CHAPTER 16

Quadratic Programming

An optimization problem with a quadratic objective function and linear constraints is called a quadratic program. Problems of this type are important in their own right, and they also arise as subproblems in methods for general constrained optimization, such as sequential quadratic programming (Chapter 18), augmented Lagrangian methods (Chapter 17), and interior-point methods (Chapter 19).

The general quadratic program (QP) can be stated as

$$\min_x \quad q(x) = \tfrac{1}{2}x^T Gx + x^T c \tag{16.1a}$$

$$\text{subject to} \quad a_i^T x = b_i, \qquad i \in \mathcal{E}, \tag{16.1b}$$

$$a_i^T x \geq b_i, \qquad i \in \mathcal{I}, \tag{16.1c}$$

where G is a symmetric $n \times n$ matrix, $\mathcal{E}$ and $\mathcal{I}$ are finite sets of indices, and c, x, and $\{a_i\}, i \in \mathcal{E} \cup \mathcal{I}$, are vectors in $\mathbb{R}^n$. Quadratic programs can always be solved (or shown to be infeasible) in a finite amount of computation, but the effort required to find a solution depends strongly on the characteristics of the objective function and the number of inequality constraints. If the Hessian matrix G is positive semidefinite, we say that (16.1) is a *convex QP*, and in this case the problem is often similar in difficulty to a linear program. (*Strictly convex QPs* are those in which G is positive definite.) *Nonconvex* QPs, in which G is an indefinite matrix, can be more challenging because they can have several stationary points and local minima.

In this chapter we focus primarily on convex quadratic programs. We start by considering an interesting application of quadratic programming.

❑ **EXAMPLE 16.1** (PORTFOLIO OPTIMIZATION)

Every investor knows that there is a tradeoff between risk and return: To increase the expected return on investment, an investor must be willing to tolerate greater risks. Portfolio theory studies how to model this tradeoff given a collection of n possible investments with returns r_i, $i = 1, 2, \ldots, n$. The returns r_i are usually not known in advance and are often assumed to be random variables that follow a normal distribution. We can characterize these variables by their expected value $\mu_i = E[r_i]$ and their variance $\sigma_i^2 = E[(r_i - \mu_i)^2]$. The variance measures the fluctuations of the variable r_i about its mean, so that larger values of σ_i indicate riskier investments. The returns are not in general independent, and we can define correlations between pairs of returns as follows:

$$\rho_{ij} = \frac{E[(r_i - \mu_i)(r_j - \mu_j)]}{\sigma_i \sigma_j}, \quad \text{for } i, j = 1, 2, \ldots, n.$$

The correlation measures the tendency of the return on investments i and j to move in the same direction. Two investments whose returns tend to rise and fall together have a positive correlation; the nearer ρ_{ij} is to 1, the more closely the two investments track each other. Investments whose returns tend to move in opposite directions have a negative correlation.

An investor constructs a portfolio by putting a fraction x_i of the available funds into investment i, for $i = 1, 2, \ldots, n$. Assuming that all available funds are invested and that short-selling is not allowed, the constraints are $\sum_{i=1}^n x_i = 1$ and $x \geq 0$. The return on the portfolio is given by

$$R = \sum_{i=1}^{n} x_i r_i. \tag{16.2}$$

To measure the desirability of the portfolio, we need to obtain measures of its expected return and variance. The expected return is simply

$$E[R] = E\left[\sum_{i=1}^{n} x_i r_i\right] = \sum_{i=1}^{n} x_i E[r_i] = x^T \mu,$$

while the variance is given by

$$\mathrm{Var}[R] = E[(R - E[R])^2] = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i x_j \sigma_i \sigma_j \rho_{ij} = x^T G x,$$

where the $n \times n$ symmetric positive semidefinite matrix G defined by

$$G_{ij} = \rho_{ij}\sigma_i\sigma_j$$

is called the *covariance matrix*.

Ideally, we would like to find a portfolio for which the expected return $x^T\mu$ is large while the variance $x^T G x$ is small. In the model proposed by Markowitz [201], we combine these two aims into a single objective function with the aid of a “risk tolerance parameter” denoted by κ, and we solve the following problem to find the optimal portfolio:

$$\max\ x^T\mu - \kappa x^T G x, \quad \text{subject to} \quad \sum_{i=1}^{n} x_i = 1, \ x \geq 0.$$

The value chosen for the nonnegative parameter κ depends on the preferences of the individual investor. Conservative investors, who place more emphasis on minimizing risk in their portfolio, would choose a large value of κ to increase the weight of the variance measure in the objective function. More daring investors, who are prepared to take on more risk in the hope of a higher expected return, would choose a smaller value of κ.

The difficulty in applying this portfolio optimization technique to real-life investing lies in defining the expected returns, variances, and correlations for the investments in question. Financial professionals often combine historical data with their own insights and expectations to produce values of these quantities. ❐