Spreadsheets in Education (eJSiE)

Volume 7 | Issue 1 Article 3

January 2014

A Regression-Based Interpretation of the Inverse of the Sample Covariance Matrix

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

A Regression-Based Interpretation of the Inverse of the Sample Covariance Matrix

Abstract

The usefulness of covariance and correlation matrices is well-known in various academic fields. Matrix inversion, if required in an analytical setting, tends to mask the intuition in interpreting the corresponding empirical or experimental results. Drawing on the finance literature in mean-variance portfolio analysis, this paper presents pedagogically a regression-based interpretation of the inverse of the sample covariance matrix. Microsoft Excel^{TM} plays an important pedagogic role in this paper. The availability of various Excel functions and computational tools for numerical illustrations provides flexibility for instructors in the delivery of the corresponding analytical materials.

Keywords

matrix inversion, sample covariance matrix, multiple linear regression

Cover Page Footnote

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

A Regression-Based Interpretation of the Inverse of the Sample Covariance Matrix

1 Introduction

For a given set of real-valued random variables, the corresponding covariance matrix, which is symmetric, contains all their variances and covariances. If the variables involved are scaled to have unit variances, the resulting matrix is a correlation matrix. The usefulness of covariance and correlation matrices is well-known across different academic disciplines. Very often, matrix inversion is involved. In finance, for example, the covariance matrix of security returns is part of the input parameters for portfolio analysis, and efficient allocations of investment funds require the inversion of such a matrix.[1]

By definition, a square matrix is invertible if there exists another square matrix of the same order, such that their product is an identity matrix. Analytically, the inverse of a square matrix can be expressed as its adjoint divided by its determinant. Thus, for a covariance matrix of order higher than three, unless the matrix is characterized to have a specific structure, to express explicitly each element of the matrix inverse in terms of the original matrix elements is a very tedious task.[2]

From a pedagogic perspective, if the analytical solution of an optimization problem requires the inversion of the covariance matrix, of interest is whether we can still go beyond the technical aspect of the task involved, by also explaining the analytical solution intuitively. To explain any analytical solutions involving matrix inversion, an explicit connection between corresponding elements of the matrix and its inverse is required. This paper is intended to provide such a

[1] See, for example, Kwan (2011) for some recent references pertaining to the use of covariance and correlation matrices, as well as their inverses, for various empirical and experimental investigations. See also Elton, Gruber, Brown, and Goetzmann (2010, Chapter 6) for a basic portfolio selection model. In that model, efficient allocations of investment funds are achieved by solving a set of linear equations, which contains elements of the covariance matrix of security returns. This is analytically equivalent to using the inverse of the covariance matrix for portfolio decisions.

[2] See, for example, Reiner (1971, Chapters 5 and 6) and Larson and Falvo (2009, Chapters 2 and 3) for definitions of various terms in matrix algebra and for basic matrix operations. See also Elton, Gruber, Brown, and Goetzmann (2010, Chapters 7-9) and Kwan (1984) for some characterizations of the covariance matrix of security returns. They include constant correlation models and various index models. Such characterizations enable the elements of each inverse matrix to be expressed explicitly, thus allowing the corresponding portfolio allocation results to be explained intuitively.

connection without having to specify a covariance structure.

In most empirical and experimental settings, as the true covariance matrix is unknown, it must be estimated with available observations. A sample covariance matrix is a covariance matrix that is estimated under the assumption of a stationary (stable) joint probability distribution of the variables considered. Under such an assumption, each sample observation can be viewed as a random draw from a given distribution. It is the stationarity assumption that allows a direct connection to be made between the inverse of the sample covariance matrix and multiple linear regression.

In a setting of portfolio analysis, Stevens (1998) has provided an innovative interpretation of the inverse of the covariance matrix of security returns. The interpretation is based on multiple linear regression models, in which the random return of each security is characterized as depending linearly on the random returns of all remaining securities considered.[3] The specific regression results that are relevant for the interpretation are the ordinary least squares (OLS) regression coefficients and the coefficient of determination, which is commonly known as R^2.

Notice that such results, unlike the standard errors of the regression coefficients (for testing their statistical significance, for example) and various other usual regression outputs, do not require any specific assumption on the underlying joint probability distribution of the variables involved.[4] Notice also that, although Stevens' interpretation pertains to a financial setting, the underlying idea is general enough to accommodate other settings. Such flexibility is important for a general interpretation of the inverse of a given sample covariance matrix.

How Stevens' interpretation can be covered effectively in the classroom, however, depends on prior mathematical and statistical knowledge of the students involved. At one extreme, students who are already familiar with both block matrix inversion and OLS regression in matrix notation

[3]In portfolio analysis that uses the sample covariance matrix of security returns as part of the input parameters, whether the matrix is invertible can be explained in terms of some return characteristics of the securities considered. If one of the securities considered is risk-free, the corresponding sample covariance matrix is not invertible. Neither is it invertible if the random returns of any two securities are perfectly correlated; nor is it invertible if the random return of any security can be replicated exactly by a portfolio of some other securities considered. All these analytical features can be detected in the multiple linear regression models under Stevens' interpretation.

[4]In regression analysis, justification for the OLS approach does require various assumptions on the probability distributions of the variables involved. However, the issue here is not about whether better parameter estimates can be achieved by using other estimation methods. Rather, it is about whether the OLS approach, which is the simplest method for regression analysis, can provide results that enable us to interpret the inverse of the sample covariance matrix.

will likely find Stevens' analytical result obvious. This is because there are close similarities between the analytical expressions of the inverse of a block matrix and some well-known OLS regression results, when a sample covariance matrix is partitioned in a specific way. At the other extreme, where students' mathematical and statistical preparedness is deemed inadequate, a less formal approach is better for presenting Stevens' interpretation. Between the two extremes is a range of situations, for which different pedagogic approaches are warranted. This paper is intended to provide, in considerable detail, the analytical materials involved for individual instructors to choose for their classroom coverage.

On the computational side, Microsoft Excel^{TM} plays an important pedagogic role in this paper.[5] Excel has various built-in functions for matrix operations. Functions such as TRANSPOSE and MMULT for matrix transposition and multiplication, respectively, are particularly useful for deducing the sample covariance matrix from a set of mean-removed observations of the variables considered. The function MINVERSE can then be used to provide the corresponding matrix inverse directly.

For the purpose of interpreting the inverse of the sample covariance matrix, Excel can be utilized in different ways to deduce the required regression results. For courses where analytical details pertaining to matrix inversion and its interpretation are covered, the corresponding regression analysis is best performed via matrix operations in Excel. For such courses, the Excel illustration in this paper is not intended to be a substitute for actually covering the corresponding analytical materials. Rather, by using matrix operations in Excel to bypass the tedious computational task, students will be able to focus on the analytical details and the corresponding concepts. In this regard, Excel can perform a similar pedagogical role as that of various statistical and computational software packages accompanying senior-level courses, such as quantitative finance, portfolio theory, and econometric methods, where analytical details of the covered topics are emphasized.

For courses where the underlying principle of regression analysis is covered without using matrix algebra, then Excel Solver is ideal for finding the regression coefficients involved. Given the regression coefficients, the corresponding R^2 can easily be computed in Excel. For courses where how regression results are generated is not a concern, the Excel function LINEST is

[5] For the rest of this paper, whenever the name Excel or any of its tools is mentioned, its trademark is implicitly recognized.

suitable for the task, as it provides the regression coefficients, R^2, and various other standard regression results. The regression tool, which is among Excel's various add-in tools for data analysis and provides more detailed regression results than LINEST does, can be used as well. The availability of these alternatives will give instructors flexibility in the Excel illustration of the analytical concepts involved.

The remainder of this paper is organized as follows: Standard textbook materials on matrix inversion are briefly covered in Section 2, in order to show the analytical challenge in connecting directly elements of a square matrix and elements of its inverse. Also covered in Section 2 is the end result of the derivation in Sections 3 and 4 pertaining to a regression-based interpretation of the inverse of the sample covariance matrix. The derivation in Sections 3 and 4 is presented pedagogically in considerable detail, except for materials that are readily accessible elsewhere for most readers.

The derivation in Section 3 is for a two-variable case, where a regression-based interpretation of the inverse of the sample covariance matrix is established without reliance on matrix algebra. A multivariate extension, which requires matrix operations, is provided in Section 4. A simple model in portfolio analysis, which focuses on the inverse of the covariance matrix of security returns, is presented in Section 5; this finance model illustrates the usefulness of the regression-based interpretation in an analytical setting. An numerical example, which illustrates how various Excel functions and computational tools can be utilized pedagogically for the topic of this paper, is presented in Section 6. Finally, Section 7 provides some concluding remarks.

2 Matrix Inversion

According to standard textbook materials in matrix algebra, if the inverse of a square matrix $\boldsymbol{V}$ exists, it can be written as

$$\boldsymbol{V}^{-1} = \frac{1}{\det \boldsymbol{V}} \operatorname{adj} \boldsymbol{V}. \tag{1}$$

Here, $\det \boldsymbol{V}$ is the determinant of $\boldsymbol{V}$ and $\operatorname{adj} \boldsymbol{V}$ is the adjoint of $\boldsymbol{V}$, which is the transpose of the matrix of cofactors of $\boldsymbol{V}$. The cofactor of the (i, j)-element of $\boldsymbol{V}$, labeled as C_{ij}, is $(-1)^{i+j}$ times the determinant of a square matrix of order $n-1$, which is the original $\boldsymbol{V}$ with its row i

and column j removed.[6] Notice that the square matrix $\boldsymbol{V}$ here need not be symmetric. As

$$\boldsymbol{V}^{-1} = \frac{1}{\det \boldsymbol{V}} \begin{bmatrix} C_{11} & C_{21} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ C_{1n} & C_{2n} & \cdots & C_{nn} \end{bmatrix}, \tag{2}$$

each (i, j)-element of $\boldsymbol{V}^{-1}$ is $C_{ji}/\det \boldsymbol{V}$.

To express explicitly each (i, j)-element of $\boldsymbol{V}^{-1}$ in terms of elements of $\boldsymbol{V}$ requires that the determinant of a square matrix be defined. If s_{ij} is the (i, j)-element of $\boldsymbol{V}$, the general definition of its determinant is

$$\det \boldsymbol{V} = \sum (\pm 1) s_{1i_1} s_{2i_2} \cdots s_{ni_n}, \tag{3}$$

where the summation is 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 interchanges to rearrange these integers as $1, 2, \ldots, n$, which involve an adjacent pair of integers for each interchange, then we use a positive multiplicative factor $(+1)$ for $s_{1i_1} s_{2i_2} \cdots s_{ni_n}$. If it takes an odd number of interchanges, then we use a negative multiplicative factor (-1) instead.

For an illustration, let us consider the case of $n = 3$, where there are $3!$ $(= 6)$ permutations of $1, 2, 3$. Specifically, let us consider the term $s_{13} s_{21} s_{32}$ in a summation of six terms. To rearrange $3, 1, 2$ as $1, 2, 3$ requires two interchanges. The first interchange (which is between 3 and 1) transforms $3, 1, 2$ into $1, 3, 2$, and the second interchange (which is between 3 and 2) transforms $1, 3, 2$ into $1, 2, 3$. As the number of interchanges is even, the corresponding multiplicative factor is $+1$.

There is an alternative expression of the determinant, which is also based on its definition in equation (3). It is commonly known as Laplace's expansion, and its expression is

$$\det \boldsymbol{V} = \sum\nolimits_{j=1}^{n} s_{ij} C_{ij} = \sum\nolimits_{i=1}^{n} s_{ij} C_{ij}. \tag{4}$$

The computational convenience of this expression notwithstanding, a direct connection between elements of $\boldsymbol{V}$ and $\boldsymbol{V}^{-1}$ remains elusive for $n > 3$.

[6] See, for example, Reiner (1971, Chapters 5 and 6) and Larson and Falvo (2009, Chapters 2 and 3) for analytical details.

2.1 A Regression-Based Version of the Inverse

To facilitate an interpretation of $\boldsymbol{V}^{-1}$, suppose that a symmetric matrix $\boldsymbol{V}$ is an $n \times n$ sample covariance matrix based on a set of m observations of the variables $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_n$. The observations are represented by $(X_{1k}, X_{2k}, \ldots, X_{nk})$, for $k = 1, 2, \ldots, m$, where the first and second subscripts are variable and observation labels, respectively. Letting

$$\overline{X}_i = \frac{1}{m} \sum\nolimits_{k=1}^{m} X_{ik}, \tag{5}$$

for $i = 1, 2, \ldots, n$, be the n sample means, we have the corresponding mean-removed variables $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$, along with a set of mean-removed observations $(x_{1k}, x_{2k}, \ldots, x_{nk})$, where

$$x_{ik} = X_{ik} - \overline{X}_i, \tag{6}$$

for $i = 1, 2, \ldots, n$ and $k = 1, 2, \ldots, m$. Each (i, j)-element of $\boldsymbol{V}$ is

$$s_{ij} = \frac{1}{m-1} \sum\nolimits_{k=1}^{m} x_{ik} x_{jk}, \tag{7}$$

for $i, j = 1, 2, \ldots, n$.

The inverse of the sample covariance matrix $\boldsymbol{V}$, as derived in Sections 3 and 4, is

$$\boldsymbol{V}^{-1} = \begin{bmatrix} 1/[s_{11}(1-R_1^2)] & -\widehat{\beta}_{12}/[s_{11}(1-R_1^2)] & \cdots & -\widehat{\beta}_{1n}/[s_{11}(1-R_1^2)] \\ -\widehat{\beta}_{21}/[s_{22}(1-R_2^2)] & 1/[s_{22}(1-R_2^2)] & \cdots & -\widehat{\beta}_{2n}/[s_{22}(1-R_2^2)] \\ \vdots & \vdots & \ddots & \vdots \\ -\widehat{\beta}_{n1}/[s_{nn}(1-R_n^2)] & -\widehat{\beta}_{n2}/[s_{nn}(1-R_n^2)] & \cdots & 1/[s_{nn}(1-R_n^2)] \end{bmatrix}. \tag{8}$$

Here, when $\boldsymbol{x}_i$ as a dependent variable is regressed on the remaining variables in a multiple linear regression, each $\widehat{\beta}_{ij}$ represents the regression coefficient for independent variable $\boldsymbol{x}_j$, and R_i^2 is the corresponding coefficient of determination. Each diagonal (i, i)-element is $1/[s_{ii}(1 - R_i^2)]$; each off-diagonal (i, j)-element, where $i \neq j$, is $-\widehat{\beta}_{ij}/[s_{ii}(1 - R_i^2)]$. Equation (8) is the key analytical result in Stevens (1998); such a result is important because it allows each element of the inverse of the sample covariance matrix to be interpreted in terms of some OLS regression results.

3 A Two-Variable Case

For a pair of random variables, $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$, suppose that m observations, labeled as (X_{11}, X_{21}), $(X_{12}, X_{22}), \ldots, (X_{1m}, X_{2m})$, are available. With $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ being a pair of mean-removed variables, we have m observations, $(x_{11}, x_{21}), (x_{12}, x_{22}), \ldots, (x_{1m}, x_{2m})$. The sample variances of $\boldsymbol{x}_1$

and $\boldsymbol{x}_2$, labeled as s_{11} and s_{22}, respectively, and the sample covariance of $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$, labeled as either s_{12} or s_{21}, are given by equation (7), where each of i and j can be 1 or 2.

The corresponding 2×2 sample covariance matrix, which is symmetric, is given by

$$\boldsymbol{V} = \begin{bmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{bmatrix}, \tag{9}$$

where $s_{12} = s_{21}$. To find the inverse of $\boldsymbol{V}$, let us label it as

$$\boldsymbol{V}^{-1} = \boldsymbol{U} = \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix}. \tag{10}$$

As the direct multiplication of $\boldsymbol{U}$ and $\boldsymbol{V}$ must result in a 2×2 identity matrix, the diagonal and off-diagonal elements of $\boldsymbol{U}$ are, respectively,

$$u_{ii} = \frac{s_{jj}}{s_{ii}s_{jj} - s_{ij}^2} \tag{11}$$

and

$$u_{ij} = u_{ji} = -\frac{s_{ij}}{s_{ii}s_{jj} - s_{ij}^2}, \tag{12}$$

where each of i and j can be 1 or 2, but $i \neq j$.

3.1 Linear Regression

To facilitate an interpretation of $\boldsymbol{U}$, we consider two linear regression models based on $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$, to be fitted by the m mean-removed observations, $(x_{11}, x_{21}), (x_{12}, x_{22}), \ldots, (x_{1m}, x_{2m})$. The roles of $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ as independent and dependent variables are reversed for the two models. Given the extensive textbook coverage of linear regression, only the part pertaining to our specific task is covered below.[7]

Suppose that $\boldsymbol{x}_i$ is the independent variable and $\boldsymbol{x}_j$ is the dependent variable, where $i \neq j$. That is, if $i = 1$, we have $j = 2$; if $i = 2$, we have $j = 1$ instead. Each regression model is

$$\boldsymbol{x}_j = \beta_j \boldsymbol{x}_i + \boldsymbol{e}_j, \tag{13}$$

where β_j is an unknown parameter to be estimated and $\boldsymbol{e}_j$ is random noise.

If $\widehat{\beta}_j$ is an estimated value of β_j, then each

$$e_{jk} = x_{jk} - \widehat{\beta}_j x_{ik}, \tag{14}$$

[7] See, for example, DeCoursey (2003, Chapter 14), Johnston (1972, Chapter 2), and Maddala (1977, Chapter 7) for textbook coverage of linear regression.

which represents the part of x_{jk} that the regression model is unable to capture, for $k = 1, 2, \ldots, m$, can be viewed as the m realizations of $\boldsymbol{e}_j$. Given that each of $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ has a zero mean, so does $\boldsymbol{e}_j$. This is because

$$\frac{1}{m}\sum\nolimits_{k=1}^{m} e_{jk} = \frac{1}{m}\left(\sum\nolimits_{k=1}^{m} x_{jk} - \widehat{\beta}_j \sum\nolimits_{k=1}^{m} x_{ik}\right) = 0, \tag{15}$$

regardless of the value of $\widehat{\beta}_j$.

Using the OLS approach, we seek a $\widehat{\beta}_j$ that minimizes

$$\sum\nolimits_{k=1}^{m} e_{jk}^2 = \sum\nolimits_{k=1}^{m} (x_{jk} - \widehat{\beta}_j x_{ik})^2 = \sum\nolimits_{i=1}^{m} x_{jk}^2 - 2\widehat{\beta}_j \sum\nolimits_{k=1}^{m} x_{jk}x_{ik} + \widehat{\beta}_j^2 \sum\nolimits_{k=1}^{m} x_{ik}^2. \tag{16}$$

It follows from

$$\frac{d}{d\widehat{\beta}_j}\sum\nolimits_{k=1}^{m} e_{jk}^2 = -2\sum\nolimits_{k=1}^{m} x_{jk}x_{ik} + 2\widehat{\beta}_j \sum\nolimits_{k=1}^{m} x_{ik}^2 = 0 \tag{17}$$

that the optimal result is

$$\widehat{\beta}_j = \frac{\sum_{k=1}^{m} x_{jk}x_{ik}}{\sum_{k=1}^{m} x_{ik}^2} = \frac{s_{ij}}{s_{ii}}. \tag{18}$$

For this value of $\widehat{\beta}_j$, equation (16) reduces to

$$\begin{aligned} \sum\nolimits_{k=1}^{m} e_{jk}^2 &= \sum\nolimits_{i=1}^{m} x_{jk}^2 - 2\widehat{\beta}_j \sum\nolimits_{k=1}^{m} x_{jk}x_{ik} + \frac{\sum_{k=1}^{m} x_{jk}x_{ik}}{\sum_{k=1}^{m} x_{ik}^2}\widehat{\beta}_j \sum\nolimits_{k=1}^{m} x_{ik}^2 \\ &= \sum\nolimits_{i=1}^{m} x_{jk}^2 - \widehat{\beta}_j \sum\nolimits_{k=1}^{m} x_{jk}x_{ik}. \end{aligned} \tag{19}$$

Notice that equation (16) is of the algebraic form $h = a\widehat{\beta}_j^2 - 2b\widehat{\beta}_j + c$, where h is to be minimized. Here, a, b, and c are of known values. As a is always positive, we can write $a(\widehat{\beta}_j - b/a)^2 + c - b^2/a = h$, by completing the squares. This simple algebraic approach allows us to bypass differential calculus for confirming that $\widehat{\beta}_j = b/a$ minimizes h. This expression of $\widehat{\beta}_j$ is the same as that in equation (18).

3.2 The Coefficient of Determination

Letting

$$\widehat{x}_{jk} = \widehat{\beta}_j x_{ik} \tag{20}$$

be the fitted value of x_{jk} according to the above linear regression model, for $k = 1, 2, \ldots, m$, we can write equation (14) equivalently as

$$x_{jk} = \widehat{x}_{jk} + e_{jk}. \tag{21}$$

Taking the sum of squares of each side leads to

$$\sum_{k=1}^{m} x_{jk}^2 = \sum_{k=1}^{m} (\widehat{x}_{jk} + e_{jk})^2 = \sum_{k=1}^{m} \widehat{x}_{jk}^2 + 2\sum_{k=1}^{m} \widehat{x}_{jk} e_{jk} + \sum_{k=1}^{m} e_{jk}^2. \quad (22)$$

As

$$\begin{aligned}\sum_{k=1}^{m} \widehat{x}_{jk} e_{jk} &= \widehat{\beta}_j \sum_{k=1}^{m} x_{ik} e_{jk} = \widehat{\beta}_j \sum_{k=1}^{m} x_{ik}(x_{jk} - \widehat{\beta}_j x_{ik}) \\ &= \widehat{\beta}_j \left(\sum_{k=1}^{m} x_{ik} x_{jk} - \widehat{\beta}_j \sum_{k=1}^{m} x_{ik}^2 \right) = 0, \end{aligned} \quad (23)$$

equation (22) reduces to

$$\sum_{k=1}^{m} x_{jk}^2 = \sum_{k=1}^{m} \widehat{x}_{jk}^2 + \sum_{k=1}^{m} e_{jk}^2. \quad (24)$$

Here, the total sum of squares (SST) of the m observed values of $\boldsymbol{x}_j$ is decomposed into two parts, with one part explained by the regression (SSR) and the remaining part unexplained by it (SSE); that is,

$$\text{SST} = \text{SSR} + \text{SSE}. \quad (25)$$

The coefficient of determination of the regression, representing the explained sum of squares as a proportion of the total sum of squares, is a goodness-of-fit measure. Defined as

$$R_j^2 = \frac{\sum_{k=1}^{m} \widehat{x}_{jk}^2}{\sum_{k=1}^{m} x_{jk}^2} \quad (26)$$

or, equivalently,

$$R_j^2 = 1 - \frac{\sum_{k=1}^{m} e_{jk}^2}{\sum_{j=1}^{m} x_{jk}^2}, \quad (27)$$

it is commonly known as R^2. Here, a subscript j has been added to indicate that this R^2 pertains to the regression model in which $\boldsymbol{x}_j$ is the dependent variable.

Combining equations (18), (19), and (27) leads to

$$R_j^2 = \frac{\widehat{\beta}_j \sum_{k=1}^{m} x_{jk} x_{ik}}{\sum_{j=1}^{m} x_{jk}^2} = \frac{\left(\sum_{k=1}^{m} x_{jk} x_{ik}\right)^2}{\left(\sum_{k=1}^{m} x_{ik}^2\right)\left(\sum_{j=1}^{m} x_{jk}^2\right)}. \quad (28)$$

This result indicates that the coefficient of determination of the regression is also the square of the sample correlation of $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$. Provided that $R_j^2 \neq 1$, equation (28) is equivalent to

$$\frac{1}{1 - R_j^2} = \frac{\left(\sum_{k=1}^{m} x_{ik}^2\right)\left(\sum_{j=1}^{m} x_{jk}^2\right)}{\left(\sum_{k=1}^{m} x_{ik}^2\right)\left(\sum_{j=1}^{m} x_{jk}^2\right) - \left(\sum_{k=1}^{m} x_{jk} x_{ik}\right)^2} = \frac{s_{ii} s_{jj}}{s_{ii} s_{jj} - s_{ij}^2}. \quad (29)$$

3.3 A Connection between Linear Regression and the Inverse of the Sample Covariance Matrix

Given equation (11), we can write equation (29) as

$$u_{jj} = \frac{1}{s_{jj}(1 - R_j^2)}. \tag{30}$$

Given equations (12) and (18), we can also write equation (29) as

$$u_{ji} = -\frac{1}{1 - R_j^2}\frac{s_{ij}}{s_{ii}s_{jj}} = -\frac{\widehat{\beta}_j}{s_{jj}(1 - R_j^2)}. \tag{31}$$

Again, each of i and j here can be 1 or 2, but $i \neq j$. Thus, the inverse of $\boldsymbol{V}$ is

$$\boldsymbol{V}^{-1} = \left[\begin{array}{cc} u_{11} & u_{12} \\ u_{21} & u_{22} \end{array}\right] = \left[\begin{array}{cc} 1/[s_{11}(1 - R_1^2)] & -\widehat{\beta}_1/[s_{11}(1 - R_1^2)] \\ -\widehat{\beta}_2/[s_{22}(1 - R_2^2)] & 1/[s_{22}(1 - R_2^2)] \end{array}\right]. \tag{32}$$

As interchanging i and j does not affect the right hand side of equation (28), R_i^2 and R_j^2 must be equal; both are the square of the correlation of the two variables considered. Further, as $\widehat{\beta}_1 = s_{21}/s_{22}$ and $\widehat{\beta}_2 = s_{12}/s_{11} = s_{21}/s_{11}$ according to equation (18), the two off-diagonal elements of $\boldsymbol{V}^{-1}$ must also be equal. That is, $\boldsymbol{V}^{-1}$ is symmetric as expected. Equation (32) is a special case of equation (8) for $n = 2$. For notational simplicity, we have omitted the second subscript in each of $\widehat{\beta}_{12}$ and $\widehat{\beta}_{21}$ in formulating the regression models for the derivation of equation (32). Thus, should the same notation as in equation (8) be followed, $\widehat{\beta}_1$ and $\widehat{\beta}_2$ in equation (32) would be $\widehat{\beta}_{12}$ and $\widehat{\beta}_{21}$, respectively.

4 A Multivariate Case

To illustrate multiple linear regression based on the values of x_{ik}, for $i = 1, 2, \ldots, n$ and $k = 1, 2, \ldots, m$, let us start with the case where the dependent variable is $\boldsymbol{x}_1$ and the $n-1$ independent variables are $\boldsymbol{x}_2, \boldsymbol{x}_3, \ldots, \boldsymbol{x}_n$. The regression model is

$$\boldsymbol{x}_1 = \beta_{12}\boldsymbol{x}_2 + \beta_{13}\boldsymbol{x}_3 + \cdots + \beta_{1n}\boldsymbol{x}_n + \boldsymbol{e}_1, \tag{33}$$

where $\beta_{12}, \beta_{13}, \ldots, \beta_{1n}$ are unknown parameters to be estimated and $\boldsymbol{e}_1$ is random noise. Given that $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$ are mean-removed variables, $\boldsymbol{e}_1$ has a zero mean as well. Denote the estimated values of $\beta_{12}, \beta_{13}, \ldots, \beta_{1n}$ as $\widehat{\beta}_{12}, \widehat{\beta}_{13}, \ldots, \widehat{\beta}_{1n}$, respectively, and let

$$\widehat{\boldsymbol{\beta}}_1 = \left[\begin{array}{cccc} \widehat{\beta}_{12} & \widehat{\beta}_{13} & \cdots & \widehat{\beta}_{1n} \end{array}\right]', \tag{34}$$

which is an $(n-1)$-element column vector. Here, the prime indicates matrix transposition.

Let also

$$\boldsymbol{y}_1 = \left[\begin{array}{cccc} x_{11} & x_{12} & \cdots & x_{1m} \end{array}\right]' \tag{35}$$

and

$$\boldsymbol{z}_1 = \left[\begin{array}{cccc} x_{21} & x_{22} & \cdots & x_{2m} \\ x_{31} & x_{32} & \cdots & x_{3m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{array}\right]' \tag{36}$$

be an m-element column vector and an $m \times (n-1)$ matrix, respectively. Further, let

$$\boldsymbol{w}_1 = \left[\begin{array}{cccc} e_{11} & e_{12} & \cdots & e_{1m} \end{array}\right]' \tag{37}$$

be an m-element column vector, where

$$e_{1k} = x_{1k} - \widehat{\beta}_{12} x_{2k} - \widehat{\beta}_{13} x_{3k} - \cdots - \widehat{\beta}_{1n} x_{nk}, \tag{38}$$

for $k = 1, 2, \ldots, m$. The subscript 1 in $\widehat{\boldsymbol{\beta}}_1$, $\boldsymbol{y}_1$, $\boldsymbol{z}_1$, and $\boldsymbol{w}_1$ is for indicating that $\boldsymbol{x}_1$ is the dependent variable.

To capture all observations of the n mean-removed variables $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$, we can write the regression model compactly as

$$\boldsymbol{y}_1 = \boldsymbol{z}_1 \widehat{\boldsymbol{\beta}}_1 + \boldsymbol{w}_1. \tag{39}$$

Noting that

$$\boldsymbol{w}_1' \boldsymbol{w}_1 = \sum_{k=1}^{m} e_{1k}^2 \tag{40}$$

is the sum of squares of the residual noise, we can determine the vector $\widehat{\boldsymbol{\beta}}_1$ corresponding to its minimization. It follows from

$$\begin{aligned} \boldsymbol{w}_1' \boldsymbol{w}_1 &= (\boldsymbol{y}_1 - \boldsymbol{z}_1 \widehat{\boldsymbol{\beta}}_1)' (\boldsymbol{y}_1 - \boldsymbol{z}_1 \widehat{\boldsymbol{\beta}}_1) \\ &= \boldsymbol{y}_1' \boldsymbol{y}_1 - \boldsymbol{y}_1' \boldsymbol{z}_1 \widehat{\boldsymbol{\beta}}_1 - \widehat{\boldsymbol{\beta}}_1' \boldsymbol{z}_1' \boldsymbol{y}_1 + \widehat{\boldsymbol{\beta}}_1' \boldsymbol{z}_1' \boldsymbol{z}_1 \widehat{\boldsymbol{\beta}}_1 \\ &= \boldsymbol{y}_1' \boldsymbol{y}_1 - 2 \widehat{\boldsymbol{\beta}}_1' \boldsymbol{z}_1' \boldsymbol{y}_1 + \widehat{\boldsymbol{\beta}}_1' \boldsymbol{z}_1' \boldsymbol{z}_1 \widehat{\boldsymbol{\beta}}_1, \end{aligned} \tag{41}$$

that

$$\frac{\partial}{\partial \widehat{\boldsymbol{\beta}}_1} \boldsymbol{w}_1' \boldsymbol{w}_1 = -2 \boldsymbol{z}_1' \boldsymbol{y}_1 + 2 \boldsymbol{z}_1' \boldsymbol{z}_1 \widehat{\boldsymbol{\beta}}_1 = \mathbf{0}, \tag{42}$$

where $\mathbf{0}$ is an $(n-1)$-element column vector of zeros. Thus, the best fit is achieved by setting

$$\widehat{\boldsymbol{\beta}}_1 = (\boldsymbol{z}_1' \boldsymbol{z}_1)^{-1} \boldsymbol{z}_1' \boldsymbol{y}_1. \tag{43}$$

4.1 The Coefficient of Determination

Given equation (39), the sum of squares of the m observations of $\boldsymbol{x}_1$ can be written as

$$\begin{aligned} \boldsymbol{y}_1'\boldsymbol{y}_1 &= (\boldsymbol{z}_1\widehat{\boldsymbol{\beta}}_1 + \boldsymbol{w}_1)'(\boldsymbol{z}_1\widehat{\boldsymbol{\beta}}_1 + \boldsymbol{w}_1) \\ &= \widehat{\boldsymbol{\beta}}_1'\boldsymbol{z}_1'\boldsymbol{z}_1\widehat{\boldsymbol{\beta}}_1 + \widehat{\boldsymbol{\beta}}_1'\boldsymbol{z}_1'\boldsymbol{w}_1 + \boldsymbol{w}_1'\boldsymbol{z}_1\widehat{\boldsymbol{\beta}}_1 + \boldsymbol{w}_1'\boldsymbol{w}_1 \\ &= \widehat{\boldsymbol{\beta}}_1'\boldsymbol{z}_1'\boldsymbol{z}_1\widehat{\boldsymbol{\beta}}_1 + 2\widehat{\boldsymbol{\beta}}_1'\boldsymbol{z}_1'\boldsymbol{w}_1 + \boldsymbol{w}_1'\boldsymbol{w}_1. \end{aligned} \tag{44}$$

As

$$\begin{aligned} \widehat{\boldsymbol{\beta}}_1'\boldsymbol{z}_1'\boldsymbol{w}_1 &= \widehat{\boldsymbol{\beta}}_1'\boldsymbol{z}_1'(\boldsymbol{y}_1 - \boldsymbol{z}_1\widehat{\boldsymbol{\beta}}_1) = \widehat{\boldsymbol{\beta}}_1'(\boldsymbol{z}_1'\boldsymbol{y}_1 - \boldsymbol{z}_1'\boldsymbol{z}_1\widehat{\boldsymbol{\beta}}_1) \\ &= \widehat{\boldsymbol{\beta}}_1'[\boldsymbol{z}_1'\boldsymbol{y}_1 - \boldsymbol{z}_1'\boldsymbol{z}_1(\boldsymbol{z}_1'\boldsymbol{z}_1)^{-1}\boldsymbol{z}_1'\boldsymbol{y}_1] = \widehat{\boldsymbol{\beta}}_1'(\boldsymbol{z}_1'\boldsymbol{y}_1 - \boldsymbol{z}_1'\boldsymbol{y}_1) = 0, \end{aligned} \tag{45}$$

equation (44) reduces to

$$\boldsymbol{y}_1'\boldsymbol{y}_1 = \widehat{\boldsymbol{\beta}}_1'\boldsymbol{z}_1'\boldsymbol{z}_1\widehat{\boldsymbol{\beta}}_1 + \boldsymbol{w}_1'\boldsymbol{w}_1 = \widehat{\boldsymbol{y}}_1'\widehat{\boldsymbol{y}}_1 + \boldsymbol{w}_1'\boldsymbol{w}_1, \tag{46}$$

where

$$\widehat{\boldsymbol{y}}_1 = \boldsymbol{z}_1\widehat{\boldsymbol{\beta}}_1 \tag{47}$$

is an m-element column vector representing the fitted values of the dependent variable $\boldsymbol{x}_1$. Analogous to the two-variable case considered earlier, the total sum of squares (SST) of the m observations of the dependent variable can also be decomposed into two parts, with one part explained by the regression (SSR) and the remaining part unexplained by it (SSE). That is, equation (25) also holds here.

The coefficient of multiple determination (or simply the coefficient of determination), denoted also by R^2, is defined as

$$R_1^2 = \widehat{\boldsymbol{y}}_1'\widehat{\boldsymbol{y}}_1(\boldsymbol{y}_1'\boldsymbol{y}_1)^{-1} \tag{48}$$

or, equivalently,

$$R_1^2 = 1 - \boldsymbol{w}_1'\boldsymbol{w}_1(\boldsymbol{y}_1'\boldsymbol{y}_1)^{-1}. \tag{49}$$

It is a goodness-of-fit measure, which provides the explained sum of squares as a proportion of the total sum of squares. Here, again, the subscript 1 is for indicating that the dependent variable in the regression is the variable $\boldsymbol{x}_1$. As

$$\widehat{\boldsymbol{y}}_1'\widehat{\boldsymbol{y}}_1 = \widehat{\boldsymbol{\beta}}_1'\boldsymbol{z}_1'\boldsymbol{z}_1\widehat{\boldsymbol{\beta}}_1 = \widehat{\boldsymbol{\beta}}_1'\boldsymbol{z}_1'\boldsymbol{z}_1(\boldsymbol{z}_1'\boldsymbol{z}_1)^{-1}\boldsymbol{z}_1'\boldsymbol{y}_1 = \widehat{\boldsymbol{\beta}}_1'\boldsymbol{z}_1'\boldsymbol{y}_1, \tag{50}$$

we can write

$$R_1^2 = \widehat{\boldsymbol{\beta}}_1' \boldsymbol{z}_1' \boldsymbol{y}_1 (\boldsymbol{y}_1' \boldsymbol{y}_1)^{-1} \tag{51}$$

and then

$$R_1^2 \boldsymbol{y}_1' \boldsymbol{y}_1 = \widehat{\boldsymbol{\beta}}_1' \boldsymbol{z}_1' \boldsymbol{y}_1. \tag{52}$$

Each side being a scalar, $\widehat{\boldsymbol{\beta}}_1' \boldsymbol{z}_1' \boldsymbol{y}_1$ is the same as $\boldsymbol{y}_1' \boldsymbol{z}_1 \widehat{\boldsymbol{\beta}}_1$.

4.2 A Connection between Multiple Linear Regression and the Inverse of the Sample Covariance Matrix

Let us partition the $n \times n$ sample covariance matrix $\boldsymbol{V}$ as

$$\boldsymbol{V} = \begin{bmatrix} \boldsymbol{V}_{11} & \boldsymbol{V}_{12} \\ \boldsymbol{V}_{21} & \boldsymbol{V}_{22} \end{bmatrix}, \tag{53}$$

where $\boldsymbol{V}_{11}$ and $\boldsymbol{V}_{22}$ are square matrices. Let us also partition its inverse conformally as

$$\boldsymbol{V}^{-1} = \begin{bmatrix} \boldsymbol{U}_{11} & \boldsymbol{U}_{12} \\ \boldsymbol{U}_{21} & \boldsymbol{U}_{22} \end{bmatrix}. \tag{54}$$

That is, $\boldsymbol{U}_{11}$, $\boldsymbol{U}_{12}$, $\boldsymbol{U}_{21}$, and $\boldsymbol{U}_{22}$ are to have the same dimensions as $\boldsymbol{V}_{11}$, $\boldsymbol{V}_{12}$, $\boldsymbol{V}_{21}$, and $\boldsymbol{V}_{22}$, respectively. As shown in the Appendix, the inverse of $\boldsymbol{V}$ can be expressed as

$$\boldsymbol{V}^{-1} = \begin{bmatrix} \boldsymbol{U}_{11} & -\boldsymbol{U}_{11}\boldsymbol{V}_{12}\boldsymbol{V}_{22}^{-1} \\ -\boldsymbol{V}_{22}^{-1}\boldsymbol{V}_{21}\boldsymbol{U}_{11} & \boldsymbol{V}_{22}^{-1} + \boldsymbol{V}_{22}^{-1}\boldsymbol{V}_{21}\boldsymbol{U}_{11}\boldsymbol{V}_{12}\boldsymbol{V}_{22}^{-1} \end{bmatrix}, \tag{55}$$

where

$$\boldsymbol{U}_{11} = (\boldsymbol{V}_{11} - \boldsymbol{V}_{12}\boldsymbol{V}_{22}^{-1}\boldsymbol{V}_{21})^{-1}. \tag{56}$$

For the task here, of interest is the case where $\boldsymbol{V}_{11}$ has only one element, which is s_{11}, the sample variance of the variable $\boldsymbol{x}_1$. As $\boldsymbol{V}^{-1}$ is partitioned conformally,

$$\boldsymbol{U}_{11} = (s_{11} - \boldsymbol{V}_{12}\boldsymbol{V}_{22}^{-1}\boldsymbol{V}_{21})^{-1} \tag{57}$$

also has only one element. Accordingly, both

$$\boldsymbol{V}_{12} = \begin{bmatrix} s_{12} & s_{13} & \cdots & s_{1n} \end{bmatrix} \tag{58}$$

and

$$\boldsymbol{U}_{12} = -\boldsymbol{U}_{11}\boldsymbol{V}_{12}\boldsymbol{V}_{22}^{-1} \tag{59}$$

are $(n-1)$-element row vectors.

It follows from equations (35) and (36) that

$$\boldsymbol{V}_{12} = \frac{1}{m-1}\boldsymbol{y}_1'\boldsymbol{z}_1 = \boldsymbol{V}_{21}' \tag{60}$$

and

$$\boldsymbol{V}_{22} = \frac{1}{m-1}\boldsymbol{z}_1'\boldsymbol{z}_1. \tag{61}$$

We can then write equation (57) as

$$\boldsymbol{U}_{11} = \left[s_{11} - \frac{1}{m-1}\boldsymbol{y}_1'\boldsymbol{z}_1(\boldsymbol{z}_1'\boldsymbol{z}_1)^{-1}\boldsymbol{z}_1'\boldsymbol{y}_1\right]^{-1}. \tag{62}$$

Further substitutions by using equations (43) and (52) lead to

$$\boldsymbol{U}_{11} = \left[s_{11} - \frac{1}{m-1}\boldsymbol{y}_1'\boldsymbol{z}_1\widehat{\boldsymbol{\beta}}_1\right]^{-1} = \left[s_{11} - \frac{R^2}{m-1}\boldsymbol{y}_1'\boldsymbol{y}_1\right]^{-1} = \frac{1}{s_{11}(1-R^2)}. \tag{63}$$

To connect $\boldsymbol{U}_{12}$ to regression results, we first write

$$\boldsymbol{V}_{12}\boldsymbol{V}_{22}^{-1} = \boldsymbol{y}_1'\boldsymbol{z}_1(\boldsymbol{z}_1'\boldsymbol{z}_1)^{-1} = \widehat{\boldsymbol{\beta}}_1'. \tag{64}$$

This requires equations (60) and (61) and, subsequently, equation (43). As

$$\boldsymbol{U}_{12} = -\boldsymbol{U}_{11}\boldsymbol{V}_{12}\boldsymbol{V}_{22}^{-1} = -\frac{1}{s_{11}(1-R_1^2)}\widehat{\boldsymbol{\beta}}_1', \tag{65}$$

the first row of $\boldsymbol{V}^{-1}$ is

$$\begin{bmatrix} \boldsymbol{U}_{11} & \boldsymbol{U}_{12} \end{bmatrix} = \begin{bmatrix} 1/[s_{11}(1-R_1^2)] & -\widehat{\beta}_{12}/[s_{11}(1-R_1^2)] & \cdots & -\widehat{\beta}_{1n}/[s_{11}(1-R_1^2)] \end{bmatrix}. \tag{66}$$

The inverse of an invertible sample covariance matrix is unique, and thus the above connection can easily be extended, for the purpose of interpreting each remaining row of the inverse of the sample covariance matrix. All that is required is to relabel the n variables, so that the variable corresponding to the row in question is treated as variable $\boldsymbol{x}_1$. Once the multiple linear regression as described above has been performed, with the corresponding R_1^2 and $\widehat{\boldsymbol{\beta}}_1$ obtained, the original variable label is restored. This procedure, which requires corresponding rearrangements of the affected rows and columns of the sample covariance matrix and its inverse, leads to equation (8).

4.3 Mean Removal

For analytical convenience, mean-removed observations have been used above for fitting linear regression models. However, the same models can be stated equivalently without requiring mean removal. To illustrate with the multivariate linear regression model where $\boldsymbol{X}_1$ is the dependent variable and $\boldsymbol{X}_2, \boldsymbol{X}_3, \ldots, \boldsymbol{X}_n$ are the independent variables, let us fit equation (33) with mean-removed observations, by writing

$$x_{1k} = \widehat{\beta}_{12} x_{2k} + \widehat{\beta}_{13} x_{3k} + \cdots + \widehat{\beta}_{1n} x_{nk} + e_{1k} \tag{67}$$

as

$$X_{1k} - \overline{X}_1 = \widehat{\beta}_{12}(X_{2k} - \overline{X}_2) + \widehat{\beta}_{13}(X_{3k} - \overline{X}_3) + \cdots + \widehat{\beta}_{1n}(X_{nk} - \overline{X}_n) + e_{1k}, \tag{68}$$

for $k = 1, 2, \ldots, m$.

Equation (68) is equivalent to

$$X_{1k} = \widehat{\alpha}_1 + \widehat{\beta}_{12} X_{2k} + \widehat{\beta}_{13} X_{3k} + \cdots + \widehat{\beta}_{1n} X_{nk} + e_{1k}, \tag{69}$$

for $k = 1, 2, \ldots, m$, where

$$\widehat{\alpha}_1 = \overline{X}_1 - \widehat{\beta}_{12}\overline{X}_2 - \widehat{\beta}_{13}\overline{X}_3 - \cdots - \widehat{\beta}_{1n}\overline{X}_n. \tag{70}$$

Thus, regardless of whether mean removal has been performed on the raw observations, the regression coefficients $\widehat{\beta}_{12}, \widehat{\beta}_{13}, \ldots, \widehat{\beta}_{1n}$ and the goodness-of-fit measure as provided by the coefficient of determination R_1^2 are unaffected. The only difference is that, without mean removal, there is an intercept term $\widehat{\alpha}_1$ in the regression model, which can be deduced from equation (70). The equivalence of the two versions of each regression model allows us to fit it without always having to use mean-removed observations, as the raw observations are already adequate for the task.

4.4 Matrix Inversion

An inspection of equation (8) reveals the conditions for $\boldsymbol{V}$ not to be invertible. Specifically, if any of $s_{11}, s_{22}, \ldots, s_{nn}$ is zero or if any of $R_1^2, R_2^2, \ldots, R_n^2$ is one, $\boldsymbol{V}^{-1}$ does not exist. As s_{ii} is the sample variance of variable $\boldsymbol{X}_i$, for $i = 1, 2, \ldots, n$, a zero s_{ii} indicates that the m observations of the variable are identical. With the entire row i (column i) of the $n \times n$ matrix $\boldsymbol{V}$ being zeros, the determinant of $\boldsymbol{V}$ is inevitably zero. Accordingly, $\boldsymbol{V}^{-1}$ does not exist. The case where $R_i^2 = 1$ pertains to the situation where there is an exact linear relationship between $\boldsymbol{X}_i$ and any of the remaining variables considered. This is the situation where row i (column i) of $\boldsymbol{V}$ can be replicated by a linear combination of some other rows (columns); it includes the special case where variable $\boldsymbol{x}_i$ is perfectly correlated with any of the remaining variables. In such a situation, as the determinant of the matrix is zero, the inverse of the matrix does not exist.

5 An Analytical Example

We now illustrate that equation (8) can improve the interpretation of a well-known analytical result of risk minimization in portfolio analysis. This illustration is much simpler than that in Stevens (1998), where risk minimization is also subject to a requirement on the portfolio's expected return. For the task here, let us consider a set of n risky securities, for which the sample covariance matrix of returns is an $n \times n$ invertible matrix $\boldsymbol{V}$. The (i, j)-element of $\boldsymbol{V}$ is s_{ij}, for $i, j = 1, 2, \ldots, n$. Suppose that a portfolio o is formed, with proportions of investment funds $a_1, a_2, \ldots, a_n$ assigned to the n securities. These proportions of investment funds are known as portfolio weights.

For analytical convenience, let us assume frictionless short sales; this is the case where the short seller provides no margin deposit for each shorted security and has immediate access to the short-sale proceeds. Under this simplifying assumption, the only constraint for the construction of a portfolio to achieve the lowest possible risk based on the set of n securities is

$$\sum\nolimits_{i=1}^{n} a_i = 1. \tag{71}$$

That is, the investment funds are intended to be fully allocated among the n securities considered.

Security i is said to be held long in portfolio o if a_i is strictly positive; it is said to be held short there if a_i is strictly negative. The portfolio risk, labeled as s_o^2 and represented by the sample variance of returns of the portfolio, is[8]

$$s_o^2 = \sum\nolimits_{i=1}^{n} \sum\nolimits_{j=1}^{n} a_i a_j s_{ij}. \tag{72}$$

[8] See Kwan (2007, Footnote 4) for a simple way to compute the variance of portfolio returns.

In matrix notation, we can write

$$s_o^2 = \boldsymbol{a}'\boldsymbol{V}\boldsymbol{a}, \tag{73}$$

where $\boldsymbol{a}$ is an n-element column vector with elements being $a_1, a_2, \ldots, a_n$.

5.1 The Sum of Elements in Each Row of the Inverse of the Covariance Matrix

As shown pedagogically in Kwan (2010), an invertible sample covariance matrix is positive definite and, given a positive definite covariance matrix $\boldsymbol{V}$, the portfolio weight vector corresponding to the lowest portfolio risk, is

$$\boldsymbol{a} = \boldsymbol{V}^{-1}\boldsymbol{\iota}\ \left(\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\iota}\right)^{-1}, \tag{74}$$

where $\boldsymbol{\iota}$ is an n-element column vector with each element being one. Here, $\boldsymbol{V}^{-1}\boldsymbol{\iota}$ is an n-element column vector with its i-th element being the sum of the n elements in row i of $\boldsymbol{V}^{-1}$; $\left(\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\iota}\right)^{-1}$, which is a positive scalar, is the reciprocal of the sum of all elements of $\boldsymbol{V}^{-1}$.[9] Thus, the portfolio weight vector $\boldsymbol{a}$ corresponding to portfolio risk minimization is proportional to $\boldsymbol{V}^{-1}\boldsymbol{\iota}$.

For ease of exposition below, let us label the i-th element of the vector $\boldsymbol{V}^{-1}\boldsymbol{\iota}$ as g_i, for $i = 1, 2, \ldots, n$. The sign of g_i is the same as the sign of a_i. The magnitude of g_i is proportional to the investment funds for security i in the portfolio.

Although the portfolio weight vector in equation (74) is easy to compute, it is not as easy to extract any intuition that underlies such a solution. In view of equation (2) and the symmetry of the covariance matrix, the i-th element of the vector $\boldsymbol{V}^{-1}\boldsymbol{\iota}$ is

$$g_i = \frac{1}{\det \boldsymbol{V}} \sum\nolimits_{j=1}^{n} C_{ij}, \tag{75}$$

for $i = 1, 2, \ldots, n$. Therefore, it is difficult to interpret intuitively the least risky portfolio weights by using standard textbook materials in matrix algebra.

[9] For an $n \times n$ covariance matrix $\boldsymbol{V}$ to be positive definite, we must have $\boldsymbol{b}'\boldsymbol{V}\boldsymbol{b} > 0$ for any n-element nonzero column vector $\boldsymbol{b}$. As we can write $0 < \boldsymbol{b}'\boldsymbol{V}\boldsymbol{b} = \boldsymbol{b}'\boldsymbol{V}\boldsymbol{V}^{-1}\boldsymbol{V}\boldsymbol{b} = (\boldsymbol{V}\boldsymbol{b})'\boldsymbol{V}^{-1}(\boldsymbol{V}\boldsymbol{b})$, where $\boldsymbol{V}\boldsymbol{b}$ is an arbitrary n-element column vector, the positive definiteness of $\boldsymbol{V}^{-1}$ is confirmed. Thus, if $\boldsymbol{V}$ is a positive definite covariance matrix, then $\boldsymbol{\iota}'\boldsymbol{V}^{-1}\boldsymbol{\iota}$ is a positive scalar.

5.2 A Regression-Based Interpretation

Suppose that the estimation of the $n \times n$ covariance matrix $\boldsymbol{V}$ is based on m return observations of the n securities considered. Equation (8) allows us to write the i-th element of the vector $\boldsymbol{V}^{-1}\boldsymbol{\iota}$ as

$$g_i = \frac{1}{s_{ii}(1 - R_i^2)} \left(1 - \sum\nolimits_{j=1,\ j \neq i}^{n} \widehat{\beta}_{ij}\right), \tag{76}$$

for $i = 1, 2, \ldots, n$, where, more explicitly,

$$\sum\nolimits_{j=1,\ j \neq i}^{n} \widehat{\beta}_{ij} = \widehat{\beta}_{i1} + \widehat{\beta}_{i2} + \cdots + \widehat{\beta}_{i,i-1} + \widehat{\beta}_{i,i+1} + \cdots + \widehat{\beta}_{in}. \tag{77}$$

The right hand side of equation (76) is unaffected by the sample means of the n security returns, and thus we can use zero sample means for expositional convenience.

As equation (76) indicates, the sign and the magnitude of each portfolio weight depend on the set of regression coefficients, and the magnitude depends also on the corresponding security's variance of returns and the coefficient of determination of the linear regression model involved. Let us examine the sign of each portfolio weight first. For the case where

$$\sum\nolimits_{j=1,\ j \neq i}^{n} \widehat{\beta}_{ij} = 1, \tag{78}$$

security i can be viewed as an inexact replication by a portfolio of securities $1, 2, \ldots, i-1, i+1, \ldots, n$, with the portfolio weights being the corresponding regression coefficients. The replication is inexact because of the presence of the residual noise $\boldsymbol{e}_i$ in the regression model. Without the noise term, however, the $n \times n$ covariance matrix $\boldsymbol{V}$ would not be invertible, and security i and the replicating portfolio would have the same variance of returns. It is the noise term that makes security i riskier (as compared to the replicating portfolio) and less attractive for investing, given that the objective here is to achieve the lowest possible variance of portfolio returns. Thus, if equation (78) holds, then no investment funds ought to be assigned to security i; this decision is consistent with $g_i = 0$ according to equation (76).[10]

For the case where

$$\sum\nolimits_{j=1,\ j \neq i}^{n} \widehat{\beta}_{ij} < 1, \tag{79}$$

[10] The idea here is based on the concept of mean-variance spanning in finance, under the assumption of frictionless short sales. As shown in Cheung, Kwan, and Mountain (2009), if each security in one set of securities can be replicated, though not exactly, by a portfolio of securities from another set of securities, it is the residual noise that makes each replicated security unattractive for portfolio holding. In such a case, the replicated securities are said to be spanned, and they are assigned zero weights in all mean-variance efficient portfolios based on these two sets of securities.

let us add a positive parameter $\widehat{\beta}_{i0}$ to the left hand side, so that

$$\sum\nolimits_{j=0,\ j\neq i}^{n} \widehat{\beta}_{ij} = 1. \tag{80}$$

Given equation (80), we can treat $\widehat{\beta}_{i0}$ and the $n-1$ regression coefficients as portfolio weights, for the purpose of replicating security i. The replicating portfolio, which is also inexact, consists of a risk-free security with a zero mean (which is the same as idle cash) and securities $1, 2, \ldots, i-1, i+1, \ldots, n$.

It is well-known in portfolio theory that, if investment funds are allocated positively between a risk-free security and a risky portfolio, the result is portfolio risk reduction. As the objective here is to achieve the lowest possible variance of portfolio returns, the replicating portfolio for security i based on equation (80) clearly has an attractive feature. Given equation (80), the term $1 - \sum_{j=1,\ j\neq i}^{n} \widehat{\beta}_{ij}$ on the right hand side of equation (76) is positive; it is the same as $\widehat{\beta}_{i0}$, which is a positive parameter representing the risk-free component of the replicating portfolio. This explains why, for the case pertaining to inequality (79), security i ought to be held long.

For the remaining case where

$$\sum\nolimits_{j=1,\ j\neq i}^{n} \widehat{\beta}_{ij} > 1, \tag{81}$$

the parameter $\widehat{\beta}_{i0}$ to be added to the left hand side, for equation (80) to hold, is negative instead. With $\widehat{\beta}_{i0}$ and the $n-1$ regression coefficients treated as portfolio weights, security i can be viewed as an inexact replication, by a portfolio with a risk-free component and a risky component based on securities $1, 2, \ldots, i-1, i+1, \ldots, n$. As $\widehat{\beta}_{i0} < 0$, risk-free borrowing at a zero interest rate is involved.

Holding security i in a long position is like using borrowed funds to invest more in a portfolio based on securities $1, 2, \ldots, i-1, i+1, \ldots, n$. Even if such risk-free borrowing is deemed feasible, it has an undesirable effect of increasing the portfolio risk. However, if security i is held short instead, the replicating portfolio (which replicates the short sale of security i) will have a positive weight on its risk-free component. If so, short selling security i will contribute to portfolio risk reduction. Thus, for the case pertaining to inequality (81), a negative portfolio weight for security i is appropriate.

Now, let us turn our attention to the magnitude of each portfolio weight. According to equation (76), g_i varies directly with $1 - \sum_{j=1,\ j\neq i}^{n} \widehat{\beta}_{ij}$ and inversely with s_{ii} and $1 - R_i^2$. It is intuitively appealing that, with everything else held constant, the greater the risk of a security, the lower is the magnitude of the portfolio weight for it. With $1 - R_i^2$ being the proportion of the total sum of squares that the regression model cannot explain, it reflects the severity of the residual noise in the regression model, which undermines the substitution effect of the replicating portfolio for security i. The higher the coefficient of determination, the less severe is the residual noise and thus the more exact is the replication. Then, with s_{ii} held constant, the more attractive is security i for portfolio holding (long or short, as the case may be) if $\sum_{j=1,\, j\neq i}^{n} \widehat{\beta}_{ij}$ departs more from 1.

6 An Excel Illustration

The illustration in the current section is on a given set of 25 observations of four variables, for which a 4×4 sample covariance matrix and its inverse are computed in Excel. The three figures below, which cover different aspects of the illustration, are based on an Excel file accompanying this paper.[11] The file contains a single worksheet covering A1:AV58, and thus cell references are common across the three figures. For ease of exposition, essential cell formulas are displayed at the bottom of each figure.

As shown in B2:E26 of Figure 1, the given observations under the headings of X1, X2, X3, and X4 are shaded. The individual sample means are displayed in B28:E28. The corresponding mean-removed observations, under the headings of x1, x2, x3, and x4, are provided in H2:K26. The sample variances of the four variables are displayed in H28:K28. The sample covariance matrix and its inverse are displayed in H32:K35 and H38:K41, respectively.[12] As expected, both matrices are symmetric. The 4×4 block containing the inverse of the sample covariance matrix is shaded.

Figure 2 shows the part of the Excel file where the inverse of the sample covariance matrix is obtained via four Solver runs. As intended, the search procedure does not require the use of matrix algebra. Nor does it involve any Excel functions for matrix operations, although the

[11] Although Excel 2007 is used for the illustration, the Excel file accompanying this paper is still a 1997-2003 version (with extension .xls), for potentially wider access by readers.

[12] An approach involving matrix operations has been used to deduce the sample covariance matrix in H32:K35. Once the corresponding cell formula has been provided for this selected 4×4 block, the keys "Shift+Ctrl+Enter" must be pressed together. An alternative approach (which is not shown in Figure 1) is to compute the individual diagonal and off-diagonal elements of the sample covariance matrix by using the Excel functions VAR and COVAR, respectively. Such an approach requires the use of a multiplicative correction factor $m/(m-1)$, where $m = 25$, for each COVAR result.

	A	B	C	D	E	F	G	H	I	J	K	L
1	Obs	X1	X2	X3	X4			x1	x2	x3	x4	
2	1	2.2	1.9	2.4	1.4			2.284	1.812	1.984	1.072	
3	2	-3.0	0.1	-0.4	-2.6			-2.916	0.012	-0.816	-2.928	
4	3	-0.1	1.4	0.2	2.9			-0.016	1.312	-0.216	2.572	
5	4	4.2	2.6	0.6	3.9			4.284	2.512	0.184	3.572	
6	5	-0.2	1.2	0.6	2.0			-0.116	1.112	0.184	1.672	
7	6	-0.9	-2.3	1.3	-0.4			-0.816	-2.388	0.884	-0.728	
8	7	-1.8	-4.0	-2.5	-3.2			-1.716	-4.088	-2.916	-3.528	
9	8	-0.9	-2.0	-0.8	-2.8			-0.816	-2.088	-1.216	-3.128	
10	9	-0.5	1.4	0.6	0.1			-0.416	1.312	0.184	-0.228	
11	10	-0.6	1.1	-0.8	-0.6			-0.516	1.012	-1.216	-0.928	
12	11	-0.5	-0.5	1.1	0.5			-0.416	-0.588	0.684	0.172	
13	12	-1.8	-3.0	0.3	-4.0			-1.716	-3.088	-0.116	-4.328	
14	13	1.5	-0.3	1.0	2.4			1.584	-0.388	0.584	2.072	
15	14	0.7	0.6	0.6	1.4			0.784	0.512	0.184	1.072	
16	15	-1.3	-2.4	-0.9	-0.9			-1.216	-2.488	-1.316	-1.228	
17	16	0.1	1.5	-0.2	-0.5			0.184	1.412	-0.616	-0.828	
18	17	2.1	0.7	2.1	2.4			2.184	0.612	1.684	2.072	
19	18	3.2	2.9	2.9	4.0			3.284	2.812	2.484	3.672	
20	19	-3.8	0.6	0.9	-1.5			-3.716	0.512	0.484	-1.828	
21	20	-2.8	-4.7	-1.8	-2.2			-2.716	-4.788	-2.216	-2.528	
22	21	-0.2	2.0	0.4	0.2			-0.116	1.912	-0.016	-0.128	
23	22	-1.9	-0.8	-1.2	-1.4			-1.816	-0.888	-1.616	-1.728	
24	23	0.3	1.8	3.2	1.4			0.384	1.712	2.784	1.072	
25	24	1.1	0.2	1.0	2.6			1.184	0.112	0.584	2.272	
26	25	2.8	2.2	-0.2	3.1			2.884	2.112	-0.616	2.772	
27												
28	Mean	-0.084	0.088	0.416	0.328		Var	3.9364	4.32777	1.88307	5.2696	
29												
30												
31								Sample Cov Mat				
32								3.9364	2.61978	1.4564	3.88078	
33								2.61978	4.32777	1.72353	3.5191	
34								1.4564	1.72353	1.88307	1.93078	
35								3.88078	3.5191	1.93078	5.2696	
36												
37								Inv of Sample Cov Mat				
38								0.92828	-0.0079	-0.0243	-0.6695	
39								-0.0079	0.55032	-0.203	-0.2873	
40								-0.0243	-0.203	0.92628	-0.1859	
41								-0.6695	-0.2873	-0.1859	0.94279	
42												
43												
44												
45												
46	B28	=AVERAGE(B2:B26)			Copied to B28:E28							
47	H2	=B2-B$28			Copied to H2:K26							
48	H28	=VAR(H2:H26)			Copied to H28:K28							
49	H32:K35		{=MMULT(TRANSPOSE(H2:K26),H2:K26)/(COUNT(H2:H26)-1)}									
50	H38:K41		{=MINVERSE(H32:K35)}									
51												
52												
53												
54												
55												
56												
57												
58												

Figure 1 The Part of an Excel Example Illustrating the Computations of the Sample Covariance Matrix and its Inverse.

	M	N	O	P	Q	R	S	T	U	V	W	X
1		x1^2	x2^2	x3^2	x4^2			e1^2	e2^2	e3^2	e4^2	
2		5.21666	3.28334	3.93626	1.14918			2.08382	0.23793	1.72076	2.22997	
3		8.50306	0.00014	0.66586	8.57318			0.61325	3.54755	0.02383	0.49021	
4		0.00026	1.72134	0.04666	6.61518			3.52072	0.0024	1.03907	4.95553	
5		18.3527	6.31014	0.03386	12.7592			2.82845	0.26812	1.43024	0.07389	
6		0.01346	1.23654	0.03386	2.79558			1.78509	0.02986	0.15386	1.90214	
7		0.66586	5.70254	0.78146	0.52998			0.08636	5.39302	2.48019	0.16393	
8		2.94466	16.7117	8.50306	12.4468			0.88248	1.31286	1.60516	0.2387	
9		0.66586	4.35974	1.47866	9.78438			2.2184	2.9E-05	0.01191	2.79699	
10		0.17306	1.72134	0.03386	0.05198			0.07157	1.87451	0.00219	0.13597	
11		0.26626	1.02414	1.47866	0.86118			0.03116	3.81205	1.53254	0.39718	
12		0.17306	0.34574	0.46786	0.02958			0.30578	0.85406	0.62292	0.26184	
13		2.94466	9.53574	0.01346	18.7316			2.05785	0.57898	2.1741	4.60321	
14		2.50906	0.15054	0.34106	4.29318			0.00604	2.91699	0.04475	0.90311	
15		0.61466	0.26214	0.03386	1.14918			3E-06	0.01608	0.02688	0.10431	
16		1.47866	6.19014	1.73186	1.50798			0.07552	1.80637	0.24242	0.42672	
17		0.03386	1.99374	0.37946	0.68558			0.61669	4.28035	0.5838	1.60657	
18		4.76986	0.37454	2.83586	4.29318			0.41011	1.25945	1.15919	6.7E-06	
19		10.7847	7.90734	6.17026	13.4836			0.29909	0.00469	1.091	4.6E-05	
20		13.8087	0.26214	0.23426	3.34158			5.83077	1.79853	0.69934	0.31266	
21		7.37666	22.9249	4.91066	6.39078			0.63072	6.82158	0.34576	1.6817	
22		0.01346	3.65574	0.00026	0.01638			0.00156	3.94576	0.16505	0.39086	
23		3.29786	0.78854	2.61146	2.98598			0.27033	0.40487	1.05449	0.02275	
24		0.14746	2.93094	7.75066	1.14918			0.22711	0.01437	4.76799	0.07368	
25		1.40186	0.01254	0.34106	5.16198			0.22163	1.70725	0.00523	1.64344	
26		8.31746	4.46054	0.37946	7.68398			0.77986	0.72356	2.92731	0.04078	
27												
28	SST	94.4736	103.866	45.1936	126.47		SSE	25.8544	43.6112	25.91	25.4562	
29							R^2	0.72633	0.58012	0.42669	0.79872	
30												
31		Initial Reg Coefficients						Reg Coefficients (Solver Results)				
32		-1	-1	-1	-1			-1	0.01431	0.02625	0.71008	
33		-1	-1	-1	-1			0.00848	-1	0.21914	0.30477	
34		-1	-1	-1	-1			0.0262	0.36886	-1	0.1972	
35		-1	-1	-1	-1			0.72118	0.52212	0.20072	-1	
36												
37								Transpose of Inv of Sample Cov Mat				
38								0.92828	-0.0079	-0.0243	-0.6695	
39								-0.0079	0.55032	-0.203	-0.2873	
40								-0.0243	-0.203	0.92628	-0.1859	
41								-0.6695	-0.2873	-0.1859	0.94279	
42												
43												
44	N2	=H2^2		Copied to N2:Q26								
45	N28	=SUM(N2:N26)		Copied to N28:Q28								
46	T2	=(H2-T$33*I2-T$34*J2-T$35*K2)^2										
47				Copied to T2:T26								
48	U2	=(I2-U$32*H2-U$34*J2-U$35*K2)^2										
49				Copied to U2:U26								
50	V2	=(J2-V$32*H2-V$33*I2-V$35*K2)^2						Four Solver Runs				
51				Copied to V2:V26								
52	W2	=(K2-W$32*H2-W$33*I2-W$34*J2)^2						Min Target T28		Change T32:T35		
53				Copied to W2:W26				Min Target U28		Change U32:U35		
54	T28	=SUM(T2:T26)		Copied to T28:W28				Min Target V28		Change V32:V35		
55	T29	=1-T28/N28		Copied to T29:W29				Min Target W28		Change W32:W35		
56	T38	=-T32/H$28/(1-T$29)										
57				Copied to T38:W41								
58												

Figure 2 The Part of an Excel Example Illustrating the Solver Approach to Implement a Regression-Based Interpretation of the Inverse of the Sample Covariance Matrix.

use of such functions will make the displayed materials more succinct. The squared deviations from individual sample means, under the headings of x1^2, x2^2, x3^2, and x4^2, are shown in N2:Q26. Each column sum in N28:Q28 represents the total sum of squares (SST) of the corresponding variable. Notice that N28 can also be reached by using directly the formula =MMULT(TRANSPOSE(H2:H26),H2:H26), where H2:H26 contains the column vector of mean-removed observations of the first variable. With N28 copied to O28:Q28, the computations of the total sums of squares as displayed there, which are based on matrix operations, do not require the data in N2:Q26.

For a Solver search of the regression coefficients, some arbitrary initial values are required. They are stored in N32:Q35 and copied to T32:W35. The three regression coefficients for each of the four regression model are placed in off-diagonal positions of the corresponding column in T32:W35. The diagonal elements in T32:W35 have no impact on the Solver search and thus do not vary during the search process. However, as having the value of -1 for each diagonal element allows the formula in T38 to be copied directly to T38:W41, for computing the inverse of the sample covariance matrix, each initial value in N32:Q35 has been set to be -1 for convenience.

Column T in Figure 2, under the heading of e1^2, pertains to the case where $\boldsymbol{x}_1$ is the dependent variable in regression analysis. The formula for T2, which computes the squared deviation of the first observation of the dependent variable from its fitted value given the regression coefficients, is copied to T2:T26. The sum of T2:T26, as displayed in T28 and labeled as SSE, is the sum of squares that is unexplained by the regression model involved. Given SST and SSE, the R^2 of the regression, as shown in T29, is computed in accordance with equation (49). The Solver results, as displayed in T32:T35, are from minimization of the target cell T28 by varying T32:T35. The formula for T38 is copied to T38:T41. As expected, the four numbers in T38:T41 match the corresponding numbers in H38:K38 (in Figure 1), which are in row 1 of the inverse of the sample covariance matrix.

The same idea also applies to each case where any remaining variable is the dependent variable in regression analysis. Cases pertaining to $\boldsymbol{x}_2$, $\boldsymbol{x}_3$, and $\boldsymbol{x}_4$ are under the headings of e2^2, e3^2, and e4^2, respectively. The formulas for T2, U2, V2, and W2 differ from each other, as each case pertains to a specific dependent variable. The formulas for U2:W2 are copied to U2:W26, and the formulas for T28:T41 are copied to U28:W41. The Solver search as described earlier is repeated for each of the remaining cases. For $\boldsymbol{x}_2$, the target cell U28 is minimized by changing U32:U35; for the remaining two cases, the target cells to be minimized are V28 and W28 instead, and the corresponding cells to change are V32:V35 and W32:W35. The 4×4 block in T38:W41, which is shaded, is the transpose of the inverse of the covariance matrix. The matrix is symmetric, it is the same as that in H38:K41 of Figure 1.

Figure 3 illustrates two further Excel-based methods to reach the required regression results for interpreting the inverse of the sample covariance matrix. They are grouped together for expositional convenience, although they differ in analytical requirements. To facilitate the regression-based computations, by using either matrix operations or the Excel function LINEST directly, the mean-removed observations are displayed in four different ways, where the dependent variable is always shown in the leading column.

Let us start with the case in Z2:AC26, under the headings of x1, x2, x3, and x4. This is the case where $\boldsymbol{x}_1$ is the dependent variable. The regression coefficients, as displayed in AA28:AC28, are based on matrix operations according to equation (43). Here, the various matrix operations have been nested, to eliminate the need for some intermediate steps during the computations. The corresponding R^2, as displayed in Z29, is based on equation (51). The variance of the dependent variable is shown in Z30. The four cells in Z33:AC33, which are shaded, are computed in accordance with the first row of the matrix in equation (8), for the case of $n = 4$. As expected, they match exactly the corresponding elements in row 1 of the inverse of the sample covariance matrix in H38:K41 (in Figure 1).

The results of the Excel function LINEST are displayed in Z37:AB41. In the case here, the dependent variable is $\boldsymbol{x}_1$ and the independent variables are $\boldsymbol{x}_2$, $\boldsymbol{x}_3$, and $\boldsymbol{x}_4$. The OLS regression coefficients $\widehat{\beta}_{14}$, $\widehat{\beta}_{13}$, and $\widehat{\beta}_{12}$ are displayed in Z37:AB37. The corresponding R^2, which is R_1^2, is shown in Z39. The explained and unexplained sums of squares (SSR and SSE) are shown in Z41:AA41. These six specific cells are shaded.

Among the LINEST results in Z37:AB41, the three cells in Z38:AB38 are the corresponding standard errors of the OLS regression coefficients in Z37:AB37. The standard error of the regression is shown in AA39. The F statistic and the degrees of freedom are shown in Z40:AA40. The three cells in AB39:AB41, with "#N/A" displayed in each case, are not utilized to show any regression results.

As the sum of SSR and SSE is SST, which is $s_{11}(m-1)$ where $m = 25$, the sample variance

	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH	AI	AJ
1		x1	x2	x3	x4			x2	x1	x3	x4	
2		2.284	1.812	1.984	1.072			1.812	2.284	1.984	1.072	
3		-2.916	0.012	-0.816	-2.928			0.012	-2.916	-0.816	-2.928	
4		-0.016	1.312	-0.216	2.572			1.312	-0.016	-0.216	2.572	
5		4.284	2.512	0.184	3.572			2.512	4.284	0.184	3.572	
6		-0.116	1.112	0.184	1.672			1.112	-0.116	0.184	1.672	
7		-0.816	-2.388	0.884	-0.728			-2.388	-0.816	0.884	-0.728	
8		-1.716	-4.088	-2.916	-3.528			-4.088	-1.716	-2.916	-3.528	
9		-0.816	-2.088	-1.216	-3.128			-2.088	-0.816	-1.216	-3.128	
10		-0.416	1.312	0.184	-0.228			1.312	-0.416	0.184	-0.228	
11		-0.516	1.012	-1.216	-0.928			1.012	-0.516	-1.216	-0.928	
12		-0.416	-0.588	0.684	0.172			-0.588	-0.416	0.684	0.172	
13		-1.716	-3.088	-0.116	-4.328			-3.088	-1.716	-0.116	-4.328	
14		1.584	-0.388	0.584	2.072			-0.388	1.584	0.584	2.072	
15		0.784	0.512	0.184	1.072			0.512	0.784	0.184	1.072	
16		-1.216	-2.488	-1.316	-1.228			-2.488	-1.216	-1.316	-1.228	
17		0.184	1.412	-0.616	-0.828			1.412	0.184	-0.616	-0.828	
18		2.184	0.612	1.684	2.072			0.612	2.184	1.684	2.072	
19		3.284	2.812	2.484	3.672			2.812	3.284	2.484	3.672	
20		-3.716	0.512	0.484	-1.828			0.512	-3.716	0.484	-1.828	
21		-2.716	-4.788	-2.216	-2.528			-4.788	-2.716	-2.216	-2.528	
22		-0.116	1.912	-0.016	-0.128			1.912	-0.116	-0.016	-0.128	
23		-1.816	-0.888	-1.616	-1.728			-0.888	-1.816	-1.616	-1.728	
24		0.384	1.712	2.784	1.072			1.712	0.384	2.784	1.072	
25		1.184	0.112	0.584	2.272			0.112	1.184	0.584	2.272	
26		2.884	2.112	-0.616	2.772			2.112	2.884	-0.616	2.772	
27												
28	Reg Coeff		0.00848	0.0262	0.72118		Reg Coeff		0.01431	0.36886	0.52212	
29	R^2	0.72633					R^2	0.58012				
30	Var	3.9364					Var	4.32777				
31												
32		Inv of Sample Cov Mat, Row 1						Inv of Sample Cov Mat, Row 2				
33		0.92828	-0.00788	-0.0243	-0.6695			0.55032	-0.0079	-0.203	-0.2873	
34								-0.0079	0.55032	-0.203	-0.2873	
35												
36		LINEST Results						LINEST Results				
37		0.72118	0.0262	0.00848				0.52212	0.36886	0.01431		
38		0.15008	0.2129	0.16415				0.25589	0.26519	0.27688		
39		0.72633	1.08407	#N/A				0.58012	1.40795	#N/A		
40		19.4632	22	#N/A				10.1321	22	#N/A		
41		68.6192	25.8544	#N/A				60.2552	43.6112	#N/A		
42												
43		Inv of Sample Cov Mat, Row 1						Inv of Sample Cov Mat, Row 2				
44		0.92828	-0.00788	-0.0243	-0.6695			0.55032	-0.0079	-0.203	-0.2873	
45								-0.0079	0.55032	-0.203	-0.2873	
46												
47	AA28:AC28	{=TRANSPOSE(MMULT(MINVERSE(MMULT(TRANSPOSE(AA2:AC26),AA2:AC26)),										
48		MMULT(TRANSPOSE(AA2:AC26),Z2:Z26)))}										
49		Copied to AG28:AI28, AL28:AO28, and AR28:AU28										
50	Z29	{=MMULT(AA28:AC28,MMULT(TRANSPOSE(AA2:AC26),Z2:Z26))/										
51		MMULT(TRANSPOSE(Z2:Z26),Z2:Z26)}										
52		Copied to AF29, AL29, and AR29										
53	Z30 =VAR(Z2:Z26)	Copied to AF30, AL30, and AR30										
54	Z33 =1/Z30/(1-Z29)	Copied to AF33, AL33, and AR33										
55	AA33 =-$Z33*AA28	Copied to AA33:AC33, AG33:AI33, AM33:AO33, and AS33:AU33										
56												
57												
58												

Figure 3 The Part of an Excel Example Illustrating Additional Approaches to Implement a Regression-Based Interpretation of the Inverse of the Sample Covariance Matrix.

	AK	AL	AM	AN	AO	AP	AQ	AR	AS	AT	AU	AV
1		x3	x1	x2	x4			x4	x1	x2	x3	
2		1.984	2.284	1.812	1.072			1.072	2.284	1.812	1.984	
3		-0.816	-2.916	0.012	-2.928			-2.928	-2.916	0.012	-0.816	
4		-0.216	-0.016	1.312	2.572			2.572	-0.016	1.312	-0.216	
5		0.184	4.284	2.512	3.572			3.572	4.284	2.512	0.184	
6		0.184	-0.116	1.112	1.672			1.672	-0.116	1.112	0.184	
7		0.884	-0.816	-2.388	-0.728			-0.728	-0.816	-2.388	0.884	
8		-2.916	-1.716	-4.088	-3.528			-3.528	-1.716	-4.088	-2.916	
9		-1.216	-0.816	-2.088	-3.128			-3.128	-0.816	-2.088	-1.216	
10		0.184	-0.416	1.312	-0.228			-0.228	-0.416	1.312	0.184	
11		-1.216	-0.516	1.012	-0.928			-0.928	-0.516	1.012	-1.216	
12		0.684	-0.416	-0.588	0.172			0.172	-0.416	-0.588	0.684	
13		-0.116	-1.716	-3.088	-4.328			-4.328	-1.716	-3.088	-0.116	
14		0.584	1.584	-0.388	2.072			2.072	1.584	-0.388	0.584	
15		0.184	0.784	0.512	1.072			1.072	0.784	0.512	0.184	
16		-1.316	-1.216	-2.488	-1.228			-1.228	-1.216	-2.488	-1.316	
17		-0.616	0.184	1.412	-0.828			-0.828	0.184	1.412	-0.616	
18		1.684	2.184	0.612	2.072			2.072	2.184	0.612	1.684	
19		2.484	3.284	2.812	3.672			3.672	3.284	2.812	2.484	
20		0.484	-3.716	0.512	-1.828			-1.828	-3.716	0.512	0.484	
21		-2.216	-2.716	-4.788	-2.528			-2.528	-2.716	-4.788	-2.216	
22		-0.016	-0.116	1.912	-0.128			-0.128	-0.116	1.912	-0.016	
23		-1.616	-1.816	-0.888	-1.728			-1.728	-1.816	-0.888	-1.616	
24		2.784	0.384	1.712	1.072			1.072	0.384	1.712	2.784	
25		0.584	1.184	0.112	2.272			2.272	1.184	0.112	0.584	
26		-0.616	2.884	2.112	2.772			2.772	2.884	2.112	-0.616	
27												
28	Reg Coeff		0.02625	0.21914	0.20072		Reg Coeff		0.71008	0.30477	0.1972	
29	R^2	0.42669					R^2	0.79872				
30	Var	1.88307					Var	5.2696				
31												
32		Inv of Sample Cov Mat, Row 3						Inv of Sample Cov Mat, Row 4				
33		0.92628	-0.0243	-0.203	-0.1859			0.94279	-0.6695	-0.2873	-0.1859	
34		-0.0243	-0.203	0.92628	-0.1859			-0.6695	-0.2873	-0.1859	0.94279	
35												
36		LINEST Results						LINEST Results				
37		0.20072	0.21914	0.02625				0.1972	0.30477	0.71008		
38		0.21079	0.15755	0.21336				0.2071	0.14937	0.14777		
39		0.42669	1.08523	#N/A				0.79872	1.07569	#N/A		
40		5.45785	22	#N/A				29.0998	22	#N/A		
41		19.2836	25.91	#N/A				101.014	25.4562	#N/A		
42												
43		Inv of Sample Cov Mat, Row 3						Inv of Sample Cov Mat, Row 4				
44		0.92628	-0.0243	-0.203	-0.1859			0.94279	-0.6695	-0.2873	-0.1859	
45		-0.0243	-0.203	0.92628	-0.1859			-0.6695	-0.2873	-0.1859	0.94279	
46												
47	AF34	=AG33		AG34	=AF33		AH34	=AH33		AI34	=AI33	
48	AL34	=AM33		AM34	=AN33		AN34	=AL33		AO34	=AO33	
49	AR34	=AS33		AS34	=AT33		AT34	=AU33		AU34	=AR33	
50	AF34:AI34			AL34:AO34				AR34:AU34				
51	Copied to AF45:AI45			Copied to AL45:AO45				Copied to AR45:AU45				
52	Z37:AB41		{=LINEST(Z2:Z26,AA2:AC26,FALSE,TRUE)}									
53				Copied to AF37:AH41, AL37:AN41, and AR37:AT41								
54	Z44	=(COUNT(Z2:Z26)-1)/((Z41+AA41)*(1-Z39))						Copied to AF44, AL44, and AR44				
55	AA44	=-$Z44*AB37		AB44	=-$Z44*AA37			AC44	=-$Z44*Z37			
56	AG44	=-$AF44*AH37		AH44	=-$AF44*AG37			AI44	=-$AF44*AF37			
57	AM44	=-$AL44*AN37		AN44	=-$AL44*AM37			AO44	=-$AL44*AL37			
58	AS44	=-$AR44*AT37		AT44	=-$AR44*AS37			AU44	=-$AR44*AR37			

Figure 3 The Part of an Excel Example Illustrating Additional Approaches to Implement a Regression-Based Interpretation of the Inverse of the Sample Covariance Matrix (Continued).

s_{11} can easily be deduced. Also shaded are the four cells in Z44:AC44, which are computed according to the four elements in the first row of the matrix in equation (8) for $n = 4$. As expected, the displayed values in Z44:AC44 match exactly those in Z33:AC33.

The computations in Z28:AC44 are repeated for AF28:AI44, AL28:AO44, and AR28:AU44, but with a different dependent variable in each case. This requires rearrangements of the mean-removed observations for AF2:AI26, AL2:AO26, and AR2:AU26, along with the corresponding headings, before repeating the computations. Subsequent to such repeated computations, rearrangements of the end results are required, so that each set of the displayed results pertains to the original order of the four variables. They are shown as shaded cells in AF34:AU34 and in AF45:AU45; they correspond to rows 2 to 4 of the inverse of the sample covariance matrix. As expected, the results here (including those in Z33:AC33 and Z44:AC44, which have already been mentioned earlier) match exactly those in H38:K41 of Figure 1.

Notice that the observations for use in the Excel function LINEST need not be mean-removed. Without mean removal, a constant term must be present in each regression model. The third argument of the function, which is "false" (to indicate the absence of the constant term) in Figure 3, becomes "true" instead. Further, each 5×3 block for displaying the LINEST result for each regression model will have to be expanded to be a 5×4 block because of the presence of an extra regression coefficient. However, there are no changes in the three regression coefficients; nor is there any change in the corresponding R^2.

Notice also that, instead of using the function LINEST, each OLS regression run can be performed equally well by selecting "Regression" in Excel's add-in tools for "Data Analysis." All that is required is to respond to the dialog box, for input ranges of the variables involved and various regression and output options. However, even for basic options, the corresponding display of the regression results contains more information than what is required to interpret the inverse of the sample covariance matrix. Thus, to avoid unnecessary digressions, no illustration based on such an add-in tool in Excel is provided here.

7 Concluding Remarks

Given the relevance of the covariance matrix and its inverse across different academic disciplines, this paper has illustrated pedagogically a regression-based interpretation of the inverse of the sample covariance matrix. The approach here draws on a finance article in mean-variance portfolio analysis by Stevens (1998). The Stevens article is innovative in connecting elements of the sample covariance matrix and its inverse, by exploiting analytical similarities between some expressions of block matrix operations and ordinary least squares (OLS) regression analysis.

The covariance matrix of security returns is part of the input parameters for mean-variance portfolio analysis. As the true values of its elements are unknown, the use of the sample covariance matrix of security returns for it, when implementing a portfolio selection model, is justified under the stationarity assumption of the return distributions. Thus, Stevens' interpretation of inverse of the sample covariance matrix of returns allows the results of mean-variance portfolio analysis to be understood much better. As indicated in this paper, Stevens' interpretation need not be confined to financial settings. In fact, the same idea holds for any empirical or experimental settings, where sample covariance matrices and their inverses are involved.

Although matrix inversion is required in many analytical settings, the corresponding textbook materials in matrix algebra do not provide any guidance to interpret the result intuitively when a square matrix is inverted. As standard textbook coverage on matrix inversion is primarily on the technical aspect of the task, it is not immediately obvious how textbook materials can improve our understanding of analytical models involving matrix inversion. In view of difficulties in interpreting intuitively the inverse of a general square matrix, the scope of this paper has been confined to the sample covariance matrix only. The objective of this paper is to make the inversion of sample covariance matrices less mysterious, from a pedagogic perspective.

Excel has played an important pedagogic role in this paper. The availability of various Excel tools for numerical illustrations will give instructors flexibility in how the analytical materials in this paper are covered in their classes. An analytical coverage, where details of the matrix operations involved are provided, can be illustrated numerically by using various Excel functions for matrix operations. A less rigorous coverage of the same materials can bypass matrix operations in the corresponding Excel illustration; Excel Solver is suitable for the task involved. For classes where analytical details are de-emphasized, the regression-based interpretation can be illustrated directly by using Excel's linear regression tools.

Notice that the expression of the inverse of the sample covariance matrix, as derived by Stevens and presented pedagogically in this paper, is not intended to be an alternative to the standard expression in terms of the determinant and cofactors of the matrix. Rather, it is intended to provide a regression-based interpretation of such a matrix inverse. Without the individual observations, the sample covariance matrix itself does not allow the regression coefficients and the goodness-of-fit measure of each regression model to be computed.

References

Cheung, C.S., Kwan, C.C.Y., and Mountain, D.C., (2009), "On the Nature of Mean-Variance Spanning," *Finance Research Letters*, 6, 106-113.

DeCoursey, W.J., (2003), *Statistics and Probability for Engineering Applications with Microsoft® Excel*, Elsevier Science, Woburn, MA.

Elton, E.J., Gruber, M.J., Brown, S.J., and Goetzmann W.N., (2010), *Portfolio Theory and Investment Analysis*, Eighth Edition, Wiley & Sons, Hoboken, NJ.

Johnston, J., (1972), *Econometric Methods*, Second Edition, McGraw-Hill Book Company, New York, NY.

Kwan, C.C.Y., (1984), "Portfolio Analysis Using Single Index, Multi-Index, and Constant Correlation Models: A Unified Treatment," *Journal of Finance*, 39(5), 1469-1483.

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.

Kwan, C.C.Y., (2010), "The Requirement of a Positive Definite Covariance Matrix of Security Returns for Mean-Variance Portfolio Analysis: A Pedagogic Illustration," *Spreadsheets in Education*, 4 (1), Article 4.

Kwan, C.C.Y., (2011), "An Introduction to Shrinkage Estimation of the Covariance Matrix: A Pedagogic Illustration," *Spreadsheets in Education*, 4 (3), Article 6.

Larson, R., and Falvo, D.C., (2009), *Elementary Linear Algebra*, Sixth Edition, Houghton Mifflin Harcourt Publishing Company, Boston, MA.

Maddala, G.S., (1977), *Econometrics*, McGraw-Hill Book Company, New York, NY.

Reiner, I., (1971), *Introduction to Matrix Theory and Linear Algebra*, Holt, Reinehart and Winston, Inc., New York, NY.

Stevens, G.V.G., (1998), "On the Inverse of the Covariance Matrix in Portfolio Analysis," *Journal of Finance*, 53(5), 1821-1827.

Appendix: Block Matrix Inversion

The product of the two matrices in equations (53) and (54) of the main text is

$$\boldsymbol{V}\boldsymbol{V}^{-1} = \boldsymbol{I} = \begin{bmatrix} \boldsymbol{I}_{11} & \mathbf{0}_{12} \\ \mathbf{0}_{21} & \boldsymbol{I}_{22} \end{bmatrix}, \tag{A1}$$

where $\boldsymbol{I}$, $\boldsymbol{I}_{11}$, and $\boldsymbol{I}_{22}$ are identity matrices with the same dimensions as $\boldsymbol{V}$, $\boldsymbol{V}_{11}$, and $\boldsymbol{V}_{22}$, respectively. The matrices $\mathbf{0}_{12}$ and $\mathbf{0}_{21}$, which are of the same dimensions as $\boldsymbol{V}_{12}$ and $\boldsymbol{V}_{21}$, respectively, have all zero elements. We are interested in expressing $\boldsymbol{U}_{11}$, $\boldsymbol{U}_{12}$, $\boldsymbol{U}_{21}$, and $\boldsymbol{U}_{22}$ in terms of $\boldsymbol{V}_{11}$, $\boldsymbol{V}_{12}$, $\boldsymbol{V}_{21}$, and $\boldsymbol{V}_{22}$.[13]

For a given invertible $\boldsymbol{V}$, its inverse is unique. However, the expression of $\boldsymbol{V}^{-1}$ in terms of block matrices is not unique; it depends on whether $\boldsymbol{V}_{11}^{-1}$ or $\boldsymbol{V}_{22}^{-1}$ is explicitly displayed. The version where $\boldsymbol{V}_{11}^{-1}$ is explicitly displayed is derived below.

Based on

$$\begin{bmatrix} \boldsymbol{V}_{11} & \boldsymbol{V}_{12} \\ \boldsymbol{V}_{21} & \boldsymbol{V}_{22} \end{bmatrix} \begin{bmatrix} \boldsymbol{U}_{11} & \boldsymbol{U}_{12} \\ \boldsymbol{U}_{21} & \boldsymbol{U}_{22} \end{bmatrix} = \begin{bmatrix} \boldsymbol{I}_{11} & \mathbf{0}_{12} \\ \mathbf{0}_{21} & \boldsymbol{I}_{22} \end{bmatrix}, \tag{A2}$$

we can write

$$\boldsymbol{V}_{11}\boldsymbol{U}_{11} + \boldsymbol{V}_{12}\boldsymbol{U}_{21} = \boldsymbol{I}_{11}, \tag{A3}$$

$$\boldsymbol{V}_{11}\boldsymbol{U}_{12} + \boldsymbol{V}_{12}\boldsymbol{U}_{22} = \mathbf{0}_{12}, \tag{A4}$$

$$\boldsymbol{V}_{21}\boldsymbol{U}_{11} + \boldsymbol{V}_{22}\boldsymbol{U}_{21} = \mathbf{0}_{21}, \tag{A5}$$

and

$$\boldsymbol{V}_{21}\boldsymbol{U}_{12} + \boldsymbol{V}_{22}\boldsymbol{U}_{22} = \boldsymbol{I}_{22}. \tag{A6}$$

[13] Notice that, for the purpose of deriving the expressions of $\boldsymbol{U}_{11}$ and $\boldsymbol{U}_{12}$ that are immediately ready to be connected to the corresponding results of multiple linear regression, we could have partitioned $\boldsymbol{V}$ in such a way that $\boldsymbol{V}_{11}$ and $\boldsymbol{V}_{12}$ would become 1×1 and $1 \times (n-1)$ matrices, respectively. However, a general partitioning of $\boldsymbol{V}$ is presented here instead, in order to make the corresponding materials in block matrix inversion more versatile.

These four matrix equations allow us to solve $\boldsymbol{U}_{11}$, $\boldsymbol{U}_{12}$, $\boldsymbol{U}_{21}$, and $\boldsymbol{U}_{22}$ in terms of $\boldsymbol{V}_{11}$, $\boldsymbol{V}_{12}$, $\boldsymbol{V}_{21}$, and $\boldsymbol{V}_{22}$. Specifically, equation (A5) gives us

$$\boldsymbol{U}_{21} = -\boldsymbol{V}_{22}^{-1}\boldsymbol{V}_{21}\boldsymbol{U}_{11}, \tag{A7}$$

which, when combined with equation (A3), leads to

$$\boldsymbol{V}_{11}\boldsymbol{U}_{11} - \boldsymbol{V}_{12}\boldsymbol{V}_{22}^{-1}\boldsymbol{V}_{21}\boldsymbol{U}_{11} = \boldsymbol{I}_{11} \tag{A8}$$

and then

$$(\boldsymbol{V}_{11} - \boldsymbol{V}_{12}\boldsymbol{V}_{22}^{-1}\boldsymbol{V}_{21})\boldsymbol{U}_{11} = \boldsymbol{I}_{11}. \tag{A9}$$

As

$$\boldsymbol{U}_{11}^{-1}\boldsymbol{U}_{11} = \boldsymbol{I}_{11}, \tag{A10}$$

it follows that

$$\boldsymbol{U}_{11} = (\boldsymbol{V}_{11} - \boldsymbol{V}_{12}\boldsymbol{V}_{22}^{-1}\boldsymbol{V}_{21})^{-1}. \tag{A11}$$

With $\boldsymbol{U}_{11}$ known, equation (A7) can be used to determine $\boldsymbol{U}_{21}$. Next, according to equation (A6), we have

$$\boldsymbol{U}_{22} = \boldsymbol{V}_{22}^{-1}(\boldsymbol{I}_{22} - \boldsymbol{V}_{21}\boldsymbol{U}_{12}) = \boldsymbol{V}_{22}^{-1} - \boldsymbol{V}_{22}^{-1}\boldsymbol{V}_{21}\boldsymbol{U}_{12}, \tag{A12}$$

which, when combined with equation (A4), leads to

$$\boldsymbol{V}_{11}\boldsymbol{U}_{12} + \boldsymbol{V}_{12}(\boldsymbol{V}_{22}^{-1} - \boldsymbol{V}_{22}^{-1}\boldsymbol{V}_{21}\boldsymbol{U}_{12}) = \boldsymbol{0}_{12}. \tag{A13}$$

Equation (A13) can be written as

$$(\boldsymbol{V}_{11} - \boldsymbol{V}_{12}\boldsymbol{V}_{22}^{-1}\boldsymbol{V}_{21})\boldsymbol{U}_{12} = -\boldsymbol{V}_{12}\boldsymbol{V}_{22}^{-1}, \tag{A14}$$

which leads to

$$\boldsymbol{U}_{12} = -\boldsymbol{U}_{11}\boldsymbol{V}_{12}\boldsymbol{V}_{22}^{-1}, \tag{A15}$$

where $\boldsymbol{U}_{11}$ is given by equation (A11). Finally, returning to equation (A12) with $\boldsymbol{U}_{12}$ determined, we can write

$$\boldsymbol{U}_{22} = \boldsymbol{V}_{22}^{-1} + \boldsymbol{V}_{22}^{-1}\boldsymbol{V}_{21}\boldsymbol{U}_{11}\boldsymbol{V}_{12}\boldsymbol{V}_{22}^{-1}. \tag{A16}$$

The combined results of $\boldsymbol{U}_{11}$, $\boldsymbol{U}_{12}$, $\boldsymbol{U}_{21}$, and $\boldsymbol{U}_{22}$ are as displayed in equations (55) and (56) of the main text.