Conic Optimization

John E. Mitchell

Department of Mathematical Sciences
RPI, Troy, NY 12180 USA

April 2018
Outline

1. Introduction
2. Duality
3. Path following methods
Conic optimization problems

We consider the conic optimization problem

\[
\begin{align*}
\min_x & \quad c^T x \\
\text{subject to} & \quad Ax = b \\
& \quad x \in K \subseteq \mathbb{R}^n
\end{align*}
\]  

(1)

where $K$ is a convex cone and $A \in \mathbb{R}^{m \times n}$.
Examples:

1. **Linear programming**: The cone $K$ is the nonnegative orthant $\mathbb{R}_+^n$.
2. **Semidefinite programming**: The cone $K$ is the set of symmetric positive semidefinite matrices $\mathbb{S}_+^n$.
3. **Second order cone programs**: The cone $K$ is the set of points satisfying
   \[ x_1^2 \geq \sum_{i=2}^n x_i^2, \quad x_1 \geq 0. \]  
   (2)
4. **$l_p$-norm for $p > 1$, so**
   \[ K = \{ x \in \mathbb{R}^n : \sum_{j=1}^q (1/p_j)|c_j - a_j^T x|^{p_j} \leq d - g^T x \} \]  
   (3)
   where $d$ and each $c_j$ are scalars, and $g$ and each $a_j$ are $n$-vectors.
5. Can have **direct products** of simpler cones. So can have, eg, a mixture of nonnegative variables and a psd matrix. Or can have multiple second order cone programs.
Examples:

1. **Linear programming**: The cone $K$ is the nonnegative orthant $\mathbb{R}_+^n$.

2. **Semidefinite programming**: The cone $K$ is the set of symmetric positive semidefinite matrices $\mathcal{S}_+^n$.

3. **Second order cone programs**: The cone $K$ is the set of points satisfying

   $$x_1^2 \geq \sum_{i=2}^{n} x_i^2, \quad x_1 \geq 0.$$  

4. **$l_p$-norm for $p > 1$**, so

   $$K = \{ x \in \mathbb{R}^n : \sum_{j=1}^{q} \left( \frac{1}{p_j} \right) |c_j - a_j^T x|^{p_j} \leq d - g^T x \}$$

   where $d$ and each $c_j$ are scalars, and $g$ and each $a_j$ are $n$-vectors.

5. **Can have direct products** of simpler cones. So can have, eg, a mixture of nonnegative variables and a psd matrix. Or can have multiple second order cones.
Examples:

1. **Linear programming**: The cone $K$ is the nonnegative orthant $\mathbb{R}^n_+$.  

2. **Semidefinite programming**: The cone $K$ is the set of symmetric positive semidefinite matrices $\mathbb{S}^n_+$.  

3. **Second order cone programs**: The cone $K$ is the set of points satisfying

$$x_1^2 \geq \sum_{i=2}^{n} x_i^2, \quad x_1 \geq 0. \quad (2)$$

4. **$l_p$-norm** for $p > 1$, so

$$K = \{x \in \mathbb{R}^n : \sum_{j=1}^{q} (1/p_j) |c_j - a_j^T x|^{p_j} \leq d - g^T x\} \quad (3)$$

where $d$ and each $c_j$ are scalars, and $g$ and each $a_j$ are $n$-vectors.

5. Can have **direct products** of simpler cones. So can have, eg, a mixture of nonnegative variables and a psd matrix. Or can have multiple second order cones.
$$x_i^2 \geq \sum_{j=2}^{\infty} x_j^2$$

Lorentz cone

“ice cream” cone

\[
\min (\text{norm of } v) + c^T v \\
\text{st. } A v = b
\]

\[
\min \quad \ell + c^T v \\
\text{st. } A v = b \\
\ell \geq \|v\|_2
\]
$x_i \geq \sum_{j=2}^{\infty} x_j^2$ is not a convex constraint:

Can write as:

$g(x) = -x_i^2 + \sum_{j=2}^{\infty} x_j^2 \leq 0$

$$
\nabla^2 g = \begin{bmatrix}
-2 & \cdots \\
\cdots & 2
\end{bmatrix}
$$

not psd

Nonetheless, can solve with interior point methods.

CPLEX, GUROBI etc handle these constraints.
Examples:

1. **Linear programming**: The cone $K$ is the nonnegative orthant $\mathbb{R}_+^n$.

2. **Semidefinite programming**: The cone $K$ is the set of symmetric positive semidefinite matrices $\mathbb{S}_+^n$.

3. **Second order cone programs**: The cone $K$ is the set of points satisfying

   $$ x_1^2 \geq \sum_{i=2}^{n} x_i^2, \quad x_1 \geq 0. $$

4. **$l_p$-norm** for $p > 1$, so

   $$ K = \{ x \in \mathbb{R}^n : \sum_{j=1}^{q} (1/p_j) |c_j - a_j^T x|^{p_j} \leq d - g^T x \} $$

   where $d$ and each $c_j$ are scalars, and $g$ and each $a_j$ are $n$-vectors.

5. Can have **direct products** of simpler cones. So can have, eg, a mixture of nonnegative variables and a psd matrix. Or can have multiple second order cones.
Examples:

1. **Linear programming**: The cone $K$ is the nonnegative orthant $\mathbb{R}_+^n$.

2. **Semidefinite programming**: The cone $K$ is the set of symmetric positive semidefinite matrices $\mathbb{S}_+^n$.

3. **Second order cone programs**: The cone $K$ is the set of points satisfying

   \[ x_1^2 \geq \sum_{i=2}^{n} x_i^2, \quad x_1 \geq 0. \]  

   (2)

4. **$l_p$-norm** for $p > 1$, so

   \[ K = \{ x \in \mathbb{R}^n : \sum_{j=1}^{q} (1/p_j)|c_j - a_j^T x|^{p_j} \leq d - g^T x \} \]  

   (3)

   where $d$ and each $c_j$ are scalars, and $g$ and each $a_j$ are $n$-vectors.

5. Can have **direct products** of simpler cones. So can have, eg, a mixture of nonnegative variables and a psd matrix. Or can have multiple second order cones.
Outline

1. Introduction
2. Duality
3. Path following methods
Lagrangian relaxation

For any \( y \in \mathbb{R}^m \) the Lagrangian relaxation is

\[
\theta(y) := \min_{x \in K} \{ c^T x - (Ax - b)^T y \} \\
= b^T y + \min_{x \in K} \{ (c - A^T y)^T x \} \\
= \begin{cases} 
    b^T y & \text{if } c - A^T y \in K^+ \\
    -\infty & \text{otherwise}
\end{cases}
\]

where \( K^+ = \{ z \in \mathbb{R}^n : z^T x \geq 0 \ \forall x \in K \} \) is the dual cone to \( K \). Hence we obtain the dual problem:

\[
\max_{y,s} \quad b^T y \\
\text{subject to} \quad A^T y + s = c \\
\quad \quad s \in K^+
\]  \hspace{1cm} (4)

The nonnegative orthant \( \mathbb{R}^n_+ \), the sdp cone \( \mathbb{S}^n_+ \), and the second order cone are all \textbf{self-dual}, so \( K^+ = K \).
\( K^+ \): makes \( \angle gb \leq 90^{\circ} \),
for any \( x \in K \),
any \( z \in K^+ \)
\( K = \text{TR}^+ \) : \( x_0 \uparrow \) \( x_1 \uparrow \)

\[ K = K^+ \]

\( K = S^+_+ \) : last time: \( K^+ = S^+_+ \)

Second order cone: also self-dual.
**Strong duality**

The set of strictly feasible solutions to (1) is

$$\mathcal{F}_P^0 := \{ x \in \text{int}(K) : Ax = b \}$$

The set of strictly feasible solutions to (4) is

$$\mathcal{F}_D^0 := \{ y \in \mathbb{R}^m, s \in \text{int}(K^+) : A^T y + s = c \}$$

**Theorem (Nesterov and Todd)**

Assume $\mathcal{F}_P^0$ and $\mathcal{F}_D^0$ are both nonempty. Then

1. the primal (1) and dual (4) problems both have optimal solutions,
2. the optimal values agree, and
3. the sets of optimal solutions are bounded.
Strong duality

The set of strictly feasible solutions to (1) is

$$\mathcal{F}_P^0 := \{x \in \text{int}(K) : Ax = b\}$$

The set of strictly feasible solutions to (4) is

$$\mathcal{F}_D^0 := \{y \in \mathbb{R}^m, s \in \text{int}(K^+) : A^T y + s = c\}$$

Theorem (Nesterov and Todd)

Assume $\mathcal{F}_P^0$ and $\mathcal{F}_D^0$ are both nonempty. Then

1. the primal (1) and dual (4) problems both have optimal solutions,
2. the optimal values agree, and
3. the sets of optimal solutions are bounded.
Strong duality

The set of strictly feasible solutions to (1) is

$$\mathcal{F}_P^0 := \{ x \in \text{int}(K) : Ax = b \}$$

The set of strictly feasible solutions to (4) is

$$\mathcal{F}_D^0 := \{ y \in \mathbb{R}^m, s \in \text{int}(K^+) : A^Ty + s = c \}$$

**Theorem (Nesterov and Todd)**

Assume $\mathcal{F}_P^0$ and $\mathcal{F}_D^0$ are both nonempty. Then

1. the primal (1) and dual (4) problems both have optimal solutions,
2. the optimal values agree, and
3. the sets of optimal solutions are bounded.
Duality

Strong duality

The set of strictly feasible solutions to (1) is

\[ F^0_P := \{ x \in \text{int}(K) : Ax = b \} \]

The set of strictly feasible solutions to (4) is

\[ F^0_D := \{ y \in \mathbb{R}^m, s \in \text{int}(K^+) : A^T y + s = c \} \]

**Theorem (Nesterov and Todd)**

Assume \( F^0_P \) and \( F^0_D \) are both nonempty. Then

1. the primal (1) and dual (4) problems both have optimal solutions,
2. the optimal values agree, and
3. the sets of optimal solutions are bounded.
Outline

1. Introduction
2. Duality
3. Path following methods
Barrier functions

The cones $\mathbb{R}_+^n$, $\mathbb{S}_+^n$, and the second order cone all possess barrier functions $f(x)$ that are logarithmically homogeneous with a parameter $\nu$, where $\nu$ satisfies

$$f(\tau x) = f(x) - \nu \ln \tau$$

for any $x$ in the interior of the cone and for any positive scalar $\tau$.

The parameter $\nu$ is equal to $n$ for $\mathbb{R}_+^n$ and $\mathbb{S}_+^n$; it is equal to 2 for a single second order cone constraint.

The barrier functions also possess a property of self-concordance and the cones possess a property defined by Nesterov and Todd as self-scaled.
Path-following algorithms

Path-following algorithms can be defined for any self-scaled cone with a $\nu$-self concordant logarithmically homogenous barrier function.

Short step path-following methods and predictor-corrector methods both converge in $O(\sqrt{\nu}L)$ iterations.

The theory can be generalized further: it is not necessary for $K \subseteq \mathbb{R}^n$, it just needs to be a closed convex cone in a finite dimensional real vector space $E$ with dual space $E^*$. 