Part V Applications

Wenxun Xing

Department of Mathematical Scieneces
Tsinghua University
wxing@tsinghua.edu.cn
Tel. 62787945
Office. The New Science Building, A416.

June, 2023

Examples and applications of linear conic programs

Content

- Weber problem
- Matrix optimization
- Approximating solutions of linear equations
- Portfolio management
- Minimum of a univariate polynomial of degree $2n$
- Stochastic queue location problem
- Robust optimization

Weber problem

In 1909, the German economist Alfred Weber introduced the problem of finding a best location for the warehouse of a company, such that the total transportation cost to serve the customers is minimum. Suppose that there are m customers needing to be served. Let the location of customer i be $a^i \in \mathbb{R}^2$, $i = 1, \ldots, m$. Suppose that customer may have different demands, to be translated as weight ω_i for customer i, $i = 1, \ldots, m$. Denote the desired location of the warehouse to be x. Then, the optimization problem is

$$\begin{array}{ll} \min & \sum_{i=1}^{m} \omega_i t_i \\ s.t. & \left[\begin{array}{c} x - a^i \\ t_i \end{array} \right] \in \mathcal{L}^3, i = 1, \ldots, m \end{array}$$

Matrix optimization I

Given $A_0, A_1, \ldots, A_m$, determine if there is $y \in \mathbb{R}^m$ such that

$$A_0 + \sum_{i=1}^m y_i A_i \preceq 0$$

which is equivalent to

$$\lambda_{\max}(A_0 + \sum_{i=1}^m y_i A_i) \quad \leq 0$$

Notice that

$$t \geq \lambda_{\max}(A_0 + \sum_{i=1}^m y_i A_i) \qquad \Leftrightarrow \qquad tI - A_0 - \sum_{i=1}^m y_i A_i \succeq 0$$

Equivalent Problem (SDP)

$$\begin{array}{ll} \min & \quad t \\ s.t. & tI - A_0 - \sum_{i=1}^m y_i A_i \succeq 0 \\ & t \in \mathbb{R}, y \in \mathbb{R}^m \end{array}$$

Yes: If its optimal value is not positive. No: Otherwise.

Matrix optimization II

- Given $A \in \mathcal{S}^n$, determine the maximum eigenvalue of A:

$$\begin{array}{rl} \min & \lambda \\ \text{s.t.} & \lambda I - A \in \mathcal{S}^n_+ \\ & \lambda \in \mathbb{R}. \end{array}$$

- For a variable $X \in \mathcal{M}(m,n)$, find its minimum spectral radius $\min \|X\|_2^2 = \min \lambda_{\max}(XX^T)$:

$$\begin{array}{rl} \min & \lambda \\ \text{s.t.} & \lambda I - XX^T \in \mathcal{S}^m_+ \\ & \lambda \in \mathbb{R}, X \in \mathcal{M}(m,n). \end{array} \quad \Leftrightarrow \quad \begin{array}{rl} \min & \lambda \\ \text{s.t.} & \begin{pmatrix} \lambda I & X \\ X^T & I \end{pmatrix} \in \mathcal{S}^{2m}_+ \\ & \lambda \in \mathbb{R}, X \in \mathcal{M}(m,n). \end{array}$$

Approximating solutions of linear equations

Given $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, solve

$$Ax = b$$

Approximation:

- l_1 norm:

$$\min_x \|Ax - b\|_1 \quad \Leftrightarrow \quad \begin{array}{ll} \min & \sum_{i=1}^m t_i \\ s.t. & -t_i \le A_{i\cdot}x - b_i \le t_i, i = 1, \ldots, m. \end{array}$$

- l_2 norm:

$$\min_x \|Ax - b\|_2 \quad \Leftrightarrow \quad \begin{array}{ll} \min & t \\ s.t. & \begin{bmatrix} Ax - b \\ t \end{bmatrix} \in \mathcal{L}^{m+1} \end{array}$$

Approximating solutions of linear equations

- l_∞ norm:

$$\min_x \|Ax-b\|_\infty \quad \Leftrightarrow \quad \begin{array}{ll} \min & t \\ s.t. & -t \le A_i.x - b_i \le t, i = 1,\dots,m. \end{array}$$

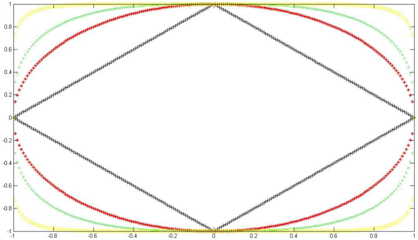
- Logarithm approximation: $(b >_{\mathbb{R}^m_+} 0)$

$$\min_x \max_{1\le i\le m} |\log(A_i.x) - \log b_i| \quad \Leftrightarrow \quad \begin{array}{ll} \min & t \\ s.t. & 1/t \le A_i.x/b_i \le t, i = 1,\dots,m, \\ & t > 0 \end{array}$$

$$\Leftrightarrow \quad \begin{array}{ll} \min & t \\ s.t. & \begin{bmatrix} t - A_i.x/b_i & 0 & 0 \\ 0 & A_i.x/b_i & 1 \\ 0 & 1 & t \end{bmatrix} \succeq 0 \end{array}$$

l_p norms

- l_2 norm. Too convex and smooth.
- l_p norm.



Black: 1-norm. Red: 2-norm. Green: 3-norm. Yellow: 8-norm.

Convex l_p-norm problems

- p-norm domain is convex ($p \geq 1$).
- For set $\{x \mid \|x\|_p \leq 1\}$, the smallest one is the domain with $p = 1$, which is the smallest convex set containing integer points $\{-1, 1\}^n$.
- For $p = 1, p = 2$, the l_p-norm problems with linear objective or linear constraints are polynomially solvable.
- Variants of l_p-norm problems should be considered.
- $0 < p < 1$?

Regularization-Sparsity of decision variables

$$\begin{array}{rl} \min & \|Ax-b\|_2 \\ s.t. & \|x\|_0 \le s \\ & x \in \mathbb{R}^n \end{array}$$

Different convex formulations

$$\begin{array}{rl} \min & \|Ax-b\|_2 \\ s.t. & \|x\|_1 \le s \\ & x \in \mathbb{R}^n \end{array}$$

$$\begin{array}{rl} \min & \|Ax-b\|_2 + \lambda\|x\|_1 \\ s.t. & x \in \mathbb{R}^n \end{array}$$

Research issues–Regularization

- Big data based methods: Newton method, the first-order method. Unconstraint optimization problem.

$$\begin{array}{ll} \min & \|Ax-b\|_2+\lambda\|x\|_1 \\ s.t. & x\in\mathbb{R}^n \end{array}$$

  - $\|x\|_1$ gets to sparsity of x, but it is non-smooth.
  - Sub-gradient methods.
- Is it an exact penalty method? The optimal solution of the primal sparsity problem is recovered from the regularization problem for a finite value of the penalty constant.
- Accelerated methods. The proximal method, augmented Lagrangian method, etc.

One accelerated method: proximal point algorithm

- Proximal point algorithm (PPA). For a given point x^k, the next point is an optimal solution of the following optimal solution.

$$\begin{array}{ll} \min & \|Ax-b\|_2 + \lambda\|x\|_1 + \mu\|x-x^k\|^2 \\ s.t. & x \in \mathbb{R}^n \end{array}$$

- Convergence rate?
- A textbook: Dimitri P. Bertsekas, Convex Optimization Algorithms, 2015.

Research topics

- Low-rank problems

$$\begin{array}{rl} \min & rank(X) \\ s.t. & \|AX - B\|_F \leq \mu \\ & X \in \mathcal{M}(m,n), \end{array}$$

$$\begin{array}{rl} \min & \|AX - B\|_F \\ s.t. & rank(X) \leq r \\ & X \in \mathcal{M}(m,n). \end{array}$$

- Use nuclear norm $\|X\|_* = \sum_{i=1}^r \sigma_i$, where $X = U\mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_r)V^T$, U, V are orthogonal matrixes. (See Ex. 5.24).
- Applications in the fields of the artificial intelligence (AI) and big data.

Research topics

- Manifold optimization problems

$$\begin{aligned} \min \quad & f(X) \\ s.t. \quad & X^T X = I_n \\ & X \in \mathcal{M}(m, n). \end{aligned}$$

- An easy problem

$$\begin{aligned} \min \quad & x^T A x \\ s.t. \quad & x^T x = 1 \\ & x \in \mathbb{M}(n, 1). \end{aligned}$$

- Newton methods and projection?
- Structure methods.

Portfolio management—-I

$$
\begin{array}{rl}
\min & x^T V x \\
\text{s.t.} & b^T x \geq \mu \\
& e^T x = 1 \\
& x \in \mathbb{R}^n_+,
\end{array}
$$

where $e = (1, 1, \ldots, 1)^T$, V correlation matrix.

$$
\begin{array}{rl}
\min & t \\
s.t. & x^T V x \leq t \\
& b^T x \geq \mu \\
& e^T x = 1 \\
& x \in \mathbb{R}^n_+, t \in \mathbb{R}.
\end{array}
$$

Portfolio management—-I

$$
\begin{array}{ll}
\min & t \\
s.t. & \begin{pmatrix} Bx \\ \frac{1-t}{2} \\ \frac{1+t}{2} \end{pmatrix} \in \mathcal{L}^{n+2} \\
 & b^T x \geq \mu \\
 & e^T x = 1 \\
 & x \in \mathbb{R}^n_+, t \in \mathbb{R},
\end{array}
$$

Portfolio management—-II

$$
\begin{aligned}
\max \quad & b^T x \\
\text{s.t.} \quad & x^T V x \le \nu \\
& e^T x = 1 \\
& x \in \mathbb{R}^n_+ .
\end{aligned}
$$

$$
\begin{aligned}
\max \quad & b^T x \\
\text{s.t.} \quad & \begin{pmatrix} Bx \\ \frac{1-\nu}{2} \\ \frac{1+\nu}{2} \end{pmatrix} \in \mathcal{L}^{n+2} \\
& e^T x = 1 \\
& x \in \mathbb{R}^n_+ .
\end{aligned}
$$

Portfolio management—-III

$$
\begin{array}{ll}
\min & \frac{x^T V x}{b^T x} \\
s.t. & b^T x \geq \mu \\
 & e^T x = 1 \\
 & x \in \mathbb{R}^n_+.
\end{array}
$$

$$
\begin{array}{ll}
\min & t \\
\text{s.t.} & \begin{pmatrix} Bx \\ \frac{t-s}{2} \\ \frac{t+s}{2} \end{pmatrix} \in \mathcal{L}^{n+2} \\
 & b^T x - s = 0 \\
 & e^T x = 1 \\
 & s \geq \mu \\
 & x \in \mathbb{R}^n_+, s, t \in \mathbb{R}.
\end{array}
$$

Minimum of a univariate polynomial

Consider the problem of finding the minimum of a univariate polynomial of degree $2n$:

$$\begin{array}{ll} \min & x^{2n} + a_1 x^{2n-1} + \cdots + a_{2n-1}x + a_{2n} \\ s.t. & x \in \mathbb{R} \end{array}$$

This problem is equivalent to

$$\begin{array}{ll} \max & t \\ s.t. & x^{2n} + a_1 x^{2n-1} + \cdots + a_{2n-1}x + a_{2n} - t \geq 0 \text{ for all } x \in \mathbb{R} \end{array}$$

Minimum of a univariate polynomial

It is well known that a univariate polynomial is nonnegative over the real domain if and only if it can be written as *sum of squares* (SOS), which is equivalent to saying that there must be a positive semidefinite matrix $X \in \mathcal{S}_+^{n+1}$ such that

$$x^{2n} + a_1 x^{2n-1} + \cdots + a_{2n-1}x + a_{2n} - t = (1, x, x^2, \cdots, x^n) X (1, x, x^2, \cdots, x^n)^T.$$

Hence, this problem can be cast as an SDP:

$$\begin{array}{ll} \max & t \\ s.t. & a_{2n} - t = X_{11} \\ & a_{2n-k} = \sum\limits_{i+j=k+2} X_{ij}, \ k = 1, \ldots, 2n-1 \\ & X_{(n+1),(n+1)} = 1 \\ & X \in \mathcal{S}_+^{n+1} \end{array}$$

Sum of squares (SOS)

For a given polynomial $p(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$, $p(x) \geq 0$ for any $x \in \mathbb{R}$ if and only if n is even and $p(x) = \sum_{i=1}^{r} q_i^2(x)$ where $q_i(x), i = 1, 2, \ldots, r$ are polynomials.

- The sufficient result is obvious.
- Necessary.
  - n is even.
  - By induction. $n = 0$, $p(x) = a_0 \geq 0$, $\Rightarrow p(x) = (\sqrt{a_0})^2$.
  - Suppose $n = 2k$, $k \geq 0$ be true. For $n = 2(k+1)$, the minimum of $p(x)$ is attainable at x_0 as $p(x)$ is continuous and $p(x) \to +\infty$ when $x \to \pm\infty$.
  - $p(x_0) \geq 0$ and $p(x) - p(x_0) = (x - x_0)^s p_1(x)$ where $p_1(x_0) \neq 0$.
  - If s is odd and $p_1(x_0) > 0$, there exists $\delta > 0$ such that $p(x) - p(x_0) < 0$ for $x_0 - \delta < x < x_0$. If s is odd and $p_1(x_0) < 0$, there exists $\delta > 0$ such that $p(x) - p(x_0) < 0$ for $x_0 < x < x_0 + \delta$. Contradictory.
  - We have s even and then $p_1(x) \geq 0$ for any $x \in \mathbb{R}$. $p_1(x) = \sum_{i=1}^{r} q_i^2(x)$ $\Rightarrow p(x) = (x - x_0)^s \sum_{i=1}^{r} q_i^2(x) + p(x_0) =$ $\sum_{i=1}^{r} [(x - x_0)^{s/2} q_i(x)]^2 + (\sqrt{p(x_0)})^2$.

Minimum of a multivariate polynomial

$$\begin{array}{ll} \min & p(x_1, x_2, \ldots, x_n) = \sum_{0 \le i_1, i_2, \ldots, i_n \le m} a_{i_1 i_2 \cdots i_n} x_1^{i_1} x_1^{i_2} \cdots x_n^{i_n} \\ s.t. & x \in \mathbb{R}^n \end{array}$$

- SOS?
- If $p(x_1, x_2, \ldots, x_n) > 0$ over a restrict domain(Ref. Dejun Chen 2011 and Christian 1987), then it can be formulated as a SOS.
- Approximation formulations.
- Enlarges the problem size: the worst case $(m+1)^n \times (m+1)^n$.

Dejun Chen, Application of polynomial optimization method in quadratic programming problems. Master degree thesis, Tsinghua Unicersity, 2011.

Berg Christian, The multidimension moment problem and semigroups. Moments in Mathematics, Proceedings of Symposia in Applied Mathematics, 37, 110-124, 1987.

Stochastic queue location problem

Background

Suppose there are m potential customers to serve in the region. Customers' demands are random, and once a customer calls for service, then the server in the service center will need to go to the customer to provide the required service. In case the server is occupied, then the customer will have to wait. The goal is to find a good location for the service center in order to minimize the expected waiting time of service.

Stochastic queue location problem

Assumptions and notations

Suppose that the service calls from the customer are identically independent distributed, and the demand process follows the Poisson distribution with overall arrival rate λ, and the probability that any service call is from customer i is assumed to be p_i for $i = 1, \ldots, m$. The queueing principle is *First Come First Service*, and there is only one server in the service center. This model can be regarded as $M/G/1$ queue as in Queueing theory, and the expected service time, including waiting time and traveling, can be explicitly computed. To this end, denote the velocity of the server to be v, and the location of customer i is a^i, $i = 1, \ldots, m$, and the location of the service center to be x.

Stochastic queue location problem

Problem formulation

The expected waiting time for customer i is given by

$$\omega_i(x) = \frac{(2\lambda/v^2)\sum_{j=1}^{m} p_j\|x-a^j\|^2}{1-(2\lambda/v)\sum_{j=1}^{m} p_j\|x-a^j\|^2} + \frac{1}{v}\|x-a^i\|,$$

where the first term in the expected term is the expected waiting time for the server to be free and the second the term is the waiting time for the server to travel after his departure at the service center.

Stochastic queue location problem

Observing the fact that

$$\|x\|_2^2/s \le t,\ s > 0 \Longleftrightarrow \left\| \begin{bmatrix} x \\ \frac{t-s}{2} \end{bmatrix} \right\|_2 \le \frac{t+s}{2}$$

We can formulate this problem as an SOCP:

$$\begin{aligned}
\min \quad & (2m\lambda/v^2) \sum_{i=1}^{m} p_i t_i + (1/v) \sum_{i=1}^{m} t_{0i} \\
s.t. \quad & \begin{bmatrix} x - a^i \\ t_{0i} \end{bmatrix} \in \mathcal{L}^3,\ \begin{bmatrix} x - a^i \\ \frac{t_i - s}{2} \\ \frac{t_i + s}{2} \end{bmatrix} \in \mathcal{L}^4,\ i = 1, ..., m \\
& s \le 1 - (2\lambda/v) \sum_{i=1}^{m} p_i s_i,\ \begin{bmatrix} x - a^i \\ s_i \end{bmatrix} \in \mathcal{L}^3,\ i = 1, ..., m.
\end{aligned}$$

Robust optimization

$$
\begin{aligned}
\min \quad & f_0(x) \\
s.t. \quad & f_i(x, u_i) \leq 0 \\
& \forall u_i \in \mathcal{U}_i, i = 1, \ldots, m.
\end{aligned}
$$

Motivation

- The parameters are inexact;
- The parameters cannot be foreseen;
- The parameters may vary with time and other environment factors;

Example: Robust linear program

$$
\begin{aligned}
\min \quad & c^T x \\
s.t. \quad & a_i^T x \le b_i \\
& \forall a_i \in \{a_i^0 + \textstyle\sum_{j=1}^{p} u_j a_i^j \mid \|u\|_2 \le 1\}, \\
& i = 1, \ldots, m
\end{aligned}
$$

For any x

$$
\begin{aligned}
\max_{a_i} a_i^T x &= \max_u (a_i^0)^T x + \textstyle\sum_{j=1}^{p} u_j (a_i^j)^T x \\
&= (a_i^0)^T x + \|((a_i^1)^T x, \ldots, (a_i^p)^T x)^T\|_2
\end{aligned}
$$

Example: Robust linear program

Therefore

$$a_i^T x \le b_i, \forall a_i \qquad \Leftrightarrow \qquad \|((a_i^1)^T x, \ldots, (a_i^p)^T x)^T\|_2 \le b_i - (a_i^0)^T x$$

Equivalent SOCP

$$\begin{aligned} \min \quad & c^T x \\ s.t. \quad & \begin{bmatrix} (a_i^1)^T x \\ \vdots \\ (a_i^p)^T x \\ b_i - (a_i^0)^T x \end{bmatrix} \in \mathcal{L}^{p+1}, i = 1, \ldots, m \end{aligned}$$

Robust convex QCQP

$$
\begin{aligned}
\min \quad & f^T x \\
s.t. \quad & x^T A^T A x \le 2b^T x + c \\
& \forall (A, b, c) \in \{(A^0, b^0, c^0) + \textstyle\sum_{j=1}^p u_j (A^j, b^j, c^j) \mid \|u\|_2 \le 1\},
\end{aligned}
$$

Let $U(x) = (A^0 x, A^1 x, \ldots, A^p x)$ and

$$
V(x) = \begin{bmatrix}
c^0 + 2(b^0)^T x & \frac{1}{2}c^1 + (b^1)^T x & \cdots & \frac{1}{2}c^p + (b^p)^T x \\
\frac{1}{2}c^1 + (b^1)^T x & 0 & & \\
\vdots & & \ddots & \\
\frac{1}{2}c^p + (b^p)^T x & & & 0
\end{bmatrix}
$$

Robust convex QCQP

$$
\begin{aligned}
& x^T A^T A x \leq 2b^T x + c \\
\Leftrightarrow \quad & \begin{bmatrix} 1 \\ u \end{bmatrix}^T U^T(x) U(x) \begin{bmatrix} 1 \\ u \end{bmatrix} - \begin{bmatrix} 1 \\ u \end{bmatrix}^T V(x) \begin{bmatrix} 1 \\ u \end{bmatrix} \leq 0, \forall \|u\|_2 \leq 1
\end{aligned}
$$

Lemma

Let $P = \begin{pmatrix} 1 & 0 \\ 0 & -I_n \end{pmatrix}$ and $Q = V(x) - U^T(x)U(x)$. Then the implication $\begin{pmatrix} 1 \\ u \end{pmatrix}^T P \begin{pmatrix} 1 \\ u \end{pmatrix} \geq 0 \quad \Rightarrow \quad \begin{pmatrix} 1 \\ u \end{pmatrix}^T Q \begin{pmatrix} 1 \\ u \end{pmatrix} \geq 0$ is valid if and only if $Q - \lambda P \succeq 0$ for some $\lambda \geq 0$.

Robust convex QCQP

From S-lemma,

$$
\begin{aligned}
& \begin{bmatrix} 1 \\ u \end{bmatrix}^T U^T(x)U(x) \begin{bmatrix} 1 \\ u \end{bmatrix} - \begin{bmatrix} 1 \\ u \end{bmatrix}^T V(x) \begin{bmatrix} 1 \\ u \end{bmatrix} \leq 0, \forall \|u\|_2 \leq 1 \\
\Leftrightarrow \quad & V(x) - U^T(x)U(x) - \lambda \begin{bmatrix} 1 & \\ & -I \end{bmatrix} \succeq 0, \exists \lambda \geq 0 \\
\Leftrightarrow \quad & \begin{bmatrix} V(x) - \lambda \begin{bmatrix} 1 & \\ & -I \end{bmatrix} & U^T(x) \\ U(x) & I \end{bmatrix} \succeq 0, \exists \lambda \geq 0 \\
& \text{(by Schur complementary theorem)}
\end{aligned}
$$

Robust convex QCQP

Equivalent SDP

$$
\begin{aligned}
\min \quad & f^T x \\
s.t. \quad & \begin{bmatrix}
c^0 + 2(b^0)^T x - \lambda & \frac{1}{2}c^1 + (b^1)^T x & \cdots & \frac{1}{2}c^p + (b^p)^T x & (A^0 x)^T \\
\frac{1}{2}c^1 + (b^1)^T x & \lambda & & & (A^1 x)^T \\
\vdots & & \ddots & & \vdots \\
\frac{1}{2}c^p + (b^p)^T x & & & \lambda & (A^p x)^T \\
A^0 x & A^1 x & \cdots & A^p x & I
\end{bmatrix} \succeq 0
\end{aligned}
$$

- An application [yu2021]

[yu2021] P. Yu, R. Gao and W. Xing, Maximizing perturbation radii for robust convex quadratically constrained quadratic programs,Vol. 293(1): 50-64, 2021.