Properties of the singular, inverse and generalized inverse partitioned Wishart distributions

Taras Bodnar, Yarema Okhrin*

Department of Statistics, European University Viadrina, Frankfurt (Oder), Germany

Abstract

In this paper we discuss the distributions and independency properties of several generalizations of the Wishart distribution. First, an analog to Muirhead [R.J. Muirhead, Aspects of Multivariate Statistical Theory, Wiley, New York, 1982] Theorem 3.2.10 for the partitioned matrix $\mathbf{A} = (\mathbf{A}_{ij})_{i,j=1,2}$ is established in the case of arbitrary partitioning for singular and inverse Wishart distributions. Second, the density of $\mathbf{A}_{21}\mathbf{A}_{11}^{-1}$ is derived in the case of singular, non-central singular, inverse and generalized inverse Wishart distributions. The importance of the derived results is illustrated with an example from portfolio theory.

Keywords: Inverse Wishart; Generalized inverse Wishart; Partitioned matrices; Singular Wishart

1. Introduction

The theory of Wishart distribution has been applied in numerous fields of applied and theoretical statistics. The most common applications are the inference procedures based on the sample covariance matrix of Gaussian observations. Let $\mathbf{X} = (\mathbf{X}_1, \ldots, \mathbf{X}_n)$ be a sample of size n from k-variate normal distribution, i.e. $\mathbf{X}_i \sim \mathcal{N}_k(\mathbf{0}, \mathbf{\Sigma})$ for $i = 1, \ldots, n$. Throughout the paper, we assume that $\mathbf{\Sigma}$ is positive definite. First consider the case when the number of observations is larger than the dimension, i.e. $k < n$. Then $\mathbf{A} = \mathbf{X}\mathbf{X}'$ follows the k-variate Wishart distribution with n degrees of freedom, denoted by $\mathcal{W}_k(n, \mathbf{\Sigma})$. Numerous important properties

* Corresponding address: Department of Statistics, European University Viadrina, PO Box 1786, D-15207 Frankfurt (Oder), Germany.

E-mail addresses: bodnar@euv-frankfurt-o.de (T. Bodnar), okhrin@euv-frankfurt-o.de (Y. Okhrin).

of Wishart matrices were established in Khatri [2], Olkin and Roy [3], Olkin and Rubin [4] with Muirhead [1] providing a detailed review.

In an application, it may by necessary to work with the inverse sample covariance matrix or to set the number of observations n less than the dimension k. This implies two important generalizations. The first leads to the inverse Wishart distribution, which we denote by $\mathcal{W}_n^{-1}(k, \mathbf{\Sigma}^{-1})$. The properties of the inverse sample covariance matrix are important for improved estimation of the precision matrices (see Tsukuma and Konno [5]). It plays a special role in Bayesian analysis, where the posterior distribution of the covariance matrix with a non-informative prior follows the inverse Wishart (see Zellner [6], Muirhead [1], Gelman et al. [7]). Also, in the portfolio theory of Markowitz [8], the inverse covariance matrix is the key component of the optimal portfolio weights (see Okhrin and Schmid [9], Bodnar and Schmid [10]). Von Rosen [11] provides many technical results related to higher order moments of the components of Wishart matrices. The second generalization leads to singular Wishart distributions introduced by Khatri [12] and Srivastava and Khatri [13]. The practical relevance of the case $k > n$ is discussed in Uhlig [14] and Ledoit and Wolf [15].

Let us now consider the partitioning of the Wishart matrix

$$\mathbf{A} = \begin{pmatrix} \bar{\mathbf{A}}_{11} & \bar{\mathbf{A}}_{12} \\ \bar{\mathbf{A}}_{21} & \bar{\mathbf{A}}_{22} \end{pmatrix} \quad \text{and} \quad \mathbf{\Sigma} = \begin{pmatrix} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} \\ \mathbf{\Sigma}_{21} & \mathbf{\Sigma}_{22} \end{pmatrix}$$

$$\text{with} \quad \dim(\bar{\mathbf{A}}_{11}) = \dim(\mathbf{\Sigma}_{11}) = m \times m, \quad m < k < n. \tag{1}$$

To distinguish between the singular and nonsingular Wishart distributions, we use bars to denote the partition of $\mathbf{A}$ in the nonsingular case ($n > k$) and use plain notation in the singular case ($n \leq k$, see (4)). Theorem 3.2.10 of Muirhead [1] provides the key properties of the components $\bar{\mathbf{A}}_{ij}$ in terms of $\mathbf{\Sigma}_{ij}$, which appear to be extremely useful in dealing with Wishart matrices. Moreover, as shown by Massam and Wesołowski [16], the independency properties from this theorem can be used for characterization of the Wishart distribution.

Consider now the quantities

$$\mathbf{Q}_1 = \mathbf{L}_2'\mathbf{A}\mathbf{L}_1(\mathbf{L}_1'\mathbf{A}\mathbf{L}_1)^{-1} \quad \text{or} \quad \mathbf{Q}_2 = \mathbf{L}_2'\mathbf{A}^{-1}\mathbf{L}_1(\mathbf{L}_1'\mathbf{A}^{-1}\mathbf{L}_1)^{-1},$$

where $\mathbf{L}_1$ and $\mathbf{L}_2$ are deterministic matrices with dimensions $\ell_1 \times k$ and $\ell_2 \times k$ respectively. The distribution of $\mathbf{Q}_1$ and $\mathbf{Q}_2$ is important in numerous applications. The first application is its clear relation to the multivariate linear regression model estimated via a generalized least squares approach. We can rewrite $\mathbf{Q}_1$ as $\mathbf{V}_{21}\mathbf{V}_{11}^{-1}$, where $\mathbf{V} = (\mathbf{L}_1\, \mathbf{L}_2)'\mathbf{A}(\mathbf{L}_1\, \mathbf{L}_2)$, $\mathbf{V}_{11} = \mathbf{L}_1'\mathbf{A}\mathbf{L}_1$ and $\mathbf{V}_{21} = \mathbf{L}_2'\mathbf{A}\mathbf{L}_1$. Thus the distribution of the quantities of the type $\bar{\mathbf{A}}_{21}\bar{\mathbf{A}}_{11}^{-1}$ is of interest. It is also possible to rewrite $\mathbf{Q}_2$ in a similar manner. The second application and motivation for this paper is to the distributional properties of optimal portfolio weights discussed in Okhrin and Schmid [9], Bodnar and Schmid [10]. It can be shown that the global minimum variance portfolio weights also possess the structure $\bar{\mathbf{A}}_{21}\bar{\mathbf{A}}_{11}^{-1}$.

The first contribution of this paper lies in the extension of properties of partitioned matrices to different generalizations of the Wishart distribution. Some important results for singular Wishart matrices and techniques also partially used in this paper can be found in Díaz-García and Gutiérrez-Jáimez [17], Díaz-García et al. [17] and Srivastava [18]. However, in these papers it is assumed that $m = n < k$. We generalize these results to partitioning with an arbitrary m. Moreover, we also provide results for the inverse Wishart and generalized inverse Wishart distributions. The second contribution of this paper is the derivation of the distribution of $\bar{\mathbf{A}}_{21}\bar{\mathbf{A}}_{11}^{-1}$

in the case of singular, non-central singular, inverse and generalized inverse Wishart distributions. In all cases we obtain explicit expressions for the density function. The importance of the results is illustrated in Section 5 with the example of mean-variance portfolio selection procedures.

2. Singular Wishart distribution

Let $\mathbf{X} = (\mathbf{X}_1, \ldots, \mathbf{X}_n)$ and $\mathbf{X}_i$ be independent identical normally distributed, with $\mathbf{X}_i \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$. It is assumed that $\boldsymbol{\Sigma}$ is positive definite. Let $k > n$. In this case, the distribution of the $k \times k$ matrix $\mathbf{A} = \mathbf{X}\mathbf{X}'$ is called a k-dimensional pseudo-Wishart distribution by Díaz-García et al. [18] and a singular Wishart by Srivastava [19]. Let us first consider the partition

$$\mathbf{A} = \begin{pmatrix} \tilde{\mathbf{A}}_{11} & \tilde{\mathbf{A}}_{12} \\ \tilde{\mathbf{A}}_{21} & \tilde{\mathbf{A}}_{22} \end{pmatrix} \quad \text{and} \quad \boldsymbol{\Sigma} = \tilde{\boldsymbol{\Sigma}} = \begin{pmatrix} \tilde{\boldsymbol{\Sigma}}_{11} & \tilde{\boldsymbol{\Sigma}}_{12} \\ \tilde{\boldsymbol{\Sigma}}_{21} & \tilde{\boldsymbol{\Sigma}}_{22} \end{pmatrix}$$
$$\text{with} \quad \dim(\tilde{\mathbf{A}}_{11}) = \dim(\tilde{\boldsymbol{\Sigma}}_{11}) = n \times n, \quad n < k. \tag{2}$$

Since $n < k$, the component $\tilde{\mathbf{A}}_{22}$ is functionally dependent on the components $\tilde{\mathbf{A}}_{11}$ and $\tilde{\mathbf{A}}_{21}$. It can be explicitly determined by $\tilde{\mathbf{A}}_{22} = \tilde{\mathbf{A}}_{21}\tilde{\mathbf{A}}_{11}^{-1}\tilde{\mathbf{A}}_{12}$. The density of $\mathbf{A}$, i.e. the joint density of the $n \times n$ dimensional matrix $\tilde{\mathbf{A}}_{11}$ and the $n \times (k-n)$ dimensional matrix $\tilde{\mathbf{A}}_{21}$, is given by (see Srivastava [19], p. 1550)

$$\frac{\pi^{n(n-k)/2}2^{-kn/2}}{\Gamma_n(\frac{n}{2})|\boldsymbol{\Sigma}|^{n/2}}|\tilde{\mathbf{A}}_{11}|^{\frac{n-k-1}{2}} \exp\left(-\frac{1}{2}\mathrm{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{A})\right). \tag{3}$$

$\Gamma_k(\cdot)$ denotes the multivariate gamma function given by

$$\Gamma_k\left(\frac{n}{2}\right) = \pi^{k(k-1)/4}\prod_{i=1}^{k}\Gamma\left(\frac{n-i+1}{2}\right).$$

In Corollary 3.4 of Srivastava [19], the author derives the conditional distribution of $\tilde{\mathbf{A}}_{21}|\tilde{\mathbf{A}}_{11}$ and the unconditional distribution of $\tilde{\mathbf{A}}_{11}$. We now consider an arbitrary partition of the matrix $\mathbf{A}$

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix} \quad \text{and} \quad \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix}$$
$$\text{with} \quad \dim(\mathbf{A}_{11}) = \dim(\boldsymbol{\Sigma}_{11}) = m \times m, \quad m < n < k. \tag{4}$$

For notational convenience, let

$$\tilde{\mathbf{A}}_{11} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12}^* \\ \mathbf{A}_{21}^* & \mathbf{A}_{22}^* \end{pmatrix} \quad \text{and} \quad \tilde{\boldsymbol{\Sigma}}_{11} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12}^* \\ \boldsymbol{\Sigma}_{21}^* & \boldsymbol{\Sigma}_{22}^* \end{pmatrix} \tag{5}$$

with $\tilde{\mathbf{A}}_{21}^*$ defined through the relation $\mathbf{A}_{21} = (\mathbf{A}_{21}^{*\prime}\tilde{\mathbf{A}}_{21}^{*\prime})'$. In Lemma 1, we extend the results of Srivastava [19] to the arbitrary partitioning in (4).

Lemma 1. *Let $k > n$ and $\mathbf{A} \sim W_k(n, \boldsymbol{\Sigma})$, where $\mathbf{A}$ and $\boldsymbol{\Sigma}$ are partitioned as in (4). Then it holds that*

(a) $\mathbf{A}_{11} \sim W_m(n, \boldsymbol{\Sigma}_{11})$.
(b) $\mathbf{A}_{21}|\mathbf{A}_{11} \sim \mathcal{N}(\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\mathbf{A}_{11}, \boldsymbol{\Sigma}_{22\cdot 1} \otimes \mathbf{A}_{11})$, *with* $\boldsymbol{\Sigma}_{22\cdot 1} = \boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}$.

Proof. (a) This part is the straightforward consequence of Corollary 3.4a of Srivastava [19] and Theorem 3.2.10 of Muirhead [1].

(b) Let $\mathbf{E}_{n\times m} = (\mathbf{I}_m \ \mathbf{0}_{m\times(n-m)})'$. From Corollary 3.4 of Srivastava [19], we obtain that

$$\tilde{\mathbf{A}}_{21}|\tilde{\mathbf{A}}_{11} \sim \mathcal{N}(\tilde{\mathbf{\Sigma}}_{21}\tilde{\mathbf{\Sigma}}_{11}^{-1}\tilde{\mathbf{A}}_{11}, \tilde{\mathbf{\Sigma}}_{22\cdot 1} \otimes \tilde{\mathbf{A}}_{11}),$$

where $\tilde{\mathbf{\Sigma}}_{22\cdot 1} = \tilde{\mathbf{\Sigma}}_{22} - \tilde{\mathbf{\Sigma}}_{21}\tilde{\mathbf{\Sigma}}_{11}^{-1}\tilde{\mathbf{\Sigma}}_{12}$. This implies that

$$\tilde{\mathbf{A}}_{21}\mathbf{E}_{n\times m}|\tilde{\mathbf{A}}_{11} \sim \mathcal{N}(\tilde{\mathbf{\Sigma}}_{21}\tilde{\mathbf{\Sigma}}_{11}^{-1}\tilde{\mathbf{A}}_{11}\mathbf{E}_{n\times m}, \tilde{\mathbf{\Sigma}}_{22\cdot 1} \otimes \mathbf{E}_{n\times m}'\tilde{\mathbf{A}}_{11}\mathbf{E}_{n\times m}).$$

From the illustration of the partitioning in (5), it follows that

$$\tilde{\mathbf{A}}_{21}\mathbf{E}_{n\times m} = \tilde{\mathbf{A}}_{21}^{*}, \qquad \mathbf{E}_{n\times m}'\tilde{\mathbf{A}}_{11}\mathbf{E}_{n\times m} = \mathbf{A}_{11}, \qquad \tilde{\mathbf{A}}_{11}\mathbf{E}_{n\times m} = \begin{pmatrix}\mathbf{A}_{11}\\ \mathbf{A}_{21}^{*}\end{pmatrix}.$$

Thus the above statement about the distribution of $\tilde{\mathbf{A}}_{21}\mathbf{E}_{n\times m}|\tilde{\mathbf{A}}_{11}$ implies that

$$\tilde{\mathbf{A}}_{21}^{*}|\tilde{\mathbf{A}}_{11} \sim \mathcal{N}\left(\tilde{\mathbf{\Sigma}}_{21}\tilde{\mathbf{\Sigma}}_{11}^{-1}\begin{pmatrix}\mathbf{A}_{11}\\ \mathbf{A}_{21}^{*}\end{pmatrix}, \tilde{\mathbf{\Sigma}}_{22\cdot 1} \otimes \mathbf{A}_{11}\right).$$

Because the right hand side of the last identity does not depend on $\mathbf{A}_{22}^{*}$, we get that $\tilde{\mathbf{A}}_{21}^{*}$ and $\mathbf{A}_{22}^{*}$ are independently distributed. Thus, $\tilde{\mathbf{A}}_{21}^{*}|\tilde{\mathbf{A}}_{11} = \tilde{\mathbf{A}}_{21}^{*}|\mathbf{A}_{11}, \mathbf{A}_{21}^{*}$. Furthermore, it holds that $\mathbf{A}_{21}^{*}|\mathbf{A}_{11} \sim \mathcal{N}(\mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1}\mathbf{A}_{11}, \mathbf{\Sigma}_{22\cdot 1} \otimes \mathbf{A}_{11})$. Putting the results together, the joint conditional density of $\mathbf{A}_{21}^{*}$ and $\tilde{\mathbf{A}}_{21}^{*}$ given $\mathbf{A}_{11}$, i.e. the conditional density of $\mathbf{A}_{21}$ given $\mathbf{A}_{11}$, is

$$\begin{aligned}
&f_{\mathbf{A}_{21}|\mathbf{A}_{11}=\mathbf{C}_{11}}(\mathbf{X}|\mathbf{C}_{11})\\
&= f_{\tilde{\mathbf{A}}_{21}^{*}|\mathbf{A}_{11}=\mathbf{C}_{11},\mathbf{A}_{21}^{*}=\mathbf{X}_1}(\mathbf{X}_2|\mathbf{C}_{11}, \mathbf{X}_1) f_{\mathbf{A}_{11}=\mathbf{C}_{11}}(\mathbf{X}_1|\mathbf{C}_{11})\\
&= \frac{\operatorname{etr}\left(-\frac{1}{2}\tilde{\mathbf{\Sigma}}_{22\cdot 1}^{-1}\left(\mathbf{X}_2 - \tilde{\mathbf{\Sigma}}_{21}\tilde{\mathbf{\Sigma}}_{11}^{-1}\begin{pmatrix}\mathbf{C}_{11}\\ \mathbf{X}_1\end{pmatrix}\right)\mathbf{C}_{11}^{-1}\left(\mathbf{X}_2 - \tilde{\mathbf{\Sigma}}_{21}\tilde{\mathbf{\Sigma}}_{11}^{-1}\begin{pmatrix}\mathbf{C}_{11}\\ \mathbf{X}_1\end{pmatrix}\right)'\right)}{(2\pi)^{\frac{(k-n)m}{2}}|\tilde{\mathbf{\Sigma}}_{22\cdot 1}|^{\frac{m}{2}}|\mathbf{C}_{11}|^{\frac{k-n}{2}}}\\
&\times \frac{\exp\left(-\frac{1}{2}\operatorname{tr}(\mathbf{\Sigma}_{22\cdot 1}^{*-1}(\mathbf{X}_1 - \mathbf{\Sigma}_{21}^{*}\mathbf{\Sigma}_{11}^{-1}\mathbf{C}_{11})\mathbf{C}_{11}^{-1}(\mathbf{X}_1 - \mathbf{\Sigma}_{21}^{*}\mathbf{\Sigma}_{11}^{-1}\mathbf{C}_{11})')\right)}{(2\pi)^{\frac{(n-m)m}{2}}|\mathbf{\Sigma}_{22\cdot 1}^{*}|^{\frac{m}{2}}|\mathbf{C}_{11}|^{\frac{n-m}{2}}}.
\end{aligned}$$

Using the properties of the determinant, we obtain that

$$|\tilde{\mathbf{\Sigma}}_{22\cdot 1}|^{\frac{m}{2}}|\mathbf{\Sigma}_{22\cdot 1}^{*}|^{\frac{m}{2}} = \frac{|\mathbf{\Sigma}|^{\frac{m}{2}}}{|\tilde{\mathbf{\Sigma}}_{11}|^{\frac{m}{2}}}\frac{|\tilde{\mathbf{\Sigma}}_{11}|^{\frac{m}{2}}}{|\mathbf{\Sigma}_{11}|^{\frac{m}{2}}} = \frac{|\mathbf{\Sigma}|^{\frac{m}{2}}}{|\mathbf{\Sigma}_{11}|^{\frac{m}{2}}} = |\mathbf{\Sigma}_{22\cdot 1}|^{\frac{m}{2}}$$

and

$$f_{\mathbf{A}_{21}|\mathbf{A}_{11}=\mathbf{C}_{11}}(\mathbf{X}|\mathbf{C}_{11}) = \frac{\exp\left(-\frac{1}{2}\operatorname{tr}(\mathbf{C}_{11}^{-1}(\mathbf{Q}_1(\mathbf{X}_1, \mathbf{X}_2) + \mathbf{Q}_2(\mathbf{X}_1)))\right)}{(2\pi)^{\frac{(k-m)m}{2}}|\mathbf{\Sigma}_{22\cdot 1}|^{\frac{m}{2}}|\mathbf{C}_{11}|^{\frac{k-m}{2}}}.$$

Let $\mathbf{B} = \mathbf{\Sigma}^{-1}$ with the partitions $(\tilde{\mathbf{B}})_{i,j=1,2}$ and $(\mathbf{B})_{i,j=1,2}$, which correspond to partitions in (2) and (4) respectively. The equalities $\tilde{\mathbf{\Sigma}}_{21}\tilde{\mathbf{\Sigma}}_{11}^{-1} = -\tilde{\mathbf{B}}_{22}^{-1}\tilde{\mathbf{B}}_{21}$ and $\tilde{\mathbf{\Sigma}}_{22\cdot 1}^{-1} = \tilde{\mathbf{B}}_{22}$ yield

$$\begin{aligned}
\mathbf{Q}_1(\mathbf{X}_1, \mathbf{X}_2) &= \left(\mathbf{X}_2 - \tilde{\mathbf{\Sigma}}_{21}\tilde{\mathbf{\Sigma}}_{11}^{-1}\begin{pmatrix}\mathbf{C}_{11}\\ \mathbf{X}_1\end{pmatrix}\right)'\tilde{\mathbf{\Sigma}}_{22\cdot 1}^{-1}\left(\mathbf{X}_2 - \tilde{\mathbf{\Sigma}}_{21}\tilde{\mathbf{\Sigma}}_{11}^{-1}\begin{pmatrix}\mathbf{C}_{11}\\ \mathbf{X}_1\end{pmatrix}\right)\\
&= \left(\mathbf{X}_2 + \tilde{\mathbf{B}}_{22}^{-1}\tilde{\mathbf{B}}_{21}\begin{pmatrix}\mathbf{C}_{11}\\ \mathbf{X}_1\end{pmatrix}\right)'\tilde{\mathbf{B}}_{22}\left(\mathbf{X}_2 + \tilde{\mathbf{B}}_{22}^{-1}\tilde{\mathbf{B}}_{21}\begin{pmatrix}\mathbf{C}_{11}\\ \mathbf{X}_1\end{pmatrix}\right)
\end{aligned}$$

$$= \mathbf{X}_2'\tilde{\mathbf{B}}_{22}\mathbf{X}_2 + \begin{pmatrix}\mathbf{C}_{11}\\ \mathbf{X}_1\end{pmatrix}' \tilde{\mathbf{B}}_{21}'\tilde{\mathbf{B}}_{22}^{-1}\tilde{\mathbf{B}}_{21}\begin{pmatrix}\mathbf{C}_{11}\\ \mathbf{X}_1\end{pmatrix}$$
$$+ \begin{pmatrix}\mathbf{C}_{11}\\ \mathbf{X}_1\end{pmatrix}' \tilde{\mathbf{B}}_{21}'\mathbf{X}_2 + \mathbf{X}_2'\tilde{\mathbf{B}}_{21}\begin{pmatrix}\mathbf{C}_{11}\\ \mathbf{X}_1\end{pmatrix},$$

where $\tilde{\mathbf{B}}_{21}'\tilde{\mathbf{B}}_{22}^{-1}\tilde{\mathbf{B}}_{21} = \tilde{\mathbf{B}}_{11} - \tilde{\mathbf{\Sigma}}_{11}^{-1}$. Similarly as for $\mathbf{A}$ let

$$\tilde{\mathbf{B}}_{11} = \begin{pmatrix}\mathbf{B}_{11} & \mathbf{B}_{12}^*\\ \mathbf{B}_{21}^* & \mathbf{B}_{22}^*\end{pmatrix} \quad \text{and} \quad \breve{\mathbf{B}} = \tilde{\mathbf{\Sigma}}_{11}^{-1} = \begin{pmatrix}\breve{\mathbf{B}}_{11} & \breve{\mathbf{B}}_{12}\\ \breve{\mathbf{B}}_{21} & \breve{\mathbf{B}}_{22}\end{pmatrix}. \tag{6}$$

Then it holds

$$\mathbf{Q}_1(\mathbf{X}_1, \mathbf{X}_2) = \mathbf{X}_2'\tilde{\mathbf{B}}_{22}\mathbf{X}_2 + 2\mathbf{X}_2'\tilde{\mathbf{B}}_{21}\begin{pmatrix}\mathbf{C}_{11}\\ \mathbf{X}_1\end{pmatrix} + \mathbf{C}_{11}(\mathbf{B}_{11} - \breve{\mathbf{B}}_{11})\mathbf{C}_{11}$$
$$+ \mathbf{X}_1'(\mathbf{B}_{21}^* - \breve{\mathbf{B}}_{21})\mathbf{C}_{11} + \mathbf{C}_{11}(\mathbf{B}_{21}^* - \breve{\mathbf{B}}_{21})'\mathbf{X}_1 + \mathbf{X}_1'(\mathbf{B}_{22}^* - \breve{\mathbf{B}}_{22})\mathbf{X}_1.$$

From $\mathbf{\Sigma}_{21}^*\mathbf{\Sigma}_{11}^{-1} = -\breve{\mathbf{B}}_{22}^{-1}\breve{\mathbf{B}}_{21}$ and $\mathbf{\Sigma}_{22\cdot 1}^{*-1} = \breve{\mathbf{B}}_{22}$, we obtain

$$\begin{aligned}
\mathbf{Q}_2(\mathbf{X}_1) &= (\mathbf{X}_1 - \mathbf{\Sigma}_{21}^*\mathbf{\Sigma}_{11}^{-1}\mathbf{C}_{11})'\mathbf{\Sigma}_{22\cdot 1}^{*-1}(\mathbf{X}_1 - \mathbf{\Sigma}_{21}^*\mathbf{\Sigma}_{11}^{-1}\mathbf{C}_{11})\\
&= (\mathbf{X}_1 + \breve{\mathbf{B}}_{22}^{-1}\breve{\mathbf{B}}_{21}\mathbf{C}_{11})'\breve{\mathbf{B}}_{22}(\mathbf{X}_1 + \breve{\mathbf{B}}_{22}^{-1}\breve{\mathbf{B}}_{21}\mathbf{C}_{11})\\
&= \mathbf{X}_1'\breve{\mathbf{B}}_{22}\mathbf{X}_1 + \mathbf{X}_1'\breve{\mathbf{B}}_{21}\mathbf{C}_{11} + \mathbf{C}_{11}\breve{\mathbf{B}}_{21}'\mathbf{X}_1 + \mathbf{C}_{11}\breve{\mathbf{B}}_{21}'\breve{\mathbf{B}}_{22}^{-1}\breve{\mathbf{B}}_{21}\mathbf{C}_{11}
\end{aligned}$$

with $\breve{\mathbf{B}}_{21}'\breve{\mathbf{B}}_{22}^{-1}\breve{\mathbf{B}}_{21} = \breve{\mathbf{B}}_{11} - \mathbf{\Sigma}_{11}^{-1}$. Hence,

$$\begin{aligned}
\mathbf{Q}_1(\mathbf{X}_1, \mathbf{X}_2) + \mathbf{Q}_2(\mathbf{X}_1) &= \mathbf{X}_2'\tilde{\mathbf{B}}_{22}\mathbf{X}_2 + 2\mathbf{X}_2'\tilde{\mathbf{B}}_{21}\begin{pmatrix}\mathbf{C}_{11}\\ \mathbf{X}_1\end{pmatrix}\\
&\quad + \mathbf{C}_{11}(\mathbf{B}_{11} - \mathbf{\Sigma}_{11}^{-1})\mathbf{C}_{11} + \mathbf{X}_1'\mathbf{B}_{21}^*\mathbf{C}_{11} + \mathbf{C}_{11}\mathbf{B}_{21}^{*\prime}\mathbf{X}_1 + \mathbf{X}_1'\mathbf{B}_{22}^*\mathbf{X}_1\\
&= \begin{pmatrix}\mathbf{X}_1\\ \mathbf{X}_2\end{pmatrix}'\mathbf{B}_{22}\begin{pmatrix}\mathbf{X}_1\\ \mathbf{X}_2\end{pmatrix} + 2\begin{pmatrix}\mathbf{X}_1\\ \mathbf{X}_2\end{pmatrix}'\mathbf{B}_{21}\mathbf{C}_{11} + \mathbf{C}_{11}\mathbf{B}_{21}'\mathbf{B}_{22}^{-1}\mathbf{B}_{21}\mathbf{C}_{11}\\
&= \left(\begin{pmatrix}\mathbf{X}_1\\ \mathbf{X}_2\end{pmatrix} + \mathbf{B}_{22}^{-1}\mathbf{B}_{21}\mathbf{C}_{11}\right)'\mathbf{B}_{22}\left(\begin{pmatrix}\mathbf{X}_1\\ \mathbf{X}_2\end{pmatrix} + \mathbf{B}_{22}^{-1}\mathbf{B}_{21}\mathbf{C}_{11}\right),
\end{aligned}$$

where we used the fact that $\mathbf{B}_{11} - \mathbf{\Sigma}_{11}^{-1} = \mathbf{B}_{21}'\mathbf{B}_{22}^{-1}\mathbf{B}_{21}$. The rest of the proof follows from the fact that $\mathbf{B}_{22}^{-1}\mathbf{B}_{21} = -\mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1}$ and $\mathbf{B}_{22} = \mathbf{\Sigma}_{22\cdot 1}^{-1}$.

In the next theorem, we establish a useful result about the distribution of the product $\mathbf{A}_{21}\mathbf{A}_{11}^{-1}$. It appears that it is a generalization of matrix-elliptical t-distribution. It is rather surprising, as Lemma 1 proves $\mathbf{A}_{21}$ to be conditionally normal and $\mathbf{A}_{11}$ to be Wishart distributed (and not its square root).

Theorem 1. *Let $k > n$ and $\mathbf{A} \sim W_k(n, \mathbf{\Sigma})$, where $\mathbf{A}$ and $\mathbf{\Sigma}$ are partitioned as in (4). Then the density of $\mathbf{A}_{21}\mathbf{A}_{11}^{-1}$ is given by*

$$f_{\mathbf{A}_{21}\mathbf{A}_{11}^{-1}}(\mathbf{X}) = \frac{|\mathbf{\Sigma}_{11}|^{\frac{k-m}{2}}\Gamma_m(\frac{n+k-m}{2})}{|\mathbf{\Sigma}_{22\cdot 1}|^{\frac{m}{2}}\pi^{\frac{(k-m)m}{2}}\Gamma_m(\frac{n}{2})}$$
$$\times |\mathbf{I} + \mathbf{\Sigma}_{11}(\mathbf{X} - \mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1})'\mathbf{\Sigma}_{22\cdot 1}^{-1}(\mathbf{X} - \mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1})|^{-\frac{1}{2}(n+k-m)}. \tag{7}$$

Proof. From Lemma 1, it follows that

$$\mathbf{A}_{21}\mathbf{A}_{11}^{-1}|\mathbf{A}_{11} \sim \mathcal{N}(\mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1}, \mathbf{\Sigma}_{22\cdot 1} \otimes \mathbf{A}_{11}^{-1})$$

and $\mathbf{A}_{11} \sim W_m(n, \mathbf{\Sigma}_{11})$. Hence,

$$\begin{aligned} f_{\mathbf{A}_{21}\mathbf{A}_{11}^{-1}}(\mathbf{X}) &= \int_{\mathbf{C}_{11}>0} f_{\mathbf{A}_{21}\mathbf{A}_{11}^{-1}|\mathbf{A}_{11}=\mathbf{C}_{11}}(\mathbf{X}|\mathbf{C}_{11}) f_{\mathbf{A}_{11}}(\mathbf{C}_{11}) \mathrm{d}\mathbf{C}_{11} \\ &= \int_{\mathbf{C}_{11}>0} \frac{\exp\left(-\frac{1}{2}\mathrm{tr}(\mathbf{\Sigma}_{22\cdot 1}^{-1}(\mathbf{X}-\mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1})\mathbf{C}_{11}(\mathbf{X}-\mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1})')\right)}{(2\pi)^{\frac{(k-m)m}{2}}|\mathbf{\Sigma}_{22\cdot 1}|^{\frac{m}{2}}|\mathbf{C}_{11}^{-1}|^{\frac{k-m}{2}}} \\ &\quad \times \frac{\exp\left(-\frac{1}{2}\mathrm{tr}(\mathbf{\Sigma}_{11}^{-1}\mathbf{C}_{11})\right)|\mathbf{C}_{11}|^{\frac{1}{2}(n-m-1)}}{2^{\frac{nm}{2}}\Gamma_m(\frac{n}{2})|\mathbf{\Sigma}_{11}|^{\frac{n}{2}}} \mathrm{d}\mathbf{C}_{11}. \end{aligned}$$

Rewriting the last integral, we obtain

$$\begin{aligned} f_{\mathbf{A}_{21}\mathbf{A}_{11}^{-1}}(\mathbf{X}) &= \frac{|\mathbf{\Sigma}_{22\cdot 1}|^{-\frac{m}{2}}|\mathbf{\Sigma}_{11}|^{-\frac{n}{2}}}{(2\pi)^{\frac{(k-m)m}{2}}2^{\frac{nm}{2}}\Gamma_m(\frac{n}{2})} \int_{\mathbf{C}_{11}>0} |\mathbf{C}_{11}|^{\frac{1}{2}(n+k-2m-1)} \\ &\quad \times \exp\left(-\frac{1}{2}\mathrm{tr}((\mathbf{\Sigma}_{11}^{-1} + (\mathbf{X}-\mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1})'\mathbf{\Sigma}_{22\cdot 1}^{-1}(\mathbf{X}-\mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1}))\mathbf{C}_{11})\right) \mathrm{d}\mathbf{C}_{11} \\ &= \frac{|\mathbf{\Sigma}_{22\cdot 1}|^{-\frac{m}{2}}|\mathbf{\Sigma}_{11}|^{-\frac{n}{2}}\Gamma_m(\frac{n+k-m}{2})}{\pi^{\frac{(k-m)m}{2}}\Gamma_m(\frac{n}{2})} \\ &\quad \times |\mathbf{\Sigma}_{11}^{-1} + (\mathbf{X}-\mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1})'\mathbf{\Sigma}_{22\cdot 1}^{-1}(\mathbf{X}-\mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1})|^{-\frac{1}{2}(n+k-m)}, \end{aligned}$$

where the computation of the matrix integral is based on the fact that under the integral we have the density function of an m-dimensional inverse Wishart distribution with $n+k-m$ degrees of freedom and parameter matrix $(\mathbf{\Sigma}_{11}^{-1} + (\mathbf{X}-\mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1})'\mathbf{\Sigma}_{22\cdot 1}^{-1}(\mathbf{X}-\mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1}))^{-1}$.

Using this theorem, we may establish further properties of the normalized product $\mathbf{A}_{21}\mathbf{A}_{11}^{-1}$. Note that in the regular case when $n > k$, or in the case of an arbitrary partitioning (4), we have an additional independency compared to the case $m = n < k$.

Corollary 1. *Let* $\mathbf{A} \sim W_k(n, \mathbf{\Sigma})$.

(a) *If* $k > n$, *then*

$$\mathbf{\Sigma}_{22\cdot 1}^{-1/2}(\mathbf{A}_{21}\mathbf{A}_{11}^{-1} - \mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1})\mathbf{A}_{11}^{1/2} \sim \mathcal{N}_{k-m\times m}(\mathbf{0}_{k-m\times m}, \mathbf{I}_{(k-m)m})$$

and is independent of $\mathbf{A}_{11}$ *and* $\mathbf{A}_{22\cdot 1}^* = \mathbf{A}_{22}^* - \mathbf{A}_{21}^*\mathbf{A}_{11}^{-1}\mathbf{A}_{12}^*$.

(b) *If* $m = n$ *then*

$$\mathbf{\Sigma}_{22\cdot 1}^{-1/2}(\mathbf{A}_{21}\mathbf{A}_{11}^{-1} - \mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1})\mathbf{A}_{11}^{1/2} \sim \mathcal{N}_{k-n\times n}(\mathbf{0}_{k-n\times n}, \mathbf{I}_{(k-n)n})$$

and is independent of $\mathbf{A}_{11}$.

(c) *If* $n > k$, *then*

$$\mathbf{\Sigma}_{22\cdot 1}^{-1/2}(\mathbf{A}_{21}\mathbf{A}_{11}^{-1} - \mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1})\mathbf{A}_{11}^{1/2} \sim \mathcal{N}_{k-m\times m}(\mathbf{0}_{k-m\times m}, \mathbf{I}_{(k-m)m})$$

and is independent of $\mathbf{A}_{11}$ *and* $\mathbf{A}_{22\cdot 1}$.

From Lemma 1(b), it follows that $\mathbf{A}_{21}\mathbf{A}_{11}^{-1}|\mathbf{A}_{11} \sim \mathcal{N}(\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}, \boldsymbol{\Sigma}_{22\cdot 1} \otimes \mathbf{A}_{11}^{-1})$. The next corollary provides an inverse statement, i.e. the distribution of $\mathbf{A}_{11}$ conditional upon the product $\mathbf{A}_{21}\mathbf{A}_{11}^{-1}$.

Corollary 2. *Let* $\mathbf{A} \sim W_k(n, \boldsymbol{\Sigma})$ *and* $k > n$. *Then it holds that*

$$\mathbf{A}_{11}|(\mathbf{A}_{21}\mathbf{A}_{11}^{-1} = \mathbf{X}) \sim W_m(n+k-m, (\boldsymbol{\Sigma}_{11}^{-1} + (\mathbf{X} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1})'\boldsymbol{\Sigma}_{22\cdot 1}^{-1}(\mathbf{X} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}))^{-1}).$$

3. Non-central singular Wishart distribution

In this section, we assume that $\mathbf{X} = (\mathbf{X}_1, \ldots, \mathbf{X}_n)$ with $k > n$ and $\mathbf{X}_i$ being independently and identically normally distributed with non-zero mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. Then the distribution of the random matrix $\mathbf{A} = \mathbf{X}\mathbf{X}'$, i.e. the joint distribution of $\tilde{\mathbf{A}}_{11}$ and $\tilde{\mathbf{A}}_{21}$, is called the non-central singular Wishart distribution and denoted by $W_k(n, \boldsymbol{\Sigma}, \boldsymbol{\Omega})$. Matrix $\boldsymbol{\Omega}$ is a noncentrality matrix and is defined as $\boldsymbol{\Omega} = \boldsymbol{\Sigma}^{-1}\mathbf{1}_n\boldsymbol{\mu}'\boldsymbol{\mu}\mathbf{1}_n'$, where $\mathbf{1}_n$ denotes the vector of length n with all elements equal to one. The density function of $\mathbf{A}$ is given by (cf. Srivastava [19], Corrollary 3.2):

$$\frac{\pi^{n(n-k)/2}2^{-kn/2}}{\Gamma_n(\frac{n}{2})|\boldsymbol{\Sigma}|^{n/2}}|\tilde{\mathbf{A}}_{11}|^{\frac{n-k-1}{2}} \exp\left(-\frac{1}{2}\mathrm{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{A})\right) {}_0F_1\left(\frac{n}{2}; \frac{1}{4}\boldsymbol{\Omega}\boldsymbol{\Sigma}^{-1}\mathbf{A}\right), \tag{8}$$

where $\tilde{\mathbf{A}}_{22} = \tilde{\mathbf{A}}_{21}\tilde{\mathbf{A}}_{11}^{-1}\tilde{\mathbf{A}}_{12}$. ${}_0F_1(.;.)$ denotes the hypergeometric function of the matrix argument (see Muirhead [1], p. 258 for the definition and properties). Unfortunately it is impossible to provide an equivalent to Muirheads [1] Theorem 3.2.10 due to the presence of the hypergeometric function outlined above. However, if $\boldsymbol{\mu} = \mathbf{0}$ then the distribution clearly reduces to the central case. The next theorem provides an explicit expression for the distribution of $\mathbf{A}_{21}\mathbf{A}_{11}^{-1}$ in the non-central case—however, only when $m = n$.

Theorem 2. *Let* $\mathbf{A} \sim \mathcal{W}_k(n, \boldsymbol{\Sigma}, \boldsymbol{\Omega})$ *and* $k > n$. $\mathbf{A}$ *is assumed to be partitioned as in* (2). *Then the density of* $\tilde{\mathbf{A}}_{21}\tilde{\mathbf{A}}_{11}^{-1}$ *is given by*

$$\begin{aligned} f_{\tilde{\mathbf{A}}_{21}\tilde{\mathbf{A}}_{11}^{-1}}(\mathbf{X}) &= \frac{|\boldsymbol{\Sigma}_{22\cdot 1}|^{-\frac{n}{2}}|\boldsymbol{\Sigma}_{11}|^{\frac{k-n}{2}}\Gamma_n(\frac{k}{2})}{\pi^{\frac{(k-n)n}{2}}\Gamma_n(\frac{n}{2})}|\mathbf{I} + \boldsymbol{\Sigma}_{11}(\mathbf{X} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1})'\boldsymbol{\Sigma}_{22\cdot 1}^{-1}(\mathbf{X} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1})|^{-\frac{k}{2}} \\ &\times {}_1F_1\left(\frac{k}{2}, \frac{n}{2}; \frac{1}{2}[\mathbf{I}_n\mathbf{X}]\boldsymbol{\Omega}\boldsymbol{\Sigma}^{-1}[\mathbf{I}_n\mathbf{X}]'(\boldsymbol{\Sigma}_{11}^{-1} + (\mathbf{X} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1})'\boldsymbol{\Sigma}_{22\cdot 1}^{-1}(\mathbf{X} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}))^{-1}\right). \end{aligned}$$

Proof. It holds that

$$\mathrm{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{A}) = \mathrm{tr}((\boldsymbol{\Sigma}_{11}^{-1} + (\tilde{\mathbf{A}}_{21}\tilde{\mathbf{A}}_{11}^{-1} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1})'\boldsymbol{\Sigma}_{22\cdot 1}^{-1}(\tilde{\mathbf{A}}_{21}\tilde{\mathbf{A}}_{11}^{-1} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}))\tilde{\mathbf{A}}_{11}).$$

Since $\mathbf{A}$ consists only of two functionally independent components $\tilde{\mathbf{A}}_{11}$ and $\tilde{\mathbf{A}}_{21}$, we can write that

$$\mathbf{A} = \left[\mathbf{I}_n(\tilde{\mathbf{A}}_{21}\tilde{\mathbf{A}}_{11}^{-1})'\right]'\tilde{\mathbf{A}}_{11}\left[\mathbf{I}_n(\tilde{\mathbf{A}}_{21}\tilde{\mathbf{A}}_{11}^{-1})'\right].$$

Then the joint density of $\tilde{\mathbf{A}}_{11}$ and $\tilde{\mathbf{A}}_{21}$ is given by

$$f_{\tilde{\mathbf{A}}_{11},\tilde{\mathbf{A}}_{21}}(\mathbf{C}_{11},\mathbf{C}_{21}) = \frac{\pi^{n(n-k)/2}2^{-kn/2}}{\Gamma_n(\frac{n}{2})|\boldsymbol{\Sigma}|^{n/2}}|\mathbf{C}_{11}|^{\frac{n-k-1}{2}}$$
$$\times {}_0F_1\left(\frac{n}{2}; \frac{1}{4}\boldsymbol{\Omega}\boldsymbol{\Sigma}^{-1}\left[\mathbf{I}_n(\mathbf{C}_{21}\mathbf{C}_{11}^{-1})'\right]'\mathbf{C}_{11}\left[\mathbf{I}_n(\mathbf{C}_{21}\mathbf{C}_{11}^{-1})'\right]\right)$$
$$\times \operatorname{etr}\left(-\frac{1}{2}(\boldsymbol{\Sigma}_{11}^{-1} + (\mathbf{C}_{21}\mathbf{C}_{11}^{-1} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1})'\boldsymbol{\Sigma}_{22\cdot 1}^{-1}(\mathbf{C}_{21}\mathbf{C}_{11}^{-1} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}))\mathbf{C}_{11}\right).$$

Making the transformation

$$\mathbf{X} = \mathbf{C}_{21}\mathbf{C}_{11}^{-1} \quad \text{and} \quad \mathbf{Y} = \mathbf{C}_{11}$$

with the Jacobian $|\mathbf{C}_{11}|^{k-n}$ yields

$$f_{\tilde{\mathbf{A}}_{11},\tilde{\mathbf{A}}_{21}\tilde{\mathbf{A}}_{11}^{-1}}(\mathbf{Y},\mathbf{X}) = \frac{\pi^{n(n-k)/2}2^{-kn/2}}{\Gamma_n(\frac{n}{2})|\boldsymbol{\Sigma}|^{n/2}}|\mathbf{Y}|^{\frac{k-n-1}{2}}\,{}_0F_1\left(\frac{n}{2}; \frac{1}{4}\boldsymbol{\Omega}\boldsymbol{\Sigma}^{-1}[\mathbf{I}_n\mathbf{X}']'\mathbf{Y}[\mathbf{I}_n\mathbf{X}']\right)$$
$$\times \operatorname{etr}\left(-\frac{1}{2}(\boldsymbol{\Sigma}_{11}^{-1} + (\mathbf{X} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1})'\boldsymbol{\Sigma}_{22\cdot 1}^{-1}(\mathbf{X} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}))\mathbf{Y}\right).$$

Integrating out $\mathbf{Y}$ leads to

$$f_{\tilde{\mathbf{A}}_{21}\tilde{\mathbf{A}}_{11}^{-1}}(\mathbf{X}) = \frac{\pi^{n(n-k)/2}2^{-kn/2}}{\Gamma_n(\frac{n}{2})|\boldsymbol{\Sigma}|^{n/2}}\int_{\mathbf{Y}>0}|\mathbf{Y}|^{\frac{k-n-1}{2}}\,{}_0F_1\left(\frac{n}{2}; \frac{1}{4}\boldsymbol{\Omega}\boldsymbol{\Sigma}^{-1}[\mathbf{I}_n\mathbf{X}']'\mathbf{Y}[\mathbf{I}_n\mathbf{X}']\right)$$
$$\times \operatorname{etr}\left(-\frac{1}{2}(\boldsymbol{\Sigma}_{11}^{-1} + (\mathbf{X} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1})'\boldsymbol{\Sigma}_{22\cdot 1}^{-1}(\mathbf{X} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}))\mathbf{Y}\right).$$

Using the definition of the hypergeometric function and the fact that the non-negative eigenvalues of the matrices

$$\boldsymbol{\Omega}\boldsymbol{\Sigma}^{-1}[\mathbf{I}_n\mathbf{X}']'\mathbf{Y}[\mathbf{I}_n\mathbf{X}'] \quad \text{and} \quad [\mathbf{I}_n\mathbf{X}']\boldsymbol{\Omega}\boldsymbol{\Sigma}^{-1}[\mathbf{I}_n\mathbf{X}']'\mathbf{Y}$$

coincide, we get

$$f_{\tilde{\mathbf{A}}_{21}\tilde{\mathbf{A}}_{11}^{-1}}(\mathbf{X}) = \frac{\pi^{n(n-k)/2}2^{-kn/2}}{\Gamma_n(\frac{n}{2})|\boldsymbol{\Sigma}|^{n/2}}\int_{\mathbf{Y}>0}|\mathbf{Y}|^{\frac{k-n-1}{2}}\,{}_0F_1\left(\frac{n}{2}; \frac{1}{4}[\mathbf{I}_n\mathbf{X}']\boldsymbol{\Omega}\boldsymbol{\Sigma}^{-1}[\mathbf{I}_n\mathbf{X}']'\mathbf{Y}\right)$$
$$\times \operatorname{etr}\left(-\frac{1}{2}(\boldsymbol{\Sigma}_{11}^{-1} + (\mathbf{X} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1})'\boldsymbol{\Sigma}_{22\cdot 1}^{-1}(\mathbf{X} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}))\mathbf{Y}\right).$$

The last integral is evaluated using Theorem 7.3.4 of Muirhead [1], and leads to

$$f_{\tilde{\mathbf{A}}_{21}\tilde{\mathbf{A}}_{11}^{-1}}(\mathbf{X})$$
$$= \frac{\pi^{n(n-k)/2}2^{-kn/2}}{\Gamma_n(\frac{n}{2})|\boldsymbol{\Sigma}|^{n/2}}2^{nk/2}\Gamma_n\left(\frac{k}{2}\right)|\boldsymbol{\Sigma}_{11}^{-1} + (\mathbf{X} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1})'\boldsymbol{\Sigma}_{22\cdot 1}^{-1}(\mathbf{X} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1})|^{-\frac{k}{2}}$$
$$\times\ {}_1F_1\left(\frac{k}{2}, \frac{n}{2}; \frac{1}{2}[\mathbf{I}_n\mathbf{X}']\boldsymbol{\Omega}\boldsymbol{\Sigma}^{-1}[\mathbf{I}_n\mathbf{X}']'(\boldsymbol{\Sigma}_{11}^{-1}\right.$$
$$\left. +\ (\mathbf{X} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1})'\boldsymbol{\Sigma}_{22\cdot 1}^{-1}(\mathbf{X} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}))^{-1}\right).$$

4. Inverse Wishart distribution

In this section, we consider the inverse Wishart distribution $\mathcal{W}_k^{-1}(n, \boldsymbol{\Sigma})$. The next theorem is analogous to Muirhead [1] Theorem 3.2.10 and to this paper's Theorem 1 for an inverse Wishart distributed random matrix.

Theorem 3. *Suppose* $\mathbf{A} \sim W_k^{-1}(n, \boldsymbol{\Sigma})$*, where* $\mathbf{A}$ *and* $\boldsymbol{\Sigma}$ *are partitioned as in* (1)*. Then*

(a) $\bar{\mathbf{A}}_{11\cdot 2} \sim W_m^{-1}(n-k+m, \boldsymbol{\Sigma}_{11\cdot 2})$ *and is independent of* $\bar{\mathbf{A}}_{22}$*;*
(b) $\bar{\mathbf{A}}_{12}|\bar{\mathbf{A}}_{22}, \bar{\mathbf{A}}_{11\cdot 2} \sim \mathcal{N}(\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\bar{\mathbf{A}}_{22}, \bar{\mathbf{A}}_{11\cdot 2} \otimes \bar{\mathbf{A}}_{22}\boldsymbol{\Sigma}_{22}^{-1}\bar{\mathbf{A}}_{22})$*;*
(c) $\bar{\mathbf{A}}_{22}$ *is* $W_{k-m}^{-1}(n-2m, \boldsymbol{\Sigma}_{22})$*;*
(d) $\bar{\mathbf{A}}_{12}\bar{\mathbf{A}}_{22}^{-1}$ *is independent of* $\bar{\mathbf{A}}_{22}$*, with the density given by*

$$f_{\bar{\mathbf{A}}_{12}\bar{\mathbf{A}}_{22}^{-1}}(\mathbf{X}) = \frac{|\boldsymbol{\Sigma}_{11\cdot 2}|^{-\frac{1}{2}(k-m)}|\boldsymbol{\Sigma}_{22}|^{\frac{1}{2}m}}{\pi^{\frac{(k-m)m}{2}}} \frac{\Gamma_m(\frac{n-m-1}{2})}{\Gamma_m(\frac{n-k-1}{2})}$$
$$\times |\mathbf{I} + \boldsymbol{\Sigma}_{11\cdot 2}^{-1}(\mathbf{X} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1})\boldsymbol{\Sigma}_{22}(\mathbf{X} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1})'|^{-\frac{1}{2}(n-m-1)};$$

(e) $\bar{\mathbf{A}}_{22}$ *is independent of* $\bar{\mathbf{A}}_{12}\bar{\mathbf{A}}_{22}^{-1}$ *and* $\bar{\mathbf{A}}_{11\cdot 2}$*;*
(f) $\bar{\mathbf{A}}_{11\cdot 2}|(\bar{\mathbf{A}}_{12}\bar{\mathbf{A}}_{22}^{-1} = \mathbf{X}) \sim W_m^{-1}(n, \boldsymbol{\Sigma}_{11\cdot 2} + (\mathbf{X} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1})\boldsymbol{\Sigma}_{22}(\mathbf{X} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1})')$.

Proof. Let

$$\mathbf{C} = \mathbf{A}^{-1} = \begin{matrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{21} & \mathbf{C}_{22} \end{matrix},$$

where $\mathbf{C}_{11}$ is $m \times m$. Then $\bar{\mathbf{A}}_{11\cdot 2} = \bar{\mathbf{A}}_{11} - \bar{\mathbf{A}}_{12}\bar{\mathbf{A}}_{22}^{-1}\bar{\mathbf{A}}_{21} = \mathbf{C}_{11}^{-1}$ and $\bar{\mathbf{A}}_{22} = \mathbf{C}_{22\cdot 1}^{-1} = (\mathbf{C}_{22} - \mathbf{C}_{12}\mathbf{C}_{11}^{-1}\mathbf{C}_{21})^{-1}$. The independence statement follows from Muirhead [1] Theorem 3.2.10 applied to the matrix $\mathbf{A}^{-1}$. From the same theorem and the relationship between the Wishart and inverse Wishart distributions, it follows that $\bar{\mathbf{A}}_{11\cdot 2} \sim W_m^{-1}(n-k+m, \boldsymbol{\Sigma}_{11\cdot 2})$ and $\bar{\mathbf{A}}_{22} \sim W_{k-m}^{-1}(n-2m, \boldsymbol{\Sigma}_{22})$. This proves parts (a) and (c).

To prove part (b) of the theorem, we recall the density of $\mathbf{A}$ given by (Muirhead [1], p. 113)

$$f(\mathbf{A}) = \frac{2^{-\frac{k(n-k-1)}{2}}}{\Gamma_k(\frac{n-k-1}{2})} |\mathbf{A}|^{-\frac{n}{2}} |\boldsymbol{\Sigma}|^{\frac{(n-k-1)}{2}} \operatorname{etr}\left(-\frac{1}{2}\mathbf{A}^{-1}\boldsymbol{\Sigma}\right. . \tag{9}$$

We make the transformation $\bar{\mathbf{A}}_{11\cdot 2} = \bar{\mathbf{A}}_{11} - \bar{\mathbf{A}}_{12}\bar{\mathbf{A}}_{22}^{-1}\bar{\mathbf{A}}_{21}$, $\mathbf{B}_{12} = \bar{\mathbf{A}}_{12}$, $\mathbf{B}_{22} = \bar{\mathbf{A}}_{22}$ with

$$\mathrm{d}\mathbf{A} = \mathrm{d}\bar{\mathbf{A}}_{11} \wedge \mathrm{d}\bar{\mathbf{A}}_{12} \wedge \mathrm{d}\bar{\mathbf{A}}_{22} = \mathrm{d}\bar{\mathbf{A}}_{11\cdot 2} \wedge \mathrm{d}\mathbf{B}_{12} \wedge \mathrm{d}\mathbf{B}_{22},$$
$$\det\mathbf{A} = \det\bar{\mathbf{A}}_{22}\det(\bar{\mathbf{A}}_{11} - \bar{\mathbf{A}}_{12}\bar{\mathbf{A}}_{22}^{-1}\bar{\mathbf{A}}_{21}) = \det\mathbf{B}_{22}\det\bar{\mathbf{A}}_{11\cdot 2},$$

and $\det\boldsymbol{\Sigma} = \det\boldsymbol{\Sigma}_{22}\det\boldsymbol{\Sigma}_{11\cdot 2}$, where $\wedge$ denotes the exterior product of two matrices (see Muirhead [1], Chapter 2).

From the proof of Muirhead [1] Theorem 3.2.10, it holds that

$$\operatorname{tr}(\mathbf{A}^{-1}\boldsymbol{\Sigma}) = \operatorname{tr} \begin{matrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{21} & \mathbf{C}_{22} \end{matrix} \begin{matrix} \boldsymbol{\Sigma}_{11\cdot 2} + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{matrix}$$
$$= \operatorname{tr}[\bar{\mathbf{A}}_{11\cdot 2}^{-1}(\boldsymbol{\Sigma}_{12} - \mathbf{B}_{12}\mathbf{B}_{22}^{-1}\boldsymbol{\Sigma}_{22})\boldsymbol{\Sigma}_{22}^{-1}(\boldsymbol{\Sigma}_{12} - \mathbf{B}_{12}\mathbf{B}_{22}^{-1}\boldsymbol{\Sigma}_{22})']$$
$$+ \operatorname{tr}(\mathbf{B}_{22}^{-1}\boldsymbol{\Sigma}_{22}) + \operatorname{tr}(\boldsymbol{\Sigma}_{11\cdot 2}\bar{\mathbf{A}}_{11\cdot 2}^{-1})$$

$$= \mathrm{tr}[\bar{\mathbf{A}}_{11\cdot 2}^{-1}(\mathbf{B}_{12} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\mathbf{B}_{22})\mathbf{B}_{22}^{-1}\boldsymbol{\Sigma}_{22}\mathbf{B}_{22}^{-1}(\mathbf{B}_{12} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\mathbf{B}_{22})'] + \mathrm{tr}(\mathbf{B}_{22}^{-1}\boldsymbol{\Sigma}_{22}) + \mathrm{tr}(\boldsymbol{\Sigma}_{11\cdot 2}\bar{\mathbf{A}}_{11\cdot 2}^{-1}),$$

where we used the equalities $\mathbf{C}_{11} = \bar{\mathbf{A}}_{11\cdot 2}^{-1}$, $\bar{\mathbf{A}}_{22}^{-1} = \mathbf{C}_{22} - \mathbf{C}_{12}\mathbf{C}_{11}^{-1}\mathbf{C}_{21}$ and $\mathbf{C}_{11}^{-1}\mathbf{C}_{12} = -\bar{\mathbf{A}}_{12}\bar{\mathbf{A}}_{22}^{-1}$. Hence, the joint density of $\mathbf{A}_{11\cdot 2}$, $\mathbf{B}_{12}$, $\mathbf{B}_{22}$ is given by

$$\frac{\exp\left(-\frac{1}{2}\mathrm{tr}(\bar{\mathbf{A}}_{11\cdot 2}^{-1}\boldsymbol{\Sigma}_{11\cdot 2})\right)|\boldsymbol{\Sigma}_{11\cdot 2}|^{\frac{1}{2}(n-k+m-m-1)}}{2^{\frac{m(n-k-1)}{2}}\Gamma_m(\frac{n-k-1}{2})|\bar{\mathbf{A}}_{11\cdot 2}|^{\frac{1}{2}(n-k+m)}}$$
$$\times \frac{\exp\left(-\frac{1}{2}\mathrm{tr}(\mathbf{B}_{22}^{-1}\boldsymbol{\Sigma}_{22})\right)|\boldsymbol{\Sigma}_{22}|^{\frac{1}{2}(n-2m-k+m-1)}}{2^{\frac{(k-m)(n-m-k-1)}{2}}\Gamma_{k-m}(\frac{n-m-k-1}{2})|\mathbf{B}_{22}|^{\frac{1}{2}(n-2m)}}$$
$$\times \frac{\exp\left(-\frac{1}{2}\bar{\mathbf{A}}_{11\cdot 2}^{-1}(\mathbf{B}_{12} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\mathbf{B}_{22})\mathbf{B}_{22}^{-1}\boldsymbol{\Sigma}_{22}\mathbf{B}_{22}^{-1}(\mathbf{B}_{12} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\mathbf{B}_{22})'\right)}{(2\pi)^{\frac{(k-m)m}{2}}|\bar{\mathbf{A}}_{11\cdot 2}|^{\frac{1}{2}(k-m)}|\mathbf{B}_{22}\boldsymbol{\Sigma}_{22}^{-1}\mathbf{B}_{22}|^{\frac{1}{2}m}}.$$

This completes the proof of part (b).

To prove (d), we show first that $\bar{\mathbf{A}}_{12}\bar{\mathbf{A}}_{22}^{-1}$ and $\bar{\mathbf{A}}_{22}$ are independently distributed. From Theorem 3(b) we obtain

$$\bar{\mathbf{A}}_{12}\bar{\mathbf{A}}_{22}^{-1}|\bar{\mathbf{A}}_{22}, \bar{\mathbf{A}}_{11\cdot 2} \sim \mathcal{N}(\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}, \bar{\mathbf{A}}_{11\cdot 2} \otimes \boldsymbol{\Sigma}_{22}^{-1}).$$

Because the conditional distribution $\bar{\mathbf{A}}_{12}\bar{\mathbf{A}}_{22}^{-1}$, given $\bar{\mathbf{A}}_{22}$ and $\bar{\mathbf{A}}_{11\cdot 2}$, does not depend on $\bar{\mathbf{A}}_{22}$, it follows that $\bar{\mathbf{A}}_{12}\bar{\mathbf{A}}_{22}^{-1}$ is independent from $\bar{\mathbf{A}}_{22}$ and $\bar{\mathbf{A}}_{12}\bar{\mathbf{A}}_{22}^{-1}|\bar{\mathbf{A}}_{22}, \bar{\mathbf{A}}_{11\cdot 2} = \bar{\mathbf{A}}_{12}\bar{\mathbf{A}}_{22}^{-1}|\bar{\mathbf{A}}_{11\cdot 2}$.

In order to obtain the density of $\bar{\mathbf{A}}_{12}\bar{\mathbf{A}}_{22}^{-1}$, we use the fact that $\bar{\mathbf{A}}_{12}\bar{\mathbf{A}}_{22}^{-1} = \mathbf{C}_{11}^{-1}\mathbf{C}_{12}$. Moreover, it holds that $\mathbf{C} \sim \mathcal{W}_k(n-k-1, \mathbf{B})$ where $\mathbf{B} = \boldsymbol{\Sigma}^{-1}$ with the partition $(\mathbf{B}_{ij})_{i,j=1,2}$ as in (1). From Theorem 1, we get

$$f_{\mathbf{C}_{11}^{-1}\mathbf{C}_{12}}(\mathbf{X}) = \frac{|\mathbf{B}_{11}|^{\frac{k-m}{2}}\Gamma_m(\frac{n-m-1}{2})}{|\mathbf{B}_{22\cdot 1}|^{\frac{m}{2}}\pi^{\frac{(k-m)m}{2}}\Gamma_m(\frac{n-k-1}{2})} \times |\mathbf{I} + \mathbf{B}_{11}(\mathbf{X} - \mathbf{B}_{11}^{-1}\mathbf{B}_{12})\mathbf{B}_{22\cdot 1}^{-1}(\mathbf{X} - \mathbf{B}_{11}^{-1}\mathbf{B}_{12})'|^{-\frac{1}{2}(n-m-1)}.$$

The rest of the proof follows from the fact that $\mathbf{B}_{11}^{-1}\mathbf{B}_{12} = \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}$, $\mathbf{B}_{22\cdot 1} = \boldsymbol{\Sigma}_{22}$, and $\mathbf{B}_{11} = \boldsymbol{\Sigma}_{11\cdot 2}^{-1}$.

For the proof of part (e), note that the proof of part (b) showed the joint density of $\bar{\mathbf{A}}_{22}$, $\bar{\mathbf{A}}_{11\cdot 2}$, and $\bar{\mathbf{A}}_{12}$ is given by

$$f_{\bar{\mathbf{A}}_{22},\bar{\mathbf{A}}_{11\cdot 2},\bar{\mathbf{A}}_{12}}(\mathbf{Y}_{22}, \mathbf{Y}_{11\cdot 2}, \mathbf{Y}_{12}) = \frac{\mathrm{etr}\left(-\frac{1}{2}\mathbf{Y}_{11\cdot 2}^{-1}\boldsymbol{\Sigma}_{11\cdot 2}\right)|\boldsymbol{\Sigma}_{11\cdot 2}|^{\frac{1}{2}(n-k-1)}}{2^{\frac{m(n-k-1)}{2}}\Gamma_m(\frac{n-k-1}{2})|\mathbf{Y}_{11\cdot 2}|^{\frac{1}{2}(n-k+m)}}$$
$$\times \frac{\mathrm{etr}\left(-\frac{1}{2}\mathbf{Y}_{22}^{-1}\boldsymbol{\Sigma}_{22}\right)|\boldsymbol{\Sigma}_{22}|^{\frac{1}{2}(n-m-k-1)}}{2^{\frac{(k-m)(n-m-k-1)}{2}}\Gamma_{k-m}(\frac{n-m-k-1}{2})|\mathbf{Y}_{22}|^{\frac{1}{2}(n-2m)}}$$
$$\times \frac{\mathrm{etr}\left(-\frac{1}{2}\mathbf{Y}_{11\cdot 2}^{-1}(\mathbf{Y}_{12} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\mathbf{Y}_{22})\mathbf{Y}_{22}^{-1}\boldsymbol{\Sigma}_{22}\mathbf{Y}_{22}^{-1}(\mathbf{Y}_{12} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\mathbf{Y}_{22})'\right)}{(2\pi)^{\frac{(k-m)m}{2}}|\mathbf{Y}_{11\cdot 2}|^{\frac{1}{2}(k-m)}|\mathbf{Y}_{22}\boldsymbol{\Sigma}_{22}^{-1}\mathbf{Y}_{22}|^{\frac{1}{2}m}}.$$

From the transformation $\mathbf{X}_{22} = \mathbf{Y}_{22}$, $\mathbf{X}_{11\cdot 2} = \mathbf{Y}_{11\cdot 2}$, and $\mathbf{X} = \mathbf{Y}_{12}\mathbf{Y}_{22}^{-1}$ with the Jacobian equal to $|\mathbf{Y}_{22}|^m$, we obtain

$$f_{\bar{\mathbf{A}}_{22},\bar{\mathbf{A}}_{11\cdot 2},\bar{\mathbf{A}}_{12}\bar{\mathbf{A}}_{22}^{-1}}(\mathbf{X}_{22}, \mathbf{X}_{11\cdot 2}, \mathbf{X}) = f_{\bar{\mathbf{A}}_{22}}(\mathbf{X}_{22}) \frac{\exp\left(-\frac{1}{2}\mathrm{tr}(\mathbf{X}_{11\cdot 2}^{-1}\mathbf{\Sigma}_{11\cdot 2})\right) |\mathbf{\Sigma}_{11\cdot 2}|^{\frac{1}{2}(n-k-1)}}{2^{\frac{m(n-k-1)}{2}}\Gamma_m(\frac{n-k-1}{2})|\mathbf{X}_{11\cdot 2}|^{\frac{1}{2}(n-k+m)}}$$
$$\times \frac{\exp\left(-\frac{1}{2}\mathbf{X}_{11\cdot 2}^{-1}(\mathbf{X} - \mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1})\mathbf{\Sigma}_{22}(\mathbf{X} - \mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1})'\right)}{(2\pi)^{\frac{(k-m)m}{2}}|\mathbf{X}_{11\cdot 2}|^{\frac{1}{2}(k-m)}|\mathbf{\Sigma}_{22}^{-1}|^{\frac{1}{2}m}},$$

Thus proving part (e). The proof of part (f) follows directly from (d).

Corollary 3. *Suppose* $\mathbf{A} \sim \mathcal{W}_k^{-1}(n, \mathbf{\Sigma})$. *Then*

$$\bar{\mathbf{A}}_{11\cdot 2}^{-1/2}(\bar{\mathbf{A}}_{12}\bar{\mathbf{A}}_{22}^{-1} - \mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1})\mathbf{\Sigma}_{22}^{1/2} \sim \mathcal{N}_{m\times(k-m)}(\mathbf{0}_{m\times(k-m)}, \mathbf{I}_{m(k-m)})$$

and is independent of $\bar{\mathbf{A}}_{22}$ *and* $\bar{\mathbf{A}}_{11\cdot 2}$.

5. Generalized inverse Wishart distribution

In this section, we extend the aforementioned inverse Wishart distribution to the case when $n < k$. It is usually referred to as a generalized inverse Wishart distribution. Let $A \sim \mathcal{W}_k(n, \mathbf{\Sigma})$ with $n < k$. Due to the singularity of $\mathbf{A}$, we cannot apply the same methods as in the case of the usual inverse Wishart distribution. Similar to Srivastava [19], we use the singular value decomposition of the matrix $\mathbf{A}$. Let

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix} = \begin{pmatrix} \mathbf{H}_{11}'\mathbf{D}\mathbf{H}_{11} & \mathbf{H}_{11}'\mathbf{D}\mathbf{H}_{12} \\ \mathbf{H}_{12}'\mathbf{D}\mathbf{H}_{11} & \mathbf{H}_{12}'\mathbf{D}\mathbf{H}_{12} \end{pmatrix} = \mathbf{H}_1'\mathbf{D}\mathbf{H}_1, \tag{10}$$

where $\mathbf{D} = \mathrm{diag}(d_1, \ldots, d_n)$ is a diagonal matrix of the positive eigenvalues of $\mathbf{A}$. The matrix $\mathbf{H}_1$ is partitioned into two submatrices $\mathbf{H}_1 = (\mathbf{H}_{11}\mathbf{H}_{12})$, where the $n \times n$ matrix $\mathbf{H}_{11}$ is nonsingular with probability one. Then $\mathbf{H}_1$ belongs to the Stiefel manifold $\mathcal{H}_{n,k}(n)$, i.e. to the set of $n \times k$ matrices such that $\mathbf{H}_1\mathbf{H}_1' = \mathbf{I}_n$. This class of matrices is also called the class of semiorthogonal matrices. Following Srivastava ([19], p. 1549) the density of the matrix $\mathbf{A}$ in terms of the matrices $\mathbf{D}$ and $\mathbf{H}_1$ is given by

$$f_{\mathbf{A}}(\mathbf{D}, \mathbf{H}_1) = K(n,k)|\mathbf{\Sigma}|^{-n/2}|\mathbf{D}|^{(k-n-1)/2}$$
$$\times \exp\left(-\frac{1}{2}\mathrm{tr}(\mathbf{\Sigma}^{-1}\mathbf{H}_1'\mathbf{D}\mathbf{H}_1)\right)\prod_{i<j}(d_i - d_j)g_{n,k}(\mathbf{H}_1)$$

with

$$K(n,k) = \frac{2^{-n}}{2^n(2\pi)^{kn/2}}2^n\pi^{n^2/2}\Gamma_n\left(\frac{n}{2}\right),$$
$$g_{n,k}(\mathbf{H}_1) = J(\mathbf{H}(d\mathbf{H}_1') \to d\mathbf{H}_1'),$$

where $\mathbf{H}' = (\mathbf{H}_1'\mathbf{H}_2')$ is an orthogonal matrix and the symbol $J(\mathbf{X} \to \mathbf{Y})$ denotes the Jacobian of the transformation from $\mathbf{X}$ to $\mathbf{Y}$.

The generalized inverse of the matrix $\mathbf{A}$ we denote by $\mathbf{A}^{(-)}$, and we define it by

$$\mathbf{A}^{(-)} = \mathbf{H}_1'\mathbf{D}^{-1}\mathbf{H}_1 = \begin{pmatrix} \mathbf{H}_{11}'\mathbf{D}^{-1}\mathbf{H}_{11} & \mathbf{H}_{11}'\mathbf{D}^{-1}\mathbf{H}_{12} \\ \mathbf{H}_{12}'\mathbf{D}^{-1}\mathbf{H}_{11} & \mathbf{H}_{12}'\mathbf{D}^{-1}\mathbf{H}_{12} \end{pmatrix}. \tag{11}$$

The density of $\mathbf{C} = \mathbf{A}^{(-)}$ is obtained in terms of $\tilde{\mathbf{D}}$ and $\mathbf{H}_1$ by making the nonsingular transformation $\tilde{\mathbf{D}} = \mathbf{D}^{-1}$ with the Jacobian $J(\mathbf{D} \to \tilde{\mathbf{D}}) = (-1)^n |\tilde{\mathbf{D}}|^{-2}$. Ignoring the sign, the density of $\mathbf{C}$ in terms of $\tilde{\mathbf{D}}$ and $\mathbf{H}_1$ is given by

$$\begin{aligned} f_{\mathbf{C}}(\tilde{\mathbf{D}}, \mathbf{H}_1) &= K(n,k)|\boldsymbol{\Sigma}|^{-n/2}|\tilde{\mathbf{D}}|^{-(k-n-1)/2}|\tilde{\mathbf{D}}|^{-2} \\ &\quad \times \exp\left(-\frac{1}{2}\mathrm{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{H}_1'\tilde{\mathbf{D}}^{-1}\mathbf{H}_1)\right)\prod_{i<j}\left(\frac{1}{\tilde{d}_i} - \frac{1}{\tilde{d}_j}\right) g_{n,k}(\mathbf{H}_1) \\ &= K(n,k)|\boldsymbol{\Sigma}|^{-n/2}|\tilde{\mathbf{D}}|^{-(k-n-1)/2}|\tilde{\mathbf{D}}|^{-2} \\ &\quad \times \exp\left(-\frac{1}{2}\mathrm{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{H}_1'\tilde{\mathbf{D}}^{-1}\mathbf{H}_1)\right)\prod_{i<j}\left(\frac{1}{\tilde{d}_i\tilde{d}_j}\right)\prod_{i<j}(\tilde{d}_i - \tilde{d}_j) g_{n,k}(\mathbf{H}_1). \end{aligned}$$

Because $\prod_{i<j}(\frac{1}{\tilde{d}_i\tilde{d}_j}) = |\tilde{\mathbf{D}}|^{-(n-1)}$, we get

$$\begin{aligned} f_{\mathbf{C}}(\tilde{\mathbf{D}}, \mathbf{H}_1) &= K(n,k)|\boldsymbol{\Sigma}|^{-n/2}|\tilde{\mathbf{D}}|^{-(k-n-1)/2}|\tilde{\mathbf{D}}|^{-(n+1)} \\ &\quad \times \exp\left(-\frac{1}{2}\mathrm{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{H}_1'\tilde{\mathbf{D}}^{-1}\mathbf{H}_1)\right)\prod_{i<j}(\tilde{d}_i - \tilde{d}_j) g_{n,k}(\mathbf{H}_1). \end{aligned}$$

Making the transformation $\tilde{\mathbf{D}}, \mathbf{H}_1 \to \mathbf{C}_{11}, \mathbf{C}_{21}$, we get the presentation of the density of $\mathbf{C}$ in terms of its functionally independent elements $\mathbf{C}_{11}$ and $\mathbf{C}_{21}$, i.e.

$$\begin{aligned} f_{\mathbf{C}}(\mathbf{C}_{11}, \mathbf{C}_{21}) &= 2^n K(n,k)|\boldsymbol{\Sigma}|^{-n/2}|\mathbf{C}_{11}|^{(n-1)/2} \\ &\quad \times |\mathbf{F}(\mathbf{C}_{11}, \mathbf{C}_{21})|^{-k/2} \exp\left(-\frac{1}{2}\mathrm{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{C}^{(-)})\right), \end{aligned}$$

where

$$\begin{aligned} \mathbf{C}^{(-)} &= \begin{bmatrix} \mathbf{H}_{11}'\tilde{\mathbf{D}}^{-1}\mathbf{H}_{11} & \mathbf{H}_{11}'\tilde{\mathbf{D}}^{-1}\mathbf{H}_{12} \\ \mathbf{H}_{12}'\tilde{\mathbf{D}}^{-1}\mathbf{H}_{11} & \mathbf{H}_{12}'\tilde{\mathbf{D}}^{-1}\mathbf{H}_{12}. \end{bmatrix} \\ &= \begin{pmatrix} \mathbf{C}_{11}\mathbf{F}(\mathbf{C}_{11}, \mathbf{C}_{21})^{-1}\mathbf{C}_{11} & \mathbf{C}_{11}\mathbf{F}(\mathbf{C}_{11}, \mathbf{C}_{21})^{-1}\mathbf{C}_{12} \\ \mathbf{C}_{21}\mathbf{F}(\mathbf{C}_{11}, \mathbf{C}_{21})^{-1}\mathbf{C}_{11} & \mathbf{C}_{21}\mathbf{F}(\mathbf{C}_{11}, \mathbf{C}_{21})^{-1}\mathbf{C}_{12} \end{pmatrix} \\ &= [\mathbf{C}_{11}\mathbf{C}_{12}]'\mathbf{F}(\mathbf{C}_{11}, \mathbf{C}_{21})^{-1}[\mathbf{C}_{11}\mathbf{C}_{12}]. \end{aligned}$$

Next, we calculate explicitly $\mathbf{F}(\mathbf{C}_{11}, \mathbf{C}_{21})$. We give the definition of the generalized inverse, i.e. $\mathbf{C} = \mathbf{C}\mathbf{C}^{(-)}\mathbf{C}$. Because the rank of the matrix $\mathbf{C}$ is equal to n, we get from Lemma 9.2.2 of Harville [20] that

$$\mathbf{C} = [\mathbf{C}_{11}\mathbf{C}_{12}]'\mathbf{C}_{11}^{-1}[\mathbf{C}_{11}\mathbf{C}_{12}].$$

Thus,

$$\mathbf{C}_{11}^{-1} = \mathbf{C}_{11}^{-1}[\mathbf{C}_{11}\mathbf{C}_{12}][\mathbf{C}_{11}\mathbf{C}_{12}]'\mathbf{F}(\mathbf{C}_{11}, \mathbf{C}_{21})^{-1}[\mathbf{C}_{11}\mathbf{C}_{12}][\mathbf{C}_{11}\mathbf{C}_{12}]'\mathbf{C}_{11}^{-1}.$$

Hence,

$$\mathbf{F}(\mathbf{C}_{11}, \mathbf{C}_{21}) = (\mathbf{C}_{11}^2 + \mathbf{C}_{12}\mathbf{C}_{21})\mathbf{C}_{11}^{-1}(\mathbf{C}_{11}^2 + \mathbf{C}_{12}\mathbf{C}_{21}).$$

Putting the above results together, we obtain the following theorem.

Theorem 4. *Let* $\mathbf{A} \sim \mathcal{W}_k(n, \mathbf{\Sigma})$*,* $k > n$*, where* $\mathbf{A}$ *and* $\mathbf{\Sigma}$ *are partitioned as in* (2)*. Then the density of the generalized inverse of the matrix* $\mathbf{A}$*, i.e.* $\mathbf{C} = \mathbf{A}^{(-)}$ *is given by*

$$f_{\mathbf{C}}(\mathbf{C}_{11}, \mathbf{C}_{21}) = \frac{\pi^{n(n-k)/2} 2^{-kn/2}}{\Gamma_n(\frac{n}{2})|\mathbf{\Sigma}|^{n/2}} |\mathbf{C}_{11}|^{(n-1)/2} \times |\mathbf{F}(\mathbf{C}_{11}, \mathbf{C}_{21})|^{-k/2} \exp\left(-\frac{1}{2}\mathrm{tr}(\mathbf{\Sigma}^{-1}\mathbf{C}^{(-)})\right), \tag{12}$$

where

$$\mathbf{F}(\mathbf{C}_{11}, \mathbf{C}_{21}) = (\mathbf{C}_{11}^2 + \mathbf{C}_{12}\mathbf{C}_{21})\mathbf{C}_{11}^{-1}(\mathbf{C}_{11}^2 + \mathbf{C}_{12}\mathbf{C}_{21}) \tag{13}$$

and

$$\mathbf{C}^{(-)} = [\mathbf{C}_{11}\mathbf{C}_{12}]'(\mathbf{C}_{11}^2 + \mathbf{C}_{12}\mathbf{C}_{21})^{-1}\mathbf{C}_{11}(\mathbf{C}_{11}^2 + \mathbf{C}_{12}\mathbf{C}_{21})^{-1}[\mathbf{C}_{11}\mathbf{C}_{12}]. \tag{14}$$

Similar to the other classes of Wishart distribution, here we also derive the density of the product $\mathbf{C}_{21}\mathbf{C}_{11}^{-1}$ in Theorem 5 and provide further distributional properties in Theorem 6.

Theorem 5. *Let* $k > n$ *and* $\mathbf{A} \sim \mathcal{W}_k(n, \mathbf{\Sigma})$*, where* $\mathbf{A}$ *and* $\mathbf{\Sigma}$ *are partitioned as in* (2)*. Then it holds that*

(a) *the density of* $\mathbf{C}_{21}\mathbf{C}_{11}^{-1}$ *is given by*

$$f_{\mathbf{C}_{21}\mathbf{C}_{11}^{-1}}(\mathbf{X}) = \frac{|\mathbf{\Sigma}_{22\cdot 1}|^{-\frac{n}{2}}|\mathbf{\Sigma}_{11}|^{\frac{k-n}{2}}\Gamma_n(\frac{k}{2})}{\pi^{\frac{(k-n)n}{2}}\Gamma_n(\frac{n}{2})} \times |\mathbf{I} + \mathbf{\Sigma}_{11}(\mathbf{X} - \mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1})'\mathbf{\Sigma}_{22\cdot 1}^{-1}(\mathbf{X} - \mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1})|^{-\frac{k}{2}};$$

(b) *The conditional distribution of* $\mathbf{C}_{11}$*, given* $\mathbf{C}_{21}\mathbf{C}_{11}^{-1} = \mathbf{X}$*, is the inverse Wishart distribution with* $n + k + 1$ *degrees of freedom and parameter matrix* $(\mathbf{I} + \mathbf{X}\mathbf{X}')^{-1}(\mathbf{\Sigma}_{11}^{-1} + (\mathbf{X} - \mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1})'\mathbf{\Sigma}_{22\cdot 1}^{-1}(\mathbf{X} - \mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1}))(\mathbf{I} + \mathbf{X}\mathbf{X}')^{-1}$*, i.e.* $\mathbf{C}_{11}|[\mathbf{C}_{21}\mathbf{C}_{11}^{-1} = \mathbf{X}] \sim \mathcal{W}_n^{-1}(n + k + 1, (\mathbf{I} + \mathbf{X}\mathbf{X}')^{-1}(\mathbf{\Sigma}_{11}^{-1} + (\mathbf{X} - \mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1})'\mathbf{\Sigma}_{22\cdot 1}^{-1}(\mathbf{X} - \mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1}))(\mathbf{I} + \mathbf{X}\mathbf{X}')^{-1})$.

Proof. (a) The statement follows from the fact that

$$\mathbf{C}_{21}\mathbf{C}_{11}^{-1} = \mathbf{H}_{12}'\mathbf{D}^{-1}\mathbf{H}_{11}(\mathbf{H}_{11}'\mathbf{D}^{-1}\mathbf{H}_{11})^{-1} = \mathbf{H}_{12}'\mathbf{H}_{11}^{-1} = \mathbf{H}_{12}'\mathbf{D}\mathbf{H}_{11}(\mathbf{H}_{11}'\mathbf{D}\mathbf{H}_{11})^{-1} = \tilde{\mathbf{A}}_{21}\tilde{\mathbf{A}}_{11}^{-1}$$

and Theorem 1 with $m = n$.

(b) From

$$\mathbf{F}(\mathbf{C}_{11}, \mathbf{C}_{21}) = \mathbf{C}_{11}(\mathbf{I} + \mathbf{C}_{11}^{-1}\mathbf{C}_{12}\mathbf{C}_{21}\mathbf{C}_{11}^{-1})\mathbf{C}_{11}(\mathbf{I} + \mathbf{C}_{11}^{-1}\mathbf{C}_{12}\mathbf{C}_{21}\mathbf{C}_{11}^{-1}),$$
$$\mathbf{C}^{(-)} = [\mathbf{I}\mathbf{C}_{11}^{-1}\mathbf{C}_{12}]'(\mathbf{I} + \mathbf{C}_{11}^{-1}\mathbf{C}_{12}\mathbf{C}_{21}\mathbf{C}_{11}^{-1})^{-1}\mathbf{C}_{11}^{-1}(\mathbf{I} + \mathbf{C}_{11}^{-1}\mathbf{C}_{12}\mathbf{C}_{21}\mathbf{C}_{11}^{-1})^{-1}[\mathbf{I}\mathbf{C}_{11}^{-1}\mathbf{C}_{12}],$$
$$|\mathbf{F}(\mathbf{C}_{11}, \mathbf{C}_{21})| = |\mathbf{C}_{11}|^3|\mathbf{I} + \mathbf{C}_{11}^{-1}\mathbf{C}_{12}\mathbf{C}_{21}\mathbf{C}_{11}^{-1}|^2$$

and making the transformation $\mathbf{X} = \mathbf{C}_{21}\mathbf{C}_{11}^{-1}$ and $\mathbf{C}_{11} = \mathbf{C}_{11}$ with the Jacobian $|\mathbf{C}_{11}|^{k-n}$, we obtain

$$f_{\mathbf{C}_{11},\mathbf{C}_{21}\mathbf{C}_{11}^{-1}}(\mathbf{C}_{11},\mathbf{X}) = \frac{\pi^{n(n-k)/2}2^{-kn/2}}{\Gamma_n(\frac{n}{2})|\mathbf{\Sigma}|^{n/2}}|\mathbf{C}_{11}|^{(n-1)/2}|\mathbf{C}_{11}|^{k-n}|\mathbf{C}_{11}|^{-3k/2}$$
$$\times |\mathbf{I}+\mathbf{X}\mathbf{X}'|^{-k}\mathrm{etr}\left(-\frac{1}{2}\mathbf{\Sigma}^{-1}\mathbf{C}^{(-)}\right).$$

From the proof of Theorem 2, we get that

$$\mathrm{tr}(\mathbf{\Sigma}^{-1}\mathbf{C}^{(-)}) = \mathrm{tr}\Big((\mathbf{I}+\mathbf{X}\mathbf{X}')^{-1}(\mathbf{\Sigma}_{11}^{-1}+(\mathbf{X}-\mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1})'\mathbf{\Sigma}_{22\cdot 1}^{-1}(\mathbf{X}-\mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1}))$$
$$\times (\mathbf{I}+\mathbf{X}\mathbf{X}')^{-1}\mathbf{C}_{11}^{-1}\Big).$$

Application of Theorem 5(a) leads to

$$f_{\mathbf{C}_{11},\mathbf{C}_{21}\mathbf{C}_{11}^{-1}}(\mathbf{C}_{11},\mathbf{X}) = \frac{2^{-kn/2}}{\Gamma_n(\frac{k}{2})}|\mathbf{C}_{11}|^{-(k+n+1)/2}$$
$$\times |(\mathbf{I}+\mathbf{X}\mathbf{X}')^{-1}(\mathbf{\Sigma}_{11}^{-1}+(\mathbf{X}-\mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1})'\mathbf{\Sigma}_{22\cdot 1}^{-1}(\mathbf{X}-\mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1}))(\mathbf{I}+\mathbf{X}\mathbf{X}')^{-1}|^{k/2}$$
$$\times \mathrm{etr}\left(-\frac{1}{2}(\mathbf{I}+\mathbf{X}\mathbf{X}')^{-1}(\mathbf{\Sigma}_{11}^{-1}+(\mathbf{X}-\mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1})'\mathbf{\Sigma}_{22\cdot 1}^{-1}(\mathbf{X}-\mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1}))\right.$$
$$\left.\times (\mathbf{I}+\mathbf{X}\mathbf{X}')^{-1}\mathbf{C}_{11}^{-1}\right)$$
$$\times \frac{|\mathbf{\Sigma}_{22\cdot 1}|^{-\frac{n}{2}}|\mathbf{\Sigma}_{11}|^{\frac{k-n}{2}}\Gamma_n(\frac{k}{2})}{\pi^{\frac{(k-n)n}{2}}\Gamma_n(\frac{n}{2})}|\mathbf{I}+\mathbf{\Sigma}_{11}(\mathbf{X}-\mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1})'\mathbf{\Sigma}_{22\cdot 1}^{-1}(\mathbf{X}-\mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1})|^{-\frac{k}{2}}$$
$$= f_{\mathbf{C}_{11}|\mathbf{C}_{21}\mathbf{C}_{11}^{-1}}(\mathbf{C}_{11}|\mathbf{X})f_{\mathbf{C}_{21}\mathbf{C}_{11}^{-1}}(\mathbf{X}),$$

where $f_{\mathbf{C}_{11}|\mathbf{C}_{21}\mathbf{C}_{11}^{-1}}(\mathbf{C}_{11}|\mathbf{X})$ is the density of the inverse Wishart distribution with $n+k+1$ degrees of freedom and the parameter matrix $(\mathbf{I}+\mathbf{X}\mathbf{X}')^{-1}(\mathbf{\Sigma}_{11}^{-1}+(\mathbf{X}-\mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1})'\mathbf{\Sigma}_{22\cdot 1}^{-1}(\mathbf{X}-\mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1}))(\mathbf{I}+\mathbf{X}\mathbf{X}')^{-1}$. This completes the proof.

Theorem 6. *Let $k > n$ and $\mathbf{A} \sim \mathcal{W}_k(n, \mathbf{\Sigma})$, where $\mathbf{A}$ and $\mathbf{\Sigma}$ are partitioned as in (2). Then it holds that*

(a) $(\mathbf{I}+\mathbf{C}_{11}^{-1}\mathbf{C}_{12}\mathbf{C}_{21}\mathbf{C}_{11}^{-1})\mathbf{C}_{11}(\mathbf{I}+\mathbf{C}_{11}^{-1}\mathbf{C}_{12}\mathbf{C}_{21}\mathbf{C}_{11}^{-1}) \sim \mathcal{W}_n^{-1}(2n+1, \mathbf{\Sigma}_{11}^{-1})$.
(b) $\mathbf{C}_{21}\mathbf{C}_{11}^{-1}|[(\mathbf{I}+\mathbf{C}_{11}^{-1}\mathbf{C}_{12}\mathbf{C}_{21}\mathbf{C}_{11}^{-1})\mathbf{C}_{11}(\mathbf{I}+\mathbf{C}_{11}^{-1}\mathbf{C}_{12}\mathbf{C}_{21}\mathbf{C}_{11}^{-1}) = \mathbf{Y}] \sim \mathcal{N}(\mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1}, \mathbf{\Sigma}_{22\cdot 1}\otimes\mathbf{Y})$.
(c) $\mathbf{C}_{11}^{-1/2}(\mathbf{I}+\mathbf{C}_{11}^{-1}\mathbf{C}_{12}\mathbf{C}_{21}\mathbf{C}_{11}^{-1})^{-1}(\mathbf{C}_{11}^{-1}\mathbf{C}_{12}-\mathbf{\Sigma}_{11}^{-1}\mathbf{\Sigma}_{12})\mathbf{\Sigma}_{22\cdot 1}^{-1/2} \sim \mathcal{N}_{n\times k-n}(\mathbf{0}_{n\times k-n}, \mathbf{I}_{n(k-n)})$.

Proof. (a) From the singular value decomposition, we get that

$$\mathbf{C}_{11}^{-1}\mathbf{C}_{12} = \mathbf{H}_{11}^{-1}\mathbf{D}\mathbf{H}_{11}'^{-1}\mathbf{H}_{11}'\mathbf{D}^{-1}\mathbf{H}_{12} = \mathbf{H}_{11}^{-1}\mathbf{H}_{12}$$

and $\mathbf{H}_{12}\mathbf{H}_{12}' = \mathbf{I}-\mathbf{H}_{11}\mathbf{H}_{11}'$. Hence

$$(\mathbf{I}+\mathbf{C}_{11}^{-1}\mathbf{C}_{12}\mathbf{C}_{21}\mathbf{C}_{11}^{-1})\mathbf{C}_{11}(\mathbf{I}+\mathbf{C}_{11}^{-1}\mathbf{C}_{12}\mathbf{C}_{21}\mathbf{C}_{11}^{-1})$$
$$= (\mathbf{I}+\mathbf{H}_{11}^{-1}\mathbf{H}_{12}\mathbf{H}_{12}'\mathbf{H}_{11}'^{-1})\mathbf{H}_{11}'\mathbf{D}^{-1}\mathbf{H}_{11}(\mathbf{I}+\mathbf{H}_{11}^{-1}\mathbf{H}_{12}\mathbf{H}_{12}'\mathbf{H}_{11}'^{-1})$$
$$= \mathbf{H}_{11}^{-1}\mathbf{H}_{11}'^{-1}\mathbf{H}_{11}'\mathbf{D}^{-1}\mathbf{H}_{11}\mathbf{H}_{11}^{-1}\mathbf{H}_{11}'^{-1} = \mathbf{H}_{11}^{-1}\mathbf{D}^{-1}\mathbf{H}_{11}'^{-1} = \mathbf{A}_{11}^{-1}.$$

The rest of the proof follows from the fact that $\mathbf{A}_{11} \sim \mathcal{W}_n(n, \mathbf{\Sigma}_{11})$ (see Srivastava [19], Corollary 3.4, and Muirhead [1], p. 113).

(b) As a partial case of Theorem 1, we get that

$$\mathbf{A}_{21}\mathbf{A}_{11}^{-1}|\mathbf{A}_{11} \sim \mathcal{N}(\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}, \boldsymbol{\Sigma}_{22\cdot 1} \otimes \mathbf{A}_{11}).$$

The statement of Theorem 6(b) follows then from the fact that $\mathbf{A}_{21}\mathbf{A}_{11}^{-1} = \mathbf{C}_{21}\mathbf{C}_{11}^{-1}$ and the proof of part (a).

(c) This part is a straightforward consequence of part (b).

6. Application to portfolio theory

The mean-variance portfolio selection developed by Markowitz [8] is a classical theory in finance. Given that the asset returns are Gaussian, it shows the investor how to choose the optimal fractions of assets in the portfolio (portfolio weights) in terms of the distribution parameters. The estimated portfolio weights are obtained by replacing these parameters with their sample estimators. The properties of the estimated weights were intensively studied recently in Okhrin and Schmid [9], Kan and Smith [21], and Kan and Zhou [22]. Here we apply the results of Section 4 to obtain some further properties of the estimated weights. Note that derivation relies on the distribution of $M'\mathbf{A}M$ and $M'\mathbf{A}^{-1}M$. This is a straightforward result for the inverse Wishart distribution; however, the extension to the generalized singular Wishart distribution is difficult.

If the investor is extremely risk averse, (s)he invests in the global minimum variance portfolio with the portfolio weights given by

$$\mathbf{w}_{GMV} = \frac{\boldsymbol{\Sigma}^{-1}\mathbf{1}_k}{\mathbf{1}_k'\boldsymbol{\Sigma}^{-1}\mathbf{1}_k}, \tag{15}$$

where $\boldsymbol{\Sigma}$ is the covariance matrix of the asset returns. This portfolio plays an important role in current portfolio decision theory (see, e.g. Jagannathan and Ma [23]).

In practice $\boldsymbol{\Sigma}$ is unknown, and should be estimated using historical values of asset returns. Let $\mathbf{X}_1, \ldots, \mathbf{X}_n$ be a sample from a k-dimensional Gaussian vector. We estimate $\boldsymbol{\Sigma}$ by

$$\hat{\boldsymbol{\Sigma}} = \frac{1}{n-1}\sum_{t=1}^{n}(\mathbf{X}_t - \bar{\mathbf{X}})(\mathbf{X}_t - \bar{\mathbf{X}})' \quad \text{with } \bar{\mathbf{X}} = \frac{1}{n}\sum_{t=1}^{n}\mathbf{X}_t. \tag{16}$$

Following Muirhead ([1], Corollary 3.2.2), the random matrix $(n-1)\hat{\boldsymbol{\Sigma}}$ follows a k-dimensional Wishart distribution with $n-1$ degrees of freedom and the parameter matrix $\boldsymbol{\Sigma}$. The estimators of the global minimum variance portfolio weights are obtained by plugging the estimator (16) in (15) instead of the unknown matrix $\boldsymbol{\Sigma}$, i.e.

$$\mathbf{w}_{GMV} = \frac{\hat{\boldsymbol{\Sigma}}^{-1}\mathbf{1}_k}{\mathbf{1}_k'\hat{\boldsymbol{\Sigma}}^{-1}\mathbf{1}_k}. \tag{17}$$

Okhrin and Schmid [9] showed that the estimator (17) is an unbiased estimator for the global minimum variance portfolio weights and derived its distribution. We are interested in linear combinations of the weights (15). Let $\mathbf{L}$ be an $m \times k$ dimensional matrix of constants with $m < k$ such that the matrix $\tilde{\mathbf{L}} = (\mathbf{L}'\mathbf{1}_k)$ is of full rank $m+1$. The vector of the linear combinations of the global minimum variance portfolio weights and its estimator are given by

$$\mathbf{w}_L = \frac{\mathbf{L}\boldsymbol{\Sigma}^{-1}\mathbf{1}_k}{\mathbf{1}_k'\boldsymbol{\Sigma}^{-1}\mathbf{1}_k} \quad \text{and} \quad \hat{\mathbf{w}}_L = \frac{\mathbf{L}\hat{\boldsymbol{\Sigma}}^{-1}\mathbf{1}_k}{\mathbf{1}_k'\hat{\boldsymbol{\Sigma}}^{-1}\mathbf{1}_k}. \tag{18}$$

For the derivation of the distribution of $\hat{\mathbf{w}}_L$, we use the results of Section 4. Let

$$\mathbf{G} = \tilde{\mathbf{L}}\,\mathbf{\Sigma}^{-1}\,\tilde{\mathbf{L}}' = \begin{bmatrix} \mathbf{L}\mathbf{\Sigma}^{-1}\mathbf{L}' & \mathbf{L}\mathbf{\Sigma}^{-1}\mathbf{1}_k \\ \mathbf{1}_k'\mathbf{\Sigma}^{-1}\mathbf{L}' & \mathbf{1}_k'\mathbf{\Sigma}^{-1}\mathbf{1}_k \end{bmatrix} = \{\mathbf{G}_{ij}\}_{i,j=1,2}$$

with $\mathbf{G}_{22} = \mathbf{1}_k'\mathbf{\Sigma}^{-1}\mathbf{1}_k$, and

$$\hat{\mathbf{G}} = \tilde{\mathbf{L}}\,\hat{\mathbf{\Sigma}}^{-1}\,\tilde{\mathbf{L}}' = \begin{bmatrix} \mathbf{L}\hat{\mathbf{\Sigma}}^{-1}\mathbf{L}' & \mathbf{L}\hat{\mathbf{\Sigma}}^{-1}\mathbf{1}_k \\ \mathbf{1}_k'\hat{\mathbf{\Sigma}}^{-1}\mathbf{L}' & \mathbf{1}_k'\hat{\mathbf{\Sigma}}^{-1}\mathbf{1}_k \end{bmatrix} = \{\hat{\mathbf{G}}_{ij}\}_{i,j=1,2}.$$

Then it holds that $\mathbf{w}_L = \mathbf{G}_{12}/\mathbf{G}_{22}$ and $\hat{\mathbf{w}}_L = \hat{\mathbf{G}}_{12}/\hat{\mathbf{G}}_{22}$. From Muirhead [1] Theorem 3.2.11, we obtain that $(n-1)\hat{\mathbf{G}}^{-1} \sim \mathcal{W}_{m+1}(n-k+m, \mathbf{G}^{-1})$. Thus, $(n-1)^{-1}\hat{\mathbf{G}} \sim \mathcal{W}_{m+1}(n-k+2m+2, \mathbf{G})$. Applying Theorem 3, we get the density of the linear combination of weights in terms of the components of $\mathbf{G}$.

$$f_{\hat{\mathbf{w}}_L}(\mathbf{x}) = \frac{|\mathbf{G}_{11\cdot 2}|^{-\frac{1}{2}}|\mathbf{G}_{22}|^{\frac{1}{2}m}}{\pi^{\frac{m}{2}}}\frac{\Gamma_m(\frac{n-k+1+m}{2})}{\Gamma_m(\frac{n-k+m}{2})} \times |\mathbf{I} + \mathbf{G}_{11\cdot 2}^{-1}(\mathbf{x} - \mathbf{G}_{12}\mathbf{G}_{22}^{-1})\mathbf{G}_{22}(\mathbf{x} - \mathbf{G}_{12}\mathbf{G}_{22}^{-1})'|^{-\frac{1}{2}(n-k+1+m)}. \tag{19}$$

Let $\mathbf{R} = \mathbf{\Sigma}^{-1} - \mathbf{\Sigma}^{-1}\mathbf{1}_k\mathbf{1}_k'\mathbf{\Sigma}^{-1}/\mathbf{1}_k'\mathbf{\Sigma}^{-1}\mathbf{1}_k$. Using the facts that $\mathbf{G}_{11\cdot 2} = \mathbf{LRL}'$,

$$\frac{\Gamma_m(\frac{n-k+1+m}{2})}{\Gamma_m(\frac{n-k+m}{2})} = \frac{\Gamma(\frac{n-k+1+m}{2})}{\Gamma(\frac{n-k}{2})},$$

and

$$|\mathbf{I} + \mathbf{G}_{11\cdot 2}^{-1}(\mathbf{x} - \mathbf{G}_{12}\mathbf{G}_{22}^{-1})\mathbf{G}_{22}(\mathbf{x} - \mathbf{G}_{12}\mathbf{G}_{22}^{-1})'| = 1 + \mathbf{G}_{22}(\mathbf{x} - \mathbf{G}_{12}\mathbf{G}_{22}^{-1})'\mathbf{G}_{11\cdot 2}^{-1}(\mathbf{x} - \mathbf{G}_{12}\mathbf{G}_{22}^{-1})$$

it follows that

$$f_{\hat{\mathbf{w}}_L}(\mathbf{x}) = \frac{|\mathbf{LRL}'|^{-\frac{1}{2}}|\mathbf{1}_k'\mathbf{\Sigma}^{-1}\mathbf{1}_k|^{\frac{1}{2}m}}{\pi^{\frac{m}{2}}}\frac{\Gamma(\frac{n-k+1+m}{2})}{\Gamma(\frac{n-k}{2})} \times (1 + \mathbf{1}_k'\mathbf{\Sigma}^{-1}\mathbf{1}_k(\mathbf{x} - \mathbf{w}_L)(\mathbf{LRL}')^{-1}(\mathbf{x} - \mathbf{w}_L)')^{-\frac{1}{2}(n-k+1+m)}. \tag{20}$$

The last expression is the density of the m-dimensional t-distribution with $n-k+1$ degrees of freedom, the mean vector $\mathbf{w}_L$ and the covariance matrix $(n-k-1)^{-1}\mathbf{LRL}'/\mathbf{1}_k'\mathbf{\Sigma}^{-1}\mathbf{1}_k$. This result was previously established by other means in Bodnar and Schmid [10].

References

[1] R.J. Muirhead, Aspects of Multivariate Statistical Theory, Wiley, New York, 1982.
[2] C.G. Khatri, On the mutual independence of certain statistics, Ann. Math. Statist. 30 (1959) 1258–1262.
[3] I. Olkin, S.N. Roy, On multivariate distribution theory, Ann. Math. Statist. 25 (1954) 329–339.
[4] I. Olkin, H. Rubin, Multivariate beta distribution and independency properties of Wishart distribution, Ann. Math. Statist. 35 (1964) 261–269.
[5] H. Tsukuma, Y. Konno, On improved estimation of normal precision matrix and discriminant coefficients, J. Multivariate Anal. 97 (2006) 1477–1500.
[6] A. Zellner, An Introduction to Bayesian Inference in Econometrics, Wiley, 1971.

[7] A. Gelman, J.B. Carlin, H.S. Stern, D.B. Rubin, Bayesian Data Analysis, Chapman & Hall, 2003.
[8] H. Markowitz, Portfolio selection, J. Finance 7 (1952) 77–91.
[9] Y. Okhrin, W. Schmid, Distributional properties of portfolio weights, J. Econometrics 134 (2006) 235–256.
[10] T. Bodnar, W. Schmid, A test for the weights of the global minimum variance portfolio in an elliptical model, Metrika 67 (2008) 127–143.
[11] D. von Rosen, Moments for the inverted Wishart distribution, Scand. J. Statist. 15 (1988) 97–109.
[12] C.G. Khatri, A note on Mitra's paper A density free approach to the matrix variate beta distribution, Sankhyā Ser. A 32 (1970) 311–317.
[13] M.S. Srivastava, C.G. Khatri, An Introduction to Multivariate Statistics, North-Holland, New York, 1979.
[14] H. Uhlig, On singular Wishart and singular multivariate beta distributions, Ann. Statist. 22 (1994) 395–405.
[15] O. Ledoit, M. Wolf, A well-conditioned estimator for large-dimensional covariance matrices, J. Multivariate Anal. 88 (2004) 356–411.
[16] H. Massam, J. Wesełowski, The Matsumoto–Yor property and the structure of the Wishart distribution, J. Multivariate Anal. 97 (2006) 103–123.
[17] J.A. Díaz-García, R. Gutiérrez-Jáimez, Proof of the conjecture of H. Uhlig on the singular multivariate beta and the jacobian of a certain matrix transformation, Ann. Statist. 25 (1997) 2018–2023.
[18] J.A. Díaz-García, R. Gutiérrez-Jáimez, K.V. Mardia, Wishart and pseudo-Wishart distributions and some applications to shape theory, J. Multivariate Anal. 63 (1997) 73–87.
[19] M.S. Srivastava, Singular Wishart and multivariate beta distributions, Ann. Statist. 31 (2003) 1537–1560.
[20] A.H. Harville, Matrix Algebra from a Statistician Perspective, Springer, 1997.
[21] R. Kan, D.R. Smith, The distribution of the sample minimum-variance frontier, Management Sci. (2007) (in press).
[22] R. Kan, G. Zhou, Optimal portfolio choice with parameter uncertainty, J. Finan. Quant. Anal. 42 (2007) 621–656.
[23] R. Jagannathan, T. Ma, Risk reduction in large portfolio: why imposing wrong constraints helps, J. Finance 58 (2003) 1651–1683.