REVSTAT – Statistical Journal
Volume 15, Number 2, April 2017, 205-222

WEIGHTED-TYPE WISHART DISTRIBUTIONS WITH APPLICATION

Authors: Mohammad Arashi
– Department of Statistics, School of Mathematical Sciences, Shahrood University of Technology, Shahrood, Iran. Department of Statistics, Faculty of Natural and Agricultural Sciences,University of Pretoria, Pretoria, South Africa. (m_arashi_stat@yahoo.com)

Andriëtte Bekker
– Department of Statistics, Faculty of Natural and Agricultural Sciences, University of Pretoria, Pretoria, South Africa. (andriette.bekker@up.ac.za)

Janet van Niekerk
– Department of Statistics, Faculty of Natural and Agricultural Sciences, University of Pretoria, Pretoria, South Africa. (janet.vanniekerk@up.ac.za)

Received: March 2016 Revised: October 2016 Accepted: November 2016

Abstract:

- In this paper, we consider a general framework for constructing new valid densities regarding a random matrix variate. However, we focus specifically on the Wishart distribution. The methodology involves coupling the density function of the Wishart distribution with a Borel measurable function as a weight. We propose three different weights by considering trace and determinant operators on matrices. The characteristics for the proposed weighted-type Wishart distributions are studied and the enrichment of this approach is illustrated. A special case of this weighted-type distribution is applied in the Bayesian analysis of the normal model in the univariate and multivariate cases. It is shown that the performance of this new prior model is competitive using various measures.

Key-Words:

- *Bayesian analysis; eigenvalues; Kummer gamma; Kummer Wishart; matrix variate; weight function; Wishart distribution.*

AMS Subject Classification:

- 62E15, 60E05, 62H12, 62F15

1. INTRODUCTION

The modeling of real world phenomena is constantly increasing in complexity and standard statistical distributions cannot model these adequately. The question arises whether we can introduce new models to compete with and enhance the standard approaches available in the literature. Various generalizations and extensions have been proposed for standard statistical models, since more complex models are needed to solve the modeling complications of real data. To mention a few: Sutradhar et al. (1989) generalized the Wishart distribution for the multivariate elliptical models, however Teng et al. (1989) considered matrix variate elliptical models in their study. Wong and Wang (1995) defined the Laplace-Wishart distribution, while Letac and Massam (2001) defined the normal quasi-Wishart distribution. In the context of graphical models, Roverato (2002) defined the hyper-inverse Wishart and Wang and West (2009) extended the inverse Wishart distribution for using hyper-Markov properties (see Dawid and Lauritzen (1993)), while Bryc (2008) proposed the compound Wishart and q-Wishart in graphical models. Abul-Magd et al. (2009) proposed a generalization to Wishart-Laguerre ensembles. Adhikari (2010) generalized the Wishart distribution for probabilistic structural dynamics, and Díaz-García et al. (2011) extended the Wishart distribution for real normed division algebras. Munilla and Cantet (2012) also formulated a special structure for the Wishart distribution to apply in modeling the maternal animal. These generalizations justify the speculative research to propose new models based on the concept of weighted distributions Rao (1965). Assuming special cases of these new models as priors for an underlying normal model in a Bayesian analysis exhibit interesting behaviour.

In this paper we propose a weighted-type Wishart distribution, making use of the mathematical mechanism frequently used in proposing weighted-type distributions, from length-biased viewpoint, and consider its applications in Bayesian analysis. The building block of our contribution is an extension of the mathematical formulation of univariate weighted-type distributions to multivariate weighted-type distributions. Specifically, if $f(x;\sigma^2)$ is the main/natural probability density function (pdf) which is imposed by a scalar weight function $h(x;\phi)$ (not necessarily positive), then the weighted-type distribution is given by

$$g(x;\boldsymbol{\theta}) = Ch(x;\phi)f(x;\sigma^2), \quad \boldsymbol{\theta} = (\sigma^2,\phi), \tag{1.1}$$

where $C^{-1} = E_{\sigma^2}[h(X;\phi)]$ and the expectation $E_{\sigma^2}[.]$ is taken over the same probability measure as $f(.)$. The parameter ϕ can be seen as an *enriching parameter*.

For the multivariate case, one can simply use the pdf of a multivariate random variable for $f(.)$ in (1.1). Further, the parameter space can be multidimensional. However, the weight function $h(.)$ should remain of scalar form. Thus the question that arises is: Why not replace $f(.)$ in (1.1) with the pdf of a matrix variate random variable? To address this issue and using (1.1) as departure, we define matrix variate weighted-type distributions, from where new matrix variate distributions originate.

Initially let $\mathcal{S}_m$ be the space of all positive definite matrices of dimension m. To set the platform for what we are proposing, consider a random matrix variate $\boldsymbol{X} \in \mathcal{S}_m$ having a pdf $f(.;\boldsymbol{\Psi})$ with parameter $\boldsymbol{\Psi}$. We construct matrix variate distributions, with pdf $g(.;\boldsymbol{\Theta})$, where $\boldsymbol{\Theta} = (\boldsymbol{\Psi}, \boldsymbol{\Phi})$ and enrichment parameter $\boldsymbol{\Phi} \in \mathcal{S}_m$, by utilizing one of the following mechanisms:

1. (Loading with a weight of trace form)

$$g(\boldsymbol{X};\boldsymbol{\Theta}) = C_1 h_1(tr[\boldsymbol{X}\boldsymbol{\Phi}]) f(\boldsymbol{X};\boldsymbol{\Psi}), \boldsymbol{\Theta} = (\boldsymbol{\Psi}, \boldsymbol{\Phi}). \tag{1.2}$$

2. (Loading with a weight of determinant form)

$$g(\boldsymbol{X};\boldsymbol{\Theta}) = C_2 h_2(|\boldsymbol{X}\boldsymbol{\Phi}|) f(\boldsymbol{X};\boldsymbol{\Psi}), \boldsymbol{\Theta} = (\boldsymbol{\Psi}, \boldsymbol{\Phi}). \tag{1.3}$$

3. (Loading with a mixture of weights of trace and determinant forms)

$$g(\boldsymbol{X};\boldsymbol{\Theta}) = C_3 h_1(tr[\boldsymbol{X}\boldsymbol{\Phi}_1]) h_2(|\boldsymbol{X}\boldsymbol{\Phi}_2|) f(\boldsymbol{X};\boldsymbol{\Psi}), \boldsymbol{\Theta} = (\boldsymbol{\Psi}, \boldsymbol{\Phi}_1, \boldsymbol{\Phi}_2), \tag{1.4}$$

where $h_i(.), i = 1, 2$ is a Borel measurable function (weight function) which admits Taylor's series expansion, C_j is a normalizing constant and $f(.)$ can be referred to as a generator.
In this paper, we consider the $f(.)$ in (1.2)-(1.4) to be the pdf of the Wishart distribution with parameters n and $\boldsymbol{\Sigma}$, i.e. $\boldsymbol{\Psi} = (n, \boldsymbol{\Sigma})$, given by

$$\frac{|\boldsymbol{\Sigma}|^{-\frac{n}{2}}}{2^{\frac{nm}{2}}\Gamma_m\left(\frac{n}{2}\right)} |\boldsymbol{X}|^{\frac{n}{2}-\frac{m+1}{2}} \operatorname{etr}\left(-\frac{1}{2}\boldsymbol{\Sigma}^{-1}\boldsymbol{X}\right), \tag{1.5}$$

with $\boldsymbol{X}, \boldsymbol{\Sigma} \in \mathcal{S}_m$, denoted by $W_m(n, \boldsymbol{\Sigma})$, and incorporate a weight function, $h_i(.)$, as given by (1.2)-(1.4). Note that $\Gamma_m(.)$ is the multivariate gamma function and $etr(.) = exp(tr(.))$.

We organize our paper as follows: In Section 2, we discuss the weighted-type Wishart distribution that originated from (1.2) and propose some of its important properties. The enrichment of this approach is illustrated by the graphical display of the joint density function of the eigenvalues of the random matrix for certain cases. In Section 3, the weighted-type Wishart distributions emanating from (1.3) and (1.4) are proposed. The significance of this approach of extending the well-known Wishart distribution, will be demonstrated in Section 4, by assuming special cases as a priors for the underlying univariate and multivariate normal model. Comparison results of these cases with well-known priors path the way for integrating these models in Bayesian analysis. Finally, some thoughts of other possible applications are given in Section 5.

2. WEIGHTED-TYPE I WISHART DISTRIBUTION

In this section we consider the construction methodology of a weighted-type I Wishart distribution according to (1.2).

Definition 2.1. The random matrix $\mathbf{X} \in \mathcal{S}_m$ is said to have a weighted-type I Wishart distribution (W1WD) with parameters $\boldsymbol{\Psi}$, $\boldsymbol{\Phi} \in \mathcal{S}_m$ and the weight function $h_1(.)$, if it has the following pdf

$$
\begin{aligned}
g(\mathbf{X};\boldsymbol{\Theta}) &= \frac{h_1(tr[\boldsymbol{X\Phi}])f(\boldsymbol{X};\boldsymbol{\Psi})}{E\left[h_1(tr[\boldsymbol{X\Phi}])\right]} \\
&= c_{n,m}(\boldsymbol{\Theta})|\boldsymbol{\Sigma}|^{-\frac{n}{2}}|\boldsymbol{X}|^{\frac{n}{2}-\frac{m+1}{2}}\operatorname{etr}\left(-\frac{1}{2}\boldsymbol{\Sigma}^{-1}\boldsymbol{X}\right)h_1(tr[\boldsymbol{X\Phi}]), \quad \boldsymbol{\Theta} = (\boldsymbol{\Psi},\boldsymbol{\Phi}),
\end{aligned}
\tag{2.1}
$$

with

$$
\{c_{n,m}(\boldsymbol{\Theta})\}^{-1} = 2^{\frac{nm}{2}}\Gamma_m\left(\frac{n}{2}\right)\sum_{k=0}^{\infty}\frac{2^k h_1^{(k)}(0)}{k!}\sum_{\kappa}\left(\frac{n}{2}\right)_{\kappa}C_{\kappa}\left(\boldsymbol{\Phi\Sigma}\right), \tag{2.2}
$$

written as $\boldsymbol{X} \sim \mathbb{W}_m^I(n,\boldsymbol{\Sigma},\boldsymbol{\Phi})$. In (2.1) $f(\boldsymbol{X};\boldsymbol{\Psi})$ is the pdf of the Wishart distribution ($W_m(n,\boldsymbol{\Sigma})$) (see 1.5) i.e. $\boldsymbol{\Psi} = (\boldsymbol{\Sigma}, n)$, $n > m-1$, $\boldsymbol{\Sigma} \in \mathcal{S}_m$ and $h_1(.)$ is a Borel measurable function that admits Taylor's series expansion, $(a)_\kappa = \frac{\Gamma_m(a,\kappa)}{\Gamma_m(a)}$ and $\Gamma_m(a,\kappa)$ is the generalized gamma function. The parameters are restricted to take those values for which the pdf is non-negative.

Remark 2.1. Note that using Taylor's series expansion for $h_1(.)$ in (2.1) it follows that

$$
h_1(tr[\boldsymbol{X\Phi}]) = \sum_{k=0}^{\infty}\frac{h_1^{(k)}(0)}{k!}\operatorname{tr}(\boldsymbol{X\Phi})^k = \sum_{k=0}^{\infty}\frac{h_1^{(k)}(0)}{k!}C_\kappa(\boldsymbol{X\Phi}), \tag{2.3}
$$

from Definition 7.21, p.228 of Muirhead (2005) where $h_1^{(k)}(0)$ is the k-th derivative of $h_1(.)$ at the point zero. Therefore using Theorem 7.2.7, p.248 of Muirhead (2005) follows from Definition 2.1 that

$$
\begin{aligned}
E\left[h_1(tr[\boldsymbol{X\Phi}])\right] &= \int_{\mathcal{S}_m} h_1(tr[\boldsymbol{X\Phi}])f(\boldsymbol{X};\boldsymbol{\Psi})d\boldsymbol{X} \\
&= \frac{|\boldsymbol{\Sigma}|^{-\frac{n}{2}}}{2^{\frac{nm}{2}}\Gamma_m\left(\frac{n}{2}\right)}\sum_{k=0}^{\infty}\frac{h_1^{(k)}(0)}{k!}\sum_{\kappa}\int_{\mathcal{S}_m}|\boldsymbol{X}|^{\frac{n}{2}-\frac{m+1}{2}}\operatorname{etr}\left(-\frac{1}{2}\boldsymbol{\Sigma}^{-1}\boldsymbol{X}\right) \\
&\qquad C_\kappa(\boldsymbol{X\Phi})d\boldsymbol{X} \\
&= \frac{|\boldsymbol{\Sigma}|^{-\frac{n}{2}}}{2^{\frac{nm}{2}}\Gamma_m\left(\frac{n}{2}\right)}\sum_{k=0}^{\infty}\frac{2^{\frac{nm}{2}+k}\Gamma_m\left(\frac{n}{2}\right)|\boldsymbol{\Sigma}|^{\frac{n}{2}}h_1^{(k)}(0)}{k!}\sum_{\kappa}\left(\frac{n}{2}\right)_{\kappa}C_\kappa\left(\boldsymbol{\Phi\Sigma}\right) \\
&= \sum_{k=0}^{\infty}\frac{2^k h_1^{(k)}(0)}{k!}\sum_{\kappa}\left(\frac{n}{2}\right)_{\kappa}C_\kappa\left(\boldsymbol{\Phi\Sigma}\right),
\end{aligned}
$$

and (2.2) follows ($C_\kappa\left(a\boldsymbol{X}\right) = a^k C_\kappa\left(\boldsymbol{X}\right)$ and $C_\kappa(.)$ is the zonal polynomial corresponding to κ (Muirhead (2005)).

Remark 2.2. Here we consider some thoughts related to Definition 2.1 and (2.1).

(1) As formerly noticed, the weight function should be a scalar function. In Definition 2.1, we used the trace operator, however any relevant operator can be used. The determinant operator will be discussed in Section 3. Another interesting operator can be the eigenvalue. In this respect one may use the result of Arashi (2013) to get closed expression for the expected value of the weight function.

(2) For $h_1(tr[\boldsymbol{X\Phi}]) = etr(\boldsymbol{X\Phi})$ in (2.1) we obtain an enriched Wishart distribution with scale matrix $\boldsymbol{\Sigma}^{-1} + \boldsymbol{\Phi}$.

(3) For $h_1(x\phi) = exp(x\phi)$ and $m = 1$ in (2.1) the pdf simplifies to

$$g(x;\theta) = c_n(\theta)(\sigma^2)^{-\frac{n}{2}} x^{\frac{n}{2}-1} \exp\left(-\left(\frac{1}{2\sigma^2} - \phi\right) x\right), \tag{2.4}$$

which is the pdf of a gamma random variable with parameters $\frac{n}{2}$ and $\frac{1}{2\sigma^2} - \phi$, with $c_n(\theta) = \frac{\left(\frac{1}{2\sigma^2}-\phi\right)^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}\right)}, \theta = (\sigma^2, \phi)$, written as $G(\alpha = \frac{n}{2}, \beta = \frac{1}{2\sigma^2} - \phi)$.

(4) For $h_1(x) = x$ and $m = 1$ in (2.1) the pdf simplifies to

$$g(x;\theta) = c_n(\theta)(\sigma^2)^{-\frac{n}{2}} x^{\frac{n}{2}-1} \exp\left(-\frac{1}{2\sigma^2}x\right) x = c_n(\theta)(\sigma^2)^{-\frac{n}{2}} x^{\frac{n}{2}} \exp\left(-\frac{1}{2\sigma^2}x\right),$$

with $c_n(\theta) = \frac{\left(\frac{1}{2\sigma^2}-\phi\right)^{\frac{n+2}{2}}}{\Gamma\left(\frac{n}{2}+1\right)}, \theta = (\sigma^2, \phi)$, hence $X \sim G(\alpha = \frac{n}{2}+1, \beta = \frac{1}{2\sigma^2} - \phi)$. This is also called the length-biased or size-biased gamma distribution (see Patil and Ord (1976)) with parameters $\frac{n}{2}$ and $\frac{1}{2\sigma^2} - \phi$.

(5) For $h_1(tr[\boldsymbol{X\Phi}]) = (1 + tr[\boldsymbol{X\Phi}])$ in (2.1) the pdf simplifies to

$$g(\boldsymbol{X};\boldsymbol{\Theta}) = c_{n,m}(\boldsymbol{\Theta})|\boldsymbol{\Sigma}|^{-\frac{n}{2}}|\boldsymbol{X}|^{\frac{n}{2}-\frac{m+1}{2}} \operatorname{etr}\left(-\frac{1}{2}\boldsymbol{\Sigma}^{-1}\boldsymbol{X}\right)(1 + tr(\boldsymbol{X\Phi})), \tag{2.5}$$

with $c_{n,m}(\boldsymbol{\Theta})$ as in (2.2), $\boldsymbol{\Theta} = (n, \boldsymbol{\Sigma}, \boldsymbol{\Phi})$, which is defined as the Kummer Wishart distribution and denoted as $KW_m(n, \boldsymbol{\Sigma}, \boldsymbol{\Phi})$.

(6) For $h_1(x\phi) = (1 + x\phi)^\gamma$, where γ is a known fixed constant, and $m = 1$ in (2.1) the pdf simplifies to

$$g(x;\theta) = c_n(\theta)(\sigma^2)^{-\frac{n}{2}} x^{\frac{n}{2}-1} \exp\left(-\frac{1}{2\sigma^2}x\right)(1 + \phi x)^\gamma \tag{2.6}$$

with $c_n(\theta) = 2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)\sum_{k=0}^{\infty}\frac{(2\phi^2\sigma^2)^k\gamma!}{(\gamma-k)!k!}\sum_\kappa\left(\frac{n}{2}\right)_\kappa$, $\theta = (\sigma^2, \phi)$. If $\phi = 1$ then this is also known as the Kummer gamma or generalized gamma distribution, written as $KG(\alpha = \frac{n}{2}, \beta = \frac{1}{2\sigma^2}, \gamma)$, by expanding the term $(1 + \phi x)^\gamma$ (see Pauw et al. (2010)).

(7) Various functional forms of $h_1(.)$ are explored and the joint density of eigenvalues of the matrix variates are graphically illustrated to show the flexibility built in by this construction, see Table 1.

2.1. Characteristics

In this section some statistical properties of the W1WD (Definition 2.1) are derived. Most of the computations here deal with the relevant use of (2.3) and Theorem 7.2.7, p.248 of Muirhead (2005), though we do not mention every time.

Theorem 2.1. *Let* $\mathbf{X} \sim \mathbb{W}_m^I(n, \boldsymbol{\Sigma}, \boldsymbol{\Phi})$*, then the* r^{th} *moment of* $|\mathbf{X}|$ *is given by*

$$E\left(|\mathbf{X}|^r\right) = \frac{c_{n,m}(\boldsymbol{\Theta})}{c_{2(\frac{n}{2}+r),m}(\boldsymbol{\Theta})}|\boldsymbol{\Sigma}|^r,$$

where $c_{n,m}(\boldsymbol{\Theta})$ *and* $c_{2(\frac{n}{2}+r),m}(\boldsymbol{\Theta})$ *as in* (2.2).

Proof: Similarly as in Remark 2.1, by using (2.1),

$$\begin{aligned} E\left(|\mathbf{X}|^r\right) &= c_{n,m}(\boldsymbol{\Theta})|\boldsymbol{\Sigma}|^{-\frac{n}{2}} \int_{\mathcal{S}_m} |\mathbf{X}|^{r+\frac{n}{2}-\frac{m+1}{2}} \operatorname{etr}\left(-\frac{1}{2}\boldsymbol{\Sigma}^{-\mathbf{1}}\mathbf{X}\right) h_1(tr[\boldsymbol{X}\boldsymbol{\Phi}])d\mathbf{X} \\ &= c_{n,m}(\boldsymbol{\Theta})|\boldsymbol{\Sigma}|^{-\frac{n}{2}} \sum_{k=0}^{\infty} \frac{h_1^{(k)}(0)}{k!} \sum_{\kappa} \int_{\mathcal{S}_m} |\boldsymbol{X}|^{r+\frac{n}{2}-\frac{m+1}{2}} \operatorname{etr}\left(-\frac{1}{2}\boldsymbol{\Sigma}^{-1}\boldsymbol{X}\right) \\ &\quad \times C_\kappa(\boldsymbol{X}\boldsymbol{\Phi})d\boldsymbol{X}, \end{aligned}$$

the result follows. □

In the following, we give the exact expression for the moment generating function (MGF) of the W1WD, provided its existence.

Theorem 2.2. *Let* $\mathbf{X} \sim \mathbb{W}_m^I(n, \boldsymbol{\Sigma}, \boldsymbol{\Phi})$*, then the moment generating function of* $\mathbf{X}$ *is given by*

$$M_{\mathbf{X}}(\mathbf{T}) = c_{n,m}(\boldsymbol{\Theta}) d_{n,m} |\boldsymbol{I}_m - 2\boldsymbol{\Sigma}\boldsymbol{T}|^{-\frac{n}{2}},$$

with $c_{n,m}(\boldsymbol{\Theta})$ *as in (2.2) and* $d_{n,m} = 2^{\frac{nm}{2}}\Gamma_m\left(\frac{n}{2}\right)\sum_{k=0}^{\infty}\sum_\kappa \frac{2^k h_1^{(k)}(0)}{k!}\left(\frac{n}{2}\right)_\kappa$ $\times C_\kappa\left(\boldsymbol{\Phi}(\boldsymbol{\Sigma}^{-1} - 2\boldsymbol{T})^{-1}\right)$.

Proof: Using equation (2.1) we have

$$\begin{aligned} M_{\boldsymbol{X}}(\boldsymbol{T}) &= E(\operatorname{etr}(\boldsymbol{T}\boldsymbol{X})) \\ &= c_{n,m}(\boldsymbol{\Theta})|\boldsymbol{\Sigma}|^{-\frac{n}{2}} \int_{\mathcal{S}_m} |\boldsymbol{X}|^{\frac{n}{2}-\frac{m+1}{2}} \operatorname{etr}\left(-\frac{1}{2}\boldsymbol{\Sigma}^{-1}\boldsymbol{X} + \boldsymbol{T}\boldsymbol{X}\right) h_1(tr[\boldsymbol{X}\boldsymbol{\Phi}])d\boldsymbol{X} \\ &= c_{n,m}(\boldsymbol{\Theta})|\boldsymbol{\Sigma}|^{-\frac{n}{2}} \sum_{k=0}^{\infty}\sum_{\kappa} \frac{2^{\frac{nm}{2}+k} h_1^{(k)}(0)(\frac{n}{2})_\kappa \Gamma_m(\frac{n}{2})\Gamma(\frac{nm}{2}+k)}{k!\Gamma(\frac{nm}{2}+k)} \\ &\quad \times |\boldsymbol{\Sigma}^{-1} - 2\boldsymbol{T}|^{-\frac{n}{2}} C_\kappa(\boldsymbol{\Phi}(\boldsymbol{\Sigma}^{-1} - 2\boldsymbol{T})^{-1}) \end{aligned}$$

and the proof is complete. □

Another important statistical characteristic is the joint pdf of the eigenvalues of $\mathbf{X}$, which is given in the next theorem.

Theorem 2.3. *Let* $\mathbf{X} \sim \mathbb{W}_m^I(n, \mathbf{\Sigma}, \mathbf{\Phi})$*, then the joint pdf of the eigenvalues* $\lambda_1 > \lambda_2 > ... > \lambda_m > 0$ *of* $\mathbf{X}$ *is*

$$\frac{c_{n,m}(\mathbf{\Theta})|\mathbf{\Sigma}|^{-\frac{n}{2}}\pi^{\frac{1}{2}m^2}}{\Gamma_m\left(\frac{m}{2}\right)}\prod_{i<j}^{m}(\lambda_i-\lambda_j)\,|\mathbf{\Lambda}|^{\frac{n}{2}-\frac{m+1}{2}}$$
$$\times\sum_{r=0}^{\infty}\sum_{k=0}^{\infty}\sum_{\rho}\sum_{\kappa}\sum_{\phi\in\rho,\kappa}\frac{h_1^{(k)}(0)C_{\phi}^{\rho,\kappa}\left(\mathbf{I}_m,\mathbf{I}_m\right)C_{\phi}^{\rho,\kappa}\left(-\frac{1}{2}\mathbf{\Sigma^{-1}},\mathbf{\Phi}\right)}{r!k!\left[C_\phi\left(\mathbf{I}_m\right)\right]^2}C_\phi\left(\mathbf{\Lambda}\right).$$

Proof: From Theorem 3.2.17, p.104 of Muirhead (2005) the pdf of $\mathbf{\Lambda} = diag(\lambda_1, ..., \lambda_m)$ is

$$\frac{\pi^{\frac{1}{2}m^2}}{\Gamma_m\left(\frac{m}{2}\right)}\prod_{i<j}^{m}(\lambda_i-\lambda_j)\int_{\mathcal{O}(m)}g(\mathbf{H\Lambda H'};\mathbf{\Theta})d\mathbf{H},$$

where $\mathcal{O}(m)$ is the space of all orthogonal matrices $\boldsymbol{H}$ of order m.

Note that

$$\begin{aligned}\mathbb{I} &= \int_{O(m)}g(\mathbf{H\Lambda H'};\mathbf{\Theta})d\mathbf{H}\\ &= c_{n,m}(\mathbf{\Theta})|\mathbf{\Sigma}|^{-\frac{n}{2}}\int_{\mathcal{O}(m)}|\mathbf{H\Lambda H'}|^{\frac{n}{2}-\frac{m+1}{2}}\,\mathrm{etr}\left(-\frac{1}{2}\mathbf{\Sigma^{-1}H\Lambda H'}\right)h_1(tr[\mathbf{H\Lambda H'\Phi}])d\mathbf{H}\end{aligned}$$

By using (2.3), we get

$$\begin{aligned}\mathbb{I} \;=\; & c_{n,m}(\mathbf{\Theta})|\mathbf{\Sigma}|^{-\frac{n}{2}}|\mathbf{\Lambda}|^{\frac{n}{2}-\frac{m+1}{2}}\sum_{r=0}^{\infty}\frac{1}{r!}\sum_{k=0}^{\infty}\frac{h_1^{(k)}(0)}{k!}\\ &\times\sum_{\rho}\sum_{\kappa}\int_{O(m)}C_\rho\left(-\frac{1}{2}\mathbf{\Sigma^{-1}H\Lambda H'}\right)C_\kappa\left(\mathbf{\Phi H\Lambda H'}\right)d\mathbf{H}.\end{aligned}$$

Note that

$$\begin{aligned}&\int_{O(m)}C_\rho\left(-\frac{1}{2}\mathbf{\Sigma H\Lambda H'}\right)C_\kappa\left(\mathbf{\Phi H\Lambda H'}\right)d\mathbf{H}\\ &=\sum_{\phi\in\rho,\kappa}\frac{C_\phi\left(\mathbf{\Lambda}\right)C_\phi^{\rho,\kappa}\left(\mathbf{I}_m,\mathbf{I}_m\right)C_\phi^{\rho,\kappa}\left(-\frac{1}{2}\mathbf{\Sigma},\mathbf{\Phi}\right)}{\left[C_\phi\left(\mathbf{I}_m\right)\right]^2}\end{aligned}$$

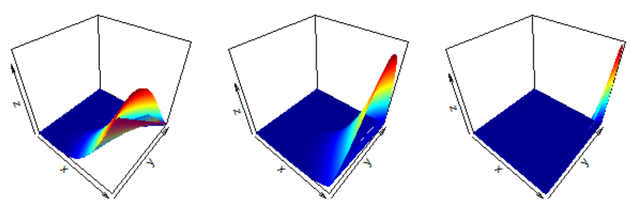
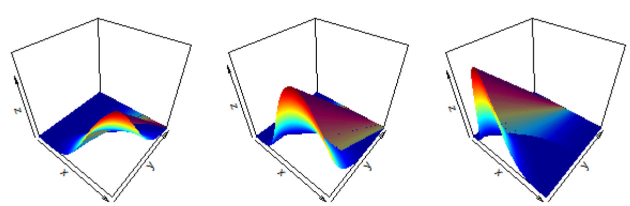
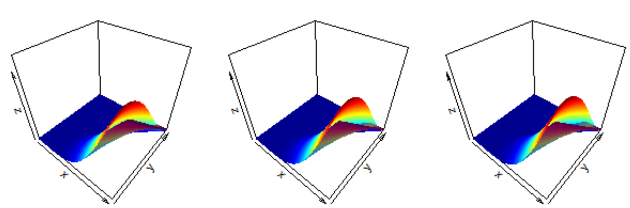
from (1.2), p.468 of Davis (1979) and the result follows. □

Remark 2.3. For $\boldsymbol{\Sigma} = c_1 \boldsymbol{I}$ and $\boldsymbol{\Phi} = c_2 \boldsymbol{I}$ the result can be obtained from Theorem 3.2.17, p.104 of Muirhead (2005) as follows:

$$\mathbb{I} = c_{n,m}(\boldsymbol{\Theta}) c_1^{-\frac{mn}{2}} |\boldsymbol{\Lambda}|^{\frac{n}{2}-\frac{m+1}{2}} \operatorname{etr}\left(-\frac{c_1}{2}\boldsymbol{\Lambda}\right) h_1\left(c_2 \boldsymbol{\Lambda}\right). \tag{2.7}$$

Based on Remark 2.3, Table 1 illustrates the joint pdf of the eigenvalues of $\boldsymbol{X}_{2\times 2}$ for specific c_1, c_2 and n and different weight functions using (2.7). It is evident that the functional form of the weight function provides increased flexibility for the user. Negative and positive correlations amongst the eigenvalues can be obtained using different weight functions, $h_1(.)$.

Table 1: Joint pdf of the eigenvalues for $n = 9, c_2 = 1$ and $c_1 = 0.1$ (Left), $c_1 = 0.5$ (Middle) and $c_1 = 1.5$ (Right)

$h_1(c_2x) = exp(c_2x)$	
$h_1(c_2x) = exp(\frac{1}{c_2x})$	
$h_1(c_2x) = 1 + c_2x$	

3. FURTHER DEVELOPMENTS

3.1. Weighted-type II Wishart distribution

In this section we focus on the construction of a weighted-type Wishart distribution for which the weight function is of determinant form (see (1.3)). Before exploring the form of the weighted-type Wishart distribution based on a weight of determinant form, let $\boldsymbol{X} \sim W_m(n, \boldsymbol{\Sigma})$ and μ_k denote the k-th moment of $|\boldsymbol{X}|$. Then from (15), p.101 of Muirhead (2005)

$$\mu_k = E\left[|\boldsymbol{X}|^k\right] = \frac{2^k \Gamma_m\left(\frac{n}{2}+k\right)}{\Gamma_m\left(\frac{n}{2}\right)} |\boldsymbol{\Sigma}|^k.$$

Thus for any Borel measurable function $h_2(.)$, making use of Taylor's series expansion, we have

$$h_2(|\boldsymbol{X}\boldsymbol{\Phi}|) = \sum_{k=0}^{\infty} \frac{h_2^{(k)}(0)}{k!} |\boldsymbol{X}\boldsymbol{\Phi}|^k. \tag{3.1}$$

Hence

$$E\left[h_2(|\boldsymbol{X}\boldsymbol{\Phi}|)\right] = \sum_{k=0}^{\infty} \frac{h_2^{(k)}(0)}{k!} |\boldsymbol{\Phi}|^k \mu_k = \sum_{k=0}^{\infty} \frac{2^k \Gamma_m\left(\frac{n}{2}+k\right) h_2^{(k)}(0)}{k! \Gamma_m\left(\frac{n}{2}\right)} |\boldsymbol{\Phi}\boldsymbol{\Sigma}|^k.$$

Accordingly, we have the following definition for a weighted-type Wishart distribution with weight of determinant form (see (1.3)).

Definition 3.1. The random matrix $\mathbf{X} \in S_m$ is said to have a weighted-type II Wishart distribution (W2WD) with parameters $\boldsymbol{\Psi}, \boldsymbol{\Phi} \in \mathcal{S}_m$ and the weight function $h_2(.)$, if it has the following pdf

$$\begin{aligned} g(\mathbf{X}; \boldsymbol{\Theta}) &= \frac{h_2(|\boldsymbol{X}\boldsymbol{\Phi}|) f(\boldsymbol{X}; \boldsymbol{\Psi})}{E\left[h_2(|\boldsymbol{X}\boldsymbol{\Phi}|)\right]} \\ &= c^*_{n,m}(\boldsymbol{\Theta}) |\boldsymbol{\Sigma}|^{-\frac{n}{2}} |\boldsymbol{X}|^{\frac{n}{2}-\frac{m+1}{2}} \operatorname{etr}\left(-\frac{1}{2}\boldsymbol{\Sigma}^{-1}\boldsymbol{X}\right) h_2(|\boldsymbol{X}\boldsymbol{\Phi}|), \quad \boldsymbol{\Theta} = (\boldsymbol{\Psi}, \boldsymbol{\Phi}) \end{aligned}$$

with

$$\begin{aligned} \left\{c^*_{n,m}(\boldsymbol{\Theta})\right\}^{-1} &= \int_{\mathcal{S}_m} |\boldsymbol{\Sigma}|^{-\frac{n}{2}} |\boldsymbol{X}|^{\frac{n}{2}-\frac{m+1}{2}} \operatorname{etr}\left(-\frac{1}{2}\boldsymbol{\Sigma}^{-1}\boldsymbol{X}\right) h_2(|\boldsymbol{X}\boldsymbol{\Phi}|) d\boldsymbol{X} \\ &= \sum_{k=0}^{\infty} \frac{h_2^{(k)}(0) 2^{\frac{(n+2k)m}{2}} \Gamma_m\left(\frac{n+2k}{2}\right)}{k!} |\boldsymbol{\Phi}\boldsymbol{\Sigma}|^k. \end{aligned}$$

and $f(\boldsymbol{X}; \boldsymbol{\Psi})$ is the pdf of the Wishart distribution ($W_m(n, \boldsymbol{\Sigma})$) i.e. $\boldsymbol{\Psi} = (\boldsymbol{\Sigma}, n)$, $n > m-1$, $\boldsymbol{\Sigma} \in \mathcal{S}_m$ and $h_2(.)$ is a Borel measurable function that admits Taylor's series expansion. The parameters are restricted to take those values for which the pdf is non-negative. We write this as $\boldsymbol{X} \sim \mathbb{W}^{II}_m(n, \boldsymbol{\Sigma}, \boldsymbol{\Phi})$.

3.2. Weighted-type III Wishart distribution

As before, in this section we give the definition of the weighted-type III Wishart distribution (W3WD). Utilizing a more extended version of (1.4) (allowing more parameters) we have the following definition:

Definition 3.2. The random matrix $\mathbf{X} \in S_m$ is said to have a weighted-type III Wishart distribution (W3WD) with parameters $\boldsymbol{\Psi}$, $\boldsymbol{\Phi}_1$ and $\boldsymbol{\Phi}_2 \in \mathcal{S}_m$ and the weight functions $h_1(.)$ and $h_2(.)$, if it has the following pdf

$$\begin{aligned} g(\mathbf{X};\boldsymbol{\Theta}) &= \frac{h_1(tr[\boldsymbol{X}\boldsymbol{\Phi}_1])h_2(|\boldsymbol{X}\boldsymbol{\Phi}_2|)f(\boldsymbol{X};\boldsymbol{\Psi})}{E\left[h_1(tr[\boldsymbol{X}\boldsymbol{\Phi}_1])h_2(|\boldsymbol{X}\boldsymbol{\Phi}_2|)\right]}, \quad \boldsymbol{\Theta} = (\boldsymbol{\Psi}, \boldsymbol{\Phi}_1, \boldsymbol{\Phi}_2) \\ &= c^{**}_{n,m}(\boldsymbol{\Theta})|\boldsymbol{\Sigma}|^{-\frac{n}{2}}|\boldsymbol{X}|^{\frac{n}{2}-\frac{m+1}{2}} \operatorname{etr}\left(-\frac{1}{2}\boldsymbol{\Sigma}^{-1}\boldsymbol{X}\right) h_1(tr[\boldsymbol{X}\boldsymbol{\Phi}_1])h_2(|\boldsymbol{X}\boldsymbol{\Phi}_2|) \end{aligned}$$

with

$$\begin{aligned} \{c^{**}_{n,m}(\boldsymbol{\Theta})\}^{-1} = \frac{1}{2^{\frac{nm}{2}}} &\sum_{k=0}^{\infty}\sum_{t=0}^{\infty} \frac{h_1^{(k)}(0)h_2^{(t)}(0)}{k!t!} 2^{mt+k}\Gamma_m\left(\frac{n}{2}+t\right)|\boldsymbol{\Phi}_1\boldsymbol{\Sigma}|^t \sum_{\kappa}\left(\frac{n}{2}+t\right)_{\kappa} \\ &\times C_{\kappa}(\boldsymbol{\Phi}_1\boldsymbol{\Sigma}). \end{aligned}$$

where $f(\boldsymbol{X};\boldsymbol{\Psi})$ is the pdf of the Wishart distribution ($W_m(n,\boldsymbol{\Sigma})$) i.e. $\boldsymbol{\Psi} = (\boldsymbol{\Sigma}, n)$, $n > m-1$, $\boldsymbol{\Sigma} \in \mathcal{S}_m$ and $h_1(.)$ and $h_2(.)$ are Borel measurable functions that admit Taylor's series expansion. We denote this as $\boldsymbol{X} \sim \mathbb{W}^{III}_m(n, \boldsymbol{\Sigma}, \boldsymbol{\Phi}_1, \boldsymbol{\Phi}_2)$. The parameters are restricted to take those values for which the pdf is non-negative.

4. APPLICATION

In this section special cases of Definition 1 are applied as priors for the normal model under the squared error loss function. First the Kummer gamma distribution ((2.6) with $\phi = 1$) as a prior for the variance of the univariate normal distribution and secondly the Kummer Wishart (2.5) as a prior for the covariance matrix of the matrix variate normal distribution. Bekker and Roux (1995) considered the Wishart prior as a competitor for the conjugate inverse-Wishart prior for the covariance matrix of the matrix variate normal distribution. Van Niekerk et al. (2016a) confirmed the value added of the latter by a numerical study. This is the stimulus to consider other possible priors.

4.1. Univariate Bayesian illustration

In this section a special univariate case of the weighted-type I Wishart distribution is applied as a prior for the variance of the normal model. Consider a random sample of size n_1 from a univariate normal distribution with unknown mean and variance, i.e. $X_i \sim N(\mu, \sigma^2), \boldsymbol{X} = (X_1, ..., X_{n_1})$. Let $h(x) = (1+x)^\gamma$, $m = 1$ and $\boldsymbol{\Sigma}_{1\times 1} = \sigma^2$ in Definition 2.1 (see (2.6)), and consider this distribution as a prior for σ^2, and an objective prior for μ. This prior model is compared with the well-known inverse gamma and gamma priors in terms of coverage and median credible interval width. The three priors under consideration are

- Inverse gamma prior ($IG(\alpha_1, \beta_1)$) with pdf

$$g(x; \alpha_1, \beta_1) = \frac{\beta_1^{\alpha_1}}{\Gamma(\alpha_1)} x^{-\alpha_1 - 1} \exp\left(-\frac{\beta_1}{x}\right), \quad x > 0$$

- Gamma prior (2.4) ($G(\alpha_2, \beta_2)$)
- Kummer gamma prior (2.6) ($KG(\alpha_3, \beta_3, \gamma = 1)$).

The marginal posterior pdf and Bayes estimator of σ^2 under the Kummer gamma prior are calculated using Remark 5 of Van Niekerk et al. (2016b) as

$$q(\sigma^2|\boldsymbol{X}) = \frac{(\sigma^2)^{\alpha_3 - \frac{n_1}{2} - \frac{1}{2}} \exp\left(-\beta_3\sigma^2\right)\left(1+\phi\sigma^2\right) \exp\left(-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{n_1} X_i^2 - \overline{X}^2\right]\right)}{\Gamma\left(\alpha_3 + \frac{1}{2}\right)\beta_3^{\alpha_3 + \frac{1}{2}} E_{\sigma_1^2}\left[(\sigma_1^2)^{-\frac{n_1}{2}}(1+\phi\sigma_1^2)\exp\left(-\frac{1}{2\sigma_1^2}\left[\sum_{i=1}^{n_1} X_i^2 - \overline{X}^2\right]\right)\right]},$$

and

$$\widehat{\sigma^2} = \frac{\beta_3\Gamma\left(\alpha_3 + \frac{3}{2}\right) E_{\sigma_2^2}\left[(\sigma_2^2)^{-\frac{n_1}{2}}(1+\phi\sigma_2^2)\exp\left(-\frac{1}{2\sigma_2^2}\left[\sum_{i=1}^{n_1} X_i^2 - \overline{X}^2\right]\right)\right]}{\Gamma\left(\alpha_3 + \frac{1}{2}\right) E_{\sigma_1^2}\left[(\sigma_1^2)^{-\frac{n_1}{2}}(1+\phi\sigma_1^2)\exp\left(-\frac{1}{2\sigma_1^2}\left[\sum_{i=1}^{n_1} X_i^2 - \overline{X}^2\right]\right)\right]},$$

where $\sigma_1^2 \sim G\left(\alpha_3 + \frac{1}{2}, \beta_3\right)$ and $\sigma_2^2 \sim G\left(\alpha_3 + \frac{3}{2}, \beta_3\right)$.

A normal sample of size 18 is simulated with mean $\mu = 0$ and variance $\sigma^2 = 1$. The hyperparameters are chosen such that $E(\sigma^2) = \sigma_0^2 = 0.9$.

Four combinations of hyperparameter values are investigated and summarized in Table 2. Note that in combination 4, the prior belief for the Kummer gamma is not 0.9 but 5.23, which is quite far from the target value of 1 and the prior information is clearly misspecified. It is clear from Table 2 that the Kummer gamma prior, with parameter combination 4, is very vague when compared to the other two priors. To evaluate the performance of the new prior structure, 1000 independent samples are simulated and for each one the posterior densities and estimates are calculated. This enables the calculation of the coverage probabilities and median credible interval width as given in Table 3.

The coverage probability obtained under the Kummer gamma prior is higher than for the inverse-gamma and gamma priors, while the median width of the credible interval (indicated in brackets) is competitive. It is interesting to note that even under total misspecification (see combination 4 in Table 2), the Kummer gamma prior is still performing well. The performance superiority of the Kummer gamma prior is clear from Table 3.

Table 2: Influence of hyperparameters on the prior pdf's (– Kummer gamma prior, - - Inverse gamma prior, ... Gamma prior)

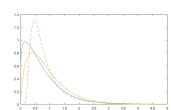
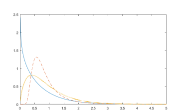
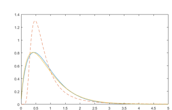
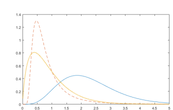
Combination	1	2	3	4
Inverse gamma prior	$\alpha_1 = 3.22$, $\beta_1 = 2$	$\alpha_1 = 4.33$, $\beta_1 = 3$	$\alpha_1 = 3.22$, $\beta_1 = 2$	$\alpha_1 = 3.22$, $\beta_1 = 2$
Gamma prior	$\alpha_2 = 1.8$, $\beta_2 = 2$	$\alpha_2 = 1.8$, $\beta_2 = 2$	$\alpha_2 = 1.8$, $\beta_2 = 2$	$\alpha_2 = 1.8$, $\beta_2 = 2$
Kummer gamma prior	$\alpha_3 = 1.2$, $\beta_3 = 2.1$	$\alpha_3 = 0.8$, $\beta_3 = 1.8$	$\alpha_3 = 1.8$, $\beta_3 = 2.5$	$\alpha_3 = 5.0$, $\beta_3 = 2.5$
				

Table 3: Coverage probabilities (median credible interval width) calculated from the posterior density functions

Combination	Inverse gamma prior	Gamma prior	Kummer gamma prior
1	74.5%(0.8)	77.4%(2.15)	90.8%(0.85)
2	72%(0.75)	78.1%(2.125)	89.2%(0.9)
3	76.8%(0.85)	77.7%(2.05)	92.1%(0.9)
4	73.2%(0.8)	77.6%(1.8)	87.8%(1.4)

4.2. Multivariate Bayesian illustration

In this section the Kummer Wishart (2.5) prior is considered for the covariance matrix of the matrix variate normal model. Consider a random sample of size n_1 from a matrix variate normal distribution with unknown mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$, i.e. $\boldsymbol{X}_i \sim N_{m,p}(\boldsymbol{\mu}, \boldsymbol{\Sigma} \otimes \boldsymbol{I}_p)$ with likelihood function

$$L(\boldsymbol{\mu}, \boldsymbol{\Sigma}|\overline{\boldsymbol{X}}, \boldsymbol{V}) \propto |\boldsymbol{\Sigma}|^{-\frac{n_1 p}{2}} etr\left[-\frac{1}{2}\boldsymbol{\Sigma}^{-1}\left[\boldsymbol{V} + n_1(\overline{\boldsymbol{X}} - \boldsymbol{\mu})(\overline{\boldsymbol{X}} - \boldsymbol{\mu})'\right]\right].$$

The three priors for $\boldsymbol{\Sigma}$ under consideration are

- Inverse Wishart prior ($IW_m(p_1, \boldsymbol{\Phi})$) with pdf

$$\begin{aligned} g(\boldsymbol{X}; p_1, \boldsymbol{\Phi}) &= \left[2^{\frac{m(p_1-m-1)}{2}}\Gamma_m\left(\frac{p_1-m-1}{2}\right)\right]^{-1} |\boldsymbol{X}|^{-\frac{p_1}{2}} |\boldsymbol{\Phi}|^{\frac{p_1-m-1}{2}} \\ &\quad \times etr\left[-\frac{1}{2}\boldsymbol{X}^{-1}\boldsymbol{\Phi}\right], \qquad \boldsymbol{X} \in S_m \end{aligned}$$

- Wishart prior (1.5) ($W_m(p_2, \boldsymbol{\Phi})$)
- Kummer Wishart prior (2.5) ($KW_m(p_3, \boldsymbol{I}, \boldsymbol{\Phi})$).

The conditional posterior pdf's of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ with a Kummer Wishart prior and objective prior for $\boldsymbol{\mu}$, necessary for the simulation of the posterior samples, are

$$q\left(\boldsymbol{\mu}|\boldsymbol{\Sigma}, \overline{\boldsymbol{X}}, \boldsymbol{V}\right) \propto etr\left[-\frac{1}{2}\boldsymbol{\Sigma}^{-1}\left[\boldsymbol{V} + n_1(\overline{\boldsymbol{X}} - \boldsymbol{\mu})(\overline{\boldsymbol{X}} - \boldsymbol{\mu})'\right]\right],$$

and

$$\begin{aligned} q\left(\boldsymbol{\Sigma}|\boldsymbol{\mu}, \overline{\boldsymbol{X}}, \boldsymbol{V}\right) \propto\ & |\boldsymbol{\Sigma}|^{-\frac{n_1 p}{2}+\frac{n}{2}-\frac{m+1}{2}} etr\left[-\frac{1}{2}\boldsymbol{\Sigma}^{-1}\left[\boldsymbol{V} + n_1(\overline{\boldsymbol{X}} - \boldsymbol{\mu})(\overline{\boldsymbol{X}} - \boldsymbol{\mu})'\right]\right] \\ & \times \mathrm{etr}\left(-\frac{1}{2}\boldsymbol{\Phi}^{-1}\boldsymbol{\Sigma}\right)(1 + tr(\boldsymbol{\Sigma}\boldsymbol{\Theta})). \end{aligned}$$

with $\boldsymbol{V} = \sum_{i=1}^{n_1}(\boldsymbol{X}_i - \overline{\boldsymbol{X}})(\boldsymbol{X}_i - \overline{\boldsymbol{X}})'$. A sample of size 10 is simulated from a multivariate normal distribution ($p = 1$) with $\boldsymbol{\mu} = \boldsymbol{0}$ and $\boldsymbol{\Sigma} = \boldsymbol{I}_m$. The hyperparameters are chosen as $\boldsymbol{\Phi} = \boldsymbol{I}_m, m = 3, p_1 = 9.5, p_2 = p_3 = 3$, according to the methodology of Van Niekerk et al. (2016a). Posterior samples of size 5000, are simulated using a Gibbs sampling scheme with an additional Metropolis-Hastings algorithm, similarly to Van Niekerk et al. (2016a).

The estimates calculated for $\boldsymbol{\Sigma}$ under the three different priors as well as the MLE are

$$\widehat{\boldsymbol{\Sigma}}_{MLE} = \begin{bmatrix} 1.8719 & 0.2168 & 0.9523 \\ 0.2168 & 2.9553 & -0.2471 \\ 0.9523 & -0.2471 & 1.0715 \end{bmatrix}, \widehat{\boldsymbol{\Sigma}}_{IW} = \begin{bmatrix} 0.6600 & 0.0772 & 0.3355 \\ 0.0772 & 1.0256 & -0.0873 \\ 0.3355 & -0.0873 & 0.3762 \end{bmatrix}$$

$$\widehat{\boldsymbol{\Sigma}}_{W} = \begin{bmatrix} 0.5547 & 0.0627 & -0.2247 \\ 0.0627 & 0.7968 & 0.0255 \\ -0.2247 & 0.0255 & 1.2348 \end{bmatrix}, \widehat{\boldsymbol{\Sigma}}_{KW} = \begin{bmatrix} 1.1389 & 0.0115 & -0.0098 \\ 0.0115 & 1.0401 & -0.0132 \\ -0.0098 & -0.0132 & 1.0763 \end{bmatrix}$$

The above estimates are obtained for one posterior sample. The Frobenius norm (see Golub and Van Loan (1996)) of the errors, defined as $||\widehat{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}||_F = \sqrt{tr\left(\widehat{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}\right)'\left(\widehat{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}\right)}$, are calculated for each estimate and given in Table 4.

The Kummer Wishart prior results in the smallest Frobenius norm of the error. For further investigation, this sampling scheme is repeated 100 times to obtain 100 estimates under each prior as well as the MLE for each of the 100 simulated samples. The Frobenius norm of the error for each estimate and every repetition is calculated and the empirical cumulative distribution function (ecdf) of each set of Frobenius norms calculated for each estimator is obtained and given in Figure 1. The ecdf which is most left in the figure is regarded as the best since for a specific value of the error norm, a higher proportion of estimates

Table 4: Frobenius norm of the error of the estimates calculated from the simulated sample

Frobenius norm	Value
$\|\|\widehat{\mathbf{\Sigma}}_{MLE} - \mathbf{\Sigma}\|\|_F$	1.2336
$\|\|\widehat{\mathbf{\Sigma}}_{IW} - \mathbf{\Sigma}\|\|_F$	0.9928
$\|\|\widehat{\mathbf{\Sigma}}_{W} - \mathbf{\Sigma}\|\|_F$	1.5766
$\|\|\widehat{\mathbf{\Sigma}}_{KW} - \mathbf{\Sigma}\|\|_F$	0.1468

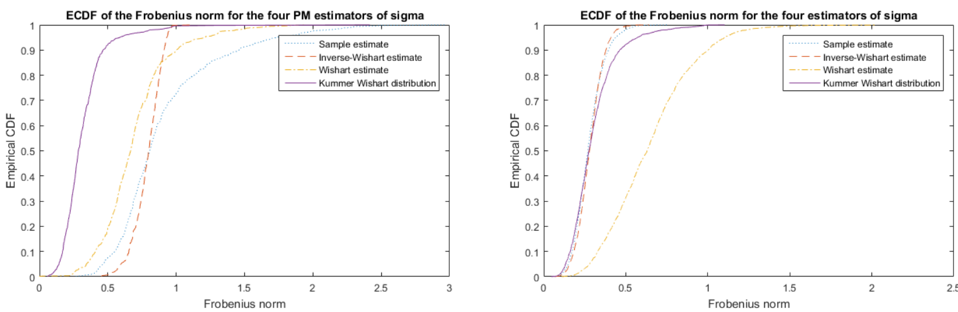


Figure 1: The empirical cumulative distribution function (ecdf) of the Frobenius norm of the estimation errors for $n = 10$(Left) and $n = 100$(Right)

from that particular prior results in less error. It is evident from Figure 1 that the performance of the sample estimate improves as the sample size increases, which is to be expected, and the performance of the Kummer Wishart prior is still competitive. From Figure 1 we conclude that the Kummer Wishart prior results in an estimate of $\mathbf{\Sigma}$, for small and larger sample sizes, with less error and preference should be given to this prior. To validate the graphical interpretation, a two-sample Kolmogorov-Smirnov test is performed for $n = 10$, pairwise, on the three different ecdf's and the p-value for some pairs are given in Table 5.

Table 5: p-values of the Kolmogorv-Smirnov two-sample test based on samples ($n = 10$) of the Frobenius norms

Pairwise comparison	p-value
MLE and IW	< 0.001
IW and KW	< 0.001
W and KW	< 0.001
MLE and KW	< 0.001

From Table 5 it is clear that the ecdf of the errors under the Kummer Wishart prior is significantly different from the other priors. Therefore, the assertion can be made that the Kummer Wishart prior structure produces an estimate that results in statistically significant less error.

5. DISCUSSION

In this paper, we proposed a construction methodology for new matrix variate distributions with specific focus on the Wishart distribution, followed by weighting the Wishart distribution with different weight functions. It was shown from Bayesian viewpoint, by simulation studies, that the Kummer gamma and Kummer Wishart priors, as special cases of the weighted-type I Wishart distribution, outperformed the well-known priors. The weighted-type III Wishart distribution gives rise to a Wishart distribution with larger degrees of freedom and scaled covariance matrix that might have application in missing value analysis.

In the following we list some thoughts that might be considered as plausible applications of the proposed distributions.

(i) Let N_1 and N_2 observations be independently and identically derived from $Z_1 \sim N_m(\mathbf{0}, \boldsymbol{\Sigma}_1)$ and $Z_2 \sim N_m(\mathbf{0}, \boldsymbol{\Sigma}_2)$, respectively. Then the statistic $\boldsymbol{T} = \sum_{j=1}^{N}(Z_1 + Z_2)(Z_1 + Z_2)^T$, has the Wishart distribution $W_m(N_1 + N_2, \boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)$. Suppose the focus of the paper is on the covariance structure $\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2$ (similar to standby systems), then, to reduce the cost of sampling, one may only consider N_1 observations from the W1WD and take $h_1(.)$ to be of exponential form in (2.1).

(ii) A weight of the form $h_2(|\boldsymbol{X}\boldsymbol{\Phi}|) = |\boldsymbol{X}|^{\frac{q}{2}}$, where q is a known fixed constant with $\boldsymbol{\Phi} = \boldsymbol{I}_m$ in Definition 3.1 has applications in missing value analysis. To see this, let $\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_{n+q}$ be a random sample from $N_m(\mathbf{0}, \boldsymbol{\Sigma})$. Define $\boldsymbol{T} = \sum_{i=1}^{n} \boldsymbol{Y}_i \boldsymbol{Y}_i^T$. Then $\boldsymbol{T} \sim W_m(n, \boldsymbol{\Sigma})$. Now, using the weight $h_2(\boldsymbol{X}) = |\boldsymbol{X}|^{\frac{q}{2}}$ one obtains the $W_m(n + q, \boldsymbol{\Sigma})$ distribution and without having the observations $n+1, \ldots, n+q$ we can find the distribution of the full sample and the relative analysis.

(iii) Finally, one may ask what is the sampling distribution regarding Definition 3.2? To answer this question, we recall that if $\boldsymbol{Y} \sim N(\mathbf{0}, \boldsymbol{I}_n \otimes \boldsymbol{\Sigma})$, then $\boldsymbol{X} = \boldsymbol{Y}^T\boldsymbol{Y} \sim W_m(n, \boldsymbol{\Sigma})$. Now, assume matrices $\boldsymbol{A} \in \mathcal{S}_m$ and $\boldsymbol{B} \in \mathcal{S}_m$ exist such that $[\boldsymbol{\Sigma} + \boldsymbol{\Phi}]^{-1} = \boldsymbol{A}^{-1} + \boldsymbol{B}^{-1}$. Then if we sample $\boldsymbol{Y}^* \sim N\left(\mathbf{0}, \boldsymbol{I}_{n+\alpha} \otimes [\boldsymbol{\Sigma} + \boldsymbol{\Phi}]\right)$, the quadratic form $\boldsymbol{X}^* = \boldsymbol{Y}^T\boldsymbol{Y}$ will have the distribution as in Definition 3.2, where $\boldsymbol{A} = \boldsymbol{\Sigma}$, $\boldsymbol{B} = \boldsymbol{\Phi}_1$ and $\boldsymbol{\Phi}_2 = \boldsymbol{I}$. In other words, if we enlarge both the covariance and number of samples in a normal population and consider the distribution of the quadratic form, we are indeed weighting a Wishart distribution with a Wishart.

ACKNOWLEDGMENTS

The authors would like to hereby acknowledge the support of the StatDisT group. This work is based upon research supported by the UP Vice-chancellor's post-doctoral fellowship programme, the National Research foundation grant (Re:CPRR3090132066 No 91497) and the vulnerable discipline-academic statistics (STAT) fund. The authors thank the reviewers and associate editor for their helpful recommendations.

REFERENCES

[1] Abul-Magd,A.Y., Akemann, G. and Vivo, P. (2009). Superstatistical generalizations of Wishart-Laguerre ensembles of random matrices, *Journal of Physics A: Mathematical and Theoretical*, **42**(17), 175–207.

[2] Adhikari, S. (2010). Generalized Wishart distribution for probabilistic structural dynamics, *Computational Mechanics*, **45**(5), 495–511.

[3] Arashi, M. (2013). On Selberg-type square matrices integrals, *Journal of Algebraic Systems*, **1**,1, 53–65.

[4] Bekker, A. and Roux, J.J.J. (1995). Bayesian multivariate normal analysis with a Wishart prior, *Communications in Statistics - Theory and Methods*, **24**(10), 2485–2497.

[5] Bryc, W. (2008). Compound real Wishart and q-Wishart matrices, *International Mathematics Research Notices*, rnn079.

[6] Davis, A.W. (1979). Invariant polynomials with two matrix arguments extending the zonal polynomials: Applications to multivariate distribution theory, *Annals of the Institute of Statistical Mathematics*, **31**(1), 465–485.

[7] Dawid, A.P. and Lauritzen, S.L. (1993). Hyper Markov laws in the statistical analysis of decomposable graphical models, *The Annals of Statistics*, 1272–1317.

[8] Díaz-García, J.A. and Gutiérrez-Jáimez, R. and others (2011). On Wishart distribution: some extensions, *Linear Algebra and its Applications*, **435**(6), 1296–1310.

[9] Golub, G. H. and Van Loan, C. F. (1996). *Matrix Computations, 3rd ed*, Baltimore, MD: Johns Hopkins.

[10] Letac, G. and Massam, H. (2001). The normal quasi-Wishart distribution, *Contemporary Mathematics*, **287**,231–240.

[11] Muirhead, R.J. (1982). *Aspects of Multivariate Statistical Theory*, John Wiley and Sons, New York.

[12] Munilla, S. and Cantet, R.J.C. (2012). Bayesian conjugate analysis using a generalized inverted Wishart distribution accounts for differential uncertainty among the genetic parameters–an application to the maternal animal model, *Journal of Animal Breeding and Genetics*, **129**(3), 173–187.

[13] Patil, G. P. and Ord, J. K. (1976). On size-biased sampling and related form-invariant weighted distributions, *Sankhyā*, **38B**, 48-61.

[14] Pauw, J., Bekker, A. and Roux, J.J.J. (2010). Densities of composite weibullized generalized gamma variables, *South African Statistical Journal*, **44**, 17–42.

[15] Rao, C. R. (1965). *On discrete distributions arising out of methods of ascertainment*, In Classical and Contagious Pergamon Press and Statistical Publishing Society, 320–332.

[16] Rènyi, A. (1961). On measures of information and entropy, *Proceedings of the fourth Berkeley Symposium on Mathematics, Statistics and Probability, 1960*, 547–561.

[17] Roverato, A. (2002). Hyper Inverse Wishart Distribution for Non-decomposable Graphs and its Application to Bayesian Inference for Gaussian Graphical Models, *Scandinavian Journal of Statistics*, **29**(3),391–411.

[18] Shannon, C.E. (1948). A Mathematical Theory of Communication, *Bell System Technical Journal*, **27**(3), 379–423.

[19] Sutradhar, B.C. and Ali, M.M. (1989). A generalization of the Wishart distribution for the elliptical model and its moments for the multivariate t model, *Journal of Multivariate Analysis*, **29**(1),155–162.

[20] Teng, C., Fang H. and Deng, W. (1989). The generalized noncentral Wishart distribution, *Journal of Mathematical Research and Exposition*, **9**(4), 479–488.

[21] Van Niekerk, J., Bekker, A., Arashi, M. and De Waal, D.J. (2016a). Estimation under the matrix variate elliptical model, *South African Statistical Journal*, **50**(1), 149–171.

[22] Van Niekerk, J., Bekker, A., and Arashi, M. (2016b). A gamma-mixture class of distributions with Bayesian application, *Communications in Statistics-Simulation and Computation*, Accepted.

[23] Wang, H. and West, M. (2009). Bayesian analysis of matrix normal graphical models, *Biometrika*, **96**(4),821–834.

[24] Wong, C.S. and Wang, T. (1995). Laplace-Wishart distributions and Cochran theorems, *Sankhyā: The Indian Journal of Statistics, Series A*, 342–359.