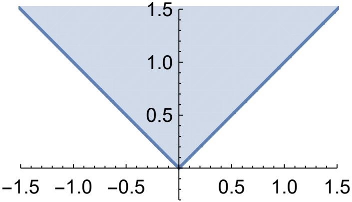
13. Cones and semidefinite constraints

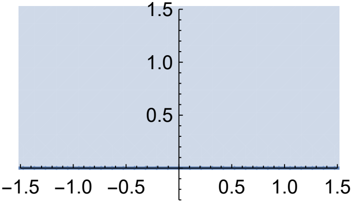
- Geometry of cones
- Second order cone programs
- Example: robust linear program
- Semidefinite constraints

Laurent Lessard (www.laurentlessard.com)

What is a cone?

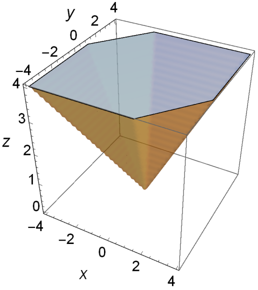
- A set of points $C \in \mathbb{R}^n$ is called a **cone** if it satisfies:
  - $\alpha x \in C$ whenever $x \in C$ and $\alpha > 0$.
  - $x + y \in C$ whenever $x \in C$ and $y \in C$.
- Similar to a subspace, but $\alpha > 0$ instead of $\alpha \in \mathbb{R}$.
  (this is a critical difference!)
- Simple examples: $|x| \le y$ and $y \ge 0$

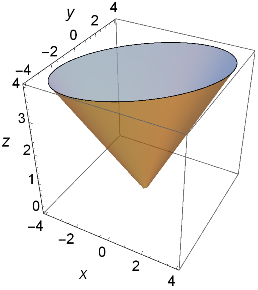


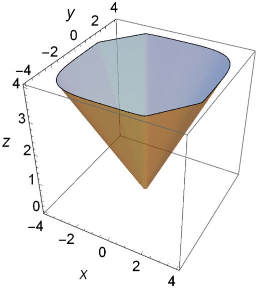


What is a cone?

- A **slice** of a cone is its intersection with a subspace.
- We are interested in **convex cones** (all slices are convex).
- Can be polyhedral, ellipsoidal, or something else...



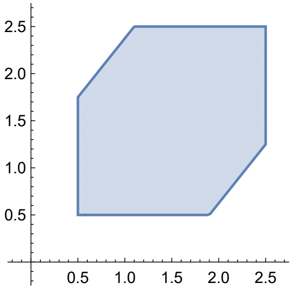


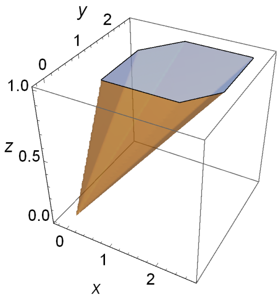


What is a cone?

Polyhedral cone recipe:

1. Begin with your favorite polyhedron $Ax \le b$ where $x \in \mathbb{R}^n$
2. $\{Ax \le bt, t \ge 0\}$ is a polyhedral cone in $(x, t) \in \mathbb{R}^{n+1}$
3. The slice $t = 1$ is the original polyhedron.

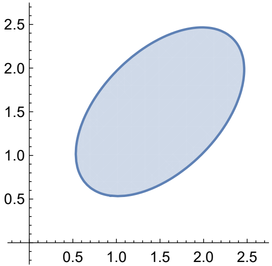


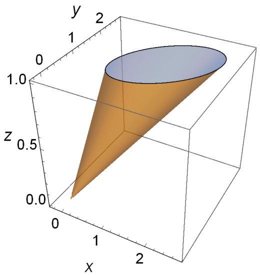


What is a cone?

Ellipsoidal cone recipe:

1. Ellipsoid $x^{\mathsf{T}} P x + q^{\mathsf{T}} x + r \leq 0$ where $P \succ 0$ and $x \in \mathbb{R}^n$
2. Complete the square $\iff$ $\|Ax + b\| \leq c$
3. $\{\|Ax + bt\| \leq ct\}$ is an ellipsoidal cone in $(x, t) \in \mathbb{R}^{n+1}$
4. The slice $t = 1$ is the original ellipsoid.





Second-order cone constraint

A **second-order cone constraint** is the set of points $x \in \mathbb{R}^n$:

$$\|Ax + b\| \le c^{\mathsf{T}} x + d$$

Every SOC constraint is a slice (set $t = 1$) of the cone $\|Ax + bt\| \le c^{\mathsf{T}} x + dt$. It's not always a cone itself!

Special cases:

- If $A = 0$, we have a linear inequality (hyperplane)
- If $c = 0$, it's a slice of an ellipsoidal cone

Every SOC constraint describes a **convex** set.

Second-order cone constraint

A **second-order cone constraint** is the set of points $x \in \mathbb{R}^n$:

$$\boxed{\|Ax + b\| \le c^{\mathsf{T}} x + d}$$

If you square both sides...

$$\|Ax + b\| \le c^{\mathsf{T}} x + d \quad \Longleftrightarrow \quad \begin{cases} \|Ax + b\|^2 \le (c^{\mathsf{T}} x + d)^2 \\ \quad c^{\mathsf{T}} x + d \ge 0 \end{cases}$$

The quadratic inequality is:

$$x^{\mathsf{T}}(A^{\mathsf{T}} A - cc^{\mathsf{T}})x + 2(b^{\mathsf{T}} A - dc^{\mathsf{T}})x + (b^{\mathsf{T}} b - d^2) \le 0$$

This may be **nonconvex**!

Second-order cone constraint

A **second-order cone constraint** is the set of points $x \in \mathbb{R}^n$:

$$\|Ax + b\| \leq c^\mathsf{T} x + d$$

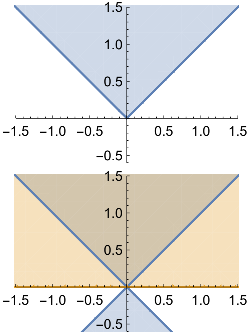
Example:

If $A = \begin{bmatrix} 1 & 0 \end{bmatrix}$ and $c = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ and $b = d = 0$:

$$|x| \leq y$$

Squaring both sides leads to:

$$x^2 - y^2 \leq 0 \quad \text{and} \quad y \geq 0$$

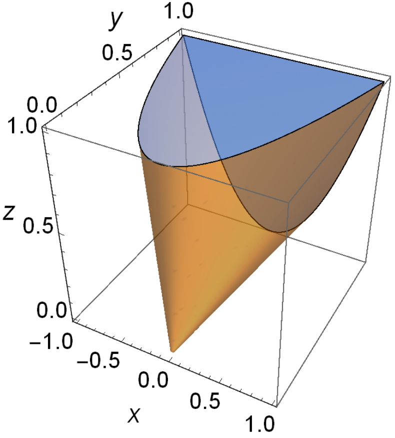


Special case: rotated second-order cone

A rotated second-order cone is the set $x \in \mathbb{R}^n$, $y, z \in \mathbb{R}$:

$$x^\mathsf{T} x \le yz, \quad y \ge 0, \quad z \ge 0$$

With $n = 1$, this looks like:



Special case: rotated second-order cone

A rotated second-order cone is the set $x \in \mathbb{R}^n$, $y, z \in \mathbb{R}$:

$$x^\mathsf{T} x \le yz, \quad y \ge 0, \quad z \ge 0$$

Can put into standard form:

$$\begin{gathered}
4x^\mathsf{T} x \le 4yz \\
4x^\mathsf{T} x + y^2 + z^2 \le 4yz + y^2 + z^2 \\
4x^\mathsf{T} x + (y-z)^2 \le (y+z)^2 \\
\sqrt{4x^\mathsf{T} x + (y-z)^2} \le y + z \\
\left\| \begin{bmatrix} 2x \\ y - z \end{bmatrix} \right\| \le y + z
\end{gathered}$$

SOCPs

A second-order cone program (SOCP) has the form:

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & c^\mathsf{T} x \\ \text{subject to:} \quad & \|A_i x + b_i\| \le c_i^\mathsf{T} x + d_i \quad \text{for } i = 1, \dots, m \end{aligned}$$

- Every LP is an SOCP (just make each $A_i = 0$)
- Every convex QP and QCQP is an SOCP
  - convert quadratic cost to epigraph form (add a variable)
  - convert quadratic constraints to SOCP (complete square)

Implementation details

A second-order cone program (SOCP) has the form:

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & c^\mathsf{T} x \\ \text{subject to:} \quad & \|A_i x + b_i\| \le c_i^\mathsf{T} x + d_i \quad \text{for } i = 1, \dots, m \end{aligned}$$

- In JuMP, you can specify SOCP using: `@constraint(m, norm(A*x+b) <= dot(c,x)+d)` works with `ECOS`, `SCS`, `Mosek`, `Gurobi`, `Ipopt`.
- Can also specify rotated cones directly in `Mosek`, `Ipopt`.

Example: robust LP

Consider a linear program with each linear constraint separately written out:

$$\begin{aligned} \underset{x}{\text{maximize}} \quad & c^{\mathsf{T}} x \\ \text{subject to:} \quad & a_i^{\mathsf{T}} x \le b_i \quad \text{for } i = 1, \dots, m \end{aligned}$$

Suppose there is **uncertainty** in some of the a_i vectors. Say for example that $a_i = \bar{a}_i + \rho u$ where $\bar{a}_i$ is a nominal value and u is the uncertainty.

- box constraint: $\|u\|_\infty \le 1$
- ball constraints: $\|u\|_2 \le 1$

Robust LP with box constraint

Substituting $a_i = \bar{a}_i + \rho u$ into $a_i^\mathsf{T} x \le b_i$, obtain:

$$\bar{a}_i^\mathsf{T} x + \rho u^\mathsf{T} x \le b_i \quad \text{for all uncertain } u$$

box constraint:

If this must hold for **all** u with $\|u\|_\infty \le 1$, then it holds for the worst-case u. Therefore:

$$u^\mathsf{T} x = \sum_{i=1}^{n} u_i x_i \le \sum_{i=1}^{n} |u_i||x_i| \le \sum_{i=1}^{n} |x_i| = \|x\|_1$$

Then we have

$$\bar{a}_i^\mathsf{T} x + \rho \|x\|_1 \le b_i$$

Robust LP with box constraint

With a box constraint $a_i = \bar{a}_i + \rho u$ with $\|u\|_\infty \leq 1$

$$\begin{array}{ll} \underset{x}{\text{maximize}} & c^\top x \\ \text{subject to:} & a_i^\top x \leq b_i \quad \text{for } i = 1, \ldots, m \end{array}$$

Is equivalent to the optimization problem

$$\begin{array}{ll} \underset{x}{\text{maximize}} & c^\top x \\ \text{subject to:} & \bar{a}_i^\top x + \rho \|x\|_1 \leq b_i \quad \text{for } i = 1, \ldots, m \end{array}$$

Robust LP with box constraint

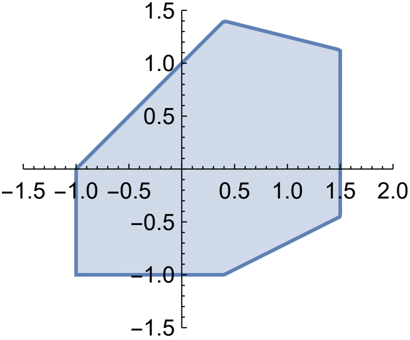
With a box constraint $a_i = \bar{a}_i + \rho u$ with $\|u\|_\infty \leq 1$

$$\begin{aligned} \underset{x}{\text{maximize}} \quad & c^\mathsf{T} x \\ \text{subject to:} \quad & a_i^\mathsf{T} x \leq b_i \quad \text{for } i = 1, \ldots, m \end{aligned}$$

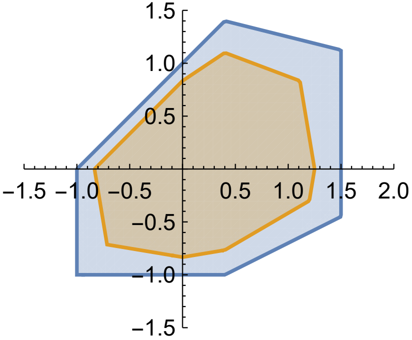
... which is equivalent to the linear program:

$$\begin{aligned} \underset{x,t}{\text{maximize}} \quad & c^\mathsf{T} x \\ \text{subject to:} \quad & \bar{a}_i^\mathsf{T} x + \rho \sum_{j=1}^{n} t_j \leq b_i \quad \text{for } i = 1, \ldots, m \\ & -t_j \leq x_j \leq t_j \quad \text{for } j = 1, \ldots, n \end{aligned}$$

Robust LP with box constraint



$a_i^T x \leq b_i$



$a_i^T x + 0.2\|x\|_1 \leq b_i$

- New region is smaller, still a polyhedron
- More robust to uncertain constraints

Robust LP with ball constraint

Substituting $a_i = \bar{a}_i + \rho u$ into $a_i^\mathsf{T} x \leq b_i$, obtain:

$$\bar{a}_i^\mathsf{T} x + \rho u^\mathsf{T} x \leq b_i \quad \text{for all uncertain } u$$

ball constraint:

If this must hold for **all** u with $\|u\|_2 \leq 1$, then it holds for the worst-case u. Using Cauchy-Schwarz inequality:

$$u^\mathsf{T} x \leq \|u\|_2 \|x\|_2 \leq \|x\|_2$$

Then we have

$$\bar{a}_i^\mathsf{T} x + \rho \|x\|_2 \leq b_i$$

(a second-order cone constraint!)

Robust LP with ball constraint

With a ball constraint $a_i = \bar{a}_i + \rho u$ with $\|u\|_2 \leq 1$

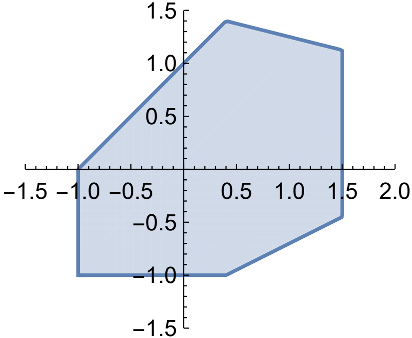
$$\begin{array}{ll} \underset{x}{\text{maximize}} & c^\mathsf{T} x \\ \text{subject to:} & a_i^\mathsf{T} x \leq b_i \quad \text{for } i = 1, \ldots, m \end{array}$$

Is equivalent to the optimization problem

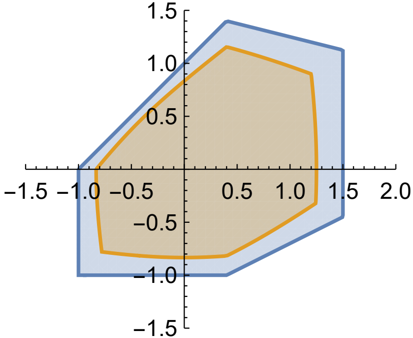
$$\begin{array}{ll} \underset{x}{\text{maximize}} & c^\mathsf{T} x \\ \text{subject to:} & \bar{a}_i^\mathsf{T} x + \rho \|x\|_2 \leq b_i \quad \text{for } i = 1, \ldots, m \end{array}$$

which is an SOCP

Robust LP with ball constraint



$a_i^\mathsf{T} x \le b_i$



$a_i^\mathsf{T} x + 0.2\|x\|_2 \le b_i$

- New region is smaller, no longer a polyhedron
- More robust to uncertain constraints

Matrix variables

Sometimes, the decision variable is a **matrix** X.

- Can always just think of $X \in \mathbb{R}^{m \times n}$ as $x \in \mathbb{R}^{mn}$.

- Linear functions:

$$\sum_{k=1}^{mn} c_k x_k = c^\mathsf{T} x$$

$$\sum_{i=1}^{m} \sum_{j=1}^{n} C_{ij} X_{ij} = \operatorname{trace}(C^\mathsf{T} X)$$

- Linear program:

$$\begin{aligned} \underset{X}{\text{maximize}} \quad & \operatorname{trace}(C^\mathsf{T} X) \\ \text{subject to:} \quad & \operatorname{trace}(A_i^\mathsf{T} X) \le b_i \quad \text{for } i = 1, \dots, k \end{aligned}$$

Matrix variables

If a decision variable is a symmetric matrix $X = X^\mathsf{T} \in \mathbb{R}^{n \times n}$, we can represent it as a vector $x \in \mathbb{R}^{n(n+1)/2}$.

$$\begin{bmatrix} x_1 & x_2 & x_3 \\ x_2 & x_4 & x_5 \\ x_3 & x_5 & x_6 \end{bmatrix} \qquad \Longleftrightarrow \qquad \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix}$$

The constraint $X \succeq 0$ is called a **semidefinite** constraint.
What does it look like geometrically?

The PSD cone

The set of matrices $X \succeq 0$ are a **convex cone** in $\mathbb{R}^{n(n+1)/2}$

Example: The set $\begin{bmatrix} x & y \\ y & z \end{bmatrix} \succeq 0$ of points in $\mathbb{R}^3$ satisfy:

$$xz \geq y^2, \quad x \geq 0, \quad z \geq 0$$

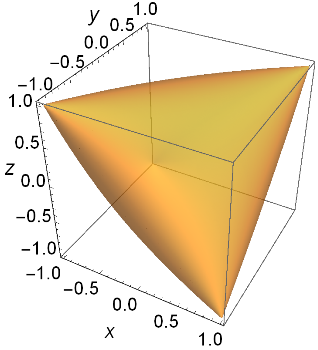
This is a rotated second-order cone! Equivalent to:

$$\left\| \begin{bmatrix} 2y \\ x - z \end{bmatrix} \right\| \leq x + z$$

More complicated example

The set of (x, y, z) satisfying $\begin{bmatrix} 1 & x & y \\ x & 1 & z \\ y & z & 1 \end{bmatrix} \succeq 0$ is the solution of:

$$\left\{ X \in \mathbb{R}^{3\times 3}, \quad X \succeq 0, \quad X_{11} = 1, \quad X_{22} = 1, \quad X_{33} = 1 \right\}$$



Spectrahedra

- Two common set representations:
  - variables $x_1, \dots, x_k$, constants $Q_i = Q_i^\mathsf{T}$, and constraint:

$$Q_0 + x_1 Q_1 + \dots x_k Q_k \succeq 0 \qquad \text{(linear matrix inequality)}$$

  - variable $X \succeq 0$ and the constraints:

$$\operatorname{trace}(A_i^\mathsf{T} X) \le b_i \qquad \text{(linear constraint form)}$$

- These sets are called **spectrahedra**.
- Very rich set, lots of possible shapes.

Semidefinite program (SDP)

Standard form #1: (looks like the standard form for an LP)

$$
\begin{aligned}
\underset{X}{\text{maximize}} \quad & \operatorname{trace}(C^\mathsf{T} X) \\
\text{subject to:} \quad & \operatorname{trace}(A_i^\mathsf{T} X) \le b_i \quad \text{for } i = 1, \dots, m \\
& X \succeq 0
\end{aligned}
$$

Standard form #2:

$$
\begin{aligned}
\underset{x}{\text{maximize}} \quad & c^\mathsf{T} x \\
\text{subject to:} \quad & Q_0 + \sum_{i=1}^{m} x_i Q_i \succeq 0
\end{aligned}
$$

Relationship with other programs

Every LP is an SDP:

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \leq \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

is the same as:

$$x_1 \begin{bmatrix} a_{11} & 0 \\ 0 & a_{21} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} & 0 \\ 0 & a_{22} \end{bmatrix} \preceq \begin{bmatrix} b_1 & 0 \\ 0 & b_2 \end{bmatrix}$$

(polyhedra are special cases of spectrahedra)

Relationship with other programs

Every SOCP is an SDP:

$$\|Ax + b\| \leq c^\mathsf{T} x + d$$

is the same as:

$$\begin{bmatrix} (c^\mathsf{T} x + d) I & Ax + b \\ (Ax + b)^\mathsf{T} & c^\mathsf{T} x + d \end{bmatrix} \succeq 0$$

This isn't obvious — proof requires use of Schur complement.
(second-order cones are special cases of spectrahedra)