorem

Theorem. Let $S \subseteq \mathbb{R}^n$ be a compact convex set. Then

$$S = \text{conv}(\text{ext}(S)).$$