FINANCE 601
Introduction to Finance

Lecture Notes

Clarence C.Y. Kwan, Professor of Finance
DeGroote School of Business
McMaster University
Hamilton, Ontario L8S 4M4

Fall 2024

Chapter 5

Mean-Variance Portfolio Theory

The topic of *mean-variance portfolio theory*, which is fundamental in modern finance, was introduced to the finance profession by Harry Markowitz in the 1950s. In the context of investments, a portfolio is a combination of financial assets, such as common stocks, intended to achieve a better risk-return trade-off, as compared to investing in a single financial asset. The term *risk-return trade-off* is about the trade-off between achieving a higher expected rate of return from an investment and achieving a lower risk exposure of the same investment. The rate of return (or, simply, the return) from an investment is random, and such randomness is considered to be the investment risk.

From a statistical perspective, given a probability distribution of random returns, we can treat the *mean* of the distribution as the expected return and the *variance* of the distribution — or its square root, known as the *standard deviation* — as a risk measure for the random returns. For analytical convenience, a normal distribution, which has the appearance of a bell-shaped curve, is treated as the underlying distribution of returns for mean-variance portfolio theory, as such a distribution is fully characterized by its mean and its standard deviation. Without using a normal distribution, justification of the mean-variance approach will require the assumption of a problematic *quadratic utility function* for each investor.

Although the worse investment outcome from holding some shares of any common stock is a -100% return (if the stock turns out to be totally worthless), a normal distribution can produce a random return that is negative infinity. As covered in this chapter, there is a remedial measure to rectify the situation, by transforming the stock return data involved. Various other practical issues pertaining to portfolio investments are covered in this chapter as well.

Portfolio theory has two separate aspects. As *normative theory*, it provides criteria for the ways in which investment decisions should be made and stipulates rules for attaining desired ends. As *positive theory*, in contrast, it attempts to explain and predict phenomena in security markets. We use the terms *security* and *asset* interchangeably below. This chapter, which introduces the mean-variance portfolio concepts, is focused on the normative aspect of the theory. The positive, or market equilibrium, aspect will be addressed in Chapter 6.

5.1 Normal Distribution and Indifference Curves

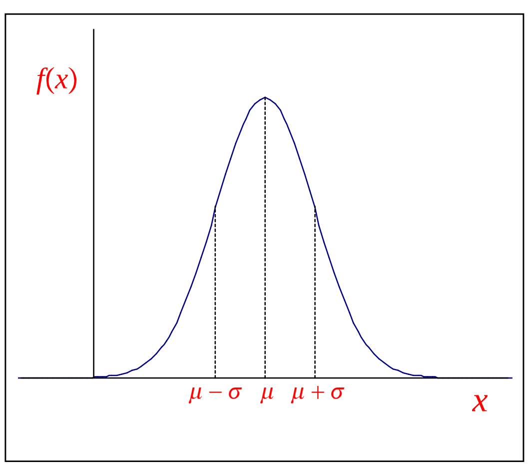
For a continuous probability density function $f(x)$, which is never negative and is defined for $-\infty < x < \infty$, the condition of

$$\int_{x=-\infty}^{\infty} f(x)dx = 1$$

must hold. The condition ensures that all potential outcomes are accommodated. The probability density function of a normal distribution is

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]. \tag{5.1}$$

The probability distribution is characterized by two parameters, μ and σ, which stand for its mean and standard deviation, respectively. The function appears graphically as a bell-shaped curve, as shown in the following:

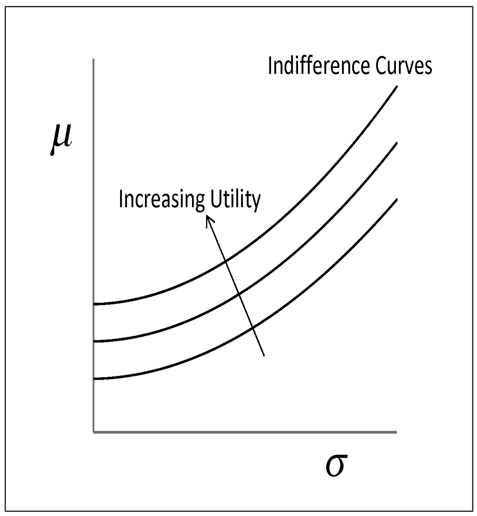


In an investment context, the variable x represents the random return given the probability distribution

In economics, given any two consumption goods, the term *indifference curve* is a graph showing their various combinations that provide the consumer involved an equal level of satisfaction (utility). When applied to an investment context, where utility is characterized by the parameters μ and σ in a normal distribution, we have, for a risk-averse investor,

$$\frac{d\mu}{d\sigma} > 0 \text{ and } \frac{d^2\mu}{d\sigma^2} > 0$$

along each indifference curve. The family of indifference curves is a set of convex curves as shown graphically in the following:



Each indifference curve has a positive slope, and the slope increases as σ increases. Intuitively, a risk-averse investor is willing to incur risk, as long as there is adequate compensation for risk taking. The greater the level of existing risk, the investor will require a greater incremental expected return for bearing a higher risk.

Suppose that (σ_1, μ_1) and (σ_2, μ_2) are points on two different indifference curve for a risk-averse investor. It is obvious that, if $\sigma_1 = \sigma_2$ and $\mu_1 < \mu_2$, the indifference curve containing (σ_2, μ_2) provides a higher utility. This means that, for two investments with the same risk exposure, a risk-averse investor always prefers the one that provides a higher expected return. It is also obvious that, if $\sigma_1 > \sigma_2$ and $\mu_1 = \mu_2$, the indifference curve containing (σ_2, μ_2) provides a higher utility. This means that, for two investments with the same expected return, a risk-averse investor always prefers the one that has a lower risk exposure.

However, if $\sigma_1 > \sigma_2$ and $\mu_1 > \mu_2$ instead, whether the investor involved has a higher utility from (σ_1, μ_1) or from (σ_2, μ_2) is unclear before we know the exact shape of each indifference curve for the investor. This ambiguity brings out a fundamental issue, which warns us about misapplications of the mean-variance approach when the random returns are not normally distributed. A simple illustrative example is provided in the next section.

5.2 A Mean-Variance Paradox: An Illustration

- A paradox is something that seems self-contradictory. In the case of an opinion, it is something that seems contrary to commonly accepted opinion.

Suppose that, for a given period of time, investment A is equally likely to earn returns of -1%, 0%, 1%, 2%, and 3%. Suppose also that, for the same period of time, investment B is equally likely to earn returns of -1%, 1%, 3%, 5%, and 7%. The task here is to verify whether there is dominance of an investment over the other investment for a rational investor.

As shown in the following payout table, the nine potential returns from the two competing investments are -1%, 0%, 1%, 2%, 3%, 4%, 5%, 6%, and 7%. in equal increments of 1%:

Returns	Investment A Probability	Investment B Probability	Investment A Cumulative Probability	Investment B Cumulative Probability
-1%	0.2	0.2	0.2	0.2
0%	0.2	0	0.4	0.2
1%	0.2	0.2	0.6	0.4
2%	0.2	0	0.8	0.4
3%	0.2	0.2	1	0.6
4%	0	0	1	0.6
5%	0	0.2	1	0.8
6%	0	0	1	0.8
7%	0	0.2	1	1

To see whether investment B is the choice by a rational investor, let us read the above table, row by row, with attention to the cumulative probabilities.

- In row 1, where both investments have a -1% return, there is a 20% probability of earning this bad return. Implicitly, there is an 80% probability of earning better returns. Thus, for this row, a rational investor must be indifferent between the two investments.

- Row 2 corresponds to a 0% return. In the case of investment A, there is a 40% probability of earning a 0% or less. Implicitly, there is a 60% probability of earning better returns. In the case of investment B, there is a 20% probability of earning a 0% or less. Implicitly, there is an 80% probability of earning more. Thus, for this row, a rational investor will choose investment B.

- Row 3 corresponds to a 1% return. In the case of investment A, there is a 60% probability of earning a 1% or less. Implicitly, there is a 40% probability of earning more. In the case of investment B, there is a 40% probability of earning a 1% or less. Implicitly, there is a 60% probability of earning more. Thus, for this row, a rational investor will choose investment B.

- Continuing with row 4 through row 8 in the same fashion, we clearly see that a rational investor will choose investment B.

- In row 9, which corresponds to a 7% return. In either investment, there is a 100% probability of earning a 7% or less. Implicitly, there is a 0% probability of earning more, regardless of which investment is chosen. Thus, for this row, a rational investor must be indifferent between the two investments.

Based on the above row-by-row comparisons, a rational investor will always choose investment B. This approach to compare risky investments is commonly known as *stochastic dominance.*

For the mean-variance approach where μ and σ represent the expected return and the standard deviation of returns, respectively, we have

$$\begin{aligned}
\mu_A &= \frac{1}{5}(-1+0+1+2+3)\% = 1\%, \\
\sigma_A &= \sqrt{\frac{1}{5}\left[(-1-1)^2+(0-1)^2+(1-1)^2+(2-1)^2+(3-1)^2\right]}\,\% = 1.41\%, \\
\mu_B &= \frac{1}{5}(-1+1+3+5+7)\% = 3\%, \\
\text{and } \sigma_B &= \sqrt{\frac{1}{5}\left[(-1-3)^2+(1-3)^2+(3-3)^2+(5-3)^2+(7-3)^2\right]}\,\% = 2.82\%.
\end{aligned}$$

With

$$\begin{aligned} \mu_A &< \mu_B \\ \text{and } \sigma_A &< \sigma_B, \end{aligned}$$

the use of the mean-variance approach will fail to reach any conclusion regarding dominance between investment A and investment B. The reason for this indecision is that the random returns here are from *discrete uniform distributions*, thus violating the assumption of normal distributions as required for the mean-variance approach.

It is important to recognize that, the approach of *stochastic dominance* as described above does not always lead to the dominance of an investment over the other investment for a rational investor. To illustrate, suppose that the worst random outcome of investment B is -2% instead of -1%. That is, investment B is equally likely to earn returns of -2%, 1%, 3%, 5%, and 7% instead. Now, there are 10 potential returns instead from the two competing investments, which are -2%, -1%, 0%, 1%, 2%, 3%, 4%, 5%, 6%, and 7%. in equal increments of 1%. We can go through the same procedure to compare the cumulative probabilities for investment A and investment B, row by row in the revised payout table. The comparisons will show that there is no dominance of an investment over the other investment.

5.3 A Remedy: Logarithmic Transformation

If the assumption of normal distributions is deemed unacceptable, is there another way to ensure that the mean-variance investment decisions be valid, or be at least practically justifiable? We provide below a remedial measure, which involves a transformation of the return data. To illustrate, let

$$P_0, P_1, P_2, \ldots, P_n$$

be the daily prices of a stock over a month, as recorded at the end of each trading day. Here, P_0 is the stock price at the end of the last day of the previous month, and P_n is the stock price at the end of the last day of the current month. Thus, ignoring any dividend component, one plus the return over the current month is

$$\begin{aligned} 1 + R &= \frac{P_n}{P_0} = \frac{P_1}{P_0} \cdot \frac{P_2}{P_1} \cdot \frac{P_3}{P_2} \cdots \frac{P_n}{P_{n-1}} \\ &= (1 + r_1)(1 + r_2)(1 + r_3) \cdots (1 + r_n), \end{aligned}$$

where r_i, for $i = 1, 2, \ldots, n$, is the return for day i.

Now, we treat each r_i as a random draw from a stationary but unknown probability distribution. We also take the natural logarithm of both sides of the above equation relating the monthly and daily returns to obtain

$$\begin{aligned} \ln(1+R) &= \ln[(1+r_1)(1+r_2)(1+r_3)\cdots(1+r_n)] \\ &= \sum\nolimits_{i=1}^{n} \ln(1+r_i). \end{aligned}$$

Each $\ln(1+r_i)$ term can also be viewed as a random draw from a stationary but unknown distribution. Then, the distribution of the sample mean

$$\frac{1}{n}\sum\nolimits_{i=1}^{n} \ln(1+r_i),$$

according to the *central limit theorem* in statistics, is approximately normal.

Therefore, the distribution of

$$\frac{1}{n}\ln(1+R)$$

is approximately normal. Assuming that the number of trading days in a month, n, is a constant, the distribution of

$$\ln(1+R)$$

is also approximately normal.

5.3.1 Logarithmic Transform: What It Means

To explore the meaning of the logarithmically transformed return data, let

$$\begin{aligned} \mathbf{R} &= \text{continuously compounded rate of return over the period,} \\ \text{and } R &= \text{effective rate of return over the same period.} \end{aligned}$$

Given that

$$\lim_{n\to\infty}\left(1+\frac{\mathbf{R}}{n}\right)^n = \exp(\mathbf{R}) = 1+R,$$

we have

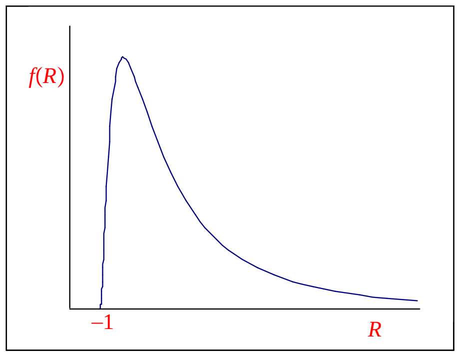
$$\mathbf{R} = \ln(1+R).$$

Therefore, if R is an effective monthly return, then

$$\ln(1+R)$$

is the corresponding continuously compounded monthly return.

Another way of stating that $\ln(1+R)$ is approximately normally distributed is that the distribution of $1+R$ is approximately *lognormal*. If $\ln(1+R)$ is normally distributed, the graph of the probability density function of R, denoted by $f(R)$, is as follows:



The lowest value of R for this density function is $R = -1$ (the worst outcome for investing in a stock). As $R \to -1$ (or, equivalently, -100%), we have $\ln(1+R) \to -\infty$, capturing the left tail of a normal distribution.

5.4 Measuring Security and Portfolio Returns

Portfolio analysis is about efficient allocations of investment funds. Before addressing any issues on risk-return trade-off, let us start with return measurements. To illustrate, suppose that an investor owns a portfolio of four stocks with the following information:

stock label	price one year ago	price today	annual dividend	number of shares owned
1	\$40	\$42	\$2	100
2	\$30	\$26	\$2	200
3	\$25	\$28	\$1	400
4	\$25	\$23	\$3	200

What is the realized annual return on such a portfolio?

Two approaches are provided here. In the first approach, we start with the portfolio value one year ago and label it as P_0. We have

$$\begin{aligned} P_0 &= 100(\$40) + 200(\$30) + 400(\$25) + 200(\$25) \\ &= \$25,000. \end{aligned}$$

Next, we compute the portfolio value today, plus any dividends received, and label them as P_1 and D_1, respectively. We have

$$P_1 = 100(\$42) + 200(\$26) + 400(\$28) + 200(\$23)$$

and

$$D_1 = 100(\$2) + 200(\$2) + 400(\$1) + 200(\$3)$$

or, simply,

$$\begin{aligned} P_1 + D_1 &= 100(\$42 + \$2) + 200(\$26 + \$2) + 400(\$28 + \$1) + 200(\$23 + \$3) \\ &= \$26,800. \end{aligned}$$

The realized annual return on the portfolio is

$$\begin{aligned} R_p &= \frac{(P_1 + D_1) - P_0}{P_0} \\ &= \frac{\$26,800 - \$25,000}{\$25,000} \\ &= 0.072 = 7.2\%. \end{aligned}$$

In the second approach, we compute the following:

stock label	(price one year ago) × (number of shares owned)			portfolio weight		
1	$40 × 100	=	$4,000	$x_1 = \$4,000/\$25,000$	=	0.16
2	$30 × 200	=	$6,000	$x_2 = \$6,000/\$25,000$	=	0.24
3	$25 × 400	=	$10,000	$x_3 = \$10,000/\$25,000$	=	0.40
4	$25 × 200	=	$5,000	$x_4 = \$5,000/\$25,000$	=	0.20
	total	=	$25,000	total $= x_1 + x_2 + x_3 + x_4$	=	1.00

stock label	annual return			
1	$R_1 =$	$(\$42 + \$2 - \$40)/\$40$	$=$	0.10
2	$R_2 =$	$(\$26 + \$2 - \$30)/\$30$	$=$	$-\frac{1}{15} = -0.0667$
3	$R_3 =$	$(\$28 + \$1 - \$25)/\$25$	$=$	0.16
4	$R_4 =$	$(\$23 + \$3 - \$25)/\$25$	$=$	0.04

The realized annual return on portfolio is

$$\begin{aligned} R_p &= \sum\nolimits_{i=1}^{4} (\text{weight})_i \times (\text{annual return})_i \\ &= \sum\nolimits_{i=1}^{4} x_i R_i \\ &= x_1R_1 + x_2R_2 + x_3R_3 + x_4R_4 \\ &= 0.16(0.10) + 0.24\left(-\frac{1}{15}\right) + 0.40(0.16) + 0.20(0.04) \\ &= 0.072 = 7.2\%. \end{aligned}$$

The results from the two approaches are the same; however, the second approach is analytically more convenient.

5.4.1 Returns over Shorter Intervals

The measurement of security returns, though generally straightforward, could still be confusing sometimes. For example, when raw price and dividend data are used to calculate stock returns, especially around the time of stock splits or stock dividends, exactly when and how to include the dividend component needs to be clarified.

Example 33 *Given the following information, find the returns for February, March, and April:*

Date	*January* 31	*February* 28	*March* 31	*April* 30
Share Price	\$40	\$39	\$41	\$43

Announcement Date:	*February (A* \$0.50 *per share dividend was announced.)*
Ex-dividend Date:	*March*
Record Date:	*March*
Payment Date:	*April*

The returns are as follows:

$$\begin{aligned} R_{\text{February}} &= \frac{39}{40} - 1 = -2.50\%; \\ R_{\text{March}} &= \frac{41 + 0.50}{39} - 1 = 6.41\%; \\ R_{\text{April}} &= \frac{43}{41} - 1 = 4.88\%. \quad \blacksquare \end{aligned}$$

In case of a stock dividend or a stock split, an adjustment to the number of shares held must be made. In the above example, if the dividend was a 1% stock dividend instead of the \$0.50 cash dividend, R_{March} would be

$$\frac{41(1.01)}{39} - 1 = 6.18\%,$$

because each share held at the beginning of the month would become 1.01 shares at the end of the month.

There are also practical issues to consider, such as proper ways to accommodate holidays, missing data, or the lack of trading activities of some securities during the sample period. In the case of international portfolio investments, non-synchronous trading hours could be issues to consider as well.

5.4.2 Practical Issues Involving Ex-Dividend and Record Dates

In North American stock markets, settlement of buy/sell transactions of stocks currently follow a **T+1 process**. The changeover from a T+3 process to a T+2 process took place on Tuesday, September 5, 2017. The changeover from a T+2 process to a T+1 process took place on Monday, May 27, 2024. Under the current T+1 process, if someone buys a stock, it takes only one business day (instead of two to three) from the transaction date (T) for the ownership of the stock certificate, as well as the money involved, to change hands between the buyer and the seller. Given the current and previous practice in stock markets, let us use a dividend announcement by Bank of Montreal (BMO) as an illustration.

BMO announced on Tuesday, February 27, 2018 a \$0.93 regular quarterly dividend per share, to be paid on Monday, May 28, 2018. The ex-dividend date was Monday, April 30, 2018 and the record date was Tuesday, May 1, 2018, which was the first business day after the ex-dividend date. Having the record date immediately following the ex-dividend date is a common practice.

Let us consider a scenario where investor A, the current owner of a round lot (i.e., 100 shares) of BMO stock, sold all these shares to investor B on Friday, April 27, 2018 or before. Given the T+2 process, the settlement of the buy/sell transactions would take place two business days later; that is, Tuesday, May 1, 2018 or earlier, as the Saturday and the Sunday between these two dates were not business days. In BMO's record as of Tuesday, May 1, 2018, investor B was an owner of 100 shares of its stock. In this scenario, investor B was entitled to receive \$93.00 in dividend from BMO on Monday, May 28, 2018.

Now, let us consider a different scenario where investor A sold the 100 BMO shares on Monday, April 30, 2018 or later instead. Given the T+2 process, the corresponding settlement dates would be Wednesday, May 2, 2018 or later. Then, on Tuesday, May 1, 2018, which was the record date, investor A was still the owner of the 100 BMO shares. In this scenario, investor A was entitled to receive the \$93.00 dividend from BMO on Monday, May 28, 2018.

According to the historical prices that are freely available from the Yahoo!Finance website, the daily closing prices of BMO from Monday, April 23, 2018 to Friday, May 4, 2018 are as follows:

Date	Closing Price
Monday, April 23, 2018	\$96.17
Tuesday, April 24, 2018	\$96.68
Wednesday, April 25, 2018	\$97.12
Thursday, April 26, 2018	\$97.93
Friday, April 27, 2018	\$98.51
Monday, April 30, 2018 (ex-dividend date)	\$97.51
Tuesday, May 1, 2018	\$97.76
Wednesday, May 2, 2018	\$97.84
Thursday, May 3, 2018	\$97.47
Friday, May 4, 2018	\$98.05

During the first week of this two-week period, the daily closing prices of BMO showed a clear upward trend, from \$96.17 to \$98.51. Investors who purchased the BMO stock on or before the end of the first week would receive the \$0.93 quarterly dividend per share. However, investors who purchased the same stock on or after the Monday that followed would not. A major contributing factor for the \$1.00 decrease in the closing share prices from \$98.51 to \$97.51, between Friday, April 27, 2018 and Monday, April 30, 2018, was the ex-dividend effect, as the dividend involved was \$0.97 per share.

Some companies pay **extra dividends** in addition to regular quarterly dividends. Some other companies pay stock dividends, instead of cash dividends. In U.S. stock markets, for example, there are specific rules for extra dividends that are 25% or more of the share value. One specific rule is that the ex-dividend date must be the first business day following the dividend payment date. As dividends are paid to shareholders who are on the company's record, the record date will have to be before the ex-dividend date. If the owner of some shares of a stock sells them, after the record date but before the ex-dividend date, the dividend received will have to be transferred from his/her account to the buyer's account. Conversely, if someone buys the same stock, after the record date but before the ex-dividend date, he/she will still receive the dividend, via fund transfer from the seller's account to his/her account.

Regardless of the rules involved (depending on the dollar amount of the dividend as a proportion of the share value), the ex-dividend date is still the dividing point in the time line to determine, between the seller and the buyer of the stock, who is entitled to receive the announced dividend. Specifically, the seller of the stock is not entitled to receive the announced dividend, if the sale is before the ex-dividend date. Neither is the buyer of the stock, if the purchase is on or after the ex-dividend date.

5.5 Portfolio Return and Risk

Let $\widetilde{R}_i$ be the random return of security i, for $i = 1, 2, \ldots, n$. The expected return of security i is

$$\mu_i = E(\widetilde{R}_i),$$

where $E(\cdot)$ stands for the expected value of any random variable $(\cdot)$. Let $x_i \geq 0$ be the proportion of investment funds allocated to security i. The random return of the portfolio is

$$\widetilde{R}_p = \sum\nolimits_{i=1}^{n} x_i \widetilde{R}_i \,,$$

where

$$\sum\nolimits_{i=1}^{n} x_i = 1.$$

The expected return of the portfolio is

$$\begin{aligned}\mu_p &= E(\widetilde{R}_p) = E\left(\sum\nolimits_{i=1}^n x_i \widetilde{R}_i\right) \\ &= \sum\nolimits_{i=1}^n x_i E(\widetilde{R}_i) = \sum\nolimits_{i=1}^n x_i \mu_i.\end{aligned}$$

To derive an expression of the variance of portfolio returns, we start with

$$\begin{aligned}\sigma_p^2 &= Var(\widetilde{R}_p) = E[(\widetilde{R}_p - \mu_p)^2] \\ &= E\left[\left(\sum\nolimits_{i=1}^n x_i \widetilde{R}_i - \sum\nolimits_{i=1}^n x_i \mu_i\right)^2\right] = E\left\{\left[\sum\nolimits_{i=1}^n x_i \left(\widetilde{R}_i - \mu_i\right)\right]^2\right\} \\ &= E\left\{\left[\sum\nolimits_{i=1}^n x_i \left(\widetilde{R}_i - \mu_i\right)\right]\left[\sum\nolimits_{j=1}^n x_j \left(\widetilde{R}_j - \mu_j\right)\right]\right\}.\end{aligned}$$

The above change in one of the subscripts from i to j is for notational clarity when two identical sums are multiplied. It follows that

$$\begin{aligned}\sigma_p^2 &= E\left[\sum\nolimits_{i=1}^n x_i \left(\widetilde{R}_i - \mu_i\right) \sum\nolimits_{j=1}^n x_j \left(\widetilde{R}_j - \mu_j\right)\right] \\ &= E\left[\sum\nolimits_{i=1}^n \sum\nolimits_{j=1}^n x_i x_j \left(\widetilde{R}_i - \mu_i\right)\left(\widetilde{R}_j - \mu_j\right)\right] \\ &= \sum\nolimits_{i=1}^n \sum\nolimits_{j=1}^n x_i x_j E\left[\left(\widetilde{R}_i - \mu_i\right)\left(\widetilde{R}_j - \mu_j\right)\right] = \boxed{\sum\nolimits_{i=1}^n \sum\nolimits_{j=1}^n x_i x_j \sigma_{ij}},\end{aligned}$$

where

$$\sigma_{ii} = \sigma_i^2 \ \text{ and } \ \sigma_{ij} = \sigma_{ji}, \ \text{for } i, j = 1, 2, \ldots, n.$$

Implicitly, we have $\sigma_i = \sqrt{\sigma_{ii}}$.

The *correlation coefficient of returns* between securities i and j or, simply, their *correlation of returns* is

$$\rho_{ij} = \frac{\sigma_{ij}}{\sigma_i \sigma_j}, \ \text{for } i, j = 1, 2, \ldots, n.$$

Unlike *covariances*, each *correlation* is a dimensionless quantity; it does not have a unit of measurement. The highest correlation of returns is 1; this is the correlation of returns between any risky security i and itself, which is ρ_{ii}. If there are two different securities, denoted as i and j, for which the condition of

$$\widetilde{R}_i = a + b\widetilde{R}_j$$

holds, where a is any constant and b is any *positive* constant, we have $\rho_{ij} = 1$ as well. The lowest correlation of returns is -1; for two different securities, denoted as i and j, this is the case where the above condition holds for a being any constant and b being any *negative* constant instead. In general, we can write

$$-1 \leq \rho_{ij} = \frac{\sigma_{ij}}{\sigma_i \sigma_j} \leq 1, \text{ for } i,j = 1,2,\ldots,n.$$

- For investment decisions involving the use of portfolio selection models for each n-security case, input parameters such as μ_i, σ_i, and ρ_{ij} — or, equivalently, μ_i and σ_{ij} — for $i,j = 1,2,\ldots,n$ are required. As the true values such parameters are unknown and need to be estimated, some relevant statistical and practical issues involved are to be addressed in Section 5.16 of this chapter. For now, we treat the values of such input parameters as known instead.

5.6 A Two-Security Illustration

Diversification is about combining securities — which are less than perfectly positively correlated in returns — into portfolios for reducing the investment risk without sacrificing the investment returns. In a two-security case, we have

$$\begin{aligned} \widetilde{R}_p &= \sum\nolimits_{i=1}^{2} x_i \widetilde{R}_i = x_1 \widetilde{R}_1 + x_2 \widetilde{R}_2, \\ \text{where } x_1 &\geq 0,\ x_2 \geq 0, \text{ and } x_1 + x_2 = 1. \end{aligned}$$

We also have

$$\begin{aligned} \mu_p &= \sum\nolimits_{i=1}^{2} x_i \mu_i = x_1 \mu_1 + x_2 \mu_2, \\ \sigma_p^2 &= \sum\nolimits_{i=1}^{2} \sum\nolimits_{j=1}^{2} x_i x_j \sigma_{ij} = x_1^2 \sigma_1^2 + x_2^2 \sigma_2^2 + 2 x_1 x_2 \sigma_{12}, \\ \text{and } \sigma_{12} &= \rho_{12} \sigma_1 \sigma_2. \end{aligned}$$

As an example, let us establish the risk-return trade-off given the following information, for $\rho_{12} = 1, -1$, and 0:

i	μ_i	σ_i
1	$5\% = 0.05$	$20\% = 0.20$
2	$15\% = 0.15$	$40\% = 0.40$

- Allocations of investment funds between the two securities in a two-security case are straightforward. For each dollar of investment funds, once a decision has been made on how much is to be invested in one security, the remainder of the dollar will automatically be allocated to the other security. Optimization tools such as those for risk minimization, if needed, are typically confined to essential univariate calculus tools. In cases where there are more than two securities, efficient allocations of each dollar of investment funds, as intended to satisfy some specific conditions, tend to require multivariate differential calculus tools, which are inevitably more sophisticated.

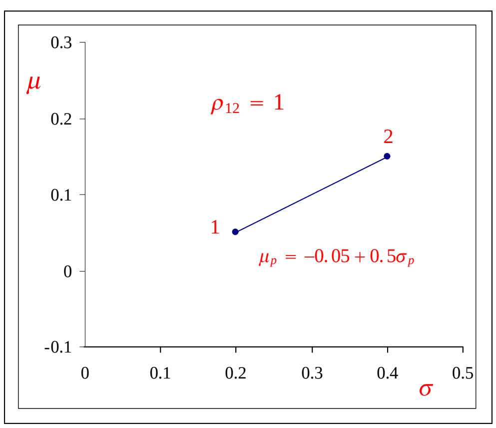
Case 1: For $\rho_{12} = 1$ (perfectly positively correlated returns), we have

$$\begin{aligned} \mu_p &= x_1\mu_1 + (1 - x_1)\mu_2 \\ &= 0.15 - 0.1x_1, \qquad (5.2) \\ \sigma_p^2 &= x_1^2\sigma_1^2 + x_2^2\sigma_2^2 + 2x_1x_2\rho_{12}\sigma_1\sigma_2 \\ &= x_1^2\sigma_1^2 + x_2^2\sigma_2^2 + 2x_1x_2\sigma_1\sigma_2 \\ &= (x_1\sigma_1 + x_2\sigma_2)^2 \\ &= (0.4 - 0.2x_1)^2, \\ \text{and } \sigma_p &= 0.4 - 0.2x_1, \quad \text{for } 0 \le x_1 \le 1. \qquad (5.3) \end{aligned}$$

Combining (5.2) and (5.3) to eliminate x_1 leads to

$$\mu_p = -0.05 + 0.5\sigma_p.$$

For $x_1 \ge 0$, $x_2 \ge 0$, and $x_1 + x_2 = 1$, the relevant line segment capturing the risk-return trade-off is between the two points as indicated.



- If we allow x_1 and x_2 to be negative without violating the condition of $x_1 + x_2 = 1$, we can extend the line segment. Doing so requires the assumption of *frictionless short* *sales.* What such an assumption is all about will soon be clear, from both analytical and practical perspectives.

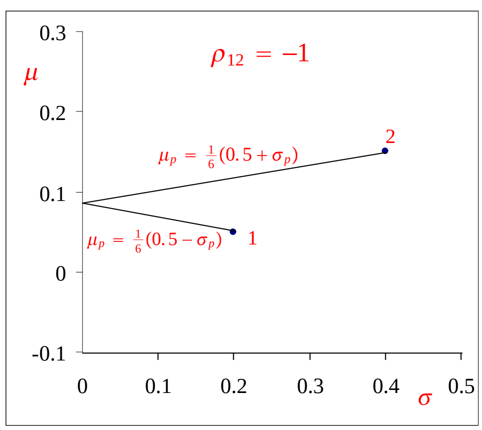
Case 2: For $\rho_{12} = -1$ (perfectly negatively correlated returns), we have

$$\begin{aligned} \mu_p &= 0.15 - 0.1x_1, && (5.4) \\ \sigma_p^2 &= x_1^2\sigma_1^2 + x_2^2\sigma_2^2 + 2x_1x_2\rho_{12}\sigma_1\sigma_2 \\ &= x_1^2\sigma_1^2 + x_2^2\sigma_2^2 - 2x_1x_2\sigma_1\sigma_2 = (x_1\sigma_1 - x_2\sigma_2)^2 \\ &= (-0.4 + 0.6x_1)^2, \\ \text{and} \quad \pm\sigma_p &= -0.4 + 0.6x_1. && (5.5) \end{aligned}$$

- It is important to recognize that σ_p, which is a standard deviation, is never negative.

Combining (5.4) and (5.5) to eliminate x_1 leads to the following two line segments:

$$\mu_p = \frac{1}{6}(0.5 \pm \sigma_p).$$



- The slopes of the two line segments are of the same magnitude and of the opposite sign.

- The two line segments meet at the point where $\sigma_p = 0$, implying that a risk-free portfolio can be reached.

- Portfolios on the negatively sloped line segment are dominated by those on the positively sloped line segment and, thus, can be disregarded.

Case 3: For $\rho_{12} = 0$ (uncorrelated returns), we have

$$\begin{aligned} \mu_p &= 0.15 - 0.1x_1 && (5.6) \\ \text{and } \sigma_p^2 &= x_1^2\sigma_1^2 + x_2^2\sigma_2^2 + 2x_1x_2\rho_{12}\sigma_1\sigma_2 = x_1^2\sigma_1^2 + x_2^2\sigma_2^2 \\ &= x_1^2(0.2)^2 + (1-x_1)^2(0.4)^2 \\ &= 0.16 - 0.32x_1 + 0.2x_1^2. && (5.7) \end{aligned}$$

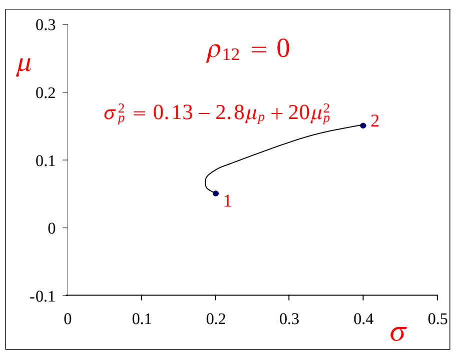
Combining (5.6) and(5.7) to eliminate x_1 leads to a branch of the following hyperbola connecting the two points:

$$\sigma_p^2 = 0.13 - 2.8\mu_p + 20\mu_p^2.$$

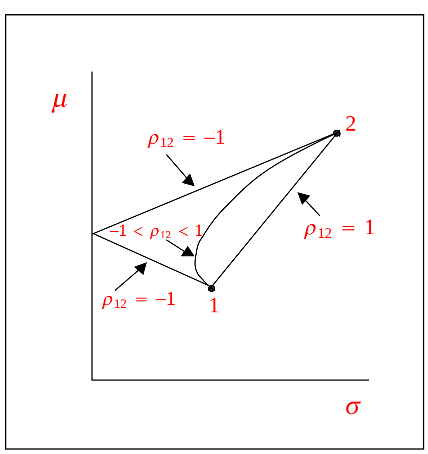
- As this equation can be expressed equivalently as

$$\frac{\sigma_p^2}{(0.1789)^2} - \frac{(\mu_p - 0.07)^2}{(0.04)^2} = 1, \tag{5.8}$$

  its graph on the (σ, μ)-plane is a branch of a hyperbola with a horizontal transverse axis.



In general, we have for each two-security case the following risk-return trade-off:



Mean-variance efficient portfolios are confined to the part of each graph on the (σ, μ)-plane with positive slopes.

Example 34 *(Risk Minimization) For case 3 of the same two-security example, where* $\mu_1 = 0.05$, $\mu_2 = 0.15$, $\sigma_1 = 0.20$, $\sigma_2 = 0.40$, *and* $\rho_{12} = 0$, *determine the least risky portfolio. Specifically, express the portfolio weights, the expected return, and the standard deviation of returns of such a portfolio in percentage terms.*

Label this portfolio as p. Given equation (5.7),

$$\sigma_p^2 = 0.16 - 0.32x_1 + 0.2x_1^2,$$

we have

$$\begin{aligned} \frac{\partial \sigma_p^2}{\partial x_1} &= -0.32 + 0.4x_1 = 0, \\ x_1 &= 0.8 = 80\%, \\ \text{and} \quad x_2 &= 1 - x_1 = 0.2 = 20\%. \end{aligned}$$

The portfolio's expected return and standard deviation of returns are, respectively,

$$\begin{aligned} \mu_p &= 0.15 - 0.1x_1 = 0.07 = 7\% \\ \text{and} \quad \sigma_p &= \sqrt{0.16 - 0.32x_1 + 0.2x_1^2} = 0.1789 = 17.89\%. \quad \blacksquare \end{aligned}$$

- In view of equation (5.8), the above results are as expected.

5.6.1 An Example of High Correlation of Returns

Let us use BMO stocks traded on the Toronto Stock Exchange (TSX) and on the New York Stock Exchange (NYSE) as an example of high correlation of returns. For this example, month-end prices and dividend data were collected online, from Yahoo! Finance, for a period from August 31, 2010 to September 30, 2015. Foreign exchange data between Canada and U.S. over the same period were also collected online from the U.S. Federal Reserve System. The price and dividend data were then used to generate 60 pairs of monthly returns, from which the correlation of returns were estimated. When accounting for the exchange rates, all U.S. prices and dividends were translated into Canadian equivalents, month by month.

Regardless of whether the U.S./Canada exchange rates are considered, a high correlation of returns is clearly noted. The correlation is 0.9144, if the return data are not corrected for changes in the exchange rates; it is 0.9036 instead, otherwise. The main reason for the correlation to be less than perfect (i.e., less than +1) is that the participants (investors) in the two markets (TSX and NYSE) are different. However, the BMO prices in the two markets, when translated into a common currency, would not be different enough to generate arbitrage opportunities for investors, when transaction costs were also considered.

- Estimation errors of correlations are considered in Appenix F at the end of this chapter.

5.7 A Three-Security Illustration

When the number of securities exceeds two, the corresponding portfolio selection problem (to determine the portfolio weights) will become more complicated if the portfolio weights are not allowed to be negative. However, if we assume frictionless short sales (to accommodate negative portfolio weights), then the problem will remain relatively simple. The following is a three-security illustration in two parts.

$$\begin{array}{llll} \text{Input Data:} & \mu_1 = 0.05 & \sigma_1 = 0.02 & \rho_{12} = 0.5 \\ & \mu_2 = 0.08 & \sigma_2 = 0.05 & \rho_{13} = 0 \\ & \mu_3 = 0.15 & \sigma_3 = 0.15 & \rho_{23} = 0.2 \end{array}$$

Part One: Construct the *global minimum variance portfolio*, which is the least risky portfolio based on the three securities (among all portfolios with different expected returns that can be formed by using the three securities) under the assumption of frictionless short sales.

Solution: We first find all the variances and covariances involved, and express the expected return μ_p and the variance of returns σ_p^2 of the portfolio, labeled as p here, in terms of the portfolio weights, x_1, x_2, and x_3. Starting with

$$\begin{aligned} \sigma_{12} &= \rho_{12}\sigma_1\sigma_2 = 0.0005 = \sigma_{21}, \\ \sigma_{13} &= \rho_{13}\sigma_1\sigma_3 = 0 = \sigma_{31}, \\ \sigma_{23} &= \rho_{23}\sigma_2\sigma_3 = 0.0015 = \sigma_{32}, \\ \mu_p &= x_1\mu_1 + x_2\mu_2 + (1 - x_1 - x_2)\mu_3 = 0.15 - 0.1x_1 - 0.07x_2, \end{aligned}$$

and

$$\begin{aligned}\sigma_p^2 &= \sum\nolimits_{i=1}^{3} \sum\nolimits_{j=1}^{3} x_i x_j \sigma_{ij} \\ &= x_1^2\sigma_1^2 + x_2^2\sigma_2^2 + x_3^2\sigma_3^2 + 2x_1x_2\sigma_{12} + 2x_1x_3\sigma_{13} + 2x_2x_3\sigma_{23},\end{aligned}$$

and letting $y = 10,000\,\sigma_p^2$, for computational convenience, we have

$$y = 4x_1^2 + 25x_2^2 + 225(1 - x_1 - x_2)^2 + 10x_1x_2 + 30x_2(1 - x_1 - x_2). \tag{5.9}$$

As there are two decision variables, x_1 and x_2, we solve them directly from the following:

$$\begin{aligned}\frac{\partial y}{\partial x_1} &= 8x_1 - 450(1 - x_1 - x_2) + 10x_2 - 30x_2 = 0, \\ x_2 &= \frac{1}{430}(450 - 458x_1), \qquad (5.10) \\ \frac{\partial y}{\partial x_2} &= 50x_2 - 450(1 - x_1 - x_2) + 10x_1 + 30(1 - x_1 - 2x_2) = 0, \\ \text{and} \quad x_2 &= \frac{1}{44}(42 - 43x_1) \qquad (5.11)\end{aligned}$$

Solving equations (5.10) and (5.11), and noting that $x_3 = 1 - x_1 - x_2$, we now have the three portfolio weights to determine μ_p and σ_p^2. The results are as follows:

$$x_1 = 1.0469 \qquad x_2 = -0.0686 \qquad x_3 = 0.0217 \qquad \sigma_p = 0.01961 \qquad \mu_p = 0.05011$$

With $x_1 > 1$ and $x_2 < 0$, an implicit assumption here is that the short seller not only provides no margin deposit, but also has immediate access to the short-sale proceeds for investing in other securities.

- This is what frictionless short sales are all about; such an assumption, which treats short selling as a way to generate investment funds, is for analytical convenience.

Part Two: Construct the least risky portfolio with a 10% expected return.

Solution: With

$$\mu_p = 0.1 = 0.15 - 0.1x_1 - 0.07x_2,$$

we have

$$\begin{aligned} x_1 &= 0.5 - 0.7x_2 \\ \text{and } x_3 &= 1 - x_1 - x_2 = 0.5 - 0.3x_2. \end{aligned}$$

Then, we can write equation (5.9) as

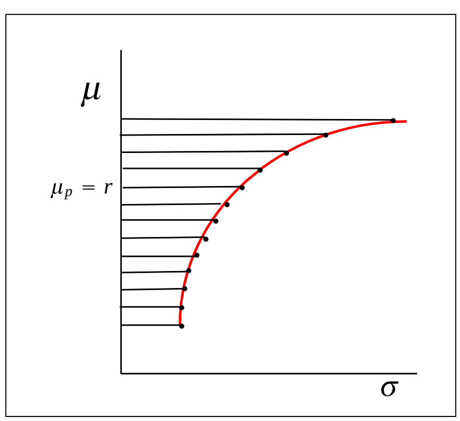
$$\begin{aligned} y &= 4(0.5 - 0.7x_2)^2 + 25x_2^2 + 225(0.5 - 0.3x_2)^2 \\ &\quad +10(0.5 - 0.7x_2)x_2 + 30x_2(0.5 - 0.3x_2). \end{aligned}$$

With x_2 being the only decision variable, we have

$$\begin{aligned} \frac{dy}{dx_2} &= [8(0.5 - 0.7x_2)(-0.7) + 50x_2 + 450(0.5 - 0.3x_2)(-0.3) \\ &\quad +10(0.5 - 1.4x_2) + 30(0.5 - 0.6x_2)] \\ &= 0, \end{aligned}$$

which give $x_2 = 0.8058$ and, accordingly, $x_1 = -0.0641$, $x_3 = 0.2583$, and $\sigma_p = 0.06082$. With $x_1 < 0$ and $x_2 + x_3 > 1$, frictionless short selling is implied here as well.

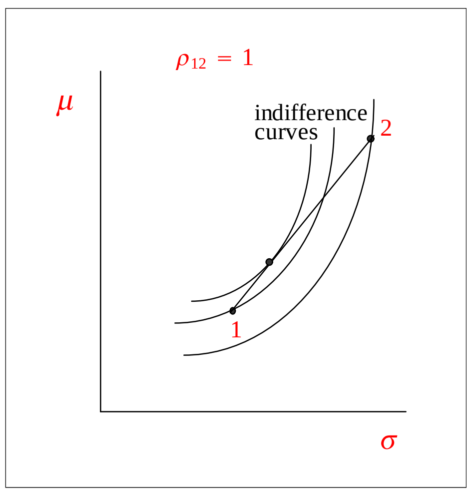
- The portfolio optimization problem considered above provides a point on the *efficient frontier* corresponding to $\mu_p = 0.1$ and $\sigma_p = 0.06082$. The efficient frontier consists of all efficient portfolios; that is, a family of minimum variance portfolios, with each corresponding to a specific expected return no lower than the expected return of the global minimum variance portfolio. To construct the entire efficient frontier, we can set $\mu_p = r$, where the parameter r is given a different value for each optimization problem (based on the same input data), as shown graphically in the following:



5.8 Implications

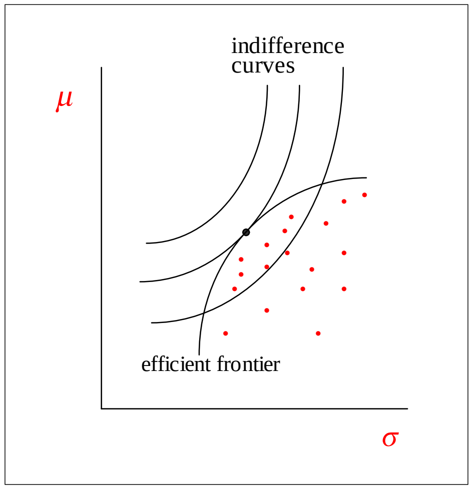
We now explore various implications of the basic analytical materials in mean-variance portfolio theory that we have covered so far. While some implications are quite obvious, some are less obvious. At this stage, our attention is on those implications not relying on the use of any sophisticated portfolio selection models.

- If the returns of two securities are perfectly positively correlated, there are no diversification benefits in terms of risk-return trade-off. However, as shown in the following graph, there can still be a portfolio based on the two securities that provides a higher utility than does an investment in either security alone.



- If the returns of two securities are perfectly negatively correlated, a risk-free portfolio can be formed by allocating investment funds between the two securities in a specific way. (This feature has been mentioned previously, in case 2 of the two-security illustration.)

- As long as the returns of the two securities are not perfectly positively correlated, there are obvious diversification benefits in terms of risk-return trade-off. An extension to portfolios based on more than two risky securities can be captured graphically as follows:



5.9 Implicit Assumptions

In the analytical materials that have been covered so far, the random return of a portfolio is

$$\widetilde{R}_p = \sum_{i=1}^{n} x_i \widetilde{R}_i \,, \tag{5.12}$$

where

$$\sum_{i=1}^{n} x_i = 1.$$

As long as the portfolio weights, $x_1, x_2, \ldots, x_n$, are all non-negative, the analytical expressions

$$\begin{aligned} \mu_p &= \sum_{i=1}^{n} x_i \mu_i \\ \text{and} \quad \sigma_p^2 &= \sum_{i=1}^{n} \sum_{j=1}^{n} x_i x_j \sigma_{ij} \end{aligned}$$

are valid. However, what happens if any of $x_1, x_2, \ldots, x_n$ is/are negative?

The short sale of security i in a portfolio setting is indicated by a negative x_i. For equation (5.12) to accommodate negative portfolio weights, the contribution of each security i to the portfolio's random return would still be $x_i \widetilde{R}_i$ even for $x_i < 0$. Then, in the same analytical setting, the random return for holding a security in a short position would be treated as the negative of the random return for holding the same security in a long position. Since

$$\sum_{i=1}^{n} x_i = 1,$$

a negative x_i would provide

$$\sum_{\substack{j=1 \\ j \neq i}}^{n} x_j > 1.$$

For example, in a four-security case, if $x_3 = -0.2$, then we would have $x_1 + x_2 + x_4 = 1.2$, given that the four portfolio weights sum to 1. If all of securities 1, 2, and 4 are to be held long, the total investment in them would be \$1.20 for each dollar of investment funds. As this example illustrates, an investor who short sells a security is assumed to be able to invest more than a dollar in other securities for each dollar of investment funds. This, obviously, is for analytical convenience. The extra funds are assumed to be generated from the immediate access of the short-sale proceeds. Also for analytical convenience, we assume that the short seller not only provides zero margin deposit but also has immediate access to the short-sale proceeds for investing in other securities. Such simplifying assumptions are usually characterized as **frictionless short sales**.

In reality, when short selling a security, the short seller not only has no immediate access to the short-sale proceeds, but also has to provide a margin deposit equal to 50% or more of the share values of any securities in short positions. There are also various other institutional features that are relevant to portfolio investments. Such practical details are considered next, before we turn our attention to portfolio strategies and models.

5.10 Short Selling of Stocks in Practice

Short selling is about selling a security — typically a common stock — that one does not own. An individual who expects a stock to drop in price may borrow a certificate of the stock from someone and sell it to someone else. If the subsequent price change turns out to be negative as expected, the individual can buy back the same stock in the market at a lower price. Once the newly acquired stock certificate is returned to the original owner, the transactions are completed. The idea is that the transactions have the potential to let the individual make a positive return by "selling high" first and "buying low" later. This sequence is the reverse of the traditional "buy low, sell high" sequence in a profitable stock investment where "buy low" is followed by "sell high." However, short selling is not always profitable. Indeed, as explained later in this section, it is highly risky.

For analytical convenience in portfolio modeling, the assumption of **frictionless short sales** is often used. Under this simplifying assumption, if an investor decides to short sell a stock, the investor can always borrow the stock from others without incurring any cost. The buyer of the borrowed stock pays the prevailing market price at the time, just like buying a stock that is sold by another investor who actually owns it. What the buyer pays for the borrowed stock is called the **short-sale proceeds**.

Under the assumption of frictionless short sales, the short seller is assumed to have immediate access to the short-sale proceeds. It is also assumed that the short seller does not have to provide any deposit — either by cash or by other financial instruments that he/she owns — to the brokerage firm involved. Thus, under the assumption of frictionless short sales, an investor can use the short-sale proceeds to augment the available investment funds, for purchasing other stocks.

However, in reality, short selling does not enable the short seller to generate additional investment funds. Here is how short selling works in reality: Stocks that are available for borrowing are typically those held in **street names** at brokerage firms. That is, stocks held at **margin accounts** of other investors. Short selling of stocks has traditionally been perceived to be highly speculative and has been subject to various regulatory restrictions and institutional biases against its application as an investment tool. The following are some familiar features of **institutional procedures for short selling** that are relevant for portfolio considerations:

1. The short seller does not have immediate access to the proceeds as provided by the buyer of the stock in a short-sale transaction. The short-sale proceeds are held as collateral for the borrowed stock.

2. The short seller usually earns no interest from the short-sale proceeds; the brokerage firm usually retains the full amount of the interest earned. However, a professional short seller with a sizable account may get, often as the result of negotiations, a substantial rebate from the brokerage firm for the interest it earns. The term "haircut" is often used to describe the action "trim off the top" by the brokerage firm, which results in only a partial interest rebate for the short seller.

3. The short seller must deposit with the brokerage firm at least 50% of the value of the shares held short to fulfill the initial margin requirements. The margin as required by individual brokerage firms can be higher.

4. Cash, interest bearing T-bills, and other securities that the short seller owns can be used to provide the required margin deposits. The short seller will earn interests on any cash deposits.

5. Any dividend payments from the issuing company of the stock, which the lender of the stock is entitled to receive, must be provided by the short seller.

6. In the event that the lender sells the stock, the short seller has to obtain the stock certificate at the market price if another loan of the same stock cannot be arranged.

If managed properly, short selling of stocks in a portfolio context may not be speculative. It can actually reduce or even neutralize investors' exposure to market and industry risks. To illustrate, suppose that an investor holds a stock and shorts another stock in the same industry. Suppose also that the long and short sides are of the same dollar value and of the same sensitivity in response to industrial movements. Here, the long side is the part of the portfolio consisting only of purchased stocks. Likewise, the short side is the part of the portfolio consisting of shorted stocks. In such a portfolio, the investor's holdings are unaffected by any industrial movements.

Of course, whether the investment is profitable depends on the investor's skill in stock selection. Long-short investing does not have to be completely market neutral or industry neutral. By combining a long-short portfolio with an investment in a market index, for example, the level of the market exposure can be adjusted to match the preference of the investor.

It must be recognized, however, that short selling, even when it is not intended to be for speculative purposes, is still very risky. Here are some risk factors to consider:

1. Long-short investing requires the portfolio manager's ability to identify both undervalued and overvalued stocks. An obvious risk is the potential misjudgment in each case.

2. Unlike long investments, short selling does not have limited liabilities. The higher the price increases subsequent to the short sale, the more losses will be incurred by the short seller.

3. Timing can be a problem. An overvalued stock may stay overvalued for a long time. Thus, the ability to identify overvalued stocks does not always translate into profits from short selling them.

4. Short squeezes can be triggered either by a company's hostility or aggressiveness towards its short sellers or by other investors' trading activities. Short-slamming tactics work better for stocks of smaller companies.

5. Sudden price increases of companies being taken over by other companies will lead to great losses to short sellers.

6. Short selling in a bullish market is highly risky as most stocks move with the rising tide.

5.11 The $1/N$ Investment Strategy: Equally Weighted Portfolios

For an n-security case, the expected return and the variance of returns of a portfolio p are

$$\mu_p = \sum_{i=1}^{n} x_i\mu_i = x_1\mu_1 + x_2\mu_2 + \cdots + x_n\mu_n$$

and

$$\begin{aligned} \sigma_p^2 &= \sum_{i=1}^{n}\sum_{j=1}^{n} x_i x_j \sigma_{ij} \\ &= x_1x_1\sigma_{11} + x_1x_2\sigma_{12} + \cdots + x_1x_n\sigma_{1n} \\ &\quad + x_2x_1\sigma_{21} + x_2x_2\sigma_{22} + \cdots + x_2x_n\sigma_{2n} \\ &\quad \vdots \\ &\quad + x_nx_1\sigma_{n1} + x_nx_2\sigma_{n2} + \cdots + x_nx_n\sigma_{nn}, \end{aligned}$$

respectively, for which

$$\sum_{i=1}^{n} x_i = x_1 + x_2 + \cdots + x_n = 1.$$

In the expression of σ_p^2, there are n^2 additive terms, with each term being of the form $x_i x_j \sigma_{ij}$. As the number of securities in the portfolio increases, the contributions of the variances of returns by individual securities to the overall portfolio risk becomes less important, and the contributions of the covariances of security returns becomes more important.

The $1/N$ investment strategy is about allocating investment funds equally among the securities considered. For an n-security case, we impose $x_i = 1/n$, for $i = 1, 2, \ldots, n$. The corresponding portfolio is an equally weighted portfolio. The expected portfolio return is

$$\mu_p = \frac{1}{n}(\mu_1 + \mu_2 + \cdots + \mu_n),$$

which is the average of the expected returns of the n securities considered. Let us denote it as $\overline{Ave}$. With all variances and covariances of returns grouped separately, we can write

$$\begin{aligned}\sigma_p^2 &= \frac{1}{n^2}\sum\nolimits_{i=1}^{n}\sum\nolimits_{j=1}^{n}\sigma_{ij} \\ &= \frac{1}{n^2}\left(\sum\nolimits_{i=1}^{n}\sigma_{ii}+\sum\nolimits_{i=1}^{n}\sum\nolimits_{\substack{j=1\\ j\neq i}}^{n}\sigma_{ij}\right).\end{aligned}$$

Let $\overline{Var}$ be the average variance of returns and $\overline{Cov}$ be the average covariance of returns; that is,

$$\overline{Var} = \frac{1}{n}\sum\nolimits_{i=1}^{n}\sigma_{ii}$$

and

$$\overline{Cov} = \frac{1}{n(n-1)}\sum\nolimits_{i=1}^{n}\sum\nolimits_{\substack{j=1\\ j\neq i}}^{n}\sigma_{ij}.$$

In the latter case, there are $(n-1)$ covariances to be considered for each security i. Thus, it follows that

$$\begin{aligned}\sigma_p^2 &= \frac{1}{n^2}\left[n\cdot\overline{Var}+n(n-1)\cdot\overline{Cov}\right] \\ &= \frac{\overline{Var}}{n}+\frac{n-1}{n}\cdot\overline{Cov}.\end{aligned}$$

- For example, if $n=4$, there are 16 $(=4^2)$ variances and covariances of returns in total. Among them are 4 variances and 12 $(=4\times 3)$ covariances. Here, we have

  $$\overline{Var} = \frac{1}{4}(\sigma_{11}+\sigma_{22}+\sigma_{33}+\sigma_{44})$$

  and

  $$\begin{aligned}\overline{Cov} &= \frac{1}{12}\left[(\sigma_{12}+\sigma_{13}+\sigma_{14})+(\sigma_{21}+\sigma_{23}+\sigma_{24})\right. \\ &\quad \left.+(\sigma_{31}+\sigma_{32}+\sigma_{34})+(\sigma_{41}+\sigma_{42}+\sigma_{43})\right].\end{aligned}$$

As $n\to\infty$, $\overline{Var}$ and $\overline{Cov}$ remain finite. Accordingly, as $n\to\infty$, we have

$$\frac{\overline{Var}}{n}\to 0$$

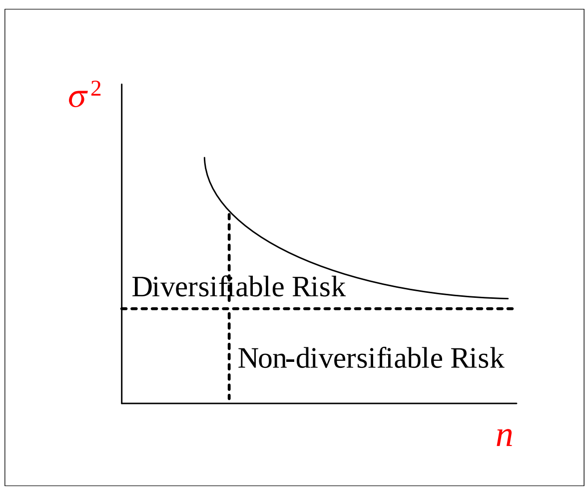
and

$$\sigma_p^2\to\overline{Cov}.$$

The following is a numerical illustration of the relative importance of variance and covariance terms in a portfolio setting:

# of securities n = # of variances	# of covariances $n(n-1)$
5	20
10	90
15	210
20	380
25	600
$\vdots$	$\vdots$
100	$9,900$
$\vdots$	$\vdots$
200	$39,800$

As more securities are added to a portfolio, the contributions to the portfolio risk by the individual variance terms tend to diminish, and the portfolio risk — which is represented by the variance of portfolio returns — asymptotically approaches the average of the covariances of security returns.



The $1/N$ investment strategy is practically relevant, primarily because of its simplicity. There is no need for the investor who use it as an investment tool to pay much attention to the quality of the input parameters, which, in contrast, is highly important for implementing any portfolio optimization models instead. The effectiveness of the $1/N$ investment strategy for the risk-return trade-off in terms of σ_p and μ_p for finite numbers of n in practice — or, equivalently, in terms of $\overline{Ave}$, $\overline{Var}$, and $\overline{Cov}$ — still hinges on the quality of security screening before implementing the strategy. If the securities considered are highly correlated in their returns, the achievable portfolio risk reductions will be limited. Further, the $1/N$ investment strategy is never intended to eliminate the investment risk; even if a large number of securities is involved, the corresponding positive $\overline{Cov}$ ensures that some risk will still remain.

5.12 A Basic Portfolio Selection Model

We now consider a basic portfolio selection model, as developed by Richard Roll [*Journal of Financial Economics*, 4(2), (1977), 129-176, Appendix]. For analytical convenience, we now use matrix notation. Suppose that, for a given set of n risky securities for portfolio investment consideration, we capture succinctly the n expected returns $\mu_1, \mu_2, \ldots, \mu_n$ — which are treated as given parameters — with an n-element column vector

$$\boldsymbol{\mu} = \begin{bmatrix} \mu_1 & \mu_2 & \cdots & \mu_n \end{bmatrix}',$$

where the prime denotes matrix transposition.

We also have given parameters σ_{ij}, for $i, j = 1, 2, \ldots, n$, called variances of returns if $i = j$ and covariances of returns if $i \neq j$. In matrix notation, we can write

$$\boldsymbol{V} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_{nn} \end{bmatrix},$$

which is an $n \times n$ symmetric matrix, commonly known as the *covariance matrix of returns* or, simply, the *covariance matrix*. It is symmetric, because

$$\sigma_{ij} = \sigma_{ji}, \text{ for } i, j = 1, 2, \ldots, n \text{ and } i \neq j.$$

Let $x_1, x_2, \ldots, x_n$ be the proportions of investment funds on the individual securities; they are commonly called *portfolio weights*. We use an n-element column vector

$$\boldsymbol{x} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}',$$

for which

$$\boldsymbol{x}'\boldsymbol{\iota} = 1$$

where $\boldsymbol{\iota}$ is an n-element column vector of ones, to capture such proportions.

Under the assumption of frictionless short sales, if a portfolio p is formed, its expected return and variance of returns are

$$\begin{aligned} \mu_p &= \sum_{i=1}^{n} x_i \mu_i = \boldsymbol{x}'\boldsymbol{\mu} \\ \text{and } \sigma_p^2 &= \sum_{i=1}^{n} \sum_{j=1}^{n} x_i x_j \sigma_{ij} = \boldsymbol{x}'\boldsymbol{V}\boldsymbol{x}, \end{aligned}$$

respectively. As derived in Appendix A of this chapter, the set of portfolio weights that minimizes the variance of portfolio returns for a given portfolio expected return μ_p is

$$\boxed{\boldsymbol{x} = \boldsymbol{V}^{-1}\boldsymbol{M}\,(\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}\boldsymbol{r}_p}, \tag{5.13}$$

where

$$\boldsymbol{M} = \begin{bmatrix} \mu_1 & 1 \\ \mu_2 & 1 \\ \vdots & \vdots \\ \mu_n & 1 \end{bmatrix}$$

is an $n \times 2$ matrix and

$$\boldsymbol{r}_p = \begin{bmatrix} \mu_p \\ 1 \end{bmatrix}$$

is a 2-element column vector.

- The case where $\mu_1 = \mu_2 = \cdots = \mu_n$ can be ignored. This is the case where there is no flexibility in specifying the expected portfolio return — as we must have $\mu_p = \mu_1 = \mu_2 = \cdots = \mu_n$ — thus rendering the resulting portfolio selection model useless.

The variance of returns of portfolio p is

$$\boxed{\sigma_p^2 = \boldsymbol{x}'\boldsymbol{V}\boldsymbol{x} = \boldsymbol{r}_p'(\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}\boldsymbol{r}_p}.$$

For the same set of n risky securities, if the set of portfolio weights

$$\boldsymbol{y} = \begin{bmatrix} y_1 & y_2 & \cdots & y_n \end{bmatrix}',$$

for which

$$\boldsymbol{y}'\boldsymbol{\iota} = 1,$$

that minimizes the variance of portfolio returns for a given portfolio expected return μ_q is

$$\boldsymbol{y} = \boldsymbol{V}^{-1}\boldsymbol{M}\,(\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}\boldsymbol{r}_q,$$

where

$$\boldsymbol{r}_q = \begin{bmatrix} \mu_q \\ 1 \end{bmatrix}.$$

the covariance of returns between portfolios p and q is

$$\boxed{\sigma_{pq} = \boldsymbol{x}'\boldsymbol{V}\boldsymbol{y} = \boldsymbol{r}_p'(\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}\boldsymbol{r}_q}.$$

Further, if we write

$$\boldsymbol{V}^{-1}\boldsymbol{M}\,(\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1} = \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \\ \vdots & \vdots \\ a_n & b_n \end{bmatrix},$$

where $a_1, a_2, \ldots, a_n$ and $b_1, b_2, \ldots, b_n$ are computed values for given $\boldsymbol{V}$ and $\boldsymbol{M}$, we must have

$$\boxed{a_1 + a_2 + \cdots + a_n = 0}$$

and

$$\boxed{b_1 + b_2 + \cdots + b_n = 1}.$$

We can use these two sums to check whether the computations, for a given set of input parameters, are performed correctly.

5.12.1 The Global Minimum Variance Portfolio

Under the assumption of frictionless short sales, the global minimum variance portfolio based on the same n risky securities is the portfolio that provides the lowest variance of portfolio returns (which is $\boldsymbol{x}'\boldsymbol{V}\boldsymbol{x}$) under the condition that investment funds be fully allocated (that is, $\boldsymbol{x}'\boldsymbol{\iota} = 1$). However, the expected portfolio return need not be specified in advance here. Let μ_o, σ_o^2, and $\boldsymbol{x}_o$ be the expected portfolio return, the variance of portfolio returns, and an n-element column vector of portfolio weights, respectively, which can be deduced for such a portfolio.

As derived also in Appendix A, we have

$$\begin{aligned} \boldsymbol{x}_o &= \boldsymbol{V}^{-1}\boldsymbol{\iota}(\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\iota})^{-1}, \\ \mu_o &= \boldsymbol{x}_o'\boldsymbol{\mu} = (\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\iota})^{-1}(\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\mu}), \\ \text{and } \sigma_o^2 &= \boldsymbol{x}_o'\boldsymbol{V}\boldsymbol{x}_o = \boxed{(\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\iota})^{-1}}. \end{aligned}$$

Notice that $\boldsymbol{V}^{-1}\boldsymbol{\iota}$ is an n-element column vector with each element i being the sum of the n elements in row i of $\boldsymbol{V}^{-1}$. Notice also that $\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\iota}$ is the sum of all elements of $\boldsymbol{V}^{-1}\boldsymbol{\iota}$. The above analytical results reveal that (1) the proportion of investment funds allocated to security i in such a portfolio is the sum of the n elements in row i of $\boldsymbol{V}^{-1}$, divided by the sum of all n^2 elements of $\boldsymbol{V}^{-1}$ and (2) the variance of returns of such a portfolio is the reciprocal of the sum of all n^2 elements of $\boldsymbol{V}^{-1}$.

Under the assumption of frictionless short sales, if the proportions of investment funds for any portfolio p, as determined according to equation (5.13), lead to $\mu_p > \mu_o$, the portfolio is an efficient portfolio. If $\mu_p < \mu_o$ instead, the portfolio is inefficient. Further, for any portfolio k based on the same set of n securities, regardless of whether it is efficient or inefficient, and regardless of whether its inefficiency is caused by not using equation (5.13) for determining its portfolio allocations, we also have

$$\sigma_{ok} = \boxed{(\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\iota})^{-1}}.$$

For the same example in Section 5.7, as

$$\boldsymbol{V} = \begin{bmatrix} 0.0004 & 0.0005 & 0 \\ 0.0005 & 0.0025 & 0.0015 \\ 0 & 0.0015 & 0.0225 \end{bmatrix},$$

we have

$$\boldsymbol{V}^{-1} = \begin{bmatrix} 3380.28169 & -704.22535 & 46.94836 \\ -704.22535 & 563.38028 & -37.55869 \\ 46.94836 & -37.55869 & 46.94836 \end{bmatrix},$$

$$\boldsymbol{V}^{-1}\boldsymbol{\iota} = \begin{bmatrix} (\text{sum of row 1 of } \boldsymbol{V}^{-1}) \\ (\text{sum of row 2 of } \boldsymbol{V}^{-1}) \\ (\text{sum of row 3 of } \boldsymbol{V}^{-1}) \end{bmatrix} = \begin{bmatrix} 2723.00470 \\ -178.40376 \\ 56.33803 \end{bmatrix},$$

$$\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\iota} = \text{ sum of all 3 elements of } \boldsymbol{V}^{-1}\boldsymbol{\iota} = 2600.93897,$$

$$\sigma_o^2 = (\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\iota})^{-1} = \frac{1}{2600.93897} = 0.00038448,$$

and

$$\sigma_o = 0.019608,$$

which has the same computed value as the σ_p displayed at the end of **Part One** of the example in Section 5.7. We also have

$$\boldsymbol{x}_o = \boldsymbol{V}^{-1}\boldsymbol{\iota}(\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\iota})^{-1} = \begin{bmatrix} 2723.00470 \\ -178.40376 \\ 56.33803 \end{bmatrix} \frac{1}{2600.93897} = \begin{bmatrix} 1.046931 \\ -0.06859 \\ 0.02166 \end{bmatrix}.$$

These three portfolio weights have the same computed values as the corresponding x_1, x_2, and x_3 displayed there as well.

Now, suppose that an equally weighted portfolio k based on the same three securities is formed. Denoting y_1, y_2, and y_3 as the corresponding equal portfolio weights. For

$$\boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1/3 \\ 1/3 \\ 1/3 \end{bmatrix},$$

we can write

$$\begin{aligned}
\sigma_{ok} &= \sigma_{ko} = \boldsymbol{y}'\boldsymbol{V}\boldsymbol{x}_o = \boldsymbol{y}'\boldsymbol{V}[\boldsymbol{V}^{-1}\boldsymbol{\iota}(\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\iota})^{-1}] = \boldsymbol{y}'\boldsymbol{\iota}(\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\iota})^{-1} \\
&= \begin{bmatrix} 1/3 & 1/3 & 1/3 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} (\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\iota})^{-1} \\
&= (\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\iota})^{-1} = \sigma_o^2 = 0.00038448.
\end{aligned}$$

5.12.2 Remarks

The portfolio k in the above illustration can also be any portfolio based on the same set of n risky securities, where $n \geq 2$. That is, it need not be an equally weighted portfolio. Regardless of how portfolio k is constructed, as the condition of $\boldsymbol{y}'\boldsymbol{\iota} = 1$ always holds under the assumption of frictionless short sales, the outcome of

$$\sigma_{ok} = (\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\iota})^{-1} = \sigma_o^2$$

is assured. In contrast, also under the assumption of frictionless short sales, the equation

$$\sigma_{pq} = \boldsymbol{r}_p'(\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}\boldsymbol{r}_q$$

holds if each of portfolios p and q has been constructed by minimizing the variance of portfolio returns given an expected return requirement instead. In such a case, each of portfolios p and q is a *minimum variance portfolio*. Each minimum variance portfolio with its expected return no lower than the expected return of the global minimum variance portfolio is an *efficient portfolio*. For a given set of securities, the graph on the (σ, μ)-plane containing all efficient portfolios is the *efficient frontier*.

At this point, we have not addressed the issue of *idle cash*, which pertains to situations where available investment funds are not fully allocated among the risky securities considered. We can address the issue as soon as we have introduced the basic portfolio concepts for including a risk-free security for portfolio considerations in the next section.

5.13 The Presence of a Risk-Free Security

Consider a risk-free security f and a portfolio p based on n risky securities. Security f being risk-free, its return R_f is a constant and its variance of returns is zero. Also, it has zero covariances of returns with all securities and all portfolios. Now, let us form a portfolio q based on security f and portfolio p. Suppose that a is the proportion of investment funds allocated to security f and the balance, $1 - a$, allocated to portfolio p. We can write

$$\begin{aligned} \widetilde{R}_q &= aR_f + (1-a)\widetilde{R}_p, \\ \mu_q &= aR_f + (1-a)\mu_p, \\ \text{and} \quad \sigma_q^2 &= E\left[\left(\widetilde{R}_q - \mu_q\right)^2\right] = E\left[(1-a)^2\left(\widetilde{R}_p - \mu_p\right)^2\right] \\ &= (1-a)^2 E\left[\left(\widetilde{R}_p - \mu_p\right)^2\right] = (1-a)^2\sigma_p^2. \end{aligned} \tag{5.14}$$

Here, $\widetilde{R}_p$ and $\widetilde{R}_q$ denote the random returns of portfolios p and q, respectively, and $E[\cdot]$ stands for the expected value of the random variable that $[\cdot]$ represents.

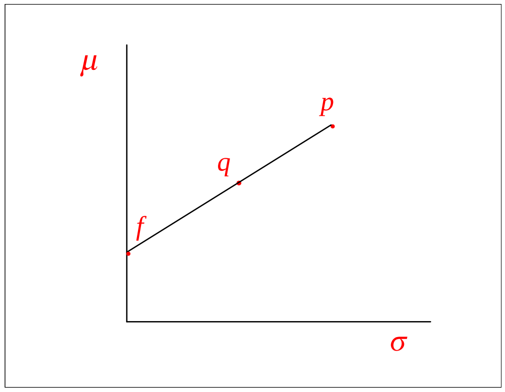
For $0 \le a \le 1$, we have

$$\sigma_q = (1-a)\sigma_p. \tag{5.15}$$

Combining (5.14) and (5.15) to eliminate a leads to

$$\mu_q = R_f + \left(\frac{\mu_p - R_f}{\sigma_p}\right)\sigma_q,$$

which is a linear relationship between μ_q and σ_q as shown graphically in the following:



5.13.1 Implications

The following are some implications of the presence of a risk-free security in portfolio investments:

1. Combining a risky portfolio and a risk-free security provides a linear risk-return trade-off. On the (σ, μ)-plane where the μ-axis is vertical and the σ-axis is horizontal as shown in the graph below, let us draw a line that passes through the point $(0, R_f)$, and let θ be the slope of the line. As the line meets the efficient frontier based on the n risky securities, the point of intersection represents a portfolio with a particular combination of μ and σ. The tangency portfolio p is the portfolio on the efficient frontier based on the n risky securities where θ is maximized.

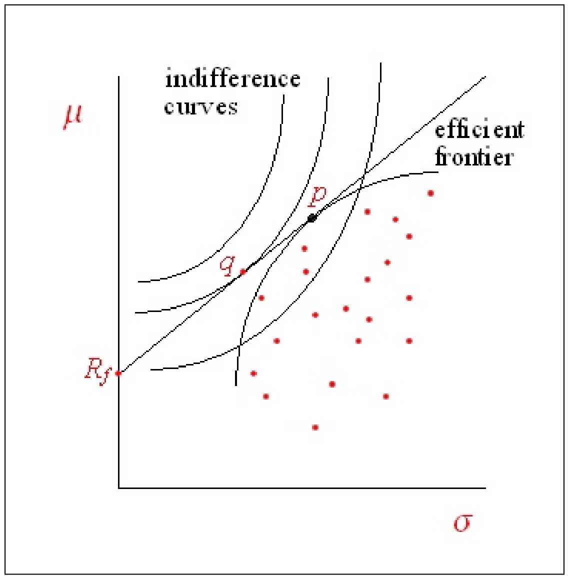
2. By assuming that the short seller has immediate access to the short-sale proceeds, the tangency portfolio p can be constructed from the following optimization:

$$\underset{\{x_1, x_2, \ldots, x_n\}}{\text{Maximize}} \theta = \frac{\mu_p - R_f}{\sigma_p} = \frac{\sum_{i=1}^{n} x_i \mu_i - R_f}{\sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} x_i x_j \sigma_{ij}}}$$

subject to $\sum_{i=1}^{n} x_i = 1.$

The corresponding portfolio selection model, which leads to a market equilibrium model, is covered in Chapter 6.

3. As shown in the following graph, the optimal allocation of investment funds between the risk-free security and the tangency portfolio p depends on the investor's utility function:



The line from the point $(0, R_f)$ on the (σ, μ)-plane that is tangent to the efficient frontier based on the n risky securities is the efficient frontier based on the risk-free security and the n risky securities. The tangency portfolio p is the optimal portfolio based on the n risky securities alone.

A portfolio q, which corresponds to the highest achievable utility for the investor, is a combination of the risk-free security and the tangency portfolio p. While the choice of portfolio q depends on the investor's utility function, the determination of the tangency portfolio p does not.

To achieve a portfolio q on the line between the point $(0, R_f)$ and the tangency portfolio p, the investor allocates his/her investment funds between the risk-free security and portfolio p. Assuming risk-free borrowing at the same interest rate R_f, a portfolio q on the line above portfolio p is the case where the investor borrows additional funds for investing more in portfolio p.

4. The separation between (i) the construction of the optimal portfolio based on the n risky securities and (ii) the decision about the allocations of investment funds between the risk-free security and the risky portfolio thus constructed makes it possible to construct portfolios based on risky securities without paying any attention to the investor's

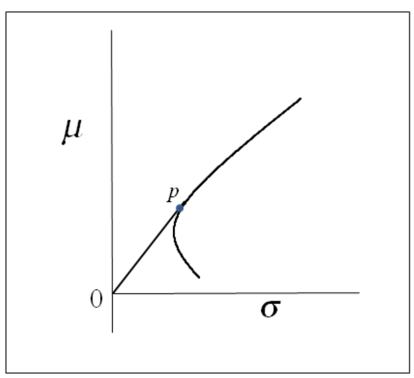
specific risk-return preferences.

5. If the interest rate for risk-free lending is lower than the interest rate for risk-free borrowing, the separation as described above will no longer hold. In such a case, the efficient frontier will consist of a non-linear segment with two linear extensions; the optimal portfolio based on the n risky securities will depend on the investor's specific risk-return preferences.

5.13.2 Idle Cash

If the risk-free security is unavailable, the effect of holding *idle cash* can be interpreted in terms of allocating some of the available investment funds to a risk-free security that earns no returns. Efficient allocations of investment funds this way can be captured graphically by drawing a tangent line from the origin of the (σ, μ)-plane to the efficient frontier. The part of the curve between the tangency portfolio and the global minimum variance portfolio — which is originally considered efficient for full allocations of investment funds — is no longer efficient. The reason is that this part of the curve, which is below the tangent line on the (σ, μ)-plane is dominated by the tangent line segment.

The line segment from the origin of the (σ, μ)-plane to point p in the following graph shows the risk-return trade-off for allocating only a proportion of the available investment funds to a risky portfolio p, with the balance held in the form of idle cash:



In the absence of a risk-free security, all minimum variance portfolios based on the same set of risky securities, as shown on the (σ, μ)-plane, are on a branch of a hyperbola under the assumption of frictionless short sales. The efficient frontier is the upward sloping part of this branch of the *hyperbola* starting from its *vertex*, which represents the global minimum variance portfolio. The efficient frontier has been extended (improved) to contain a tangent line segment, drawn from the origin to the tangency portfolio p.

Whether this linear part of the extended efficient frontier is relevant for the investor depends on the investor's risk-return preference, as captured by a set of indifference curves not shown in the graph. If the highest achievable utility corresponds to holding an efficient portfolio with an expected return greater than that of the tangency portfolio p, then the investor will not hold any idle cash. Otherwise, holding idle cash will be rational.

5.14 Tangency Portfolio Approach for Portfolio Selection

The same efficient frontier on the (σ, μ)-plane according to the basic portfolio selection model (based on Roll's work) in Section 5.12 can also be reached by drawing tangent lines with different values of the μ-intercept. Let us denote the μ-intercept of each tangent line as r. To determine the corresponding n-element column vector $\boldsymbol{x}$ of efficient portfolio weights, we follow John Lintner (1916-1983) to maximize the slope of the tangent line

$$\theta = \frac{\mu_p - r}{\sigma_p},$$

where

$$\mu_p = \sum\nolimits_{i=1}^{n} x_i \mu_i$$

and

$$\sigma_p = \sqrt{\sum\nolimits_{i=1}^{n} \sum\nolimits_{j=1}^{n} x_i x_j \sigma_{ij}},$$

subject to the condition of

$$\sum\nolimits_{i=1}^{n} x_i = 1.$$

5.14.1 The Main Analytical result

As derived in Appendix B, the main analytical result is

$$x_i = \frac{z_i}{\sum_{j=1}^{n} z_j}, \text{ for } i = 1, 2, \ldots, n,$$

where z_i is element i of the n-element column vector

$$\boldsymbol{z} = \boldsymbol{V}^{-1}(\boldsymbol{\mu} - r\boldsymbol{\iota}).$$

With $z_1, z_2, \ldots, z_n$ and $x_1, x_2, \ldots, x_n$ established, the expected portfolio return and standard deviation of returns can be computed from

$$\mu_p = \sum\nolimits_{i=1}^{n} x_i \mu_i$$

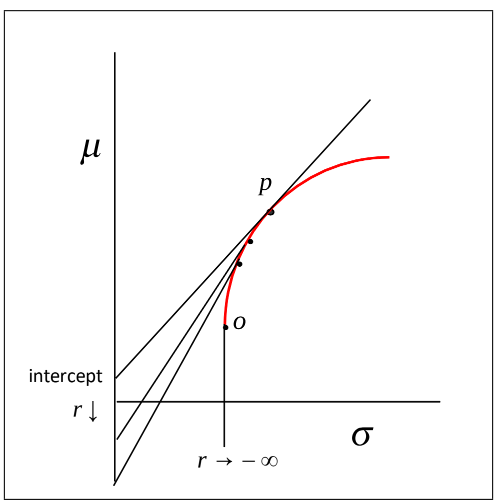
and

$$\sigma_p = \sqrt{\frac{\mu_p - r}{\sum\nolimits_{j=1}^{n} z_j}},$$

respectively.

For the above analytical result to be meaningful, r must be lower than the expected return μ_o of the global minimum variance portfolio based on the n securities. If $r > \mu_o$, we have minimization of $\theta = (\mu_p - r)/\sigma_p$ instead. If $r = \mu_o$, portfolio optimization cannot be performed.

Thus, by allowing r to decrease gradually from an initial value that is slightly less than μ_o, the entire efficient frontier can be established. To reach the global minimum variance portfolio, we let $r = -N$, where N is a very large positive number. The whole idea can be captured graphically as follows:



As $r \to -\infty$ (that is, as $N \to \infty$), the corresponding tangency portfolio approaches the one obtained from

$$\boldsymbol{x}_o = \boldsymbol{V}^{-1}\boldsymbol{\iota}\,(\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\iota})^{-1}.$$

- To verify this, we allow $N\boldsymbol{\iota}$ to become so large that $\boldsymbol{\mu}$ becomes very small by comparison. As

  $$\boldsymbol{\mu} - r\boldsymbol{\iota} = \boldsymbol{\mu} + N\boldsymbol{\iota} \approx N\boldsymbol{\iota}$$

  and

  $$\boldsymbol{z} = \boldsymbol{V}^{-1}(\boldsymbol{\mu} - r\boldsymbol{\iota}) = \boldsymbol{V}^{-1}(\boldsymbol{\mu} + N\boldsymbol{\iota}) \approx \boldsymbol{V}^{-1}N\boldsymbol{\iota} = N\boldsymbol{V}^{-1}\boldsymbol{\iota},$$

  we can write

  $$\boldsymbol{x}_o = \boldsymbol{z}(\boldsymbol{\iota}'\boldsymbol{z})^{-1} = (N\boldsymbol{V}^{-1}\boldsymbol{\iota})(\boldsymbol{\iota}'N\boldsymbol{V}^{-1}\boldsymbol{\iota})^{-1} = \boldsymbol{V}^{-1}\boldsymbol{\iota}\,(\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\iota})^{-1},$$

  with the scalar N eliminated from the expression.

5.14.2 Relevance of the Tangency Portfolio Approach

A special case of the tangency portfolio approach is where r has the same value as the risk-free interest rate R_f. In such a case, we look for a specific efficient portfolio based on a given set of risky securities; investors then allocate their available investment funds between this efficient portfolio and the risk-free security to achieve utility maximization. The case of $r = R_f$ also has asset-pricing implications to be addressed in Chapter 6.

A derivation of the Capital Asset Pricing Model (CAPM) based on the Lintner approach, to be covered in Chapter 6, is by specifying that the μ-intercept of the tangent line be the risk-free interest rate R_f. This seminal work by John Linter was published in *Review of Economics and Statistics*, (1965), 47(1), 13-37. The analytical result of slope maximization, when combined with market equilibrium considerations, is crucial in the Lintner derivation. The corresponding equilibrium concepts, as well as the significance of the CAPM in modern finance will be covered in Chapter 6.

In the above tangency portfolio formulation, which is intended to be equivalent to the Roll model for efficient portfolio selection, the common condition of

$$\sum\nolimits_{i=1}^{n} x_i = 1$$

implies that additional investment funds are made available by short selling a security. Such a condition in a normative setting has inevitably weakened the practical appeal of either formulation of the same model. However, such a condition can be substituted by

$$\sum_{i=1}^{n} |x_i| = 1$$

in the tangency portfolio formulation with only minor changes to the analytical expressions involved. Under the latter condition, which is known as Lintner's treatment of short sales, a short seller not only has no access to the short-sale proceeds, but also is required to provide to an escrow account a deposit that is equal to the full amount of short-sale proceeds; however, the short seller earns interests on both the deposit and the short-sale proceeds. Although Lintner's treatment of short sales still departs from reality, it is a closer characterization of reality without sacrificing analytical tractability.

- Under Lintner's treatment of short sales, the focus is on $r = R_f$. As

  $$x_i = \frac{z_i}{\sum_{j=1}^{n} |z_j|}, \text{ for } i = 1, 2, \ldots, n,$$

  where z_i is element i of the n-element column vector

  $$\boldsymbol{z} = \boldsymbol{V}^{-1}(\boldsymbol{\mu} - R_f \boldsymbol{\iota}),$$

  the expected portfolio return μ_p and standard deviation of returns σ_p can be computed from

  $$\mu_p - R_f = \sum_{i=1}^{n} x_i(\mu_i - R_f)$$

  and

  $$\sigma_p = \sqrt{\frac{\mu_p - R_f}{\sum_{j=1}^{n} |z_j|}},$$

  respectively. A derivation of the above analytical result is also provided in Appendix B.

5.15 A Brief Description of the Markowitz Critical Line Method

For analytical convenience, we have considered portfolio selection models under the assumption of frictionless short sales. Though analytically convenient, such models often produce impractical results, where there is heavy reliance on the short-sale proceeds as additional investment funds. Going from frictionless short sales to Lintner's treatment of short sales in portfolio modeling is an improvement. However, as most professionally managed investment portfolios do not involve short sales, either by choice or due to institutional restrictions, portfolio selection models without short sales are more practical.

From an analytical perspective, going from frictionless short sales to disallowance of short sales is far from being a minor extension, even in the absence of any other constraints for portfolio construction. This is because, for a given expected portfolio return, while the former case involves only direct substitutions of the input parameters into a derived formula, an iterative procedure is required in the latter case to determine which of the securities considered are to be selected. A brief description of the iterative procedure, known as the *Markowitz critical line method*, is provided below.

In the absence of short sales, let $x_i \geq 0$ be the proportion of investment funds for security i. Under the condition of $\sum_{i=1}^{n} x_i = 1$ for full allocations of investment funds among the n securities considered, the decision problem here can be formulated as constrained minimization of

$$\sum_{i=1}^{n} \sum_{j=1}^{n} x_i x_j \sigma_{ij} - \lambda \sum_{i=1}^{n} x_i \mu_i,$$

where $\lambda \geq 0$ is a parameter that quantifies the attitude towards risk of the investor involved. The higher the value of λ, the more risk tolerant is the investor involved. The efficient frontier is a graph of standard deviations of returns (on the horizontal axis) versus expected returns (on the vertical axis) of the best achievable portfolios for all $\lambda \geq 0$.

A security considered is either *in* (selected) or *out* (not selected), depending on whether it receives a positive or zero proportion of investment funds. The Markowitz critical line method identifies iteratively the *in/out* status of each security for any given value of $\lambda \geq 0$. It also allocates investment funds efficiently among the securities that are *in*. As the efficient frontier can be viewed as a collection of portfolios for investors with different levels of risk tolerance, decreases in λ from an infinitely large initial value correspond to downward movements on the efficient frontier. Such movements will encounter a series of corner portfolios, each of which is an efficient portfolio where an *in/out* status change of a security occurs. Between any two adjacent corner portfolios on the efficient frontier, the *in/out* status of each security remains unchanged.

The Markowitz algorithm starts with the portfolio decision by an extremely risk tolerant investor, for whom λ is infinitely large. For such an investor, the security with the highest expected return among the n securities considered is the only security that is *in*. Decreasing λ from this initial value will eventually lead to a critical value that calls for the selection of a second security, which has the effect of reducing both the expected return and the variance of returns of the portfolio. A further decrease of λ will lead to another critical value, at which the status change of yet another security is warranted. As the procedure continues, a selected security may have to exit from the portfolio at some critical value of λ. The procedure, therefore, will provide the efficient frontier and a series of critical values of λ, with each involving an *in/out* status change of a security. The procedure stops when the portfolio corresponding to $\lambda = 0$ is reached.

- An alternative to the Markowitz approach under a simple characterization of the covariance matrix of returns is considered in Appendix E.

5.16 Input Parameters for Portfolio Selection Models

To implement any portfolio selection model, the true values of the input parameters μ_i and σ_{ij}, for $i, j = 1, 2, \ldots, n$ are unknown and must be estimated. As the usefulness of each portfolio selection model is affected by the quality of its input parameters, parameter estimation is an important component of the work involved. Typically, the estimation involves the use of historical return data. The choice of the sample period is often a trade-off between having a longer sample to reduce estimation errors and having a shorter sample to avoid unmanageable changes in the probability distributions of security returns. This section covers a basic estimation method, addresses some practical issues, and provides a sketch of a remedial measure for a potential problem.

5.16.1 Sample Means, Variances, and Covariances of Returns

Consider a set of n securities, for which we have T monthly return observations. For each security i, let R_{it} be the return of security i as observed in month t, for $i = 1, 2, \ldots, n$ and

$t = 1, 2, \ldots, T$. The observations can be captured as follows:

$$\begin{array}{ll} R_{11}, R_{12}, \ldots, R_{1T} & \text{for security 1} \\ R_{21}, R_{22}, \ldots, R_{2T} & \text{for security 2} \\ \vdots & \\ R_{n1}, R_{n2}, \ldots, R_{nT} & \text{for security } n \end{array}$$

The simplest method to estimate input parameters for portfolio models is that we use the sample mean return,

$$\overline{R}_i = \frac{1}{T} \sum\nolimits_{t=1}^{T} R_{it},$$

as an estimate of the true expected return μ_i, for $i = 1, 2, \ldots, n$. We also use the sample covariance of returns,

$$\widehat{\sigma}_{ij} = \frac{1}{T-1} \sum\nolimits_{t=1}^{T} \left(R_{it} - \overline{R}_i \right) \left(R_{jt} - \overline{R}_j \right),$$

as an estimate of the true covariance of returns between securities i and j, for $i, j = 1, 2, \ldots, n$. As we have learned in statistics, the use of $1/(T-1)$ instead of $1/T$ is to correct estimation bias. Implicitly, if $i = j$, we use the sample variance of returns,

$$\widehat{\sigma}_{ii} = \widehat{\sigma}_i^2 = \frac{1}{T-1} \sum\nolimits_{t=1}^{T} \left(R_{it} - \overline{R}_i \right)^2,$$

as an estimate of the true variance of returns σ_i^2. The sample correlation of returns between securities i and j is estimated from

$$\widehat{\rho}_{ij} = \frac{\widehat{\sigma}_{ij}}{\widehat{\sigma}_i \widehat{\sigma}_j}.$$

In matrix notation, the sample mean return vector $\overline{\boldsymbol{R}}$ is an n-element vector with it elements being $\overline{R}_1, \overline{R}_2, \ldots, \overline{R}_n$. The sample covariance matrix $\widehat{\boldsymbol{V}}$ is an $n \times n$ matrix with each (i,j)-element being $\widehat{\sigma}_{ij}$, for $i, j = 1, 2, \ldots, n$. The sample correlation matrix $\widehat{\boldsymbol{\rho}}$ is an $n \times n$ matrix with each (i,j)-element being $\widehat{\rho}_{ij}$, for $i, j = 1, 2, \ldots, n$, where $\widehat{\rho}_{ii} = 1$. The n-element vector $\widehat{\boldsymbol{\sigma}}$ with its elements being $\widehat{\sigma}_1, \widehat{\sigma}_2, \ldots, \widehat{\sigma}_n$ consists of the square root of the individual diagonal elements of $\widehat{\boldsymbol{V}}$. Both $\widehat{\boldsymbol{V}}$ and $\widehat{\boldsymbol{\rho}}$ are symmetric matrices. We use $\overline{\boldsymbol{R}}$ and $\widehat{\boldsymbol{V}}$ as estimates of the expected return vector and the the covariance matrix of returns. As $\widehat{\boldsymbol{\sigma}}$ and $\widehat{\boldsymbol{\rho}}$ can be used to deduce $\widehat{\boldsymbol{V}}$, we can also use $\overline{\boldsymbol{R}}$, $\widehat{\boldsymbol{\sigma}}$, and $\widehat{\boldsymbol{\rho}}$ as input parameters instead. In either case, we implicitly assume that the probability distribution will continue to remain the same at the time when portfolio decisions are to be made. The choice of the length of the sample period is often a trade-off between (i) having a long enough sample to reduce estimation errors and (ii) having a short enough sample to avoid potential violations the stationary assumption.

5.16.2 Revisions of Sample Estimates of Input Parameters

In practice, the use of past return observations to form expectations can also serve as a starting point in generating a set of acceptable input parameters. This is because insights of financial analysts are often deemed necessary for revising the estimated values of the input parameters. For an n-security case, the individual input parameters to be estimated include n expected returns, n standard deviations of returns, and $n(n-1)/2$ correlations of returns, given the the symmetry of covariance and correlation matrices. A relevant question is whether it is practical for financial analysts to revise the entire set of estimated input parameters or only part of it.

Take, for example, the case where $n = 50$, which is not a large number for the size of a professionally managed stock portfolio. The insights of a group of analysts who track the financial data of the 50 companies considered, collectively, can lead to some improvements of the 50-element vector of sample average returns — as proxies for the corresponding expected returns — and the 50-element vector of sample standard deviations of returns. However, to revise the $1,225$ individual correlation coefficients, even partially, based on analysts' insights is a highly burdensome task. For any larger n, the enormity of the task will be even more overwhelming. Such practical concerns have made it highly unlikely that estimations of the correlation matrix can be based on approaches other than those relying on past return observations.

Still, if there are any changes to the elements of $\widehat{\boldsymbol{V}}$ for potential improvements in the quality of the estimated input parameters, whether the revised $\widehat{\boldsymbol{V}}$ — which is no longer the sample covariance matrix — will remain positive semi-definite cannot be assured. Likewise, if the revised $\widehat{\boldsymbol{V}}$ is invertible, there is no guarantee that it is positive definite. Thus, after any changes to $\widehat{\boldsymbol{V}}$, its positive definiteness must be verified, before it can be used for implementing a basic portfolio selection model. The task can be performed by using *Sylvester's criterion*, which states that a real symmetric matrix is positive definite if and only if all of its *leading principal minors* — each of which is the determinant of a square sub-matrix containing the $(1,1)$-element — are positive.

- To illustrate Sylvester's criterion, let us return to the same three-security example in

Section 5.7. In that example, the covariance matrix of returns is

$$\begin{bmatrix} 0.0004 & 0.0005 & 0.0000 \\ 0.0005 & 0.0025 & 0.0015 \\ 0.0000 & 0.0015 & 0.0225 \end{bmatrix}.$$

The three leading principal minors — which are all determinants — are

$$|0.0004| = 0.0004 > 0,$$

$$\begin{vmatrix} 0.0004 & 0.0005 \\ 0.0005 & 0.0025 \end{vmatrix} = 7.500 \times 10^{-7} > 0,$$

and

$$\begin{vmatrix} 0.0004 & 0.0005 & 0.0000 \\ 0.0005 & 0.0025 & 0.0015 \\ 0.0000 & 0.0015 & 0.0225 \end{vmatrix} = 1.5975 \times 10^{-8} > 0,$$

thus confirming the positive definiteness of the covariance matrix.

- The issue as to whether the sample covariance matrix of returns is positive definite is considered in Appendix C.

5.16.3 A Remedial Measure for Insufficient Return Observations: Shrinkage Estimation

Momentum investing is an investment strategy that pays special attention to the market momentum. Suppose that a security exhibits over a period of time some persistent upward price movements, which have resulted in observations of many repeated positive returns. The pursuit of market momentum being of considerable interest to many investment practitioners, they may be tempted to rely on short samples of security returns to estimate input parameters for implementing practical portfolio selection models. The problems with using short samples, however, are two-fold: (*i*) severe estimation errors and (*ii*) the failure to satisfy the positive definiteness requirement for the covariance matrix of returns.

Shrinkage estimation is a good remedy for the above problems. In essence, shrinkage estimation is about taking a weighted average of the sample covariance matrix of returns and a structured matrix of the same dimensions. A suitable structured matrix is a covariance matrix of returns where all covariances (excluding the variances) are characterized by a constant correlation, such as the sample mean of all pairwise correlations between different securities. The use of such a structured matrix will ensure that the resulting covariance matrix of returns be always positive definite. A sketch of shrinkage estimation is provided in Appendix D.

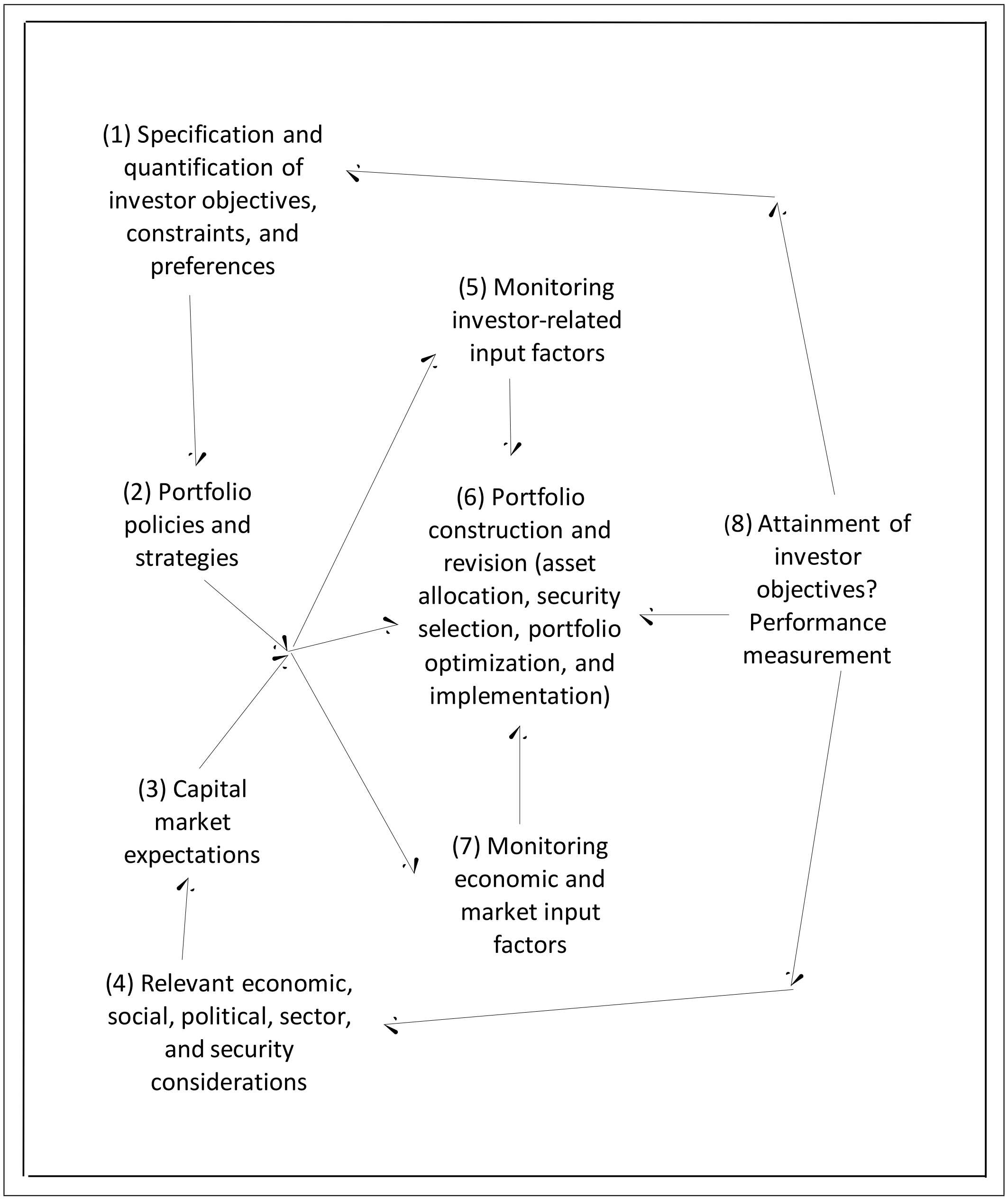
5.17 The Portfolio Management Process

The portfolio management process starts with a detailed knowledge of the investment attitude and preferences of the investor, the institutional environment, and the current and anticipated economic conditions. Information pertaining to the risk-return preference of the investor and relevant institutional constraints allows us to formulate an appropriate analytical framework. Information pertaining to market conditions enables us to generate the necessary input data for the analysis.

An important part of the portfolio management process is to identify the specific assets to be included in an investment portfolio and to determine the proportions of investment funds to be allocated to the individual assets considered. In a quantitative approach, we rely on optimization of certain criterion for such a task. Noting that investor preferences can change, and the market conditions do change, sometimes drastically, the investment portfolios need to be assessed periodically. The issue as to whether the investor objectives have been achieved has to be addressed. Objective measures of portfolio performance, in terms of risk-return trade-off, are required for the assessment.

The process can have its focus on a specific asset class, such as stocks only, or combinations of some asset classes, such as bonds and stocks. If more than one asset class is involved, a portfolio revision will require changes, not only in the allocations of investment funds across different asset classes, but also in the allocations among the individual assets within the same asset class. The process requires the participation of professionals with different skill sets. For example, to generate input data for the analysis requires expertise in economic forecasting; to perform portfolio analysis with practically relevant models using such data requires expertise in portfolio modeling.

The portfolio management process as described here can be captured as follows:



In a quantitative approach, the portfolio management process as described in the above chart has three components: The **framework** — (1), (2), and (5) — pertains to portfolio modeling under relevant constraints; the **input** — (3), (4), and (7) — pertains to the generation of the required input data, which involves econometric modeling for capital market forecasts, as well as industry and company considerations; the **analysis** — (6) and (8) — pertains to portfolio optimization and revisions.

5.17.1 Economic Analysis and Forecasting

Let us start with the part of the portfolio management process that is directly related to the generation of input data for portfolio analysis. There are three levels of the economy to consider. That is, we consider the economy-wide level, the industry level, and then the company level. At the economy-wide level, we examine the Gross National Product (GNP), which represents the total value of the final output of goods and services as produced in the economy. The general idea is that, for example, an adequate increase in GNP suggests an optimistic economic climate.

Here are some GNP components:

1. Government spending
2. Business investment in plant and equipment
3. Residential construction
4. Inventory investment
5. Durable consumer goods
6. Nondurable consumer goods and services

Short-term forecasting techniques include the following:

1. Anticipatory surveys

   (including surveys of expert opinions and general surveys)

2. Economic indicators

   (for example, average workweek in manufacturing measured in hours, as a leading indicator of productive activities)

3. Diffusion indices

   (summaries of various economic indicators)

4. Econometric models

   (applications of mathematical and statistical techniques)

5.17.2 Industry Analysis

For industry analysis, we generally ask questions in the following categories:

1. Past sales and earnings performance
    - Stable or unstable performance?
    - Cost structure of the industry?
2. Relative permanence of the industry
    - Major technological changes?
    - Declining or growing industry?
3. Societal and governmental attitudes towards the industry
    - Is the government hostile or protective?
    - Any legal restrictions?
4. Labour conditions
    - Labour intensive or capital intensive?
    - Strong unions?
    - Strikes likely?
5. Competitive conditions
    - Any barriers to entry?
    - Any product differentiation edge?
    - Cost advantage?
6. Industry share prices relative to earnings
    - Do future earnings prospects justify the share prices?

5.17.3 Company Analysis

For company analysis, we pay special attention to company earnings. Earnings movements can be attributed to economy-wide and industrial influences, as well as factors pertaining to the individual companies under consideration. The stability or volatility of earnings of a company depends also on its cost structure (e.g., fixed costs versus variable costs) and its capital structure (e.g., equity versus debt component of its capital). Available earnings forecast figures, along with other public information, allow stocks to be priced in capital markets.

Some commonly used categories of stocks

In the case of stock investment, the idea of diversification is to invest in more than just a very small number of stocks. A practical issue is how best to diversify. Here are some familiar categories of stocks:

1. Blue chip stocks: high quality stocks, usually major companies with long, uninterrupted history of dividend payments.
2. Income stocks: companies with high dividend payments (i.e., with higher than average in the percentage of their earnings for dividend payments).
3. Cyclical stocks: companies whose earnings are greatly affected by the economy (i.e., they follow the business cycle).
4. Defensive stocks: companies whose earnings tend to be immune to changes in the economy; companies with low betas.
5. Growth stocks: companies that tend to retain earnings for reinvestments; companies with little or no dividends.
6. Speculative stocks: growth stocks at the far end of the risk spectrum.
7. Penny stocks: risky stocks with very low share values.

If the investor chooses to diversify only among speculative stocks, for example, the portfolio will likely be speculative in terms of its risk-return trade-off.

Security Screening

Since it is very costly and time consuming to evaluate a large number of stocks, an important first step in equity portfolio investment is security screening. Publications by Value Line, Standard & Poor's, and Moody's, for example, as well as information from full-service brokerage firms, are useful for security screening. Investors without access to these publications can still rely on financial press and public sources such as the internet for obtaining some basic information for security screening. Here are some examples of such information:

1. Price/earnings ratio. (The figures as reported in financial sections of newspapers are trailing P/E ratios — i.e., P/E ratios based on past earnings — but what really matter are future earnings.)

2. Dividend yield. (Some investors find dividend income highly attractive.)

3. Share price. (Investors prefer round-lot trading. Therefore, for a small investment capital, the purchase of stocks with high share prices may result in odd-lot holdings. Stocks with very low share prices are also unattractive to many investors for risk considerations.)

4. Exchange listing. (Many investors prefer big board stocks because of the stringent listing requirements from big boards.)

5. Familiarity. (Some companies tend to be followed very closely by the investment community and the media and thus are familiar to investors. Information about such companies is more easily accessible than companies that are unfamiliar to most investors.)

6. Current share price as compared to the 52-week high and low. (This information seems to provide a rough reference range of potential future prices, under the implicit assumption that history is likely going to repeat itself. However, this information alone does not reveal how the stock price was moving, and why it was moving up or down, during the past 52 weeks.)

7. Availability of options. (Stocks when combined with their corresponding derivative securities can enhance the versatility of the portfolio involved.)

5.17.4 Asset Allocation: Alternative Strategic Directions

There are three different strategies for asset allocation, with one being *passive* and the remaining two being *active* in nature. Here they are:

1. **Static Strategy** (passive)

   Constant mix of asset classes

2. **Reactive Strategy** (active)

   Reactions based on events that have already occurred

   Portfolio rebalancing: the process of periodically adjusting the portfolio to maintain its original conditions

3. **Anticipatory Strategy** (active)

   Decisions based on anticipation of market movements and conditions

Under a static strategy, once the way the investment funds are allocated across different asset classes is determined (based on the information available at the time of the decision), no further changes to the allocation are deemed necessary. This strategy, if followed over an extended period, would suggest that the cost of changing the asset mix is much too high to justify the potential benefits from the change.

Formula Plans

To implement a reactive strategy requires a set of specific rules. Such rules are often called *formula plans*. Here are some basic formula plans for portfolio rebalancing within the stock component of a portfolio:

1. **Constant Proportions**

   Adjustments are made to maintain the relative weighting of the portfolio components as their prices change.

   The adjustments often involve selling some shares of stocks in the portfolio with price increases and buying more shares of stocks in the portfolio with price decreases.

2. **Constant Beta**

   Adjustments are made to maintain the portfolio beta as stock prices or stock betas change.

   - Notice that beta is a measure of systematic risk, the risk associated with the co-movement of individual security or portfolio returns with the market returns.
   - In general, this is how a constant portfolio beta can be maintained:

   $$\begin{array}{rl} \text{Portfolio beta} \uparrow: & \text{sell some high-beta stocks} \\ & \text{buy some low-beta stocks} \\ \text{Portfolio beta} \downarrow: & \text{sell some low-beta stocks} \\ & \text{buy some high-beta stocks} \end{array}$$

   - Notice also that there are different ways to revise the portfolio to maintain a constant portfolio beta.

3. **Indexing**

   It involves investing in a portfolio with beta being as close to one as possible. The idea of indexing is to allow the portfolio returns to mimic the returns of the market index.

Here are some basic formula plans, also under the category of reactive strategy, for portfolio rebalancing across different asset classes:

1. **Constant Dollar Values**

   The dollar value of the stock component of the portfolio is to remain constant. Here is how:

   $$\begin{array}{rl} \text{stock value} \uparrow: & \text{sell some stocks} \\ & \text{buy some bonds} \\ \text{stock value} \downarrow: & \text{sell some bonds} \\ & \text{buy some stocks} \end{array}$$

2. **Constant Ratio**

The ratio of stock value to bond value is to remain constant. Here is how:

Ratio $\uparrow$:	sell some stocks
	buy some bonds
Ratio $\downarrow$:	sell some bonds
	buy some stocks

3. **Variable Ratio**

   The plan specifies a schedule of stock/bond mix for various levels of stock prices.

Rationales for Formula Plans

A reactive strategy, if followed properly, will allow the timing of investment activities to be separated from any emotional issues that the investor might have regarding such activities. The following are some rationales for formula plans:

1. Some investors are hesitant to buy when prices are low for fear that prices may fall lower.

2. Some investors are hesitant to sell when prices are high because they feel that prices may rise further.

3. Formula plans can minimize the emotions involved in investing.

4. Stock prices tend to fluctuate up and down in cycles. But the direction of the next fluctuation is difficult to predict. That is, market timing is difficult.

Notice that there is also an investment strategy called **Dollar-Cost Averaging (DCA)**. This strategy is based on the same rationales as formula plans; under DCA, a constant dollar amount is invested in a specified stock or portfolio at periodic dates, regardless of the price movements. Here is what DCA entails:

Price high:	buy less number of shares
Price low:	buy more number of shares

5.18 Closing Remarks

We have now completed the main text of this chapter. The assumption of frictionless short sales is required for analytical convenience, thus weakening the practical appeal of the portfolio selection models involved. However, the tangency portfolio approach as covered in Section 5.14 has turned out to be more versatile; it not only is able to relax to some extent the unrealistic assumption of frictionless short sale in a normative setting, but also provides a good starting point to develop a major asset pricing model in an equilibrium setting, to be covered in Chapter 6. The four appendices below are intended to strengthen significantly students' understanding of various analytical materials underlying the coverage in the main text.

5.19 Appendix A: Derivation of a Basic Portfolio Selection Model

In matrix notation, the optimization problem can be stated as

$$\begin{aligned} \text{minimize } \sigma_p^2 &= \boldsymbol{x}'\boldsymbol{V}\boldsymbol{x}, \\ \text{subject to } \mu_p &= \boldsymbol{x}'\boldsymbol{\mu} \ \text{ and } \boldsymbol{x}'\boldsymbol{\iota} = 1. \end{aligned}$$

Here, $\boldsymbol{\iota}$ is an n-element column vector where each element is 1. To verify the correspondence between the matrix version and the long-hand version, we now carry out the individual matrix multiplications. We can confirm directly that

$$\begin{aligned} \boldsymbol{x}'\boldsymbol{\mu} &= \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{bmatrix} = x_1\mu_1 + x_2\mu_2 + \cdots + x_n\mu_n \\ &= \sum\nolimits_{i=1}^{n} x_i\mu_i \end{aligned}$$

and

$$\begin{aligned}
\boldsymbol{x}'\boldsymbol{\iota} &= \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} = x_1 + x_2 + \cdots + x_n \\
&= \sum_{i=1}^{n} x_i.
\end{aligned}$$

For the expansion of

$$\boldsymbol{x}'\boldsymbol{V}\boldsymbol{x} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix},$$

we expand $\boldsymbol{x}'\boldsymbol{V}$ first, although expanding $\boldsymbol{V}\boldsymbol{x}$ first instead will still lead to the same end result. The detail of expanding $\boldsymbol{x}'\boldsymbol{V}$ first is as follows:

$$\begin{aligned}
\boldsymbol{x}'\boldsymbol{V}\boldsymbol{x} &= \begin{bmatrix} \left(\sum_{i=1}^{n} x_i\sigma_{i1}\right) & \left(\sum_{i=1}^{n} x_i\sigma_{i2}\right) & \cdots & \left(\sum_{i=1}^{n} x_i\sigma_{in}\right) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \\
&= \sum_{i=1}^{n} x_i\sigma_{i1}x_1 + \sum_{i=1}^{n} x_i\sigma_{i2}x_2 + \cdots + \sum_{i=1}^{n} x_i\sigma_{in}x_n \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} x_i x_j \sigma_{ij}.
\end{aligned}$$

As the two constraints in the optimization problem are linear, we can incorporate them into the objective function that σ_p^2 represents, by using a *Lagrangian* approach. For a predetermined μ_p, the Lagrangian is

$$L = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i x_j \sigma_{ij} - \phi\left(\sum_{i=1}^{n} x_i\mu_i - \mu_p\right) - \theta\left(\sum_{i=1}^{n} x_i - 1\right)$$

or, equivalently,

$$L = \boldsymbol{x}'\boldsymbol{V}\boldsymbol{x} - \phi(\boldsymbol{x}'\boldsymbol{\mu} - \mu_p) - \theta(\boldsymbol{x}'\boldsymbol{\iota} - 1), \tag{5.16}$$

where the portfolio weight vector $\boldsymbol{x}$ and the *Lagrange multipliers* ϕ and θ are the decision variables. Minimizing L eventually leads to

$$\boldsymbol{x} = \boldsymbol{V}^{-1}\boldsymbol{M}\,(\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}\boldsymbol{r}_p. \tag{5.17}$$

Here,

$$\boldsymbol{M} = \begin{bmatrix} \boldsymbol{\mu} & \boldsymbol{\iota} \end{bmatrix}$$

is an $n \times 2$ matrix and

$$\boldsymbol{r}_p = \begin{bmatrix} \mu_p & 1 \end{bmatrix}'$$

is a 2-element column vector. Using the expression of $\boldsymbol{x}$ in equation (5.17), we can also derive the variance of returns of the minimum variance portfolio (for a predetermined μ_p) as

$$\sigma_p^2 = \boldsymbol{x}'\boldsymbol{V}\,\boldsymbol{x} = \boldsymbol{r}_p{}'(\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}\boldsymbol{r}_p. \tag{5.18}$$

We now show how equation (5.17) is derived. For this task, let us write

$$\frac{\partial L}{\partial \boldsymbol{x}} = \begin{bmatrix} \partial L/\partial x_1 \\ \partial L/\partial x_2 \\ \vdots \\ \partial L/\partial x_n \end{bmatrix} = \begin{bmatrix} \left(2\sum_{j=1}^n x_j\sigma_{1j} - \phi\mu_1 - \theta\right) \\ \left(2\sum_{j=1}^n x_j\sigma_{2j} - \phi\mu_2 - \theta\right) \\ \vdots \\ \left(2\sum_{j=1}^n x_j\sigma_{nj} - \phi\mu_n - \theta\right) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

or, equivalently,

$$\frac{\partial L}{\partial \boldsymbol{x}} = 2\begin{bmatrix} \sum_{j=1}^n \sigma_{1j}x_j \\ \sum_{j=1}^n \sigma_{2j}x_j \\ \vdots \\ \sum_{j=1}^n \sigma_{nj}x_j \end{bmatrix} - \phi\begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{bmatrix} - \theta\begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Thus, we have

$$\frac{\partial L}{\partial \boldsymbol{x}} = 2\boldsymbol{V}\,\boldsymbol{x} - \phi\boldsymbol{\mu} - \theta\boldsymbol{\iota} = \boldsymbol{0}, \tag{5.19}$$

where $\boldsymbol{0}$ is the above n-element column vector of zeros. We can write, equivalently,

$$2\boldsymbol{V}\,\boldsymbol{x} = \phi\boldsymbol{\mu} + \theta\boldsymbol{\iota} = \begin{bmatrix} (\phi\mu_1 + \theta) \\ (\phi\mu_2 + \theta) \\ \vdots \\ (\phi\mu_n + \theta) \end{bmatrix} = \begin{bmatrix} \mu_1 & 1 \\ \mu_2 & 1 \\ \vdots & \vdots \\ \mu_n & 1 \end{bmatrix}\begin{bmatrix} \phi \\ \theta \end{bmatrix}.$$

With

$$\boldsymbol{M} = \begin{bmatrix} \boldsymbol{\mu} & \boldsymbol{\iota} \end{bmatrix} = \begin{bmatrix} \mu_1 & 1 \\ \mu_2 & 1 \\ \vdots & \vdots \\ \mu_n & 1 \end{bmatrix}$$

and

$$\boldsymbol{G} = \begin{bmatrix} \phi \\ \theta \end{bmatrix}$$

by definition, we can also write

$$2\boldsymbol{V}\boldsymbol{x} = \boldsymbol{M}\,\boldsymbol{G}.$$

This expression leads to

$$\boldsymbol{x} = \frac{1}{2}\boldsymbol{V}^{-1}\boldsymbol{M}\,\boldsymbol{G}.$$

To eliminate $\boldsymbol{G}$, we combine the two constraints as

$$\boldsymbol{M}'\boldsymbol{x} = \begin{bmatrix} \mu_1 & \mu_2 & \cdots & \mu_n \\ 1 & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} x_i\mu_i \\ \sum_{i=1}^{n} x_i \end{bmatrix} = \begin{bmatrix} \mu_p \\ 1 \end{bmatrix} = \boldsymbol{r}_p.$$

Given that

$$\boldsymbol{r}_p = \boldsymbol{M}'\boldsymbol{x} = \boldsymbol{M}'\left(\frac{1}{2}\boldsymbol{V}^{-1}\boldsymbol{M}\,\boldsymbol{G}\right) = \boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M}\left(\frac{1}{2}\boldsymbol{G}\right),$$

we can write

$$\frac{1}{2}\boldsymbol{G} = (\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}\boldsymbol{r}_p.$$

Further, given that

$$\boldsymbol{x} = \frac{1}{2}\boldsymbol{V}^{-1}\boldsymbol{M}\,\boldsymbol{G} = \boldsymbol{V}^{-1}\boldsymbol{M}\left(\frac{1}{2}\boldsymbol{G}\right),$$

we finally reach

$$\boldsymbol{x} = \boldsymbol{V}^{-1}\boldsymbol{M}\,(\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}\boldsymbol{r}_p,$$

which is equation (5.17).

To find the variance of returns of the least risky portfolio corresponding to μ_p being the portfolio's expected return, we can express

$$\sigma_p^2 = \boldsymbol{x}'\boldsymbol{V}\boldsymbol{x}$$

in terms of the input parameters. Let us start with

$$\boldsymbol{x}' = \left[\boldsymbol{V}^{-1}\boldsymbol{M}\,(\boldsymbol{M}\,'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}\boldsymbol{r}_p\right]'.$$

As $\boldsymbol{V}$ is symmetric, so is $\boldsymbol{V}^{-1}$. It follows that

$$\left(\boldsymbol{V}^{-1}\right)' = \boldsymbol{V}^{-1}$$

and that

$$\left[(\boldsymbol{M}\,'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}\right]' = \left[(\boldsymbol{M}\,'\boldsymbol{V}^{-1}\boldsymbol{M})'\right]^{-1} = (\boldsymbol{M}\,'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}.$$

Thus, we can write

$$\boldsymbol{x}' = \left[\boldsymbol{V}^{-1}\boldsymbol{M}\,(\boldsymbol{M}\,'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}\boldsymbol{r}_p\right]' = \boldsymbol{r}_p'(\boldsymbol{M}\,'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}\boldsymbol{M}\,'\boldsymbol{V}^{-1}$$

and then

$$\begin{aligned}
\sigma_p^2 &= \boldsymbol{x}'\boldsymbol{V}\boldsymbol{x} \\
&= \left[\boldsymbol{r}_p'(\boldsymbol{M}\,'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}\boldsymbol{M}\,'\boldsymbol{V}^{-1}\right]\boldsymbol{V}\left[\boldsymbol{V}^{-1}\boldsymbol{M}\,(\boldsymbol{M}\,'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}\boldsymbol{r}_p\right] \\
&= \boldsymbol{r}_p'(\boldsymbol{M}\,'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}\left(\boldsymbol{M}\,'\boldsymbol{V}^{-1}\boldsymbol{M}\right)(\boldsymbol{M}\,'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}\boldsymbol{r}_p \\
&= \boldsymbol{r}_p'(\boldsymbol{M}\,'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}\boldsymbol{r}_p.
\end{aligned}$$

As $\boldsymbol{r}_p$ is a two-element column vector, $\boldsymbol{r}_p'$ is a two-element row vector. Further, as the 2×2 matrix that $\boldsymbol{M}\,'\boldsymbol{V}^{-1}\boldsymbol{M}$ represents is symmetric, so is the 2×2 matrix that $(\boldsymbol{M}\,'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}$ represents. The matrix product $\boldsymbol{r}_p'(\boldsymbol{M}\,'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}\boldsymbol{r}_p$ is a scalar.

5.19.1 An analytical Exercise

Consider two minimum variance portfolios, p and q, based on the same set of n securities with random returns being $\widetilde{R}_1, \widetilde{R}_2, \ldots, \widetilde{R}_n$. The corresponding portfolio weights are $x_1, x_2, \ldots, x_n$ for portfolio p and $y_1, y_2, \ldots, y_n$ for portfolio q, satisfying the conditions of

$$x_1 + x_2 + \cdots + x_n = 1$$

and

$$y_1 + y_2 + \cdots + y_n = 1.$$

Let $\widetilde{R}_p$ and $\widetilde{R}_q$ be the random returns of the two portfolios. Let μ_p and μ_q be the corresponding expected returns.

Verify that the covariance of returns between portfolios p and q, defined as

$$\sigma_{pq} = Cov(\widetilde{R}_p, \widetilde{R}_q) = E[(\widetilde{R}_p - \mu_p)(\widetilde{R}_q - \mu_q)],$$

is

$$\sigma_{pq} = \boldsymbol{x}'\boldsymbol{V}\boldsymbol{y} = \boldsymbol{y}'\boldsymbol{V}\boldsymbol{x},$$

where $\boldsymbol{y}$ is an n-element column vector with elements being $y_1, y_2, \ldots, y_n$. Verify also that

$$\sigma_{pq} = \boldsymbol{r}_p'(\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}\boldsymbol{r}_q = \boldsymbol{r}_q'(\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}\boldsymbol{r}_p,$$

where

$$\boldsymbol{r}_q = \begin{bmatrix} \mu_q \\ 1 \end{bmatrix}.$$

Solution: This exercise has two parts. The first part is to find the covariance of returns of portfolios p and q. Using the same approach as shown in Section 5.5, we have

$$\begin{aligned}
\sigma_{pq} &= E\left[\left(\widetilde{R}_p - \mu_p\right)\left(\widetilde{R}_q - \mu_q\right)\right] \\
&= E\left[\sum\nolimits_{i=1}^{n} x_i \left(\widetilde{R}_i - \mu_i\right) \sum\nolimits_{j=1}^{n} y_j \left(\widetilde{R}_j - \mu_j\right)\right] \\
&= E\left[\sum\nolimits_{i=1}^{n} \sum\nolimits_{j=1}^{n} x_i y_j \left(\widetilde{R}_i - \mu_i\right)\left(\widetilde{R}_j - \mu_j\right)\right] \\
&= \sum\nolimits_{i=1}^{n} \sum\nolimits_{j=1}^{n} x_i y_j E\left[\left(\widetilde{R}_i - \mu_i\right)\left(\widetilde{R}_j - \mu_j\right)\right] \\
&= \sum\nolimits_{i=1}^{n} \sum\nolimits_{j=1}^{n} x_i y_j \sigma_{ij} = \boldsymbol{x}'\boldsymbol{V}\boldsymbol{y} = \boldsymbol{y}'\boldsymbol{V}\boldsymbol{x}. \blacksquare
\end{aligned}$$

For the second part, as portfolios p and q are minimum variance portfolios, we have

$$\boldsymbol{x} = \boldsymbol{V}^{-1}\boldsymbol{M}\left(\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M}\right)^{-1}\boldsymbol{r}_p$$

and

$$\boldsymbol{y} = \boldsymbol{V}^{-1}\boldsymbol{M}\left(\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M}\right)^{-1}\boldsymbol{r}_q.$$

As

$$\begin{aligned} \boldsymbol{x}' &= \left[\boldsymbol{V}^{-1}\boldsymbol{M}\,(\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}\boldsymbol{r}_p\right]' \\ &= \boldsymbol{r}_p'(\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}\boldsymbol{M}'\boldsymbol{V}^{-1}, \end{aligned}$$

we can write

$$\begin{aligned} \sigma_{pq} &= \boldsymbol{x}'\boldsymbol{V}\boldsymbol{y} \\ &= \left[\boldsymbol{r}_p'(\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}\boldsymbol{M}'\boldsymbol{V}^{-1}\right]\boldsymbol{V}\left[\boldsymbol{V}^{-1}\boldsymbol{M}\,(\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}\boldsymbol{r}_q\right] \\ &= \boldsymbol{r}_p'(\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}\left(\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M}\right)(\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}\boldsymbol{r}_q \\ &= \boldsymbol{r}_p'(\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}\boldsymbol{r}_q, \end{aligned}$$

which is equivalent to

$$\sigma_{qp} = \boldsymbol{r}_q'(\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}\boldsymbol{r}_p$$

by interchanging p and q.

5.19.2 Another Analytical Exercise

Under the assumption of frictionless short sales, let μ_o, σ_o^2, and $\boldsymbol{x}_o$ be the expected return, the variance of returns, and an n-element column vector of the corresponding portfolio weights, respectively, pertaining to the global minimum variance portfolio. Express them in terms of the $n \times n$ covariance matrix of returns $\boldsymbol{V}$ and, if necessary, also the n-element column vector of expected returns $\boldsymbol{\mu}$. Again, it is important that we do the exercise on our own before reading the solution that is provided below.

Solution: For a portfolio optimization that minimizes $\boldsymbol{x}'\boldsymbol{V}\boldsymbol{x}$ subject to $\boldsymbol{\iota}'\boldsymbol{x} = 1$, the Lagrangian is

$$L = \boldsymbol{x}'\boldsymbol{V}\boldsymbol{x} - \theta(\boldsymbol{x}'\boldsymbol{\iota} - 1),$$

where θ is a Lagrange multiplier. From

$$\frac{\partial L}{\partial \boldsymbol{x}} = 2\boldsymbol{V}\boldsymbol{x} - \theta\boldsymbol{\iota} = \boldsymbol{0},$$

where $\boldsymbol{0}$ is an n-element column vector of zeros, we have

$$\boldsymbol{x} = \frac{\theta}{2}\boldsymbol{V}^{-1}\boldsymbol{\iota} = \boldsymbol{V}^{-1}\boldsymbol{\iota}\frac{\theta}{2}.$$

As $\theta/2$ is a scalar, it can be placed in front of, or at the back of, $\boldsymbol{V}^{-1}\boldsymbol{\iota}$. To eliminate $\theta/2$, we use

$$\boldsymbol{\iota}'\boldsymbol{x} = 1.$$

Specifically, as

$$1 = \boldsymbol{\iota}'\boldsymbol{x} = (\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\iota})\frac{\theta}{2},$$

we have

$$\frac{\theta}{2} = (\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\iota})^{-1}$$

and then

$$\boldsymbol{x} = \boldsymbol{V}^{-1}\boldsymbol{\iota}(\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\iota})^{-1}.$$

Let us denote this $\boldsymbol{x}$ as $\boldsymbol{x}_o$. The expected returns of the global minimum variance portfolio is

$$\begin{aligned} \mu_o &= \boldsymbol{x}_o'\boldsymbol{\mu} = [\boldsymbol{V}^{-1}\boldsymbol{\iota}(\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\iota})^{-1}]'\boldsymbol{\mu} \\ &= [(\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\iota})^{-1}\boldsymbol{\iota}'\boldsymbol{V}^{-1}]\boldsymbol{\mu} = (\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\iota})^{-1}(\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\mu}) \end{aligned}$$

or, equivalently,

$$\mu_o = \boldsymbol{\mu}'\boldsymbol{x}_o = \boldsymbol{\mu}'[\boldsymbol{V}^{-1}\boldsymbol{\iota}(\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\iota})^{-1}] = (\boldsymbol{\mu}'\boldsymbol{V}^{-1}\boldsymbol{\iota})(\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\iota})^{-1},$$

given that $\boldsymbol{V}^{-1}$ is symmetric. The corresponding variance of returns is

$$\sigma_o^2 = \boldsymbol{x}_o'\boldsymbol{V}\boldsymbol{x}_o = \boldsymbol{x}_o'\boldsymbol{V}\boldsymbol{V}^{-1}\boldsymbol{\iota}(\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\iota})^{-1} = \boldsymbol{x}_o'\boldsymbol{\iota}(\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\iota})^{-1} = (\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\iota})^{-1},$$

as $\boldsymbol{V}\boldsymbol{V}^{-1}$ is an identity matrix.

- **An extension of the above exercise:** Consider a portfolio k based on the same set of n risky securities. This portfolio need not be efficient. Let $\boldsymbol{w}$ be an n-element column vector, which represents the portfolio weights of portfolio k. Under the assumption of frictionless short sales, as

  $$\boldsymbol{w}'\boldsymbol{\iota} = 1,$$

  the covariance of returns between portfolios o and k is

  $$\sigma_{ok} = \boldsymbol{w}'\boldsymbol{V}\boldsymbol{x}_o = \boldsymbol{w}'\boldsymbol{V}\boldsymbol{V}^{-1}\boldsymbol{\iota}(\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\iota})^{-1} = \boldsymbol{w}'\boldsymbol{\iota}(\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\iota})^{-1} = (\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\iota})^{-1},$$

which is the same as σ_o^2.

5.20 Appendix B: A Derivation for the Tangency Portfolio Approach

For the tangency portfolio approach, the optimization problem to be solved is via maximizing

$$\theta = \frac{\mu_p - r}{\sigma_p} = \frac{\sum_{i=1}^n x_i \mu_i - r}{\sqrt{\sum_{i=1}^n \sum_{j=1}^n x_i x_j \sigma_{ij}}}$$

subject to

$$\sum\nolimits_{i=1}^n x_i = 1.$$

Here, the n-element column vector of portfolio weights

$$\boldsymbol{x} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}',$$

where the prime stands for matrix transposition, is to be determined in terms of the μ-intercept of the tangent line r on the (σ, μ)-plane. The n-element column vector of expected returns is

$$\boldsymbol{\mu} = \begin{bmatrix} \mu_1 & \mu_2 & \cdots & \mu_n \end{bmatrix}',$$

and the $n \times n$ covariance matrix of returns

$$\boldsymbol{V} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_{nn} \end{bmatrix},$$

is positive definite.

As

$$\sum\nolimits_{i=1}^n x_i = 1,$$

we can write, equivalently,

$$\theta = \frac{\sum_{i=1}^n x_i \mu_i - r \sum_{i=1}^n x_i}{\sqrt{\sum_{i=1}^n \sum_{j=1}^n x_i x_j \sigma_{ij}}} = \frac{\sum_{i=1}^n x_i \mu_i - \sum_{i=1}^n x_i r}{\sqrt{\sum_{i=1}^n \sum_{j=1}^n x_i x_j \sigma_{ij}}} = \frac{\sum_{i=1}^n x_i (\mu_i - r)}{\sqrt{\sum_{i=1}^n \sum_{j=1}^n x_i x_j \sigma_{ij}}}.$$

For $s > 0$ being an arbitrary constant, we can also write

$$\theta = \frac{s \sum_{i=1}^{n} x_i(\mu_i - r)}{\sqrt{s^2 \sum_{i=1}^{n} \sum_{j=1}^{n} x_i x_j \sigma_{ij}}} = \frac{\sum_{i=1}^{n} (sx_i)(\mu_i - r)}{\sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} (sx_i)(sx_j)\sigma_{ij}}}.$$

Letting

$$y_i = sx_i, \text{ for } i = 1, 2, \ldots, n,$$

leads to

$$\theta = \frac{\sum_{i=1}^{n} y_i(\mu_i - r)}{\sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} y_i y_j \sigma_{ij}}}.$$

As

$$\sum\nolimits_{i=1}^{n} y_i = s \sum\nolimits_{i=1}^{n} x_i = s,$$

which is arbitrary, the original optimization problem to determine $x_1, x_2, \ldots, x_n$ now becomes an unconstrained case, where the variables to be solved are $y_1, y_2, \ldots, y_n$ instead.

Letting

$$H = \sum\nolimits_{i=1}^{n} y_i(\mu_i - r)$$

and

$$K = \sqrt{\sum\nolimits_{i=1}^{n} \sum\nolimits_{j=1}^{n} y_i y_j \sigma_{ij}}$$

for notational convenience, we have

$$\frac{\partial}{\partial y_i}\left(\frac{H}{K}\right) = \frac{K\,(\partial H/\partial y_i) - H\,(\partial K/\partial y_i)}{K^2} = 0, \quad \text{for} \quad i = 1, 2, \ldots, n.$$

As $K^2 > 0$, this step leads to

$$K\,\frac{\partial H}{\partial y_i} = H\,\frac{\partial K}{\partial y_i}, \quad \text{for} \quad i = 1, 2, \ldots, n.$$

Noting that $\boldsymbol{V}$ is symmetric, from

$$K\,\frac{\partial H}{\partial y_i} = K\,\frac{\partial}{\partial y_i} \sum\nolimits_{h=1}^{n} (\mu_h - r) y_h = K\,(\mu_i - r)$$

and

$$\begin{aligned}
H\,\frac{\partial K}{\partial y_i} &= H\,\frac{\partial}{\partial y_i}\left(\sum\nolimits_{h=1}^{n}\sum\nolimits_{k=1}^{n} y_h y_k \sigma_{hk}\right)^{1/2} \\
&= \frac{H}{2}\left(\sum\nolimits_{h=1}^{n}\sum\nolimits_{k=1}^{n} y_h y_k \sigma_{hk}\right)^{-1/2} \frac{\partial}{\partial y_i}\left(\sum\nolimits_{h=1}^{n}\sum\nolimits_{k=1}^{n} y_h y_k \sigma_{hk}\right) \\
&= H\left(\sum\nolimits_{h=1}^{n}\sum\nolimits_{k=1}^{n} y_h y_k \sigma_{hk}\right)^{-1/2} \sum\nolimits_{h=1}^{n} y_h \sigma_{hi} \\
&= \frac{H}{K}\sum\nolimits_{h=1}^{n} \sigma_{ih} y_h, \quad \text{for} \quad i = 1, 2, \ldots, n,
\end{aligned}$$

we have

$$\mu_i - r = \frac{H}{K^2}\sum\nolimits_{h=1}^{n} \sigma_{ih} y_h, \quad \text{for} \quad i = 1, 2, \ldots, n,$$

and then

$$\begin{aligned}
\mu_i - r &= \frac{\sum_{h=1}^{n}(\mu_h - r) y_h}{\sum_{h=1}^{n}\sum_{k=1}^{n} y_h y_k \sigma_{hk}}\sum\nolimits_{h=1}^{n} \sigma_{ih} y_h \\
&= \frac{\sum_{h=1}^{n}(\mu_h - r)(s x_h)}{\sum_{h=1}^{n}\sum_{k=1}^{n} (s x_h)(s x_k) \sigma_{hk}}\sum\nolimits_{h=1}^{n} \sigma_{ih} (s x_h) \\
&= \frac{s\sum_{h=1}^{n}(\mu_h - r) x_h}{s^2\sum_{h=1}^{n}\sum_{k=1}^{n} x_h x_k \sigma_{hk}}\left(s\sum\nolimits_{h=1}^{n} \sigma_{ih} x_h\right) \\
&= \frac{\sum_{h=1}^{n}(\mu_h - r) x_h}{\sum_{h=1}^{n}\sum_{k=1}^{n} x_h x_k \sigma_{hk}}\left(\sum\nolimits_{h=1}^{n} \sigma_{ih} x_h\right), \quad \text{for} \quad i = 1, 2, \ldots, n.
\end{aligned}$$

Letting

$$z_i = \left[\frac{\sum_{h=1}^{n}(\mu_h - r) x_h}{\sum_{h=1}^{n}\sum_{k=1}^{n} x_h x_k \sigma_{hk}}\right] x_i = \left(\frac{\mu_p - r}{\sigma_p^2}\right) x_i, \text{ for } i = 1, 2, \ldots, n, \tag{5.20}$$

we can write

$$\mu_i - r = \sum\nolimits_{h=1}^{n} \sigma_{ih} z_h, \text{ for } \; i = 1, 2, \ldots, n,$$

which represents n simultaneous linear equations, where the n unknown variables are $z_1, z_2, \ldots, z_n$. In matrix notation, these equation can be stated succinctly as

$$\boldsymbol{\mu} - r\boldsymbol{\iota} = \boldsymbol{V}\boldsymbol{z},$$

where

$$\boldsymbol{\iota} = \begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix}'$$

and

$$\boldsymbol{z} = \begin{bmatrix} z_1 & z_2 & \cdots & z_n \end{bmatrix}'$$

are n-element column vectors.

If $\boldsymbol{V}$ is positive definite, $\boldsymbol{V}^{-1}$ exists and we can solve $\boldsymbol{z}$ from

$$\boldsymbol{z} = \boldsymbol{V}^{-1}(\boldsymbol{\mu} - r\boldsymbol{\iota}).$$

As

$$\sum\nolimits_{i=1}^{n} z_i = \boldsymbol{z}'\boldsymbol{\iota}$$

and

$$\begin{aligned}
\sum\nolimits_{i=1}^{n} z_i &= \sum\nolimits_{i=1}^{n} \left[\frac{\sum_{h=1}^{n}(\mu_h - r)x_h}{\sum_{h=1}^{n}\sum_{k=1}^{n} x_h x_k \sigma_{hk}}\right] x_i \\
&= \left[\frac{\sum_{h=1}^{n}(\mu_h - r)x_h}{\sum_{h=1}^{n}\sum_{k=1}^{n} x_h x_k \sigma_{hk}}\right] \sum\nolimits_{i=1}^{n} x_i \\
&= \frac{\sum_{h=1}^{n}(\mu_h - r)x_h}{\sum_{h=1}^{n}\sum_{k=1}^{n} x_h x_k \sigma_{hk}},
\end{aligned}$$

we can determine $x_1, x_2, \ldots, x_n$ from

$$x_i = \frac{z_i}{\sum_{h=1}^{n} z_h}, \quad \text{for} \quad i = 1, 2, \ldots, n,$$

thus completing the task. In matrix notation, we can write

$$\boldsymbol{x} = \boldsymbol{z}(\boldsymbol{\iota}'\boldsymbol{z})^{-1} = \boldsymbol{V}^{-1}(\boldsymbol{\mu} - r\boldsymbol{\iota})[\boldsymbol{\iota}'\boldsymbol{V}^{-1}(\boldsymbol{\mu} - r\boldsymbol{\iota})]^{-1}.$$

With $\boldsymbol{x}$ known, the determination of μ_p and σ_p is straightforward.

5.20.1 An Analytical Exercise

Verify that, for each tangency portfolio p corresponding to r being the μ-intercept of the tangent line,

$$\sigma_p = \sqrt{\frac{\mu_p - r}{\sum_{h=1}^{n} z_h}},$$

where σ_p, μ_p, and z_j, for $j = 1, 2, \ldots, n$, are as defined previously. For computational purposes, this expression is more convenient than the conventional approach via $\sigma_p^2 = \boldsymbol{x}'\boldsymbol{V}\boldsymbol{x}$. This convenient expression was pointed out in class by a former student.

Solution: As

$$\sum_{i=1}^{n} z_i = \frac{\sum_{h=1}^{n}(\mu_h - r)x_h}{\sum_{h=1}^{n}\sum_{k=1}^{n} x_h x_k \sigma_{hk}} = \frac{\mu_p - r}{\sigma_p^2},$$

we can write

$$\sigma_p^2 = \frac{\mu_p - r}{\sum_{h=1}^{n} z_h}.$$

Noting that, by definition, σ_p is never negative, we have

$$\sigma_p = \sqrt{\frac{\mu_p - r}{\sum_{h=1}^{n} z_h}},$$

5.20.2 Another Analytical Exercise

In this exercise, we focus on a special case where $r = R_f$, which is the risk-free interest rate. In Lintner's formulation of the tangency portfolio approach, optimization is via maximizing

$$\theta = \frac{\sum_{i=1}^{n} x_i \mu_i - R_f}{\sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} x_i x_j \sigma_{ij}}}$$

subject to

$$\sum_{i=1}^{n} |x_i| = 1.$$

Determine $x_1, x_2, \ldots, x_n$ in terms of R_f and elements of $\boldsymbol{\mu}$ and $\boldsymbol{V}$ as defined earlier. Determine also μ_p and σ_p.

- The idea is that, regardless of whether a security is held long or short, an amount equal to the value of the security must be provided. To illustrate, suppose that an investor has \$1 to invest in two securities, security 1 in a long position and security 2 in a short position. If \$0.60 is used for the purchase of security 1, the remaining amount, \$0.40, will provide the cash deposit for short selling security 2. In this case, we have $x_1 = 0.6$, $x_2 = -0.4$, and thus $|x_1| + |x_2| = 1$.

Solution: We first write

$$\theta = \frac{\sum_{i=1}^{n} x_i \mu_i - R_f}{\sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} x_i x_j \sigma_{ij}}} = \frac{\sum_{i=1}^{n} x_i \mu_i - R_f \sum_{i=1}^{n} |x_i|}{\sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} x_i x_j \sigma_{ij}}}.$$

Suppose that k of the n securities are held long and the remaining $n-k$ securities are held short in a portfolio, where k is arbitrary. For notational convenience, let us relabel the n securities in such a way that securities $1, 2, \ldots, k$ are held long and securities $k+1, k+2, \ldots, n$ are held short. Then, we can write

$$R_f = R_f \sum_{i=1}^{n} |x_i| = R_f \sum_{i=1}^{k} x_i - R_f \sum_{i=k+1}^{n} x_i,$$

and, as the short seller earns interests from both the deposit and the short-sale proceeds,

$$\begin{aligned}\mu_p - R_f &= \sum_{i=1}^{n} x_i \mu_i - 2R_f \sum_{i=k+1}^{n} x_i - R_f \sum_{i=1}^{k} x_i + R_f \sum_{i=k+1}^{n} x_i \\ &= \sum_{i=1}^{n} x_i (\mu_i - R_f).\end{aligned}$$

Regardless of how the individual securities are labeled, the expression for

$$\theta = \frac{\mu_p - R_f}{\sigma_p} = \frac{\sum_{i=1}^{n} x_i (\mu_i - R_f)}{\sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} x_i x_j \sigma_{ij}}}$$

is unaffected. Letting

$$y_i = s x_i, \text{ for } i = 1, 2, \ldots, n,$$

where $s > 0$ is an arbitrary constant, leads to

$$\sum_{i=1}^{n} |y_i| = s \sum_{i=1}^{n} |x_i| = s,$$

which is also arbitrary. As the original optimization of θ now becomes an unconstrained case, tracing the same analytical steps will give us equation (5.20) where r is substituted by R_f and then

$$\sum_{i=1}^{n} |z_i| = \frac{\sum_{h=1}^{n} (\mu_h - R_f) x_h}{\sum_{h=1}^{n} \sum_{k=1}^{n} x_h x_k \sigma_{hk}},$$

thus leading to

$$x_i = \frac{z_i}{\sum_{h=1}^{n} |z_h|}, \quad \text{for} \quad i = 1, 2, \ldots, n,$$

$$\mu_p = R_f + \sum_{i=1}^{n} x_i (\mu_i - R_f)$$

and

$$\sigma_p = \sqrt{\frac{\mu_p - R_f}{\sum_{h=1}^{n} |z_h|}}.$$

5.21 Appendix C: Positive Definite and Semi-Definite Matrices

In this appendix, we first confirm that the $n \times n$ sample covariance matrix of returns $\widehat{\boldsymbol{V}}$ is always *positive semi-definite*. By definition, for any n-element column vector $\boldsymbol{w}$, if $\boldsymbol{w}'\widehat{\boldsymbol{V}}\boldsymbol{w}$ is never negative, $\widehat{\boldsymbol{V}}$ is positive semi-definite. For the task, let

$$u_{it} = \frac{R_{it} - \overline{R}_i}{\sqrt{T-1}}$$

be the (i,t)-element of an $n \times T$ matrix $\boldsymbol{U}$, for $i = 1, 2, \ldots, n$ and $t = 1, 2, \ldots, T$. With the term $\sqrt{T-1}$ in the expression of u_{it} being common for all i and all t, we can write

$$\boldsymbol{U} = \frac{1}{\sqrt{T-1}} \begin{bmatrix} (R_{11} - \overline{R}_1) & (R_{12} - \overline{R}_1) & \cdots & (R_{1T} - \overline{R}_1) \\ (R_{21} - \overline{R}_2) & (R_{22} - \overline{R}_2) & \cdots & (R_{2T} - \overline{R}_2) \\ \vdots & \vdots & \ddots & \vdots \\ (R_{n1} - \overline{R}_n) & (R_{n2} - \overline{R}_n) & \cdots & (R_{nT} - \overline{R}_n) \end{bmatrix}.$$

With its transpose being

$$\boldsymbol{U}' = \frac{1}{\sqrt{T-1}} \begin{bmatrix} (R_{11} - \overline{R}_1) & (R_{21} - \overline{R}_2) & \cdots & (R_{n1} - \overline{R}_n) \\ (R_{12} - \overline{R}_1) & (R_{22} - \overline{R}_2) & \cdots & (R_{n2} - \overline{R}_n) \\ \vdots & \vdots & \ddots & \vdots \\ (R_{1T} - \overline{R}_1) & (R_{2T} - \overline{R}_2) & \cdots & (R_{nT} - \overline{R}_n) \end{bmatrix},$$

we have

$$\begin{aligned} \boldsymbol{U}\boldsymbol{U}' &= \frac{1}{T-1} \begin{bmatrix} \sum_{t=1}^{T} \left(R_{1t} - \overline{R}_1\right)\left(R_{1t} - \overline{R}_1\right) & \cdots & \sum_{t=1}^{T} \left(R_{1t} - \overline{R}_1\right)\left(R_{nt} - \overline{R}_n\right) \\ \vdots & \ddots & \vdots \\ \sum_{t=1}^{T} \left(R_{nt} - \overline{R}_n\right)\left(R_{1t} - \overline{R}_1\right) & \cdots & \sum_{t=1}^{T} \left(R_{nt} - \overline{R}_n\right)\left(R_{nt} - \overline{R}_n\right) \end{bmatrix} \\ &= \widehat{\boldsymbol{V}}. \end{aligned}$$

The matrix product $\boldsymbol{w}'\widehat{\boldsymbol{V}}\boldsymbol{w}$, for any n-element column vector $\boldsymbol{w}$ is therefore the same as

$$\boldsymbol{w}'\boldsymbol{U}\boldsymbol{U}'\boldsymbol{w} = (\boldsymbol{U}'\boldsymbol{w})'(\boldsymbol{U}'\boldsymbol{w}).$$

We now let

$$\boldsymbol{v} = \boldsymbol{U}'\boldsymbol{w}$$

be a T-element column vector and label its elements as $v_1, v_2, \cdots, v_T$. It follows that

$$\boldsymbol{w}'\widehat{\boldsymbol{V}}\,\boldsymbol{w} = \boldsymbol{v}'\boldsymbol{v} = \sum\nolimits_{t=1}^{T} v_t^2,$$

which is never negative. With $\boldsymbol{w}$ and, consequently, $\boldsymbol{v}$ being arbitrary, the positive semi-definiteness of the sample covariance matrix $\widehat{\boldsymbol{V}}$ is confirmed.

For the sample covariance matrix $\widehat{\boldsymbol{V}}$ to be usable as part of the input parameters for implementing a basic portfolio selection model, it must be invertible. An invertible positive semi-definite matrix is positive definite. Thus, causes for $\widehat{\boldsymbol{V}}$ to have a zero determinant must be ruled out. Specifically, none of the n securities considered can be risk-free; otherwise, the corresponding $\widehat{\boldsymbol{V}}$ will have a zero determinant. Further, no two securities considered can be perfectly correlated in their returns, and none of the n securities considered can be replicated exactly by a linear combination of any of the remaining securities.

For $\widehat{\boldsymbol{V}}$ to be invertible, we must rule out the use of insufficient return data for its estimation. As $\widehat{\boldsymbol{V}}$ is an $n \times n$ matrix, it has n^2 elements. If $T < n$, then there are only nT data points of R_{it}, for $i = 1, 2, \cdots, n$ and $t = 1, 2, \ldots, T$, for estimating the n^2 elements. The situation where $T < n$ is also where $nT < n^2$, and thus the return data are insufficient for estimating $\widehat{\boldsymbol{V}}$.

If $T = n$ instead, the return data are still insufficient. This is because, if $T = n$, then any of the n rows (columns) of $\widehat{\boldsymbol{V}}$ can always be replicated by a linear combination of the remaining $n - 1$ rows (columns), thus causing .the failure to invert $\widehat{\boldsymbol{V}}$.

5.22 Appendix D: A Sketch of Shrinkage Estimation

For an n-security case, let $\boldsymbol{R}_1, \boldsymbol{R}_2, \ldots, \boldsymbol{R}_n$ be the random returns of the individual securities. Let also $R_{i1}, R_{i2}, \ldots, R_{iT}$ be the returns observed at times $1, 2, \ldots, T$, respectively, for $i = 1, 2, \ldots, n$. For each security i, the sample mean return and the sample variance of returns are

$$\overline{\boldsymbol{R}}_i = \frac{1}{T}\sum\nolimits_{t=1}^{T} R_{it}$$

and

$$s_i^2 = \frac{1}{T-1}\sum\nolimits_{t=1}^{T}(R_{it} - \overline{\boldsymbol{R}}_i)^2,$$

respectively, where s_i is the sample standard deviation of returns. The sample covariance of returns between securities i and j, for $i, j = 1, 2, \ldots, n$, is

$$s_{ij} = \frac{1}{T-1}\sum\nolimits_{t=1}^{T}(R_{it} - \overline{\boldsymbol{R}}_i)(R_{jt} - \overline{\boldsymbol{R}}_j) = s_{ji}.$$

It is implicit that

$$s_i^2 = s_{ii}, \text{ for } i = 1, 2, \ldots, n.$$

The sample correlation of returns between securities i and j, for $i, j = 1, 2, \ldots, n$, is

$$r_{ij} = \frac{s_{ij}}{s_i s_j} = r_{ji}.$$

It is also implicit that

$$r_{ii} = 1, \text{ for } i = 1, 2, \ldots, n.$$

The sample covariance matrix of returns is

$$\widehat{\boldsymbol{V}} = \begin{bmatrix} s_{11} & s_{12} & s_{13} & \cdots & s_{1n} \\ s_{21} & s_{22} & s_{23} & \cdots & s_{2n} \\ s_{31} & s_{32} & s_{33} & \cdots & s_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ s_{n1} & s_{n2} & s_{n3} & \cdots & s_{nn} \end{bmatrix} = \begin{bmatrix} s_1^2 & r_{12}s_1s_2 & r_{13}s_1s_3 & \cdots & r_{1n}s_1s_n \\ r_{21}s_2s_1 & s_2^2 & r_{23}s_2s_3 & \cdots & r_{2n}s_2s_n \\ r_{31}s_3s_1 & r_{32}s_3s_2 & s_3^2 & \cdots & r_{3n}s_3s_n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r_{n1}s_ns_1 & r_{n2}s_ns_2 & r_{n3}s_ns_3 & \cdots & s_n^2 \end{bmatrix}.$$

There are $n(n-1)/2$ off-diagonal elements in the upper triangle of the sample correlation matrix of returns. Thus, the average of the sample correlations of returns can be written as

$$\begin{aligned} \overline{r} &= \frac{2}{n(n-1)}[(r_{12} + r_{13} + \cdots + r_{1n}) + (r_{23} + r_{24} + \cdots + r_{2n}) \\ &\quad +(r_{34} + r_{35} + \cdots + r_{3n}) + \cdots + (r_{n-2,n-1} + r_{n-2,n}) + (r_{n-1,n})] \end{aligned}$$

or, equivalently,

$$\overline{r} = \frac{2}{n(n-1)}\sum\nolimits_{i=1}^{n-1}\sum\nolimits_{j=i+1}^{n} r_{ij}.$$

In practice, we have $\overline{r} \geq 0$. The structured covariance matrix of returns, where each r_{ij} in the off-diagonal elements of $\widehat{\boldsymbol{V}}$ is substituted by the average $\overline{r}$, is

$$\widehat{\boldsymbol{A}} = \begin{bmatrix} s_1^2 & \overline{r}s_1s_2 & \overline{r}s_1s_3 & \cdots & \overline{r}s_1s_n \\ \overline{r}s_2s_1 & s_2^2 & \overline{r}s_2s_3 & \cdots & \overline{r}s_2s_n \\ \overline{r}s_3s_1 & \overline{r}s_3s_2 & s_3^2 & \cdots & \overline{r}s_3s_n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \overline{r}s_ns_1 & \overline{r}s_ns_2 & \overline{r}s_ns_3 & \cdots & s_n^2 \end{bmatrix},$$

which is called the *shrinkage target*.

The idea of shrinkage estimation of the covariance matrix of returns is to take a weighted average of $\widehat{\boldsymbol{V}}$ and $\widehat{\boldsymbol{A}}$. With λ being a positive weight — known as the *shrinkage intensity* — assigned to $\widehat{\boldsymbol{A}}$, we can write the weighted average as

$$\widehat{\boldsymbol{C}} = (1-\lambda)\widehat{\boldsymbol{V}} + \lambda\widehat{\boldsymbol{A}},$$

where λ is a weighting factor satisfying the condition of $0 < \lambda < 1$. The determination of λ is a technical issue that is beyond the scope of FINANCE 601. The corresponding materials are more suitable for an advanced investment course instead.

To establish that $\widehat{\boldsymbol{C}}$ is positive definite for $0 < \lambda < 1$, we first decompose $\widehat{\boldsymbol{A}}$ into the sum of two matrices as follows:

$$\widehat{\boldsymbol{A}} = \widehat{\boldsymbol{B}} + \widehat{\boldsymbol{D}},$$

where

$$\widehat{\boldsymbol{B}} = \begin{bmatrix} \overline{r}s_1s_1 & \overline{r}s_1s_2 & \overline{r}s_1s_3 & \cdots & \overline{r}s_1s_n \\ \overline{r}s_2s_1 & \overline{r}s_2s_2 & \overline{r}s_2s_3 & \cdots & \overline{r}s_2s_n \\ \overline{r}s_3s_1 & \overline{r}s_3s_2 & \overline{r}s_3s_3 & \cdots & \overline{r}s_3s_n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \overline{r}s_ns_1 & \overline{r}s_ns_2 & \overline{r}s_ns_3 & \cdots & \overline{r}s_ns_n \end{bmatrix}$$

and

$$\widehat{\boldsymbol{D}} = \begin{bmatrix} (1-\overline{r})s_1s_1 & 0 & 0 & \cdots & 0 \\ 0 & (1-\overline{r})s_2s_2 & 0 & \cdots & 0 \\ 0 & 0 & (1-\overline{r})s_3s_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & (1-\overline{r})s_ns_n \end{bmatrix}.$$

Then, with $\widehat{\boldsymbol{B}}$ expressed as

$$\widehat{\boldsymbol{B}} = \widehat{\boldsymbol{g}}\widehat{\boldsymbol{g}}' = \begin{bmatrix} \sqrt{\bar{r}}s_1 \\ \sqrt{\bar{r}}s_2 \\ \sqrt{\bar{r}}s_3 \\ \vdots \\ \sqrt{\bar{r}}s_n \end{bmatrix} \begin{bmatrix} \sqrt{\bar{r}}s_1 & \sqrt{\bar{r}}s_2 & \sqrt{\bar{r}}s_3 & \cdots & \sqrt{\bar{r}}s_n \end{bmatrix},$$

we can write, for any n-element non-zero column vector $\boldsymbol{x}$,

$$\begin{aligned} \boldsymbol{x}'\widehat{\boldsymbol{C}}\,\boldsymbol{x} &= \boldsymbol{x}'[(1-\lambda)\widehat{\mathbf{V}} + \lambda\widehat{\mathbf{A}}]\boldsymbol{x} = (1-\lambda)\boldsymbol{x}'\widehat{\mathbf{V}}\,\boldsymbol{x} + \lambda\boldsymbol{x}'\widehat{\mathbf{A}}\,\boldsymbol{x} \\ &= (1-\lambda)\boldsymbol{x}'\widehat{\mathbf{V}}\,\boldsymbol{x} + \lambda\boldsymbol{x}'[\widehat{\boldsymbol{B}} + \widehat{\boldsymbol{D}}]\,\boldsymbol{x} = (1-\lambda)\boldsymbol{x}'\widehat{\mathbf{V}}\,\boldsymbol{x} + \lambda\boldsymbol{x}'[\widehat{\boldsymbol{g}}\widehat{\boldsymbol{g}}' + \widehat{\boldsymbol{D}}]\,\boldsymbol{x} \\ &= (1-\lambda)\boldsymbol{x}'\widehat{\mathbf{V}}\,\boldsymbol{x} + \lambda\boldsymbol{x}'\widehat{\boldsymbol{g}}\widehat{\boldsymbol{g}}'\boldsymbol{x} + \lambda\boldsymbol{x}'\widehat{\mathbf{D}}\,\boldsymbol{x} = (1-\lambda)\boldsymbol{x}'\widehat{\mathbf{V}}\,\boldsymbol{x} + \lambda(\widehat{\boldsymbol{g}}'\boldsymbol{x})'(\widehat{\boldsymbol{g}}'\boldsymbol{x}) + \lambda\boldsymbol{x}'\widehat{\mathbf{D}}\,\boldsymbol{x}. \end{aligned}$$

As $\widehat{\boldsymbol{V}}$ is positive semi-definite, we must have

$$\boldsymbol{x}'\widehat{\mathbf{V}}\,\boldsymbol{x} \geq 0,$$

regardless of the number of return observations used for its estimation. Next, as $(\widehat{\boldsymbol{g}}'\boldsymbol{x})$ is a 1×1 matrix, so is its transpose $(\widehat{\boldsymbol{g}}'\boldsymbol{x})'$. It follows that $(\widehat{\boldsymbol{g}}'\boldsymbol{x})'(\widehat{\boldsymbol{g}}'\boldsymbol{x})$, which is the square of a number, is always non-negative. Finally, as $\widehat{\boldsymbol{D}}$ is a diagonal matrix with all positive elements, $\boldsymbol{x}'\widehat{\boldsymbol{D}}\,\boldsymbol{x}$ is always positive. Therefore, for $0 < \lambda < 1$, as

$$\boldsymbol{x}'\widehat{\boldsymbol{C}}\,\boldsymbol{x} > 0,$$

the positive definiteness of $\widehat{\boldsymbol{C}}$ is confirmed.

5.23 Appendix E: The Constant Correlation Model for Portfolio Selection

For a given set of risky securities for mean-variance portfolio analysis, the required input data typically include the expected returns, the variances of returns, and the correlation matrix of returns for these securities. Given the dependence of the portfolio allocation results on the input data provided, their estimation is a relevant issue for finance academics and practitioners alike. As Elton, Gruber, and Urich [*Journal of Finance*, 33, (1978), pp.

1375-1384] indicate, "while analysts may be capable of providing estimates of returns and variances, the development of estimates of correlation coefficients from anything other than models utilizing historical data is highly unlikely." (p.1375). Thus, with the correlation matrix estimated from historical data, a practical question is whether there is an estimation method that dominates others in terms of estimation accuracy.

Surprisingly, Elton, Gruber, and Urich observe that sample averages of correlations outperform various more sophisticated models in forecasting the correlation matrix. These models include cases where security returns are related to some market or group indices, as well as a full historical model where forecasts of individual correlations of security returns are based on their corresponding historical values. Likewise, Chan, Karceski, and Lakonishok [*Review of Financial Studies*, 12, (1999), pp. 937-974] report that their sample average of correlations generates the lowest average of absolute errors in forecasting correlations.

The highly influential textbook by Elton, Gruber, Brown, and Goetzman, *Modern Portfolio Theory and Investment Analysis*, 9th Edition, 2014, John Wiley & Sons, covers in Chapter 9 a simple portfolio selection model known as the constant correlation model. Such a model is based on the work by Elton, Gruber, and Padberg [*Journal of Finance*, 11 (1976), pp. 1341-1357]. The attractiveness of such a model is that it is able to bypass the usual technical complications in efficient portfolio selection without short sales.

5.23.1 Efficient Portfolio Selection with Frictionless Short Sales

Consider a set of n securities, labeled as $i = 1, 2, \ldots, n$, where

$$\begin{aligned} \sigma_{ij} &= \rho\sigma_i\sigma_j, \quad \text{for} \quad i \neq j, \\ \text{and} \quad \sigma_i^2 &= \sigma_{ii} = \rho\sigma_i\sigma_i + (1-\rho)\sigma_i^2. \end{aligned}$$

Here, ρ is the constant correlation of security returns, and σ_i is the standard deviation of returns of security i. The purpose of decomposing σ_i^2 into two additive terms is for maintaining the form of $\rho\sigma_i\sigma_j$ when $i = j$. As $\sigma_{ii} \neq \rho\sigma_i\sigma_i$, the term $(1-\rho)\sigma_i^2$ is required to account for the difference between σ_{ii} and $\rho\sigma_i\sigma_i$.

Let us assume for now that frictionless short sales are allowed. For the tangency portfolio approach, let r be the μ-intercept of the tangent line on the (σ, μ)-plane. Let $\mu_1, \mu_2, \ldots, \mu_n$ be the expected returns of the n securities considered. Let also μ_p and σ_p^2 be the expected return and the variance of returns of the tangency portfolio p, respectively. Efficient allocations of investment funds are based on

$$\sum_{j=1}^{n} \sigma_{ij} z_j = \mu_i - r, \text{ for } i = 1, 2, \ldots, n, \tag{5.21}$$

where

$$z_j = \left(\frac{\mu_p - r}{\sigma_p^2}\right) x_j, \text{ for } i = 1, 2, \ldots, n,$$

under the condition of

$$\sum_{j=1}^{n} x_j = 1.$$

The efficient portfolio weights $x_1, x_2, \ldots, x_n$ can be deduced from the solved values of $z_1, z_2, \ldots, z_n$.

As we can write each σ_{ij} explicitly in terms of σ_i, σ_j, and ρ, depending on whether $i = j$ or $i \neq j$. Specifically, if $i = j$, there will be the extra term $(1-\rho)\sigma_i^2$. Thus, we have, according to equation (5.21),

$$\begin{aligned}
\sum_{j=1}^{n} \rho\sigma_i\sigma_j z_j + (1-\rho)\sigma_i^2 z_i &= \mu_i - r, \\
\underbrace{\rho\sum_{j=1}^{n} \sigma_j z_j}_{\phi} + (1-\rho)\sigma_i z_i &= \frac{\mu_i - r}{\sigma_i}, \\
\phi + (1-\rho)\sigma_i z_i &= \frac{\mu_i - r}{\sigma_i}, \qquad (5.22)\\
\text{and } \rho\phi + (1-\rho)\rho\sigma_i z_i &= \rho\left(\frac{\mu_i - r}{\sigma_i}\right), \quad \text{for } \; i = 1, 2, ..., n.
\end{aligned}$$

Summing up the n cases of i gives us

$$\begin{aligned}
n\rho\phi + (1-\rho)\rho\sum_{i=1}^{n} \sigma_i z_i &= \rho\sum_{i=1}^{n} \frac{\mu_i - r}{\sigma_i}, \\
n\rho\phi + (1-\rho)\phi &= \rho\sum_{i=1}^{n} \frac{\mu_i - r}{\sigma_i}, \\
\text{and } \; (1 + n\rho - \rho)\phi &= \rho\sum_{i=1}^{n} \frac{\mu_i - r}{\sigma_i}. \qquad (5.23)
\end{aligned}$$

Combining equations (5.22) and (5.23) leads to

$$\begin{aligned}
z_i &= \frac{1}{(1-\rho)\sigma_i}\left(\frac{\mu_i - r}{\sigma_i} - \phi\right), \quad \text{for } \; i = 1, 2, \ldots, n, \qquad (5.24)\\
\text{where } \phi &= \frac{\rho}{1 + n\rho - \rho}\sum_{j=1}^{n} \frac{\mu_j - r}{\sigma_j} \qquad (5.25)
\end{aligned}$$

Once $z_1, z_2, \ldots, z_n$ are known, the portfolio weights $x_1, x_2, \ldots, x_n$, as well as the expected return and the standard deviation of returns of the portfolio, can easily be determined.

Implications of the Analytical Solution

1. The n securities can be ranked and labeled according to a performance measure, the ratio of excess return to risk, such that

$$\frac{\mu_1 - r}{\sigma_1} \geq \frac{\mu_2 - r}{\sigma_2} \geq \cdots \geq \frac{\mu_n - r}{\sigma_n}.$$

   Here, ϕ is the cutoff rate of security performance. Notice that the case where $r = R_f$, the risk-free interest rate, can be interpreted intuitively.

2. If security i is held long, so are securities $1, 2, \ldots, i-1$, for $i > 1$.

3. If security i is held short, so are securities $i+1$, $i+2, \ldots, n$, for $i < n$.

4. The higher (lower) the ratio $(\mu_i - r)/\sigma_i$, the more attractive is security i for purchasing (short selling).

5. Each security i must have the ratio $(\mu_i - r)/\sigma_i$ above (below) a computed selection standard ϕ for it to be held long (short).

An Illustrative Example

Suppose that, for $n = 6$, where all returns are measured in %, we have the following:

$$r = 8, \ \rho = 0.4,$$

and

i	μ_i	$\mu_i - r$	σ_i	$(\mu_i - r)/\sigma_i$
1	29	21	14	1.50
2	23	15	12	1.25
3	20	12	12	1.00
4	16	8	10	0.80
5	20	12	16	0.75
6	12	4	8	0.50

From

$$\begin{aligned}\phi &= \frac{\rho}{1+n\rho-\rho}\sum_{j=1}^{n}\frac{\mu_j-r}{\sigma_j}\\ &= \frac{0.4}{1+6(0.4)-0.4}(1.50+1.25+1.00+0.80+0.75+0.50)=0.77333\end{aligned}$$

and

$$z_i=\frac{1}{(1-\rho)\sigma_i}\left(\frac{\mu_i-r}{\sigma_i}-\phi\right),\quad \text{for}\quad i=1,2,\ldots,6,$$

we have

i	z_i	x_i
1	0.086508	0.6692
2	0.066204	0.5122
3	0.031481	0.2435
4	0.004444	0.0344
5	−0.002431	−0.0188
6	−0.056944	−0.4405
	$(\mu_p-r)/\sigma_p^2=\sum_{j=1}^{6}z_j=0.129262$	$\sum_{j=1}^{6}x_j=1$

which leads to

$$\begin{aligned}\mu_p &= \sum_{i=1}^{6}x_i\mu_i=30.946\\ \text{and }\sigma_p &= \sqrt{\frac{\mu_p-r}{\sum_{j=1}^{6}z_j}}=13.324.\end{aligned}$$

5.23.2 Efficient Portfolio Selection without Short Sales.

Given the constant correlation characterization, as securities can be ranked according to their expected performance in terms of risk-return trade-off, here are some simple questions with simple answers: If we are to choose a single security for the portfolio, which security ought to be selected? Security 1 is the obvious choice. If we are to choose two securities instead for the portfolio, which two securities ought to be selected? Securities 1 and 2 are the obvious choices. Now we see a pattern. If we are to choose k securities for the portfolio, they ought to be securities $1,2,\ldots,k$.

However, how do we determine the exact number of securities for the portfolio? The illustration below based on the same 6-security example will shed some light on this issue. In the illustration, let us use $\phi_{(i)}$ to represent the value of ϕ for a portfolio consisting of securities $1, 2, \ldots, i$. For each i, we compare $(\mu_i - r)/\sigma_i$ and $\phi_{(i)}$. If

$$\frac{\mu_i - r}{\sigma_i} > \phi_{(i)},$$

we must also have, given how the individual securities are ranked and labeled,

$$\frac{\mu_1 - r}{\sigma_1} > \frac{\mu_2 - r}{\sigma_2} > \cdots > \frac{\mu_i - r}{\sigma_i} > \phi_{(i)}.$$

We simply look for the maximum i for which the condition

$$\frac{\mu_i - r}{\sigma_i} > \phi_{(i)}$$

holds.

In the same 6-security example, as we have

$$r = 8, \quad \rho = 0.4,$$

and

i	μ_i	$\mu_i - r$	σ_i	$(\mu_i - r)/\sigma_i$	$\phi_{(i)}$		
1	29	21	14	1.50	0.6000	based on	$i = 1$
2	23	15	12	1.25	0.7857		$i = 1$ to 2
3	20	12	12	1.00	$\boxed{0.8333}$		$i = 1$ to 3
4	16	8	10	0.80	0.8273		$i = 1$ to 4
5	20	12	16	0.75	0.8154		$i = 1$ to 5
6	12	4	8	0.50	0.7733		$i = 1$ to 6

In the example, when securities 1, 2, and 3 are selected for the portfolio, optimality is reached. The results are as follows:

i	z_i	x_i
1	0.079365	0.4949
2	0.057870	0.3608
3	0.023148	0.1443
	$(\mu_p - r)/\sigma_p^2 = \sum_{j=1}^{3} z_j = 0.1603836$	$\sum_{j=1}^{3} x_j = 1$

$$\begin{aligned}
x_4 &= x_5 = x_6 = 0, \\
\mu_p &= \sum\nolimits_{i=1}^{3} x_i \mu_i = 25.536, \\
\text{and } \sigma_p &= \sqrt{\frac{\mu_p - r}{\sum_{j=1}^{3} z_j}} = 10.457. \quad \blacksquare
\end{aligned}$$

Notice that, for the purpose of establishing the efficient frontier without short sales on the (σ, μ)-plane, we can simply repeat the same approach with different values of $r < \max(\mu_1, \mu_2, \ldots, \mu_6)$, including negative values of r.

5.24 Appendix F: Fisher's z-Transformation

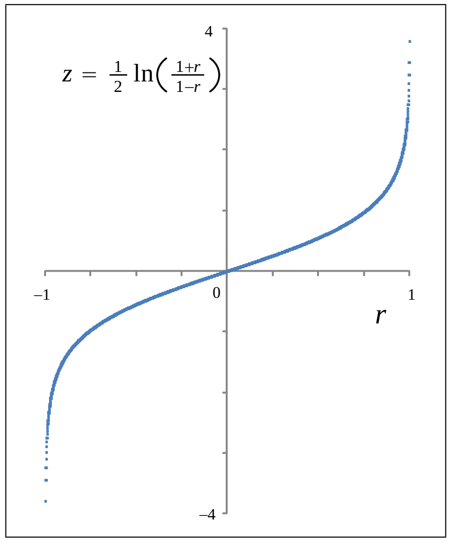
There are simple statistical tests for comparing correlations One of them requires the use of Fisher's z-transformation, which is a statistical tool in honour of Sir Ronald Aylmer Fisher who introduced it in 1921 (*Metron*, Volume 1, pp. 3-32). Before describing any statistical tests, we first provide below a sketch of Fisher's z-transformation and its adaptation for comparing a pair of correlations.

Let ρ be the true but unknown correlation of two random variables, and r be the sample correlation based on N observations. As correlations are necessarily in the range of -1 to 1, the probability distribution of r, as a random variable, is inevitably not symmetric when the magnitude of ρ is high. Suppose that, for example, the true correlation ρ is 0.9. A sample correlation that is greater than 0.9 must be in the narrow range of 0.9 to somewhere below 1.0; however, a sample correlation that is less than 0.9 can be over a very wide range of values. The presence of skewness in the probability distribution of r, which is also confined to the range of -1 to 1, will make it difficult to establish a confidence interval for ρ.

Following Fisher, we define

$$z = \frac{1}{2} \ln \left(\frac{1+r}{1-r} \right).$$

The corresponding graph on the (r, z)-plane is as follows:



The graph is symmetric about the origin; if we change the sign of r, the corresponding z will change its sign without changing its magnitude.

The distribution of z follows approximately a normal distribution with its mean being

$$\mu = \frac{1}{2}\ln\left(\frac{1+\rho}{1-\rho}\right)$$

and its variance being

$$\sigma^2 = \frac{1}{N-3}.$$

For a 95% confidence interval for the standard normal distribution, the critical values of the two tails, each with a 2.5% probability, are ± 1.960. For a 99% confidence interval, the critical values of the two tails, each with a 0.5% probability, are ± 2.576 instead.

To compare the sample correlations r_1 and r_2 for two pairs of variables in a two-tailed test based on N_1 and N_2 observations, respectively, where a variable is common in each pair of variables, the test statistic is

$$Z = \frac{z_1 - z_2}{\sigma_{1,2}},$$

where

$$z_1 = \frac{1}{2}\ln\left(\frac{1+r_1}{1-r_1}\right),$$

$$z_2 = \frac{1}{2}\ln\left(\frac{1+r_2}{1-r_2}\right),$$

and

$$\sigma_{1,2} = \sqrt{\frac{1}{N_1 - 3} + \frac{1}{N_2 - 3}}.$$

It is implicit that the expression of the standard error $\sigma_{1,2}$ is equivalent to

$$\sigma_{1,2}^2 = \sigma_1^2 + \sigma_2^2,$$

where

$$\sigma_1^2 = \frac{1}{N_1 - 3}$$

and

$$\sigma_2^2 = \frac{1}{N_2 - 3}.$$

This expression of $\sigma_{1,2}^2$ requires that the observations of the two variables involved be independent, so that there is no need to include a covariance term when decomposing $\sigma_{1,2}^2$. For a given confidence interval, if it turns out that the test statistic Z is beyond the corresponding critical values, then the null hypothesis of the two correlations being equal is rejected. Otherwise, the null hypothesis is not rejected.

5.25 Exercises

1. (C14) Stocks A and B are selling for \$25 and \$75 per share and are expected to be worth \$30 and \$100 per share, respectively, at the end of the period. The standard deviation of returns for each share is 40% and the correlation coefficient of returns between A and B is 0.5.

   (a) What is the variance of returns of a portfolio comprised of one share of A and one share of B?

   (b) What is the expected return of a portfolio comprised of 3 shares of A and 2 shares of B?

2. (C15) Given

$$\begin{array}{lll} \mu_1 = 0.1 & \sigma_1 = 0.05 & \rho_{12} = 0 \\ \mu_2 = 0.2 & \sigma_2 = 0.1 & \rho_{13} = 0 \\ \mu_3 = 0.3 & \sigma_3 = 0.2 & \rho_{23} = 0 \end{array}$$

   (a) What is the composition of the least risky portfolio, based on these three securities, that yields an expected return of 0.22?

(b) You have \$1,000 to invest in these three securities and have decided to invest \$400 in security 3. How do you divide the remaining \$600 between securities 1 and 2 to obtain a portfolio with the lowest risk?

3. (D19) For portfolio selection with frictionless short sales based on three risky securities, it is known that the expected return vector is $\boldsymbol{\mu} = \begin{bmatrix} 0.09 & 0.05 & 0.03 \end{bmatrix}'$ (with the prime indicating matrix transposition) and the lower triangle of the inverse of the covariance matrix $\boldsymbol{V}$, which is symmetric, is

$$\boldsymbol{V}^{-1} = \begin{bmatrix} 400 & & \\ 0 & 1,200 & \\ -300 & 0 & 2,000 \end{bmatrix}.$$

It is also known that the determinant of $\boldsymbol{V}^{-1}$ is 8.52×10^8. The efficient frontier based on the three securities is constructed by maximizing the slope of each tangent line with a predetermined μ-intercept on the (σ, μ)-plane, where μ (expected return) is the vertical axis and σ (standard deviation of returns) is the horizontal axis.

(a) Verify whether the above $\boldsymbol{V}^{-1}$ is positive definite according to Sylvester's criterion.

(b) Determine the range of possible values of the μ-intercept of the tangent line that ensures the success of the intended task of slope maximization. [Hint: The portfolio weight vector of the global minimum variance portfolio is given by $\boldsymbol{V}^{-1}\boldsymbol{\iota}\,(\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\iota})^{-1}$, where $\boldsymbol{\iota}$ is a column vector of ones.]

(c) It is known that the efficient portfolio weight vector $\boldsymbol{x}$ can be expressed as

$$\boldsymbol{x} = \boldsymbol{V}^{-1}\boldsymbol{M}\,(\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}\boldsymbol{r}_p.$$

For the three-security case here, $\boldsymbol{M}$ is a 3×2 matrix, where its first column is $\boldsymbol{\mu}$ and each element of its second column is 1, and $\boldsymbol{r}_p$ is a two-element column vector $\begin{bmatrix} \mu_p & 1 \end{bmatrix}'$, where $\mu_p = 0.06$ is the expected return of the corresponding efficient portfolio. Given that the transpose of $\boldsymbol{V}^{-1}\boldsymbol{M}\,(\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}$ is

$$\begin{bmatrix} 14.1975 & 7.4074 & ? \\ ? & 0.1037 & 1.4309 \end{bmatrix}$$

(where two elements, as indicated by question marks, are not provided), find the corresponding three-element vector $\boldsymbol{x}$.

4. (D21) Under the assumption of frictionless short sales, consider a portfolio selection problem based on three risky securities, for which $\boldsymbol{\mu}$ is a 3-element column vector of expected returns and $\boldsymbol{V}$ and a 3×3 symmetric covariance matrix of returns. Let $\boldsymbol{M}$ be a 3×2 matrix, where its first column is $\boldsymbol{\mu}$ and each element of its second column is 1.

   The 3-element column vector of efficient portfolio weights is $\boldsymbol{x} = \boldsymbol{V}^{-1}\boldsymbol{M}(\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}\boldsymbol{r}$, which can be expressed equivalently as $\boldsymbol{x}' = \boldsymbol{r}'(\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}\boldsymbol{M}'\boldsymbol{V}^{-1}$. The covariance of returns of any two portfolios with portfolio weight (column) vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ is $\boldsymbol{x}'\boldsymbol{V}\boldsymbol{y}$, which is the same as $\boldsymbol{y}'\boldsymbol{V}\boldsymbol{x}$. Here, the prime stands for matrix transposition, and $\boldsymbol{r}$ is a 2-element column vector, where the first element is the expected return of the efficient portfolio involved and the second element is 1. The portfolio weight (column) vector of the global minimum variance portfolio o is given by $\boldsymbol{x}_o = \boldsymbol{V}^{-1}\boldsymbol{\iota}\,(\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\iota})^{-1}$, where $\boldsymbol{\iota}$ is a 3-element column vector of ones. It is known that

$$
\begin{aligned}
\boldsymbol{x}'\boldsymbol{\iota} &= \boldsymbol{y}'\boldsymbol{\iota} = \boldsymbol{x}_o'\boldsymbol{\iota} = 1, \\
\boldsymbol{\mu} &= \begin{bmatrix} 0.07 & 0.10 & 0.12 \end{bmatrix}', \\
(\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1} &= \begin{bmatrix} 3.534562 & -0.287051 \\ -0.287051 & 0.025228 \end{bmatrix}, \\
\boldsymbol{V}^{-1}\boldsymbol{\iota} &= \begin{bmatrix} 357.883994 & 117.366899 & 46.606566 \end{bmatrix}', \\
\text{and } \boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\iota} &= 521.857459.
\end{aligned}
$$

   (a) According to Sylvestor's criterion, a symmetric matrix with all real elements is positive definite if all its leading principal minors are positive. It is known that the determinant of $\boldsymbol{V}$ is 1.4698×10^{-7}. Some partial information of $\boldsymbol{V}$ is also given as follows:

$$
\begin{bmatrix} 0.0025 & ? & ? \\ 0.0005 & 0.0064 & ? \\ 0.0010 & ? & 0.0100 \end{bmatrix}.
$$

   Determine the **correlation** of returns between securities 1 and 3. Can a risk-free portfolio be constructed by using the three securities? Explain clearly with computational support.

   (b) An efficient portfolio h with expected return of $\mu_h = 0.10$ is constructed. Given that

$$
\begin{bmatrix} 3.534562 & -0.287051 \\ -0.287051 & 0.025228 \end{bmatrix} \begin{bmatrix} 0.10 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.0664055 \\ -0.0034767 \end{bmatrix},
$$

determine the **correlation** of returns between portfolio h and the global minimum variance portfolio.

(c) For portfolio construction based on the three securities, by maximizing the slope of the tangent line on the (σ, μ)-plane, determine the range of possible values of the μ-intercept of the tangent line that ensures the success of the intended task of slope maximization.

(d) Suppose that, for a tangency portfolio p based on the three securities, the corresponding tangent line has an intercept of $r = 0.05$ on the axis of expected return. Let μ_p and σ_p be the expected return and the standard deviation of returns of portfolio p, respectively. It is known that

$$\boldsymbol{z} = \left(\frac{\mu_p - r}{\sigma_p^2} \right) \boldsymbol{x}_p = \boldsymbol{V}^{-1}(\boldsymbol{\mu} - r\boldsymbol{\iota}),$$

$$\mu_p = 0.098582, \ \boldsymbol{z} = \begin{bmatrix} 4.521177 & 6.140500 & 5.626807 \end{bmatrix}', \text{ and } \boldsymbol{z}'\boldsymbol{\iota} = 16.288484,$$

where $\boldsymbol{x}_p$ is a 3-element column vector of the tangency portfolio weights, for which $\boldsymbol{x}_p'\boldsymbol{\iota} = 1$. Determine σ_p.

5. (H03) Under the assumption of frictionless short sales, consider a portfolio selection problem based on three securities, for which $\boldsymbol{\mu} = \begin{bmatrix} 0.07 & 0.10 & 0.12 \end{bmatrix}'$ is a 3-element column vector of expected returns and $\boldsymbol{V}$ is a 3×3 covariance matrix of returns. Let $\boldsymbol{M}$ be a 3×2 matrix, where the first column is $\boldsymbol{\mu}$ and each element of the second column is 1. The 3-element column vector of efficient portfolio weights is $\boldsymbol{x} = \boldsymbol{V}^{-1}\boldsymbol{M}(\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}\boldsymbol{r}$, for which $\boldsymbol{x}' = \boldsymbol{r}'(\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}\boldsymbol{M}'\boldsymbol{V}^{-1}$ and $\boldsymbol{x}'\boldsymbol{\iota} = 1$, where $\boldsymbol{\iota}$ is a 3-element column vector of ones. Here, the prime stands for matrix transposition, and $\boldsymbol{r}$ is a 2-element column vector, where the first element is the expected return of the efficient portfolio involved and the second element is 1. The covariance of returns of any two portfolios with portfolio weight (column) vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ is $\boldsymbol{x}'\boldsymbol{V}\boldsymbol{y}$ or, equivalently, $\boldsymbol{y}'\boldsymbol{V}\boldsymbol{x}$, where $\boldsymbol{x}'\boldsymbol{\iota} = \boldsymbol{y}'\boldsymbol{\iota} = 1$. The portfolio weight (column) vector of the global minimum variance portfolio o is $\boldsymbol{x}_o = \boldsymbol{V}^{-1}\boldsymbol{\iota}\,(\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\iota})^{-1}$, where $\boldsymbol{x}_o'\boldsymbol{\iota} = 1$. It

is known that

$$\begin{aligned}
(\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1} &= \begin{bmatrix} 3.45620 & -0.28191 \\ -0.28191 & 0.02490 \end{bmatrix}, \\
\boldsymbol{V}^{-1}\boldsymbol{M}(\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1} &= \begin{bmatrix} ? & 9.0708 & 14.5575 \\ ? & -0.5133 & -1.0920 \end{bmatrix}', \\
\boldsymbol{V}^{-1}\boldsymbol{\iota} &= \begin{bmatrix} 356.1533 & 119.0316 & 50.1009 \end{bmatrix}', \\
\text{and } \boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\iota} &= 525.2858.
\end{aligned}$$

(a) Find the two missing elements in the 3×2 matrix $\boldsymbol{V}^{-1}\boldsymbol{M}(\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}$ above, as indicated by the two question marks there.

(b) Given that the determinant of $\boldsymbol{V}$ is 1.4870×10^{-7}, that the correlation of returns between securities 2 and 3 is 0.15, and that the lower triangle of $\boldsymbol{V}$ is

$$\begin{bmatrix} 0.0025 & & \\ 0.0005 & 0.0064 & \\ 0.0010 & ? & 0.0100 \end{bmatrix},$$

can a risk-free portfolio be constructed by using the three securities? **Explain clearly** with computational support. Then, find the missing element in the lower triangle of $\boldsymbol{V}$, as indicated by the question mark there.

(c) An efficient portfolio h with expected return of $\mu_h = 0.10$ is constructed. Given that

$$\begin{bmatrix} 3.45620 & -0.28191 \\ -0.28191 & 0.02490 \end{bmatrix} \begin{bmatrix} 0.10 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.063708 \\ -0.003293 \end{bmatrix},$$

determine the **correlation** of returns between portfolio h and the global minimum variance portfolio.

(d) For portfolio construction based on the three securities, by maximizing the slope of the tangent line on the (σ, μ)-plane, determine the range of possible values of the μ-intercept of the tangent line that ensures the success of the intended task of slope maximization.

(e) Suppose that, for a tangency portfolio p based on the three securities, the corresponding tangent line has an intercept of $r = 0.05$ on the axis of expected return. Let μ_p and σ_p be the expected return and the standard deviation of returns of portfolio p, respectively. It is known that

$$\boldsymbol{z} = \left(\frac{\mu_p - r}{\sigma_p^2}\right) \boldsymbol{x}_p = \boldsymbol{V}^{-1}(\boldsymbol{\mu} - r\boldsymbol{\iota}) = \begin{bmatrix} 4.406187 & 6.381977 & 5.793544 \end{bmatrix}',$$

$\mu_p = 0.099016$, and $\boldsymbol{z}'\boldsymbol{\iota} = 16.581708$, for which $\boldsymbol{x}_p'\boldsymbol{\iota} = 1$. Determine σ_p.

6. (I03) Under the assumption of frictionless short sales, consider a portfolio selection problem based on three securities, for which $\boldsymbol{\mu} = \begin{bmatrix} 0.07 & 0.10 & 0.12 \end{bmatrix}'$ is a 3-element column vector of expected returns and $\boldsymbol{V}$ is a 3×3 covariance matrix of returns. Let $\boldsymbol{M}$ be a 3×2 matrix, where the first column is $\boldsymbol{\mu}$ and each element of the second column is 1. The 3-element column vector of efficient portfolio weights is $\boldsymbol{x} = \boldsymbol{V}^{-1}\boldsymbol{M}(\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}\boldsymbol{r}$, for which $\boldsymbol{x}'\boldsymbol{\iota} = 1$, where $\boldsymbol{\iota}$ is a 3-element column vector of ones. Here, the prime stands for matrix transposition, and $\boldsymbol{r}$ is a 2-element column vector, where the first element is the expected return of the efficient portfolio involved and the second element is 1. The covariance of returns of any two portfolios with portfolio weight (column) vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ is $\boldsymbol{x}'\boldsymbol{V}\boldsymbol{y}$ or, equivalently, $\boldsymbol{y}'\boldsymbol{V}\boldsymbol{x}$, where $\boldsymbol{x}'\boldsymbol{\iota} = \boldsymbol{y}'\boldsymbol{\iota} = 1$. The portfolio weight (column) vector of the global minimum variance portfolio o is $\boldsymbol{x}_o = \boldsymbol{V}^{-1}\boldsymbol{\iota}\,(\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\iota})^{-1}$, for which $\boldsymbol{x}_o'\boldsymbol{\iota} = 1$. Here is some partial information:

$$\begin{aligned} (\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1} &= \begin{bmatrix} 3.49825 & -0.28540 \\ -0.28540 & 0.02516 \end{bmatrix}, \\ \boldsymbol{V}^{-1}\boldsymbol{\iota} &= \begin{bmatrix} 360.4713 & 124.5313 & 49.0091 \end{bmatrix}', \text{ and } \boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\iota} = 534.0118. \end{aligned}$$

(a) Given that the determinant of $\boldsymbol{V}$ is 1.4936×10^{-7}, that the correlation of returns between securities 2 and 3 is 0.15, and that the upper triangle of $\boldsymbol{V}$ is

$$\begin{bmatrix} 0.0025 & 0.0004 & 0.0010 \\ & 0.0064 & ? \\ & & 0.0100 \end{bmatrix},$$

can a risk-free portfolio be constructed by using the three securities? Explain clearly with computational support. Then, find the missing element in the upper triangle of $\boldsymbol{V}$, as indicated by the question mark there.

(b) An efficient portfolio k with expected return of $\mu_k = 0.10$ is constructed. It is known that

$$(\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}\begin{bmatrix} 0.10 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.064421 \\ -0.003383 \end{bmatrix}.$$

Determine the correlation of returns between portfolio k and the global minimum variance portfolio.

(c) Suppose that, for a tangency portfolio p based on the three securities, the corresponding tangent line has an intercept of $r = 0.05$ on the axis of expected return. Let μ_p and σ_p be the expected return and the standard deviation of returns of portfolio p, respectively. It is known that

$$\boldsymbol{z} = \left(\frac{\mu_p - r}{\sigma_p^2}\right)\boldsymbol{x}_p = \boldsymbol{V}^{-1}(\boldsymbol{\mu} - r\boldsymbol{\iota}) = \begin{bmatrix} 4.665238 & 6.440814 & 5.760578 \end{bmatrix}',$$

$\mu_p = 0.098533$, and $\boldsymbol{z}'\boldsymbol{\iota} = 16.866631$, for which $\boldsymbol{x}_p'\boldsymbol{\iota} = 1$. Determine σ_p.

(d) For portfolio construction based on the three securities, by maximizing the slope of the tangent line on the (σ, μ)-plane, determine the range of possible values of the μ-intercept of the tangent line that ensures the success of the intended task of slope maximization.

7. (J03) Under the assumption of frictionless short sales, consider a portfolio selection problem based on three securities, for which $\boldsymbol{\mu} = \begin{bmatrix} 0.07 & 0.10 & 0.12 \end{bmatrix}'$ is a 3-element column vector of expected returns and $\boldsymbol{V}$ is a 3×3 covariance matrix of returns. The three diagonal elements of $\boldsymbol{V}$ are $\sigma_{11} = 0.0025$, $\sigma_{22} = 0.0064$, and $\sigma_{33} = 0.0100$. The correlation of returns between any two different securities is 0.20. Let $\boldsymbol{M}$ be a 3×2 matrix, where the first column is $\boldsymbol{\mu}$ and each element of the second column is 1. The 3-element column vector of efficient portfolio weights is $\boldsymbol{x} = \boldsymbol{V}^{-1}\boldsymbol{M}(\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}\boldsymbol{r}$, for which $\boldsymbol{x}'\boldsymbol{\iota} = 1$, where $\boldsymbol{\iota}$ is a 3-element column vector of ones. Here, the prime stands for matrix transposition, and $\boldsymbol{r}$ is a 2-element column vector, where the first element is the expected return of the efficient portfolio involved and the second element is 1. The covariance of returns of any two portfolios with portfolio weight (column) vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ is $\boldsymbol{x}'\boldsymbol{V}\boldsymbol{y}$ or, equivalently, $\boldsymbol{y}'\boldsymbol{V}\boldsymbol{x}$, where $\boldsymbol{x}'\boldsymbol{\iota} = \boldsymbol{y}'\boldsymbol{\iota} = 1$. The portfolio weight (column) vector of the global minimum variance portfolio o is $\boldsymbol{x}_o = \boldsymbol{V}^{-1}\boldsymbol{\iota}\,(\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\iota})^{-1}$, for which $\boldsymbol{x}_o'\boldsymbol{\iota} = 1$. It is known that

$$\boldsymbol{V}^{-1}\boldsymbol{\iota} = \begin{bmatrix} 348.2143 & 100.4464 & 49.1071 \end{bmatrix}' \text{ and } \boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\iota} = 497.7679.$$

(a) Consider a portfolio q based on the three risky securities. Portfolio q, which has a portfolio weight (column) vector $\boldsymbol{w}$ satisfying the condition of $\boldsymbol{w}'\boldsymbol{\iota} = 1$, need not be efficient. Derive analytically an expression of the correlation of returns between portfolio o and portfolio q in terms of σ_o and σ_q.

(b) An efficient portfolio k with expected return of $\mu_k = 0.10$ is constructed. It is known that

$$(\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}\begin{bmatrix} 0.10 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.065231 \\ -0.003274 \end{bmatrix}.$$

Determine the correlation of returns between portfolio o and portfolio k. Determine also the correlation of returns between portfolio o and security 3.

(c) Suppose that, for a tangency portfolio p based on the three securities, the corresponding tangent line has an intercept of $r = 0.04$ on the axis of expected return. Let μ_p and σ_p be the expected return and the standard deviation of returns of portfolio p, respectively. It is known that

$$\boldsymbol{z} = \left(\frac{\mu_p - r}{\sigma_p^2}\right)\boldsymbol{x}_p = \boldsymbol{V}^{-1}(\boldsymbol{\mu} - r\boldsymbol{\iota}) = \begin{bmatrix} 7.321429 & 6.919643 & 6.160714 \end{bmatrix}',$$

$\mu_p = 0.095274$, and $\boldsymbol{z}'\boldsymbol{\iota} = 20.401786$, for which $\boldsymbol{x}_p'\boldsymbol{\iota} = 1$. Determine σ_p.

(d) For portfolio construction based on the three securities, by maximizing the slope of the tangent line on the (σ, μ)-plane, determine the range of possible values of the μ-intercept of the tangent line that ensures the success of the intended task of slope maximization.

8. (K01) For efficient portfolio selection based on three risky security and a risk-free security, it is known that the expected return vector, the covariance matrix of returns, and the risk-free interest rate are as follows:

$$\boldsymbol{\mu} = \begin{bmatrix} 5 \\ 8 \\ 15 \end{bmatrix}\%, \quad \boldsymbol{V} = \begin{bmatrix} 4 & 5 & 0 \\ 5 & 25 & 15 \\ 0 & 15 & 225 \end{bmatrix}(\%)^2, \quad \text{and} \quad R_f = 4.5\%.$$

Determine the allocation of investment funds, the expected portfolio return, and the standard deviation of portfolio returns under

(a) the assumption of frictionless short sales.

(b) Lintner's treatment of short sales.

(c) What implications can be deduced from comparing the results in **Part (a)** and **Part (b)** under these two alternative assumptions? Explain clearly.

5.26 Highlights of This Chapter

- A normal distribution is characterized by two parameters, μ and σ, which stand for its mean and standard deviation, respectively. The probability density function of the probability distribution

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]$$

  appears graphically as a bell-shaped curve, for $-\infty < x < \infty$.

- Given any two consumption goods, the term *indifference curve* is a graph showing their various combinations that provide the consumer involved an equal level of satisfaction (utility). When applied to an investment context, where utility is characterized by the parameters μ and σ in a normal distribution, we have, for a risk-averse investor,

$$\frac{d\mu}{d\sigma} > 0 \text{ and } \frac{d^2\mu}{d\sigma^2} > 0$$

  along each indifference curve.

- The mean-variance approach, where investment decisions are based on μ and σ^2 (or, equivalently, μ and σ), is inappropriate if the underlying probability distribution is not a normal distribution, unless a problematic assumption about the investor's utility function is imposed.

- The worst investment outcome for holding a stock is a 100% return, not a $-\infty$ return. Thus, the normality assumption of the return distribution, which allows random returns to approach $-\infty$, can be problematic. A simple remedy is to perform a logarithmic transformation of each effective return R over any given period of time. By using the *central limit theorem* in statistics, we can establish that the distribution of $\ln(1 + R)$ is approximately normal. The resulting return, $\mathbf{R} = \ln(1 + R)$, is the corresponding continuously compounded return over the same period of time.

- Under the current T+1 process, if someone buys a stock, it takes a business day from the transaction date (T) for the ownership of the stock certificate, as well as the money involved, to change hands between the buyer and the seller.

- For each dividend payment, there are four relevant dates, which include the *announcement date*, the *ex-dividend date*, the *record date*, and the *payment date* in a chronological order. Under the current T+1 process, the ex-dividend date and the record date are on two adjacent business days. If the stock ownership is transferred via a buy/sell transaction prior to the ex-dividend date, the buyer is entitled to receive the dividend involved. If the transaction is on or after the ex-dividend date instead, the seller is entitled to receive the dividend involved, as the seller is still the owner of the stock on the record date. In the latter case, as the buyer will not receive the dividend involved, the change in the stock price — from the date just before the ex-dividend date to the ex-dividend date — has a negative component.

- Suppose that a portfolio p is formed. Under the assumption of frictionless short sales, where the short seller not only provides zero margin deposit but also has immediate access to the short-sale proceeds for investing in other securities, the expected portfolio return is

$$\mu_p = \sum_{i=1}^{n} x_i \mu_i,$$

and the variance of portfolio returns is

$$\sigma_p^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i x_j \sigma_{ij},$$

for which

$$\sum_{i=1}^{n} x_i = 1.$$

Here, x_i is the proportion of investment funds allocated to security i, μ_i is the expected return of security i, and σ_{ij} is the covariance of returns between securities i and j, for $i, j = 1, 2, \ldots, n$. Implicitly, we have $\sigma_{ii} = \sigma_i^2$ and $\sigma_{ij} = \sigma_{ji}$. The correlation of returns between securities i and j is ρ_{ij}, for which the following condition always holds regardless of whether there are any assumptions about short sales:

$$-1 \leq \rho_{ij} = \frac{\sigma_{ij}}{\sigma_i \sigma_j} \leq 1, \text{ for } i, j = 1, 2, \ldots, n.$$

- In matrix notation, letting

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{bmatrix}, \quad \boldsymbol{\iota} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}, \quad \text{and} \quad \boldsymbol{V} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_{nn} \end{bmatrix},$$

we can write

$$\mu_p = \boldsymbol{x}'\boldsymbol{\mu}, \quad \sigma_p^2 = \boldsymbol{x}'\boldsymbol{V}\boldsymbol{x}, \quad \text{and} \quad \boldsymbol{x}'\boldsymbol{\iota} = 1$$

for

$$\mu_p = \sum\nolimits_{i=1}^{n} x_i \mu_i, \quad \sigma_p^2 = \sum\nolimits_{i=1}^{n} \sum\nolimits_{j=1}^{n} x_i x_j \sigma_{ij}, \quad \text{and} \quad \sum\nolimits_{i=1}^{n} x_i = 1,$$

respectively, where the prime stands for matrix transposition.

- The assumption of frictionless short sales is for analytical convenience. In practice, however, under institutional procedures for short selling, the short seller does not have immediate access to the proceeds as provided by the buyer of the stock in a short-sale transaction, and the short-sale proceeds are held as collateral for the borrowed stock. The short seller earns either no interest or only a partial interest rebate from the brokerage firm for the interest it earns. The short seller must deposit with the brokerage firm at least 50% of the value of the shares held short to fulfill the initial margin requirements; cash, interest bearing T-bills, and other securities that the short seller owns can be used to provide the required margin deposits, and the short seller will earn interests on any cash deposits. Further, any dividend payments from the issuing company of the stock, which the lender of the stock is entitled to receive, must be provided by the short seller. In the event that the lender sells the stock, the short seller has to obtain the stock certificate at the market price if another loan of the same stock cannot be arranged.

- Short selling in practice, even when it is not intended to be for speculative purposes, is still very risky. Long-short investing requires the portfolio manager's ability to identify both undervalued and overvalued stocks. An obvious risk is the potential misjudgment. Unlike long investments, short selling does not have limited liabilities. An overvalued stock may stay overvalued for a long time. Short squeezes can be triggered either by a company's hostility or aggressiveness towards its short sellers or by other investors' trading activities. Sudden price increases of companies being taken over by other companies will lead to great losses to short sellers. Further, short selling in a bullish market is highly risky as most stocks move with the rising tide.

- The $1/N$ investment strategy is about allocating investment funds equally among the securities considered. The variance of returns of an equally weighted portfolio p based on n securities is

$$\sigma_p^2 = \frac{\overline{Var}}{n} + \frac{n-1}{n} \cdot \overline{Cov},$$

where $\overline{Var}$ is the average variance of individual security returns and $\overline{Cov}$ be the average covariance of individual security returns. Given that $\overline{Var}$ and $\overline{Cov}$ are finite as n increases, $\overline{Var}/n$ approaches zero and σ_p^2 approaches $\overline{Cov}$ as n approaches infinity.

As more securities are used to implement the $1/N$ investment strategy, the contributions of the covariances of individual security returns to the portfolio risk will become more important, as compared to the contributions of the variances of individual security returns. If the securities considered are highly correlated in their returns, the achievable portfolio risk reductions will be limited.

- Under the assumption of frictionless short sales, Roll has established that, for an n-security case, the portfolio weight vector $\boldsymbol{x}$ corresponding to the minimum variance portfolio p requiring an expected return μ_p is

$$\boldsymbol{x} = \boldsymbol{V}^{-1}\boldsymbol{M}\,(\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}\boldsymbol{r}_p,$$

where

$$\boldsymbol{M} = \begin{bmatrix} \boldsymbol{\mu} & \boldsymbol{\iota} \end{bmatrix} \quad \text{and} \quad \boldsymbol{r}_p = \begin{bmatrix} \mu_p \\ 1 \end{bmatrix}.$$

The variance of returns of portfolio p is

$$\sigma_p^2 = \boldsymbol{r}_p'(\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}\boldsymbol{r}_p,$$

and the covariance of returns between portfolio p and another minimum variance portfolio q is

$$\sigma_{pq} = \boldsymbol{r}_p'(\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}\boldsymbol{r}_q,$$

where

$$\boldsymbol{r}_q = \begin{bmatrix} \mu_q \\ 1 \end{bmatrix}.$$

Denote portfolio o as the global minimum variance portfolio. The corresponding portfolio weight vector, expected return, and variance of returns are

$$\boldsymbol{x}_o = \boldsymbol{V}^{-1}\boldsymbol{\iota}(\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\iota})^{-1}, \quad \mu_o = (\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\iota})^{-1}(\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\mu}), \quad \text{and} \quad \sigma_o^2 = (\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\iota})^{-1},$$

respectively. Each portfolio weight x_i in portfolio o is the sum of all n elements in row i of $\boldsymbol{V}^{-1}$, divided by the sum of all elements in $\boldsymbol{V}^{-1}$. The variance of returns of portfolio o is the reciprocal of the sum of all elements in $\boldsymbol{V}^{-1}$.

- Under the assumption of frictionless short sales, let portfolio o be the global minimum variance portfolio based on n securities. Let k be any portfolio based on the same n securities. The covariance of returns between portfolios o and k is the same as the variance of returns of portfolio o, regardless of whether portfolio k is efficient.

- To verify whether the computations of portfolio weights based on

$$\boldsymbol{x} = \boldsymbol{V}^{-1}\boldsymbol{M}\,(\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}\boldsymbol{r}_p$$

have been performed correctly under the assumption of frictionless short sales, we write

$$\begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \\ \vdots & \vdots \\ a_n & b_n \end{bmatrix} = \boldsymbol{V}^{-1}\boldsymbol{M}\,(\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}.$$

Correct computational results satisfy the following two conditions:

$$a_1 + a_2 + \cdots + a_n = 0 \;\text{ and }\; b_1 + b_2 + \cdots + b_n = 1.$$

- Under the condition of full allocations of investment funds among the n securities considered, each minimum variance portfolio p corresponding to a given expected return μ_p is an efficient portfolio if $\mu_p \geq \mu_o$, where μ_o is the expected return of the global minimum variance portfolio. If $\mu_p < \mu_o$ instead, portfolio p is not an efficient portfolio.

- Combinations of a risk-free security f and a portfolio p based on n risky securities are captured by a line on the (σ, μ)-plane. The expected performance of all portfolios on the line is $\theta = (\mu_p - R_f)/\sigma_p$, which is the slope of the line, where μ_p is the expected return of portfolio p, σ_p is the standard deviation of returns of portfolio p, and R_f is the risk-free interest rate. A specific portfolio p is the tangency portfolio; it is the portfolio that maximizes θ.

- Under the assumption of frictionless short sales, Lintner has shown that the tangency portfolio weights can be computed from

$$x_i = \frac{z_i}{\sum_{j=1}^{n} z_j}, \text{ for } i = 1, 2, \ldots, n,$$

where z_i is element i of the n-element column vector

$$\boldsymbol{z} = \boldsymbol{V}^{-1}(\boldsymbol{\mu} - R_f\boldsymbol{\iota}).$$

With $z_1, z_2, \ldots, z_n$ and $x_1, x_2, \ldots, x_n$ established, the expected portfolio return and standard deviation of returns can be computed from

$$\mu_p = \sum\nolimits_{i=1}^{n} x_i \mu_i$$

and

$$\sigma_p = \sqrt{\frac{\mu_p - R_f}{\sum_{j=1}^{n} z_j}},$$

respectively. For the results to be meaningful, R_f must be lower than the expected return μ_o of the global minimum variance portfolio based on the n risky securities. If $R_f > \mu_o$, we have minimization of θ instead. If $R_f = \mu_o$, portfolio optimization cannot be performed.

- A portfolio q, which corresponds to the highest achievable utility for the investor, is a combination of the risk-free security and the tangency portfolio p. While the choice of portfolio q depends on the investor's utility function, the determination of the tangency portfolio p does not. To achieve a portfolio q on the line on the (σ, μ)-plane between point $(0, R_f)$ and the tangency portfolio p, the investor allocates investment funds between the risk-free security and portfolio p. If risk-free borrowing is available at the same interest rate R_f, a portfolio q on the line above portfolio p is the case where the investor borrows additional funds for investing more in portfolio p.

- Under the assumption of frictionless short sales, Roll's approach to construct the efficient frontier based only on n risky securities is by varying the required expected return μ_p in
$$\boldsymbol{x} = \boldsymbol{V}^{-1}\boldsymbol{M}\,(\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}\boldsymbol{r}_p.$$
Lintner's approach for the same task is by substituting the risk-free interest rate with different values of r, which is the μ-intercept of the tangent line for maximization of $\theta = (\mu_p - r)/\sigma_p$. For the task to be successful, each attempted value of r must be lower than μ_o, the expected return of the global minimum variance portfolio.

- The separation between (i) the construction of the optimal portfolio based on the n risky securities considered and (ii) the decision about the allocations of investment funds between the risk-free security and the risky portfolio thus constructed makes it possible to construct portfolios based on risky securities without paying any attention to the investor's specific risk-return preferences.

- The effect of holding *idle cash* can be interpreted in terms of allocating some of the available investment funds to a risk-free security that earns no returns. Efficient allocations of investment funds this way can be captured graphically by drawing a tangent line from the origin of the (σ, μ)-plane to the efficient frontier. The part of the curve between the tangency portfolio and the global minimum variance portfolio — which is originally considered efficient for full allocations of investment funds — is no longer efficient. The reason is that this part of the curve, which is below the tangent line on the (σ, μ)-plane is dominated by the tangent line segment.

- As a remedy for the shortcomings of the assumption of frictionless short sales, Lintner's treatment of short sale provides a closer characterization of reality without sacrificing analytical tractability. In Lintner's formulation — which focuses on the μ-intercept of the tangent line on the (σ, μ)-plane being the risk-free interest rate R_f — the short seller not only has no access to the short-sale proceeds, but also is required to provide to an escrow account a deposit that is equal to the full amount of short-sale proceeds; however, the short seller earns interests from both the deposit and the short-sale proceeds.

  For an n-security case, the tangency portfolio weights are determined from

$$x_i = \frac{z_i}{\sum_{j=1}^{n} |z_j|}, \text{ for } i = 1, 2, \ldots, n,$$

  where z_i is still element i of the n-element column vector

$$\boldsymbol{z} = \boldsymbol{V}^{-1}(\boldsymbol{\mu} - R_f \boldsymbol{\iota}).$$

  The expected portfolio return μ_p and standard deviation of returns σ_p can be computed from

$$\mu_p - R_f = \sum_{i=1}^{n} x_i(\mu_i - R_f)$$

  and

$$\sigma_p = \sqrt{\frac{\mu_p - R_f}{\sum_{j=1}^{n} |z_j|}},$$

  respectively.

- All covariance matrices are *positive semi-definite*, and invertible covariance matrices are *positive definite*. To illustrate, let us denote an $n \times n$ covariance matrix of returns as $\boldsymbol{V}$. The matrix is positive semi-definite, as $\boldsymbol{x}'\boldsymbol{V}\boldsymbol{x}$ is never negative for any n-element column vector $\boldsymbol{x}$, including the case where all n elements of $\boldsymbol{x}$ are zeros. The positive definiteness of $\boldsymbol{V}$ requires that $\boldsymbol{x}'\boldsymbol{V}\boldsymbol{x}$ be positive for any n-element column vector $\boldsymbol{x}$ whose elements are not all zeros. If $\boldsymbol{V}$ is invertible, it is also positive definite.

- The positive definiteness of any $n \times n$ real symmetric matrix $\boldsymbol{A}$ can be verified by using *Sylvester's criterion*, which states that a real symmetric matrix is positive definite if and only if all of its *leading principal minors* — each of which is the determinant of a square sub-matrix containing the $(1,1)$-element — are positive. The $n \times n$ matrix $\boldsymbol{A}$ has n leading principal minors.

- To implement a portfolio selection model requires reliable values of the input parameters involved. The use of past return observations for their estimation is a good starting point in practice. To illustrate, suppose that R_{it} is the return of security i as observed in month t, for $i = 1, 2, \ldots, n$ and $t = 1, 2, \ldots, T$. Under the assumption of stationary (stable) probability distributions of returns, we can use the sample mean return,
$$\overline{R}_i = \frac{1}{T} \sum\nolimits_{t=1}^{T} R_{it},$$
as an estimate of the true expected return μ_i, for $i = 1, 2, \ldots, n$. We can also use the sample covariance of returns,
$$\widehat{\sigma}_{ij} = \frac{1}{T-1} \sum\nolimits_{t=1}^{T} \left(R_{it} - \overline{R}_i\right) \left(R_{jt} - \overline{R}_j\right),$$
as an estimate of the true covariance of returns between securities i and j, for $i, j = 1, 2, \ldots, n$. Implicitly, if $i = j$, we can use the sample variance of returns,
$$\widehat{\sigma}_{ii} = \widehat{\sigma}_i^2 = \frac{1}{T-1} \sum\nolimits_{t=1}^{T} \left(R_{it} - \overline{R}_i\right)^2,$$
as an estimate of the true variance of returns σ_i^2. The sample correlation of returns between securities i and j can be estimated from
$$\widehat{\rho}_{ij} = \frac{\widehat{\sigma}_{ij}}{\widehat{\sigma}_i \widehat{\sigma}_j}.$$

- The sample covariance matrix is always positive semi-definite, and an invertible sample covariance matrix is positive definite. For a sample covariance matrix of returns to be usable for implementing portfolio selection models, it must be invertible. If the matrix has been estimated with long enough series of past return observations, all that is required (for the matrix to be invertible) is essentially the removal of securities that can be replicated by portfolios of some other securities for the same portfolio consideration.

- In practice, the use of past return observations to form expectations can also serve as a starting point in generating a set of acceptable input parameters. Insights of financial analysts are often deemed necessary for revising the estimated values of the input parameters. If there are changes to the elements of the sample covariance matrix of returns for potential improvements in the quality of the estimated input parameters, whether the resulting matrix will remain positive definite cannot be assured. Thus, its positive definiteness must be verified, before it can be used for implementing a portfolio selection model.

- The choice of the sample period is often a trade-off between having a longer sample to reduce estimation errors and having a shorter sample to avoid unmanageable changes in the probability distributions of security returns. The pursuit of market momentum being of considerable interest to many investment practitioners, they may be tempted to rely on short samples of security returns to estimate input parameters for implementing practical portfolio selection models. The problems with using short samples, however, are two-fold: (*i*) severe estimation errors and (*ii*) the failure to satisfy the positive definiteness requirement for the covariance matrix of returns.

  Shrinkage estimation, which is a remedy for the above problems, is about taking a positively weighted average of the sample covariance matrix of returns and a structured matrix of the same dimensions. A suitable structured matrix is a covariance matrix of returns where all covariances (excluding the variances) are characterized by a constant correlation, such as the sample mean of all pairwise correlations between different securities. Such a structured matrix is positive definite. A positively weighted average of the sample covariance matrix — which is always positive semi-definite regardless of the number of past return observations for its estimation — and such a structured matrix is positive definite.