Conic Optimization Refresher

Alex Hahn
from mosek cookbook

November 2, 2022

1 Types of Cones

Conic optimization is a class of convex optimization problems. The geneneral form of a conic optimization problem is:

$$\text{maximize } \mathbf{c}^T\mathbf{x}$$
$$\text{subject to } \mathbf{Ax} + \mathbf{b} \in \mathcal{K}$$

where $\mathcal{K}$ is a product of the following basic types of cones:

- **Linear cone:**

$$\mathbb{R}, \mathbb{R}^n_+, 0$$

- **Quadratic cone and rotated quadratic cone:**

The quadratic cone is the set

$$\mathcal{Q}^n = \left\{ x \in \mathbb{R} \middle| x_1 \geq \sqrt{x_2^2 + \cdots + x_n^2} \right\}$$

The rotated quadratic cone is the set

$$\mathcal{Q}^n_r = \left\{ x \in \mathbb{R} \middle| 2x_1x_2 \geq x_3^2 + \cdots + x_n^2, x_1, x_2 \geq 0 \right\}$$

Together the union of these two cones covers the class of SOCO (second-order cone optimization) problems which includes all QO (quadratic optimization) and QCQO (quadratically constrained quadratic optimization) problems as well.

- **Primal power cone:**

$$\mathcal{P}_n^{\alpha, 1-\alpha} = \left\{ x \in \mathcal{R}^n \middle| x_1^\alpha x_2^{1-\alpha} \geq \sqrt{x_3^2 + \cdots + x_n^2}, x_1, x_2 \geq 0 \right\}$$

- **Primal exponential cone:**

$$K_{\exp} = \left\{ x \in \mathcal{R}^3 \middle| x_1 \geq x_2 \exp\left(\frac{x_3}{x_2}\right), x_1, x_2 \geq 0 \right\}$$

- **Semidefinite cone:**

$$\mathcal{S}_+^n = \{X \in \mathbb{R}^{n\times n} | X \text{ is symmetric positive semidefinite}$$

Semidefinite cones model SDO problems.

Each of these cones allow formulating different types of convex constraints.

2 Selection of Conic Constraints

Examples of real world constraints (financial)and how to convert them to conic form.

2.1 Maximum function

Model the maximum constraint $\max(x_1, x_2, ..., x_n) \leq c$ using n linear constraints introduces with an auxiliary variable t:

$$\begin{aligned} t &\leq c, \\ t &\geq x_1, \\ &\vdots \\ t &\geq x_n. \end{aligned} \tag{1}$$

For example we can write the constraints $\max(x_i, 0) \leq c_i, 1, ..., n$ as

$$\mathbf{t} \leq \mathbf{c}, \mathbf{t} \geq \mathbf{x}, \mathbf{t} \geq \mathbf{0},$$

where $\mathbf{t}$ is an n-dimensional vector.

2.2 Positive and negative part

A special case of modeling the maximum function is to model the positive part x^+ and the negative part x^- of a variable x. We define these as $x^+ = \max(x, 0)$ and $x^- = \max(-x, 0)$. Model them explicitly with the above methodology and the inequatlities $x^+ = \max(x, 0)$ and $x^- = \max(-x, 0)$, or implicitly with the constraints:

$$\begin{aligned} x = x^+ - x^-, \\ x^+, x^- \geq 0. \end{aligned}$$

Note, there is still a degree of freedom in both the implicit and explicit magnitudes of x^+ and x^-. In the explicit case we have inequalities and in the implicit case only the difference of the variables is constrained. Ultimately it is possible for both x^+ and x^- to be positive, allowing optimal solutions where $x^+ = \max(x, 0)$ and $x^-1 = \max(-x, 0)$ does not hold.

Theoretically we can introduce a complimentarity contraint $x^+x^- = 0$ (or $\langle \mathbf{x}^+, \mathbf{x}^- \rangle = 0$) but that is non-convex/ can't be modeled. There are two workarounds to ensure that $x^+ = \max(x, 0)$ and $x^-1 = \max(-x, 0)$ hold the optimal solution: 1) penalize the magnitude of the two solutions, so if both are positive in any one solution, the solver could always improve the objective by reducing them until either one becomes zero. 2) formulate a mixed integer problem.

2.3 Absolute value

We can model the absolute value constraint $|x| \leq c$ by using the maximum function observing that $|x| = \max(x, -x)$:

$$-c \leq x \leq c$$

Or you could use a quadratic cone:

$$(c, x) \in \mathcal{Q}^2$$

2.4 Sum of largest elements

The sum of the m largest elements of a vector $\mathbf{x}$ is the optimal solution of the LO problem:

$$\text{maximize } \mathbf{x}^T\mathbf{z}$$

$$\text{subject to } \mathbf{1}^T\mathbf{z} = m,$$

$$\mathbf{0} \leq \mathbf{z} \leq \mathbf{1}$$

Here $\mathbf{x}$ cannot be a variable as this would result in a nonlinear objective. Looking at the dual of the problem:

$$\text{minimize } mt + \mathbf{1}^T\mathbf{u}$$

$$\text{subject to } \mathbf{u} + t \geq \mathbf{x},$$

$$\mathbf{u} \geq 0.$$

This is the same problem as $\min_t \; mt + \sum_i \max(0, x_i - t)$, in which x can be a variable and thus optimized.

2.5 Linear combination of largest elements

Selecting $\mathbf{z}$ tot have an upper bound $\mathbf{c} \geq \mathbf{0}$, and a real number $0 \leq b \leq c_{\text{sum}}$ instead of integer m, where $c_{\text{sum}} = \sum_i c_i$:

$$\text{maxmize } \mathbf{x}^T\mathbf{z}$$

$$\text{subject to } \mathbf{1}^T\mathbf{Z} = c_{\text{sum}} - b,$$

$$\mathbf{0} \leq \mathbf{z} \leq \mathbf{c}.$$

This has the optimal objective $c^{\text{frac}} x_{ib} + \sum_{i > i_b} c_i x_i$ where i_b is such that $\sum_{i=1}^{i_b - 1} c_i < b \leq \sum_{i=1}^{i_b} c_i$, and $c_{i_b}^{\text{frac}} = \sum_{i=1}^{i_b} c_i - b < c_{i_b}$. The dual of this problem:

$$\text{minimize } (c_{\text{sum}} - b)t + \mathbf{c}^T\mathbf{u}$$

$$\text{subject to } \mathbf{u} + t \geq \mathbf{x},$$

$$\mathbf{u} \geq \mathbf{0}.$$

which is the same as $\min_t (c_{\text{sum}} - b)t + \sum_i c_i \max(0, x_i - t)$.

2.6 Manhattan Norm

Let $\mathbf{x} \in \mathbb{R}^n$ with the standard 1-norm / Manhattan norm. A 1-norm constrant $||\mathbf{x}||_1 \leq c||$ can be formed by modeling the absolute value for each coordinate:

$$\mathbf{z} \leq x \leq \mathbf{z}, \sum_{i=1}^{n} z_i = c,$$

where $\mathbf{z}$ is an auxilliary variable

2.7 Euclidean Norm

Let $\mathbf{x} \in \mathbb{R}^n$ with the standard 2-norm / Euclidean norm definition $||\mathbf{x}||_2 = \sqrt{x_1^2 + \cdots + x_n^2}$. A Euclidean norm constraint $||\mathbf{x}||_2 \leq c$ can be modelled using a quadratic cone:

$$(c, \mathbf{x}) \in \mathcal{Q}^{n+1}$$

2.8 Squared Euclidean Norm

Let $\mathbf{x} \in \mathbb{R}^n$ take the square of the Euclidean norm $||\mathbf{x}||_2^2 = \mathbf{x}^T\mathbf{x} = x_1^2 + \cdots + x_n^2$. We can model the squared Euclidean norm or sum-of-squared constraint $||\mathbf{x}||_2^2 \leq c$ using a rotated quadratic cone:

$$(c, \frac{1}{2}, \mathbf{x}) \in \mathcal{Q}_r^{n+2}$$

2.9 Quadratic Form

Let $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{Q} \in S_+^n$ ie a symmetric positive semidefinite matrix. We can model a quadratic form constraint $\frac{1}{2}\mathbf{x}^T\mathbf{Q}\mathbf{x} \leq c$ using either a quadratic cone or a rotated quadratic cone. Observe that there exists a (non unique) matrix $\mathbf{G} \in \mathbb{R}^{n \times k}$ such that $\mathbf{Q} = \mathbf{G}\mathbf{G}^T$. Common ways of computing $\mathbf{G}$ are:

- Cholesky decomposition: $\mathbf{Q} = \mathbf{C}\mathbf{C}^T$, where $\mathbf{C}$ is a lower triangular matrix with nonnegative entries on the diagonal. With this decompostion we have $\mathbf{G} = \mathbf{C} \in \mathbf{R}^{n \times n}$.
- Eigenvalue decomposition: $\mathbf{Q} = \mathbf{V}\mathbf{D}\mathbf{V}^T$, where the diagonal matrix $\mathbf{D}$ contains the (nonnegative) eigenvalues of $\mathbf{Q}$ and the unitary matriix $\mathbf{V}$ contains the corresponding eigenvectors in its columns. From this decomposition we have $\mathbf{G} = \mathbf{V}\mathbf{D}^{\frac{1}{2}} \in \mathbb{R}^{n \times n}$.
- Matrix square root: $\mathbf{Q} = \mathbf{Q}^{1/2}\mathbf{Q}^{1/2}$, where $\mathbf{Q}^{1/2}$ is the symmetric positive semidefinite "square root" matrix of $\mathbf{Q}$. From this decomposition we have $\mathbf{G} = \mathbf{Q}^{1/2} \in \mathbb{R}^{n \times n}$.
- Factor model: if $\mathbf{Q}$ is a covariance matrix of some date, then we can approximate the data series with the combination of $k \ll n$ common factors. We have the decomposition $\mathbf{Q} = \beta\mathbf{Q}_F\beta^T + \mathbf{D}$, where $\mathbf{Q}_F \in \mathbb{R}^{k \times k}$ is the covariance of the factors, $\beta \in \mathbb{R}^{n \times k}$ is the exposure of the data series to each factor, and $\mathbf{D}$ is diagonal. From this, by computing the Cholesky decomposition $\mathbf{Q}_F = \mathbf{F}\mathbf{F}^T$ we have $\mathbf{G} = [\beta\mathbf{F}, \mathbf{D}^{1/2}] \in \mathbb{R}^{n \times (n+k)}$. The advantage of factor models is that $\mathbf{G}$ is very sparse and the factors have a direct financial interpretation.

After obtaining $\mathbf{G}$, we can write the quadratic form constraint as a sum-of-squares $\frac{1}{2}\mathbf{x}^T G^T\mathbf{x} \leq c$ which is a squared Euclidean norm constraint $\frac{1}{2}||\mathbf{G}^T\mathbf{x}||_2^2 \leq c$. We can choose to model this using the rotated quadratic cone as

$$(c, 1, \mathbf{G}^T\mathbf{x}) \in \mathcal{Q}^{k+2},$$

or we can choose to model its square root using the quadratic cone as

$$(\sqrt{c}, \mathbf{G}^T\mathbf{x}) \in \mathcal{Q}^{k+1}$$

Usually the quadratic cone is used to model 2-norm constraints while the rotated quadratic cone is used to model quadratic functions.

2.10 Power

Let $x \in \mathbb{R}$ and $\alpha > 1$, we can model a power constraint $c \geq |x|^\alpha$ or equivalently $c^{1/\alpha} \geq |x|$ using a power cone:

$$(c, 1, x) \in \mathcal{P}_3^{1/\alpha,(\alpha-1)/\alpha}$$

2.11 Exponential

Let $x \in \mathbb{R}$ and $\alpha > 1$, we can model an exponential constraint $t \geq e^x$ using the exponential cone:

$$(t, 1, x) \in K_{\exp}$$