Conic Programming

Ranajit Saha
MTech CSE, 2nd Year
Roll No. 20172119
IIIT, Hyderabad
Hyderabad, India
ranajit.saha@students.iiit.ac.in

***Abstract*—This report discusses the find outs of the study on Conic Programming. The report is divided into four parts. Part I explains about Conic Linear Programming with examples. Part II discusses about Conic Optimization and the duality in Conic Programming. Part III describes Quadratic Cones and Part IV discusses about Power Cones and Exponential Cones. The example codes are available at https://github.com/rnjtsh/Conic-Programming.**

***Index Terms*—optimization, conic programming**

I. Conic Linear Programming

A subset C of a vector space V is a *cone* (also called a *linear cone*) if for each $x \in C$ and positive scalars α, the product αx is in C. A cone C is a *convex cone* if αx + βy belongs to C, for any positive scalars α, β, and any x, y in C.

Given a closed convex cone $K \subset \xi$, $C \in \xi$, $A_i \in \xi$, $i = 1, 2, ..., m$, and $b \in \mathbb{R}^m$, the *conic linear programming* (CLP) problem is to find a matrix X for the optimization problem in a canonical form.

$$\begin{gathered} Inf \quad C \bullet X \\ s.t. \quad A_i \bullet X = b_i, \;\; i = 1, 2, ..., m \,, X \in \xi \\ X \succeq \mathbf{0} \end{gathered} \tag{1}$$

The minimum value of the object may exist but may not be obtained at a finite solution. Hence we use *Inf*. The $\bullet$ operation is the standard inner product defined as

$$A \bullet B := tr(A^T B)$$

When $K = \mathbb{R}^n_+$, an element in K is conventionally written as $x \geq 0$ or x is componentwise nonnegative; while when $K = \mathbb{S}^n_+$, an element in K is conventionally written as $X \succeq 0$ or X is a positive semi-definite matrix. Furthermore, $X \succeq 0$ means that X is a positive definite matrix. If a point X is in the interior of K and satisfies all equations in *CLP*, it is called a (primal) strictly or interior feasible solution.

The characteristics of *CLP* are listed as:

- Linear Objective Functions/Constraints
- Variables are in a pointed and closed convex cone
- Conic interior, boundary, extreme point (corner)
- Every local optimizer is global, and they form a convex (optimizer) set
- Most of them possess efficient algorithms in both practice and theory (polynomial-time)

Conic Programming (CP) is a problem of optimizing a linear objective over a convex set, and thus is a convex problem. The examples of conic programming include the following class of problems.

A. Linear Programming (LP)

In *LP*, the variables form a vector which is required to be component-wise non-negative, while in CLP they are points in a pointed convex cone. When $X, A_i, C \in \mathbb{R}^n$ and $K \in \mathbb{R}^n_+$, *CLP* reduces to *LP*.

B. Second Order Conic Programming (SOCP)

A simple type of closed convex pointed cone that captures many optimization problems of interest is the second order cone, also called the *Lorentz Cone* or "*ice cream cone*". Mathematically, the *Second Order Cone* (*SOC*) is defined as $\mathcal{Q}^n = \{X \in \mathbb{R}^n \mid X_n^2 \geq \sum_{i=1}^{n-1} X_i^2, X_n \geq 0\}$. When $\mathcal{Q}^n$ is used, the problem reduces to SOCP.

Let $\xi = \mathbb{R}^{n_1+1} \times \mathbb{R}^{n_2+1} \times ... \times \mathbb{R}^{n_l+1}$ and $K = \mathcal{Q}^{n_1+1} \times \mathcal{Q}^{n_2+1} \times ... \times \mathcal{Q}^{n_l+1}$ The standard form of SOCP is reduced as

$$\begin{gathered} Inf \quad \sum_{j=1}^{l} c_j^T x_j \\ s.t. \quad \sum_{j=1}^{l} a_{ij}^T x_j = b_i, \quad i = 1, 2, ..., m, \\ x_i \succeq_{\mathcal{Q}^{n_i+1}} \mathbf{0}, \qquad i = 1, 2, ..., l, \end{gathered} \tag{2}$$

where $c_j \in \mathbb{R}^{n_j+1}$ and $a_{ij} \in \mathbb{R}^{n_j+1}$, $for \; i = 1, 2, ...m,$ $and \; j = 1, 2, ..., l.$

C. Semi-Definite Programming (SDP)

SDP is a kind of CLP, where the variable points are symmetric matrices constrained to be positive semi-definite. When $X, A_i, C \in \mathbb{S}^n$ and $\xi \in \mathbb{S}^n_+$, *CLP* reduces to *SDP*. The standard form of *SDP* is written as

$$\begin{gathered} Inf \quad C \bullet X \\ s.t. \quad A_i \bullet X = b_i, \; i = 1, 2, ..., m \,, X \in \xi \\ X \succeq \mathbf{0}, \end{gathered} \tag{3}$$

where $C, A_i \in \mathbb{S}^n$ and $b_i \in \mathbb{R}$.

The relation among these classes of problems is $LP \subset SOCP \subset SDP$.

Example 1. *Antenna array weight design:*
The problem here is to determine how to combine the outputs from each antenna in an array of n antennae so that the resulting aggregate output signal has certain desired characteristics. For concreteness, we assume that the antennae are evenly spaced along the y-axis (in a 2-dimensional world). We also assume that the input signal has a certain fixed wavelength λ but that it can arrive from any direction θ. For certain values of θ, the so-called *side-lobe S*, we wish to attenuate the signal as much as possible whereas for other values, those in the *main-lobe M*, we wish to preserve some given desired output profile. The optimization problem then is

$$\begin{aligned} minimize \quad & s \\ s.t. \quad & |G(\theta)| \leq s, \quad \theta \in S \\ & |G(\theta)| \leq u(\theta)^2, \quad \theta \in M \\ & G(\theta) = G_0(\theta), \quad \theta \in P \\ G(\theta) = \sum_{k=1}^{n} w_k \ & exp(-i\frac{2\pi}{\lambda} y_k \sin\theta), \quad \theta \in S \cup M \end{aligned}$$

Here, i = $\sqrt{-1}$, y_k is the position of the k-th antenna on the y-axis, the $u(\theta)$s are given upper bounds for the $|G(\theta)|$s when $\theta \in M$, and the $G_0(\theta)$s are given desired values for the $G(\theta)$s when θ belongs to a small finite set P of angles in M. The variables in the model are the scalar s, the weights w_k, $k = 1, ..., n$ and the array outputs $G(\theta)$, $\theta \in S \cup M$.

II. Conic Optimization and the Duality

Conic optimization is a subfield of convex nonlinear optimization that studies problems consisting of minimizing a convex function over the intersection of an affine subspace and a convex cone. These class of problems lies between linear programming (LP) problems and general convex nonlinear problems. A conic optimization problem can be written as an LP – with a linear objective and linear constraints – plus one or more cone constraints. A cone constraint specifies that the vector formed by a set of decision variables is constrained to lie within a closed convex pointed cone. The simplest example of such a cone is the *non-negative orthant*, the region where all variables are non-negative – the normal situation in an LP. But conic optimization allows for more general cones.

The standard (*primal*) form of the *Conic Optimization* problem (*CLP*) can be written as

$$\begin{aligned} minimize \quad & C \bullet X \\ s.t. \quad & AX = b, \ X \in K \end{aligned}$$

where K is the convex cone. K can be written as a product $K = K_1 \times K_2 \times ... \times K_m$ of smaller cones corresponding to actual constraints in the problem formulation.

The *dual* form of conic optimization (*CLD*) is

$$\begin{aligned} maximize \quad & b^T y \\ s.t. \quad & \sum_{i=1}^{n} y_i A_i + S = C, \ S \in K^* \end{aligned}$$

where K^* is the dual cone of K and defined as

$$K^* = \{S \in \mathbb{R}^n \mid S^T X \geq 0, \ \forall X \in K\}.$$

When $K = K^*$, the cone is called a *self dual*.

Example 1. The following example is a *SOCP* primal problem:

$$\begin{aligned} minimize \quad & 2x_1 + x_2 + x_3 \\ s.t. \quad & x_1 + x_2 + x_3 = 1, \\ & \sqrt{x_2^2 + x_3^2} \leq x_1. \end{aligned}$$

The dual of this problem is:

$$\begin{aligned} maximize \quad & 1.y \\ s.t. \quad & y \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + S = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}, \\ S = (S_1, & S_2, S_3) \in K^* = \mathcal{N}_3^2 \end{aligned}$$

Theorem II.1 (Weak Duality for Conic Optimization)**.** *Let $\mathcal{F}_p$ and $\mathcal{F}_d$ denote the feasible regions of (CLP) and (CLD), respectively, and be non-empty. Then,*

$$C \bullet X \geq b^T y, \ \forall X \in \mathcal{F}_p, \ (y, S) \in \mathcal{F}_p.$$

The quantity, $C \bullet X - b^T y$ is called the *duality* or *complementarity gap* of *CLP* and *CLD*, as in *Linear Programming*.

Proof. By direct calculation we can prove the *weak duality*.

$$\begin{aligned} C \bullet X - b^T y &= \left(\sum_{i=1}^{m} y_i A_i + S \right) \bullet X - b^T y \\ &= \sum_{i=1}^{m} y_i (A_i \bullet X) + S \bullet X - b^T y \\ &= \sum_{i=1}^{m} y_i b_i + S \bullet X - b^T y \\ &= S \bullet X \\ &\geq 0 \, [\because X \in K, \ S \in K^*] \\ \therefore C \bullet X \geq b^T y. & \end{aligned}$$

□

Corollary II.1.1. *If CLP or CLD is feasible but unbounded (i.e. its objective value is unbounded) then the other is infeasible or has no feasible solution.*

Corollary II.1.2. *If a pair of feasible solutions can be found to the primal and dual problems with an equal objective value, then they are both optimal.*

Linear programming admits a strong duality theorem: when both $\mathcal{F}_p$ and $\mathcal{F}_d$ are nonempty, then there is no gap at optimality. But such *strong duality* theorem doesn't hold for conic linear programming in general. However, under certain technical conditions, there would be no duality gap at optimality. The version of the strong duality theorem for *CLP* and *CLD* is stated in Theorem II.2.

Theorem II.2 (Strong Duality for Conic Optimization)**.**

i *Let CLP or CLD be infeasible, and furthermore the other be feasible and has an interior. Then the other is unbounded.*

ii *Let CLP and CLD be both feasible, and furthermore one of them has an interior. Then there is no duality gap at optimality between CLP and CLD.*

iii *Let CLP and CLD be both feasible and have interior. Then, both have optimal solutions with no duality gap.*

If one of *CLP* and *CLD* has no interior feasible solution, the common optimal objective value may not be attainable. For example,

$$C = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix},\ A_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},\ b_1 = 2.$$

The dual is feasible but has no interior, while the primal has an interior. The common objective value equals 0, but no primal solution attaining the infimum value.

Example 2. *Application of Conic Duality in Linear Regression:*

Least-squares linear regression is the problem of minimizing $\|Ax - b\|_2$ over $x \in \mathbb{R}^n$, where $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$ are fixed. This problem can be posed in conic form as

$$\begin{aligned} &minimize \quad t \\ &s.t. \quad (t, Ax - b) \in \mathcal{Q}^{m+1}, \end{aligned}$$

where $\mathcal{Q}^{m+1}$ is the $(m+1)$ - dimensional *quadratic cone*. The Lagrangian of the problem is
$\mathcal{L}(t, x, u, v) = t - tu - v^T(Ax - b) = t(1-u) - x^T A^T v + b^T v$.
The constraints in the dual problem are:

$$u = 1, A^T v = 0, (u, v) \in \mathcal{Q}^{m+1}.$$

III. Quadratic Cones

We define the n-dimensional *quadratic cone* as

$$\mathcal{Q}^n = \{x \in \mathbb{R}^n \mid x_n^2 \geq \sum_{i=1}^{n-1} x_i^2,\ x_n \geq 0\}.$$

The geometric representation of the quadratic cones is shown in Figure 1 for a cone with three variables and illustrates how the boundary of the cone resembles an ice-cream cone. The 1-dimensional quadratic cone simply states non-negativity $x_1 \geq 0$.
The examples of *quadratic cones* are:

- Epigraph of absolute value: $|x| \leq t \iff (t, x) \in \mathcal{Q}^2$
- Epigraph of Euclidean norm: $\|x\|_2 \leq t \iff (t, x) \in \mathcal{Q}^{n-1}$ where $x \in \mathbb{R}^n$ and $\|x\|_2$ is the *L2-norm* or the *Euclidean norm* of x.

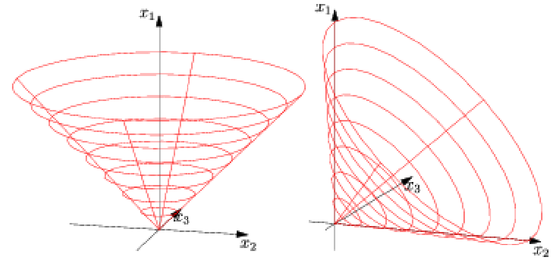

Figure 1. Quadratic Cones: Boundary of quadratic cone (left) $x_1 \geq \sqrt{x_2^2 + x_3^2}$ and rotated quadratic cone (right) $2x_1x_2 \geq x_3^2$, $x_1, x_2 \geq 0$

- Second-order cone inequality: $\|Ax + b\|_2 \leq c^T x + d \iff (c^T x + d, Ax + b) \in \mathcal{Q}^{m+1}$ where $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$, $d \in \mathbb{R}$
- Ellipsoidal set: The set $\xi = \{x \in \mathbb{R}^n \mid \|P(x - c)\|_2 \leq 1\}$ describes an ellipsoid centred at c. It has a natural conic quadratic representation, i.e., $x \in \xi$ if and only if $x \in \xi \iff (1, P(x - c)) \in \mathcal{Q}^{n+1}$.

An n-dimensional *rotated quadratic cone* is defined as $\mathcal{Q}_r^n = \{x \in \mathbb{R}^n \mid 2x_1x_2 \geq \sum_{i=3}^n x_i^2,\ x_1, x_2 \geq 0\}$.
The examples of *quadratic cones* are:

- Epigraph of squared Euclidean norm: The epigraph of the squared Euclidean norm can be described as the intersection of a rotated quadratic cone with an affine hyperplane, $\|x\|_2 \leq t \iff (1/2, t, x) \in \mathcal{Q}_r^{n+2}$.
- Square roots: The set of square roots $\sqrt{x} \geq t, x \geq 0 \iff (1/2, x, t) \in \mathcal{Q}_r^3$ forms a 3-D *rotated quadratic cone*.
- Convex hyperbolic function: $1/x \leq t, x > 0 \iff (x, t, \sqrt{2}) \in \mathcal{Q}_r^3$
- Convex positive rational power: $x_{2/3} \leq t, x \geq 0 \iff (s, t, x), (x, 1/8, s), \in \mathcal{Q}_r^3$
- Convex negative rational power: $1/x^2 \leq t, x \geq 0 \iff (t, 1/2, s), (s, x, \sqrt{2}) \in \mathcal{Q}_r^3$

IV. Power Cones

The quadratic and rotated quadratic cone family can be expanded to *Power Cones* which expresses models involving powers other than 2. The power cones include the quadratic cones as special cases. But solving the power cone problems require more advanced and less efficient algorithms than solving quadratic cones at the current state-of-the-art.
The n-dimensional *power cones* form a family of convex cones parameterized by a real number $0 < \alpha < 1$:

$$\begin{aligned} \mathcal{P}_n^{\alpha,1-\alpha} = \{x \in \mathbb{R}^n \mid x_1^\alpha x_2^{1-\alpha} \geq \sqrt{x_3^2 + x_4^2 + ... + x_n^2}, \\ x_1, x_2 \geq 0\}. \end{aligned}$$

The constraint in the definition of $P_n^{\alpha,1-\alpha}$ can be expressed as a composition of two constraints, one of which is a quadratic cone:

$$\begin{aligned} x_1^\alpha x_2^{1-\alpha} &\geq |z|, \\ z &\geq \sqrt{x_3^2 + x_4^2 + ... + x_n^2}. \end{aligned}$$

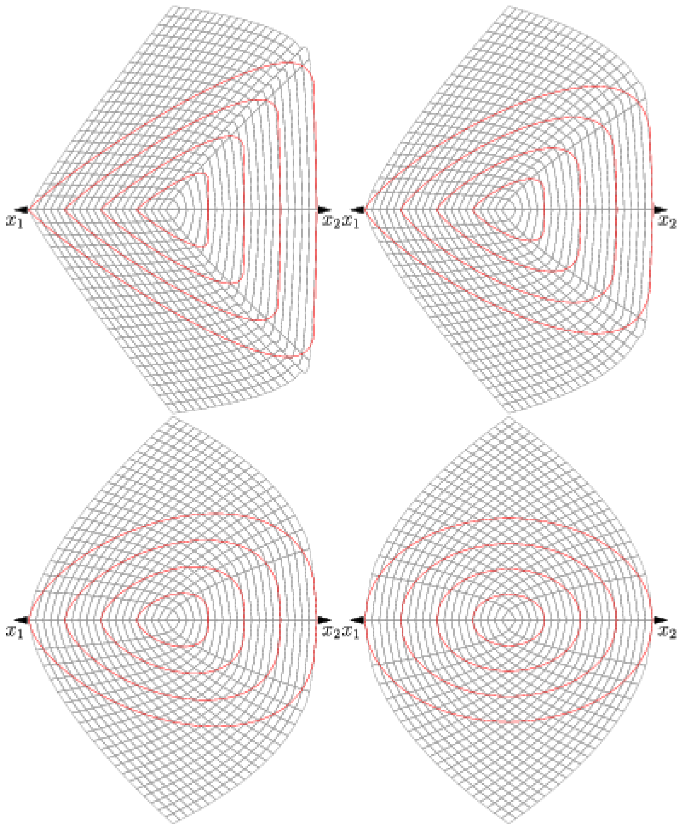

Figure 2. Power Cones: The boundary of $\mathcal{P}_3^{\alpha,1-\alpha}$ seen from a point inside the cone for α = 0.1, 0.2, 0.35, 0.5

This means that the basic building block is the three-dimensional power cone

$$\mathcal{P}_3^{\alpha,1-\alpha} = \{x \in \mathbb{R}^3 \mid x_1^{\alpha} x_2^{1-\alpha} \geq |x_3|,\ x_1, x_2 \geq 0\}.$$

More generally, *power cones* are defined as

$$\mathcal{P}_n^{\alpha_1,\ldots,\alpha_m} = \{x \in \mathbb{R}^n \mid \prod_{i=1}^{m} x_i^{\alpha_i} \geq \sqrt{\sum_{i=m+1}^{n} x_i^2}, \quad x_1, ..., x_m \geq 0\},$$

for $m < n$ and a sequence of exponents $\alpha_1, \alpha_2, ..., \alpha_m$. The examples of constraints which can be expressed as *power cones* are:

- *Powers*: For all values of $p \neq 0, 1$ we can bound x^p depending on the convexity of $f(x) = x^p$.
  - For $p > 1$ the inequality $t \geq |x|_p$ is equivalent to $t^{1/p} \geq |x|$ and hence corresponds to $t \geq |x|^p \iff (t, 1, x) \in \mathcal{P}_3^{1/p, 1-1/p}$.
  - For $0 < p < 1$ the function $f(x) = x^p$ is concave for $x \geq 0$ and hence $t \leq |x|^p, x \geq 0 \iff (x, 1, t) \in \mathcal{P}_3^{p,1-p}$.
  - For $p < 0$ the function $f(x) = x^p$ is convex for $x > 0$ and in this range the inequality $t \geq x^p$ is equivalent to $t \geq |x|^p \iff t^{1/(1-p)} x^{-p/(1-p)} \geq 1 \iff (t, x, 1) \in \mathcal{P}_3^{1/(1-p), -p/(1-p)}$. For example $t \geq 1/\sqrt{x}$ is the same as $(t, x, 1) \in \mathcal{P}_3^{2/3, 1/3}$.
- *p-norm cones*: The *p-norm cone* is an example of the power cone. For $p \geq 1$, the p-norm cone is a convex set, defined as $\{(t, x) \in \mathbb{R}^{n+1} \mid t \geq \|x\|_p\}$. For $p = 2$, this reduces to *quadratic cone*. When $0 < p < 1$ or $p < 0$ the formula for $\|x\|_p$ gives a concave set and the set is then modeled as $\{(t, x) \in \mathbb{R}^{n+1} \mid 0 \leq t \leq \|x\|_p,\ x_i \geq 0\},\ p < 1, p \neq 0$.

V. Exponential Cones

The *exponential cone* is a convex subset of $\mathbb{R}^3$ defined as

$$K_{exp} = \{x \in \mathbb{R}^3 \mid x_1 \geq x_2 e^{x_3/x_2},\ x_2 > 0\} \cup \{\mathbb{R}_+ \times \{0\} \times \mathbb{R}_-\}$$

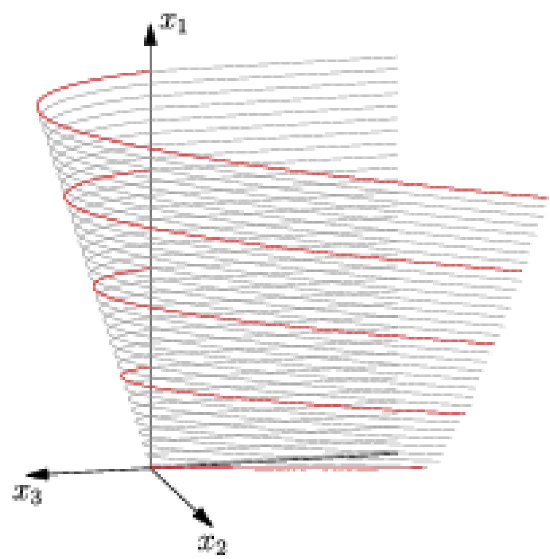

Figure 3. Exponential Cones: The boundary of the exponential cone K_{exp}. The red lines are graphs of $x_2 \to x_2 \log(x_1/x_2)$ for fixed x_1

The *exponential cone* can be used to model a variety of constraints involving exponentials and logarithms. The simple examples of *exponential cones* are:

- *Epigraph of exponential*: $t \geq e^x \iff (t, 1, x) \in K_{exp}$
- *Epigraph of negative logarithm*:

$$-\log x \leq t \iff (x, 1, -t) \in K_{exp}$$

- Epigraph of negative entropy:

$$x \log x \leq t \iff (1, x, -t) \in K_{exp}$$

- *Softplus function*: In neural networks the function $f(x) = \log(1 + e^x)$, known as the *softplus function*, is used as an analytic approximation to the rectifier activation function $r(x) = x^+ = max(0, x)$. The softplus function is convex and we can express its epigraph $t \geq \log(1 + e^x)$ by combining two exponential cones.

$$t \geq \log(1 + e^x) \iff e^{x-t} + e^{-t} \leq 1$$

Therefore $t \geq \log(1 + e^x)$ is equivalent to the following set of conic constraints:

$$
\begin{gathered}
u + v \leq 1, \\
(u, 1, x - t) \in K_{exp}, \\
(v, 1, -t) \in K_{exp}.
\end{gathered}
$$

REFERENCES

[1] Convex Cone Wikipedia: https://en.wikipedia.org/wiki/Convex_cone

[2] Book titled Conic Linear Programming by Yinyu Ye

[3] mosek modelling coookbook: https://docs.mosek.com/modeling-cookbook/intro.html

[4] R.J. Vanderbei and H. Yurttan, "Using LOQO to solve second-order cone programming problems," Report SOR 98-09, Princeton University, 1998.