Introduction to quadratic programming

Ben Rosenberg

January 18, 2026

Overview

- Review: expected value, variance, and covariance
- Motivation
- Problem formulation and canonical form
- Comparison with LP/ILP
- Visualization of objective function and feasible region

Review: expected value, variance, and covariance

We need to go over these so that the portfolio optimization example makes more sense. This is only a brief overview.

The *expected value* of a random variable X is denoted $\mathbb{E}[X]$, and is the average value taken on by X.

For example, if X represents the value of a coin flip, then $\mathbb{E}[X]$ is going to be $\frac{\sum_{i=1}^{6} i}{6} = 3.5$ (because X is a discrete, uniformly distributed random variable with range $[1, \ldots, 6]$).

Review: expected value, variance, and covariance (cont.)

The *variance* of a random variable X is denoted by $\mathrm{Var}(X)$, and is the expected value of the squared difference between X and the expected value of X:

$$\mathrm{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2]$$

Note that if we add a constant to the RV inside an expectation, the effect is linear:

$$\mathbb{E}[aX] = a\mathbb{E}[X]$$

But for variance, the effect is quadratic:

$$\mathrm{Var}(aX) = a^2\mathrm{Var}(X)$$

Review: expected value, variance, and covariance (cont.)

The *covariance* of a random variable X and a random variable Y is denoted by $\text{Cov}(X, Y)$, and appears in the formula for the variance of the sum of two random variables (here, X and Y), as a term to counterbalance the interaction between X and Y:

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y)$$

For the sum of more than two RVs, we just need to add all the pairwise covariances: $\text{Var}(X, Y, Z) =$
$\text{Var}(X)+\text{Var}(Y)+\text{Var}(Z)+2\text{Cov}(X, Y)+2\text{Cov}(X, Z)+2\text{Cov}(Y, Z)$

With constants, we get this general form:

$$\text{Var}(a_1X_1, \ldots, a_nX_n) = \sum_{i=1}^{n} a_i^2\text{Var}(X_i) + 2\sum_{1\le i<j\le n} a_i a_j \text{Cov}(X_i, X_j)$$

Motivation

Sometimes a hyperplanar objective is not enough.

Hyperplanes are straight, but we may want something *quadratic* instead.

We see this in portfolio optimization, where we need to choose assets for a portfolio with an objective that includes their correlation matrix (a "quadratic" term). We will go through this example later.

Example: Convex Optimization (Boyd), p. 153

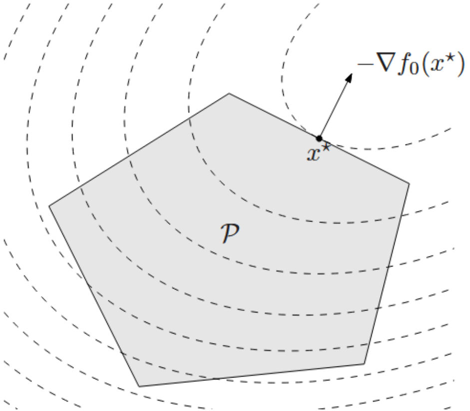

Figure 4.5 Geometric illustration of QP. The feasible set $\mathcal{P}$, which is a polyhedron, is shown shaded. The contour lines of the objective function, which is convex quadratic, are shown as dashed curves. The point $x^\star$ is optimal.

Problem formulation and canonical form

A quadratic program can be expressed in the following form:

$$\begin{aligned} \min \quad & \frac{1}{2} x^{\mathsf{T}} Q x + c^{\mathsf{T}} x \\ s.t. \quad & Ax \leq b \\ & x \geq 0 \end{aligned}$$

The only change from linear programming is the addition of $x^{\mathsf{T}} Q x$ to the objective, where Q is a positive semidefinite matrix (which we will discuss in the next lecture on QP requirements).

Example: portfolio optimization

Suppose you are an investor and you have 1,000 dollars to invest in 3 stocks, where R_i is the return of stock i over a 1-year period. You want to construct a portfolio using these stocks that attains an expected annual return of at least 18%, and minimizes the variance on the return of your portfolio.

You are given the following data for the expected value of R:

$$\mathbb{E}[R] = \langle 0.1, 0.2, 0.3 \rangle$$

This means that we expect stock 1 to turn \$1 into \$1.1, stock 2 to turn \$1 into \$1.2, and stock 3 to turn \$1 into \$1.3, over a 1-year period.

Example: portfolio optimization (cont.)

However, these stocks may not give the exact return that we expect. There is some *variance* to the stocks' returns, $\mathrm{Var}(R_i)$, that we are also given:

$$\mathrm{Var}(R) = \langle 0.12, 0.16, 0.14 \rangle$$

Additionally, these stocks' variances could be *correlated* with each other to some degree: if the market is volatile (e.g., due to the imposition of tariffs or some other event), the magnitude of the reaction in one stock may be similar to the magnitude of the reaction in another stock.

This "variance correlation" of the returns of two stocks is called *covariance*, and is denoted $\mathrm{Cov}(R_i, R_j)$ for stocks i and j. We are given these as well:

$$\mathrm{Cov}(R_1, R_2) = 0.02, \quad \mathrm{Cov}(R_1, R_3) = 0.05, \quad \mathrm{Cov}(R_2, R_3) = 0.04$$

Example: portfolio optimization (cont.)

Denote the amount (out of your \$1,000) that you invest in stock i by x_i.

Then, the annual return for your chosen portfolio will be given by $\sum_i x_i R_i$.

The expected return of your portfolio will then be (by linearity of expectation):

$$\sum_i x_i \mathbb{E}[R_i] = 0.1x_1 + 0.2x_2 + 0.3x_3$$

Example: portfolio optimization (cont.)

The total variance of your portfolio's return is then:

$$\mathrm{Var}(\sum_i x_i R_i) = x_i^2 \mathrm{Var}(R_i) + 2 \sum_{1 \le i < j \le 3} \mathrm{Cov}(R_i, R_j)$$
$$= 0.1x_1^2 + 0.2x_2^2 + 0.3x_3^2 + 0.02 \cdot 2x_1x_2 + 0.05 \cdot 2x_1x_3 + 0.04 \cdot 2x_2x_3$$
$$= 0.1x_1^2 + 0.2x_2^2 + 0.3x_3^2 + 0.04x_1x_2 + 0.1x_1x_3 + 0.08x_2x_3$$

We want to minimize this variance, subject to the constraints on expected return and total budget:

$$\begin{aligned} \min \quad & 0.1x_1^2 + 0.2x_2^2 + 0.3x_3^2 + 0.04x_1x_2 + 0.1x_1x_3 + 0.08x_2x_3 \\ s.t. \quad & 0.1x_1 + 0.2x_2 + 0.3x_3 \ge 1,180 \\ & x_1 + x_2 + x_3 = 1,000 \\ & x_1, x_2, x_3 \ge 0 \end{aligned}$$

Example: portfolio optimization (cont.)

This isn't in the specified form, however. To get this in canonical form, we can take the covariances of the stocks' returns ($\mathrm{Cov}(R_1, R_2) = 0.02, \mathrm{Cov}(R_1, R_3) = 0.05, \mathrm{Cov}(R_2, R_3) = 0.04$) and their variances ($\mathrm{Var}(R_1) = 0.12, \mathrm{Var}(R_2) = 0.16, \mathrm{Var}(R_3) = 0.14$) and put them into a matrix called the *covariance matrix*, where entry i, j is the covariance of the return of stock i with that of stock j. (The covariance of the return of a stock with itself is just its variance.)

We will call this matrix Q, for consistency:

$$Q = \begin{bmatrix} 0.12 & 0.02 & 0.05 \\ 0.02 & 0.16 & 0.04 \\ 0.05 & 0.04 & 0.14 \end{bmatrix}$$

Then, the objective can be written as $x^\intercal Q x$, as given in the QP format.

Comparison with LP/ILP

Quadratic programming is more similar to linear programming than integer linear programming:

Feature	LP	ILP	QP
Feasible region	Convex	Non-convex	Convex
Algorithmic complexity	Polynomial	Exponential	Polynomial
Constraint shape	Hyperplane	Hyperplane	Hyperplane
Objective shape	Hyperplane	Hyperplane	Quadric surface

Differences from LP

Some main differences:

- In QP, the optimum could lie inside of the feasible region, rather than at a corner or boundary as in LP
- Because of this, the simplex algorithm is not usable for QP. Instead, interior point methods are typically used by solvers (among other methods)