12. Cones and semidefinite constraints

- Geometry of cones
- Second order cone programs
- Example: robust linear program
- Semidefinite constraints
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: $|x| \leq y$ and $y \geq 0$
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 $Ax \leq b$ where $x \in \mathbb{R}^n$

2. $\{Ax \leq bt, t \geq 0\}$ 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 \( x^T P x + q^T x + r \leq 0 \) 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

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

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

Special cases:

- If $A = 0$, we have a linear constraint (hyperplane)
- If $c = 0$, can square both sides, (ellipsoid)

In general, you **cannot** just square both sides!
Second-order cone

A **second-order cone** is the set of points \( x \in \mathbb{R}^n \) satisfying

\[
\|Ax - b\| \leq c^T x + d
\]

Counterexample:

If \( A = \begin{bmatrix} 1 & 0 \end{bmatrix} \) and \( c = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \) and \( b = d = 0 \), we have:

\[
|x| \leq y
\]

Square both sides, we get a nonconvex quadratic constraint!

\[
x^2 - y^2 \leq 0
\]
A rotated second-order cone is the set \( x \in \mathbb{R}^n, y, z \in \mathbb{R} \):
\[
x^T x \leq yz, \quad y \geq 0, \quad z \geq 0
\]

With \( n = 1 \), this looks like:
A rotated second-order cone is the set $x \in \mathbb{R}^n$, $y, z \in \mathbb{R}$:

$$x^T x \leq yz, \quad y \geq 0, \quad z \geq 0$$

Can put into standard form:

$$4x^T x \leq 4yz$$

$$4x^T x + y^2 + z^2 \leq 4yz + y^2 + z^2$$

$$4x^T x + (y - z)^2 \leq (y + z)^2$$

$$\sqrt{4x^T x + (y - z)^2} \leq y + z$$

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

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

\[
\begin{align*}
\text{minimize} & \quad c^T x \\
\text{subject to:} & \quad \|A_i x + b_i\| \leq c_i^T x + d_i \quad \text{for } i = 1, \ldots, m
\end{align*}
\]

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

\[
\begin{align*}
\text{minimize} & \quad c^T x \\
\text{subject to:} & \quad \|A_i x + b_i\| \leq c_i^T x + d_i \quad \text{for } i = 1, \ldots, m
\end{align*}
\]

- 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:

\[
\begin{align*}
\text{maximize} & \quad c^T x \\
\text{subject to:} & \quad a_i^T x \leq b_i \quad \text{for } i = 1, \ldots, m
\end{align*}
\]

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: \(\|u\|_\infty \leq 1\)
- ball constraints: \(\|u\|_2 \leq 1\)
Example: robust LP

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

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

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

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

Then we have

$$\bar{a}_i^T x + \rho \|x\|_1 \leq b_i$$
Robust LP with box constraint

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

\[
\begin{align*}
\text{maximize} & \quad c^T x \\
\text{subject to:} & \quad a_i^T x \leq b_i \quad \text{for } i = 1, \ldots, m
\end{align*}
\]

Is equivalent to the optimization problem

\[
\begin{align*}
\text{maximize} & \quad c^T x \\
\text{subject to:} & \quad \bar{a}_i^T x + \rho \|x\|_1 \leq b_i \quad \text{for } i = 1, \ldots, m
\end{align*}
\]
Robust LP with box constraint

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

\[
\begin{align*}
\text{maximize} & \quad c^T x \\
\text{subject to:} & \quad a_i^T x \leq b_i \quad \text{for } i = 1, \ldots, m
\end{align*}
\]

... which is equivalent to the linear program:

\[
\begin{align*}
\text{maximize} & \quad c^T x \\
\text{subject to:} & \quad \bar{a}_i^T x + \rho \sum_{j=1}^{n} t_j \leq b_i \quad \text{for } i = 1, \ldots, m \\
& \quad -t_j \leq x_j \leq t_j \quad \text{for } j = 1, \ldots, n
\end{align*}
\]
Example: robust LP

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

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

**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:

\[
u^T x \leq \|u\|_2 \|x\|_2 \leq \|x\|_2
\]

Then we have

\[
\bar{a}_i^T x + \rho \|x\|_2 \leq b_i
\]

(a second-order cone constraint!)
Robust LP with ball constraint

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

\[
\begin{align*}
\text{maximize} & \quad c^T x \\
\text{subject to:} & \quad a_i^T x \leq b_i \quad \text{for } i = 1, \ldots, m
\end{align*}
\]

Is equivalent to the optimization problem

\[
\begin{align*}
\text{maximize} & \quad c^T x \\
\text{subject to:} & \quad \bar{a}_i^T x + \rho \|x\|_2 \leq b_i \quad \text{for } i = 1, \ldots, m
\end{align*}
\]

which is an SOCP
Ball constraint example

\[ a_i^T x \leq b_i \]

\[ a_i^T x + 0.2\|x\|_2 \leq b_i \]

- 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:

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

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

- Linear program:

$$
\begin{align*}
\text{maximize} \quad & \text{trace}(C^T X) \\
\text{subject to} \quad & \text{trace}(A_i^T X) \leq b_i \quad \text{for } i = 1, \ldots, k
\end{align*}
$$
Matrix variables

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

\[
\begin{bmatrix}
  x_1 & x_2 & x_3 \\
  x_2 & x_4 & x_5 \\
  x_3 & x_5 & x_6 \\
\end{bmatrix} \quad \iff \quad \begin{bmatrix}
  x_1 \\
  x_2 \\
  x_3 \\
  x_4 \\
  x_5 \\
  x_6 \\
\end{bmatrix}
\]

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:

$$\begin{bmatrix} 2y \\ x - z \end{bmatrix} \leq x + z$$
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:

\[
\{ X \in \mathbb{R}^{3\times3}, \ X \succeq 0, \ X_{11} = 1, \ X_{22} = 1, \ X_{33} = 1 \}
\]
Spectrahedra

- Two common set representations:
  - Variables $x_1, \ldots, x_k$, constants $Q_i = Q_i^T$, and constraint:
    \[ Q_0 + x_1 Q_1 + \ldots x_k Q_k \succeq 0 \]  
    (linear matrix inequality)
  - Variable $X \succeq 0$ and the constraints:
    \[ \text{trace}(A_i^T X) \leq b_i \]  
    (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)

\[
\begin{align*}
\text{maximize} & \quad \text{trace } C^T X \\
\text{subject to:} & \quad \text{trace}(A_i^T X) \leq b_i \quad \text{for } i = 1, \ldots, m \\
& \quad X \succeq 0
\end{align*}
\]

Standard form #2:

\[
\begin{align*}
\text{maximize} & \quad c^T x \\
\text{subject to:} & \quad Q_0 + \sum_{i=1}^{m} x_i Q_i \succeq 0
\end{align*}
\]
Relationship with other programs

Every LP is an SDP:

\[
\begin{bmatrix}
a_{11} & a_{12} \\
a_{21} & a_{22}
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2
\end{bmatrix}
\preceq
\begin{bmatrix}
b_1 \\
b_2
\end{bmatrix}
\]

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)
Every SOCP is an SDP:

\[ \|Ax + b\| \leq c^T x + d \]

is the same as:

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

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