Chapter 6
Conic Linear Programming

6.1 Convex Cones

Conic Linear Programming, hereafter CLP, is a natural extension of Linear programming (LP). In LP, the variables form a vector which is required to be componentwise nonnegative, while in CLP they are points in a pointed convex cone (see Appendix B.1) of an Euclidean space, such as vectors as well as matrices of finite dimensions. For example, Semidefinite programming (SDP) is a kind of CLP, where the variable points are symmetric matrices constrained to be positive semidefinite. Both types of problems may have linear equality constraints as well. Although CLPs have long been known to be convex optimization problems, no efficient solution algorithm was known until about two decades ago, when it was discovered that interior-point algorithms for LP discussed in Chap. 5, can be adapted to solve certain CLPs with both theoretical and practical efficiency. During the same period, it was discovered that CLP, especially SDP, is representative of a wide assortment of applications, including combinatorial optimization, statistical computation, robust optimization, Euclidean distance geometry, quantum computing, optimal control, etc. CLP is now widely recognized as a powerful mathematical computation model of general importance.

First, we illustrate several convex cones popularly used in conic linear optimization.

Example 1. The followings are all (closed) convex cones.

- The n-dimensional non-negative orthant, $E^n_+ = \{\mathbf{x} \in E^n : \mathbf{x} \geq \mathbf{0}\}$, is a convex cone.
- The set of all n-dimensional symmetric positive semidefinite matrices, denoted by $\mathcal{S}^n_+$, is a convex cone, called the *positive semidefinite matrix cone*. When $\mathbf{X}$ is positive semidefinite (positive definite), we often write the property as $\mathbf{X} \succeq (\succ)\mathbf{0}$.

D.G. Luenberger, Y. Ye, *Linear and Nonlinear Programming*, International Series in Operations Research & Management Science 228,
DOI 10.1007/978-3-319-18842-3_6

- The set $\{(u;\mathbf{x}) \in E^{n+1} \, : \, u \geq |\mathbf{x}|_p\}$ is a convex cone in E^{n+1}, called the *p-order cone* where $1 \leq p < \infty$. When $p = 2$, the cone is called second-order cone or "Ice-cream" cone.

Sometimes, we use the notion of conic inequalities $\mathbf{P} \succeq_K \mathbf{Q}$ or $\mathbf{Q} \preceq_K \mathbf{P}$, in which cases we simply mean $\mathbf{P} - \mathbf{Q} \in K$.

Suppose $\mathbf{A}$ and $\mathbf{B}$ are $k \times n$ matrices. We define the inner product

$$\mathbf{A} \bullet \mathbf{B} = \text{trace}(\mathbf{A}^T\mathbf{B}) = \sum_{i,j} a_{ij}b_{ij}.$$

When $k = 1$, they become n-dimensional vectors and the inner product is the standard dot product of two vectors. In SDP, this definition is almost always used for the case where the matrices are both square and symmetric. The matrix norm associated with the inner product is called *Frobenius norm*:

$$|\mathbf{X}|_f = \sqrt{\mathbf{X} \bullet \mathbf{X}}\,.$$

For a cone K, the dual of K is the cone

$$K^* := \{\mathbf{Y} : \; \mathbf{X} \bullet \mathbf{Y} \geq 0 \quad \text{for all} \quad \mathbf{X} \in K\}.$$

It is not difficult to see that the dual cones of the first two cones in Example 1 are all them self, respectively; while the dual cone of the p-order cone is the q-order cone where

$$\frac{1}{p} + \frac{1}{q} = 1.$$

One can see that when $p = 2$, $q = 2$ as well; that is, they are both 2-order cones. For a closed convex cone K, the dual of the dual cone is itself.

6.2 Conic Linear Programming Problem

Now let $\mathbf{C}$ and $\mathbf{A}_i$, $i = 1, 2, \ldots, m$, be given matrices of $E^{k\times n}$, $\mathbf{b} \in E^m$, and $\mathcal{K}$ be a closed convex cone in $E^{k\times n}$. And let $\mathbf{X}$ be an unknown matrix of $E^{k\times n}$. Then, the standard form (primal) conic linear programming problem is

$$\begin{aligned} \text{(CLP)} \quad & \text{minimize } \mathbf{C} \bullet \mathbf{X} \\ & \text{subject to } \mathbf{A}_i \bullet \mathbf{X} = b_i, \; i = 1, 2, \ldots, m, \; \mathbf{X} \in K. \end{aligned} \tag{6.1}$$

Note that in CLP we minimize a linear function of the decision matrix constrained in cone K and subject to linear equality constraints.

For convenience, we define an operator from a symmetric matrix to a vector:

$$\mathcal{A}\mathbf{X} = \begin{pmatrix} \mathbf{A}_1 \bullet \mathbf{X} \\ \mathbf{A}_2 \bullet \mathbf{X} \\ \cdots \\ \mathbf{A}_m \bullet \mathbf{X} \end{pmatrix}. \tag{6.2}$$

Then, CLP can be written in a compact form:

$$\begin{array}{ll}\text{(CLP)} & \text{minimize } \mathbf{C} \bullet \mathbf{X} \\ & \text{subject to } \mathcal{A}\mathbf{X} = \mathbf{b},\ \mathbf{X} \in K.\end{array}$$

When cone K is the non-negative orthant E^n_+, CLP reduces to linear programming (LP) in the standard form, where $\mathcal{A}$ becomes the constraint matrix $\mathbf{A}$. When K is the positive semidefinite cone $\mathcal{S}^n_+$, CLP is called semidefinite programming (SDP); and when K is the p-order cone, it is called p-order cone programming. In particular, when $p = 2$, the model is called second-order cone programming (SOCP). Frequently, we write variable $\mathbf{X}$ in (CLP) as $\mathbf{x}$ if it is indeed a vector, such as when K is the nonnegative orthant or p-order cone.

One can see that the problem (SDP) (that is, (6.1) with the semidefinite cone) generalizes classical linear programming in standard form:

$$\begin{array}{l}\text{minimize } \mathbf{c}^T\mathbf{x}, \\ \text{subject to } \mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \geq \mathbf{0}.\end{array}$$

Define $\mathbf{C} = \text{Diag}[c_1,\ c_2,\ \dots,\ c_n]$, and let $\mathbf{A}_i = \text{Diag}[a_{i1},\ a_{i2},\ \dots,\ a_{in}]$ for $i = 1, 2,\ \dots m$. The unknown is the $n \times n$ symmetric matrix $\mathbf{X}$ which is constrained by $\mathbf{X} \succeq \mathbf{0}$. Since the trace of $\mathbf{C} \bullet \mathbf{X}$ and $\mathbf{A}_i \bullet \mathbf{X}$ depend only on the diagonal elements of $\mathbf{X}$, we may restrict the solutions $\mathbf{X}$ to diagonal matrices. It follows that in this case the SDP problem is equivalent to a linear program, since a diagonal matrix is positive semidefinite is and only if its all diagonal elements are nonnegative.

One can further see the role of cones in the following examples.

Example 1. Consider the following optimization problems with three variables.

- This is a linear programming problem in standard form:

$$\begin{array}{ll}\text{minimize} & 2x_1 + x_2 + x_3 \\ \text{subject to} & x_1 + x_2 + x_3 = 1, \\ & (x_1; x_2; x_3) \geq \mathbf{0}.\end{array}$$

- This is a semidefinite programming problem where the dimension of the matrix is two:

$$\begin{array}{ll}\text{minimize} & 2x_1 + x_2 + x_3 \\ \text{subject to} & x_1 + x_2 + x_3 = 1, \\ & \begin{bmatrix} x_1 & x_2 \\ x_2 & x_3 \end{bmatrix} \succeq \mathbf{0},\end{array}$$

Let

$$\mathbf{C} = \begin{bmatrix} 2 & .5 \\ .5 & 1 \end{bmatrix} \quad \text{and} \quad \mathbf{A}_1 = \begin{bmatrix} 1 & .5 \\ .5 & 1 \end{bmatrix}.$$

Then, the problem can be written in a standard SDP form

$$\begin{array}{ll} \text{minimize} & \mathbf{C} \bullet \mathbf{X} \\ \text{subject to} & \mathbf{A}_1 \bullet \mathbf{X} = 1, \; \mathbf{X} \in \mathcal{S}^2_+. \end{array}$$

- This is a second-order cone programming problem:

$$\begin{array}{ll} \text{minimize} & 2x_1 + x_2 + x_3 \\ \text{subject to} & x_1 + x_2 + x_3 = 1, \\ & \sqrt{x_2^2 + x_3^2} \leq x_1. \end{array}$$

We present several application examples to illustrate the flexibility of this formulation.

Example 2 (Binary Quadratic Optimization). Consider a binary quadratic maximization problem

$$\begin{array}{ll} \text{maximize} & \mathbf{x}^T\mathbf{Q}\mathbf{x} + 2\mathbf{c}^T\mathbf{x} \\ \text{subject to} & x_j = \{1, \; -1\}, \text{ for all } j = 1, \; \ldots, \; n, \end{array}$$

which is a difficult nonconvex optimization problem. The problem can be rewritten as

$$\begin{array}{rl} z^* \equiv \text{maximize} & \begin{bmatrix} \mathbf{x} \\ 1 \end{bmatrix}^T \begin{bmatrix} \mathbf{Q} & \mathbf{c} \\ \mathbf{c}^T & 0 \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ 1 \end{bmatrix} \\ \text{subject to} & (x_j)^2 = 1, \text{ for all } j = 1, \; \ldots, \; n, \end{array}$$

which can be also written as a homogeneous quadratic binary problem

$$\begin{array}{rl} z^* \equiv \text{maximize} & \begin{bmatrix} \mathbf{Q} & \mathbf{c} \\ \mathbf{c}^T & \mathbf{0} \end{bmatrix} \bullet \begin{bmatrix} \mathbf{x} \\ x_{n+1} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ x_{n+1} \end{bmatrix}^T \\ \text{subject to} & \mathbf{I}_j \bullet \begin{bmatrix} \mathbf{x} \\ x_{n+1} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ x_{n+1} \end{bmatrix}^T = 1, \text{ for all } j = 1, \; \ldots, \; n+1, \end{array}$$

where $\mathbf{I}_j$ is the $(n + 1) \times (n + 1)$ matrix whose components are all zero except at the jth position on the main diagonal where it is 1. Let $(\mathbf{x}^*; \; x^*_{n+1})$ be an optimal solution for the homogeneous problem. Then, one can see that $\mathbf{x}^*/x^*_{n+1}$ would be an optimal solution to the original problem.

Since $\begin{bmatrix} \mathbf{x} \\ x_{n+1} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ x_{n+1} \end{bmatrix}^T$ forms a positive-semidefinite matrix (with rank equal to 1), a *semidefinite relaxation* of the problem is defined as

$$
\begin{aligned}
z^{SDP} \equiv \text{maximize} & \begin{bmatrix} \mathbf{Q} & \mathbf{c} \\ \mathbf{c}^T & 0 \end{bmatrix} \bullet \mathbf{Y} \\
\text{subject to} & \ \mathbf{I}_j \bullet \mathbf{Y} = 1, \text{ for all } j = 1, \ldots, n+1, \\
& \ \mathbf{Y} \in \mathcal{S}^{n+1}_+,
\end{aligned}
\tag{6.3}
$$

where the symmetric matrix $\mathbf{Y}$ has dimension $n + 1$. Obviously, z^{SDP} is a upper bound of z^*, since the rank-1 requirement is not enforced in the relaxation.

Let's see how to use the relaxation. For simplicity, assuming $z^{SDP} > 0$, it has been shown that in many cases of this problem an optimal SDP solution either constitutes an exact solution or can be rounded to a good approximate solution of the original problem. In the former case, one can show that a rank-1 optimal solution matrix $\mathbf{Y}$ exists for the semidefinite relaxation and it can be found by using a rank-reduction procedure. For the latter case, one can, using a randomized rank-reduction procedure or the principle components of $\mathbf{Y}$, find a rank-1 feasible solution matrix $\hat{\mathbf{Y}}$ such that

$$
\begin{bmatrix} \mathbf{Q} & \mathbf{c} \\ \mathbf{c}^T & \mathbf{0} \end{bmatrix} \bullet \hat{\mathbf{Y}} \geq \alpha \cdot Z^{SDP} \geq \alpha \cdot Z^*
$$

for a provable factor $0 < \alpha \leq 1$. Thus, one can find a feasible solution to the original problem whose objective value is no less than a factor α of the true maximal objective cost.

Example 3 (Sensor Localization). This problem is that of determining the location of sensors (for example, several cell phones scattered in a building) when measurements of some of their separation Euclidean distances can be determined, but their specific locations are not known. In general, suppose there are n unknown points $\mathbf{x}_j \in E^d$, $j = 1, \ldots, n$. We consider an edge to be a path between two points, say, i and j. There is a known subset N_e of pairs (edges) ij for which the separation distance d_{ij} is known. For example, this distance might be determined by the signal strength or delay time between the points. Typically, in the cell phone example, N_e contains those edges whose lengths are small so that there is a strong radio signal. Then, the localization problem is to find locations $\mathbf{x}_j$, $j = 1, \ldots, n$, such that

$$
|\mathbf{x}_i - \mathbf{x}_j|^2 = (d_{ij})^2, \text{ for all } (i, j) \in N_e,
$$

subject to possible rotation and translation. (If the locations of some of the sensors are known, these may be sufficient to determine the rotation and translation as well.)

Let $\mathbf{X} = [\mathbf{x}_1 \ \mathbf{x}_2 \ \ldots \ \mathbf{x}_n]$ be the $d \times n$ matrix to be determined. Then

$$
|\mathbf{x}_i - \mathbf{x}_j|^2 = (\mathbf{e}_i - \mathbf{e}_j)^T \mathbf{X}^T \mathbf{X} (\mathbf{e}_i - \mathbf{e}_j),
$$

where $\mathbf{e}_i \in E^n$ is the vector with 1 at the i**th** position and zero everywhere else. Let $\mathbf{Y} = \mathbf{X}^T \mathbf{X}$. Then the semidefinite relaxation of the localization problem is to find $\mathbf{Y}$ such that

$$(\mathbf{e}_i - \mathbf{e}_j)(\mathbf{e}_i - \mathbf{e}_j)^T \bullet \mathbf{Y} = (d_{ij})^2, \text{ for all } (i,\ j) \in N_e,$$
$$\mathbf{Y} \succeq \mathbf{0}.$$

This problem is one of finding a feasible solution; the objective function is null. But if the distance measurements have noise, one can add additional variables and an error objective to minimize. For example,

$$\begin{array}{ll} \text{minimize} & \sum_{(i,j)\in N_e} |z_{ij}| \\ \text{subject to} & (\mathbf{e}_i - \mathbf{e}_j)(\mathbf{e}_i - \mathbf{e}_j)^T \bullet \mathbf{Y} + z_{ij} = (d_{ij})^2, \text{ for all } (i,\ j) \in N_e, \\ & \mathbf{Y} \succeq \mathbf{0}. \end{array}$$

This problem can be converted into a conic linear program with mixed nonnegative orthant and semidefinite cones.

Under certain graph structure, an optimal SDP solution $\mathbf{Y}$ of the formulation would be guaranteed rank-d so that it constitutes an exact solution of the original problem. Also, in general $\mathbf{Y}$ can be rounded to a good approximate solution of the original problem. For example, one can, using a randomized rank-reduction procedure or the d principle components of $\mathbf{Y}$, find a rank-d solution matrix $\hat{\mathbf{Y}}$.

6.3 Farkas' Lemma for Conic Linear Programming

We first introduce the notion of "interior" of cones.

Definition 1. We call $\mathbf{X}$ an interior point of cone K if and only if, for any point $\mathbf{Y} \in K^*$, $\mathbf{Y} \bullet \mathbf{X} = 0$ implies $\mathbf{Y} = \mathbf{0}$.

The set of interior points of K is denoted by $\overset{\circ}{K}$.

Theorem 1. *The interior of the followings convex cones are given as:*

- *The interior of the non-negative orthant cone is the set of all vectors where every entry is positive.*
- *The interior of the positive semidefinite cone is the set of all positive definite matrices.*
- *The interior of p-order cone is the set of* $\{(u; \mathbf{x}) \in E^{n+1} :\ u > |\mathbf{x}|_p\}$.

We give a sketch of the proof for the second order cone, i.e., $p = 2$. Let $(\bar{u}; \bar{\mathbf{x}}) \neq \mathbf{0}$ be any second-order cone point but $\bar{u} = |\bar{\mathbf{x}}|$. Then, we can choose a dual cone (also the second-order cone) point $(v; \mathbf{y})$ such that

$$v = \alpha\bar{u},\ \mathbf{y} = -\alpha\bar{\mathbf{x}},$$

for a positive α. Note that

$$(\bar{u}; \bar{\mathbf{x}}) \bullet (v; \mathbf{y}) = \alpha\bar{v}^2 - \alpha|\bar{\mathbf{x}}|^2 = 0.$$

Then, one can let $\alpha \to \infty$ so that $(v; \mathbf{y})$ cannot be bounded.

Now let $(\bar{u}; \bar{\mathbf{x}})$ be any given second-order cone point with $\bar{u} > |\bar{\mathbf{x}}|$. We like to prove that, for any dual cone (also the second-order cone) point $(v; \mathbf{y})$,

$$(\bar{u}; \bar{\mathbf{x}}) \bullet (v; \mathbf{y}) = 0$$

implies that $(v; \mathbf{y})$ is bounded. Note that

$$0 = (\bar{u}; \bar{\mathbf{x}}) \bullet (v; \mathbf{y}) = \bar{u}v + \bar{\mathbf{x}} \bullet \mathbf{y}$$

or

$$\bar{u}v \leq -\bar{\mathbf{x}} \bullet \mathbf{y} \leq |\bar{\mathbf{x}}||\mathbf{y}|.$$

If $v = 0$, we must have $\mathbf{y} = \mathbf{0}$; otherwise,

$$\bar{u} \leq |\bar{\mathbf{x}}||\mathbf{y}|/v \leq |\mathbf{x}|,$$

which contradicts $\bar{u} > |\bar{\mathbf{x}}|$.

We leave the proof of the following proposition as an exercise.

Proposition 1. *Let* $\mathbf{X} \in \overset{\circ}{K}$ *and* $\mathbf{Y} \in K^*$. *Then For any nonnegative constant* κ, $Y \bullet X \leq \kappa$ *implies that* $\mathbf{Y}$ *is bounded.*

Let us now consider the feasible region of (CLP) (6.1):

$$\mathcal{F} := \{\mathbf{X} : \ \mathcal{A}\mathbf{X} = \mathbf{b}, \ \mathbf{X} \in K\};$$

where the interior of the feasible region is

$$\overset{\circ}{\mathcal{F}} := \{\mathbf{X} : \ \mathcal{A}\mathbf{X} = \mathbf{b}, \ \mathbf{X} \in \overset{\circ}{K}\}.$$

If $\mathcal{F}$ is empty with $K = E^n_+$, from Farkas' lemma for linear programming, a vector $\mathbf{y} \in E^m$, with $\mathbf{y}^T A \leq \mathbf{0}$ and $\mathbf{y}^T\mathbf{b} > 0$, always exists and is called an infeasibility certificate for the system $\{\mathbf{x} : \ A\mathbf{x} = \mathbf{b}, \ \mathbf{x} \geq \mathbf{0}\}$.

Does this alternative relations hold for K being a general closed convex one? Let us rigorousize the question. Let us define the reverse operator of (6.2) from a vector to a matrix:

$$\mathbf{y}^T\mathcal{A} = \sum_{i=1}^{m} \mathbf{A}_i y_i. \tag{6.4}$$

Note that, by the definition, for any matrix $\mathbf{X} \in E^{k \times n}$

$$\mathbf{y}^T\mathcal{A} \bullet \mathbf{X} = \mathbf{y}^T(\mathcal{A}\mathbf{X}),$$

that is, the association property holds. Also, $(\mathbf{y}^T\mathcal{A})^T = \mathcal{A}^T\mathbf{y}$, that is, the transpose operation applies here as well.

Then, the question becomes: when $\mathcal{F}$ is empty, does there exist a vector $\mathbf{y} \in E^m$ such that $-\mathbf{y}^T\mathcal{A} \in K^*$ and $\mathbf{y}^T\mathbf{b} > 0$? Similarly, one can ask: when set

$\{\mathbf{y}: \mathbf{C}^T - \mathbf{y}^T\mathcal{A} \in K\}$ is empty, does there exist a matrix $\mathbf{X} \in K^*$ such that $\mathcal{A}\mathbf{X} = \mathbf{0}$ and $\mathbf{C} \bullet \mathbf{X} < 0$? Note that the answer to the second question is also "yes" when $K = E^n_+$.

Example 1. The answer to either question is "not necessarily"; see example below.

- For the first question, consider $K = \mathcal{S}^2_+$ and

$$\mathbf{A}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \mathbf{A}_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

  and

$$\mathbf{b} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$$

- For the second question, consider $K = \mathcal{S}^2_+$ and

$$C = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \text{and} \quad A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

However, if the data set $\mathcal{A}$ satisfies additional conditions, the answer would be "yes"; see theorem below.

Theorem 2 (Farkas' Lemma for CLP). *We have*

- *Consider set*

$$\mathcal{F}_p := \{\mathbf{X}: \mathcal{A}\mathbf{X} = \mathbf{b},\ \mathbf{X} \in K\}.$$

  Suppose that there exists a vector $\overset{\circ}{\mathbf{y}}$ *such that* $-\overset{\circ}{\mathbf{y}}^T \mathcal{A} \in \overset{\circ}{K^*}$. *Then,*

  1. *Set* $C := \{\mathcal{A}\mathbf{X} \in E^m : \mathbf{X} \in K\}$ *is a closed convex set;*
  2. $\mathcal{F}_p$ *has a (feasible) solution if and only if set* $\{\mathbf{y} : -\mathbf{y}^T\mathcal{A} \in K^*,\ \mathbf{y}^T\mathbf{b} > 0\}$ *has no feasible solution.*

- *Consider set*

$$\mathcal{F}_d := \{\mathbf{y}: \mathbf{C}^T - \mathbf{y}^T\mathcal{A} \in K\}.$$

  Suppose that there exists a vector $\overset{\circ}{\mathbf{X}} \in \overset{\circ}{K^*}$ *such that* $\mathcal{A}\overset{\circ}{\mathbf{X}} = \mathbf{0}$. *Then,*

  1. *Set* $C := \{\mathbf{S} - \mathbf{y}^T\mathcal{A} : \mathbf{S} \in K\}$ *is a closed convex set;*
  2. $\mathcal{F}_d$ *has a (feasible) solution if and only if set* $\{\mathbf{X}: \mathcal{A}\mathbf{X} = \mathbf{0},\ \mathbf{X} \in K^*,\ \mathbf{C} \bullet \mathbf{X} < 0\}$ *has no feasible solution.*

Proof. We prove the first statement of the theorem. We prove the first part. It is clear that C is a convex set. To prove that C is a closed set, we need to show that if $\mathbf{y}^k := \mathcal{A}\mathbf{X}^k \in E^m$ for $\mathbf{X}^k \in K$, $k = 1, \ldots$, converges to a vector $\bar{\mathbf{y}}$, then $\bar{\mathbf{y}} \in C$ or there is $\bar{\mathbf{X}} \in K$ such that $\bar{\mathbf{y}} := \mathcal{A}\bar{\mathbf{X}}$. Without loss of generality, we assume that $\mathbf{y}^k$ is a bounded sequence. Then, we have, for a positive constant c,

$$c \geq -(\overset{\circ}{\mathbf{y}})^T\mathbf{y}^k = -(\overset{\circ}{\mathbf{y}})^T(\mathcal{A}\mathbf{X}^k) = -(\overset{\circ}{\mathbf{y}})^T\mathcal{A} \bullet \mathbf{X}^k, \forall k.$$

Since $-(\overset{\circ}{\mathbf{y}})^T\mathcal{A} \in \overset{\circ}{K^*}$, by definition, the sequence of $\mathbf{X}^k$ is also bounded. Then there is at least an accumulate point $\bar{\mathbf{X}} \in K$ because K is a closed cone. Thus, we must have $\bar{\mathbf{y}} := \mathcal{A}\bar{\mathbf{X}}$.

We now prove the second part. If $\mathcal{F}_p$ has a feasible solution $\bar{\mathbf{X}}$. Then, let $\mathbf{y}$ make $-\mathbf{y}^T\mathcal{A} \in K^*$

$$-\mathbf{y}^T\mathbf{b} = -\mathbf{y}^T(\mathcal{A}\bar{\mathbf{X}}) = -\mathbf{y}^T\mathcal{A} \bullet \bar{\mathbf{X}} \geq 0.$$

Thus, it must be true $\mathbf{y}^T\mathbf{b} \leq 0$, that is, $\{\mathbf{y} : -\mathbf{y}^T\mathcal{A} \in K^*, \ \mathbf{y}^T\mathbf{b} > 0\}$ must be empty.

On the other hand, let $\mathcal{F}_p$ has no feasible solution, or equivalently, $\mathbf{b} \notin C$. We now show that $\{\mathbf{y} : -\mathbf{y}^T\mathcal{A} \in K^*, \ \mathbf{y}^T\mathbf{b} > 0\}$ must be nonempty.

Since C is a closed convex set, from the separating hyperplane theorem, there must exist a $\bar{\mathbf{y}} \in E^m$ such that

$$\bar{\mathbf{y}}^T\mathbf{b} > \bar{\mathbf{y}}^T\mathbf{y}, \ \forall \mathbf{y} \in C,$$

or, from $\mathbf{y} = \mathcal{A}\mathbf{X}, \ \mathbf{X} \in K$, we have

$$\bar{\mathbf{y}}^T\mathbf{b} > \bar{\mathbf{y}}^T(\mathcal{A}\mathbf{X}) = \bar{\mathbf{y}}^T\mathcal{A} \bullet \mathbf{X}, \ \forall \mathbf{X} \in K.$$

That is, $\bar{\mathbf{y}}^T\mathcal{A} \bullet \mathbf{X}$ is bounded above for all $\mathbf{X} \in K$.

Immediately, we see $\bar{\mathbf{y}}^T\mathbf{b} > 0$ since $\mathbf{0} \in K$. Next, it must be true $-\bar{\mathbf{y}}^T\mathcal{A} \in K^*$. Otherwise, we must be able to find an $\bar{\mathbf{X}} \in K$ such that $-\bar{\mathbf{y}}^T\mathcal{A} \bullet \bar{\mathbf{X}} < 0$ by the definition of K and its dual K^*. For any positive constant α we maintain $\alpha\bar{\mathbf{X}} \in K$ and let α go to ∞. Then, $\bar{\mathbf{y}}^T\mathcal{A} \bullet (\alpha\bar{\mathbf{X}})$ goes to ∞, contradicting the fact that $\bar{\mathbf{y}}^T\mathcal{A} \bullet \mathbf{X}$ is bounded above for all $\mathbf{X} \in K$. Thus, $\bar{\mathbf{y}}$ is a feasible solution in $\{\mathbf{y} : -\mathbf{y}^T\mathcal{A} \in K^*, \ \mathbf{y}^T\mathbf{b} > 0\}$. ∎

Note that C may not be a closed set if the interior condition of Theorem 2 is not met. Consider $\mathbf{A}_1$, $\mathbf{A}_2$ and $\mathbf{b}$ in Example 1, and we have

$$C = \left\{\mathcal{A}\mathbf{X} = \begin{bmatrix} \mathbf{A}_1 \bullet \mathbf{X} \\ \mathbf{A}_2 \bullet \mathbf{X} \end{bmatrix} : \ \mathbf{X} \in \mathcal{S}^2_+ \right\}.$$

Let

$$\mathbf{X}^k = \begin{bmatrix} \frac{1}{k} & 1 \\ 1 & k \end{bmatrix} \in \mathcal{S}^2_+, \ \forall k = 1, \ldots.$$

Then we see

$$\mathbf{y}^k = \mathcal{A}\mathbf{X}^k = \begin{bmatrix} \frac{1}{k} \\ 2 \end{bmatrix}.$$

As $k \to \infty$ we see $\mathbf{y}^k$ converges $\mathbf{b}$, but $\mathbf{b}$ is *not* in C.

6.4 Conic Linear Programming Duality

Because conic linear programming is an extension of classical linear programming, it would seem that there is a natural dual to the primal problem, and that this dual is itself a conic linear program. This is indeed the case, and it is related to the primal in much the same way as primal and dual linear programs are related. Furthermore, the primal and dual together lead to the formation a primal-dual solution method, which is discussed later in this chapter.

The dual of the (primal) CLP (6.1) is

$$\begin{array}{ll}\text{(CLD)} & \text{maximize } \mathbf{y}^T\mathbf{b} \\ & \text{subject to } \sum_i^m y_i\mathbf{A}_i + \mathbf{S} = \mathbf{C}^T,\ \mathbf{S} \in K^*.\end{array} \tag{6.5}$$

On written in a compact form:

$$\begin{array}{ll}\text{(CLD)} & \text{maximize } \mathbf{y}^T\mathbf{b} \\ & \text{subject to } \mathbf{y}^T\mathcal{A} + \mathbf{S} = \mathbf{C}^T,\ \mathbf{S} \in K^*.\end{array}$$

Notice that $\mathbf{S}$ represents a slack matrix, and hence the problem can alternatively be expressed as

$$\begin{array}{l}\text{maximize } \mathbf{y}^T\mathbf{b} \\ \text{subject to } \sum_i^m y_i\mathbf{A}_i \preceq_{K^*} \mathbf{C}^T.\end{array} \tag{6.6}$$

Recall that conic inequality $\mathbf{Q} \preceq_K \mathbf{P}$ means $\mathbf{P} - \mathbf{Q} \in K$.

Again, just like linear programming, the dual of (CLD) will be (CLP), and they form a primal and dual pair. Whichever is the primal, then the other will be the dual. We would see more primal and dual relations later.

Example 1. Here are dual problems to the three instances in Example 1 where $\mathbf{y}$ is just a scalar.

- The dual to the linear programming instance:

$$\begin{array}{ll}\text{maximize} & y \\ \text{subject to} & y(1,\ 1,\ 1) + (s_1,\ s_2,\ s_3) = (2,\ 1,\ 1), \\ & \mathbf{s} = (s_1,\ s_2,\ s_3) \in K^* = E^3_+.\end{array}$$

- The dual to semidefinite programming instance:

$$\begin{array}{ll}\text{maximize} & y \\ \text{subject to} & y\mathbf{A}_1 + \mathbf{S} = \mathbf{C}, \\ & \mathbf{S} \in K^* = \mathcal{S}^2_+,\end{array}$$

where recall

$$\mathbf{C} = \begin{bmatrix} 2 & .5 \\ .5 & 1 \end{bmatrix} \quad \text{and} \quad \mathbf{A}_1 = \begin{bmatrix} 1 & .5 \\ .5 & 1 \end{bmatrix}.$$

- The dual to the second-order cone instance:

$$\begin{array}{ll} \text{maximize} & y \\ \text{subject to} & y(1,\ 1,\ 1) + (s_1,\ s_2,\ s_3) = (2,\ 1,\ 1), \\ & \sqrt{s_2^2 + s_3^2} \le s_1,\ \text{or } \mathbf{s} = (s_1,\ s_2,\ s_3) \text{ in second-order cone.} \end{array}$$

Let us consider a couple of more dual examples of the problems we posted earlier.

Example 2 (The Dual of Binary Quadratic Maximization). Consider the semidefinite relaxation (6.3) for the binary quadratic maximization problem. It's dual is

$$\begin{array}{ll} \text{minimize} & \sum_{j=1}^{n+1} y_j \\ \text{subject to} & \sum_{j=1}^{n+1} y_j \mathbf{I}_j - \mathbf{S} = \begin{bmatrix} \mathbf{Q} & \mathbf{c} \\ \mathbf{c}^T & 0 \end{bmatrix},\ \mathbf{S} \succeq \mathbf{0}. \end{array}$$

Note that

$$\sum_{j=1}^{n+1} y_j \mathbf{I}_j - \begin{bmatrix} \mathbf{Q} & \mathbf{c} \\ \mathbf{c}^T & 0 \end{bmatrix}$$

is exactly the Hessian matrix of the Lagrange function of the quadratic maximization problem; see Chap. 11. Therefore, there is a close connection between the Lagrange and conic dualities. The problems is to find a diagonal matrix $\text{Diag}[(y_1; \ldots; y_{n+1})]$ such that the Lagrange Hessian is positive semidefinite and its sum of diagonal elements is minimized.

Example 3 (The Dual of Sensor Localization). Consider the semidefinite programming relaxation for the sensor localization problem (with no noises). It's dual is

$$\begin{array}{ll} \text{maximize} & \sum_{(i,j) \in N_e} y_{ij} \\ \text{subject to} & \sum_{(i,j) \in N_e} y_{ij} (\mathbf{e}_i - \mathbf{e}_j)(\mathbf{e}_i - \mathbf{e}_j)^T + \mathbf{S} = \mathbf{0},\ \mathbf{S} \succeq \mathbf{0}. \end{array}$$

Here, y_{ij} represents an internal force or tension on edge $(i,\ j)$. Obviously, $y_{ij} = 0$ for all $(i,\ j) \in N_e$ is a feasible solution for the dual. However, finding non-trivial internal forces is a fundamental problem in network and structure design, and the maximization of the dual would help to achieve the goal.

Many optimization problems can be directly cast in the CLD form.

Example 4 (Euclidean Facility Location). This problem is to determine the location of a facility serving n clients placed in a Euclidean space, whose known locations are denoted by $\mathbf{a}_j \in E^d,\ j = 1,\ \ldots,\ n$. The location of the facility would minimize the sum of the Euclidean distances from the facility to each of the clients. Let the location decision be vector $\mathbf{f} \in E^d$. Then the problem is

$$\text{minimize } \sum_{j=1}^{n} |\mathbf{f} - \mathbf{a}_j| .$$

The problem can be reformulated as

$$\begin{array}{ll} \text{minimize} & \sum_{j=1}^{n} \delta_j \\ \text{subject to} & \mathbf{s}_j + \mathbf{f} = \mathbf{a}_j, \ \forall j = 1, \ldots, n, \\ & |\mathbf{s}_j| \le \delta_j, \quad \forall j = 1, \ldots, n. \end{array}$$

This is a conic formulation in the (CLD) form. To see it clearly, let $d = 2$ and $n = 3$ in the example, and let

$$\mathbf{A}^T = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 \end{bmatrix} \in E^{9\times5}, \ \mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \in E^5, \ \mathbf{c} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \mathbf{a}_1 \\ \mathbf{a}_2 \\ \mathbf{a}_3 \end{bmatrix} \in E^9,$$

and variable vector

$$\mathbf{y} = [\delta_1; \ \delta_2; \ \delta_3; \ \mathbf{f}] \in E^5.$$

Then, the facility location problem becomes

$$\begin{array}{ll} \text{minimize} & \mathbf{y}^T\mathbf{b} \\ \text{subject to} & \mathbf{y}^T\mathbf{A} + \mathbf{s}^T = \mathbf{c}^T, \ \mathbf{s} \in K; \end{array}$$

where K is the product of three second-order cones each of which has dimension 3. More precisely, the first three elements of $\mathbf{s} \in E^9$ are in the 3-dimensional second-order cone; and so are the second three elements and the third three elements of $\mathbf{s}$. In general, the product of (possibly mixed) cones, say K_1, K_2 and K_3, is denoted by $K_1 \oplus K_2 \oplus K_3$, and $\mathbf{X} \in K_1 \oplus K_2 \oplus K_3$ means that $\mathbf{X}$ is divided into three components such that

$$\mathbf{X} = (\mathbf{X}_1; \ \mathbf{X}_2; \ \mathbf{X}_3), \ \text{where } \mathbf{X}_1 \in K_1, \ \mathbf{X}_2 \in K_2, \ \text{and } \mathbf{X}_3 \in K_3.$$

The dual of the facility location problem would be in the (CLP) form:

$$\begin{array}{ll} \text{minimize} & \mathbf{c}^T\mathbf{x} \\ \text{subject to} & \mathbf{A}\mathbf{x} = \mathbf{b}, \ \mathbf{x} \in K^*; \end{array}$$

where

$$K^* = (K_1 \oplus K_2 \oplus K_3)^* = K_1^* \oplus K_2^* \oplus K_3^*.$$

That is, in this particular problem, the first three elements of $\mathbf{x} \in E^9$ are in the 3-dimensional second-order cone; and so are the second three elements and the third three elements of $\mathbf{x}$.

Consider further the equality constraints, the dual can be simplified as

$$\begin{array}{lll} \text{maximize} & \sum_{j=1}^{3} \mathbf{a}_j^T \mathbf{x}_j & \\ \text{subject to} & \sum_{j=1}^{3} \mathbf{x}_j & = \mathbf{0} \in E^2, \\ & |\mathbf{x}_j| & \leq 1, \ \forall\ j = 1, 2, 3. \end{array}$$

Example 5 (Quadratic Constraints). Quadratic constraints can be transformed to linear semidefinite form by using the concept of *Schur complements*. Let $\mathbf{A}$ be a (symmetric) m-dimension positive definite matrix, $\mathbf{C}$ be a symmetric n-dimension matrix, and $\mathbf{B}$ be an $m \times n$ matrix. Then, matrix

$$\mathbf{S} = \mathbf{C} - \mathbf{B}^T \mathbf{A}^{-1} \mathbf{B}$$

is called the Schur complement of $\mathbf{A}$ in the matrix

$$\mathbf{Z} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^T & \mathbf{C} \end{bmatrix}.$$

Moreover, $\mathbf{Z}$ is positive semidefinite if and only if $\mathbf{S}$ is positive semidefinite.

Now consider a general quadratic constraint of the form

$$\mathbf{y}^T \mathbf{B}^T \mathbf{B} \mathbf{y} - \mathbf{c}^T \mathbf{y} - d \leq 0. \tag{6.7}$$

This is equivalent to

$$\begin{bmatrix} \mathbf{I} & \mathbf{B}\mathbf{y} \\ \mathbf{y}^T \mathbf{B}^T & \mathbf{c}^T \mathbf{y} + d \end{bmatrix} \succeq \mathbf{0} \tag{6.8}$$

because the Schur complement of this matrix with respect to $\mathbf{I}$ is the negative of the left side of the original constraint (6.7). Note that in this larger matrix, the variable $\mathbf{y}$ appears only affinely, not quadratically.

Indeed, (6.8) can be written as

$$\mathbf{P}(\mathbf{y}) = \mathbf{P}_0 + y_1 \mathbf{P}_1 + y_2 \mathbf{P}_2 + \cdots y_n \mathbf{P}_n \succeq \mathbf{0}, \tag{6.9}$$

where

$$\mathbf{P}_0 = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & d \end{bmatrix}, \ \mathbf{P}_i = \begin{bmatrix} \mathbf{0} & \mathbf{b}_i \\ \mathbf{b}_i^T & c_i \end{bmatrix} \text{ for } i = 1, 2, \ \ldots n$$

with $\mathbf{b}_i$ being the i**th** column of $\mathbf{B}$ and c_i being the ith component of $\mathbf{c}$. The constraint (6.9) is of the form that appears in the dual form of a semidefinite program.

There is a more efficient mixed semidefinite and second-order cone formulation of the inequality (6.7) to reduce the dimension of semidefinite cone. We first introduce slack variable $\mathbf{s}$ and s_0 by linear constraints:

$$\mathbf{B}\mathbf{y} - \mathbf{s} = \mathbf{0}$$

Then, we let $|\mathbf{s}| \le s_0$ (or $(s_0; \mathbf{s})$ in the second-order cone) and

$$\begin{bmatrix} 1 & s_0 \\ s_0 & \mathbf{c}^T\mathbf{y} + d \end{bmatrix} \succeq \mathbf{0}.$$

Again, the matrix constraint is of the dual form of a semidefinite cone, but its dimension is fixed at 2.

Suppose the original optimization problem has a quadratic objective: minimize $q(\mathbf{x})$. The objective can be written instead as: minimize t subject to $q(\mathbf{x}) \le t$, and then this constraint as well as any number of other quadratic constraints can be transformed to semidefinite constraints, and hence the entire problem converted to a mixed second-order cone and semidefinite program. This approach is useful in many applications, especially in various problems of financial engineering and control theory.

The duality is manifested by the relation between the optimal values of the primal and dual programs. The weak form of this relation is spelled out in the following lemma, the proof of which, like the weak form of other duality relations we have studied, is essentially an accounting issue.

Weak Duality in CLP. *Let* $\mathbf{X}$ *be feasible for (CLP) and* $(\mathbf{y}, \mathbf{S})$ *feasible for (CLD). Then,*

$$\mathbf{C} \bullet \mathbf{X} \ge \mathbf{y}^T\mathbf{b}.$$

Proof. By direct calculation

$$\begin{aligned} \mathbf{C} \bullet \mathbf{X} - \mathbf{y}^T\mathbf{b} &= \left(\sum_{i=1}^m y_i\mathbf{A}_i + \mathbf{S}\right) \bullet \mathbf{X} - \mathbf{y}^T\mathbf{b} \\ &= \sum_{i=1}^m y_i(\mathbf{A}_i \bullet \mathbf{X}) + \mathbf{S} \bullet \mathbf{X} - \mathbf{y}^T\mathbf{b} \\ &= \sum_{i=1}^m y_i b_i + \mathbf{S} \bullet \mathbf{X} - \mathbf{y}^T\mathbf{b} \\ &= \mathbf{S} \bullet \mathbf{X} \ge 0, \end{aligned}$$

where the last inequality comes from $\mathbf{X} \in K$ and $\mathbf{S} \in K^*$. ∎

As in other instances of duality, the strong duality of conic linear programming is weak unless other conditions hold. For example, the duality gap may not be zero at optimality in the following SDP instance.

Example 6. The following semidefinite program has a duality gap:

$$\mathbf{C} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \ \mathbf{A}_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \ \mathbf{A}_2 = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

and

$$\mathbf{b} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}.$$

The primal minimal objective value is 0 achieved by

$$\mathbf{X} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

and the dual maximal objective value is -2 achieved by $\mathbf{y} = [0, \ -1]$; so the duality gap is 2.

However, under certain technical conditions, there would be no duality gap. One condition is related to weather or not the primal feasible region $\mathcal{F}_p$ or dual feasible region has an interior feasible solution. We say $\mathcal{F}_p$ has an interior (feasible solution) if and only if

$$\overset{\circ}{\mathcal{F}}_p := \{\mathbf{X} : \ \mathcal{A}\mathbf{X} = \mathbf{b}, \ \mathbf{X} \in \overset{\circ}{K}\}$$

is non-empty, and $\mathcal{F}_d$ has an interior feasible solution if and only if

$$\overset{\circ}{\mathcal{F}}_d := \{(\mathbf{y}, \mathbf{S}) : \ \mathbf{y}^T \mathcal{A} + \mathbf{S} = \mathbf{C}, \ \mathbf{S} \in \overset{\circ}{K}{}^*\}$$

is non-empty. We state here a version of the strong duality theorem.

Strong Duality in (CLP).

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 between (CLP) and (CLD).*

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

Proof. We let cone $H = K \oplus E^1_+$ in the following proof.

i) Suppose $\mathcal{F}_d$ is empty and $\mathcal{F}_p$ is feasible and has an interior feasible solution. Then, we have an $\bar{\mathbf{X}} \in \overset{\circ}{K}$ and $\bar{\tau} = 1$ that is an interior feasible solution to (homogeneous) conic system:

$$\mathcal{A}\bar{\mathbf{X}} - \mathbf{b}\bar{\tau} = \mathbf{0}, \ (\bar{\mathbf{X}}, \bar{\tau}) \in \overset{\circ}{H}.$$

Now, for any z^*, we form an alternative system pair based on Farkas' Lemma (Theorem 2):

$$\{(\mathbf{X}, \tau) : \ \mathcal{A}\mathbf{X} - \mathbf{b}\tau = \mathbf{0}, \ \mathbf{C} \bullet \mathbf{X} - z^*\tau < 0, \ (\mathbf{X}, \tau) \in H\},$$

and

$$\{(\mathbf{y}; \mathbf{S}, \kappa) : \ \mathcal{A}^T\mathbf{y} + \mathbf{S} = \mathbf{C}, \ -\mathbf{b}^T\mathbf{y} + \kappa = -z^*, \ (\mathbf{S}, \kappa) \in H^*\}.$$

But the latter is infeasible, so that the former has a feasible solution $(\mathbf{X}, \tau)$. At such a feasible solution, if $\tau > 0$, we have $\mathbf{C} \bullet (\mathbf{X}/\tau) < z^*$ for any z^*. Otherwise, $\tau = 0$ implies that a new solution $\bar{\mathbf{X}} + \alpha\mathbf{X}$ is feasible for (CLP) for any positive α; and, as $\alpha \to \infty$, the objective value of the new solution goes to $-\infty$. Hence, either way we have a feasible solution for (CLP) whose objective value is unbounded from below.

ii) Let $\mathcal{F}_p$ be feasible and have an interior feasible solution, and let z^* be its objective infimum. Again, we have an alternative system pair as listed in the proof of i). But now the former is infeasible, so that we have a solution for the latter. From the Weak Duality theorem $\mathbf{b}^T\mathbf{y} \le z^*$, thus we must have $\kappa = 0$, that is, we have a solution $(\mathbf{y}, \mathbf{S})$ such that

$$\mathcal{A}^T\mathbf{y} + \mathbf{S} = \mathbf{C}, \ \mathbf{b}^T\mathbf{y} = z^*, \ \mathbf{S} \in K^*.$$

iii) We only need to prove that there exist a solution $\mathbf{X} \in \mathcal{F}_p$ such that $\mathbf{C} \bullet \mathbf{X} = z^*$, that is, the infimum of (CLP) is attainable. But this is just the other side of the proof given that $\mathcal{F}_d$ is feasible and has an interior feasible solution, and z^* is also the supremum of (CLD). ∎

Again, if one of (CLP) and (CLD) has no interior feasible solution, the common objective value may not be attainable. For example,

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

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.

Most of these examples that make the strong duality failed are superficial, and a small perturbation would overcome the failure. Thus, in real applications and in the rest of the chapter, we may assume that both (CLP) and (CLD) have interior when they are feasible. Consequently, any primal and dual optimal solution pair must satisfy the optimality conditions:

$$\begin{array}{r} \mathbf{C} \bullet \mathbf{X} - \mathbf{y}^T\mathbf{b} = 0 \\ \mathcal{A}\mathbf{X} = \mathbf{b} \\ \mathbf{y}^T\mathcal{A} + \mathbf{S} = \mathbf{C}^T \\ \mathbf{X} \in K, \quad \mathbf{S} \in K^* \end{array} ; \tag{6.10}$$

or

$$\begin{aligned} \mathbf{X} \bullet \mathbf{S} &= 0 \\ \mathcal{A}\mathbf{X} &= \mathbf{b} \\ \mathbf{y}^T \mathcal{A} + \mathbf{S} &= \mathbf{C}^T \\ \mathbf{X} \in K, \quad & \mathbf{S} \in K^* \end{aligned} \quad . \tag{6.11}$$

We now present an application of the strong duality theorem.

Example 7 (Robust Portfolio Design). The Markowitz portfolio design model (also see 5) is

$$\begin{aligned} &\text{minimize} \;\; \mathbf{x}^T \boldsymbol{\Sigma} \mathbf{x} \\ &\text{subject to} \;\; \mathbf{1}^T \mathbf{x} = 1, \; \boldsymbol{\pi}^T \mathbf{x} \ge \pi, \end{aligned}$$

where $\boldsymbol{\Sigma}$ is the covariance matrix and $\boldsymbol{\pi}$ is the expect return rate vector of a set of stocks, and π is the desired return rate of the portfolio. The problem can be equivalently written as a mixed conic problem

$$\begin{aligned} &\text{minimize} \; \boldsymbol{\Sigma} \bullet \mathbf{X} \\ &\text{subject to} \; \mathbf{1}^T \mathbf{x} = 1, \; \boldsymbol{\pi}^T \mathbf{x} \ge \pi, \\ &\qquad\qquad \mathbf{X} - \mathbf{x}\mathbf{x}^T \succeq \mathbf{0}. \end{aligned}$$

Now suppose $\boldsymbol{\Sigma}$ is incomplete and/or uncertain, and it is expressed by

$$\boldsymbol{\Sigma}_0 + \sum_{i=1}^{m} y_i \boldsymbol{\Sigma}_i (\succeq \mathbf{0}),$$

for some variables y_i's. Then, we like to solve a robust model

$$\begin{aligned} &\text{minimize} \left\{ \begin{array}{l} \max_{\mathbf{y}} \left(\boldsymbol{\Sigma}_0 + \sum_{i=1}^m y_i \boldsymbol{\Sigma}_i \right) \bullet \mathbf{X} \\ \quad \text{s.t.} \; \boldsymbol{\Sigma}_0 + \sum_{i=1}^m y_i \boldsymbol{\Sigma}_i \succeq \mathbf{0} \end{array} \right\} \\ &\text{subject to} \; \mathbf{1}^T \mathbf{x} = 1, \; \boldsymbol{\pi}^T \mathbf{x} \ge \pi, \\ &\qquad\qquad \mathbf{X} - \mathbf{x}\mathbf{x}^T \succeq \mathbf{0}. \end{aligned}$$

The inner problem is an SDP problem. Assuming strong duality holds, we replace it by its dual, and have

$$\begin{aligned} &\text{minimize} \left\{ \begin{array}{l} \min_{\mathbf{Y}} \boldsymbol{\Sigma}_0 \bullet (\mathbf{Y} + \mathbf{X}) \\ \quad \text{s.t.} \; \boldsymbol{\Sigma}_i \bullet (\mathbf{Y} + \mathbf{X}) = 0, \; \forall i = 1, \ldots, m, \\ \qquad \mathbf{Y} \succeq \mathbf{0} \end{array} \right\} \\ &\text{subject to} \; \mathbf{1}^T \mathbf{x} = 1, \; \boldsymbol{\pi}^T \mathbf{x} \ge \pi, \\ &\qquad\qquad \mathbf{X} - \mathbf{x}\mathbf{x}^T \succeq \mathbf{0}. \end{aligned}$$

Then, we can integrate the two minimization problems together and form

$$\begin{aligned} &\text{minimize } \boldsymbol{\Sigma}_0 \bullet (\mathbf{Y} + \mathbf{X}) \\ &\text{subject to } \mathbf{1}^T\mathbf{x} = 1,\ \boldsymbol{\pi}^T\mathbf{x} \geq \pi, \\ &\qquad \boldsymbol{\Sigma}_i \bullet (\mathbf{Y} + \mathbf{X}) = 0,\ \forall i = 1, \ldots, m, \\ &\qquad \mathbf{Y} \succeq \mathbf{0},\ \mathbf{X} - \mathbf{x}\mathbf{x}^T \succeq \mathbf{0}. \end{aligned}$$

6.5 Complementarity and Solution Rank of SDP

In linear programming, since $\mathbf{x} \geq \mathbf{0}$ and $\mathbf{s} \geq \mathbf{0}$,

$$0 = \mathbf{x} \bullet \mathbf{s} = \mathbf{x}^T\mathbf{s} = \sum_{j=1}^{n} x_j s_j$$

implies that $x_j s_j = 0$ for all $j = 1, \ldots, n$. This property is often called complementarity. Thus, besides feasibility, and optimal linear programming solution pair must satisfy complementarity.

Now consider semidefinite cone $\mathcal{S}^n_+$. Since $\mathbf{X} \succeq \mathbf{0}$ and $\mathbf{S} \succeq \mathbf{0}$, $0 = \mathbf{X} \bullet \mathbf{S}$ implies $\mathbf{XS} = \mathbf{0}$, that is, the regular matrix product of the two is a zero matrix. In other words, every column (or row) of $\mathbf{X}$ is orthogonal to every column (or row) of $\mathbf{X}$. We also call such property complementarity. Thus, besides feasibility, an optimal semidefinite programming solution pair must satisfy complementarity.

Proposition 1. *Let $\mathbf{X}^*$ and $(\mathbf{y}^*, \mathbf{S}^*)$ be any optimal SDP solution pair with zero duality gap. Then complementarity of $\mathbf{X}^*$ and $\mathbf{S}^*$ implies*

$$\text{rank}(\mathbf{X}^*) + \text{rank}(\mathbf{S}^*) \leq n.$$

Furthermore, is there an optimal (dual) $\mathbf{S}^*$ such that $\text{rank}\mathbf{S}^* \geq d$, *then the rank of any optimal (primal) $\mathbf{X}^*$ is bounded above by $n - d$, where integer $0 \leq d \leq n$; and the converse is also true.*

In certain SDP problems, one may be interested in finding an optimal solution whose rank is minimal, while the interior-point algorithm for SDP (developed later) typically generates solution whose rank is maximal for primal and dual, respectively. Thus, a rank reduction method sometimes is necessary to achieve this goal. For linear programming in the standard form, it is known that if there is an optimal solution, then there is an optimal *basic* solution $\mathbf{x}^*$ whose positive entries have at most m many. Is there a similar structural fact for semidefinite programming? In deed, we have

Proposition 2. *If there is an optimal solution for SDP, then there is an optimal solution of SDP whose rank r satisfies $\frac{r(r+1)}{2} \leq m$.*

The proposition resembles the linear programming fundamental theorem of Carathéodory in Sect. 2.4. We now give a sketch of similar constructive proof, as well as several other rank-reduction methods.

Null-Space Rank Reduction

Let $\mathbf{X}^*$ be an optimal solution of SDP with rank r. If $r(r+1)/2 > m$, we orthonormally factorize $\mathbf{X}^*$

$$\mathbf{X}^* = (\mathbf{V}^*)^T\mathbf{V}^*, \quad \mathbf{V}^* \in E^{r\times n}.$$

Then we consider a related SDP problem

$$\begin{array}{ll} \text{minimize} & \mathbf{V}^*\mathbf{C}(\mathbf{V}^*)^T \bullet \mathbf{U} \\ \text{subject to} & \mathbf{V}^*\mathbf{A}_i(\mathbf{V}^*)^T \bullet \mathbf{U} = b_i,\ i = 1,\ldots,m \\ & \mathbf{U} \in \mathcal{S}^r_+. \end{array} \tag{6.12}$$

Note that, for any feasible solution of (6.12) one can construct a feasible solution for original SDP using

$$\mathbf{X}(\mathbf{U}) = (\mathbf{V}^*)^T\mathbf{U}\mathbf{V}^* \quad \text{and} \quad \mathbf{C} \bullet \mathbf{X}(\mathbf{U}) = \mathbf{V}^*\mathbf{C}(\mathbf{V}^*)^T \bullet \mathbf{U}.$$

Thus, the minimal value of (6.12) is also z^*, and in particular $\mathbf{U} = \mathbf{I}$ (the identity matrix) is an minimizer of (6.12), since

$$\mathbf{V}^*\mathbf{C}(\mathbf{V}^*)^T \bullet \mathbf{I} = \mathbf{C} \bullet (\mathbf{V}^*)^T\mathbf{V}^* = \mathbf{C} \bullet \mathbf{X}^* = z^*.$$

Also, one can show that any feasible solution $\mathbf{U}$ of (6.12) is its minimizer, so that $\mathbf{X}(\mathbf{U})$ is a minimizer of original SDP.

Consider the system of homogeneous linear equations:

$$\mathbf{V}^*\mathbf{A}_i(\mathbf{V}^*)^T \bullet \mathbf{W} = 0,\ i = 1,\ldots,m.$$

where $\mathbf{W} \in \mathcal{S}^r$ (i.e., a $r \times r$ symmetric matrices that does not need to be semidefinite). This system has $r(r+1)/2$ real variables and m equations. Thus, as long as $r(r+1)/2 > m$, we must be able to find a symmetric matrix $\mathbf{W} \neq \mathbf{0}$ to satisfy all the m equations. Without loss of generality, let $\mathbf{W}$ be either indefinite or negative semidefinite (if it is positive semidefinite, we take $-\mathbf{W}$ as $\mathbf{W}$), that is, $\mathbf{W}$ have at least one negative eigenvalue. Then we consider

$$\mathbf{U}(\alpha) = \mathbf{I} + \alpha\mathbf{W}.$$

Choosing a α^* sufficiently large such that $\mathbf{U}(\alpha^*) \succeq \mathbf{0}$ and it has at least one 0 eigenvalue (or rank$\mathbf{U}(\alpha^*) < r$). Note that

$$\mathbf{V}^*\mathbf{A}_i(\mathbf{V}^*)^T \bullet \mathbf{U}(\alpha^*) = \mathbf{V}^*\mathbf{A}_i(\mathbf{V}^*)^T \bullet (\mathbf{I} + \alpha^*\mathbf{W}) = \mathbf{V}^*\mathbf{A}_i(\mathbf{V}^*)^T \bullet I = b_i,\ i = 1,\ldots,m.$$

That is, $\mathbf{U}(\alpha^*)$ is feasible and also optimal for (6.12). Thus, $\mathbf{X}(\mathbf{U}(\alpha^*))$ is a new minimizer for the original SDP, and its rank is strictly less than r. This process can be repeated till the system of homogeneous linear equations has only all-zero solution, which is necessary when $r(r+1)/2 \le m$. Such a solution rank reduction procedure is called the Null-space reduction, which is deterministic.

To see an application of Proposition 2, consider a general quadratic minimization with sphere constraint

$$z^* \equiv \text{minimize } \mathbf{x}^T\mathbf{Q}\mathbf{x} + 2\mathbf{c}^T\mathbf{x}$$
$$\text{subject to } |\mathbf{x}|^2 = 1, \ \mathbf{x} \in E^n,$$

where Q is general. The problem has an SDP relaxation:

$$\begin{aligned} z^{SDP} \equiv \text{maximize } & \begin{bmatrix} \mathbf{Q} & \mathbf{c} \\ \mathbf{c}^T & 0 \end{bmatrix} \bullet \mathbf{Y} \\ \text{subject to } & \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0}^T & 0 \end{bmatrix} \bullet \mathbf{Y} = 1, \\ & \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0}^T & 1 \end{bmatrix} \bullet \mathbf{Y} = 1, \\ & \mathbf{Y} \in \mathcal{S}_+^{n+1}. \end{aligned}$$

Note that the relaxation and its dual both have interior so that the strong duality theorem holds, and it must have a rank-1 optimal SDP solution because $m = 2$. But a rank-1 optimal SDP solution would be optimal to the original quadratic minimization with sphere constraint. Thus, we must have $z^* = z^{SDP}$.

Gaussian Projection Reduction

There is also a randomized procedure to produce an approximate SDP solution with a desired low rank d. Again, let $\mathbf{X}^*$ be an optimal solution of SDP with rank $r > d$ and we factorize $\mathbf{X}^*$ as

$$\mathbf{X}^* = (\mathbf{V}^*)^T\mathbf{V}^*, \quad \mathbf{V}^* \in E^{r \times n}.$$

We then generate i.i.d. Gaussian random variables ξ_i^j with mean 0 and variance $1/d$, $i = 1, \ldots, r;\ j = 1, \ldots, d$, and form random vectors $\xi^j = (\xi_1^j; \ldots; \xi_r^j)$, $j = 1, \ldots, d$. Finally, we let

$$\hat{\mathbf{X}} = (\mathbf{V}^*)^T \left[\sum_{j=1}^d \xi^j(\xi^j)^T \right] \mathbf{V}^*.$$

Note that the rank of $\hat{\mathbf{X}}$ is d and

$$\mathrm{E}(\hat{\mathbf{X}}) = (\mathbf{V}^*)^T \mathrm{E}\left[\sum_{j=1}^d \xi^j(\xi^j)^T \right] \mathbf{V}^* = (\mathbf{V}^*)^T\mathbf{I}\mathbf{V}^* = \mathbf{X}^*.$$

One can further show that $\hat{\mathbf{X}}$ would be a good rank-d approximate SDP solution in many cases.

Randomized Binary Reduction

As discussed in the binary QP optimization, we like to produce a vector $\mathbf{x}$ where each entry is either 1 or -1. A procedure to achieve this is as follows. Let $\mathbf{X}^*$ be any optimal solution of SDP and we factorize $\mathbf{X}^*$ as

$$\mathbf{X}^* = (\mathbf{V}^*)^T\mathbf{V}^*, \quad \mathbf{V}^* \in E^{n\times n}.$$

Then, we generate a random n-dimensional vector ξ where each entry is a i.i.d. Gaussian random variable with mean 0 and variance 1. Then we let

$$\hat{\mathbf{x}} = \text{sign}((\mathbf{V}^*)^T\xi)$$

where

$$\text{sign}(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ -1 & \text{otherwise.} \end{cases}$$

It was proved by Sheppard [228]:

$$\text{E}[\hat{x}_i\hat{x}_j] = \frac{2}{\pi}\arcsin(\mathbf{X}^*_{ij}), \quad i, j = 1, 2, \ldots, n.$$

Obviously, each entry of $\hat{\mathbf{x}}$ is either 1 or -1.

One can further show $\hat{\mathbf{x}}$ would be a good approximate solution to the original binary QP. Let us consider the (homogeneous) binary quadratic maximization problem

$$\begin{aligned} z^* := \text{maximize } & \mathbf{x}^T\mathbf{Q}\mathbf{x} \\ \text{subject to } & x_j = \{1, -1\}, \text{ for all } j = 1, \ldots, n, \end{aligned}$$

where we assume Q is positive semidefinite. Then, the SDP relaxation would be

$$\begin{aligned} z^{SDP} := \text{maximize } & \mathbf{Q} \bullet \mathbf{X} \\ \text{subject to } & \mathbf{I}_j \bullet \mathbf{X} = 1, \text{ for all } j = 1, \ldots, n, \\ & \mathbf{X} \in \mathcal{S}^n_+; \end{aligned}$$

and let $\mathbf{X}^*$ be any optimal solution, from which we produced a random binary vector $\hat{\mathbf{x}}$. Let us evaluate the expected objective value

$$\text{E}(\hat{\mathbf{x}}^T\mathbf{Q}\hat{\mathbf{x}}) = \text{E}(\mathbf{Q} \bullet \hat{\mathbf{x}}\hat{\mathbf{x}}^T) = \mathbf{Q} \bullet \text{E}(\hat{\mathbf{x}}\hat{\mathbf{x}}^T) = Q \bullet \frac{2}{\pi}\arcsin[\mathbf{X}^*] = \frac{2}{\pi}(Q \bullet \arcsin[\mathbf{X}^*]),$$

where $\arcsin[\mathbf{X}^*] \in \mathcal{S}^n$ whose (i, j) the entry equals $\arcsin(\mathbf{X}^*_{ij})$. One can further show

$$\arcsin[\mathbf{X}^*] - \mathbf{X}^* \geq \mathbf{0}$$

so that (from $Q \geq \mathbf{0}$)

$$Q \bullet \arcsin[\mathbf{X}^*] \geq Q \bullet \mathbf{X}^* = z^{SDP} \geq z^*,$$

that is, the expected objective value of $\hat{\mathbf{x}}$ is no less than factor $\frac{2}{\pi}$ of the maximal value of the binary QP.

The randomized binary reduction can be extended to quadratic optimization with simple bound constraints such as $x_j^2 \leq 1$.

6.6 Interior-Point Algorithms for Conic Linear Programming

Since (CLP) is a convex minimization problem, many optimization algorithms are applicable for solving it. However, the most natural conic linear programming algorithm seems to be an extension of the interior-point linear programming algorithm described in Chap. 5. We describe what it is now.

To develop efficient interior-point algorithms, the key is to find a suitable barrier or potential function. There is a general theory on selection of barrier functions for (CLP), depending on the convex cone involved. We present few for the convex cones listed in Example 1.

Example 1. The following are barrier function for each of the convex cones.

- The n-dimensional non-negative orthant E^n_+:

$$B(\mathbf{x}) = -\sum_{j=1}^{n} \log(x_j).$$

- The n-dimensional semidefinite cone $\mathcal{S}^n_+$:

$$B(\mathbf{X}) = -\log(\det \mathbf{X}).$$

- The $(n+1)$-dimensional second-order cone $\{(u; \mathbf{x}) : \ u \geq |\mathbf{x}|\}$:

$$B(u; \mathbf{x}) = -\log(u^2 - |\mathbf{x}|^2).$$

In the rest of the section, we devote our discussion on solving (SDP). Similar to LP, we consider (SDP) with the barrier function added in the objective:

$$\begin{array}{ll}(SDPB) & \text{minimize } \mathbf{C} \bullet \mathbf{X} - \mu \log\det(\mathbf{X}) \\ & \text{subject to } \mathbf{X} \in \overset{\circ}{\mathcal{F}}_p,\end{array}$$

or (SDD) with the barrier function added in the objective:

$$\begin{array}{ll}(SDDB) & \text{maximize } \mathbf{y}^T\mathbf{b} + \mu \log\det(\mathbf{S}) \\ & \text{subject to } (\mathbf{y}, \mathbf{S}) \in \overset{\circ}{\mathcal{F}}_d,\end{array}$$

where again $\mu > 0$ is called the barrier weight parameter. For a given μ, the minimizers of (SDPB) and (SDDB) satisfy conditions:

$$\begin{aligned} \mathbf{XS} &= \mu\mathbf{I} \\ \mathcal{A}\mathbf{X} &= \mathbf{b} \\ \mathcal{A}^T\mathbf{y} + \mathbf{S} &= \mathbf{C} \\ \mathbf{X} \succ \mathbf{0}, &\quad \mathbf{S} \succ \mathbf{0} \end{aligned} \tag{6.13}$$

Since

$$\mu = \frac{\text{trace}(\mathbf{XS})}{n} = \frac{X \bullet S}{n} = \frac{\mathbf{C} \bullet \mathbf{X} - \mathbf{y}^T\mathbf{b}}{n},$$

so that μ equals the average of complementarity or duality gap. And, these minimizers, denoted by $(\mathbf{X}(\mu), \mathbf{y}(\mu), \mathbf{S}(\mu))$, form the central path of SDP for $mu \in (0, \infty)$. It is known that when $\mu \to 0$, $(\mathbf{X}(\mu), \mathbf{y}(\mu), \mathbf{S}(\mu))$ tends to an optimal solution pair whose rank is maximal (Exercise 11).

We can also extend the primal-dual potential function from LP to SDP as a descent merit function:

$$\psi_{n+\rho}(\mathbf{X},\ \mathbf{S}) = (n + \rho)\log(\mathbf{X} \bullet \mathbf{S}) - \log(\det(\mathbf{X}) \cdot \det(\mathbf{S}))$$

where $\rho \geq 0$. Note that if $\mathbf{X}$ and $\mathbf{S}$ are diagonal matrices, these definitions reduce to those for linear programming.

Once we have an interior feasible point $(\mathbf{X},\ \mathbf{y},\ \mathbf{S})$, we can generate a new iterate $(\mathbf{X}^+,\ \mathbf{y}^+,\ \mathbf{S}^+)$ by solving for $(\mathbf{D_x},\ \mathbf{d_y},\ \mathbf{D_s})$ from the primal-dual system of linear equations

$$\begin{aligned} \mathbf{D}^{-1}\mathbf{D_x}\mathbf{D}^{-1} + \mathbf{D_s} &= \frac{n}{n+\rho}\mu\mathbf{X}^{-1} - \mathbf{S}, \\ \mathbf{A}_i \bullet \mathbf{D_x} &= 0, \text{ for all } i, \\ \sum\nolimits_i^m (\mathbf{d_y})_i\mathbf{A}_i + \mathbf{D_s} &= \mathbf{0}, \end{aligned} \tag{6.14}$$

where $\mathbf{D}$ is the (scaling) matrix

$$\mathbf{D} = \mathbf{X}^{\frac{1}{2}}\ (\mathbf{X}^{\frac{1}{2}}\ \mathbf{S}\mathbf{X}^{\frac{1}{2}})^{-\frac{1}{2}}\mathbf{X}^{\frac{1}{2}}$$

and $\mu = \mathbf{X} \bullet \mathbf{S}/n$. Then one assigns $\mathbf{X}^+ = \mathbf{X} + \alpha\mathbf{D_x}$, $\mathbf{y}^+ = \mathbf{y} + \alpha\mathbf{d_y}$, and $\mathbf{S}^+ = \mathbf{s} + \alpha\mathbf{D_s}$ for a step size $\alpha > 0$. Furthermore, it can be shown that there exists a step size $\alpha = \bar{\alpha}$ such that

$$\psi_{n+\rho}(\mathbf{X}^+,\ \mathbf{S}^+) - \psi_{n+\rho}(\mathbf{X},\ \mathbf{S}) \leq -\delta$$

for a constant $\delta > 0.2$.

We outline the algorithm here

Step 1. Given $(\mathbf{X}^0, \mathbf{y}^0, \mathbf{S}^0) \in \overset{\circ}{\mathcal{F}}$. Set $\rho \geq \sqrt{n}$ and $k := 0$.
Step 2. Set $(\mathbf{X}, \mathbf{S}) = (\mathbf{X}^k, \mathbf{S}^k)$ and compute $(\mathbf{D_x}, \mathbf{d_y}, \mathbf{D_s})$ from (6.14).

Step 3. Let $\mathbf{X}^{k+1} = \mathbf{X}^k + \bar{\alpha}\mathbf{D_x}$, $\mathbf{y}^{k+1} = \mathbf{y}^k + \bar{\alpha}\mathbf{d_y}$, and $\mathbf{S}^{k+1} = \mathbf{S}^k + \bar{\alpha}\mathbf{D_s}$, where

$$\bar{\alpha} = \arg\min_{\alpha \geq 0} \psi_{n+\rho}(\mathbf{X}^k + \alpha\mathbf{D_x}, \mathbf{S}^k + \alpha\mathbf{D_s}).$$

Step 4. Let $k := k + 1$. If $\frac{\mathbf{X}^k \bullet \mathbf{S}^k}{\mathbf{X}^0 \bullet \mathbf{S}^0} \leq \epsilon$, Stop. Otherwise return to Step 2.

Theorem 3. *Let* $\psi_{n+\rho}(\mathbf{X}^0, \mathbf{S}^0) \leq \rho \log(\mathbf{X}^0 \bullet \mathbf{S}^0) + n \log n$. *Then, the algorithm terminates in at most* $O(\rho \log(n/\epsilon))$ *iterations.*

Initialization: The HSD Algorithm

The linear programming Homogeneous Self-Dual Algorithm is also extendable to conic linear programming. Consider the minimization problem Homogeneous self-dual algorithm! for conic linear programming

$$\begin{array}{lllllll}
(HSDCLP) & \min & & & & (n+1)\theta & \\
& \text{s.t.} & & \mathcal{A}\mathbf{X} & -\mathbf{b}\tau & +\bar{\mathbf{b}}\theta & = \mathbf{0}, \\
& & -\mathcal{A}^T\mathbf{y} & & +\mathbf{C}\tau & -\bar{\mathbf{C}}\theta & = \mathbf{S} \in K^*, \\
& & \mathbf{b}^T\mathbf{y} & -\mathbf{C}\bullet\mathbf{X} & & +\bar{z}\theta & = \kappa \geq 0, \\
& & -\bar{\mathbf{b}}^T\mathbf{y} & +\bar{\mathbf{C}}\bullet\mathbf{X} & -\bar{z}\tau & & = -(n+1), \\
& & \mathbf{y} \text{ free}, & \mathbf{X} \in K, & \tau \geq 0, & \theta \text{ free}, &
\end{array}$$

where

$$\bar{\mathbf{b}} = \mathbf{b} - \mathcal{A}\mathbf{X}^0, \quad \bar{\mathbf{C}} = \mathbf{C} - \mathbf{S}^0, \quad \bar{z} = \mathbf{C} \bullet \mathbf{X}^0 + 1$$

Here $\mathbf{X}^0$ and $\mathbf{S}^0$ are any pair of interior points in the interior of K and K^* such that they form a central path point with $\mu = 1$. Note that $\mathbf{X}^0$ and $\mathbf{S}^0$ don't need to satisfy other equality constraint, so that they can be easily identified. For examples, $\mathbf{x}^0 = \mathbf{y}^0 = \mathbf{1}$ for the nonnegative orthant cone; $\mathbf{x}^0 = \mathbf{y}^0 = (1; \mathbf{0})$ for the p-order cone; and $\mathbf{X}^0 = \mathbf{X}^0 = \mathbf{I}$ for the semidefinite cone.

Let $\mathcal{F}$ be the set of all feasible points $(\mathbf{y}, \mathbf{X} \in K, \tau \geq 0, \theta, \mathbf{S} \in K^*, \kappa \geq 0)$. Then $\overset{\circ}{\mathcal{F}}$ is the set of interior feasible points $(\mathbf{y}, \mathbf{X} \in \overset{\circ}{K}, \tau > 0, \theta, \mathbf{S} \in \overset{\circ}{K}{}^*, \kappa > 0)$.

Theorem 4. *Consider the conic optimization (HSDCLP).*

i) (HSDCLP) is self-dual, that is, its dual has an identical form of (HSDCLP).
ii) (HSDCLP) has an optimal solution and its optimal solution set is bounded.
iii) (HSDCLP) has an interior feasible point

$$\mathbf{y} = \mathbf{0}, \quad \mathbf{X} = \mathbf{X}^0, \quad \tau = 1, \quad \theta = 1, \quad \mathbf{S} = \mathbf{S}^0, \quad \kappa = 1.$$

iv) For any feasible point $(\mathbf{y}, \mathbf{X}, \tau, \theta, \mathbf{S}, \kappa) \in \mathcal{F}$

$$\mathbf{S}^0 \bullet \mathbf{X} + \mathbf{X}^0 \bullet \mathbf{S} + \tau + \kappa - (n+1)\theta = (n+1),$$

and

$$\mathbf{X} \bullet \mathbf{S} + \tau\kappa = (n+1)\theta.$$

v) The optimal objective value of (HSDCLP) is zero, that is, any optimal solution of (HSDCLP) has

$$\mathbf{X}^* \bullet \mathbf{S}^* + \tau^*\kappa^* = (n+1)\theta^* = 0.$$

Now we are ready to apply the interior-point algorithm, starting from a available initial interior-point feasible solution, to solve (HSDCLP). The question is: how is an optimal solution of (HSDCLP) related to optimal solutions of original (CLP) and (CLD)? We present the next theorem, and leave this proof as an exercise.

Theorem 5. *Let* $(\mathbf{y}^*, \mathbf{X}^*, \tau^*, \theta^* = 0, \mathbf{S}^*, \kappa^*)$ *be a (maximal rank) optimal solution of (HSDCLP) (as it is typically computed by interior-point algorithms).*

i) (CLP) and (CLD) have an optimal solution pair if and only if $\tau^* > 0$*. In this case,* $\mathbf{X}^*/\tau^*$ *is an optimal solution for (CLP) and* $(\mathbf{y}^*/\tau^*, \mathbf{S}^*/\tau^*)$ *is an optimal solution for (CLD).*

ii) (CLP) or (CLD) has an infeasibility certificate if and only if $\kappa^* > 0$*. In this case,* $\mathbf{X}^*/\kappa^*$ *or* $\mathcal{S}^*/\kappa^*$ *or both are certificates for proving infeasibility; see Farkas' lemma for CLP.*

iii) For all other cases, $\tau^* = \kappa^* = 0$.

6.7 Summary

A relatively new class of mathematical programming problems, Conic linear programming (hereafter CLP), is a natural extension of Linear programming that is a central decision model in Management Science and Operations Research. In CLP, the unknown is a vector or matrix in a closed convex cone while its entries are also restricted by some linear equalities and/or inequalities.

One of cones is the semidefinite cone, that is, the set of all symmetric positive semidefinite matrices in a given dimension. There is a variety of interesting and important practical problems that can be naturally cast in this form. Because many problems which appear nonlinear (such as quadratic problems) become essentially linear in semidefinite form. We have described some of these applications and selected results in Combinatory Optimization, Robust Optimization, and Engineering Sensor Network. We have also illustrated some analyses to show why CLP is an effective model to tackle these difficult optimization problems.

We present fundamental theorems underlying conic linear programming. These theorems include Farkas' lemma, weak and strong dualities, and solution rank structure. We show the common features and differences of these theorems between LP and CLP.

The efficient interior-point algorithms for linear programming can be extended to solving these problems as well. We describe these extensions applied to general conic programming problems. These algorithms closely parallel those for linear programming. There is again a central path and potential functions, and Newton's method is a good way to follow the path or reduce the potential function. The homogeneous and self-dual algorithm, which is popularly used for linear programming, is also extended to CLP.

6.8 Exercises

1. Prove that

   i) The dual cone of E^n_+ is itself.
   ii) The dual cone of $\mathcal{S}^n_+$ is itself.
   iii) The dual cone of p-order cone is the q-order cone where $\frac{1}{p} + \frac{1}{q} = 1$ and $1 \le p \le \infty$.

2. When both K_1 and K_2 are closed convex cones. Show

   i) $(K_1^*)^* = K_1$.
   ii) $K_1 \subset K_2 \Longrightarrow K_2^* \subset K_1^*$.
   iii) $(K_1 \oplus K_2)^* = K_1^* \oplus K_2^*$.
   iv) $(K_1 + K_2)^* = K_1^* \cap K_2^*$.
   v) $(K_1 \cap K_2)^* = K_1^* + K_2^*$.

   Note: by definition $S + T = \{\mathbf{s} + \mathbf{t} : \mathbf{s} \in S,\ \mathbf{t} \in T\}$.
3. Prove the following:

   i) Theorem 1.
   ii) Proposition 1.
   iii) Let $\mathbf{X} \in \overset{\circ}{K}$ and $\mathbf{Y} \in \overset{\circ}{K^*}$. Then $\mathbf{X} \bullet \mathbf{Y} > 0$.

4. Guess an optimal solution and the optimal objective value of each instance of Example 1.
5. Prove the second statement of Theorem 2.
6. Verify the weak duality theorem of the three CLP instances in Example 1 in Sect. 6.2 and Example 1 in Sect. 6.4.
7. Consider the SDP relaxation of the sensor network localization problem with four sensors:

$$(\mathbf{e}_i - \mathbf{e}_j)(\mathbf{e}_i - \mathbf{e}_j)^T \bullet \mathbf{X} = 1,\ \forall i < j = 1, 2, 3, 4,$$
$$\mathbf{X} \in \mathcal{S}^4_+,$$

   in which $m = 6$. Show that the SDP problem has the solution with rank 3, which reaches the bound of Proposition 2.
8. Let $\mathbf{A}$ and $\mathbf{B}$ be two symmetric and positive semidefinite matrices. Prove that $\mathbf{A} \bullet \mathbf{B} \ge 0$, and $\mathbf{A} \bullet \mathbf{B} = 0$ implies $\mathbf{AB} = \mathbf{0}$.
9. Let $\mathbf{X}$ and $\mathbf{S}$ both be positive definite. Prove that

$$n \log(\mathbf{X} \bullet \mathbf{S}) - \log(\det(\mathbf{X}) \cdot \det(\mathbf{S})) \ge n \log n.$$

10. Consider a SDP and the potential level set

$$\mathbf{\Psi}(\delta) = \{(\mathbf{X},\ \mathbf{y},\ \mathbf{S}) \in \overset{\circ}{\mathcal{F}} : \psi_{n+\rho}(\mathbf{X},\ \mathbf{S}) \le \delta\}.$$

Prove that

$$\Psi(\delta^1) \subset \Psi(\delta^2) \quad \text{if} \quad \delta^1 \le \delta^2,$$

and for every δ, $\mathbf{\Psi}(\delta)$ is bounded and its closure $\overline{\mathbf{\Psi}}(\delta)$ has non-empty intersection with the SDP solution set.

11. Let both (SDP) and (SDD) have interior feasible points. Then for any $0 < \mu < \infty$, the central path point $(\mathbf{X}(\mu), \mathbf{y}(\mu), \mathbf{S}(\mu))$ exists and is unique. Moreover,

    i) the central path point $(\mathbf{X}(\mu), \mathbf{y}(\mu), \mathbf{S}(\mu))$ is bounded where $0 < \mu \le \mu^0$ for any given $0 < \mu^0 < \infty$.
    ii) For $0 < \mu' < \mu$,

    $$\mathbf{C} \bullet \mathbf{X}(\mu') < \mathbf{C} \bullet \mathbf{X}(\mu) \quad \text{and} \quad \mathbf{b}^T\mathbf{y}(\mu') > \mathbf{b}^T\mathbf{y}(\mu)$$

    if $\mathbf{X}(\mu) \ne \mathbf{X}(\mu')$ and $\mathbf{y}(\mu) \ne \mathbf{y}(\mu')$.
    iii) $(\mathbf{X}(\mu), \mathbf{y}(\mu), \mathbf{S}(\mu))$ converges to an optimal solution pair for (SDP) and (SDD), and the rank of the limit of $\mathbf{X}(\mu)$ is maximal among all optimal solutions of (SDP) and the rank of the limit $\mathbf{S}(\mu)$ is maximal among all optimal solutions of (SDD).

12. Prove the logarithmic approximation lemma for SDP. Let $\mathbf{D} \in \mathcal{S}^n$ and $|\mathbf{D}|_\infty < 1$. Then,

$$\text{trace}(\mathbf{D}) \ge \log\det(\mathbf{I} + \mathbf{D}) \ge \text{trace}(\mathbf{D}) - \frac{|\mathbf{D}|^2}{2(1 - |\mathbf{D}|_\infty)}.$$

13. Let $\mathbf{V} \in \overset{\circ}{\mathcal{S}}{}^n_+$ and $\rho \ge \sqrt{n}$. Then,

$$\frac{|\mathbf{V}^{-1/2} - \frac{n+\rho}{\mathbf{I}\bullet\mathbf{V}}\mathbf{V}^{1/2}|}{|\mathbf{V}^{-1/2}|_\infty} \ge \sqrt{3/4}.$$

14. Prove both Theorems 4 and 5.

References

6.1 Most of the materials presented can be found from convex analysis, such as Rockeafellar [219].

6.2 Semidefinite relaxations have appeared in relation to relaxations discrete optimization problems. In Lovasz and Shrijver [159], a "lifting" procedure is presented to obtain a problem in $\mathfrak{R}^{n^2}$; and then the problem is projected back to obtain tighter inequalities; see also Balas et al. [12]. Then, there have been several remarkable results of SDP relaxations for combinatorial optimization. The binary QP, a generalized Max-Cut problem, was studied by Goemans and Williamson [G8] and Nesterov [189]. Other SDP relaxations can be found in the survey by Luo et al. [171] and references therein. More CLP applications can be found in Boyd et al [B22], Vandenberghe and Boyd [V2], and Lobo, Vandenberghe and Boyd [156], Lasserre [150], Parrilo [204], etc.

The sensor localization problem described here is due to Biswas and Ye [B17]. Note that we can view the Sensor Network Localization problem as a Graph Realization or Embedding problem in Euclidean spaces, see So and Ye [231] and references therein; and it is related to the Euclidean Distance Matrix Completion Problems, see Alfakih et al. [3] and Laurent [151].

6.3 Farkas' lemma for conic linear constraints are closely linked to convex analysis (i.e, Rockeafellar [219]) and the CLP duality theorems commented next.

6.4 The conic formulation of the Euclidean facility location problem was due to Xue and Ye [264]. For discussion of Schur complements see Boyd and Vandenberghe [B23]. Robust optimization models using SDP can be found in Ben-Tal and Nemirovski [26] and Goldfarb and Iyengar [112], and etc. The SDP duality theory was studied by Barvinok [16], Nesterov and Nemirovskii [N2], Ramana [214], Ramana e al. [215], etc. The SDP example with a duality gap was constructed by R. Freund (private communication).

6.5 Complementarity and rank. The exact rank theorem described here is due to Pataki [205], also see Barvinok [15]. A analysis of the Gaussian projection was presented by So et al. [232] which can be sees as a generalization of the Johnson and Lindenstrauss theorem [137]. The expectation of the randomized binary reduction is due to Sheppard [228] in 1900, and it was extensively used in Goemans and Williamson [G8] and Nesterov [189], Ye [265], and Bertsimas and Ye, [31].

6.6 In interior-point algorithms, the search direction $(\mathbf{D_x}, \mathbf{d_y}, \mathbf{D_s})$ can be determined by Newton's method with three different scalings: primal, dual and primal-dual. A primal-scaling (potential reduction) algorithm for semidefinite programming is due to Alizadeh [A4, A3] where Yinyu Ye "suggested studying the primal-dual potential function for this problem" and "looking at symmetric preserving scalings of the form $X_0^{-1/2}XX_0^{-1/2}$", and to Nesterov and Nemirovskii [N2]. A dual-scaling algorithm was developed by Benson et al. [25] which exploits the sparse structure of the dual SDP. The primal-dual SDP algorithm described here is due to Nesterov and Todd [N3] and references therein.

Efficient interior-point algorithms are also developed for optimization over the second-order cone; see Nesterov and Nemirovskii [N2] and Xue and Ye [264]. These algorithms have established the best approximation complexity results for certain combinatorial location problems.

The homogeneous and self-dual initialization model was originally developed by Ye, Todd and Mizuno for LP [Y2], and for SDP by de Klerk et al. [72], Luo et al. [L18], and Nesterov et al. [191], and it became the foundational algorithm implemented in Sturm [S11] and Andersen [6].