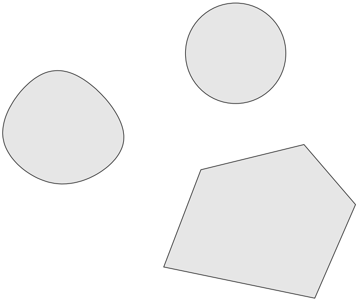
Lecture 6 - Convex Sets

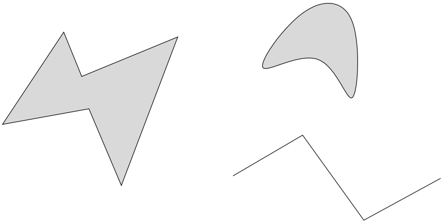
Definition A set $C \subseteq \mathbb{R}^n$ is called convex if for any $\mathbf{x}, \mathbf{y} \in C$ and $\lambda \in [0, 1]$, the point $\lambda \mathbf{x} + (1 - \lambda)\mathbf{y}$ belongs to C.

- The above definition is equivalent to saying that for any $\mathbf{x}, \mathbf{y} \in C$, the line segment $[\mathbf{x}, \mathbf{y}]$ 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
$$L = \{\mathbf{z} + t\mathbf{d} : t \in \mathbb{R}\},$$
where $\mathbf{z}, \mathbf{d} \in \mathbb{R}^n$ and $\mathbf{d} \neq \mathbf{0}$.
- $[\mathbf{x}, \mathbf{y}], (\mathbf{x}, \mathbf{y})$ for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n (\mathbf{x} \neq \mathbf{y})$.
- $\emptyset, \mathbb{R}^n$.
- A hyperplane is a set of the form
$$H = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{a}^T\mathbf{x} = b\} \quad (\mathbf{a} \in \mathbb{R}^n \backslash \{\mathbf{0}\}, b \in \mathbb{R})$$
The associated half-space is the set
$$H^- = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{a}^T\mathbf{x} \leq b\}$$
Both hyperplanes and half-spaces are convex sets.

Convexity of Balls

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

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

and the closed ball

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

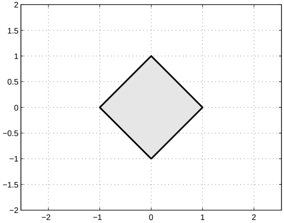
are convex.

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

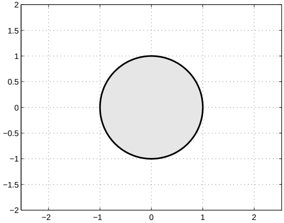
Proof. In class

l_1, l_2 and l_∞ balls

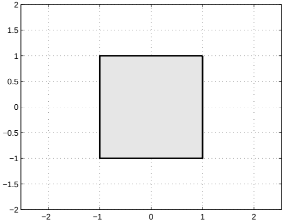
l_1



l_2



l_∞



Convexity of Ellipsoids

An ellipsoid is a set of the form

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

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

Lemma: E is convex.

Proof.

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

$$\begin{aligned} f(\mathbf{z}) &= \mathbf{z}^T\mathbf{Q}\mathbf{z} + 2\mathbf{b}^T\mathbf{z} + c \\ &\le \lambda\mathbf{x}^T\mathbf{Q}\mathbf{x} + (1-\lambda)\mathbf{y}^T\mathbf{Q}\mathbf{y} + 2\lambda\mathbf{b}^T\mathbf{x} + 2(1-\lambda)\mathbf{b}^T\mathbf{y} + \lambda c + (1-\lambda)c \\ &= \lambda f(\mathbf{x}) + (1-\lambda)f(\mathbf{y}) \le 0, \end{aligned}$$

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

compare this definition of convex polyhedron with that on page 20. (I will help you connect the two seemingly different formulations in the HW.)

Example: Consider the set

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

where $\mathbf{A} \in \mathbb{R}^{m \times n}$ and $\mathbf{b} \in \mathbb{R}^m$. 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
$$C_1 \times C_2 \times \cdots \times C_m = \{(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m) : \mathbf{x}_i \in C_i, i = 1, 2, \ldots, m\}$$
is convex.
3. Let $M \subseteq \mathbb{R}^n$ be a convex set and let $\mathbf{A} \in \mathbb{R}^{m \times n}$. Then the set
$$\mathbf{A}(M) = \{\mathbf{A}\mathbf{x} : \mathbf{x} \in M\}$$
is convex.
4. Let $D \subseteq \mathbb{R}^m$ be convex and let $\mathbf{A} \in \mathbb{R}^{m \times n}$. Then the set
$$\mathbf{A}^{-1}(D) = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{A}\mathbf{x} \in D\}$$
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 $\lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2 + \cdots + \ldots + \lambda_m \mathbf{x}_m$, 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 $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m \in C$. Then for any $\boldsymbol{\lambda} \in \Delta_m$, the relation $\sum_{i=1}^m \lambda_i \mathbf{x}_i \in C$ holds.

Proof by induction on m.

- For $m = 1$ the result is obvious.
- The induction hypothesis is that for any m vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m \in C$ and any $\boldsymbol{\lambda} \in \Delta_m$, the vector $\sum_{i=1}^m \lambda_i \mathbf{x}_i$ belongs to C. We will now prove the theorem for $m + 1$ vectors.
- Suppose that $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{m+1} \in C$ and that $\boldsymbol{\lambda} \in \Delta_{m+1}$. We will show that $\mathbf{z} \equiv \sum_{i=1}^{m+1} \lambda_i \mathbf{x}_i \in C$.
- If $\lambda_{m+1} = 1$, then $\mathbf{z} = \mathbf{x}_{m+1} \in C$ and the result obviously follows.
- If $\lambda_{m+1} < 1$ then

$$\mathbf{z} = \sum_{i=1}^m \lambda_i \mathbf{x}_i + \lambda_{m+1}\mathbf{x}_{m+1} = (1 - \lambda_{m+1}) \underbrace{\sum_{i=1}^{m} \frac{\lambda_i}{1 - \lambda_{m+1}} \mathbf{x}_i}_{\mathbf{v}} + \lambda_{m+1}\mathbf{x}_{m+1}.$$

- $\mathbf{v} \in C$ and hence $\mathbf{z} = (1 - \lambda_{m+1})\mathbf{v} + \lambda_{m+1}\mathbf{x}_{m+1} \in C$.

note that the m weights defining v add up to 1

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:

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

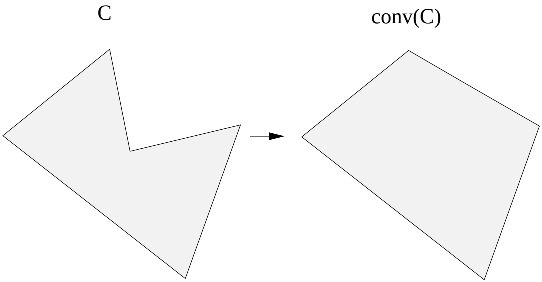


Figure: A nonconvex set and its convex hull

The Convex Hull

In fact, we can characterize the convex hull of S as the intersection of all the convex sets that contain S. (And this would imply the following lemma immediately.)

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 $\mathbf{z} \in$ conv(S).
- There exist $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k \in S \subseteq T$ (where k is a positive integer), and $\boldsymbol{\lambda} \in \Delta_k$ such that $\mathbf{z} = \sum_{i=1}^k \lambda_i \mathbf{x}_i$.
- Since $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k \in T$, it follows that $\mathbf{z} \in T$, showing the desired result.

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 $\boldsymbol{\lambda} \in \Delta_{n+1}$ 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 $\boldsymbol{\lambda} \in \Delta_k$ 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.

$$\textstyle\sum_{i=2}^{k} \mu_i (\mathbf{x}_i - \mathbf{x}_1) = \mathbf{0}.$$

Proof of Carathéodory Theorem Contd.

After (1), the idea is to tune alpha so that all lambda_i+alpha*mu_i are non-negative and one of them is exactly zero.

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

$$\sum_{i=1}^k \mu_i \mathbf{x}_i = \mathbf{0},$$

- 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

$$\mathbf{x} = \sum_{i=1}^k \lambda_i \mathbf{x}_i = \sum_{i=1}^k \lambda_i \mathbf{x}_i + \alpha \sum_{i=1}^k \mu_i \mathbf{x}_i = \sum_{i=1}^k (\lambda_i + \alpha\mu_i)\mathbf{x}_i. \tag{1}$$

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

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

.

- 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 $j \in \operatorname*{argmin}_{i:\mu_i<0} \left\{ -\frac{\mu_i}{\lambda_i} \right\}$.
- This means that we found a representation of $\mathbf{x}$ as a convex combination of $k-1$ (or less) vectors.
- This process can be carried on until a representation of $\mathbf{x}$ as a convex combination of no more than $n+1$ vectors is derived.

The concept of simplex is relevant here. A simplex is a generalization of points (0-d), line segments (1-d), triangles (2-D), tetrahedrons (3-d). Precisely, a (k-)simplex is the convex hull of k+1 points x0, x1,..., xk in R^n so that x1-x0, ..., xk-x0 are linearly independent (so k<=n). In this case, we also call x0,...,xk the vertices of the simplex. The Caratheodory theorem can be restated as follows: any point x in conv(S) is contained in a simplex with vertices in S.

Example

For $n = 2$, consider the four vectors

$$\mathbf{x}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \mathbf{x}_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \mathbf{x}_3 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \mathbf{x}_4 = \begin{pmatrix} 2 \\ 2 \end{pmatrix},$$

and let $\mathbf{x} \in \text{conv}(\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4\})$ be given by

$$\mathbf{x} = \frac{1}{8}\mathbf{x}_1 + \frac{1}{4}\mathbf{x}_2 + \frac{1}{2}\mathbf{x}_3 + \frac{1}{8}\mathbf{x}_4 = \begin{pmatrix} \frac{13}{8} \\ \frac{11}{8} \end{pmatrix}.$$

Find a representation of $\mathbf{x}$ as a convex combination of no more than 3 vectors.

In class

Every point in the square conv(x1,x2,x3,x4), one of the following must be true:
(i) it is one of the four points x1,...,x4,
(ii) it lies on the line segment of xi, xj (there are 6 such line segments)
(iii) it lies on a triangle with vertices xi,xj,xk (there are 4 such triangles)

Convex Cones

- A set S is called a cone if it satisfies the following property: for any $\mathbf{x} \in S$ and $\lambda \geq 0$, the inclusion $\lambda \mathbf{x} \in S$ 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. $\mathbf{x}, \mathbf{y} \in S \Rightarrow \mathbf{x} + \mathbf{y} \in S$.

B. $\mathbf{x} \in S, \lambda \geq 0 \Rightarrow \lambda \mathbf{x} \in S$.

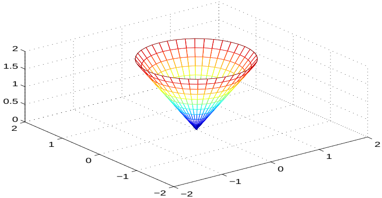
Simple exercise

Note: a cone always contains the origin. (A convex set doesn't have this property.)

means S is both a convex set and a cone

Examples of Convex Cones

- The convex polytope
$$C = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{A}\mathbf{x} \leq \mathbf{0}\},$$
where $\mathbf{A} \in \mathbb{R}^{m \times n}$.
- Lorentz Cone The Lorenz cone, or *ice cream cone* is given by
$$L^n = \left\{ \begin{pmatrix} \mathbf{x} \\ t \end{pmatrix} \in \mathbb{R}^{n+1} : \|\mathbf{x}\| \leq t, \mathbf{x} \in \mathbb{R}^n, t \in \mathbb{R} \right\}.$$
- nonnegative polynomials. set consisting of all possible coefficients of polynomials of degree $n-1$ which are nonnegative over $\mathbb{R}$:
$$K^n = \{\mathbf{x} \in \mathbb{R}^n : x_1 t^{n-1} + x_2 t^{n-2} + \ldots + x_{n-1} t + x_n \geq 0 \forall t \in \mathbb{R}\}$$



The Conic Hull

Definition. Given m points $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m \in \mathbb{R}^n$, a conic 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 $\lambda \in \mathbb{R}^m_+$.

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:

$$\text{cone}(S) \equiv \left\{ \sum_{i=1}^{k} \lambda_i \mathbf{x}_i : \mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k \in S, \boldsymbol{\lambda} \in \mathbb{R}^k_+ \right\}.$$

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 $\text{cone}(S) \subseteq T$.

Representation Theorem for Conic Hulls

a similar result to Carathéodory theorem

Conic Representation Theorem. Let $S \subseteq \mathbb{R}^n$ and let $\mathbf{x} \in \text{cone}(S)$. Then there exist k linearly independent vector $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k \in S$ such that $\mathbf{x} \in \text{cone}(\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\})$, that is, there exist $\boldsymbol{\lambda} \in \mathbb{R}^k_+$ such that

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

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

E.g. if A=[1,1,1], b=[1], then P is a triangle -- a 2-dimensional convex polygon -- in R^3. (In general, equality constraints take away degrees of freedom.)

- Consider the convex polyhedron.

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

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

Example. Consider the linear system:

We shall see later that bfs are the extreme points (a.k.a. vertices) of the convex polyhedron.

$$\begin{aligned} x_1 + x_2 + x_3 &= 6 \\ x_2 + x_4 &= 3 \\ x_1, x_2, x_3, x_4 &\geq 0. \end{aligned}$$

how does this convex set look like? (a quadrilaterial in R^4?)

Find all the basic feasible solutions. In class

Existence of bfs's

feasible/constraint set of a linear program

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

Proof.

- $P \neq \emptyset \Rightarrow \mathbf{b} \in \text{cone}(\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n\})$ where $\mathbf{a}_i$ denotes the i-th column of $\mathbf{A}$.
- 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 $\mathbf{b} = \sum_{j=1}^k y_{i_j}\mathbf{a}_{i_j}$ and $\mathbf{a}_{i_1}, \mathbf{a}_{i_2}, \ldots, \mathbf{a}_{i_k}$ are linearly independent.
- Denote $\bar{\mathbf{x}} = \sum_{j=1}^k y_{i_j}\mathbf{e}_{i_j}$. Then obviously $\bar{\mathbf{x}} \geq \mathbf{0}$ and in addition

$$\mathbf{A}\bar{\mathbf{x}} = \sum_{j=1}^k y_{i_j}\mathbf{A}\mathbf{e}_{i_j} = \sum_{j=1}^k y_{i_j}\mathbf{a}_{i_j} = \mathbf{b}.$$

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

A typical convex polyhedron in high-dimension contains a gazillion of vertices/bfs/extreme points. (We will see in Theorem 6.34 that bfs is equivalent to extreme points/ vertices.)

Topological Properties of Convex Sets

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

Proof.

- Let $\mathbf{x}, \mathbf{y} \in \mathrm{cl}(C)$ and let $\lambda \in [0,1]$.
- There exist sequences $\{\mathbf{x}_k\}_{k\geq 0} \subseteq C$ and $\{\mathbf{y}_k\}_{k\geq 0} \subseteq C$ for which $\mathbf{x}_k \to \mathbf{x}$ and $\mathbf{y}_k \to \mathbf{y}$ as $k \to \infty$.
- (*) $\lambda\mathbf{x}_k + (1-\lambda)\mathbf{y}_k \in C$ for any $k \geq 0$.
- (**) $\lambda\mathbf{x}_k + (1-\lambda)\mathbf{y}_k \to \lambda\mathbf{x} + (1-\lambda)\mathbf{y}$.
- (*)+(**) $\Rightarrow \lambda\mathbf{x} + (1-\lambda)\mathbf{y} \in \mathrm{cl}(C)$.

The Line Segment Principle

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

Proof.

- There exists $\varepsilon > 0$ such that $B(\mathbf{x}, \varepsilon) \subseteq C$.
- Let $\mathbf{z} = (1-\lambda)\mathbf{x} + \lambda\mathbf{y}$. We will show that $B(\mathbf{z}, (1-\lambda)\varepsilon) \subseteq C$.
- Let $\mathbf{w} \in B(\mathbf{z}, (1-\lambda)\varepsilon)$. Since $\mathbf{y} \in \text{cl}(C)$, $\exists \mathbf{w}_1 \in C$ s.t.

$$\|\mathbf{w}_1 - \mathbf{y}\| < \frac{(1-\lambda)\varepsilon - \|\mathbf{w} - \mathbf{z}\|}{\lambda}. \tag{3}$$

- Set $\mathbf{w}_2 = \frac{1}{1-\lambda}(\mathbf{w} - \lambda\mathbf{w}_1)$. Then

$$\begin{aligned} \|\mathbf{w}_2 - \mathbf{x}\| &= \left\| \frac{\mathbf{w} - \lambda\mathbf{w}_1}{1-\lambda} - \mathbf{x} \right\| = \frac{1}{1-\lambda}\|(\mathbf{w} - \mathbf{z}) + \lambda(\mathbf{y} - \mathbf{w}_1)\| \\ &\leq \frac{1}{1-\lambda}(\|\mathbf{w} - \mathbf{z}\| + \lambda\|\mathbf{w}_1 - \mathbf{y}\|) \overset{(3)}{<} \varepsilon, \end{aligned}$$

- Hence, since $B(\mathbf{x}, \varepsilon) \subseteq C$, it follows that $\mathbf{w}_2 \in C$. Finally, since $\mathbf{w} = \lambda\mathbf{w}_1 + (1-\lambda)\mathbf{w}_2$ with $\mathbf{w}_1, \mathbf{w}_2 \in C$, we have that $\mathbf{w} \in C$.

Convexity of the Interior

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

Proof.

- If $\mathrm{int}(C) = \emptyset$, then the theorem is obviously true.
- Otherwise, let $\mathbf{x}_1, \mathbf{x}_2 \in \mathrm{int}(C)$, and let $\lambda \in (0, 1)$.
- By the LSP, $\lambda \mathbf{x}_1 + (1 - \lambda)\mathbf{x}_2 \in \mathrm{int}(C)$, establishing the convexity of $\mathrm{int}(C)$.

Combination of Closure and Interior

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

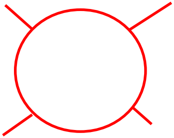
1. $\mathrm{cl}(\mathrm{int}(C)) = \mathrm{cl}(C)$.
2. $\mathrm{int}(\mathrm{cl}(C)) = \mathrm{int}(C)$.

Proof of 1.

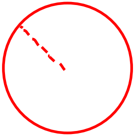
- Obviously, $\mathrm{cl}(\mathrm{int}(C)) \subseteq \mathrm{cl}(C)$ holds.
- To prove that opposite, let $\mathbf{x} \in \mathrm{cl}(C), \mathbf{y} \in \mathrm{int}(C)$.
- Then $\mathbf{x}_k = \frac{1}{k}\mathbf{y} + \left(1 - \frac{1}{k}\right)\mathbf{x} \in \mathrm{int}(C)$ for any $k \geq 1$.
- Since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

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

an example that fails 1 but satisfies 2.



an example that fails 2 but satisfies 1.



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 $\|\mathbf{x}\| \leq M$ for any $\mathbf{x} \in S$.
- Let $\mathbf{y} \in \text{conv}(S)$. Then there exist $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{n+1} \in S$ and $\boldsymbol{\lambda} \in \Delta_{n+1}$ for which $\mathbf{y} = \sum_{i=1}^{n+1} \lambda_i \mathbf{x}_i$ and therefore

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

  establishing the boundedness of $\text{conv}(S)$.
- To prove the closedness of $\text{conv}(S)$, let $\{\mathbf{y}_k\}_{k \geq 1} \subseteq \text{conv}(S)$ be a sequence converging to $\mathbf{y} \in \mathbb{R}^n$.
- There exist $\mathbf{x}_1^k, \mathbf{x}_2^k, \ldots, \mathbf{x}_{n+1}^k \in S$ and $\boldsymbol{\lambda}^k \in \Delta_{n+1}$ such that

$$\mathbf{y}_k = \sum_{i=1}^{n+1} \lambda_i^k \mathbf{x}_i^k. \tag{4}$$

Proof Contd.

- By the compactness of S and Δ_{n+1}, it follows that $\{(\boldsymbol{\lambda}^k, \mathbf{x}_1^k, \mathbf{x}_2^k, \ldots, \mathbf{x}_{n+1}^k)\}_{k\geq 1}$ has a convergent subsequence $\{(\boldsymbol{\lambda}^{k_j}, \mathbf{x}_1^{k_j}, \mathbf{x}_2^{k_j}, \ldots, \mathbf{x}_{n+1}^{k_j})\}_{j\geq 1}$ whose limit will be denoted by

$$(\boldsymbol{\lambda}, \mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{n+1})$$

  with $\boldsymbol{\lambda} \in \Delta_{n+1}, \mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{n+1} \in S$
- Taking the limit $j \to \infty$ in

$$\mathbf{y}_{k_j} = \sum_{i=1}^{n+1} \lambda_i^{k_j} \mathbf{x}_i^{k_j},$$

  we obtain that $\mathbf{y} = \sum_{i=1}^{n+1} \lambda_i \mathbf{x}_i \in \text{conv}(S)$ 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 $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k \in \mathbb{R}^n$. Then $\text{cone}(\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k\})$ is closed.

Proof.

- By the conic representation theorem, each element of $\text{cone}(\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k\})$ can be represented as a conic combination of a linearly independent subset of $\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k\}$.
- Therefore, if $S_1, S_2, \ldots, S_N$ are all the subsets of $\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k\}$ comprising linearly independent vectors, then

$$\text{cone}(\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k\}) = \bigcup_{i=1}^{N} \text{cone}(S_i).$$

- 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

$$S_i = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_m\},$$

  where $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_m$ are linearly independent.
- $\text{cone}(S_i) = \{\mathbf{B}\mathbf{y} : \mathbf{y} \in \mathbb{R}^m_+\}$, where $\mathbf{B}$ is the matrix whose columns are $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_m$.

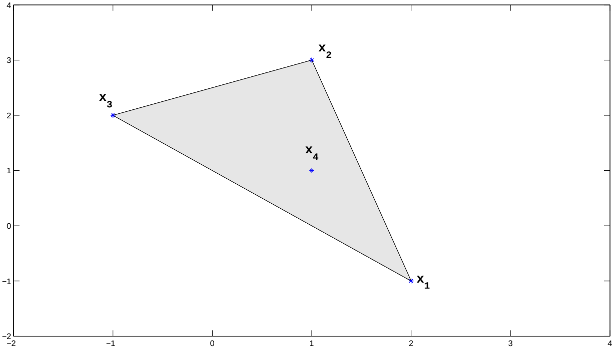
Proof Contd.

- Suppose that $\mathbf{x}_k \in \text{cone}(S_i)$ for all $k \geq 1$ and that $\mathbf{x}_k \to \bar{\mathbf{x}}$.
- $\exists \mathbf{y}_k \in \mathbb{R}^m_+$ such that
$$\mathbf{x}_k = \mathbf{B}\mathbf{y}_k. \tag{5}$$
- $$\mathbf{y}_k = (\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T\mathbf{x}_k.$$
- Taking the limit as $k \to \infty$ in the last equation, we obtain that $\mathbf{y}_k \to \bar{\mathbf{y}}$ where $\bar{\mathbf{y}} = (\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T\bar{\mathbf{x}}$.
- $\bar{\mathbf{y}} \in \mathbb{R}^m_+$.
- Thus, taking the limit in (5), we conclude that $\bar{\mathbf{x}} = \mathbf{B}\bar{\mathbf{y}}$ with $\bar{\mathbf{y}} \in \mathbb{R}^m_+$, and hence $\bar{\mathbf{x}} \in \text{cone}(S_i)$.

Extreme Points

Definition. Let $S \subseteq \mathbb{R}^n$ be a convex set. A point $\mathbf{x} \in S$ is called an extreme point of S if there do not exist $\mathbf{x}_1, \mathbf{x}_2 \in S (\mathbf{x}_1 \neq \mathbf{x}_2)$ and $\lambda \in (0, 1)$, such that $\mathbf{x} = \lambda \mathbf{x}_1 + (1 - \lambda)\mathbf{x}_2$.

- The set of extreme point is denoted by $\mathrm{ext}(S)$.
- 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 $P = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$, where $\mathbf{A} \in \mathbb{R}^{m \times n}$ has linearly independent rows and $\mathbf{b} \in \mathbb{R}^m$. The $\bar{\mathbf{x}}$ 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

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