On the Distribution of Matrix Quadratic Forms

Martin Ohlson and Timo Koski

Linköping University Post Print

N.B.: When citing this work, cite the original article.

ON THE DISTRIBUTION OF MATRIX QUADRATIC FORMS

Martin Ohlson* and Timo Koski†

*Department of Mathematics,

Linköping University,

SE–581 83 Linköping, Sweden.

E-mail: martin.ohlson@liu.se.

†Department of Mathematics,

Royal Institute of Technology,

SE–100 44 Stockholm, Sweden.

Key Words: Quadratic form; Spectral decomposition; Eigenvalues; Singular matrix normal distribution; Non-central Wishart distribution.

ABSTRACT

A characterization of the distribution of the multivariate quadratic form given by $\mathbf{X}\mathbf{A}\mathbf{X}'$, where $\mathbf{X}$ is a $p \times n$ normally distributed matrix and $\mathbf{A}$ is an $n \times n$ symmetric real matrix, is presented. We show that the distribution of the quadratic form is the same as the distribution of a weighted sum of non-central Wishart distributed matrices. This is applied to derive the distribution of the sample covariance between the rows of $\mathbf{X}$ when the expectation is the same for every column and is estimated with the regular mean.

1. INTRODUCTION

Univariate and multivariate quadratic forms play an important role in the theory of statistical analysis, specially when we are dealing with sample variances and covariance matrices. The univariate quadratic form

$$q = \boldsymbol{x}'\mathbf{A}\boldsymbol{x},$$

where $\boldsymbol{x}$ has a multivariate normal distribution, $\boldsymbol{x} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ are commonly used in the univariate theory. Results of the distribution of q can be found in many books of statisti-

cal analysis, see for example Rao (1973); Muirhead (1982); Srivastava and Khatri (1979); Graybill (1976).

One characterization of the univariate quadratic q is that it has the same distribution as the weighted sum of independent non-central χ^2 variables, see for example Baldessari (1967); Tan (1977) for more details. This and the fact that the sum of independent non-central χ^2 variables is again non-central χ^2 distributed give several results on the distribution of the univariate quadratic form as special cases.

One of the first to discuss independence between quadratic forms was Cochran (1934). Many authors have generalized Cochran's theorem, see for example Chipman and Rao (1964); Styan (1970); Tan (1977) and the references therein.

Assume that $\mathbf{X}$ follows a matrix normal distribution with a separable covariance matrix, i.e., $\mathbf{X} \sim N_{p,n}(\mathbf{M}, \boldsymbol{\Sigma}, \boldsymbol{\Psi})$, where $\boldsymbol{\Sigma} : p \times p$ and $\boldsymbol{\Psi} : n \times n$ are the covariance matrices not necessary positive definite and $N_{p,n}(\bullet, \bullet, \bullet)$ stands for the matrix normal distribution. We are interested in a characterization for the distribution of the quadratic form

$$\mathbf{Q} = \mathbf{X}\mathbf{A}\mathbf{X}', \tag{1}$$

where $\mathbf{A} : n \times n$ is a symmetric and real matrix. Several authors have investigated the conditions under which the quadratic form $\mathbf{Q}$ has a Wishart distribution. Rao (1973) showed that $\mathbf{Q}$ is Wishart if and only if $\boldsymbol{l}'\mathbf{Q}\boldsymbol{l}$ is χ^2 distributed, for any fixed vector $\boldsymbol{l}$. Hence, the theory of univariate quadratic forms can be applied to the multivariate case.

Khatri (1962) extended Cochran's theorem to the multivariate case by discussing conditions for Wishartness and independence of second degree polynomials. Other have also generalized Cochran's theorem for the multivariate case, see for example Rao and Mitra (1971); Khatri (1980); Vaish and Chaganty (2004); Tian and Styan (2005); Hu (2008). More generally, Wong and Wang (1993); Mathew and Nordström (1997); Masaro and Wong (2003) discussed Wishartness for the quadratic form when the covariance matrix is non-separable, i.e., when the covariance matrix cannot be written as a Kronecker product.

Khatri (1966) derived the density for $\mathbf{Q}$ in the central case, i.e., when $\mathbf{M} = \mathbf{0}$. The density function involves the hypergeometric function of matrix argument and is cumbersome to

handle. The hypergeometric function can be expanded in terms of zonal polynomials which is slowly convergent and the expansion in terms of Laguerre polynomials may be preferable for computational purpose.

The probability density function given by Khatri (1966) is written as the product of a Wishart density function and a generalized hypergeometric function. This form is not always convenient for studying properties of $\mathbf{Q}$. For $\mathbf{M} = \mathbf{0}$ and $\mathbf{\Psi} = \mathbf{I}_n$, both Hayakawa (1966) and Shah (1970) derived the probability density function for $\mathbf{Q}$. Using the moment generating function, Shah (1970) expressed the density function of $\mathbf{Q}$ in terms of Laguerre polynomials with matrix argument. Hayakawa (1966) also showed that any quadratic form $\mathbf{Q}$ can be decomposed to a linear combination of independent central Wishart or pseudo Wishart matrices with coefficients equal to the eigenvalues of $\mathbf{A}$.

In the non-central case, when $\mathbf{X} \sim N_{p,n}(\mathbf{M}, \mathbf{\Sigma}, \mathbf{\Psi})$ and $\mathbf{M} \neq \mathbf{0}$, Gupta and Nagar (2000) derived the non-central density for $\mathbf{Q}$ in terms of generalized Hayakawa polynomials, which are expectation of certain zonal polynomials. Gupta and Nagar (2000) also computed the moment generating function which they used for proving Wishartness and independence of quadratic forms. In (Khatri, 1977) the Laplace transform was used to generalize the results of Shah (1970) to the non-central case. When $\mathbf{A} = \mathbf{I}$, Khatri (1977) also obtained a similar representation of the non-central Wishart density in terms of the generalized Laguerre polynomial with matrix argument.

In this paper a characterization of the distribution of the quadratic form $\mathbf{Q}$ when $\mathbf{X} \sim N_{p,n}(\mathbf{M}, \mathbf{\Sigma}, \mathbf{\Psi})$ is given. Instead of representing it in terms of a hypergeometric function of matrix argument and an expansion in zonal polynomials as in (Khatri, 1966) and (Hayakawa, 1966) we show that the distribution of $\mathbf{Q}$ coincide with the distribution of a weighted sum of non-central Wishart distributed matrices, similar as in the case when $\mathbf{M} = \mathbf{0}$ and $\mathbf{\Psi} = \mathbf{I}$ done by Hayakawa (1966). We also discuss the complex normal case and show that the same properties hold.

The organization of this paper is as follows. In Section 2 the main theorem is proved. The characteristic function for the multivariate quadratic form is derived and the distribution of

3

the quadratic form is characterized. Some properties of the quadratic form which can be proved through this new characterization of the distribution is shown. In Section 3 the complex case is discussed and in Section 4 an example where this new characterization of the distribution can be used is presented.

2. DISTRIBUTION OF MULTIVARIATE QUADRATIC FORMS

Assume that the matrix $\mathbf{X}$ follows matrix normal distribution with a separable covariance matrix, i.e., the covariance matrix can be written as $\boldsymbol{\Psi} \otimes \boldsymbol{\Sigma}$, where $\otimes$ is the Kronecker product between the matrices $\boldsymbol{\Psi} : n \times n$ and $\boldsymbol{\Sigma} : p \times p$. This is written $\mathbf{X} \sim N_{p,n}(\mathbf{M}, \boldsymbol{\Sigma}, \boldsymbol{\Psi})$ and is equivalent to

$$\text{vec}\, \mathbf{X} \sim N_{pn}(\text{vec}\, \mathbf{M}, \boldsymbol{\Psi} \otimes \boldsymbol{\Sigma}),$$

where vec $\cdot$ is the vectorization operator, see Kollo and von Rosen (2005) for more details. Consider the matrix quadratic form

$$\mathbf{Q} = \mathbf{X}\mathbf{A}\mathbf{X}'.$$

We will use the characteristic function of $\mathbf{Q}$ for a characterization of the distribution. Start with the following theorem for the independent column case.

Theorem 1 *Let $\mathbf{Y} \sim N_{p,n}(\mathbf{M}, \boldsymbol{\Sigma}, \mathbf{I}_n)$, $\boldsymbol{\Sigma} \geq \mathbf{0}$ and let $\mathbf{A} : n \times n$ be a symmetric real matrix. The characteristic function of $\mathbf{Q} = \mathbf{Y}\mathbf{A}\mathbf{Y}'$ is then*

$$\varphi_{\mathbf{Q}}(\mathbf{T}) = \prod_{j=1}^{r} |\mathbf{I} - i\lambda_j \boldsymbol{\Gamma}\boldsymbol{\Sigma}|^{-1/2} \text{etr}\left\{\frac{1}{2}i\lambda_j \boldsymbol{\Omega}_j (\mathbf{I} - i\lambda_j \boldsymbol{\Gamma}\boldsymbol{\Sigma})^{-1}\boldsymbol{\Gamma}\right\},$$

where $\mathbf{T} = (t_{ij})$, where $i, j = 1, \ldots, p$ and $\boldsymbol{\Gamma} = (\gamma_{ij}) = ((1 + \delta_{ij})\,t_{ij})$, $t_{ij} = t_{ji}$ and δ_{ij} is the Kronecker delta. The non-centrality parameters are $\boldsymbol{\Omega}_j = \boldsymbol{m}_j \boldsymbol{m}_j'$, where $\boldsymbol{m}_j = \mathbf{M}\boldsymbol{a}_j$. The vectors $\boldsymbol{a}_j$ and the value λ_j are the orthonormal eigenvectors and eigenvalues of $\mathbf{A}$, respectively.

Proof Since $\mathbf{A}$ is symmetric, there exist $\boldsymbol{\Delta} \in O(n)$ ($O(n)$ is the orthogonal group, $O(n) = \{\boldsymbol{\Delta} : n \times n | \boldsymbol{\Delta}'\boldsymbol{\Delta} = \mathbf{I}_n\}$) and $\mathbf{D}_\lambda = \text{diag}(\lambda_1, \ldots, \lambda_r, 0, \ldots, 0)$ with $r = \text{rank}(\mathbf{A})$ such that

$\mathbf{A} = \boldsymbol{\Delta} \mathbf{D}_\lambda \boldsymbol{\Delta}'$. We have then

$$\mathbf{Q} = \mathbf{Y}\mathbf{A}\mathbf{Y}' = \mathbf{Y}\boldsymbol{\Delta}\mathbf{D}_\lambda\boldsymbol{\Delta}'\mathbf{Y}' = \mathbf{Z}\mathbf{D}_\lambda\mathbf{Z}',$$

where

$$\mathbf{Z} = \mathbf{Y}\boldsymbol{\Delta} \sim N_{p,n}\left(\mathbf{M}\boldsymbol{\Delta}, \boldsymbol{\Sigma}, \boldsymbol{\Delta}'\boldsymbol{\Delta}\right),$$

which implies that

$$\mathbf{Z} \sim N_{p,n}\left(\mathbf{M}\boldsymbol{\Delta}, \boldsymbol{\Sigma}, \mathbf{I}\right). \tag{2}$$

We can now rewrite $\mathbf{Q}$ as

$$\mathbf{Q} = \mathbf{Z}\mathbf{D}_\lambda\mathbf{Z}' = \sum_{j=1}^{r} \lambda_j \boldsymbol{z}_j \boldsymbol{z}_j', \tag{3}$$

where $\mathbf{Z} = (\boldsymbol{z}_1, \ldots, \boldsymbol{z}_n)$. Furthermore, from (2) we have that the column vectors of $\mathbf{Z}$ are independently distributed as $\boldsymbol{z}_j \sim N_p(\boldsymbol{m}_j, \boldsymbol{\Sigma})$, $j = 1, \ldots, r$, where $\boldsymbol{m}_j = \mathbf{M}\boldsymbol{a}_j$ and $\boldsymbol{\Delta} = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n)$.

The characteristic function of $\mathbf{Q}$ is given by

$$\varphi_{\mathbf{Q}}\left(\mathbf{T}\right) = \mathrm{E}\left(\exp\left\{i\sum_{j \leqslant k}^{p} t_{jk}q_{kj}\right\}\right),$$

where $\mathbf{Q} = (q_{ij})$, for $i, j = 1, \ldots, p$. Using the matrices $\boldsymbol{\Gamma}$ and $\mathbf{Q}$ the characteristic function can be written as

$$\varphi_{\mathbf{Q}}(\mathbf{T}) = \mathrm{E}\left(\exp\left\{\frac{1}{2}i\mathrm{tr}\left\{\boldsymbol{\Gamma}\mathbf{Q}\right\}\right\}\right) = \mathrm{E}\left(\exp\left\{\frac{1}{2}i\mathrm{tr}\left\{\boldsymbol{\Gamma}\mathbf{Z}\mathbf{D}_\lambda\mathbf{Z}'\right\}\right\}\right)$$
$$= \prod_{j=1}^{r} \mathrm{E}\left(\exp\left\{\frac{1}{2}i\lambda_j\boldsymbol{z}_j'\boldsymbol{\Gamma}\boldsymbol{z}_j\right\}\right).$$

Since the matrix $\boldsymbol{\Sigma} \geq \mathbf{0}$ we have $\mathrm{rank}(\boldsymbol{\Sigma}) = l \leq p$. The variable $\boldsymbol{z}_j$ is singular or non-singular multivariate normal $\boldsymbol{z}_j \sim N_p(\boldsymbol{m}_j, \boldsymbol{\Sigma}|l)$ and we have $\boldsymbol{z}_j = \mathbf{L}\boldsymbol{s}_j + \boldsymbol{m}_j$, for some matrix $\mathbf{L} : p \times l$, $\mathbf{L}\mathbf{L}' = \boldsymbol{\Sigma}$ with $\mathrm{rank}(\mathbf{L}) = l$ and with $\boldsymbol{s}_j \sim N_l(\mathbf{0}, \mathbf{I})$. Then the characteristic function can be

written as

$$\varphi_{\mathbf{Q}}(\mathbf{T}) = \prod_{j=1}^{r} \mathrm{E}\left(\exp\left\{\frac{1}{2}i\lambda_j(\mathbf{L}s_j + m_j)'\mathbf{\Gamma}(\mathbf{L}s_j + m_j)\right\}\right)$$

$$= \prod_{j=1}^{r} \mathrm{E}\left(\exp\left\{\frac{1}{2}i\lambda_j(s_j'\mathbf{L}'\mathbf{\Gamma}\mathbf{L}s_j + 2m_j'\mathbf{\Gamma}\mathbf{L}s_j + m_j'\mathbf{\Gamma}m_j)\right\}\right).$$

Let $\mathbf{H} \in O(l)$, such that $\mathbf{HL}'\mathbf{\Gamma}\mathbf{LH}' = \mathrm{diag}(\eta_1, \ldots, \eta_l) = \mathbf{D}_\eta$, where $\eta_j \ j = 1, \ldots, l$, are the eigenvalues of $\mathbf{L}'\mathbf{\Gamma}\mathbf{L}$. Now, let

$$u_j = \mathbf{H}s_j \sim N_l(\mathbf{0}, \mathbf{I}).$$

Furthermore $u_j = (u_{j1}, \ldots, u_{jl})'$, so

$$\varphi_{\mathbf{Q}}(\mathbf{T}) = \prod_{j=1}^{r} \mathrm{E}\left(\exp\left\{\frac{1}{2}i\lambda_j\left(u_j'\mathbf{D}_\eta u_j + 2m_j'\mathbf{\Gamma}\mathbf{LH}'u_j + m_j'\mathbf{\Gamma}m_j\right)\right\}\right)$$

$$= \prod_{j=1}^{r} \exp\left\{\frac{1}{2}i\lambda_j m_j'\mathbf{\Gamma}m_j\right\} \mathrm{E}\left(\exp\left\{\frac{1}{2}i\lambda_j \sum_{k=1}^{l}(\eta_k u_{jk}^2 + 2\theta_{jk}u_{jk})\right\}\right)$$

$$= \prod_{j=1}^{r} \exp\left\{\frac{1}{2}i\lambda_j m_j'\mathbf{\Gamma}m_j\right\} \prod_{k=1}^{l} \mathrm{E}\left(\exp\left\{\frac{1}{2}i\lambda_j(\eta_k u_{jk}^2 + 2\theta_{jk}u_{jk})\right\}\right), \qquad (4)$$

where

$$\boldsymbol{\theta}_j' = m_j'\mathbf{\Gamma}\mathbf{LH}' = (\theta_{j1}, \ldots, \theta_{jl})'.$$

The expectations in (4) can easily be calculated using the fact that $u_{jk} \sim N(0,1)$; they are

$$\mathrm{E}\left(\exp\left\{\frac{1}{2}i\lambda_j(\eta_k u_{jk}^2 + 2\theta_{jk}u_{jk})\right\}\right)$$

$$= (1 - i\lambda_j\eta_k)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}\lambda_j^2\theta_{jk}^2(1 - i\lambda_j\eta_k)^{-1}\right\}, \quad k = 1, \ldots, l.$$

Hence, the final expression is

$$\varphi_{\mathbf{Q}}(\mathbf{T}) = \prod_{j=1}^{r}\left[\exp\left\{\frac{1}{2}i\lambda_j m_j'\mathbf{\Gamma}m_j\right\}\right.$$

$$\left. \times \prod_{k=1}^{l}(1 - i\lambda_j\eta_k)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}\lambda_j^2\theta_{jk}^2(1 - i\lambda_j\eta_k)^{-1}\right\}\right], \qquad (5)$$

6

which can be simplified through rewriting the two factors in last product in (5) as

$$\prod_{j=1}^{r}\prod_{k=1}^{l}(1-i\lambda_j\eta_k)^{-1/2} = \prod_{j=1}^{r}|\mathbf{I}-i\lambda_j\mathbf{D}_\eta|^{-1/2} = \prod_{j=1}^{r}|\mathbf{I}-i\lambda_j\mathbf{\Gamma\Sigma}|^{-1/2} \tag{6}$$

and

$$\prod_{j=1}^{r}\prod_{k=1}^{l}\exp\left\{-\frac{1}{2}\lambda_j^2\theta_{jk}^2(1-i\lambda_j\eta_k)^{-1}\right\}$$

$$= \prod_{j=1}^{r}\exp\left\{-\frac{1}{2}\lambda_j^2\boldsymbol{\theta}_j'(\mathbf{I}-i\lambda_j\mathbf{D}_\eta)^{-1}\boldsymbol{\theta}_j\right\}$$

$$= \prod_{j=1}^{r}\exp\left\{-\frac{1}{2}\lambda_j^2\boldsymbol{m}_j'\mathbf{\Gamma L}(\mathbf{I}-i\lambda_j\mathbf{L'\Gamma L})^{-1}\mathbf{L'\Gamma}\boldsymbol{m}_j\right\}. \tag{7}$$

Together with the constant term the second part becomes

$$\exp\left\{\frac{1}{2}i\lambda_j\boldsymbol{m}_j'\mathbf{\Gamma}\boldsymbol{m}_j\right\}\exp\left\{-\frac{1}{2}\lambda_j^2\boldsymbol{m}_j'\mathbf{\Gamma L}(\mathbf{I}-i\lambda_j\mathbf{L'\Gamma L})^{-1}\mathbf{L'\Gamma}\boldsymbol{m}_j\right\}$$

$$= \exp\left\{\frac{1}{2}\boldsymbol{m}_j'\left(i\lambda_j\mathbf{\Gamma}-\lambda_j^2\mathbf{\Gamma L}(\mathbf{I}-i\lambda_j\mathbf{L'\Gamma L})^{-1}\mathbf{L'\Gamma}\right)\boldsymbol{m}_j\right\}$$

$$= \text{etr}\left\{\frac{1}{2}i\lambda_j\mathbf{\Omega}_j(\mathbf{I}-i\lambda_j\mathbf{\Gamma\Sigma})^{-1}\mathbf{\Gamma}\right\}, \tag{8}$$

where $\mathbf{\Omega}_j = \boldsymbol{m}_j\boldsymbol{m}_j'$. Insertion of (6), (7) and Equation (8) in (5) results in the final expression for the characteristic function of $\mathbf{Q}$,

$$\varphi_{\mathbf{Q}}(\mathbf{T}) = \prod_{j=1}^{r}|\mathbf{I}-i\lambda_j\mathbf{\Gamma\Sigma}|^{-1/2}\text{etr}\left\{\frac{1}{2}i\lambda_j\mathbf{\Omega}_j(\mathbf{I}-i\lambda_j\mathbf{\Gamma\Sigma})^{-1}\mathbf{\Gamma}\right\}.$$

This complete the proof of the theorem. □

We are now ready to give a theorem for the distribution of $\mathbf{Q} = \mathbf{YAY}'$. Let $W_p(\bullet,\bullet,\bullet)$ stand for the non-central Wishart distribution. We have the following theorem.

Theorem 2 *Assume* $\mathbf{Y} \sim N_{p,n}(\mathbf{M},\mathbf{\Sigma},\mathbf{I})$, $\mathbf{\Sigma} \geq \mathbf{0}$ *and let* $\mathbf{Q}$ *be the quadratic form* $\mathbf{Q} = \mathbf{YAY}'$, *where* $\mathbf{A}: n \times n$ *is a symmetric real matrix of rank* r. *Then the distribution of* $\mathbf{Q}$ *is that of*

$$\mathbf{W} = \sum_{j=1}^{r}\lambda_j\mathbf{W}_j,$$

where λ_j are the nonzero eigenvalues of $\mathbf{A}$ and $\mathbf{W}_j$ are independent non-central Wishart, i.e.,

$$\mathbf{W}_j \sim W_p(1, \mathbf{\Sigma}, \boldsymbol{m}_j \boldsymbol{m}_j'),$$

where $\boldsymbol{m_j} = \mathbf{M}\boldsymbol{a}_j$ and $\boldsymbol{a}_j$ are the corresponding orthonormal eigenvectors, i.e., $\boldsymbol{a}_j'\boldsymbol{a}_j = 1$ for $j = 1, \ldots, r$. In case of singular covariance matrix $\mathbf{\Sigma}$ the non-central Wishart distributions are singular.

Proof The characteristic function of $\mathbf{W}_j$ is given in Muirhead (1982) page 444 as

$$\varphi_{\mathbf{W}_j}(\mathbf{T}) = |\mathbf{I} - i\mathbf{\Gamma}\mathbf{\Sigma}|^{-1/2}\mathrm{etr}\left(\frac{1}{2}i\mathbf{\Omega}_j\left(\mathbf{I} - i\mathbf{\Gamma}\mathbf{\Sigma}\right)^{-1}\mathbf{\Gamma}\right),$$

and the characteristic function of $\mathbf{W} = \sum_j \lambda_j \mathbf{W}_j$ is (using the fact that $\mathbf{W}_j$ are independent)

$$\varphi_{\mathbf{W}}(\mathbf{T}) = \prod_{j=1}^{r} \varphi_{\mathbf{W}_j}(\lambda_j \mathbf{T})$$

$$= \prod_{j=1}^{r} |\mathbf{I} - i\lambda_j\mathbf{\Gamma}\mathbf{\Sigma}|^{-1/2}\mathrm{etr}\left\{\frac{1}{2}i\lambda_j\mathbf{\Omega}_j(\mathbf{I} - i\lambda_j\mathbf{\Gamma}\mathbf{\Sigma})^{-1}\mathbf{\Gamma}\right\}.$$

Using Theorem 1 we conclude that $\varphi_{\mathbf{W}}(\mathbf{T}) = \varphi_{\mathbf{Q}}(\mathbf{T})$, i.e., the characteristic function of $\mathbf{W}$ and $\mathbf{Q}$ are equal. Hence, the distribution of $\mathbf{Q}$ is the same as of $\mathbf{W} = \sum_j \lambda_j \mathbf{W}_j$ (Durrett, 1996). $\square$

The matrix quadratic form $\mathbf{Q} = \mathbf{Y}\mathbf{A}\mathbf{Y}'$ has the same distribution as $\mathbf{W} = \sum_j \lambda_j \mathbf{W}_j$, where λ_j are the nonzero eigenvalues of $\mathbf{A}$ and $\mathbf{W}_j$ are independent non-central Wishart, i.e., $\mathbf{W}_j \sim W_p(1, \mathbf{\Sigma}, \boldsymbol{m}_j \boldsymbol{m}_j')$, where $\boldsymbol{m_j} = \mathbf{M}\boldsymbol{a}_j$ and $\boldsymbol{a}_j$ are the corresponding eigenvectors. The fact that the two variables have the same distribution is here denoted by $\mathbf{Q} \overset{d}{=} \mathbf{W}$.

If $\mathbf{A}$ is idempotent the characterization above can be used to prove some properties. The Wishartness is just given here for the sake of completeness.

Corollary 3 Let $\mathbf{Y} \sim N_{p,n}(\mathbf{M}, \mathbf{\Sigma}, \mathbf{I})$, $\mathbf{\Sigma} \geq \mathbf{0}$ and let $\mathbf{A} : n \times n$ be a symmetric real matrix. If $\mathbf{A}$ is idempotent, then $\mathbf{Q} = \mathbf{Y}\mathbf{A}\mathbf{Y}' \sim W_p(r, \mathbf{\Sigma}, \mathbf{M}\mathbf{A}\mathbf{M}')$, where $r = \mathrm{rank}(\mathbf{A})$.

Proof If $\mathbf{A}$ is idempotent and $\text{rank}(\mathbf{A}) = r$, then $\lambda_j = 1$ for $j = 1, \ldots, r$ and zero otherwise. Using the fact that the sum of non-central Wishart distributed matrices again is non-central Wishart and Theorem 2, the proof follows. $\square$

Definition 1 *If* $\mathbf{Y} \sim N_{p,n}(\mathbf{M}, \mathbf{\Sigma}, \mathbf{I})$, $\mathbf{\Sigma} \geq \mathbf{0}$ *and if* $\mathbf{A} : n \times n$ *is a symmetric real matrix we define the distribution of the multivariate quadratic form* $\mathbf{Q} = \mathbf{Y}\mathbf{A}\mathbf{Y}'$ *to be* $Q_p(\mathbf{A}, \mathbf{M}, \mathbf{\Sigma})$.

If we transform a Wishart distributed matrix $\mathbf{W}$ as $\mathbf{B}\mathbf{W}\mathbf{B}'$, we have a new Wishart distributed matrix. In the same way we can transform our quadratic form.

Theorem 4 *Let* $\mathbf{Q} \sim Q_p(\mathbf{A}, \mathbf{M}, \mathbf{\Sigma})$ *and* $\mathbf{B} : q \times p$ *real matrix. Then*

$$\mathbf{B}\mathbf{Q}\mathbf{B}' \sim Q_q(\mathbf{A}, \mathbf{B}\mathbf{M}, \mathbf{B}\mathbf{\Sigma}\mathbf{B}').$$

Proof $\mathbf{Q} \stackrel{d}{=} \mathbf{W} = \sum_i \lambda_i \mathbf{W}_i$, where λ_i are the eigenvalues of $\mathbf{A}$ and $\mathbf{W}_i$ are independent non-central Wishart as $\mathbf{W}_i \sim W_p(1, \mathbf{\Sigma}, \mathbf{M}\boldsymbol{a}_i\boldsymbol{a}_i'\mathbf{M}')$, where $\boldsymbol{a}_i$ are the corresponding eigenvectors. Hence,

$$\mathbf{B}\mathbf{Q}\mathbf{B}' \stackrel{d}{=} \sum_i \lambda_i \mathbf{B}\mathbf{W}_i\mathbf{B}'$$

and since $\mathbf{B}\mathbf{W}_i\mathbf{B}' \sim W_q(n, \mathbf{B}\mathbf{\Sigma}\mathbf{B}', \mathbf{B}\mathbf{M}\boldsymbol{a}_i\boldsymbol{a}_i'\mathbf{M}'\mathbf{B}')$ we have

$$\mathbf{B}\mathbf{Q}\mathbf{B}' \sim Q_q(\mathbf{A}, \mathbf{B}\mathbf{M}, \mathbf{B}\mathbf{\Sigma}\mathbf{B}').$$

$\square$

Several other properties for the Wishart distribution, see for example Muirhead (1982) chapter 3.2 and 10.3, can be established for the distribution of the multivariate quadratic form.

Theorem 5 *If the matrices* $\mathbf{Q}_1, \ldots, \mathbf{Q}_r$ *are independently distributed as*

$$\mathbf{Q}_i \sim Q_p(\mathbf{A}_i, \mathbf{M}, \mathbf{\Sigma}), \quad i = 1, \ldots, r,$$

then $\sum_{i=1}^r \mathbf{Q}_i \sim Q_p(\mathbf{A}, \mathbf{M}, \mathbf{\Sigma})$, *where* $\mathbf{A} = \sum_{i=1}^r \mathbf{A}_i$.

Proof The proof follows directly from the definition. □

Theorem 6 *If* $\mathbf{Q}$ *is* $Q_p(\mathbf{A}, \mathbf{M}, \boldsymbol{\Sigma})$ *and* $\mathbf{Q}$, $\boldsymbol{\Sigma}$ *and* $\mathbf{M}$ *are partitioned as*

$$\mathbf{Q} = \begin{pmatrix} \mathbf{Q}_{11} & \mathbf{Q}_{12} \\ \mathbf{Q}_{21} & \mathbf{Q}_{22} \end{pmatrix}, \quad \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix} \quad \text{and} \quad \mathbf{M} = \begin{pmatrix} \mathbf{M}_1 \\ \mathbf{M}_2 \end{pmatrix},$$

where $\mathbf{Q}_{11}$ *and* $\boldsymbol{\Sigma}_{11}$ *are* $k \times k$ *and* $\mathbf{M}_1$ *is* $k \times n$, *then* $\mathbf{Q}_{11}$ *is* $Q_k(\mathbf{A}, \mathbf{M}_1, \boldsymbol{\Sigma}_{11})$.

Proof Put $\mathbf{B} = \begin{pmatrix} \mathbf{I}_k & : & \mathbf{0} \end{pmatrix} : k \times p$ in Theorem 4 and the result follows. □

Corollary 7 *Assume* $\mathbf{Q} \sim Q_p(\mathbf{A}, \mathbf{M}, \boldsymbol{\Sigma})$ *and* $\boldsymbol{\alpha}(\neq \mathbf{0})$ *is an* $p \times 1$ *fixed vector, then*

$$\frac{\boldsymbol{\alpha}'\mathbf{Q}\boldsymbol{\alpha}}{\boldsymbol{\alpha}'\boldsymbol{\Sigma}\boldsymbol{\alpha}} \stackrel{d}{=} \sum_{j=1}^{r} \lambda_j \chi_1^2(\delta_j),$$

where $\delta_j = \dfrac{\boldsymbol{\alpha}'\mathbf{M}\boldsymbol{a}_j\boldsymbol{a}_j'\mathbf{M}'\boldsymbol{\alpha}}{\boldsymbol{\alpha}'\boldsymbol{\Sigma}\boldsymbol{\alpha}}$, *and* λ_j *and* $\boldsymbol{a}_j$ *are the eigenvalues and orthonormal eigenvectors of* $\mathbf{A}$, *respectively.*

Proof Follows directly from Theorem 4. □

In the literature two types of multivariate beta distribution are discussed, see Kollo and von Rosen (2005) for more details. The multivariate beta distribution is closely connected to the multivariate normal distribution and Wishart distribution. We will also have a version of a sum of weighted multivariate beta distributions as follows.

Theorem 8 *Assume* $\mathbf{Q} \sim Q_p(\mathbf{A}, \mathbf{0}, \mathbf{I})$ *and* $\mathbf{W} \sim W_p(m, \mathbf{I})$ *independently distributed. Then*

$$\mathbf{W}^{-1/2}\mathbf{Q}\mathbf{W}^{-1/2} \stackrel{d}{=} \sum_{i=1}^{r} \lambda_i \mathbf{B}_i,$$

where λ_i, $i = 1, \ldots, r$ *are the eigenvalues of* $\mathbf{A}$ *and the* $\mathbf{B}_i$'s, *are multivariate beta distributed of type II with one degree of freedom, i.e.,* $\mathbf{B}_i \sim M\beta_{II}(p, m, 1)$.

See Kollo and von Rosen (2005) page 250 for the definition of multivariate beta distribution. Now, suppose that $\mathbf{X} \sim N_{p,n}(\mathbf{M}, \boldsymbol{\Sigma}, \boldsymbol{\Psi})$, $\boldsymbol{\Sigma} \geq \mathbf{0}$ and $\boldsymbol{\Psi} > \mathbf{0}$ i.e., the columns are dependent as well.

Corollary 9 *Let $\mathbf{X} \sim N_{p,n}(\mathbf{M}, \mathbf{\Sigma}, \mathbf{\Psi})$, $\mathbf{\Sigma} \geq \mathbf{0}$, $\mathbf{\Psi} > \mathbf{0}$ and let $\mathbf{A} : n \times n$ be a symmetric real matrix of rank r, then $\mathbf{Q} = \mathbf{X}\mathbf{A}\mathbf{X}'$ is distributed as*

$$\mathbf{W} = \sum_{j=1}^{r} \lambda_j \mathbf{W}_j,$$

where λ_j are the nonzero eigenvalues of $\mathbf{\Psi}^{1/2}\mathbf{A}\mathbf{\Psi}^{1/2}$ and $\mathbf{W}_j$ are independent non-central Wishart distributed as follows

$$\mathbf{W}_j \sim W_p(1, \mathbf{\Sigma}, \boldsymbol{m}_j \boldsymbol{m}'_j), \quad j = 1, \ldots, r,$$

where $\boldsymbol{m}_j = \mathbf{M}\mathbf{\Psi}^{-1/2}\boldsymbol{a}_j$ and $\boldsymbol{a}_j$ are the corresponding orthonormal eigenvectors, i.e., $\boldsymbol{a}'_j \boldsymbol{a}_j = 1$ for $j = 1, \ldots, r$. In the case of a singular covariance matrix $\mathbf{\Sigma}$ the non-central Wishart distributions are singular.

Hence, we see that the distribution of the matrix quadratic form $\mathbf{Q} = \mathbf{X}\mathbf{A}\mathbf{X}'$ is given by $Q_p(\mathbf{\Psi}^{1/2}\mathbf{A}\mathbf{\Psi}^{1/2}, \mathbf{M}\mathbf{\Psi}^{-1/2}, \mathbf{\Sigma})$.

Using Corollary 9 we can characterize the distribution of $\mathrm{tr}\mathbf{Q}$.

Theorem 10 *Assume $\mathbf{Q} \sim Q_p(\mathbf{A}, \mathbf{M}, \mathbf{\Sigma})$, then*

$$\mathrm{tr}\mathbf{Q} \stackrel{d}{=} \sum_{i=1}^{n} \sum_{j=1}^{r} \omega_i \lambda_j \chi_1^2(\delta_{ij}),$$

where $\delta_{ij} = \boldsymbol{\sigma}'_i \mathbf{\Sigma}^{-1/2}\mathbf{M}\boldsymbol{a}_j \boldsymbol{a}'_j \mathbf{M}'\mathbf{\Sigma}^{-1/2}\boldsymbol{\sigma}_i$, λ_j and $\boldsymbol{a}_j$ are the eigenvalues and eigenvectors of $\mathbf{A}$, ω_i and $\boldsymbol{\sigma}_i$ are the eigenvalues and eigenvectors of $\mathbf{\Sigma}$, respectively.

Proof From Equation (3) it follows that $\mathrm{tr}\mathbf{Q} = \sum_{j=1}^{r} \lambda_j \boldsymbol{z}'_j \boldsymbol{z}_j$, where the random vectors $\boldsymbol{z}'_j \sim N_{1,p}(\boldsymbol{m}'_j, 1, \mathbf{\Sigma})$, $\boldsymbol{m}'_j = (\mathbf{M}\boldsymbol{a}_j)'$ for $j = 1, \ldots, r$. Using Corollary 9 for the quadratic form $\boldsymbol{z}'_j \boldsymbol{z}_j$, i.e., $\mathbf{A} = \mathbf{I}$, the proof follows. $\qquad\square$

In the independent column case the Wishartness for a quadratic form was given in Corollary 3 and is well known in the literature. Here again for the sake of completeness the Wishartness is given for the dependent column case.

Corollary 11 *Let* $\mathbf{X} \sim N_{p,n}(\mathbf{M}, \boldsymbol{\Sigma}, \boldsymbol{\Psi})$, $\boldsymbol{\Sigma} \geq \mathbf{0}$, $\boldsymbol{\Psi} > \mathbf{0}$ *and let* $A : n \times n$ *be a symmetric real matrix. If* $\mathbf{A}\boldsymbol{\Psi}$ *is idempotent, then*

$$\mathbf{Q} = \mathbf{X}\mathbf{A}\mathbf{X}' \sim W_p(r, \boldsymbol{\Sigma}, \mathbf{M}\mathbf{A}\mathbf{M}'),$$

where $r = \mathrm{rank}(\mathbf{A})$.

Proof If $\mathbf{A}\boldsymbol{\Psi}$ is idempotent, then $\boldsymbol{\Psi}^{1/2}\mathbf{A}\boldsymbol{\Psi}^{1/2}$ is idempotent as well and $\lambda_j = 1$ for $j = 1, \ldots, r$ and zero otherwise. Using the fact that the sum of non-central Wishart distributed matrices again is non-central Wishart and Theorem 2 completes the proof. $\square$

We will also give a corollary for the case $\boldsymbol{\Psi} \geq \mathbf{0}$ and $\mathbf{M} = \mathbf{0}$.

Corollary 12 *Let* $\mathbf{X} \sim N_{p,n}(\mathbf{0}, \boldsymbol{\Sigma}, \boldsymbol{\Psi})$, $\boldsymbol{\Sigma} \geq \mathbf{0}$, $\boldsymbol{\Psi} \geq \mathbf{0}$ *and let* $\mathbf{A} : n \times n$ *be a symmetric real matrix of rank* r, *then* $\mathbf{Q} = \mathbf{X}\mathbf{A}\mathbf{X}'$ *is distributed as*

$$\mathbf{W} = \sum_{j=1}^{r_0} \lambda_j \mathbf{W}_j,$$

where $r_0 = \mathrm{rank}(\mathbf{L}'\mathbf{A}\mathbf{L})$, λ_j *are the nonzero eigenvalues of* $\mathbf{L}'\mathbf{A}\mathbf{L}$ *and the matrix* $\mathbf{L} : n \times l$ *is such that* $\boldsymbol{\Psi} = \mathbf{L}\mathbf{L}'$ *with* $\mathrm{rank}(\mathbf{L}) = l \leq n$. *Furthermore,* $\mathbf{W}_j$ *are independent Wishart as*

$$\mathbf{W}_j \sim W_p(1, \boldsymbol{\Sigma}).$$

In case of singular covariance matrix $\boldsymbol{\Sigma}$ *the Wishart distributions are singular.*

Using Corollary 12, we see that the distribution of the matrix quadratic form $\mathbf{Q} = \mathbf{X}\mathbf{A}\mathbf{X}'$, when $\mathbf{X} \sim N_{p,n}(\mathbf{0}, \boldsymbol{\Sigma}, \boldsymbol{\Psi})$, $\boldsymbol{\Sigma} \geq \mathbf{0}$, $\boldsymbol{\Psi} = \mathbf{L}\mathbf{L}' \geq \mathbf{0}$, is given by $Q_p(\mathbf{L}'\mathbf{A}\mathbf{L}, \mathbf{0}, \boldsymbol{\Sigma})$.

3. COMPLEX MATRIX QUADRATIC FORMS

In this section will we consider complex multivariate normal distributions denoted by $CN_{p,n}(\bullet, \bullet, \bullet)$. The complex matrix quadratic form is $\mathbf{Q} = \mathbf{Y}\mathbf{A}\mathbf{Y}^*$, where $\mathbf{Y}^*$ is the conjugate transpose of the matrix $\mathbf{Y}$. For more details about the complex multivariate normal and complex Wishart distributions, see for example Goodman (1963); Khatri (1965); Srivastava (1965). Let $\bar{t}$ be the complex conjugate of t, we have then the following theorem.

Theorem 13 *Let* $\mathbf{Y} \sim CN_{p,n}(\mathbf{M}, \boldsymbol{\Sigma}, \mathbf{I})$, *where* $\mathbf{M} = \mathbf{M}_1 + i\mathbf{M}_2$ *and* $\boldsymbol{\Sigma}$ *is positive definite and Hermitian. Let also* $\mathbf{A} : n \times n$ *be a Hermitian matrix. The characteristic function of* $\mathbf{Q} = \mathbf{Y}\mathbf{A}\mathbf{Y}^*$ *is then*

$$\varphi_{\mathbf{Q}}(\mathbf{T}) = \prod_{j=1}^{r} |\mathbf{I} - i\lambda_j \boldsymbol{\Gamma}\boldsymbol{\Sigma}|^{-1} \mathrm{etr} \left\{ \frac{1}{2} i\lambda_j \boldsymbol{\Omega}_j (\mathbf{I} - i\lambda_j \boldsymbol{\Gamma}\boldsymbol{\Sigma})^{-1} \boldsymbol{\Gamma} \right\},$$

where $\mathbf{T} = (t_{jk})$, $t_{jk} = t_{jk} + it_{jk}$ *for* $j, k = 1, \dots, p$ *and* $\boldsymbol{\Gamma} = (\gamma_{jk}) = ((1 + \delta_{jk}) t_{jk})$, $t_{jk} = \bar{t}_{kj}$ *and* δ_{jk} *is the Kronecker delta. The non-centrality parameters are* $\boldsymbol{\Omega}_j = \boldsymbol{m}_j \boldsymbol{m}_j^*$, *where* $\boldsymbol{m}_j = \mathbf{M}\boldsymbol{a}_j$. *The vectors* $\boldsymbol{a}_j$ *and the value* λ_j *are the orthonormal eigenvectors and eigenvalues of* $\mathbf{A}$ *respectively.*

Proof The proof is similar to the proof of Theorem 1. $\qquad\square$

Furthermore, we also have a characterization for the distribution of the complex matrix quadratic form, which is analogous to that obtained for the non-complex case.

Theorem 14 *Assume* $\mathbf{Y} \sim CN_{p,n}(\mathbf{M}, \boldsymbol{\Sigma}, \mathbf{I})$, *where* $\mathbf{M} = \mathbf{M}_1 + i\mathbf{M}_2$ *and* $\boldsymbol{\Sigma}$ *is positive definite and Hermitian. Let* $\mathbf{Q}$ *be the quadratic form* $\mathbf{Q} = \mathbf{Y}\mathbf{A}\mathbf{Y}^*$, *where* $\mathbf{A} : n \times n$ *is a Hermitian matrix of rank* r. *Then the distribution of* $\mathbf{Q}$ *is that of*

$$\mathbf{W} = \sum_{j=1}^{r} \lambda_j \mathbf{W}_j,$$

where λ_j *are the nonzero eigenvalues of* $\mathbf{A}$ *and* $\mathbf{W}_j$ *are independent non-central complex Wishart, i.e.,*

$$\mathbf{W}_j \sim CW_p(1, \boldsymbol{\Sigma}, \boldsymbol{m}_j \boldsymbol{m}_j^*),$$

where $\boldsymbol{m_j} = \mathbf{M}\boldsymbol{a}_j$ *and* $\boldsymbol{a}_j$ *are the corresponding orthonormal eigenvectors, i.e.,* $\boldsymbol{a}_j^* \boldsymbol{a}_j = 1$ *for* $j = 1, \dots, r$.

Proof The characteristic function of $\mathbf{W}_j$ is given by Goodman (1963) as

$$\varphi_{\mathbf{W}_j}(\mathbf{T}) = |\mathbf{I} - i\boldsymbol{\Gamma}\boldsymbol{\Sigma}|^{-1} \mathrm{etr} \left(\frac{1}{2} i\boldsymbol{\Omega}_j (\mathbf{I} - i\boldsymbol{\Gamma}\boldsymbol{\Sigma})^{-1} \boldsymbol{\Gamma} \right),$$

and the proof parallels that of Theorem 2. $\qquad\square$

4. A SPECIAL SAMPLE COVARIANCE MATRIX

The example discussed in this section is similar to the example considered in Vaish and Chaganty (2004).

Let $\mathbf{X} \sim N_{p,n}\left(\mathbf{M}, \boldsymbol{\Sigma}, \boldsymbol{\Psi}\right)$ where $\mathbf{M} = \boldsymbol{\mu}\mathbf{1}'$ and suppose that covariance matrix $\boldsymbol{\Psi}$ is known. We want to estimate the covariance $\boldsymbol{\Sigma}$. The matrix normal density function is given by

$$f(\mathbf{X}) = (2\pi)^{-\frac{1}{2}pn}\,|\boldsymbol{\Sigma}|^{-n/2}|\boldsymbol{\Psi}|^{-p/2}\mathrm{etr}\left\{-\frac{1}{2}\boldsymbol{\Sigma}^{-1}\left(\mathbf{X}-\mathbf{M}\right)\boldsymbol{\Psi}^{-1}\left(\mathbf{X}-\mathbf{M}\right)'\right\}$$

and the maximum likelihood estimators of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are given by

$$\widehat{\boldsymbol{\mu}}_{ml} = (\mathbf{1}'\boldsymbol{\Psi}^{-1}\mathbf{1})^{-1}\mathbf{X}\boldsymbol{\Psi}^{-1}\mathbf{1},$$

$$n\widehat{\boldsymbol{\Sigma}}_{ml} = \mathbf{X}\left(\boldsymbol{\Psi}^{-1} - \boldsymbol{\Psi}^{-1}\mathbf{1}\left(\mathbf{1}'\boldsymbol{\Psi}^{-1}\mathbf{1}\right)^{-1}\mathbf{1}'\boldsymbol{\Psi}^{-1}\right)\mathbf{X}' = \mathbf{X}\mathbf{H}\mathbf{X}',$$

where

$$\mathbf{H} = \boldsymbol{\Psi}^{-1} - \boldsymbol{\Psi}^{-1}\mathbf{1}\left(\mathbf{1}'\boldsymbol{\Psi}^{-1}\mathbf{1}\right)^{-1}\mathbf{1}'\boldsymbol{\Psi}^{-1}.$$

Since $\mathbf{H}\boldsymbol{\Psi}$ is idempotent and $\mathrm{rank}(\mathbf{H}\boldsymbol{\Psi}) = n - 1$ we have that $\mathbf{X}\mathbf{H}\mathbf{X}' \sim W_p\left(n - 1, \boldsymbol{\Sigma}\right)$, see Corollary 11.

Now for some reason we estimate the expectation $\boldsymbol{\mu}$ with the regular mean $\hat{\boldsymbol{\mu}} = \frac{1}{n}\mathbf{X}\mathbf{1} = \bar{\boldsymbol{x}}$, i.e., we use the same estimator as if $\boldsymbol{\Psi} = \mathbf{I}$. This can be done for several reasons. For example, the estimator $\hat{\boldsymbol{\mu}}$ is more robust than $\hat{\boldsymbol{\mu}}_{ml}$ for large number of observations, i.e., for large n. Another reason could be that we only know the centralized observations, $\mathbf{X} - \hat{\boldsymbol{\mu}}\mathbf{1}'$. However, when we estimate the covariance matrix $\boldsymbol{\Sigma}$, we use the dependent model with $\boldsymbol{\Psi}$. The estimator of $\boldsymbol{\Sigma}$ is then

$$n\widehat{\boldsymbol{\Sigma}} = \left(\mathbf{X} - \hat{\boldsymbol{\mu}}\mathbf{1}'\right)\boldsymbol{\Psi}^{-1}\left(\mathbf{X} - \hat{\boldsymbol{\mu}}\mathbf{1}'\right)' = \mathbf{X}\mathbf{C}\boldsymbol{\Psi}^{-1}\mathbf{C}\mathbf{X}',$$

where $\mathbf{C}$ is the centralization matrix

$$\mathbf{C} = \mathbf{I} - \mathbf{1}(\mathbf{1}'\mathbf{1})^{-1}\mathbf{1}'. \tag{9}$$

14

Using Corollary 9 we have that the distribution of $\mathbf{X}\mathbf{C}\boldsymbol{\Psi}^{-1}\mathbf{C}\mathbf{X}'$ is the same as the distribution of $\mathbf{W} = \mathbf{W}_1 + \tilde{\lambda}\widetilde{\mathbf{W}}$, where $\mathbf{W}_1$ and $\widetilde{\mathbf{W}}$ are independently distributed as

$$\mathbf{W}_1 \sim W_p\left(n-2, \boldsymbol{\Sigma}\right),$$
$$\widetilde{\mathbf{W}} \sim W_p\left(1, \boldsymbol{\Sigma}\right)$$

and $\tilde{\lambda} = 1 - \frac{1}{n}\mathbf{1}'\boldsymbol{\Psi}\mathbf{C}\boldsymbol{\Psi}^{-1}\mathbf{1}$.

The expectation of $\mathbf{X}\mathbf{C}\boldsymbol{\Psi}^{-1}\mathbf{C}\mathbf{X}'$ which can be computed straightforwardly is given by

$$\mathrm{E}\left(\mathbf{X}\mathbf{C}\boldsymbol{\Psi}^{-1}\mathbf{C}\mathbf{X}'\right) = \left(n-1-\frac{1}{n}\mathbf{1}'\boldsymbol{\Psi}\mathbf{C}\boldsymbol{\Psi}^{-1}\mathbf{1}\right)\boldsymbol{\Sigma}.$$

Hence, an unbiased estimator of $\boldsymbol{\Sigma}$ is

$$\widehat{\boldsymbol{\Sigma}} = \left(n-1-\frac{1}{n}\mathbf{1}'\boldsymbol{\Psi}\mathbf{C}\boldsymbol{\Psi}^{-1}\mathbf{1}\right)^{-1}\mathbf{X}\mathbf{C}\boldsymbol{\Psi}^{-1}\mathbf{C}\mathbf{X}'.$$

ACKNOWLEDGEMENT

BIBLIOGRAPHY

Baldessari, B. (1967). The distribution of a quadratic form of normal random variables. *The Annals of Mathematical Statistics*, 38(6):1700–1704.

Chipman, J. S. and Rao, M. M. (1964). Projections, generalized inverses, and quadratic forms. *Journal of Mathematical Analysis and Applications*, 9(1):1–11.

Cochran, W. G. (1934). The distribution of quadratic forms in a normal system, with applications to the analysis of covariance. *In Mathematical Proceedings of the Cambridge Philosophical Society*, 30(2):178–191.

Durrett, R. (1996). *Probability: Theory and Examples*. Duxbury Press, Belmont, California, 2nd edition.

Goodman, N. R. (1963). Statistical analysis based on a certain multivariate complex Gaussian distribution (an introduction). *The Annals of Mathematical Statistics*, 34(1):152–177.

Graybill, F. A. (1976). *Theory and Application of the Linear Model.* Duxbury Press, North Scituate, Massachusetts.

Gupta, A. K. and Nagar, D. K. (2000). *Matrix Variate Distributions.* Chapman and Hall, Boca Raton, Florida.

Hayakawa, T. (1966). On the distribution of a quadratic form in multivariate normal sample. *Annals of the Institute of Statistical Mathematics*, 18(1):191–201.

Hu, J. (2008). Wishartness and independence of matrix quadratic forms in a normal random matrix. *Journal of Multivariate Analysis*, 99(3):555–571.

Khatri, C. G. (1962). Conditions for Wishartness and independence of second degree polynomials in a normal vector. *The Annals of Mathematical Statistics*, 33(3):1002–1007.

Khatri, C. G. (1965). Classical statistical analysis based on a certain multivariate complex Gaussian distribution. *The Annals of Mathematical Statistics*, 38(1):98–114.

Khatri, C. G. (1966). On certain distribution problems based on positive definite quadratic functions in normal vectors. *The Annals of Mathematical Statistics*, 37(2):468–479.

Khatri, C. G. (1977). Distribution of a quadratic form in non-central normal vectors using generalised Laguerre polynomials. *South African Statistical Journal*, 11:167–179.

Khatri, C. G. (1980). Statistical inference for covariance matrices with linear structure. *In Handbook of Statistics (P. R. Krishnaiah, ed.). North-Holland, Amsterdam*, pages 443–469.

Kollo, T. and von Rosen, D. (2005). *Advanced Multivariate Statistics with Matrices.* Springer, Dordrecht.

Masaro, J. and Wong, C. S. (2003). Wishart distributions associated with matrix quadratic forms. *Journal of Multivariate Analysis*, 85(1):1–9.

Mathew, T. and Nordström, K. (1997). Wishart and Chi-square distributions associated with matrix quadratic forms. *Journal of Multivariate Analysis*, 61(1):129–143.

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

Rao, C. R. (1973). *Linear Statistical Inference and Its Applications*. John Wiley & Sons, New York, 2nd edition.

Rao, C. R. and Mitra, S. K. (1971). *Generalized Inverse of a Matrix and its Applications*. Jon Wiley & Sons, New York.

Shah, B. K. (1970). Distribution theory of a positive definite quadratic form with matrix argument. *The Annals of Mathematical Statistics*, 41(2):692–697.

Srivastava, M. S. (1965). On the complex Wishart distribution. *The Annals of Mathematical Statistics*, 36(1):313–315.

Srivastava, M. S. and Khatri, C. G. (1979). *An Introduction to Multivariate Statistics*. North Holland, New York.

Styan, G. P. H. (1970). Notes on the distribution of quadratic forms in singular normal variables. *Biometrika*, 57(3):567–572.

Tan, W. Y. (1977). On the distribution of quadratic forms in normal random variables. *The Canadian Journal of Statistics/La Revue Canadienne de Statistique*, 5(2):241–250.

Tian, Y. and Styan, G. P. H. (2005). Cochran's statistical theorem for outer inverses of matrices and matrix quadratic forms. *Linear and Multilinear Algebra*, 53(5):387–392.

Vaish, A. K. and Chaganty, N. R. (2004). Wishartness and independence of matrix quadratic forms for Kronecker product covariance structures. *Linear Algebra and Its Applications*, 388:379–388.

Wong, C. S. and Wang, T. (1993). Multivariate versions of Cochran's theorems II. *Journal of Multivariate Analysis*, 44(1):146–159.