EE/ACM 150 - Applications of Convex Optimization in Signal Processing and Communications
Lecture 6

Andre Tkacenko

Signal Processing Research Group
Jet Propulsion Laboratory

April 19, 2012

Outline

1 Relevant Sets in Convex Optimization
- Affine Sets
- Convex Sets
- Cones

2 Important Example Sets

3 Operations That Preserve Convexity

4 Generalized Inequalities
- Proper Cones
- Minimum and Minimal Elements

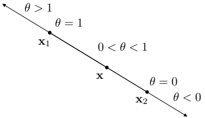
5 Separating and Supporting Hyperplanes

6 Dual Cones
- Role in Generalized Inequalities
- Dual Characterization of Minimum and Minimal Elements

Affine Sets: Definition and Properties

A set $\mathcal{C} \in \mathbb{R}^n$ is *affine* if the line between any two distinct points in $\mathcal{C}$ is also in $\mathcal{C}$. In other words, if $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{C}$, then all points $\mathbf{x}$ of the form

$$\mathbf{x} = \theta \mathbf{x} + (1-\theta)\,\mathbf{x}_2 = \mathbf{x}_2 + \theta\,(\mathbf{x}_1 - \mathbf{x}_2)\ , \text{ are in } \mathcal{C} \text{ for all } \theta \in \mathbb{R}\,.$$



This can be generalized to more than two points. Specifically, if $\mathbf{x}_1, \ldots, \mathbf{x}_k \in \mathcal{C}$, then $\mathcal{C}$ is affine if and only if any *affine combination* $\mathbf{x}$ of the points $\mathbf{x}_1, \ldots, \mathbf{x}_k$ is also in $\mathcal{C}$, where $\mathbf{x}$ is given by

$$\mathbf{x} = \theta_1 \mathbf{x}_1 + \cdots + \theta_k \mathbf{x}_k\ , \text{ where } \theta_1 + \cdots + \theta_k = 1\,.$$

Properties:

- Any affine set $\mathcal{C}$ is simply a linear subspace $\mathcal{V}$ shifted spatially by any offset point $\mathbf{x}_0 \in \mathcal{C}$, i.e., can be represented by

$$\mathcal{C} = \mathcal{V} + \mathbf{x}_0 = \{\mathbf{v} + \mathbf{x}_0 : \mathbf{v} \in \mathcal{V}\}\ .$$

- A set $\mathcal{C} \in \mathbb{R}^n$ is affine if and only if it can be expressed as the solution set of a system of linear equations, i.e., can be represented in the form

$$\mathcal{C} = \{\mathbf{x} : \mathbf{A}\mathbf{x} = \mathbf{b}\}\ , \text{ where } \mathbf{A} \in \mathbb{R}^{m \times n} \text{ and } \mathbf{b} \in \mathbb{R}^m.$$

Affine Hull, Affine Dimension, and Relative Interior

Affine Hull:
The *affine hull* of any set $\mathcal{C} \subseteq \mathbb{R}^n$ is the set of all affine combinations of points in $\mathcal{C}$ and denoted $\text{aff}(\mathcal{C})$. In other words, we have

$$\text{aff}(\mathcal{C}) \triangleq \{\theta_1 \mathbf{x}_1 + \cdots + \theta_k \mathbf{x}_k : \mathbf{x}_1, \ldots, \mathbf{x}_k \in \mathcal{C}\,,\ \theta_1 + \cdots + \theta_k = 1\}\ .$$

The affine hull of $\mathcal{C}$ is the smallest affine set that contains $\mathcal{C}$. Namely, if $\mathcal{S}$ is any affine set with $\mathcal{C} \subseteq \mathcal{S}$, then $\text{aff}(\mathcal{C}) \subseteq \mathcal{S}$.

Affine Dimension:
The *affine dimension* of a set $\mathcal{C}$ is the dimension of its affine hull. For example, the affine hull of a set of coplanar points in $\mathbb{R}^3$ is a plane with affine dimension 2.

Relative Interior:
If the affine dimension of a set $\mathcal{C} \subseteq \mathbb{R}^n$ is less than n, then $\mathcal{C}$ lies in an affine set $\text{aff}(\mathcal{C}) \neq \mathbb{R}^n$. In this case, the interior of $\mathcal{C}$ relative to $\mathbb{R}^n$ is empty, but the interior relative to the affine hull, defined as

$$\text{relint}(\mathcal{C}) \triangleq \{\mathbf{x} \in \mathbb{C} : \mathcal{B}(\mathbf{x}, r) \cap \text{aff}(\mathcal{C}) \subseteq \mathcal{C} \text{ for some } r > 0\}\ ,$$

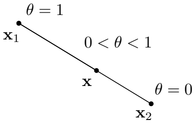
is in general not empty. Here, $\mathcal{B}(\mathbf{x}, r) \triangleq \{\mathbf{y} : ||\mathbf{y} - \mathbf{x}|| \leq r\}$ is a closed ball of radius r centered at $\mathbf{x}$ (corresponding to some norm $||\cdot||$). From this we can define a *relative boundary* of a set $\mathcal{C}$ as

$$\text{relbd}(\mathcal{C}) \triangleq \text{cl}(\mathcal{C}) \setminus \text{relint}(\mathcal{C})\ .$$

Convex Sets: Definitions and Examples

A set $\mathcal{C} \in \mathbb{R}^n$ is *convex* if the line segment between any two distinct points in $\mathcal{C}$ is also in $\mathcal{C}$. In other words, if $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{C}$, then all points $\mathbf{x}$ of the form

$$\mathbf{x} = \theta \mathbf{x} + (1 - \theta)\, \mathbf{x}_2 = \mathbf{x}_2 + \theta\, (\mathbf{x}_1 - \mathbf{x}_2) \ , \text{ are in } \mathcal{C} \text{ for all } \theta \text{ with } 0 \leq \theta \leq 1 \, .$$



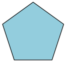
This can be generalized to more than two points. Specifically, if $\mathbf{x}_1, \ldots, \mathbf{x}_k \in \mathcal{C}$, then $\mathcal{C}$ is convex if and only if any *convex combination* $\mathbf{x}$ of the points $\mathbf{x}_1, \ldots, \mathbf{x}_k$ is also in $\mathcal{C}$, where $\mathbf{x}$ is given by

$$\mathbf{x} = \theta_1 \mathbf{x}_1 + \cdots + \theta_k \mathbf{x}_k \ , \text{ where } \theta_1 + \cdots + \theta_k = 1 \text{ and } \theta_\ell \geq 0 \text{ for } \ell = 1, \ldots, k \, .$$

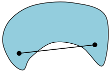
In the most general form, if $p : \mathbb{R}^n \to \mathbb{R}$ is some function which satisfies $p(\mathbf{x}) \geq 0$ for all $\mathbf{x} \in \mathcal{C}$ and $\int_{\mathcal{C}} p(\mathbf{x})\, d\mathbf{x} = 1$ (i.e., $p(\mathbf{x})$ is a *probability density function* (pdf) for $\mathbf{x} \in \mathcal{C}$), then $\mathcal{C}$ is convex if and only if

$$\int_C p(\mathbf{x})\, \mathbf{x}\, d\mathbf{x} = E[\mathbf{x}] \in \mathcal{C} \, .$$

Examples:



convex



nonconvex

nonconvex

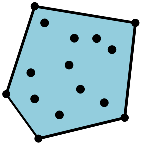
Convex Hull: Definition and Examples

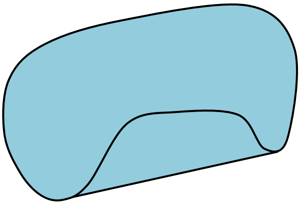
The *convex hull* of any set $\mathcal{C} \subseteq \mathbb{R}^n$ is the set of all convex combinations of points in $\mathcal{C}$ and denoted $\mathrm{conv}(\mathcal{C})$. In other words, we have

$$\mathrm{conv}(\mathcal{C}) \triangleq \{\theta_1 \mathbf{x}_1 + \cdots + \theta_k \mathbf{x}_k : \mathbf{x}_1, \ldots, \mathbf{x}_k \in \mathcal{C} \, , \; \theta_\ell \geq 0 \, , \ell = 1, \ldots, k \, , \; \theta_1 + \cdots + \theta_k = 1\} \; .$$

The convex hull of $\mathcal{C}$ is the smallest convex set that contains $\mathcal{C}$. Namely, if $\mathcal{S}$ is any convex set with $\mathcal{C} \subseteq \mathcal{S}$, then $\mathrm{conv}(\mathcal{C}) \subseteq \mathcal{S}$.

Examples:





Cones: Definitions and Examples

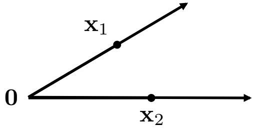
A set $\mathcal{C}$ is a *cone* if, for every $\mathbf{x} \in \mathcal{C}$ and $\theta \geq 0$, we have $\theta \mathbf{x} \in \mathcal{C}$. Similarly, a set $\mathcal{C}$ is said to be a *convex cone* if it is convex and a cone, meaning that for any $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{C}$ and $\theta_1, \theta_2 \geq 0$, we have

$$\theta_1 \mathbf{x}_1 + \theta_2 \mathbf{x}_2 \in \mathcal{C} \,.$$

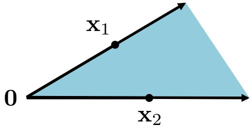
This can be generalized to more than two points. Specifically, if $\mathbf{x}_1, \ldots, \mathbf{x}_k \in \mathcal{C}$, then $\mathcal{C}$ is a convex cone if and only if any *conic combination* $\mathbf{x}$ of the points $\mathbf{x}_1, \ldots, \mathbf{x}_k$ is also in $\mathcal{C}$, where $\mathbf{x}$ is given by

$$\mathbf{x} = \theta_1 \mathbf{x}_1 + \cdots + \theta_k \mathbf{x}_k \,, \text{ where } \theta_\ell \geq 0 \text{ for } \ell = 1, \ldots, k \,.$$

Examples:



nonconvex cone



convex cone

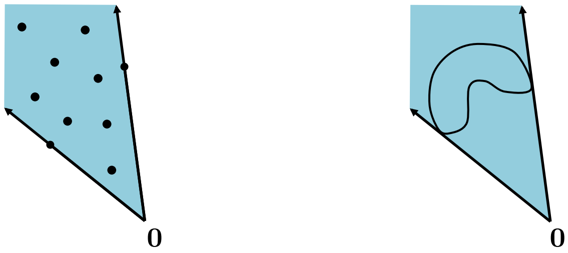
Conic Hull: Definition and Examples

The *conic hull* of any set $\mathcal{C} \subseteq \mathbb{R}^n$ is the set of all conic combinations of points in $\mathcal{C}$ and denoted $\mathrm{conc}(\mathcal{C})$. In other words, we have

$$\mathrm{conc}(\mathcal{C}) \triangleq \{\theta_1 \mathbf{x}_1 + \cdots + \theta_k \mathbf{x}_k : \mathbf{x}_1, \ldots, \mathbf{x}_k \in \mathcal{C}\,,\ \theta_\ell \geq 0\,, \ell = 1, \ldots, k\}\ .$$

The conic hull of $\mathcal{C}$ is the smallest convex cone that contains $\mathcal{C}$. Namely, if $\mathcal{S}$ is any convex cone with $\mathcal{C} \subseteq \mathcal{S}$, then $\mathrm{conc}(\mathcal{C}) \subseteq \mathcal{S}$.

Examples:



Hyperplanes and Halfspaces

A *hyperplane* is a set of the form

$$\left\{ \mathbf{x} : \mathbf{a}^T \mathbf{x} = b \right\} ,$$

where $\mathbf{a} \in \mathbb{R}^n$ with $\mathbf{a} \neq \mathbf{0}$ and $b \in \mathbb{R}$. Here, $\mathbf{a}$ represents the *normal vector* to the hyperplane and b determines the offset of the hyperplane from the origin. Equivalently, the hyperplane can be expressed in the form

$$\left\{ \mathbf{x} : \mathbf{a}^T \left(\mathbf{x} - \mathbf{x}_0 \right) = 0 \right\} ,$$

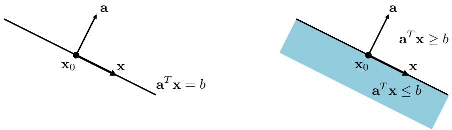
where $\mathbf{x}_0$ is any point in the hyperplane (i.e., any point for which $\mathbf{a}^T \mathbf{x}_0 = b$). This can be expressed in turn as

$$\left\{ \mathbf{x}_0 + \mathbf{a}^{\perp} : \mathbf{a}^{\perp} = \left\{ \mathbf{v} : \mathbf{a}^T \mathbf{v} = 0 \right\} \right\} .$$

In other words, the hyperplane is the orthogonal complement of $\mathbf{a}$ offset by a vector $\mathbf{x}_0$.
Any hyperplane divides $\mathbb{R}^n$ into two *halfspaces*. A halfspace is a set of the form

$$\left\{ \mathbf{x} : \mathbf{a}^T \mathbf{x} \leq b \right\} \text{ (closed) } , \left\{ \mathbf{x} : \mathbf{a}^T \mathbf{x} < b \right\} \text{ (open) } .$$

Hyperplanes are affine and convex, whereas halfspaces are only convex.



Euclidean Balls and Ellipsoids

A *Euclidean ball* in R^n with center at $\mathbf{x}_c \in \mathbb{R}^n$ and radius $r > 0$ is defined as

$$\mathcal{B}(\mathbf{x}_c, r) \triangleq \{\mathbf{x} : ||\mathbf{x} - \mathbf{x}_c||_2 \leq r\} = \left\{\mathbf{x} : (\mathbf{x} - \mathbf{x}_c)^T (\mathbf{x} - \mathbf{x}_c) \leq r^2\right\} .$$

It can be equivalently represented as

$$\mathcal{B}(\mathbf{x}_c, r) = \{\mathbf{x}_c + r\mathbf{u} : ||\mathbf{u}||_2 \leq 1\} .$$

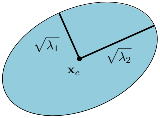
An *ellipsoid* is a set of the form

$$\mathcal{E} \triangleq \left\{\mathbf{x} : (\mathbf{x} - \mathbf{x}_c)^T \mathbf{P}^{-1} (\mathbf{x} - \mathbf{x}_c) \leq 1\right\} ,$$

where $\mathbf{P} \in \mathbb{S}^n_{++}$ (i.e., $\mathbf{P} \succ \mathbf{0}$). Here, $\mathbf{x}_c$ denotes the center of the ellipsoid $\mathcal{E}$ and the lengths of the semi-axes are $\sqrt{\lambda_k}$, where λ_k are the eigenvalues of $\mathbf{P}$. Another representation is given by

$$\mathcal{E} = \{\mathbf{x}_c + \mathbf{A}\mathbf{u} : ||\mathbf{u}||_2 \leq 1\} ,$$

where $\mathbf{A}$ is square and nonsingular. If we allow $\mathbf{A}$ to be singular, we obtain a *degenerate ellipsoid* for which $\dim(\mathrm{aff}(\mathcal{E})) = \mathrm{rank}(\mathbf{A})$.



Euclidean balls, ellipsoids, and degenerate ellipsoids are all convex.

Norm Balls and Norm Cones

A *norm ball* for the norm $||\cdot||$ with center $\mathbf{x}_c$ and radius $r > 0$ is a set $\mathcal{B}(\mathbf{x}_c, r)$ of the form

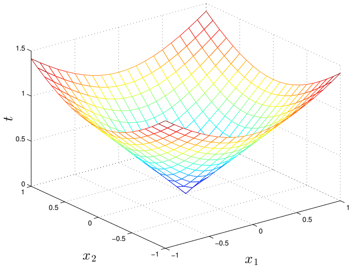
$$\mathcal{B}(\mathbf{x}_c, r) = \{\mathbf{x} : ||\mathbf{x} - \mathbf{x}_c|| \le r\} \ .$$

By exploiting the homogeneity property and triangle inequality of norms, it can be easily shown that $\mathcal{B}(\mathbf{x}_c, r)$ is always convex.

For a norm $||\cdot||$ on $\mathbb{R}^n$, the *norm cone* associated with $||\cdot||$ is the set

$$\mathcal{C} \triangleq \{(\mathbf{x}, t) : ||\mathbf{x}|| \le t\} \subseteq \mathbb{R}^{n+1} \ .$$

Any norm cone is a convex cone. A norm cone based off of the Euclidean norm is called the *second-order cone* and is shown below for $\mathbb{R}^3$.

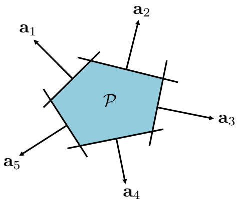


Polyhedra

A *polyhedron* $\mathcal{P} \in \mathbb{R}^n$ is the solution set of a finite number of linear inequalities and equalities of the form

$$\mathcal{P} = \{\mathbf{x} : \mathbf{A}\mathbf{x} \preceq \mathbf{b}, \mathbf{C}\mathbf{x} = \mathbf{d}\} \ ,$$

where $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{b} \in \mathbb{R}^m$, $\mathbf{C} \in \mathbb{R}^{p \times n}$, and $\mathbf{d} \in \mathbb{R}^p$. Here, $\preceq$ denotes *componentwise inequality*, i.e., $\mathbf{u} \preceq \mathbf{v}$ means $[\mathbf{u}]_k \leq [\mathbf{v}]_k$ for all k. As any polyhedron is the intersection of a finite number of halfspaces and hyperplanes, it is convex.



A special type of polyhedra are *simplexes*. If $\mathbf{v}_0, \ldots, \mathbf{v}_k \in \mathbb{R}^n$ are *affinely independent*, meaning $\mathbf{v}_1 - \mathbf{v}_0, \ldots, \mathbf{v}_k - \mathbf{v}_0$ are linearly independent, the simplex $\mathcal{S}$ determined be them is given by

$$\mathcal{S} \triangleq \text{conv}(\{\mathbf{v}_0, \ldots, \mathbf{v}_k\}) = \left\{\theta_0\mathbf{v}_0 + \cdots + \theta_k\mathbf{v}_k : \boldsymbol{\theta} \succeq \mathbf{0}, \mathbf{1}^T\boldsymbol{\theta} = 1\right\} .$$

The most common simplex is the *probability simplex* given by

$$\mathcal{S}_{\text{prob}} = \left\{\mathbf{x} : \mathbf{x} \succeq \mathbf{0}, \mathbf{1}^T\mathbf{x} = 1\right\} .$$

Positive Semidefinite Cone

One important cone is the *positive semidefinite cone* $\mathbb{S}_+^n$ defined as

$$\mathbb{S}_+^n \triangleq \{\mathbf{X} \in \mathbb{S}^n : \mathbf{X} \succeq \mathbf{0}\} \ .$$

This can be easily shown to be a convex cone. To do this, suppose $\theta_1, \theta_2 \geq 0$ and $\mathbf{A}, \mathbf{B} \in \mathbb{S}_+^n$. Then we have, for any $\mathbf{x} \in \mathbb{R}^n$,

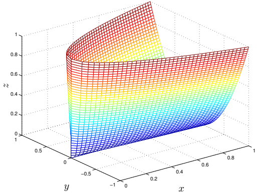
$$\mathbf{x}^T (\theta_1 \mathbf{A} + \theta_2 \mathbf{B}) \mathbf{x} = \theta_1 \mathbf{x}^T \mathbf{A} \mathbf{x} + \theta_2 \mathbf{x}^T \mathbf{B} \mathbf{x} \geq 0 \, ,$$

and so the conic combination $(\theta_1 \mathbf{A} + \theta_2 \mathbf{B}) \in \mathbb{S}_+^n$.

Example: For $\mathbf{S}_+^2$, we have

$$\mathbf{X} = \begin{bmatrix} x & y \\ y & z \end{bmatrix} \in \mathbb{S}_+^2 \iff x \geq 0 \, , \ z \geq 0 \, , \ xz \geq y^2 \, .$$

The boundary of this cone is shown below as plotted in $\mathbb{R}^3$.



The Calculus of Convex Sets

In most cases, determining whether or not a set $\mathcal{C}$ is convex can be quite cumbersome if we try to do so using the definition:

$$\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{C} \Longrightarrow \theta \mathbf{x}_1 + (1-\theta)\,\mathbf{x}_2 \in \mathcal{C}\,,\ 0 \leq \theta \leq 1\,.$$

Typically, what is done in a practical sense is to see if $\mathcal{C}$ can be represented in terms of simple convex sets (such as hyperplanes, halfspaces, norm balls, etc.) using operations that preserve convexity. These operations lead to a *calculus of convex sets*, which can be used, in many cases, to easily determine whether or not a given set is convex.

Convexity Preserving Operations:

- intersection
- affine functions
- perspective and linear-fractional functions

Intersection

The *intersection* of any number (finite or infinite) of convex sets is convex. Mathematically, if $\mathcal{I}$ is any index set with finite or infinite cardinality, and $\mathcal{S}_i$ is convex for every $i \in \mathcal{I}$, then the set $\mathcal{S} \triangleq \bigcap_{i \in \mathcal{I}} \mathcal{S}_i$ is convex.

Examples:

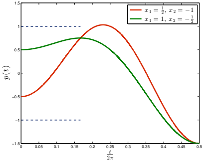
- Any polyhedron is the intersection of a finite number of halfspaces and hyperplanes and as such is convex.
- Consider the set

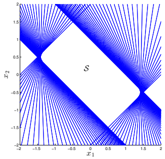
$$\mathcal{S} \triangleq \left\{ \mathbf{x} \in \mathbb{R}^m : |p(t)| \leq 1 \text{ for } |t| \leq \frac{\pi}{3} \right\} \text{, where } p(t) = \sum_{k=1}^{m} x_k \cos(kt) \ .$$

  This set can be expressed as the intersection of an infinite number of *slabs*:

$$\mathcal{S} = \bigcap_{|t| \leq \frac{\pi}{3}} \mathcal{S}_t \text{, where } \mathcal{S}_t = \left\{ \mathbf{x} : -1 \leq \begin{bmatrix} \cos(t) & \cdots & \cos(mt) \end{bmatrix}^T \mathbf{x} \leq 1 \right\} .$$

  For $m = 2$, we have the following plots.





Affine Functions

Suppose $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$ is an *affine function*, i.e., it is of the form $\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$, where $\mathbf{A} \in \mathbb{R}^{m \times n}$ and $\mathbf{b} \in \mathbb{R}^m$. Then, we have the following:

- The *image* of a convex set under $\mathbf{f}$ is convex:

$$\mathcal{S} \subseteq \mathbb{R}^n \text{ is convex} \Longrightarrow \mathbf{f}(\mathcal{S}) \triangleq \{\mathbf{f}(\mathbf{x}) : \mathbf{x} \in \mathcal{S}\} \text{ is convex}.$$

- The *inverse image* of a convex set under $\mathbf{f}$ is convex:

$$\mathcal{C} \subseteq \mathbb{R}^m \text{ is convex} \Longrightarrow \mathbf{f}^{-1}(\mathcal{C}) \triangleq \{\mathbf{x} \in \mathbb{R}^n : \mathbf{f}(\mathbf{x}) \in \mathcal{C}\} \text{ is convex}.$$

Examples:

- Scaling and translation:
  If $\mathcal{S}$ is convex, then so are $\alpha\mathcal{S} \triangleq \{\alpha\mathbf{x} : \mathbf{x} \in \mathcal{S}\}$ and $\mathcal{S} + \mathbf{a} \triangleq \{\mathbf{x} + \mathbf{a} : \mathbf{x} \in \mathcal{S}\}$.
- Projection:
  The projection of a convex set onto some of its coordinates is convex. Namely if $\mathcal{S} \subseteq \mathbb{R}^m \times \mathbb{R}^n$ is convex, then $\mathcal{T} \triangleq \{\mathbf{x}_1 \in \mathbb{R}^m : (\mathbf{x}_1, \mathbf{x}_2) \in \mathcal{S} \text{ for some } \mathbf{x}_2 \in \mathbb{R}^n\}$ is convex.
- Solution set to a linear matrix inequality (LMI):
  An LMI is of the following form

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

  where $\mathbf{A}_k, \mathbf{B} \in \mathbb{S}^m$. The solution set of the LMI, namely $\{\mathbf{x} : \mathbf{A}(\mathbf{x}) \preceq \mathbf{B}\}$, is convex as it is the inverse image of the positive semidefinite cone $\mathbb{S}_+^m$ under the affine function $\mathbf{f} : \mathbb{R}^n \to \mathbb{S}^m$ given by $\mathbf{f}(\mathbf{x}) = \mathbf{B} - \mathbf{A}(\mathbf{x})$.
- Hyperbolic cone:
  The set $\left\{\mathbf{x} : \mathbf{x}^T\mathbf{P}\mathbf{x} \leq \left(\mathbf{c}^T\mathbf{x}\right)^2,\ \mathbf{c}^T\mathbf{x} \geq 0\right\}$, where $\mathbf{P} \in \mathbb{S}_+^n$ and $\mathbf{c} \in \mathbb{R}^n$ is convex as it is the inverse image of the second-order cone under the affine function $\mathbf{f}(\mathbf{x}) = \left(\mathbf{P}^{\frac{1}{2}}\mathbf{x}, \mathbf{c}^T\mathbf{x}\right)$.

Perspective and Linear-Fractional Functions

The Perspective Function:
The *perspective function* scales or normalizes a given vector so that its last component is unity and then drops the last component. More specifically, the perspective function $\mathbf{P} : \mathbb{R}^{n+1} \rightarrow \mathbb{R}^n$ is defined as

$$\mathbf{P}(\mathbf{x}, t) = \frac{\mathbf{x}}{t}, \ \mathrm{dom}(\mathbf{P}) = \{(\mathbf{x}, t) : \mathbf{x} \in \mathbb{R}^n, t > 0\} .$$

It can be shown that the images and inverse images of convex sets under the perspective function are convex. The intuition is that a convex object, when viewed through a pin-hole camera, yields a convex image.

Linear-Fractional Functions:
A *linear-fractional function* is formed by the composition of the perspective function with an affine function. More specifically, a linear-fractional function $\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is of the form

$$\mathbf{f}(\mathbf{x}) = \frac{\mathbf{A}\mathbf{x} + \mathbf{b}}{\mathbf{c}^T\mathbf{x} + d}, \ \mathrm{dom}(\mathbf{f}) = \left\{\mathbf{x} : \mathbf{c}^T\mathbf{x} + d > 0\right\} .$$

Here, $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{b} \in \mathbb{R}^m$, $\mathbf{c} \in \mathbb{R}^n$, and $d \in \mathbb{R}$. Images and inverse images of convex sets under linear-fractional functions are convex.

Proper Cones: Definition and Examples

A cone $\mathcal{K} \subseteq \mathbb{R}^n$ is a *proper cone* if the following hold true:

- $\mathcal{K}$ is convex and closed (meaning it contains its boundary).
- $\mathcal{K}$ is solid (meaning it has a nonempty interior).
- $\mathcal{K}$ is pointed (meaning it contains no line).

Examples:

- Nonnegative orthant:
  The set of vectors in $\mathbb{R}^n$ whose components are nonnegative, given by

  $$\mathcal{K} = \mathbb{R}_+^n = \{\mathbf{x} \in \mathbb{R}^n : x_k \geq 0\,,\ k = 1, \ldots, n\}\ ,$$

  is a proper cone.
- Positive semidefinite cone:
  The set $\mathcal{K} = \mathbb{S}_+^n$ is a proper cone.
- Nonnegative polynomials on $[0, 1]$:
  The set defined as

  $$\mathcal{K} \triangleq \left\{\mathbf{c} \in \mathbb{R}^n : c_1 + c_2 t + c_3 t^2 + \cdots + c_n t^n \geq 0 \text{ for } t \in [0, 1]\right\}\ ,$$

  is a proper cone.

Proper Cones and Generalized Inequalities

A proper cone can be used to define a *generalized inequality*, which is a partial ordering similar to the standard ordering on $\mathbb{R}$. Specifically, for a proper cone $\mathcal{K}$, nonstrict and strict generalized inequalities $\preceq_{\mathcal{K}}$ and $\prec_{\mathcal{K}}$, respectively, are defined as follows:

$$\mathbf{x} \preceq_{\mathcal{K}} \mathbf{y} \Longleftrightarrow \mathbf{y} - \mathbf{x} \in \mathcal{K} \, , \; \mathbf{x} \prec_{\mathcal{K}} \mathbf{y} \Longleftrightarrow \mathbf{y} - \mathbf{x} \in \operatorname{int}(\mathcal{K}) \; .$$

Examples:

- Componentwise inequality ($\mathcal{K} = \mathbb{R}_+^n$):

$$\mathbf{x} \preceq_{\mathbb{R}_+^n} \mathbf{y} \Longleftrightarrow x_k \leq y_k \, , \; k = 1, \ldots, n, .$$

- Matrix inequality ($\mathcal{K} = \mathbb{S}_+^n$):

$$\mathbf{X} \preceq_{\mathbb{S}_+^n} \mathbf{Y} \Longleftrightarrow \mathbf{Y} - \mathbf{X} \text{ is positive semidefinite} \, .$$

These two types are the so common that the subscripts is typically dropped in $\preceq_{\mathcal{K}}$.

Properties:

- If $\mathbf{x} \preceq_{\mathcal{K}} \mathbf{y}$ and $\mathbf{u} \preceq_{\mathcal{K}} \mathbf{v}$, then $\mathbf{x} + \mathbf{u} \preceq_{\mathcal{K}} \mathbf{y} + \mathbf{v}$.
- If $\mathbf{x} \preceq_{\mathcal{K}} \mathbf{y}$ and $\mathbf{y} \preceq_{\mathcal{K}} \mathbf{z}$, then $\mathbf{x} \preceq_{\mathcal{K}} \mathbf{z}$.
- If $\mathbf{x} \preceq_{\mathcal{K}} \mathbf{y}$ and $\mathbf{y} \preceq_{\mathcal{K}} \mathbf{x}$, then $\mathbf{x} = \mathbf{y}$.
- If $\mathbf{x} \prec_{\mathcal{K}} \mathbf{y}$ and $\mathbf{u} \preceq_{\mathcal{K}} \mathbf{v}$, then $\mathbf{x} + \mathbf{u} \prec_{\mathcal{K}} \mathbf{y} + \mathbf{v}$.

Minimum/Minimal Elements: Definition and Examples

The generalized inequality $\preceq_{\mathcal{K}}$ is not a *linear ordering* in general. Namely, we can have both $\mathbf{x} \npreceq_{\mathcal{K}} \mathbf{y}$ and $\mathbf{y} \npreceq_{\mathcal{K}} \mathbf{x}$, so that $\mathbf{x}$ and $\mathbf{y}$ are not *comparable*. However, we can define the notion of the *minimum* element and a *minimal* element.

Minimum element:
For some set $\mathcal{S}$, we say that $\mathbf{x} \in \mathcal{S}$ is *the minimum element* of $\mathcal{S}$ with respect to $\preceq_{\mathcal{K}}$ if

$$\mathbf{y} \in \mathcal{S} \Longrightarrow \mathbf{x} \preceq_{\mathcal{K}} \mathbf{y} \, .$$

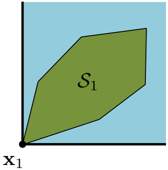
In other words, $\mathbf{x}$ is the minimum element if it beats all other points in $\mathcal{S}$ with respect to $\preceq_{\mathcal{K}}$.

Minimal element:
For some set $\mathcal{S}$, we say that $\mathbf{x} \in \mathcal{S}$ is *a minimal element* of $\mathcal{S}$ with respect to $\preceq_{\mathcal{K}}$ if

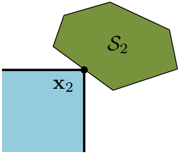
$$\mathbf{y} \in \mathcal{S} \, , \; \mathbf{y} \preceq_{\mathcal{K}} \mathbf{x} \Longrightarrow \mathbf{y} = \mathbf{x} \, .$$

In other words, $\mathbf{x}$ is a minimal element if no other points in $\mathcal{S}$ beat it with respect to $\preceq_{\mathcal{K}}$.

Example:



the minimum element



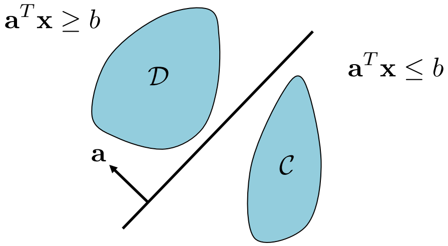
a minimal element

Separating Hyperplane Theorem

If $\mathcal{C}$ and $\mathcal{D}$ are disjoint convex sets, then there exists an $\mathbf{a} \neq \mathbf{0}, b$ such that

$$\mathbf{a}^T\mathbf{x} \leq b \text{ for } \mathbf{x} \in \mathcal{C}\,,\ \mathbf{a}^T\mathbf{x} \geq b \text{ for } \mathbf{x} \in \mathcal{D}\,.$$

In other words, the hyperplane $\mathcal{H} \triangleq \left\{\mathbf{x} : \mathbf{a}^T\mathbf{x} = b\right\}$ *separates* $\mathcal{C}$ and $\mathcal{D}$.



The separating hyperplane theorem can be used to develop a *theorem of alternatives* for strict linear inequalities. Specifically, if $\mathbf{A} \in \mathbb{R}^{m \times n}$ and $\mathbf{b} \in \mathbb{R}^m$ represent some data, then exactly one of the following systems is solvable.

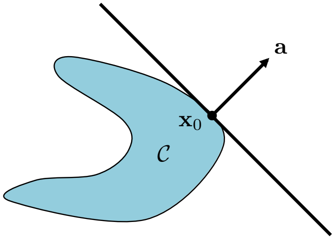
- $\mathbf{A}\mathbf{x} \prec \mathbf{b}$, where $\mathbf{x} \in \mathbb{R}^n$.
- $\boldsymbol{\lambda} \neq \mathbf{0}$, $\boldsymbol{\lambda} \succeq \mathbf{0}$, $\mathbf{A}^T\boldsymbol{\lambda} = \mathbf{0}$, $\boldsymbol{\lambda}^T\mathbf{b} \leq 0$, where $\boldsymbol{\lambda} \in \mathbb{R}^m$.

Supporting Hyperplane Theorem

A *supporting hyperplane* $\mathcal{H}$ to a set $\mathcal{C}$ at a boundary point $\mathbf{x}_0$ is one for which

$$\mathcal{H} = \left\{ \mathbf{x} : \mathbf{a}^T \mathbf{x} = \mathbf{a}^T \mathbf{x}_0 \right\} ,$$

where $\mathbf{a} \neq \mathbf{0}$ and $\mathbf{a}^T \mathbf{x} \leq \mathbf{a}^T \mathbf{x}_0$ for all $\mathbf{x} \in \mathcal{C}$.



The *supporting hyperplane theorem* states that if $\mathcal{C}$ is convex, then there exists a supporting hyperplane at every boundary point of $\mathcal{C}$.

Dual Cones and Generalized Inequalities

The *dual cone* $\mathcal{K}^{\circledast}$ of a cone $\mathcal{K}$ is defined as follows:

$$\mathcal{K}^{\circledast} \triangleq \{\mathbf{y} : \mathrm{Re}[\langle \mathbf{x}, \mathbf{y} \rangle] \geq 0 \text{ for all } \mathbf{x} \in \mathcal{K}\} \ .$$

The dual cone is always a convex cone, even if the primal cone $\mathcal{K}$ is not. If $\mathcal{K}$ is a closed, convex cone, then $\mathcal{K}^{\circledast\circledast} = \mathcal{K}$.

Examples:

- $\mathcal{K} = \mathbb{R}_{+}^{n}$, $\mathcal{K}^{\circledast} = \mathbb{R}_{+}^{n}$.
- $\mathcal{K} = \mathbb{S}_{+}^{n}$, $\mathcal{K}^{\circledast} = \mathbb{S}_{+}^{n}$.
- $\mathcal{K} = \left\{(\mathbf{x}, t) : ||\mathbf{x}||_2 \leq t\right\}$, $\mathcal{K}^{\circledast} = \left\{(\mathbf{x}, t) : ||\mathbf{x}||_2 \leq t\right\}$.
- $\mathcal{K} = \left\{(\mathbf{x}, t) : ||\mathbf{x}||_1 \leq t\right\}$, $\mathcal{K}^{\circledast} = \left\{(\mathbf{x}, t) : ||\mathbf{x}||_\infty \leq t\right\}$.
- $\mathcal{K} = \{(\mathbf{x}, t) : ||\mathbf{x}|| \leq t\}$, $\mathcal{K}^{\circledast} = \left\{(\mathbf{x}, t) : ||\mathbf{x}||_{\circledast} \leq t\right\}$.

The first three examples are *self-dual* cones and are the most commonly encountered.

If $\mathcal{K}$ is a proper cone, then so is $\mathcal{K}^{\circledast}$ and thus they both induce generalized inequalities $\preceq_{\mathcal{K}}$ and $\preceq_{\mathcal{K}^{\circledast}}$, respectively.

Properties relating a generalized inequality to its dual:

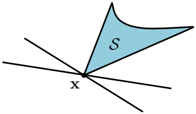
- $\mathbf{x} \preceq_{\mathcal{K}} \mathbf{y}$ if and only if $\mathrm{Re}[\langle \mathbf{x}, \boldsymbol{\lambda} \rangle] \leq \mathrm{Re}[\langle \mathbf{y}, \boldsymbol{\lambda} \rangle]$ for all $\boldsymbol{\lambda} \succeq_{\mathcal{K}^{\circledast}} \mathbf{0}$.
- $\mathbf{x} \prec_{\mathcal{K}} \mathbf{y}$ if and only if $\mathrm{Re}[\langle \mathbf{x}, \boldsymbol{\lambda} \rangle] < \mathrm{Re}[\langle \mathbf{y}, \boldsymbol{\lambda} \rangle]$ for all $\boldsymbol{\lambda} \succeq_{\mathcal{K}^{\circledast}} \mathbf{0}$, $\boldsymbol{\lambda} \neq \mathbf{0}$.
- Since $\mathcal{K}^{\circledast\circledast} = \mathcal{K}$, the above properties hold if $\mathcal{K}$ and $\mathcal{K}^{\circledast}$ are swapped.

Minimum and Minimal Elements Via Dual Inequalities

The minimum element and minimal elements of a set $\mathcal{S}$ with respect to the generalized inequality $\preceq_{\mathcal{K}}$ can be elegantly expressed in terms of the dual inequality $\preceq_{\mathcal{K}^\circledast}$.

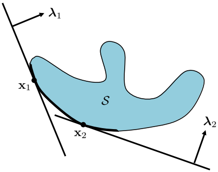
Minimum element:

A point $\mathbf{x}$ is the minimum element of $\mathcal{S}$ if and only if for all $\boldsymbol{\lambda} \succ_{\mathcal{K}^\circledast} \mathbf{0}$, $\mathbf{x}$ is the unique minimizer of $\boldsymbol{\lambda}^T\mathbf{z}$ over $\mathcal{S}$.



Minimal element:

- If $\mathbf{x}$ minimizes $\boldsymbol{\lambda}^T\mathbf{z}$ over $\mathcal{S}$ for some $\boldsymbol{\lambda} \succ_{\mathcal{K}^\circledast} \mathbf{0}$, then $\mathbf{x}$ is minimal.



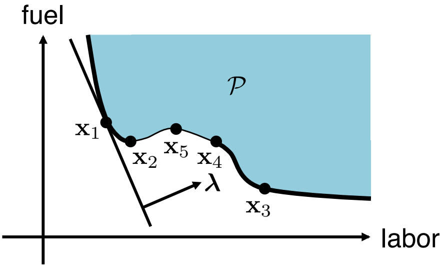
- If $\mathbf{x}$ is a minimal element of a convex set $\mathcal{S}$, then there exists a nonzero $\boldsymbol{\lambda} \succeq_{\mathcal{K}^\circledast} \mathbf{0}$ such that $\mathbf{x}$ minimizes $\boldsymbol{\lambda}^T\mathbf{z}$ over $\mathcal{S}$.

Pareto Optimal/Efficient Production Frontier

To illustrate the concepts of minimal elements and the interpretation with respect to dual generalized inequalities, we consider its use in determining efficient production methods for a particular good. We consider a product which requires n resources (such as labor, fuel, electricity, etc.) to manufacture. This leads to the following:

- Different production methods require different amounts of resources $\mathbf{x} \in \mathbb{R}^n$.
- The production set $\mathcal{P}$ is the set of all resource vectors $\mathbf{x}$ for all production methods.
- Those resource vectors $\mathbf{x}$ that are minimal with respect to $\mathbb{R}^n_+$ and said to be *Pareto optimal* or *Pareto efficient*.

Example:



Here, $\mathbf{x}_1$, $\mathbf{x}_2$, and $\mathbf{x}_3$ are efficient, whereas $\mathbf{x}_4$ and $\mathbf{x}_5$ are not.