MATH: Operations Research **2014-15 First Term**

Handout 6: A Brief Introduction to Conic Linear Programming

Instructor: Junfeng Yang *October 12, 2014*

Contents

6.1 Conic linear programming **6-2**

6.2 Semidefinite programming **6-2**

6.3 Second order cone programming **6-4**

6.4 Connections among LP, SOCP and SDP **6-5**

6.5 Convex optimization is a CLP **6-5**

6.6 General constrained optimization is a CLP **6-6**

6.7 Short summary **6-6**

6.8 Dual problem of CLP and duality **6-7**

6.1 Conic linear programming

Let $\mathcal{X}$ and $\mathcal{Y}$ be, respectively, n and m dimensional real Euclidean spaces. The following problem is called conic linear programming:

$$\begin{aligned} \textbf{CLP}: \quad p^* := \min_{x \in \mathcal{X}} \ & \langle c, x \rangle \\ s.t.\ & \mathcal{A}x = b \\ & x \in K, \end{aligned}$$

where $\mathcal{A} : \mathcal{X} \to \mathcal{Y}$ is a linear mapping, $K \subset \mathcal{X}$ is a cone, $b \in \mathcal{Y}$ and $c \in \mathcal{X}$. Clearly, the objective function is linear, the constraint set is the intersection of an affine set with a cone. All difficulties are hidden in the cone.

Linear programming is a special case of CLP with $\mathcal{X} = R^n$, $\mathcal{Y} = R^m$, $K = R^n_+$ and $\langle x, y \rangle = x^T y$.

6.2 Semidefinite programming

Semidefinite programming is another special case of CLP. Let S^n and S^n_+ be, respectively, the sets of all symmetric and symmetric positive semidefinite matrices of order n, i.e.,

$$\begin{aligned} S^n &:= \{X \in R^{n \times n} : X^T = X\} \\ S^n_+ &:= \{X \in R^{n \times n} : X \succeq 0\}. \end{aligned}$$

The notation "$X \succeq 0$" means that $X \in S^n$ and X is positive semidefinite. Clearly, S^n_+ is a cone. Let S^n be endowed with an inner product

$$\langle X, Y \rangle = \sum_{i,j=1}^{n} X_{ij} Y_{ij} = \mathrm{tr}(X^T Y), \ \ \forall\, X, Y \in S^n.$$

Here "tr" means trace. Given a set of symmetric matrices $\{A_1, \ldots, A_m\} \subset S^n$, the following defines a linear mapping $\mathcal{A}$ from S^n to R^m:

$$\mathcal{A}X := \begin{pmatrix} \langle A_1, X \rangle \\ \vdots \\ \langle A_m, X \rangle \end{pmatrix}, \ \ \forall\, X \in S^n.$$

Let $C \in S^n$ and $b \in R^m$. The following problem is a generalization of LP and is called semidefinite programming:

$$\begin{aligned} \textbf{SDP}: \quad \min_{X \in S^n} \ & \langle C, X \rangle \\ s.t.\ & \mathcal{A}X = b \\ & X \succeq 0. \end{aligned}$$

SDP is a special case of **CLP** with $\mathcal{X} = S^n$, $\mathcal{Y} = R^m$, $K = S^n_+$ and $\langle X, Y \rangle = \mathrm{tr}(X^T Y)$.

Example 6.2.1 (The Max-Cut Problem) *Let $G = (V, E)$ be a graph with $n = |V|$ vertices, $w : E \to R$ be an edge weight function on G, i.e.,*

$$w_{ij} = \begin{cases} w((i,j)), & \textit{if } (i,j) \in E; \\ 0, & \textit{o.w.} \end{cases}$$

Let $W = (w_{ij})_{i,j=1,2,\ldots,n}$ be the matrix whose ij-th entry is w_{ij}. A cut is a partition of V into two sets $S \subset V$ and $V \setminus S$. The size of the cut $(S, V \setminus S)$ is

$$\texttt{size}(S) := \sum_{i \in S,\, j \in V \setminus S} w_{ij}.$$

The max-cut problem aims to find a cut $(S, V \setminus S)$ of a weighted graph (V, E, w) such that $\texttt{size}(S)$ is maximized among all cuts. The max-cut problem is well known to be NP-hard!

Now we give a reformulation of the max-cut problem. For any $S \subset V$, define

$$x = (x_1, \ldots, x_n)^T, \text{ where } x_i = \begin{cases} 1, & \text{if } i \in S; \\ -1, & \text{if } i \notin S. \end{cases}$$

Then the max-cut problem can be formulated as

$$\begin{aligned} \max t &:= \frac{1}{2} \sum_{i<j} w_{ij}(1 - x_i x_j) \\ s.t. \quad & x_i \in \{-1, 1\}, \ \forall\, i \in V. \end{aligned}$$

Let $\mathbf{1} := (1, 1, \ldots, 1)^T \in R^n$ and define $C := \frac{1}{4}(Diag(W\mathbf{1}) - W)$. It can be shown that

$$t = x^T C x.$$

In fact,

$$\begin{aligned} t = \frac{1}{2} \sum_{i<j} w_{ij}(1 - x_i x_j) &= \frac{1}{4} \left(2 \sum_{i<j} w_{ij} x_i^2 - 2 \sum_{i<j} w_{ij} x_i x_j \right) \\ &= \frac{1}{4} \left(\sum_i \sum_j w_{ij} x_i^2 - \sum_i \sum_j w_{ij} x_i x_j \right) \\ &= \frac{1}{4} \left(\sum_i \sum_j w_{ij} x_i^2 - x^T W x \right) \\ &= \frac{1}{4} \left(\sum_i (W\mathbf{1})_i x_i^2 - x^T W x \right) = \frac{1}{4} \left(x^T Diag(W\mathbf{1}) x - x^T W x \right) \\ &= \frac{1}{4} x^T \left(Diag(W\mathbf{1}) - W \right) x = x^T C x \left(= constant - \frac{1}{4} x^T W x \right). \end{aligned}$$

If we define $X = xx^T \in S_+^n$, then $t = x^T C x = tr(x^T C x) = tr(C x x^T) = \langle C, X \rangle$. Also, the constraint $x_i \in \{1, -1\}$ is equivalent to $Diag(X) = \mathbf{1}$.

Lemma 6.1 *$X \in S_+^n$ and $rank(X) = 1 \Longleftrightarrow X = xx^T$ for some nonzero $x \in R^n$.*

According to the above lemma, the max-cup problem can be reexpressed as

$$\begin{aligned} \max \quad & \langle C, X \rangle \\ s.t. \ & Diag(X) = \mathbf{1} \\ & {\color{red} rank(X) = 1} \\ & X \succeq 0. \end{aligned}$$

By considering the following SDP relaxation problem, we can get an upper bound on the optimal value of the max-cut problem:

$$\begin{aligned} \max \quad & \langle C, X\rangle \\ s.t.\ & Diag(X) = \mathbf{1} \\ & X \succeq 0. \end{aligned}$$

Remark 6.2.1 *SDP problems have a particular structure that makes their solution computationally tractable by interior-point methods. Some selected applications among many:*

- *linear matrix inequalities (LMI) arising from systems and control;*
- *engineering applications such as structural optimization and signal processing;*
- *SDP is now widely used in the relaxations of NP-hard combinatorial optimization problems (max cut, max clique, 0-1 integer LP, etc);*
- *polynomial optimization where, under certain mild assumptions, a sequence of increasing order of SDP relaxations can converge to a globally optimal solution;*
- *robust optimization with uncertain data;*

The widespread applications of SDP have stimulated great demands on robust SDP solvers. Now SDP problems of moderate size (roughly, $n \le 5000$) can be solved relatively efficiently on PC. SDP solvers: SDPT3, Sedumi, SDPA, ...

6.3 Second order cone programming

Definition 6.2 (Second order cone) *The second-order cone or Lorentz cone of order n is defined by*

$$Q^n := \{(t, x) \in R \times R^n : \|x\|_2 \le t\}.$$

The following problem often arises in robust optimization:

$$\mathbf{SOCP}: \quad \min_{x \in R^n} \quad c^T x \tag{6.1}$$
$$s.t.\ \|A_i x - b_i\|_2 \le c_i^T x + d_i, \tag{6.2}$$
$$i = 1, 2, \ldots, m, \tag{6.3}$$

where $c \in R^n$, $A_i \in R^{n_i \times n}$, $b_i \in R^{n_i}$, $c_i \in R^n$, $d_i \in R$, $i = 1, 2, \ldots, m$. By introducing auxiliary variables, this problem can be transformed to one with linear objective function, linear equality constraints and second order cone constraints.

In robust linear programming, the following problem is considered:

$$\begin{aligned} \min \quad & c^T x \\ s.t. \quad & a_i^T x \le b_i, \quad \forall\, a_i \in \mathcal{E}_i := \{\bar{a}_i + P_i u : \|u\|_2 \le 1\}, \\ & i = 1, \ldots, m, \end{aligned}$$

where $P_i \succeq 0$. The robust linear constraint can be expressed as

$$\max\{a_i^T x : a_i \in \mathcal{E}_i\} = \bar{a}_i^T x + \|P_i x\|_2 \le b_i,$$

which is evidently a second-order cone constraint. Hence, the robust linear programming can be expressed as the following SOCP

$$\begin{aligned} \min \quad & c^T x \\ s.t.\ & \bar{a}_i^T x + \|P_i x\|_2 \le b_i \\ & i = 1, 2, \ldots, m. \end{aligned}$$

6.4 Connections among LP, SOCP and SDP

Clearly, SOCP reduces to LP by letting $A_i \equiv 0$ and $b_i \equiv 0$, $i = 1, 2, \ldots, m.$
Note that

$$\|x\|_2 \le t \Longleftrightarrow \begin{pmatrix} t & x^T \\ x & tI \end{pmatrix} \succeq 0.$$

Therefore, the SOCP problem (6.1) is equivalent to an SDP:

$$\begin{aligned} \min_{x \in R^n} \quad & c^T x \\ s.t. \quad & \begin{pmatrix} c_i^T x + d_i & (A_i x - b_i)^T \\ A_i x - b_i & (c_i^T x + d_i) I \end{pmatrix} \succeq 0, \\ & i = 1, 2, \ldots, m. \end{aligned}$$

It is also obvious that SDP reduces to LP by restricting X to be a diagonal matrix.

6.5 Convex optimization is a CLP

General convex optimization (CO) minimizes a convex function over a convex set. An important property of CO is that any local optimal solution is also global optimal. We consider CO in the following form

$$\begin{aligned} \mathbf{CO}: \quad \min \quad & f(x) \\ s.t. \quad & c_i(x) \le 0 \\ & i = 1, 2, \ldots, m, \end{aligned}$$

where f and c_i, $i = 1, 2, \ldots, m$, are all convex functions on R^n. Clearly, **CO** is equivalent to

$$\begin{aligned} \min_{x,\alpha} \quad & \alpha \\ s.t. \quad & f(x) - \alpha \le 0 \\ & c_i(x) \le 0 \\ & i = 1, 2, \ldots, m. \end{aligned}$$

The new constraint functions remain convex in (x, α). Thus, it is sufficient to consider CO with linear objective function:

$$\begin{aligned} \mathbf{CO}: \quad \min \quad & \langle c, x \rangle \\ s.t. \quad & c_i(x) \le 0 \\ & i = 1, 2, \ldots, m. \end{aligned}$$

For $i = 1, 2, \ldots, m$, we define $C_i := \{(t, x) \in R \times R^n : t > 0, c_i(x/t) \leq 0\}$. It can be shown that each C_i is a convex cone (not closed in general). By noting that $c_i(x) \leq 0$ is equivalent to $(1, x) \in C_i$, the above CO can be rewritten as a CLP:

$$\begin{array}{rc} \min & \langle (0, c), (t, x) \rangle \\ s.t. & \langle (1, 0), (t, x) \rangle = 1 \\ & (t, x) \in C_1 \cap C_2 \cap \ldots \cap C_m. \end{array}$$

6.6 General constrained optimization is a CLP

General constrained optimization takes the form

$$\min\{f(x) : s.t.\, x \in Q \subset R^n\},$$

where Q is generally described by a set of (nonlinear) equality or inequality constraints. Clearly, the above problem is equivalent to

$$\min\{\alpha : s.t.\, x \in Q, f(x) \leq \alpha\},$$

Thus, it is sufficient to consider general constrained optimization with linear objective function

$$\min\{\langle c, x \rangle : s.t.\, x \in Q\}.$$

It is clear that

$$\begin{aligned} x \in Q &\Longleftrightarrow (x, 1) \in \{(y, 1) : y \in Q\} \subset R^{n+1} \\ &\Longleftrightarrow (x, 1) \in \{(ty, t) : y \in Q, t = 1\} \\ &\Longleftrightarrow (x, t) \in K \cap \{(y, t) : y \in R^n, t = 1\}, \end{aligned}$$

where $K := \{(ty, t) : y \in Q, t \geq 0\}$ is a cone. Therefore, the general constrained problem

$$\min\{\langle c, x \rangle : s.t.\, x \in Q\}$$

can be rewritten as

$$\begin{array}{rl} \min & \langle (c, 0), (x, t) \rangle \\ s.t. & \langle (x, t), (0, 1) \rangle = 1 \\ & (x, t) \in K. \end{array}$$

6.7 Short summary

The standard form of conic linear programming is

$$\textbf{CLP}: \quad p^* := \min_{x \in \mathcal{X}} \ \langle c, x \rangle \quad s.t.\ \mathcal{A}x = b, \quad x \in K.$$

CLP is a very general model for mathematical programming. It consists linear programming, semidefinite programming, second order cone programming as special cases. General convex optimization, or even nonlinear optimization problems, can be reformulated as CLP. As a result, the study on algorithms for CLP is usually case by case. The cone K could be cross product of several types of cones. Many others cones exist, e.g., matrix norm cone. All difficulties of CLP are hidden in the cone. Different properties of K can be exploited, e.g., symmetry, homogeneity, convexity, etc.

6.8 Dual problem of CLP and duality

Definition 6.3 (Dual cone) *Let K be a cone in $\mathcal{X}$. The set $K^* := \{y \in \mathcal{X} : \langle y, x\rangle \geq 0, \quad \forall x \in K\}$ is a cone and is called the dual cone of K.*

Theorem 6.4 *Let K be any cone. The dual cone K^* is convex.*

Definition 6.5 (Self-dual cone) *A cone K is called self-dual if $K^* = K$.*

Homework. Prove that the nonnegative orthant R^n_+, the SDP cone S^n_+ and the second order cone $Q^n = \{(t, x) : \|x\|_2 \leq t\}$ are all self-dual.

Define $\mathcal{F} := \{x \in \mathcal{X} : \mathcal{A}x = b, x \in K\}$, and $\mathcal{L} : \mathcal{X} \times R^m \times \mathcal{X}$ by

$$\mathcal{L}(x, y, \mu) := \langle c, x\rangle - y^T(\mathcal{A}x - b) - \langle \mu, x\rangle.$$

For any $x \in \mathcal{F}$, $y \in R^m$ and $\mu \in K^*$, it holds that

$$\mathcal{L}(x, y, \mu) \leq \langle c, x\rangle.$$

Therefore, for any $y \in R^m$ and $\mu \in K^*$, it holds that

$$\inf_{x\in\mathcal{X}} \mathcal{L}(x, y, \mu) \leq \inf_{x\in\mathcal{F}} \mathcal{L}(x, y, \mu) \leq \inf_{x\in\mathcal{F}} \langle c, x\rangle = p^*,$$

i.e., $\inf_{x\in\mathcal{X}} \mathcal{L}(x, y, \mu)$ is a lower bound of p^*. The best lower bound of p^* that can be so obtained is

$$\begin{aligned} & \sup_{y\in R^m, \mu\in K^*} \inf_{x\in\mathcal{X}} \mathcal{L}(x, y, \mu) \\ = & \sup_{y\in R^m, \mu\in K^*} \begin{cases} b^T y, & \text{if } c - \mathcal{A}^* y - \mu = 0; \\ -\infty, & \text{o.w.} \end{cases} \\ = & \sup_{y\in R^m} \{b^T y : c - \mathcal{A}^* y \in K^*\}. \end{aligned}$$

As a result, we call the following conic linear programming the dual problem of **CLP**:

$$\text{Dual-CLP}: \quad d^* := \max_{y\in R^m} \{b^T y : \ s.t.\ c - \mathcal{A}^* y \in K^*\}.$$

Theorem 6.6 (Weak duality) *Let p^* and d^* be, respectively, defined for* **CLP** *and* **Dual-CLP**. *Then, $d^* \leq p^*$.*

Weak duality implies that a feasible solution to either problem yields a bound on the optimal value of the other problem. We call $\langle c, x\rangle - b^T y$ the duality gap.

Corollary 6.7 *Let x and y be, respectively, feasible for* **CLP** *and* **Dual-CLP**. *If $\langle c, x\rangle = b^T y$, then x and y are optimal for respective problems.*

A big question is that whether the reverse is also true. That is the strong duality: given x optimal for **CLP**, is there a y that is feasible for **Dual-CLP** and also satisfies $\langle c, x\rangle = b^T y$?

Let $(K, K^*) = (R^n_+, R^n_+)$. According to Farkas lemma, the following systems are alternative:

$$\{x \in R^n : Ax = b, x \in K\} \quad v.s. \quad \{y \in R^m : -A^T y \in K^*, b^T y > 0\}.$$

$$\{x \in R^n : Ax = 0, x \in K, \langle c, x\rangle < 0\} \quad v.s. \quad \{y \in R^m : c - A^T y \in K^*\}.$$

For general closed convex cone K and its dual cone K^*, are the following systems still alternative

$$\{x \in \mathcal{X} : \mathcal{A}x = b, x \in K\} \quad v.s. \quad \{y \in R^m : -\mathcal{A}^* y \in K^*, b^T y > 0\}?$$

$$\{x \in \mathcal{X} : \mathcal{A}x = 0, x \in K, \langle c, x\rangle < 0\} \quad v.s. \quad \{y \in R^m : c - \mathcal{A}^* y \in K^*\}?$$

Here $\mathcal{A}^*$ represents the adjoint operator of $\mathcal{A}$, i.e., $\mathcal{A}^*$ satisfies

$$\langle \mathcal{A}x, y\rangle = \langle x, \mathcal{A}^* y\rangle, \ \forall x \in \mathcal{X}, y \in \mathcal{Y}.$$

Let K be a **general closed convex cone** and K^* be its dual cone. Define

$$\begin{aligned} S_1 &:= \{x \in \mathcal{X} : \mathcal{A}x = b, x \in K\} \\ S_2 &:= \{y \in R^m : -\mathcal{A}^* y \in K^*, b^T y > 0\}; \\ T_1 &:= \{x \in \mathcal{X} : \mathcal{A}x = 0, x \in K, \langle c, x\rangle < 0\} \\ T_2 &:= \{y \in R^m : c - \mathcal{A}^* y \in K^*\}. \end{aligned}$$

Then we have

$$\begin{aligned} &S_1 \neq \emptyset \Longrightarrow S_2 = \emptyset \quad \text{but} \quad S_1 \neq \emptyset \not\Leftarrow S_2 = \emptyset; \\ &T_1 \neq \emptyset \Longrightarrow T_2 = \emptyset \quad \text{but} \quad T_1 \neq \emptyset \not\Leftarrow T_2 = \emptyset, \end{aligned}$$

which implies that S_1 and S_2 cannot be simultaneously nonempty but can be simultaneously empty, and similarly for T_1 and T_2. This is because $\{\mathcal{A}x : x \in K\}$ and $\{\mathcal{A}^* y + s : y \in R^m, s \in K^*\}$ are, though convex, not necessarily closed.

Example 6.8.1 *Let $\varepsilon > 0$ and*

$$A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \ A_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \ b = \begin{pmatrix} 0 \\ 2\varepsilon \end{pmatrix},$$

A linear mapping $\mathcal{A} : S^2 \to R^2$ is defined by

$$\mathcal{A}X = \begin{pmatrix} \langle A_1, X\rangle \\ \langle A_2, X\rangle \end{pmatrix} = \begin{pmatrix} x_{11} \\ 2x_{12} \end{pmatrix}.$$

Note that $\mathcal{A}^* : R^2 \to S^2$ is given by

$$\mathcal{A}^* y = y_1 A_1 + y_2 A_2.$$

It can be shown that (1) for any $X \succeq 0$, $\mathcal{A}X \neq b$, i.e., $S_1 := \{X \in S^2 : \mathcal{A}X = b, X \succeq 0\} = \emptyset$; and (2) $S_2 := \{y \in R^2 : -\mathcal{A}^* y \succeq 0, b^T y > 0\} = \emptyset$. In this example, the set $\{\mathcal{A}X : X \succeq 0\}$ is not closed, and b is on the boundary of this set. Therefore, b cannot be strictly separated from this set. This is why strong duality can fail for general conic linear programming.

Example 6.8.2 (SDP example with duality gap)

$$A_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, A_2 = \begin{pmatrix} 0 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}, C = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, b = \begin{pmatrix} 0 \\ 2 \end{pmatrix}.$$

It can be shown that $-2 = d^* < p^* = 0$.

Theorem 6.8 (Farkas Lemma for general closed convex cone) *Let K be a* **general closed convex cone** *and K^* be its dual cone. Define*

$$\begin{aligned}
S_1 &:= \{x \in \mathcal{X} : \mathcal{A}x = b, x \in K\} \\
S_2 &:= \{y \in R^m : -\mathcal{A}^* y \in K^*, b^T y > 0\}; \\
T_1 &:= \{x \in \mathcal{X} : \mathcal{A}x = 0, x \in K, \langle c, x\rangle < 0\} \\
T_2 &:= \{y \in R^m : c - \mathcal{A}^* y \in K^*\}.
\end{aligned}$$

Then, we have

1. *$S_1 \neq \emptyset \implies S_2 = \emptyset$, $T_1 \neq \emptyset \implies T_2 = \emptyset$;*
2. *If there exists $y \in R^m$ such that $-\mathcal{A}^* y \in int(K^*)$ (the interior of K^*), then $S_1 \neq \emptyset \;\Leftarrow\; S_2 = \emptyset$ (or equivalently, $S_1 = \emptyset \implies S_2 \neq \emptyset$);*
3. *If there exists $x \in int(K)$ such that $\mathcal{A}x = 0$, then $T_1 \neq \emptyset \Leftarrow T_2 = \emptyset$ (or equivalently, $T_1 = \emptyset \implies T_2 \neq \emptyset$).*

Theorem 6.9 (Strong duality) *Let $\mathcal{F}_P$ and $\mathcal{F}_D$ be the feasible sets of* **CLP** *and* **Dual-CLP**, *respectively. Assume that both $\mathcal{F}_P$ and $\mathcal{F}_D$ are nonempty and at least one of them has an interior, i.e.,*

$$\exists x \in int(K) \text{ such that } \mathcal{A}x = b$$

or

$$\exists y \in R^m \text{ such that } c - \mathcal{A}^* y \in int(K^*).$$

Then, $x \in \mathcal{F}_P$ is optimal for **CLP** *and $y \in \mathcal{F}_D$ is optimal for* **Dual-CLP** *if any only if $\langle c, x\rangle = b^T y$.*

Theorem 6.10

1. *If one of* **CLP** *or* **Dual-CLP** *is unbounded, then the other has no feasible solution.*
2. *If* **CLP** *and* **Dual-CLP** *are both feasible, then both have bounded optimal objective values and the optimal objective values may have a duality gap.*
3. *If one of* **CLP** *and* **Dual-CLP** *has a strictly or interior feasible solution and it has an optimal solution, then the other is feasible and has an optimal solution with the same optimal value.*

Theorem 6.11 (Optimality system for CLP) *Assume strong duality holds. Then, x and y are optimal for respective problems if and only if there exists $s \in K^*$ such that the following conditions are satisfied:*

$$\begin{aligned}
\textbf{Primal feasibility}: &\quad \mathcal{A}x = b, \quad x \in K \\
\textbf{Dual feasibility}: &\quad \mathcal{A}^* y + s = c, \quad s \in K^* \\
\textbf{Complementarity}: &\quad \langle c, x\rangle - b^T y = 0 \text{ (or, } \langle s, x\rangle = 0).
\end{aligned}$$