16

Portfolio Selection

16.1 Trading Off Expected Return and Risk

How should we invest our wealth? Portfolio theory provides an answer to this question based upon two principles:

- we want to maximize the expected return; and
- we want to minimize the risk, which we define in this chapter to be the standard deviation of the return, though we may ultimately be concerned with the probabilities of large losses.

These goals are somewhat at odds because riskier assets generally have a higher expected return, since investors demand a reward for bearing risk. The difference between the expected return of a risky asset and the risk-free rate of return is called the *risk premium*. Without risk premiums, few investors would invest in risky assets.

Nonetheless, there are optimal compromises between expected return and risk. In this chapter we show how to maximize expected return subject to an upper bound on the risk, or to minimize the risk subject to a lower bound on the expected return. One key concept that we discuss is reduction of risk by diversifying the portfolio.

16.2 One Risky Asset and One Risk-Free Asset

We start with a simple example with one risky asset, which could be a portfolio, for example, a mutual fund. Assume that the expected return is 0.15 and the standard deviation of the return is 0.25. Assume that there is a *risk-free asset*, such as, a 90-day T-bill, and the risk-free rate is 6 %, so the return on the risk-free asset is 6 %, or 0.06. The standard deviation of the return on the

risk-free asset is 0 by definition of "risk-free." The rates and returns here are annual, though all that is necessary is that they be in the same time units.

We are faced with the problem of constructing an investment portfolio that we will hold for one time period, which is called the *holding period* and which could be a day, a month, a quarter, a year, 10 years, and so forth. At the end of the holding period we might want to readjust the portfolio, so for now we are only looking at returns over one time period. Suppose that a fraction w of our wealth is invested in the risky asset and the remaining fraction $1 - w$ is invested in the risk-free asset. Then the expected return is

$$E(R) = w(0.15) + (1 - w)(0.06) = 0.06 + 0.09w, \qquad (16.1)$$

the variance of the return is

$$\sigma_R^2 = w^2 (0.25)^2 + (1 - w)^2 (0)^2 = w^2 (0.25)^2,$$

and the standard deviation of the return is

$$\sigma_R = 0.25 \, |w|. \qquad (16.2)$$

As will be discussed later, w is negative if the risky asset is sold short, so we have $|w|$ rather than w in (16.2).

To decide what proportion w of one's wealth to invest in the risky asset, one chooses either the expected return $E(R)$ one wants or the amount of risk σ_R with which one is willing to live. Once either $E(R)$ or σ_R is chosen, w can be determined.

Although σ is a measure of risk, a more direct measure of risk is actual monetary loss. In the next example, w is chosen to control the maximum size of the loss.

Example 16.1. Finding w to achieve a targeted value-at-risk

Suppose that a firm is planning to invest \$1,000,000 and has capital reserves that could cover a loss of \$150,000 but no more. Therefore, the firm would like to be certain that, if there is a loss, then it is no more than 15 %, that is, that R is greater than -0.15. Suppose that R is normally distributed. Then the only way to guarantee that R is greater than -0.15 with probability equal to 1 is to invest entirely in the risk-free asset. The firm might instead be more modest and require only that $P(R < -0.15)$ be small, for example, 0.01. Therefore, the firm should find the value of w such that

$$P(R < -0.15) = \Phi \left(\frac{-0.15 - (0.06 + 0.09 \, w)}{0.25 \, w} \right) = 0.01.$$

The solution is

$$w = \frac{-0.21}{0.25 \, \Phi^{-1}(0.01) + 0.9} = 0.4264.$$

The value of $\Phi^{-1}(0.01)$ is calculated by `qnorm(0.01)` and is -2.33.

In Chap. 19, \$150,000 is called the value-at-risk (= VaR) and $1 - 0.01 = 0.99$ is called the confidence coefficient. What was done in this example is to find the portfolio that has a VaR of \$150,000 with 0.99 confidence.

We saw in Chap. 5 that the distributions of stock returns usually have much heavier tails than a normal distribution. In Chap. 19, VaR is estimated under more realistic assumptions, e.g., that the returns are t-distributed.   □

More generally, if the expected returns on the risky and risk-free assets are μ_1 and μ_f and if the standard deviation of the risky asset is σ_1, then the expected return on the portfolio is $w\mu_1 + (1 - w)\mu_f$ while the standard deviation of the portfolio's return is $|w|\,\sigma_1$.

This model is simple but not as useless as it might seem at first. As discussed later, finding an optimal portfolio can be achieved in two steps:

1. finding the "optimal" portfolio of risky assets, called the "tangency portfolio," and
2. finding the appropriate mix of the risk-free asset and the tangency portfolio.

So we now know how to do the second step. What we still need to learn is how find the tangency portfolio.

16.2.1 Estimating $E(R)$ and σ_R

The value of the risk-free rate, μ_f, will be known since Treasury bill rates are published in sources providing financial information.

What should we use as the values of $E(R)$ and σ_R? If returns on the asset are assumed to be stationary, then we can take a time series of past returns and use the sample mean and standard deviation. Whether the stationarity assumption is realistic is always debatable. If we think that $E(R)$ and σ_R will be different from the past, we could subjectively adjust these estimates upward or downward according to our opinions, but we must live with the consequences if our opinions prove to be incorrect. Also, the sample mean and standard deviation are not particularly accurate and could be replaced by estimates from a factor model such as the CAPM or the Fama-French model; see Chaps. 17, 18, and 20.

Another question is how long a time series to use, that is, how far back in time one should gather data. A long series, say 10 or 20 years, will give much less variable estimates. However, if the series is not stationary but rather has slowly drifting parameters, then a shorter series (maybe 1 or 2 years) will be more representative of the future. Almost every time series of returns is nearly stationary over short enough time periods.

Even if the time series is stationary, it is likely to exhibit volatility clustering. In that case one might use a GARCH estimate of the conditional standard deviation of the return over the holding period. See Chap. 14.

16.3 Two Risky Assets

16.3.1 Risk Versus Expected Return

The mathematics of mixing risky assets is most easily understood when there are only two risky assets. This is where we start.

Suppose the two risky assets have returns R_1 and R_2 and that we mix them in proportions w and $1 - w$, respectively. The return on the portfolio is $R_p = wR_1 + (1 - w)R_2$. The expected return on the portfolio is $E(R_P) = w\mu_1 + (1 - w)\mu_2$. Let ρ_{12} be the correlation between the returns on the two risky assets. The variance of the return on the portfolio is

$$\sigma_R^2 = w^2\sigma_1^2 + (1 - w)^2\sigma_2^2 + 2w(1 - w)\rho_{12}\,\sigma_1\sigma_2. \qquad (16.3)$$

Note that $\rho_{12}\sigma_1\sigma_2 = \sigma_{R_1,R_2}$.

Example 16.2. The expectation and variance of the return on a portfolio with two risky assets

Suppose that $\mu_1 = 0.14$, $\mu_2 = 0.08$, $\sigma_1 = 0.2$, $\sigma_2 = 0.15$, and $\rho_{12} = 0$. Then

$$E(R_P) = 0.08 + 0.06w,$$

and because $\rho_{12} = 0$ in this example,

$$\sigma_{R_P}^2 = (0.2)^2\,w^2 + (0.15)^2\,(1 - w)^2.$$

Using differential calculus, one can easily show that the portfolio with the minimum risk is $w = 0.045/0.125 = 0.36$. For this portfolio $E(R_P) = 0.08 + (0.06)(0.36) = 0.1016$ and $\sigma_{R_P} = \sqrt{(0.2)^2(0.36)^2 + (0.15)^2(0.64)^2} = 0.12$.

The somewhat parabolic curve[1] in Fig. 16.1 is the locus of values of $(\sigma_R, E(R))$ when $0 \le w \le 1$. The leftmost point on this locus achieves the minimum value of the risk and is called the *minimum variance portfolio*. The points on this locus that have an expected return at least as large as the minimum variance portfolio are called the *efficient frontier*. Portfolios on the efficient frontier are called *efficient portfolios* or, more precisely, *mean-variance efficient portfolios*.[2] The points labeled R_1 and R_2 correspond to $w = 1$ and $w = 0$, respectively. The other features of this figure are explained in Sect. 16.4.   □

---

[1] In fact, the curve would be parabolic if σ_R^2 were plotted on the x-axis instead of σ_R.

[2] When a risk-free asset is available, then the efficient portfolios are no longer those on the efficient frontier but rather are characterized by Result 16.1 ahead.
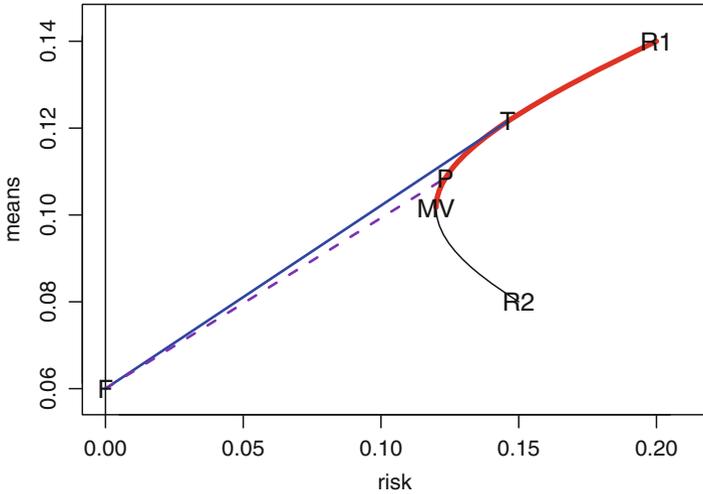
Fig. 16.1. *Expected return versus risk for Example 16.2. F = risk-free asset. T = tangency portfolio. R1 is the first risky asset. R2 is the second risky asset. MV is the minimum variance portfolio. The efficient frontier is the red curve. All points on the curve connecting R2 and R1 are attainable with $0 \leq w \leq 1$, but the ones on the black curve are suboptimal. P is a typical portfolio on the efficient frontier.*

In practice, the mean and standard deviations of the returns can be estimated as discussed in Sect. 16.2.1 and the correlation coefficient can be estimated by the sample correlation coefficient. Alternatively, in Chaps. 18 and 20 factor models are used to estimate expected returns and the covariance matrix of returns.

16.4 Combining Two Risky Assets with a Risk-Free Asset

Our ultimate goal is to find optimal portfolios combining many risky assets with a risk-free asset. However, many of the concepts needed for this task can be first understood most easily when there are only two risky assets.

16.4.1 Tangency Portfolio with Two Risky Assets

As mentioned in Sect. 16.3.1, each point on the efficient frontier in Fig. 16.1 is $(\sigma_{R_P}, E(R_p))$ for some value of w between 0 and 1. If we fix w, then we have a fixed portfolio of the two risky assets. Now let us mix that portfolio of risky assets with the risk-free asset. The point F in Fig. 16.1 gives $(\sigma_{R_P}, E(R))$ for the risk-free asset; of course, $\sigma_{R_P} = 0$ at F. The possible values of $(\sigma_{R_P}, E(R_p))$ for a portfolio consisting of the fixed portfolio of two

risky assets and the risk-free asset is a line connecting the point F with a point on the efficient frontier, for example, the dashed purple line.

Notice that the solid blue line connecting F with the point labeled T lies above the dashed purple line connecting F and the typical portfolio. This means that for any value of σ_{R_P}, the solid blue line gives a higher expected return than the dashed purple line. The slope of each line is called its *Sharpe's ratio*, named after William Sharpe, whom we will meet again in Chap. 17. If $E(R_P)$ and σ_{R_P} are the expected return and standard deviation of the return on a portfolio and μ_f is the risk-free rate, then

$$\frac{E(R_P) - \mu_f}{\sigma_{R_P}} \tag{16.4}$$

is Sharpe's ratio of the portfolio. Sharpe's ratio can be thought of as a "reward-to-risk" ratio. It is the ratio of the reward quantified by the excess expected return[3] to the risk as measured by the standard deviation.

A line with a larger slope gives a higher expected return for a given level of risk, so the larger Sharpe's ratio, the better regardless of what level of risk one is willing to accept. The point T on the efficient frontier is the portfolio with the highest Sharpe's ratio. It is the optimal portfolio for the purpose of mixing with the risk-free asset. This portfolio is called the *tangency portfolio* since its line is tangent to the efficient frontier.

Result 16.1 *The optimal or* efficient *portfolios mix the tangency portfolio with the risk-free asset. Each efficient portfolio has two properties:*

- *it has a higher expected return than any other portfolio with the same or smaller risk, and*
- *it has a smaller risk than any other portfolio with the same or higher expected return.*

Thus we can only improve (reduce) the risk of an efficient portfolio by accepting a worse (smaller) expected return, and we can only improve (increase) the expected return of an efficient portfolio by accepting worse (higher) risk.

Note that all efficient portfolios use the same mix of the two risky assets, namely, the tangency portfolio. Only the proportion allocated to the tangency portfolio and the proportion allocated to the risk-free asset vary.

Given the importance of the tangency portfolio, you may be wondering "how do we find it?" Again, let μ_1, μ_2, and μ_f be the expected returns on the two risky assets and the return on the risk-free asset. Let σ_1 and σ_2 be the standard deviations of the returns on the two risky assets and let ρ_{12} be the correlation between the returns on the risky assets.

---

[3] Here "excess" means in excess of the risk-free rate.

Define $V_1 = \mu_1 - \mu_f$ and $V_2 = \mu_2 - \mu_f$, the excess expected returns. Then the tangency portfolio uses weight

$$w_T = \frac{V_1\sigma_2^2 - V_2\rho_{12}\,\sigma_1\sigma_2}{V_1\sigma_2^2 + V_2\sigma_1^2 - (V_1 + V_2)\rho_{12}\,\sigma_1\sigma_2} \tag{16.5}$$

for the first risky asset and weight $(1 - w_T)$ for the second.

Let R_T, $E(R_T)$, and σ_T be the return, expected return, and standard deviation of the return on the tangency portfolio. Then $E(R_T)$ and σ_T can be found by first finding w_T using (16.5) and then using the formulas

$$E(R_T) = w_T\mu_1 + (1 - w_T)\mu_2$$

and

$$\sigma_T = \sqrt{w_T^2\sigma_1^2 + (1 - w_T)^2\sigma_2^2 + 2w_T(1 - w_T)\rho_{12}\sigma_1\sigma_2}\,.$$

Example 16.3. The tangency portfolio with two risky assets

Suppose as before that $\mu_1 = 0.14$, $\mu_2 = 0.08$, $\sigma_1 = 0.2$, $\sigma_2 = 0.15$, and $\rho_{12} = 0$. Suppose as well that $\mu_f = 0.06$. Then $V_1 = 0.14 - 0.06 = 0.08$ and $V_2 = 0.08 - 0.06 = 0.02$. Plugging these values into formula (16.5), we get $w_T = 0.693$ and $1 - w_t = 0.307$. Therefore,

$$E(R_T) = (0.693)(0.14) + (0.307)(0.08) = 0.122,$$

and

$$\sigma_T = \sqrt{(0.693)^2(0.2)^2 + (0.307)^2(0.15)^2} = 0.146.$$

□

16.4.2 Combining the Tangency Portfolio with the Risk-Free Asset

Let R_p be the return on the portfolio that allocates a fraction ω of the investment to the tangency portfolio and $1 - \omega$ to the risk-free asset. Then $R_p = \omega R_T + (1 - \omega)\mu_f = \mu_f + \omega(R_T - R_f)$, so that

$$E(R_p) = \mu_f + \omega\{E(R_T) - \mu_f\} \quad \text{and} \quad \sigma_{R_p} = \omega\sigma_T.$$

Example 16.4. (Continuation of Example 16.2 and 16.3)

In this example, we will find the optimal investment with $\sigma_{R_p} = 0.05$.

The maximum expected return with $\sigma_{R_p} = 0.05$ mixes the tangency portfolio and the risk-free asset such that $\sigma_{R_p} = 0.05$. Since $\sigma_T = 0.146$, we have that $0.05 = \sigma_{R_p} = \omega\,\sigma_T = 0.146\,\omega$, so that $\omega = 0.05/0.146 = 0.343$ and $1 - \omega = 0.657$.

So 65.7 % of the portfolio should be in the risk-free asset, and 34.3 % should be in the tangency portfolio. Thus $(0.343)(69.3\%) = 23.7\%$ should be in the first risky asset and $(0.343)(30.7\%) = 10.5\%$ should be in the second risky asset. The total is not quite 100 % because of rounding.   □

Example 16.5. (Continuation of Examples 16.2–16.4)

Now suppose that you want a 10 % expected return. In this example we will compare

- the best portfolio of only risky assets, and
- The best portfolio of the risky assets and the risk-free asset.

The best portfolio of only risky assets uses w solving $0.1 = w(0.14)+(1-w)(0.08)$, which implies that $w = 1/3$. This is the *only* portfolio of risky assets with $E(R_p) = 0.1$, so by default it is best. Then

$$\sigma_{R_P} = \sqrt{w^2(0.2)^2 + (1-w)^2(0.15)^2} = \sqrt{(1/9)(0.2)^2 + 4/9(0.15)^2} = 0.120.$$

The best portfolio of the two risky assets and the risk-free asset can be found as follows. First, $0.1 = E(R) = \mu_f + \omega\{E(R_T) - \mu_f\} = 0.06 + 0.062\,\omega = 0.06 + 0.425\,\sigma_R$, since $\sigma_{R_P} = \omega\,\sigma_T$ or $\omega = \sigma_{R_P}/\sigma_T = \sigma_{R_P}/0.146$. This implies that $\sigma_{R_P} = 0.04/0.425 = 0.094$ and $\omega = 0.04/0.062 = 0.645$. So combining the risk-free asset with the two risky assets reduces σ_{R_P} from 0.120 to 0.094 while maintaining $E(R_p)$ at 0.1. The reduction in risk is $(0.120 - 0.094)/0.094 = 28\%$, which is substantial.   □

16.4.3 Effect of ρ_{12}

Positive correlation between the two risky assets increases risk. With positive correlation, the two assets tend to move together which increases the volatility of the portfolio. Conversely, negative correlation is beneficial since it decreases risk. If the assets are negatively correlated, a negative return of one tends to occur with a positive return of the other so the volatility of the portfolio decreases. Figure 16.2 shows the efficient frontier and tangency portfolio when $\mu_1 = 0.14$, $\mu_2 = 0.09$, $\sigma_1 = 0.2$, $\sigma_2 = 0.15$, and $\mu_f = 0.03$. The value of ρ_{12} is varied from 0.5 to −0.7. Notice that Sharpe's ratio of the tangency portfolio returns increases as ρ_{12} decreases. This means that when ρ_{12} is small, then efficient portfolios have less risk for a given expected return compared to when ρ_{12} is large.
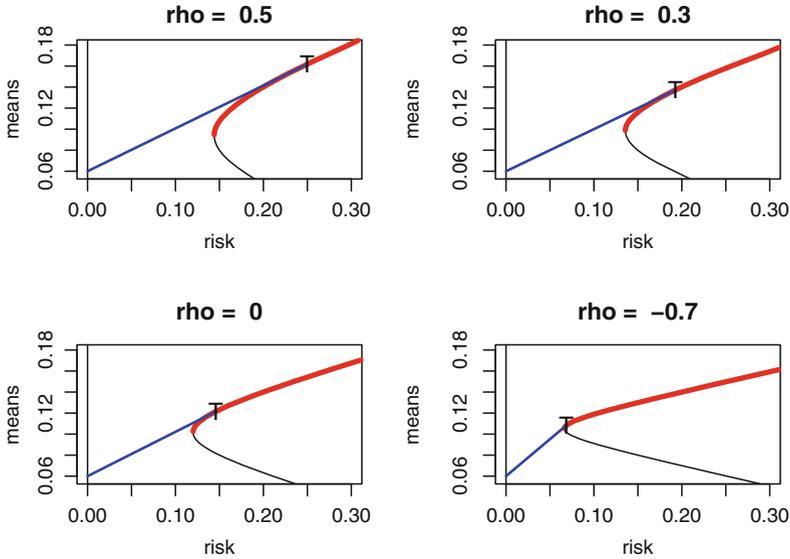
Fig. 16.2. *Efficient frontier (red) and tangency portfolio (T) when* $\mu_1 = 0.14$, $\mu_2 = 0.09$, $\sigma_1 = 0.2$, $\sigma_2 = 0.15$, *and* $\mu_f = 0.03$. *The value of* ρ_{12} *is varied from 0.5 to* -0.7.

16.5 Selling Short

Often some of the weights in an efficient portfolio are negative. A negative weight on an asset means that this asset is sold short. *Selling short* is a way to profit if a stock price goes *down*. To sell a stock short, one sells the stock without owning it. The stock must be borrowed from a broker or another customer of the broker. At a later point in time, one buys the stock and gives it back to the lender. This closes the short position.

Suppose a stock is selling at \$25/share and you sell 100 shares short. This gives you \$2500. If the stock goes down to \$17/share, you can buy the 100 shares for \$1700 and close out your short position. You made a profit of \$800 (ignoring transaction costs) because the stock went down 8 points. If the stock had gone up, then you would have had a loss.

Suppose now that you have \$100 and there are two risky assets. With your money you could buy \$150 worth of risky asset 1 and sell \$50 short of risky asset 2. The net cost would be exactly \$100. If R_1 and R_2 are the returns on risky assets 1 and 2, then the return on your portfolio would be

$$\frac{3}{2}R_1 + \left(-\frac{1}{2}\right)R_2.$$

Your portfolio weights are $w_1 = 3/2$ and $w_2 = -1/2$. Thus, you hope that risky asset 1 rises in price and risky asset 2 falls in price. Here, again, we have ignored transaction costs.

If one sells a stock short, one is said to have a *short position* in that stock, and owning the stock is called a *long position*.

16.6 Risk-Efficient Portfolios with N Risky Assets

In this section, we use quadratic programming to find efficient portfolios with an arbitrary number of assets. An advantage of quadratic programming is that it allows one to impose constraints such as limiting short sales. With no constraints on the allocation vector $\boldsymbol{w}$, analytic formulas for the tangency portfolio can be derived using Lagrange multipliers, but this approach does not generalize to constrained $\boldsymbol{w}$.

Assume that we have N risky assets and that the return on the ith risky asset is R_i and has expected value μ_i. Define

$$\boldsymbol{R} = \begin{pmatrix} R_1 \\ \vdots \\ R_N \end{pmatrix}$$

to be the random vector of returns,

$$E(\boldsymbol{R}) = \boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_N \end{pmatrix},$$

and $\boldsymbol{\Sigma}$ to be the covariance matrix of $\boldsymbol{R}$.

Let

$$\boldsymbol{w} = \begin{pmatrix} w_1 \\ \vdots \\ w_N \end{pmatrix}$$

be a vector of portfolio weights so that $w_1 + \cdots + w_N = \boldsymbol{1}^{\mathsf{T}} \boldsymbol{w} = 1$, where

$$\boldsymbol{1} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$

is a column of N ones. The expected return on the portfolio is

$$\sum_{i=1}^{N} w_i \mu_i = \boldsymbol{\omega}^{\mathsf{T}} \boldsymbol{\mu}. \tag{16.6}$$

Suppose there is a target value, μ_P, of the expected return on the portfolio. When $N = 2$, the target expected returns is achieved by only one portfolio and its w_1-value solves $\mu_P = w_1 \mu_1 + w_2 \mu_2 = \mu_2 + w_1(\mu_1 - \mu_2)$. For $N \geq 3$,

there will be an infinite number of portfolios achieving the target μ_P. The one with the smallest variance is called the "efficient" portfolio. Our goal is to find the efficient portfolio.

The variance of the return on the portfolio with weights $\boldsymbol{w}$ is

$$\boldsymbol{w}^{\mathsf{T}} \boldsymbol{\Sigma} \boldsymbol{w}. \tag{16.7}$$

Thus, given a target μ_P, the efficient portfolio minimizes (16.7) subject to

$$\boldsymbol{w}^{\mathsf{T}} \boldsymbol{\mu} = \mu_P \tag{16.8}$$

and

$$\boldsymbol{w}^{\mathsf{T}} \boldsymbol{1} = 1. \tag{16.9}$$

Quadratic programming is used to minimize a quadratic objective function subject to linear constraints. In applications to portfolio optimization, the objective function is the variance of the portfolio return. The objective function is a function of N variables, such as the weights of N assets, that are denoted by an $N \times 1$ vector $\boldsymbol{x}$. Suppose that the quadratic objective function to be minimized is

$$\frac{1}{2}\boldsymbol{x}^{\mathsf{T}} \boldsymbol{D} \boldsymbol{x} - \boldsymbol{d}^{\mathsf{T}} \boldsymbol{x}, \tag{16.10}$$

where $\boldsymbol{D}$ is an $N \times N$ matrix and $\boldsymbol{d}$ is an $N \times 1$ vector. The factor of $1/2$ is not essential but is used here to keep our notation consistent with R. There are two types of linear constraints on $\boldsymbol{x}$, inequality and equality constraints. The linear inequality constraints are

$$\boldsymbol{A}_{\mathrm{neq}}^{\mathsf{T}} \boldsymbol{x} \geq \boldsymbol{b}_{\mathrm{neq}}, \tag{16.11}$$

where $\boldsymbol{A}_{\mathrm{neq}}$ is an $m \times N$ matrix, $\boldsymbol{b}_{\mathrm{neq}}$ is an $m \times 1$ vector, and m is the number of inequality constraints. The equality constraints are

$$\boldsymbol{A}_{\mathrm{eq}}^{\mathsf{T}} \boldsymbol{x} = \boldsymbol{b}_{\mathrm{eq}}, \tag{16.12}$$

where $\boldsymbol{A}_{\mathrm{eq}}$ is an $n \times N$ matrix, $\boldsymbol{b}_{\mathrm{eq}}$ is an $n \times 1$ vector, and n is the number of equality constraints. Quadratic programming minimizes the quadratic objective function (16.10) subject to linear inequality constraints (16.11) and linear equality constraints (16.12).

To apply quadratic programming to find an efficient portfolio, we use $\boldsymbol{x} = \boldsymbol{w}$, $\boldsymbol{D} = 2\boldsymbol{\Sigma}$, and $\boldsymbol{d}$ equal to an $N \times 1$ vector of zeros so that (16.10) is $\boldsymbol{w}^{\mathsf{T}} \boldsymbol{\Sigma} \boldsymbol{w}$, the return variance of the portfolio. There are two equality constraints, one that the weights sum to 1 and the other that the portfolio return is a specified target μ_P. Therefore, we define

$$\boldsymbol{A}_{\mathrm{eq}}^{\mathsf{T}} = \begin{pmatrix} \boldsymbol{1}^{\mathsf{T}} \\ \boldsymbol{\mu}^{\mathsf{T}} \end{pmatrix}$$

and

$$b_{\text{eq}} = \begin{pmatrix} 1 \\ \mu_P \end{pmatrix},$$

so that (16.12) becomes

$$\begin{pmatrix} \mathbf{1}^\mathsf{T} w \\ \boldsymbol{\mu}^\mathsf{T} w \end{pmatrix} = \begin{pmatrix} 1 \\ \mu_P \end{pmatrix},$$

which is the same as constraints (16.8) and (16.9). So far, inequality constraints have not been used.

Investors often wish to impose additional inequality constraints. If an investor cannot or does not wish to sell short, then the constraint

$$w \geq \mathbf{0}$$

can be used. Here $\mathbf{0}$ is a vector of N zeros. In this case A_{neq} is the $N \times N$ identical matrix and $b_{\text{neq}} = \mathbf{0}$.

To avoid concentrating the portfolio in just one or a few stocks, an investor may wish to constrain the portfolio so that no w_i exceeds a bound λ, for example, $\lambda = 1/4$ means that no more than $1/4$ of the portfolio can be in any single stock. In this case, $w \leq \lambda \mathbf{1}$ or equivalently $-w \geq -\lambda \mathbf{1}$, so that A_{neq} is minus the $N \times N$ identity matrix and $b_{\text{neq}} = -\lambda \mathbf{1}$. One can combine these constraints with those that prohibit short selling.

To find the efficient frontier, one uses a grid of values of μ_P and finds the corresponding efficient portfolios. For each portfolio, σ_P^2, which is the minimized value of the objective function, can be calculated. Then one can find the minimum variance portfolio by finding the portfolio with the smallest value of the σ_P^2. The efficient frontier is the set of efficient portfolios with expected return above the expected return of the minimum variance portfolio. One can also compute Sharpe's ratio for each portfolio on the efficient frontier and the tangency portfolio is the one maximizing Sharpe's ratio.

Example 16.6. Finding the efficient frontier, tangency portfolio, and minimum variance portfolio using quadratic programming

The following R program uses the returns on three stocks, GE, IBM, and Mobil, in the CRSPday data set in the Ecdat package. The function solve.QP() in the quadprog package is used for quadratic programming. solve.QP() combines A_{eq}^T and $A_{\text{neq}}^\mathsf{T}$ into a single matrix Amat by stacking A_{eq}^T on top of $A_{\text{neq}}^\mathsf{T}$. The parameter meq is the number of rows of A_{eq}^T. b_{eq} and b_{neq} are handled analogously. In this example, there are no inequality constraints, so $A_{\text{neq}}^\mathsf{T}$ and b_{neq} are not needed, but they are used in the next example.

The efficient portfolio is found for each of 300 target values of μ_P between 0.05 and 0.14. For each portfolio, Sharpe's ratio is found at line 28 and the

logical vector `ind` at line 29 indicates which portfolio is the tangency portfolio maximizing Sharpe's ratio. Similarly, `ind2` at line 34 indicates the minimum variance portfolio. Also, `ind3` at line 36 indicates the points on the efficient frontier. It is assumed that the risk-free rate is 1.3 %/year; see line 26.

```
1  llibrary(Ecdat)
2  library(quadprog)
3  data(CRSPday)
4  R = 100*CRSPday[ ,4:6]  #  convert to percentages
5  mean_vect = apply(R, 2 ,mean)
6  cov_mat = cov(R)
7  sd_vect = sqrt(diag(cov_mat))
8  Amat = cbind(rep(1, 3), mean_vect)  # set the constraints matrix
9  muP = seq(0.05, 0.14, length = 300)  # target portfolio means
10 # for the expect portfolio return
11 sdP = muP # set up storage for std dev's of portfolio returns
12 weights = matrix(0, nrow = 300, ncol = 3) # storage for weights
13 for (i in 1:length(muP))  # find the optimal portfolios
14 {
15   bvec = c(1, muP[i])  # constraint vector
16   result =
17     solve.QP(Dmat = 2 * cov_mat, dvec = rep(0, 3),
18       Amat = Amat, bvec = bvec, meq = 2)
19   sdP[i] = sqrt(result$value)
20   weights[i,] = result$solution
21 }
22 pdf("quad_prog_plot.pdf", width = 6, height = 5)
23 plot(sdP, muP, type = "l", xlim = c(0, 2.5),
24    ylim = c(0, 0.15), lty = 3)  # plot efficient frontier (and
25              # inefficient portfolios below the min var portfolio)
26 mufree = 1.3 / 253 # input value of risk-free interest rate
27 points(0, mufree, cex = 4, pch = "*")  # show risk-free asset
28 sharpe = (muP - mufree) / sdP # compute Sharpe's ratios
29 ind = (sharpe == max(sharpe)) # Find maximum Sharpe's ratio
30 weights[ind, ] #  print the weights of the tangency portfolio
31 lines(c(0, 2), mufree + c(0, 2) * (muP[ind] - mufree) / sdP[ind],
32    lwd = 4, lty = 1, col = "blue")  # show line of optimal portfolios
33 points(sdP[ind], muP[ind], cex = 4, pch = "*") #  tangency portfolio
34 ind2 = (sdP == min(sdP)) # find minimum variance portfolio
35 points(sdP[ind2], muP[ind2], cex = 2, pch = "+") # min var portfolio
36 ind3 = (muP > muP[ind2])
37 lines(sdP[ind3], muP[ind3], type = "l", xlim = c(0, 0.25),
38    ylim = c(0, 0.3), lwd = 3, col = "red")  # plot efficient frontier
39 text(sd_vect[1], mean_vect[1], "GE", cex = 1.15)
40 text(sd_vect[2], mean_vect[2], "IBM", cex = 1.15)
41 text(sd_vect[3], mean_vect[3], "Mobil", cex = 1.15)
42 graphics.off()
```
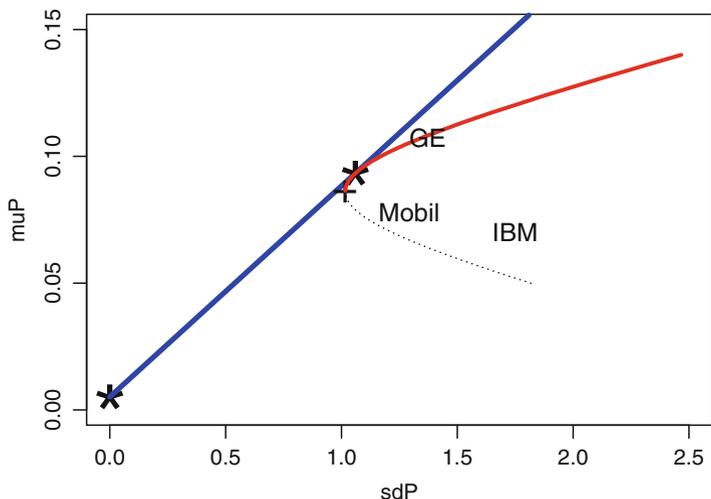
Fig. 16.3. *Efficient frontier (solid), line of efficient portfolios (dashed) connecting the risk-free asset and tangency portfolio (asterisks), and the minimum variance portfolio (plus) with three stocks (GE, IBM, and Mobil). The three stocks are also shown on reward-risk space.*

The plot produced by this program is Fig. 16.3. The program prints the weights of the tangency portfolio, which are

```
> weights[ind,] # Find tangency portfolio
[1] 0.5512 0.0844 0.3645
```

□

Example 16.7. Finding the efficient frontier, tangency portfolio, and minimum variance portfolio with no short selling using quadratic programming

In this example, Example 16.6 is modified so that short sales are not allowed. Only three lines of code need to be changed. When short sales are prohibited, the target expected return on the portfolio must lie between the smallest and largest expected returns on the stocks. To prevent numerical errors, the target expected returns will start 0.0001 above the smallest expected stock return and end 0.0001 below the largest expected stock return. This is enforced by the following change:

```
muP = seq(min(mean_vect) + 0.0001, max(mean_vect) - 0.0001,
    length = 300)
```

To enforce no short sales, an $\boldsymbol{A}_{\mathrm{neq}}$ matrix is needed and is set equal to a 3×3 identity matrix:

```
Amat = cbind(rep(1, 3), mean_vect, diag(1, nrow = 3))
```

Also, $\boldsymbol{b}_{\mathrm{neq}}$ is set equal to a three-dimensional vector of zeros:

```
bvec = c(1, muP[i], rep(0, 3))
```

The new plot is shown in Fig. 16.4. Since the tangency portfolio in Example 16.6 had all weights positive, the tangency portfolio is unchanged by the prohibition of short sales. The efficient frontier is changed since without short sales, it is impossible to have expected returns greater than the expected return of GE, the stock with the highest expected return. In contrast, when short sales are allowed, there is no upper bound on the expected return (or on the risk). In Fig. 16.4 the red curve is the entire efficient frontier, but in Fig. 16.3 the efficient frontier is the red curve extended to $(+\infty, +\infty)$.   □

16.7 Resampling and Efficient Portfolios

The theory of portfolio optimization assumes that the expected returns and the covariance matrix of the returns is known. In practice, one must replace these quantities with estimates as in the previous examples. However, the effects of estimation error, especially with smaller values of N, can result in portfolios that only appear efficient. This problem will be investigated in this section using the bootstrap to quantify the effects of estimation error.

Example 16.8. The global asset allocation problem

One application of optimal portfolio selection is allocation of capital to different market segments. For example, Michaud (1998) discusses a global asset allocation problem where capital must be allocated to "U.S. stocks and government/corporate bonds, euros, and the Canadian, French, German, Japanese, and U.K. equity markets." Here we look at a similar example where we allocate capital to the equity markets of 10 different countries. Monthly returns for these markets were calculated from MSCI Hong Kong, MSCI Singapore, MSCI Brazil, MSCI Argentina, MSCI UK, MSCI Germany, MSCI Canada, MSCI France, MSCI Japan, and the S&P 500. "MSCI" means "Morgan Stanley Capital Index." The data are from January 1988 to January 2002, inclusive, so there are 169 months of data.

Assume that we want to find the tangency portfolio that maximizes Sharpe's ratio. The tangency portfolio was estimated using sample means and the sample covariance as in Example 16.6, and its Sharpe's ratio is estimated to be 0.3681. However, we should suspect that 0.3681 must be an overestimate since this portfolio only maximizes Sharpe's ratio using estimated parameters,
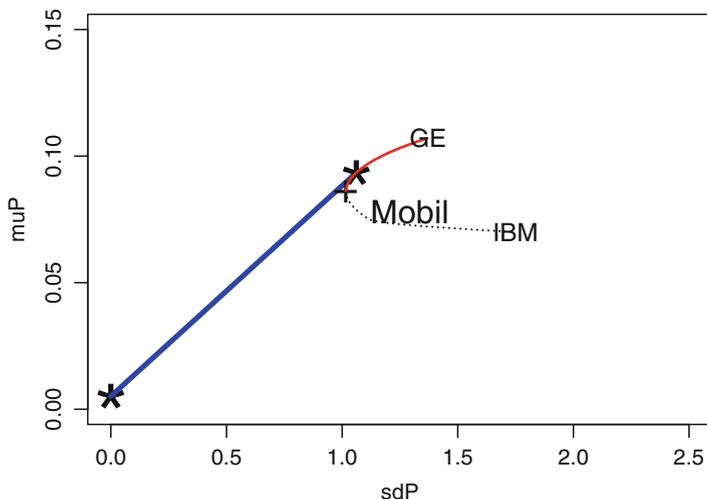
Fig. 16.4. *Efficient frontier (solid), line of efficient portfolios (dashed) connecting the risk-free asset and tangency portfolio (asterisks), and the minimum variance portfolio (plus) with three stocks (GE, IBM, and Mobil) with short sales prohibited.*

not the true means and covariance matrix. To evaluate the possible amount of overestimation, one can use the bootstrap. As discussed in Chap. 6, in the bootstrap simulation experiment, the sample is the "true population" so that the sample mean and covariance matrix are the "true parameters," and the resamples mimic the sampling process. Actual Sharpe's ratios are calculated with the sample means and covariance matrix, while estimated Sharpe's ratio use the means and covariance matrix of the resamples.

First, 250 resamples were taken and for each the tangency portfolio was estimated. Resampling was done by sampling rows of the data matrix as discussed in Sect. 7.11. For each of the 250 tangency portfolios estimated from the resamples, the actual and estimated Sharpe's ratios were calculated. Boxplots of the 250 actual and 250 estimated Sharpe's ratios of the estimated tangency portfolios are in Fig. 16.5a. "Estimated" means calculated from the resample and "true" means calculated from the sample. In this figure, there is a dashed horizontal line at height 0.3681, the actual Sharpe's ratio of the true tangency portfolio. One can see that all 250 estimated tangency portfolios have actual Sharpe's ratios below this value, as they must since the actual Sharpe's ratio is maximized by the true tangency portfolio, not the estimated tangency portfolios.

From the boxplot on the right-hand side of (a), one can see that the estimated Sharpe's ratios overestimate not only the actual Sharpe's ratios of the estimated tangency portfolios but also the somewhat larger (and unattainable) actual Sharpe's ratio of the true (but unknowable) tangency portfolio. □
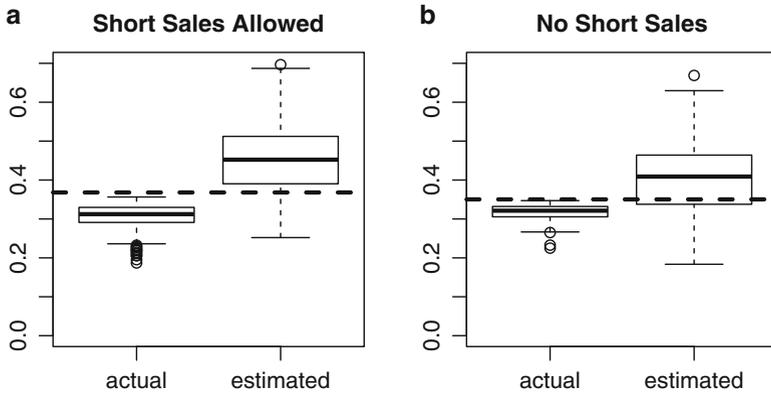
Fig. 16.5. *Bootstrapping estimation of the tangency portfolio and its Sharpe's ratio. (a) Short sales allowed. The left-hand boxplot is of the actual Sharpe's ratios of the estimated tangency portfolios for 250 resamples. The right-hand boxplot contains the estimated Sharpe's ratios for these portfolios. The horizontal dashed line indicates Sharpe's ratio of the true tangency portfolio. (b) Same as (a) but with short sales not allowed.*

There are several ways to alleviate the problems caused by estimation error when attempting to find a tangency portfolio. One can try to find more accurate estimators; the factor models of Chap. 18 and Bayes estimators of Chap. 20 (see especially Example 20.12) do this. Another possibility is to restrict short sales.

Portfolios with short sales aggressively attempt to maximize Sharpe's ratio by selling short those stocks with the smallest estimated mean returns and having large long positions in those stocks with the highest estimated mean returns. The weakness with this approach is that it is particularly sensitive to estimation error. Unfortunately, expected returns are estimated with relatively large uncertainty. This problem can be seen in Table 16.1, which has 95 % confidence intervals for the mean returns. The percentile method is used for the confidence intervals, so the endpoints are the 2.5 and 97.5 bootstrap percentiles. Notice for Singapore and Japan, the confidence intervals include both positive and negative values. In the table, the returns are expressed as percentage returns.

Example 16.9. The global asset allocation problem: short sales prohibited

This example repeats the bootstrap experimentation of Example 16.8 with short sales prohibited by using inequality constraints such as in Example 16.7. With short sales not allowed, the actual Sharpe's ratio of the true tangency portfolio is 0.3503, which is only slightly less than when short sales are allowed.

Table 16.1. *95 % percentile-method bootstrap confidence intervals for the mean returns of the* 10 *countries.*

Country	2.5 %	97.5 %
Hong Kong	0.186	2.709
Singapore	−0.229	2.003
Brazil	0.232	5.136
Argentina	0.196	6.548
UK	0.071	1.530
Germany	0.120	1.769
Canada	0.062	1.580
France	0.243	2.028
Japan	−0.884	0.874
U.S.	0.636	1.690

Boxplots of actual and apparent Sharpe's ratios are in Fig. 16.5b. Comparing Fig. 16.5a and b, one sees that prohibiting short sales has two beneficial effects—Sharpe's ratios actually achieved are slightly higher with no short sales allowed compared to having no constraints on short sales. In fact, the mean of the 250 actual Sharpe's ratios is 0.3060 with short sales allowed and 0.3169 with short sales prohibited. Moreover, the overestimation of Sharpe's ratio is reduced by prohibiting short sales—the mean apparent Sharpe's ratio is 0.4524 [with estimation error $(0.4524 − 0.3681) = 0.0843$] with short sales allowed but only 0.4038 [with estimation error $(0.4038 − 0.3503) = 0.0535$] with short sales prohibited. However, these effects, though positive, are only modest and do not entirely solve the problem of overestimation of Sharpe's ratio.   □

Example 16.10. The global asset allocation problem: Shrinkage estimation and short sales prohibited

In Example 16.9, we saw that prohibiting short sales can increase Sharpe's ratio of the estimated tangency portfolio, but the improvement is only modest. Further improvement requires more accurate estimation of the mean vector or the covariance matrix of the returns.

This example investigates possible improvements from shrinking the 10 estimated means toward each other. Specifically, if $\overline{Y}_i$ is the sample mean of the ith country, $\overline{Y} = (\sum_{i=1}^{10} \overline{Y}_i)/10$ is the grand mean (mean of the means), and α is a tuning parameter between 0 and 1, then the estimated mean return for the ith country is

$$\widehat{\mu}_i = \alpha\overline{Y}_i + (1 - \alpha)\overline{Y}. \tag{16.13}$$

The purpose of shrinkage is to reduce the variance of the estimator, though the reduced variance comes at the expense of some bias. Since it is the mean of

10 means, $\overline{Y}$ is much less variable than any of $\overline{Y}_1, \ldots, \overline{Y}_{10}$. Therefore, $\mathrm{Var}(\widehat{\mu}_i)$ decreases as α is decreased toward 0. However,

$$E(\widehat{\mu}_i) = \alpha\mu_i + \frac{1-\alpha}{10}\sum_{i=1}^{10}\mu_i \qquad (16.14)$$

so that, for any $\alpha \neq 1$, $\widehat{\mu}_i$ is biased, except under the very likely circumstance that $\mu_1 = \cdots = \mu_{10}$. The parameter α controls the bias–variance tradeoff. In this example, $\alpha = 1/2$ will be used for illustration and short sales will not be allowed.
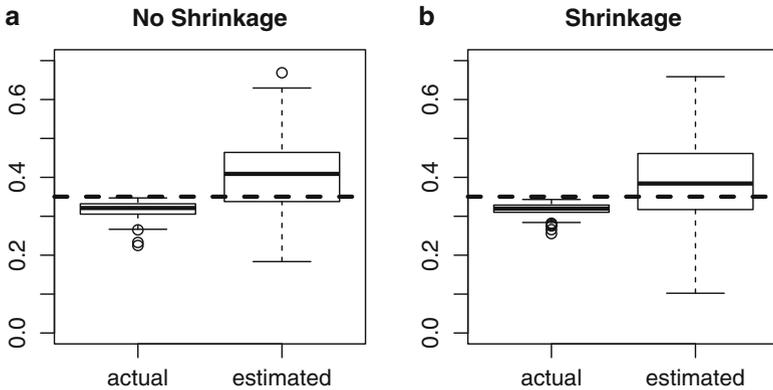

Fig. 16.6. *Bootstrapping estimation of the tangency portfolio and its Sharpe's ratio. Short sales not allowed. (a) No shrinkage. The left-hand boxplot is of the actual Sharpe's ratios of the estimated tangency portfolios for 250 resamples. The right-hand boxplot contains the estimated Sharpe's ratios for these portfolios. The horizontal dashed line indicates Sharpe's ratio of the true tangency portfolio. (b) Same as (a) but with shrinkage.*

Figure 16.6 compares the performance of shrinkage versus no shrinkage. Panel (a) contains the boxplots that we saw in panel (b) of Fig. 16.5 where $\alpha = 1$. Panel (b) has the boxplots when the tangency portfolio is estimated using $\alpha = 1/2$. Compared to panel (a), in panel (b) the actual Sharpe's ratios are somewhat closer to the dashed line indicating Sharpe's ratio of the true tangency portfolio; the means of the actual Sharpe's ratios are 0.317 and 0.318 with and without shrinkage, respectively. These values should be compared with the Sharpe ratio of the true (but unknown) tangency portfolio of 0.34.

Moreover, the estimated Sharpe's ratios in (b) are smaller and closer to the true Sharpe's ratios, so there is less overoptimization—shrinkage has helped in two ways. The mean estimated Sharpe's ratios are 0.390 and 0.404 with and without shrinkage.

The next step might be selection of α to optimize performance of shrinkage estimation. Doing this need not be difficult, since different values of α can be compared by bootstrapping.   □

There are other methods for improving the estimation of the mean vector and estimation of the covariance matrix can be improved as well, for example, by using the factor models in Chap. 18 or Bayesian estimation as in Chap. 20. Moreover, one need not focus on the tangency portfolio but could, for example, estimate the minimum variance portfolio. Whatever the focus of estimation, the bootstrap can be used to compare various strategies for improving the estimation of the optimal portfolio.

16.8 Utility

Economists generally do not model economic decisions in terms of the mean and variance of the return but rather by using a *utility function*. The utility of an amount X of money is said to be $U(X)$ where the utility function U generally has the properties:

1. $U(0) = 0$;
2. U is strictly increasing;
3. the first derivative $U'(X)$ is strictly decreasing.

Assumption 1 is not necessary but is reasonable and states that the utility of 0 dollars is 0. Assumption 2 merely states that more money is better than less. Assumption 3 implies that the more money we have the less we value an extra dollar and is called *risk aversion*. Assumption 3 implies that we would decline a bet that pays $\pm\Delta$ with equal probabilities. In fact, Assumption 3 implies that we would decline any bet with a payoff that is symmetrically distributed about 0, because the expected utility of our wealth would be reduced if we accepted the bet. Mathematically, Assumption 3 implies that U is strictly concave. If the second derivative U'' exists then Assumption 3 is equivalent to the assumption that $U''(X) < 0$ for all X.

It is assumed that a rational person will make investment decisions so as to maximize

$$E\{U(X)\} = E[U\{X_0(1 + R)\}] \tag{16.15}$$

where X is that person's final wealth, X_0 is the person's initial wealth, and R is the return from the investments. In economics this is almost a part of the definition of a rational person, with another component of the definition being that a rational person will update probabilities using Bayes' law (see Chap. 20). Each individual is assumed to have his or her own utility function and two different rational people may make different decisions because they have different utility functions.

How different are mean-variance efficient portfolios and portfolios that maximize expected utility? In the case that returns are normally distributed, this question can be answered.

Result 16.2 *If returns on all portfolios are normally distributed and if U satisfies Assumptions 3, then the portfolio that maximizes expected utility is on the efficient frontier.*

So, if one chose a portfolio to maximize expected utility, then a mean-variance efficient portfolio would be selected. Exactly which portfolio on the efficient frontier one chooses would depend on one's utility function.

Proof of Result 16.2

This result can be proven by proving the following fact: if R_1 and R_2 are normally distributed with the same means and with standard deviations σ_1 and σ_2 such that $\sigma_1 < \sigma_2$, then $E\{U(R_1)\} > E\{U(R_2)\}$. We will show that this follows from Jensen's inequality which states that if U is concave function and X is any random variable, then $E\{U(X)\} \leq U\{E(X)\}$. The inequality is strict if U is strictly convex and X is nondegenerate.[4]

Let $X = R_1 + e$ where e is independent of R_1 and normally distributed with mean 0 and variance $\sigma_2^2 - \sigma_1^2$. Then X has the same distribution as R_2, e is nondegenerate, and, using the law of iterated expectations and then Jensen's inequality and Assumption 3, we have

$$E\{U(R_2)\} = E\{U(X)\} = E[E\{U(X)|R_1\}] < E[\{U\{E(X|R_1)\}\}] = E\{U(R_1)\}, \tag{16.16}$$

since $E(X|R_1) = R_1$.   □

The assumption in Result 16.2 that the returns are normally distributed can be weakened to the more realistic assumption that the vector of returns on the assets is a multivariate scale mixture, e.g., has a multivariate t-distribution. To prove this extension, one conditions on the mixing variable so that the returns have a conditional multivariate normal distribution. Then (16.16) holds conditionally for all values of the mixing variable and therefore holds unconditionally.

A common class of utility functions is

$$U(x; \lambda) = 1 - \exp(-\lambda x), \tag{16.17}$$

where $\lambda > 0$ determines the amount of risk aversion. Note that $U'(x; \lambda) = \lambda \exp(\lambda x)$ and $U''(x; \lambda) = -\lambda^2 \exp(-\lambda x)$ (differentiation is with respect to x).

---

[4] Jensen's inequality is usually stated for convex functions with the inequality reversed. If U is concave then $-U$ is convex so that the two forms of Jensen's inequality are equivalent. A random variable X is degenerate if there is a constant a such that $P(X = a) = 1$. Otherwise, it is nondegenerate.
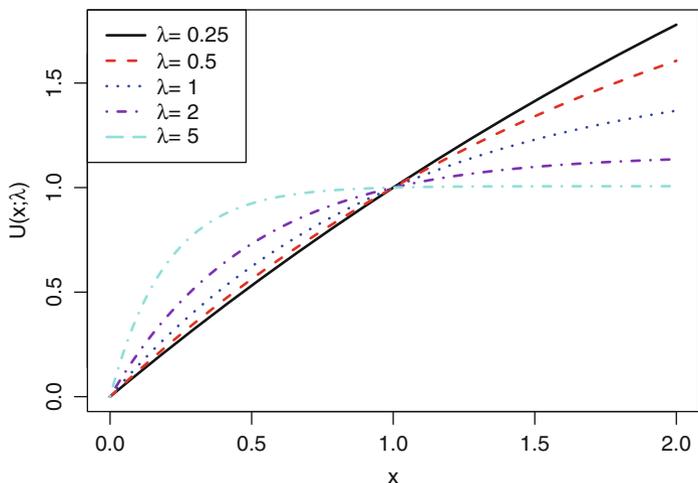
Fig. 16.7. *The utility functions $\widetilde{U}(x; \lambda) = U(x; \lambda)/U(0, \lambda)$ where $U(x; \lambda) = 1 - \exp(-\lambda x)$.*

Thus, $U''(x; \lambda)$ is negative for all x so Assumption 3 is met; it is easy to see that Assumptions 1 and 2 also hold. As $x \to \infty$, $U(x : \lambda) \to 1$.

Multiplying a utility function by a positive constant will not affect which decision maximizes utility and can standardize utility functions to make them more comparable. In Fig. 16.7, $\widetilde{U}(x; \lambda) := U(x; \lambda)/U(0; \lambda)$ is plotted for λ = 0.25, 0.5, 1, 2 and 5. Since $\widetilde{U}(1; \lambda) = 1$ for all λ, these utility functions have been standardized so that the utility corresponding to a return of 0 is always 1. Stated differently, a return equal to 0 has the same utility for all degrees of risk aversion.

Adding a constant to the utility function does not effect the optimal decision, so one could work with the slightly simpler utility function $- \exp(-\lambda x)$ instead of $U(x; \lambda)$ or $\widetilde{U}(x; \lambda)$.

When the utility function is given by (16.17) and R is normally distributed, then (16.15) becomes

$$1 - \exp\left[-\lambda X_0 \{1 + E(R)\} + (\lambda X_0)^2 \frac{\text{var}(R)}{2} \right] \tag{16.18}$$

by properties of the lognormal distribution; see Appendix A.9.4. For given values of λ and X_0, the expected utility is maximized by maximizing

$$E(R) - (\lambda X_0) \frac{\text{var}(R)}{2}. \tag{16.19}$$

Therefore, using the notation of Sect. 16.6, one selects the allocation vector $\boldsymbol{w}$ of the portfolio to maximize

$$\boldsymbol{w}^{\mathsf{T}}\boldsymbol{\mu} - (\lambda X_0)\frac{\boldsymbol{w}^{\mathsf{T}}\boldsymbol{\Sigma}\boldsymbol{w}}{2} \tag{16.20}$$

subject to $\boldsymbol{w}^{\mathsf{T}}\mathbf{1} = 1$.

Maximizing (16.20) subject to linear constraints is a quadratic programming problem. As $\lambda \to 0$, the expected return and standard deviation of the return converge to ∞. Conversely, as $\lambda \to \infty$, the solution converges to the minimum variance portfolio. Therefore, as λ is varied from ∞ to 0, one finds all of the portfolios on the efficient frontier from left to right. This behavior is illustrated in the next example; see Fig. 16.8.

Example 16.11. Finding portfolios the maximize expected utility

We will use stock price data in the file Stock_Bond.csv. This data set was discussed in Sect. 2.4.1. For simplicity of notation, we subsume X_0 into λ. The R code below solves the quadratic program (16.20) for 250 values of $\log(\lambda)$ equally spaced from 2 to 8; this range was selected by trial-and-error.

```
1  library(quadprog)
2  dat = read.csv("Stock_Bond.csv")
3  y = dat[, c(3, 5, 7, 9, 11, 13, 15, 17, 19, 21)]
4  n = dim(y)[1]
5  m = dim(y)[2] - 1
6  r = y[-1,] / y[-n,] - 1
7  mean_vect = as.matrix(colMeans(r))
8  cov_mat = cov(r)
9  nlambda = 250
10 loglambda_vect = seq(2, 8, length = nlambda)
11 w_matrix = matrix(nrow = nlambda, ncol = 10)
12 mu_vect = matrix(nrow = nlambda, ncol = 1)
13 sd_vect = mu_vect
14 ExUtil_vect = mu_vect
15 conv_vect = mu_vect
16 for (i in 1:nlambda)
17 {
18    lambda = exp(loglambda_vect[i])
19    opt = solve.QP(Dmat = as.matrix(lambda^2 * cov_mat),
20        dvec = lambda * mean_vect, Amat = as.matrix(rep(1,10)),
21        bvec = 1, meq = 1)
22    w = opt$solution
23    mu_vect[i] = w %*% mean_vect
24    sd_vect[i] = sqrt(w %*% cov_mat %*% w)
25    w_matrix[i,] = w
26    ExUtil_vect[i] = opt$value
27 }
```

Next, the expected return and the standard deviation of the return are plotted against λ and then the efficient frontier is drawn by plotting the expect return again the standard deviation of the return. The plots are in Fig. 16.8. □
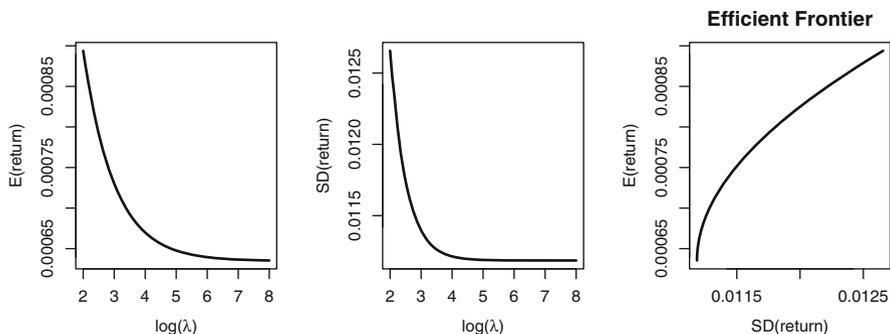
Fig. 16.8. *The expected portfolio return versus* $\log(\lambda$ *(left), the standard deviation of the return versus* $\log(\lambda)$ *(center) and the efficient frontier (right).*

16.9 Bibliographic Notes

Markowitz (1952) was the original paper on portfolio theory and was expanded into the book Markowitz (1959). Bodie and Merton (2000) provide an elementary introduction to portfolio selection theory. Bodie, Kane, and Marcus (1999) and Sharpe, Alexander, and Bailey (1999) give a more comprehensive treatment. See also Merton (1972). Formula (16.5) is derived in Example 5.10 of Ruppert (2004).

Jobson and Korkie (1980) and Britten-Jones (1999) discuss the statistical issue of estimating the efficient frontier; see the latter for additional recent references. Britten-Jones (1999) shows that the tangency portfolio can be estimated by regression analysis and hypotheses about the tangency portfolio can be tested by regression F-tests. Jagannathan and Ma (2003) discuss how imposing constraints such as no short sales can reduce risk.

16.10 R Lab

16.10.1 Efficient Equity Portfolios

This section uses daily stock prices in the data set Stock Bond.csv that is posted on the book's website and in which any variable whose name ends with "AC" is an adjusted closing price. As the name suggests, these prices have been adjusted for dividends and stock splits, so that returns can be calculated without further adjustments. Run the following code which will read the data, compute the returns for six stocks, create a scatterplot matrix of these returns, and compute the mean vector, covariance matrix, and vector of standard deviations of the returns. Note that returns will be percentages.

```
dat = read.csv("Stock_Bond.csv", header = T)
prices = cbind(dat$GM_AC, dat$F_AC, dat$CAT_AC, dat$UTX_AC,
   dat$MRK_AC, dat$IBM_AC)
n = dim(prices)[1]
returns =  100 * (prices[2:n, ] / prices[1:(n-1), ] - 1)
pairs(returns)
mean_vect = colMeans(returns)
cov_mat = cov(returns)
sd_vect = sqrt(diag(cov_mat))
```

Problem 1 *Write an R program to find the efficient frontier, the tangency portfolio, and the minimum variance portfolio, and plot on "reward-risk space" the location of each of the six stocks, the efficient frontier, the tangency portfolio, and the line of efficient portfolios. Use the constraints that $-0.1 \leq w_j \leq 0.5$ for each stock. The first constraint limits short sales but does not rule them out completely. The second constraint prohibits more than 50 % of the investment in any single stock. Assume that the annual risk-free rate is 3 % and convert this to a daily rate by dividing by 365, since interest is earned on trading as well as nontrading days.*

Problem 2 *If an investor wants an efficient portfolio with an expected daily return of 0.07 %, how should the investor allocate his or her capital to the six stocks and to the risk-free asset? Assume that the investor wishes to use the tangency portfolio computed with the constraints $-0.1 \leq w_j \leq 0.5$, not the unconstrained tangency portfolio.*

Problem 3 *Does this data set include Black Monday?*

16.10.2 Efficient Portfolios with Apple, Exxon-Mobil, Target, and McDonald's Stock

This section constructs portfolios with stocks from four companies: Apple Inc. (AAPL), Exxon-Mobil (XOM), Target Corp. (TGT), and McDonalds (MCD). Run the following code to get 2013 returns in terms of percentage for each of the 4 companies:

```
dat = read.csv("FourStocks_Daily2013.csv", header = TRUE)
head(dat)
prices = dat[,-1]
n = dim(prices)[1]
returns = 100*(prices[-1,] / prices[-n,] - 1)
```

Problem 4 *Write an* R *program to plot the efficient frontier and to find the allocation weight vector $\boldsymbol{w}$ corresponding to the tangency portfolio. Use the sample mean vector and sample covariance matrix of the returns to estimate $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$. Assume that the annual risk free rate is 1.3 %. Use the constraints that no w_j can be less than -0.5 or greater than 0.5. Let μ_P range from 0.045 to 0.06 %. Report both the Sharpe's Ratio and $\boldsymbol{w}$ for the tangency portfolio.*

Problem 5 *Write an* R *program to minimize*

$$\boldsymbol{w}^{\mathsf{T}}\boldsymbol{\mu} - \lambda\frac{\boldsymbol{w}^{\mathsf{T}}\boldsymbol{\Sigma}\boldsymbol{w}}{2} \tag{16.21}$$

over $\boldsymbol{w}$, subject to $\boldsymbol{w}^{\mathsf{T}}\boldsymbol{1} = 1$, for each λ on a log-spaced grid. Plot the expected return and standard deviation of the return for the portfolios found this way and show that the curve coincides with the efficient frontier found in Problem 4. Select the range of the grid of log-λ values by trial and error to cover an interesting range of the efficient frontier. What value of λ yields a portfolio with $\mu_P = 0.046$? What value of λ yields to the tangency portfolio? What value of λ yields to the minimum variance portfolio?

16.10.3 Finding the Set of Possible Expected Returns

In Sect. 16.6 when we found the efficient frontier by quadratic programming, it was necessary to set up a grid of possible values of the expected returns on the portfolios. When there are no constraints on the allocation vector $\boldsymbol{w}$ except that its elements sum to 1, any expected return is feasible.[5] We saw in Example 16.7, that if short sales are prohibited by the constraints $0 \le w_i \le 1$ for all i, the the feasible expected portfolio returns lie between the smallest and largest expected returns on the individual assets.

When more complex constraints are placed on the w_i, the set of feasible expected portfolio returns can be found by linear programming. In this section, we use the same data as used in Sect. 16.10.1. We will impose the constraints that $w_i \le$ B1 and $-$B2 $\le w_i$ for all i.

The function `solveLP()` in the `linprog` package minimizes (or maximizes) over N-dimensional variable $\boldsymbol{x}$ the objection function $\boldsymbol{c}^{\mathsf{T}}\boldsymbol{x}$ subject to $\boldsymbol{A}\boldsymbol{x} \le \boldsymbol{b}$ and $\boldsymbol{x} \ge \boldsymbol{0}$. Here $\boldsymbol{c}$ is an $N \times 1$ constant vector, $\boldsymbol{A}$ is an $N \times k$ constant matrix, and $\boldsymbol{b}$ is a $k \times 1$ constant vector for some integers N and k. Also, $\boldsymbol{0}$ is a k-dimensional zero vector.

Since $\boldsymbol{x} \ge \boldsymbol{1}$ we cannot let $\boldsymbol{w}$ be $\boldsymbol{x}$ unless we are prohibiting short sale. When short sales are allowed, we can instead let $\boldsymbol{w}$ equal $\boldsymbol{x}_1 - \boldsymbol{x}_2$ and $\boldsymbol{x}^{\mathsf{T}} = (\boldsymbol{x}_1^{\mathsf{T}}, \boldsymbol{x}_2^{\mathsf{T}})$. Then the constraints are that each element of $\boldsymbol{x}_1$ is at most B1 and each element of $\boldsymbol{x}_2$ is at most B2. The objective function $\boldsymbol{w}^{\mathsf{T}}\boldsymbol{\mu}$ is equal to $(\boldsymbol{\mu}^{\mathsf{T}}, -\boldsymbol{\mu}^{\mathsf{T}})\boldsymbol{x}$.

---

[5] "Feasible" means that there exists a vector $\boldsymbol{w}$ achieving that expected return.

The constraints in $\boldsymbol{Ax} \leq \boldsymbol{b}$ can be a mixture of equality and inequality constraints. The argument `const.dir` specifies the directions of the constraints; see line 17. In the program below, there are $2M + 1$ constraints where M is the number of assets. The first M constraints are that $w_i \leq$ B1 for all i, the next M constraints are that $-$B2 $\leq w_i$ for all i, and the last constraint is that $\boldsymbol{w}^\mathsf{T}\boldsymbol{1} = 1$.

The function `solveLP()` is used twice, once at lines 18 and 19 to find the smallest feasible expected portfolio return and then at lines 20 and 21 to find the largest possible expected return.

```
1  dat = read.csv("Stock_Bond.csv", header = T)
2  prices = cbind(dat$GM_AC, dat$F_AC, dat$CAT_AC, dat$UTX_AC,
3    dat$MRK_AC, dat$IBM_AC)
4  n = dim(prices)[1]
5  returns =  100 * (prices[2:n, ] / prices[1:(n-1), ] - 1)
6  mean_vect = colMeans(returns)
7  M = length(mean_vect)
8  B1 = 0.3
9  B2 = 0.1
10 library(linprog)
11 AmatLP1 = cbind(diag(1, nrow = M), matrix(0, nrow = M, ncol = M))
12 AmatLP2 = cbind(matrix(0, nrow = M, ncol = M), diag(1, nrow = M))
13 AmatLP3 = c(rep(1, M), rep(-1, M))
14 AmatLP = rbind(AmatLP1, AmatLP2, AmatLP3)
15 bvecLP = c(rep(B1, M), rep(B2, M), 1)
16 cLP =  c(mean_vect, -mean_vect)
17 const.dir = c(rep("<=", 2 * M), "=")
18 resultLP_min = solveLP(cvec = cLP, bvec = bvecLP, Amat = AmatLP,
19   lpSolve=T, const.dir = const.dir, maximum = FALSE)
20 resultLP_max = solveLP(cvec = cLP, bvec = bvecLP,
21   Amat = AmatLP, lpSolve = TRUE, maximum = TRUE)
```

Problem 6 *What is the set of feasible expected portfolio returns when $-0.1 \leq w_i \leq 0.3$ for all i? What allocation vector $\boldsymbol{w}$ achieve the smallest possible expected portfolio return? What allocation vector $\boldsymbol{w}$ achieve the largest possible expected portfolio return?*

Problem 7 *Would it be possible to use* B1$=$ *0.15 and* B2 $=$ *0.15? Explain your answer.*

16.11 Exercises

1. Suppose that there are two risky assets, A and B, with expected returns equal to 2.3 % and 4.5 %, respectively. Suppose that the standard deviations of the returns are $\sqrt{6}$ % and $\sqrt{11}$ % and that the returns on the assets have a correlation of 0.17.

  (a) What portfolio of A and B achieves a 3 % rate of expected return?

  (b) What portfolios of A and B achieve a $\sqrt{5.5}\,\%$ standard deviation of return? Among these, which has the largest expected return?

2. Suppose there are two risky assets, C and D, the tangency portfolio is 65 % C and 35 % D, and the expected return and standard deviation of the return on the tangency portfolio are 5 % and 7 %, respectively. Suppose also that the risk-free rate of return is 1.5 %. If you want the standard deviation of your return to be 5 %, what proportions of your capital should be in the risk-free asset, asset C, and asset D?

3. (a) Suppose that stock A shares sell at \$75 and stock B shares at \$115. A portfolio has 300 shares of stock A and 100 of stock B. What are the weights w and $1 - w$ of stocks A and B in this portfolio?

  (b) More generally, if a portfolio has N stocks, if the price per share of the jth stock is P_j, and if the portfolio has n_j shares of stock j, then find a formula for w_j as a function of $n_1, \ldots, n_N$ and $P_1, \ldots, P_N$.

4. Let $\mathcal{R}_P$ be a return of some type on a portfolio and let $\mathcal{R}_1, \ldots, \mathcal{R}_N$ be the same type of returns on the assets in this portfolio. Is

$$\mathcal{R}_P = w_1 \mathcal{R}_1 + \cdots + w_N \mathcal{R}_N$$

true if $\mathcal{R}_P$ is a net return? Is this equation true if $\mathcal{R}_P$ is a gross return? Is it true if $\mathcal{R}_P$ is a log return? Justify your answers.

5. Suppose one has a sample of monthly log returns on two stocks with sample means of 0.0032 and 0.0074, sample variances of 0.017 and 0.025, and a sample covariance of 0.0059. For purposes of resampling, consider these to be the "true population values." A bootstrap resample has sample means of 0.0047 and 0.0065, sample variances of 0.0125 and 0.023, and a sample covariance of 0.0058.

  (a) Using the resample, estimate the efficient portfolio of these two stocks that has an expected return of 0.005; that is, give the two portfolio weights.

  (b) What is the estimated variance of the return of the portfolio in part (a) using the resample variances and covariances?

  (c) What are the actual expected return and variance of return for the portfolio in (a) when calculated with the true population values (e.g., with using the original sample means, variances, and covariance)?

6. Stocks 1 and 2 are selling for \$100 and \$125, respectively. You own 200 shares of stock 1 and 100 shares of stock 2. The weekly returns on these stocks have means of 0.001 and 0.0015, respectively, and standard deviations of 0.03 and 0.04, respectively. Their weekly returns have a correlation of 0.35. Find the correlation matrix of the weekly returns on the two stocks and the mean and standard deviation of the weekly returns on the portfolio.

References

Bodie, Z., and Merton, R. C. (2000) *Finance*, Prentice-Hall, Upper Saddle River, NJ.

Bodie, Z., Kane, A., and Marcus, A. (1999) *Investments*, 4th ed., Irwin/ McGraw-Hill, Boston.

Britten-Jones, M. (1999) The sampling error in estimates of mean-variance efficient portfolio weights. *Journal of Finance*, **54**, 655–671.

Jagannathan, R. and Ma, T. (2003) Risk reduction in large portfolios: Why imposing the wrong constraints helps. *Journal of Finance*, **58**, 1651–1683.

Jobson, J. D., and Korkie, B. (1980) Estimation for Markowitz efficient portfolios. *Journal of the American Statistical Association*, **75**, 544–554.

Markowitz, H. (1952) Portfolio Selection. *Journal of Finance*, **7**, 77–91.

Markowitz, H. (1959) *Portfolio Selection: Efficient Diversification of Investment*, Wiley, New York.

Merton, R. C. (1972) An analytic derivation of the efficient portfolio frontier. *Journal of Financial and Quantitative Analysis*, **7**, 1851–1872.

Michaud, R. O. (1998) *Efficient Asset Management: A Practical Guide to Stock Portfolio Optimization and Asset Allocation*, Harvard Business School Press, Boston.

Ruppert, D. (2004) *Statistics and Finance: An Introduction*, Springer, New York.

Sharpe, W. F., Alexander, G. J., and Bailey, J. V. (1999) *Investments*, 6th ed., Prentice-Hall, Upper Saddle River, NJ.