A linear test for the global minimum variance portfolio for small sample and singular covariance

Bodnar, Taras; Mazur, Stepan; Podgorski, Krzysztof

2015

Working Papers in Statistics
No 2015:10

Department of Statistics
School of Economics and Management
Lund University

A linear test for the global minimum variance portfolio for small sample and singular covariance

TARAS BODNAR, STOCKHOLM UNIVERSITY

STEPAN MAZUR, LUND UNIVERSITY

KRZYSZTOF PODGÓRSKI, LUND UNIVERSITY

A linear test for the global minimum variance portfolio for small sample and singular covariance

Taras Bodnar[a], Stepan Mazur[b] and Krzysztof Podgórski[b] [1]

[a] *Department of Mathematics, Stockholm University, Roslagsvägen 101, SE-10691 Stockholm, Sweden*

[b] *Department of Statistics, Lund University, Tycho Brahes väg 1, SE-22007 Lund, Sweden*

Abstract

Bodnar and Schmid (2008) derived the distribution of the global minimum variance portfolio weights and obtained the distribution of the test statistics for the general linear hypothesis. Their results are obtained in the case when the number of observations n is bigger or equal than the size of portfolio k. In the present paper, we extend the result by analyzing the portfolio weights in a small sample case of $n < k$, with the singular covariance matrix. The results are illustrated using actual stock returns. A discussion of practical relevance of the model is presented.

ASM Classification: 91G10, 62H12

Keywords: global minimum variance portfolio, singular Wishart distribution, singular covariance matrix, small sample problem

[1] The authors appreciate the financial support of the Swedish Research Council Grant Dnr: 2013-5180 and Riksbankens Jubileumsfond Grant Dnr: P13-1024:1

1 Introduction

Starting from 1952, when H. Markowitz published his seminal paper on portfolio selection and provided the mathematical foundation for the problem (see Markowitz (1952)), the portfolio theory has become a well established branch of finance. The variance was introduced as a measurement of risk to complement the expected return as the main criteria for portfolio construction. The global minimum variance (GMV) portfolio was defined as the asset portfolio with the lowest return variance for a given covariance matrix, i.e.

$$\mathbf{w}_{GMV} = \text{argmin}\{\mathbf{w}^T\mathbf{\Sigma}\mathbf{w}; \mathbf{w}^T\mathbf{1}_k = 1\}, \tag{1}$$

where $\mathbf{w} = (w_1, ..., w_k)^T$ is the vector of portfolio weights, $\mathbf{1}_k$ denotes the vector whose components are all equal to 1, and $\mathbf{\Sigma}$ stands for the covariance matrix of the asset returns. If $\mathbf{\Sigma}$ is a positive definite matrix, then

$$\mathbf{w}_{GMV} = \frac{\mathbf{\Sigma}^{-1}\mathbf{1}_k}{\mathbf{1}_k^T\mathbf{\Sigma}^{-1}\mathbf{1}_k}.$$

The importance of the GMV portfolio in the financial applications is well motivated by Merton (1980), Best and Grauer (1991), Chopra and Ziemba (1993), Glombek (2014) and others. Note that the GMV portfolio is the limiting case of the expected quadratic utility portfolio, when the risk aversion increases without bound, i.e. the case of a fully risk-averse investor.

A number of papers were written on this topic for the case when the number of observations is larger than or equal to the dimension of portfolio, see Okhrin and Schmid (2006), Bodnar and Schmid (2008), Bodnar and Okhrin (2011) among others. This assumption realistically may be not valid for portfolios which consist of a large number of assets, in particular, when historical data on assets are scarce due to the assumption of historical independence of observations. For these reasons the case when the number of observations is less than the size of portfolio assets is of importance. We also note the large portfolio case rises also the issues of singularity of the covariance matrix – large portfolios bring together assets that are dependent and thus singularities become intrinsically present in the data. However, these particular cases of singular covariance and small sample relatively to the portofolio size have not been given sufficient attention in the portfolio theory. This motivates our work in which we extend the results obtained by Bodnar and Schmid (2008), where the case of the non-singular covariance matrix and the sample size is larger than the number of portfolio assets has been worked out.

More precisely, we work under two 'singular' conditions. Namely, we assume that $\mathbf{\Sigma}$ is singular and that the sample size n is smaller than the portfolio size k. Under these conditions, one has two distinct cases: $\text{rank}(\mathbf{\Sigma}) = r < n$ and $\text{rank}(\mathbf{\Sigma}) = r \geq n$. While the

latter is currently an open question in this theory, we solve completely the former using previously obtained results of Bodnar et al. (2015). The presented results are random scale invariant, i.e. they are valid under the assumption that the matrix of returns has a matrix normal distribution scaled by an arbitrary random variable. In particular, it is valid for the elliptically contoured distributions that are very useful in finance, see, for example, Owen and Rabinovitch (1983), Zhou (1993), Berk (1997), Bodnar and Schmid (2007).

The main contribution of the paper is derivation of a test for the hypothesis of linear combinations between GMV portfolio weights for the case of the portfolio size exceeding the sample size and the covariance matrix with rank smaller than the sample size. For a financial market analyst, such a test is an important tool for setting investment strategies and we illustrate this using actual return data of diversified stocks coming from several industries.

2 The global minimal variance portfolio

We consider a portfolio consisting of k assets and $\mathbf{x}_t$ denotes the k-dimensional vector of the log-returns of these assets at time t. Throughout the paper it is assumed that $\mathbf{x}_1, .., \mathbf{x}_n$ are vectors with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$, which is a non-negative definite matrix with $\text{rank}(\boldsymbol{\Sigma}) = r < n$.

Since $\boldsymbol{\Sigma}$ is singular, the Moore-Penrose inverse $\mathbf{A}^+$ of a matrix $\mathbf{A}$ will be employed as an important tool of analysis. Next, we revise the definition of the Moore-Penrose inverse. A matrix $\mathbf{A}^+$ is the Moore-Penrose of $\mathbf{A}$ if the following conditions hold (see Horn and Johnsohn (1985))

(I) $\mathbf{A}\mathbf{A}^+\mathbf{A} = \mathbf{A}$,

(II) $\mathbf{A}^+\mathbf{A}\mathbf{A}^+ = \mathbf{A}^+$,

(III) $(\mathbf{A}\mathbf{A}^+)^T = \mathbf{A}\mathbf{A}^+$,

(IV) $(\mathbf{A}^+\mathbf{A})^T = \mathbf{A}^+\mathbf{A}$.

The optimization problem (1) has an infinite number of solutions, since in our setup $\boldsymbol{\Sigma}$ does not have a full rank. One of the solutions can be expressed as

$$\mathbf{w}_{GMV} = \frac{\boldsymbol{\Sigma}^+\mathbf{1}_k}{\mathbf{1}_k^T\boldsymbol{\Sigma}^+\mathbf{1}_k}, \tag{2}$$

which is the unique minimal Euclidean norm solution given that $\mathbf{1}_k^T\boldsymbol{\Sigma}^+\mathbf{1}_k \neq 0$, see Pappas et al. (2010). However, if the latter is not true, then equation (2) cannot be used. Therefore it is important to test from the data if the condition $\mathbf{1}_k^T\boldsymbol{\Sigma}^+\mathbf{1}_k = 0$ is occurring.

At the end of this section we discuss how this problem can be addressed although formal statistical tests are not know at the moment.

In the reality, the unknown parameters $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ have to be estimated. The sample mean vector and the sample covariance matrix are used for this purpose

$$\overline{\mathbf{x}} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_i \quad \text{and} \quad \mathbf{S} = \frac{1}{n-1}\sum_{i=1}^{n}(\mathbf{x}_i - \overline{\mathbf{x}})(\mathbf{x}_i - \overline{\mathbf{x}})^T. \tag{3}$$

Replacing $\boldsymbol{\Sigma}$ with $\mathbf{S}$ in (2) we obtain a sample estimator for GMV portfolio weights

$$\widehat{\mathbf{w}}_{GMV} = \frac{\mathbf{S}^+\mathbf{1}_k}{\mathbf{1}_k^T\mathbf{S}^+\mathbf{1}_k}. \tag{4}$$

We consider the linear transformation of the GMV portfolio weights

$$\boldsymbol{\theta} = \mathbf{L}\mathbf{w}_{GMV} = \frac{\mathbf{L}\boldsymbol{\Sigma}^+\mathbf{1}_k}{\mathbf{1}_k^T\boldsymbol{\Sigma}^+\mathbf{1}_k}, \tag{5}$$

where $\mathbf{L}$ is the $p \times k$ matrix of constants with $\text{rank}(\mathbf{L}) = p < r$. The corresponding sample estimator of (5) is given by

$$\widehat{\boldsymbol{\theta}} = \mathbf{L}\widehat{\mathbf{w}}_{GMV} = \frac{\mathbf{L}\mathbf{S}^+\mathbf{1}_k}{\mathbf{1}_k^T\mathbf{S}^+\mathbf{1}_k}. \tag{6}$$

The density function of $\widehat{\boldsymbol{\theta}}$ under the assumption of normality is summarized in the following theorem, the proof of which can be found in Bodnar et al. (2015).

Theorem 1. *Let $\mathbf{x}_1, ..., \mathbf{x}_n$ be i.i.d. random vectors with $\mathbf{x}_1 \sim \mathcal{N}_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $k \geq n$ and with $rank(\boldsymbol{\Sigma}) = r < n$. Consider a $p \times k$ matrix $\mathbf{L}$ of the full rank p with the rows linearly independent of $\mathbf{1}_k^T$. Then the density function of $\widehat{\boldsymbol{\theta}} = \mathbf{L}\widehat{\mathbf{w}}_{GMV}$ is given by*

$$\widehat{\boldsymbol{\theta}} \sim t_p\left(n-r+1; \boldsymbol{\theta}, \frac{1}{n-r+1}\frac{\mathbf{L}\mathbf{R}\mathbf{L}^T}{\mathbf{1}_k^T\boldsymbol{\Sigma}^+\mathbf{1}_k}\right),$$

where $\mathbf{R} = \boldsymbol{\Sigma}^+ - \boldsymbol{\Sigma}^+\mathbf{1}_k\mathbf{1}_k^T\boldsymbol{\Sigma}^+/\mathbf{1}_k^T\boldsymbol{\Sigma}^+\mathbf{1}_k$. The symbol $t_p(d; \mathbf{a}, \mathbf{A})$ stands for the p-dimensional multivariate t-distribution with d degrees of freedom, the location parameter $\mathbf{a}$, and the dispersion matrix $\mathbf{A}$.

This result is similar to the one obtained under the assumption of the non-singularity (see Bodnar and Schmid (2008)). The only difference is in the degrees of freedom of the t-distribution. Applying the properties of the multivariate t-distribution we obtain

$$E(\widehat{\boldsymbol{\theta}}) = \boldsymbol{\theta} \quad \text{and} \quad Var(\widehat{\boldsymbol{\theta}}) = \frac{1}{n-r-1}\frac{\mathbf{L}\mathbf{R}\mathbf{L}^T}{\mathbf{1}_k^T\boldsymbol{\Sigma}^+\mathbf{1}_k}. \tag{7}$$

Next, by using the above result, we derive a confidence interval for a linear combination of the GMV portfolio weights. Without loss of generality we deal with the first weight of the GMV portfolio only and note that the confidence intervals for other weights can be obtained similarly. Let $\mathbf{L} = \mathbf{e}_1^T = (1, 0, ..., 0)$, then the distribution for $\widehat{\theta} = \mathbf{e}_1^T \mathbf{S}^+ \mathbf{1}_k / \mathbf{1}_k^T \mathbf{S}^+ \mathbf{1}_k$ is expressed as

$$\widehat{\theta} \sim t\left(n - r + 1; \frac{\mathbf{e}_1^T \mathbf{\Sigma}^+ \mathbf{1}_k}{\mathbf{1}_k^T \mathbf{\Sigma}^+ \mathbf{1}_k}, \frac{1}{n - r + 1} \frac{\mathbf{e}_1^T \mathbf{R} \mathbf{e}_1}{\mathbf{1}_k^T \mathbf{\Sigma}^+ \mathbf{1}_k}\right). \tag{8}$$

The application of (8) leads to the $(1 - \alpha)$-confidence interval for the first weight of the GMV portfolio given by

$$\frac{\mathbf{e}_1^T \mathbf{\Sigma}^+ \mathbf{1}_k}{\mathbf{1}_k^T \mathbf{\Sigma}^+ \mathbf{1}_k} \pm \frac{1}{\sqrt{n - r + 1}} \frac{\sqrt{\mathbf{e}_1^T \mathbf{R} \mathbf{e}_1}}{\sqrt{\mathbf{1}_k^T \mathbf{\Sigma}^+ \mathbf{1}_k}} t_{n-r+1;\alpha/2},$$

where $t_{m;\beta}$ denotes the β-quantile of the t-distribution with m degrees of freedom.

We conclude this section with three remarks.

Remark 1. There is an interesting property which follows directly from Theorem 1. Namely, it yields that $\widehat{\mathbf{w}}_{GMV}$ also has a multivariate t-distribution but it is a singular one. This happens in the case when the dispersion matrix $\mathbf{A}$ in the definition of t-distribution is singular. Formally such a distribution has a linear subspace $U \subset \mathbb{R}^k$ as its support and on this support it has a regular multivariate t-distribution. As a singular distribution, it does not have a density with respect to volume measure in $\mathbb{R}^k$, but it does have a density with respect to the (lower-dimension) volume measure on U. For more properties of singular multivariate t-distributions see Gupta and Nagar (2000). Thus the singularity of $\widehat{\mathbf{w}}_{GMV}$ follows directly from its characteristic function $\varphi_{\widehat{\mathbf{w}}_{GMV}}(\mathbf{t}) = \varphi_{\mathbf{t}^T \widehat{\mathbf{w}}_{GMV}}(1)$, and the fact that $\mathbf{R}$ is a singular matrix since $\mathbf{R}\mathbf{1}_k = \mathbf{0}$.

Remark 2. We note that Theorem 1, as well as all the result in this work, is scale invariant in the following sense. Let $[\tilde{\mathbf{x}}_1 \dots \tilde{\mathbf{x}}_n] = R[\mathbf{x}_1 \dots \mathbf{x}_n]$, where R is an arbitrary, possible random scale. Then $\widehat{\boldsymbol{\theta}}$ evaluated for the scaled matrix is exactly the same as the one original one. In particular, by a proper choice of R, the matrix $R[\mathbf{x}_1 \dots \mathbf{x}_n]$ can be elliptically contoured. However, practical consequences of such an extension are limited due to the fact that only a single value of scaling R is involved and thus such data are equivalent to non-random scaling normal matrix variates.

Remark 3. The case $\mathbf{1}_k^T \mathbf{\Sigma}^+ \mathbf{1}_k = 0$ is only trivially encountered in practice. To see this note that this condition is equivalent to $\mathbf{1}_k^T \mathbf{\Sigma} \mathbf{1}_k = 0$. Indeed, since $\mathbf{\Sigma}$ is the singular covariance matrix with $\text{rank}(\mathbf{\Sigma}) = r$, we obtain that $\mathbf{\Sigma} = \mathbf{H}\mathbf{D}\mathbf{H}^T$, where $\mathbf{H}$ is an $k \times r$ orthogonal matrix such that $\mathbf{H}^T\mathbf{H} = \mathbf{I}_r$, and $\mathbf{D}$ is a $r \times r$ positive definite diagonal matrix. Then it holds that $\mathbf{\Sigma}^+ = \mathbf{H}\mathbf{D}^{-1}\mathbf{H}^T$. As a result, we get that $\mathbf{1}^T \mathbf{\Sigma}^+ \mathbf{1} = 0$ if and only if $\mathbf{1}^T \mathbf{H}\mathbf{D}^{-1}\mathbf{H}^T \mathbf{1} = 0$. From the last equality we get that $\mathbf{H}^T \mathbf{1} = \mathbf{0}$ because $\mathbf{D}^{-1}$ is the

positive definite matrix. On the another hand, the equlity $\mathbf{1}_k^T\boldsymbol{\Sigma}\mathbf{1}_k = 0$ means that $\mathbf{x}_t^T\mathbf{1}_k$, $t = 1, \ldots, n$, are deterministic, i.e. for $\mathbf{X} = (\mathbf{x}_1, ..., \mathbf{x}_n)$ is the $k \times n$ matrix of returns, $\mathbf{X}^T\mathbf{1}_k = (\boldsymbol{\mu}^T\mathbf{1}_k, ..., \boldsymbol{\mu}^T\mathbf{1}_k)^T$.

In practice, $\mathbf{X}^T\mathbf{1}_k$ will typically be random and only its small variation can be an indication of a problem. Investigating the variance of this vector becomes equivalent to $\mathbf{1}_k^T\boldsymbol{\Sigma}^+\mathbf{1}_k \approx 0$. One can thus choose a small value δ and test if $Var(\mathbf{x}_t^T\mathbf{1}_k) \leq \delta$. Since $\mathbf{x}_t^T\mathbf{1}_k$ are independent identically distributed variables it leads to the classical test for the variance of iid variables. The choice of δ is an interesting problem that deserve a separate study.

3 The main result

Our results by assuming that the portfolio size is larger than the sample size can be useful for a large portfolio size. For such large portfolios, an investor maybe interested in knowing whether the weights of the GMW portfolio fulfill some linear restrictions. Such a need can arise, for example, in a decision of changing investing strategy from the one that was previously established under different market conditions. This can be formulate as a testing hypothesis problem in the following way

$$H_0 : \mathbf{L}\mathbf{w}_{GMV} = \mathbf{r} \quad \text{against} \quad H_1 : \mathbf{L}\mathbf{w}_{GMV} \neq \mathbf{r} \tag{9}$$

with the following test statistics which extends the one introduced by Bodnar and Schmid (2008) to the case of singular covariance matrix

$$T = \frac{n-r}{p}(\mathbf{1}_k^T\mathbf{S}^+\mathbf{1}_k)(\widehat{\boldsymbol{\theta}} - \mathbf{r})^T\widehat{\mathbf{R}}_{\mathbf{L}}^{-1}(\widehat{\boldsymbol{\theta}} - \mathbf{r}), \tag{10}$$

where $\mathbf{r} \in R^p$ is a vector of constants and $\widehat{\mathbf{R}}_{\mathbf{L}} = \mathbf{L}\widehat{\mathbf{R}}\mathbf{L}^T$ with $\widehat{\mathbf{R}} = \mathbf{S}^+ - \mathbf{S}^+\mathbf{1}_k\mathbf{1}_k^T\mathbf{S}^+/\mathbf{1}_k^T\mathbf{S}^+\mathbf{1}_k$. It is noted that the test statistic (10) is a generalization of the multivariate test for the mean vector (see Muirhead (1982)).

Let $\mathcal{F}_{i,j}$ stand for the $\mathcal{F}$-distribution with i and j degrees of freedom and its density function we denote by $f_{\mathcal{F}_{i,j}}$. Also, let ${}_2F_1(a, b, c; x)$ be the hypergeometric function (see Chapter 15 of Abramowitz and Stegun (1984)), that is,

$${}_2F_1(a, b, c; x) = \frac{\Gamma(c)}{\Gamma(a)\Gamma(b)} \sum_{i=0}^{\infty} \frac{\Gamma(a+i)\Gamma(b+i)}{\Gamma(c+i)} \frac{x^i}{i!}.$$

The density function of the test statistic (10) is presented in Theorem 2.

Theorem 2. *Let $\mathbf{x}_1, ..., \mathbf{x}_n$ be i.i.d. random vectors with $\mathbf{x}_1 \sim \mathcal{N}_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $k \geq n$ and with $rank(\boldsymbol{\Sigma}) = r < n$. Consider $\mathbf{L}$ a $p \times k$ non-random matrix with $rank(\mathbf{L}^T, \mathbf{1}_k) = p+1 \leq r$.*

Then the density function of T is given by

$$\begin{aligned} f_T(x) &= f_{\mathcal{F}_{p,n-r}}(x)(1+\lambda)^{-(n-r+p)/2} \\ &\quad \times {}_2F_1\left(\frac{n-r+p}{2}, \frac{n-r+p}{2}, \frac{p}{2}; \frac{px}{n-r+px}\frac{\lambda}{1+\lambda}\right), \end{aligned}$$

with $\lambda = \mathbf{1}_k^T \mathbf{\Sigma}^+ \mathbf{1}_k (\boldsymbol{\theta} - \mathbf{r})^T (\mathbf{LRL}^T)^{-1} (\boldsymbol{\theta} - \mathbf{r})$. *Moreover, under the null hypothesis it holds that* $T \sim \mathcal{F}_{p,n-r}$

Proof. First we demonstrate that

$$T|(n-1)\widehat{\mathbf{R}}_{\mathbf{L}}^{-1} = \mathbf{C} \sim F_{p,n-r,\lambda(\mathbf{C})},$$

where $\widehat{\mathbf{R}}_{\mathbf{L}}$ is defined through

$$\begin{aligned} \widetilde{\mathbf{S}} &= \begin{pmatrix} \mathbf{LS}^+\mathbf{L}^T & \mathbf{LS}^+\mathbf{1}_k \\ \mathbf{1}_k^T\mathbf{S}^+\mathbf{L}^T & \mathbf{1}_k^T\mathbf{S}^+\mathbf{1}_k \end{pmatrix} = \begin{pmatrix} \widetilde{\mathbf{S}}_{11} & \widetilde{\mathbf{S}}_{12} \\ \widetilde{\mathbf{S}}_{21} & \widetilde{S}_{22} \end{pmatrix}, \\ \widetilde{\mathbf{S}}^{-1} &= \begin{pmatrix} \widehat{\mathbf{R}}_{\mathbf{L}}^{-1} & -\widehat{\mathbf{R}}_{\mathbf{L}}^{-1}\widehat{\boldsymbol{\theta}} \\ -\widehat{\boldsymbol{\theta}}^T\widehat{\mathbf{R}}_{\mathbf{L}}^{-1} & (\widetilde{S}_{22} - \widetilde{\mathbf{S}}_{21}\widetilde{\mathbf{S}}_{11}^{-1}\widetilde{\mathbf{S}}_{12})^{-1} \end{pmatrix}. \end{aligned}$$

Because $(n-1)\mathbf{S} \sim \mathcal{W}_k(n-1, \mathbf{\Sigma})$, $k > n-1$, has a singular Wishart distribution (see Theorem 4 of Bodnar *et al.* (2015)) and we get from Theorem 1 of Bodnar *et al.* (2015) and Theorem 3.4.1 of Gupta and Nagar (2000) that the random matrix $\widetilde{\mathbf{S}} = \{\widetilde{\mathbf{S}}_{ij}\}_{i,j=1,2}$ has the $(p+1)$-variate inverse Wishart distribution with $(n-r+2p+2)$ degrees of freedom and the non-singular covariance matrix $\widetilde{\mathbf{\Sigma}}$, i.e. $\widetilde{\mathbf{S}} \sim \mathcal{W}_{p+1}(n-r+2p+2, \widetilde{\mathbf{\Sigma}})$, where

$$\widetilde{\mathbf{\Sigma}} = \begin{pmatrix} \mathbf{L\Sigma}^+\mathbf{L}^T & \mathbf{L\Sigma}^+\mathbf{1}_k \\ \mathbf{1}_k^T\mathbf{\Sigma}^+\mathbf{L}^T & \mathbf{1}_k^T\mathbf{\Sigma}^+\mathbf{1}_k \end{pmatrix} = \begin{pmatrix} \widetilde{\mathbf{\Sigma}}_{11} & \widetilde{\mathbf{\Sigma}}_{12} \\ \widetilde{\mathbf{\Sigma}}_{21} & \widetilde{\Sigma}_{22} \end{pmatrix}.$$

It is easy to see that $\widehat{\boldsymbol{\theta}} = \widetilde{S}_{22}^{-1}\widetilde{\mathbf{S}}_{12}$ and $\boldsymbol{\theta} = \widetilde{\Sigma}_{22}^{-1}\widetilde{\mathbf{\Sigma}}_{12}$.

We can factorize T as follows

$$T = \frac{n-r}{p}\frac{(\mathbf{1}_k^T\mathbf{\Sigma}^+\mathbf{1}_k)(\widehat{\boldsymbol{\theta}} - \mathbf{r})^T\widehat{\mathbf{R}}_{\mathbf{L}}^{-1}(\widehat{\boldsymbol{\theta}} - \mathbf{r})}{\mathbf{1}_k^T\mathbf{\Sigma}^+\mathbf{1}_k / \mathbf{1}_k^T\mathbf{S}^+\mathbf{1}_k}$$

From Theorem 3.2.10 (ii) of Muirhead (1982) we get

$$\sqrt{n-1}\widetilde{\Sigma}_{22}^{1/2}\widehat{\mathbf{R}}_{\mathbf{L}}^{-1/2}(\widehat{\boldsymbol{\theta}} - \mathbf{r})\Big|\,(n-1)\widehat{\mathbf{R}}_{\mathbf{L}}^{-1} = \mathbf{C} \sim \mathcal{N}_p\left(\widetilde{\Sigma}_{22}^{1/2}\mathbf{C}^{1/2}(\mathbf{r} - \boldsymbol{\theta}), \mathbf{I}_p\right),$$

so that

$$(n-1)\widetilde{\Sigma}_{22}(\widehat{\boldsymbol{\theta}} - \mathbf{r})^T\widehat{\mathbf{R}}_{\mathbf{L}}^{-1}(\widehat{\boldsymbol{\theta}} - \mathbf{r})\Big|\,(n-1)\widehat{\mathbf{R}}_{\mathbf{L}}^{-1} = \mathbf{C} \sim \chi^2_{p,\lambda(\mathbf{C})},$$

where $\lambda(\mathbf{C}) = \widetilde{\Sigma}_{22}(\boldsymbol{\theta} - \mathbf{r})^T \mathbf{C}(\boldsymbol{\theta} - \mathbf{r})$. Moreover, from Corollary 1 of Bodnar et al. (2015) we know that

$$(n-1)\frac{\mathbf{1}_k^T \boldsymbol{\Sigma}^+ \mathbf{1}_k}{\mathbf{1}_k^T \mathbf{S}^+ \mathbf{1}_k} \sim \chi^2_{n-r}.$$

Also, from Theorem 3.2.10 of Muirhead (1982) we have that $\widetilde{S}_{22}$ is independent of $\widehat{\mathbf{R}}_{\mathbf{L}}^{-1}$ and $\widehat{\mathbf{R}}_{\mathbf{L}}^{-1}\widehat{\boldsymbol{\theta}}$. Thus, $\widetilde{S}_{22}$ is independent of $\widehat{\mathbf{R}}_{\mathbf{L}}^{-1}$ and $(\widehat{\boldsymbol{\theta}} - \mathbf{r})^T \widehat{\mathbf{R}}_{\mathbf{L}}^{-1}(\widehat{\boldsymbol{\theta}} - \mathbf{r})$. Putting all above results together we get

$$T|(n-1)\widehat{\mathbf{R}}_{\mathbf{L}}^{-1} = \mathbf{C} \sim F_{p,n-r,\lambda(\mathbf{C})}.$$

Using the fact that $(n-1)\widehat{\mathbf{R}}_{\mathbf{L}}^{-1} \sim \mathcal{W}_p(n-r+p, \mathbf{R}_{\mathbf{L}}^{-1})$ with $\mathbf{R}_{\mathbf{L}}^{-1} = \mathbf{L}\mathbf{R}\mathbf{L}^T$, we obtain the unconditional density function of T which is given by

$$f_T(x) = \int_{\mathbf{C}>0} f_{\mathcal{F}_{p,n-r,\lambda(\mathbf{C})}}(x) f_{\mathcal{W}_p(n-r+p,\mathbf{R}_{\mathbf{L}}^{-1})}(\mathbf{C})\mathbf{dC},$$

where $f_{\mathcal{F}_{i,j,\lambda}}$ denotes the density of the non-central $\mathcal{F}$-distribution with i and j degrees of freedom and noncentrality parameter λ, and $f_{\mathcal{W}_p(\cdot)}$ stands for the density of the Wishart distribution. For a case when $\lambda = 0$ we write $f_{\mathcal{F}_{i,j}}$.

Application of Theorem 1.3.6 of Muirhead (1982) leads us to

$$\begin{aligned} f_{\mathcal{F}_{p,n-r,\lambda(\mathbf{C})}}(x) &= f_{\mathcal{F}_{p,n-r}}(x) \exp\left(-\frac{\lambda(\mathbf{C})}{2}\right) \\ &\quad \times \sum_{i=0}^{\infty} \frac{((n-r+p)/2)^i}{(p/2)^i} \frac{(\lambda(\mathbf{C}))^i}{i!} \left(\frac{px}{2(n-r+px)}\right)^i. \end{aligned}$$

Let

$$\kappa(i) = \frac{1}{i!} \frac{[(n-r+p)/2]^i}{(p/2)^i} \left[\frac{px}{2(n-r+px)}\right]^i.$$

Then

$$\begin{aligned} f_T(x) &= f_{\mathcal{F}_{p,n-r}}(x) \sum_{i=0}^{\infty} \kappa(i) \int_{\mathbf{C}>0} (\lambda(\mathbf{C}))^i \exp\left(-\frac{\lambda(\mathbf{C})}{2}\right) \\ &\quad \times \frac{|\mathbf{R}_{\mathbf{L}}|^{(n-r+p)/2} |\mathbf{C}|^{(n-r-1)/2}}{2^{p(n-r+p)/2}\Gamma_p((n-r+p)/2)} \mathrm{etr}\left\{-\frac{1}{2}\mathbf{R}_{\mathbf{L}}\mathbf{C}\right\} \mathbf{dC} \\ &= f_{\mathcal{F}_{p,n-r}}(x) |\mathbf{R}_{\mathbf{L}}|^{(n-r+p)/2} |\mathbf{R}_{\mathbf{L}} + \mathbf{1}_k^T \boldsymbol{\Sigma}^+ \mathbf{1}_k (\mathbf{r} - \boldsymbol{\theta})(\mathbf{r} - \boldsymbol{\theta})^T|^{-(n-r+p)/2} \\ &\quad \times \sum_{i=0}^{\infty} \kappa(i) (\mathbf{1}_k^T \boldsymbol{\Sigma}^+ \mathbf{1}_k)^i E\left[\left((\mathbf{r} - \boldsymbol{\theta})^T \widetilde{\mathbf{C}} (\mathbf{r} - \boldsymbol{\theta})\right)^i\right], \end{aligned}$$

where $\widetilde{\mathbf{C}} \sim \mathcal{W}_p(n-r+p, \widetilde{\mathbf{R}}_{\mathbf{L}})$ with $\widetilde{\mathbf{R}}_{\mathbf{L}} = (\mathbf{R}_{\mathbf{L}} + \mathbf{1}_k^T \boldsymbol{\Sigma}^+ \mathbf{1}_k (\mathbf{r} - \boldsymbol{\theta})(\mathbf{r} - \boldsymbol{\theta})^T)^{-1}$. The symbol

$\Gamma_p(\cdot)$ denotes the multivariate gamma function (see Muirhead (1982)), while etr stands denotes the exponential of the trace of a matrix.

Using Theorem 3.2.8 of Muirhead (1982) we get

$$E\left[\left((\mathbf{r}-\boldsymbol{\theta})^T\widetilde{\mathbf{C}}(\mathbf{r}-\boldsymbol{\theta})\right)^i\right] = 2^i((n-r+p)/2)^i\left[\frac{(\mathbf{r}-\boldsymbol{\theta})^T\mathbf{R}_{\mathbf{L}}^{-1}(\mathbf{r}-\boldsymbol{\theta})}{1+\mathbf{1}_k^T\boldsymbol{\Sigma}^+\mathbf{1}_k(\mathbf{r}-\boldsymbol{\theta})^T\mathbf{R}_{\mathbf{L}}^{-1}(\mathbf{r}-\boldsymbol{\theta})}\right]^i.$$

Finally, putting all above together we get the statement of the theorem. □

Under the null hypothesis it holds that $T \sim \mathcal{F}_{p,n-r}$ so that the null hypothesis is rejected if $T > \mathcal{F}_{p,n-r;1-\alpha}$, where $\mathcal{F}_{p,n-r;1-\alpha}$ denotes for the $(1-\alpha)$ quantile of the central $\mathcal{F}$-distribution. Moreover, from Theorem 2 it is easy to get the elliptical confidence set for $\widehat{\boldsymbol{\theta}}$ which is expressed as

$$\left\{\mathbf{r} \in R^p : \frac{n-r}{p}(\mathbf{1}_k^T\mathbf{S}^+\mathbf{1}_k)(\widehat{\boldsymbol{\theta}}-\mathbf{r})^T\widehat{\mathbf{R}}_{\mathbf{L}}^{-1}(\widehat{\boldsymbol{\theta}}-\mathbf{r}) \leq \mathcal{F}_{p,n-r,1-\alpha}\right\}.$$

4 Empirical Illustration

The test derived in Theorem 2 can be an useful tool for a financial market analyst. To illustrate this, in this section, we presented an example of analysis in which we use the test to decide if a strategy adopted in the past should be modified based on new data that inform about current market conditions. More specifically, we consider the financial quality control setup in which we assume that based on information from the past T_0 periods we have built an investment strategy expressed as a portfolio $\mathbf{w}_0$. We assume that within each period we have approximately constant volatility so the the model can be viewed Gaussian within such a period. In general the periods do not need to be of the same length but rather should split the data over ranges during which volatility is approximately constant. For $\mathbf{r}$ we take the average of the portfolio estimates obtained for each of the past periods. The data from a new 'current' period are used to evaluate a new estimate $\mathbf{L}\mathbf{w}_{GMV}$ for the weights $\mathbf{L}\widehat{\mathbf{w}}_{GMV}$. The goal is to determine if $\mathbf{L}\mathbf{w}_{GMV}$ differs from $\mathbf{r}$ so a new strategy of investment should be evaluated and implemented.

We consider the log return monthly data of $k = 24$ stocks for leading oil, insurance, car, and IT companies that are listed in NYSE and NASDAQ stock exchanges for the period from September 2013 to July 2015, i.e. $n = 23$. Their abbreviated symbolic names are `COP, MRO, VLO, CVX, XOM, GSK, AZN, MRK, NVS, RHHBY, PFE, JNJ, PG, LLY, FCAU, GM, F, HMC, TM, AAPL, FB, GOOGL, YHOO, MSFT`. Similar data are used by Bodnar et al. (2015) for the same period.

First we discuss the distribution of the estimated global minimum variance portfolio weights. This analysis is based on Theorem 1. A quite important issue in applications is that we can not observe the singularity of the data in the strict sense due to the

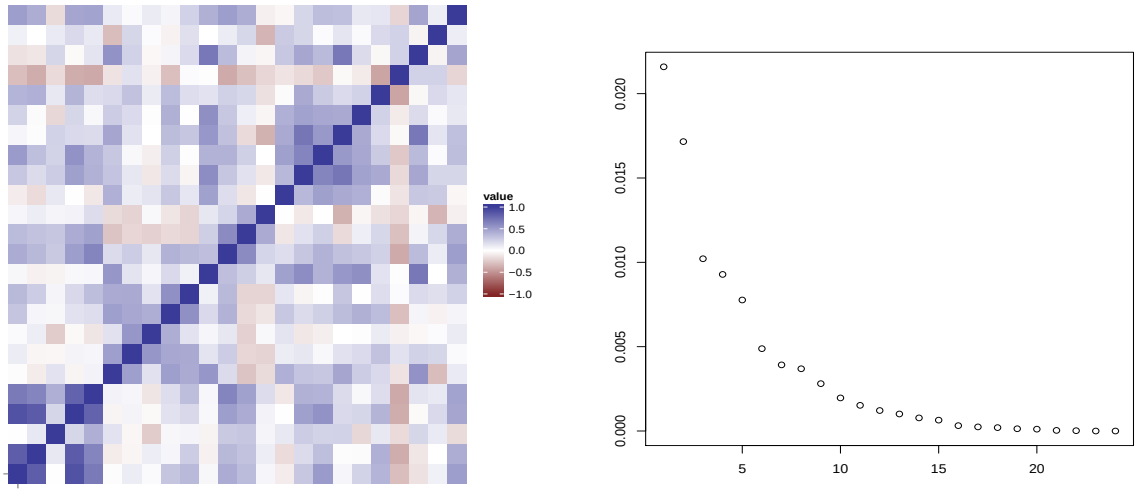

Figure 1: Heatmap of the estimated correlation matrix of 24 stocks (left) and its eigenvalues (right)

observational noise. For this reason, we should rather examine small eigenvalues (see Srivastava (2007)) and determine the rank of the covariance matrix as the number of its 'significantly' large eigenvalues. In Figure 1, we present the heatmap of the estimated covariance matrix of stocks as well as its eigenvalues. We do not attempt here to provide a formal test for singularity, but from the plot of eigenvalues one can say that the rank of the covariance matrix is $r = 5$ or $r = 9$.

Using the above setup, we get the estimated GMV portfolio weights

$$\widehat{\mathbf{w}}_{GMV} = [\text{-0.33 -0.13 0.05 0.5 0.06 0.13 -0.24 0.22 -0.14 0.3 0.24 0.18} \\ \text{-0.12 0.04 -0.07 -0.1 0.18 -0.08 0.1 0.17 0.18 -0.21 -0.02 0.08}]^T$$

with the estimated standard deviation based on (7) and with $r = 9$

$$\widehat{\mathbf{s}}_{GMV} = [0.16\ 0.08\ 0.08\ 0.28\ 0.17\ 0.23\ 0.14\ 0.07\ 0.20\ 0.25\ 0.23\ 0.31\ 0.36 \\ 0.14\ 0.10\ 0.08\ 0.15\ 0.24\ 0.11\ 0.09\ 0.07\ 0.11\ 0.12\ 0.10]^T.$$

We can see that some weights are negative that means a short sales for corresponding assets.

In Figure 2, we plot the densities of the estimation error centered at the estimated GMV portfolio weights that can be used to construct confidence intervals arround the true weights. Here we use the result which is obtained in Theorem 2. Thus, we plotted the estimated GMV portfolio weights of Conoco, Pfizer, General Motors, and Facebook for different values of $r = \{5, 9\}$. We observe larger variances when $r = 9$ and higher peaks $r = 5$.

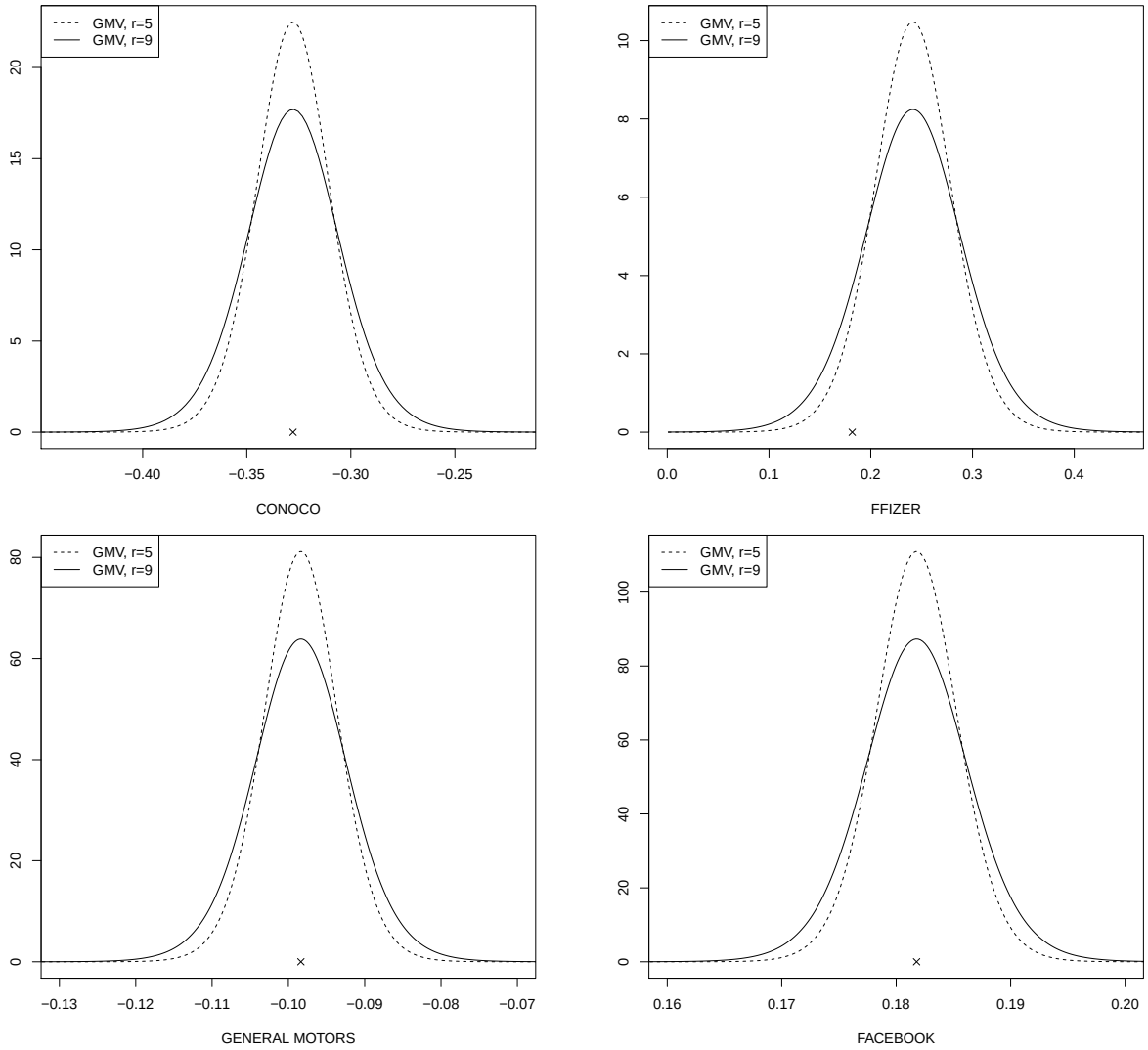

Figure 2: The densities of the GMV portfolio weights (two cases: $r = 5$ and $r = 9$) for four stocks: Conoco, Pfizer, General Motors, and Facebook. The estimated GMV portfolio weights are marked by '$\times$'.

In the second part of our analysis, we apply to the data the test in order to determine if a change of original investment allocation is justified. We consider the test statistics T for four stocks: Conoco, Pfizer, General Motors, and Facebook. A given value $\mathbf{r} = \mathbf{L}\widehat{\mathbf{w}}^*_{GMV}$ is a reference vector from a previous time period, namely, from December 2011 to August 2013. This period is divided into 5 sub-periods of monthly log returns. The average vector of the GMV portfolio weights $\widehat{\mathbf{w}}^*_{GMV}$ from the weights of the GMV portfolio of five

sub-periods is

$$\widehat{\mathbf{w}}^*_{GMV} = [0.01\ 0.06\ 0.03\ 0.04\ 0.03\ 0.03\ 0.06\ 0.06\ 0.05\ 0.05\ 0.02\ 0.02\ 0.02\ 0.05\ 0.00\ 0.1\ 0.08\ 0.03\ 0.02\ 0.07\ 0.09\ 0.02\ 0.00\ 0.07\ ]^T$$

In Table 4, we present the values of the test statitics T of four stocks using the setup above. Moreover, we have that $(1-\alpha)$ quantiles of the central $\mathcal{F}$- distribution for $r = 5$ and $r = 9$ are 4.413873 and 4.60011, respectively, with $\alpha = 5\%$. Using the fact that the null hypothesis is rejected if $T > \mathcal{F}_{p,n-r;1-\alpha}$, and results in Table 4, we can say that the null hypothesis for Conoco and General Motors should be rejected. As a consequence, an investor should consider a new allocation of these assets.

$r \setminus$ Company	Conoco	Pfizer	General Motors	Facebook
5	6.1836	1.2471	7.3289	2.0888
9	4.8094	0.9699	5.7002	1.6246

Table 1: Values of the test statistics T for four stocks: Conoco, Pfizer, General Motors, and Facebook. Two cases: $r = 5$ and $r = 9$.

5 Summary

In this paper, we extended the results obtained by Bodnar and Schmid (2008). We discuss the distribution of estimated GMV portfolio weights for the case when the portfolio size exceeding the sample size and the covariance matrix has the rank smaller than the sample size. A test for the general linear hypothesis is given as well as the distribution of the test statistics under the null and alternative hypothesis. These results are applied to the empirical data of certain popular stocks traded on NYSE and NASDAQ stock exchanges. Through this we illustrate the utility of such results for a financial investor.

References

Abramowitz, M. and Stegun, I. A. (1984). Pocketbook of mathematical functions. Verlag Harri Deutsch, Frankfurt (Main).

Berk, J. B. (1997). Necessary conditions for the CAPM. Journal of Economic Theory, 73:245–257.

Best, M. J. and Grauer, R. R. (1991). On the sensitivity of mean-variance-effecient portfolios to changes in asset means: some analytical and computational results. Review of Financial Studies, 4:315–342.

Bodnar, T., Mazur, S., and Podgórski, K. (2015). Singular inverse wishart distribution with application to portfolio theory. Working papers in statistics, Lund University.

Bodnar, T. and Okhrin, Y. (2011). On the product of inverse wishart and normal distributions with applications to discriminant analysis and portfolio theory. Scandinavian Journal of Statistics, 38:311–331.

Bodnar, T. and Schmid, W. (2007). The distribution of the sample variance of the global minimum variance portfolio in elliptical models. Statistics: A Journal of Theoretical and Applied Statistics, 44(1):65–75.

Bodnar, T. and Schmid, W. (2008). A test for the weights of the global minimum variance portfolio in an elliptical model. Metrika, 67(2):127–143.

Chopra, V. K. and Ziemba, W. T. (1993). The effect of errors in means, variances and covariances on optimal portfolio choice. Journal of Portfolio Management, 19(2):6–11.

Glombek, K. (2014). Statistical inference for high-dimensional global minimum variance portfolios. Scandinavian Journal of Statistics, 41:845–865.

Gupta, A. and Nagar, D. (2000). Matrix Variate Distributions. Chapman and Hall/CRC, Boca Raton.

Horn, R. A. and Johnsohn, C. R. (1985). Matrix Analysis. Cambridge University Press, Cambridge.

Markowitz, H. (1952). Portfolio selection. Journal of Finance, 7:77–91.

Merton, R. C. (1980). On estimating the expected return on the market: an exploratory investigation. Journal of Financial Economics, 8:323–361.

Muirhead, R. J. (1982). Aspects of Multivariate Statistical Theory. Wiley, New York.

Okhrin, Y. and Schmid, W. (2006). Distributional properties of portfolio weights. Journal of Econometrics, 134:235–256.

Owen, J. and Rabinovitch, R. (1983). On the class of elliptical distributions and their applications to the theory of portfolio choice. The Journal of Finance, 38:745–752.

Pappas, D., Kiriakopoulos, K., and Kaimakamis, G. (2010). Optimal portfolio selection with singular covariance matrix. International Mathematical Forum, 5(47):2305–2318.

Srivastava, M. S. (2007). Multivariate theory for analyzing high dimensional data. J. Japan Statist. Soc, pages 53–86.

Zhou, G. (1993). Asset-pricing tests under alternative distributions. The Journal of Finance, 48:1927–1942.

http://journals.lub.lu.se/stat

LUND UNIVERSITY
School of Economics and Management

Working Papers in Statistics 2015
LUND UNIVERSITY
SCHOOL OF ECONOMICS AND MANAGEMENT
Department of Statistics
Box 743
220 07 Lund, Sweden