Operations that preserve convexity

- Examples

1. The polyhedron $\{\mathbf{x} | \mathbf{A}\mathbf{x} \preceq \mathbf{b}, \mathbf{C}\mathbf{x} = \mathbf{d}\}$ can be expressed as the inverse image of the Cartesian product of the nonnegative orthant and the origin under the affine function $f(\mathbf{x}) = (\mathbf{b} - \mathbf{A}\mathbf{x},\ \mathbf{d} - \mathbf{C}\mathbf{x})$:

$$\{\mathbf{x} | \mathbf{A}\mathbf{x} \preceq \mathbf{b}, \mathbf{C}\mathbf{x} = \mathbf{d}\} = \{\mathbf{x} | f(\mathbf{x}) \in \mathbb{R}^m_+ \times \{0\}\}.$$

2. The condition

$$A(\mathbf{x}) = x_1 \mathbf{A}_1 + ... + x_n \mathbf{A}_n \preceq \mathbf{B},$$

where $\mathbf{B}, \mathbf{A}_i \in \mathbb{S}^m$, is called a *linear matrix inequality (LMI)* in $\mathbf{x}$.

The solution set of a linear matrix inequality, $\{\mathbf{x} | A(\mathbf{x}) \preceq \mathbf{B}\}$, is convex. Indeed, it is the inverse image of the positive semidefinite cone under the affine function $f : \mathbb{R}^n \to \mathbb{S}^m$ given by $f(\mathbf{x}) = \mathbf{B} - A(\mathbf{x})$.

Operations that preserve convexity

- Examples

3. The set

$$\{\mathbf{x} | \mathbf{x}^T \mathbf{P} \mathbf{x} \leq (\mathbf{c}^T \mathbf{x})^2,\ \mathbf{c}^T \mathbf{x} \geq 0\}$$

where $\mathbf{P} \in \mathbb{S}_+^n$ and $\mathbf{c} \in \mathbb{R}^n$, is convex, since it is the inverse image of the second-order cone,

$$\{(\mathbf{z}, t) |\ \mathbf{z}^T \mathbf{z} \leq t^2, t \geq 0\},$$

under the affine function $f(\mathbf{x}) = (\mathbf{P}^{1/2}\mathbf{x},\ \mathbf{c}^T \mathbf{x})$.

4. The ellipsoid

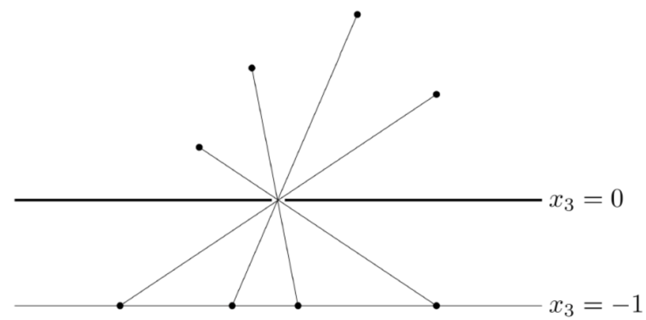
$$\epsilon = \{\mathbf{x} | (\mathbf{x} - \mathbf{x}_c)^T \mathbf{P}^{-1} (\mathbf{x} - \mathbf{x}_c) \leq 1\},$$

where $\mathbf{P} \in \mathbb{S}_{++}^n$, is the image of the unit Euclidean ball $\{\mathbf{u} |\ \|\mathbf{u}\|_2 \leq 1\}$ under the affine mapping $f(\mathbf{u}) = \mathbf{P}^{1/2}\mathbf{u} + \mathbf{x}_c$. (It is also the inverse image of the unit ball under the affine mapping $g(\mathbf{x}) = \mathbf{P}^{-1/2}(\mathbf{x} - \mathbf{x}_c)$.)

Operations that preserve convexity

- Perspective functions

We define the *perspective function* $P : \mathbb{R}^{n+1} \to \mathbb{R}^n$, with domain $\mathbf{dom} \quad P = \mathbb{R}^n \times \mathbb{R}_{++}$, as $P(\mathbf{z}, t) = \mathbf{z}/t$.



The inverse image of a convex set under the perspective function is also convex: if $C \subseteq \mathbb{R}^n$ is convex, then

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

is convex.

Question: If function f preserves convexity: if C_1 is convex then $f(C_1)$ is also convex, does f^{-1} also perserve convexity?

Operations that preserve convexity

- Linear-fractional functions

A linear-fractional function is formed by composing the perspective function with an affine function. Suppose $g : \mathbb{R}^n \to \mathbb{R}^{m+1}$ is affine, i.e.,

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

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

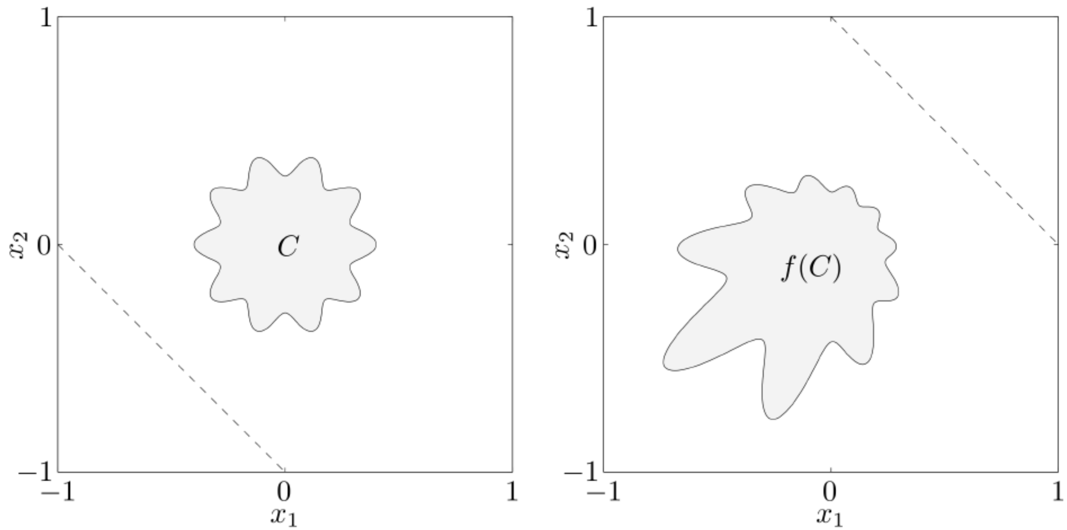
$$f(\mathbf{x}) = (\mathbf{A}\mathbf{x} + \mathbf{b})/(\mathbf{c}^T\mathbf{x} + d), \operatorname{dom} f = \{\mathbf{x} | \mathbf{c}^T\mathbf{x} + d > 0\},$$

is called a *linear-fractional* (or projective) function.

Operations that preserve convexity

- Linear-fractional functions

$$f(\mathbf{x}) = \frac{1}{\mathbf{x}_1 + \mathbf{x}_2 + 1}\mathbf{x}, \ \ \mathbf{dom}\ f = \{(x_1, x_2) | x_1 + x_2 + 1 > 0\}.$$



Conditional probabilities: Let $p_{ij} = \mathbb{P}(u = i, v = j)$. Then the conditional probability

$$f_{ij} = \frac{p_{ij}}{\sum_{k=1}^{n} p_{kj}}$$

is obtained by a linear-fractional mapping from $\mathbf{p}$.

Generalized inequalities

- Proper cones and generalized inequalities

A cone $K \subseteq \mathbb{R}^n$ is called a *proper cone* if it satisfies the following:

- K is convex.
- K is closed.
- K is solid, which means it has nonempty interior.
- K is pointed, which means that it contains no line (or equivalently, $\mathbf{x} \in K, -\mathbf{x} \in K \implies \mathbf{x} = 0$).

We associate the proper cone K with the partial ordering on $\mathbb{R}^n$ defined by

$$\mathbf{x} \preceq_K \mathbf{y} \iff \mathbf{y} - \mathbf{x} \in K.$$

We also write $\mathbf{x} \succeq_K \mathbf{y}$ for $\mathbf{y} \preceq_K \mathbf{x}$. Similarly, we define an associated strict partial ordering by

$$\mathbf{x} \prec_K \mathbf{y} \iff \mathbf{y} - \mathbf{x} \in K^\circ,$$

and write $\mathbf{x} \succ_K \mathbf{y}$ for $\mathbf{y} \prec_K \mathbf{x}$.

Generalized inequalities

- Examples

1. When $K = \mathbb{R}_+$, the partial ordering $\preceq_K$ is the usual ordering $\leq$ on $\mathbb{R}$, and the strict partial ordering $\prec_K$ is the same as the usual strict ordering $<$ on $\mathbb{R}$.

2. *Nonnegative orthant and componentwise inequality:* The nonnegative orthant $K = \mathbb{R}^n_+$ is a proper cone. The associated generalized inequality $\preceq_K$ corresponds to componentwise inequality between vectors: $\mathbf{x} \preceq_K \mathbf{y}$ means that $\mathbf{x}_i \leq \mathbf{y}_i, i = 1, ..., n$. The associated strict inequality corresponds to componentwise strict inequality: $\mathbf{x} \prec_K \mathbf{y}$ means that $\mathbf{x}_i < \mathbf{y}_i, i = 1, ..., n$.

$\boxed{\text{For simplicity, we write } \mathbf{x} \leq \mathbf{y} \text{ and } \mathbf{x} < \mathbf{y} \text{ instead of } \mathbf{x} \preceq_{\mathbb{R}^n_+} \mathbf{y} \text{ and } \mathbf{x} \prec_{\mathbb{R}^n_{++}} \mathbf{y}}$

3. *Positive semidefinite cone and matrix inequality*: The positive semidefinite cone S^n_+ is a proper cone in S^n. The associated generalized inequality $\preceq_K$ is the usual matrix inequality: $\mathbf{X} \preceq_K \mathbf{Y}$ means $\mathbf{Y} - \mathbf{X}$ is positive semidefinite. The interior of $\mathbb{S}^n_+$ (in $\mathbb{S}^n$) consists of the positive definite matrices, so the strict generalized inequality also agrees with the usual strict inequality between symmetric matrices: $\mathbf{X} \prec_K \mathbf{Y}$ means $\mathbf{Y} - \mathbf{X}$ is positive definite.

$\boxed{\text{For simplicity, we write } \mathbf{X} \preceq \mathbf{Y} \text{ and } \mathbf{X} \prec \mathbf{Y} \text{ instead of } \mathbf{X} \preceq_{\mathbb{S}^n_+} \mathbf{Y} \text{ and } \mathbf{X} \prec_{\mathbb{S}^n_+} \mathbf{Y}}$

Generalized inequalities

- Examples

4. *Cone of polynomials nonnegative on [0, 1]*: Let K be defined as

$$K = \{\mathbf{c} \in \mathbb{R}^n |\ c_1 + c_2 t + ... + c_n t^{n-1} \geq 0 \text{ for } t \in [0, 1]\},$$

i.e., K is the cone of (coefficients of) polynomials of degree $n - 1$ that are nonnegative on the interval [0, 1]. It can be shown that K is a proper cone, its interior is the set of coefficients of polynomials that are positive on the interval [0, 1].

Two vectors $\mathbf{c},\ \mathbf{d} \in \mathbb{R}^n$ satisfy $\mathbf{c} \prec_K \mathbf{d}$ if and only if

$$c_1 + c_2 t + ... + c_n t^{n-1} \leq d_1 + d_2 t + ... + d_n t^{n-1}$$

for all $t \in [0, 1]$.

Generalized inequalities

- Properties of generalized inequalities

- $\preceq_K$ is preserved under addition: if $\mathbf{x} \preceq_K \mathbf{y}$ and $\mathbf{u} \preceq_K \mathbf{v}$, then $\mathbf{x} + \mathbf{u} \preceq_K \mathbf{y} + \mathbf{v}$.
- $\preceq_K$ is transitive: if $\mathbf{x} \preceq_K \mathbf{y}$ and $\mathbf{y} \preceq_K \mathbf{z}$ then $\mathbf{x} \preceq_K \mathbf{z}$.
- $\preceq_K$ is preserved under nonnegative scaling: if $\mathbf{x} \preceq_K \mathbf{y}$ and $\alpha \geq 0$ then $\alpha\mathbf{x} \preceq_K \alpha\mathbf{y}$.
- $\preceq_K$ is reflexive: $\mathbf{x} \preceq_K \mathbf{x}$.
- $\preceq_K$ is antisymmetric: if $\mathbf{x} \preceq_K \mathbf{y}$ and $\mathbf{y} \preceq_K \mathbf{x}$, then $\mathbf{x} = \mathbf{y}$.
- $\preceq_K$ is preserved under limits: if $\mathbf{x}_i \preceq_K \mathbf{y}_i$ for $i = 1, 2, ..., \mathbf{x}_i \to \mathbf{x}$ and $\mathbf{y}_i \to \mathbf{y}$ as $i \to \infty$, then $\mathbf{x} \preceq_K \mathbf{y}$.

Generalized inequalities

- Properties of generalized inequalities

- if $\mathbf{x} \prec_K \mathbf{y}$ then $\mathbf{x} \preceq_K \mathbf{y}$.
- if $\mathbf{x} \prec_K \mathbf{y}$ and $\mathbf{u} \preceq_K \mathbf{v}$ then $\mathbf{x} + \mathbf{u} \prec_K \mathbf{y} + \mathbf{v}$.
- if $\mathbf{x} \prec_K \mathbf{y}$ and $\alpha > 0$ then $\alpha\mathbf{x} \prec_K \alpha\mathbf{y}$.
- $\mathbf{x} \not\prec_K \mathbf{x}$.
- if $\mathbf{x} \prec_K \mathbf{y}$, then for $\mathbf{u}$ and $\mathbf{v}$ small enough, $\mathbf{x} + \mathbf{u} \prec_K \mathbf{y} + \mathbf{v}$.

Generalized inequalities

- Minimum and minimal elements

We say that $\mathbf{x} \in S$ is the *minimum element* of S (with respect to the generalized inequality $\preceq_K$) if for every $\mathbf{y} \in S$ we have $\mathbf{x} \preceq_K \mathbf{y}$. We define the *maximum element* of a set S, with respect to a generalized inequality, in a similar way. If a set has a minimum (maximum) element, then it is unique. A related concept is minimal element. We say that $\mathbf{x} \in S$ is a *minimal element* of S (with respect to the generalized inequality $\preceq_K$) if $\mathbf{y} \in S$, $\mathbf{y} \preceq_K \mathbf{x}$ only if $\mathbf{y} = \mathbf{x}$. We define *maximal element* in a similar way. A set can have many different minimal (maximal) elements.

total ordering vs. *partial ordering*:

1. Reflexivity: $a \preceq a$, for all $a \in \mathcal{A}$;
2. Antisymmetry: $a \preceq b$ and $b \preceq a$ imply $a = b$;
3. Transitivity: $a \preceq b$ and $b \preceq c$ imply $a \preceq c$;
4. Comparibility: for all a and b in $\mathcal{A}$, either $a \preceq b$ or $b \preceq a$.

Generalized inequalities

- Minimum and minimal elements: set notation

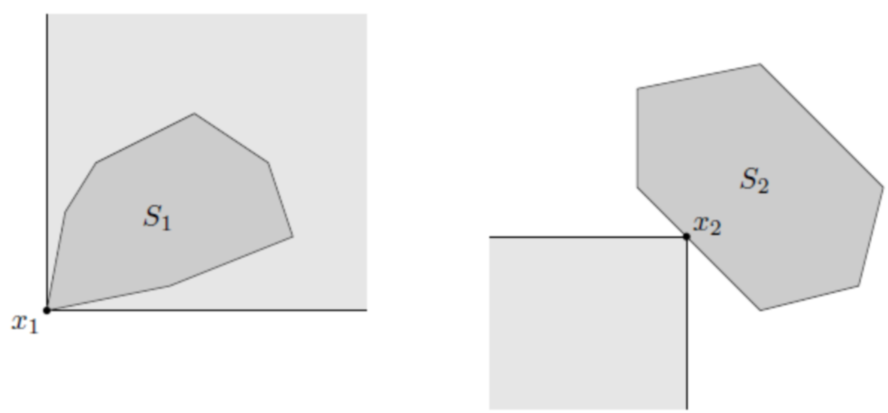
A point $\mathbf{x} \in S$ is the *minimum element* of S if and only if

$$S \subseteq \mathbf{x} + K.$$

Here $\mathbf{x} + K$ denotes all the points that are comparable to $\mathbf{x}$ and greater than or equal to $\mathbf{x}$ (according to $\preceq_K$). A point $\mathbf{x} \in S$ is a *minimal element* if and only if

$$(\mathbf{x} - K) \cap \; S = \{\mathbf{x}\}.$$

Here $\mathbf{x} - K$ denotes all the points that are comparable to $\mathbf{x}$ and less than or equal to $\mathbf{x}$ (according to $\preceq_K$), the only point in common with S is $\mathbf{x}$.



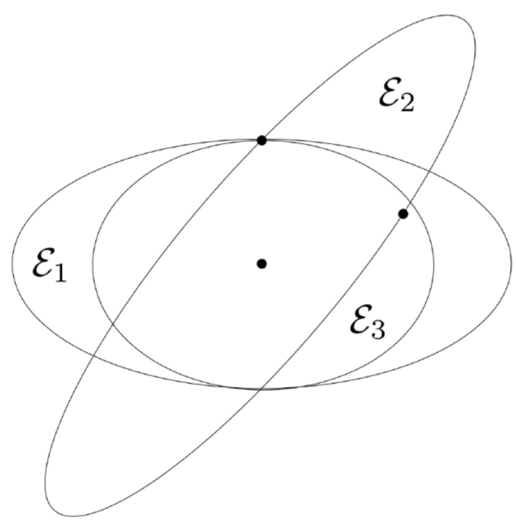
Generalized inequalities

- Minimum and minimal elements: examples

We associate with each $\mathbf{A} \in \mathbb{S}^n_{++}$ an ellipsoid centered at the origin, given by

$$\varepsilon_{\mathbf{A}} = \{\mathbf{x} |\ \mathbf{x}^T \mathbf{A}^{-1} \mathbf{x} \leq 1\}.$$

We have $\mathbf{A} \preceq \mathbf{B}$ if and only if $\varepsilon_{\mathbf{A}} \subseteq \varepsilon_{\mathbf{B}}$. Let $\mathbf{v}_1, \ldots, \mathbf{v}_k \in \mathbb{R}^n$ be given and define S to be the set of ellipsoids that contain these points. The set S does not have a minimum element: for any ellipsoid that contains the points $\mathbf{v}_1, \ldots, \mathbf{v}_k$ we can find another one that contains the points, and is not comparable to it. An ellipsoid is minimal if it contains the points, but no smaller ellipsoid does.

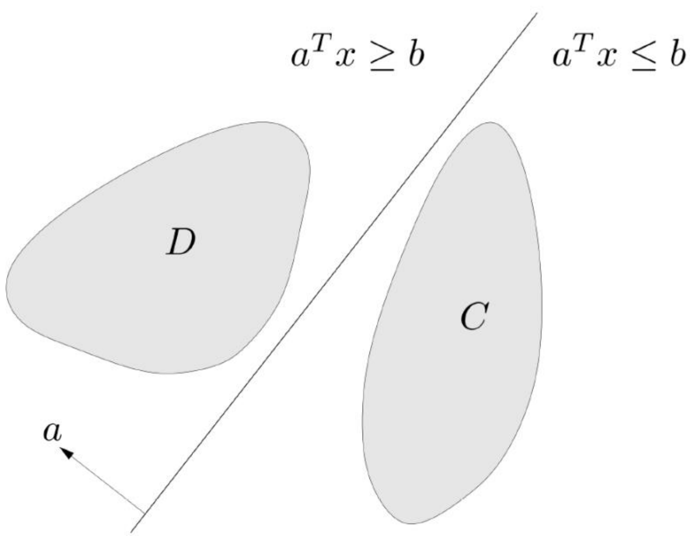


Separating and supporting hyperplanes

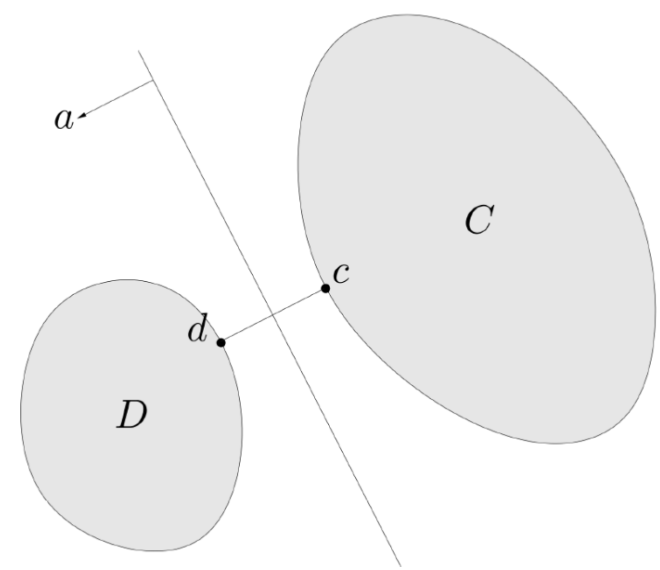
- Separating hyperplane theorem

Theorem 1. *Suppose C and D are two convex sets that do not intersect, i.e., $C \cap D = \emptyset$. Then there exist $\mathbf{a} \neq \mathbf{0}$ and b such that $\mathbf{a}^T\mathbf{x} \leq b$ for all $\mathbf{x} \in C$ and $\mathbf{a}^T\mathbf{x} \geq b$ for all $\mathbf{x} \in D$. In other words, the affine function $\mathbf{a}^T\mathbf{x} - b$ is nonpositive on C and nonnegative on D.*

The hyperplane $\{\mathbf{x} | \mathbf{a}^T\mathbf{x} = b\}$ is called a *separating hyperplane* for the sets C and D, or is said to separate the sets C and D.



Proof:



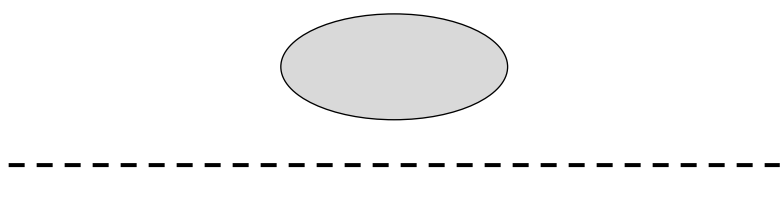
Separating and supporting hyperplanes

- Separating hyperplane theorem: example

Suppose C is convex and D is affine, i.e., $D = \{\mathbf{F}\mathbf{u} + \mathbf{g} \mid \mathbf{u} \in \mathbb{R}^m\}$, where $\mathbf{F} \in \mathbb{R}^{n \times m}$. Suppose C and D are disjoint, so by the separating hyperplane theorem there are $\mathbf{a} \neq 0$ and b such that $\mathbf{a}^T\mathbf{x} \leq b$ for all $\mathbf{x} \in C$ and $\mathbf{a}^T\mathbf{x} \geq b$ for all $\mathbf{x} \in D$.

Now $\mathbf{a}^T\mathbf{x} \geq b$ for all $\mathbf{x} \in D$ means $\mathbf{a}^T\mathbf{F}\mathbf{u} \geq b - \mathbf{a}^T\mathbf{g}$ for all $\mathbf{u} \in \mathbb{R}^m$. But a linear function is bounded below on $\mathbb{R}^m$ only when it is zero, so we conclude $\mathbf{a}^T\mathbf{F} = \mathbf{0}$ (and hence, $b \leq \mathbf{a}^T\mathbf{g}$).

Thus we conclude that there exists $\mathbf{a} \neq 0$ such that $\mathbf{F}^T\mathbf{a} = \mathbf{0}$ and $\mathbf{a}^T\mathbf{x} \leq \mathbf{a}^T\mathbf{g}$ for all $\mathbf{x} \in C$.

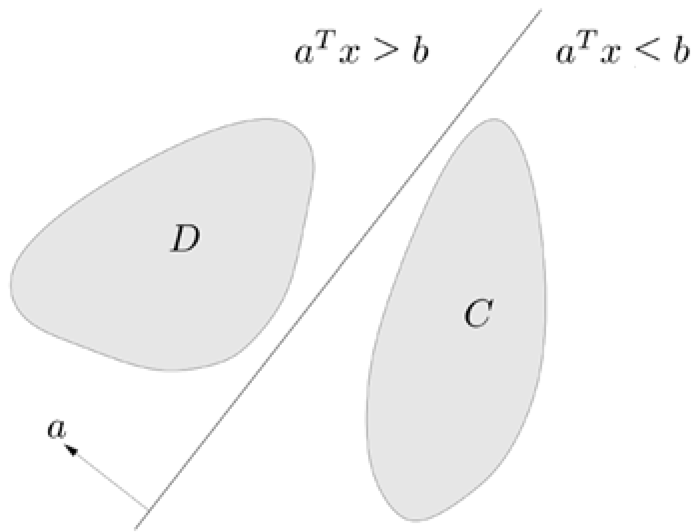


When is the separation hyperplane unique?

Separating and supporting hyperplanes

- Strict separation

If the separating hyperplane satisfies the stronger condition that $\mathbf{a}^T\mathbf{x} < b$ for all $\mathbf{x} \in C$ and $\mathbf{a}^T\mathbf{x} > b$ for all $\mathbf{x} \in D$, then the sets C and D are called *strictly separated*.



Disjoint convex sets need not be strictly separable by a hyperplane (even when the sets are closed)

Example: a point and a closed convex set

Separating and supporting hyperplanes

- Converse separating hyperplane theorems

Theorem 1. *Any two convex sets C and D, at least one of which is open, are disjoint if and only if there exists a separating hyperplane.*

Example: (Theorem of alternatives for strict linear inequalities) We derive the necessary and sufficient conditions for solvability of a system of strict linear inequalities $\boxed{\mathbf{A}\mathbf{x} < \mathbf{b}.}$

pair of alternatives

These inequalities are infeasible if and only if the (convex) sets

$$C = \{\mathbf{b} - \mathbf{A}\mathbf{x} | \mathbf{x} \in \mathbb{R}^n\}, \quad D = \mathbb{R}^m_{++} = \{\mathbf{y} \in \mathbb{R}^m | \mathbf{y} \succ \mathbf{0}\}$$

do not intersect. The set D is open, C is an affine set. Hence by the above theorem, C and D are disjoint iff there exists a separating hyperplane, i.e., a nonzero $\boldsymbol{\lambda} \in \mathbb{R}^m$ and $\mu \in \mathbb{R}$ such that $\boldsymbol{\lambda}^T\mathbf{y} \leq \mu$ on C and $\boldsymbol{\lambda}^T\mathbf{y} \geq \mu$ on D.

$\mu \leq 0$ and $\boldsymbol{\lambda} \geq \mathbf{0}$, $\boldsymbol{\lambda} \neq \mathbf{0}$. $\Rightarrow$ $\boxed{\exists \boldsymbol{\lambda} \text{ s.t. } \boldsymbol{\lambda} \neq \mathbf{0}, \boldsymbol{\lambda} \geq \mathbf{0}, \mathbf{A}^T\boldsymbol{\lambda} = \mathbf{0}, \boldsymbol{\lambda}^T\mathbf{b} \leq 0.}$

Separating and supporting hyperplanes

- Converse separating hyperplane theorems

Theorem 1 (Theorem of the Alternative (Fakas' Lemma))**.** *For* $\mathbf{A} \in \mathbb{R}^{m \times n}$ *and* $\mathbf{b} \in \mathbb{R}^m$ *the following are strong alternatives:*

1. $\exists \mathbf{x} \in \mathbb{R}^n_+$ *such that* $\mathbf{A}\mathbf{x} = \mathbf{b}$,
2. $\exists \mathbf{y} \in \mathbb{R}^m$ *such that* $\mathbf{A}^T\mathbf{y} \geq \mathbf{0}$ *and* $\mathbf{b}^T\mathbf{y} < 0$.

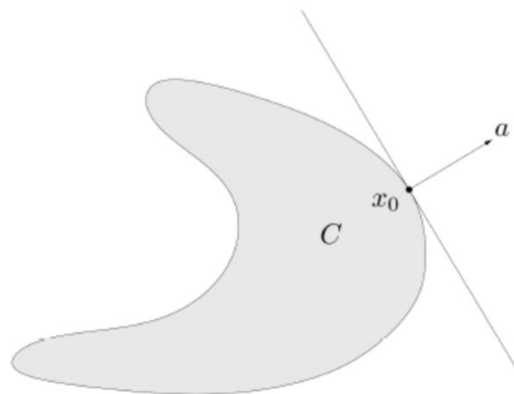
Proof. 1) $\implies \neg$2): For $\mathbf{x} \in \mathbb{R}^n_+$ with $\mathbf{A}\mathbf{x} = \mathbf{b}$ and $\mathbf{y} \in \mathbb{R}^m$ with $\mathbf{A}^T\mathbf{y} \geq 0$ we have $\mathbf{b}^T\mathbf{y} = \mathbf{x}^T\mathbf{A}^T\mathbf{y} \geq 0$.

$\neg$1) $\implies$ 2): $C := cone(\mathbf{A})$ is a closed convex cone which does not contain the vector $\mathbf{b}$: by the Separating Hyperplane Theorem there exists a $\mathbf{y} \in \mathbb{R}^m$ with $\langle \mathbf{y}, \mathbf{x} \rangle \geq 0 > \langle \mathbf{y}, \mathbf{b} \rangle$ for all $\mathbf{x} \in C$, in particular $\mathbf{A}_i^T\mathbf{y} = \langle \mathbf{y}, \mathbf{A}_i \rangle \geq 0$, $\forall i$, that is, $\mathbf{A}^T\mathbf{y} \geq \mathbf{0}$.

Separating and supporting hyperplanes

- Supporting hyperplanes

Suppose $C \subseteq \mathbb{R}^n$, and $\mathbf{x}_0$ is a point in its boundary ∂C. If $\mathbf{a} \neq \mathbf{0}$ satisfies $\mathbf{a}^T\mathbf{x} \leq \mathbf{a}^T\mathbf{x}_0$ for all $\mathbf{x} \in C$, then the hyperplane $\{\mathbf{x} | \ \mathbf{a}^T\mathbf{x} = \mathbf{a}^T\mathbf{x}_0\}$ is called a *supporting hyperplane* to C at the point $\mathbf{x}_0$.



Theorem 1 (Supporting Hyperplane Theorem). *For any nonempty convex set C, and any $\mathbf{x}_0 \in \partial C$, there exists a supporting hyperplane to C at $\mathbf{x}_0$.*