and the τ-day log return rate as

$$r_{\log}(t:\tau) = \log\left(\frac{S(t+1)}{S(t)}\frac{S(t+2)}{S(t+1)}\cdots\frac{S(t+\tau-1)}{S(t+\tau-2)}\frac{S(t+\tau)}{S(t+\tau-1)}\right)$$

$$= \underbrace{\sum_{\Delta=0}^{\tau-1}\log\left(\frac{S(t+\Delta+1)}{S(t+\Delta)}\right)}_{\text{sum of future 1-day log returns}} \tag{5.10}$$

$$= \sum_{\Delta=0}^{\tau-1} r_{\log}(t+\Delta:1).$$

The multi-period log formula (5.10) is an appealing one; if we are able to model the evolution of the daily log return then the model for the multi-period return can easily be constructed. We shall return to this property later in the book.

In Summary

There is no universal agreement regarding which measure of return should be used, both have their strengths and weaknesses. We collect together our main observations.

- **Standard**
  The standard return possesses the linearity property (5.5) and this makes it a good candidate for portfolio analysis. The perceived drawback of the standard return is that we have to be careful when making assumptions regarding its distribution function; it is bounded below by -1 and this rules out many of the popular choices such as the normal distribution.
- **Logarithmic**
  The log return does not possess the linearity property, although an approximate linear relationship exists over a small holding period. One of the most interesting properties of the log return is that a longer-period return can be expressed as a sum of future daily returns. Under the right conditions this property can be exploited; insight into the evolution of daily returns can be used so solve problems involving multi-period returns.
  Unlike the standard return, the log return can theoretically take any value on $\mathbb{R}$ and so can easily be fit to a whole host of popular probability distributions. This fact is used by many academics and practitioners who aim to model the way a stock price evolves through time.

5.2 SETTING UP THE OPTIMAL PORTFOLIO PROBLEM

A financial portfolio is a fundamental investment, it is manufactured from a collection of basic financial assets and its composition depends upon the investors' preferences and requirements. In mathematical terms our intention is to build a portfolio from a set of n assets denoted by $\{S_1,\ldots,S_n\}$. Given that we intend to hold our portfolio for a total τ-days, the corresponding standard return rates for this period are given by

$$\{r_1,\ldots,r_n\} \quad \text{where} \quad r_i = \frac{S_i(t+\tau)-S_i(t)}{S_i(t)}, \quad i=1,\ldots,n,$$

and thus, for a collection of portfolio weights $\{w_1, \ldots, w_n\}$ that satisfy

$$\sum_{i=1}^{n} w_i = 1, \tag{5.11}$$

the corresponding τ-day portfolio return is given by

$$r_p = \sum_{i=1}^{n} w_i r_i. \tag{5.12}$$

We can employ vector notation and write

$$\mathbf{r} = \begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix}, \mathbf{w} = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} \text{ and } \mathbf{1} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}, \tag{5.13}$$

then equations (5.11) and (5.12) can be expressed as

$$r_p = \mathbf{w}^T\mathbf{r} \quad \text{and} \quad \mathbf{w}^T\mathbf{1} = 1 \quad \text{respectively.}$$

We remark that some of the entries of the weight vector $\mathbf{w}$ may be negative; this corresponds to a short-selling strategy for the assets in question.

The portfolio is completely determined by its portfolio weights: the potential reward and the (unavoidable) risk attached to a given portfolio can be altered by a modification of the portfolio weights. Our aim is to find the weights that somehow provide the optimal balance between risk and expected reward.

- **Reward**
  To measure the potential reward we can use the expected return rate on the portfolio given by

$$\mu_p = \mathbb{E}[r_p] = \sum_{i=1}^{n} w_i \mathbb{E}[r_i] = \sum_{i=1}^{n} w_i \mu_i.$$

  In matrix vector form we write

$$\mu_p = \mathbf{w}^T\mathbf{e} \text{ where } \mathbf{e} = \mathbb{E}[\mathbf{r}] = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix} \tag{5.14}$$

- **Risk**
  The variance of the portfolio return serves as a measure of portfolio risk. To compute this we need the covariance information of all the return rates. We recall from (3.16) that the covariance matrix can be expressed as

$$\mathbf{V} = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \ldots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \ldots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \ldots & \sigma_{nn} \end{pmatrix} = \mathbb{E}[(\mathbf{r} - \mathbf{e})(\mathbf{r} - \mathbf{e})^T].$$

The variance of the portfolio with weight vector $\mathbf{w}$ is then given by

$$\sigma_p^2 = \mathbf{w}^T \mathbf{V} \mathbf{w} = \sum_{i=1}^{n} \sum_{j=1}^{n} w_i w_j \sigma_{ij}. \tag{5.15}$$

Every portfolio weight vector (whose component weights sum to one) defines a feasible portfolio and for each feasible portfolio we can compute its expected return μ_p and its portfolio volatility σ_p. In order to investigate the way these values depend upon the portfolio weights we can display the information in (σ, μ)-space, i.e., a space where we can plot the coordinates (σ_p, μ_p) for a given feasible portfolio p. We begin by performing a simple experiment with the following steps:

1. Generate a feasible weight vector $\mathbf{w}_p \in \mathbb{R}^n$.
2. Plug $\mathbf{w}_p$ into equations (5.14) and (5.15) to generate the expected return μ_p and volatility σ_p for the feasible portfolio p.
3. Plot the point (σ_p, μ_p) in (σ, μ)-space and repeat many times for many different weight vectors.

As the above experiment evolves we begin to see a distinct area of (σ, μ)-space that is occupied by feasible portfolios, this area is bullet shaped and appears to have a very definite boundary (see Figure 5.2). If we take a closer look at Figure 5.2 we see that we have randomly generated four portfolios, each with different feasible weights, for which the expected return is 7%. If, as investors, we are happy with 7% as a potential rate of return then we would always choose the portfolio with the least risk. In a nutshell we want to be on the perceived boundary or frontier of the diagram. We want to achieve the 7% with minimum risk!

Our task can now be defined. Let's suppose our investor is aiming to achieve a return on his portfolio of $100 \times \mu\%$, then in order to do this feasibly (i.e., with a portfolio whose

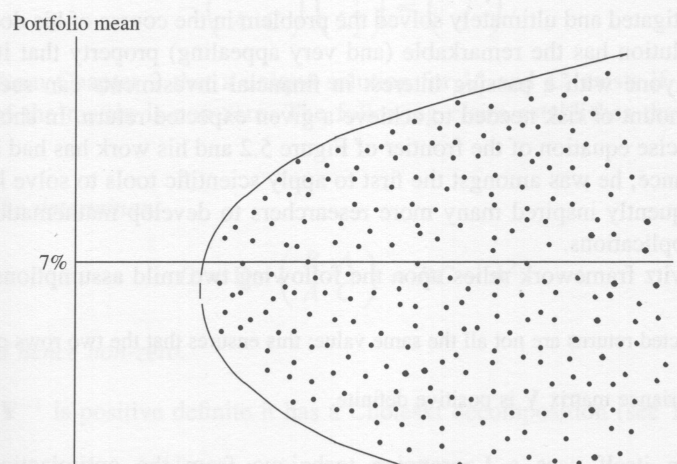

Figure 5.2 Filling (σ, μ)-space with feasible portfolios.

weights sum to one) and with the least risk we must find the portfolio weight vector that solves the following optimization problem:

$$\begin{aligned} &\text{minimize } \frac{1}{2}\mathbf{w}^T\mathbf{V}\mathbf{w} \\ &\text{subject to } \mathbf{w}^T\mathbf{e} = \mu_p \quad \text{expected return matches desired rate} \\ &\qquad\qquad \mathbf{w}^T\mathbf{1} = 1 \quad \text{achieved with a feasible portfolio.} \end{aligned} \tag{5.16}$$

We remark that scaling the variance by a factor of $1/2$ does not affect the location of the optimal solution; it is merely to ensure a cleaner mathematical solution.

We can write the two constraints in matrix form as

$$\mathbf{P}\mathbf{w} = \begin{pmatrix} \mu_1 & \mu_2 & \dots & \mu_n \\ 1 & 1 & \dots & 1 \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} = \begin{pmatrix} \mu \\ 1 \end{pmatrix} = \mathbf{d},$$

and so our optimal portfolio problem can be stated neatly as

$$\text{minimize } \frac{1}{2}\mathbf{w}^T\mathbf{V}\mathbf{w} \quad \text{subject to} \quad \mathbf{P}\mathbf{w} = \mathbf{d}. \tag{5.17}$$

5.3 SOLVING THE OPTIMAL PORTFOLIO PROBLEM

The mathematical formulation of the optimal portfolio problem (5.17) dates back to the 1950s and is attributed to Harry Markowitz who, as a PhD student at the University of Chicago, investigated and ultimately solved the problem in the course of his doctoral studies. Markowitz' solution has the remarkable (and very appealing) property that it can be visualized; thus anyone with a passing interest in financial investments can see, at a glance, the smallest amount of risk needed to achieve a given expected return. In short, Markowitz derived the precise equation of the frontier of Figure 5.2 and his work has had a huge impact on modern finance; he was amongst the first to apply scientific tools to solve hard problems and has subsequently inspired many more researchers to develop mathematical techniques for financial applications.

The Markowitz framework relies upon the following two mild assumptions:

- **A1**. The expected returns are not all the same value; this ensures that the two rows of **P** are linearly independent.
- **A2**. The covariance matrix **V** is positive definite.

The solution itself uses a Lagrangian technique from the optimization toolbox of Chapter 4; specifically we seek the solution $(\mathbf{w}_p, \boldsymbol{\lambda}^*)^T \in \mathbb{R}^{n+2}$ to the linear system

$$\begin{pmatrix} \mathbf{V} & \mathbf{P}^T \\ \mathbf{P} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{w}_p \\ \boldsymbol{\lambda}^* \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{d} \end{pmatrix}$$

i.e.,

$$\mathbf{V}\mathbf{w}_p + \mathbf{P}^T\boldsymbol{\lambda}^* = \mathbf{0} \;\Rightarrow\; \mathbf{V}\mathbf{w}_p = -\lambda_1^*\mathbf{e} - \lambda_2^*\mathbf{1}, \tag{5.18}$$

where $\boldsymbol{\lambda}^* = (\lambda_1^*, \lambda_2^*)^T$ is the vector of Lagrange multipliers and $\mathbf{w}_p$ is the optimal weight vector which satisfies

$$\mathbf{P}\mathbf{w}_p = \mathbf{d} \Rightarrow \begin{cases} \mathbf{e}^T\mathbf{w}_p = \mu_p, \\ \mathbf{1}^T\mathbf{w}_p = 1. \end{cases} \tag{5.19}$$

Since $\mathbf{V}$ is positive definite it is invertible, thus we can solve (5.18) to give

$$\mathbf{w}_p = -\lambda_1^*\mathbf{V}^{-1}\mathbf{e} - \lambda_2^*\mathbf{V}^{-1}\mathbf{1}. \tag{5.20}$$

We can then substitute this expression into the constraint equations (5.19) to give

$$\begin{aligned} -(\mathbf{e}^T\mathbf{V}^{-1}\mathbf{e})\lambda_1^* - (\mathbf{e}^T\mathbf{V}^{-1}\mathbf{1})\lambda_2^* &= \mu_p, \\ -(\mathbf{1}^T\mathbf{V}^{-1}\mathbf{e})\lambda_1^* - (\mathbf{1}^T\mathbf{V}^{-1}\mathbf{1})\lambda_2^* &= 1. \end{aligned} \tag{5.21}$$

For a cleaner presentation we follow the approach of Huand and Litzenberger (1988) and set

$$\begin{pmatrix} B & A \\ A & C \end{pmatrix} = \begin{pmatrix} \mathbf{e}^T\mathbf{V}^{-1}\mathbf{e} & \mathbf{e}^T\mathbf{V}^{-1}\mathbf{1} \\ \mathbf{1}^T\mathbf{V}^{-1}\mathbf{e} & \mathbf{1}^T\mathbf{V}^{-1}\mathbf{1} \end{pmatrix}, \tag{5.22}$$

then the constraint equations can be rewritten in matrix form as

$$-\begin{pmatrix} B & A \\ A & C \end{pmatrix}\begin{pmatrix} \lambda_1^* \\ \lambda_2^* \end{pmatrix} = \begin{pmatrix} \mu_p \\ 1 \end{pmatrix}. \tag{5.23}$$

We know from Chapter 2 that a unique solution for λ_1^* and λ_2^* exists if and only if the determinant of the matrix is non-zero. The following claim establishes that this is indeed the case.

Claim 5.1. *The determinant*

$$D = \mathrm{Det}\begin{pmatrix} B & A \\ A & C \end{pmatrix} = BC - A^2 \tag{5.24}$$

is positive and hence non-zero.

Proof. Since $\mathbf{V}^{-1}$ is positive definite it has a Choleski decomposition (see Theorem 2.6)

$$\mathbf{V}^{-1} = \mathbf{R}\mathbf{R}^T, \quad \text{where} \quad \mathbf{R} \in \mathbb{R}^{n\times n} \text{ is lower triangular.}$$

Using this we can consider the vectors

$$\mathbf{u} = \mathbf{R}^T\mathbf{e} \quad \text{and} \quad \mathbf{v} = \mathbf{R}^T\mathbf{1}.$$

We note that, from these choices, we can deduce

$$\mathbf{u}^T\mathbf{v} = \mathbf{e}^T\mathbf{R}\mathbf{R}^T\mathbf{1} = \mathbf{e}^T\mathbf{V}^{-1}\mathbf{1} = A,$$

$$\mathbf{u}^T\mathbf{u} = \mathbf{e}^T\mathbf{R}\mathbf{R}^T\mathbf{e} = \mathbf{e}^T\mathbf{V}^{-1}\mathbf{e} = B,$$

$$\mathbf{v}^T\mathbf{v} = \mathbf{1}^T\mathbf{R}\mathbf{R}^T\mathbf{1} = \mathbf{1}^T\mathbf{V}^{-1}\mathbf{1} = C.$$

We now evoke the Cauchy–Schwarz inequality, a geometric result which tells us that for any two vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ we have

$$(\mathbf{u}^T\mathbf{v})^2 \leq (\mathbf{u}^T\mathbf{u})(\mathbf{v}^T\mathbf{v})$$

with equality if and only if $\mathbf{u}$ and $\mathbf{v}$ are linearly dependent. Applying this to our vectors yields

$$A^2 \leq BC.$$

We recall that we have assumed that the expected returns of our assets are not all the same, this means that $\mathbf{e}$ and $\mathbf{1}$ are linearly independent and thus, so are $\mathbf{u}$ and $\mathbf{v}$. We can conclude then that the inequality above is strict, that is

$$D = BC - A^2 > 0.$$

□

We can now establish that the Lagrange multipliers λ_1^* and λ_2^* are given by

$$\begin{pmatrix} \lambda_1^* \\ \lambda_2^* \end{pmatrix} = \frac{1}{D}\begin{pmatrix} C & -A \\ -A & B \end{pmatrix}\begin{pmatrix} \mu_p \\ 1 \end{pmatrix},$$

and so

$$\lambda_1^* = \frac{A - C\mu_p}{D} \quad \text{and} \quad \lambda_2^* = \frac{A\mu_p - B}{D}. \tag{5.25}$$

The weight vector $\mathbf{w}_p$ whose corresponding feasible portfolio p provides an expected return of $100 \times \mu_p\%$ with minimum risk is now available; we simply substitute the Lagrange multipliers above into equation (5.20) to discover that

$$\begin{aligned} \mathbf{w}_p &= \left(\frac{C\mu_p - A}{D}\right)\mathbf{V}^{-1}\mathbf{e} + \left(\frac{B - A\mu_p}{D}\right)\mathbf{V}^{-1}\mathbf{1} \\ &= \frac{1}{D}\left(B\mathbf{V}^{-1}\mathbf{1} - A\mathbf{V}^{-1}\mathbf{e}\right) + \frac{1}{D}\left(C\mathbf{V}^{-1}\mathbf{e} - A\mathbf{V}^{-1}\mathbf{1}\right)\mu_p. \end{aligned} \tag{5.26}$$

To simplify this expression we let

$$\mathbf{g} = \frac{1}{D}\left(B\mathbf{V}^{-1}\mathbf{1} - A\mathbf{V}^{-1}\mathbf{e}\right) \quad \text{and} \quad \mathbf{h} = \frac{1}{D}\left(C\mathbf{V}^{-1}\mathbf{e} - A\mathbf{V}^{-1}\mathbf{1}\right), \tag{5.27}$$

and so we can write the optimal vector of portfolio weights as

$$\mathbf{w}_p = \mathbf{g} + \mathbf{h}\mu_p. \tag{5.28}$$

We note that the vectors $\mathbf{g}$ and $\mathbf{h}$ in formulae (5.27) depend only upon the covariance information contained in the matrix $\mathbf{V}$ and the vector of expected returns $\mathbf{e}$. Crucially these vectors are independent of the desired level of expected return μ_p and thus the formula (5.28) provides the optimal weight vector for the whole range of expected returns. This means that the portfolio whose weight vector is given by (5.28) is guaranteed to be the feasible portfolio that provides an expected return $100 \times \mu\%$ with minimum risk. This is clearly an extremely useful discovery, however, it is only part of the story; what is missing is the actual value of the risk attached to these optimal portfolios. The fact that the risk is minimized is not enough for a potential investor; the risk needs to be quantified. The variance of the optimal portfolio corresponding to an expected return of $100 \times \mu\%$ is given by

$$\begin{aligned}
\sigma^2 &= \mathbf{w}^T\mathbf{V}\mathbf{w} \\
&= \mathbf{w}^T\mathbf{V}(\mathbf{g}+\mathbf{h}) \to w^T V(g+h_\mu) \\
&= \mathbf{w}^T\mathbf{V}\left(\frac{1}{D}\left[B\mathbf{V}^{-1}\mathbf{1} - A\mathbf{V}^{-1}\mathbf{e}\right] + \mu\left[C\mathbf{V}^{-1}\mathbf{e} - A\mathbf{V}^{-1}\mathbf{1}\right]\right) \\
&= \frac{1}{D}\mathbf{w}^T\left([B\mathbf{1} - A\mathbf{e}] + \mu\,[C\mathbf{e} - A\mathbf{1}]\right) \quad \text{[follows since } \mathbf{V}^{-1}\mathbf{V} = \mathbf{I}_n\text{]} \\
&= \frac{1}{D}\left(C\mu^2 - 2A\mu + B\right) \quad \text{[follows since } \mathbf{w}^T\mathbf{e} = \mu \text{ and } \mathbf{w}^T\mathbf{1} = 1\text{]} \qquad (5.29) \\
&= \frac{C}{D}\left(\mu^2 - \frac{2A}{C}\mu + \frac{B}{C}\right) \\
&= \frac{C}{D}\left[\left(\mu - \frac{A}{C}\right)^2 + \frac{BC - A^2}{C^2}\right] \\
&= \frac{C}{D}\left(\mu - \frac{A}{C}\right)^2 + \frac{1}{C} \quad \text{[follows since } BC - A^2 = D\text{].}
\end{aligned}$$

The strength of the above formula is that it is completely general. An investor whose aim is to construct a portfolio with an expected return of μ_p can use the formula to deduce that the least amount of risk involved in hitting this target is given by

$$\sigma_p^2 = \frac{C}{D}\left(\mu_p - \frac{A}{C}\right)^2 + \frac{1}{C}. \qquad (5.30)$$

If the investor is comfortable with this level of risk then his required portfolio weights are given by (5.28).

We can display the relationship between risk and expected return in (σ, μ)-space. A minor rearrangement of (5.29) shows that the risk–reward coordinates of any optimal portfolio are related by

$$\frac{\sigma^2}{1/C} - \frac{(\mu - A/C)^2}{D/C^2} = 1. \qquad (5.31)$$

We recognize that this formula describes a familiar curve studied in high-school geometry; it represents a hyperbola with centre $(0, A/C)$, asymptotes

$$\mu = \pm\sqrt{\frac{D}{C}}\sigma + \frac{A}{C} \tag{5.32}$$

and vertex

$$V = \left(\sqrt{\frac{1}{C}}, \frac{A}{C}\right) \tag{5.33}$$

We note that this simple curve is precisely the boundary of the region of feasible portfolios, it is commonly called the optimal frontier (see Figure 5.3).

It is remarkable that a seemingly complicated problem has such an elegant solution. A portfolio manager need only compute the constants A, B, C and D and, with this information alone, the whole optimal frontier can be plotted and visualized. A potential investor can consult the optimal frontier to find, at a glance, the level of expected return that suits his appetite for risk; he can then use formula (5.28) to determine the required composition of his desired portfolio.

We remark that the problem we have solved assumes that short selling of assets is allowed and is unrestricted. We have already observed, in Chapter 1, that short selling is a high-risk strategy and, for this reason, it is common for short selling to be restricted; indeed, in September 2008, in response to the global credit crisis, the UK and USA imposed a temporary ban on short selling in an attempt to stabilize their markets. In the case where

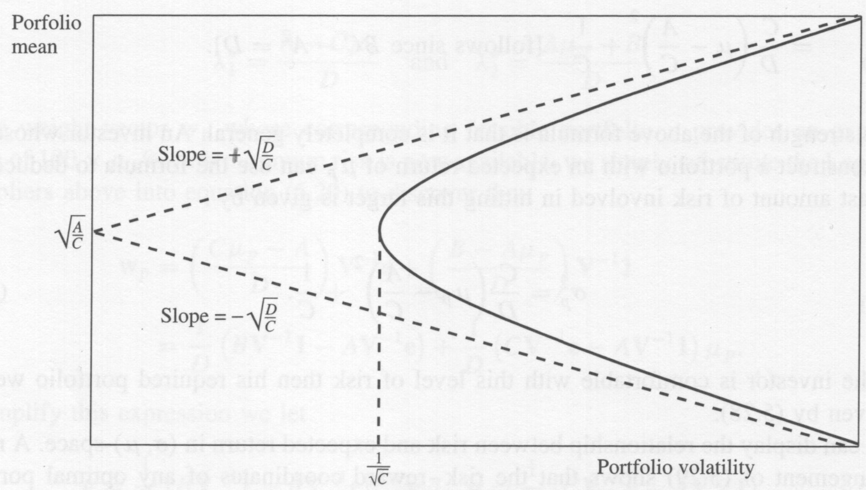

Figure 5.3 The general optimal frontier.

restrictions on short selling are enforced, the corresponding optimal portfolio problem takes the form

$$\begin{aligned} &\text{minimize } \frac{1}{2}\mathbf{w}^T\mathbf{V}\mathbf{w} \quad \text{subject to} \quad \mathbf{P}\mathbf{w} = \mathbf{d} \\ &\qquad\qquad\qquad\quad \text{and} \quad w_i \geq l_i \text{ for } i = 1, \ldots, n. \end{aligned} \tag{5.34}$$

This new problem can be solved by using an extension of the Lagrange function approach, however its solution, in general, will not have a neat closed form. Instead, the solution can be found by employing an appropriate numerical optimization algorithm; the reader who is interested in discovering how such algorithms are constructed is advised to consult Nash and Sofer (1996) and/or Gill, Murray and Wright (1982).