Conic optimization

Definition

Conic optimization The problem of minimizing a *linear* function over a set defined by *linear equalities* and *cone membership constraints*:

\[
\min_{x} c^T x \\
Ax = b \\
x \in C
\]

where \( C \) is a *closed convex cone* in a finite-dimensional vector space.
Conic optimization

Definition

Conic optimization The problem of minimizing a linear function over a set defined by linear equalities and cone membership constraints:

$$\min_x c^T x$$

$$Ax = b$$

$$x \in C$$

where $C$ is a closed convex cone in a finite-dimensional vector space

Question

What is a cone?
# Cones

## Definition

**Cone**
- A cone is a set that is closed under *positive* scalar multiplication.
- $C$ is a cone if $\lambda x \in C$, for all $\lambda \geq 0$ and $x \in C$
- A cone is called pointed if it does not include any lines

## Examples

- $C_{\text{r}} = \{ x \in \mathbb{R}^n : x \geq 0 \}$, the non-negative orthant
- Any set in the form $C = \{ x \in \mathbb{R}^n : Ax \geq 0 \}$ (polyhedral cone)
- $C_{\text{q}} = \{ x = (x_0, \ldots, x_n) \in \mathbb{R}^{n+1} : x_0 \geq \| (x_1, \ldots, x_n) \| \}$, the second order cone
- $C_{\text{s}} = \{ X \in \mathbb{R}^{n \times n} : X = X^T, X \text{ is positive semidefinite} \}$, the cone of symmetric, positive semidefinite matrices
## Cones

### Definition

**Cone**
- A cone is a set that is closed under *positive* scalar multiplication.
- $C$ is a cone if $\lambda x \in C$, for all $\lambda \geq 0$ and $x \in C$
- A cone is called pointed if it does not include any lines

### Examples

- $C_I = \{x \in \mathbb{R}^n : x \geq 0\}$, the non negative orthant
- Any set in the form $C = \{x \in \mathbb{R}^n : A x \geq 0\}$ (polyhedral cone)
Cone

• A cone is a set that is closed under *positive* scalar multiplication.
• $C$ is a cone if $\lambda x \in C$, for all $\lambda \geq 0$ and $x \in C$
• A cone is called pointed if it does not include any lines

Examples

• $C_l = \{x \in \mathbb{R}^n : x \geq 0\}$, the non negative orthant
• Any set in the form $C = \{x \in \mathbb{R}^n : Ax \geq 0\}$ (polyhedral cone)
• $C_q = \{x = (x_0, \ldots, x_n) \in \mathbb{R}^{n+1} : x_0 \geq \|(x_1, \ldots, x_n)\|\}$, the *second order cone*
Cones

**Definition**

**Cone**

- A cone is a set that is closed under *positive* scalar multiplication.
- \( C \) is a cone if \( \lambda x \in C \), for all \( \lambda \geq 0 \) and \( x \in C \)
- A cone is called pointed if it does not include any lines

**Examples**

- \( C_l = \{ x \in \mathbb{R}^n : x \geq 0 \} \), the non negative orthant
- Any set in the form \( C = \{ x \in \mathbb{R}^n : Ax \geq 0 \} \) (polyhedral cone)
- \( C_q = \{ x = (x_0, \ldots, x_n) \in \mathbb{R}^{n+1} : x_0 \geq \| (x_1, \ldots, x_n) \| \} \), the second order cone
- \( C_s = \{ X \in \mathbb{R}^{n \times n} : X = X^T, X \text{ is positive semidefinite} \} \), the cone of symmetric, positive semidefinite matrices.
norm of the remaining elements. This corresponds to the case where $C$ is the
cone $C_q$ obtained through a 45 degree rotation (Why?). This is why the cone $C_r_q$ is called the
rotated quadratic cone.

The second order cone $C$ is also often called the ice-cream cone.

Example

The second order cone

$$\{(x_1, x_2, x_3) : x_1 \geq \|x_2, x_3\|\}$$
The cone of symmetric, positive semidefinite matrices in $\mathbb{R}^{2\times 2}$
Example

Question

*Can you give an example of a non-convex cone?*
Can you give an example of a non-convex cone?

- The graph of $f(x) = |x|$
Question

Can you give an example of a non-convex cone?

• The graph of $f(x) = |x|$
Example

Question

Can you give an example of a non-convex cone?

- The graph of $f(x) = |x|
- The union of the first and third quadrants of the plane (also non-pointed)
Conic optimization

\[
\min_x c^T x
\]
\[
Ax = b
\]
\[
x \in C
\]

where \( C \) is a closed convex cone in a finite-dimensional vector space \( X \).

- When \( X = \mathbb{R}^n \) and \( C = \mathbb{R}^n_+ \) this is a standard form LP.

Conic programming provides a unifying framework for:
- LP;
- second-order cone programming (SOCP);
- semidefinite programming (SDP);
Second-order cone programming (SOCP)

The second order cone:

\[ C_q = \{ x = (x_0, \ldots, x_n) \in \mathbb{R}^{n+1} : x_0 \geq \|(x_1, \ldots, x_n)\| \} \]

- By intersecting the second-order cone with a hyperplane, we obtain ellipsoidal sets.
- Any convex quadratic constraint can be expressed using the second-order cone (or its rotations) and one or more hyperplanes.
Ellipsoidal uncertainty for linear constraints

Common use of SOCP in financial applications

Modeling and treatment of *parameter uncertainties*;

- SOCP allows to model the uncertainties in the constraints/objective function of an optimization problem;
- the resulting model will find a solution to the original optimization problem that will perform well under many scenarios (i.e., many values of the parameters)
Consider the following single-constraint LP:

\[
\begin{align*}
\text{min } & \quad c^T x \\
\text{s.t. } & \quad a^T x + b \geq 0
\end{align*}
\]

- We are sure about \( c \), but have uncertainties about \( a \) and \( b \).  
- We know that \([a; b]\) belong to the following ellipsoidal uncertainty set:

\[
U = \left\{ [a; b] = [a^0; b^0] + \sum_{j=1}^{k} u_j[a^j; b^j], \|u\| \leq 1 \right\}
\]
Goal

Find a solution that minimizes the objective function among the vectors that are feasible for all \([a; b] \in \mathcal{U}\).

I.e., solve

\[
\begin{align*}
\min & \quad c^T x \\
\text{s.t.} & \quad a^T x + b \geq 0, \forall [a; b] \in \mathcal{U}
\end{align*}
\]
Goal

Find a solution that minimizes the objective function among the vectors that are feasible for all $[a; b] \in U$. I.e., solve

$$\min c^T x \\
\text{s.t. } a^T x + b \geq 0, \forall [a; b] \in U$$

Fix $x$ and let

$$\alpha = (a^0)^T x + b^0 \quad \text{and} \quad \beta = (\beta_1, \ldots, \beta_k) \quad \text{with} \quad \beta_j = (a^j)^T x + b^j$$

For each fixed $x$, the constraint is satisfied if and only if:

$$0 \leq \min_{[a; b] \in U} a^T x + b = \min_{u: \|u\| \leq 1} \alpha + u^T \beta$$

Let's now look at the minimization problem on the r.h.s.
Recall: \( x \) is fixed and
\[
\alpha = (a^0)^T x + b^0 \quad \text{and} \quad \beta = (\beta_1, \ldots, \beta_k) \text{ with } b_j = (a_j)^T x + b^j
\]
We want to minimize
\[
\min_{u: \|u\| \leq 1} \alpha + u^T \beta
\]

Observations
- W.r.t. \( u \), \( \alpha \) is...
Recall: $x$ is fixed and

$$
\alpha = (a^0)^T x + b^0 \quad \text{and} \quad \beta = (\beta_1, \ldots, \beta_k) \text{ with } b_j = (a^j)^T x + b^j
$$

We want to minimize

$$
\min_{u: \|u\| \leq 1} \alpha + u^T \beta
$$

Observations

- W.r.t. $u$, $\alpha$ is... constant.
Recall: $x$ is fixed and

$$\alpha = (a^0)^T x + b^0 \quad \text{and} \quad \beta = (\beta_1, \ldots, \beta_k) \text{ with } b_j = (a^j)^T x + b^j$$

We want to minimize

$$\min_{u: \|u\| \leq 1} \alpha + u^T \beta$$

**Observations**

- W.r.t. $u$, $\alpha$ is... constant.
- W.r.t. $u$, $\beta$ is...
Recall: $x$ is fixed and

$$\alpha = (a^0)^T x + b^0 \quad \text{and} \quad \beta = (\beta_1, \ldots, \beta_k) \text{ with } b_j = (a^j)^T x + b^j$$

We want to minimize

$$\min_{u: \|u\| \leq 1} \alpha + u^T \beta$$

Observations

- W.r.t. $u$, $\alpha$ is... constant.
- W.r.t. $u$, $\beta$ is... also constant.

So we really want:

$$\min_{u: \|u\| \leq 1} u^T \beta$$
Question

How can we rewrite $u^T \beta$?

(hint: it involves the angle $\theta$ between $u$ and $\beta$)
min_{u: \|u\| \leq 1} u^T \beta

**Question**

*How can we rewrite* $u^T \beta$?

*(hint: it involves the angle $\theta$ between $u$ and $\beta$)*

$$u^T \beta = \|u\| \|\beta\| \cos \theta$$

**Observations**

- $\|\beta\|$ is...
\[
\min_{u: \|u\| \leq 1} u^T \beta
\]

**Question**

*How can we rewrite* \(u^T \beta\)?

*(hint: it involves the angle \(\theta\) between \(u\) and \(\beta\))*

\[
\begin{align*}
    u^T \beta &= \|u\| \|\beta\| \cos \theta
\end{align*}
\]

**Observations**

- \(\|\beta\|\) is... constant
\[
\min_{u: \|u\| \leq 1} u^T \beta
\]

**Question**

*How can we rewrite \( u^T \beta \)?*  
*\( \text{(hint: it involves the angle } \theta \text{ between } u \text{ and } \beta) \)*

\[
u^T \beta = \|u\| \|\beta\| \cos \theta
\]

**Observations**

- \( \|\beta\| \) is... constant
- \( \|u\| \) is...
$$\min_{u: \|u\| \leq 1} u^T \beta$$

**Question**

*How can we rewrite* $u^T \beta$?  
*(hint: it involves the angle $\theta$ between $u$ and $\beta$)*

$$u^T \beta = \|u\| \|\beta\| \cos \theta$$

**Observations**

- $\|\beta\|$ is... constant
- $\|u\|$ is... $\leq 1$ and $\geq 0$
\[ \min_{u: \|u\| \leq 1} u^T \beta \]

**Question**

*How can we rewrite \( u^T \beta \)?*  
*(hint: it involves the angle \( \theta \) between \( u \) and \( \beta \))*

\[ u^T \beta = \| u \| \| \beta \| \cos \theta \]

**Observations**

- \( \| \beta \| \) is... constant
- \( \| u \| \) is... \( \leq 1 \) and \( \geq 0 \)
- \( \cos \theta \) is...
$$\min_{u: \|u\| \leq 1} u^T \beta$$

**Question**

*How can we rewrite $u^T \beta$?*

*(hint: it involves the angle $\theta$ between $u$ and $\beta$)*

$$u^T \beta = \|u\| \|\beta\| \cos \theta$$

**Observations**

- $\|\beta\|$ is... constant
- $\|u\|$ is... $\leq 1$ and $\geq 0$
- $\cos \theta$ is... $\leq 1$ and $\geq -1$

Hence, to minimize $u^T \beta$, we want: $\|u\| =$
\[
\min_{u: \|u\| \leq 1} u^T \beta
\]

**Question**

*How can we rewrite \(u^T \beta\)?*

*(hint: it involves the angle \(\theta\) between \(u\) and \(\beta\))*

\[u^T \beta = \|u\| \|\beta\| \cos \theta\]

**Observations**

- \(\|\beta\|\) is... constant
- \(\|u\|\) is... \(\leq 1\) and \(\geq 0\)
- \(\cos \theta\) is... \(\leq 1\) and \(\geq -1\)

Hence, to minimize \(u^T \beta\), we want: \(\|u\| = 1\) and \(\cos \theta =\)
\[
\min_{u: \|u\| \leq 1} u^T \beta
\]

**Question**

*How can we rewrite \( u^T \beta \)?*(hint: it involves the angle \( \theta \) between \( u \) and \( \beta \))

\[
 u^T \beta = \|u\| \|\beta\| \cos \theta
\]

**Observations**

- \( \|\beta\| \) is... constant
- \( \|u\| \) is... \( \leq 1 \) and \( \geq 0 \)
- \( \cos \theta \) is... \( \leq 1 \) and \( \geq -1 \)

Hence, to minimize \( u^T \beta \), we want: \( \|u\| = 1 \) and \( \cos \theta = -1 \);

I.e., \( u = \ldots \) (hint: is a function of \( \beta \))
\[
\min_{u: \|u\| \leq 1} u^T \beta
\]

Question

How can we rewrite \( u^T \beta \)?
(hint: it involves the angle \( \theta \) between \( u \) and \( \beta \))

\[
u^T \beta = \|u\| \|\beta\| \cos \theta
\]

Observations

- \( \|\beta\| \) is...constant
- \( \|u\| \) is...\( \leq 1 \) and \( \geq 0 \)
- \( \cos \theta \) is...\( \leq 1 \) and \( \geq -1 \)

Hence, to minimize \( u^T \beta \), we want: \( \|u\| = 1 \) and \( \cos \theta = -1 \);
I.e., \( u = \ldots \) (hint: is a function of \( \beta \))

\[
u = -\beta / \|\beta\|
\]
\[ u = -\beta / \| \beta \| \]

Hence \( \min_u : \| u \| \leq 1 \alpha + u^T \beta = \alpha - \| \beta \| \) We then have

\[
\min_{[a;b] \in U} a^T x + b = \alpha - \| \beta \| = (a^0)^T + b^0 - \sqrt{\sum_{j=1}^{k} ((a_j)^T x + b_j)^2}
\]
We can then rewrite our original LP in the following *robust* version:

\[
\begin{align*}
\min & \quad c^T x \\
\text{s.t.} & \quad (a^0)^T + b^0 - \sqrt{\sum_{j=1}^{k} ((a^j)^T x + b^j)^2}
\end{align*}
\]
We can then rewrite our original LP in the following \textit{robust} version:

\[
\begin{align*}
\text{min } & \quad c^T x \\
\text{s.t. } & \quad (a^0)^T + b^0 - \sqrt{k \sum_{j=1}^{k} ((aj)^T x + bj)^2} 
\end{align*}
\]

Can we rewrite the constraint to involve the second-order cone

\[C_q = \{y = (y_1, \ldots, y_n) \in \mathbb{R}^n : : y_1 \geq \| (x_2, \ldots, x_n) \| \}\]?
We can then rewrite our original LP in the following *robust* version:

\[
\begin{align*}
& \min c^T x \\
& \text{s.t. } (a^0)^T + b^0 - \sqrt{\sum_{j=1}^{k} ((a^j)^T x + b^j)^2} 
\end{align*}
\]

Can we rewrite the constraint to involve the second-order cone

\[ C_q = \{ y = (y_1, \ldots, y_n) \in \mathbb{R}^n : y_1 \geq \|(x_2, \ldots, x_n)\| \} \]?

Introduce the variables

\[ z_j = (a^j)^T x + b^j, j = 0, \ldots, k \]

and add the constraint

\[ (z_0, z_1, \ldots, z_k) \in C_q \]
Conic, robust version of the original LP optimization problem:

$$\begin{align*}
\min & \quad c^T x \\
\text{s.t.} & \quad z_j = (a^j)^T x + b^j, j = 0, \ldots, k \\
& \quad (z_0, \ldots, z_k) \in C_q
\end{align*}$$
Conic, robust version of the original LP optimization problem:

\[
\begin{align*}
\min & \quad c^T x \\
\text{s.t.} & \quad z_j = (a^j)^T x + b^j, j = 0, \ldots, k \\
& \quad (z_0, \ldots, z_k) \in C_q
\end{align*}
\]

The approach we just described generalizes to multiple constraints, when the uncertainties sets of parameters in different constraints are unrelated.
Converting quadratic constraints into second-order cone constraints

Fact

Any convex quadratic constraint can be rewritten using second-order cone membership constraint.

Consider the quadratic constraint

\[ x^T Q x + 2 p^T x + \gamma \leq 0 \]

- Assume \( Q \) is positive definite;
- Then there exists an invertible matrix such that \( Q = RR^T \)

We can write the above as:

\[ (R^T x)^T (R^T x) + 2 p^T x + \gamma \leq 0 \]
\[(R^T x)(R^T x) + 2p^T x + \gamma \leq 0\]

Let \( y = (y_1, \ldots, y_k)^T = R^T x + R^{-1} p \). Then

\[ y^T y = (R^T x)^T (R^T x) + 2p^T x + p^T Q^{-1} p \]

Then the constraint is equivalent to require that

\[ \exists y \text{ s.t. } y = R^T x + R^{-1} p, y^T y \leq p^T Q^{-1} p - \gamma \]

(assume r.h.s is \( \geq 0 \))
\[ \exists y \text{ s.t. } y = R^T x + R^{-1} p, y^T y \leq p^T Q^{-1} p - \gamma \]

**Observations**

- \( y^T y \) is...
\[ \exists y \text{ s.t. } y = R^T x + R^{-1} p, y^T y \leq p^T Q^{-1} p - \gamma \]

Observations

- \( y^T y \) is... \( \|y\|^2 \)
\[ \exists y \text{ s.t. } y = R^T x + R^{-1} p, y^T y \leq p^T Q^{-1} p - \gamma \]

**Observations**

- \( y^T y \) is... \( \|y\|^2 \)
- We want \( \sqrt{p^T Q^{-1} p - \gamma} \geq \|y\| \)

Hence we can rewrite the constraint as the following set of linear equations coupled with a second-order constraint:

\[
\begin{bmatrix}
  y_1 \\
  \vdots \\
  y_k
\end{bmatrix} = R^T x + R^{-1} p
\]

\[
y_0 = \sqrt{p^T Q^{-1} p - \gamma}
\]

\((y_0, y_1, \ldots, y_k) \in C_q\)
Application: approximating covariance matrices

- We often need to approximate the covariance matrix of a vector of random variables.
- E.g., for forecasting, time-series modeling, etc.
- Various ways exist to approximate the covariance matrix, but often they do not impose problem-dependent constraints on the approximations, or even general constraints (e.g., symmetry, positive semidefiniteness);
- Typically, one is interested in finding the smallest distortion of the original estimate that satisfies the desired constraints;
• Let $\hat{\Sigma} \in S^n$ be an estimate of a covariance matrix

• $\hat{\Sigma}$ is symmetric ($\in S^n$) but not positive semidefinite.

Goal: find the positive semidefinite matrix that is closest to $\hat{\Sigma}$ w.r.t. the Frobenius norm:

$$d_F(\Sigma, \hat{\Sigma}) = \sqrt{\sum_{i,j} (\Sigma_{ij} - \hat{\Sigma}_{ij})^2}$$

Formally, we want to solve

$$\min_{\Sigma} d_F(\Sigma, \hat{\Sigma})$$

$$\Sigma \in C_s^n$$

where $C_s^n$ is the cone of $n \times n$ symmetric and positive semidefinite matrices.
\[
\min_{\Sigma} d_F(\Sigma, \hat{\Sigma})
\]
\[
\Sigma \in C_s^n
\]

Let’s rewrite it as a quadratic optimization problem
\[
\min_{\Sigma} d_F(\Sigma, \hat{\Sigma})
\]
\[
\Sigma \in C^n_s
\]

Let’s rewrite it as a quadratic optimization problem

We can introduce a dummy variable \( t \) and rewrite the problem as

\[
\min t
\]
\[
d_F(\Sigma, \hat{\Sigma}) \leq t
\]
\[
\Sigma \in C^n_s
\]

The first constraint can be written as a second-order cone constraint, so the problem is be transformed into a conic optimization problem.