INDEPENDENCE PROPERTIES OF WISHART MATRICES

JACEK WESOLOWSKI

1. Bartlett's decomposition

Let Ω_n denotes the cone of $n\times n$ real positive definite matrices. Let $\mathbf{X}\in\Omega_n$ be a Wishart matrix with the density

$$f(x)\propto |x|^{p-\frac{n+1}{2}}\,\exp\left(-\tfrac{\operatorname{tr} x}{2}\right)\,I_{\Omega_n}(x) \tag{1}$$

where $p>\frac{n-1}{2}$. Then we write $\mathbf{X}\sim \mathrm{W}_n(p,\mathbf{I})$. In general, we say that $\mathbf{X}$ is Wishart $\mathrm{W}_n(p,\mathbf{\Sigma})$, where $\mathbf{\Sigma}$ is a positive definite matrix, if $\mathbf{\Sigma}^{-1/2}\mathbf{X}\mathbf{\Sigma}^{-1/2}$ is Wishart, $\mathrm{W}_n(p,\mathbf{I})$.

Let $\mathcal{LT}_n$ and $\mathcal{V}_n$ denote, respectively, linear spaces of lower triangular and symmetric $n\times n$ real matrices. By Cholesky decomposition any $\mathbf{a}\in\mathcal{V}_n$ has a unique representation

$$\mathbf{a}=\mathbf{b}\,\mathbf{b}^*,\quad \mathbf{b}\in\mathcal{LT}_n \tag{2}$$

The following result goes back to Bartlett (1933).

Theorem 1.1. *Let $\mathbf{X}$ be a Wishart matrix with the density (1). Define the random matrix $\mathbf{T}\in\mathcal{LT}_n$ by the Cholesky decomposition of $\mathbf{X}$, that is $\mathbf{X}=\mathbf{T}\,\mathbf{T}^*$.*

Then the matrix $\mathbf{T}=[T_{ij}]$ has independent components such that

$$T_{ii}^2\sim\chi^2\left(2p-i+1\right),\quad i=1,\dots,n,\quad \textit{and}\quad T_{ij}\sim \mathrm{N}(0,1),\quad 1\le j<i\le n.$$

1.1. **Determinants of certain endomorphisms.**

Lemma 1.2. *Let $\mathbf{a}\in\mathcal{M}_n$ (the space of $n\times n$ real matrices). Let $\mathbb{P}(\mathbf{a}):\mathcal{V}_n\to\mathcal{V}_n$ be an endomorphism defined by*

$$\mathbb{P}(\mathbf{a})\mathbf{x}=\mathbf{a}\mathbf{x}\mathbf{a}^*.$$

Then

$$\operatorname{Det}\mathbb{P}(\mathbf{a})=\pm|\mathbf{a}|^{n+1}. \tag{3}$$

Proof. Consider first diagonal $\mathbf{a}=\operatorname{diag}(a_1,\dots,a_n)$. Let $\mathbf{e}^{ij}\in\mathcal{V}_n$ be a matrix with only non-zero elements, equal 1, at entries (i,j) and (j,i), $i,j\in\{1,\dots,n\}$. Then $\mathbb{P}(\mathbf{a})\mathbf{e}^{ij}=a_ia_j\mathbf{e}^{ij}$. Consequently, a_ia_j is the eigenvalue of $\mathbb{P}(\mathbf{a})$. Therefore,

$$\operatorname{Det}\mathbb{P}(\mathbf{a})=\prod_{1\le i\le j\le n}a_ia_j=\prod_{i=1}^n a_i^{n+1}=|\mathbf{a}|^{n+1}.$$

For a general $\mathbf{a}\in\mathcal{M}_n$ we use polar decomposition: $\mathbf{a}=\mathbf{u}\,\mathbf{d}\,\mathbf{v}^*$, where $\mathbf{u}$, $\mathbf{v}\in\mathcal{M}_n$ are orthogonal and $\mathbf{d}\in\mathcal{M}_n$ is diagonal with non-negative entries. Consequently,

$$|\mathbf{a}|=\pm|\mathbf{d}|.$$

Now we note that $\mathbb{P}(\mathbf{ab})=\mathbb{P}(\mathbf{a})\circ\mathbb{P}(\mathbf{b})$ and thus $\operatorname{Det}\mathbb{P}(\mathbf{ab})=\operatorname{Det}\mathbb{P}(\mathbf{a})\operatorname{Det}\mathbb{P}(\mathbf{b})$.

Therefore, using the SVD for $\mathbf{a}$ we get

$$\operatorname{Det}\mathbb{P}(\mathbf{a})=\operatorname{Det}\mathbb{P}(\mathbf{u})\operatorname{Det}\mathbb{P}(\mathbf{d})\operatorname{Det}\mathbb{P}(\mathbf{v})=\pm\operatorname{Det}\mathbb{P}(\mathbf{d})=\pm|\mathbf{d}|^{n+1}=\pm|\mathbf{a}|^{n+1}.$$

Date: June 20, 2016.

$\square$

Lemma 1.3. *Let $v : \mathcal{LT}_n \to \mathcal{V}_n$ be an endomorphism defined by*

$$v(\mathbf{x}) = \mathbf{x} + \mathbf{x}^*.$$

Then

$$\mathrm{Det}\, v = 2^n. \tag{4}$$

Proof. Since Det of a linear operator does not depend on the choice of the basis we identify $\mathcal{V}_n \cong \mathcal{LT}_n$. Under this identification

$$v(\mathbf{x}) = \mathbf{x} + \mathrm{diag}(x_{11}, \ldots, x_{nn}) \in \mathcal{LT}_n.$$

Then we see that

$$v(\mathbf{e}^{ii}) = 2\mathbf{e}^{ii}, \quad i = 1, \ldots, n,$$

so the eigenvalue $\lambda = 2$ is of multiplicity n.

Moreover, for $\mathbf{f}^{ij} \in \mathcal{LT}_n$ with only non-zero element equal 1 at the (i, j)th entry, $i > j$, we have

$$v(\mathbf{f}^{ij}) = \mathbf{f}^{ij},$$

so the eigenvalue $\lambda = 1$ is of multiplicity $\frac{n(n-1)}{2}$. $\square$

Lemma 1.4. *For $\mathbf{a} = [a_{ij}] \in \mathcal{LT}_n$ let $m_{\mathbf{a}} : \mathcal{LT}_n \to \mathcal{LT}_n$ be an endomorphism defined by*

$$m_{\mathbf{a}}\mathbf{x} = \mathbf{a}\mathbf{x}.$$

Then

$$\mathrm{Det}\, m_{\mathbf{a}} = \prod_{i=1}^{n} a_{ii}^{i}. \tag{5}$$

Proof. Assume first that a_{ii}, $i = 1, \ldots, n$, are distinct numbers. For any pair $i \in \{1, \ldots, n\}$ and any $j \in \{1, \ldots, i\}$ consider a matrix $\mathbf{g}^{(ij)}$ with all entries equal 0 except $g_{ij}^{(ij)} = 1$ and $g_{kj}^{(ij)} = x_{kj}$, $k = i+1, \ldots, n$, satisfying

$$a_{li} + \sum_{k=i+1}^{l} a_{lk}x_{kj} = a_{ii}x_{lj}, \quad l = i+1, \ldots, n. \tag{6}$$

This system has a unique solution since its determinant is $\prod_{l=i+1}^{n} (a_{ii} - a_{ll}) \neq 0$. We also note that $\mathbf{g}^{(ij)}$, $1 \leq j \leq i \leq n$ are linearly independent elements of $\mathcal{LT}_n$. Moreover, equations (6) imply

$$m_{\mathbf{a}}\left(\mathbf{g}^{(ij)}\right) = a_{ii}\mathbf{g}^{(ij)}, \quad j = 1, \ldots, i \leq n.$$

That is a_{ii} is an eigenvalue of $m_{\mathbf{a}}$ having multiplicity i, $i = 1, \ldots, n$. Consequently, in this case (5) holds true.

If a_{ii}, $i = 1, \ldots, n$, are not distinct, we can consider for $\epsilon > 0$ perturbed matrices $\mathbf{a}_\epsilon = [a_{ij}(\epsilon)]$ such that $a_{ii}(\epsilon)$, $i = 1, \ldots, n$, are distinct and $a_{ij}(\epsilon) \to a_{ij}$ as $\epsilon \to 0$, $1 \leq j \leq i \leq n$. Then

$$\mathrm{Det}\, m_{\mathbf{a}} = \lim_{\epsilon \to 0} \mathrm{Det}\, m_{\mathbf{a}_\epsilon} = \lim_{\epsilon \to 0} \prod_{i=1}^{n} (a_{ii}(\epsilon))^i = \prod_{i=1}^{n} a_{ii}^{i}.$$

$\square$

Lemma 1.5. *For non-singular $\mathbf{a} = [a_{ij}] \in \mathcal{LT}_n$ let $s_{\mathbf{a}} : \mathcal{LT}_n \to \mathcal{LT}_n$ be an endomorphism defined by*

$$s_{\mathbf{a}}(\mathbf{x}) = \mathbf{a}\mathbf{x}^* + \mathbf{x}\mathbf{a}^*.$$

Then

$$\mathrm{Det}\, s_{\mathbf{a}} = 2^n \prod_{i=1}^{n} a_{ii}^{n-i+1}.$$

Proof. Note that

$$\mathbf{a}\mathbf{x}^* + \mathbf{x}\mathbf{a}^* = \mathbf{a}\left(\mathbf{x}^*(\mathbf{x}^{-1})^* + \mathbf{a}^{-1}\mathbf{x}\right)\mathbf{a}^* = \mathbf{a}\left((\mathbf{a}^{-1}\mathbf{x})^* + \mathbf{a}^{-1}\mathbf{x}\right)\mathbf{a}^*$$

Therefore,

$$s_{\mathbf{a}} = \mathbb{P}(\mathbf{a}) \circ v \circ m_{\mathbf{a}^{-1}}.$$

Consequently,

$$\mathrm{Det}\, s_{\mathbf{a}} = \mathrm{Det}\,\mathbb{P}(\mathbf{a})\,\mathrm{Det}\,v\,\mathrm{Det}\,m_{\mathbf{a}^{-1}} = |\mathbf{a}|^{n+1}\, 2^n \prod_{i=1}^{n} a_{ii}^{-i}$$

and the result follows since $|\mathbf{a}| = \prod_{i=1}^{n} a_{ii}$. $\square$

1.2. **Proof of the Bartlett decomposition theorem.** Note that the Cholesky decomposition of a matrix $\mathbf{X} \in \mathcal{V}_n$ is a bijection between $\mathcal{V}_n$ and $\mathcal{LT}_n \cong \mathcal{V}_n$. Therefore if $\mathbf{X}$ is a Wishart random matrix then the random matrix $\mathbf{T}$ satisfying $\mathbf{X} = \mathbf{T}\mathbf{T}^*$ has a density of the form

$$f_{\mathbf{T}}(\mathbf{t}) = J\ \ (\mathbf{t})\, f_{\mathbf{X}}(\mathbf{t}\mathbf{t}^*). \tag{7}$$

That is we need to find the Jacobian of the transformation $\quad : \mathcal{LT}_n \to \mathcal{V}_n \cong \mathcal{LT}_n$ defined by

$$\ \ (\mathbf{t}) = \mathbf{t}\mathbf{t}^*.$$

Note that its derivative $D\ \ (\mathbf{t})$ is an endomorphism of $\mathcal{LT}_n$ of the form

$$D\ \ (\mathbf{t})\mathbf{h} = \mathbf{t}\mathbf{h}^* + \mathbf{h}\mathbf{t}^* = s_{\mathbf{t}}(\mathbf{h}), \quad \mathbf{h} \in \mathcal{LT}_n.$$

Therefore, by Lemma 1.5,

$$J\ \ (\mathbf{t}) = \mathrm{Det}\, D\ \ (\mathbf{t}) = \mathrm{Det}\, s_{\mathbf{t}} = 2^n \prod_{i=1}^{n} t_{ii}^{n-i+1}. \tag{8}$$

Inserting the above into (7) we get

$$f_{\mathbf{T}}(\mathbf{t}) \propto \prod_{i=1}^{n} t_{ii}^{n-i+1} |\mathbf{t}\mathbf{t}^*|^{p-\frac{n+1}{2}}\, e^{-\frac{1}{2}\mathbf{t}\mathbf{t}^*}.$$

Since

$$|\mathbf{t}\mathbf{t}^*| = \prod_{i=1}^{n} t_{ii}^2$$

we can write

$$f_{\mathbf{T}}(\mathbf{t}) \propto \prod_{i=1}^{n} t_{ii}^{2p-i}\, e^{-\frac{t_{ii}^2}{2}} \prod_{1 \le j < i \le n} e^{-\frac{t_{ij}^2}{2}}.$$

Consequently, the random variables T_{ii}, T_{ij}, $1 \le j \le i \le n$, are independent, $T_{ij} \sim \mathrm{N}(0,1)$, $1 \le j \le i \le n$, and the density of T_{ii}, $1 \le i \le j$ has the form

$$f_{T_{ii}}(x) \propto x^{2p-i} e^{-\frac{x^2}{2}}.$$

Therefore, the density of T_{ii}^2 has the form

$$f_{T_{ii}^2}(y) \propto (\sqrt{y})^{2p-i}\, e^{-\frac{y}{2}}\, \frac{1}{\sqrt{y}} = y^{\frac{2p-i+1}{2}-1} e^{-\frac{y}{2}}.$$

1.3. **Multivariate gamma function.** The multivariate gamma function Γ_n is defined by

$$\Gamma_n(a) = \int_{\Omega_n} e^{-\operatorname{tr}\mathbf{x}}\,|\mathbf{x}|^{a-\frac{n+1}{2}}\,\mathrm{d}\mathbf{x}, \quad a > \tfrac{n-1}{2}. \tag{9}$$

Proposition 1.6.

$$\Gamma_n(a) = \pi^{\frac{n(n-1)}{4}} \prod_{i=1}^{n} \Gamma\left(a - \tfrac{i-1}{2}\right). \tag{10}$$

Proof. In the integral from the definition (9) we make the change of variable by the Cholesky decomposition $\mathbf{x} = \mathbf{t}\mathbf{t}^*$, where $\mathbf{t} = [t_{ij}] \in \mathcal{LT}_n^+$,where the superscript $+$ denotes that $t_{ii} > 0$, $i = 1, \ldots, n$. Using the form of the Jacobian as given in (8) we obtain

$$\Gamma_n(a) = 2^n \int_{\mathcal{LT}_n^+} \left.\prod_{i=1}^{n} t_{ii}^{n-i+1}\right)\, e^{-\sum_{1\le j\le i\le n} t_{ij}^2} \prod_{i=1}^{n} t_{ii}^{2a-n-1}\,\mathrm{d}\mathbf{t}$$

$$= \left.\prod_{i=1}^{n} \int_0^\infty \left(t_{ii}^2\right)^{a-\frac{i+1}{2}} e^{-t_{ii}^2}\, 2t_{ii}\mathrm{d}t_{ii}\right)\left(\prod_{1\le j<i\le n} \int_{-\infty}^{\infty} e^{-t_{ij}^2}\,\mathrm{d}t_{ij}\right).$$

The result follows by elementary integrals

$$\int_0^\infty \left(t_{ii}^2\right)^{a-\frac{i-1}{2}-1} e^{-t_{ii}^2}\,\mathrm{d}t_{ii}^2 = \Gamma\left(a - \tfrac{i-1}{2}\right)$$

and

$$\int_{-\infty}^{\infty} e^{-t_{ij}^2}\,\mathrm{d}t_{ij} = \sqrt{\pi}.$$

□

1.4. **Applications for asymptotics of Wishart determinants.**

Theorem 1.7. *Let* $\mathbf{X}_n \sim \mathrm{W}_n(n/2, \mathbf{\Sigma}_n)$, $n \ge 1$. *Then*

$$\left(\tfrac{|\mathbf{X}_n|}{|\mathbf{\Sigma}_n|\,(n-1)!}\right)^{\frac{1}{\sqrt{2\log n}}} \xrightarrow{d} e^N, \tag{11}$$

where $N \sim \mathrm{N}(0,1)$.

Proof. Define

$$\mathbf{Y}_n = \mathbf{\Sigma}_n^{-1/2}\mathbf{X}_n\mathbf{\Sigma}_n^{-1/2} \sim \mathrm{W}_n(n/2, \mathbf{I}_n).$$

By the Bartlett theorem we conclude that

$$|\mathbf{Y}_n| = \prod_{i=1}^{n} T_{n,ii}^2,$$

where $\mathbf{Y}_n = \mathbf{T}_n\mathbf{T}_n^*$ and $T_{n,ii}^2 \sim \chi^2(n-i+1)$.

Consequently, for row-wise iid double array $(Z_{kj})_{j=1,\ldots,k}$, where $Z_{11} \sim \chi^2(1)$ we have

$$|\mathbf{Y}_n| = \prod_{k=1}^{n} S_k,$$

where $S_k = \sum_{k=1}^{i} Z_{ki} \sim \chi^2(k)$.

Now we will use the following result on asymptotic distribution of products of independent sums from Remapala and Wesolowski (2005) (important improvements regarding the moments assumption in Kosiński, 2009).

Theorem 1.8. *Let $(X_{ki})_{i=1,\dots,k}$, $k=1,2,\dots$, be a double array of iid p-integrable, $p>2$, random variables with $\mathbb{E}\,X_{ki}=\mu>0$, $\mathbb{V}\mathrm{ar}\,X_{ki}=\sigma^2>0$ and $\gamma=\sigma/\mu$. Then for $S_k=X_{k1}+\ldots+X_{kk}$, $k\geq 1$,*

$$\left(n^{\frac{\gamma^2}{2}}\frac{\prod_{k=1}^n S_k}{n!\,\mu^n}\right)^{\frac{1}{\gamma\sqrt{\log\, n}}}\xrightarrow{d} e^N, \tag{12}$$

where $N\sim \mathrm{N}(0,1)$.

Since $\mathbb{E}\,Z_{kj}=\mu=1$ and $\mathbb{V}\mathrm{ar}\,Z_{kj}=\sigma^2=2$ then $\gamma^2=2$ and by Theorem 1.8 we get

$$\left(n\frac{|\mathbf{Y}_n|}{n!}\right)^{\frac{1}{\gamma\sqrt{\log\, n}}}\xrightarrow{d} e^N.$$

Now the result follows since $|\mathbf{Y}_n|=\frac{|\mathbf{X}_n|}{|\mathbf{\Sigma}_n|}$.

□

The above asymptotics should be compared to Anderson (1958) result:

Theorem 1.9. *If $\mathbf{X}_n\sim \mathrm{W}_m(n/2,\frac{1}{n}\mathbf{\Sigma})$ then*

$$\left(\frac{|\mathbf{X}_n|}{|\mathbf{\Sigma}|}\right)^{\sqrt{\frac{n}{2m}}}\xrightarrow{d} e^N,$$

where $N\sim\mathrm{N}(0,1)$.

Proof. By the Bartlett theorem

$$n^m\frac{|\mathbf{X}_n|}{|\mathbf{\Sigma}|}=\prod_{i=1}^m T^2_{n,ii},$$

where $T^2_{n,ii}\sim\chi^2(n-i+1)$, $i=1,\dots,m$, are independent. Consequently,

$$\log\,\frac{|\mathbf{X}_n|}{|\mathbf{\Sigma}|}=\sum_{i=1}^m\left(T^2_{n,ii}-\log\, n\right),$$

We will prove that

$$\sqrt{\tfrac{n}{2}}\left(T^2_{n,ii}-\log\, n\right)\xrightarrow{d} N_i, \tag{13}$$

where $N_i\sim\mathrm{N}(0,1)$, $i=1,\dots,m$, are iid. To prove (13) we first note that

$$\sqrt{\tfrac{n}{2}}\left(T^2_{n,ii}-\log\, n\right)=\sqrt{\tfrac{n}{n-i+1}}\sqrt{\tfrac{n-i+1}{2}}\left(T^2_{n,ii}-\log(n-i+1)\right)-\sqrt{\tfrac{n}{2}}\log\left(1-\tfrac{i-1}{n}\right).$$

Since $\sqrt{\frac{n}{n-i+1}}\to 1$ and

$$\sqrt{n}\log\left(1-\tfrac{i-1}{n}\right)=\tfrac{1}{\sqrt{n}}\log\,\left(1-\tfrac{i-1}{n}\right)^n\to 0$$

it suffices to prove the result for $i=1$. Then by the δ-method:

$$\sqrt{n}\left(g(\bar{X}_n)-g(\mathbb{E}\,X)\right)\xrightarrow{d}|g'(\mathbb{E}\,X)|\sqrt{\mathbb{V}\mathrm{ar}\,X}N$$

we have with $g=\log$:

$$\sqrt{n}\log\,\bar{X}_n\xrightarrow{d}\sqrt{2}N,$$

since $X\sim\chi^2(1)$ and thus $\mathbb{E}\,X=1$, $\mathbb{V}\mathrm{ar}\,X=2$.

Consequently

$$\sqrt{\tfrac{n}{2}}\log\,\frac{|\mathbf{X}_n|}{|\mathbf{\Sigma}|}\sim\sum_{i=1}^m N_i\sim\mathrm{N}(0,m).$$

□

2. Generalized Bartlett's decomposition: Independencies of blocks of the Wishart matrix

Note that for any $n \times n$ matrix $\mathbf{A}$ with the block decomposition

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

according to dimensions r and s, $r+s=n$ and such that $\mathbf{A}_{11}$ is invertible

$$\mathbf{A} = \begin{bmatrix} \mathbf{I}_r & \mathbf{0}_{rs} \\ \mathbf{A}_{21}\left(\mathbf{A}_{11}\right)^{-1} & \mathbf{I}_s \end{bmatrix} \begin{bmatrix} \mathbf{A}_{11} & \mathbf{0}_{rs} \\ \mathbf{0}_{sr} & \mathbf{A}_{2\cdot 1} \end{bmatrix} \begin{bmatrix} \mathbf{I}_r & \left(\mathbf{A}_{11}\right)^{-1}\mathbf{A}_{12} \\ \mathbf{0}_{sr} & \mathbf{I}_s \end{bmatrix},$$

where $\mathbf{A}_{2\cdot 1} = \mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}$ is called the Schur complement.

This decomposition gives immediately the following decomposition of the determinant

$$|\mathbf{A}| = |\mathbf{A}_{11}|\,|\mathbf{A}_{2\cdot 1}|. \tag{14}$$

For a non-singular matrix $\mathbf{A}$ the block decomposition of its inverse has the form

$$\mathbf{A}^{-1} = \begin{bmatrix} \left(\mathbf{A}^{-1}\right)_{11} & \left(\mathbf{A}^{-1}\right)_{12} \\ \left(\mathbf{A}^{-1}\right)_{21} & \left(\mathbf{A}^{-1}\right)_{22} \end{bmatrix} = \begin{bmatrix} \left(\mathbf{A}_{1\cdot 2}\right)^{-1} & -\left(\mathbf{A}_{11}\right)^{-1}\mathbf{A}_{12}\left(\mathbf{A}_{2\cdot 1}\right)^{-1} \\ -\left(\mathbf{A}_{22}\right)^{-1}\mathbf{A}_{21}\left(\mathbf{A}_{1\cdot 2}\right)^{-1} & \left(\mathbf{A}_{2\cdot 1}\right)^{-1} \end{bmatrix}. \tag{15}$$

If $\mathbf{A}$ is a symmetric matrix, then $\mathbf{A}^{-1}$ is also symmetric and thus (14) gives

$$\left(\mathbf{A}_{11}\right)^{-1}\mathbf{A}_{12}\left(\mathbf{A}_{2\cdot 1}\right)^{-1} = \left(\mathbf{A}_{1\cdot 2}\right)^{-1}\mathbf{A}_{12}\left(\mathbf{A}_{22}\right)^{-1}. \tag{16}$$

Now we are ready to state the main result, see Muirhead (1994).

Theorem 2.1. *Let* $\mathbf{X}$ *be a Wishart matrix* $\mathrm{W}_n(p,\mathbf{\Sigma})$, $p > \frac{n-1}{2}$, $\mathbf{\Sigma} \in \Omega$. *Let*

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_{11} & \mathbf{X}_{12} \\ \mathbf{X}_{21} & \mathbf{X}_{22} \end{bmatrix} \quad \textit{and} \quad \mathbf{\Sigma} = \begin{bmatrix} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} \\ \mathbf{\Sigma}_{21} & \mathbf{\Sigma}_{22} \end{bmatrix}$$

be block decomposition of $\mathbf{X}$ *and* $\mathbf{\Sigma}$ *according to dimensions* r *and* s, $r+s=n$, *respectively.*

Then

$$\mathbf{X}_{22} \sim \mathrm{W}_s(p,\mathbf{\Sigma}_{22}); \tag{17}$$

$$\mathbf{X}_{12}|\mathbf{X}_{22} \sim \mathrm{N}_{r\times s}\left(\mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}\mathbf{X}_{22},\, \mathbf{\Sigma}_{1\cdot 2}\otimes\mathbf{X}_{22}\right); \tag{18}$$

$$\mathbf{X}_{1\cdot 2}|\left(\mathbf{X}_{12},\, \mathbf{X}_{22}\right) \sim \mathrm{W}_r\left(p-\tfrac{s}{2},\mathbf{\Sigma}_{1\cdot 2}\right); \tag{19}$$

Proof. Note that the Jacobian of the transformation $(\mathbf{x}_{11},\mathbf{x}_{12},\mathbf{x}_{22}) \to (\mathbf{x}_{1\cdot 2},\,\mathbf{x}_{12},\,\mathbf{x}_{22})$ is equal 1. Therefore the density of $(\mathbf{X}_{1\cdot 2},\mathbf{X}_{12},\mathbf{X}_{22})$ has the form

$$f(\mathbf{x}_{1\cdot 2},\mathbf{x}_{12},\mathbf{x}_{22}) \propto f_{\mathrm{W}_n(p,\mathbf{\Sigma})}(\mathbf{x}_{1\cdot 2}+\mathbf{x}_{12}\mathbf{x}_{22}^{-1}\mathbf{x}_{21},\,\mathbf{x}_{12},\mathbf{x}_{22}). \tag{20}$$

Note that

$$\mathrm{tr}(\mathbf{\Sigma}^{-1}\mathbf{x}) = \mathrm{tr}\begin{bmatrix} \left(\mathbf{\Sigma}^{-1}\right)_{11} & \left(\mathbf{\Sigma}^{-1}\right)_{12} \\ \left(\mathbf{\Sigma}^{-1}\right)_{21} & \left(\mathbf{\Sigma}^{-1}\right)_{22} \end{bmatrix} \begin{bmatrix} \mathbf{x}_{1\cdot 2}+\mathbf{x}_{12}\mathbf{x}_{22}^{-1}\mathbf{x}_{21} & \mathbf{x}_{12} \\ \mathbf{x}_{21} & \mathbf{x}_{22} \end{bmatrix}$$

$$= \mathrm{tr}\left\{\left(\mathbf{\Sigma}^{-1}\right)_{11}\mathbf{x}_{1\cdot 2}\right\} + \mathrm{tr}\left\{\left(\mathbf{\Sigma}^{-1}\right)_{11}\mathbf{x}_{12}\mathbf{x}_{22}^{-1}\mathbf{x}_{21}\right\} + 2\mathrm{tr}\left\{\left(\mathbf{\Sigma}^{-1}\right)_{12}\mathbf{x}_{21}\right\} + \mathrm{tr}\left\{\left(\mathbf{\Sigma}^{-1}\right)_{22}\mathbf{x}_{22}\right\} \tag{21}$$

Hence

$$\mathrm{tr}(\mathbf{\Sigma}^{-1}\mathbf{x}) = \mathrm{tr}\left\{\left(\mathbf{\Sigma}^{-1}\right)_{11}\left[\mathbf{x}_{12}+\left(\left(\mathbf{\Sigma}^{-1}\right)_{11}\right)^{-1}\left(\mathbf{\Sigma}^{-1}\right)_{12}\mathbf{x}_{22}\right]\mathbf{x}_{22}^{-1}\left[\mathbf{x}_{12}+\left(\left(\mathbf{\Sigma}^{-1}\right)_{11}\right)^{-1}\left(\mathbf{\Sigma}^{-1}\right)_{12}\mathbf{x}_{22}\right]^*\right\}$$
$$+\mathrm{tr}\left\{\left[\left(\mathbf{\Sigma}^{-1}\right)_{22}-\left(\mathbf{\Sigma}^{-1}\right)_{21}\left(\left(\mathbf{\Sigma}^{-1}\right)_{11}\right)^{-1}\left(\mathbf{\Sigma}^{-1}\right)_{12}\right]\mathbf{x}_{22}\right\} + \mathrm{tr}\left\{\left(\mathbf{\Sigma}^{-1}\right)_{11}\mathbf{x}_{1\cdot 2}\right\}.$$

Since, by (15),

$$\left(\mathbf{\Sigma}^{-1}\right)_{11} = \left(\mathbf{\Sigma}_{1\cdot 2}\right)^{-1}$$

$$\left(\mathbf{\Sigma}^{-1}\right)_{22} - \left(\mathbf{\Sigma}^{-1}\right)_{21}\left(\left(\mathbf{\Sigma}^{-1}\right)_{11}\right)^{-1}\left(\mathbf{\Sigma}^{-1}\right)_{12} = \left(\mathbf{\Sigma}_{22}\right)^{-1}$$

and

$$\left(\left(\mathbf{\Sigma}^{-1}\right)_{11}\right)^{-1}\left(\mathbf{\Sigma}^{-1}\right)_{12} = -\mathbf{\Sigma}_{12}\left(\mathbf{\Sigma}_{22}\right)^{-1},$$

we see that

$$\operatorname{tr}(\mathbf{\Sigma}^{-1}\mathbf{x}) = \operatorname{tr}\left\{\left(\mathbf{\Sigma}_{1\cdot 2}\right)^{-1}\left[\mathbf{x}_{12} - \mathbf{\Sigma}_{12}\left(\mathbf{\Sigma}_{22}\right)^{-1}\mathbf{x}_{22}\right]\mathbf{x}_{22}^{-1}\left[\mathbf{x}_{12} - \mathbf{\Sigma}_{12}\left(\mathbf{\Sigma}_{22}\right)^{-1}\mathbf{x}_{22}\right]^{*}\right\}$$

$$+\operatorname{tr}\left\{\left(\mathbf{\Sigma}_{22}\right)^{-1}\mathbf{x}_{22}\right\} + \operatorname{tr}\left\{\left(\mathbf{\Sigma}_{1\cdot 2}\right)^{-1}\mathbf{x}_{1\cdot 2}\right\}. \tag{22}$$

Moreover, (14) gives

$$|\mathbf{x}| = |\mathbf{x}_{1\cdot 2}|\,|\mathbf{x}_{22}| \quad \text{and} \quad |\mathbf{\Sigma}| = |\mathbf{\Sigma}_{1\cdot 2}|\,|\mathbf{\Sigma}_{22}|. \tag{23}$$

Plugging (22) and (23) into (20) we see that

$$f(\mathbf{x}_{1\cdot 2}, \mathbf{x}_{12}, \mathbf{x}_{22}) \propto |\mathbf{x}_{1\cdot 2}|^{p-\frac{s}{2}-\frac{r+1}{2}}\, e^{-\frac{1}{2}\operatorname{tr}\left\{\left(\mathbf{\Sigma}_{1\cdot 2}\right)^{-1}\mathbf{x}_{1\cdot 2}\right\}}\cdot$$
$$\cdot|\mathbf{x}_{22}|^{p-\frac{s+1}{2}}\, e^{-\frac{1}{2}\operatorname{tr}\left\{\left(\mathbf{\Sigma}_{22}\right)^{-1}\mathbf{x}_{22}\right\}}\cdot$$
$$\cdot\,|\mathbf{x}_{22}|^{-\frac{r}{2}}\, e^{-\frac{1}{2}\operatorname{tr}\left\{\left(\mathbf{\Sigma}_{1\cdot 2}\right)^{-1}\left[\mathbf{x}_{12}-\mathbf{\Sigma}_{12}\left(\mathbf{\Sigma}_{22}\right)^{-1}\mathbf{x}_{22}\right]\mathbf{x}_{22}^{-1}\left[\mathbf{x}_{12}-\mathbf{\Sigma}_{12}\left(\mathbf{\Sigma}_{22}\right)^{-1}\mathbf{x}_{22}\right]^{*}\right\}}.$$

We recall that an $r\times s$ matrix $\mathbf{Y}$ has the matrix variate normal distribution $\mathrm{N}_{r\times s}(\mathbf{m},\, \mathbf{C}\otimes\mathbf{D})$, where $\mathbf{m}\in\mathcal{M}_{r\times s}$, $\mathbf{C}\in\Omega_r$ and $\mathbf{D}\in\Omega_s$ if its density is of the form

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{1}{(2\pi)^{\frac{rs}{2}}\,|\mathbf{C}|^{\frac{s}{2}}\,|\mathbf{D}|^{\frac{r}{2}}}\, e^{-\frac{1}{2}\operatorname{tr}\left(\mathbf{C}^{-1}(\mathbf{y}-\mathbf{m})\mathbf{B}^{-1}(\mathbf{y}-\mathbf{m})^{*}\right)}, \quad \mathbf{y}\in\mathcal{M}_{r\times s}.$$

Consequently, (17), (19) and (18) follow. □

2.1. **Another proof of the Bartlett decomposition.** The alternative proof of the Bartlett decomposition is by induction wrt to the dimension n.

First we note that for $\mathbf{X}\sim\mathrm{W}_n(p,\mathbf{I}_n)$ Theorem 2.1 yields

$$\mathbf{X}_{22}\sim\mathrm{W}_s(p,\mathbf{I}_s),$$
$$\mathbf{X}_{1\cdot 2}|\left(\mathbf{X}_{12},\, \mathbf{X}_{22}\right)\sim\mathrm{W}_r\left(p-\tfrac{s}{2},\mathbf{I}_r\right)$$

and

$$\mathbf{X}_{12}|\mathbf{X}_{22}\sim\mathrm{N}_{r\times s}\left(\mathbf{0}_{r\times s},\, \mathbf{I}_r\otimes\mathbf{X}_{22}\right). \tag{24}$$

Lemma 2.2. *Let $\mathbf{a}\in\Omega_q$. The determinant of the linear map $L_{\mathbf{a}}:\mathcal{M}_{p\times q}\to\mathcal{M}_{p\times q}$ defined by*

$$L_{\mathbf{a}}(\mathbf{x}) = \mathbf{x}\mathbf{a}$$

has the form

$$\operatorname{Det}(L_{\mathbf{a}}) = |\mathbf{a}|^{p}.$$

Proof. Consider first $\mathbf{a}$ which is diagonal. Then we have

$$L_{\mathbf{a}}\mathbf{F}_{ij} = a_{jj}\mathbf{F}_{ij}, \quad i=1,\dots,p,\ j=1,\dots,q.$$

That is a_{jj} is an eigenvalue of multiplicity p, $j=1,\dots,q$. Thus the result follows.

For any $\mathbf{a}\in\Omega_q$ we can write $\mathbf{a}=\mathbf{o}\mathbf{d}\mathbf{o}^{*}$, where $\mathbf{o}$ is orthogonal and $\mathbf{d}$ is diagonal. Then $L_{\mathbf{a}} = L_{\mathbf{o}}\otimes L_{\mathbf{d}}\otimes L_{\mathbf{o}^{*}}$. Consequently

$$\operatorname{Det}(L_{\mathbf{a}}) = \operatorname{Det}(L_{\mathbf{o}})\operatorname{Det}(L_{\mathbf{d}})\operatorname{Det}(L_{\mathbf{o}^{*}}).$$

Note that $\operatorname{Det}(L_{\mathbf{d}}) = |\mathbf{d}|^{p} = |\mathbf{a}|^{p}$. Moreover,

$$\operatorname{tr}(L_{\mathbf{o}}(\mathbf{x})\left(L_{\mathbf{o}}(\mathbf{x})\right)^{*}) = \operatorname{tr}\mathbf{x}\mathbf{o}\mathbf{o}^{*}\mathbf{x} = \operatorname{tr}\mathbf{x}\mathbf{x}^{*}.$$

Consequently, $L_{\mathbf{o}}$ is a unitary transformation and thus $\operatorname{Det}L_{\mathbf{o}}=\pm 1$. □

The above Lemma, in view of (24) implies

$$\mathbf{X}_{12}\left(\mathbf{X}_{22}\right)^{-\frac{1}{2}}|\mathbf{X}_{22} \sim \mathrm{N}_{r\times s}\left(\mathbf{0}_{r\times s},\,\mathbf{I}_r\otimes\mathbf{I}_s\right).$$

Therefore, by symmetry we also have

$$\mathbf{X}_{11} \sim \mathrm{W}_r(p,\mathbf{I}_r),$$

$$\mathbf{X}_{2\cdot 1}|\left(\mathbf{X}_{21},\,\mathbf{X}_{11}\right) \sim \mathrm{W}_s\left(p-\tfrac{r}{2},\mathbf{I}_s\right)$$

and

$$\mathbf{X}_{21}\left(\mathbf{X}_{11}\right)^{-\frac{1}{2}}|\mathbf{X}_{11} \sim \mathrm{N}_{s\times r}\left(\mathbf{0}_{s\times r},\,\mathbf{I}_s\otimes\mathbf{I}_r\right).$$

Thus

$$\left(\mathbf{X}_{11},\,\mathbf{X}_{21}\left(\mathbf{X}_{11}\right)^{-\frac{1}{2}},\,\mathbf{X}_{2\cdot 1}\right) \sim \mathrm{W}_r(p,\mathbf{I}_r)\otimes\mathrm{N}_{s\times r}\left(\mathbf{0}_{s\times r},\,\mathbf{I}_s\otimes\mathbf{I}_r\right)\otimes\mathrm{W}_s\left(p-\tfrac{r}{2},\mathbf{I}_s\right). \tag{25}$$

For $n=1$ we see that $\mathrm{W}_1(p,\mathbf{I}_1)=\chi^2(2p)$ and thus $\mathbf{X}=\mathbf{T}^2\sim\chi^2(2p)$.
For $n=2$ we have $\mathbf{X}\sim\mathrm{W}_2(p,\mathbf{I}_2)$. Then

$$\mathbf{T}=\left[\begin{array}{cc}\sqrt{X_{11}} & 0\\ \frac{X_{12}}{\sqrt{X_{11}}} & \sqrt{X_{22}-\frac{X_{12}^2}{X_{11}}}\end{array}\right].$$

Now, (25) implies that

$$t_{11}^2=X_{11}=\mathbf{X}_{11}=\sim\chi^2(2p),\quad t_{22}^2=X_{22}-\frac{X_{12}^2}{X_{11}}=\mathbf{X}_{1\cdot 2}\sim\chi^2(2p-1),\quad t_{12}=\frac{X_{12}}{\sqrt{X_{11}}}=\mathbf{X}_{21}\left(\mathbf{X}_{11}\right)^{-\frac{1}{2}}\sim\mathrm{N}(0,1)$$

and that they are independent.

Consider now $\mathbf{X}\sim\mathrm{W}_{n+1}(p,\mathbf{I}_{n+1},\,p>\frac{n}{2}$ and assume that the Bartlett decomposition holds for any Wishart matrix with distribution $\mathrm{W}_n(q,\mathbf{I}_n)$ with $q>\frac{n-1}{2}$. Consider the block decomposition of $\mathbf{X}$ according to dimensions $r=1$ and $s=n$ as above. Then $\mathbf{X}=\mathbf{T}\mathbf{T}^*$, where

$$\mathbf{T}=\left[\begin{array}{cc}\sqrt{X_{11}} & 0\\ \frac{\mathbf{X}_{12}}{\sqrt{X_{11}}} & \mathbf{T}_{2\cdot 1}\end{array}\right]\in\mathcal{LT}_{n+1},$$

with $\mathbf{T}_{2\cdot 1}\in\mathcal{LT}_n$ defined by $\mathbf{X}_{2\cdot 1}=\mathbf{T}_{2\cdot 1}\mathbf{T}_{2\cdot 1}^*$. By (25)

$$X_{11}\sim\mathrm{W}_1(p,\mathbf{I}_1),\quad \frac{\mathbf{X}_{12}}{\sqrt{X_{11}}}\sim\mathrm{N}_n(0,\mathbf{I}_n),\quad \mathbf{X}_{2\cdot 1}\sim\mathrm{W}_n\left(p-\tfrac{1}{2},\mathbf{I}_n\right)$$

are independent. Since $p-\frac{1}{2}>\frac{n-1}{2}$ by the induction assumption,

$$t_{2\cdot 1,ii}^2\sim\chi^2(2p-i),\ i=1,\dots,n,\quad t_{2\cdot 1,ij}\sim\mathrm{N}(0,1),\ 1\le i<j\le n,$$

are independent jointly. Finally, we conclude that

$$\begin{aligned}t_{11}^2=X_{11}&\sim\chi^2(2p),\\ t_{ii}^2=t_{2\cdot 1,(i-1)(i-1)}^2&\sim\chi^2(2p-i+1),\quad i=2,\dots,n+1,\\ t_{1j}=\frac{X_{1j}}{\sqrt{X_{11}}}&\sim\mathrm{N}(0,1),\quad j=2,\dots,n+1,\\ t_{ij}=t_{2\cdot 1,(i-1)(j-1)}&\sim\mathrm{N}(0,1),\quad 2\le i<j\le n+1\end{aligned}$$

are jointly independent.

3. Matsumoto-Yor property for Wishart matrices

To move deeper in the structure of Wishart matrices we need to introduce the matrix variate generalized inverse Gaussian distribution $\mathrm{GIG}_n(-q, \mathbf{A}, \mathbf{B})$ with the density

$$f(\mathbf{x}) \propto |\mathbf{x}|^{q-\frac{n+1}{2}} e^{-\frac{1}{2}\left(\mathrm{tr}\,\mathbf{A}\mathbf{x}+\mathrm{tr}\mathbf{B}\mathbf{x}^{-1}\right)} I_{\Omega_n}(\mathbf{x}. \tag{26}$$

This is a well defined density only in the following cases (Letac, 2003)

- $\mathbf{A}$, $\mathbf{B} \in \Omega_n$ and $q \in \mathbb{R}$;
- $\mathbf{A} \in \delta(\Omega_n)$ with $\mathrm{rank}\,\mathbf{A} = m \in \{0, 1, \ldots, n-1\}$, $\mathbf{B} \in \Omega_n$ and $q < -\frac{n-m-1}{2}$;
- $\mathbf{A} \in \Omega_n$, $\mathbf{B} \in \delta(\Omega_n)$ with $\mathrm{rank}\,\mathbf{B} = m \in \{0, 1, \ldots, n-1\}$ and $q > \frac{n-m-1}{2}$.

The following observation is due to Butler (1998)

Theorem 3.1. *Let* $\mathbf{X}$ *be a Wishart matrix* $\mathrm{W}_n(p, \mathbf{\Sigma})$, $p > \frac{n-1}{2}$, $\mathbf{\Sigma} \in \Omega$. *Let*

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_{11} & \mathbf{X}_{12} \\ \mathbf{X}_{21} & \mathbf{X}_{22} \end{bmatrix} \quad \textit{and} \quad \mathbf{\Sigma} = \begin{bmatrix} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} \\ \mathbf{\Sigma}_{21} & \mathbf{\Sigma}_{22} \end{bmatrix}$$

be block decomposition of $\mathbf{X}$ *and* $\mathbf{\Sigma}$ *according to dimensions* r *and* s, $r + s = n$, *respectively.*

Then

$$\mathbf{X}_{22}|\mathbf{X}_{12} \sim \mathrm{GIG}_s\left(p - \tfrac{r}{2},\, \mathbf{\Sigma}_{2\cdot 1},\, \mathbf{X}_{21}\mathbf{\Sigma}_{1\cdot 2}\mathbf{X}_{12}\right). \tag{27}$$

Proof. Using the decomposition (21) and (23) in the joint density of $\mathbf{X}$ we see that it can be written as

$$f(\mathbf{x}_{1\cdot 2}, \mathbf{x}_{12}, \mathbf{x}_{22}) \propto |\mathbf{x}_{1\cdot 2}|^{p-\frac{s}{2}-\frac{r+1}{2}} e^{-\frac{1}{2}\mathrm{tr}\left\{(\mathbf{\Sigma}_{1\cdot 2})^{-1}\mathbf{x}_{1\cdot 2}\right\}} \cdot e^{-\mathrm{tr}\left\{\left(\mathbf{\Sigma}^{-1}\right)_{12}\mathbf{x}_{21}\right\}}$$
$$|\mathbf{x}_{22}|^{p-\frac{r}{2}-\frac{s+1}{2}} e^{-\frac{1}{2}\left(\mathrm{tr}\left\{\left(\mathbf{\Sigma}^{-1}\right)_{11}\mathbf{x}_{12}\mathbf{x}_{22}^{-1}\mathbf{x}_{21}\right\}+\mathrm{tr}\left\{\left(\mathbf{\Sigma}^{-1}\right)_{22}\mathbf{x}_{22}\right\}\right)}.$$

Integrating out $\mathbf{x}_{1\cdot 2}$ we see that the conditional density

$$f_{\mathbf{X}_{22}|\mathbf{X}_{12}=\mathbf{x}_{12}}(\mathbf{x}_{22}) \propto |\mathbf{x}_{22}|^{p-\frac{r}{2}-\frac{s+1}{2}} e^{-\frac{1}{2}\left(\mathrm{tr}\left\{\mathbf{x}_{21}\mathbf{\Sigma}_{1\cdot 2}\mathbf{x}_{12}\mathbf{x}_{22}^{-1}\right\}+\mathrm{tr}\left\{\mathbf{\Sigma}_{2\cdot 1}\mathbf{x}_{22}\right\}\right)}.$$

□

Theorem 3.2. *Let* $\mathbf{X}$ *be a Wishart matrix* $\mathrm{W}_n(p, \mathbf{\Sigma})$, $p > \frac{n-1}{2}$, $\mathbf{\Sigma} \in \Omega$. *Let*

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_{11} & \mathbf{X}_{12} \\ \mathbf{X}_{21} & \mathbf{X}_{22} \end{bmatrix} \quad \textit{and} \quad \mathbf{\Sigma} = \begin{bmatrix} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} \\ \mathbf{\Sigma}_{21} & \mathbf{\Sigma}_{22} \end{bmatrix}$$

be block decomposition of $\mathbf{X}$ *and* $\mathbf{\Sigma}$ *according to dimensions* r *and* s, $r + s = n$, *respectively.*

Then the density of the conditional distribution of $(\mathbf{X}_{11},\, \mathbf{X}_{22})\,|\mathbf{X}_{12} = \mathbf{x}_{12}$ *has the form*

$$f_{(\mathbf{X}_{11},\,\mathbf{X}_{22})|\mathbf{X}_{12}=\mathbf{x}_{12}}(\mathbf{x}_{11}, \mathbf{x}_{22}) \propto |\mathbf{x}|^{p-\frac{n+1}{2}} e^{-\frac{1}{2}\mathrm{tr}(\mathbf{\Sigma}_{1\cdot 2}\mathbf{x}_{11}+\mathbf{\Sigma}_{2\cdot 1}\mathbf{x}_{22})} I_{K(\mathbf{x}_{12})}(\mathbf{x}_{11}, \mathbf{x}_{22}), \tag{28}$$

where

$$K(\mathbf{x}_{12}) = \left\{\mathbf{x} = \begin{bmatrix} \mathbf{x}_{11} & \mathbf{x}_{12} \\ \mathbf{x}_{12}^T & \mathbf{x}_{22} \end{bmatrix} :\ \mathbf{x} \in \Omega_n\right\}.$$

Proof. From Theorems 2.1 and 3.1 we conclude that if $\mathbf{X} \sim \mathrm{W}_n(p, \mathbf{\Sigma})$ then

$$(\mathbf{X}_{1\cdot 2},\, \mathbf{X}_{22})\,|\mathbf{X}_{12} \sim \mathrm{W}_r\left(p - \tfrac{s}{2}, \mathbf{\Sigma}_{1\cdot 2}\right) \otimes \mathrm{GIG}_s\left(p - \tfrac{r}{2}, \mathbf{\Sigma}_{2\cdot 1}, \mathbf{X}_{21}\mathbf{\Sigma}_{1\cdot 2}\mathbf{X}_{12}\right) \tag{29}$$

Since the jacobian of the transformation

$$K(\mathbf{x}_{12}) \ni (\mathbf{x}_{11}, \mathbf{x}_{22}) \mapsto (\mathbf{x}_{11} - \mathbf{x}_{12}\mathbf{x}_{22}^{-1}\mathbf{x}_{21},\, \mathbf{x}_{22}) = (\mathbf{x}_{1\cdot 2},\, \mathbf{x}_{22}) \in \Omega_r \times \Omega_s$$

equals 1 for $(\mathbf{x}_{11}, \mathbf{x}_{22}) \in K(\mathbf{x}_{12})$ we get

$$f_{(\mathbf{X}_{11},\,\mathbf{X}_{22})|\mathbf{X}_{12}=\mathbf{x}_{12}}(\mathbf{x}_{11}, \mathbf{x}_{22}) = f_{\mathbf{X}_{1\cdot 2}}(\mathbf{x}_{11} - \mathbf{x}_{12}\mathbf{x}_{22}^{-1}\mathbf{x}_{21})\, f_{\mathbf{X}_{22}|\mathbf{X}_{12}=\mathbf{x}_{12}}(\mathbf{x}_{22})$$
$$\propto |\mathbf{x}_{1\cdot 2}|^{p-\frac{s}{2}-\frac{r+1}{2}} e^{-\frac{1}{2}\mathrm{tr}\left\{\mathbf{\Sigma}_{1\cdot 2}\left(\mathbf{x}_{11}-\mathbf{x}_{12}\mathbf{x}_{22}^{-1}\mathbf{x}_{21}\right)\right\}} \cdot |\mathbf{x}_{22}|^{p-\frac{r}{2}-\frac{s+1}{2}} e^{-\frac{1}{2}\mathrm{tr}\left\{\mathbf{\Sigma}_{2\cdot 1}\mathbf{x}_{22}+\mathbf{x}_{21}\mathbf{\Sigma}_{1\cdot 2}\mathbf{x}_{12}\mathbf{x}_{22}^{-1}\right\}}$$

$$= |\mathbf{x}|^{p-\frac{n+1}{2}}\, e^{-\frac{1}{2}\mathrm{tr}(\mathbf{\Sigma}_{1\cdot 2}\mathbf{x}_{11}+\mathbf{\Sigma}_{2\cdot 1}\mathbf{x}_{22})}.$$

□

Note that by symmetry the dual to (29) also holds,

$$(\mathbf{X}_{11},\, \mathbf{X}_{2\cdot 1})\, |\mathbf{X}_{12} \sim \mathrm{GIG}_r\left(p-\tfrac{s}{2}, \mathbf{\Sigma}_{1\cdot 2}, \mathbf{X}_{12}\mathbf{\Sigma}_{2\cdot 1}\mathbf{X}_{21}\right) \otimes \mathrm{W}_s\left(p-\tfrac{r}{2}, \mathbf{\Sigma}_{2\cdot 1}\right). \tag{30}$$

Since

$$\mathbf{X}_{11} = \mathbf{X}_{1\cdot 2} + \mathbf{X}_{12}\mathbf{X}_{22}^{-1}\mathbf{X}_{12}$$

and

$$\mathbf{X}_{2\cdot 1} = \mathbf{X}_{22} - \mathbf{X}_{21}\left(\mathbf{X}_{1\cdot 2} + \mathbf{X}_{12}\mathbf{X}_{22}^{-1}\mathbf{X}_{12}\right)^{-1}\mathbf{X}_{12}$$

we conclude that (29) $\Leftrightarrow$ (30) (without assumption $\mathbf{X} \sim \mathrm{W}_n(p, \mathbf{\Sigma})$).

Consider now a special case of $r = s = 1$ and conditioning on $X_{12} = 1$ and denote $a = \mathbf{\Sigma}_{1\cdot 2}$, $b = \mathbf{\Sigma}_{2\cdot 1}$, $q = p - \frac{1}{2}$. Then from the equivalence (29) $\Leftrightarrow$ (30) we get

$$\left(K_1, K_2 - \tfrac{1}{K_1}\right) \sim \mathrm{GIG}(q,a,b) \otimes \mathrm{G}(q,b) \quad \Leftrightarrow \quad \left(K_1 - \tfrac{1}{K_2}, K_2\right) \sim \mathrm{G}(q,a) \otimes \mathrm{GIG}(q,b,a).$$

Further let us denote $X = 1/K_1$ and $Y = K_2 - 1/K_1$ then from the above we conclude that

$$(X,Y) \sim \mathrm{GIG}(-q,b,a) \otimes \mathrm{G}(q,b) \quad \Leftarrow \left(\tfrac{1}{X+Y},\, \tfrac{1}{X} - \tfrac{1}{X+Y}\right) \sim \mathrm{GIG}(-q,a,b) \otimes \mathrm{G}(q,a) \tag{31}$$

which is known as the Matsumoto-Yor property. It was observed in the study of conditional structure of the exponential Brownian motion in Matsumoto and Yor (2001). Later on, in Matsumoto and Yor (2003) it was related to hitting times of the Brownian motion:

Let B be a Brownian motion, $a, b > 0$. Let

$$\tau_b^a(B) = \inf\{t > 0 :\, B_t + at = b\} \quad \text{and} \quad \sigma_b^a(B) = \sup\{t > 0 :\, B_t + at = b\}$$

Then

$$(X,Y) = (\tau_b^a(B),\, \sigma_b^a(B) - \tau_b^a(B)) \sim \mathrm{GIG}(\tfrac{1}{2}, a^2, b^2) \otimes \mathrm{G}(\tfrac{1}{2}, a^2).$$

Define $\tilde{B}_t = -tB_{1/t}$, $t > 0$ and $\tilde{B}_0 = 0$. Then, by the previous observation

$$(\tau_a^b(\tilde{B}),\, \sigma_a^b(\tilde{B}) - \tau_a^b(\tilde{B})) \sim \mathrm{GIG}(\tfrac{1}{2}, b^2, a^2) \otimes \mathrm{G}(\tfrac{1}{2}, b^2).$$

But

$\tau_a^b(\tilde{B}) = \inf\{t > 0 :\, \tilde{B}_t + bt = a\} = \inf\{t > 0 :\, -tB_{1/t} + bt = a\} = \inf\{t > 0 : B_{1/t} + a/t = b\} = \frac{1}{\sup\{t>0:\, B_t + at = b\}}$.

Thus

$$\tau_a^b(\tilde{B}) = \tfrac{1}{\sigma_b^a(B)} = \tfrac{1}{X+Y}.$$

Similarly,

$$\sigma_a^b(\tilde{B}) - \tau_a^b(\tilde{B}) = \tfrac{1}{\tau_b^a(B)} - \tfrac{1}{\sigma_b^a(B)} = \tfrac{1}{X} - \tfrac{1}{X+Y}.$$

The equivalence (29) $\Leftrightarrow$ (30) implies the following simple fact

Proposition 3.3. *Let $\mathbf{z} \in \mathcal{M}_{r\times s}$, $\mathbf{a} \in \Omega_r$, $\mathbf{b} \in \Omega_s$ and $p - \frac{r-1}{2} = q - \frac{s-1}{2} > 0$. For $\mathbf{x} \in \Omega_r$ and $\mathbf{y} \in \Omega_s$ define*

$$(\mathbf{x}, \mathbf{y}) = \left((\mathbf{z}^*\mathbf{x}\mathbf{z} + \mathbf{y})^{-1},\, \mathbf{x}^{-1} - \mathbf{z}\,(\mathbf{z}^*\mathbf{x}\mathbf{z} + \mathbf{y})^{-1}\,\mathbf{z}^*\right) \tag{32}$$

Then

$$(\mathbf{X}, \mathbf{Y}) \sim \mathrm{GIG}_r\,(-p, \mathbf{z}\mathbf{b}\mathbf{z}^*, \mathbf{a}) \otimes \mathrm{W}_s(q, b) \tag{33}$$

if and only if

$$(\mathbf{U}, \mathbf{V}) = \quad (\mathbf{X},\, \mathbf{Y}) \sim \mathrm{GIG}_s\,(-q, \mathbf{z}^*\mathbf{a}\mathbf{z}, \mathbf{b}) \otimes \mathrm{W}_r(q, a). \tag{34}$$

Note that for $r = s = 1$ and $\mathbf{z} = 1$ Proposition 3.3 is equivalent to (31).

It appears that the independencies of Proposition 3.3 give characterization of GIG and Wishart matrices as given in Massam and JW (2006).

3.1. **A characterization through MY property.** As a warm up we consider the univariate case.

Theorem 3.4 (Letac, JW, 2000)**.** *Let X and Y be real, positive, non-degenerate and independent. Assume that $U = \frac{1}{X+Y}$ and $V = \frac{1}{X} - \frac{1}{X+Y}$ are also independent. Then there exist $a, b, p > 0$ such that $X \sim \mathrm{GIG}(p, a, b)$ and $Y \sim \mathrm{G}(p, a)$.*

Proof. Note that $Y/X = V/U$. From the independencies it follows that for any $s, t < 0$ and $\alpha > 0$

(35)
$$\mathbb{E}\, Y^{\alpha} e^{sY}\, \mathbb{E}\, X^{-\alpha} e^{sX+tX^{-1}} = \mathbb{E}\,(Y/X)^{\alpha}\, e^{s(X+Y)+tX^{-1}} = \mathbb{E}\,(V/U)^{\alpha} e^{sU^{-1}+t(V+U)} = \mathbb{E}\, V^{\alpha} e^{tV}\, \mathbb{E}\, U^{-\alpha} e^{tU+sU^{-1}}.$$

Take logarithm and derivatives $\frac{\partial^2}{\partial s\, \partial t}$ of both sides above to get

$$\frac{\mathbb{E}\, X^{1-\alpha} e^{sX+tX^{-1}}\, \mathbb{E}\, X^{-1-\alpha} e^{sX+tX^{-1}}}{\left(\mathbb{E}\, X^{-\alpha} e^{sX+tX^{-1}}\right)^2} = \frac{\mathbb{E}\, U^{1-\alpha} e^{tU+sU^{-1}}\, \mathbb{E}\, U^{-1-\alpha} e^{tU+sU^{-1}}}{\left(\mathbb{E}\, U^{-\alpha} e^{tU+sU^{-1}}\right)^2}.$$

We take $\alpha = 1$ in the formula above and compare it with (35) with $\alpha = 0, 1, 2$ which gives

$$\frac{\mathbb{E}\, Y^2 e^{sY}\, \mathbb{E}\, e^{sY}}{(\mathbb{E}\, Y e^{sY})^2} = \frac{\mathbb{E}\, V^2 e^{tY}\, \mathbb{E}\, e^{tY}}{(\mathbb{E}\, Y e^{tY})^2} = p + 1 > 1.$$

Consequently

$$L_Y''(s) L_Y(s) = (1+p)(L_Y'(s))^2 \quad \text{and} \quad L_V''(t) L_V(t) = (1+p)(L_V'(t))^2.$$

Therefore, $Y \sim \mathrm{G}(p, a)$ and $V \sim \mathrm{G}(p, b)$ for some $a, b > 0$.

To identify the laws of X and U we can now use densities since if Y is absolutely continuous then $X + Y$ also has a density.

Alternatively, one can apply an identification of GIG distribution through the identity in law

$$X = \frac{1}{V+U} \stackrel{d}{=} \frac{1}{Y' + \frac{1}{Y+X}},$$

where $Y \sim \mathrm{G}(p, b)$ and $Y' \sim \mathrm{G}(p, a)$ are independent - see Letac and Seshadri (1983). □

Remark 3.1 (A problem by Marc Yor)**.** *We know that for $X \sim \mathrm{GIG}(-p, a, a)$ and $Y \sim \mathrm{G}(p, a)$ the random variable $\frac{1}{X} - \frac{1}{X+Y} = \frac{Y}{X(X+Y)} \sim \mathrm{G}(p, a)$. Alternatively, we can write*

$$ZX^2 + X \stackrel{d}{=} Z \tag{36}$$

where $Z^{-1} \sim \mathrm{G}(p, a)$.

Question: Assume that X and Z in (36) are positive independent and $Z^{-1} \sim \mathrm{G}(p, a)$. Is it true that $X \sim \mathrm{GIG}(-p, a, a)$.

Now we present an extension of Theorem 3.6 to the case of matrices of different dimensions, which will play a crucial role in the characterization of Wishart matrices by its independence structures. Here we need smooth strictly positive densities. We present a version with differentiable densities, however continuity should be sufficient - see Kolodziejek (2015).

Theorem 3.5. *Let $\mathbf{X}$ and $\mathbf{Y}$ be independent random matrices assuming values in Ω_r and Ω_s, respectively, and having strictly positive differentiable densities on Ω_r and Ω_s. Let $(\mathbf{U}, \mathbf{V}) = \ (\mathbf{X}, \mathbf{Y})$ with defined in (32). Assume that $\mathbf{U}$ and $\mathbf{V}$ are also independent. Then there exist $\mathbf{z} \in \mathcal{M}_{r\times s}$, $\mathbf{a} \in \Omega_r$, $\mathbf{b} \in \Omega_s$ and $p - \frac{r-1}{2} = q - \frac{s-1}{2} > 0$ such that (33), and thus (34), holds.*

Proof. (in the case $r = s$ and $\mathbf{z} = \mathbf{I}_r$) Then

$$(\mathbf{x}, \mathbf{y}) = ((\mathbf{x}+\mathbf{y})^{-1},\, \mathbf{x}^{-1} - (\mathbf{x}+\mathbf{y})^{-1}).$$

Note that $\ = \ {}^{-1}$, that is is an involution. To find its Jacobian we write it as $\ = \phi_2 \circ \phi_1$, where

$$\phi_1(\mathbf{x}, \mathbf{y}) = (\mathbf{x}, \mathbf{x}+\mathbf{y}) \quad \text{and} \quad \phi_2(\mathbf{a}, \mathbf{b}) = (\mathbf{b}^{-1}, \mathbf{a}^{-1} - \mathbf{b}^{-1}).$$

That is

$$J_{\ {}^{-1}}(\mathbf{u}, \mathbf{v}) = J_{\phi_1}(\mathbf{u}, \mathbf{v})\, J_{\phi_2}(\mathbf{u}, \mathbf{u}+\mathbf{v}).$$

Note that $J_{\phi_1} \equiv 1$. Since the derivative of $k : \Omega_r \to \Omega_r$ defined by $k(\mathbf{x}) = \mathbf{x}^{-1}$ has the form

$$D\, k(\mathbf{x})\mathbf{h} = -\mathbf{x}^{-1}\mathbf{h}\mathbf{x}^{-1},$$

therefore $D\, k(\mathbf{x}) = -\mathbb{P}(\mathbf{x}^{-1})$. Thus by Lemma 1.2 it follows that $\mathrm{Det}(D\, k(\mathbf{x})) = \pm|\mathbf{x}|^{-r-1}$. Consequently,

$$J_{\ {}^{-1}}(\mathbf{u}, \mathbf{v}) = J_{\phi_2}(\mathbf{u}, \mathbf{u}+\mathbf{v}) = (|\mathbf{u}|\, |(\mathbf{u}+\mathbf{v})|)^{-r-1}.$$

Consequently, the independence property and the smoothness assumption yield

$$f_{\mathbf{U}}(\mathbf{u})\, f_{\mathbf{V}}(\mathbf{v}) = (|\mathbf{u}|\, |(\mathbf{u}+\mathbf{v})|)^{-r-1}\ f_{\mathbf{X}}\left((\mathbf{u}+\mathbf{v})^{-1}\right)\ f_{\mathbf{Y}}\left(\mathbf{u}^{-1} - (\mathbf{u}+\mathbf{v})^{-1}\right)$$

for any $\mathbf{u}, \mathbf{v} \in \Omega_r$. Upon taking logarithms the above equation can be rewritten as

$$\phi_1(\mathbf{u}) + \phi_2(\mathbf{v}) = \phi_3(\mathbf{u}+\mathbf{v}) + \phi_4(\mathbf{u}^{-1} - (\mathbf{u}+\mathbf{v})^{-1}) \tag{37}$$

with

$$\phi_1(\mathbf{u}) = \log\, f_{\mathbf{U}}(\mathbf{u}) + (r+1)\log|\mathbf{u}|, \quad \phi_2 = \log\, f_{\mathbf{V}},$$

$$\phi_3(\mathbf{u}) = \log\, f_{\mathbf{X}}(\mathbf{u}^{-1}) - (r+1)\log|\mathbf{u}| \quad \text{and} \quad \phi_4 = \log\, f_{\mathbf{Y}}.$$

Since the derivative $D(\mathbf{x}^{-1} = -\mathbb{P}(\mathbf{x}^{-1})$ then differentiating (37) wrt $\mathbf{u}$ and $\mathbf{v}$ separately we obtain two equations

$$\phi_1'(\mathbf{u}) = \phi_3'(\mathbf{u}+\mathbf{v}) + \left(\mathbb{P}((\mathbf{u}+\mathbf{v})^{-1}) - \mathbb{P}(\mathbf{u}^{-1})\right) \circ \phi_4'(\mathbf{u}^{-1} - (\mathbf{u}+\mathbf{v})^{-1}) \tag{38}$$

and

$$\phi_2'(\mathbf{v}) = \phi_3'(\mathbf{u}+\mathbf{v}) + \mathbb{P}\left((\mathbf{u}+\mathbf{v})^{-1}\right) \circ \phi_4'(\mathbf{u}^{-1} - (\mathbf{u}+\mathbf{v})^{-1}). \tag{39}$$

Eliminating $\phi_4'(\mathbf{u}^{-1} - (\mathbf{u}+\mathbf{v})^{-1})$ from (38) and (39) leads to

$$\mathbb{P}(\mathbf{u}) \circ \phi_1'(\mathbf{u}) - (\mathbb{P}(\mathbf{u}) - \mathbb{P}(\mathbf{u}+\mathbf{v})) \circ \phi_2'(\mathbf{v}) = \mathbb{P}(\mathbf{u}+\mathbf{v}) \circ \phi_3'(\mathbf{u}+\mathbf{v}). \tag{40}$$

Now we will use the following result

Theorem 3.6 (JW, 2002)**.** *Let $A, B : \Omega \to \mathcal{V}$ and $C : \Omega^2 \to \mathcal{V}$ and $C(\mathbf{x}, \mathbf{y}) = C(\mathbf{y}, \mathbf{x})$ for any $\mathbf{x}, \mathbf{y} \in \Omega$. Assume that*

$$A(\mathbf{u}) + (\mathbb{P}(\mathbf{u}+\mathbf{v}) - \mathbb{P}(\mathbf{u})) \circ B(\mathbf{v}) = C(\mathbf{u}, \mathbf{v}), \quad \mathbf{u},\, \mathbf{v} \in \Omega.$$

Then there exist $\mathbf{a},\, \mathbf{b} \in \mathcal{V}$ and $\lambda \in \mathbb{R}$ such that

$$A(\mathbf{u}) = \mathbf{a} + \lambda\mathbf{u} + \mathbb{P}(\mathbf{u})\mathbf{b}, \quad B(\mathbf{u}) = \mathbf{b} - \lambda\mathbf{u}^{-1}$$

and

$$C(\mathbf{u}, \mathbf{v}) = \mathbf{a} - \lambda(\mathbf{u}+\mathbf{v}) + \mathbb{P}(\mathbf{u}+\mathbf{v})\mathbf{b}.$$

Applying Theorem 3.6 to equation (40) we see that the form of A implies

$$\phi_1'(\mathbf{u}) = \mathbb{P}(\mathbf{u}^{-1})\mathbf{a} - \lambda\mathbf{u}^{-1} - \mathbf{b}$$

for some $\mathbf{a}, \mathbf{b} \in \mathcal{V}$ and $\lambda \in \mathbb{R}$. Consequently,

$$\phi_1(\mathbf{u}) = K_1 - \lambda \log|\mathbf{u}| - \operatorname{tr}\left(\mathbf{bu} + \mathbf{au}^{-1}\right),$$

since $D(\log|\mathbf{u}|) = \mathbf{u}^{-1}$ and $D(\operatorname{tr}(\mathbf{au}^{-1})) = -\mathbb{P}(\mathbf{u}^{-1})\mathbf{a}$. We write $\lambda = p - \frac{r-1}{2}$ and from the definition of ϕ_1 we see that

$$\log f_{\mathbf{U}}(\mathbf{u}) = K_1 - \left(p - \tfrac{r-1}{2}\right)\log|\mathbf{u}| - \operatorname{tr}\left(\mathbf{bu} + \mathbf{au}^{-1}\right) - (r+1)\log|\mathbf{u}|.$$

Consequently, $\mathbf{a}, \mathbf{b} \in \Omega$ and $\mathbf{U} \sim \mathrm{GIG}_r(-p, \mathbf{b}, \mathbf{a})$.

Finally, the form of B from Theorem 3.6 gives

$$\phi_2'(\mathbf{v}) = -\mathbf{b} + \lambda\mathbf{v}^{-1},$$

and thus

$$\log f_{\mathbf{V}}(\mathbf{v}) = \phi_2(\mathbf{v}) = K_2 + \left(p - \tfrac{r+1}{2}\right)|\mathbf{v}| - \operatorname{tr}\mathbf{bv}.$$

Consequently, $p > \frac{r-1}{2}$ and $\mathbf{V} \sim \mathrm{W}_r(p, \mathbf{b})$. □

3.2. **Characterization of the Wishart matrix by independencies of blocks.** Note that from the part (19) of Theorem 2.1 it follows that for any block decomposition of the Wishart matrix $\mathbf{X}$ it follows that

$$\mathbf{X}_{1\cdot 2} \quad \text{and} \quad (\mathbf{X}_{12}, \mathbf{X}_{22}) \quad \text{are independent.} \tag{41}$$

Geiger and Heckerman (*Ann. Statist.*, 2002) proved the converse result, which is a characterization of the Wishart matrix by independencies of blocks.

Theorem 3.7 (Geiger and Heckerman, 2002). *Let $\mathbf{X}$ be a random matrix assuming values in Ω_n, $n \geq 3$, having a density f. If* (41) *is satisfied for* **any** *block decomposition of $\mathbf{X}$ then $\mathbf{X}$ is a Wishart matrix.*

Their proof essentially is a solution of a system of functional equations for the density f. This is based on a very complicated method designed by Jarai (1986) which allows to gradually improve regularity properties of unknown functions satisfying certain type of functional equations. In particular, it allows to move from measurability of unknown functions to their continuous differentiability of any order functions. However, this method assumes that the functional equation is satisfied everywhere on an open set. But independence property for densities allows to write the respective functional equation only almost everywhere on Ω_n. Therefore, the above characterization can be proved by such method when we additionally assume that $f > 0$ and f continuous on Ω_n.

We will improve this result by assuming only block independencies for three pairs of block independencies and the proof will be based on the Matsumoto-Yor type characterization given in Theorem 3.5.

For $\mathbf{x} \in \Omega_n$ and $i \in \{1, \ldots, n\}$ consider the block partitioning $(\mathbf{x}_{11}^{(i)}, \mathbf{x}_{12}^{(i)}, \mathbf{x}_{22}^{(i)})$, where

$$\mathbf{x}_{11}^{(i)} = [x_{ii}] \in \Omega_1, \quad \mathbf{x}_{12}^{(i)} = [x_{ij}]_{j \neq i} = \left(\mathbf{x}_{21}^{(i)}\right)^* \in \mathcal{M}_{1\times n}, \quad \mathbf{x}_{22}^{(i)} = [x_{lm}]_{l\neq i \wedge m \neq i} \in \Omega_{n-1}.$$

For the Schur complements we will use the notation

$$\mathbf{x}_{1\cdot 2}^{(i)} = \mathbf{x}_{11}^{(i)} - \mathbf{x}_{12}^{(i)}\left(\mathbf{x}_{22}^{(i)}\right)^{-1}\mathbf{x}_{21}^{(i)}$$

and

$$\mathbf{x}_{2\cdot 1}^{(i)} = \mathbf{x}_{22}^{(i)} - \mathbf{x}_{21}^{(i)}\left(\mathbf{x}_{11}^{(i)}\right)^{-1}\mathbf{x}_{12}^{(i)}.$$

Theorem 3.8. *Let $\mathbf{X} \in \Omega_n$ be a random matrix with strictly positive differentiable density on Ω_n. If for three different values of $i \in \{1, \ldots, n\}$*

$$\mathbf{X}_{2\cdot 1}^{(i)} \quad \textit{and} \quad \left(\mathbf{X}_{11}^{(i)}, \mathbf{X}_{12}^{(i)}\right) \quad \textit{are independent} \tag{42}$$

and

$$\mathbf{X}_{1\cdot 2}^{(i)} \quad \textit{and} \quad \left(\mathbf{X}_{22}^{(i)}, \mathbf{X}_{21}^{(i)}\right) \quad \textit{are independent} \tag{43}$$

then $\mathbf{X}$ is a Wishart matrix.

Proof. Without any loss of generality we assume that the three values of i are $i = 1, 2, 3$.

Since independence of X and (Y, Z) implies conditional independence of X and Y given Z from (42) and (43) it follows that

$$\mathbf{H} := \mathbf{X}_{2\cdot 1}^{(i)} \quad \text{and} \quad \mathbf{G}^{-1} := \mathbf{X}_{11}^{(i)} \quad \text{are conditionally independent given} \quad \mathbf{X}_{12}^{(i)}$$

and

$$\mathbf{G}^{-1} - \mathbb{P}(\mathbf{X}_{12}^{(i)}) \left(\mathbb{P}(\mathbf{X}_{21}^{(i)})\mathbf{G} + \mathbf{H}\right)^{-1} = \mathbf{X}_{1\cdot 2}^{(i)} \quad \text{and} \quad \left(\mathbb{P}(\mathbf{X}_{21}^{(i)})\mathbf{G} + \mathbf{H}\right)^{-1} = \left(\mathbf{X}_{22}^{(i)}\right)^{-1}$$

$$\text{are conditionally independent given} \quad \mathbf{X}_{12}^{(i)}.$$

By Theorem 3.5 the conditional distributions of

$$\left(\mathbf{X}_{11}^{(i)}, \mathbf{X}_{2\cdot 1}^{(i)}\right) \quad \text{given} \quad \mathbf{X}_{12}^{(i)}$$

and

$$\left(\mathbf{X}_{22}^{(i)}, \mathbf{X}_{1\cdot 2}^{(i)}\right) \quad \text{given} \quad \mathbf{X}_{12}^{(i)}$$

are uniquely determined up to constants $a^{(i)}, b^{(i)}, p^{(i)}$. Denote now

$$\mathbf{X}_{-1} = (X_{23}, \ldots, X_{2n}, X_{34}, \ldots, X_{3n}, \ldots, X_{n-1,n})$$

$$\mathbf{X}_{-2} = (X_{1,3}, \ldots, X_{1n}, X_{34}, \ldots, X_{3n}, \ldots, X_{n-1,n})$$

$$\mathbf{X}_{-3} = (X_{12}, X_{14}, \ldots, X_{1n}, X_{24}, \ldots, X_{2n}, X_{45}, \ldots, X_{4n}, \ldots, X_{n-1,n}).$$

Note that $(\mathbf{X}_{12}^{(i)}, \mathbf{X}_{-i}) = [X_{lm},\, 1 \le l < m]$ the off-diagonal elements in the upper triangular part of $\mathbf{X}$ and

$$\mathbf{X}_{12}^{(1)} = (X_{12}, X_{13}, \ldots, X_{1n}),$$

$$\mathbf{X}_{12}^{(2)} = (X_{12}, X_{23}, \ldots, X_{2n}),$$

$$\mathbf{X}_{12}^{(3)} = (X_{13}, X_{23}, X_{34}, \ldots, X_{3n}).$$

By c_i denote the conditional density of $\left(\mathbf{X}_{11}^{(i)}, \mathbf{X}_{2\cdot 1}^{(i)}\right)$ given $\mathbf{X}_{12}^{(i)} = \mathbf{x}_{12}^{(i)}$, as function of the whole $\mathbf{x}$, $i = 1, 2, 3$.

Writing the joint density of $\mathbf{X}$ for each $i = 1, 2, 3$, as the product of c_i times the marginal density f_i of $\mathbf{X}_{12}^{(i)}$ we get

$$c_1(\mathbf{x}) f_1(\mathbf{x}_{12}^{(1)}) = c_2(\mathbf{x}) f_2(\mathbf{x}_{12}^{(2)}) = c_3(\mathbf{x}) f_3(\mathbf{x}_{12}^{(1)}). \tag{44}$$

In the first equality of (44) we set $\mathbf{x}_{12}^{(1)} = 0$. Then we get

$$c_1(\mathbf{x} :\, \mathbf{x}_{12}^{(1)} = \mathbf{0}) f_1(\mathbf{0}) = c_2(\mathbf{x} :\, \mathbf{x}_{12}^{(1)} = \mathbf{0}) f_2(0, x_{23}, \ldots, x_{2n}).$$

Setting $x_{12} = 0$ in the second equality of (44) we get

$$c_2(\mathbf{x} :\, x_{12} = 0) f_2(0, x_{23}, \ldots, x_{2n}) = c_3(\mathbf{x} :\, x_{12} = 0) f_3(\mathbf{x}_{12}^{(3)}).$$

Combining two above equations (using the fact that the density is non-zero everywhere we get

$$f_3(\mathbf{x}_{12}^{(3)}) = \tfrac{c_2(\mathbf{x}:\, x_{12}=0)}{c_3(\mathbf{x}:\, x_{12}=0)}\, \tfrac{c_1(\mathbf{x}:\, \mathbf{x}_{12}^{(1)}=\mathbf{0})}{c_2(\mathbf{x}:\, \mathbf{x}_{12}^{(1)}=\mathbf{0})}\, f_1(\mathbf{0}).$$

That is f_3 (i.e. the law of $\mathbf{X}_{12}^{(3)}$) is uniquely defined by c_i, $i=1,2,3$ and thus by $q_i > \frac{n-2}{2}$, $\mathbf{a}_i \in \Omega_{n-1}$, $b_i > 0$, $i=1,2,3$. Consequently, the joint distribution of $\left(\mathbf{X}_{22}^{(3)}, \mathbf{X}_{1\cdot 2}^{(3)}, \mathbf{X}_{12}^{(3)}\right)$ is also uniquely defined. It gives uniqueness of the law of $\mathbf{X}$ in terms of the parameters. They can be identified to come from a single $p > \frac{n-1}{2}$ and a single matrix $\mathbf{\Sigma} \in \Omega_n$ through three representations of the joint distribution of $\mathbf{X}$ as given in (44). Now the result follows by the fact that Wishart distribution with parameters p and $\mathbf{\Sigma}$ has the same conditional distributions of blocks. □

References

Anderson, T.W. (1958) *An Introduction to Multivariate Statistical Analysis.* Wiley, New York.

Bartlett, M.S. (1933) On the theory of statistical regression. *Proc. Royal. Soc. Edinburgh* **53**, 260-283.

Butler, R.W. (1998) Generalized inverse Gaussian distribution and its Wishart connection. *Scand. J. Statist.* **25**, 69-75.

Dermer, W.L., Olkin, I. (1951) Jacobians of certain matrix transformations useful in multivariate analysis. *Biometrika* **38**, 345-367.

Geiger, D., Heckerman, D. (1998) A characterization of the bivariate Wishart distribution. *Probab. Math. Statist.* **18**, 119-131.

Geiger, D., Heckerman, D. (2002) Parameter priors for directed acyclic graphical models and the characterization of several probability distributions. *Ann. Statist.* **30**, 1412-1440.

Kolodziejek, B. (2015) The Matsumoto-Yor property and its converse on symmetric cones. *J. Theor. Probab.*, DOI: 10.1007/s10959-015-0648-2, 1-12.

Kosiński, K. (2009) Asymptotics for sums of a function of normalized independent sums. *Statist. Probab. Lett.* **79**, 415-419.

Letac, G. (2003) Convergence de l'integrale GIG. *unpublished note*, 1-2.

Letac, G., Massam, H. (2001) The normal quasi-Wishart distribution. In: *Algebraic Methods in Statistics and Probability* AMS Contemp. Math. **287**, 231-239.

Letac, G., Seshadri, V. (1983) A characterization of the generalized inverse Gaussian distribution by continued fraction. *Z. Wahrsh. Verw. Geb.* **62**, 485-489.

Letac, G., Wesolowski, J. (2000) An independence property for the product of GIG and gamma laws. *Ann. Probab.* **28**, 1371-1383.

Massam, H., Wesolowski, J. (2006) The Matsumoto-Yor property and the structure of the Wishart distribution. *J. Multivar. Anal.* **97**, 103-123.

Matsumoto, H., Yor, M. (2001) An analogue of Pitman's $2M - X$ theorem for exponential Wiener functionals, part II: the role of the generalized inverse Gaussian laws. *Nagoya Math. J.* **162**, 65-86.

Matsumoto, H., Yor, M. (2003) Interpretation via Brownian motion of some independence properties between GIG and gamma variables.*Statist. Probab. Lett.* **61**, 253-259.

Muirhead, R.J. (1994) *Aspects of Multivariate Statistical Analysis.* Wiley, New York.

Rempala, G., Wesolowski, J. (2005) Asymptotics for products of independent sums with applications to Wishart determinants. *Statist. Probab. Lett.* **74**, 129-138.

Seshadri, V., Wesolowski, J. (2008) More on connections between Wishart and matrix GIG distributions. *Metrika* **68**, 219-232.

Wesolowski, J. (2002) The Matsumoto-Yor independence property for GIG and gamma laws, revisited. *Math. Proc. Cambridge Philos. Soc.* **133**, 153-161.

Wydzial Matematyki i Nauk Informacyjnych, Politechnika Warszawska, Warszawa, Poland
E-mail address: wesolo mini.pw.edu,pl