REVSTAT – Statistical Journal
Volume 0, Number 0, Month 0000, 000–000
https://doi.org/00.00000/revstat.v00i0.000

On the independence of linear and quadratic forms in matrix normal distribution and Wishart distribution

Authors: Jiyuan Tao ✉
– Department of Mathematics and Statistics, Loyola University Maryland,
Baltimore, Maryland 21210, U.S.A.
jtao@loyola.edu

Lingchen Kong
– Department of Applied Mathematics, Beijing Jiaotong University,
Beijing 100044, P.R. China
konglchen@126.com

Zhihua Su
– Department of Statistics, University of Florida,
Gainesville, FL 32611, U.S.A.
zhihuasu@stat.ufl.edu

Guoqiang Wang
– School of Mathematics, Physics and Statistics, Shanghai University of Engineering Science,
Shanghai 201620, P.R. China
guoq_wang@hotmail.com

Received: Month 0000 Revised: Month 0000 Accepted: Month 0000

Abstract:

- It is well-known that the Craig-Sakamoto theorem establishes the independence of two linear forms and two quadratic forms in normal variates. Replacing the random normal vectors by the random normal matrices and Wishart variates, in this paper, we investigate interconnections between the independence of linear forms, quadratic forms, trace forms in matrix normal distribution and Wishart distribution. We show that the Craig-Sakamoto theorem still establishes the independence of both two linear forms and two quadratic forms in matrix normal distribution, but it fails establishing the independence of two linear forms and two quadratic forms in Wishart variates.

Keywords:

- *Matrix normal distribution; Wishart distribution; Craig-Sakamoto theorem; Vectorization; Kronecker product.*

AMS Subject Classification:

- 62E15, 62H10.

✉ Corresponding author

1. INTRODUCTION

In statistics and matrix theory, the famous Craig-Sakamoto theorem asserts that for two $n \times n$ real symmetric matrices A and B and for a random vector $x \sim N_n(0, I_n)$, where I_n to denote an $n \times n$ identity matrix, the following are equivalent:

(a) $\det(I_n - \alpha A - \beta B) = \det(I_n - \alpha A)\det(I_n - \beta B)$ for any α, $\beta \in \mathbb{R}$.

(b) $AB = 0$.

(c) Two quadratic forms $x^T Ax$ and $x^T Bx$ are independent.

This theorem has drawn the attention of many researchers since its first appearance, the interested reader may look at Driscoll and Gundberg (1986), Driscoll and Krasnicka (1995), Ogawa and Olkin (2008) for the history of this result. Also, many different proofs of this result can be found in Driscoll and Krasnicka (1995), Li (2000), Olkin (1997), Zhang and Yi (2012). When the covariance matrix Σ is a general positive definite matrix (non-identity), the transformation $x \to \Sigma^{-\frac{1}{2}}x$, $A \to \Sigma^{\frac{1}{2}}A\Sigma^{\frac{1}{2}}$, $B \to \Sigma^{\frac{1}{2}}B\Sigma^{\frac{1}{2}}$ reduces the problem to the case that the $x's$ are independent, identically distributed standard normal random variables. The condition $AB = 0$ then becomes $A\Sigma B = 0$ (see Ogawa and Olkin (2008)).

Replacing normal random variates by Wishart random variates on symmetric cones, Letac and Massam Letac and Massam (1995) extended the Craig-Sakamoto theorem from $\mathbb{R}^n$ to simple Euclidean Jordan algebras. Recently, Tao and Wang Tao and Wang (2016) provided another proof for an extension of the Craig-Sakamoto theorem by reformulating the result in terms of rank and determinant equalities and by proving the result in each of the simple Euclidean Jordan algebras. When their results are specialized to the cone of symmetric positive semidefinite matrices, we have the following theorem.

Throughout, we use the notation tr to denote the trace of a matrix, the notation $X \succeq 0$ ($X \succ 0$) to denote that X is a symmetric positive semidefinite (positive definite) matrix, the notation I_n to denote an $n \times n$ identity matrix, and the notation $\otimes$ to denote the Kronecker product.

Theorem 1.1. (see Theorem 4.1 in Letac and Massam (1995) and Theorem 3.2 in Tao and Wang (2016)) *Let A and B be $p \times p$ symmetric matrices and let S be a random symmetric matrix with Wishart distribution $W(I_p, p, n)$* (see Definition 2.2), *where I_p is the $p \times p$ identity matrix. Then the following are equivalent:*

(a) $\mathrm{tr}(AS)$ *and* $\mathrm{tr}(BS)$ *are independent.*

(b) $AB = 0$.

(c) *The nonzero eigenvalues of $A + B$ are just the nonzero eigenvalues of A and B.*

(d) $\mathrm{rank}(A+B) = \mathrm{rank}(A) + \mathrm{rank}(B)$ *and* $\mathrm{ndet}(A+B) = \mathrm{ndet}(A)\mathrm{ndet}(B)$, *where* $\mathrm{ndet}(X)$ *denotes the products of all nonzero eigenvalues of X and* $\mathrm{rank}(X)$ *denotes the rank of X.*

(e) $P_A P_B = 0$, *where* $P_A(X) := AXA$ *for all* $n \times n$ *symmetric matrices* X.

The Craig-Sakamoto theorem also establishes the independence of two linear forms. In a recent paper Zhang (2017), Zhang provided an elegant proof for the following results.

Theorem 1.2. *Let* $x \sim N_p(\mu, \Sigma)$ *with* $\Sigma \succ 0$ *(*Σ *is a symmetric positive definite matrix), and* A *and* B *be* $p \times p$ *symmetric matrices. Then the following are equivalent:*

(a) $A\Sigma B = 0$.

(b) Ax *and* Bx *are independent.*

(c) $x^T Ax$ *and* $x^T Bx$ *are independent.*

Motivated by Theorem 1.2, as a random matrix is a generalization of a random vector and the Wishart distribution is the generalization of chi-square distribution, it is natural to ask the following questions:

- Question 1: Replacing the random normal vectors by random normal matrices, does the Craig-Sakamoto theorem still give a characterization of the independence of two linear forms/quadratic forms?
- Question 2: Replacing the random normal vector by Wishart variates, does the Craig-Sakamoto theorem still give a characterization of the independence of two linear forms/two quadratic forms?

The main objective of this paper is to answer these two questions.

2. PRELIMINARIES

In this section, we recall some concepts, properties, and results used in this paper.

Let $M_{n\times p}(\mathbb{R})$ be the space of $n \times p$ real matrices. Then the *vectorization* of a matrix $(a_{ij}) \in M_{n\times p}(\mathbb{R})$, denoted by $\mathrm{vec}(A)$, is the $np \times 1$ column vector obtained by the following way:

$$\mathrm{vec}(A) = [a_{11}, \ldots, a_{n1}, a_{12}, \ldots, a_{n2}, \ldots, a_{1p}, \ldots, a_{np}]^T.$$

It is well-known that the mapping $X \mapsto \mathrm{vec}(X)$ creates a linear isomorphism between $M_{n\times p}(\mathbb{R})$ and $\mathbb{R}^{np}$. The definition of vec with the Kronecker product immediately yields the following proposition.

Proposition 2.1. ((1.15), Page 340 and Theorem 16.21, Harville (2012)) *For matrices* A, B, *and* X *of appropriate sizes,*

(a) $\mathrm{vec}(AXB) = (B^T \otimes A)\,\mathrm{vec}(X)$.

(b) $(A \otimes B)^T = A^T \otimes B^T$.

As a random vector is a generalization of a random variable, a random matrix is a generalization of a random vector. The following definition is a random matrix with normal distribution.

Definition 2.1. (Definition 2.2.1, Van Perlo-ten Kleij (2004)) Let X be an $n \times p$ matrix of random elements such that

$$\text{vec}(X^T) \sim N_{np}(\text{vec}(M^T), \Omega),$$

where $M = E(X)$ is an $n \times p$ matrix and Ω is the $np \times np$ convariance matrix of $\text{vec}(X^T)$. It is then written

$$X \sim N_{n\times p}(M, \Omega),$$

where $N_{n\times p}$ denotes the matrix normal distribution.

Proposition 2.2. (Proposition 2.2.1, Van Perlo-ten Kleij (2004)) *If $X = [X_1, \ldots, X_n]^T$ is an $n \times p$ random matrix such that $X_1, \ldots, X_n$ are independently and identically distributed as $N_p(0, \Sigma)$, then*

$$X \sim N_{n\times p}(0, I_n \otimes \Sigma).$$

Proposition 2.3. (Proposition 2.2.3, Van Perlo-ten Kleij (2004)) *If $X \sim N_{n\times p}(M, \Omega)$, A is an $m \times n$ matrix, and B is a $p \times h$ matrix, then*

$$AXB \sim N_{m\times h}(AMB, (A \otimes B^T)\Omega(A^T \otimes B)).$$

Proposition 2.3 immediately yields the following corollary.

Corollary 2.1. *If $X \sim N_{n\times p}(0, I_n \otimes \Sigma)$ with $\Sigma \succ 0$, then*

$$X\Sigma^{-\frac{1}{2}} \sim N_{n\times p}(0, I_n \otimes I_p).$$

Definition 2.2. (Definition 8.1, Eaton (2007), Definition 2.3.1, Van Perlo-ten Kleij (2004)) If $X = [X_1, \ldots, X_n]^T$ is an $n \times p$ random matrix with $X \sim N_{n\times p}(0, I_n \otimes \Sigma)$, then $S = X^T X$ has the Wishart distribution, denoted by $W(\Sigma, p, n)$.

The Wishart distribution is the generalization of chi-square distribution (when $p = 1$ and $\Sigma = 1$ in Definition 2.2). It is widely applied in many models of the multivariate statistical analysis (see Anderson (2003), Dempster (1969), Eaton (2007), Muirhead (2005)).

Proposition 2.4. (Proposition 2.3.1, Van Perlo-ten Kleij (2004)) *If $S \sim W(\Sigma, p, n)$ and A is an $m \times p$ matrix, then*

$$ASA^T \sim W(A\Sigma A^T, m, n).$$

Proposition 2.4 immediately yields the following corollary.

Corollary 2.2. *If* $S \sim W(\Sigma, p, n)$ *with* $\Sigma \succ 0$, *then*

$$\Sigma^{-\frac{1}{2}} S \Sigma^{-\frac{1}{2}} \sim W(I_p, p, n).$$

Definition 2.3. (Definition 5.2.3, Ravishanker and Dey (2002)) Let $y_1 = A_1 x$ and $y_2 = A_2 x$ denote two linear functions of a random vector $x \sim N_p(\mu, \Sigma)$. Then the cross-covariance matrix between y_1 and y_2 is given by $\text{cov}(y_1, y_2) = A_1 \Sigma A_2^T$.

Definition 2.4. (Definition 5.2.1, Ravishanker and Dey (2002)) A random vector z with $z \sim N_p(0, I_p)$ if and only if its density function is

$$f(z) = \frac{1}{(2\pi)^{p/2}} \exp\{-\frac{1}{2} z^T z\}.$$

Proposition 2.5. *Let* x *be a random vector with* $x \sim N_p(0, \Sigma)$ *and* $\Sigma \succ 0$. *Let* A_1 *be an* $m \times p$ *matrix and* A_2 *be an* $n \times p$ *matrix. If* $y_1 = A_1 x$ *and* $y_2 = A_2 x$, *then* y_1 *and* y_2 *are independent if and only if* $A_1 \Sigma A_2^T = 0$.

Proof: The "Only if" part follows from Definition 2.3.

"If" part. Without loss of generality, we assume that $\Sigma = I_p$ (this can be done by setting $y = \Sigma^{-\frac{1}{2}} x$ such that $y \sim N_p(0, I_p)$). Now, let $z = (A_1 x, A_2 x)^T$. Then the moment generating function (MGF) of z is $M_z(s,t) = E[\exp(s^T A_1 x + t^T A_2 x)]$, where $s = [t_1, \ldots, t_m]^T$ and $t = [t'_1, \ldots, t'_n]^T$. Thus,

$$\begin{aligned}
M_z(s,t) &= E[\exp(s^T A_1 x + t^T A_2 x)] \\
&= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp(s^T A_1 x + t^T A_2 x) f(x) dx \\
&= \frac{1}{(2\pi)^{p/2}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp(s^T A_1 x + t^T A_2 x - \frac{1}{2} x^T x) dx \\
&= \frac{1}{(2\pi)^{p/2}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp\{-\frac{1}{2}[(x - A_1^T s - A_2^T t)^T (x - A_1^T s - A_2^T t)] \\
&+ \frac{1}{2}(s^T A_1 A_1^T s + t^T A_2 A_2^T t)\} dx \\
&= \exp\{\frac{1}{2}(s^T A_1 A_1^T s + t^T A_2 A_2^T t)\} \\
&= \exp\{\frac{1}{2}(s^T A_1 A_1^T s)\} \exp\{\frac{1}{2}(t^T A_2 A_2^T t)\} \\
&= M_{A_1 x}(s) M_{A_2 x}(t),
\end{aligned}$$

where $M_{A_1 x}(s)$ is MGF of $A_1 x$ and $M_{A_2 x}(t)$ is MGF of $A_2 x$. Thus, y_1 and y_2 are independent.

□

Remark 2.1. By slightly modifying the proof in Proposition 2.5, we can prove that the equivalence of (a) and (b) in Theorem 1.2.

It is well-known that if two random variables x and y are independent, then for any two continuous functions f and g, $f(x)$ and $g(y)$ are independent too. In what follows, we state and prove a general result for this.

First we introduce some notations.

- $(\mathcal{V}_i, \langle\cdot,\cdot\rangle_i)$ denotes an inner product space (vector space $\mathcal{V}_i$ and inner product $\langle\cdot,\cdot\rangle_i$) $(i = 1, 2)$.
- $(\mathcal{W}_i, [\cdot,\cdot]_i)$ denotes an inner product space (vector space $\mathcal{W}_i$ and inner product $[\cdot,\cdot]_i$) $(i = 1, 2)$.
- $\mathfrak{B}(\mathcal{V})$ is the Borel σ-algebra of $(\mathcal{V}, \langle\cdot,\cdot\rangle)$.
- $Q_i(B_i)$, $B_i \in \mathfrak{B}(\mathcal{V}_i)$ $(i = 1, 2)$ is the distribution of X and Y, respectively.
- $Q(B_1 \times B_2)$ is the joint distribution of $(X, Y) \in \mathcal{V}_1 \oplus \mathcal{V}_2$ defined on $\mathfrak{B}(\mathcal{V}_1 \oplus \mathcal{V}_2)$, where $\oplus$ denotes the direct sum and $B_1 \times B_2 \in \mathfrak{B}(\mathcal{V}_1 \oplus \mathcal{V}_2)$.
- $\phi_1(v)$ is the characteristic functions of X defined as
$$\phi_1(v) = E(\exp(i\langle v, X\rangle_1)) = \int_{\mathcal{V}_1} \exp(i\langle v, x\rangle_1)Q_1 \mathrm{d}x, v \in \mathcal{V}_1.$$
- $\phi_2(w)$ is the characteristic functions of Y defined as
$$\phi_2(w) = E(\exp(i\langle w, Y\rangle_2)) = \int_{\mathcal{V}_2} \exp(i\langle w, y\rangle_2)Q_2 \mathrm{d}y, w \in \mathcal{V}_2.$$
- $\phi(v, w)$ is the characteristic function of $(X, Y) \in \mathcal{V}_1 \oplus \mathcal{V}_2$ defined as
$$\begin{aligned}\phi(v, w)) &= E(\exp(i\langle v, X\rangle_1 + i\langle w, Y\rangle_2)) \\ &= \int_{\mathcal{V}_1\oplus\mathcal{V}_2} \exp(i\langle v, x\rangle_1 + i\langle w, y\rangle_2)Q \,\mathrm{d}x \,\mathrm{d}y, v \in \mathcal{V}_1, w \in \mathcal{V}_2,\end{aligned}$$

Recall a result from Eaton (2007).

Proposition 2.6. (Proposition 2.9, Eaton (2007)) *Let random vectors $X \in (\mathcal{V}_1, \langle\cdot,\cdot\rangle_1)$ and $Y \in (\mathcal{V}_2, \langle\cdot,\cdot\rangle_2)$. Then the following are equivalent:*

(a) *X and Y are independent.*

(b) *$Q(B_1 \times B_2) = Q_1(B_1)Q_2(B_2)$ for all $B_1 \in \mathfrak{B}(\mathcal{V}_1)$ and $B_2 \in \mathfrak{B}(\mathcal{V}_2)$.*

(c) *$\phi(v, w) = \phi_1(v)\phi_2(w)$ for all $v \in \mathcal{V}_1$ and $w \in \mathcal{V}_2$.*

Proposition 2.6 immediately yields the following proposition.

Proposition 2.7. *Suppose that functions $g : \mathcal{V}_1 \mapsto \mathcal{W}_1$ and $h : \mathcal{V}_2 \mapsto \mathcal{W}_2$ are continuous. If random vectors $X \in (\mathcal{V}_1, \langle\cdot,\cdot\rangle_1)$ and $Y \in (\mathcal{V}_2, \langle\cdot,\cdot\rangle_2)$ are independent, then $g(X) \in (\mathcal{W}_1, [\cdot,\cdot]_1)$ and $h(Y) \in (\mathcal{W}_2, [\cdot,\cdot]_2)$ are independent.*

Proof: Letting $\psi(v, w)$ be the characteristic function of $(g(X), h(Y)) \in \mathcal{W}_1 \oplus \mathcal{W}_2$, $\psi_1(v)$ be the characteristic function of $g(X)$, and $\psi_2(w)$ be the characteristic function of $h(Y)$), then for $v \in \mathcal{W}_1$ and $w \in \mathcal{W}_2$.

$$\begin{aligned}
\psi(v, w) &= E(\exp(i[v, g(X)]_1 + i[w, h(Y)]_2)) \\
&= \int_{\mathcal{V}_1 \oplus \mathcal{V}_2} \exp(i[v, g(x)]_1 + i[w, h(y)]_2) Q \, \mathrm{d}x \, \mathrm{d}y \\
&= \int_{\mathcal{V}_1} \int_{\mathcal{V}_2} \exp(i[v, g(x)]_1) \exp(i[w, h(y)]_2) Q_1 \mathrm{d}x \, Q_2 \mathrm{d}y \\
&= \int_{\mathcal{V}_1} \exp(i[v, g(x)]_1) Q_1 \mathrm{d}x \int_{\mathcal{V}_2} \exp(i[w, h(y)]_2) Q_2 \mathrm{d}y \\
&= \psi_1(v)\psi_2(w).
\end{aligned}$$

Note that the third equality is from independence of X and Y (see Proposition 2.6). Thus, by Proposition 2.6, $g(X)$ and $h(Y)$ are independent. □

3. MAIN RESULTS

3.1. The independence of linear and quadratic forms in matrix normal distribution

In this subsection, we answer *"Question 1"* in Introduction. First, we recall two known results. $A = AA^-A$ and $A = A(A^TA)^-A^TA$, where A^- is a generalized inverse of matrix A.

Theorem 3.1. *Let a random matrix $X \sim N_{n\times p}(0, I_n \otimes \Sigma)$ with $\Sigma \succ 0$ and let A and B be $p \times p$ symmetric real matrices. Then the following are equivalent:*

(a) $A\Sigma B = 0$.

(b) *XA and XB are independently distributed.*

(c) *XA and XBX^T are independently distributed.*

(d) *XAX^T and XBX^T are independently distributed.*

(e) *$\mathrm{tr}(AS)$ and $\mathrm{tr}(BS)$ are independently distributed, where $S = X^TX$.*

Proof: Since $X \sim N_{n\times p}(0, I_n \otimes \Sigma)$, $\mathrm{vec}(X^T) \sim N_{np}(0, I_n \otimes \Sigma)$ by Definition 2.1. Without loss of generality, we assume that $\Sigma = I_p$ (this can be done by setting $Y = X\Sigma^{-\frac{1}{2}}$

such that $Y \sim N_{n\times p}(0, I_n \otimes I_p)$ by Corollary 2.1 and the transformation $A \to \Sigma^{\frac{1}{2}} A \Sigma^{\frac{1}{2}}$, $B \to \Sigma^{\frac{1}{2}} B \Sigma^{\frac{1}{2}}$).

(a) $\Rightarrow$ (b): From Proposition 2.1,

$$\mathrm{vec}(AX^T) = (I_n \otimes A)\,\mathrm{vec}(X^T) \text{ and } \mathrm{vec}(BX^T) = (I_n \otimes B)\,\mathrm{vec}(X^T).$$

Thus,

$$(I_n \otimes A)(I_n \otimes I_p)(I_n \otimes B)^T = (I_n \otimes A)(I_n \otimes I_p)(I_n \otimes B) = (I_n \otimes AB) = 0.$$

Hence, from Proposition 2.5, $\mathrm{vec}(AX^T)$ and $\mathrm{vec}(BX^T)$ are independent. Since there exist linear (isomorphism) functions f and g such that $f(\mathrm{vec}(AX^T)) := AX^T$ and $g(\mathrm{vec}(BX^T)) := BX^T$, AX^T and BX^T are independent by Proposition 2.7, and hence $XA = (AX^T)^T$ and $XB = (BX^T)^T$ are independent by Proposition 2.7.

(b) $\Rightarrow$ (c): From Proposition 2.7, XA is independent of

$$f_B(XB) := XBB^-(XB)^T = XBB^-BX^T = XBX^T,$$

where B^- is a generalized inverse of B.

(c) $\Rightarrow$ (d): From Proposition 2.7, XBX^T is independent of

$$f_A(XA) := XAA^-(XA)^T = XAA^-AX^T = XAX^T,$$

where A^- is a generalized inverse of A.

(d) $\Rightarrow$ (e): From Proposition 2.7, $f(XAX^T) := \mathrm{tr}(XAX^T) = \mathrm{tr}(AX^TX) = \mathrm{tr}(AS)$ is independent of $g(XBX^T) := \mathrm{tr}(XBX^T) = \mathrm{tr}(BX^TX) = \mathrm{tr}(BS)$, where $S = X^TX$.

Since $S \sim W(I_p, p, n)$, the implication of (e) $\Rightarrow$ (a) follows from Theorem 1.1. □

Remark 3.1. Hogg (Theorem 4, Hogg (1963)) proved that XA^T and XBX^T are independently distributed if and only if $A\Sigma B = 0$, where A is not necessary to be a square matrix. In the next theorem (see Theorem 3.2), we give a simple proof for Hogg's results by using Theorem 3.1.

Theorem 3.2. *Let a random matrix $X \sim N_{n\times p}(0, I_n \otimes \Sigma)$ with $\Sigma \succ 0$ and let A be a $q \times p$ matrix and B be $p \times p$ a symmetric matrix. Then the following are equivalent:*

(a) $A\Sigma B = 0$.

(b) *XA^T and XB are independently distributed.*

(c) *XA^T and XBX^T are independently distributed.*

Proof: Since $X \sim N_{n\times p}(0, I_n \otimes \Sigma)$, $\mathrm{vec}(X^T) \sim N_{np}(0, I_n \otimes \Sigma)$ by Definition 2.1.

(a) $\Rightarrow$ (b): From Proposition 2.1,

$$\mathrm{vec}(AX^T) = (I_n \otimes A)\,\mathrm{vec}(X^T) \text{ and } \mathrm{vec}(BX^T) = (I_n \otimes B)\,\mathrm{vec}(X^T).$$

Thus,

$$(I_n \otimes A)(I_n \otimes \Sigma)(I_n \otimes B)^T = (I_n \otimes A)(I_n \otimes \Sigma)(I_n \otimes B) = (I_n \otimes A\Sigma B) = 0.$$

Hence, from Proposition 2.5, $\mathrm{vec}(AX^T)$ and $\mathrm{vec}(BX^T)$ are independent. Repeating the same argument as the proof of (a) $\Rightarrow$ (b) in Theorem 3.1, we have that XA^T and XB are independent.

(b) $\Rightarrow$ (c): From Proposition 2.7, XA^T is independent of

$$f_B(XB) = XBB^-(XB)^T = XBB^-BX^T = XBX^T,$$

where B^- is a generalized inverse of B.

(c) $\Rightarrow$ (a): Since XA^T is independent of XBX^T, $(XA^T)(XA^T)^T = XA^TAX^T$ is independent of XBX^T by Proposition 2.7. Thus, from Theorem 3.1, $(A^TA)\Sigma B = 0$. Hence, $A(A^TA)^-(A^TA)\Sigma B = 0$. Since $A = A(A^TA)^-(A^TA)$, we have $A\Sigma B = 0$. □

3.2. The independence of linear and quadratic forms in Wishart variates

In this subsection, we study interconnections between the independence of linear forms, quadratic forms, trace forms in Wishart variates. In particular, we answer *"Question 2"* in Introduction.

Theorem 3.3. *Let $S \sim W(\Sigma, p, n)$ with $\Sigma \succ 0$ and let A and B be $p \times p$ symmetric matrices. Then the following are equivalent:*

(a) $(A\Sigma A)\Sigma(B\Sigma B) = 0$.

(b) $A\Sigma B = 0$.

(c) *$P_A(S) := ASA$ and $P_B(S) := BSB$ are independently distributed.*

(d) *$P_A(S)$ and $\mathrm{tr}(P_B(S))$ are independently distributed.*

(e) *$\mathrm{tr}(P_A(S))$ and $\mathrm{tr}(P_B(S))$ are independently distributed.*

(f) *The nonzero eigenvalues of $(\Sigma^{\frac{1}{2}}A\Sigma^{\frac{1}{2}})^2 + (\Sigma^{\frac{1}{2}}B\Sigma^{\frac{1}{2}})^2$ are just the nonzero eigenvalues of $(\Sigma^{\frac{1}{2}}A\Sigma^{\frac{1}{2}})^2$ and $(\Sigma^{\frac{1}{2}}B\Sigma^{\frac{1}{2}})^2$.*

(g) *The nonzero eigenvalues of $\Sigma^{\frac{1}{2}}(A+B)\Sigma^{\frac{1}{2}}$ are just the nonzero eigenvalues of $\Sigma^{\frac{1}{2}}A\Sigma^{\frac{1}{2}}$ and $\Sigma^{\frac{1}{2}}B\Sigma^{\frac{1}{2}}$.*

(h) $P_{\Sigma^{\frac{1}{2}}A\Sigma^{\frac{1}{2}}}P_{\Sigma^{\frac{1}{2}}B\Sigma^{\frac{1}{2}}} = 0$.

(i) $\mathrm{tr}(AS)$ *and* $\mathrm{tr}(BS)$ *are independently distributed.*

Proof: Without loss of generality, we assume that $\Sigma = I_p$ (this can be done by setting $S' = \Sigma^{-\frac{1}{2}} S \Sigma^{-\frac{1}{2}}$ such that $S' \sim W(I_p, p, n)$ by Corollar 2.2 and the transformation $A \to \Sigma^{\frac{1}{2}} A \Sigma^{\frac{1}{2}}$, $B \to \Sigma^{\frac{1}{2}} B \Sigma^{\frac{1}{2}}$).

(a) $\Rightarrow$ (b): $A^2B^2 = 0 \Rightarrow \mathrm{tr}(A^2B^2) = 0$. Since

$$\mathrm{tr}(A^2B^2) = \mathrm{tr}(A(AB)B) = \mathrm{tr}(ABBA) = \mathrm{tr}[(AB)(AB)^T],$$

we have $\mathrm{tr}[(AB)(AB)^T] = 0$. Thus, $AB = 0$.

(b) $\Rightarrow$ (c): Since $S \sim W(I_p, p, n)$, $S = X^TX$ with $\mathrm{vec}(X^T) \sim N_{np}(0, I_n \otimes I_p)$ by Definition 2.2. Thus, $ASA = AX^TXA$. Let $Y = AX^T$. Then from Proposition 2.1, we have $\mathrm{vec}(Y) = (I_n \otimes A)\mathrm{vec}(X^T)$. Similarly, $BSB = BX^TXB$. Let $Z = BX^T$. Then $\mathrm{vec}(Z) = (I_n \otimes B)\mathrm{vec}(X^T)$. Thus,

$$(I_n \otimes A)(I_n \otimes I_p)(I_n \otimes B)^T = (I_n \otimes A)(I_n \otimes I_p)(I_n \otimes B) = (I_n \otimes AB) = 0.$$

Hence, from Proposition 2.5, $\mathrm{vec}(Y)$ and $\mathrm{vec}(Z)$ are independent. Since there exist linear (isomorphism) functions f and g such that $f(\mathrm{vec}(Y)) := Y$ and $g(\mathrm{vec}(Z)) := Z$, Y and Z are independent by Proposition 2.7. Hence,

$$YY^T = AX^TXA = ASA \text{ and } ZZ^T = BX^TXB = BSB$$

are also independently distributed by Proposition 2.7, i.e., $P_A(S)$ and $P_B(S)$ are independent.

(c) $\Rightarrow$ (d): (c) implies that ASA is independent of $f(BSB) := \mathrm{tr}(BSB)$.

(d) $\Rightarrow$ (e): (d) implies that $\mathrm{tr}(BSB)$ is independent of $g(ASA) := \mathrm{tr}(ASA)$.

(e) $\Rightarrow$ (f): Since $\mathrm{tr}(ASA) = \mathrm{tr}(SA^2)$ and $\mathrm{tr}(BSB) = \mathrm{tr}(SB^2)$, from Theorem 1.1, we have $A^2B^2 = 0$. Also, the equivalence (f) and (a) follows from Theorem 1.1. Thus, (f) holds.

The equivalence of (g) and (b) follows from Theorem 1.1.

The equivalence of (h) and (b) follows from Theorem 1.1.

The equivalence of (i) and (b) follows from Theorem 3.1. □

The following theorem shows that the items (b) and (c) in Theorem 3.3 are equivalent when A is not a square matrix.

Theorem 3.4. *Let $S \sim W(\Sigma, p, n)$ with $\Sigma \succ 0$ and let A be a $q \times p$ matrix and B be a $p \times p$ symmetric matrix. Then $P_A(S) := ASA^T$ and $P_B(S) := BSB$ are independently distributed if and only if $A\Sigma B = 0$.*

Proof: Since $S \sim W(\Sigma, p, n)$, $S = X^TX$ with $\mathrm{vec}(X^T) \sim N_{np}(0, I_n \otimes \Sigma)$ by Definition 2.2.

"If part": From $ASA^T = AX^TXA^T$, we define $Y = AX^T$. Then we have

$$\text{vec}(Y) = (I_n \otimes A)\text{vec}(X^T).$$

Similarly, we can define $Z = BX^T$. Then $\text{vec}(Z) = (I_n \otimes B)\text{vec}(X^T)$. Thus,

$$(I_n \otimes A)(I_n \otimes \Sigma)(I_n \otimes B)^T = (I_n \otimes A)(I_n \otimes \Sigma)(I_n \otimes B) = (I_n \otimes A\Sigma B) = 0.$$

Hence, from Proposition 2.5, $\text{vec}(Y)$ and $\text{vec}(Z)$ are independent. Since there exist linear (isomorphism) functions f and g such that $f(\text{vec}(Y)) := Y$ and $g(\text{vec}(Z)) := Z$, Y and Z are independent by Proposition 2.7. Hence,

$$YY^T = AX^TXA^T = ASA^T \text{ and } ZZ^T = BX^TXB = BSB$$

are also independently distributed by Proposition 2.7, i.e., $P_A(S)$ and $P_B(S)$ are independent.

"Only if" part: Since $P_A(S) = ASA^T$ and $P_B(S) = BSB$ are independent,

$$f(ASA^T) := A^T(ASA^T)A = (A^TA)S(A^TA) \text{ and } P_B(S)$$

are also independent. From Theorem 3.3, $(A^TA)\Sigma B = 0$. Hence, $A(A^TA)^-(A^TA)\Sigma B = 0$. Since $A = A(A^TA)^-(A^TA)$, we have $A\Sigma B = 0$. □

Theorem 3.5. *Let $S \sim W(\Sigma, p, n)$ with $\Sigma \succ 0$ and let A and B be $p \times p$ symmetric matrices. Consider the following statements:*

(a) *SA and SB are independently distributed.*

(b) *SA and SBS are independently distributed.*

(c) *SAS and SBS are independently distributed.*

(d) *$A\Sigma B = 0$.*

Then we have (a) $\Rightarrow$ (b) $\Rightarrow$ (c).

If, in addition, $A \succeq 0$ and $B \succeq 0$, then (c) $\Rightarrow$ (d).

Proof: (a) $\Rightarrow$ (b): From Proposition 2.7, SA is independent of

$$f_B(SB) := SBB^-(SB)^T = SBB^-BS = SBS,$$

where B^- is a generalized inverse of B.

(b) $\Rightarrow$ (c): From Proposition 2.7, SBS is independent of

$$f_A(SA) := SAA^-(SA)^T = SAA^-AS^T = SAS,$$

where A^- is a generalized inverse of A.

(c) $\Rightarrow$ (d): Suppose that $A \succeq 0$ and $B \succeq 0$. Then $A^{1/2}$ and $B^{1/2}$ exist. From Proposition 2.7,

$$\begin{aligned} f(SAS) := A^{1/2}SASA^{1/2} &= A^{1/2}SA^{1/2}A^{1/2}SA^{1/2} \\ &= (A^{1/2}SA^{1/2})(A^{1/2}SA^{1/2}) \\ &= (A^{1/2}SA^{1/2})^2 \end{aligned}$$

and

$$\begin{aligned} g(SBS) := B^{1/2}SBSB^{1/2} &= B^{1/2}SB^{1/2}B^{1/2}SB^{1/2} \\ &= (B^{1/2}SB^{1/2})(B^{1/2}SB^{1/2}) \\ &= (B^{1/2}SB^{1/2})^2 \end{aligned}$$

are independent. Since $S \succeq 0$, we have $A^{1/2}SA^{1/2} \succeq 0$ and $B^{1/2}SB^{1/2} \succeq 0$. Again, from Proposition 2.7, $[(A^{1/2}SA^{1/2})^2]^{1/2} = A^{1/2}SA^{1/2}$ and $[(B^{1/2}SB^{1/2})^2]^{1/2} = B^{1/2}SB^{1/2}$ are independent. Hence, from Theorem 3.3, $A^{1/2}\Sigma B^{1/2} = 0 \Rightarrow A\Sigma B = 0$. □

Corollary 3.1. *Let $S \sim W(\Sigma, p, n)$ with $\Sigma \succ 0$ and let $A \succeq 0$ and $B \succeq 0$. Consider the following statements:*

(a) $P_S(A) := SAS$ *and* $P_S(B) := SBS$ *are independently distributed.*

(b) $P_A(S) := ASA$ *and* $P_B(S) := BSB$ *are independently distributed.*

(c) $\mathrm{tr}(P_A(S))$ *and* $\mathrm{tr}(P_B(S))$ *are independently distributed.*

(d) $\mathrm{tr}(AS)$ *and* $\mathrm{tr}(BS)$ *are independently distributed.*

(e) $A\Sigma B = 0$.

Then we have (a) $\Rightarrow$ (b) $\Leftrightarrow$ (c) $\Leftrightarrow$ (d) $\Leftrightarrow$ (e).

Proof: (a) $\Rightarrow$ (b): From the proof of (c) $\Rightarrow$ (d) in Theorem 3.5, we have $A\Sigma B = 0$. Thus, by Theorem 3.3, we have that $P_A(S)$ and $P_B(S)$ are independent.

The equivalence of (b)-(e) follows from Theorem 3.3. □

The following example shows that Item (d) in Theorem 3.5 implies neither Item (a) nor (c).

Example 3.1. Let

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \text{ and } X = [X_1, X_2, X_3]^T,$$

where

$$X_1 = [x_{11}\, x_{12}]^T, \; X_2 = [x_{21}\, x_{22}]^T, \text{ and } X_3 = [x_{31}\, x_{32}]^T$$

are independent with $X_i \sim N_2(0, I_2)$ $(i = 1, 2, 3)$. Then $\mathrm{vec}(X^T) \sim N_6(0, I_3 \otimes I_2)$ and $S \sim W(I_2, 2, 3)$, where $S = X^T X$. It is easy to verify that $AB = 0$,

$$SA = \begin{bmatrix} x_{11}^2 + x_{21}^2 + x_{31}^2 & 0 \\ x_{11}x_{12} + x_{21}x_{22} + x_{31}x_{32} & 0 \end{bmatrix}, \qquad SB = \begin{bmatrix} 0 & x_{11}x_{12} + x_{21}x_{22} + x_{31}x_{32} \\ 0 & x_{12}^2 + x_{22}^2 + x_{32}^2 \end{bmatrix},$$

$$SAS = \begin{bmatrix} (x_{11}^2 + x_{21}^2 + x_{31}^2)^2 & (x_{11}^2 + x_{21}^2 + x_{31}^2)(x_{11}x_{12} + x_{21}x_{22} + x_{31}x_{32}) \\ (x_{11}^2 + x_{21}^2 + x_{31}^2)(x_{11}x_{12} + x_{21}x_{22} + x_{31}x_{32}) & (x_{11}x_{12} + x_{21}x_{22} + x_{31}x_{32})^2 \end{bmatrix},$$

and

$$SBS = \begin{bmatrix} (x_{11}x_{12} + x_{21}x_{22} + x_{31}x_{32})^2 & (x_{11}x_{12} + x_{21}x_{22} + x_{31}x_{32})(x_{12}^2 + x_{22}^2 + x_{32}^2) \\ (x_{11}x_{12} + x_{21}x_{22} + x_{31}x_{32})(x_{12}^2 + x_{22}^2 + x_{32}^2) & (x_{12}^2 + x_{22}^2 + x_{32}^2)^2 \end{bmatrix}.$$

Thus, if SA and SB are independent, then $\mathrm{vec}(SA)$ and $\mathrm{vec}(SB)$ are independent by Proposition 2.7. Now, letting $z := x_{11}x_{12} + x_{21}x_{22} + x_{31}x_{32}$, we have $E[(z - E(z))^2] > 0$. Thus, the cross-covariance matrix of $\mathrm{vec}(SA)$ and $\mathrm{vec}(SB)$

$$E[(\mathrm{vec}(SA) - E(\mathrm{vec}(SA)))((\mathrm{vec}(SB) - E(\mathrm{vec}(SB)))^T] \neq 0.$$

Hence, SA and SB are not independent.

Similarly, if SAS and SBS are independent, then $\mathrm{vec}(SAS)$ and $\mathrm{vec}(SBS)$ are independent by Proposition 2.7. Now letting $y := (x_{11}x_{12} + x_{21}x_{22} + x_{31}x_{32})^2$, we have $E[(y - E(y))^2] > 0$. Thus, the cross-covariance matrix of $\mathrm{vec}(SAS)$ and $\mathrm{vec}(SBS)$

$$E[(\mathrm{vec}(SAS) - E(\mathrm{vec}(SAS)))((\mathrm{vec}(SBS) - E(\mathrm{vec}(SBS)))^T] \neq 0.$$

Hence, SAS and SBS are not independent.

Remark 3.2. Example 3.1 shows that the Craig-Sakamoto theorem fails in establishing the independence of two linear forms and the independence of two quadratic forms in Wishart variates.

4. CONCLUDING REMARKS

In this paper, we have investigated interconnections between the independence of linear forms, quadratic forms, trace forms in matrix normal distribution and Wishart distribution. We have showed that the Craig-Sakamoto theorem still establishes the independence of both two linear forms and two quadratic forms in matrix normal distribution, but it establishes neither the independence of two linear forms nor the independence of two quadratic forms in Wishart variates. An interesting future research project is to find some applications for the theorems obtained by this paper.

ACKNOWLEDGMENTS

We would like to thank two anonymous referees for their very constructive suggestions and comments, which helped to improve the presentation of this paper. Part of the work was done when the first author was visiting the Department of Applied Mathematics, Hong Kong Polytechnic University during summer of 2023. He wishes to thank Professor Defeng Sun for the invitation and hospitality. The work was supported in part by the Simons Foundation (632688), the 111 Project of China (B16002), and the National Natural Science Foundation of China (11971302, 12171307, and 12071022).

REFERENCES

Anderson, T. W. (2003). *An Introduction to Multivariate Statistical Analysis*. New York, John Wiely & Sons.

Dempster, A. P. (1969). *Elements of Continuous Multivariate Statistical Analysis*. Mass Academic Press, Reading.

Driscoll, M. F. and Gundberg, W. R. (1986). A history of the development of craig's theorem. *Journal of the American Statistical Association*, 40(1):65–70.

Driscoll, M. F. and Krasnicka, B. (1995). An accessible proof of craig's theorem in the general case. *Journal of the American Statistical Association*, 49(1):59–62.

Eaton, M. L. (2007). *Multivariate Statistics – A Vector Space Approach, Institute of Mathematical Statistics*. Lecture Notes-Monograph Series, Vol. 53.

Harville, D. A. (2012). *Matrix Algebra From a Statistical's Perspective*. Springer.

Hogg, R. V. (1963). On the independence of certain wishart variables. *Annals of Mathematical Statistics*, 34(3):935–939.

Letac, G. and Massam, H. (1995). Craig-sakamoto's theorem for the wishart distributions on symmetric cones. *Annals of the Institute of Statistical Mathematics*, 47(4):785–799.

Li, C. K. (2000). A simple proof of the craig-sakamoto theorem. *Linear Algebra and its Applications*, 321(1-3):281–283.

Muirhead, R. J. (2005). *Aspects of Multivariate Statistical Theory*. New York, John Wiely & Sons.

Ogawa, J. and Olkin, I. (2008). A tale of two countries: The craig–sakamoto–matusita theorem. *Journal of Statistical Planning and Inference*, 138(11):3419–3428.

Olkin, I. (1997). A determinantal proof of craig-sakamoto theorem. *Linear Algebra and its Applications*, 264(Sixth Special Issue):217–223.

Ravishanker, N. and Dey, D. K. (2002). *A First Course in Linear Model Theory*. Texts in Statistical Science, Chapman & Hall CRC.

Tao, J. and Wang, G. (2016). A generalization of the craig-sakamoto theorem to euclidean jordan algebras. *Linear Algebra and its Applications*, 493(1):134–145.

Van Perlo-ten Kleij, F. (2004). *Contributions to multivariate analysis with Applications in marketing*. Ph.D. thesis, University of Groningen, Groningen, Netherlands.

Zhang, J. (2017). On the independence of linear and quadratic forms in normal variate. *Communications in Statistics-Theory and Methods*, 46(17):8493–8496.

Zhang, J. and Yi, J. (2012). A simple proof of the generalized craig-sakamoto theorem. *Linear Algebra and its Applications*, 437(3):781–782.