Stochastic Mean-Variance Optimization in Portfolio Analysis

Midori Munechika*

Contents

1. Introduction

The Markowitz mean-variance analysis that laid the foundations of modern portfolio theory (MPT) shows how rational investors should construct their optimal portfolios under conditions of uncertainty. Why is his theory called mean-variance analysis? It is because only two parameters, mean return and return variance (or standard deviation), are taken into account to construct the optimal portfolio. Essentially, the aim of investors is to maximize the terminal value of their investment. The mean-variance approach needs to justify the switch from maximization of utility of wealth to utility consisting of two characteristics: mean return and return variance. Theoretically, this can be done by making two hypotheses behind the theory: the normal distribution hypothesis about return distributions and a quadratic utility function (see Munechika [2002]).

Mean-variance analysis is clearly classified as a normative theory rather than a positive one, while it approaches normative issues in a positive context. Sharpe [1963] summarizes the process of Markowitz's portfolio selection into three steps: (1) making probabilistic estimates of

*Financial support from Grant-in-Aid for Scientific Research (C-No.15530227) of Japan Society for the Promotion of Science (JSPS) under the Ministry of Education, Culture, Sports, Science and Technology (MEXT) and a grant from Toyo University are gratefully acknowledged.

the future performances of asset returns, (2) analyzing those estimates to determine an efficient set of portfolios, and (3) selecting from that set the portfolios best suited to the investor's preferences. Corresponding to these three steps, a portfolio optimization methodology is composed of three key ingredients: a return forecast, an optimizer (i.e., a software program used in the computational procedure) and a utility function. The first step of the process (i.e., the normal distribution hypothesis) makes the Markowitz model into stochastic one.

The major concern of this article is to introduce a method of stochastic optimization by applying Monte Carlo simulation to the second step and to examine its usefulness and its limitations. In Section 2, I define portfolio return and risk by using matrix notation and point out that a covariance matrix is a concise form for the information of volatilities and correlations of assets and that is useful for constructing optimal portfolios. In Section 3, the mean-variance optimization problem is mathematically formulated and its solution is provided by quadratic programming. In Section 4, I implement the mean-variance optimization problem by using the method of Monte Carlo simulation. In Section 5, the stylized facts about return distributions are discussed in the context of stochastic optimization.

2. Matrix Approach to Portfolio Risk

I first formulate the problem of Markowitz's mean-variance optimization in a formal mathematical context. Suppose that a portfolio is composed of n risky assets. The return on a portfolio is a weighted average of individual asset returns.

$$R_p = w_1 \cdot R_1 + w_2 \cdot R_2 + \cdots + w_n \cdot R_n = \sum_{i=1}^{n} w_i R_i \tag{1}$$

where $\sum_{i=1}^{n} w_i = 1$. The return on the i-th risky asset is R_i and the portion of the i-th asset held in the portfolio is w_i.

Portfolio risk is defined as portfolio returns variance.

$$Var\left[R_p\right] = \sigma_p^2 \tag{2}$$

The variance of the portfolio return is the mixture of variability of returns for respective assets and their co-movement, which can be expressed in a matrix format as in the following table.

	Asset 1	Asset 2	$\cdots$	Asset n
Asset 1	$w_1^2\sigma_1^2$	$w_1w_2\sigma_{12}$	$\cdots$	$w_1w_n\sigma_{1n}$
Asset 2	$w_2w_1\sigma_{21}$	$w_2^2\sigma_2^2$	$\cdots$	$w_2w_n\sigma_{2n}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
Asset n	$w_nw_1\sigma_{n1}$	$w_nw_2\sigma_{n2}$	$\cdots$	$w_n^2\sigma_n^2$

The diagonal terms contain the variances of the individual assets and the off-diagonal terms contain the covariances. The covariance between returns on the i th asset and the j th asset is given by:

$$Cov(R_i, R_j) = E[(R_i - \mu_i)(R_j - \mu_j)] = \sigma_{ij} \tag{3}$$

where μ_i and μ_j are the mean returns of R_i and R_j respectively. The sign of the covariance will indicate the direction of covariance R_i and R_j. Thus, the variance of a portfolio's return can be calculated as the sum of all the cells of the table.

$$\sigma_p^2 = w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + \cdots + w_n^2\sigma_n^2 + \sum_{i=1}^{1}\sum_{\substack{j=1\\ i\neq j}}^{n} w_1w_j\sigma_{1j} + \sum_{i=2}^{2}\sum_{\substack{j=1\\ i\neq j}}^{n} w_2w_j\sigma_{2j} + \cdots + \sum_{i=n}^{n}\sum_{\substack{j=1\\ i\neq j}}^{n} w_nw_j\sigma_{nj}$$

$$= \underbrace{\sum_{i=1}^{n} w_i^2\sigma_i^2}_{\text{variance term}} + \underbrace{\sum_{i=1}^{n}\sum_{\substack{j=1\\ i\neq j}}^{n} w_iw_j\sigma_{ij}}_{\text{covariance term}} \tag{4}$$

In the double summation $i \neq j$ of the covariance term, if $i = j$, then the term would be $w_iw_i\sigma_{ii} = w_i^2\sigma_i^2$ since $\sigma_{ii} = E[(R_i - \mu_i)(R_i - \mu_j)] = E(R_i - \mu_i)^2 = \sigma_{ii}$. This is the exactly the variance term in the first summation. The portfolio risk can also be written as:

$$\sigma_p^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} w_iw_j\sigma_{ij} \tag{5}$$

It is convenient to present the portfolio return and risk in the form of matrix notation as shown in the above explanation. The set of asset returns is expressed as a column vector consisting of the random variable $R_i, \cdots, R_n$:

$$\sum_{i=1}^{n} R_i = \mathbf{r} = \begin{bmatrix} R_1 \\ \vdots \\ R_n \end{bmatrix} \tag{6}$$

The set of portfolio weights is:

$$\sum_{i=1}^{n} w_i = \mathbf{w} = \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix} \tag{7}$$

The return on a portfolio of equation (1) may be written as:

$$R_p = \sum_{i=1}^{n} w_i R_i = \begin{bmatrix} w_1 R_1 & \cdots & w_n R_n \end{bmatrix} = \begin{bmatrix} w_1 & \cdots & w_n \end{bmatrix} \begin{bmatrix} R_1 \\ \vdots \\ R_n \end{bmatrix} = \mathbf{w}^T \mathbf{r} \tag{8}$$

where $\mathbf{w}^T$ is a transpose of $\mathbf{w}$:

$$\mathbf{w}^T = \begin{bmatrix} w_1 & \cdots & w_n \end{bmatrix} \tag{9}$$

The portfolio variance of equation (5) is expressed as:

$$\sigma_p^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} w_i w_j \sigma_{ij} = \begin{bmatrix} w_1^2 \sigma_1^2 & w_1 w_2 \sigma_{12} & \cdots & w_1 w_n \sigma_{1n} \\ w_2 w_1 \sigma_{21} & w_2^2 \sigma_2^2 & \cdots & w_2 w_n \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ w_n w_1 \sigma_{n1} & w_n w_2 \sigma_{n2} & \cdots & w_n^2 \sigma_n^2 \end{bmatrix} \tag{10}$$

which can be broken down into the following matrix multiplications:

$$\sigma_p^2 = \begin{bmatrix} w_1 & w_2 & \cdots & w_n \end{bmatrix} \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_n^2 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} = \mathbf{w}^T \mathbf{V} \mathbf{w} \tag{11}$$

where $\mathbf{V}$ is a variance-covariance matrix with the variance terms on the diagonal and the covariance terms on the off-diagonal. The variance-covariance matrix is also referred to as the covariance matrix. The covariance matrix V is always square and symmetric. In fact, the covariance σ_{ij} between risky asset i and risky asset j will be equal to the covariance between risky asset j and risky asset i:

$$\sigma_{ij} = \sigma_{ji} \tag{12}$$

Then, $\mathbf{V}$ can be arranged in the following square matrix:

$$\mathbf{V} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_n^2 \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{12} & \sigma_2^2 & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1n} & \sigma_{2n} & \cdots & \sigma_n^2 \end{bmatrix} \tag{13}$$

Therefore, $\mathbf{V}$ is also symmetric.

The covariance matrix demonstrates how to reduce the portfolio risk through portfolio diversification. As the number of assets n increases, the total number of elements in the covariance matrix becomes n^2, the number of variance terms becomes n, and the number of covariance terms thus becomes $(n^2 - n)$. For example, a portfolio of 100 stocks has 100 variance terms and 9900 covariance terms. It is clear that the risk of a portfolio with many assets is more dependent on the covariances between the individual assets than on the variances of the individual assets. Therefore, the degree of co-movements between different pairs of stocks in a portfolio is crucial to estimate and reduce the portfolio risk.

There is another typical measure of the degree of co-movement between two variables: correlation. The magnitude of covariance, σ_{ij} depends not only on the degree of co-movement among the returns but also on their sizes. For instance, the covariance of monthly returns will normally be greater than the covariance of any daily returns in the same market because monthly returns are of a much greater order of magnitude than daily returns. The scales of measurement will affect the magnitude of covariance. Therefore, a preferable measure to make comparisons is correlation, which is the covariance divided by the product of the standard deviations:

$$\rho_{ij} = \frac{\sigma_{ij}}{\sigma_i \sigma_j} \tag{14}$$

The correlation coefficient ρ_{ij} has the same sign as the covariance, but its number always lies between -1 and $+1$, which is unaffected by any scaling of the variables. We obtain the correlation matrix by dividing σ_{ij} by $\sigma_i \sigma_j$:

$$\mathbf{C} = \begin{bmatrix} 1 & \rho_{12} & \cdots & \rho_{1n} \\ \rho_{21} & 1 & \cdots & \rho_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{n1} & \rho_{n2} & \cdots & 1 \end{bmatrix} \tag{15}$$

When the returns on the i-th asset and the j-th asset are independent random variables, they are not correlated to each other. The covariance matrix becomes a diagonal matrix, that is, a square matrix in which elements are all zero except the ones on the diagonal.

$$\mathbf{V} = \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_n^2 \end{bmatrix} \tag{16}$$

This means that the portfolio risk stems only from the variances of the individual assets. In this case, the correlation matrix becomes an identity matrix, $\mathbf{I}$, which is a scalar matrix with ones on the diagonal.

$$\mathbf{C} = \mathbf{I} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \tag{17}$$

It is important to note that the covariance matrix is a concise form for information on the two key determinants of a portfolio risk, volatilities and correlations. Volatility is a measure of the dispersion in a probability distribution of the asset returns. The most common measure of dispersion is the standard deviation, σ_i of a random variable, that is, the square root of its variance, σ_i^2. Therefore, a succinct form for information on all the volatilities and correlations in a portfolio can be obtained through simple mathematical operations on the elements of the covariance matrix.

3. Mean-Variance Optimization Problem

The Markowitz mean-variance optimization problem can be solved by quadratic programming. Quadratic programming is a mathematical programming problem that has a quadratic objective function and linear constraints. The optimization problem that investors face is equivalent to a constrained optimization problem minimizing the portfolio variance for a given portfolio return (or maximizing the portfolio return for a given portfolio variance). An optimization model consists of three major elements: decision variables, constraints, and an objective.

In its simplest version, the model is written as follows:

$$\text{Min: } \sigma_p^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \sigma_{ij} \tag{18}$$

subject to

$$E[R_p] = \sum_{i=1}^{n} w_i E[R_i] = T \tag{19}$$

$$\sum_{i=1}^{n} w_i = 1 \tag{20}$$

$$w_i \geq 0 \qquad i = 1, \cdots, n \tag{21}$$

The objective of equation (18) is to minimize the risk of the portfolio, σ_p^2 and the decision variables are the percentage of the portfolio invested in each asset, w_i. The constraints are represented in the three equations of (19) to (21). Equation (19) represents return target T that we have to meet, and equation (20) shows 100% of budget invested. Equation (21) indicates that no short sales are allowed since investors cannot invest a negative amount of w_i. Moreover, the analysis has been simplified by the assumption that no risk-free asset exists, that is, there are no cases of riskless lending and borrowing. Varying the desired level of the expected return, T and repeatedly solving the quadratic program identifies the minimum variance portfolio for each value of T. These are the efficient portfolios that compose the efficient set. In general, the efficient frontier can be traced by plotting the corresponding values of the objective function and T, variance and return respectively.

Let us start with a simple numerical example where only three stocks are considered as candidates for constructing portfolios. To implement a mean-variance optimization, we use the data of annual returns for three randomly selected stocks (Itochu: I, Nisseki-Mitsubishi: NM, Toyota: T) in the first section of the Tokyo Stock Exchange from 1987 to 2001 (estimates of the inputs are presented in Table 1). During the 15 years, the stock of Toyota has the highest expected return, 10.467% and the lowest standard deviation, 19.148% among the three stocks. Clearly, Toyota dominates the other two stocks, with a lower risk and a higher return. At first glance, it seems that even a risk averse investor would like to invest all his money in Toyota, which means no portfolio diversification.

In a mean-variance optimization, the degree of co-movement of the returns for each stock plays an important role in minimizing portfolio risk. In the covariance matrix in Table 1, the entries off the main diagonal represent covariances between different pairs of stocks. Algebraically, the model for this problem is given as:

$$Min: \sigma_p^2 = 0.077 w_I^2 + 0.039 w_{NM}^2 + 0.037 w_T^2 + 2(0.040 w_I w_{NM} + 0.034 w_I w_T + 0.015 w_{NM} w_T)$$

subject to :

$$0.02527 w_I - 0.001 w_{NM} + 0.10467 w_T = T$$

$$w_I + w_{NM} + w_T = 1$$

$$w_I, w_{NM}, w_T \geq 0$$

Table 1 Data Set for the Illustrative Example

Period:1987-2001 <Annual>	Stock		
	Itochu	Nisseki-Mitsubishi	Toyota
ER (%)	2.527	-0.100	10.467
SD (%)	27.661	19.780	19.148
Information ratio (=ER/SD)	0.0913	-0.0051	0.5466
	Covariance Matrix		
	I	NM	T
I	0.077		
NM	0.040	0.039	
T	0.034	0.015	0.037
	Correlation Matrix		
	I	NM	T
I	1		
NM	0.739	1	
T	0.636	0.386	1

Source: Statistics are calculated on the basis of data drawn from Japan Securities Research Institute, Kabushiki toshi shuekiritsu (Stock Return Statistics).

Using matrix notation, the objective function of the model is stated as:

$$Min: \sigma_p^2 = \begin{bmatrix} w_I & w_{NM} & w_T \end{bmatrix} \begin{bmatrix} 0.077 & 0.040 & 0.034 \\ 0.040 & 0.039 & 0.015 \\ 0.034 & 0.015 & 0.037 \end{bmatrix} \begin{bmatrix} w_I \\ w_{NM} \\ w_T \end{bmatrix} \tag{22}$$

The covariance matrix in equation (22) provides concise information about key determinants of portfolio risk, statistics of variances and covariances.

Investment decision-making in the context of mean-variance analysis is essentially a problem of an optimal trade-off between risk and returns. That is, an individual investor faces two conflicting objectives simultaneously: minimizing risk and maximizing expected returns. One way of dealing with these conflicting objectives is to solve the following problem.

$$Max: (1-\lambda)E\left[R_p\right] - \lambda(\sigma_p^2) = (1-\lambda)\sum_{i=1}^{n} w_i E\left[R_i\right] - \lambda \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \sigma_{ij} \tag{23}$$

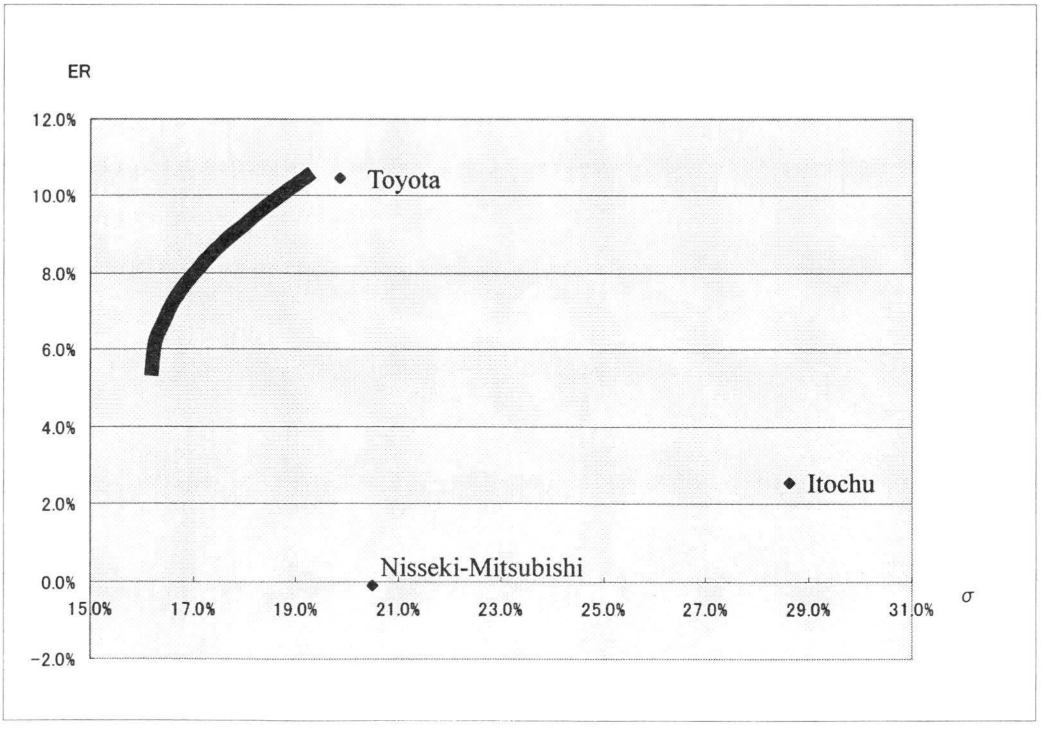
subject to

$$\sum_{i=1}^{n} w_i = 1$$

$$w_i \geq 0$$

Here, for modeling the risk-return trade-off, the above objective function involves the parameter λ, $0 \leq \lambda \leq 1$ which represents the investor's aversion to risk (λ is called the risk aversion value in Ragsdale [2001], p.375). The risk aversion value, λ lies between 0 and 1. When $\lambda = 1$, that indicates maximum risk aversion of the investors, the objective function seeks to minimize the portfolio risk. This solution exhibits the smallest possible portfolio variance, which is called the global minimum variance portfolio. Conversely, when $\lambda = 0$, that indicates a total disregard of risk, the objective function seeks to maximize the expected portfolio return. This solution exhibits the maximum return portfolio. To demonstrate the relationship between the minimum-variance portfolio with a given targeted return and the degree of the investor's aversion to risk, varying the risk aversion value λ from 0 to 1, and repeatedly solving the objective function of equation (23) identifies the minimum variance portfolio for each value of λ. Plotting the corresponding values of portfolio returns and risk respectively traces the efficient frontier (Figure 1). Therefore, the efficient frontier represents the set of the trade-off between risk and return faced by a risk-averse investor when constructing his portfolio. This type of optimization clarifies the relationship of the optimal portfolio selected by an individual investor and the degree of his aversion to risk.

Figure 1 Efficient Frontier



Source: Author's compilation.

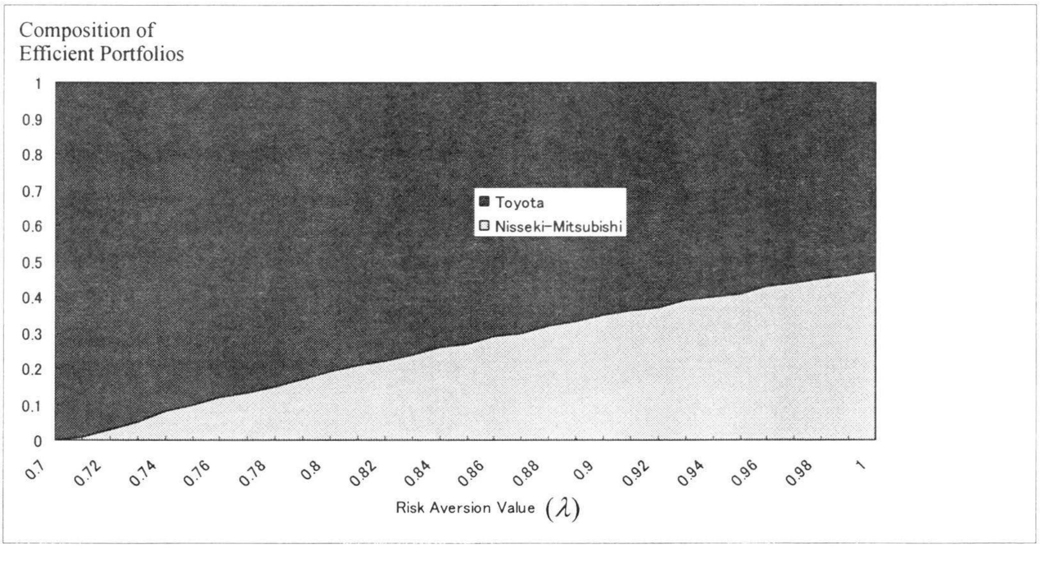
Table 2　Efficient Set of Portfolio

Risk aversion value(λ)	Portfolio return	Portfolio risk (σ)	Percentage			Total
			Itochu	Nisseki-Mitsubishi	Toyota	
0.0	10.47%	19.16%	0%	0%	100%	100%
0.1	10.47%	19.16%	0%	0%	100%	100%
0.2	10.47%	19.16%	0%	0%	100%	100%
0.3	10.47%	19.16%	0%	0%	100%	100%
0.4	10.47%	19.16%	0%	0%	100%	100%
0.5	10.47%	19.16%	0%	0%	100%	100%
0.6	10.47%	19.16%	0%	0%	100%	100%
0.7	10.47%	19.16%	0%	0%	100%	100%
0.8	8.46%	17.32%	0%	19.01%	80.99%	100%
0.9	6.79%	16.43%	0%	34.76%	65.24%	100%
1.0	5.46%	16.19%	0%	47.36%	52.64%	100%

Source: Author's calculation.

Table 2 shows the various efficient portfolios with each pair of portfolio return and risk including the ratio of each stock in the portfolio corresponding to the degree of aversion to risk (i.e. selective risk aversion values from 0 to 1). In the global minimum variance portfolio ($\lambda=1$), the solution by quadratic programming places 47.36% of the investor's money in Nisseki-Mitsubishi and 52.64% in Toyota. On the other hand, in the maximum return portfolio ($\lambda=0$), the solution places 100% of the investor's money in Toyota.

Figure 2　Risk Aversion and Portfolio Choice



Source: Author's compilation.

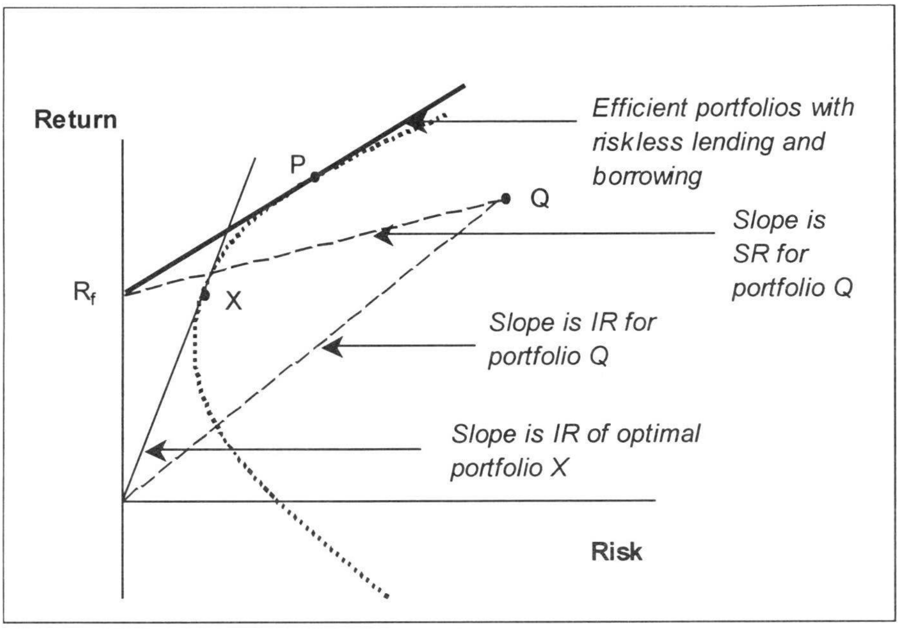
According to Table 2, the moderate risk averse investor allocates all his money to Toyota and the very risk averse investor ($\lambda > 0.7$) diversifies his money into Nisseki-Mitsubishi. The higher the risk aversion value, the larger the fraction of Nisseki-Mitsubishi in the efficient portfolios (Figure 2). Interestingly, Itochu is not included in the efficient portfolios, even though Nisseki-Mitsubishi with a negative expected return is included in the efficient portfolios selected by the very risk averse investors. This stems from the fact that the stock returns of Itochu were relatively highly correlated to those of Nisseki-Mitsubishi and Toyota as shown in the correlation matrix in Table 1. The correlation coefficient between Toyota and Nisseki-Mitsubishi is approximately 0.386, much less than those between Toyota and Itochu, 0.636 and Itochu and Nisseki-Mitsubishi, 0.739. It is note that the benefits of diversification are essentially due to the combination between assets with low (or, if possible, negative) correlation. In this point, the covariance matrix that concisely contains all necessary information about volatility and correlation statistics plays a key role in forming efficient portfolios. To construct an efficient portfolio, it is necessary to include inefficient assets because the risk of an individual asset should be of little importance, but its contribution to the portfolio's risk as a whole should be taken into account for the investor.

There are some risk-adjusted performance measures (RAPMs) which take account of both risk and return characteristics for portfolio construction. The Sharp ratio and the information ratio are two of the standard RAPMs for investment analysis. The most common RAPM is the Sharp ratio, which is computed as the excess return over the risk-free rate, R_f divided by the volatility of the asset (or portfolio). Mathematically,

$$SR = \frac{(R_q - R_f)}{\sigma_q} \tag{24}$$

where R_q and σ_q are an arbitrarily chosen portfolio return and risk. This is given graphically as the slope of the dotted line from the point of R_f on the vertical axis in Figure 3. The Sharp ratio is applied in the case of allowing unlimited riskless lending and borrowing as a risk-free rate. If risk-free returns are assumed to be zero (or, no risk-free asset exists), the appropriate RAPM is the information ratio given by the slope of the dotted line from the origin to the point Q,

Figure 3 Risk Adjusted Performance Measures



Source: Alexander [2001], p.193.

$$IR = \frac{R_q}{\sigma_q} \tag{25}$$

These RAPMs indicate the measures of reward per unit of risk. How can these ratios help us in constructing an optimal portfolio? According to the mean-variance optimization rule, investors seek to maximize the portfolio return for a given portfolio risk. In either case, for a given volatility (risk), σ_q , investors want the rate of return to be as great as possible, that is, they need to choose the solid line with the greatest possible slope. In the case of the above numerical example, the maximum information ratio is 0.5466, in which the portfolio is solely composed of the stock of Toyota (100%).

Finally, in our simple numerical example, the stock of Toyota has a huge effect on the shape of the efficient frontier and the construction of optimal portfolios, particularly for an investor who is not very risk-averse. This is because Toyota has a much higher return than that of the other two stocks. In the practical application of mean-variance optimization it is quite common that optimal portfolio will be dominated by just a few assets with high-return, high-risk characteristics[1].

---

[1] Alexander [2001] indicates the predominant effect of a few high-risk, high-return assets on the shape of the efficient frontier among 35 assets. p.199.

4. Stochastic Optimization by Monte Carlo Simulation

In the previous section, I explained the mathematical setting of mean-variance optimization problem and provided a simple numerical example for depicting the efficient frontier in the context of a trade-off between risk and return. The solution to the quadratic programming problem is to find a set of values for the decision variables, w_i that optimizes the associated objective. All data (expected returns, volatility and correlation statistics) used in the model were calculated from historical performances of the individual stock returns, that is, they are inputted as constant variables. This makes the model deterministic. However, decision making in portfolio analysis essentially involves ex ante returns (i.e., future performances) and uncertainty of these returns has to be quantified in the optimization process. In this regard, ex ante (in the original meaning of the "expected" and uncertain) returns can only be described probabilistically.

It is worthwhile to note that the method of optimization should be selected out of those reflecting the original theoretical insight in the Markowitz mean-variance analysis. Markowitz [1952] pointed out that "Our suggestion as to tentative μ_i, σ_{ij} is to use the observed μ_i, σ_{ij} for some period of the past. I believe that better methods, which take into account more information, can be found. I believe that what is needed is essentially a probabilistic reformulation of security analysis" (emphasis mine) in the last part of his seminal article.

Monte Carlo simulation is a powerful technique for analyzing models involving probabilistic assumptions. In a stochastic optimization model, the simulation assumptions capture the uncertainty of ex ante returns using probability distribution and forecasts of the objective will also have some probability distributions of possible results for the model. The central idea behind Monte Carlo simulation is based on repeated random sampling from a given probability distribution that is assumed to model inputs to characterize the distributions of model outputs. Crystal Ball, which I will employ as an optimizer in this section, is one of the popular programs of Monte Carlo simulation. The process of Monte Carlo simulation using Crystal Ball is roughly divided into three steps: formulating the spread-sheet model to solve the problem, identifying probability distributions of input variables to generate random numbers, and implementing Monte Carlo simulation to evaluate the outcome from the distribution of model output.

The spread-sheet model used in Crystal Ball is the same as the model of quadratic programming in Section 3. The historical data used in the simulation is the same three stocks as in Table 1. Probability distributions of input variables to generate random numbers are normal distributions following the normal distribution hypothesis behind the Markowitz mean-variance analysis. In a deterministic optimization, three major elements of the model were decision variables, constraints and an objective. A stochastic optimization model has additional elements: the simulation assumptions about probability distributions used to generate model data and the forecasts expressed as the frequency distributions of possible results for the model.

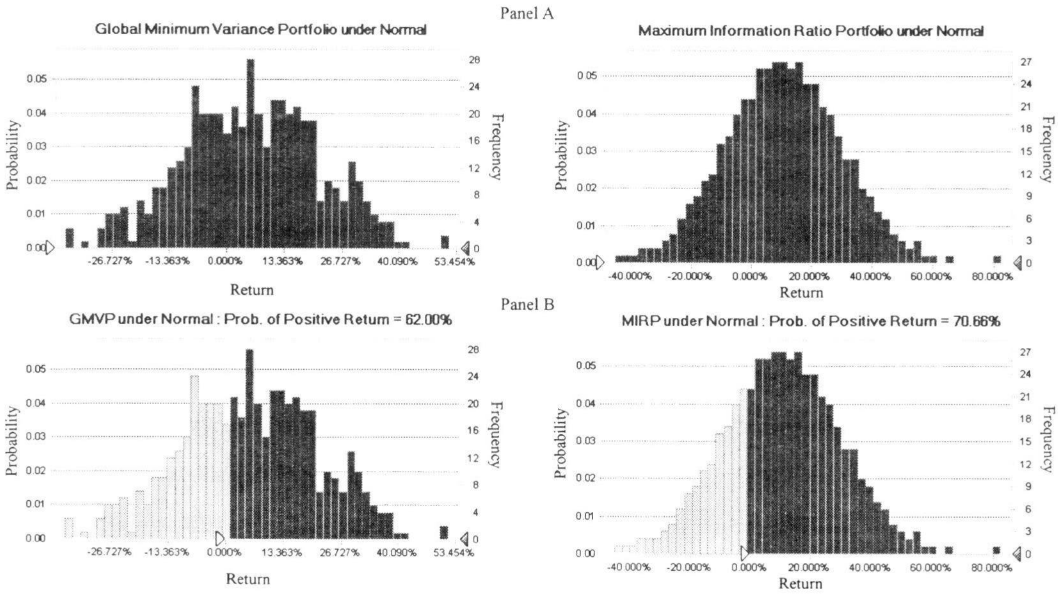
Table 3 Summary of Mean-Variance Optimization

	Deterministic Optimization <Quadratic Programming>	Stochastic Optimization <Monte Carlo Simulation>			
		Normal		Student-t	
Global minimun variance portfolio					
Return (Mean)	5.463%	5.448%		5.450%	
Risk (SD)	16.186%	16.194%		16.194%	
Composition					
Itochu	0.00%	0.00%		0.00%	
Nisseki-Mitsubishi	47.36%	47.49%		47.47%	
Toyota	52.64%	52.51%		52.53%	
Probability of a positive return	*	62.00%		61.75%	
		Min	Max	Min	Max
90% Certainty Range	*	-20.892%	31.247%	-23.271%	33.391%
95% Certainty Range	*	-26.878%	34.694%	-30.273%	38.102%
99% Certainty Range	*	-36.100%	40.870%	-41.445%	46.254%
100% Certainty Range	*	-37.376%	51.700%	-44.657%	57.890%
Downside 10% Range	*	-14.533%		-15.830%	
Downside 5% Range	*	-20.892%		-23.271%	
Downside 1% Range	*	-29.786%		-33.659%	
Maximum information ratio portfolio					
Information ratio	0.5466	0.5466		0.5466	
Return (Mean)	10.467%	10.467%		10.467%	
Risk (SD)	19.148%	19.148%		19.148%	
Composition					
Itochu	0.00%	0.00%		0.00%	
Nisseki-Mitsubishi	0.00%	0.00%		0.00%	
Toyota	100.00%	100.00%		100.00%	
Probability of a positive return	*	70.66%		70.21%	
		Min	Max	Min	Max
90% Certainty Range	*	-21.394%	41.651%	-23.690%	43.825%
95% Certainty Range	*	-27.264%	47.675%	-30.857%	51.133%
99% Certainty Range	*	-39.990%	57.788%	-48.210%	64.599%
100% Certainty Range	*	-45.318%	82.160%	-56.441%	106.960%
Downside 10% Range	*	-14.133%		-15.355%	
Downside 5% Range	*	-21.394%		-23.690%	
Downside 1% Range	*	-35.260%		-41.415%	

Source: Author's calculation.

The results of the experiments are summarized in Table 3, and the outcomes of the simulation are provided as frequency charts of model output (Figure 4). In contrast, deterministic optimization for the global minimum variance portfolio (GMVP) provides only one value of portfolio return 5.463% and risk 16.186%, since all data are inputted as constant variables. The solution to the global minimum variance portfolio under stochastic optimization has a portfolio risk of 16.194% (standard deviation) and a portfolio return of 5.448%, in which 0% of the investor's money is allocated to Itochu, 47.49% to Nisseki-Mitsubishi and 52.51% to Toyota. By manipulating the end-point grabbers or by changing the range and certainty values in the boxes in the frequency chart (Panel B of Figure 4), we can specify a certainty level or a probability interval of realizing the global minimum variance portfolio return. For example, the probability of a positive return is 62.00%. In case of the result of changing the certainty level to 100%, the range centered about the mean is from -37.376% to 51.7% compared to that of the maximum information ratio portfolio from -45.318% to 82.160% (Table 3). It is clear that the global minimum variance portfolio has a smaller range of expected return compared to the maximum information ratio portfolio (MIRP).

Figure 4 Forecast of Monte Carlo Simulation



To sum up, stochastic optimization provides us with forecasts of our objectives probabilistically. In addition, we can get some insight about alternative measures of financial risk, such as value at risk (VaR) and the expected tail loss (ETL) from our probabilistic forecasts of Monte Carlo simulation (for a more detailed explanation about VaR and ETL, see Dowd [2002]). In particular, it might be possible to provide a much better approach to allow the return distribution with non-normality to be less restricted.

5. Stylized Facts about Return Distributions

The result of Monte Carlo simulation depends crucially on the probability distributions that will be assumed to generate the data of random sampling. If some unrealistic assumptions have been made in the data generating process, the simulation experiments will not give a precise answer to the problem. In the context of mean-variance optimization, the optimal asset allocation obtained from a simulation will not be accurate if the data generating process assumed normal distribution while the actual returns series is not normally distributed.

The normal distribution hypothesis behind mean-variance analysis was based on the fact that asset returns are influenced by many different independent facts[2]. However, since the early 1960's empirical research on returns distributions has almost universally found that such distributions are characterized by the features of the fat tails and high peakedness -excess kurtosis- and are often skewed. Those features are known as stylized facts about financial return series, especially high frequency data.

Table 4 represents descriptive statistics of historical performances of annual and monthly return series (Itochu, Nisseki-Mitsubishi, Toyota and market) for a fifty year period from 1955 to 2004 and a fifteen year period from 1987 to 2001. One of the features which stands out most prominently from the last columns is that the kurtosis of the four series is much higher than the normal value, 3. This reflects the fact that the tails of the distributions of these series are fatter than the tails of the normal distribution. Put differently, large outlying (i.e. very small and very large) observations occur with rather high-frequency.

Next, three individual stock return series have positive skewness. The skewness of the normal distribution is zero because its distribution is symmetric. Positive skewness implies that the right tail of the distribution is fatter than the left tail. That is, large positive returns tend to occur more often than large negative ones. Positive skewness has an important impact on portfolio choice because it means that these stocks have a larger probability of very large payoffs,

---

[2] According to probability theory and statistics, any phenomenon made up of a large number of independent or weakly dependent variables has a normal distribution. See Focardi & Fabozz [2004], p.194.

Table 4 Descriptive Statistics of Stock Return Series

〔%〕

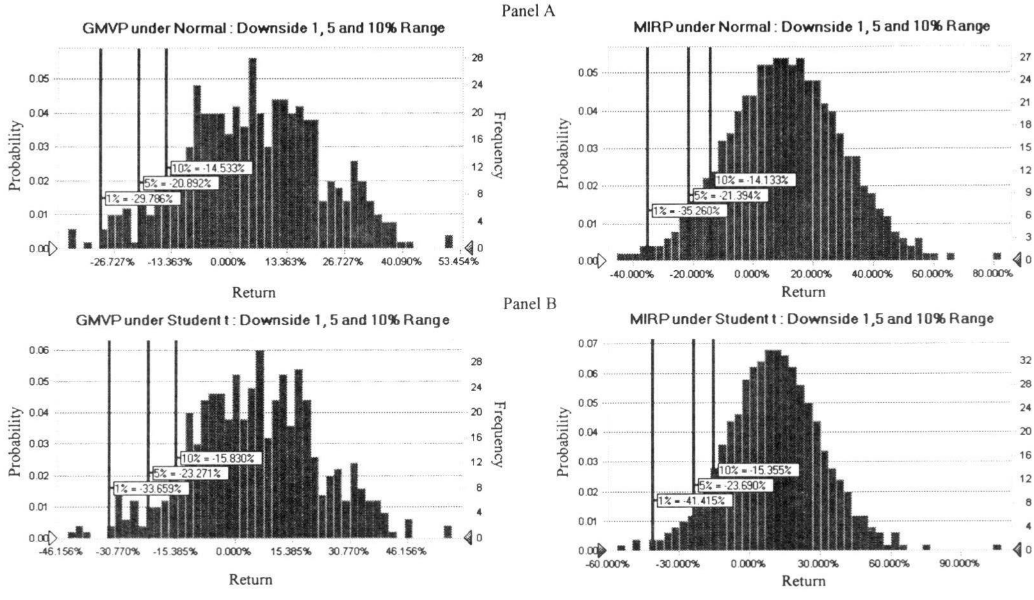
	Mean	Median	Maximum	Minimum	Std. Dev.	Skewness	Kurtosis
1955-2004: Annual							
Itochu	17.628	10.700	240.100	-43.500	46.093	2.527	12.364
Nisseki-Mitsubishi	15.994	7.300	144.800	-30.500	34.982	1.374	5.478
Toyota	23.254	19.200	110.600	-37.900	33.087	0.617	2.942
Market	14.776	15.950	72.100	-24.800	19.956	0.356	3.271
1955-2004: Monthly							
Itochu	1.433	0.000	72.700	-38.100	11.017	1.412	9.220
Nisseki-Mitsubishi	1.368	0.000	52.000	-28.900	10.214	1.026	6.330
Toyota	1.891	1.200	46.800	-25.000	9.287	0.772	5.040
Market	0.909	0.800	17.500	-19.800	5.058	-0.130	3.934
1987-2001: Monthly							
Itochu	0.381	-0.800	60.000	-38.100	12.982	1.049	6.968
Nisseki-Mitsubishi	0.043	-0.950	36.800	-28.900	9.657	0.299	4.617
Toyota	0.778	0.500	43.100	-20.000	8.128	1.002	7.028
Market	0.022	-0.400	17.500	-19.800	6.037	0.092	3.534

Source: Calculated on the basis of data drawn from Japan Securities Research Institute, Kabushiki toshi shuekiritsu (Stock Return Statistics).

thus, they should have a preference for positive skewness. The monthly market return series for the period of 1955 to 2004 has only slightly negative skewness. It implies that large negative returns tend to occur more often than large positive ones. It is clear that the distributions of all stock return series listed in Table 4 diverge considerably from the normal distribution, have fatter tails, are more highly peaked, and are often skewed. The use of a normal distribution assumed for the data generation process in the simulation is likely to lead to a systematic underestimate of the occurrences of both sides of extreme values of returns. This is because they are more likely in practice than would arise under a normal distribution.

One approach to remedies for stylized facts, especially for the tail's parts of return distributions, is the replacement of the distribution assumed in the simulation from a normal distribution by a Student t-distribution. A Student t-distribution is a symmetric and bell-shaped distribution similar to a normal distribution, but with fatter tails and a smaller peak at the mean. We have summarized the results of Monte Carlo simulation under the different assumed distributions in Table 3 and Figure 5. The information demonstrates that the use of a normal distribution under the simulations leads to a systematic underestimate of the tail's parts of forecasts about the global minimum variance portfolio(GMVP) and the maximum information ratio portfolio(MIRP), as extremely large positive and negative returns are more likely in practice than would arise under a normal distribution.

Figure 5 Forecast of Tail's Part



An alternative approach to overcome stylized facts about returns distribution would be to use bootstrapping. In Monte Carlo simulation, the data are generated completely artificially from the assumed distribution. On the other hand, bootstrapping does not assume some preset distribution but uses the actual data themselves. However, Brooks [2002] points out that there are at least two situations where the bootstrap will not work well. First, if there are outliers in the data, the conclusions of the bootstrap may be affected. Second, use of the bootstrap implicitly assumes that the data are independent of one another. If there were autocorrelation in the data, this would obviously not hold.

6. Concluding Remarks and Future Direction of Research

In this article, we have considered the implementation of the mean-variance optimization in portfolio analysis. The mean-variance analysis suggested by Markowitz is theoretically based on the expected utility theory and the normal distribution hypothesis about return distributions. The worth of the analysis rests on revealing normative rules for optimal portfolio choice by an individual. The theory is relatively straightforward, however its implementation can get quite complicated. Recently, remarkable progress has occurred in the area of risk management. Breakthroughs in its implementation fall into two categories: an optimizer and a return forecast. In order to get robust results, these issues are crucial and closely related to each other.

I began with formulating the problem of mean-variance optimization in a formal mathematical context and then demonstrated that a covariance matrix is a cornerstone of risk reduction through portfolio diversification. Next, we implemented mean-variance optimization by using the methods of Monte Carlo simulation, which is consistent with the original idea of Markowitz's approach in an aspect of probabilistic setting about return forecast. The result of simulation relies mainly upon the assumed probability distribution to generate return data in the model. Its robustness depends on whether the actual return distribution is fitting to the assumed distribution under simulation. According to the historical performance of return series, the return distributions are not normally distributed. To rectify the results of simulation to tail's parts of return distributions, I added to implement the simulation under the assumption of Student t-distribution and pointed out the possibility of using the method of bootstrapping. These methods are directed to improve the assumed distribution under simulation in an aspect of curve fitting to actual return distributions. In other words, I attempt to fit an assumed distribution to historical data unconditionally.

More fundamentally, practical problems with mean-variance optimization lie on the use of return data. Three key ingredients in portfolio optimization are a return forecast, an optimizer and a utility function. Although those are inseparable from each other, an accurate return forecast is of paramount importance on data input in the optimization process. The covariance matrix provides concise and precise information about volatilities and correlations as discussed in Section 2. Volatility and correlation are parameters of the stochastic process that are used to model the variation in asset prices. The estimation and forecasting is at the heart of mean-variance optimization. In practice, they are not directly observable, unlike asset prices. They can only be estimated in the context of a statistical model, and those estimates depend on the choice of model applied to historical return series. Campbell, Lo and MacKinlay [1997] state that nonlinear characteristics in economic behavior might be found in financial markets, for instance, investor's attitudes toward risk and expected return, strategic interactions among market participants. The information stemming from nonlinear characteristics is incorporated into asset prices, the dynamics of which are embodied in the stylized facts after all. The stylized facts implying leptokurtosis and volatility clustering lead us to consider the introduction of nonlinear models such as GARCH models and so on to describe the observed patterns in stock return series.

References

Alexander, C. [2001], *Market Models: A guide to Financial Data Analysis*, John Wiley, Chichester.

Amenc, N. and Sourd, V. L. [2003], *Portfolio Theory and Performance Analysis*, John Wiley, Chichester.

Brooks, C. [2002], *Introductory Econometrics for Finance*, Cambridge University Press, Cambridge.

Camm, J. D. and Evans, J. R. [2000], *Management Science and Decision Technology*, South-Western College Publishing.

Campbell, J. Y., Lo, A. W. and MacKinlay, A. C. [1997], *The Econometrics of Financial Markets*, Princeton University Press, New Jersey.

Dowd, K. [2002], *Measuring Market Risk*, John Wiley, Chichester.

Elton, E. J. and Gruber, M. J. [1995], *Modern Portfolio Theory and Investment Analysis*, 5th ed., John Wiley, Chichester.

Fabozzi, F. J., Gupta, F. and Markowitz, H. M. [2002], "The Legacy of Modern Portfolio Theory," *Journal of Investing*, Fall, pp.7-22.

Facardi, S. M. and Fabozzi, F. J. [2004], *The Mathematics of Financial Modeling and Invetment management*, John Wiley, Chichester.

Franses, P. H. and Dijk, D. V. [2000], *Non-linear Time Series Models in Empirical Finance*, Cambridge University Press, Cambridge.

Glantz, M. [2000], *Scientific Financial Management*, AMACOM, New York.

Greene, W. H. [2000], *Econometric Analysis*, 4th ed., Prentice-Hall, Inc, New Jersey.

Jocobsen, B. and Dannenburg, D. [2003], "Volatility Clustering in Monthly Stock Returns," *Journal of Empirical Finance*, Vol.10, pp.479-503.

Jobst, N. J., Horniman, M.D., Lucas, C.A. and Mitra, G., [2001], "Computational Aspects of Alternative Portfolio Selection Models in the Presence of Discrete Asset Choice Constraints," *Quantitative Finance*, Vol.1, No.5, pp.489-501.

Kallberg, J.G. and Ziemba, W. T. [1983], "Comparision of Alternative Utility Functions in Portfolio Selection Problem," *Management Science*, Vol.29, pp.1257-1276.

Laws, J. [2003], "Portfolio Analysis Using Excel," in *Applied Quantitative Methods for Trading and Investment*, C. Dunis, J. Laws and P. Naïm (eds), John Wiley, Chichester, pp. 293-311.

Ledoit, O. and Wolf, M. [2003], "Improved Estimation of the Covariance Matrix of Stock Returns with an Application to Portfolio Selection," *Journal of Empirical Finance*, Vol.10, Issue 5, pp.603-621.

Lhabitant, F-S. [2002], *Hedge Funds*, John Wiley, Chichester.

Markowitz, H. [1952], "Portfolio Selection," *Journal of Finance*, Vol.7, pp.77-91.

Munechika, M. [2002], "Foundations of the Mean-Variance Analysis: Normal Distribution Hypothesis and Expected Utility," *Economic Review of Toyo University*, Vol.28, No.1, pp.51-87.

Mills, T. C. [1999], *The Econometric Modelling of Financial Time Series*, 2nd. ed., Cambridge University Press, Cambridge.

Mun, J. [2004], *Applied Risk Analysis*, John Wiley, Chichester.

Ragsdale, C. T. [2001], *Spreadsheet Modeling and Decision Analysis*, 3rd. ed., South-Western College Publishing, New York.

Ross, S. A., Westerfield, R. W. and Jaffe, J. [1996], *Corporate Finance*, 4th ed., McGraw-Hill Companies, Inc. Chicago.

Ruppert, D. [2004], *Statistics and Finance: An Introduction*, Springer-Verlag, New York.

Sharpe, W. F. [1963], "A Simplified Model for Portfolio Analysis," *Management Science*, Vol.9, pp.277-293.

Stevens, G. V. [1998], "On the Inverse Covariance Matrix in Porfolio Analysis," *Journal of Finance*, Vol.LIII, No.5, October, pp.1821-1827.

Teall, J. L. and Hasan, I. [2002], *Quantitative Methods for Finance and Investments*, Blackwell, Oxford.

Teall, H. [1983], "Linear Algebra and Matrix Methods in Econometrics," in *Handbook of Econometrics*, Vol. 1, Z. Griliches and M. D. Intriligator (eds.), North-Holland, Amsterdam, pp.5-65.