Spreadsheets in Education (eJSiE)

Volume 4 | Issue 1 Article 4

7-6-2010

The Requirement of a Positive Definite Covariance Matrix of Security Returns for Mean-Variance Portfolio Analysis: A Pedagogic Illustration

Clarence C. Y. Kwan
McMaster University, kwanc@mcmaster.ca

The Requirement of a Positive Definite Covariance Matrix of Security Returns for Mean-Variance Portfolio Analysis: A Pedagogic Illustration

Abstract
This study considers, from a pedagogic perspective, a crucial requirement for the covariance matrix of security returns in mean-variance portfolio analysis. Although the requirement that the covariance matrix be positive definite is fundamental in modern finance, it has not received any attention in standard investment textbooks. Being unaware of the requirement could cause confusion for students over some strange portfolio results that are based on seemingly reasonable input parameters. This study considers the requirement both informally and analytically. Electronic spreadsheet tools for constrained optimization and basic matrix operations are utilized to illustrate the various concepts involved.

Keywords
Mean-Variance Portfolio Analysis, Positive Definite Covariance Matrix, Excel Illustration

The Requirement of a Positive Definite Covariance Matrix of Security Returns for Mean-Variance Portfolio Analysis: A Pedagogic Illustration

Clarence C.Y. Kwan
DeGroote School of Business
McMaster University
Hamilton, Ontario L8S 4M4
Canada
E-mail: kwanc@mcmaster.ca

July 2010

Abstract

This study considers, from a pedagogic perspective, a crucial requirement for the covariance matrix of security returns in mean-variance portfolio analysis. Although the requirement that the covariance matrix be positive definite is fundamental in modern finance, it has not received any attention in standard investment textbooks. Being unaware of the requirement could cause confusion for students over some strange portfolio results that are based on seemingly reasonable input parameters. This study considers the requirement both informally and analytically. Electronic spreadsheet tools for constrained optimization and basic matrix operations are used to illustrate the various concepts involved.

Submitted May 2010, revised and accepted July 2010.

Key words: mean-variance portfolio analysis, positive definite covariance matrix, Excel spreadsheet illustration

1 Introduction

Mean-variance portfolio theory is an important part of the core curriculum of modern finance in business education. The theory, which provides the foundation of investment decisions, captures the risk of an investment with the variance of the probability distribution of the investment's random rates of returns. A practical aspect of the theory is that, with the expected returns (the means), the variances of returns, and the covariances of returns of individual financial securities being input parameters for portfolio models, it provides guidance for allocating investment funds among the securities considered to achieve the best risk-return trade-off.

This study considers, from a pedagogic perspective, a crucial analytical requirement on input parameters for mean-variance portfolio analysis. Being traditionally considered to be beyond the scope of the standard finance curriculum, the requirement is seldom brought to the attention of students, even in advanced portfolio investment courses. As a result, there is an undesirable gap between what is viewed as fundamental in the academic finance literature and what students learn in investment courses. Being unaware of the requirement could lead to confusion over some strange portfolio results that are based on seemingly reasonable input parameters.

Specifically, the requirement is that the covariance matrix of security returns, which contains all variances and covariances of returns of the securities considered, be positive definite.[1] In the context of portfolio investments under the assumption of frictionless short sales, the covariance matrix is positive semidefinite if the variance of portfolio returns is always non-negative, regardless of how investment funds are allocated among the securities considered.[2] If the portfolio variance is always strictly positive, the covariance matrix is also positive definite. A positive definite covariance matrix is invertible; however, a covariance matrix that is positive semidefinite but not positive definite is not invertible.

At first glance, as the variance of a random variable, by definition, cannot be negative, the attainment of a positive definite covariance matrix seems to be assured if individual securities or their combinations that can lead to risk-free investments are excluded from portfolio consideration. As shown in Appendix A, the sample covariance matrix — the covariance matrix estimated from a sample of past return observations — is always positive semidefinite. Once the various situations causing the sample covariance matrix to be non-invertible are ruled out, the positive definiteness requirement will be satisfied. Then, why is the requirement still a relevant issue to consider? That is because the variance of a linear combination of some random variables, as expressed *directly* in terms of the variances and covariances of these variables, can be negative if not all individual variances and covariances correspond to their sample estimates. Here are some specific examples:

In classroom settings, input parameters for numerical illustrations of portfolio analysis are often generated artificially. Although the covariance matrix for more than two securities thus generated is invertible, with the implied correlations of returns between different securities being always in the permissible range of -1 to 1, whether it is positive definite is not immediately obvious. In practical settings, security analysts' insights are often required to revise the sample estimates of input parameters for portfolio analysis in order to recognize changes in the economic environment. Likewise, for a high-dimensional portfolio selection problem, if the covariance matrix is estimated with insufficient ob-

[1] The requirement is explicitly stated in Merton (1972), Roll (1977), and Jobson and Korkie (1989), among others.

[2] Under the assumption of frictionless short sales, if an investor short sells a security, the investor not only provides no cash deposit, but also has immediate access to the short-sale proceeds for investing in other securities. The individual portfolio weights — the proportions of investment funds as allocated to the individual securities considered — can be of either sign, as long as the sum of all portfolio weights is unity.

servations (for concerns about outdated return data from a long sample period), some matrix elements will have to be revised to make the resulting matrix invertible.[3] In either case, it is not immediately obvious whether the resulting matrix is positive definite. Further, if estimation of the covariance matrix is based on models that can accommodate time-varying volatility of security returns, the positive definiteness of the resulting covariance matrix may not be assured.

In this study, we consider the positive definiteness requirement first informally and then analytically, in order to accommodate different pedagogic approaches — which correspond to different levels of analytical rigor — in the delivery of the mean-variance portfolio concepts to the classroom.[4] In an informal approach, we illustrate with a three-security case that having all correlations of security returns in the permissible range of -1 to 1 alone does not ensure the validity of the covariance matrix. Specifically, by using Microsoft Excel^{TM}, we compute the covariances and correlations of returns of various portfolios. Further, we use *Solver*, a numerical tool in Excel, to search for the allocation of investment funds (under the assumption of frictionless short sales) that corresponds to the lowest variance of portfolio returns. If it turns out that any correlations of portfolio returns are outside the permissible range or the lowest variance is negative, then the covariance matrix in question cannot be acceptable.

In our analytical approach to consider the positive definiteness requirement, the algebraic and statistical tools involved are confined to those known to most business students. By using a basic portfolio selection model, in which frictionless short sales are assumed, we can directly reach the portfolio solution in terms of the input parameters provided. We can also address the issue of positive definiteness without being encumbered by the algorithmic details that are often associated with solution methods for more sophisticated portfolio selection models. For ease of exposition, matrix notation is used. The required matrix operations, which are basic, can easily be performed on spreadsheets.

[3] There are various empirical approaches to ensure the invertibility of a covariance matrix, even if there are insufficient observations for its estimation. A simple approach is to impose a particular covariance structure. In the case of the constant correlation model, for example, all correlations of security returns are characterized to be the same. In such a case, the estimated covariances are based on the individual sample variances and a common correlation. The use of the average of all sample correlations for the securities considered as the common correlation will ensure that the resulting covariance matrix be positive definite. [See Kwan (2006) for some analytical properties of the constant correlation model.] A more sophisticated approach to ensure positive definiteness is shrinkage estimation, which takes a weighted average of the sample covariance matrix and a structured matrix. [See Ledoit and Wolf (2004) and Kwan (2008) for analytical details of the shrinkage approach in which the structured matrix is based on the constant correlation model.]

[4] In advanced investment courses, it is also useful to justify analytically the mean-variance approach, to discuss the limitations of the approach, and to provide alternative risk measures. [See, for example, Cheung, Kwan, and Miu (2007) for an alternative risk measure in view of the limitations of mean-variance.] Even if the conditions to justify mean-variance are deemed acceptable for the securities considered, having a positive definite covariance matrix of security returns as part of the input parameters for portfolio analysis does not automatically ensure good quality of the portfolio selection results. As the expected returns and the covariance matrix have to be estimated, estimation errors are inevitable. Such errors, if large, would weaken the results from portfolio analysis. [See, for example, Kwan (2009) for a pedagogic illustration of estimation errors in sample variances, covariances, and correlations.]

They include matrix transposition, addition, multiplication, and inversion, as well as finding the determinant. As this study is intended to be self-contained for pedagogic purposes, the analytical materials involved are derived. Some proofs are provided in footnotes or appendices, in order to avoid digressions.

The rest of this study is organized as follows: Section 2 provides a two-security exposition of basic portfolio concepts. An extension to a three-security case is presented in Section 3. An Excel example there illustrates the need for addressing the positive definiteness issue. Section 4 presents a basic portfolio selection model. As the covariance matrix must be invertible for the model to provide any portfolio allocation results, situations leading to its non-invertibility are first identified in Section 5. The central theme of Section 5 is that, for the portfolio allocation results to be meaningful, the covariance matrix must be positive definite. It implications are also considered there. With the aid of an Excel example, Section 6 illustrates some consequences for violating the positive definiteness requirement. Finally, Section 7 provides some concluding remarks.

2 Portfolio concepts based on two securities

In introductory finance courses, the delivery of mean-variance portfolio concepts typically starts with a two-security case, where the security with a higher expected return also has a higher variance of returns. With the random returns of the two securities being R_1 and R_2 and their expected returns being μ_1 and μ_2, each of the two variances of returns, labeled as σ_1^2 and σ_2^2, is the expected value of the squared deviation of the corresponding random return from its mean. The covariance of returns of R_1 and R_2, labeled as σ_{12}, is the expected value of $(R_1-\mu_1)(R_2-\mu_2)$. The correlation of returns ρ_{12}, defined as $\sigma_{12}/(\sigma_1\sigma_2)$, is always in the range of -1 to 1. It is implicit that $\sigma_{11}=\sigma_1^2$, $\sigma_{22}=\sigma_2^2$, $\sigma_{12}=\sigma_{21}$, $\rho_{12}=\rho_{21}$, $\rho_{11}=\sigma_{11}/(\sigma_1\sigma_1)=1$, and $\rho_{22}=\sigma_{22}/(\sigma_2\sigma_2)=1$.

Suppose that a portfolio p is formed, with the proportions of investment funds as allocated to the two securities — the portfolio weights — being x_1 and $x_2=1-x_1$. Under the assumption of frictionless short sales, each portfolio weight can be of either sign. The variance of portfolio returns, labeled as σ_p^2, is the expected value of $(R_p-\mu_p)^2$, where

$$R_p = x_1R_1 + x_2R_2 \tag{1}$$

$$\text{and } \mu_p = x_1\mu_1 + x_2\mu_2. \tag{2}$$

It follows from

$$\begin{aligned}(R_p-\mu_p)^2 &= [x_1(R_1-\mu_1)+x_2(R_2-\mu_2)]^2\\ &= x_1^2(R_1-\mu_1)^2 + x_2^2(R_2-\mu_2)^2 + 2x_1x_2(R_1-\mu_1)(R_2-\mu_2)\end{aligned} \tag{3}$$

that

$$\sigma_p^2 = x_1^2\sigma_1^2 + x_2^2\sigma_2^2 + 2x_1x_2\sigma_{12} = x_1^2\sigma_1^2 + x_2^2\sigma_2^2 + 2x_1x_2\rho_{12}\sigma_1\sigma_2. \tag{4}$$

Equations (2) and (4), when combined to eliminate the portfolio weights, will allow σ_p to be determined directly for any given μ_p. If the returns of the two securities considered are perfectly correlated (i.e., $\rho_{12} = \pm 1$), the relationship between σ_p and μ_p is linear or piecewise linear; otherwise, it is nonlinear. Figure 1 shows some graphs on the (σ, μ)-plane — where standard deviation of returns σ and expected return μ are the horizontal and vertical axes, respectively — that are commonly used in introductory finance courses for describing the risk-return trade-off from investing in two risky securities without short sales.

These graphs reveal some basic portfolio concepts. Specifically, if $\rho_{12} = 1$, as the risk-return trade-off in portfolio investments is captured by the line joining points (σ_1, μ_1) and (σ_2, μ_2), there is no diversification effect. If $\rho_{12} < 1$ instead, portfolio investments in the two securities will lead to risk reductions for any expected return requirements (as compared to the case where $\rho_{12} = 1$). The lower the correlation, the greater is the risk-reduction effect. If $\rho_{12} = -1$, a risk-free portfolio can be reached. Notice that, under the assumption of frictionless short sales, the line joining points (σ_1, μ_1) and (σ_2, μ_2) for the case where $\rho_{12} = 1$ can be extended to reach its μ-intercept; that is, if $\rho_{12} = 1$, a risk-free portfolio can also be reached.

If $-1 < \rho_{12} < 1$, no investments in the two securities can completely eliminate the portfolio risk. To see this, let us rewrite equation (4) as

$$\sigma_p^2 = (x_1\sigma_1 + x_2\rho_{12}\sigma_2)^2 + x_2^2(1 - \rho_{12}^2)\sigma_2^2 \tag{5}$$

by completing the square. With $1 - \rho_{12}^2$ being positive, there are no portfolio weights of either sign that can result in σ_p^2 being zero or negative.

In linear algebra, it is known as Sylvester's Criterion that a real symmetric matrix is positive definite if and only if all of its leading principal minors are positive. For an $n \times n$ matrix, there are n leading principal minors, each of which is the determinant of the submatrix containing the first k rows and the first k columns, for $k = 1, 2, \ldots, n$. A proof of this matrix property is provided in Appendix B. The 2×2 covariance matrix $\begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix}$ under the condition of $-1 < \rho_{12} < 1$ is positive definite because both of its leading principal minors, σ_{11} $(= \sigma_1^2)$ and $\sigma_{11}\sigma_{22} - \sigma_{21}\sigma_{12}$ $[= \sigma_1^2\sigma_2^2(1 - \rho_{12}^2)]$, are positive. Thus, in a two-security case, if the returns of the two securities are not perfectly correlated, the positive definiteness requirement is automatically satisfied.

3 A three-security illustration of the positive definiteness requirement

Depending on the finance courses involved, extensions to portfolio investments in more than two risky securities can differ considerably. Regardless of the pedagogic approach that is followed, the efficient frontier on the (σ, μ)-plane, which captures the best risk-return trade-off from portfolio investments in the set of risky securities considered, is described as a concave curve. In essence, the differences among the various pedagogic approaches are in whether the analytical details are covered and, if so, in how sophisticated the required analytical tools are.

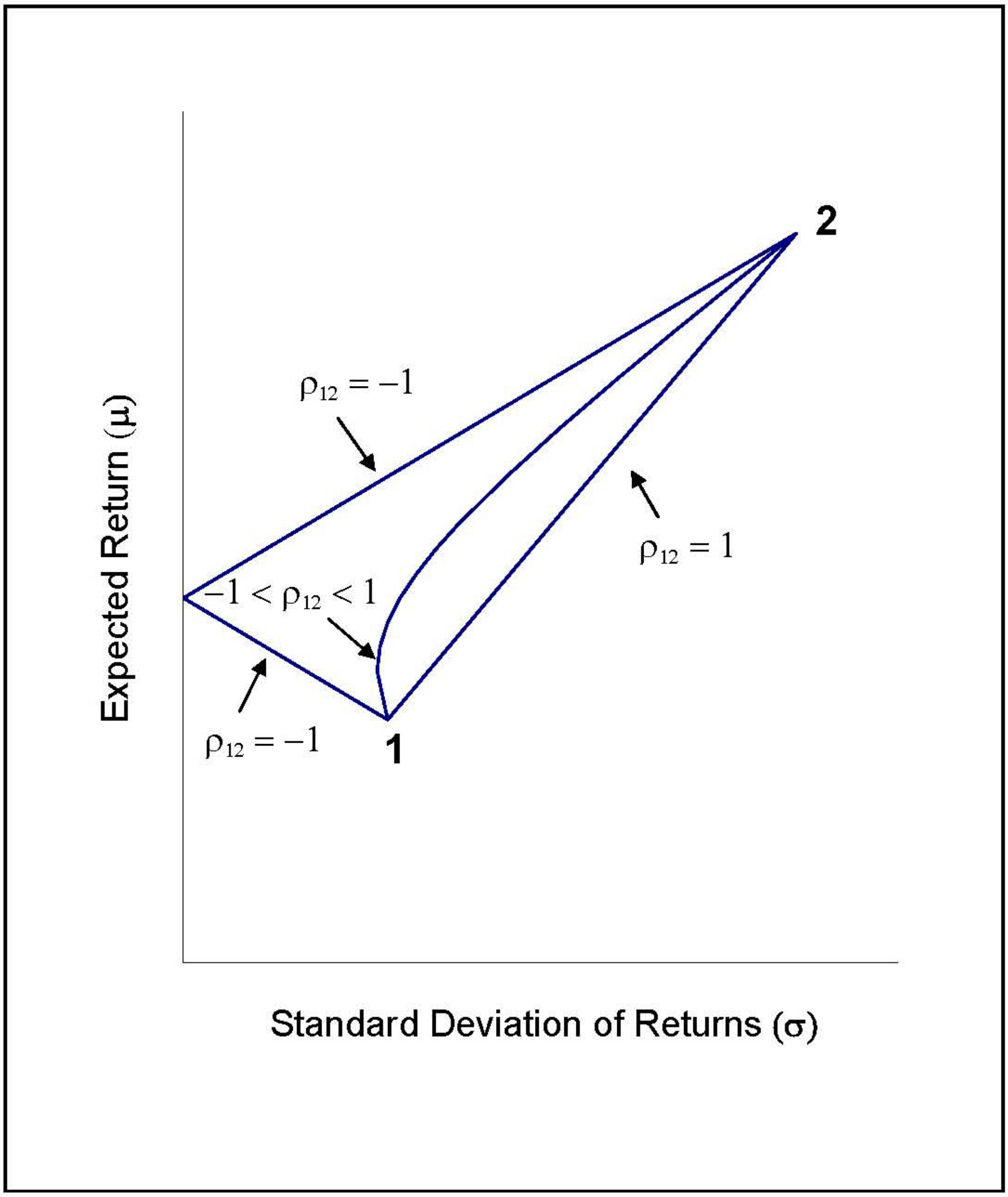

Figure 1: A two-security illustration of risk-return trade-off for different correlations of security returns.

A common message from the various pedagogic approaches is that, as long as the risky securities considered are less than perfectly, positively correlated, there will be portfolio diversification effects. To ensure the attainment of a meaningful efficient frontier, however, a relevant question now is whether there is any other requirement for the covariance matrix beyond having all correlations of security returns in the permissible range of -1 to 1. As investment textbooks are silent on the issue, we illustrate below with a simple three-security example that a further requirement is warranted.

To facilitate the illustration, let R_1, R_2, R_3, μ_1, μ_2, and μ_3 be the random and expected returns of the three-risky securities considered. The three variances ($\sigma_1^2 = \sigma_{11}$, $\sigma_2^2 = \sigma_{22}$, and $\sigma_3^2 = \sigma_{33}$) and the six covariances ($\sigma_{12} = \sigma_{21}$, $\sigma_{13} = \sigma_{31}$, and $\sigma_{23} = \sigma_{32}$) of security returns can be arranged as elements of a 3×3 covariance matrix. With each element (i, j) being σ_{ij}, which is the same as σ_{ji}, for $i, j = 1, 2$, and 3, the covariance matrix is symmetric. A 3×3 correlation matrix can be inferred directly from the covariance matrix.

We now allocate investment funds in two different ways, with the corresponding portfolios labeled as p and q. For portfolio p, let x_1, x_2, and x_3 be the portfolio weights satisfying the condition of $x_1 + x_2 + x_3 = 1$. The random return and the expected return of the portfolio are $R_p = x_1R_1 + x_2R_2 + x_3R_3$ and $\mu_p = x_1\mu_1 + x_2\mu_2 + x_3\mu_3$, respectively. The variance of returns of the portfolio, σ_p^2, which is the expected value of $(R_p - \mu_p)^2 = [x_1(R_1 - \mu_1) + x_2(R_2 - \mu_2) + x_3(R_3 - \mu_3)]^2$, can be expressed as the sum of nine terms of the form $x_ix_j\sigma_{ij}$, for $i, j = 1, 2$, and 3; that is, $\sigma_p^2 = \sum_{i=1}^3 \sum_{j=1}^3 x_ix_j\sigma_{ij}$.

For portfolio q, which is based on a different set of portfolio weights, y_1, y_2, and y_3, satisfying the condition of $y_1 + y_2 + y_3 = 1$, the expected return, μ_q, and the variance of returns, σ_q^2, can be computed in an analogous manner. The covariance of returns of portfolios p and q, labeled as σ_{pq}, is the expected value of $(R_p - \mu_p)(R_q - \mu_q)$. When expressed in terms of the variances and covariances of the individual securities, σ_{pq} is the sum of nine terms of the form $x_iy_j\sigma_{ij}$, for $i, j = 1, 2$, and 3; that is, $\sigma_{pq} = \sum_{i=1}^3 \sum_{j=1}^3 x_iy_j\sigma_{ij}$.

3.1 An Excel example

Figure 2 shows an Excel worksheet for a three-security case, where all returns are measured in percentage terms. The variances and covariances of returns provided for the example consist of $\sigma_1^2 = 100$, $\sigma_2^2 = 36$, $\sigma_3^2 = 16$, $\sigma_{12} = 33$, $\sigma_{13} = -20$, and $\sigma_{23} = 12$. Due to symmetry, σ_{21}, σ_{31}, and σ_{32} can be inferred directly. The 3×3 covariance matrix of returns is shown in cells `B5:D7` of the worksheet. The implied correlations of returns, as shown in cells `B9:D11`, are all in the permissible range of -1 to 1. Notice that cell formulas are listed at the bottom of Figure 2. Likewise, whenever an Excel worksheet is displayed in any of the subsequent figures, cell formulas are listed as well.

To see whether having all pairwise correlations in the permissible range alone is sufficient to ensure the validity of the covariance matrix, we have attempted different allocations of investment funds for portfolios p and q. In each case, we have checked whether the correlation of portfolio returns, defined as $\rho_{pq} = \sigma_{pq}/(\sigma_p\sigma_q)$, is in the permissible range as well. The portfolio weights, labeled as x_1, x_2, x_3, y_1, y_2, and y_3, are stored in cells B13:D13 and B15:D15. Whenever there are changes in these portfolio weights, the corresponding computational results are automatically updated in the worksheet.

The results in Figure 2 are from one of the many sets of portfolio weights that invalidate the covariance matrix in cells B5:D7. Specifically, for $x_1 = 0.3$, $x_2 = 0$, $x_3 = 0.7$, $y_1 = 0.1$, $y_2 = 0.7$, and $y_3 = 0.2$, the computed values of $x_i x_j \sigma_{ij}$, $y_i y_j \sigma_{ij}$, and $x_i y_j \sigma_{ij}$, for $i, j = 1, 2, 3$, are displayed in cells B17:D19, B21:D23, and B25:D27, respectively. Summing the individual 3×3 blocks gives us $\sigma_p^2 = 8.44$, $\sigma_q^2 = 26.46$, and $\sigma_{pq} = 15.45$, as shown in cells G19, G23, and G27, respectively. The implied correlation $\rho_{pq} = 1.0339$, as shown in cell K27, is outside the permissible range, thus indicating that the covariance matrix in question is unacceptable.

This example points out that there must be a requirement for the covariance matrix of security returns in addition to, or encompassing, the obvious requirement that all pairwise correlations be in the permissible range of -1 to 1. The requirement is that the covariance matrix be positive definite. In the context of portfolio investments under the assumption of frictionless short sales, the requirement amounts to the variance of portfolio returns being always strictly positive, regardless of how investment funds are allocated among the risky securities considered. A simple way to find out whether the 3×3 covariance matrix is positive definite is by searching, with the numerical tool *Solver* in Excel, for the set of portfolio weights that minimizes the variance of portfolio returns. If the numerical search produces a strictly positive variance of portfolio returns, the covariance matrix is positive definite; otherwise, it is not.

Figure 3 shows the *Solver* results. Initially, with the contents of cells B13:D13 — which provide equal portfolio weights (of 1/3 each) for the three securities — copied to cells B15:D15, the corresponding variance of portfolio returns is displayed in cell G19, by following the same computational steps for portfolio p in Figure 2. The initial results are not shown in Figure 3, as any changes in cells B15:D15 will automatically allow cell G19 to be updated. By using *Solver*, we minimize the target cell G19, by changing the portfolio weights in cells B15:D15, subject to the constraint that cell G15, the sum of the portfolio weights, be equal to 1. As the lowest variance from the numerical search, displayed in cell G19 is -3.9893, is a negative number, the covariance matrix in cells B5:D7 is not positive definite.

An alternative way to use Excel to find out whether the covariance matrix is positive definite is by checking the signs of all leading principal minors of the covariance matrix. In a three-security case, the three leading principal minors are the determinants of the $h \times h$ matrices consisting of the first h columns and the first h rows of the covariance matrix, for $h = 1$, 2, and 3. As shown in cells B22, C23, and D24, the three leading principal minors are 100, 2511, and -4464, respectively. With the presence of a negative number here, the 3×3 covariance matrix is not positive definite and is thus invalid.

As the validity of a given covariance matrix as part of the input parameters for portfolio analysis is not readily noticeable, the above example illustrates the importance of verifying its positive definiteness. A positive definite covariance matrix can sometimes differ from a non-positive definite case in just a few matrix elements. For example, if we have $\sigma_{13} = \sigma_{31} = -12$ instead of -20, while keeping all other elements of the covariance

	A	B	C	D	E	F	G	H	I	J	K
1	Sec Label	1	2	3							
2											
3	St Dev	10	6	4							
4						St Dev		Wgt p		Wgt q	
5	Cov Mat	100	33	-20		10		0.3		0.1	
6		33	36	12		6		0		0.7	
7		-20	12	16		4		0.7		0.2	
8											
9	Corr Mat	1	0.55	-0.5							
10		0.55	1	0.5							
11		-0.5	0.5	1							
12											
13	Wgt p	0.3	0	0.7		Total Wgt	1				
14											
15	Wgt q	0.1	0.7	0.2		Total Wgt	1				
16											
17	Wgt p * Cov * Wgt p	9	0	-4.2							
18		0	0	0							
19		-4.2	0	7.84		Var p	8.44				
20						St Dev p	2.9052				
21	Wgt q * Cov * Wgt q	1	2.31	-0.4							
22		2.31	17.64	1.68							
23		-0.4	1.68	0.64		Var q	26.46				
24						St Dev q	5.1439				
25	Wgt p * Cov * Wgt q	3	0	-1.4							
26		6.93	0	5.88						Implied	
27		-1.2	0	2.24		Cov pq	15.45			Corr pq	1.0339
28											
29	Cell Formulas	B3		=SQRT(B5)							
30		C3		=SQRT(C6)							
31		D3		=SQRT(D7)							
32		B6		=C5							
33		B7		=D5							
34		C7		=D6							
35		F5:F7		{=TRANSPOSE(A3:D3)}							
36		H5:H7		{=TRANSPOSE(A13:D13)}							
37		J5:J7		{=TRANSPOSE(A15:D15)}							
38		B9		=B5/B$3/$F5				Pasted to B9:D11			
39		G13		=SUM(B13:D13)				Pasted to G15			
40		B17		=B$13*B5*$H5				Pasted to B17:D19			
41		G19		=SUM(B17:D19)				Pasted to G23 and G27			
42		G20		=SQRT(G19)				Pasted to G24			
43		B21		=B$15*B5*$J5				Pasted to B21:D23			
44		B25		=B$13*B5*$J5				Pasted to B25:D27			
45		K27		=G27/G20/G24							

Figure 2: An Excel example illustrating the invalidity of a covariance matrix of security returns based on the correlation of portfolio returns.

	A	B	C	D	E	F	G	H
1	Sec Label	1	2	3				
2								
3	St Dev	10	6	4				
4						St Dev		Wgt p
5	Cov Mat	100	33	-20		10		0.6971
6		33	36	12		6		-1.2761
7		-20	12	16		4		1.5791
8								
9	Corr Mat	1	0.55	-0.5				
10		0.55	1	0.5				
11		-0.5	0.5	1				
12								
13	Initial Weight	0.33333	0.33333	0.33333		Total Wgt	1	
14								
15	Wgt p	0.69705099	-1.276139873	1.57909		Total Wgt	1	
16								
17	Wgt p * Cov * Wgt p	48.58800828	-29.35464055	-22.014				
18		-29.35464055	58.62718715	-24.182				
19		-22.01410939	-24.18165945	39.8963		Var p	-3.9893	
20								
21	Leading Prin Minors							
22	(1x1)	100						
23	(2x2)		2511					
24	(3x3)			-4464				
25								
26	Cell Formulas	B3	=SQRT(B5)					
27		C3	=SQRT(C6)					
28		D3	=SQRT(D7)					
29		B6	=C5					
30		B7	=D5					
31		C7	=D6					
32		F5:F7	{=TRANSPOSE(A3:D3)}					
33		H5:H7	{=TRANSPOSE(A15:D15)}					
34		B9	=B5/B$3/$F5				Pasted to B9:D11	
35		G13	=SUM(B13:D13)				Pasted to G15	
36		B17	=B$13*B5*$H5				Pasted to B17:D19	
37		G19	=SUM(B17:D19)					
38		B22	=MDETERM($B$5:B5)				Pasted to C23 and D24	
39								
40	Solver	Min		$G$19				
41		Changing Cells		$B$15:$D$15				
42		Constraint		$G$15=1				

Figure 3: An Excel example illustrating the invalidity of the same covariance matrix of security returns in Figure 2 based on the Solver result of portfolio variance minimization and the signs of the leading principal minors.

matrix unchanged, the same computations for the worksheets in Figures 2 and 3 will confirm that the revised covariance matrix is positive definite.

Specifically, regardless of how investment funds are allocated in portfolios p and q, the implied correlations of portfolio returns, as displayed in cell K27 of the worksheet in Figure 2, are always in the permissible range. The *Solver* result of the lowest possible portfolio variance, as displayed in cell G19 of the worksheet in Figure 3, becomes 7.2995, a positive number instead. Further, the three leading principal minors, as displayed in cells B22, C23, and D24 of the same worksheet are 100, 2511, and 11088, respectively; that is, they are all positive.

4 A basic portfolio selection model

The above Excel example illustrates that, for the portfolio solution to be meaningful, the covariance matrix involved must be positive definite. However, this being an illustration, the positive definiteness requirement still has to be justified properly. The task requires the use of a portfolio selection model. Following Roll (1977), we minimize the variance of portfolio returns, for a given set of n risky securities, subject to an expected return requirement. Under the assumption of frictionless short sales, if a portfolio p is formed, its expected return and variance of returns are

$$\mu_p = \sum\nolimits_{i=1}^{n} x_i \mu_i \tag{6}$$

$$\text{and } \sigma_p^2 = \sum\nolimits_{i=1}^{n} \sum\nolimits_{j=1}^{n} x_i x_j \sigma_{ij}, \tag{7}$$

respectively, satisfying the condition of $\sum_{i=1}^{n} x_i = 1$. In matrix notation, let $\boldsymbol{\mu}$ and $\boldsymbol{x}$ be n-element column vectors of expected returns and portfolio weights, with the corresponding elements being μ_i and x_i, for $i = 1, 2, \ldots, n$. Let $\boldsymbol{V}$ be a symmetric $n \times n$ covariance matrix of returns, with its element (i, j) being σ_{ij}, for $i, j = 1, 2, \ldots, n$. Let also $\boldsymbol{\iota}$ be an n-element column vector where each element is 1. Then, equations (6) and (7) can be written more compactly as $\mu_p = \boldsymbol{x}'\boldsymbol{\mu}$ and $\sigma_p^2 = \boldsymbol{x}'\boldsymbol{V}\boldsymbol{x}$, with $\boldsymbol{x}'\boldsymbol{\iota} = 1$.

For a predetermined μ_p, the Lagrangean is

$$L = \boldsymbol{x}'\boldsymbol{V}\boldsymbol{x} - \phi(\boldsymbol{x}'\boldsymbol{\mu} - \mu_p) - \theta(\boldsymbol{x}'\boldsymbol{\iota} - 1), \tag{8}$$

where the portfolio weight vector $\boldsymbol{x}$ and the Lagrange multipliers ϕ and θ are decision variables.[5] It is implicit that the elements of $\boldsymbol{\mu}$ are not all equal; otherwise, as all portfolios based on the n securities will have the same expected return, the use of a predetermined μ_p that differs from this expected return will inevitably fail to provide any portfolio allocation results. Minimizing L leads to

$$\boldsymbol{x} = \boldsymbol{V}^{-1}\boldsymbol{M}\,(\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}\boldsymbol{r}_p, \tag{9}$$

[5] See Kwan (2007) for an equivalent formulation of the same portfolio selection problem and for an intuitive explanation of the Lagrangian approach.

where $\boldsymbol{M} = [\ \boldsymbol{\mu} \quad \boldsymbol{\iota}\ ]$ is an $n \times 2$ matrix and $\boldsymbol{r}_p = [\ \mu_p \quad 1\ ]'$ is a 2-element column vector.[6] The variance of returns of the minimum variance portfolio (for a predetermined μ_p) is

$$\sigma_p^2 = \boldsymbol{x}'\boldsymbol{V}\boldsymbol{x} = \boldsymbol{r}_p{}'(\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}\boldsymbol{r}_p. \tag{10}$$

By repeating the computations based on equation (9) for different values of μ_p, we can establish a family of minimum variance portfolios. The least risky portfolio of the family is called the global minimum variance portfolio. Such a portfolio can also be reached by minimizing $\boldsymbol{x}'\boldsymbol{V}\boldsymbol{x}$ subject only to $\boldsymbol{x}'\boldsymbol{\iota} = 1$. Similar to the derivation of equation (9), minimizing the Lagrangean $L = \boldsymbol{x}'\boldsymbol{V}\boldsymbol{x} - \theta(\boldsymbol{x}'\boldsymbol{\iota} - 1)$ leads to

$$\boldsymbol{x}_o = \boldsymbol{V}^{-1}\boldsymbol{\iota}\ (\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\iota})^{-1}. \tag{11}$$

Here, the vector of portfolio weights, $\boldsymbol{x}$, has been labeled as $\boldsymbol{x}_o$ instead for notational clarity. The expected return and the variance of returns of the global minimum variance portfolio are

$$\mu_o = \boldsymbol{\mu}'\boldsymbol{x}_o = \boldsymbol{\mu}'\boldsymbol{V}^{-1}\boldsymbol{\iota}\ (\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\iota})^{-1} \tag{12}$$

$$\text{and } \sigma_o^2 = \boldsymbol{x}_o{}'\boldsymbol{V}\boldsymbol{x}_o = (\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\iota})^{-1}, \tag{13}$$

respectively.

As equation (13) shows, the variance of returns of the global minimum variance portfolio is the reciprocal of the sum of all elements of the inverse of the covariance matrix. The proportion of investment funds for each security in this portfolio, according to equation (11), is the sum of all elements of the corresponding row of the inverse of the covariance matrix, divided by the sum of all elements of the same inverse matrix. Thus, the sum of portfolio weights being unity is assured.

5 The positive definiteness requirement and its implications

From an analytical perspective, equations (9)-(13) are results of the first-order conditions for optimization. Whether such results actually correspond to variance minimization as intended, rather than variance maximization or the presence of saddle points, still requires confirmation. Of interest, therefore, is whether the input parameters $\boldsymbol{\mu}$ and $\boldsymbol{V}$ are required to satisfy some specific conditions in order to ensure variance minimization.

As an initial step in the search of conditions on the input parameters for equations (9)-(13) to work properly, we must ensure that the covariance matrix be invertible. Noting that the invertibility of the covariance matrix requires its determinant to be nonzero, we must rule out the following three situations. First, if a security considered is

[6] To derive equation (9), we use $\partial L/\partial \boldsymbol{x} = 2\boldsymbol{V}\boldsymbol{x} - \phi\boldsymbol{\mu} - \theta\boldsymbol{\iota} = \boldsymbol{0}$, where $\boldsymbol{0}$ is an n-element column vector of zeros. This equation can be rewritten as $\boldsymbol{x} = \frac{1}{2}\boldsymbol{V}^{-1}\boldsymbol{M}\boldsymbol{G}$, where $\boldsymbol{G} = [\ \phi \quad \theta\ ]'$. As $\boldsymbol{M}'\boldsymbol{x} = \boldsymbol{r}_p$, we have $\frac{1}{2}\boldsymbol{G} = (\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}\boldsymbol{r}_p$. Equation (9) follows directly. See Merton (1972) for a non-matrix version of the same portfolio selection model.

risk-free, it will have a zero variance of returns and zero covariances of returns with all other securities considered; this situation will produce a row (column) of zeros in the covariance matrix. Second, if the random return of a security is perfectly correlated with that of another security considered, two rows (columns) of the covariance matrix will be proportional to each other. Third, if the random return of a security is a linear combination of the random returns of some other securities considered, a row (column) of the covariance matrix can be replicated exactly by combining other rows (columns). In each situation, the covariance matrix is non-invertible, as the corresponding determinant is zero. The above three situations arising from sample estimates of the covariance matrix are considered in more detail in Appendix A.

Students in portfolio investment courses, most likely, are made aware of the above three situations and their implications. The presence of a risk-free security transforms the portfolio selection problem into a two-part problem, with the first part pertaining to the selection of a portfolio based only on the risky securities and the second part pertaining to the allocation of investment funds between the risk-free security and the risky portfolio. In the remaining two situations, the contribution of a security to the risk-return trade-off of the portfolio can be replicated exactly by that of another security or a combination of other securities. Thus, the portfolio selection results will no longer be unique, and equations (9)-(13) will fail to perform their intended tasks.

With the above three situations ruled out, we now show that, for equations (9)-(13) to work properly, the covariance matrix must be positive definite. Indeed, if the covariance matrix is positive definite, the above three situations will automatically be ruled out, and these equations will work as intended. In the language of matrix algebra, an $n \times n$ matrix $\boldsymbol{V}$ is positive semidefinite if $\boldsymbol{x}'\boldsymbol{V}\boldsymbol{x}$, which is a scalar, is always non-negative for any n-element column vector $\boldsymbol{x}$. It is also positive definite if $\boldsymbol{x}'\boldsymbol{V}\boldsymbol{x}$ is strictly positive for any non-zero vector $\boldsymbol{x}$. In the context of portfolio analysis, $\boldsymbol{V}$ is the covariance matrix and $\boldsymbol{x}$ is the portfolio weight vector. As it will soon be clear, with the positive definiteness requirement satisfied, the invertibility of $\boldsymbol{V}$, $\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M}$, and $\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\iota}$ is assured. However, the converse is not true; the invertibility of these matrices does not ensure that the covariance matrix be positive definite.

We now turn our attention to the condition for a minimized Lagrangean in the basic portfolio selection model. In that model, we seek to minimize the Lagrangean L in equation (8) with the portfolio weights $x_1, x_2, \ldots, x_n$ and the Lagrange multipliers ϕ and θ being the decision variables. Now, for ease of exposition, let us treat ϕ and θ as x_{n+1} and x_{n+2}, respectively. As the Lagrangean L is a quadratic function of the $n+2$ decision variables, its partial derivatives beyond the second order are all zeros. Then, with L^* being the Lagrangean as evaluated at $x_1 = x_1^*$, $x_2 = x_2^*$, $\ldots$, $x_{n+2} = x_{n+2}^*$, for which its first partial derivative with respect to each of the $n+2$ variables is set to be zero, we can express L by means of a second-order Taylor expansion. That is, we can write

$$L = L^* + \frac{1}{2} \sum\nolimits_{i=1}^{n+2} \sum\nolimits_{j=1}^{n+2} (x_i - x_i^*)(x_j - x_j^*) \frac{\partial^2 L}{\partial x_i \partial x_j}, \tag{14}$$

where each second partial derivative is evaluated at $x_1 = x_1^*, x_2 = x_2^*, \ldots, x_{n+2} = x_{n+2}^*$.[7]

The $(n+2) \times (n+2)$ symmetric matrix consisting of the above second partial derivatives is commonly called the bordered Hessian. In such a matrix, we have $\partial^2 L/(\partial x_i \partial x_j) = 2\sigma_{ij}$, $\partial^2 L/(\partial x_i \partial \phi) = -\mu_i$, $\partial^2 L/(\partial x_i \partial \theta) = -1$, and $\partial^2 L/\partial \phi^2 = \partial^2 L/\partial \theta^2 = \partial^2 L/(\partial \phi \partial \theta) = 0$, for $i, j = 1, 2, \ldots, n$. Noting that $\sum_{i=1}^n x_i \mu_i = \sum_{i=1}^n x_i^* \mu_i = \mu_p$ and $\sum_{i=1}^n x_i = \sum_{i=1}^n x_i^* = 1$, we can write equation (14) as

$$L - L^* = \sum\nolimits_{i=1}^n \sum\nolimits_{j=1}^n (x_i - x_i^*)(x_j - x_j^*)\sigma_{ij} = (\triangle \boldsymbol{x})' \boldsymbol{V} (\triangle \boldsymbol{x}), \tag{15}$$

where $\triangle \boldsymbol{x}$ is an n-element column vector with elements being $x_i - x_i^*$, for $i = 1, 2, \ldots, n$. For L^* to be a minimum, we must have $L - L^* > 0$. That is, we must have $(\triangle \boldsymbol{x})' \boldsymbol{V} (\triangle \boldsymbol{x}) > 0$. With $\triangle \boldsymbol{x}$ being arbitrary, we confirm that minimization of the Lagrangean requires the covariance matrix to be positive definite.[8]

The sub-sections below show various implications of the covariance matrix being positive definite, along with some cautionary notes.

5.1 The inverse of the covariance matrix and the global minimum variance portfolio

If the covariance matrix $\boldsymbol{V}$ is positive definite, so is $\boldsymbol{V}^{-1}$. The reason is that, as $\boldsymbol{V}$ is invertible and as $\boldsymbol{x}' \boldsymbol{V} \boldsymbol{x}$ is positive for any non-zero column vector $\boldsymbol{x}$,

$$\boldsymbol{x}' \boldsymbol{V} \boldsymbol{x} = \boldsymbol{x}' \boldsymbol{V} (\boldsymbol{V}^{-1} \boldsymbol{V})\, \boldsymbol{x} = (\boldsymbol{V} \boldsymbol{x})' \boldsymbol{V}^{-1} (\boldsymbol{V} \boldsymbol{x}) \tag{16}$$

is also positive. With $\boldsymbol{V}\boldsymbol{x}$ being an arbitrary column vector and $(\boldsymbol{V}\boldsymbol{x})'$ being its transpose, the positive definiteness of $\boldsymbol{V}^{-1}$ is assured. If an invertible $\boldsymbol{V}$ is not positive definite, a non-zero column vector $\boldsymbol{x}$ corresponding to $\boldsymbol{x}' \boldsymbol{V} \boldsymbol{x}$ being non-positive must exist. For such a vector $\boldsymbol{x}$, we have a corresponding column vector $\boldsymbol{V}\boldsymbol{x}$. Then, in view of equation (16), $\boldsymbol{V}^{-1}$ is also not positive definite.

Now, suppose that $\boldsymbol{V}$ is invertible but its positive definiteness has not been confirmed. For $\boldsymbol{V}^{-1}$ to be positive definite, it must not have any negative diagonal elements. To verify this requirement, suppose instead that $\boldsymbol{V}^{-1}$ has at least a negative diagonal element, with element (i, i) being negative. Let $\boldsymbol{c}$ be an n-element column vector with

[7] For students who are unfamiliar with multivariate Taylor series, an explanation of equation (14) is necessary. The idea is similar to the expansion of a quadratic function $L(x)$. In this univariate case, we have $L(x) = L(x^*) + (x - x^*)L'(x^*) + \frac{1}{2}(x - x^*)^2 L''(x^*)$, with the derivatives evaluated at $x = x^*$. To extend to a multivariate case, we must account for all second derivatives of L; that is, $\partial^2 L/(\partial x_i \partial x_j)$, for all i, j. This is captured by the terms under the double summation on the right hand side of equation (14).

[8] The above proof encompasses, as a special case, the Lagrangian that is intended for the global minimum variance portfolio. In this special case with $n + 1$ decision variables, the only Lagrange multiplier, θ, can be treated as x_{n+1}. Equation (15) can be rewritten analogously. For $x_1^*, x_2^*, \ldots, x_{n+1}^*$ from the first-order conditions to provide a minimized Lagrangian, we also require $(\triangle \boldsymbol{x})' \boldsymbol{V} (\triangle \boldsymbol{x}) > 0$ or, equivalently, $\boldsymbol{V}$ to be positive definite.

element i being its only non-zero element. In such a case, $\boldsymbol{c}'\boldsymbol{V}^{-1}\boldsymbol{c}$ is negative.[9] As $\boldsymbol{V}^{-1}$ is not positive definite, its inverse — which is $\boldsymbol{V}$ — cannot be positive definite. The implication is that, if the inverse of the covariance matrix has any negative diagonal elements, the covariance matrix itself cannot be positive definite.

The positive definiteness of $\boldsymbol{V}$, which implies the same for $\boldsymbol{V}^{-1}$, ensures that the scalar $\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\iota}$ be positive. If $\boldsymbol{V}$ is positive definite, then, with $(\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\iota})^{-1}$ being positive, equations (11)-(13) do provide portfolio allocation results for the global minimum variance portfolio as intended. However, without first confirming that $\boldsymbol{V}$ is positive definite, we cannot confirm the validity of the results from these equations even if $(\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\iota})^{-1}$ turns out to be positive.

What is crucial here is that, although the set of portfolio weights from equation (11) allows us to compute the corresponding variance of portfolio returns as $(\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\iota})^{-1}$, whether this variance is the lowest possible variance still depends on the positive definiteness of $\boldsymbol{V}$. If $\boldsymbol{V}$ is not positive definite, the set of portfolio weights from equation (11) does not give us a minimum of the Lagrangean L; what we get can be a maximum or a saddle point instead.

5.2 The inverse of a 2×2 matrix in the basic portfolio selection model and minimum variance portfolios

The invertibility of the 2×2 matrix $\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M}$ in equations (9) and (10) requires that its determinant — labeled as k here — be non-zero. As the elements $(1,1)$, $(1,2)$, $(2,1)$, and $(2,2)$ of $\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M}$ are $\boldsymbol{\mu}'\boldsymbol{V}^{-1}\boldsymbol{\mu}$, $\boldsymbol{\mu}'\boldsymbol{V}^{-1}\boldsymbol{\iota}$, $\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\mu}$, and $\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\iota}$, respectively, we have

$$k = (\boldsymbol{\mu}'\boldsymbol{V}^{-1}\boldsymbol{\mu})(\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\iota}) - (\boldsymbol{\mu}'\boldsymbol{V}^{-1}\boldsymbol{\iota})(\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\mu}), \tag{17}$$

where $\boldsymbol{\mu}'\boldsymbol{V}^{-1}\boldsymbol{\iota} = \boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\mu}$. For the proof below that k is positive and thus $\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M}$ is invertible, we draw on the matrix property that a symmetric positive definite matrix can be written as a product of a square matrix and its transpose. The proof of this matrix property is provided in Appendix C.

Specifically, with $\boldsymbol{V}^{-1}$ being symmetric and positive definite, we can write $\boldsymbol{V}^{-1} = \boldsymbol{L}\boldsymbol{L}'$, where $\boldsymbol{L}$ is an $n \times n$ matrix. It follows that

$$\boldsymbol{\mu}'\boldsymbol{V}^{-1}\boldsymbol{\mu} = (\boldsymbol{L}'\boldsymbol{\mu})'(\boldsymbol{L}'\boldsymbol{\mu}) = \sum_{i=1}^{n} a_i^2 > 0, \tag{18}$$

where a_i is element i of the column vector $\boldsymbol{L}'\boldsymbol{\mu}$. Likewise, we can write

$$\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\iota} = (\boldsymbol{L}'\boldsymbol{\iota})'(\boldsymbol{L}'\boldsymbol{\iota}) = \sum_{i=1}^{n} b_i^2 > 0, \tag{19}$$

where b_i is element i of the column vector $\boldsymbol{L}'\boldsymbol{\iota}$. We can also write $\boldsymbol{\mu}'\boldsymbol{V}^{-1}\boldsymbol{\iota}$ and $\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\mu}$ as

$$(\boldsymbol{L}'\boldsymbol{\mu})'(\boldsymbol{L}'\boldsymbol{\iota}) = \sum_{i=1}^{n} a_i b_i. \tag{20}$$

[9] To see this, let us label element i of $\boldsymbol{c}$ as c_i and element (i,i) of $\boldsymbol{V}^{-1}$ as g_{ii}. Further, let all other elements of $\boldsymbol{c}$ be zeros. As element i of the n-element column vector $\boldsymbol{V}^{-1}\boldsymbol{c}$ is $g_{ii}c_i$, it follows that $\boldsymbol{c}'\boldsymbol{V}^{-1}\boldsymbol{c} = c_i g_{ii} c_i = c_i^2 g_{ii} < 0$ if $g_{ii} < 0$.

Drawing on Cauchy-Schwarz inequality, we have

$$k = \left(\sum\nolimits_{i=1}^{n} a_i^2\right)\left(\sum\nolimits_{i=1}^{n} b_i^2\right) - \left(\sum\nolimits_{i=1}^{n} a_i b_i\right)^2 \geq 0. \tag{21}$$

Strict inequality holds if the case of $a_i = cb_i$, with c being any constant, for $i = 1, 2, \ldots, n$, can be ruled out. A simple algebraic proof of Cauchy-Schwarz inequality is provided in Appendix D. As it is implicit in the model formulation in Section 4 that the elements of $\boldsymbol{\mu}$ are not all equal, strict inequality is assured; that is, k is strictly positive and, accordingly, $\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M}$ is invertible.

The positive definiteness of $\boldsymbol{V}$ ensures that the two leading principal minors of the 2×2 matrix $\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M}$ — which are $\boldsymbol{\mu}'\boldsymbol{V}^{-1}\boldsymbol{\mu}$ and k — are both positive. That is, if $\boldsymbol{V}$ is positive definite, so are $\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M}$ and $(\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}$. Then, regardless of the value of μ_p in the 2-element column vector $\boldsymbol{r}_p$ in equations (9) and (10), these equations will always provide minimum variance portfolios, all with positive variances of portfolio returns, as intended. However, if $\boldsymbol{V}$ is invertible but not positive definite, the 2×2 matrix $\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M}$ based on some given input parameters $\boldsymbol{\mu}$ and $\boldsymbol{V}$ can still be positive definite, as indicated by both $\boldsymbol{\mu}'\boldsymbol{V}^{-1}\boldsymbol{\mu}$ and k being positive. In such a case, although equation (10) will still provide positive variances of portfolio returns for all values of μ_p, the results are invalid, as they do not correspond to minimization of the Lagrangean L. Therefore, it is important to confirm the positive definiteness of $\boldsymbol{V}$ before accepting the results from equations (9) and (10).

5.3 Graphs of minimum variance portfolios on the $(\boldsymbol{\sigma}, \boldsymbol{\mu})$-plane

Let us label elements $(1,1)$, $(1,2)$, and $(2,2)$ of the symmetric matrix $(\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}$ as α, γ, and β, respectively. With $\alpha = \boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\iota}\,/k$, $\gamma = -\boldsymbol{\mu}'\boldsymbol{V}^{-1}\boldsymbol{\iota}\,/k$, and $\beta = \boldsymbol{\mu}'\boldsymbol{V}^{-1}\boldsymbol{\mu}\,/k$, we can write equation (10) as

$$\sigma_p^2 = \alpha\mu_p^2 + 2\gamma\mu_p + \beta, \tag{22}$$

which, by completing the square, becomes

$$\sigma_p^2 = \alpha\left(\mu_p + \frac{\gamma}{\alpha}\right)^2 + \frac{\alpha\beta - \gamma^2}{\alpha}. \tag{23}$$

Noting that the determinants of $\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M}$ and $(\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}$ are reciprocals of each other, we have $\alpha\beta - \gamma^2 = 1/k$.

If $\boldsymbol{V}$ is positive definite, both $\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\iota}$ and k are positive. Accordingly, α and $\alpha\beta - \gamma^2$ are also positive. Thus, with $(\alpha\beta - \gamma^2)/\alpha$ being positive, equation (23) is a hyperbola with a horizontal transverse axis on the (σ, μ)-plane. The centre of the hyperbola is the point $(0, -\gamma/\alpha)$ on the (σ, μ)-plane and its vertices are the points $\left(\pm\sqrt{(\alpha\beta - \gamma^2)/\alpha}, -\gamma/\alpha\right)$. With $-\gamma/\alpha = \mu_o$ and $(\alpha\beta - \gamma^2)/\alpha = \sigma_o^2$, representing the expected return and the variance of returns of the global minimum variance portfolio, respectively, equation (23) can also be written as

$$\sigma_p^2 = \alpha\left(\mu_p - \mu_o\right)^2 + \sigma_o^2. \tag{24}$$

By definition, standard deviations are never negative. Thus, the relevant branch of the hyperbola is the one where $\sigma_p > 0$. With the transverse axis of the hyperbola being horizontal, this branch can accommodate all values of μ_p; that is, there will not be any value of μ_p that can lead to a negative σ_p^2. The efficient frontier is the upper half of this branch, starting from the global minimum variance portfolio. It is a concave curve on the (σ, μ)-plane, as indicated by $d\mu_p/d\sigma_p > 0$ and $d^2\mu_p/d\sigma_p^2 < 0$. This characterization of the efficient frontier is what students learn in investment courses.

However, if $\boldsymbol{V}$ is not positive definite, the graph of equation (23) on the (σ, μ)-plane can be any conic section, depending on the signs of α and $\alpha\beta - \gamma^2$. More specifically, we have the following potential cases:

1. If both α and $\alpha\beta - \gamma^2$, as computed from the input parameters $\boldsymbol{\mu}$ and $\boldsymbol{V}$, are positive, the graph is also a hyperbola with the same characteristics as described above.

2. If α is positive, but $\alpha\beta - \gamma^2$ is negative instead, the resulting hyperbola will have the μ-axis as its transverse axis.

3. If both α and $\alpha\beta - \gamma^2$ are negative, the graph is an ellipse.[10]

4. If $\alpha = 0$, equation (22) reduces to $\sigma^2 = 2\gamma\mu + \beta$, which is a parabola with the μ-axis being its axis of symmetry.[11]

Among the four cases here, only the first case provides positive portfolio variances for all predetermined values of μ_p. In the remaining three cases, if the predetermined μ_p does not have a corresponding σ_p on the graph, a negative σ_p^2 will be produced. For example, in the second case, if μ_p is set at values in the gap between the two vertices of the hyperbola, there cannot be any corresponding real values of σ_p. Obviously, none of graphs of risk-return trade-off for these three cases resemble what students learn in investment courses.

Then, is the first case above acceptable? Although it provides a seemingly reasonable graph on the (σ, μ)-plane, with $d\mu_p/d\sigma_p > 0$ and $d^2\mu_p/d\sigma_p^2 < 0$ for all portfolios with expected returns greater than μ_o, the graph does not correspond to the efficient frontier. The point (σ_o, μ_o) does not even correspond to the global minimum variance portfolio. The reason is that none of the portfolios on the graph are results of minimization of the Lagrangean L; instead, they correspond to saddle points of L as a function of the $n + 2$ decision variables, $x_1, x_2, \ldots, x_n, \phi$, and θ.

---

[10] If α is negative and $\alpha\beta - \gamma^2$ is positive instead, equation (23) cannot be graphed on the (σ, μ)-plane.

[11] For equation (9) to produce any portfolio allocation results, $\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M}$ must be invertible. Thus, the determinant of $\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M}$ cannot be zero. The determinant of $(\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}$, which is $\alpha\beta - \gamma^2$, is the reciprocal of the determinant of $\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M}$. Thus, with $\alpha\beta - \gamma^2$ not being zero, α and γ cannot both be zeros.

5.4 Implied correlations of returns

Satisfaction of the positive definiteness requirement for the $n \times n$ matrix ensures that its first two leading principal minors be positive. That is, we must have $-1 < \rho_{12} < 1$. There are $n!$ ways to label the n securities as $1, 2, \ldots, n$ and to arrange the corresponding variances and covariances of security returns in an $n \times n$ matrix. As any pair of securities can be labeled as securities 1 and 2, the positive definiteness requirement ensures that $-1 < \rho_{ij} < 1$, for all $i \neq j$. However, as illustrated in the Excel example earlier, the converse is not true; having all pairwise correlations in the permissible range does not imply a positive definite covariance matrix. This explains why a covariance matrix where all implied pairwise correlations of returns are in the permissible range can still give us strange portfolio results.

As $\alpha\beta - \gamma^2$ is the determinant of $(\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}$, its being positive ensures that the correlation of returns between any two minimum variance portfolios be in the range of -1 to 1. To see this, let σ_{pq} be the covariance of returns between two arbitrary minimum variance portfolios p and q with expected returns μ_p and μ_q, respectively. Given the corresponding portfolio weight vectors $\boldsymbol{x}$ and $\boldsymbol{y}$, the covariance is $\sigma_{pq} = \boldsymbol{x}'\boldsymbol{V}\,\boldsymbol{y}$. According to equation (9), we have

$$\boldsymbol{x}'\boldsymbol{V}\,\boldsymbol{y} = \boldsymbol{r}'\,(\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}\boldsymbol{s}, \tag{25}$$

where $\boldsymbol{s} = \begin{bmatrix} \mu_q & 1 \end{bmatrix}'$ is a 2-element column vector. It is implicit that $\sigma_{pp} = \boldsymbol{x}'\boldsymbol{V}\,\boldsymbol{x}$, $\sigma_{qq} = \boldsymbol{y}'\boldsymbol{V}\,\boldsymbol{y}$, and $\sigma_{qp} = \boldsymbol{y}'\boldsymbol{V}\,\boldsymbol{x} = \boldsymbol{x}'\boldsymbol{V}\,\boldsymbol{y} = \sigma_{pq}$. The 2×2 matrix

$$\boldsymbol{A} = \begin{bmatrix} \sigma_{pp} & \sigma_{pq} \\ \sigma_{qp} & \sigma_{qq} \end{bmatrix} \tag{26}$$

can also be written as

$$\boldsymbol{A} = \begin{bmatrix} \mu_p & 1 \\ \mu_q & 1 \end{bmatrix} (\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1} \begin{bmatrix} \mu_p & \mu_q \\ 1 & 1 \end{bmatrix}. \tag{27}$$

Thus, the determinant of $\boldsymbol{A}$ is the product of the determinants of the three 2×2 matrices on the right hand side of equation (27); that is,

$$\sigma_{pp}\sigma_{qq} - \sigma_{qp}\sigma_{pq} = (\mu_p - \mu_q)^2(\alpha\beta - \gamma^2). \tag{28}$$

For $\boldsymbol{V}$ being positive definite, noting that $\alpha\beta - \gamma^2$ is positive, we have $(\sigma_{pq})^2 < \sigma_{pp}\sigma_{qq}$. That is, the correlation of returns between any two different minimum variance portfolios p and q based on the same set of risky securities is always in the range of -1 and 1, as required for the portfolio results to be valid.

Among the four cases identified in Sub-section 5.3, where $\boldsymbol{V}$ is not positive definite, only the first case has a positive $\alpha\beta - \gamma^2$. This is the only case where the correlation of returns between any two different portfolios with portfolio weights given by equation (9) is always in the permissible range. Obviously, the remaining three cases are all invalid; as $\alpha\beta - \gamma^2 < 0$, the implied correlations of portfolio returns are outside the permissible range.

6 An Excel illustration of the positive definiteness requirement

As shown in the previous section, if the graph of equation (22) on the (σ, μ)-plane is not a hyperbola with a horizontal transverse axis, the input parameters for the analysis are invalid. However, the attainment of such a hyperbola does not automatically ensure the validity of the input parameters; the requirement is still the positive definiteness of the covariance matrix. A crucial point is that a positive definite covariance matrix implies such a hyperbola, but not the other way round.

We now use two four-security cases to illustrate the positive definiteness requirement. The only difference between the two cases is in the covariance of returns between two specific securities. The covariance matrix in the first case is positive definite; the one in the second case is not. Figures 4a and 4b show part of an Excel worksheet for the first case. Here, the input parameters, consisting of

$$\boldsymbol{\mu}' = \begin{bmatrix} 12.4 & 10.8 & 9.0 & 8.5 \end{bmatrix}'$$
$$\text{and } \boldsymbol{V} = \begin{bmatrix} 100 & -32 & 36 & -25 \\ -32 & 64 & -24 & 12 \\ 36 & -24 & 36 & 3 \\ -25 & 12 & 3 & 25 \end{bmatrix},$$

as measured in percentage terms, are provided in cells `B3:E3` and `B8:E11`, respectively.

The positive definiteness of $\boldsymbol{V}$ is confirmed by the positive sign of its four leading principal minors, as shown in the cells `J8:J11`. As expected, all implied pairwise correlations of security returns, as shown in cells `B13:E16`, are in the permissible range of -1 to 1. To find $\boldsymbol{x}_o$, σ_o^2, and μ_o of the global minimum variance portfolio, we start with $\boldsymbol{V}^{-1}$, which is shown in cells `B18:E21`. We find each element of the row vector $\boldsymbol{x}_o'$ by dividing the corresponding 4-element column sum of the symmetric $\boldsymbol{V}^{-1}$ by the sum of all its 16 elements; the result is displayed in cells `B24:E24`. The reciprocal of the sum of all 16 elements of $\boldsymbol{V}^{-1}$, which is σ_o^2, is positive as expected; it is shown in cell `G24`. Its square root, which is σ_o, is shown in cell `I24`. With $\boldsymbol{x}_o$ determined, $\mu_o = \boldsymbol{x}_o'\boldsymbol{\mu}$ is provided in cell `J24`.

We also use *Solver* to find $\boldsymbol{x}_o$, σ_o^2, and μ_o. With the target cell `G26` being for $\sigma_o^2 = \boldsymbol{x}_o'\boldsymbol{V}\boldsymbol{x}_o$, we arbitrarily set equal weights of 1/4 for all four securities and paste these initial values from cells `B25:E25` to cells `B26:E26`. By changing the four portfolio weights in cells `B26:E26`, *Solver* minimizes the target cell subject to the only constraint that the sum of these portfolio weights, as captured by cell `F26`, be unity. As confirmed by the corresponding numbers in rows 24 and 26 of the Excel worksheet, the *Solver* results are numerically the same as those based on equations (11)-(13).

To facilitate the computations for any predetermined expected return μ_p based on equations (9) and (10), we show in cells `B31:C34`, `B37:C38`, and `B41:C42` the computed values of $\boldsymbol{V}^{-1}\boldsymbol{M}$, $\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M}$, and $(\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}$, respectively. As expected, the two leading principal minors of $\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M}$, as shown in cells `G37:G38`, are positive. So are the two principal minors of $(\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}$, as shown in cells `G41:G42`. The optimal

	A	B	C	D	E	F	G	H	I	J	K
1	Sec Label	1	2	3	4						
2											
3	Exp Ret	12.4	10.8	9.0	8.5						
4	One	1	1	1	1						
5											
6	St Dev	10	8	6	5						
7							St Dev		Lead Prin Minors		
8	V	100	-32	36	-25		10		det (1x1)	100	
9		-32	64	-24	12		8		det (2x2)	5376	
10		36	-24	36	3		6		det (3x3)	108288	
11		-25	12	3	25		5		det (4x4)	1099584	
12											
13	Corr Mat	1	-0.4	0.6	-0.5						
14		-0.4	1	-0.5	0.3						
15		0.6	-0.5	1	0.1						
16		-0.5	0.3	0.1	1						
17											
18	V inv	0.0324777	-0.006351	-0.040073	0.0403353						
19		-0.006351	0.0261917	0.0256461	-0.022001						
20		-0.040073	0.0256461	0.0902159	-0.063209						
21		0.0403353	-0.022001	-0.063209	0.0984809						
22											
23						Sum Xo	Var o		St Dev o	Exp Ret o	
24	Xo	0.2273696	0.2023571	0.1083877	0.4618857	1	8.616349		2.9353618	9.9063566	
25	Xo Initial	0.25	0.25	0.25	0.25	1					
26	Xo Solver	0.2273695	0.2023572	0.1083878	0.4618855	1	8.616349		2.9353618	9.9063566	
27											
28											
29											
30	(V inv) M										
31		0.3163174	0.0263882								
32		0.2479178	0.0234852								
33		0.054732	0.0125793								
34		0.5307491	0.0536057								
35											
36	M' (V inv) M					Lead Prin Minors					
37		11.603803	1.1497163			det (1x1)	11.603803				
38		1.1497163	0.1160584			det (2x2)	0.0248718				
39											
40	[M' (V inv) M] inv					Lead Prin Minors					
41		4.6662718	-46.22575			det (1x1)	4.6662718				
42		-46.22575	466.54513			det (2x2)	40.206226				
43											
44						Sum X	Var		St Dev	Exp Ret	One
45	X	0.5075726	0.2802557	-0.248242	0.4604141	1	14.197471		3.7679532	11	1
46	X Initial	0.25	0.25	0.25	0.25	1					
47	X Solver	0.5075728	0.2802558	-0.248243	0.460414	1	14.197481		3.7679545	11.000001	
48											

Figure 4a: An Excel example illustrating the consistency of the results of portfolio variance minimization and the corresponding Solver results for a positive definite covariance matrix of security returns.

	L	M	N	O	P	Q	R	S	T	U
1										
2	Cell Formulas	B6	=SQRT(B8)							
3		C6	=SQRT(C9)							
4		D6	=SQRT(D10)							
5		E6	=SQRT(E11)							
6		G8:G11	{=TRANSPOSE(A6:E6)}							
7										
8		J8	=MDETERM($B$8:B8)							
9		J9	=MDETERM($B$8:C9)							
10		J10	=MDETERM($B$8:D10)							
11		J11	=MDETERM($B$8:E11)							
12										
13		B13	=B8/B$6/$G8						Pasted to B13:E16	
14		B18:E21	{=MINVERSE(B8:E11)}							
15										
16		B24	=SUM(B18:B21)/SUM($B18:$E21)						Pasted to B24:E24	
17		F24	=SUM(B24:E24)						Pasted to F24:F26	
18		G24	=1/SUM(B18:E21)							
19		I24	=IF(G24>0,SQRT(G24),"")						Pasted to I26	
20		J24	{=MMULT(B24:E24,TRANSPOSE(B$3:E$3))}						Pasted to J26	
21		G26	{=MMULT(B26:E26,MMULT(B8:E11,TRANSPOSE(B26:E26)))}							
22										
23		B31:C34	{=MMULT(B18:E21,TRANSPOSE(B3:E4))}							
24		B37:C38	{=MMULT(B3:E4,B31:C34)}							
25		G37	=MDETERM($B$37:B37)							
26		G38	=MDETERM($B$37:C38)							
27										
28		B41:C42	{=MINVERSE(B37:C38)}							
29		G41	=MDETERM($B$41:B41)							
30		G42	=MDETERM($B$41:C42)							
31										
32		B45:E45	{=TRANSPOSE(MMULT(B31:C34,MMULT(B41:C42,TRANSPOSE(J45:K45))))}							
33		F45	=SUM(B45:E45)						Pasted to F45:F47	
34		G45	{=MMULT(B45:E45,MMULT(B$8:E$11,TRANSPOSE(B45:E45)))}						Pasted to G47	
35		I45	=IF(G45>0,SQRT(G45),"")						Pasted to I47	
36		J47	{=MMULT(B47:E47,TRANSPOSE(B3:E3))}							
37										
38										
39	Solver for Min Var o			Min		$G$26				
40				Changing Cells		$B$26:$E$26				
41				Constraint		$F$26=1				
42										
43										
44	Solver for Min Var with Given Exp Ret			Min		$G$47				
45				Changing Cells		$B$47:$E$47				
46				Constraints		$F$47=1				
47						$J$47=$J$45				
48										

Figure 4b: An Excel example illustrating the consistency of the results of portfolio variance minimization and the corresponding Solver results for a positive definite covariance matrix of security returns (continued).

portfolio weight vector, $\boldsymbol{x}$, and the corresponding variance and standard deviation of portfolio returns, σ_p^2 and σ_p, for $\mu_p = 11$ as an example, are shown in cells `B45:E45`, `G45`, and `I45`, respectively.

The above numerical results are also confirmed by using *Solver*. The approach is similar to that in the computations for the global minimum variance portfolio. The only difference is the additional constraint that the portfolio's expected return, $\boldsymbol{x}'\boldsymbol{\mu}$, as shown in cell `J47`, be equal to its predetermined value in cell `J45`, which is 11. Except for some minor rounding errors, the *Solver* results, as shown in the corresponding cells in row 47, are identical to the results based on equations (9) and (10).

By repeating the computations based on these two equations for different values of μ_p, we are able to provide corresponding values of σ_p and μ_p for a graph on the (σ, μ)-plane. The graph, as shown in Figure 5, is generated by using Excel's graphic feature *Charts, X Y (Scatter)*. The individual points of (σ_i, μ_i), for $i = 1, 2, 3$, and 4, are also shown. As expected, the graph is a branch of a hyperbola with a horizontal transverse axis. The upward-sloping part of the graph, starting from the global minimum variance portfolio, is the efficient frontier.

The input parameters for the Excel worksheet in Figure 6 differ from those in Figure 4a only in the covariance of returns between securities 3 and 4; we have $\sigma_{34} = \sigma_{43} = 12$ instead of 3. All implied correlations of returns, as shown in cells `B13:E16`, are still in the permissible range of -1 to 1. Now, let us first consider the results in Figure 6 that are based on equations (9)-(13). Specifically, the computed value of σ_o^2 based on equation (13) is positive, as shown in cell `G24`. The 2×2 matrix $(\boldsymbol{M}'\boldsymbol{V}^{-1}\boldsymbol{M})^{-1}$ has two positive leading principal minors, as shown in cells `G41:G42`. The two portfolio variances as shown in cells `G24` and `G45` are positive. Further, as shown in Figure 7, the graph on the (σ, μ)-plane for different values of σ_p and μ_p is still a hyperbola with a horizontal transverse axis. At first glance, as the graph has captured all the characteristics of minimum variance portfolios, equations (10)-(13) seem to have performed their intended tasks.

However, a closer inspection of Figure 6 reveals two problems. First, as shown in cell J11, the fourth leading principal minor — which is the determinant of the covariance matrix itself — is negative. Second, the inverse of the covariance matrix, as shown in cells `B18:E21`, has three negative diagonal elements. Both are indications that the covariance matrix is not positive definite. To show that the graph in Figure 7 does not correspond to minimum variance portfolios, we first repeat the same *Solver* runs in Figure 4a with initially equal portfolio weights and with the default setting of *Solver* options. The two portfolio variances as shown in cells `G26` and `G47`, intended for σ_o^2 and σ_p^2, are -9.3×10^{15} and -4.8×10^{15}, respectively. These results suggest that the true value in each case is minus infinity. Regardless of their true values, what is clear is that, with the covariance matrix not being positive definite, equations (10)-(13) do not correspond to any minimum variance portfolios.

As portfolio weights with extremely large magnitudes are unrealistic, we now restrict such magnitudes in the *Solver* runs in order to illustrate that, if the covariance matrix fails to be positive definite, some reasonable portfolio weights can still lead to portfolio

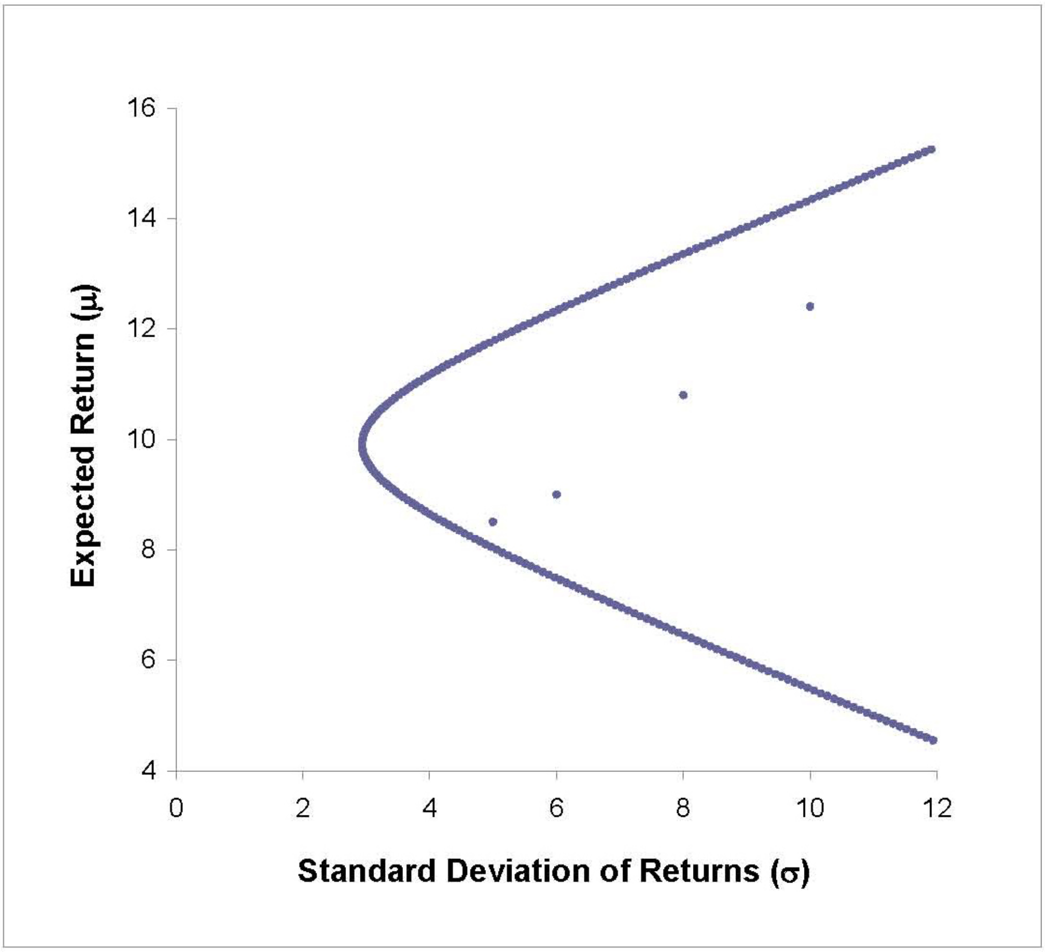

Figure 5: The graph of expected return versus standard deviation of returns for a basic portfolio selection model based on the input parameters in Figure 4a.

	A	B	C	D	E	F	G	H	I	J	K
1	Sec Label	1	2	3	4						
2											
3	Exp Ret	12.4	10.8	9.0	8.5						
4	One	1	1	1	1						
5											
6	St Dev	10	8	6	5						
7							St Dev		Lead Prin Minors		
8	V	100	-32	36	-25		10		det (1x1)	100	
9		-32	64	-24	12		8		det (2x2)	5376	
10		36	-24	36	12		6		det (3x3)	108288	
11		-25	12	12	25		5		det (4x4)	-586944	
12											
13	Corr Mat	1	-0.4	0.6	-0.5						
14		-0.4	1	-0.5	0.3						
15		0.6	-0.5	1	0.4						
16		-0.5	0.3	0.4	1						
17											
18	V inv	-0.03729	0.035083	0.093719	-0.09912						
19		0.035083	0.001533	-0.05418	0.060353						
20		0.093719	-0.05418	-0.16901	0.200851						
21		-0.09912	0.060353	0.200851	-0.18449						
22											
23						Sum Xo	Var o		St Dev o	Exp Ret o	
24	Xo	-0.09037	0.508462	0.848166	-0.26626	1	11.88242		3.447089	9.741096	
25	Xo Initial	0.25	0.25	0.25	0.25	1					
26	Xo Solver	0.323379	0.134765	-0.1	0.641856	1	7.966032		2.822416	10.02114	
27						Sum Abs Xo					
28	Abs Xo	0.323379	0.134765	0.1	0.641856	1.2					
29											
30	(V inv) M										
31		-0.08253	-0.00761								
32		0.476987	0.042791								
33		0.763117	0.07138								
34		-0.33778	-0.02241								
35											
36	M' (V inv) M					Lead Prin Minors					
37		8.124975	0.819791			det (1x1)	8.124975				
38		0.819791	0.084158			det (2x2)	0.011725				
39											
40	[M' (V inv) M] inv					Lead Prin Minors					
41		7.17794	-69.921			det (1x1)	7.17794				
42		-69.921	692.9897			det (2x2)	85.2913				
43											
44						Sum X	Var		St Dev	Exp Ret	One
45	X	-0.16669	1.05204	1.46082	-1.34617	1	23.25829		4.822685	11	1
46	X Initial	0.25	0.25	0.25	0.25	1					
47	X Solver	0.453243	0.340153	-0.1	0.306604	1	13.97939		3.738902	11	
48						Sum Abs X					
49	Abs X	0.453243	0.340153	0.1	0.306604	1.2					
50											
51	Additional Cell Formulas			B28	=ABS(B26)		Pasted to B28:E28 and B49:E49				
52				F28	=SUM(B28:E28)		Pasted to F49				
53	Additional Constraint in Solver for Min Var o						$F$28=1.2				
54	Additional Constraint in Solver for Min Var with Given Exp Re						$F$49=1.2				
55											

Figure 6: An Excel example illustrating the inconsistency of the intended results of portfolio variance minimization and the corresponding Solver results for an invertible covariance matrix of security returns that is not positive definite.

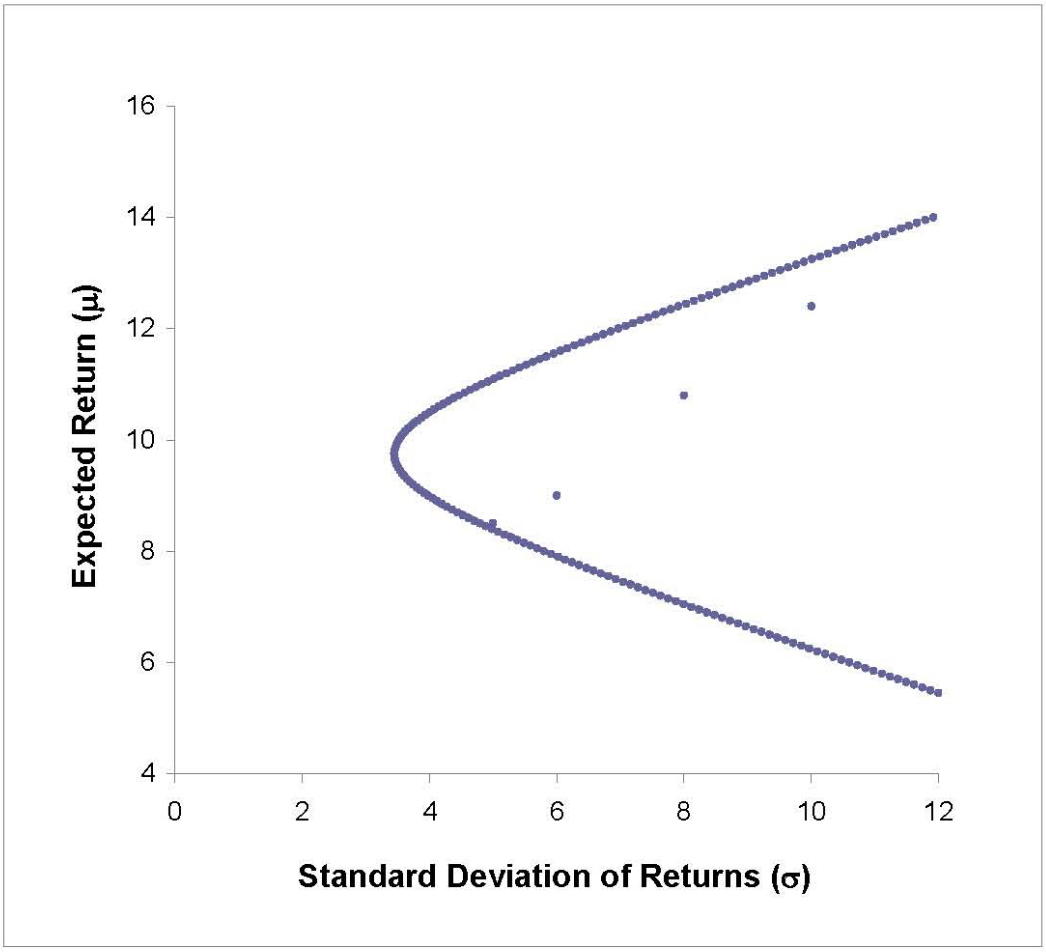

Figure 7: The graph of expected return versus standard deviation of returns for a basic portfolio selection model based on the input parameters in Figure 6.

variances lower than those based on equations (10)-(13). For the illustration, we arbitrarily restrict the sum of the absolute values of the four portfolio weights to be 1.2. Thus, there is an additional constraint in each *Solver* run. The *Solver* results as shown in row 26 of the Excel worksheet in Figure 6 indicate that a different set of portfolio weights can lead to a much lower variance of returns than that according to equation (13). Likewise, as shown in row 47, a different portfolio has a much lower variance of returns than that according to equation (10) for $\mu_p = 11$. Thus, it is clear that, unless the covariance matrix is positive definite, the portfolio results based on equations (10)-(13) do not correspond to minimum variance portfolios.

7 Concluding remarks

In finance courses that are part of the core curriculum of business education, portfolio concepts are typically conveyed in terms of the correlations of security returns. For a given set of risky securities for portfolio investments, it is obvious that all pairwise correlations of security returns must be in the permissible range of -1 to 1. However, standard investment textbooks are silent on whether there are any other requirements on the variances and covariances of security returns.

In contrast, it is well-known in the academic finance literature that, when the variances and covariances are arranged as elements of a symmetric matrix, called the covariance matrix of security returns, such a matrix must satisfy a specific requirement in order to be meaningful. This study has considered the requirement from a pedagogic perspective. In the language of matrix algebra, the requirement is that the covariance matrix be positive definite. In the context of portfolio investments under the assumption of frictionless short sales, a positive definite covariance matrix ensures that the variance of portfolio returns be always positive, regardless of how investment funds are allocated among the securities considered.

The positive definiteness requirement is crucial because its violation will prevent portfolio selection models from performing their intended tasks. With the aid of Excel tools, this study has revealed some consequences of its violation. Of special relevance is that its violation does not preclude the attainment of a hyperbola — which is intended to capture the achievable risk-return trade-off from portfolio investments — with a horizontal transverse axis on the plane of standard deviation of returns (σ) and expected return (μ), where σ is the horizontal axis. As students are taught that the efficient frontier, starting from the global minimum variance portfolio, is a concave curve on the (σ, μ)-plane, violation of the positive definiteness requirement may appear to be inconsequential at first glance. Nevertheless, it is such a situation that can cause serious confusion for students.

This study has shown that, if the positive definiteness requirement is violated, the analytical results will not correspond to constrained minimization of the variance of portfolio returns as intended. That is, some other allocations of investment funds satisfying the same constraints can lead to lower variances of portfolio returns. If so, Excel *Solver*, which is highly flexible in accommodating different optimization settings, is ideal for providing numerical illustrations. This study has provided various *Solver* illustrations, thus helping students understand better the positive definiteness requirement and implications for its violations.

Besides *Solver*, Excel functions for basic matrix operations are also useful for teaching mean-variance portfolio analysis. They include matrix transposition, multiplication, and inversion, as well as finding the determinant of a matrix. To allow the positive definiteness requirement and implications of its violations to be covered effectively in the classroom, we as instructors must ensure that students are completely at ease with basic matrix operations. As matrix operations in Excel are easy to follow, students can focus on relevant analytical issues without being encumbered by the attendant computational chores. This study has illustrated with examples how various Excel features can be used for pedagogic purposes.

In order to make the analytical materials self-contained, this study has included several algebraic proofs that are set at levels accessible to business students with general analytical skills. The materials as presented in this study can be used in different ways, depending on the pedagogic approaches of the investment courses involved. For classes where an informal approach is intended for illustrating the positive definiteness requirement, neither the basic portfolio selection model nor the various proofs involved are essential. For advanced investment classes, however, such proofs — in the form of either lecture materials or exercises for students — can facilitate a better understanding of the basic portfolio selection model by students.

Acknowledgement

The author wishes to thank the anonymous reviewers for helpful comments and suggestions.

References

[1] Cheung, C.S., Kwan, C.C.Y., and Miu, P.C. (2007). "Mean-Gini Portfolio Analysis: A Pedagogic Illustration," *Spreadsheets in Education*, **2**(2), Article 3.

[2] Gilbert, G.T. (1991). "Positive Definite Matrices and Sylvester's Criterion," *American Mathematical Monthly*, **98**(1), 44–46.

[3] Jobson, J.D., and Korkie, B. (1989). "A Performance Interpretation of Multivariate Tests of Asset Set Intersection, Spanning, and Mean-Variance Efficiency," *Journal of Financial and Quantitative Analysis*, **24**(2), 185–204.

[4] Johnson, C.R. (1970). "Positive Definite Matrices," *American Mathematical Monthly*, **77**(3), 259–264.

[5] Kwan, C.C.Y. (2006). "Some Further Analytical Properties of the Constant Correlation Model for Portfolio Selection," *International Journal of Theoretical and Applied Finance,* **9**(7), 1071–1091.

[6] Kwan, C.C.Y. (2007). "A Simple Spreadsheet-Based Exposition of the Markowitz Critical Line Method for Portfolio Selection," *Spreadsheets in Education,* **2**(3), Article 2.

[7] Kwan, C.C.Y. (2008). "Estimation Error in the Average Correlation of Security Returns and Shrinkage Estimation of Covariance and Correlation Matrices," *Finance Research Letters,* **5**, 236–244.

[8] Kwan, C.C.Y. (2009). "Estimation Error in the Correlation of Two Random Variables: A Spreadsheet-Based Exposition," *Spreadsheets in Education,* **3**(2), Article 2.

[9] Ledoit, I., and Wolf, M. (2004). "Honey, I Shrunk the Sample Covariance Matrix," *Journal of Portfolio Management* (Summer), 110–119.

[10] Martin, R.S., Peters, G., and Wilkinson, J.H. (1965). "Symmetric Decomposition of a Positive Definite Matrix," *Numerische Mathematik*, **7**, 362–383.

[11] Merton, R.C. (1972). "An Analytic Derivation of the Efficient Portfolio Frontier," *Journal of Financial and Quantitative Analysis*, **7**(4), 1851–1872.

[12] Roll, R. (1977). "A Critique of the Asset Pricing Theory's Tests, Part I: On Past and Potential Testability of the Theory," *Journal of Financial Economics*, **4**, 129–176.

Appendix A

This appendix shows that the sample covariance matrix is always positive semidefinite. The proof starts with defining R_{it} as the return of security i observed at time t, for $i = 1, 2, \ldots, n$ and $t = 1, 2, \ldots, T$, where T is the number of return observations. Let also

$$\widehat{\sigma}_{ij} = \frac{1}{T-1} \sum\nolimits_{t=1}^{T} \left(R_{it} - \overline{R}_i\right) \left(R_{jt} - \overline{R}_j\right) \tag{A1}$$

be the sample covariance of returns between securities i and j, where $\overline{R}_i$ and $\overline{R}_j$ are the corresponding sample mean returns, for $i, j = 1, 2, \ldots, n$, including $i = j$. The sample covariance matrix $\widehat{\boldsymbol{V}}$ is an $n \times n$ matrix with each element (i, j) being $\widehat{\sigma}_{ij}$.

To show that $\widehat{\boldsymbol{V}}$ is positive semidefinite, let

$$u_{it} = \frac{\left(R_{it} - \overline{R}_i\right)}{\sqrt{T-1}} \tag{A2}$$

be element (i, t) of an $n \times T$ matrix $\boldsymbol{U}$. Accordingly, we have $\widehat{\boldsymbol{V}} = \boldsymbol{U}\boldsymbol{U}'$. The matrix product $\boldsymbol{x}'\widehat{\boldsymbol{V}}\,\boldsymbol{x}$, for any n-element column vector $\boldsymbol{x}$, is therefore $\boldsymbol{x}'\boldsymbol{U}\boldsymbol{U}'\ \boldsymbol{x} = (\boldsymbol{U}'\boldsymbol{x})'(\boldsymbol{U}'\boldsymbol{x})$. Let $\boldsymbol{w} = \boldsymbol{U}'\boldsymbol{x}$ be a T-element column vector and label its elements as $w_1, w_2, \cdots, w_T$. It follows that $\boldsymbol{x}'\widehat{\boldsymbol{V}}\,\boldsymbol{x} = \boldsymbol{w}'\,\boldsymbol{w} = \sum_{t=1}^{T} w_t^2$, which is never negative. With $\boldsymbol{x}$ and, consequently, $\boldsymbol{w}$ being arbitrary, the positive semidefiniteness of the sample covariance matrix is confirmed.

Notice that, if there are insufficient observations (with $T < n$) for the estimation of the covariance matrix, we can write $\widehat{\boldsymbol{V}} = \boldsymbol{U}\,\boldsymbol{U}' = \boldsymbol{Z}\boldsymbol{Z}'$, where $\boldsymbol{Z} = \left[\begin{array}{cc} \boldsymbol{U} & \boldsymbol{0} \end{array}\right]$ is an $n \times n$ matrix formed by appending the $n \times T$ matrix $\boldsymbol{U}$ with an $n \times (n-T)$ matrix $\boldsymbol{0}$ with all zero elements. The determinant of $\widehat{\boldsymbol{V}}$ is zero, as it is the product of the determinants of $\boldsymbol{Z}$ and $\boldsymbol{Z}'$, both of which are zeros. Although $\widehat{\boldsymbol{V}}$ in this situation is not invertible, it is still positive semidefinite. Whether the sample covariance matrix is invertible is irrelevant in the proof of its positive semidefiniteness.

Notice also the following, regardless of whether there are sufficient observations:

1. If $R_{i1} = R_{i2} = \cdots = R_{iT}$, security i is risk-free. As $R_{i1} - \overline{R}_i$, $R_{i2} - \overline{R}_i$, $\ldots$, $R_{iT} - \overline{R}_i$ are zeros, row i of $\boldsymbol{U}$ has all zero elements. A vector $\boldsymbol{x}$ where element i is its only non-zero element will result in $\boldsymbol{w}$ being a vector of zeros.

2. For two risky securities i and j, if $R_{it} = a + bR_{jt}$ for $t = 1, 2, \ldots, T$, where a and b are constants, the returns of the two securities are perfectly correlated. In such a case, knowing one of R_{it} and R_{jt} will enable us to determine the remaining one directly for any t. As $R_{it} - \overline{R}_i = b(R_{jt} - \overline{R}_j)$, a vector $\boldsymbol{x}$ with its only non-zero elements being -1 and b for elements i and j, respectively, will make $\boldsymbol{w}$ consist only of zero elements.

3. If R_{it} as observed at any t can be replicated exactly by the same linear combination of some of the remaining returns among $R_{1t}, R_{2t}, \ldots, R_{nt}$, in the form of

$$R_{it} = a + \sum\nolimits_{k=1}^{i-1} b_k R_{kt} + \sum\nolimits_{k=i+1}^{n} b_k R_{kt}, \tag{A3}$$

where $a, b_1, b_2, \ldots, b_{i-1,} b_{i+1}, b_{i+2}, \ldots, b_n$ are constants, including some zeros, $\boldsymbol{w}$ can be made a vector of zeros as well. The reason is that, with

$$R_{it} - \overline{R}_i = \sum\nolimits_{k=1}^{i-1} b_k (R_{kt} - \overline{R}_k) + \sum\nolimits_{k=i+1}^{n} b_k (R_{kt} - \overline{R}_k), \tag{A4}$$

we can set

$$\boldsymbol{x} = \left[\begin{array}{ccccccccc} b_1 & b_2 & \cdots & b_{i-1} & -1 & b_{i+1} & b_{i+2} & \cdots & b_n \end{array}\right] \tag{A5}$$

to ensure that $\boldsymbol{w}$ be a vector of zeros. Equation (A3) encompasses, as a special case, the situation where security i is a portfolio of some other securities among the n securities; in such a case, we have $a = 0$ and $\sum_{k=1}^{i-1} b_k + \sum_{k=i+1}^{n} b_k = 1$.

Appendix B

For more elegant proofs of Sylvester's criterion, see, for example, Johnson (1970) and Gilbert (1991). The proof below requires only matrix properties that are accessible to business students with general algebraic skills. Let us start with the definition of determinant. The determinant of an $n \times n$ matrix $\boldsymbol{V}$, where each element (i,j) is σ_{ij}, is defined as $\sum(\pm 1)\sigma_{1i_1}\sigma_{2i_2}\cdots\sigma_{ni_n}$, with the summation over all $n!$ permutations of

$i_1, i_2, \ldots, i_n$. The n different integers that $i_1, i_2, \ldots, i_n$ represent can be any permutation of $1, 2, \ldots, n$. If it takes an even number of inversions (interchanges), each involving an adjacent pair of integers in the sequence, to rearrange these integers as $1, 2, \ldots, n$, the sign factor for $\sigma_{1i_1}\sigma_{2i_2}\cdots\sigma_{ni_n}$ is $(+1)$. If it takes an odd number of inversions, the sign factor is (-1) instead.[12] Here, we label the determinant of $\boldsymbol{V}$ as det $\boldsymbol{V}$.

Suppose that we wish to find a parameter λ and a corresponding n-element non-zero column vector $\boldsymbol{x}$ satisfying the equation $\boldsymbol{V}\boldsymbol{x} = \lambda\boldsymbol{x}$. In the language of matrix algebra, λ is an eigenvalue and the corresponding $\boldsymbol{x}$ is an eigenvector. As the equation can be written as $(\boldsymbol{V} - \lambda\boldsymbol{I})\boldsymbol{x} = \boldsymbol{0}$, where $\boldsymbol{I}$ is an $n \times n$ identity matrix and $\boldsymbol{0}$ is an n-element column vector of zeros, the existence of a solution requires that the determinant of $\boldsymbol{V} - \lambda\boldsymbol{I}$ be zero. The equation $\det(\boldsymbol{V} - \lambda\boldsymbol{I}) = 0$ is called the characteristic equation. The idea is that, if $\det(\boldsymbol{V} - \lambda\boldsymbol{I}) \neq 0$, the matrix $\boldsymbol{V} - \lambda\boldsymbol{I}$ is invertible and the resulting vector $\boldsymbol{x} = (\boldsymbol{V} - \lambda\boldsymbol{I})^{-1}\boldsymbol{0}$ will inevitably be a vector of zeros. The characteristic equation being a polynomial equation of degree n, there are n potential values of λ, which need not be all distinct.

As the product of the n diagonal elements of $\boldsymbol{V} - \lambda\boldsymbol{I}$, which is $(\sigma_{11} - \lambda)(\sigma_{22} - \lambda)\cdots(\sigma_{nn} - \lambda)$, provides the highest power of λ in the expression of $\det(\boldsymbol{V} - \lambda\boldsymbol{I})$, we can write the expression as a polynomial function $P(\lambda)$. The coefficient of λ^n, the highest power term in $P(\lambda)$, is $(-1)^n$. The existence of n eigenvalues of λ, labeled as λ_i, for $i = 1, 2, \ldots, n$, implies that

$$P(\lambda) = (-1)^n(\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_n) = 0. \tag{B1}$$

The value of this polynomial function at $\lambda = 0$,

$$P(0) = (-1)^n(-\lambda_1)(-\lambda_2)\cdots(-\lambda_n) = \lambda_1\lambda_2\cdots\lambda_n, \tag{B2}$$

is also det $\boldsymbol{V}$. Thus, the determinant of an $n \times n$ matrix is the product of its n eigenvalues.

A crucial matrix property as required for proving Sylvester's criterion is that all eigenvalues of a real symmetric matrix are real. To prove this property, let us assume that, contrary to the assertion, each λ is of the form $a + b\sqrt{-1}$, where both a and b are real. The complex conjugate of $a + b\sqrt{-1}$ is $a - b\sqrt{-1}$, by changing the sign of the imaginary part of the expression. Notice that $(\sqrt{-1})^2 = -1$ and that $\overline{(a + b\sqrt{-1})(c + d\sqrt{-1})} = \overline{(a + b\sqrt{-1})}\ \overline{(c + d\sqrt{-1})}$, where c and d are also real and each algebraic expression with a bar above it represents its complex conjugate. Thus, for any specific eigenvalue λ, we can write the equation $\boldsymbol{V}\boldsymbol{x} = \lambda\boldsymbol{x}$ as $\boldsymbol{V}\overline{\boldsymbol{x}} = \overline{\lambda}\overline{\boldsymbol{x}}$, by taking the complex conjugates of both sides of the equation. Suppose that the corresponding eigenvector $\boldsymbol{x}$ is a non-zero vector, with each element written generally in the form of $a + b\sqrt{-1}$. We can also write $\overline{\boldsymbol{x}}'\boldsymbol{x} = (\overline{\boldsymbol{x}}'\boldsymbol{x})' = \boldsymbol{x}'\overline{\boldsymbol{x}} > 0$, because the three matrix products here, each being the sum of n positive terms of the form $a^2 + b^2$ per term, will all result in the same scalar.

[12] For example, in the case of $n = 3$, the sequence $3, 2, 1$ requires three inversions to reach $1, 2, 3$. The inversions can be first between 3 and 2 (to reach $2, 3, 1$); then between 3 and 1 in the revised sequence (to reach $2, 1, 3$); and finally between 2 and 1 (to reach $1, 2, 3$). In this example, as the number of inversions is odd, the sign factor is (-1).

Combining $\overline{\boldsymbol{x}}'\boldsymbol{V}\boldsymbol{x} = \lambda\overline{\boldsymbol{x}}'\boldsymbol{x}$ and $\boldsymbol{x}'\boldsymbol{V}\overline{\boldsymbol{x}} = \overline{\lambda}\boldsymbol{x}'\overline{\boldsymbol{x}}$ leads to

$$\overline{\boldsymbol{x}}'\boldsymbol{V}\boldsymbol{x} - \boldsymbol{x}'\boldsymbol{V}\overline{\boldsymbol{x}} = (\lambda - \overline{\lambda})\overline{\boldsymbol{x}}'\boldsymbol{x}. \tag{B3}$$

As $\boldsymbol{V}$ is symmetric, the left hand side of equation (B3) is zero. Thus, we must have $\lambda = \overline{\lambda}$; that is, each eigenvalue of $\boldsymbol{V}$ must be real. Notice that, with all elements of the matrix $\boldsymbol{V} - \lambda\boldsymbol{I}$ being real for each specific eigenvalue λ, we can choose the corresponding eigenvector $\boldsymbol{x}$ that satisfies the equation $(\boldsymbol{V} - \lambda\boldsymbol{I})\boldsymbol{x} = \boldsymbol{0}$ to be a real vector.

To prove the part of Sylvester's criterion that, if $\boldsymbol{V}$ is positive definite, its leading principal minors are all positive, we first draw on the result that $\boldsymbol{x}'\boldsymbol{V}\boldsymbol{x} = \lambda\boldsymbol{x}'\boldsymbol{x}$ is positive for any real eigenvector $\boldsymbol{x}$ associated with each specific eigenvalue λ. With $\boldsymbol{x}'\boldsymbol{x}$ being positive, it follows that the corresponding λ is positive. As the determinant of a matrix is the product of its eigenvalues according to equation (B2), $\det \boldsymbol{V}$, which is the n-th leading principal minor of $\boldsymbol{V}$, is positive as well. Further, the $n \times n$ matrix $\boldsymbol{V}$ being positive definite, the scalar $\boldsymbol{x}'\boldsymbol{V}\boldsymbol{x}$ for any n-element non-zero column vector $\boldsymbol{x}$ is positive. For $m = 1, 2, \ldots, n-1$, if the last $n - m$ elements of $\boldsymbol{x}$ are set to be zeros, the condition of $\boldsymbol{x}'\boldsymbol{V}\boldsymbol{x} > 0$ is equivalent to $\boldsymbol{x}_m'\boldsymbol{V}_m\boldsymbol{x}_m > 0$, where $\boldsymbol{V}_m$ is the m-th leading principal submatrix of $\boldsymbol{V}$ and $\boldsymbol{x}_m$ is an arbitrary m-element non-zero column vector. With each $\boldsymbol{V}_m$ being positive definite, it follows that the corresponding $\det \boldsymbol{V}_m$ is positive.

To prove, by induction, the part of Sylvester's criterion that, if all leading principal minors of $\boldsymbol{V}$ are positive, $\boldsymbol{V}$ is positive definite, we start with the first leading principal submatrix of $\boldsymbol{V}$, labeled as $\boldsymbol{V}_1$, which consists of a single element, σ_{11}. Obviously, if $\sigma_{11} > 0$, any single-element non-zero vector $\boldsymbol{x}_1$ will give us $\boldsymbol{x}_1'\boldsymbol{V}_1\boldsymbol{x}_1 > 0$, confirming that $\boldsymbol{V}_1$ is positive definite. Now, suppose that, by inductive hypothesis, $\det \boldsymbol{V}_1$, $\det \boldsymbol{V}_2$, $\ldots$, $\det \boldsymbol{V}_n$ are all positive and $\boldsymbol{V}_n$ is positive definite. The task now is to show that, if $\det \boldsymbol{V}_{n+1}$ is positive, $\boldsymbol{V}_{n+1}$ is positive definite.

Letting $\boldsymbol{v} = \begin{bmatrix} \sigma_{1,n+1} & \sigma_{2,n+1} & \cdots & \sigma_{n,n+1} \end{bmatrix}'$, an n-element column vector consisting of the first n elements of the last column of $\boldsymbol{V}_{n+1}$, we can write $\boldsymbol{V}_{n+1} = \boldsymbol{Q}'\boldsymbol{B}\boldsymbol{Q}$, a product of three $(n+1) \times (n+1)$ matrices. Here, $\boldsymbol{B}$ is constructed by augmenting the $n \times n$ matrix $\boldsymbol{V}_n$ with both a row $n+1$ and a column $n+1$ of zeros, except for element $(n+1, n+1)$, which is an unspecified parameter g. The matrix $\boldsymbol{Q}$ is constructed by substituting the first n elements of the last column of an $(n+1) \times (n+1)$ identity matrix with n elements of the column vector $\boldsymbol{V}_n^{-1}\boldsymbol{v}$. As it will soon be clear, although g can be solved in terms of the elements of $\boldsymbol{V}_{n+1}$, there is no need to do so for the purpose here. We can write $\det \boldsymbol{V}_{n+1} = (\det \boldsymbol{Q}')(\det \boldsymbol{B})(\det \boldsymbol{Q}) = \det \boldsymbol{B}$, as $\det \boldsymbol{Q} = \det \boldsymbol{Q}' = 1$. Given that $\det \boldsymbol{V}_{n+1} > 0$ by inductive hypothesis, we also have $\det \boldsymbol{B} > 0$. With $\det \boldsymbol{B} = g(\det \boldsymbol{V}_n)$ and $\det \boldsymbol{V}_n > 0$, we have $g > 0$ as well.

For any $(n+1)$-element non-zero column vector $\boldsymbol{x}_{n+1}$, we can write

$$\boldsymbol{x}_{n+1}'\boldsymbol{V}_{n+1}\boldsymbol{x}_{n+1} = (\boldsymbol{Q}\boldsymbol{x}_{n+1})'\boldsymbol{B}(\boldsymbol{Q}\boldsymbol{x}_{n+1}), \tag{B4}$$

where $\boldsymbol{Q}\boldsymbol{x}_{n+1}$ is an $(n+1)$-element column vector. Let us label the vector consisting of the first n elements of $\boldsymbol{Q}\boldsymbol{x}_{n+1}$ as $\boldsymbol{y}_n$ and the last element as h. Equation (B4) now becomes

$$\boldsymbol{x}_{n+1}'\boldsymbol{V}_{n+1}\boldsymbol{x}_{n+1} = \boldsymbol{y}_n'\boldsymbol{V}_n\boldsymbol{y}_n + gh^2. \tag{B5}$$

As $\boldsymbol{x}_{n+1}$ is arbitrary, so is $\boldsymbol{y}_n$. With the right hand side of equation (B5) being positive, the positive definiteness of $\boldsymbol{V}_{n+1}$ is confirmed. This completes the inductive proof.

Appendix C

This appendix shows that a symmetric positive definite matrix can be written as the product of a square matrix and its transpose. To do so, consider an $n \times n$ symmetric positive-definite matrix $\boldsymbol{A}$. As indicated in Appendix B, if we write $\boldsymbol{As} = \lambda \boldsymbol{s}$, λ is an eigenvalue and $\boldsymbol{s}$ is a corresponding eigenvector. There are n (not necessarily all distinct) values of λ. With $\boldsymbol{A}$ being positive definite and thus $\boldsymbol{s}'\boldsymbol{As} = \lambda \boldsymbol{s}'\boldsymbol{s}$ being positive, it follows that each of the n values of λ is positive. We can choose the n corresponding eigenvectors to be orthogonal unit vectors. That is, we can scale the individual vectors to ensure that the sum of the squares of the elements of each vector be unity and that the sum of the products of the corresponding elements of any two vectors be zero.

Let $\boldsymbol{\Lambda}$ be an $n \times n$ diagonal matrix with its diagonal elements being the n eigenvectors. Let also $\boldsymbol{S}$ be an $n \times n$ matrix containing these n eigenvectors as columns in the same order. With $\boldsymbol{S}'\boldsymbol{S} = \boldsymbol{I}$ by construction, it follows that $\boldsymbol{AS} = \boldsymbol{S\Lambda}$ and $\boldsymbol{A} = \boldsymbol{ASS}' = \boldsymbol{S\Lambda S}'$. This decomposition of a symmetric matrix is often called spectral decomposition. As the diagonal elements of $\boldsymbol{\Lambda}$ are all positive, we can write $\boldsymbol{\Lambda} = \boldsymbol{DD}'$, where $\boldsymbol{D}$ is an $n \times n$ diagonal matrix with each diagonal element being the square root of the the corresponding eigenvalue in $\boldsymbol{\Lambda}$. Thus, we can also write $\boldsymbol{A} = \boldsymbol{SDD}'\boldsymbol{S}' = (\boldsymbol{SD})(\boldsymbol{SD})'$, which is the product of a square matrix and its transpose.

Notice that there is a familiar matrix property, called Cholesky decomposition, which allows each symmetric positive definite matrix to be written as the product of a lower triangular matrix and its transpose. A lower triangular matrix is a square matrix with all elements above its diagonal being zeros. See, for example, Martin, Peters, and Wilkinson (1965), for the algorithmic detail.

Appendix D

This appendix provides a simple algebraic proof of Cauchy-Schwarz inequality. We start with

$$(u_1 z_2 - u_2 z_1)^2 \geq 0, \tag{D1}$$

for any real u_1 and u_2 and for any $z_1 > 0$ and $z_2 > 0$. The inequality can be written as

$$2u_1 u_2 \leq \frac{u_1^2 z_2}{z_1} + \frac{u_2^2 z_1}{z_2}. \tag{D2}$$

After adding $u_1^2 + u_2^2$ to both sides and dividing the resulting expressions by $z_1 + z_2$, we obtain

$$\frac{(u_1 + u_2)^2}{z_1 + z_2} \leq \frac{u_1^2}{z_1} + \frac{u_2^2}{z_2}. \tag{D3}$$

Substituting $u_2 + u_3$ for u_2, substituting $z_2 + z_3$ (where $z_3 > 0$) for z_2, and noting the analogous inequality

$$\frac{(u_2 + u_3)^2}{z_2 + z_3} \leq \frac{u_2^2}{z_2} + \frac{u_3^2}{z_3}, \tag{D4}$$

we have

$$\frac{(u_1 + u_2 + u_3)^2}{z_1 + z_2 + z_3} \leq \frac{u_1^2}{z_1} + \frac{u_2^2}{z_2} + \frac{u_3^2}{z_3}. \tag{D5}$$

Further substitutions in the same manner will eventually lead to

$$\frac{(\sum_{i=1}^m u_i)^2}{\sum_{i=1}^m z_i} \leq \sum\nolimits_{i=1}^m \frac{u_i^2}{z_i}. \tag{D6}$$

Letting $u_i = a_i b_i$ and $z_i = b_i^2$, we have

$$\left(\sum\nolimits_{i=1}^m a_i b_i\right)^2 \leq \left(\sum\nolimits_{i=1}^m a_i^2\right)\left(\sum\nolimits_{i=1}^m b_i^2\right). \tag{D7}$$

Notice that, with each z_i being positive, the corresponding b_i is non-zero. To accommodate cases where $b_i = 0$, if any, we can label them as $b_{m+1}, b_{m+2}, \ldots, b_n$, along with the corresponding $a_{m+1}, a_{m+2}, \ldots, a_n$. With

$$\left(\sum\nolimits_{i=1}^m a_i b_i\right)^2 = \left(\sum\nolimits_{i=1}^n a_i b_i\right)^2 \tag{D8}$$

$$\text{and } \left(\sum\nolimits_{i=1}^m a_i^2\right)\left(\sum\nolimits_{i=1}^m b_i^2\right) \leq \left(\sum\nolimits_{i=1}^n a_i^2\right)\left(\sum\nolimits_{i=1}^n b_i^2\right), \tag{D9}$$

it follows that

$$\left(\sum\nolimits_{i=1}^n a_i b_i\right)^2 \leq \left(\sum\nolimits_{i=1}^n a_i^2\right)\left(\sum\nolimits_{i=1}^n b_i^2\right). \tag{D10}$$

Strict inequality holds if $a_1, a_2, \ldots, a_n$ cannot be duplicated exactly by $cb_1, cb_2, \ldots, cb_n$ for any constant c.