Linear Conic Optimization

Wenxun Xing

Department of Mathematical Sciences Tsinghua University wxing@tsinghua.edu.cn Office hour: Thur. 4:00-5:00 Based on OR706 of Prof. Fang at NCSU

April, 2023

Content

- Part I: Introduction
- Part II: Preliminary and Convex Cone Structure
- Part III: Duality Theory of Linear Conic Programs
- *Part IV: Interior Point Method
- Part V: Examples and Applications
- Part VI: Quadratically Constrained Quadratic Programs

Linear Conic Optimization Part I

Wenxun Xing

Department of Mathematical Sciences Tsinghua University wxing@tsinghua.edu.cn Office hour: Thur. 4:00-5:00 Based on OR706 of Prof. Fang at NCSU

April, 2023

Introduction

Content

- Linear Conic Programs
- Applications
- Duality Theory and Algorithms
- References

Linear Conic Program: Standard Form

Min
$$C \bullet X$$

s.t. $A_i \bullet X = B_i, i = 1, 2, ..., m$ (LCoP)
 $X \in K$

where K is a closed, convex cone; C, A and B are in the space of interests with \bullet being an appropriate linear operator.

Cone K: $\forall X \in K$ and $\alpha > 0$, we have $\alpha X \in K$.

Linear Programming: An example

Nutrition table (mg/g)		
	Food 1	Food 2
Vitamin B1	0.005	0.004
Phosphorus	0.027	0.060
Iron	0.046	0.039

- Daily demand: B1 1.5 mg, Phosphorus 8 mg, Iron 12 mg.
- Cost: Food 1, 0.40 yuan/g, Food 2, 0.3 yuan/g.
- Aim: Use less (minimum) money to buy foods.

$$\begin{array}{ll} \min & 0.4x_1 + 0.3x_2 \\ s.t. & 0.005x_1 + 0.004x_2 \geq 1.5 \\ & 0.027x_1 + 0.06x_2 \geq 8 \\ & 0.046x_1 + 0.039x_2 \geq 12 \\ & x_1 \geq 0, x_2 \geq 0. \end{array}$$

Standard form

$$\begin{array}{ll} \min & 0.4x_1 + 0.3x_2 \\ s.t. & 0.005x_1 + 0.004x_2 - x_3 = 1.5 \\ & 0.027x_1 + 0.06x_2 - x_4 = 8 \\ & 0.046x_1 + 0.039x_2 - x_5 = 12 \\ & x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_4 \geq 0, x_5 \geq 0. \end{array}$$

Linear Programming-Inequivalent Form

When $K = \mathbb{R}^n_+ = \{x \in \mathbb{R}^n | x_i \geq 0, \ i = 1, ..., n\}$, symmetric form of LP.

$$\begin{array}{ll} Min & c^Tx \\ s.t. & Ax \geq b \\ & x \geq_{\mathbb{R}^n_+} 0 \end{array} \tag{LP}$$

where $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ and $c \in \mathbb{R}^n$.

Dual of LP:

$$\begin{array}{ll}
Max & b^T y \\
s.t. & A^T y \le c \\
& y \ge_{\mathbb{R}^m} 0
\end{array} \tag{LD}$$

Linear Programming—Standard Form:

$$K = \mathbb{R}^n_+$$

When $K = \mathbb{R}^n_+ = \{x \in \mathbb{R}^n | x_i \geq 0, \ i = 1,...,n\}$, LCoP becomes LP.

$$\begin{array}{ll} Min & c^T x \\ s.t. & Ax = b \\ & x \geq_{\mathbb{R}^n_+} 0 \end{array} \tag{LP}$$

where $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ and $c \in \mathbb{R}^n$.

Dual of LP:

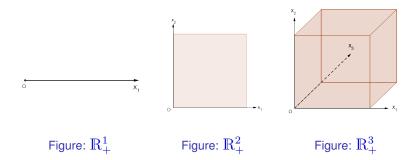
$$Max \quad b^{T}y$$
s.t.
$$A^{T}y + s = c$$

$$s \ge_{\mathbb{R}^{n}} 0$$
 (LD)

where $y \in \mathbb{R}^m$ and $s \in \mathbb{R}^n$.



$K = \mathbb{R}^n_+$



Second-Order Cone (SOC) Programming:

$$K = \mathcal{L}^n$$

When $K=\mathcal{L}^n=\{x\in\mathbb{R}^n|\sqrt{x_1^2+\cdots+x_{n-1}^2}\leq x_n\}$, LCoP becomes SOCP.

$$\begin{array}{ll}
Min & c^T x \\
s.t. & Ax = b \\
 & x \ge_{\mathcal{L}^n} 0
\end{array}$$
(SOCP)

where $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ and $c \in \mathbb{R}^n$.

Dual of SOCP:

$$Max \quad b^{T}y$$

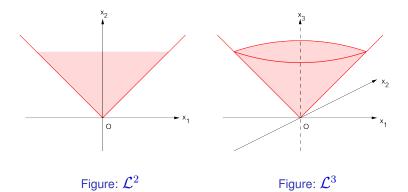
$$s.t. \quad A^{T}y + s = c$$

$$s >_{\mathcal{L}^{n}} 0$$
(SOCD)

where $y \in \mathbb{R}^m$ and $s \in \mathbb{R}^n$.



$K = \mathcal{L}^n$



Application of SOCP

Torricelli Point Problem

The problem is proposed by Pierre de Fermat in 17th century. Given three points $a,\ b$ and c on the \mathbb{R}^2 plane, find the point in the plane that minimizes the total distance to the three given points. The solution method was found by Torricelli, hence know as Torricelli point.

SOCP Formulation

Min
$$t_1 + t_2 + t_3$$

s.t. $\begin{bmatrix} x - a \\ t_1 \end{bmatrix} \in \mathcal{L}^3, \begin{bmatrix} x - b \\ t_2 \end{bmatrix} \in \mathcal{L}^3, \begin{bmatrix} x - c \\ t_3 \end{bmatrix} \in \mathcal{L}^3$

Question:

$$C = ?$$
 $A = ?$ $X = ?$



Standard Formulation of Torricelli Point Problem

Let

$$y = x - a$$
, $z = x - b$, $w = x - c$.

Standard Formulation

Min
$$t_1 + t_2 + t_3$$

s.t. $y - z = b - a$
 $y - w = c - a$
 $\begin{bmatrix} y \\ t_1 \end{bmatrix} \in \mathcal{L}^3, \begin{bmatrix} z \\ t_2 \end{bmatrix} \in \mathcal{L}^3, \begin{bmatrix} w \\ t_3 \end{bmatrix} \in \mathcal{L}^3$

$$\begin{bmatrix} y \\ t_1 \end{bmatrix} \in \mathcal{L}^3, \begin{bmatrix} z \\ t_2 \end{bmatrix} \in \mathcal{L}^3, \begin{bmatrix} w \\ t_3 \end{bmatrix} \in \mathcal{L}^3 \\ \iff (y_1, y_2, t_1, z_1, z_2, t_2, w_1, w_2, t_3)^T \in \mathcal{L}^3 \times \mathcal{L}^3 \times \mathcal{L}^3.$$

Application of SOCP

Robust Portfolio Design

Assume returns r are known within an ellipsoid

$$\mathcal{E} = \{ r = \hat{r} + \kappa \Sigma^{1/2} u : ||u||_2 \le 1 \}.$$

where \hat{r} is the expected return, Σ is the empirical covariance matrix, $0 < \kappa < 1$ is a given constant.

robust counterpart: (optimize the worst case)

$$\max_{\omega} \min_{r \in \mathcal{E}} \{ r^T \omega : e^T \omega = 1, \ \omega \ge 0 \}.$$

Application of SOCP

SOCP Formulation

Notice that

$$\begin{aligned} & & \min_{r \in \mathcal{E}} \ r^T \omega \\ &= & \min_{\|u\|_2 \le 1} \left\{ \hat{r}^T \omega + \kappa u^T \Sigma^{1/2} \omega \right\} \\ &= & \hat{r}^T \omega - \kappa \|\Sigma^{1/2} \omega\|_2 \end{aligned}$$

Robust portfolio problem is an SOCP

$$\begin{array}{lll} Max & \hat{r}^T\omega - \kappa \|\Sigma^{1/2}\omega\|_2 \\ s.t. & e^T\omega = 1, \ \omega \geq 0 \end{array} \iff \begin{array}{lll} Max & t \\ s.t. & e^T\omega = 1, \ \omega \geq 0 \\ & \left[\begin{array}{c} \kappa \Sigma^{1/2}\omega \\ \hat{r}^T\omega - t \end{array}\right] \in \mathcal{L}^{n+1} \end{array}$$

Question:

$$C = ?$$

$$C = ?$$
 $A = ?$ $X = ?$

$$X = ?$$

Other Applications - QCQP ⇒ SOCP

The popularity of SOCP is also due to the fact that it is a generalized form of convex QCQP (Quadratically Constrained Quadratic Programming). Specifically, consider the following QCQP:

Min
$$x^T A_0 x + 2b_0^T x + c_0$$

s.t. $x^T A_i x + 2b_i^T x + c_i \le 0, i = 1, ..., m$

where $A_i \succeq 0$ for $i=0,1,\ldots,m$ (Notice the assumption $A_i \succ 0$ for one i in papers). Note that

$$t \ge \sum_{i=1}^{n} x_i^2 \Longleftrightarrow \left\| \begin{bmatrix} x_1 \\ \vdots \\ x_n \\ (t-1)/2 \end{bmatrix} \right\|_2 \le \frac{t+1}{2} \Longleftrightarrow \begin{bmatrix} x_1 \\ \vdots \\ x_n \\ (t-1)/2 \\ (t+1)/2 \end{bmatrix} \in \mathcal{L}^{n+2}$$

Other Applications - QCQP ⇒ SOCP

Therefore, for each $i = 1, \ldots, m$

$$x^{T} A_{i} x + 2b_{i}^{T} x + c_{i} \leq 0 \iff \begin{bmatrix} A_{i}^{1/2} x \\ -1/2 - b_{i}^{T} x - c_{i}/2 \\ 1/2 - b_{i}^{T} x - c_{i}/2 \end{bmatrix} \in \mathcal{L}^{n+2}$$

QCQP can be equivalently written as

Min
$$u$$

s.t.
$$\begin{bmatrix} A_0^{1/2}x \\ -1/2 - b_0^T x + u/2 - c_0/2 \\ 1/2 - b_0^T x + u/2 - c_0/2 \end{bmatrix} \in \mathcal{L}^{n+2}$$

$$\begin{bmatrix} A_i^{1/2}x \\ -1/2 - b_i^T x - c_i/2 \\ 1/2 - b_i^T x - c_i/2 \end{bmatrix} \in \mathcal{L}^{n+2}, i = 1, \dots, m.$$

Semi-Definite Programming (SDP):

$$K = \mathcal{S}^n_+$$

When $K = \mathcal{S}^n_+ = \{X \in \mathbb{R}^{n \times n} | X = X^T \succeq 0\}$, LCoP becomes SDP.

Min
$$C \bullet X$$

s.t. $A_i \bullet X = b_i, i = 1, ..., m$ (SDP)
 $X \succeq 0$

where $C,A_1,...,A_m$ are given $n\times n$ symmetric matrices and $b_1,...,b_m$ are given scalars, and

$$M \bullet X = \sum_{i,j} M_{ij} X_{ij} = \operatorname{tr}(M^T X).$$

Dual Form

Dual of SDP:

$$Max \quad b^{T}y$$

$$s.t. \quad \sum_{i=1}^{m} y_{i}A_{i} + S = C$$

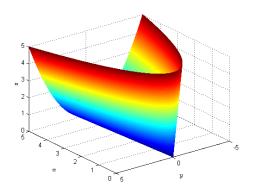
$$S \succeq 0$$
(SDD)

where $y=(y_1,...,y_m)^T$ is a vector in \mathbb{R}^m and S is an $n\times n$ symmetric matrix.

- How to get the above form?
- What are the properties of the semi-definite positive cone?

$$K = \mathcal{S}^n_+$$

$$\mathcal{S}_{+}^{2} = \left\{ (x, y, z) \in \mathbb{R}^{3} | \begin{bmatrix} x & y \\ y & z \end{bmatrix} \succeq 0. \right\} \Longleftrightarrow x \geq 0, z \geq 0, xz \geq y^{2}.$$



21 / 38

THU Conic Programming

Application of SDP

Correlation Matrix Verification

Consider three random variables $A,\ B$ and C. By definition, their correlation coefficients ρ_{AB},ρ_{AC} and ρ_{BC} are valid if and only if

$$\begin{bmatrix} 1 & \rho_{AB} & \rho_{AC} \\ \rho_{AB} & 1 & \rho_{BC} \\ \rho_{AC} & \rho_{BC} & 1 \end{bmatrix} \succeq 0$$

Suppose we know from some prior knowledge (e.g. empirical results of experiments) that $-0.2 \le \rho_{AB} \le -0.1$ and $0.4 \le \rho_{BC} \le 0.5$. What are the smallest and largest values that ρ_{AC} can take?

Covariance Matrix

Suppose X_1, X_2, \dots, X_n be n random variables and k_1, k_2, \dots, k_n be n coefficients. Then the expectation and variance are

$$E(\sum_{i=1}^{n} k_i X_i) = \sum_{i=1}^{n} k_i E X_i,$$

$$Var(\sum_{i=1}^{n} k_i X_i) = E(\sum_{i=1}^{n} k_i X_i - E(\sum_{i=1}^{n} k_i X_i))^2$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} k_i E(X_i - E(X_i))(X_j - E(X_j))k_j$$

$$= (k_1, k_2, \dots, k_n) (E(X_i - E(X_i))(X_j - E(X_j)))_{n \times n} (k_1, k_2, \dots, k_n)^T$$

Define

$$Cov(X) = (E(X_i - E(X_i))(X_j - E(X_j)))_{n \times n}.$$

Then Cov(X) is a semi-definite positive matrix.

Correlation Matrix

Suppose X_1, X_2, \dots, X_n be n random variables. Define

$$DX_i = E(X_i - E(X_i))(X_i - E(X_i)) = E(X_i - EX_i)^2.$$

Define the correlation between X_i and X_j as

$$r_{ij} = \frac{E((X_i - E(X_i))(X_j - E(X_j)))}{\sqrt{DX_i}\sqrt{DX_j}}$$

and the correlation matrix as $Cor(X) = (r_{ij})_{n \times n}$. Then

- $r_{ii} = 1$,
- $-1 \le r_{ij} \le 1$,
- $Cov(X) = diag(\sqrt{DX_1}, \sqrt{DX_2}, \dots, \sqrt{DX_n})Cor(X)diag(\sqrt{DX_1}, \sqrt{DX_2}, \dots, \sqrt{DX_n})$
- Cor(X) is semi-definite positive.

Application of SDP

SDP formulation

The above problem can be formulated as following problem:

$$\begin{aligned} Min/Max & \rho_{AC} \\ s.t. & -0.2 \leq \rho_{AB} \leq -0.1 \\ & 0.4 \leq \rho_{BC} \leq 0.5 \\ & \rho_{AA} = \rho_{BB} = \rho_{CC} = 1 \\ & \begin{bmatrix} \rho_{AA} & \rho_{AB} & \rho_{AC} \\ \rho_{AB} & \rho_{BB} & \rho_{BC} \\ \rho_{AC} & \rho_{BC} & \rho_{CC} \end{bmatrix} \succeq 0 \end{aligned}$$

SDP formulation

In order to formulate the problem as in standard form, we handle the inequality constraints by augmenting the variable matrix and introducing slack variables, for example

$$= \rho_{AB} + s_1 = -0.$$



SDP formulation

$$X = \begin{bmatrix} 1 & \rho_{AB} & \rho_{AC} & 0 & 0 & 0 & 0 \\ \rho_{AB} & 1 & \rho_{BC} & 0 & 0 & 0 & 0 \\ \rho_{AC} & \rho_{BC} & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & s_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & s_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & s_3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & s_4 \end{bmatrix}$$

$$A_2 = ?, A_3 = ? \text{ and } A_4 = ?$$



SDP standard form

$$\begin{aligned} Min/Max & x_{13} \\ s.t. & x_{12} + x_{44} = -0.1 \\ & x_{12} - x_{55} = -0.2 \\ & x_{23} + x_{66} = 0.5 \\ & x_{23} - x_{77} = 0.4 \\ & x_{11} = x_{22} = x_{33} = 1 \\ & x_{ij} = 0, 1 \leq i \leq 3 \land 4 \leq j \leq 7 \\ & x_{ij} = 0, 4 \leq i \leq 7 \land 1 \leq j \leq 3 \\ & x_{ij} = 0, 4 \leq i, j \leq 7 \land i \neq j \\ & (x_{ij}) \in \mathcal{S}_{+}^{7}. \end{aligned}$$

Other Applications - SOCP ⇒ SDP

 \mathcal{L}^{n+1} can be easily embedded into \mathcal{S}_{+}^{n+1} by observing the fact that

$$\left[\begin{array}{c} x \\ t \end{array}\right] \in \mathcal{L}^{n+1} \Longleftrightarrow \left[\begin{array}{cc} t & x^T \\ x & tI_n \end{array}\right] \in \mathcal{S}^{n+1}_+$$

Based on this, we will focus on the theorems and algorithms for SDP. But this does not mean that SOCP is useless or we should transform SOCP to SDP in any case.

Max-Cut Problem

An undirect graph G=(N,E), vertex set $N=\{1,2,\ldots,n\}$, edge set $E=\{(i,j)\mid i,j\in N=\{1,2,\cdots,n\}\}$, weight $w_{ij}\geq 0$ for $(i,j)\in E$. Find a partition S,S' of $N,S\bigcup S'=N,S\bigcap S'=\emptyset$, to maximize the weight over S and S'.

If $i \in S$, let $x_i = 1$, otherwise $x_i = -1$. Define $w_{ij} = 0, (i, j) \notin E$, then the objective function is

$$\frac{1}{2} \left(\sum_{(i,j) \in E} w_{ij} - \sum_{(i,j) \in E} w_{ij} x_i x_j \right)
= \frac{1}{2} \left(\frac{1}{2} \sum_{i,j=1}^n w_{ij} - \frac{1}{2} \sum_{i,j=1}^n w_{ij} x_i x_j \right)
= \frac{1}{4} \sum_{i,j=1}^n w_{ij} (1 - x_i x_j).$$

Max-Cut Problem

$$\max_{\substack{1 \\ \text{s.t.}}} \frac{1}{4} \sum_{i,j=1}^{n} w_{ij} (1 - x_i x_j)$$

s.t.
$$x_i^2 = 1, i = 1, 2, \dots, n$$

$$x \in \mathbb{R}^n.$$

A quadratically constrained quadratic programming

$$v = \max \quad \frac{1}{2}x^{T}Ax$$

s.t.
$$x_{i}^{2} = 1, i = 1, 2, \dots, n,$$
$$x \in \mathbb{R}^{n}$$

where
$$A=rac{\sum_{i,j=1}^n w_{ij}}{2n}I-rac{1}{2}(w_{ij}).$$



Lagrangian function

$$L(x,\lambda) = \frac{1}{2}x^{T}(A+\Lambda)x - \frac{1}{2}\sum_{i=1}^{n}\lambda_{i},$$

where
$$\Lambda = \operatorname{Diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$$
.

$$v_D = \min$$
 σ
s.t. $L(x, \lambda) \le \sigma$, $\forall x \in \{-1, 1\}^n$
 $\sigma \in \mathbb{R}, \lambda \in \mathbb{R}^n$.



Dual of max-cut

$$v_{D} = \min \sigma$$
s.t.
$$\begin{pmatrix} 1 \\ x \end{pmatrix}^{T} \begin{pmatrix} -\sum_{i=1}^{n} \lambda_{i} - 2\sigma & 0 \\ 0 & A + \Lambda \end{pmatrix} \begin{pmatrix} 1 \\ x \end{pmatrix} \leq 0, \forall x \in \{-1, 1\}^{n}$$

$$\sigma \in \mathbb{R}, \lambda \in \mathbb{R}^{n}.$$

$$v_{D} = \min \sigma$$
s.t.
$$U = -\begin{pmatrix} -\sum_{i=1}^{n} \lambda_{i} - 2\sigma & 0 \\ 0 & A + \Lambda \end{pmatrix}$$

$$U \in \mathcal{D}_{\{-1, 1\}^{n}}, \lambda \in \mathbb{R}^{n}, \sigma \in \mathbb{R}.$$

A Hard Cone

$$\mathcal{D}_{\{-1,1\}^n} = \left\{ U \middle| \left(\begin{array}{c} 1 \\ x \end{array} \right)^T U \left(\begin{array}{c} 1 \\ x \end{array} \right) \ge 0, \forall x \in \{-1,1\}^n \right\}.$$

- $\mathcal{D}_{\{-1,1\}^n}$ is a cone.
- How to check a matrix $U \in \mathcal{D}_{\{-1,1\}^n}$ is hard.

Duality Theorems of LP

Theorem (Weak Duality Theorem of LP)

If x is primal feasible and y is dual feasible, then $c^Tx \ge b^Ty$.

Theorem (Strong Duality Theorem of LP)

- If either LP or LD has a finite optimal solution, then so does the other and they achieve the same optimal objective value.
- If either LP or LD has an unbounded objective value, then the other has no feasible solution.

How about the duality theorems of LCoP?

Algorithms for LP

Simplex Method for LP

- Starting from one vertex
- Check whether current vertex is optimal or not. If yes, stop.
 Otherwise, go to the next step.
- Move to a neighbor vertex, go to above step.

The complexity of simplex method is not polynomial.

Polynomial-time Algorithms

- Ellipsoid Method
- Karmarkar's Projective Scaling Algorithm
- Affine Scaling Algorithm: Primal, Dual and Primal-Dual

How about the algorithms for LCoP?



References

Books

- Bertsekas D.P., Nedić A. and Ozdaglar A.E., Convex Analysis and Optimization, Athena Scientific: Belmont, MA USA 2003
- Boyd S. and Vandenberghe L., Convex Optimization, Cambridge University Press: Cambridge, UK 2004
- Fang S. and Puthenpura S., Linear Optimization and Extensions: Theory and Algorithms, Prentice-Hall Inc.: Englewood Cliffs, NJ USA 1993
- Nemirovski A., Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications, Society for Industrial and Applied Mathematics: Philadelphia, PA USA 2001
- Renegar J., A Mathematical View of Interior-point Methods in Convex Optimization, Society for Industrial and Applied Mathematics: Philadelphia, PA USA 2001
- Handbook of Semidefinite Programming: Theory, Algorithms, and Applications, edited by Wolkowicz H., Saigal R. and Vandenberghe L., Kluwer Academic Publisher: Norwell, MA USA 2000

THU

References

- Rockafellar R.T., Convex Analysis, Princeton University Press: Princeton, NJ USA 1970
- Wright. S., Primal-Dual Interior-Point Methods, Society for Industrial and Applied Mathematics: Philadelphia, PA USA 1997

Others

- Xing W. and Fang S.-C., Introduction to Linear Conic Optimization, Tsinghua University Press, 2020.
- Fang S.-C. and Xing W., Linear Conic Optimization, Science Press, China, 2013.
- Ye Y., Linear Conic Programming, lecture notes online: http://www.stanford.edu/class/msande314/sdpmain.pdf

A very popular general purpose SDP solver, CVX can be found in: http://cvxr.com/cvx/

