Convexity and Polyhedra

Carlo Mannino
(from Geir Dahl notes on convexity)

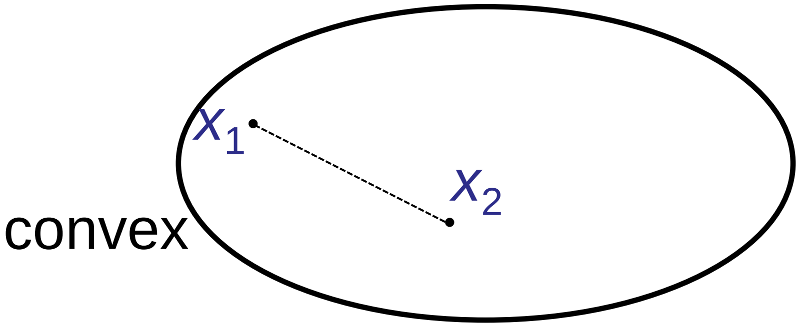
University of Oslo, INF-MAT5360 - Autumn 2011 (Mathematical optimization)

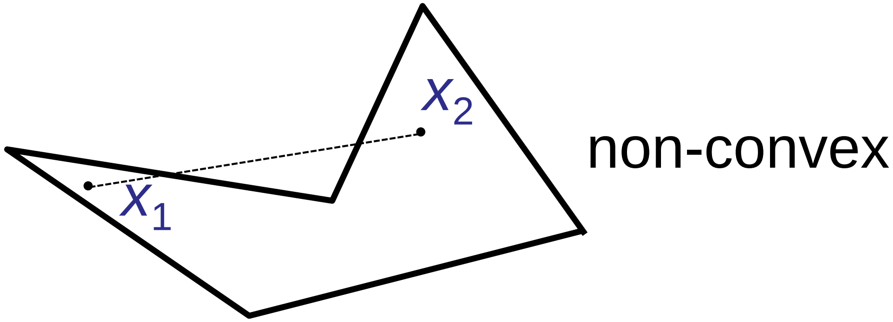
Convex Sets

- Set $C \subseteq R^n$ *is* convex if $(1-\lambda)x_1 + \lambda x_2 \in C$ whenever

  $x_1, x_2 \in C \quad 0 \leq \lambda \leq 1$

(the segment joining x_1 , x_2 is contained in C)

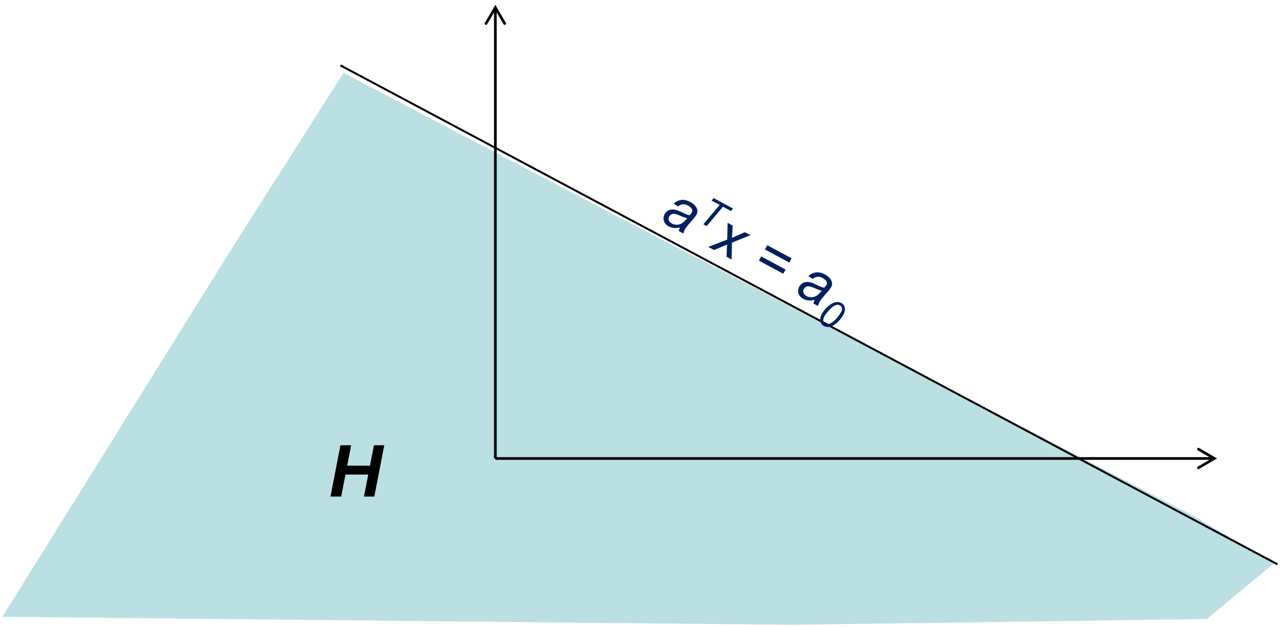




- Show that the unit ball $B = \{x \in R^n : ||x|| \leq 1\}$ is convex. (Hint use the triangle inequality $||x+y|| \leq ||x|| + ||y||$)

Half-spaces

- Example of convex sets: half-spaces $H = \{x \in R^n : a^T x \leq a_0\}$



$x_1 \in H \rightarrow a^T x_1 \leq a_0 \rightarrow (1-\lambda)\, a^T x_1 \leq (1-\lambda)\, a_0$

$x_2 \in H \rightarrow a^T x_2 \leq a_0 \rightarrow \lambda a^T x_2 \leq \lambda\, a_0$

$0 \leq \lambda \leq 1$

summing up $x_2 \in H \rightarrow a^T ((1-\lambda)\, x_1 + \lambda\, x_2) \leq a_0$

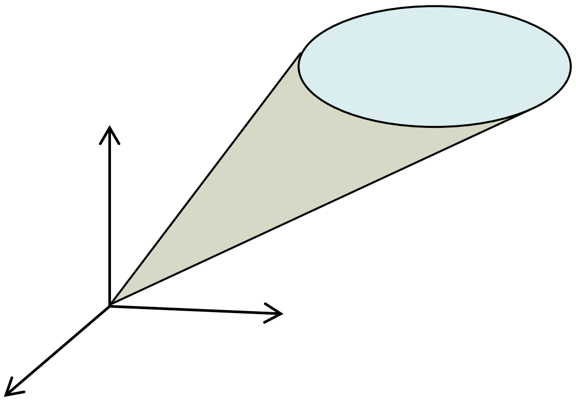
$\Rightarrow$ $(1-\lambda)\, x_1 + \lambda\, x_2 \in H$

Convex Cones

- The set of solutions to a linear system of equation is a polyhedron.

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

> *Convex Cone*: $C \subseteq \mathbb{R}^n$ if $\lambda_1 x_1 + \lambda_2 x_2 \in C$ whenever $x_1, x_2 \in C$ and $\lambda_1 , \lambda_2 \geq 0$.



- Each convex cone is a convex set. (show)

- Let $A \in \mathbb{R}^{m,n}$. Then $C = \{x \in \mathbb{R}^n : Ax \leq 0\}$ is a convex cone (show).

> Let $x_1,\ldots,x_t \in \mathbb{R}^n$, and $\lambda_1,\ldots,\lambda_t \geq 0$. The vector $x = \sum_{j=1}^{t} \lambda_j x_j$ is a *nonnegative* (or *conical*) *combination* of $x_1,\ldots,x_t$

- The set $C(x_1,\ldots,x_t)$ of all nonnegative combinations of $x_1,\ldots,x_t \in \mathbb{R}^n$ is a convex cone (show), called *finitely generated cone*.

Linear Programming

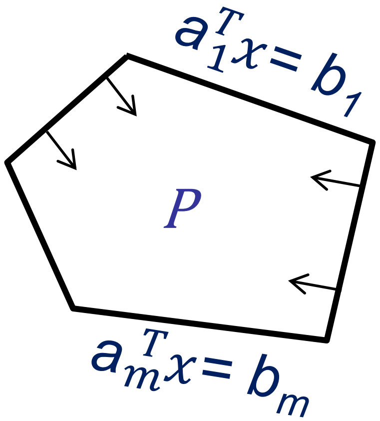
- Property: C_1 , C_2 convex sets → $C_1 \cap C_2$ convex (show!)

Linear programming: $x = (x_1, \ldots, x_n)$

$$\text{maximize} \quad c_1x_1 + \ldots + c_n x_n$$

Subject to

$$\begin{aligned} a_{11}x_1 + \quad \ldots \quad + a_{1n}x_n &\le b_1 \\ &\vdots \\ a_{m1}x_1 + \qquad\quad + a_{mn}x_n &\le b_m \\ x_1, \ldots, x_n &\ge 0 \end{aligned}$$



$\max\{c^T x : x \in P\}$, with $P = \{x \in R^n : Ax \le b, x \ge 0\}$

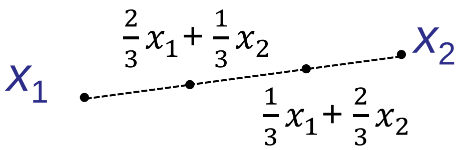
Find the optimum solution in P

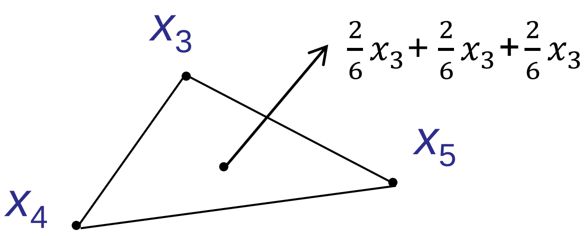
P intersection of a finite number of half-spaces: convex set (polyhedron)

- The set of optimal solutions to a linear program is a polyhedron (show!)

Convex Combinations

Let $x_1, \ldots, x_t \in \mathbb{R}^n$, and $\lambda_1, \ldots, \lambda_t \geq 0$, such that $\sum_{j=1}^{t} \lambda_j = 1$. The vector $x = \sum_{j=1}^{t} \lambda_j x_j$ is called *convex combination* of $x_1, \ldots, x_t$





Convex Combinations

Theorem: a set C is convex if and only if it contains all convex combinations of its points.

⇒ If C contains all convex combinations → it contains all convex combinations of any 2 points → C is convex

⇐ Suppose C contains all convex combinations of t-1 points. True if $t \leq 3$ (since C convex).

Let $x_1, \ldots, x_t \in \mathbb{R}^n$, and let $x = \sum_{j=1}^{t} \lambda_j x_j$ where $\lambda_1, \ldots, \lambda_t > 0$, $\sum_{j=1}^{t} \lambda_j = 1$

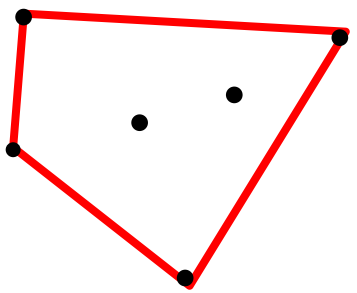
$$x = \lambda_1 x_1 + \sum_{j=2}^{t} \lambda_j x_j = \lambda_1 x_1 + (1 - \lambda_1) \sum_{j=2}^{t} (\lambda_j / (1 - \lambda_1))\, x_j$$

$$\sum_{j=1}^{t} \lambda_j = 1 \rightarrow \sum_{j=2}^{t} (\lambda_j / (1 - \lambda_1)) = 1 \Rightarrow \sum_{j=2}^{t} (\lambda_j / (1 - \lambda_1))\, x_j = y \in C$$

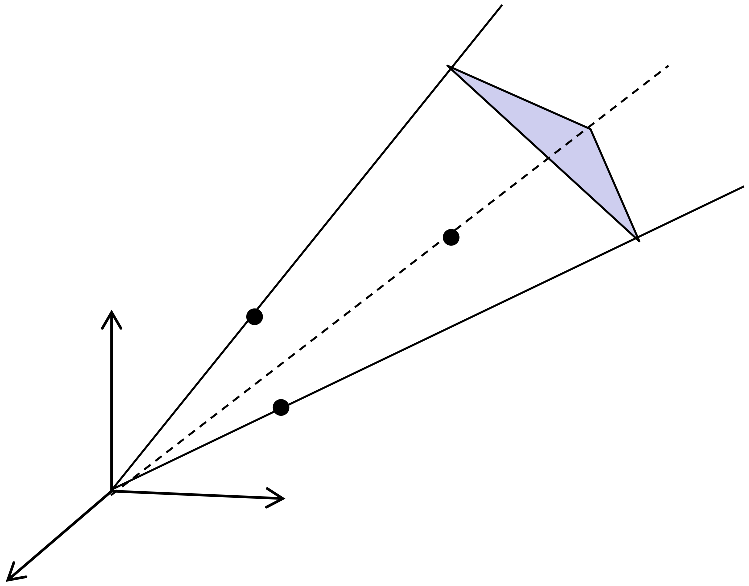
$$\Rightarrow x = \lambda_1 x_1 + (1 - \lambda_1)\, y \in C$$

■

Convex and Conical Hull

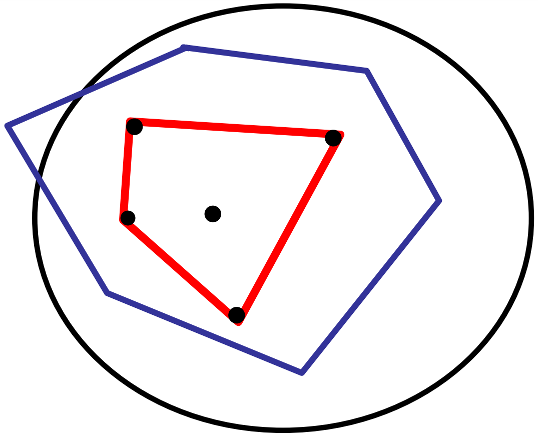


- There many convex sets containing a given set of points *S*.
- The smallest is the set conv(*S*) of all convex combinations of the points in S.
- conv(*S*) is called *convex hull* of *S*



- The set *cone*(*S*) of all nonnegative (conical) combinations of points in *S* is called *conical hull*

Convex Hull



Proposition 2.2.1 (Convex hull). Let $S \subseteq \mathbb{R}^n$. Then conv(S) is equal to the intersection of all convex sets containing S.

- If S is finite, conv(S) is called *polytope*.
- Consider the following optimization problem:

  max $\{c^T x: x \in P\}$, with P = conv(S), $S = \{x_1, \ldots, x_t\}$

Let x^*: $c^T x^* = \max \{c^T x: x \in S\} = v$ (x^* optimum in S)

For any $y \in P$ there exist $\lambda_1, \ldots, \lambda_t \geq 0$, $\sum_{j=1}^{t} \lambda_j = 1$, such that $y = \sum_{j=1}^{t} \lambda_j x_j$

⇨ $c^T y = c^T \sum_{j=1}^{t} \lambda_j x_j = \sum_{j=1}^{t} \lambda_j c^T x_j \leq \sum_{j=1}^{t} \lambda_j c^T x^* = \sum_{j=1}^{t} \lambda_j v = v$

⇨ x^* optimum in P

Affine independence

A set of vectors $x_1,\ldots,x_t \in \mathbb{R}^n$, are *affinely independent* *if* $\sum_{j=1}^{t} \lambda_j x_j = 0$ and $\sum_{j=1}^{t} \lambda_j = 0$ imply $\lambda_1 = \ldots = \lambda_t = 0$.

Proposition 2.3.1 (Affine independence). The vectors $x_1,\ldots,x_t \in \mathbb{R}^n$ are affinely independent if and only if the t-1 vectors $x_2 - x_1,\ldots,x_t - x_1$ are linearly independent.

Only if. $x_1,\ldots,x_t \in \mathbb{R}^n$ affinely independent and assume $\lambda_2,\ldots,\lambda_t \in \mathbb{R}^n$ with

$$\sum_{j=2}^{t} \lambda_j (x_j - x_1) = 0 \quad \Rightarrow \quad -\left(\sum_{j=2}^{t} \lambda_j\right) x_1 + \sum_{j=2}^{t} \lambda_j x_j = 0$$

$-\left(\sum_{j=2}^{t} \lambda_j\right) + \sum_{j=2}^{t} \lambda_j = 0$ and $x_1,\ldots,x_t$ affinely independent

$\Rightarrow$ $\lambda_2 = \ldots = \lambda_t = 0$ $\Rightarrow$ $x_2 - x_1,\ldots,x_t - x_1$ linearly independent. ∎

Affine independence

Proposition 2.3.1 (Affine independence). The vectors $x_1,\ldots,x_t \in \mathbb{R}^n$ are affinely independent if and only if the t-1 vectors x_2-$x_1,\ldots,x_t$-x_1 are linearly independent.

if. x_2-$x_1,\ldots,x_t$-x_1 linearly independent.

Assume $\sum_{j=1}^{t} \lambda_j x_j = 0$ and $\sum_{j=1}^{t} \lambda_j = 0$. Then $\lambda_1 = -\sum_{j=2}^{t} \lambda_j$

$$0 = \sum_{j=1}^{t} \lambda_j x_j = -(\sum_{j=2}^{t} \lambda_j)x_1 + \sum_{j=2}^{t} \lambda_j x_j = \sum_{j=2}^{t} \lambda_j (x_j - x_1)$$

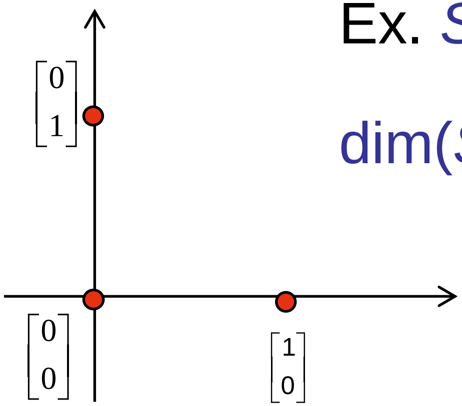
As x_2-$x_1,\ldots,x_t$-x_1 linearly independent $\lambda_2=\ldots=\lambda_t = 0$

Also $\lambda_1 = -\sum_{j=2}^{t} \lambda_j = 0$ ▪ ■

Corollary. There are at mots n+1 affinely independent vectors in $\mathbb{R}^n$.

Dimension

The *dimension* dim(S) *of a set* $S \subseteq \mathbb{R}^n$ is the maximal number of affinely independent points of S minus 1.



Ex. $S = \{x_1 = (0,0),\ x_2 = (0,1),\ x_3 = (1,0)\}$.

dim(S) = 2 ($x_2 - x_1$, $x_3 - x_1$ are linearly independent)

A *simplex* $P \subseteq \mathbb{R}^n$ is the convex hull of a set S of affinely independent vectors in $\mathbb{R}^n$

Caratheodory's theorem

Theorem. 2.5.1 (Caratheodory's theorem) Let $S \subseteq \mathbb{R}^n$. Then each $x \in \mathrm{conv}(S)$ is the convex combination of m affinely independent points in S, with $m \leq n+1$.

x can be obtained as a convex combination of points in S

Choose one with smallest number of points:

$x = \sum_{j=1}^{t} \lambda_j x_j$ with $\lambda_1, \ldots, \lambda_t > 0$, $\sum_{j=1}^{t} \lambda_j = 1$ and t smallest possible

Then $x_1, \ldots, x_t$ are affinely independent (with $t \leq n+1$). Suppose not.

There are $\mu_1, \ldots, \mu_t$ not all 0 such that $\sum_{j=1}^{t} \mu_j x_j = 0$ and $\sum_{j=1}^{t} \mu_j = 0$

Then there is at least one positive coefficient, say μ_1

Caratheodory's theorem

$\sum_{j=1}^{t} \mu_j x_j = 0$, $\sum_{j=1}^{t} \mu_j = 0$, $\mu_1 > 0$

Combining $x = \sum_{j=1}^{t} \lambda_j x_j$ and $\alpha \sum_{j=1}^{t} \mu_j x_j = 0$ for $\alpha \geq 0$

$\Rightarrow$ $x = \sum_{j=1}^{t} (\lambda_j - \alpha\mu_j) x_j$

Increase α from 0 to α_0 until the first coefficient becomes 0, say the *r-th*.

$\lambda_j - \alpha\mu_j \geq 0 \quad j = 1, \ldots, t$ and $\lambda_r - \alpha\mu_r = 0$

$\sum_{j=1}^{t} (\lambda_j - \alpha\mu_j) = \sum_{j=1}^{t} \lambda_j - \sum_{j=1}^{t} \alpha\mu_j = \sum_{j=1}^{t} \lambda_j = 1$

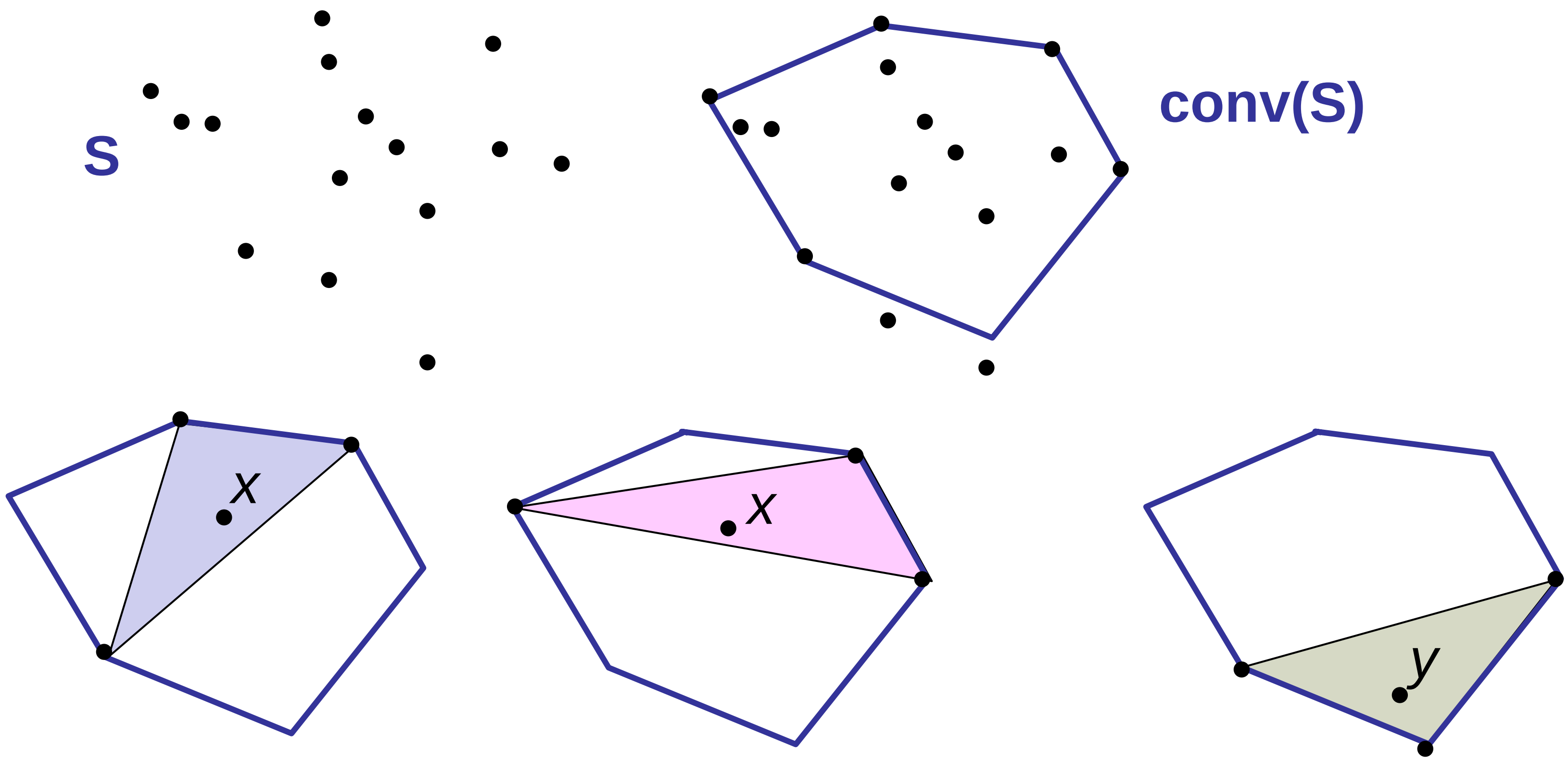
Then x is obtained as a convex combination of t-1 point in S, contrad. ■

A similar result for conical hulls.

Theorem. 2.5.2. (Caratheodory's theorem for conical hulls). Let $S \subseteq \mathbb{R}^n$. Then each $x \in \text{cone}(S)$ is the conical combination of m linearly independent points in S, with $m \leq n$.

Caratheodory's theorem for cones

- Any point in conv(S) $\subseteq R^n$ can be *generated* by (at most) $n+1$ points of S.
- The generators of a point x are not necessarily unique.
- The generators of different points may be different.



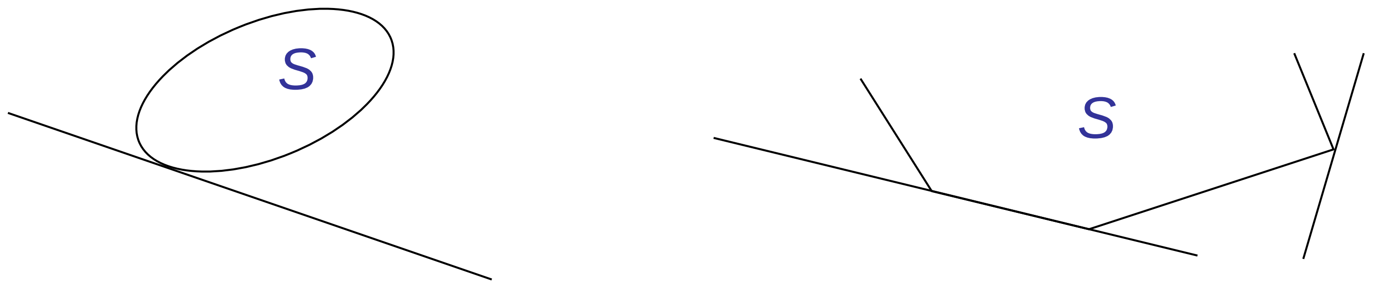
Caratheodory's theorem and LP

- Consider LP: $\max \{c^T x : x \in P\}$, with $P = \{x \in R^n : Ax = b, x \geq 0\}$
- $A \in R^{m,n}$, $m \leq n$. Let $a_1, \ldots, a_n \in R^m$ be the columns of A
- Ax can be written as $\sum_{j=1}^{n} x_j a_j$, $x_1, \ldots, x_n \in R_+$
- $P \neq \emptyset$ if and only if $b \in \text{cone}(\{a_1, \ldots, a_n\})$
- Caratheodory: b can be obtained conical combination of $t \leq m$ *linearly independent* a_j's.
- *Equivalently*: there exists a non-negative $x \in R^n$ with at least n-t components being 0 and $Ax = b$ …
- … and the non-zeros of x correspond to linearly independent columns of A *(basic fesible solution)*
- Fundamental result*: if an LP is non-empty then it contains a basic feasible solution*

Supporting Hyperplanes

A *hyperplane* is a set $H \subset \mathbb{R}^n$ of the form $H = \{x \in \mathbb{R}^n : a^T x = \alpha\}$ for some nonzero vector a and a real number α.

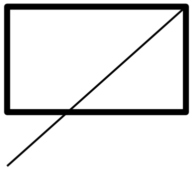
- Let $H^- = \{x \in \mathbb{R}^n : a^T x \leq \alpha\}$ and $H^+ = \{x \in \mathbb{R}^n : a^T x \geq \alpha\}$ be the two halfspaces identified by H.
- H is a convex set ($H = H^- \cap H^+$).
- If $S \subset \mathbb{R}^n$ is contained in one of the two halfspaces H^- and H^+, and $S \cap H$ is non-empty, then H is a *supporting hyperplane* of S.
- H supports S at x for $x \in S \cap H$. If S is convex, then $S \cap H$ is called *exposed face* of S, which is convex (S and H are convex).



Faces

Let C be a convex set. A convex subset F of C is a *face* if $x_1, x_2 \in C$ and $(1-\lambda)\, x_1 + \lambda\, x_2 \in F$ for some $0 < \lambda < 1$, then $x_1, x_2 \in F$

(if a relative interior point of the line segment between two points of C lies in F then the whole line segment lies in F)



The sides and the vertices of the square are faces.

The diagonal is not a face (show it!)

A face F with $\dim(F) = 0$ is called *extreme point*. The set of all extreme points of C is $\mathrm{ext}(C)$. A bounded face F with $\dim(F) = 1$ is called *edge*.

An unbounded face F with $\dim(F) = 1$ is either a line or a halfline (*ray*). {i.e. a set $\{x_0 + \lambda z : \lambda \geq 0\}$) and is called *extreme halfline (ray).*

The set of all extreme halflines of C is $\mathrm{exthl}(C)$.

Exposed Faces are Faces

Proposition 4.1.1 Let C be a nonempty convex set. Each exposed face F of C is also a face of C.

Let $H = \{x \in \mathbb{R}^n: c^T x = v\}$ and $F = C \cap H$. H supporting C implies (say) $C \subseteq H^- = \{x \in \mathbb{R}^n: c^T x \leq v\}$ and $v = max\ \{c^T x : x \in C\}$.

So F is the set of points of C maximizing $c^T x$.

Let $x_1, x_2 \in C$ and suppose $(1-\lambda)\ x_1 + \lambda\ x_2 \in F$ for some $0 < \lambda < 1$.

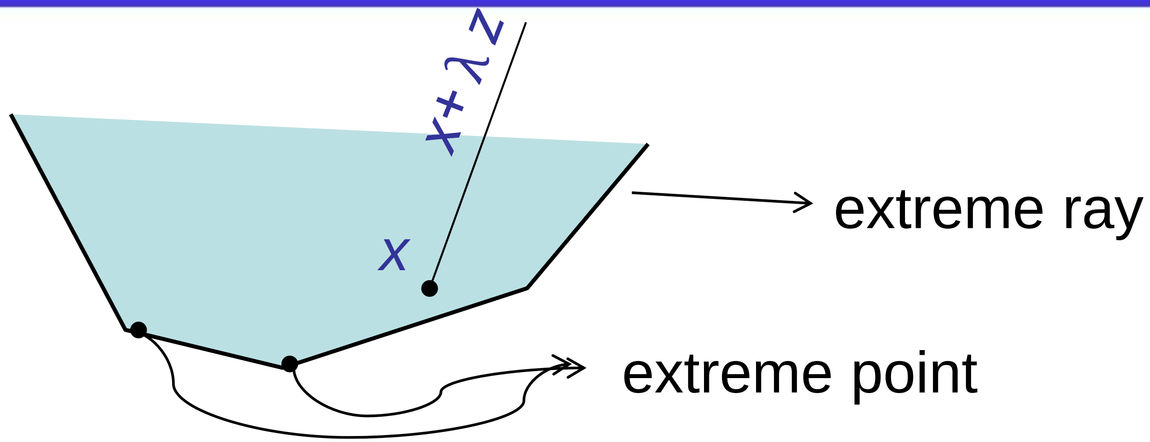
$x_1, x_2 \in C$ imply (i) $c^T x_1 \leq v$ and (ii) $c^T x_2 \leq v$. Suppose $x_1 \notin F$. Then $c^T x_1 < v$.

λ, $1-\lambda > 0$ implies $(1-\lambda)\ c^T x_1 < (1-\lambda) v$ and $\lambda\ c^T x_2 \leq \lambda\ v$.

$$v > (1-\lambda)\ c^T x_1 + \lambda\ c^T x_2 = c^T(\ \lambda\ (1-\lambda)\ x_1 + \lambda\ x_2) = v,$$ contraddiction.

■

Recession Cone



Let C be a *closed convex set*. The set of directions of halflines from x that lie in C are denoted by $rec(C,x) = \{z \in \mathbb{R}^n : x + \lambda z \in C \text{ for all } \lambda \geq 0\}$

One can show the following:

Proposition 4.2.1 $rec(C,x)$ does not depend on x.

Let $rec(C) = rec(C,x)$ ($x \in C$) be the *recession cone* of C

- Show that $\mathrm{rec}\{x \in R^n : Ax \leq b\} = \{x \in R^n : Ax \leq O\}$

Inner Description

- Let C be a closed convex set.
- Let Z be the set of directions of the extreme rays (halflines) of C.
- One can show that the recession cone of C is the conical combination of the directions in Z, namely $\mathrm{rec}(C) = \mathrm{cone}(Z)$.

Corollary 4.3.3 (Inner description). Let $C \subseteq \mathbb{R}^n$ be a nonempty and line-free (*pointed*) closed convex set. Choose a direction vector z for each extreme halfline of C and let Z be the set of these direction vectors. Then we have that

$$C = \mathrm{conv}(\mathrm{ext}(C)) + \mathrm{rec}(C) = \mathrm{conv}(\mathrm{ext}(C)) + \mathrm{cone}(Z).$$

Corollary 4.3.4 (Minkowsky theorem). Let $C \subseteq \mathbb{R}^n$ be a bounded closed (compact) convex (set, then C is the convex hull of its extreme points: $C = \mathrm{conv}(\mathrm{ext}(C))$

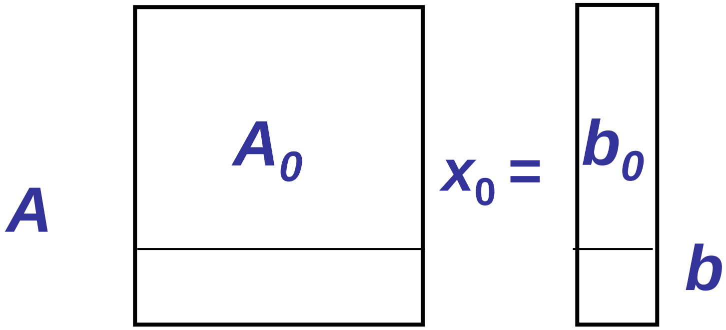
Polytopes and Polyhedra

- We consider a non-empty, line-free polyhedron $P = \{x \in \mathbb{R}^n : Ax \leq b\}$, where $A \in \mathbb{R}^{m,n}$, $b \in \mathbb{R}^m$.
- P pointed implies $rank(A) = n$ and $m \geq n$.

(If $rank(A) < n$ then there exists a non zero vector z: $Az = 0$; then for any $x_0 \in P$ we have $Ax_0 \leq b$ and $A(x_0 + \lambda z) \leq b$ for any $\lambda \in R$ and P contains the line through x_0 having direction z).

> A point $x_0 \in P$ is called a *vertex* if it is the unique solution to n linear independent equations from the system $Ax = b$.

- x_0 vertex of P: there exists an $n \times n$ non-singular sub-matrix A_0 of A, such that $A_0 x_0 = b_0$, with b_0 sub-vector of b corresponding to A_0



Vertices and extreme points

Lemma 4.4.1. A point $x_0 \in P = \{x \in \mathbb{R}^n: Ax \leq b\}$ is a *vertex of* P if and only if it is an extreme point of P.

only if. By contradiction. x_0 *vertex* but not extreme point

⇒ $x_0 = \frac{1}{2}x_1 + \frac{1}{2}x_2$ with $x_1, x_2 \in P$ and $A_0 x_0 = b_0$ with A_0 nonsingular

Let a_i be any row of A_0 (treated as a row vector)

since $x_1, x_2 \in P$, $a_i x_1 \leq b_i$ and $a_i x_2 \leq b_i$

if $a_i x_1 < b_i$ then $a_i x_0 = \frac{1}{2}a_i x_1 + \frac{1}{2}a_i x_2 < b_i$, contradiction.

Since a_i is any row , we have $A_0 x_1 = b_0$, and $A_0 x_2 = b_0$.

A_0 nonsigular implies $x_1 = x_2 = x_0$ ■

Vertices and extreme points

Lemma 4.4.1. A point $x_0 \in P = \{x \in \mathbb{R}^n: Ax \leq b\}$ is a *vertex of* P if and only if it is an extreme point of P.

if. Suppose x_0 is not a vertex.

Consider all i for which $a_i x_0 = b_i$ and let A_0 be the associated submatrix

Let $A_0 x_0 = b_0$ corresponding system. x_0 non vertex → $\text{rank}(A_0) < n$

$\text{rank}(A_0) < n$ → there is a nonzero vector z such that $A_0 z = O$

There is small $\varepsilon > 0$ such that $x_1 = x_0 + \varepsilon \cdot z \in P$ and $x_2 = x_0 - \varepsilon \cdot z \in P$

(since if a_i not in A_0 then $a_i x_0 < b_i$)

Then $x_0 = \frac{1}{2} x_1 + \frac{1}{2} x_2$ with $x_1, x_2 \in P$ and $x_1 \neq x_2$ ■

■

Extreme Halflines (rays)

A face F of P is a halfline if $F = x_0 + \text{cone}(\{z\}) = \{x_0 + \lambda z\colon \lambda \geq 0\}$

F <u>extreme</u> if there are not two distinct $z_1, z_2 \in \text{rec}(P)$ with $z = z_1 + z_2$

Lemma 4.4.2 (extreme halfline). $R = x_0 + \text{cone}(\{z\}) \subseteq P$ is an extreme halfline of P if and only if $A_0 z = O$ for some $(n-1)\times n$ submatrix of A with $\text{rank}(A_0) = n-1$.

- Since there are only $\binom{m}{n-1}$ ways of choosing n-1 rows of A, the number of extreme halflines is finite.
- Similarly, the number of extreme points is finite.

■

The main theorem for Polyhedra

Theorem 4.4.4 Each polyhedron P may be written as $P = \text{conv}(V) + \text{cone}(Z)$ for finite sets $V, Z \subset \mathbb{R}^n$. In particular, if P is pointed, we may here let V be the set of vertices and let Z consist of a direction vector of each extreme halfline of P.

Conversely, if V and Z are finite sets in $\mathbb{R}^n$, then $P=\text{conv}(V)+\text{cone}(Z)$ is a polyhedron. i.e., there is a matrix $A \in \mathbb{R}^{m,n}$ and a vector $b \in \mathbb{R}^m$ for some m such that $\text{conv}(V) + \text{cone}(Z) = \{x \in \mathbb{R}^n : Ax \leq b\}$.

Corollary 4.4.5 A set is a polytope if and only if it is a bounded polyhedron.

Exercises

- Show that the unit ball $B = \{x \in R^n : ||x|| \leq 1\}$ is convex. (Hint use the triangle inequality $||x+y|| \leq ||x||+ ||y||$)
- Show that C_1 , C_2 convex sets $\rightarrow C_1 \cap C_2$ is a convex set
- The set of optimal solutions to a linear program is a polyhedron
- Each convex cone is a convex set.
- Show that 2 distinct points are affinely independent
- Show that the diagonal of the square is not a face
- Let *x* be an extreme point of a convex set *C*, then there do not exist two distinct points of C *such that x is the convex combination of such points*
- What is rec(C) when C is a polytope?
- Show that $\text{rec}\{x \in R^n : Ax \leq b\} = \{x \in R^n : Ax \leq O\}$