Convex Optimization and Modeling

Convex sets

Second lecture, 21.04.2010

Jun.-Prof. Matthias Hein

Convex Sets:

- affine and convex sets
- cones
- operations that preserve convexity
- generalized inequalities
- separating and supporting hyperplane theorems
- extreme points - fundamental theorem of linear programming

Why are convex sets important ?

The domain of a convex function has to be convex

$$f\big(\theta x + (1-\theta)y\big) \quad \leq \quad \theta\, f(x) + (1-\theta)\, f(y), \quad \forall\, \theta \in [0,1].$$

Definition 1. *A set $C \subset \mathbb{R}^n$ is* ***affine*** *if for any $x_1, x_2 \in C$ and $\theta \in \mathbb{R}$ we have*

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

A set is **affine** if the line through every two points inside the set is contained in the set.

Affine set: subspace with some offset,

$$C = x_0 + V = \{x_0 + v \,|\, v \in V\},$$

where V is a subspace of $\mathbb{R}^n$.

Affine dimension of an affine set: dimension of the subspace V.

Definition 2. *The* ***affine hull*** *of a set of points $x_1, \dots, x_k$ is defined as,*

$$\mathit{aff}\ C = \Big\{ \sum_{i=1}^{k} \theta_i x_i \,\Big|\, \sum_{i=1}^{k} \theta_i = 1 \Big\}.$$

Important note: If the affine hull of a set of points is not equal to $\mathbb{R}^n$ the interior of an affine set is always empty.

Definition 3. *The* ***relative interior*** *relint C of the set C is defined as,*

$$\mathit{relint}\ C = \{ x \in C \mid B(x,r) \cap \mathit{aff}\ C \subseteq C \ \textit{for some}\ r > 0 \},$$

where $B(x,r) = \{ y \in \mathbb{R}^n \mid \|x - y\| \leq r \}$ is the ball of radius r around x with respect to some norm on $\mathbb{R}^n$. The ***relative boundary*** *of C is the set $\overline{C} \setminus \mathit{relint}\ C$.*

Convex sets: A convex set C is a set such that every line segment connecting any two points in C is again contained in C.

Definition 4. *A set C is* ***convex*** *if for any $x_1, x_2 \in C$ and for any θ with $0 \le \theta \le 1$ we have*

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

$\Longrightarrow$ Every affine set is clearly convex.

Definition 5. *A point $z = \sum_{i=1}^k \theta_i x_i$ where $\sum_{i=1}^k \theta_i = 1$ and $\theta_i \ge 0$ is a* ***convex combination*** *of $x_1, \ldots, x_k$. The* ***convex hull*** *of a set C is defined as*

$$\operatorname{conv} C = \Big\{ \sum_{i=1}^k \theta_i x_i \,\Big|\, x_1, \ldots, x_k \in C,\ \theta_i \ge 0,\ \sum_{i=1}^k \theta_i = 1,\ k \in \mathbb{N} \Big\}.$$

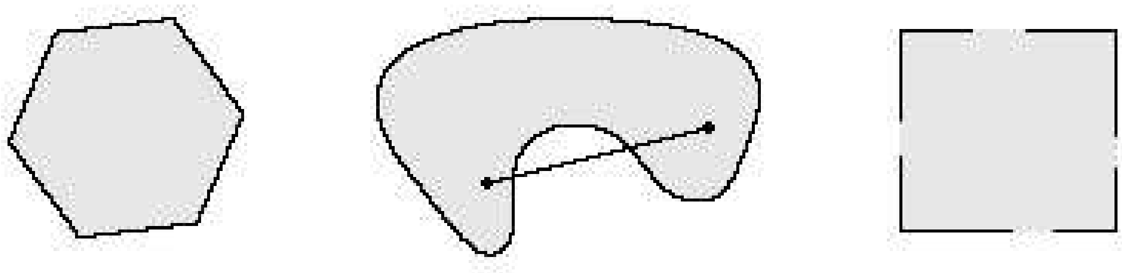
The convex combination can be seen as the weighted average of the points.

Convex Sets II

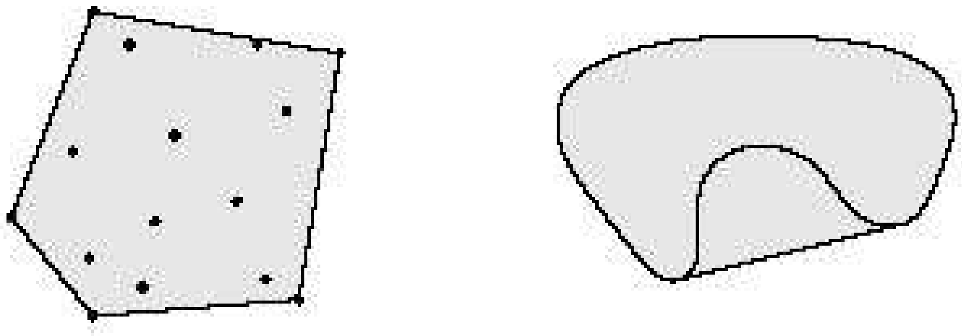
- The convex hull $\operatorname{conv} C$ of a set C is convex. It is the smallest convex set containing C.
- extension to continuous weights

$$\operatorname{conv} C = \left\{ \int_C x\, p(x)\, dx \;\middle|\; p(x) \geq 0,\; \int_C p(x)\, dx = 1 \right\}.$$

Note: $\int_C xp(x)\, dx = \mathbb{E}[X]$.
Discrete version: use an atomic measure, where $\mathrm{P}(X = x_i) = p_i$.



Left: Convex set, Middle: Not Convex, Right: Not Convex.



Left: convex hull of a set of points, Right: convex hull of a non-convex set.

Definition 6. *A set C is called a* **cone** *if for every $x \in C$ and $\theta \geq 0$ we have $\theta x \in C$. A set C is a* **convex cone** *if it is convex and a cone, which means that for any $x_1, x_2 \in C$ and $\theta_1, \theta_2 \geq 0$ we have*

$$\theta_1 x_1 + \theta_2 x_2 \in C.$$

Note: a cone always contains the origin in $\mathbb{R}^n$.

Definition 7. *The* **conic hull** *of a set C is defined as*

$$\Big\{ \sum_{i=1}^{k} \theta_i x_i \,\Big|\, x_1, \ldots, x_k \in C, \theta_i \geq 0,\; k \in \mathbb{N} \Big\}.$$

- The conic hull is the smallest convex cone containing C.
- A cone is convex if and only if it contains all conic combinations of its elements.

Definition 8. *A **polyhedron** P is the set*

$$P = \{x \mid \langle a_j, x\rangle \leq b_j,\ j = 1, \ldots, m, \quad \langle c_i, x\rangle = d_i,\ i = 1, \ldots, p\}.$$

- more compact form of a polyhedron P:

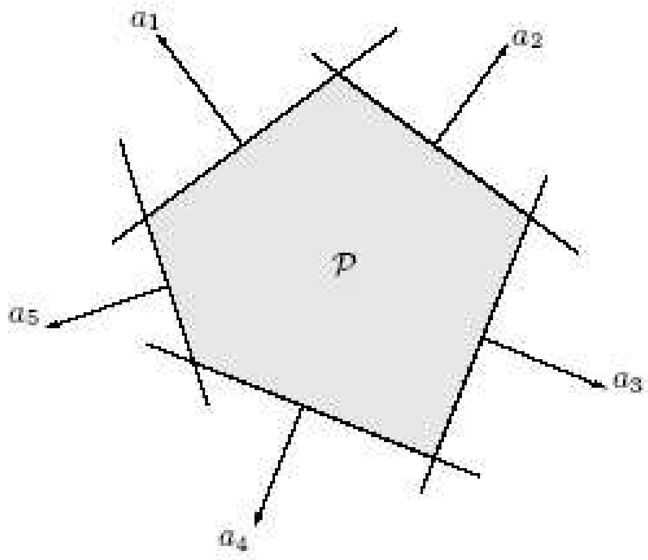
$$P = \{x \mid Ax \preceq b,\ Cx = d\}.$$

  where $\preceq$ indicates componentwise inequality and the j-th row of A is a_j and the i-th row of C is c_i.

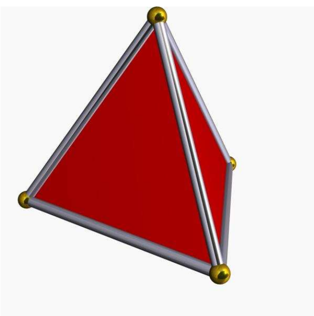
Definition 9. *Let $x_0, \ldots, x_k$ be $k+1$ points and assume that $x_1 - x_0, \ldots, x_k - x_0$ are linearly independent. Then the **simplex** determined by them is given by*

$$C = \mathrm{conv}\{x_0, \ldots, x_k\} = \Big\{ \sum_{i=0}^{k} \theta_i x_i \,\Big|\, \theta \succeq 0,\ \mathbf{1}^T\theta = 1 \Big\}.$$

Polyhedron



k-Simplex


A 3-simplex in $\mathbb{R}^3$.

- 1-simplex: line segment,
- 2-simplex: triangle,
- 3-simplex: tetrahedron.

The positive semidefinite cone

The symbol S^n denotes the set of **symmetric $n \times n$ matrices**,

$$S^n = \{X \in \mathbb{R}^{n \times n} \mid X = X^T\},$$

The set of **positive semi-definite matrices** S^n_+,

$$S^n_+ = \{X \in S^n \mid X \succeq 0\},$$

where $X \succeq 0 \quad \Longleftrightarrow \quad \forall\, w \in \mathbb{R}^n, \quad \langle w, Xw \rangle \geq 0.$

The set of **positive-definite matrices** S^n_{++},

$$S^n_{++} = \{X \in S^n \mid X \succ 0\},$$

where $X \succ 0 \quad \Longleftrightarrow \quad \forall\, w \in \mathbb{R}^n, w \neq 0, \quad \langle w, Xw \rangle > 0.$

Note: **S^n_+ is a convex cone but not S^n_{++}.**

Definition 10. *An **affine function** $f : \mathbb{R}^n \to \mathbb{R}^m$ has the form $f(x) = Ax + b$ with $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$.*

Definition 11. *The **perspective function** $P : \mathbb{R}^{n+1} \to \mathbb{R}^n$ with domain $\mathrm{dom}\, P = \mathbb{R}^n \times \mathbb{R}_{++}$ is defined as $P(z,t) = \frac{z}{t}$.*

Definition 12. *A **linear fractional function** is a composition of the perspective with an affine function. Let $g : \mathbb{R}^{n+1} \to \mathbb{R}^{m+1}$ be affine,*

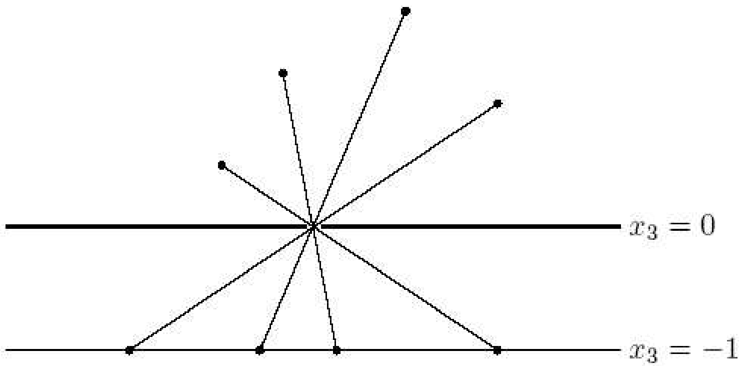
$$g(x) = \begin{pmatrix} A \\ c^T \end{pmatrix} x + \begin{pmatrix} b \\ d \end{pmatrix},$$

where $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, $d \in \mathbb{R}$ and $c \in \mathbb{R}^n$. The function $f : \mathbb{R}^n \to \mathbb{R}^m$ given by $f = P \circ g$, that is

$$f(x) = \frac{Ax + b}{c^T x + d},$$

*with domain $\mathrm{dom}\, f = \{x \,|\, c^T x + d > 0\}$ is **linear fractional**.*

Pin-hole camera interpretation of the perspective function:



Important example for a linear fractional function:

Let X and Y be random variables on $\{1, \ldots, n\}$ and $\{1, \ldots, m\}$. and denote by $p_{ij} = \mathrm{P}(X = i, Y = j)$ the probability that $X = i$ and $Y = j$. The **conditional probability**

$$f_{ij} = \mathrm{P}(X = i | Y = j) = \frac{p_{ij}}{\sum_{i=1}^{n} p_{ij}}.$$

is a linear-fractional function of p_{ij}. Given a convex set of probabilities on X and Y also the conditional probabilities will be convex.

Collection of operations that conserve convexity:

Theorem 1.

- *Let $\{C_\alpha\}_{\alpha \in I}$ be an arbitrary set of convex sets, then their* ***intersection*** *$\cap_{\alpha \in I} C_\alpha$ is convex.*
- *The* ***closure and relative interior*** *of a convex set are convex.*
- *Let $S \subseteq \mathbb{R}^n$ be a convex set and f an affine function, then the image of S under f, The* ***image of a convex set*** *S* ***under an affine function*** *$f : \mathbb{R}^n \to \mathbb{R}^m$,*

$$f(S) = \{f(x) \in \mathbb{R}^m \mid x \in S\},$$

is convex. Similarly, the ***pre-image of a convex set*** *S* ***under an affine function*** *$f : \mathbb{R}^k \to \mathbb{R}^n$,*

$$f^{-1}(S) = \{x \in \mathbb{R}^k \mid f(x) \in S\},$$

is convex.

Prominent examples: ***scaling***, ***translation*** *and* ***projection***.

A more indirect example of the image of an affine function:

The **cartesian or direct product** $C_1 \times C_2$ of two convex sets $C_1 \subseteq \mathbb{R}^n, C_2 \subseteq \mathbb{R}^m$ (turned into a vector space) is convex

$$\lambda\Big(x_1, x_2\Big) + (1-\lambda)\Big(y_1, y_2\Big) = \Big(\lambda x_1 + (1-\lambda)x_2,\ \lambda y_1 + (1-\lambda)y_2\Big) \in C_1 \times C_2.$$

Then the function $f : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ defined as $f(x_1, x_2) = x_1 + x_2$ preserves convexity, therefore **the sum** $C_1 + C_2$ defined as

$$C_1 + C_2 = \{x + y \mid x \in C_1,\ y \in C_2\}$$

is convex.

Collection of operations that conserve convexity (continued):

Theorem 2. • *If $C \subseteq \operatorname{dom} P$ is convex, then its* ***image*** *$P(C) = \{P(x) \,|\, x \in C\}$ is convex. The* ***pre-image of a convex set under the perspective function*** *is also convex. If $C \subseteq \mathbb{R}^n$ is convex, then*

$$P^{-1}(C) = \{(x,t) \in \mathbb{R}^{n+1} \mid x/t \in C, \quad t > 0\},$$

is convex.

- *Using the result about the perspective and affine function, it follows immediately that the* ***image of a convex set C under a linear fractional function*** *f is convex given that C lies in the domain of f (that means $c^T x + d > 0$ for all $x \in C$).*

A proper cone

Definition 13. *A cone $K \subseteq \mathbb{R}^n$ is called a* ***proper cone*** *if the following holds*

- *K is convex,*
- *K is closed,*
- *K has nonempty interior (K is solid),*
- *K is pointed, that is it contains no line or equivalently $x \in K$ and $-x \in K$ imply $x = 0$.*

Note: The last property excludes $\mathbb{R}^n$, the last and second last excludes hyperplanes and half-spaces.

Definition 14. *A proper cone K defines a* ***generalized inequality*** *which is a partial ordering on $\mathbb{R}^n$,*

$$x \preceq_K y \quad \Longleftrightarrow \quad y - x \in K.$$

A strict partial ordering can be defined via

$$x \prec_K y \quad \Longleftrightarrow \quad y - x \in \operatorname{int} K.$$

Examples:

- $K = \{x \in \mathbb{R}^n \,|\, x_i \geq 0,\ i = 1, \ldots, n\}$, (positive orthant) generalized inequality $\preceq_K$: componentwise inequality,
- for the set of symmetric matrices S^n, $K = S^n_+$ the set of positive semi-definite matrices,
  $A \preceq_K B \quad \Longleftrightarrow \quad B - A \in S^n_+$

Definition 15. *A* ***partial order*** $\leq$ *over a set* A *is a binary relation on* A*, which fulfills for all* $a, b, c \in A$*,*

- *reflexivity:* $a \leq a$*,*
- *antisymmetry:* $a \leq b$ *and* $b \leq a \Rightarrow a = b$*,*
- *transitivity:* $a \leq b$ *and* $b \leq c \Rightarrow a \leq c$*.*

A set with a partial order is called ***partially ordered set*** *or* ***poset****.*

A **generalized inequality** induces a **partial order** and fulfills additionally

- preserved under addition, $x \preceq_K y$ and $u \preceq_K v \;\Rightarrow\; x + u \preceq_K y + v$,
- preserved under scaling, $x \preceq_K y \;\Rightarrow\; \alpha\, x \preceq_K \alpha\, y$,
- preserved under limits, $x_t \preceq_K y_t \;\Rightarrow\; x \preceq_K y$,
  where $x = \lim_{t\to\infty} x_t$, $y = \lim_{t\to\infty} y_t$.

Minimum and minimal element

In a partial order not any two elements of a set can be compared !

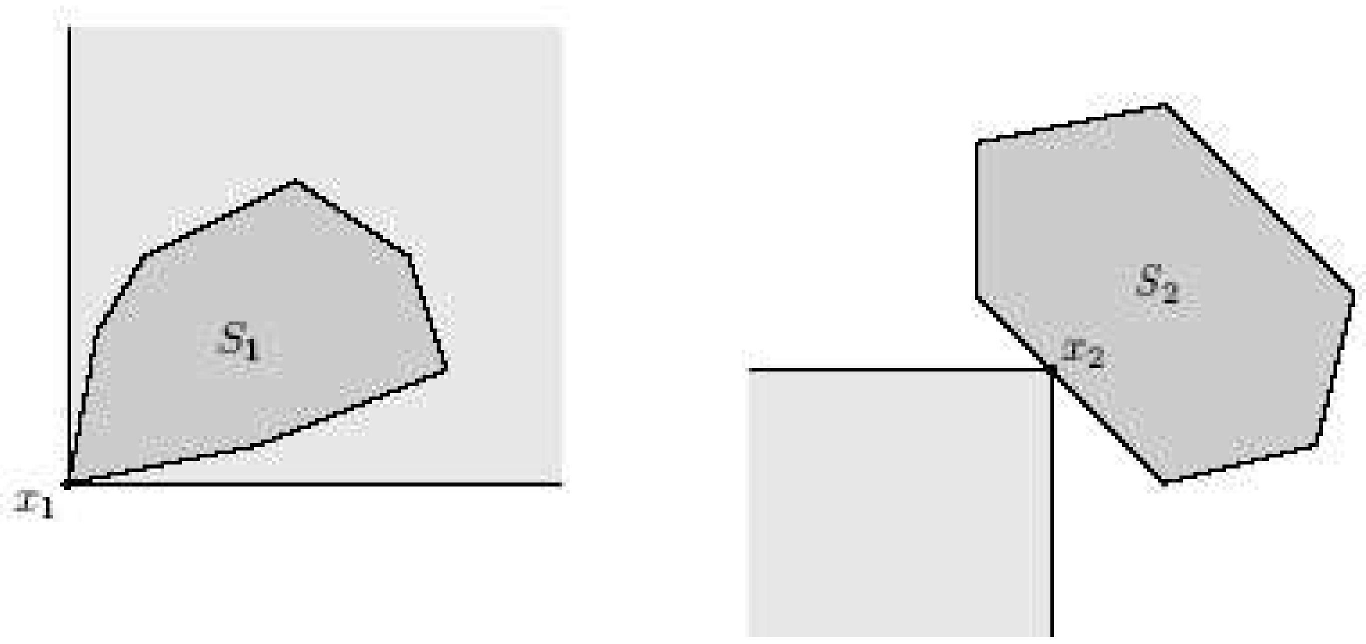
Definition 16. *If for every $y \in C$, we have $x \preceq_K y$, then we say that x is the* ***minimum element*** *of C with respect to the generalized inequality $\preceq_K$ (the maximum element can be defined analogously).*

$$C \subseteq \{x\} + K.$$

If for every $y \in C$ we have $y \preceq_K x$ only if $y = x$, then we say that x is a ***minimal element*** *of C with respect to the generalized inequality.*

$$(x - K) \cap C = \{x\}.$$

Minimum and minimal element



Left: The minimum of a convex set, Right: A minimal element of a convex set.

Theorem 3. *Let C_1 and C_2 be two nonempty and disjoint convex subsets of $\mathbb{R}^n$, then there exists* ***a hyperplane that separates*** *them:*

$$\exists\, a \in \mathbb{R}^n \text{ such that } \langle a, x_1 \rangle \leq \langle a, x_2 \rangle \quad \forall\, x_1 \in C_1,\, x_2 \in C_2.$$

Theorem 4. *Let C_1 and C_2 be two nonempty and disjoint convex subsets of $\mathbb{R}^n$, then there exists* ***a hyperplane that separates*** *them:*

$$\exists\, a \in \mathbb{R}^n \text{ such that } \langle a, x_1 \rangle \leq \langle a, x_2 \rangle \quad \forall\, x_1 \in C_1,\, x_2 \in C_2.$$

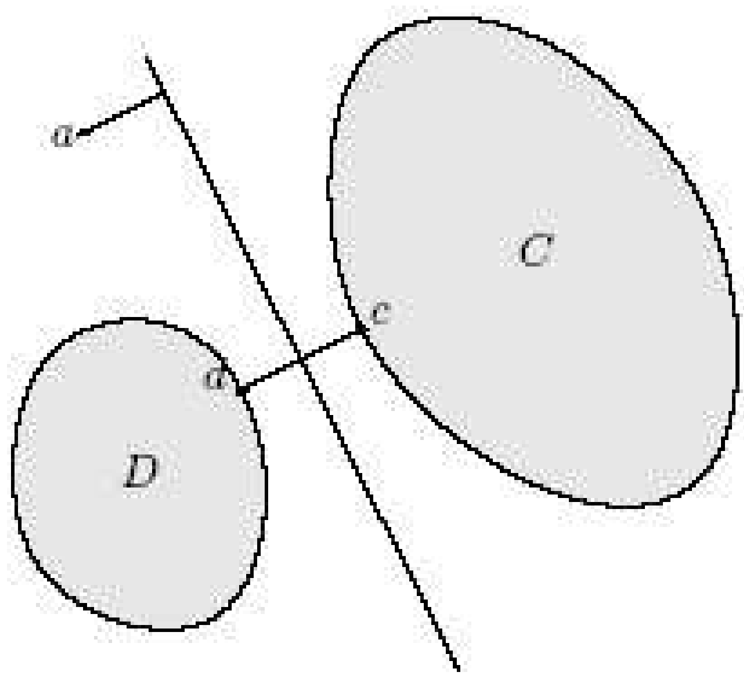
- C_1 be compact, C_2 cld, convex and disj. and $(c_1, c_2) = \arg\min_{x \in C_1,\, y \in C_2} \|x - y\|$,
- Define $w = c_2 - c_1$ and $b = \frac{\|c_2\|^2 - \|c_1\|^2}{2}$ and the function $f(x) = \langle w, x \rangle + b = \big\langle c_2 - c_1, x - \frac{1}{2}(c_2 + c_1) \big\rangle$.
- to show: f is negative on C_1 and positive on C_2.
- Suppose there exists a point z on C_2 such that $f(z) < 0$. Now

$$f(z) = \langle c_2 - c_1, z - c_2 \rangle + \frac{1}{2} \|c_2 - c_1\|^2 < 0 \quad \Longrightarrow \quad \langle c_2 - c_1, z - c_2 \rangle < 0,$$

$$\frac{\partial}{\partial t} \|c_2 + t(z - c_2) - c_1\|^2 \Big|_{t=0} = 2 \langle z - c_2, c_2 - c_1 \rangle < 0,$$

thus $t \ll 1$, $\|c_2 + t(z - c_2) - c_1\| \leq \|c_2 - c_1\|$ and $c_2 + t(z - c_2) \in C_2 \Rightarrow ↯$

Operations that conserve convexity



The construction of the separating hyerplane of two convex sets.

Proof continued

- general case: take $S = C_1 - C_2$ and $\{0\}$ $\quad (C_1 \cap C_2 = \emptyset \;\Rightarrow\; \{0\} \notin S)$,
- we assume $\{0\} \in \overline{S} \;\Rightarrow\; 0 \in \partial S$,
- If S has empty interior $\Rightarrow$ S contained in hyperplane through origin,
- S non-empty interior: consider $S_{-\varepsilon} = \{z \,|\, B(z, \varepsilon) \subset S\}$ which is convex and $\{0\} \notin S_{-\varepsilon}$,
- $\Rightarrow$ separating hyperplane a_ε which strictly separates 0 from $\overline{S_{-\varepsilon}}$.
- assume that $\|a_\varepsilon\| = 1$. The bounded sequence a_ε contains a convergent subsequence with limit a'.
- Since $\langle a_\varepsilon, z\rangle > 0$ for all $z \in S_{-\varepsilon}$, we have that $\langle a', z\rangle > 0$ for all $z \in \operatorname{int} S$ and $\langle a', z\rangle \geq 0$ for all $z \in S$ and thus $\langle a', x_1\rangle \geq \langle a', x_2\rangle$.

Separating Hyperplane Theorem III

Corollary 1 (Strict Separating Hyperplane Theorem). *Let C_1 and C_2 be two nonempty and disjoint convex subsets of $\mathbb{R}^n$ such that C_1 is closed and C_2 is compact, then there exists a hyperplane that strictly separates them:*

$$\exists\, a \in \mathbb{R}^n,\ b \in \mathbb{R},\ \textit{such that}\ \langle a, x_1\rangle < b < \langle a, x_2\rangle \quad \forall\, x_1 \in C_1,\, x_2 \in C_2.$$

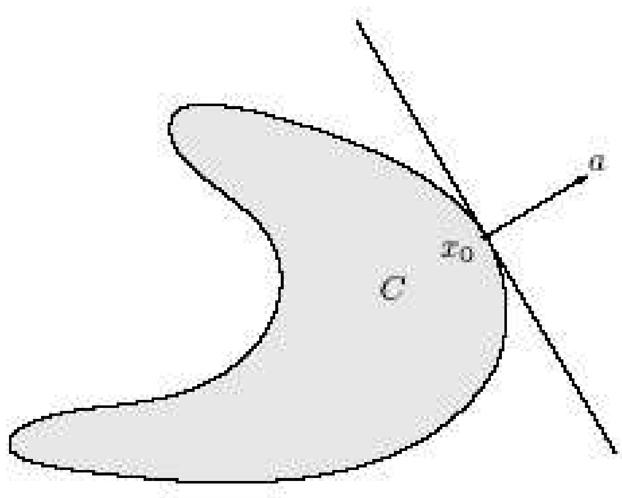
Proposition 1. *A closed convex set is the intersection of all half-spaces that contain it.*

Proof:

- S is the intersection of all half-spaces containing the closed convex set C,
- $x \in C \Rightarrow x \in S$.
- Assume $x \in S$ but $x \notin C$. By the strict separation theorem there exists a halfspace containing C but not x which implies that x cannot lie in S.

Theorem 5 (Supporting Hyperplane Theorem)**.** *Let C be a convex set in $\mathbb{R}^n$ and x' a point that does not belong to the interior of C. Then there exists a vector $a \neq 0$ such that*

$$\langle a, x \rangle \geq \langle a, x' \rangle, \quad \forall\, x \in C.$$

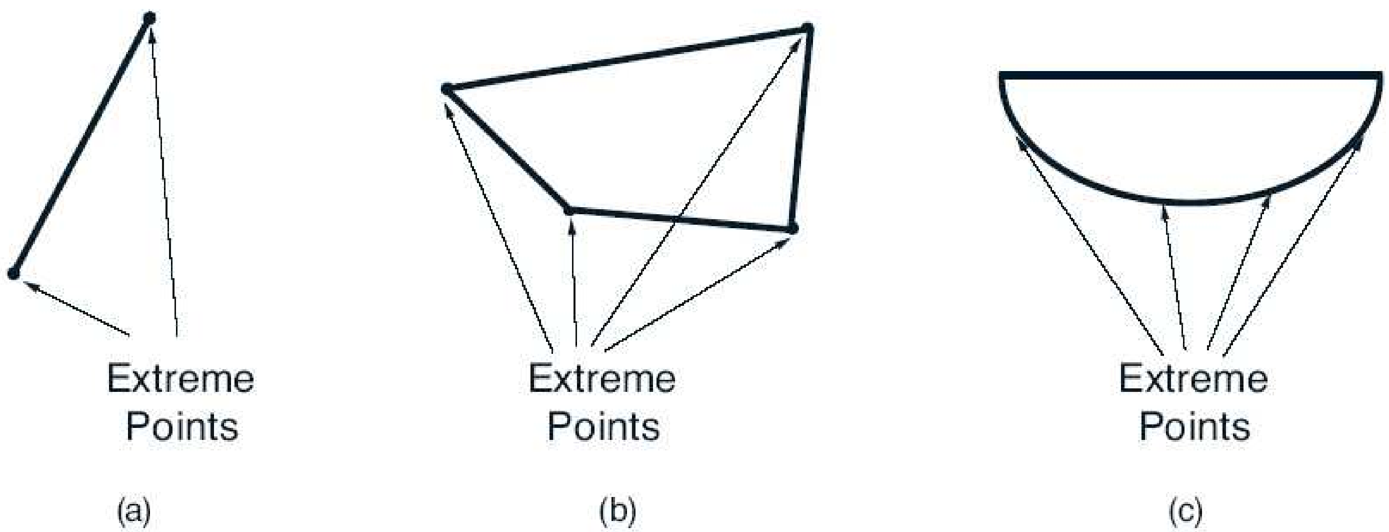


The supporting hyperplane of a (non-convex) set at x_0.

Extreme Points:

Definition 17. *Let C be a convex set. Then $x \in C$ is an* ***extreme point*** *if there exist no $y, z \in C$ with $y \neq x$ and $z \neq x$ such that for $0 < \alpha < 1$,*

$$x = \alpha y + (1 - \alpha) z.$$

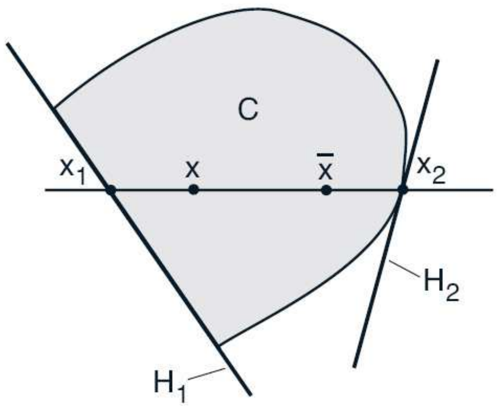


Properties of extreme points:

Proposition 2. *Let C be a nonempty, closed convex set in $\mathbb{R}^n$. Then C has at least one extreme point if and only if it does not contain a line, that is, a set L of the form $L = \{x + \alpha\, d \,|\, \alpha \in \mathbb{R}\}$ with $d \neq 0$.*

The next theorem shows that extreme points are "generators" of convex sets.

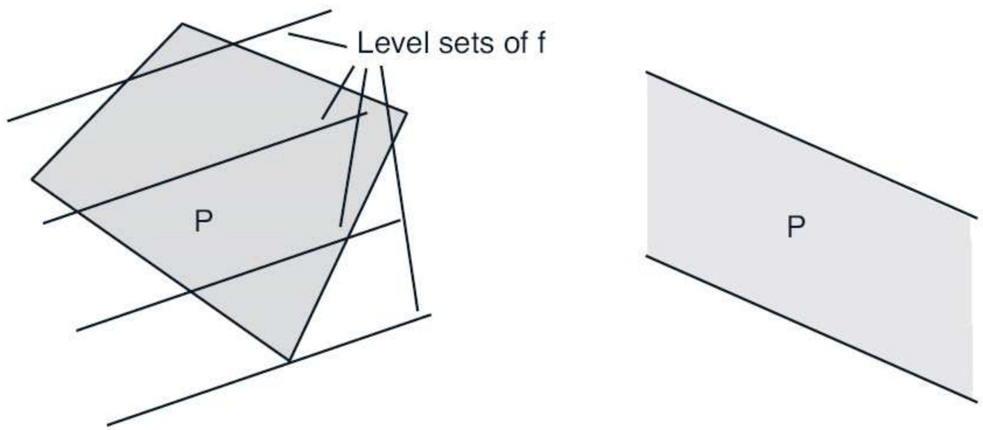
Theorem 6 (Krein-Milman-Theorem)**.** *A convex and compact set C is equal to the convex hull of its extreme points.*



Extreme points as minima of concave functions:

Proposition 3. *Let C be a convex subset of $\mathbb{R}^n$ and let C^* be the set of minima of a concave function $f : \mathbb{R}^n \to \mathbb{R}$ with* $\operatorname{dom} f = C$. *Then*

- *If C^* contains a relative interior point of C, then f must be constant over $C \implies C^* = C$.*
- *If C is closed and contains at least one extreme point, and C^* is nonempty, then C^* contains some extreme point of C.*



Basis for the simplex algorithm in Linear Programming

Proposition 4. *Let P be a polyhedral subset of $\mathbb{R}^n$. If P has the form*

$$P = \{x \,|\, \langle a_j, x\rangle \le b_j, \quad j = 1, \ldots, r\},$$

where $a_j \in \mathbb{R}^n$ and $b_j \in \mathbb{R}$, then a vector $v \in P$ is an extreme point if and only if the set

$$\{a_j \,|\, \langle a_j, v\rangle = b_j, \quad j = 1, \ldots, r\},$$

contains n linearly independent vectors.

Proposition 5 (Fundamental Theorem of Linear Programming)**.** *Let P be a polyhedral set that has at least one extreme point. Then, if a linear function attains a minimum over P, it attains a minimum at some extreme point of P.*