Working Papers in Statistics
No 2015:2

Department of Statistics
School of Economics and Management
Lund University

Singular Inverse Wishart Distribution with Application to Portfolio Theory

TARAS BODNAR, STOCKHOLM UNIVERSITY

STEPAN MAZUR, LUND UNIVERSITY

KRZYSZTOF PODGÓRSKI, LUND UNIVERSITY

Singular Inverse Wishart Distribution with Application to Portfolio Theory

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 Brahe 1, SE-22007 Lund, Sweden*

Abstract

The inverse of the standard estimate of covariance matrix is frequently used in the portfolio theory to estimate the optimal portfolio weights. For this problem, the distribution of the linear transformation of the inverse is needed. We obtain this distribution in the case when the sample size is smaller than the dimension, the underlying covariance matrix is singular, and the vectors of returns are independent and normally distributed. For the result, the distribution of the inverse of covariance estimate is needed and it is derived and referred to as the singular inverse Wishart distribution. We use these results to provide an explicit stochastic representation of an estimate of the mean-variance portfolio weights as well as to derive its characteristic function and the moments of higher order.

ASM Classification: 62H10, 62H12, 91G10

Keywords: singular Wishart distribution, mean-variance portfolio, sample estimate of precision matrix, Moore-Penrose inverse

[1]Corresponding author. E-mail address: Krzysztof.Podgorski@stat.lu.se. 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

Analyzing multivariate data having fewer observations than their dimension is an important problem in the multivariate data analysis. For example, in the mathematical finance, due to the dependence in historical data the sample size of portfolio assets should be often considered effectively smaller than the portfolio size. In contrast to the covariance estimation problem, for which the singularities due to both the small sample size and the linear dependence between variables have been considered, see Díaz-García et al. (1997), in the portfolio theory, where the linear transformations of the inverse of covariance estimates need to be considered, the singularity problems have not been tackled. In particular, the problem of finding the distribution of the mean-variance (MV) portfolio weights was only discussed for the non-singular covariance of the vector of returns and when the sample size of assets is larger than the portfolio size, see Bodnar and Schmid (2011). Our goal is to fill this gap and to provide results for this portfolio theory problem, when the small sample size and the singular covariance matrix are both present. One important reason for considering the singular covariance matrix case in the portfolio theory is that often in for a given set of assets, there maybe strong stochastic dependence between them. This is due to some natural interrelation between asset prices. For example, valuation of assets within a specific industry branch often are highly correlated. If the dimension of portfolio is relatively large there is a possibility of (approximate) singularity and the problem needs to be addressed in the theory.

The paper has two major contributions. The first one lies in deriving the distributional properties of the generalized inverse Wishart (GIW) random matrix under singularity of the covariance matrix. This singular covariance case is referred to as the singular inverse Wishart distribution (SIW). In particular, we show that under the linear transformations the family of the SIW distributions remains within the GIW distributions. The notable special case is when the rank of the linear transformation is smaller than the rank of the covariance matrix. Under this assumption the distribution becomes a regular inverse Wishart distribution. This is used in our second main contribution that gives a stochastic representation of a linear transformation of the estimated MV portfolio weights under the singularity conditions as well as their characteristic function and the moments of higher order. These results are complementary to the ones obtained in Okhrin and Schmid (2006), Bodnar and Schmid (2011).

The paper is structured as follows. First, in Section 2, we introduce basic notation and review known facts about (inverse) Wishart distributions and their generalizations. In Section 3, we consider the distributional properties for the linear symmetric transformations of the SIW random matrix. In Theorem 1, we prove that for a SIW matrix $\mathbf{A}$, its linear symmetric transformation $\mathbf{LAL}^T$, for a detereministic matrix $\mathbf{L}$, remains generalized Wishart distributed. Theorem 1 is then used to obtain Theorem 2 and Corollary 1 that show independence on random linear transformation. The results can be utilized for developing test statistics in the multivariate singular problems, see Srivastava (2007) and Muirhead (1982). In Section 4 we consider estimation of the optimal portfolio weights under the singularity. In Theorem 4, we show the independence between the sample mean vector and the sample covariance matrix and derive their distributions when the sample size is smaller than the dimension of portfolio. In Theorem 7, we present a stochastic representation of the distribution of a linear transformation for the estimated MV portfolio weights. Finally, in Corollary 3 and Corollary 4, the expressions of the characteristic function and the moments of higher order are provided.

2 Notation and basic facts

The Wishart matrix distribution is a multivariate generalization of the chi-square distribution and has been applied in numerous fields of applied and theoretical statistics. The distributional properties of the Wishart matrices, the inverse Wishart matrices and related quantities were established by Olkin and Roy (1954), Khatri (1959), Díaz-García et al. (1997), Bodnar and Okhrin (2008), Drton et al. (2008), Bodnar et al. (2013) among others. In this section, we collect some basic facts about the Wishart and inverse Wishart distributions as well as about some of their generalizations.

Let $\mathbf{X} \sim \mathcal{N}_{k,n}(\mathbf{0}, \mathbf{\Sigma} \otimes \mathbf{I}_n)$, i.e. the columns of the random $k \times n$ matrix $\mathbf{X}$ represent an iid sample of size n from the k-dimensional normal distribution with zero mean vector and non-singular covariance matrix $\mathbf{\Sigma}$. If the sample size n is greater than the dimension k, then $\mathbf{A} = \mathbf{X}\mathbf{X}^T$ has the k-dimensional Wishart distribution with n degrees of freedom and the matrix parameter $\mathbf{\Sigma}$.

In Srivastava (2003), a generalization of the Wishart distribution was studied by considering the quadratic form $\mathbf{A} = \mathbf{X}\mathbf{X}^T$ in the case of the sample size being smaller than the dimension, i.e. for $k > n$. In this case, the distribution is called the singular Wishart in Srivastava (2003) and the k-dimensional pseudo-Wishart distribution in Díaz-García et al. (1997). The distribution is residing on the singular $n \times n$ dimensional subspace of non-negatively definite matrices $\mathbf{A}$ that for the following partitioned forms

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} \tag{1}$$

have the $n \times n$ matrix $\mathbf{A}_{11}$ non-singular and $\mathbf{A}_{22} = \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}$. In an abbreviated form we simply write $\mathbf{A} \sim \mathcal{W}_k(n, \mathbf{\Sigma})$ both when $n \geq k$ and $n < k$. The characteristic function of the singular Wishart distribution is presented in Bodnar et al. (2014).

The additional source of 'singularity' can be due to a singular matrix parameter $\mathbf{\Sigma}$. Here terminology is not uniquely established but in Díaz-García et al. (1997) they refer to this case as a singular matrix Wishart distribution and distinguishing the case of the rank of $\mathbf{\Sigma}$ bigger than n by adding the prefix *pseudo-*. We continue to use notation $\mathcal{W}_k(n, \mathbf{\Sigma})$ to cover this case.

In the case of the sample size n greater or equal to the dimension k and a non-singular covariance $\mathbf{\Sigma}$, the inverse Wishart distribution is defined as the distribution of the inverse of $\mathbf{X}\mathbf{X}^T$. The number of degrees of freedom is set to $n+k+1$ and the parameter is taken as the precision matrix $\mathbf{\Psi} = \mathbf{\Sigma}^{-1}$. We abbreviate this to $\mathcal{IW}_k(n+k+1, \mathbf{\Psi})$.

The inverse Wishart distribution with a nonsingular $\mathbf{\Psi}$ can be extended to the singular case of $n < k$. For this we need some basic facts about the generalized inverse matrices. The generalized (Moore-Penrose) inverse $\mathbf{A}^+$ of a $k \times k$ non-negatively defined matrix $\mathbf{A}$ of the rank $n \le k$ can be explicitly defined through its spectral representation $\mathbf{A} = \mathbf{P}\mathbf{\Lambda}\mathbf{P}^T$, where $\mathbf{\Lambda}$ is the $n \times n$ diagonal matrix of positive eigenvalues and $\mathbf{P}$ is the $k \times n$ matrix having the corresponding eigenvectors as columns. With this notation we have that $\mathbf{A}^+ = \mathbf{P}\mathbf{\Lambda}^{-1}\mathbf{P}^T$. This can be also written as

$$\mathbf{A}^+ = \begin{bmatrix} \mathbf{P}_1\mathbf{\Lambda}^{-1}\mathbf{P}_1^T & \mathbf{P}_1\mathbf{\Lambda}^{-1}\mathbf{P}_2^T \\ \mathbf{P}_2\mathbf{\Lambda}^{-1}\mathbf{P}_1^T & \mathbf{P}_2\mathbf{\Lambda}^{-1}\mathbf{P}_2^T \end{bmatrix} = \begin{bmatrix} \mathbf{A}_{11}^+ & \mathbf{A}_{12}^+ \\ \mathbf{A}_{21}^+ & \mathbf{A}_{22}^+ \end{bmatrix},$$

where the second equality serves as the definition of $\mathbf{A}_{ij}^+$, $i, j = 1, 2$, while the $n \times n$ non-singular matrix $\mathbf{P}_1$ is made of the first n rows of $\mathbf{P}$, while the $k - n \times n$ matrix $\mathbf{P}_2$ is made of the remaining $k - n$ rows of $\mathbf{P}$.

For $\mathbf{A} \sim \mathcal{W}_k(n, \mathbf{\Sigma})$ and $n < k$, the generalized inverse Wishart distribution $\mathcal{IW}_k(n + k + 1, \mathbf{\Psi})$ is extended as the distribution of $\mathbf{B} = \mathbf{A}^+$. Note that the distribution is residing on the same subspace of non-negative matrices as for the Wishart distribution, i.e. matrices $\mathbf{B}$ such that the $n \times n$ upper-left 'corner' $\mathbf{B}_{11}$ is non-singular and $\mathbf{B}_{22} = \mathbf{B}_{21}{\mathbf{B}_{11}}^{-1}\mathbf{B}_{12}$. For more properties see Bodnar and Okhrin (2008).

In this work we consider the singular inverse Wishart distribution that is defined as the distribution of the Moore-Penrose inverse of a Wishart distributed matrix $\mathbf{A}$ with $n < k$ and a singular matrix $\mathbf{\Sigma}$.

3 Linear transformations of singular inverse Wishart distribution

In Theorem 1 we derive the distribution of linear form of a singular inverse Wishart distributed random matrix. The results are obtained when the covariance matrix $\mathbf{\Sigma}$ is assumed to be singular. Since terminology for the singular cases of inverse Wishart matrices is not that well-established, to avoid confusion we express the result using the generalized Moore-Penrose inverses and utilizing our standard notation for matrix Wishart distributions.

Theorem 1. *Let* $\mathbf{W} \sim \mathcal{W}_k(n, \mathbf{\Sigma})$*,* $k > n$ *with* $rank(\mathbf{\Sigma}) = r \leq n$ *and let* $\mathbf{L} : p \times k$ *be a matrix of constants of rank* p*. If* $m = rank(\mathbf{L\Sigma}) = min(r, p)$*, then*

$$(\mathbf{L}\mathbf{W}^{+}\mathbf{L}^T)^{+} \sim \mathcal{W}_p\left(n - r + m, (\mathbf{L}\mathbf{\Sigma}^{+}\mathbf{L}^T)^{+}\right).$$

Moreover, if $m = p$*, then both* $\mathbf{L}\mathbf{W}^{+}\mathbf{L}^T$ *and* $\mathbf{L}\mathbf{\Sigma}^{+}\mathbf{L}^T$ *are of the full rank* p *and thus their Moore-Penrose inverses becomes the regular inverses.*

Proof. From Srivastava (2003) we get the stochastic representation of $\mathbf{W}$ expressed as

$$\mathbf{W} \stackrel{d}{=} \mathbf{X}\mathbf{X}^T \quad \text{with} \quad \mathbf{X} \sim \mathcal{N}_{k,n}(\mathbf{0}, \mathbf{\Sigma} \otimes \mathbf{I}_n)\,, \tag{2}$$

where the symbol $\stackrel{d}{=}$ denotes the equality in distribution.

Let $\mathbf{\Sigma} = \mathbf{Q\Lambda Q}^T$ be the singular value decomposition of $\mathbf{\Sigma}$ where $\mathbf{\Lambda} : r \times r$ is the matrix of non-zero eigenvalues and $\mathbf{Q} : k \times r$ is the orthogonal matrix of the corresponding eigenvectors. Then the stochastic representation of $\mathbf{X}$ is given by

$$\mathbf{X} \stackrel{d}{=} \mathbf{Q}\mathbf{\Lambda}^{1/2}\mathbf{Z} \quad \text{with} \quad \mathbf{Z} \sim \mathcal{N}_{r,n}(\mathbf{0}, \mathbf{I}_r \otimes \mathbf{I}_n)\,. \tag{3}$$

From (2) and (3), we obtain

$$\mathbf{W} \stackrel{d}{=} \mathbf{Q}\mathbf{\Lambda}^{1/2}\mathbf{Z}\mathbf{Z}^T\mathbf{\Lambda}^{1/2}\mathbf{Q}^T\,, \tag{4}$$

where $\mathbf{Z}\mathbf{Z}^T \sim \mathcal{W}_r(n, \mathbf{I}_r)$.

Since $\mathbf{Q}\mathbf{\Lambda}^{1/2}$ is the full column-rank matrix and $\mathbf{\Lambda}^{1/2}\mathbf{Q}^T$ is the full row-rank matrix, we get

$$\begin{aligned}
\mathbf{L}\mathbf{W}^{+}\mathbf{L}^T &\stackrel{d}{=} \mathbf{L}\left(\mathbf{Q}\mathbf{\Lambda}^{1/2}\mathbf{Z}\mathbf{Z}^T\mathbf{\Lambda}^{1/2}\mathbf{Q}^T\right)^{+}\mathbf{L}^T \\
&= \mathbf{L}\mathbf{Q}\mathbf{\Lambda}^{-1/2}\left(\mathbf{Z}\mathbf{Z}^T\right)^{+}\mathbf{\Lambda}^{-1/2}\mathbf{Q}^T\mathbf{L}^T \\
&= \mathbf{L}\mathbf{Q}\mathbf{\Lambda}^{-1/2}\left(\mathbf{Z}\mathbf{Z}^T\right)^{-1}\mathbf{\Lambda}^{-1/2}\mathbf{Q}^T\mathbf{L}^T\,,
\end{aligned} \tag{5}$$

because $\mathbf{Z}\mathbf{Z}^T$ is non-singular (cf., Greville (1966)). Finally, the identity $\mathbf{Z}\mathbf{Z}^T \sim \mathcal{W}_r(n, \mathbf{I}_r)$ and the assumption that $m = p$ after the application of Theorem 3.2.11 in Muirhead (1982) lead to

$$(\mathbf{L}\mathbf{W}^{+}\mathbf{L}^T)^{-1} \sim \mathcal{W}_p\left(n - r + p, \left(\mathbf{L}\mathbf{Q}\mathbf{\Lambda}^{-1/2}\mathbf{I}_r\mathbf{\Lambda}^{-1/2}\mathbf{Q}^T\mathbf{L}^T\right)^{-1}\right) = \mathcal{W}_p\left(n - r + p, (\mathbf{L}\mathbf{\Sigma}^{+}\mathbf{L}')^{-1}\right).$$

This proves the case $m = p$.

For the proof in the case when $m = r$, note that here it is assumed that $rank(\mathbf{L}) = p > r$ and $rank(\mathbf{L\Sigma}) = r$. Then $\mathbf{L}\mathbf{Q}\mathbf{\Lambda}^{-1/2}$ has a full column-rank and $\mathbf{\Lambda}^{-1/2}\mathbf{Q}^T\mathbf{L}^T$ has a full row-rank. Applying the last property to (5) and using Theorem 2.4.2 of Gupta and

Nagar (2000) we get

$$\begin{aligned}(\mathbf{L}\mathbf{W}^{+}\mathbf{L}^{T})^{+} &\stackrel{d}{=} (\mathbf{L}\mathbf{Q}\mathbf{\Lambda}^{-1/2}\left(\mathbf{Z}\mathbf{Z}^{T}\right)^{-1}\mathbf{\Lambda}^{-1/2}\mathbf{Q}^{T}\mathbf{L}^{T})^{+}\\ &= (\mathbf{\Lambda}^{-1/2}\mathbf{Q}^{T}\mathbf{L}^{T})^{+}\mathbf{Z}\mathbf{Z}^{T}(\mathbf{L}\mathbf{Q}\mathbf{\Lambda}^{-1/2})^{+}\\ &= \widetilde{\mathbf{Z}}\widetilde{\mathbf{Z}}^{T} \qquad (6)\end{aligned}$$

with $\widetilde{\mathbf{Z}} \sim \mathcal{N}_{p,n}(\mathbf{0}, (\mathbf{L}\mathbf{\Sigma}^{+}\mathbf{L}^{T})^{+} \otimes \mathbf{I}_n)$.

Thus, if $p < n$ then $\widetilde{\mathbf{Z}}\widetilde{\mathbf{Z}}^{T}$ has the Wishart distribution with singular covariance matrix, otherwise, i.e. $p > n$, it has the pseudo-Wishart distribution (see Theorem 5.2 of Srivastava (2003)). □

An application of Theorem 1 leads to Theorem 2 and ensuing Corollary 1.

Theorem 2. *Let* $\mathbf{W} \sim \mathcal{W}_k(n, \mathbf{\Sigma})$, $k > n$ *with* $rank(\mathbf{\Sigma}) = r \leq n$ *and let* $\mathbf{Y} : p \times k$ *be a random matrix such that with probability one* $rank(\mathbf{Y}\mathbf{\Sigma}) = p$, $p \leq r$, *and which is independent of* $\mathbf{W}$. *Then*

$$(\mathbf{Y}\mathbf{\Sigma}^{+}\mathbf{Y}^{T})^{-1/2}(\mathbf{Y}\mathbf{W}^{+}\mathbf{Y}^{T})^{-1}(\mathbf{Y}\mathbf{\Sigma}^{+}\mathbf{Y}^{T})^{-1/2} \sim \mathcal{W}_p\left(n - r + p, \mathbf{I}_p\right) \qquad (7)$$

and it is independent of $\mathbf{Y}$.

Proof. Since $\mathbf{Y}$ and $\mathbf{W}$ are independent, we get that the conditional distribution of $(\mathbf{Y}\mathbf{W}^{+}\mathbf{Y}^{T})^{-1}$ given $\mathbf{Y} = \mathbf{Y}_0$ is equal to the distribution of $(\mathbf{Y}_0\mathbf{W}^{+}\mathbf{Y}_0^{T})^{-1}$. The application of Theorem 1 leads to

$$(\mathbf{Y}_0\mathbf{\Sigma}^{+}\mathbf{Y}_0^{T})^{1/2}(\mathbf{Y}_0\mathbf{W}^{+}\mathbf{Y}_0^{T})^{-1}(\mathbf{Y}_0\mathbf{\Sigma}^{+}\mathbf{Y}_0^{T})^{1/2} \sim \mathcal{W}_p\left(n - r + p, \mathbf{I}_p\right),$$

which does not depend on $\mathbf{Y}_0$. Hence, it is also the unconditional distribution of $(\mathbf{Y}\mathbf{\Sigma}^{+}\mathbf{Y}^{T})^{1/2}(\mathbf{Y}\mathbf{W}^{+}\mathbf{Y}^{T})^{-1}(\mathbf{Y}\mathbf{\Sigma}^{+}\mathbf{Y}^{T})^{1/2}$ which appears to be independent of $\mathbf{Y}$. □

One consequence of Theorem 2 is the following corollary, where the important case of $p = 1$ is considered.

Corollary 1. *If* $\mathbf{W} \sim W_k(n, \mathbf{\Sigma})$, $k > n$ *with* $rank(\mathbf{\Sigma}) = r \leq n$ *and* $\mathbf{y}$ *is any* k*-dimensional random vector distributed independently of* $\mathbf{W}$ *such that* $\mathbf{y}^T\mathbf{\Sigma}$ *is non-zero with probability one, then*

$$\frac{\mathbf{y}^T\mathbf{\Sigma}^{+}\mathbf{y}}{\mathbf{y}^T\mathbf{W}^{+}\mathbf{y}} \sim \chi^2_{n-r+1},$$

and is independent of $\mathbf{y}$.

It should be noted that in Theorem 1, the assumption that $rank(\mathbf{\Sigma})$ is smaller than the sample size n is essential. The problem of finding the distribution of the linear transformation of the generalized inverse Wishart distribution in the general case seems to be difficult and remains open. Some special case in a general case can be obtained as shown in the next result where we consider the orthogonal transformation of the generalized inverse Wishart random matrix.

Theorem 3. *Let* $\mathbf{W} \sim \mathcal{W}_k(n, \mathbf{\Sigma})$*,* $k > n$ *and let* $\mathbf{L} : k \times k$ *be an orthogonal matrix. Then*

$$(\mathbf{L}\mathbf{W}^{+}\mathbf{L}^{T})^{+} \sim \mathcal{W}_k(n, \mathbf{L}\mathbf{\Sigma}\mathbf{L}^{T}).$$

Proof. It follows from general properties of the Moore-Penrose inverse matrices (see Boullion and Odell (1971)) that for an orthogonal matrix $\mathbf{L}$:

$$(\mathbf{L}\mathbf{W}^{+}\mathbf{L}^{T})^{+} = \mathbf{L}\mathbf{W}\mathbf{L}^{T}$$

and the result follows since $\mathbf{L}\mathbf{W}\mathbf{L}^{T} = \mathbf{L}\mathbf{X}(\mathbf{L}\mathbf{X})^{T}$ and $\mathbf{L}\mathbf{X} \sim \mathcal{N}_{k,n}(\mathbf{0}, \mathbf{L}\mathbf{\Sigma}\mathbf{L}^{T} \otimes \mathbf{I}_n)$. □

4 Application to portfolio theory

In this section, using the properties of the singular inverse Wishart distribution shown in Section 3, we derive the stochastic representation of the linear transformation of the mean-variance portfolio weights under the assumption of normality for the case when the number of observations n from k-variate Gaussian distribution is smaller than the dimension k and a singular covariance matrix $\mathbf{\Sigma}$.

We consider the vector of portfolio weights $\mathbf{w} = (w_1, ..., w_k)$ of k assets, i.e. $\mathbf{w}^T\mathbf{1}_k = 1$. We assume that the asset log-returns are normally and identically distributed with mean vector $\boldsymbol{\mu}$ and covariance matrix $\mathbf{\Sigma}$. Let $\mathbf{\Sigma}$ be a nonnegative definite matrix with with $rank(\mathbf{\Sigma}) = r \leq n$.

The MV portfolio, $\mathbf{w}_{MV}$ is the solution of the following optimization problem

$$\max_{\mathbf{w}:\mathbf{w}^T\mathbf{1}_k=1} \mathbf{w}^T\boldsymbol{\mu} - \frac{\alpha}{2}\mathbf{w}^T\mathbf{\Sigma}\mathbf{w}, \tag{8}$$

where $\mathbf{1}_k$ be the k-dimensional vector of ones. The symbol $\alpha > 0$ describes the risk aversion of an investor.

Since $\mathbf{\Sigma}$ is singular, the optimization problem (8) has an infinite number of solutions. In Pappas et al. (2010), a solution was expressed as

$$\mathbf{w}_{MV} = \frac{\mathbf{\Sigma}^{+}\mathbf{1}_k}{\mathbf{1}_k^T\mathbf{\Sigma}^{+}\mathbf{1}_k} + \alpha^{-1}\mathbf{R}\boldsymbol{\mu} \quad \text{with} \quad \mathbf{R} = \mathbf{\Sigma}^{+} - \mathbf{\Sigma}^{+}\mathbf{1}_k\mathbf{1}_k^T\mathbf{\Sigma}^{+}/\mathbf{1}_k^T\mathbf{\Sigma}^{+}\mathbf{1}_k, \tag{9}$$

which appears to be unique solution with the minimal Euclidean norm. Relation (9) can be used only under the constrain $\mathbf{1}_k^T\mathbf{\Sigma}^{+}\mathbf{1}_k \neq 0$, which is assumed throughout the paper.

Finally, we point out that if we have the fully risk-averse investor, i.e. $\alpha \to \infty$, then the global minimum variance portfolio is the limit case of the MV portfolio.

In practice $\mathbf{\Sigma}$ is an unknown matrix and should be estimated using historical values of asset returns. Given a sample of n independent observations $\mathbf{x}_1, ..., \mathbf{x}_n$ of log-returns on k assets we calculate the sample covariance matrix by

$$\mathbf{S} = \frac{1}{n-1} \sum_{i=1}^{n} (\mathbf{x}_i - \overline{\mathbf{x}})(\mathbf{x}_i - \overline{\mathbf{x}})^T, \tag{10}$$

where $\overline{\mathbf{x}} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{x}_i$. Replacing $\boldsymbol{\mu}$ and $\mathbf{\Sigma}$ with $\overline{\mathbf{x}}$ and $\mathbf{S}$, respectively, in (9) we obtain the sample estimator of the MV portfolio weights given by

$$\widehat{\mathbf{w}}_{MV} = \frac{\mathbf{S}^+ \mathbf{1}_k}{\mathbf{1}_k^T \mathbf{S}^+ \mathbf{1}_k} + \alpha^{-1} \widehat{\mathbf{R}} \overline{\mathbf{x}} \quad \text{with} \quad \widehat{\mathbf{R}} = \mathbf{S}^+ - \frac{\mathbf{S}^+ \mathbf{1}_k \mathbf{1}_k^T \mathbf{S}^+}{\mathbf{1}_k^T \mathbf{S}^+ \mathbf{1}_k}. \tag{11}$$

The distribution of $\widehat{\mathbf{w}}_{MV}$ is of obvious interest for the portfolio theory and was discussed for the non-singular case, i.e. $k \leq n-1$, by Okhrin and Schmid (2006), Bodnar and Schmid (2011). The following result completes the MV portfolio theory by providing the distribution in the singular case. We consider a more general case, namely the distribution of a linear transformation of $\widehat{\mathbf{w}}_{MV}$ is derived. Let

$$\boldsymbol{\theta}_{MV} = \mathbf{L}\mathbf{w}_{MV} = \frac{\mathbf{L}\mathbf{\Sigma}^+ \mathbf{1}_k}{\mathbf{1}_k^T \mathbf{\Sigma}^+ \mathbf{1}_k} + \alpha^{-1} \mathbf{L}\mathbf{R}\boldsymbol{\mu}. \tag{12}$$

where $\mathbf{L}$ is a non-random $p \times k$ matrix of rank $p < r$ such that $\text{rank}(\mathbf{L}\mathbf{\Sigma}) = p$. Applying the estimator (10) we obtain

$$\widehat{\boldsymbol{\theta}}_{MV} = \mathbf{L}\widehat{\mathbf{w}}_{MV} = \frac{\mathbf{L}\mathbf{S}^+ \mathbf{1}_k}{\mathbf{1}_k^T \mathbf{S}^+ \mathbf{1}_k} + \alpha^{-1} \mathbf{L}\widehat{\mathbf{R}}\overline{\mathbf{x}}. \tag{13}$$

The following theorem shows that the sample mean vector $\overline{\mathbf{x}}$ and the sample covariance matrix $\mathbf{S}$ are independently distributed.

Theorem 4. *Let* $\mathbf{X} \sim \mathcal{N}_{k,n}(\boldsymbol{\mu}\mathbf{1}_n^T, \mathbf{\Sigma} \otimes \mathbf{I}_n)$, $k > n$ *with* $rank(\mathbf{\Sigma}) = r \leq n$. *Then*

(a) $(n-1)\mathbf{S} \sim \mathcal{W}_k(n-1, \mathbf{\Sigma})$,

(b) $\overline{\mathbf{x}} \sim \mathcal{N}_k\left(\boldsymbol{\mu}, \frac{1}{n}\mathbf{\Sigma}\right)$,

(c) $\overline{\mathbf{x}}$ *and* $\mathbf{S}$ *are independently distributed.*

Proof. From Theorem 2.1 of Díaz-García et al. (1997) the density function of $\mathbf{X}$ is given by

$$f(\mathbf{X}) = \frac{1}{(2\pi)^{rn/2}\left(\prod_{i=1}^{r}\lambda_i\right)^{n/2}}\mathrm{etr}\left(-\frac{1}{2}(\mathbf{X}-\boldsymbol{\mu}\mathbf{1}_n^T)^T\boldsymbol{\Sigma}^+(\mathbf{X}-\boldsymbol{\mu}\mathbf{1}_n^T)\right), \tag{14}$$

where λ_i are the non-zero eigenvalues of $\boldsymbol{\Sigma}$.

Let $\mathbf{V} = \mathbf{X}\mathbf{F}^T$ with the Jacobian of transformation equals to 1, where $\mathbf{F}$ is an orthogonal $n \times n$ matrix with elements in the last row which are equal to $n^{-1/2}$. The matrix $\mathbf{V}$ is partitioned as $\mathbf{V} = (\mathbf{Z}, \mathbf{v})$ where $\mathbf{Z}$ is $k \times (n-1)$ and $\mathbf{v}$ is $k \times 1$. Then it holds that

$$\mathbf{X}\mathbf{X}^T = \mathbf{V}\mathbf{V}^T = \mathbf{Z}\mathbf{Z}^T + \mathbf{v}\mathbf{v}^T. \tag{15}$$

Because the first $(n-1)$ rows of $\mathbf{F}$ are orthogonal to $\mathbf{1}_n$, i.e. $\mathbf{F}\mathbf{1}_n = (0, ..., 0, n^{1/2})^T$, we have that

$$\mathbf{X}\mathbf{1}_n\boldsymbol{\mu}^T = \mathbf{V}\mathbf{F}\mathbf{1}_n\boldsymbol{\mu}^T = n^{1/2}\mathbf{v}\boldsymbol{\mu}^T. \tag{16}$$

Using (15) and (16) the term $(\mathbf{X}-\boldsymbol{\mu}\mathbf{1}_n^T)(\mathbf{X}-\boldsymbol{\mu}\mathbf{1}_n^T)^T$ which is presented in (14) can be rewritten as

$$\begin{aligned}(\mathbf{X}-\boldsymbol{\mu}\mathbf{1}_n^T)(\mathbf{X}-\boldsymbol{\mu}\mathbf{1}_n^T)^T &= \mathbf{Z}\mathbf{Z}^T + \mathbf{v}\mathbf{v}^T - n^{1/2}\boldsymbol{\mu}\mathbf{v}^T - n^{1/2}\mathbf{v}\boldsymbol{\mu}^T + n\boldsymbol{\mu}\boldsymbol{\mu}^T \\ &= \mathbf{Z}\mathbf{Z}^T + (\mathbf{v}-n^{1/2}\boldsymbol{\mu})(\mathbf{v}-n^{1/2}\boldsymbol{\mu})^T. \end{aligned} \tag{17}$$

Hence, we obtain the joint density function of $\mathbf{Z}$ and $\mathbf{v}$:

$$\begin{aligned} f(\mathbf{Z},\mathbf{v}) &= \frac{1}{(2\pi)^{r(n-1)/2}\left(\prod_{i=1}^{r}\lambda_i\right)^{(n-1)/2}}\mathrm{etr}\left(-\frac{1}{2}\boldsymbol{\Sigma}^+\mathbf{Z}\mathbf{Z}^T\right) \\ &\times \frac{1}{(2\pi)^{r/2}\left(\prod_{i=1}^{r}\lambda_i\right)^{1/2}}\exp\left(-\frac{1}{2}(\mathbf{v}-n^{1/2}\boldsymbol{\mu})^T\boldsymbol{\Sigma}^+(\mathbf{v}-n^{1/2}\boldsymbol{\mu})\right), \end{aligned}$$

where $\mathbf{Z} \sim \mathcal{N}_{k,n}(\mathbf{0}, \boldsymbol{\Sigma}\otimes\mathbf{I}_n)$ and $\mathbf{v} \sim \mathcal{N}_k(n^{1/2}\boldsymbol{\mu}, \boldsymbol{\Sigma})$ which are independently distributed. It leads to the fact that $\overline{\mathbf{x}} \sim \mathcal{N}_k(\boldsymbol{\mu}, 1/n\boldsymbol{\Sigma})$ and is independent of $\mathbf{Z}$ since $\mathbf{v} = n^{-1/2}\mathbf{X}^T\mathbf{1}_n = n^{1/2}\overline{\mathbf{x}}$. Also, after the transformation $\mathbf{S} = \mathbf{Z}\mathbf{Z}^T$ and the application of Theorem 5.2 of Srivastava (2003) we obtain that $\mathbf{S} \sim \mathcal{W}_k(n-1, \boldsymbol{\Sigma})$. □

From Theorem 4 we have that $\mathbf{S}$ and $\overline{\mathbf{x}}$ are independent, then the conditional distribution of $\widehat{\boldsymbol{\theta}}_{EU}$ given $\overline{\mathbf{x}} = \overline{\mathbf{x}}^*$ is expressed as

$$\begin{aligned} \widehat{\boldsymbol{\theta}}_{MV}(\overline{\mathbf{x}}^*) &= \frac{\mathbf{L}\mathbf{S}^+\mathbf{1}_k}{\mathbf{1}_k^T\mathbf{S}^+\mathbf{1}_k} + \alpha^{-1}\mathbf{L}\widehat{\mathbf{R}}\overline{\mathbf{x}}^* \\ &= \frac{\mathbf{L}\mathbf{S}^+\mathbf{1}_k}{\mathbf{1}_k^T\mathbf{S}^+\mathbf{1}_k} + \alpha^{-1}(n-1)\frac{\overline{\mathbf{x}}^{*T}\widehat{\mathbf{R}}\overline{\mathbf{x}}^*}{(n-1)\overline{\mathbf{x}}^{*T}\mathbf{R}\overline{\mathbf{x}}^*}\overline{\mathbf{x}}^{*T}\mathbf{R}\overline{\mathbf{x}}^*\frac{\mathbf{L}\widehat{\mathbf{R}}\overline{\mathbf{x}}^*}{\overline{\mathbf{x}}^{*T}\widehat{\mathbf{R}}\overline{\mathbf{x}}^*} \\ &= \widehat{\boldsymbol{\theta}}_{MV;1} + \widetilde{\alpha}^{-1}\hat{s}^{*-1}\widehat{\boldsymbol{\theta}}_{MV;2}(\overline{\mathbf{x}}^*), \end{aligned} \tag{18}$$

where $\widehat{\boldsymbol{\theta}}_{MV;1} = \mathbf{L}\mathbf{S}^{+}\mathbf{1}_k/\mathbf{1}_k^T\mathbf{S}^{+}\mathbf{1}_k$, $\widetilde{\alpha} = \alpha/(n-1)$, $\hat{s}^* = (n-1)\overline{\mathbf{x}}^{*T}\mathbf{R}\overline{\mathbf{x}}^*/\overline{\mathbf{x}}^{*T}\widehat{\mathbf{R}}\overline{\mathbf{x}}^*$, and $\widehat{\boldsymbol{\theta}}_{MV;2}(\overline{\mathbf{x}}^*) = \overline{\mathbf{x}}^{*T}\mathbf{R}\overline{\mathbf{x}}^*\mathbf{L}\widehat{\mathbf{R}}\overline{\mathbf{x}}^*/\overline{\mathbf{x}}^{*T}\widehat{\mathbf{R}}\overline{\mathbf{x}}^*$.

In Theorem 5, we present the density function of $\widehat{\boldsymbol{\theta}}_{MV;1}$, which is the sample estimator of the linear transformation for the weights of the global minimum variance portfolio and plays an important role in the portfolio theory.

Theorem 5. *Let $\mathbf{x}_1, ..., \mathbf{x}_n$ be i.i.d. random vectors with $\mathbf{x}_1 \sim N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $k > n-1$ and let $rank(\boldsymbol{\Sigma}) = r \le n-1$. Consider $\mathbf{L}$ a $p \times k$ non-random matrix with $rank(\mathbf{L}^T, \mathbf{1}_k) = p+1 \le r$ and set $\boldsymbol{\theta}_{MV;1} = \mathbf{L}\boldsymbol{\Sigma}^{+}\mathbf{1}_k/\mathbf{1}_k^T\boldsymbol{\Sigma}^{+}\mathbf{1}_k$. Then the density function of $\widehat{\boldsymbol{\theta}}_{MV;1}$ is given by*

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

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 $\mathbf{A}$.

Proof. Let $\widetilde{\mathbf{L}} = (\mathbf{L}^T, \mathbf{1}_k)^T$ and define $\widetilde{\mathbf{S}} = \widetilde{\mathbf{L}}\mathbf{S}^{+}\widetilde{\mathbf{L}}^T = \left\{\widetilde{\mathbf{S}}_{ij}\right\}_{i,j=1,2}$ with $\widetilde{\mathbf{S}}_{11} = \mathbf{L}\mathbf{S}^{+}\mathbf{L}^T$, $\widetilde{\mathbf{S}}_{12} = \mathbf{L}\mathbf{S}^{+}\mathbf{1}_k$, $\widetilde{\mathbf{S}}_{21} = \mathbf{1}_k^T\mathbf{S}^{+}\mathbf{L}^T$ and $\widetilde{S}_{22} = \mathbf{1}_k^T\mathbf{S}^{+}\mathbf{1}_k$. Similarly, let $\widetilde{\boldsymbol{\Sigma}} = \widetilde{\mathbf{L}}\boldsymbol{\Sigma}^{+}\widetilde{\mathbf{L}}^T = \{\widetilde{\boldsymbol{\Sigma}}_{ij}\}_{i,j=1,2}$ with $\widetilde{\boldsymbol{\Sigma}}_{11} = \mathbf{L}\boldsymbol{\Sigma}^{+}\mathbf{L}^T$, $\widetilde{\boldsymbol{\Sigma}}_{12} = \mathbf{L}\boldsymbol{\Sigma}^{+}\mathbf{1}_k$, $\widetilde{\boldsymbol{\Sigma}}_{21} = \mathbf{1}_k^T\boldsymbol{\Sigma}^{+}\mathbf{L}^T$ and $\widetilde{\Sigma}_{22} = \mathbf{1}_k^T\boldsymbol{\Sigma}^{+}\mathbf{1}_k$. Then it holds that $\widehat{\boldsymbol{\theta}}_{MV;1} = \widetilde{S}_{22}^{-1}\widetilde{\mathbf{S}}_{12}$ and $\boldsymbol{\theta}_{MV;1} = \widetilde{\Sigma}_{22}^{-1}\widetilde{\boldsymbol{\Sigma}}_{12}$. Because $(n-1)\mathbf{S} \sim \mathcal{W}_k(n-1, \boldsymbol{\Sigma})$ and $\text{rank}(\widetilde{\mathbf{L}}) = p+1 \le r$ we get from Theorem 1 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 $\widetilde{\boldsymbol{\Sigma}}$, i.e. $(n-1)^{-1}\widetilde{\mathbf{S}} \sim \mathcal{IW}_{p+1}(n-r+2p+2, \widetilde{\boldsymbol{\Sigma}})$. Using Theorem 3 (d) of Bodnar and Okhrin (2008) we get the density function of $\widehat{\boldsymbol{\theta}}_{MV;1}$ through

$$\begin{aligned} f_{\widehat{\boldsymbol{\theta}}_{MV;1}}(\mathbf{x}) &\sim \left[1 + \widetilde{\Sigma}_{22}(\mathbf{x} - \widetilde{\Sigma}_{22}^{-1}\widetilde{\boldsymbol{\Sigma}}_{12})^T\widetilde{\boldsymbol{\Sigma}}_{11\cdot 2}^{-1}(\mathbf{x} - \widetilde{\Sigma}_{22}^{-1}\widetilde{\boldsymbol{\Sigma}}_{12})\right]^{-(n-r+p+1)/2}, \\ &= [1 + \mathbf{1}_k^T\boldsymbol{\Sigma}^{+}\mathbf{1}_k(\mathbf{x} - \boldsymbol{\theta}_{MV;1})^T(\mathbf{L}\mathbf{R}\mathbf{L}^T)^{-1}(\mathbf{x} - \boldsymbol{\theta}_{MV;1})]^{-(n-r+p+1)/2}. \end{aligned}$$

where $\widetilde{\boldsymbol{\Sigma}}_{11\cdot 2} = \widetilde{\boldsymbol{\Sigma}}_{11} - \widetilde{\boldsymbol{\Sigma}}_{12}\widetilde{\boldsymbol{\Sigma}}_{21}/\widetilde{\Sigma}_{22}$. This concludes the argument. □

Applying the distributional properties of the multivariate t-distribution we have that

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

Theorem 5 says that $\widehat{\boldsymbol{\theta}}_{MV;1}$ belongs to the same class of distribution and has the same mathematical expectation as in the non-singular case (see Bodnar and Schmid (2008)). The difference is present in the degrees of freedom of the t-distribution only.

Let $\mathbf{b}^* = \mathbf{L}\mathbf{R}\mathbf{L}^T - \mathbf{L}\mathbf{R}\overline{\mathbf{x}}^*\overline{\mathbf{x}}^{*T}\mathbf{R}\mathbf{L}^T/\overline{\mathbf{x}}^{*T}\mathbf{R}\overline{\mathbf{x}}^*$ and let $\mathbf{M}^T = (\mathbf{L}^T, \overline{\mathbf{x}}^*, \mathbf{1}_k)$ with $rank(\mathbf{M}) = p + 2 \leq r$. In Theorem 6 we derived the joint density function of $\widehat{\boldsymbol{\theta}}_{MV;1}$, $\widehat{\boldsymbol{\theta}}_{MV;2}(\overline{\mathbf{x}}^*)$ and $\hat{s}^*$.

Theorem 6. *Let $\mathbf{x}_1, ..., \mathbf{x}_n$ be i.i.d. random vectors with $\mathbf{x}_1 \sim N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $k > n-1$ and with $rank(\boldsymbol{\Sigma}) = r \leq n-1$. Consider $\mathbf{L}$ a $p \times k$ non-random matrix such that $rank(\mathbf{L}^T, \overline{\mathbf{x}}^*, \mathbf{1}_k) = p+2 \leq r$ and $\mathbf{R}$ that is defined in Theorem 5. Then $\widehat{\boldsymbol{\theta}}_{MV;1}$, $\widehat{\boldsymbol{\theta}}_{MV;2}(\overline{\mathbf{x}}^*)$, and $\hat{s}^*$ are mutually independently distributed according to*

$$\begin{aligned}
\widehat{\boldsymbol{\theta}}_{MV;1} &\sim t_p\left(n-r+1, \boldsymbol{\theta}_{MV;1}, \frac{1}{n-r+1}\frac{\mathbf{L}\mathbf{R}\mathbf{L}^T}{\mathbf{1}_k^T\boldsymbol{\Sigma}^+\mathbf{1}_k}\right),\\
\widehat{\boldsymbol{\theta}}_{MV;2}(\overline{\mathbf{x}}^*) &\sim t_p\left(n-r+2, \mathbf{L}\mathbf{R}\overline{\mathbf{x}}^*, \frac{1}{n-r+2}\overline{\mathbf{x}}^{*T}\mathbf{R}\overline{\mathbf{x}}^*\mathbf{b}^*\right),\\
\hat{s}^* &\sim \chi^2_{n-r+1}.
\end{aligned}$$

Proof. Let $\mathbf{H} = \mathbf{M}\boldsymbol{\Sigma}^+\mathbf{M}^T = \{\mathbf{H}_{ij}\}_{i,j=1,2}$ with $H_{22} = \mathbf{1}_k^T\boldsymbol{\Sigma}^+\mathbf{1}_k$ and let $\widehat{\mathbf{H}} = \mathbf{M}\mathbf{S}^+\mathbf{M}^T = \{\widehat{\mathbf{H}}_{ij}\}_{i,j=1,2}$ with $\widehat{H}_{22} = \mathbf{1}_k^T\mathbf{S}^+\mathbf{1}_k$. Similarly, let $\mathbf{G} = \mathbf{H}_{11} - \mathbf{H}_{12}\mathbf{H}_{21}/H_{22} = \{\mathbf{G}_{ij}\}_{i,j=1,2}$ with $G_{22} = \overline{\mathbf{x}}^{*T}\boldsymbol{\Sigma}^+\overline{\mathbf{x}}^* - (\overline{\mathbf{x}}^{*T}\boldsymbol{\Sigma}^+\mathbf{1}_k)^2/\mathbf{1}_k^T\boldsymbol{\Sigma}^+\mathbf{1}_k = \overline{\mathbf{x}}^{*T}\mathbf{R}\overline{\mathbf{x}}^*$ and let $\widehat{\mathbf{G}} = \widehat{\mathbf{H}}_{11} - \widehat{\mathbf{H}}_{12}\widehat{\mathbf{H}}_{21}/\widehat{H}_{22} = \{\widehat{\mathbf{G}}_{ij}\}_{i,j=1,2}$ with $\widehat{G}_{22} = \overline{\mathbf{x}}^{*T}\mathbf{S}^+\overline{\mathbf{x}}^* - (\overline{\mathbf{x}}^{*T}\mathbf{S}^+\mathbf{1}_k)^2/\mathbf{1}_k^T\mathbf{S}^+\mathbf{1}_k = \overline{\mathbf{x}}^{*T}\widehat{\mathbf{R}}\overline{\mathbf{x}}^*$.

Then

$$\widehat{\boldsymbol{\theta}}_{MV}(\overline{\mathbf{x}}^*) = \frac{\mathbf{E}\widehat{\mathbf{H}}_{12}}{\widehat{H}_{22}} + \alpha^{-1}\widetilde{\mathbf{G}}_{12},$$

where $\mathbf{E} = (\mathbf{e}_1, ..., \mathbf{e}_p, \mathbf{0}_k)$ with $(\mathbf{e}_1, ..., \mathbf{e}_p)$ being the usual basis in R^p and $\mathbf{0}_k$ is the k-dimensional zero vector. Additionally, let $\widehat{\mathbf{b}}^* = \widehat{\mathbf{H}}_{11} - \widehat{\mathbf{H}}_{12}\widehat{\mathbf{H}}_{21}/\widehat{H}_{22}$ and $\mathbf{b}^* = \mathbf{H}_{11} - \mathbf{H}_{12}\mathbf{H}_{21}/H_{22}$.

The unconditional distribution of $\widehat{\boldsymbol{\theta}}_{MV;1}$ has already been derived in Theorem 5. Next, we prove that $\widehat{\boldsymbol{\theta}}_{MV;1}$, $\widehat{\boldsymbol{\theta}}_{MV;2}(\overline{\mathbf{x}}^*)$, and $\hat{s}^*$ are mutually independently distributed and derive the distribution of $\widehat{\boldsymbol{\theta}}_{MV;2}(\overline{\mathbf{x}}^*)$ and $\hat{s}^*$. Using Theorem 1 we obtain that

$$(n-1)\widehat{\mathbf{H}}^{-1} = (n-1)(\mathbf{M}\mathbf{S}^+\mathbf{M}^T)^{-1} \sim \mathcal{W}_{p+2}(n-r+p+1, (\mathbf{M}\boldsymbol{\Sigma}^+\mathbf{M})^{-1}). \tag{19}$$

From (19) and Theorem 3.4.1 of Gupta and Nagar (2000) we get

$$(n-1)^{-1}\widehat{\mathbf{H}} \sim \mathcal{IW}_{p+2}(n-r+2p+4, \mathbf{H}). \tag{20}$$

Applying Theorem 3 of Bodnar and Okhrin (2008) we obtain that

$$\begin{aligned}
(n-1)^{-1}\widehat{\mathbf{G}} &\sim \mathcal{IW}_{p+1}(n-r+2p+3,\mathbf{G}),\\
(n-1)^{-1}\widehat{H}_{22} &\sim \mathcal{IW}_{1}(n-r+2,H_{22}),\\
(n-1)^{-1}\widehat{\mathbf{H}}_{12}|(n-1)^{-1}\widehat{H}_{22},\ (n-1)^{-1}\widehat{\mathbf{G}} &\sim \mathcal{N}\left((n-1)^{-1}\mathbf{H}_{12}H_{22}^{-1}\widehat{H}_{22},(n-1)^{-3}\frac{\widehat{H}_{22}^{2}}{H_{22}}\widehat{\mathbf{G}}\right).
\end{aligned}$$

It leads to

$$\frac{(n-1)^{-1}\mathbf{E}\widehat{\mathbf{H}}_{12}}{(n-1)^{-1}\widehat{H}_{22}}|(n-1)^{-1}\widehat{H}_{22},\ (n-1)^{-1}\widehat{\mathbf{G}}\sim\mathcal{N}\left(\mathbf{E}\frac{\mathbf{H}_{12}}{H_{22}},(n-1)^{-1}\frac{\mathbf{E}\widehat{\mathbf{G}}\mathbf{E}^{T}}{H_{22}}\right).$$

Using the fact that $\mathbf{E}\widehat{\mathbf{G}}\mathbf{E}^{T}=\widehat{\mathbf{G}}_{11}$ we obtain that the conditional distribution of $\widehat{\boldsymbol{\theta}}_{MV;1}$ does not depend on $\widehat{\mathbf{G}}_{12}$, $\widehat{G}_{22}$ and $\widehat{H}_{22}$. As a result, the unconditional distribution is independent of $\widehat{\mathbf{G}}_{12}$ and $\widehat{\mathbf{G}}_{22}$. From Theorem 3 of Bodnar and Okhrin (2008) it follows that $\widehat{\mathbf{G}}_{12}/\widehat{G}_{22}$ and $\widehat{G}_{22}$ are independent. Moreover, we get

$$(n-1)^{-1}\widehat{G}_{22}\sim\mathcal{IW}_{1}(n-r+3,G_{22}).$$

Finally, from the proof of Theorem 5 it holds that

$$\frac{\widehat{\mathbf{G}}_{12}}{\widehat{G}_{22}}\sim t_{p}\left(n-r+2,\mathbf{LR}\overline{\mathbf{x}}^{*},\frac{1}{n-r+2}\overline{\mathbf{x}}^{*T}\mathbf{R}\overline{\mathbf{x}}^{*}\mathbf{b}^{*}\right).$$

Putting all together we obtain the statement of Theorem 6 . □

The stochastic representation of $\widehat{\boldsymbol{\theta}}_{MV}$ is derived in the following theorem.

Theorem 7. *Let $\mathbf{x}_1,...,\mathbf{x}_n$ be i.i.d. random vectors with $\mathbf{x}_1\sim N_k(\boldsymbol{\mu},\boldsymbol{\Sigma})$, $k>n-1$ and with rank$(\boldsymbol{\Sigma})=r\le n-1$. Consider $\mathbf{L}$ a $p\times k$ non-random matrix with rank$(\mathbf{L}^T,\mathbf{1}_k)=p+1\le r$ and $\mathbf{R}$ that is defined in Theorem 5. Additionally, let $\mathbf{S}_1=(\mathbf{LRL}^T)^{-1/2}\mathbf{LR}^{1/2}$ and $\mathbf{Q}_1=\mathbf{S}_1^T\mathbf{S}_1$. Then the stochastic representation of $\widehat{\boldsymbol{\theta}}_{MV}$ is given by*

$$\begin{aligned}
\widehat{\boldsymbol{\theta}}_{MV} &\stackrel{d}{=} \widehat{\boldsymbol{\theta}}_{MV;1}+\widetilde{\alpha}^{-1}\hat{s}^{*-1}\mathbf{LR}\overline{\mathbf{x}}+\frac{\widetilde{\alpha}^{-1}\hat{s}^{*-1}}{\sqrt{n-r+2}}(\mathbf{LRL}^{T})^{1/2}\\
&\times\left[\sqrt{\overline{\mathbf{x}}^{T}\mathbf{R}\overline{\mathbf{x}}}\mathbf{I}_{p}-\frac{\sqrt{\overline{\mathbf{x}}^{T}\mathbf{R}\overline{\mathbf{x}}}-\sqrt{\overline{\mathbf{x}}^{T}(\mathbf{R}-\mathbf{Q}_1)\overline{\mathbf{x}}}}{\overline{\mathbf{x}}^{T}\mathbf{Q}_1\overline{\mathbf{x}}}\mathbf{S}_1\overline{\mathbf{x}}\overline{\mathbf{x}}^{T}\mathbf{S}_1^{T}\right]\mathbf{t}_0,
\end{aligned}$$

where $\hat{s}^{*}\sim\chi^2_{n-r+1}$, $\widehat{\boldsymbol{\theta}}_{MV;1}\sim t_p\left(n-r+1,\boldsymbol{\theta}_{MV;1},\frac{1}{n-r+1}\frac{\mathbf{LRL}^T}{\mathbf{1}_k^T\boldsymbol{\Sigma}^{+}\mathbf{1}_k}\right)$, $\overline{\mathbf{x}}\sim\mathcal{N}_k\left(\boldsymbol{\mu},\frac{1}{n}\boldsymbol{\Sigma}\right)$, and $\mathbf{t}_0\sim t_p(n-r+2,\mathbf{0},\mathbf{I}_p)$; moreover, $\hat{s}^{*}$, $\widehat{\boldsymbol{\theta}}_{MV;1}$, $\mathbf{t}_0$, and $\overline{\mathbf{x}}$ are mutually independent.

Proof. From (18) and Theorem 6 we obtain the stochastic representation of $\widehat{\boldsymbol{\theta}}_{MV}$:

$$\widehat{\boldsymbol{\theta}}_{MV} \stackrel{d}{=} \widehat{\boldsymbol{\theta}}_{MV;1} + \widetilde{\alpha}^{-1}\hat{s}^{*-1}\mathbf{LR}\overline{\mathbf{x}} + \frac{\widetilde{\alpha}^{-1}\hat{s}^{*-1}}{\sqrt{n-r+2}}\left[\overline{\mathbf{x}}^T\mathbf{R}\overline{\mathbf{x}}\mathbf{LRL}^T - \mathbf{LR}\overline{\mathbf{x}}\overline{\mathbf{x}}^T\mathbf{RL}^T\right]^{1/2}\mathbf{t}_0,$$

where $\hat{s}^* \sim \chi^2_{n-r+1}$, $\widehat{\boldsymbol{\theta}}_{MV;1} \sim t_p\left(n-r+1, \boldsymbol{\theta}_{MV;1}, \frac{1}{n-r+1}\frac{\mathbf{LRL}^T}{\mathbf{1}_k^T\boldsymbol{\Sigma}^+\mathbf{1}_k}\right)$, $\overline{\mathbf{x}} \sim \mathcal{N}_k\left(\boldsymbol{\mu}, \frac{1}{n}\boldsymbol{\Sigma}\right)$, and $\mathbf{t}_0 \sim t_p(n-r+2, \mathbf{0}, \mathbf{I}_p)$; moreover, $\hat{s}^*$, $\widehat{\boldsymbol{\theta}}_{MV;1}$, $\mathbf{t}_0$, and $\overline{\mathbf{x}}$ are mutually independent.

Now we calculate the square root of $(\overline{\mathbf{x}}^T\mathbf{R}\overline{\mathbf{x}}\mathbf{LRL}^T - \mathbf{LR}\overline{\mathbf{x}}\overline{\mathbf{x}}^T\mathbf{RL}^T)$ using the following equality

$$(\mathbf{B} - \mathbf{c}\mathbf{c}^T)^{1/2} = \mathbf{B}^{1/2}(\mathbf{I}_p - d\mathbf{B}^{-1/2}\mathbf{c}\mathbf{c}^T\mathbf{B}^{-1/2})$$

with $d = \frac{1-\sqrt{1-\mathbf{c}^T\mathbf{B}^{-1}\mathbf{c}}}{\mathbf{c}^T\mathbf{B}^{-1}\mathbf{c}}$, $\mathbf{c} = \mathbf{LR}\overline{\mathbf{x}}$, and $\mathbf{B} = \overline{\mathbf{x}}^T\mathbf{R}\overline{\mathbf{x}}\mathbf{LRL}^T$ that leads to

$$\begin{aligned}\widehat{\boldsymbol{\theta}}_{MV} &\stackrel{d}{=} \widehat{\boldsymbol{\theta}}_{MV;1} + \widetilde{\alpha}^{-1}\hat{s}^{*-1}\mathbf{LR}\overline{\mathbf{x}} + \frac{\widetilde{\alpha}^{-1}\hat{s}^{*-1}}{\sqrt{n-r+2}}(\mathbf{LRL}^T)^{1/2} \\ \times &\left[\sqrt{\overline{\mathbf{x}}^T\mathbf{R}\overline{\mathbf{x}}}\mathbf{I}_p - \frac{\sqrt{\overline{\mathbf{x}}^T\mathbf{R}\overline{\mathbf{x}}} - \sqrt{\overline{\mathbf{x}}^T(\mathbf{R}-\mathbf{Q}_1)\overline{\mathbf{x}}}}{\overline{\mathbf{x}}^T\mathbf{Q}_1\overline{\mathbf{x}}}\mathbf{S}_1\overline{\mathbf{x}}\overline{\mathbf{x}}^T\mathbf{S}_1^T\right]\mathbf{t}_0\end{aligned}$$

with $\mathbf{S}_1 = (\mathbf{LRL}^T)^{-1/2}\mathbf{LR}^{1/2}$ and $\mathbf{Q}_1 = \mathbf{S}_1^T\mathbf{S}_1$. □

In the next corollary, we consider the special case of $p = 1$ and $\mathbf{L} = \mathbf{l}^T$.

Corollary 2. *Let $\mathbf{x}_1, ..., \mathbf{x}_n$ be i.i.d. random vectors with $\mathbf{x}_1 \sim N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $k > n-1$ and with rank$(\boldsymbol{\Sigma}) = r \le n-1$. Also, let $\mathbf{l}$ be a k-dimensional vector of constants and rank$(\mathbf{l}^T, \mathbf{1}_k) = 2 \le r$. Then the stochastic representation of $\widehat{\theta}_{MV}$ is given by*

$$\widehat{\theta}_{MV} \stackrel{d}{=} \widehat{\theta}_{MV;1} + \frac{\widetilde{\alpha}^{-1}}{\hat{s}^*}\left(\mathbf{l}^T\mathbf{R}\boldsymbol{\mu} + \sqrt{\frac{\left(1+\frac{r-2}{n-r+2}u_2\right)\mathbf{l}^T\mathbf{R}\mathbf{l}}{n}}u_1\right),$$

where $\widehat{\theta}_{MV;1} \sim t(n-r+1, \theta_{MV;1}, \frac{1}{n-r+1}\mathbf{l}^T\mathbf{R}\mathbf{l}/\mathbf{1}_k^T\boldsymbol{\Sigma}^+\mathbf{1}_k)$, $\hat{s}^* \sim \chi^2_{n-r+1}$, $u_1 \sim \mathcal{N}(0,1)$, and $u_2 \sim F\left(\frac{r-2}{2}, \frac{n-r+2}{2}, n\Lambda\right)$ with $\Lambda = \boldsymbol{\mu}^T\mathbf{R}\boldsymbol{\mu} - (\mathbf{l}^T\mathbf{R}\boldsymbol{\mu})^2/\mathbf{l}^T\mathbf{R}\mathbf{l}$. Here $F(k_1, k_2, \lambda)$ denotes the non-central F-distribution with k_1 and k_2 degrees of freedom and non-centrality parameter λ. Moreover, the random variables $\widehat{\theta}_{MV;1}$, $\hat{s}^*$, u_1 and u_2 are mutually independently distributed.

The application of Theorem 7 and Corollary 2 leads to the expression of the characteristic function of $\widehat{\mathbf{w}}_{MV}$ which is given in the following corollary.

Corollary 3. *Let $\mathbf{x}_1, ..., \mathbf{x}_n$ be i.i.d. random vectors with $\mathbf{x}_1 \sim N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $k > n-1$ and with rank$(\boldsymbol{\Sigma}) = r \le n-1$. Additionally, let $p = 1$, $r \ge 2$, and rank$(\mathbf{M}) = 3$. Then with* *the notation of the previous corollary, the characteristic function of $\widehat{\mathbf{w}}_{MV}$ is given by*

$$\begin{aligned}
\varphi_{\widehat{\mathbf{w}}_{MV}}(\mathbf{t}) &= \varphi_{t(n-r+1,\theta_{MV;1},\frac{1}{n-r+1}\mathbf{t}^T\mathbf{R}\mathbf{t}/\mathbf{1}_k^T\mathbf{\Sigma}^+\mathbf{1}_k)}(1)\exp\left(-\frac{n\Lambda}{2}\right)\\
&\times \int_0^\infty \exp\left(i\frac{\mathbf{t}^T\mathbf{R}\boldsymbol{\mu}}{\widetilde{\alpha}v}-\frac{\mathbf{t}^T\mathbf{R}\mathbf{t}}{2n\widetilde{\alpha}^2v^2}\right)f_{\chi^2_{n-r+1}}(v)\\
&\times \sum_{i=0}^\infty \frac{\left(\frac{n\Lambda}{2}\right)^i}{i!}\,{}_1F_1\left(\frac{r-2}{2}+j,-\frac{n-r+2}{2},\frac{\tilde{u}\mathbf{t}^T\mathbf{R}\mathbf{t}}{2n\widetilde{\alpha}^2v^2}\right)dv,
\end{aligned}$$

where ${}_1F_1(\cdot,\cdot,\cdot)$ is the confluent hypergeometric function (see Andrews et al. (2000)).

Proof. From Corollary 2 the density function of $\widehat{\theta}_{MV}=\mathbf{t}^T\widehat{\mathbf{w}}_{MV}$ is given by

$$\begin{aligned}
f_{\widehat{\theta}_{MV}}(y) &= \widetilde{\alpha}\frac{n-r+2}{r-2}\int_{-\infty}^\infty f_{t(n-r+1,\theta_{MV;1},\frac{1}{n-r+1}\mathbf{t}^T\mathbf{R}\mathbf{t}/\mathbf{1}_k^T\mathbf{\Sigma}^+\mathbf{1}_k)}(y-\omega_1)\\
&\times \int_0^\infty \omega_2 f_{\chi^2_{n-r+1}}(\omega_2)\int_0^\infty f_{\mathcal{N}(\mathbf{t}^T\mathbf{R}\boldsymbol{\mu},(1+\omega_3)\mathbf{t}^T\mathbf{R}\mathbf{t}/n)}(\widetilde{\alpha}\omega_2\omega_1)\\
&\times f_{F\left(\frac{r-2}{2},\frac{n-r+2}{2},n\Lambda\right)}\left(\frac{n-r+2}{r-2}\omega_3\right)d\omega_1 d\omega_2 d\omega_3, \qquad (21)
\end{aligned}$$

where f subindexed by a distribution stands for the density of this distribution.

Since $\varphi_{\widehat{\mathbf{w}}_{MV}}(\mathbf{t})=\varphi_{\widehat{\theta}_{MV}}(1)$, the conclusion follows from the proof of Corollary 3.5 of Bodnar and Schmid (2011). □

Another important application of Theorem 7 leads to the conditional and unconditional moments of higher order of $\widehat{\boldsymbol{\theta}}_{MV}$. Let the symbol $m_{i_1,\ldots,i_p}(\boldsymbol{\mu},\mathbf{\Sigma})$ denote the mixed moment of the p-dimensional normal distribution with parameters $\boldsymbol{\mu}$ and $\mathbf{\Sigma}$, and let

$$C_{n_1,\ldots,n_p}^{k_1,\ldots,k_p}(s)=\prod_{i=1}^p C_{n_i}^{k_i}\frac{\Gamma\left(\frac{s}{2}-\sum_{i_1}^p(n_i-k_i)\right)}{\Gamma(s/2)},$$

where $C_{n_i}^{k_i}=n_i!/k_i!(n_i-k_i)!$ is a binomial coefficient. The statement of the corollary follows from Theorem 7 and the binomial formula which is applied three times and we omit the proof details.

Corollary 4. *Let $\mathbf{x}_1,\ldots,\mathbf{x}_n$ be i.i.d. random vectors with $\mathbf{x}_1\sim N_k(\boldsymbol{\mu},\mathbf{\Sigma})$, $k>n-1$ and with $rank(\mathbf{\Sigma})=r\le n-1$. Consider $\mathbf{L}$ a $p\times k$ non-random matrix such that $rank(\mathbf{L}^T,\overline{\mathbf{x}}^*,\mathbf{1}_k)=p+2\le r$. Then the conditional mixed moments of the $\widehat{\boldsymbol{\theta}}_{MV}(\overline{\mathbf{x}}^*)$ are* *given by*

$$\begin{aligned}
E_{n_1,\dots,n_p} &= E\left(\prod_{i=1}^{p}(\mathbf{e}_i^T\widehat{\boldsymbol{\theta}}_{MV;1}+\widetilde{\alpha}^{-1}\hat{s}^{*-1}\mathbf{e}_i^T\widehat{\boldsymbol{\theta}}_{MV;2})^{n_i}|\overline{\mathbf{x}}=\overline{\mathbf{x}}^*\right)\\
&= \sum_{i=1}^{p}\sum_{j_i=0}^{n_i}\widetilde{\alpha}^{-\sum_{i=1}^{p}(n_i-j_i)}C_{n_1,\dots,n_p}^{j_1,\dots,j_p}(n-r+1)\\
&\times \left(\sum_{i=1}^{p}\sum_{k_i=0}^{j_i}C_{j_1,\dots,j_p}^{k_1,\dots,k_p}(n-r+1)\right.\\
&\times \left.\prod_{i=1}^{p}(\mathbf{l}_i^T\boldsymbol{\theta}_{MV})^{k_i}m_{j_1-k_1,\dots,j_p-k_p}\left(\mathbf{0},\frac{1}{n-r+1}\frac{\mathbf{L}\mathbf{R}\mathbf{L}^T}{\mathbf{1}_k^T\boldsymbol{\Sigma}^+\mathbf{1}_k}\right)\right)\\
&\times \left(\sum_{i=1}^{p}\sum_{q_i=0}^{n_i-j_i}C_{n_1-j_1,\dots,n_p-j_p}^{q_1,\dots,q_p}(n-r+2)\right.\\
&\times \left.\prod_{i=1}^{p}(\mathbf{l}_i^T\mathbf{R}\overline{\mathbf{x}}^*)^{q_i}m_{n_1-j_1-q_1,\dots,n_p-j_p-q_p}\left(\mathbf{0},\frac{\mathbf{b}^*\overline{\mathbf{x}}^{*T}\mathbf{R}\overline{\mathbf{x}}^*}{n-r+2}\right)\right),
\end{aligned}$$

where $\mathbf{L}^T=(\mathbf{l}_1,\dots,\mathbf{l}_p)$.

The above result can be used to obtain the formula for the unconditional mean and variance of the estimator

$$E(\widehat{\boldsymbol{\theta}}_{MV})=\frac{\mathbf{L}\boldsymbol{\Sigma}^+\mathbf{1}_k}{\mathbf{1}_k^T\boldsymbol{\Sigma}^+\mathbf{1}_k}+\frac{n-1}{n-r-1}\alpha^{-1}\mathbf{L}\mathbf{R}\boldsymbol{\mu} \tag{22}$$

and

$$\begin{aligned}
Var(\widehat{\boldsymbol{\theta}}_{MV}) &= \frac{1}{n-r-1}\frac{\mathbf{L}\mathbf{R}\mathbf{L}^T}{\mathbf{1}_k^T\boldsymbol{\Sigma}^+\mathbf{1}_k}+\alpha^{-2}(c_1\boldsymbol{\mu}^T\mathbf{R}\boldsymbol{\mu}\mathbf{L}\mathbf{R}\mathbf{L}^T+c_2\mathbf{L}\mathbf{R}\boldsymbol{\mu}\boldsymbol{\mu}^T\mathbf{R}\mathbf{L}^T)\\
&+ \frac{\alpha^{-2}}{n}\left(c_2+c_1(r-1)+\frac{(n-1)^2}{(n-r+1)^2}\right)\mathbf{L}\mathbf{R}\mathbf{L}^T. \qquad (23)
\end{aligned}$$

Indeed, from Corollary 4 it holds that

$$E(\widehat{\boldsymbol{\theta}}_{MV}|\overline{\mathbf{x}}=\overline{\mathbf{x}}^*)=\frac{\mathbf{L}\boldsymbol{\Sigma}^+\mathbf{1}_k}{\mathbf{1}_k^T\boldsymbol{\Sigma}^+\mathbf{1}_k}+\frac{n-1}{n-r-1}\alpha^{-1}\mathbf{L}\mathbf{R}\overline{\mathbf{x}}^*$$

and

$$Var(\widehat{\boldsymbol{\theta}}_{MV}|\overline{\mathbf{x}}=\overline{\mathbf{x}}^*)=\frac{1}{n-r-1}\frac{\mathbf{L}\mathbf{R}\mathbf{L}^T}{\mathbf{1}_k^T\boldsymbol{\Sigma}^+\mathbf{1}_k}+\alpha^{-2}(c_1\overline{\mathbf{x}}^{*T}\mathbf{R}\overline{\mathbf{x}}^*\mathbf{L}\mathbf{R}\mathbf{L}^T+c_2\mathbf{L}\mathbf{R}\overline{\mathbf{x}}^*\overline{\mathbf{x}}^{*T}\mathbf{R}\mathbf{L}^T)$$

with

$$c_1=\frac{(n-1)^2}{(n-r)(n-r-1)(n-r-3)} \quad\text{and}\quad c_2=\frac{(n-1)^2(n-r+1)}{(n-r)(n-r-1)^2(n-r-3)}.$$

The final form of the conditional mean and variance, then follow easily from the following standard relations

$$E(\widehat{\boldsymbol{\theta}}_{MV}) = E\left(E(\widehat{\boldsymbol{\theta}}_{MV}|\overline{\mathbf{x}} = \overline{\mathbf{x}}^*)\right)$$

and

$$Var(\widehat{\boldsymbol{\theta}}_{MV}) = E\left(Var(\widehat{\boldsymbol{\theta}}_{MV}|\overline{\mathbf{x}} = \overline{\mathbf{x}}^*)\right) + Var\left(E(\widehat{\boldsymbol{\theta}}_{MV}|\overline{\mathbf{x}} = \overline{\mathbf{x}}^*)\right)$$

See also Bodnar and Schmid (2011) for more details.

5 Summary

Distributional properties of the linear symmetric transformations of the inverse sample covariance matrix are very important tool for derivation of the distribution of the estimated optimal portfolio weights. In the present paper we provide its distribution when the sample size is smaller than the size of portfolio and the covariance matrix is singular. Several important special cases of the transformations are considered and can be utilize in the portfolio theory. Assuming independent and multivariate normally distributed returns we prove stochastic independence between the sample mean vector and the sample covariance matrix, and derive their distributions under the singularity. Moreover, we extend results which are obtained by Bodnar and Schmid (2011) by providing a stochastic representation of the estimated MV portfolio weights. Additionally, we obtain the expressions of the characteristic function and the moments of higher order.

References

Andrews, G. E., Askey, R., and Roy, R. (2000). Special functions. Cambridge University Press.

Bodnar, T., Mazur, S., and Okhrin, Y. (2013). On exact and approximate distributions of the product of the Wishart matrix and normal vector. Journal of Multivariate Analysis, 122:70–81.

Bodnar, T., Mazur, S., and Okhrin, Y. (2014). Distribution of the product of singular Wishart Matrix and normal vector. Theory of Probability and Mathematical Statistics, (91):1–15.

Bodnar, T. and Okhrin, Y. (2008). Properties of the singular, inverse and generalized inverse partitioned Wishart distributions. Journal of Multivariate Analysis, 99:2389–2405.

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.

Bodnar, T. and Schmid, W. (2011). On the exact distribution of the estimated expected utility portfolio weights: Theory and applications. Statistics & Risk Modeling, 28:319–342.

Boullion, T. L. and Odell, P. L. (1971). Generalized inverse matrices. Wiley, New York.

Díaz-García, J. A., Gutiérrez-Jáimez, R., and Mardia, K. V. (1997). Wishart and pseudo-Wishart distributions and some applications to shape theory. Journal of Multivariate Analysis, 63:73–87.

Drton, M., Massam, H., and Olkin, I. (2008). Moments of minors of Wishart matrices. Annals of Statistics, 36:2261–2283.

Greville, T. N. E. (1966). Note on the generalized inverse of a product matrix. SIAM Review, 8(4):518–521.

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

Khatri, C. G. (1959). On the mutual independence of certain statistics. Annals of Mathematical Statistics, 30:1258–1262.

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.

Olkin, I. and Roy, S. N. (1954). On multivariate distribution theory. Annals of Mathematical Statistics, 25:329–339.

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. (2003). Singular wishart and multivariate beta distributions. Annals of Statistics, 31:1537–1560.

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