Chapter 3 LP, QP, SDP, SOCP Problems and Software

In this chapter we examine several specific classes of convex constrained problems. For each class, examples are provided to illustrate its usefulness by explaining how practical problems can be formulated as that type of optimization. A popular CP solver known as **CVX** is introduced to facilitate computing numerical solution of CP problems.

3.1 Linear Programming

3.1.1 Examples

Example 3.1 *Transportation problem*

Quantities q_1, q_2, …, q_m of a certain product are produced by m manufacturing divisions of a company, which are at distinct locations. The product is to be shipped to n destinations that require quantities b_1, b_2, …, b_n. Assume that the cost of shipping a unit from manufacturing division i to destination j is $c_{i,j}$ with $i = 1, 2, …, m$ and $j = 1, 2, …, n$.

Let $x_{i,j}$ to be the quantity shipped from division i to destination j so as to minimize the total cost of transportation. There are several constraints on variables $x_{i,j}$. First, each division can provide only a fixed quantity of the product, hence

$$\sum_{j=1}^{n} x_{i,j} = q_i \quad \text{for} \quad i = 1, 2, ..., m$$

Second, the quantity to be shipped to a specific destination has to meet the need of that destination and so

$$\sum_{i=1}^{m} x_{i,j} = b_j \quad \text{for} \quad j = 1, 2, ..., n$$

In addition, the variables $x_{i,j}$ are nonnegative and thus, we have

$$x_{i,j} \geq 0 \quad \text{for} \quad i = 1, 2, ..., m \quad \text{and} \quad j = 1, 2, ..., n$$

The transportation problem is to find quantities $\{x_{i,j}\}$ that minimizes $C = \sum_{i=1}^{m}\sum_{j=1}^{n} c_{i,j}x_{i,j}$ subject to the above constraints. If we let

$$\boldsymbol{c} = \begin{bmatrix} c_{11} & \cdots & c_{1n} & c_{21} & \cdots & c_{2n} & \cdots & c_{m1} & \cdots & c_{mn} \end{bmatrix}^T$$

$$\boldsymbol{x} = \begin{bmatrix} x_{11} & \cdots & x_{1n} & x_{21} & \cdots & x_{2n} & \cdots & x_{m1} & \cdots & x_{mn} \end{bmatrix}^T$$

$$\boldsymbol{A} = \begin{bmatrix} 1 & 1 & \cdots & 1 & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 1 & 1 & \cdots & 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & \cdots & 1 & 1 & \cdots & 1 \\ 1 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & 1 & \cdots & 0 & \cdots & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 1 & \cdots & 0 & 0 & \cdots & 1 \end{bmatrix}$$

$$\boldsymbol{b}=\begin{bmatrix}q_1 & \cdots & q_m & b_1 & \cdots & b_n\end{bmatrix}^T$$

then the transportation problem can be stated as

$$\begin{aligned}&\text{minimize} \quad C=\boldsymbol{c}^T\boldsymbol{x}\\&\text{subject to: } \boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}\\&\qquad\qquad\quad \boldsymbol{x}\geq\boldsymbol{0}\end{aligned}$$

Since both the objective function and constraints are linear, the problem is known as a *linear programming (LP) problem*. ■

Example 3.2 *Chebyshev centre of a polyhedron*

A polyhedron is a set defined by

$$\mathcal{P}=\{\boldsymbol{x}:\boldsymbol{a}_i^T\boldsymbol{x}\leq b_i,\, i=1,2,\ldots,q\}$$

The Chebeshev centre of $\mathcal{P}$ is the centre $\boldsymbol{x}_c$ of the largest ball $\mathcal{B}=\{\boldsymbol{x}_c+\boldsymbol{u}:\|\boldsymbol{u}\|_2\leq r\}$ that is inscribed in $\mathcal{P}$.

To find the largest inscribed ball, the centre $\boldsymbol{x}_c$ and radius r of ball $\mathcal{B}$ are treated as *variables* such that r is to be maximized among the balls inscribed in $\mathcal{P}$. Note that

$$\boldsymbol{a}_i^T\boldsymbol{x}\leq b_i \text{ for } \boldsymbol{x}\in\mathcal{B} \text{ for } i=1,2,\ldots,q$$

$$\Updownarrow$$

$$\sup_{\|\boldsymbol{u}\|_2\leq r}\{\boldsymbol{a}_i^T(\boldsymbol{x}_c+\boldsymbol{u})\leq b_i\} \text{ for } i=1,2,\ldots,q$$

$$\Updownarrow$$

$$\boldsymbol{a}_i^T\boldsymbol{x}_c+\sup_{\|\boldsymbol{u}\|_2\leq r}\{\boldsymbol{a}_i^T\boldsymbol{u}\}\leq b_i \text{ for } i=1,2,\ldots,q$$

$$\Updownarrow$$

$$\boldsymbol{a}_i^T\boldsymbol{x}_c+r\|\boldsymbol{a}_i\|_2\leq b_i \text{ for } i=1,2,\ldots,q$$

Therefore, the problem at hand can be formulated as

$$\begin{aligned}&\text{minimize} \quad -r\\&\text{subject to: } \boldsymbol{a}_i^T\boldsymbol{x}_c+r\|\boldsymbol{a}_i\|_2\leq b_i \text{ for } i=1,2,\ldots,q\end{aligned}$$

which is obviously an LP problem. ■

Example 3.3 *Piecewise-linear minimization*

It is well known that the piecewise linear function $f(\boldsymbol{x})=\max_{1\leq i\leq q}\{\boldsymbol{a}_i^T\boldsymbol{x}+b_i\}$ itself is convex, and is a simple model to approximate sophisticated convex functions. Here we consider minimizing such a function, i.e.,

$$\text{minimize} \quad \max_{1\leq i\leq q}\{\boldsymbol{a}_i^T\boldsymbol{x}+b_i\}$$

By introducing an auxiliary scalar variable t, the above problem can be cast as an LP problem:

$$\begin{aligned}&\text{minimize} \quad t\\&\text{subject to: } \boldsymbol{a}_i^T\boldsymbol{x}+b_i\leq t \quad \text{for} \quad i=1,2,\ldots,q\end{aligned}$$

■

3.1.2 Primal and Dual of LP Problems

Consider the standard LP problem

$$\begin{aligned} &\text{minimize} \quad \boldsymbol{c}^T\boldsymbol{x} \\ &\text{subject to:} \ \ \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \\ &\qquad\qquad\quad \boldsymbol{x} \geq \boldsymbol{0} \end{aligned} \tag{3.1}$$

where matrix $\boldsymbol{A} \in R^{p\times n}$ is of full row rank as the *primal* problem. From Chapter 2, the *dual* problem to (3.1) is given by

$$\begin{aligned} &\underset{\boldsymbol{\lambda},\,\boldsymbol{\mu}}{\text{maximize}} \quad -\boldsymbol{b}^T\boldsymbol{\lambda} \\ &\text{subject to:} \quad -\boldsymbol{A}^T\boldsymbol{\lambda} + \boldsymbol{\mu} = \boldsymbol{c} \\ &\qquad\qquad\qquad \boldsymbol{\mu} \geq \boldsymbol{0} \end{aligned} \tag{3.2}$$

It is known from the KKT conditions that $\boldsymbol{x}^*$ is a minimizer of the problem in (3.1) if and only if there exist $\boldsymbol{\lambda}^*$ and $\boldsymbol{\mu}^*$ with $\boldsymbol{\mu}^* \geq \boldsymbol{0}$ such that

$$-\boldsymbol{A}^T\boldsymbol{\lambda}^* + \boldsymbol{\mu}^* = \boldsymbol{c} \tag{3.3a}$$

$$\boldsymbol{A}\boldsymbol{x}^* = \boldsymbol{b} \tag{3.3b}$$

$$x_i^* \cdot \mu_i^* = 0 \quad \text{for } \ 1 \leq i \leq n \tag{3.3c}$$

$$\boldsymbol{x}^* \geq \boldsymbol{0}, \quad \boldsymbol{\mu}^* \geq \boldsymbol{0} \tag{3.3d}$$

It follows that $\boldsymbol{x}^*$ is a feasible point of (3.1) and $(\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$ is a feasible pair of (3.2). Actually, $\boldsymbol{x}^*$ solves problem (3.1) and $(\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$ solves the dual problem (3.2). Therefore we shall call $\{\boldsymbol{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*\}$ satisfying (3.3) a primal-dual solution.

Let $\{\boldsymbol{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*\}$ satisfies (3.3), then $\boldsymbol{x}$* is a solution of (3.1) and $(\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$ are the Lagrange multipliers of (3.1). Furthermore, $(\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$ is a solution of (3.2) and $\boldsymbol{x}^*$ may be interpreted as the Lagrange multiplier of the dual problem (3.2).

The *duality gap*, defined as the difference between the cost of the primal and the cost of the dual, is given by

$$\delta(\boldsymbol{x}, \boldsymbol{\lambda}) = \boldsymbol{c}^T\boldsymbol{x} - \left(-\boldsymbol{b}^T\boldsymbol{\lambda}\right) \tag{3.4}$$

which is always nonnegative, namely,

$$\delta(\boldsymbol{x}, \boldsymbol{\lambda}) \geq 0 \tag{3.5}$$

This is because from (3.1) and (3.2) we have

$$\boldsymbol{c}^T \boldsymbol{x} = \left(-\boldsymbol{A}^T \boldsymbol{\lambda} + \boldsymbol{\mu}\right)^T \boldsymbol{x} = -\boldsymbol{\lambda}^T \boldsymbol{A}\boldsymbol{x} + \boldsymbol{\mu}^T \boldsymbol{x} \geq -\boldsymbol{b}^T \boldsymbol{\lambda} \tag{3.6}$$

Furthermore, at the primal-dual solution $\{\boldsymbol{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*\}$, the duality gap is reduced to its minimum – zero, i.e.,

$$\delta(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) = 0 \tag{3.7}$$

This follows from a line of argument similar to (3.6) in conjunction with the complementarity condition (3.3c):

$$\boldsymbol{c}^T \boldsymbol{x}^* = \left(-\boldsymbol{A}^T \boldsymbol{\lambda}^* + \boldsymbol{\mu}^*\right)^T \boldsymbol{x}^* = -\boldsymbol{\lambda}^{*T} \boldsymbol{A}\boldsymbol{x}^* + \boldsymbol{\mu}^{*T} \boldsymbol{x}^* = -\boldsymbol{b}^T \boldsymbol{\lambda}^*$$

· Another important concept related to primal-dual solutions is *central path*. By (3.3), set $\{\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}\}$ is a primal-dual solution if it satisfies the conditions

$$\begin{aligned} &\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} && \text{with } \boldsymbol{x} \geq \boldsymbol{0} \\ &-\boldsymbol{A}^T \boldsymbol{\lambda} + \boldsymbol{\mu} = \boldsymbol{c} && \text{with } \boldsymbol{\mu} \geq \boldsymbol{0} \\ &\boldsymbol{X}\boldsymbol{\mu} = \boldsymbol{0} && \end{aligned} \tag{3.8}$$

where $\boldsymbol{X} = \text{diag}\{x_1, x_2, \ldots, x_n\}$.

3.2 Quadratic Programming

General quadratic programming (QP) problems assume the form

$$\begin{aligned} &\text{minimize} \quad \tfrac{1}{2}\boldsymbol{x}^T \boldsymbol{H}\boldsymbol{x} + \boldsymbol{p}^T \boldsymbol{x} + \kappa \\ &\text{subject to:} \quad \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \\ &\qquad\qquad\quad\ \boldsymbol{C}\boldsymbol{x} \leq \boldsymbol{d} \end{aligned} \tag{3.9a-c}$$

The QP problem is convex if matrix $\boldsymbol{H}$ is positive semidefinite.

Example 3.4 *Unconstrained least-squares (LS) problem*

$$\text{minimize} \quad \tfrac{1}{2}\|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|_2^2$$

where $\boldsymbol{A} \in R^{p \times n}$ with $p > n$. The objective function is quadratic and convex because

$$\tfrac{1}{2}\|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|_2^2 = \tfrac{1}{2}\boldsymbol{x}^T \boldsymbol{A}^T \boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}^T \boldsymbol{A}\boldsymbol{x} + \tfrac{1}{2}\|\boldsymbol{b}\|_2^2$$

If $\boldsymbol{A}$ has full (column) rank, the LS problem has unique and global solution $\boldsymbol{x}^* = \left(\boldsymbol{A}^T \boldsymbol{A}\right)^{-1} \boldsymbol{A}^T \boldsymbol{b}$.

■

Example 3.5 *QP with equality constraints*

Consider convex QP problem

$$\begin{aligned} &\text{minimize} \quad \tfrac{1}{2}\boldsymbol{x}^T \boldsymbol{H}\boldsymbol{x} + \boldsymbol{p}^T \boldsymbol{x} + \kappa \\ &\text{subject to:} \quad \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \end{aligned} \tag{3.10a-b}$$

where $\boldsymbol{A} \in R^{p\times n}$. An effective approach to solve (3.10) is to use the QR decomposition of $\boldsymbol{A}^T$, i.e.,

$$\boldsymbol{A}^T = \boldsymbol{Q}\begin{bmatrix} \boldsymbol{R} \\ \boldsymbol{0} \end{bmatrix} \tag{3.11}$$

where $\boldsymbol{Q}$ is an $n\times n$ orthogonal and $\boldsymbol{R}$ is a $p\times p$ upper triangular matrix (see Appendix). Using (3.11), the constraints in (3.10b) can be expressed as

$$\boldsymbol{R}^T\hat{\boldsymbol{x}}_1 = \boldsymbol{b}$$

where $\hat{\boldsymbol{x}}_1$ is the vector composed of the first p elements of $\hat{\boldsymbol{x}} = \boldsymbol{Q}^T\boldsymbol{x}$. If we denote

$$\hat{\boldsymbol{x}} = \begin{bmatrix} \hat{\boldsymbol{x}}_1 \\ \boldsymbol{\phi} \end{bmatrix} \text{ and } \boldsymbol{Q} = [\boldsymbol{Q}_1 \quad \boldsymbol{Q}_2]$$

with $\boldsymbol{\phi} \in R^{(n-p)\times 1}$, $\boldsymbol{Q}_1 \in R^{n\times p}$, and $\boldsymbol{Q}_2 \in R^{n\times(n-p)}$, then we obtain

$$\boldsymbol{x} = \boldsymbol{Q}_2\boldsymbol{\phi} + \boldsymbol{Q}_1\boldsymbol{R}^{-T}\boldsymbol{b} \tag{3.12}$$

The parameterized solutions in (3.12) can be used to convert the problem in (3.10) into a unconstrained problem which leads to the unique global minimizer of the problem in (3.10) as

$$\boldsymbol{x}^* = \boldsymbol{Q}_2\boldsymbol{\phi}^* + \boldsymbol{Q}_1\boldsymbol{R}^{-T}\boldsymbol{b} \tag{3.13a}$$

where $\boldsymbol{\phi}^*$ is the solution of the linear system

$$\left(\boldsymbol{Q}_2^T\boldsymbol{H}\boldsymbol{Q}_2\right)\boldsymbol{\phi} = -\boldsymbol{Q}_2^T\left(\boldsymbol{H}\boldsymbol{Q}_1\boldsymbol{R}^{-T}\boldsymbol{b} + \boldsymbol{p}\right) \tag{3.13b}$$

Alternatively, problem (3.10) can also be solved by using the first-order necessary conditions described in Theorem 1.1, which are given by

$$\begin{aligned} &\boldsymbol{H}\boldsymbol{x}^* + \boldsymbol{p} + \boldsymbol{A}^T\boldsymbol{\lambda}^* = \boldsymbol{0} \\ &\boldsymbol{A}\boldsymbol{x}^* - \boldsymbol{b} = \boldsymbol{0} \end{aligned}$$

i.e.,

$$\begin{bmatrix} \boldsymbol{H} & \boldsymbol{A}^T \\ \boldsymbol{A} & \boldsymbol{0} \end{bmatrix}\begin{bmatrix} \boldsymbol{x}^* \\ \boldsymbol{\lambda}^* \end{bmatrix} = \begin{bmatrix} -\boldsymbol{p} \\ \boldsymbol{b} \end{bmatrix} \tag{3.14}$$

see Example 1.3. If $\boldsymbol{H}$ is positive definite and $\boldsymbol{A}$ has full row rank, then the system matrix in (3.14) is nonsingular (explain why?) and the solution $\boldsymbol{x}^*$ from (3.14) is the unique global minimizer of the problem in (3.10). Rather than solving linear system (3.14) of size (n + p), the solution $\boldsymbol{x}^*$ and Lagrange multipliers $\boldsymbol{\lambda}^*$ can be expressed as

$$\boldsymbol{\lambda}^* = -(\boldsymbol{A}\boldsymbol{H}^{-1}\boldsymbol{A}^T)^{-1}(\boldsymbol{A}\boldsymbol{H}^{-1}\boldsymbol{p} + \boldsymbol{b}) \tag{3.15a}$$

$$x^* = -H^{-1}(A^T \lambda^* + p) \tag{3.15b}$$

■

Example 3.6 *LP with random cost*

Consider an LP problem of the form

$$\begin{array}{ll} \text{minimize} & \boldsymbol{c}^T \boldsymbol{x} \\ \text{subject to:} & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b},\ \boldsymbol{x} \geq \boldsymbol{0} \end{array}$$

where $\boldsymbol{c}$ is a *random* vector with mean $\overline{\boldsymbol{c}}$ and covariance $\Sigma = E[(\boldsymbol{c}-\overline{\boldsymbol{c}})(\boldsymbol{c}-\overline{\boldsymbol{c}})^T]$. Consequently, the objective $\boldsymbol{c}^T\boldsymbol{x}$ is *random* with mean $\overline{\boldsymbol{c}}^T\boldsymbol{x}$ (representing expected cost) and variance $\boldsymbol{x}^T\Sigma\,\boldsymbol{x}$ (representing risk). A deterministic way to deal with such LP problems is to examine the convex QP problem

$$\begin{array}{ll} \text{minimize} & \overline{\boldsymbol{c}}^T \boldsymbol{x} + \gamma \boldsymbol{x}^T \Sigma\, \boldsymbol{x} \\ \text{subject to:} & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b},\ \boldsymbol{x} \geq \boldsymbol{0} \end{array}$$

instead, where $\gamma > 0$ is a parameter that controls the trade-off between expected cost and variance (risk). ■

3.3 Semidefinite Programming (SDP)

Semidefinite programming (SDP) is a branch of convex programming that has been a subject of intensive research since the early 1990's. The continued interest in SDP has been motivated mainly by two reasons. First, many important classes of optimization problems such as LP and convex QP problems can be viewed as SDP problems, and many CP problems of practical usefulness that are neither LP nor QP problems can also be formulated as SDP problems. Second, several interior-point methods that have proven efficient for LP and convex QP problems have been extended to SDP.

SDP refers to a class of convex problems of the form

$$\begin{array}{ll} \text{minimize} & \boldsymbol{c}^T \boldsymbol{x} \\ \text{subject to:} & F(\boldsymbol{x}) \succeq \boldsymbol{0} \end{array} \tag{3.16a-b}$$

where

$$\boldsymbol{F}(\boldsymbol{x}) = \boldsymbol{F}_0 + \sum_{i=1}^{p} x_i \boldsymbol{F}_i$$

with $\boldsymbol{F}_i \in S^n$ for $0 \leq i \leq p$. Note that the positive semidefinite constraint on matrix $\boldsymbol{F}(\boldsymbol{x})$ in (3.16b) is dependent on vector $\boldsymbol{x}$ in an *affine* manner. In the literature, the type of problems described by (3.16) are often referred to as *convex optimization* problems with *linear matrix inequality* (LMI) constraints, and have found many applications in science and engineering. Since