# The Wishart distribution with multiple degrees of freedom The Wishart distribution with multiple degrees of freedom

**Preprint** · October 2022

DOI: 10.13140/RG.2.2.12315.34089



**1 author:**

Haruhiko Ogasawara

Otaru University of Commerce

# The Wishart distribution with multiple degrees of freedom

October 31, 2022

Haruhiko Ogasawara*

*Otaru University of Commerce, 3-5-21 Midori, Otaru 047-8501 Japan; Email: emt-hogasa@emt.otaru-uc.ac.jp; Web: https://www.otaru-uc.ac.jp/~emt-hogasa/

**The Wishart distribution with multiple degrees of freedom**

Abstract: Some extended Wishart distribution including different degrees of freedom (df's) are obtained with their probability density function. The multiple df's are due to the sum of independent Wishart distributions with different row/column sizes of the random matrices.

Keywords: multivariate normality, known or unknown parameter, cross product, multiple regression, projection matrix.

## 1. Introduction

An essential expression of the bivariate Wishart distribution was given by Fisher (1915, p. 510). Wishart (1928) obtained the corresponding result for general multivariate cases. The derivation tends to be involved with geometric viewpoints (see e.g., Anderson, 2003, Section 7.2) or not self-contained (for the references of derivations see Anderson, 2003, pp. 256-257 and Ghosh & Sinha, 2002). A simple, expository and self-contained derivation was obtained by Ghosh and Sinha (2002). A Wishart distributed $p \times p$ cross product $\mathbf{U}^*$ under multivariate normality is denoted by $\mathbf{U}^* \sim \mathrm{W}_p(\mathbf{\Sigma}, n)$ with scale parameter $\mathbf{\Sigma}$ and $n$ degrees of freedom (df). The probability density function (pdf) of $\mathbf{U}^*$ at $\mathbf{U}$ is

$$w_p(\mathbf{U} \mid \mathbf{\Sigma}, n) = \frac{\exp\{-\mathrm{tr}(\mathbf{\Sigma}^{-1}\mathbf{U})/2\} \mid \mathbf{U} \mid^{(n-p-1)/2}}{2^{np/2} \mid \mathbf{\Sigma} \mid^{n/2} \Gamma_p(n/2)} \quad (n \geq p \geq 1)$$

(see e.g., Anderson, 2003, Theorem 7.2.2), where $\Gamma_p(t) = \pi^{p(p-1)/4} \prod_{i=1}^{p} \Gamma\left(t - \frac{i-1}{2}\right)$ is the multivariate gamma function (Anderson, 2003, Definition 7.2.1; Subsection 7.2, Equation (19); see also DLMF, 2021, Section 35.3, https://dlmf.nist.gov/35.3).

By definition, the Wishart distribution has a single df. In this paper, extended Wishart distributions with multiple df's are considered. In Section 2, elementary examples for the extended Wishart are presented to show the motivation. Section 3 gives a bivariate case. In Section 4, some multivariate cases are discussed.

## 2. Motivation

One of the most typical cases of $\mathbf{U}^*$ in practice takes $\mathbf{U}^* = (\mathbf{X} - \overline{\mathbf{X}}\mathbf{1}_N^{\mathrm{T}})(\mathbf{X} - \overline{\mathbf{X}}\mathbf{1}_N^{\mathrm{T}})^{\mathrm{T}}$, where $\mathbf{X} = \{X_{ij}\}$ $(i = 1, ..., p; j = 1, ..., N)$, whose $j$-th column is multivariate normally distributed as $\mathrm{N}_p(\boldsymbol{\mu}, \mathbf{\Sigma})$ independent of the other columns with population mean vector $\boldsymbol{\mu}$ and covariance matrix $\mathbf{\Sigma}$; $\overline{\mathbf{X}}$ is the $p \times 1$ mean vector of the $N$ columns of $\mathbf{X}$; and $\mathbf{1}_N$ is the $N \times 1$ vector of 1's. It is well known that $\mathbf{U}^* = \{U_{ij}^*\}$ $(i, j = 1, ..., p)$ can be re-expressed as $\mathbf{U}^* = (\mathbf{X}^* - \boldsymbol{\mu}\mathbf{1}_n^{\mathrm{T}})(\mathbf{X}^* - \boldsymbol{\mu}\mathbf{1}_n^{\mathrm{T}})^{\mathrm{T}}$, where each column of the $p \times n$ matrix $\mathbf{X}^*$ with $n = N - 1$ is obtained from $\mathbf{X}$ by the Helmert transformation (see e.g., Farhadian &

Clarke, 2022) and follows $\mathrm{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ as for $\mathbf{X}$. That is, when $\boldsymbol{\mu}$ is unknown $\mathbf{U}^* = (\mathbf{X} - \bar{\mathbf{X}}\mathbf{1}_N^{\mathrm{T}})(\mathbf{X} - \bar{\mathbf{X}}\mathbf{1}_N^{\mathrm{T}})^{\mathrm{T}} \sim \mathrm{W}_p(\boldsymbol{\Sigma}, n)$ with reduced df $n = N - 1$ while when $\boldsymbol{\mu}$ is known or available $\mathbf{U}^* = (\mathbf{X} - \boldsymbol{\mu}\mathbf{1}_N^{\mathrm{T}})(\mathbf{X} - \boldsymbol{\mu}\mathbf{1}_N^{\mathrm{T}})^{\mathrm{T}} \sim \mathrm{W}_p(\boldsymbol{\Sigma}, N)$ with unreduced $N$ df.

It is obvious that the latter case has an advantage over the former in that the $\{p(p+1)/2\} \times \{p(p+1)/2\}$ Fisher information matrix for $\boldsymbol{\Sigma} = \{\sigma_{ij}\}\,(i, j = 1,\ldots, p)$ is $N/(N-1)$ times that of the former. Equivalently, in the former and latter cases, $\mathrm{cov}(U_{ij}^*, U_{kl}^*) = N(\sigma_{ik}\sigma_{jl} + \sigma_{il}\sigma_{jk})$ and $\mathrm{cov}(U_{ij}^*, U_{kl}^*) = n(\sigma_{ik}\sigma_{jl} + \sigma_{il}\sigma_{jk})$ $(i, j = 1,\ldots, p; k, l = 1,\ldots, p)$, respectively. When some elements of $\boldsymbol{\mu}$ e.g., the last $p_2$ with $p = p_1 + p_2$ are known while those of the remaining $p_1$ are unknown with $N$ independent normal vectors in $\mathbf{X}$, a typical practice may be to use $\mathbf{U}^* = (\mathbf{X} - \bar{\mathbf{X}}\mathbf{1}_N^{\mathrm{T}})(\mathbf{X} - \bar{\mathbf{X}}\mathbf{1}_N^{\mathrm{T}})^{\mathrm{T}}$, where some amount of information is lost since the known $p_2$ population means are not used. Then, it is natural to consider the distribution of

$$\begin{aligned}
\mathbf{U}^* &= \begin{pmatrix} \mathbf{U}_{11}^* \ \mathbf{U}_{12}^* \\ \mathbf{U}_{21}^* \ \mathbf{U}_{22}^* \end{pmatrix} \\
&= \{\mathbf{X} - (\bar{X}_1,\ldots,\bar{X}_{p_1}, \mu_{p_1+1},\ldots,\mu_{p_1+p_2})^{\mathrm{T}}\mathbf{1}_N^{\mathrm{T}}\}\{\mathbf{X} - (\bar{X}_1,\ldots,\bar{X}_{p_1}, \mu_{p_1+1},\ldots,\mu_{p_1+p_2})\mathbf{1}_N^{\mathrm{T}}\}^{\mathrm{T}} \\
&= \{\mathbf{X} - (\bar{\mathbf{X}}_1^{\mathrm{T}}, \boldsymbol{\mu}_2^{\mathrm{T}})^{\mathrm{T}}\mathbf{1}_N^{\mathrm{T}}\}\{\mathbf{X} - (\bar{\mathbf{X}}_1^{\mathrm{T}}, \boldsymbol{\mu}_2^{\mathrm{T}})^{\mathrm{T}}\mathbf{1}_N^{\mathrm{T}}\},
\end{aligned}$$

where $\bar{X}_i$ and $\mu_i$ are the $i$-th elements of $\bar{\mathbf{X}}$ and $\boldsymbol{\mu}$, respectively $(i = 1,\ldots, p)$; $\bar{\mathbf{X}} = (\bar{\mathbf{X}}_1^{\mathrm{T}}, \bar{\mathbf{X}}_2^{\mathrm{T}})^{\mathrm{T}}$; $\boldsymbol{\mu} = (\boldsymbol{\mu}_1^{\mathrm{T}}, \boldsymbol{\mu}_2^{\mathrm{T}})^{\mathrm{T}}$; and

$$\begin{aligned}
\mathbf{U}_{11}^* &= (\mathbf{X}_1 - \bar{\mathbf{X}}_1\mathbf{1}_N^{\mathrm{T}})(\mathbf{X}_1 - \bar{\mathbf{X}}_1\mathbf{1}_N^{\mathrm{T}})^{\mathrm{T}} = \mathbf{X}_1\mathbf{X}_1^{\mathrm{T}} - N\bar{\mathbf{X}}_1\bar{\mathbf{X}}_1^{\mathrm{T}}, \\
\mathbf{U}_{12}^* &= (\mathbf{X}_1 - \bar{\mathbf{X}}_1\mathbf{1}_N^{\mathrm{T}})(\mathbf{X}_2 - \boldsymbol{\mu}_2\mathbf{1}_N^{\mathrm{T}})^{\mathrm{T}} = \mathbf{X}_1\mathbf{X}_2^{\mathrm{T}} - N\bar{\mathbf{X}}_1\bar{\mathbf{X}}_2^{\mathrm{T}} = \mathbf{U}_{21}^{*\mathrm{T}}, \\
\mathbf{U}_{22}^* &= (\mathbf{X}_2 - \boldsymbol{\mu}_2\mathbf{1}_N^{\mathrm{T}})(\mathbf{X}_2 - \boldsymbol{\mu}_2\mathbf{1}_N^{\mathrm{T}})^{\mathrm{T}} \\
&= \{\mathbf{X}_2 - \bar{\mathbf{X}}_2\mathbf{1}_N^{\mathrm{T}} + (\bar{\mathbf{X}}_2 - \boldsymbol{\mu}_2)\mathbf{1}_N^{\mathrm{T}}\}\{\mathbf{X}_2 - \bar{\mathbf{X}}_2\mathbf{1}_N^{\mathrm{T}} - (\bar{\mathbf{X}}_2 - \boldsymbol{\mu}_2)\mathbf{1}_N^{\mathrm{T}}\}^{\mathrm{T}} \\
&= \mathbf{X}_2\mathbf{X}_2^{\mathrm{T}} - N\bar{\mathbf{X}}_2\bar{\mathbf{X}}_2^{\mathrm{T}} + N(\bar{\mathbf{X}}_2 - \boldsymbol{\mu}_2)(\bar{\mathbf{X}}_2 - \boldsymbol{\mu}_2)^{\mathrm{T}}.
\end{aligned}$$

It is found that in the above case of $\mathbf{U}^* = \begin{pmatrix} \mathbf{U}_{11}^* \ \mathbf{U}_{12}^* \\ \mathbf{U}_{21}^* \ \mathbf{U}_{22}^* \end{pmatrix}$, only the submatrix $\mathbf{U}_{22}^*$ is different from the corresponding one of $(\mathbf{X} - \bar{\mathbf{X}}\mathbf{1}_N^{\mathrm{T}})(\mathbf{X} - \bar{\mathbf{X}}\mathbf{1}_N^{\mathrm{T}})^{\mathrm{T}}$ and is larger than the latter by

$N(\bar{\mathbf{X}}_2 - \boldsymbol{\mu}_2)(\bar{\mathbf{X}}_2 - \boldsymbol{\mu}_2)^{\mathrm{T}}$, which is seen as an independent single observation using the Helmert transformation.

Another example is multivariate multiple regression when

$$\mathbf{Y} = \mathbf{XB} + \mathbf{E},$$

where $\mathbf{Y}$ is a $N \times q$ random response matrix for $q$ dependent variables, $\mathbf{X}$ is a $N \times p$ design matrix, $\mathbf{B}$ is a $p \times q$ matrix of population regression coefficients, where $p$ is the number of independent variables possibly including intercepts; and each transposed row of $\mathbf{E}$ independently follows $\mathrm{N}_q(\mathbf{0}, \boldsymbol{\Sigma})$. It can be shown that when $\mathbf{B}$ is unknown and is replaced by the maximum likelihood estimator $\hat{\mathbf{B}} = (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{Y}$, $(\mathbf{Y} - \hat{\mathbf{Y}})^{\mathrm{T}}(\mathbf{Y} - \hat{\mathbf{Y}})$ with $\hat{\mathbf{Y}} = \mathbf{X}\hat{\mathbf{B}} = \mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{Y} \equiv \mathbf{P}_{\mathbf{X}}\mathbf{Y}$ follows $\mathrm{W}_q(\boldsymbol{\Sigma}, N-p)$, where $\mathbf{P}_{\mathrm{X}}$ is a projection matrix onto the space spanned by the columns of $\mathbf{X}$. On the other hand, if $\mathbf{B}$ is known, $(\mathbf{Y} - \mathbf{XB})^{\mathrm{T}}(\mathbf{Y} - \mathbf{XB}) \sim \mathrm{W}_q(\boldsymbol{\Sigma}, N)$.

As in the earlier case, suppose that in $\mathbf{B} = (\mathbf{B}_1\ \mathbf{B}_2)$, where $\mathbf{B}_i$ is a $p \times q_i$ matrix with $q = q_1 + q_2$, $\mathbf{B}_1$ is unknown but $\mathbf{B}_2$ is known. Then, it is natural to consider the distribution of

$$\mathbf{U}^* = \begin{pmatrix} \mathbf{U}_{11}^* \ \mathbf{U}_{12}^* \\ \mathbf{U}_{21}^* \ \mathbf{U}_{22}^* \end{pmatrix}$$

$$= \{\mathbf{Y} - (\mathbf{P}_{\mathbf{X}}\mathbf{Y}_1\ \ \mathbf{XB}_2)\}^{\mathrm{T}}\{\mathbf{Y} - (\mathbf{P}_{\mathbf{X}}\mathbf{Y}_1\ \ \mathbf{XB}_2)\} = \{\mathbf{Y} - (\hat{\mathbf{Y}}_1\ \ \mathbf{XB}_2)\}^{\mathrm{T}}\{\mathbf{Y} - (\hat{\mathbf{Y}}_1\ \ \mathbf{XB}_2)\}$$

rather than $(\mathbf{Y} - \hat{\mathbf{Y}})^{\mathrm{T}}(\mathbf{Y} - \hat{\mathbf{Y}}) = \mathbf{Y}^{\mathrm{T}}(\mathbf{I}_N - \mathbf{P}_{\mathbf{X}})\mathbf{Y}$, where $\mathbf{Y} = (\mathbf{Y}_1\ \mathbf{Y}_2)$ and $\hat{\mathbf{Y}} = (\hat{\mathbf{Y}}_1\ \hat{\mathbf{Y}}_2)$ with obvious partitions and $\mathbf{I}_N$ is the $N \times N$ identity matrix. Noting that in the above expression

$$\mathbf{Y} - (\mathbf{P}_{\mathbf{X}}\mathbf{Y}_1\ \ \mathbf{XB}_2) = \mathbf{Y} - \mathbf{P}_{\mathbf{X}}\mathbf{Y} + (\mathbf{O}\ \ \mathbf{P}_{\mathbf{X}}\mathbf{Y}_2 - \mathbf{XB}_2) = (\mathbf{I}_N - \mathbf{P}_{\mathbf{X}})\mathbf{Y} + \{\mathbf{O}\ \ \mathbf{P}_{\mathbf{X}}(\mathbf{Y}_2 - \mathbf{B}_2)\},$$

$\mathbf{U}^*$ becomes $\mathbf{Y}^{\mathrm{T}}(\mathbf{I}_N - \mathbf{P}_{\mathbf{X}})\mathbf{Y} + \{\mathbf{O}\ \ \mathbf{P}_{\mathbf{X}}(\mathbf{Y}_2 - \mathbf{B}_2)\}^{\mathrm{T}}\{\mathbf{O}\ \ \mathbf{P}_{\mathbf{X}}(\mathbf{Y}_2 - \mathbf{B}_2)\}$. That is, among the four submatrices in $\mathbf{U}^*$ only $\mathbf{U}_{22}^*$ is different from that of $(\mathbf{Y} - \hat{\mathbf{Y}})^{\mathrm{T}}(\mathbf{Y} - \hat{\mathbf{Y}})$ and is larger than the latter by $(\mathbf{Y}_2 - \mathbf{B}_2)^{\mathrm{T}}\mathbf{P}_{\mathbf{X}}(\mathbf{Y}_2 - \mathbf{B}_2)$, whose rank is $\min(p, q_2)$, in Lőwner's sense.

## 3. A bivariate case

Consider the bivariate case with $n_1 \ge 2$ and $n_2 \ge 1$. Define $\mathbf{V}^* = \mathbf{V}^{(1)*} + \mathbf{V}^{(2)*}$, where

$$\mathbf{V}^{(1)*} = \begin{pmatrix} V_{11}^{(1)*} & V_{12}^{(1)*} \\ V_{21}^{(1)*} & V_{22}^{(1)*} \end{pmatrix} = \mathbf{X}^{(1)}\mathbf{X}^{(1)\mathrm{T}} = \begin{pmatrix} \sum_{i=1}^{n_1} X_{1i}^2 & \sum_{i=1}^{n_1} X_{1i}X_{2i} \\ \sum_{i=1}^{n_1} X_{2i}X_{1i} & \sum_{i=1}^{n_1} X_{2i}^2 \end{pmatrix} \text{ and}$$

$$\mathbf{V}^{(2)*} = \begin{pmatrix} 0 & 0 \\ 0 & V_{22}^{(2)*} \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & \sum_{i=n_1+1}^{n_1+n_2} X_{2i}^2 \end{pmatrix}.$$

Suppose that each column of $\mathbf{X}^{(1)}$ i.e., $(X_{1i}\ X_{2i})^{\mathrm{T}}$ independently follows $\mathrm{N}_2(\mathbf{0}, \mathbf{\Sigma})$ $(i = 1,\ldots,n_1)$ and is independent of $X_{2i}$ following $\mathrm{N}(0, \sigma_{22})$ $(i = n_1+1,\ldots,n_1+n_2)$, which are also mutually independent. Then, we have the following results.

**Theorem 1**. *The pdf of* $\mathbf{V}^*$ *at* $\mathbf{V} = \begin{pmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{pmatrix}$ *is*

(i) $n_1 = 2$

$$w_2\{\mathbf{V} \mid \mathbf{\Sigma}, \mathbf{n} = (n_1, n_1+n_2)^{\mathrm{T}}\}$$
$$= 2C_{\mathrm{V}} \mid \mathbf{V} \mid^{-1/2} v_{22}^{n_2/2} \sum_{k=0}^{\infty} \frac{(1/2)_k}{k!} (\mid \mathbf{V} \mid^{-1} v_{11}v_{22})^k \sum_{l=0}^{\infty} \frac{\{(\sigma^{22} - \sigma_{22}^{-1})v_{22}/2\}^l}{l!(n_2 + 2k + 2l)},$$

*where* $C_{\mathrm{V}} = \dfrac{\exp[\{-\mathrm{tr}(\mathbf{\Sigma}^{-1}\mathbf{V})\}/2]}{2^{(2n_1+n_2)/2} \mid \mathbf{\Sigma} \mid^{n_1/2} \sigma_{22}^{n_2/2} \Gamma_2(n_1/2)\Gamma(n_2/2)}$ ; $(a)_k = a(a+1)\cdots(a+k-1)$ *when k is a non-negative integer with* $(a)_0 = 1\ (a \ne 0)$ *and* $(a)_k = \Gamma(a+k)/\Gamma(a)$ $(a > 0, k > 0)$ *is the rising or ascending factorial using the Pochhammer symbol;* $\mathbf{\Sigma}^{-1} = \{\sigma^{ij}\}$ $(i, j = 1, 2)$.

(ii) $n_1 = 3, 5, \ldots$

$$w_2\{\mathbf{V} \mid \mathbf{\Sigma}, \mathbf{n} = (n_1, n_1+n_2)^{\mathrm{T}}\}$$
$$= 2C_{\mathrm{V}} \mid \mathbf{V} \mid^{(n_1-3)/2} v_{22}^{n_2/2} \sum_{k=0}^{(n_1-3)/2} \binom{(n_1-3)/2}{k} (-\mid \mathbf{V} \mid^{-1} v_{11}v_{22})^k \sum_{l=0}^{\infty} \frac{\{(\sigma^{22} - \sigma_{22}^{-1})v_{22}/2\}^l}{l!(n_2 + 2k + 2l)}.$$

*where* $\binom{0}{0} \equiv 1$.

(iii) $n_1 = 4, 6, \ldots$

$$
\begin{aligned}
& w_2\{\mathbf{V} \mid \boldsymbol{\Sigma}, \mathbf{n}=(n_1, n_1+n_2)^{\mathrm{T}}\} \\
& = 2C_{\mathrm{V}} \sum_{k=0}^{(n_1-3)/2} \binom{(n_1-2)/2}{k} (-1)^k \sum_{l=0}^{\infty} |\mathbf{V}|^{(n_1-2k-2l-3)/2}\, v_{11}^{k+l} \frac{(1/2)_l}{l!} \\
& \quad \times \sum_{m=0}^{\infty} \frac{\{(\sigma^{22}-\sigma_{22}^{-1})/2\}^m v_{22}^{(n_2+2k+2l+2m)/2}}{m!(n_2+2k+2l+2m)}.
\end{aligned}
$$

Proof. Under the condition stated above, $\mathbf{V}^{(1)*} \sim \mathrm{W}_2(\boldsymbol{\Sigma}, n_1)$ and $V_{22}^{(2)*} \sim \mathrm{W}_1(\sigma_{22}, n_2)$. Then, the joint pdf of $\mathbf{V}^{(1)*}$ and $V_{22}^{(2)*}$ at $\mathbf{V}^{(1)} = \begin{pmatrix} v_{11}^{(1)} & v_{12}^{(1)} \\ v_{21}^{(1)} & v_{22}^{(1)} \end{pmatrix}$ and $v_{22}^{(2)}$, respectively is

$$
\begin{aligned}
& w_2\{\mathbf{V}^{(1)}, v_{22}^{(2)} \mid \boldsymbol{\Sigma}, \mathbf{n}=(n_1, n_1+n_2)^{\mathrm{T}}\} \\
& = \frac{\exp\{-\mathrm{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{V}^{(1)})/2\} |\mathbf{V}^{(1)}|^{(n_1-3)/2}}{2^{n_1} |\boldsymbol{\Sigma}|^{n_1/2} \Gamma_2(n_1/2)} \times \frac{\exp\{-v_{22}^{(2)}/(2\sigma_{22})\} v_{22}^{(2)(n_2-2)/2}}{(2\sigma_{22})^{n_2/2}\Gamma(n_2/2)}.
\end{aligned}
$$

Employ the change of variable from $\mathbf{V}^{(1)*}$ to $\mathbf{V}^* = \mathbf{V}^{(1)*} + \mathbf{V}^{(2)*}$ with unchanged $\mathbf{V}^{(2)*}$ or $V_{22}^{(2)*}$. Since the Jacobian is unity, the joint pdf of $\mathbf{V}^*$ and $V_{22}^{(2)*}$ at $\mathbf{V} = \begin{pmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{pmatrix}$ and $v_{22}^{(2)}$, respectively becomes

$$
\begin{aligned}
& w_2\{\mathbf{V}, v_{22}^{(2)} \mid \boldsymbol{\Sigma}, \mathbf{n}=(n_1, n_1+n_2)^{\mathrm{T}}\} \\
& = \frac{\exp\left[-\mathrm{tr}\{\boldsymbol{\Sigma}^{-1}(\mathbf{V}-\mathbf{V}^{(2)})/2\}\right] |\mathbf{V}-\mathbf{V}^{(2)}|^{(n_1-3)/2}}{2^{n_1} |\boldsymbol{\Sigma}|^{n_1/2} \Gamma_2(n_1/2)} \times \frac{\exp\{-v_{22}^{(2)}/(2\sigma_{22})\} v_{22}^{(2)(n_2-2)/2}}{(2\sigma_{22})^{n_2/2}\Gamma(n_2/2)} \\
& = \frac{\exp\{-\mathrm{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{V})/2\} \exp\{(\sigma^{22}-\sigma_{22}^{-1}) v_{22}^{(2)}/2\} (|\mathbf{V}| - v_{11}v_{22}^{(2)})^{(n_1-3)/2} v_{22}^{(2)(n_2-2)/2}}{2^{(2n_1+n_2)/2} |\boldsymbol{\Sigma}|^{n_1/2} \sigma_{22}^{n_2/2} \Gamma_2(n_1/2)\Gamma(n_2/2)} \\
& = C_{\mathrm{V}} \exp\{(\sigma^{22}-\sigma_{22}^{-1}) v_{22}^{(2)}/2\} (|\mathbf{V}| - v_{11}v_{22}^{(2)})^{(n_1-3)/2} v_{22}^{(2)(n_2-2)/2}
\end{aligned}
$$

where $\mathbf{V}^{(2)} = \begin{pmatrix} 0 & 0 \\ 0 & v_{22}^{(2)} \end{pmatrix}$. The pdf of $\mathbf{V}^*$ at $\mathbf{V}$ is given by integrating out $v_{22}^{(2)}$ in the above density over its support $(0, v_{22}]$. Denote $v_{22}^{(2)}$ by $u$ for simplicity of notation. Then,

$$
\begin{aligned}
& w_2\{\mathbf{V} \mid \boldsymbol{\Sigma}, \mathbf{n}=(n_1, n_1+n_2)^{\mathrm{T}}\} \\
& = C_{\mathrm{V}} \int_0^{v_{22}} \exp\{(\sigma^{22}-\sigma_{22}^{-1})u/2\} (|\mathbf{V}| - v_{11}u)^{(n_1-3)/2} u^{(n_2-2)/2} \mathrm{d}u \\
& = C_{\mathrm{V}} |\mathbf{V}|^{(n_1-3)/2} \int_0^{v_{22}} \exp\{(\sigma^{22}-\sigma_{22}^{-1})u/2\} (1 - |\mathbf{V}|^{-1} v_{11}u)^{(n_1-3)/2} u^{(n_2-2)/2} \mathrm{d}u,
\end{aligned}
$$

where $\sigma^{22} - \sigma_{22}^{-1} = \sigma_{22}^{-1}\sigma_{12}^2 / |\boldsymbol{\Sigma}| \geq 0$ and by construction $1 - |\mathbf{V}|^{-1} v_{11}u > 0$ since

$|\mathbf{V}| - v_{11}u = |\mathbf{V}^{(1)}| > 0$. The following results are given by cases

(i) $n_1 = 2$

$$
\begin{aligned}
& w_2\{\mathbf{V} \mid \boldsymbol{\Sigma}, \mathbf{n} = (n_1, n_1 + n_2)^{\mathrm{T}}\} \\
&= C_{\mathbf{V}} \, |\mathbf{V}|^{-1/2} \int_0^{v_{22}} \exp\{(\sigma^{22} - \sigma_{22}^{-1})u/2\}(1 - |\mathbf{V}|^{-1} \, v_{11}u)^{-1/2} u^{(n_2-2)/2} \mathrm{d}u \\
&= C_{\mathbf{V}} \, |\mathbf{V}|^{-1/2} \sum_{k=0}^{\infty} \int_0^{v_{22}} \exp\{(\sigma^{22} - \sigma_{22}^{-1})u/2\} \frac{(1/2)_k}{k!} (|\mathbf{V}|^{-1} \, v_{11}u)^k u^{(n_2-2)/2} \mathrm{d}u \\
&= C_{\mathbf{V}} \sum_{k=0}^{\infty} |\mathbf{V}|^{-(2k+1)/2} \, v_{11}^k \frac{(1/2)_k}{k!} \int_0^{v_{22}} \exp\{(\sigma^{22} - \sigma_{22}^{-1})u/2\} u^{(n_2+2k-2)/2} \mathrm{d}u \\
&= C_{\mathbf{V}} \sum_{k=0}^{\infty} |\mathbf{V}|^{-(2k+1)/2} \, v_{11}^k \frac{(1/2)_k}{k!} \sum_{l=0}^{\infty} \frac{\{(\sigma^{22} - \sigma_{22}^{-1})/2\}^l 2}{l!(n_2 + 2k + 2l)} v_{22}^{(n_2+2k+2l)/2} \\
&= 2C_{\mathbf{V}} \, |\mathbf{V}|^{-1/2} \, v_{22}^{n_2/2} \sum_{k=0}^{\infty} \frac{(1/2)_k}{k!} (|\mathbf{V}|^{-1} \, v_{11}v_{22})^k \sum_{l=0}^{\infty} \frac{\{(\sigma^{22} - \sigma_{22}^{-1})v_{22}/2\}^l}{l!(n_2 + 2k + 2l)}.
\end{aligned}
$$

(ii) $n_1 = 3, 5, \ldots$

$$
\begin{aligned}
& w_2\{\mathbf{V} \mid \boldsymbol{\Sigma}, \mathbf{n} = (n_1, n_1 + n_2)^{\mathrm{T}}\} \\
&= C_{\mathbf{V}} \int_0^{v_{22}} \exp\{(\sigma^{22} - \sigma_{22}^{-1})u/2\}(|\mathbf{V}| - v_{11}u)^{(n_1-3)/2} u^{(n_2-2)/2} \mathrm{d}u \\
&= C_{\mathbf{V}} \, |\mathbf{V}|^{(n_1-3)/2} \int_0^{\infty} \exp\{(\sigma^{22} - \sigma_{22}^{-1})u/2\}(1 - |\mathbf{V}|^{-1} \, v_{11}u)^{(n_1-3)/2} u^{(n_2-2)/2} \mathrm{d}u \\
&= C_{\mathbf{V}} \sum_{k=0}^{(n_1-3)/2} \binom{(n_1-3)/2}{k} |\mathbf{V}|^{(n_1-3)/2} \, (-|\mathbf{V}|^{-1} \, v_{11})^k \int_0^{v_{22}} \exp\{(\sigma^{22} - \sigma_{22}^{-1})u/2\} u^{(n_2+2k-2)/2} \mathrm{d}u \\
&= C_{\mathbf{V}} \sum_{k=0}^{(n_1-3)/2} \binom{(n_1-3)/2}{k} |\mathbf{V}|^{(n_1-3)/2} \, (-|\mathbf{V}|^{-1} \, v_{11})^k \sum_{l=0}^{\infty} \frac{\{(\sigma^{22} - \sigma_{22}^{-1})/2\}^l 2}{l!(n_2 + 2k + 2l)} v_{22}^{(n_2+2k+2l)/2} \\
&= 2C \, |\mathbf{V}|^{(n_1-3)/2} \, v_{22}^{n_2/2} \sum_{k=0}^{(n_1-3)/2} \binom{(n_1-3)/2}{k} (-|\mathbf{V}|^{-1} \, v_{11}v_{22})^k \sum_{l=0}^{\infty} \frac{\{(\sigma^{22} - \sigma_{22}^{-1})v_{22}/2\}^l}{l!(n_2 + 2k + 2l)}.
\end{aligned}
$$

(iii) $n_1 = 4, 6, \ldots$

$$
\begin{aligned}
& w_2\{\mathbf{V} \mid \boldsymbol{\Sigma}, \mathbf{n} = (n_1, n_1 + n_2)^{\mathrm{T}}\} \\
&= C_{\mathbf{V}} \, |\mathbf{V}|^{(n_1-3)/2} \int_0^{v_{22}} \exp\{(\sigma^{22} - \sigma_{22}^{-1})u/2\}(1 - |\mathbf{V}|^{-1} \, v_{11}u)^{(n_1-3)/2} u^{(n_2-2)/2} \mathrm{d}u \\
&= C_{\mathbf{V}} \, |\mathbf{V}|^{(n_1-3)/2} \int_0^{v_{22}} \exp\{(\sigma^{22} - \sigma_{22}^{-1})u/2\}(1 - |\mathbf{V}|^{-1} \, v_{11}u)^{(n_1-2)/2} (1 - |\mathbf{V}|^{-1} \, v_{11}u)^{-1/2} u^{(n_2-2)/2} \mathrm{d}u \\
&= C_{\mathbf{V}} \, |\mathbf{V}|^{(n_1-3)/2} \sum_{k=0}^{(n_1-2)/2} \binom{(n_1-2)/2}{k} \\
&\quad \times \sum_{l=0}^{\infty} \int_0^{v_{22}} \exp\{(\sigma^{22} - \sigma_{22}^{-1})u/2\}(-|\mathbf{V}|^{-1} \, v_{11}u)^k \frac{(1/2)_l}{l!} (|\mathbf{V}|^{-1} \, v_{11}u)^l u^{(n_2-2)/2} \mathrm{d}u
\end{aligned}
$$

$$= C_{\mathrm{V}} \sum_{k=0}^{(n_1-3)/2} \sum_{l=0}^{\infty} \binom{(n_1-2)/2}{k} (-1)^k \, |\, \mathbf{V} \,|^{(n_1-2k-2l-3)/2} \; v_{11}^{k+l} \, \frac{(1/2)_l}{l!}$$

$$\times \int_0^{v_{22}} \exp\{(\sigma^{22} - \sigma_{22}^{-1}) u / 2\} u^{(n_2+2k+2l-2)/2} \mathrm{d}u$$

$$= 2C_{\mathrm{V}} \sum_{k=0}^{(n_1-3)/2} \binom{(n_1-2)/2}{k} (-1)^k \sum_{l=0}^{\infty} |\, \mathbf{V} \,|^{(n_1-2k-2l-3)/2} \; v_{11}^{k+l} \, \frac{(1/2)_l}{l!}$$

$$\times \sum_{m=0}^{\infty} \frac{\{(\sigma^{22} - \sigma_{22}^{-1})/2\}^m v_{22}^{(n_2+2k+2l+2m)/2}}{m!(n_2+2k+2l+2m)}$$

The above cases give the required results. Q.E.D.

**Remark 1**. For e.g., (i) of Theorem 1, we have an alternative expression

$$w_2\{\mathbf{V} \,|\, \mathbf{\Sigma}, \mathbf{n} = (n_1, n_1+n_2)^{\mathrm{T}}\}$$

$$= 2C_{\mathrm{V}} \,|\, \mathbf{V} \,|^{-1/2} \; v_{22}^{n_2/2}$$

$$\times {}_1F_{0\mathrm{W}} \left[ 1/2;;|\, \mathbf{V} \,|^{-1} \, v_{11} v_{22}; \, {}_0F_{0\mathrm{W}} \left\{ ;;(\sigma^{22} - \sigma_{22}^{-1}) v_{22}/2; (n_2+2k+2l)^{-1}, l = 0,1.,... \right\}, k = 0,1,... \right],$$

where ${}_1F_{0\mathrm{W}}(a;;z;w_k, k=0,1,...) = \sum_{k=0}^{\infty} (a)_k z^k w_k / k! \; (a>0)$ is the weighted negative binomial series or hypergeometric function (hgf) as used by Ogasawara (2022) (for similar weighted Gauss and Kummer confluent hgf's see Ogasawara, 2021, Equations (4) and (7)) where $w_k$ is the weight for the *k*-th term; and in this case $w_k = {}_0F_{0\mathrm{W}}(;;z;w_l^{(k)}, l=0,1,...)$ $= \sum_{l=0}^{\infty} z^l w_l^{(k)} / l!$, which seems to be new to the author's knowledge and may be called a weighted exponential series or hgf. Then, the above alternative expression is seen as a weighted bivariate or 2-fold hgf. Note that the usual unweighted cases are

$${}_1F_{0\mathrm{W}}(a;;z;w_k = 1, k=0,1,...) = {}_1F_0(a;;z) = (1-z)^{-a} \quad \text{and}$$

$${}_0F_{0\mathrm{W}}(;;z;w_l^* = 1, l=0,1,...) = {}_0F_0(;;z) = e^z .$$

## 4. Some multivariate cases

In this section multivariate cases are discussed.

**Remark 2**. Consider an extended case of Theorem 1 with $p_1+p_2$ variables, where $\mathbf{V}^{(1)*} \sim \mathrm{W}_{p_1+p_2}(\mathbf{\Sigma}, n_1)(n_1 \ge p_1+p_2)$ independent of $\mathbf{V}_{22}^{(2)*} \sim \mathrm{W}_{p_2}(\mathbf{\Sigma}, n_2)(n_2 \ge p_2)$. Let $\mathbf{V}^* = \mathbf{V}^{(1)*} + \mathbf{V}^{(2)*}$, where $\mathbf{V}^{(2)*} = \begin{pmatrix} \mathbf{O} & \mathbf{O} \\ \mathbf{O} & \mathbf{V}_{22}^{(2)*} \end{pmatrix}$. Then, using an integral expression, the pdf of

$\mathbf{V}^*$ at $\mathbf{V} = \begin{pmatrix} \mathbf{V}_{11} & \mathbf{V}_{12} \\ \mathbf{V}_{21} & \mathbf{V}_{22} \end{pmatrix}$ is obtained in the following way. The joint pdf of $\mathbf{V}^{(1)*}$ and $\mathbf{V}_{22}^{(2)*}$

at $\mathbf{V}^{(1)} = \begin{pmatrix} \mathbf{V}_{11}^{(1)} & \mathbf{V}_{12}^{(1)} \\ \mathbf{V}_{21}^{(1)} & \mathbf{V}_{22}^{(1)} \end{pmatrix}$ and $\mathbf{V}_{22}^{(2)}$, respectively is

$$w_{p_1+p_2}\{\mathbf{V}^{(1)}, \mathbf{V}_{22}^{(2)} \mid \mathbf{\Sigma}, \mathbf{n} = (n_1, n_1+n_2)^{\mathrm{T}}\}$$
$$= \frac{\exp\{-\mathrm{tr}(\mathbf{\Sigma}^{-1}\mathbf{V}^{(1)})/2\} \mid \mathbf{V}^{(1)} \mid^{(n_1-p_1-p_2-1)/2}}{2^{(p_1+p_2)n_1/2} \mid \mathbf{\Sigma} \mid^{n_1/2} \Gamma_{p_1+p_2}(n_1/2)} \times \frac{\exp\{-\mathrm{tr}(\mathbf{\Sigma}_{22}^{-1}\mathbf{V}_{22}^{(2)})/2\} \mid \mathbf{V}_{22}^{(2)} \mid^{(n_2-p_2-1)/2}}{2^{p_2n_2/2} \mid \mathbf{\Sigma}_{22} \mid^{n_2/2} \Gamma_{p_2}(n_2/2)},$$

where $\mathbf{\Sigma} = \begin{pmatrix} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} \\ \mathbf{\Sigma}_{21} & \mathbf{\Sigma}_{22} \end{pmatrix}$ with $\mathbf{\Sigma}_{ij}$ being the $p_i \times p_j$ submatrix $(i, j = 1, 2)$

Employ the change of variable from $\mathbf{V}^{(1)*}$ to $\mathbf{V}^* = \mathbf{V}^{(1)*} + \mathbf{V}^{(2)*}$ with unchanged $\mathbf{V}^{(2)*}$ or $\mathbf{V}_{22}^{(2)*}$. Since the Jacobian is unity, the joint pdf of $\mathbf{V}^*$ and $\mathbf{V}_{22}^{(2)*}$ at $\mathbf{V}$ and $\mathbf{V}_{22}^{(2)}$, respectively becomes

$$w_{p_1+p_2}\{\mathbf{V}, \mathbf{V}_{22}^{(2)} \mid \mathbf{\Sigma}, \mathbf{n} = (n_1, n_1+n_2)^{\mathrm{T}}\}$$
$$= \frac{\exp\left[-\mathrm{tr}\{\mathbf{\Sigma}^{-1}(\mathbf{V} - \mathbf{V}^{(2)})\}/2\right] \mid \mathbf{V} - \mathbf{V}^{(2)} \mid^{(n_1-p_1-p_2-1)/2}}{2^{(p_1+p_2)n_1/2} \mid \mathbf{\Sigma} \mid^{n_1/2} \Gamma_{p_1+p_2}(n_1/2)}$$
$$\times \frac{\exp\{-\mathrm{tr}(\mathbf{\Sigma}_{22}^{-1}\mathbf{V}_{22}^{(2)})/2\} \mid \mathbf{V}_{22}^{(2)} \mid^{(n_2-p_2-1)/2}}{2^{p_2n_2/2} \mid \mathbf{\Sigma}_{22} \mid^{n_2/2} \Gamma_{p_2}(n_2/2)}$$
$$= \frac{\exp\{-\mathrm{tr}(\mathbf{\Sigma}^{-1}\mathbf{V})/2\} \mid \mathbf{V} \mid^{(n_1-p_1-p_2-1)/2}}{2^{(p_1+p_2)n_1/2} \mid \mathbf{\Sigma} \mid^{n_1/2} \Gamma_{p_1+p_2}(n_1/2)}$$
$$\times \frac{\exp[\mathrm{tr}\{(\mathbf{\Sigma}^{22} - \mathbf{\Sigma}_{22}^{-1})\mathbf{V}_{22}^{(2)}/2](1 - \mid \mathbf{V} \mid^{-1} \mid \mathbf{V}_{11} \mid\mid \mathbf{V}_{22}^{(2)} \mid)^{(n_1-p_1-p_2-1)/2} \mid \mathbf{V}_{22}^{(2)} \mid^{(n_2-p_2-1)/2}}{2^{p_2n_2/2} \mid \mathbf{\Sigma}_{22} \mid^{n_2/2} \Gamma_{p_2}(n_2/2)}.$$

The pdf of $\mathbf{V}$ is formally given by integrating out $\mathbf{V}_{22}^{(2)}$ in the above result over $\mathbf{O} < \mathbf{V}_{22}^{(2)} \le \mathbf{V}_{22}$:

$$w_{p_1+p_2}\{\mathbf{V} \mid \mathbf{\Sigma}, \mathbf{n} = (n_1, n_1+n_2)^{\mathrm{T}}\}$$
$$= \frac{\exp\{-\mathrm{tr}(\mathbf{\Sigma}^{-1}\mathbf{V})/2\} \mid \mathbf{V} \mid^{(n_1-p_1-p_2-1)/2}}{2^{(p_1+p_2)n_1/2} \mid \mathbf{\Sigma} \mid^{n_1/2} \Gamma_{p_1+p_2}(n_1/2)}$$
$$\times \frac{\int_{\mathbf{O}<\mathbf{V}_{22}^{(2)} \le \mathbf{V}_{22}} \exp[\mathrm{tr}\{(\mathbf{\Sigma}^{22} - \mathbf{\Sigma}_{22}^{-1})\mathbf{V}_{22}^{(2)}\}/2](1 - \mid \mathbf{V} \mid^{-1} \mid \mathbf{V}_{11} \mid\mid \mathbf{V}_{22}^{(2)} \mid)^{(n_1-p_1-p_2-1)/2} \mid \mathbf{V}_{22}^{(2)} \mid^{(n_2-p_2-1)/2} \mathrm{d}\mathbf{V}_{22}^{(2)}}{2^{p_2n_2/2} \mid \mathbf{\Sigma}_{22} \mid^{n_2/2} \Gamma_{p_2}(n_2/2)},$$

where the inequalities are used in Lőwner's sense and $\mathbf{\Sigma}^{-1}=\begin{pmatrix}\mathbf{\Sigma}^{11} & \mathbf{\Sigma}^{12}\\ \mathbf{\Sigma}^{21} & \mathbf{\Sigma}^{22}\end{pmatrix}$ with $\mathbf{\Sigma}^{22}=\mathbf{\Sigma}_{22}^{-1}+\mathbf{\Sigma}_{22}^{-1}\mathbf{\Sigma}_{21}(\mathbf{\Sigma}_{11}-\mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}\mathbf{\Sigma}_{21})^{-1}\mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}\geq\mathbf{\Sigma}_{22}^{-1}$. However, the explicit result will become complicated possibly including zonal polynomials (Muirhead, 1982, Chapter 7; Takemura, 1984) when $p_2>1$. The case with $p_2=1$ is given in Remark 3

**Remark 3**. The pdf in Remark 2 with $\mathbf{V}_{22}=v_{p_1+1,p_1+1}$ when $p_2=1$ is obtained as follows:

$$w_{p_1+1}\{\mathbf{V}\,|\,\mathbf{\Sigma},\mathbf{n}=(n_1,n_1+n_2)^{\mathrm{T}}\}$$
$$=\frac{\exp\{-\mathrm{tr}(\mathbf{\Sigma}^{-1}\mathbf{V})/2\}\,|\,\mathbf{V}\,|^{(n_1-p_1-2)/2}}{2^{(p_1+1)n_1/2}\,|\,\mathbf{\Sigma}\,|^{n_1/2}\,\Gamma_{p_1+1}(n_1/2)}$$
$$\times\frac{\int_0^{v_{p_1+1,p_1+1}}\exp\{(\sigma^{p_1+1,p_1+1}-\sigma_{p_1+1,p_1+1}^{-1})u/2\}(1-|\,\mathbf{V}\,|^{-1}|\,\mathbf{V}_{11}\,|u)^{(n_1-p_1-2)/2}u^{(n_2-2)/2}\mathrm{d}u}{(2\sigma_{p_1+1,p_1+1})^{n_2/2}\Gamma(n_2/2)},$$

whose results can be shown by cases for $n_1-p_1\,(\geq 1)$ as in Theorem 1.

**Remark 4**. In Section 2, the example of known and unknown means was shown. When, the number of variables for known means is more than 1, the corresponding Wishart distribution to be added to that for unknown means become a singular Wishart with one degree of freedom (Srivastava, 2003), which is another problem to be investigated with other singular cases (Bodnar & Okhrin, 2008; Yonenaga, 2022).

## Declarations

The author states that there is no conflict of interest.

## References


Anderson, T. W. (2003). *An introduction to multivariate statistical analysis* (3rd ed.). New York: Wiley.

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

DLMF (2021). *NIST Digital Library of Mathematical Functions*. https://dlmf.nist.gov/, Release 1.1.3 of 2021-09-15. F. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I.

Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller, B. V. Saunders, H. S. Cohl, and M. A. McClain (Eds.).

Farhadian, R., & Clarke, B. R. (2022). A note on the Helmert transformation. *Communications in Statistics – Theory and Methods, 51*, 5258-5264.

Fisher, R. A. (1915). Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. *Biometrika, 10*, 507-521.

Ghosh, M., & Sinha, B. K. (2002). A simple derivation of the Wishart distribution. *The American Statistician, 56*, 100-101.

Muirhead, R. J. (1982). *Aspects of multivariate statistical theory*. New York: Wiley.

Ogasawara, H. (2021). A non-recursive formula for various moments of the multivariate normal distribution with sectional truncation. *Journal of Multivariate Analysis* (online published). https://doi.org/10.1016/j.jmva.2021.104729.

Ogasawara, H. (2022). *The distribution of the sample correlation coefficient under variance-truncated normality*. Preprint at https://doi.org/10.13140/RG.2.2.19080.90883.

Srivastava, M. S. (2003). Singular Wishart and multivariate beta distributions. *The Annals of Statistics*, *31*, 1537-1560.

Takemuta, A. (1984). *Zonal polynomials*. Hayward, CA: Institute of Mathematical Statistics.

Wishart, J. (1928). The generalised product moment distribution in samples from a normal multivariate population. *Biometrika*, 32-52.

Yonenaga, K. (2022). *Functionals of a Wishart matrix and a normal vector and its application to linear discriminant analysis*. Doctoral dissertation, Graduate School of Economics and Business, Hokkaido University, Japan.