Chapter 4
The Matrix-Variate Gaussian Distribution

4.1. Introduction

This chapter relies on various results presented in Chap. 1. We will introduce a class of integrals called the real matrix-variate Gaussian integrals and complex matrix-variate Gaussian integrals wherefrom a statistical density referred to as the matrix-variate Gaussian density and, as a special case, the multivariate Gaussian or normal density will be obtained, both in the real and complex domains.

The notations introduced in Chap. 1 will also be utilized in this chapter. Scalar variables, mathematical and random, will be denoted by lower case letters, vector/matrix variables will be denoted by capital letters, and complex variables will be indicated by a tilde. Additionally, the following notations will be used. All the matrices appearing in this chapter are $p \times p$ real positive definite or Hermitian positive definite unless stated otherwise. $X > O$ will mean that that the $p \times p$ real symmetric matrix X is positive definite and $\tilde{X} > O$, that the $p \times p$ matrix $\tilde{X}$ in the complex domain is Hermitian, that is, $\tilde{X} = \tilde{X}^*$ where $\tilde{X}^*$ denotes the conjugate transpose of $\tilde{X}$ and $\tilde{X}$ is positive definite. $O < A < X < B$ will indicate that the $p \times p$ real positive definite matrices are such that $A > O,\ B > O,\ X > O,\ X - A > O,\ B - X > O$. $\int_X f(X)\mathrm{d}X$ represents a real-valued scalar function $f(X)$ being integrated out over all X in the domain of X where $\mathrm{d}X$ stands for the wedge product of differentials of all distinct elements in X. If $X = (x_{ij})$ is a real $p \times q$ matrix, the x_{ij}'s being distinct real scalar variables, then $\mathrm{d}X = \mathrm{d}x_{11} \wedge \mathrm{d}x_{12} \wedge \ldots \wedge \mathrm{d}x_{pq}$ or $\mathrm{d}X = \wedge_{i=1}^{p} \wedge_{j=1}^{q} \mathrm{d}x_{ij}$. If $X = X'$, that is, X is a real symmetric matrix of dimension $p \times p$, then $\mathrm{d}X = \wedge_{i \ge j=1}^{p} \mathrm{d}x_{ij} = \wedge_{i \le j=1}^{p} \mathrm{d}x_{ij}$, which involves only $p(p+1)/2$ differential elements $\mathrm{d}x_{ij}$. When taking the wedge product, the elements x_{ij}'s may be taken in any convenient order to start with. However, that order has to be maintained until the computations are completed. If $\tilde{X} = X_1 + iX_2$, where X_1 and X_2 are real $p \times q$ matrices, $i = \sqrt{(-1)}$, then $\mathrm{d}\tilde{X}$ will be defined as $\mathrm{d}\tilde{X} = \mathrm{d}X_1 \wedge \mathrm{d}X_2$. $\int_{A < \tilde{X} < B} f(\tilde{X})\mathrm{d}\tilde{X}$ represents the real-valued scalar function f of complex matrix argument $\tilde{X}$ being integrated out over

A. M. Mathai et al., *Multivariate Statistical Analysis in the Real and Complex Domains*,
https://doi.org/10.1007/978-3-030-95864-0_4

all $p \times p$ matrix $\tilde{X}$ such that $A > O,\ \tilde{X} > O,\ B > O,\ \tilde{X} - A > O,\ B - \tilde{X} > O$ (all Hermitian positive definite), where A and B are constant matrices in the sense that they are free of the elements of $\tilde{X}$. The corresponding integral in the real case will be denoted by $\int_{A<X<B} f(X)\mathrm{d}X = \int_A^B f(X)\mathrm{d}X,\ A > O,\ X > O,\ X - A > O,\ B > O,\ B - X > O$, where A and B are constant matrices, all the matrices being of dimension $p \times p$.

4.2. Real Matrix-variate and Multivariate Gaussian Distributions

Let $X = (x_{ij})$ be a $p \times q$ matrix whose elements x_{ij} are distinct real variables. For any real matrix X, be it square or rectangular, $\mathrm{tr}(XX') = \mathrm{tr}(X'X) =$ sum of the squares of all the elements of X. Note that XX' need not be equal to $X'X$. Thus, $\mathrm{tr}(XX') = \sum_{i=1}^p \sum_{j=1}^q x_{ij}^2$ and, in the complex case, $\mathrm{tr}(\tilde{X}\tilde{X}^*) = \sum_{i=1}^p \sum_{j=1}^q |\tilde{x}_{ij}|^2$ where if $\tilde{x}_{rs} = x_{rs1} + ix_{rs2}$ where x_{rs1} and x_{rs2} are real, $i = \sqrt{(-1)}$, with $|\tilde{x}_{rs}| = +[x_{rs1}^2 + x_{rs2}^2]^{\frac{1}{2}}$. Consider the following integrals over the real rectangular $p \times q$ matrix X:

$$
\begin{aligned}
I_1 &= \int_X \mathrm{e}^{-\mathrm{tr}(XX')}\mathrm{d}X = \int_X \mathrm{e}^{-\sum_{i=1}^p \sum_{j=1}^q x_{ij}^2}\mathrm{d}X = \prod_{i,j} \int_{-\infty}^{\infty} \mathrm{e}^{-x_{ij}^2}\mathrm{d}x_{ij} \\
&= \prod_{i,j} \sqrt{\pi} = \pi^{\frac{pq}{2}}, \qquad (i) \\
I_2 &= \int_X \mathrm{e}^{-\frac{1}{2}\mathrm{tr}(XX')}\mathrm{d}X = (2\pi)^{\frac{pq}{2}}. \qquad (ii)
\end{aligned}
$$

Let $A > O$ be $p \times p$ and $B > O$ be $q \times q$ constant positive definite matrices. Then we can define the unique positive definite square roots $A^{\frac{1}{2}}$ and $B^{\frac{1}{2}}$. For the discussions to follow, we need only the representations $A = A_1A_1',\ B = B_1B_1'$ with A_1 and B_1 nonsingular, a prime denoting the transpose. For an $m \times n$ real matrix X, consider

$$
\begin{aligned}
\mathrm{tr}(AXBX') = \mathrm{tr}(A^{\frac{1}{2}}A^{\frac{1}{2}}XB^{\frac{1}{2}}B^{\frac{1}{2}}X') &= \mathrm{tr}(A^{\frac{1}{2}}XB^{\frac{1}{2}}B^{\frac{1}{2}}X'A^{\frac{1}{2}}) \\
&= \mathrm{tr}(YY'),\ Y = A^{\frac{1}{2}}XB^{\frac{1}{2}}. \qquad (iii)
\end{aligned}
$$

In order to obtain the above results, we made use of the property that for any two matrices P and Q such that PQ and QP are defined, $\mathrm{tr}(PQ) = \mathrm{tr}(QP)$ where PQ need not be equal to QP. As well, letting $Y = (y_{ij})$, $\mathrm{tr}(YY') = \sum_{i=1}^p \sum_{j=1}^q y_{ij}^2$. YY' is real positive definite when Y is $p \times q,\ p \leq q$, is of full rank p. Observe that any real square matrix U that can be written as $U = VV'$ for some matrix V where V may be square or rectangular, is either positive definite or at least positive semi-definite. When V is a $p \times q$ matrix, $q \geq p$, whose rank is p, VV' is positive definite; if the rank of V is less than p, then VV' is positive semi-definite. From Result 1.6.4,

$$Y = A^{\frac{1}{2}} X B^{\frac{1}{2}} \Rightarrow \mathrm{d}Y = |A|^{\frac{q}{2}} |B|^{\frac{p}{2}} \mathrm{d}X$$
$$\Rightarrow \mathrm{d}X = |A|^{-\frac{q}{2}} |B|^{-\frac{p}{2}} \mathrm{d}Y \qquad (iv)$$

where we use the standard notation $|(\cdot)| = \det(\cdot)$ to denote the determinant of $(\cdot)$ in general and $|\det(\cdot)|$ to denote the absolute value or modulus of the determinant of $(\cdot)$ in the complex domain. Let

$$f_{p,q}(X) = \frac{|A|^{\frac{q}{2}} |B|^{\frac{p}{2}}}{(2\pi)^{\frac{pq}{2}}} \mathrm{e}^{-\frac{1}{2}\mathrm{tr}(AXBX')},\ A > O,\ B > O \qquad (4.2.1)$$

for $X = (x_{ij}),\ -\infty < x_{ij} < \infty$ for all i and j. From the steps *(i)* to *(iv)*, we see that $f_{p,q}(X)$ in (4.2.1) is a statistical density over the real rectangular $p \times q$ matrix X. This function $f_{p,q}(X)$ is known as the *real matrix-variate Gaussian density*. We introduced a $\frac{1}{2}$ in the exponent so that particular cases usually found in the literature agree with the real p-variate Gaussian distribution. Actually, this $\frac{1}{2}$ factor is quite unnecessary from a mathematical point of view as it complicates computations rather than simplifying them. In the complex case, the factor $\frac{1}{2}$ does not appear in the exponent of the density, which is consistent with the current particular cases encountered in the literature.

Note 4.2.1. If the factor $\frac{1}{2}$ is omitted in the exponent, then 2π is to be replaced by π in the denominator of (4.2.1), namely,

$$f_{p,q}(X) = \frac{|A|^{\frac{q}{2}} |B|^{\frac{p}{2}}}{(\pi)^{\frac{pq}{2}}} \mathrm{e}^{-\mathrm{tr}(AXBX')},\ A > O,\ B > O. \qquad (4.2.2)$$

When $p = 1$, the matrix X is $1 \times q$ and we let $X = (x_1, \ldots, x_q)$ where X is a row vector whose components are $x_1, \ldots, x_q$. When $p = 1$, A is 1×1 or a scalar quantity. Letting $A = 1$ and $B = V^{-1}$, $V > O$, be of dimension $q \times q$, then in the real case,

$$f_{1,q}(X) = \frac{|\frac{1}{2}V^{-1}|^{\frac{1}{2}}}{\pi^{\frac{q}{2}}} \mathrm{e}^{-\frac{1}{2}XV^{-1}X'},\ X = (x_1, \ldots, x_q),$$
$$= \frac{1}{(2\pi)^{\frac{q}{2}} |V|^{\frac{1}{2}}} \mathrm{e}^{-\frac{1}{2}XV^{-1}X'}, \qquad (4.2.3)$$

which is the usual real nonsingular Gaussian density with parameter matrix V, that is, $X' \sim N_q(O, V)$. If a location parameter vector $\mu = (\mu_1, \ldots, \mu_q)$ is introduced or, equivalently, if X is replaced by $X - \mu$, then we have

$$f_{1,q}(X) = [(2\pi)^{\frac{q}{2}} |V|^{\frac{1}{2}}]^{-1} \mathrm{e}^{-\frac{1}{2}(X-\mu)V^{-1}(X-\mu)'},\ V > O. \qquad (4.2.4)$$

On the other hand, when $q = 1$, a real p-variate Gaussian or normal density is available from (4.2.1) wherein $B = 1$; in this case, $X \sim N_p(\mu, A^{-1})$ where X and the location parameter vector μ are now $p \times 1$ column vectors. This density is given by

$$f_{p,1}(X) = \frac{|A|^{\frac{1}{2}}}{(2\pi)^{\frac{p}{2}}} \mathrm{e}^{-\frac{1}{2}(X-\mu)'A(X-\mu)}, \ A > O. \tag{4.2.5}$$

Example 4.2.1. Write down the exponent and the normalizing constant explicitly in a real matrix-variate Gaussian density where

$$X = \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \end{bmatrix}, \ E[X] = M = \begin{bmatrix} 1 & 0 & -1 \\ -1 & -2 & 0 \end{bmatrix},$$
$$A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}, \ B = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 3 \end{bmatrix},$$

where the x_{ij}'s are real scalar random variables.

Solution 4.2.1. Note that $A = A'$ and $B = B'$, the leading minors in A being $|(1)| = 1 > 0$ and $|A| = 1 > 0$ so that $A > O$. The leading minors in B are $|(1)| = 1 > 0$, $\begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix} = 1 > 0$ and

$$|B| = (1)\begin{vmatrix} 2 & 1 \\ 1 & 3 \end{vmatrix} - (1)\begin{vmatrix} 1 & 1 \\ 1 & 3 \end{vmatrix} + (1)\begin{vmatrix} 1 & 2 \\ 1 & 1 \end{vmatrix} = 2 > 0,$$

and hence $B > O$. The density is of the form

$$f_{p,q}(X) = \frac{|A|^{\frac{q}{2}}|B|^{\frac{p}{2}}}{(2\pi)^{\frac{pq}{2}}} \mathrm{e}^{-\frac{1}{2}\mathrm{tr}(A(X-M)B(X-M)')}$$

where the normalizing constant is $\frac{(1)^{\frac{3}{2}}(2)^{\frac{2}{2}}}{(2\pi)^{\frac{(2)(3)}{2}}} = \frac{2}{(2\pi)^3} = \frac{1}{4\pi^3}$. Let X_1 and X_2 be the two rows of X and let $Y = X - M = \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}$. Then $Y_1 = (y_{11}, y_{12}, y_{13}) = (x_{11} - 1, x_{12}, x_{13} + 1)$, $Y_2 = (y_{21}, y_{22}, y_{23}) = (x_{21} + 1, x_{22} + 2, x_{23})$. Now

$$(X - M)B(X - M)' = \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} B[Y_1', Y_2'] = \begin{bmatrix} Y_1BY_1' & Y_1BY_2' \\ Y_2BY_1' & Y_2BY_2' \end{bmatrix},$$

$$\begin{aligned} A(X - M)B(X - M)' &= \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} Y_1BY_1' & Y_1BY_2' \\ Y_2BY_1' & Y_2BY_2' \end{bmatrix} \\ &= \begin{bmatrix} Y_1BY_1' + Y_2BY_1' & Y_1BY_2' + Y_2BY_2' \\ Y_1BY_1' + 2Y_2BY_1' & Y_1BY_2' + 2Y_2BY_2' \end{bmatrix}. \end{aligned}$$

Thus,

$$\begin{aligned}\mathrm{tr}[A(X-M)B(X-M)'] &= Y_1BY_1' + Y_2BY_1' + Y_1BY_2' + 2Y_2BY_2'\\ &= Y_1BY_1' + 2Y_1BY_2' + 2Y_2BY_2', \equiv Q, \end{aligned}\qquad (i)$$

noting that Y_1BY_2' and Y_2BY_1' are equal since both are real scalar quantities and one is the transpose of the other. Here are now the detailed computations of the various items:

$$Y_1BY_1' = y_{11}^2 + 2y_{11}y_{12} + 2y_{11}y_{13} + 2y_{12}^2 + 2y_{12}y_{13} + 3y_{13}^2 \qquad (ii)$$

$$Y_2BY_2' = y_{21}^2 + 2y_{21}y_{22} + 2y_{21}y_{23} + 2y_{22}^2 + 2y_{22}y_{23} + 3y_{23}^2 \qquad (iii)$$

$$\begin{aligned}Y_1BY_2' &= y_{11}y_{21} + y_{11}y_{22} + y_{11}y_{23} + y_{12}y_{21} + 2y_{12}y_{22} + y_{12}y_{23}\\ &\quad + y_{13}y_{21} + y_{13}y_{22} + 3y_{13}y_{33}\end{aligned} \qquad (iv)$$

where the y_{1j}'s and y_{2j}'s and the various quadratic and bilinear forms are as specified above. The density is then

$$f_{2,3}(X) = \frac{1}{4\pi^3}\,\mathrm{e}^{-\frac{1}{2}(Y_1BY_1'+2Y_1BY_2'+Y_2BY_2')}$$

where the terms in the exponent are given in (*ii*)-(*iv*). This completes the computations.

4.2a. The Matrix-variate Gaussian Density, Complex Case

In the following discussion, the absolute value of a determinant will be denoted by $|\det(A)|$ where A is a square matrix. For example, if $\det(A) = a + ib$ with a and b real scalar and $i = \sqrt{(-1)}$, the determinant of the conjugate transpose of A is $\det(A^*) = a - ib$. Then the absolute value of the determinant is

$$|\det(A)| = +\sqrt{(a^2+b^2)} = +[(a+ib)(a-ib)]^{\frac{1}{2}} = +[\det(A)\det(A^*)]^{\frac{1}{2}} = +[\det(AA^*)]^{\frac{1}{2}}. \qquad (4.2a.1)$$

The matrix-variate Gaussian density in the complex case, which is the counterpart to that given in (4.2.1) for the real case, is

$$\tilde{f}_{p,q}(\tilde{X}) = \frac{|\det(A)|^q|\det(B)|^p}{\pi^{pq}}\mathrm{e}^{-\mathrm{tr}(A\tilde{X}B\tilde{X}^*)} \qquad (4.2a.2)$$

for $A > O,\ B > O,\ \tilde{X} = (\tilde{x}_{ij})$, $|(\cdot)|$ denoting the absolute value of $(\cdot)$. When $p = 1$ and $A = 1$, the usual multivariate Gaussian density in the complex domain is obtained:

$$\tilde{f}_{1,q}(\tilde{X}) = \frac{|\det(B)|}{\pi^q}\mathrm{e}^{-(\tilde{X}-\mu)B(\tilde{X}-\mu)^*},\ \tilde{X}' \sim \tilde{N}_q(\tilde{\mu}', B^{-1}) \qquad (4.2a.3)$$

where $B > O$ and $\tilde{X}$ and μ are $1 \times q$ row vectors, μ being a location parameter vector. When $q = 1$ in (4.2*a*.1), we have the p-variate Gaussian or normal density in the complex case which is given by

$$\tilde{f}_{p,1}(\tilde{X}) = \frac{|\det(A)|}{\pi^p} e^{-(\tilde{X}-\mu)^* A(\tilde{X}-\mu)}, \ \tilde{X} \sim \tilde{N}_p(\mu, A^{-1}) \tag{4.2a.4}$$

where $\tilde{X}$ and the location parameter also denoted by μ are now $p \times 1$ vectors.

Example 4.2a.1. Consider a 2×3 complex matrix-variate Gaussian density. Write down the normalizing constant and the exponent explicitly if

$$\tilde{X} = \begin{bmatrix} \tilde{x}_{11} & \tilde{x}_{12} & \tilde{x}_{13} \\ \tilde{x}_{21} & \tilde{x}_{22} & \tilde{x}_{23} \end{bmatrix}, \ E[\tilde{X}] = \tilde{M} = \begin{bmatrix} i & -i & 1+i \\ 0 & 1-i & 1 \end{bmatrix},$$
$$A = \begin{bmatrix} 3 & 1+i \\ 1-i & 2 \end{bmatrix}, \ B = \begin{bmatrix} 4 & 1+i & i \\ 1-i & 2 & 1-i \\ -i & 1+i & 3 \end{bmatrix},$$

where the $\tilde{x}_{ij}$'s are scalar complex random variables.

Solution 4.2a.1. Let us verify the definiteness of A and B. It is obvious that $A = A^*$, $B = B^*$ and hence they are Hermitian. The leading minors of A are $|(3)| = 3 > 0$, $|A| = 4 > 0$ and hence $A > O$. The leading minors of B are $|(4)| = 4 > 0$, $\begin{vmatrix} 4 & 1+i \\ 1-i & 2 \end{vmatrix} = 6 > 0$,

$$|B| = 4 \begin{vmatrix} 2 & 1-i \\ 1+i & 3 \end{vmatrix} - (1+i) \begin{vmatrix} 1-i & 1-i \\ -i & 3 \end{vmatrix} + i \begin{vmatrix} 1-i & 2 \\ -i & 1+i \end{vmatrix} = 8 > 0,$$

and hence $B > O$. The normalizing constant is then

$$\frac{|\det(A)|^q |\det(B)|^p}{\pi^{pq}} = \frac{(4^3)(8^2)}{\pi^6}.$$

Let the two rows of $\tilde{X}$ be $\tilde{X}_1$ and $\tilde{X}_2$. Let $(\tilde{X} - \tilde{M}) = \tilde{Y} = \begin{bmatrix} \tilde{Y}_1 \\ \tilde{Y}_2 \end{bmatrix}$,

$$\tilde{Y}_1 = (\tilde{y}_{11}, \tilde{y}_{12}, \tilde{y}_{13}) = (\tilde{x}_{11} - i, \tilde{x}_{12} + i, \tilde{x}_{13} - (1+i))$$
$$\tilde{Y}_2 = (\tilde{y}_{21}, \tilde{y}_{22}, \tilde{y}_{23}) = (\tilde{x}_{21}, \tilde{x}_{22} - (1-i), \tilde{x}_{23} - 1).$$

$$(\tilde{X}-\tilde{M})B(\tilde{X}-\tilde{M})^{*}=\tilde{Y}B\tilde{Y}^{*}=\begin{bmatrix}\tilde{Y}_1\\ \tilde{Y}_2\end{bmatrix}B[\tilde{Y}_1^{*},\tilde{Y}_2^{*}]$$

$$=\begin{bmatrix}\tilde{Y}_1B\tilde{Y}_1^{*} & \tilde{Y}_1B\tilde{Y}_2^{*}\\ \tilde{Y}_2B\tilde{Y}_1^{*} & \tilde{Y}_2B\tilde{Y}_2^{*}\end{bmatrix}.$$

Then,

$$\begin{aligned}\mathrm{tr}[A(\tilde{X}-\tilde{M})B(\tilde{X}-\tilde{M})^{*}]&=\mathrm{tr}\left\{\begin{bmatrix}3 & 1+i\\ 1-i & 2\end{bmatrix}\begin{bmatrix}\tilde{Y}_1B\tilde{Y}_1^{*} & \tilde{Y}_1B\tilde{Y}_2^{*}\\ \tilde{Y}_2B\tilde{Y}_1^{*} & \tilde{Y}_2B\tilde{Y}_2^{*}\end{bmatrix}\right\}\\ &=3\tilde{Y}_1B\tilde{Y}_1^{*}+(1+i)(\tilde{Y}_2B\tilde{Y}_1^{*})+(1-i)(\tilde{Y}_1B\tilde{Y}_2^{*})+2\tilde{Y}_2B\tilde{Y}_2^{*}\\ &\equiv Q \qquad (i)\end{aligned}$$

where

$$\begin{aligned}\tilde{Y}_1B\tilde{Y}_1^{*}&=4\tilde{y}_{11}\tilde{y}_{11}^{*}+2\tilde{y}_{12}\tilde{y}_{12}^{*}+3\tilde{y}_{13}\tilde{y}_{13}^{*}\\ &\quad+(1+i)\tilde{y}_{11}\tilde{y}_{12}^{*}+i\,\tilde{y}_{11}\tilde{y}_{13}^{*}+(1-i)\tilde{y}_{12}\tilde{y}_{11}^{*}\\ &\quad+(1-i)\tilde{y}_{12}\tilde{y}_{13}^{*}-i\,\tilde{y}_{13}\tilde{y}_{11}^{*}+(1+i)\tilde{y}_{13}\tilde{y}_{12}^{*} \qquad (ii)\end{aligned}$$

$$\begin{aligned}\tilde{Y}_2B\tilde{Y}_2^{*}&=4\tilde{y}_{21}\tilde{y}_{21}^{*}+2\tilde{y}_{22}\tilde{y}_{22}^{*}+3\tilde{y}_{23}\tilde{y}_{23}^{*}\\ &\quad+(1+i)\tilde{y}_{21}\tilde{y}_{22}^{*}+i\,\tilde{y}_{21}\tilde{y}_{23}^{*}+(1-i)\tilde{y}_{22}\tilde{y}_{21}^{*}\\ &\quad+(1-i)\tilde{y}_{22}\tilde{y}_{23}^{*}-i\,\tilde{y}_{23}\tilde{y}_{21}^{*}+(1+i)\tilde{y}_{23}\tilde{y}_{22}^{*} \qquad (iii)\end{aligned}$$

$$\begin{aligned}\tilde{Y}_1B\tilde{Y}_2^{*}&=4\tilde{y}_{11}\tilde{y}_{21}^{*}+2\tilde{y}_{12}\tilde{y}_{22}^{*}+3\tilde{y}_{13}\tilde{y}_{23}^{*}\\ &\quad+(1+i)\tilde{y}_{11}\tilde{y}_{22}^{*}+i\,\tilde{y}_{11}\tilde{y}_{23}^{*}+(1-i)\tilde{y}_{12}\tilde{y}_{21}^{*}\\ &\quad+(1-i)\tilde{y}_{12}\tilde{y}_{23}^{*}-i\,\tilde{y}_{13}\tilde{y}_{21}^{*}+(1+i)\tilde{y}_{13}\tilde{y}_{22}^{*} \qquad (iv)\end{aligned}$$

$$\tilde{Y}_2B\tilde{Y}_1^{*}=(iv)\text{ with }\tilde{y}_{1j}\text{ and }\tilde{y}_{2j}\text{ interchanged.} \qquad (v)$$

Hence, the density of $\tilde{X}$ is given by

$$\tilde{f}_{2,3}(\tilde{X})=\frac{(4^3)(8^2)}{\pi^6}\,\mathrm{e}^{-Q}$$

where Q is given explicitly in (i)-(v) above. This completes the computations.

4.2.1. Some properties of a real matrix-variate Gaussian density

In order to derive certain properties, we will need some more Jacobians of matrix transformations, in addition to those provided in Chap. 1. These will be listed in this section as basic results without proofs. The derivations as well as other related Jacobians are available from Mathai (1997).

Theorem 4.2.1. *Let X be a $p\times q$, $q\geq p$, real matrix of rank p, that is, X has full rank, where the pq elements of X are distinct real scalar variables. Let $X=TU_1$ where T is a $p\times p$ real lower triangular matrix whose diagonal elements are positive and U_1 is a semi-orthonormal matrix such that $U_1U_1'=I_p$. Then*

$$\mathrm{d}X=\left\{\prod_{j=1}^{p}t_{jj}^{q-j}\right\}\mathrm{d}T\;h(U_1) \tag{4.2.6}$$

where $h(U_1)$ is the differential element corresponding to U_1.

Theorem 4.2.2. *For the differential elements $h(U_1)$ in (4.2.6), the integral is over the Stiefel manifold $V_{p,q}$ or over the space of $p\times q$, $q\geq p$, semi-orthonormal matrices and the integral over the full orthogonal group O_p when $q=p$ are respectively*

$$\int_{V_{p,q}}h(U_1)=\frac{2^p\pi^{\frac{pq}{2}}}{\Gamma_p(\frac{q}{2})}\text{ and }\int_{O_p}h(U_1)=\frac{2^p\pi^{\frac{p^2}{2}}}{\Gamma_p(\frac{p}{2})} \tag{4.2.7}$$

where $\Gamma_p(\alpha)$ is the real matrix-variate gamma function given by

$$\Gamma_p(\alpha)=\pi^{\frac{p(p-1)}{4}}\Gamma(\alpha)\Gamma(\alpha-1/2)\cdots\Gamma(\alpha-(p-1)/2),\ \Re(\alpha)>\tfrac{p-1}{2}, \tag{4.2.8}$$

$\Re(\cdot)$ denoting the real part of $(\cdot)$.

For example,

$$\Gamma_3(\alpha)=\pi^{\frac{3(2)}{4}}\Gamma(\alpha)\Gamma(\alpha-1/2)\Gamma(\alpha-1)=\pi^{\frac{3}{2}}\Gamma(\alpha)\Gamma(\alpha-1/2)\Gamma(\alpha-1),\ \Re(\alpha)>1.$$

With the help of Theorems 4.2.1, 4.2.2 and 1.6.7 of Chap. 1, we can derive the following result:

Theorem 4.2.3. *Let X be a real $p\times q$, $q\geq p$, matrix of rank p and $S=XX'$. Then, $S>O$ (real positive definite) and*

$$\mathrm{d}X=\frac{\pi^{\frac{pq}{2}}}{\Gamma_p(\frac{q}{2})}|S|^{\frac{q}{2}-\frac{p+1}{2}}\mathrm{d}S, \tag{4.2.9}$$

after integrating out over the Stiefel manifold.

4.2a.1. Some properties of a complex matrix-variate Gaussian density

The corresponding results in the complex domain follow.

Theorem 4.2a.1. *Let $\tilde{X}$ be a $p\times q,\ q\ge p$, matrix of rank p in the complex domain and $\tilde{T}$ be a $p\times p$ lower triangular matrix in the complex domain whose diagonal elements $t_{jj}>0,\ j=1,\ldots,p$, are real and positive. Then, letting $\tilde{U}_1$ be a semi-unitary matrix such that $\tilde{U}_1\tilde{U}_1^{*}=I_p$,*

$$\tilde{X}=\tilde{T}\tilde{U}_1\Rightarrow \mathrm{d}\tilde{X}=\Big\{\prod_{j=1}^{p}t_{jj}^{2(q-j)+1}\Big\}\,\mathrm{d}\tilde{T}\ \tilde{h}(\tilde{U}_1) \tag{4.2a.5}$$

where $\tilde{h}(\tilde{U}_1)$ is the differential element corresponding to $\tilde{U}_1$.

When integrating out $\tilde{h}(\tilde{U}_1)$, there are three situations to be considered. One of the cases is $q>p$. When $q=p$, the integration is done over the full unitary group $\tilde{O}_p$; however, there are two cases to be considered in this instance. One case occurs where all the elements of the unitary matrix $\tilde{U}_1$, including the diagonal ones, are complex, in which case $\tilde{O}_p$ will be denoted by $\tilde{O}_p^{(1)}$, and the other one, wherein the diagonal elements of $\tilde{U}_1$ are real, in which instance the unitary group will be denoted by $\tilde{O}_p^{(2)}$. When unitary transformations are applied to Hermitian matrices, this is our usual situations when Hermitian matrices are involved, then the diagonal elements of the unique $\tilde{U}_1$ are real and hence the unitary group is $\tilde{O}_p^{(2)}$. The integral of $\tilde{h}(\tilde{U}_1)$ under these three cases are given in the next theorem.

Theorem 4.2a.2. *Let $\tilde{h}(\tilde{U}_1)$ be as defined in equation (4.2a.5). Then, the integral of $\tilde{h}(\tilde{U}_1)$, over the Stiefel manifold $\tilde{V}_{p,q}$ of semi-unitary matrices for $q>p$, and when $q=p$, the integrals over the unitary groups $\tilde{O}_p^{(1)}$ and $\tilde{O}_p^{(2)}$ are the following:*

$$\int_{\tilde{V}_{p,q}}\tilde{h}(\tilde{U}_1)=\frac{2^p\pi^{pq}}{\tilde{\Gamma}_p(q)},\ q>p;$$

$$\int_{\tilde{O}_p^{(1)}}\tilde{h}(\tilde{U}_1)=\frac{2^p\pi^{p^2}}{\tilde{\Gamma}_p(p)},\quad \int_{\tilde{O}_p^{(2)}}\tilde{h}(\tilde{U}_1)=\frac{\pi^{p(p-1)}}{\tilde{\Gamma}_p(p)}, \tag{4.2a.6}$$

the factor 2^p being omitted when $\tilde{U}_1$ is uniquely specified; $\tilde{O}_p^{(1)}$ is the case of a general $\tilde{X}$, $\tilde{O}_p^{(2)}$ is the case corresponding to $\tilde{X}$ Hermitian, and $\tilde{\Gamma}_p(\alpha)$ is the complex matrix-variate gamma, given by

$$\tilde{\Gamma}_p(\alpha)=\pi^{\frac{p(p-1)}{2}}\Gamma(\alpha)\Gamma(\alpha-1)\cdots\Gamma(\alpha-p+1),\ \Re(\alpha)>p-1. \tag{4.2a.7}$$

For example,

$$\tilde{\Gamma}_3(\alpha) = \pi^{\frac{3(2)}{2}}\Gamma(\alpha)\Gamma(\alpha-1)\Gamma(\alpha-2) = \pi^3\Gamma(\alpha)\Gamma(\alpha-1)\Gamma(\alpha-2),\ \Re(\alpha) > 2.$$

Theorem 4.2a.3. *Let $\tilde{X}$ be $p \times q,\ q \geq p$, matrix of rank p in the complex domain and $\tilde{S} = \tilde{X}\tilde{X}^* > O$. Then after integrating out over the Stiefel manifold,*

$$\mathrm{d}\tilde{X} = \frac{\pi^{pq}}{\tilde{\Gamma}_p(q)}|\mathrm{det}(\tilde{S})|^{q-p}\mathrm{d}\tilde{S}. \tag{4.2a.8}$$

4.2.2. Additional properties in the real and complex cases

On making use of the above results, we will establish a few results in this section as well as additional ones later on. Let us consider the matrix-variate Gaussian densities corresponding to (4.2.1) and (4.2a.2) with location matrices M and $\tilde{M}$, respectively, and let the densities be again denoted by $f_{p,q}(X)$ and $\tilde{f}_{p,q}(\tilde{X})$ respectively, where

$$f_{p,q}(X) = \frac{|A|^{\frac{q}{2}}|B|^{\frac{p}{2}}}{(2\pi)^{\frac{pq}{2}}}\mathrm{e}^{-\frac{1}{2}\mathrm{tr}[A(X-M)B(X-M)']} \tag{4.2.10}$$

and

$$\tilde{f}_{p,q}(\tilde{X}) = \frac{|\mathrm{det}(A)|^q|\mathrm{det}(B)|^p}{\pi^{pq}}\mathrm{e}^{-\mathrm{tr}[A(\tilde{X}-\tilde{M})B(\tilde{X}-\tilde{M})^*]}. \tag{4.2a.9}$$

Then, in the real case the expected value of X or the mean value of X, denoted by $E(X)$, is given by

$$E(X) = \int_X Xf_{p,q}(X)\,\mathrm{d}X = \int_X (X-M)f_{p,q}(X)\,\mathrm{d}X + M\int_X f_{p,q}(X)\,\mathrm{d}X. \tag{i}$$

The second integral in *(i)* is the total integral in a density, which is 1, and hence the second integral gives M. On making the transformation $Y = A^{\frac{1}{2}}(X-M)B^{\frac{1}{2}}$, we have

$$E[X] = M + A^{-\frac{1}{2}}\frac{1}{(2\pi)^{\frac{np}{2}}}\int_Y Y\mathrm{e}^{-\frac{1}{2}\mathrm{tr}(YY')}\mathrm{d}Y\,B^{-\frac{1}{2}}. \tag{ii}$$

But $\mathrm{tr}(YY')$ is the sum of squares of all elements in Y. Hence $Y\mathrm{e}^{-\frac{1}{2}\mathrm{tr}(YY')}$ is an odd function and the integral over each element in Y is convergent, so that each integral is zero. Thus, the integral over Y gives a null matrix. Therefore $E(X) = M$. It can be shown in a similar manner that $E(\tilde{X}) = \tilde{M}$.

Theorem 4.2.4, 4.2a.4. *For the densities specified in (4.2.10) and (4.2a.9),*

$$E(X) = M \text{ and } E(\tilde{X}) = \tilde{M}. \tag{4.2.11}$$

Theorem 4.2.5, 4.2a.5. *For the densities given in (4.2.10), (4.2a.9)*

$$E[(X-M)B(X-M)'] = qA^{-1},\ E[(X-M)'A(X-M)] = pB^{-1} \tag{4.2.12}$$

and

$$E[(\tilde{X}-\tilde{M})B(\tilde{X}-\tilde{M})^*] = qA^{-1},\ E[(\tilde{X}-\tilde{M})^*A(\tilde{X}-\tilde{M})] = pB^{-1}. \tag{4.2a.10}$$

***Proof*:** Consider the real case first. Let $Y = A^{\frac{1}{2}}(X-M)B^{\frac{1}{2}} \Rightarrow A^{-\frac{1}{2}}Y = (X-M)B^{\frac{1}{2}}$. Then

$$E[(X-M)B(X-M)'] = \frac{A^{-\frac{1}{2}}}{(2\pi)^{\frac{pq}{2}}}\int_Y YY'\mathrm{e}^{-\frac{1}{2}\mathrm{tr}(YY')}\mathrm{d}Y A^{-\frac{1}{2}}. \tag{i}$$

Note that Y is $p\times q$ and YY' is $p\times p$. The non-diagonal elements in YY' are dot products of the distinct row vectors in Y and hence linear functions of the elements of Y. The diagonal elements in YY' are sums of squares of elements in the rows of Y. The exponent has all sum of squares and hence the convergent integrals corresponding to all the non-diagonal elements in YY' are zeros. Hence, only the diagonal elements need be considered. Each diagonal element is a sum of squares of q elements of Y. For example, the first diagonal element in YY' is $y_{11}^2 + y_{12}^2 + \cdots + y_{1q}^2$ where $Y = (y_{ij})$. Let $Y_1 = (y_{11}, \ldots, y_{1q})$ be the first row of Y and let $s = Y_1Y_1' = y_{11}^2 + \cdots + y_{1q}^2$. It follows from Theorem 4.2.3 that when $p = 1$,

$$\mathrm{d}Y_1 = \frac{\pi^{\frac{q}{2}}}{\Gamma(\frac{q}{2})}s^{\frac{q}{2}-1}\mathrm{d}s. \tag{ii}$$

Then

$$\int_{Y_1} Y_1Y_1'\mathrm{e}^{-\frac{1}{2}Y_1Y_1'}\mathrm{d}Y_1 = \int_{s=0}^{\infty} s\frac{\pi^{\frac{q}{2}}}{\Gamma(\frac{q}{2})}s^{\frac{q}{2}-1}\mathrm{e}^{-\frac{1}{2}s}\mathrm{d}s. \tag{iii}$$

The integral part over s is $2^{\frac{q}{2}+1}\Gamma(\frac{q}{2}+1) = 2^{\frac{q}{2}+1}\frac{q}{2}\Gamma(\frac{q}{2}) = 2^{\frac{q}{2}}q\Gamma(\frac{q}{2})$. Thus $\Gamma(\frac{q}{2})$ is canceled and $(2\pi)^{\frac{q}{2}}$ cancels with $(2\pi)^{\frac{pq}{2}}$ leaving $(2\pi)^{\frac{(p-1)q}{2}}$ in the denominator and q in the numerator. We still have $p-1$ such sets of q, y_{ij}^2's in the exponent in *(i)* and each such integrals is of the form $\int_{-\infty}^{\infty} e^{-\frac{1}{2}z^2}dz = \sqrt{(2\pi)}$ which gives $(2\pi)^{\frac{(p-1)q}{2}}$ and thus the factor containing π is also canceled leaving only q at each diagonal position in YY'. Hence the integral $\frac{1}{(2\pi)^{\frac{pq}{2}}}\int_Y YY' e^{-\frac{1}{2}\text{tr}(YY')}dY = qI$ where I is the identity matrix, which establishes one of the results in (4.2.12). Now, write

$$\text{tr}[A(X-M)B(X-M)'] = \text{tr}[(X-M)'A(X-M)B] = \text{tr}[B(X-M)'A(X-M)].$$

This is the same structure as in the previous case where B occupies the place of A and the order is now q in place of p in the previous case. Then, proceeding as in the derivations from (i) to (iii), the second result in (4.2.12) follows. The results in (4.2a.10) are established in a similar manner.

From (4.2.10), it is clear that the density of Y, denoted by $g(Y)$, is of the form

$$g(Y) = \frac{1}{(2\pi)^{\frac{pq}{2}}} e^{-\frac{1}{2}\text{tr}(YY')},\ Y = (y_{ij}),\ -\infty < y_{ij} < \infty, \tag{4.2.13}$$

for all i and j. The individual y_{ij}'s are independently distributed and each y_{ij} has the density

$$g_{ij}(y_{ij}) = \frac{1}{\sqrt{(2\pi)}} e^{-\frac{1}{2}y_{ij}^2},\ -\infty < y_{ij} < \infty. \tag{iv}$$

Thus, we have a real standard normal density for y_{ij}. The complex case corresponding to (4.2.13), denoted by $\tilde{g}(\tilde{Y})$, is given by

$$\tilde{g}(\tilde{Y}) = \frac{1}{\pi^{pq}} e^{-\text{tr}(\tilde{Y}\tilde{Y}^*)}. \tag{4.2a.11}$$

In this case, the exponent is $\text{tr}(\tilde{Y}\tilde{Y}^*) = \sum_{i=1}^{p}\sum_{j=1}^{q} |\tilde{y}_{ij}|^2$ where $\tilde{y}_{rs} = y_{rs1} + iy_{rs2},\ y_{rs1},$ y_{rs2} real, $i = \sqrt{(-1)}$ and $|\tilde{y}_{rs}|^2 = y_{rs1}^2 + y_{rs2}^2$.

For the real case, consider the probability that $y_{ij} \le t_{ij}$ for some given t_{ij} and this is the distribution function of y_{ij}, which is denoted by $F_{y_{ij}}(t_{ij})$. Then, let us compute the density of y_{ij}^2. Consider the probability that $y_{ij}^2 \le u,\ u > 0$ for some u. Let $u_{ij} = y_{ij}^2$. Then, $Pr\{u_{ij} \le v_{ij}\}$ for some v_{ij} is the distribution function of u_{ij} evaluated at v_{ij}, denoted by $F_{u_{ij}}(v_{ij})$. Consider

$$Pr\{y_{ij}^2 \le t,\ t > 0\} = Pr\{|y_{ij}| \le \sqrt{t}\} = Pr\{-\sqrt{t} \le y_{ij} \le \sqrt{t}\} = F_{y_{ij}}(\sqrt{t}) - F_{y_{ij}}(-\sqrt{t}). \tag{v}$$

Differentiate throughout with respect to t. When $Pr\{y_{ij}^2 \le t\}$ is differentiated with respect to t, we obtain the density of $u_{ij} = y_{ij}^2$, evaluated at t. This density, denoted by $h_{ij}(u_{ij})$, is given by

$$
\begin{aligned}
h_{ij}(u_{ij})|_{u_{ij}=t} &= \frac{\mathrm{d}}{\mathrm{d}t}F_{y_{ij}}(\sqrt{t}) - \frac{\mathrm{d}}{\mathrm{d}t}F(-\sqrt{t})\\
&= g_{ij}(y_{ij}=t)\tfrac{1}{2}t^{\frac{1}{2}-1} - g_{ij}(y_{ij}=t)(-\tfrac{1}{2}t^{\frac{1}{2}-1})\\
&= \frac{1}{\sqrt{(2\pi)}}[t^{\frac{1}{2}-1}\mathrm{e}^{-\frac{1}{2}t}] = \frac{1}{\sqrt{(2\pi)}}[u_{ij}^{\frac{1}{2}-1}\mathrm{e}^{-\frac{1}{2}u_{ij}}] \qquad (vi)
\end{aligned}
$$

evaluated at $u_{ij} = t$ for $0 \le t < \infty$. Hence we have the following result:

Theorem 4.2.6. *Consider the density $f_{p,q}(X)$ in (4.2.1) and the transformation $Y = A^{\frac{1}{2}}XB^{\frac{1}{2}}$. Letting $Y = (y_{ij})$, the y_{ij}'s are mutually independently distributed as in (iv) above and each y_{ij}^2 is distributed as a real chi-square random variable having one degree of freedom or equivalently a real gamma with parameters $\alpha = \frac{1}{2}$ and $\beta = 2$ where the usual real scalar gamma density is given by*

$$
f(z) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)}z^{\alpha-1}\mathrm{e}^{-\frac{z}{\beta}}, \qquad (vii)
$$

for $0 \le z < \infty,\ \Re(\alpha) > 0,\ \Re(\beta) > 0$ and $f(z) = 0$ elsewhere.

As a consequence of the y_{ij}^2's being independently gamma distributed, $\sum_{j=1}^{q} y_{ij}^2$ is real gamma distributed with the parameters $\alpha = \frac{q}{2}$ and $\beta = 2$. Then $\mathrm{tr}(YY')$ is real gamma distributed with the parameters $\alpha = \frac{pq}{2}$ and $\beta = 2$ and each diagonal element in YY' is real gamma distributed with parameters $\frac{q}{2}$ and $\beta = 2$ or a real chi-square variable with q degrees of freedom and an expected value $2\frac{q}{2} = q$. This is an alternative way of proving (4.2.12). Proofs for the other results in (4.2.12) and (4.2a.10) are parallel and hence are omitted.

4.2.3. Some special cases

Consider the real $p \times q$ matrix-variate Gaussian case where the exponent in the density is $-\frac{1}{2}\mathrm{tr}(AXBX')$. On making the transformation $A^{\frac{1}{2}}X = Z \Rightarrow \mathrm{d}Z = |A|^{\frac{q}{2}}\mathrm{d}X$, Z has a $p \times q$ matrix-variate Gaussian density of the form

$$
f_{p,q}(Z) = \frac{|B|^{\frac{p}{2}}}{(2\pi)^{\frac{pq}{2}}}\mathrm{e}^{-\frac{1}{2}\mathrm{tr}(ZBZ')}. \qquad (4.2.14)
$$

If the distribution has a $p \times q$ constant matrix M as location parameter, then replace Z by $Z - M$ in (4.2.14), which does not affect the normalizing constant. Letting $Z_1, Z_2, \ldots, Z_p$ denote the rows of Z, we observe that Z_j has a q-variate multinormal distribution with the null vector as its mean value and B^{-1} as its covariance matrix for each $j = 1, \ldots, p$. This can be seen from the considerations that follow. Let us consider the transformation $Y = ZB^{\frac{1}{2}} \Rightarrow \mathrm{d}Z = |B|^{-\frac{p}{2}}\mathrm{d}Y$. The density in (4.2.14) then reduces to the following, denoted by $f_{p,q}(Y)$:

$$f_{p,q}(Y) = \frac{1}{(2\pi)^{\frac{pq}{2}}}\mathrm{e}^{-\frac{1}{2}\mathrm{tr}(YY')}. \tag{4.2.15}$$

This means that each element y_{ij} in $Y = (y_{ij})$ is a real univariate standard normal variable, $y_{ij} \sim N_1(0, 1)$ as per the usual notation, and all the y_{ij}'s are mutually independently distributed. Letting the p rows of Y be $Y_1, \ldots, Y_p$, then each Y_j is a q-variate standard normal vector for $j = 1, \ldots, p$. Letting the density of Y_j be denoted by $f_{Y_j}(Y_j)$, we have

$$f_{Y_j}(Y_j) = \frac{1}{(2\pi)^{\frac{q}{2}}}\mathrm{e}^{-\frac{1}{2}(Y_jY_j')}.$$

Now, consider the transformation $Z_j = Y_jB^{-\frac{1}{2}} \Rightarrow \mathrm{d}Y_j = |B|^{\frac{1}{2}}\mathrm{d}Z_j$ and $Y_j = Z_jB^{\frac{1}{2}}$. That is, $Y_jY_j' = Z_jBZ_j'$ and the density of Z_j denoted by $f_{Z_j}(Z_j)$ is as follows:

$$f_{Z_j}(Z_j) = \frac{|B|^{\frac{1}{2}}}{(2\pi)^{\frac{q}{2}}}\mathrm{e}^{-\frac{1}{2}(Z_jBZ_j')},\ B > O, \tag{4.2.16}$$

which is a q-variate real multinormal density with the covariance matrix of Z_j given by B^{-1}, for each $j = 1, \ldots, p$, and the Z_j's, $j = 1, \ldots, p$, are mutually independently distributed. Thus, the following result:

Theorem 4.2.7. *Let $Z_1, \ldots, Z_p$ be the p rows of the $p \times q$ matrix Z in (4.2.14). Then each Z_j has a q-variate real multinormal distribution with the covariance matrix B^{-1}, for $j = 1, \ldots, p$, and $Z_1, \ldots, Z_p$ are mutually independently distributed.*

Observe that the exponent in the original real $p \times q$ matrix-variate Gaussian density can also be rewritten in the following format:

$$\begin{aligned}-\frac{1}{2}\mathrm{tr}(AXBX') &= -\frac{1}{2}\mathrm{tr}(X'AXB) = -\frac{1}{2}\mathrm{tr}(BX'AX)\\ &= -\frac{1}{2}\mathrm{tr}(U'AU) = -\frac{1}{2}\mathrm{tr}(ZBZ'),\ A^{\frac{1}{2}}X = Z,\ XB^{\frac{1}{2}} = U.\end{aligned}$$

Now, on making the transformation $U = XB^{\frac{1}{2}} \Rightarrow \mathrm{d}X = |B|^{-\frac{p}{2}}\mathrm{d}U$, the density of U, denoted by $f_{p,q}(U)$, is given by

$$f_{p,q}(U) = \frac{|A|^{\frac{q}{2}}}{(2\pi)^{\frac{pq}{2}}}\mathrm{e}^{-\frac{1}{2}\mathrm{tr}(U'AU)}. \tag{4.2.17}$$

Proceeding as in the derivation of Theorem 4.2.7, we have the following result:

Theorem 4.2.8. *Consider the $p \times q$ real matrix U in (4.2.17). Let $U_1, \ldots, U_q$ be the columns of U. Then, $U_1, \ldots, U_q$ are mutually independently distributed with U_j having a p-variate multinormal density, denoted by $f_{U_j}(U_j)$, given as*

$$f_{U_j}(U_j) = \frac{|A|^{\frac{1}{2}}}{(2\pi)^{\frac{p}{2}}} \mathrm{e}^{-\frac{1}{2}(U_j' A U_j)}. \tag{4.2.18}$$

The corresponding results in the $p \times q$ complex Gaussian case are the following:

Theorem 4.2a.6. *Consider the $p \times q$ complex Gaussian matrix $\tilde{X}$. Let $A^{\frac{1}{2}}\tilde{X} = \tilde{Z}$ and $\tilde{Z}_1, \ldots, \tilde{Z}_p$ be the rows of $\tilde{Z}$. Then, $\tilde{Z}_1, \ldots, \tilde{Z}_p$ are mutually independently distributed with $\tilde{Z}_j$ having a q-variate complex multinormal density, denoted by $\tilde{f}_{\tilde{Z}_j}(\tilde{Z}_j)$, given by*

$$\tilde{f}_{\tilde{Z}_j}(\tilde{Z}_j) = \frac{|\mathrm{det}(B)|}{\pi^q} \mathrm{e}^{-(\tilde{Z}_j B \tilde{Z}_j^*)}. \tag{4.2a.12}$$

Theorem 4.2a.7. *Let the $p \times q$ matrix $\tilde{X}$ have a complex matrix-variate distribution. Let $\tilde{U} = \tilde{X}B^{\frac{1}{2}}$ and $\tilde{U}_1, \ldots, \tilde{U}_q$ be the columns of $\tilde{U}$. Then $\tilde{U}_1, \ldots, \tilde{U}_q$ are mutually independently distributed as p-variate complex multinormal with covariance matrix A^{-1} each, the density of $\tilde{U}_j$, denoted by $\tilde{f}_{\tilde{U}_j}(\tilde{U}_j)$, being given as*

$$\tilde{f}_{\tilde{U}_j}(\tilde{U}_j) = \frac{|\mathrm{det}(A)|}{\pi^p} \mathrm{e}^{-(\tilde{U}_j^* A \tilde{U}_j)}. \tag{4.2a.13}$$

Exercises 4.2

4.2.1. Prove the second result in equation (4.2.12) and prove both results in (4.2*a*.10).

4.2.2. Obtain (4.2.12) by establishing first the distribution of the row sum of squares and column sum of squares in Y, and then taking the expected values in those variables.

4.2.3. Prove (4.2*a*.10) by establishing first the distributions of row and column sum of squares of the absolute values in $\tilde{Y}$ and then taking the expected values.

4.2.4. Establish 4.2.12 and 4.2a.10 by using the general polar coordinate transformations.

4.2.5. First prove that $\sum_{j=1}^{q} |\tilde{y}_{ij}|^2$ is a $2q$-variate real gamma random variable. Then establish the results in (4.2*a*.10) by using the those on real gamma variables, where $\tilde{Y} = (\tilde{y}_{ij})$, the $\tilde{y}_{ij}$'s in (4.2*a*.11) being in the complex domain and $|\tilde{y}_{ij}|$ denoting the absolute value or modulus of $\tilde{y}_{ij}$.

4.2.6. Let the real matrix $A > O$ be 2×2 with its first row being $(1,\ 1)$ and let $B > O$ be 3×3 with its first row being $(1, 1, -1)$. Then complete the other rows in A and B so that $A > O,\ B > O$. Obtain the corresponding 2×3 real matrix-variate Gaussian density when (1): $M = O$, (2): $M \neq O$ with a matrix M of your own choice.

4.2.7. Let the complex matrix $A > O$ be 2×2 with its first row being $(1, 1 + i)$ and let $B > O$ be 3×3 with its first row being $(1, 1 + i, -i)$. Complete the other rows with numbers in the complex domain of your own choice so that $A = A^* > O,\ B = B^* > O$. Obtain the corresponding 2×3 complex matrix-variate Gaussian density with (1): $\tilde{M} = O$, (2): $\tilde{M} \neq O$ with a matrix $\tilde{M}$ of your own choice.

4.2.8. Evaluate the covariance matrix in (4.2.16), which is $E(Z_j' Z_j)$, and show that it is B^{-1}.

4.2.9. Evaluate the covariance matrix in (4.2.18), which is $E(U_j U_j')$, and show that it is A^{-1}.

4.2.10. Repeat Exercises 4.2.8 and 4.2.9 for the complex case in (4.2a.12) and (4.2a.13).

4.3. Moment Generating Function and Characteristic Function, Real Case

Let $T = (t_{ij})$ be a $p \times q$ parameter matrix. The matrix random variable $X = (x_{ij})$ is $p \times q$ and it is assumed that all of its elements x_{ij}'s are real and distinct scalar variables. Then

$$\mathrm{tr}(TX') = \sum_{i=1}^{p}\sum_{j=1}^{q} t_{ij}x_{ij} = \mathrm{tr}(X'T) = \mathrm{tr}(XT'). \qquad (i)$$

Note that each t_{ij} and x_{ij} appear once in *(i)* and thus, we can define the moment generating function (mgf) in the real matrix-variate case, denoted by $M_f(T)$ or $M_X(T)$, as follows:

$$M_f(T) = E[\mathrm{e}^{\mathrm{tr}(TX')}] = \int_X \mathrm{e}^{\mathrm{tr}(TX')} f_{p,q}(X)\mathrm{d}X = M_X(T) \qquad (ii)$$

whenever the integral is convergent, where E denotes the expected value. Thus, for the $p \times q$ matrix-variate real Gaussian density,

$$M_X(T) = M_f(T) = \frac{|A|^{\frac{q}{2}}|B|^{\frac{p}{2}}}{(2\pi)^{\frac{pq}{2}}} \int_X \mathrm{e}^{\mathrm{tr}(TX') - \frac{1}{2}\mathrm{tr}(A^{\frac{1}{2}}XBX'A^{\frac{1}{2}})}\mathrm{d}X$$

where A is $p \times p$, B is $q \times q$ and A and B are constant real positive definite matrices so that $A^{\frac{1}{2}}$ and $B^{\frac{1}{2}}$ are uniquely defined. Consider the transformation $Y = A^{\frac{1}{2}}XB^{\frac{1}{2}} \Rightarrow \mathrm{d}Y = |A|^{\frac{q}{2}}|B|^{\frac{p}{2}}\mathrm{d}X$ by Theorem 1.6.4. Thus, $X = A^{-\frac{1}{2}}YB^{-\frac{1}{2}}$ and

$$\mathrm{tr}(TX') = \mathrm{tr}(TB^{-\frac{1}{2}}Y'A^{-\frac{1}{2}}) = \mathrm{tr}(A^{-\frac{1}{2}}TB^{-\frac{1}{2}}Y') = \mathrm{tr}(T_{(1)}Y')$$

where $T_{(1)} = A^{-\frac{1}{2}} T B^{-\frac{1}{2}}$. Then

$$M_X(T) = \frac{1}{(2\pi)^{\frac{pq}{2}}} \int_Y \mathrm{e}^{\mathrm{tr}(T_{(1)}Y') - \frac{1}{2}\mathrm{tr}(YY')} \mathrm{d}Y.$$

Note that $T_{(1)}Y'$ and YY' are $p \times p$. Consider $-2\mathrm{tr}(T_{(1)}Y') + \mathrm{tr}(YY')$, which can be written as

$$-2\mathrm{tr}(T_{(1)}Y') + \mathrm{tr}(YY') = -\mathrm{tr}(T_{(1)}T'_{(1)}) + \mathrm{tr}[(Y - T_{(1)})(Y - T_{(1)})'].$$

Therefore

$$\begin{aligned} M_X(T) &= \mathrm{e}^{\frac{1}{2}\mathrm{tr}(T_{(1)}T'_{(1)})} \frac{1}{(2\pi)^{\frac{pq}{2}}} \int_Y \mathrm{e}^{-\frac{1}{2}\mathrm{tr}[(Y - T_{(1)})(Y' - T'_{(1)})]} \mathrm{d}Y \\ &= \mathrm{e}^{\frac{1}{2}\mathrm{tr}(T_{(1)}T'_{(1)})} = \mathrm{e}^{\frac{1}{2}\mathrm{tr}(A^{-\frac{1}{2}} T B^{-1} T' A^{-\frac{1}{2}})} = \mathrm{e}^{\frac{1}{2}\mathrm{tr}(A^{-1} T B^{-1} T')} \end{aligned} \tag{4.3.1}$$

since the integral is 1 from the total integral of a matrix-variate Gaussian density.

In the presence of a location parameter matrix M, the matrix-variate Gaussian density is given by

$$f_{p,q}(X) = \frac{|A|^{\frac{q}{2}} |B|^{\frac{p}{2}}}{(2\pi)^{\frac{pq}{2}}} \mathrm{e}^{-\frac{1}{2}\mathrm{tr}(A^{\frac{1}{2}}(X-M)B(X-M)'A^{\frac{1}{2}})} \tag{4.3.2}$$

where M is a constant $p \times q$ matrix. In this case, $TX' = T(X - M + M)' = T(X - M)' + TM'$, and

$$\begin{aligned} M_X(T) = M_f(T) &= E[\mathrm{e}^{\mathrm{tr}(TX')}] = \mathrm{e}^{\mathrm{tr}(TM')} E[\mathrm{e}^{\mathrm{tr}(T(X-M)')}] \\ &= \mathrm{e}^{\mathrm{tr}(TM')} \mathrm{e}^{\frac{1}{2}\mathrm{tr}(A^{-1}TB^{-1}T')} = \mathrm{e}^{\mathrm{tr}(TM') + \mathrm{tr}(\frac{1}{2}A^{-1}TB^{-1}T')}. \end{aligned} \tag{4.3.3}$$

When $p = 1$, we have the usual q-variate multinormal density. In this case, A is 1×1 and taken to be 1. Then the mgf is given by

$$M_X(T) = \mathrm{e}^{TM' + \frac{1}{2}TB^{-1}T'} \tag{4.3.4}$$

where $T,\ M$ and X are $1\times q$ and $B>O$ is $q\times q$. The corresponding characteristic function when $p=1$ is given by

$$\phi(T)=\mathrm{e}^{iTM'-\frac{1}{2}TB^{-1}T'}. \tag{4.3.5}$$

Example 4.3.1. Let X have a 2×3 real matrix-variate Gaussian density with the following parameters:

$$X=\begin{bmatrix}x_{11}&x_{12}&x_{13}\\x_{21}&x_{22}&x_{23}\end{bmatrix},\ E[X]=M=\begin{bmatrix}1&0&-1\\-1&1&0\end{bmatrix},\ A=\begin{bmatrix}1&1\\1&2\end{bmatrix},$$
$$B=\begin{bmatrix}3&-1&1\\-1&2&1\\1&1&3\end{bmatrix}.$$

Consider the density $f_{2,3}(X)$ with the exponent preceded by $\frac{1}{2}$ to be consistent with p-variate real Gaussian density. Verify whether A and B are positive definite. Then compute the moment generating function (mgf) of X or that associated with $f_{2,3}(X)$ and write down the exponent explicitly.

Solution 4.3.1. Consider a 2×3 parameter matrix $T=(t_{ij})$. Let us compute the various quantities in the mgf. First,

$$TM'=\begin{bmatrix}t_{11}&t_{12}&t_{13}\\t_{21}&t_{22}&t_{23}\end{bmatrix}\begin{bmatrix}1&-1\\0&1\\-1&0\end{bmatrix}=\begin{bmatrix}t_{11}-t_{13}&-t_{11}+t_{12}\\t_{21}-t_{23}&-t_{21}+t_{22}\end{bmatrix},$$

so that

$$\mathrm{tr}(TM')=t_{11}-t_{13}-t_{21}+t_{22}. \tag{i}$$

Consider the leading minors in A and B. Note that $|(1)|=1>0,\ |A|=1>0,\ |(3)|=3>0,\ \begin{vmatrix}3&-1\\-1&2\end{vmatrix}=5>0,\ |B|=8>0$; thus both A and B are positive definite. The inverses of A and B are obtained by making use of the formula $C^{-1}=\frac{1}{|C|}(\mathrm{Cof}(C))'$; they are

$$A^{-1}=\begin{bmatrix}2&-1\\-1&1\end{bmatrix},\ B^{-1}=\frac{1}{8}\begin{bmatrix}5&4&-3\\4&8&-4\\-3&-4&5\end{bmatrix}.$$

For determining the exponent in the mgf, we need $A^{-1}T$ and $B^{-1}T'$, which are

$$A^{-1}T=\begin{bmatrix}2&-1\\-1&1\end{bmatrix}\begin{bmatrix}t_{11}&t_{12}&t_{13}\\t_{21}&t_{22}&t_{23}\end{bmatrix}$$
$$=\begin{bmatrix}2t_{11}-t_{21}&2t_{12}-t_{22}&2t_{13}-t_{23}\\-t_{11}+t_{21}&-t_{12}+t_{22}&-t_{13}+t_{23}\end{bmatrix}$$

$$B^{-1}T'=\frac{1}{8}\begin{bmatrix}5&4&-3\\4&8&-4\\-3&-4&5\end{bmatrix}\begin{bmatrix}t_{11}&t_{21}\\t_{12}&t_{22}\\t_{13}&t_{23}\end{bmatrix}$$
$$=\frac{1}{8}\begin{bmatrix}5t_{11}+4t_{12}-3t_{13}&5t_{21}+4t_{22}-3t_{23}\\4t_{11}+8t_{12}-4t_{13}&4t_{21}+8t_{22}-4t_{23}\\-3t_{11}-4t_{12}+5t_{13}&-3t_{21}-4t_{22}+5t_{23}\end{bmatrix}.$$

Hence,

$$\begin{aligned}\frac{1}{2}\text{tr}[A^{-1}TB^{-1}T']=\frac{1}{16}[&(2t_{11}-t_{21})(5t_{11}+4t_{12}-3t_{13})\\&+(2t_{12}-t_{22})(4t_{11}+8t_{12}-4t_{13})+(2t_{13}-t_{23})(-3t_{11}-4t_{12}+5t_{13})\\&+(-t_{11}+t_{21})(5t_{21}+4t_{22}-3t_{23})+(-t_{12}+t_{22})(4t_{21}+8t_{22}-4t_{23})\\&+(-t_{13}+t_{23})(-3t_{21}-4t_{22}+5t_{23})].\qquad (ii)\end{aligned}$$

Thus, the mgf is $M_X(T)=\text{e}^{Q(T)}$ where

$$Q(T)=\text{tr}(TM')+\frac{1}{2}\text{tr}(A^{-1}TB^{-1}T'),$$

these quantities being given in *(i)* and *(ii)*. This completes the computations.

4.3a. Moment Generating and Characteristic Functions, Complex Case

Let $\tilde{X}=(\tilde{x}_{ij})$ be a $p\times q$ matrix where the $\tilde{x}_{ij}$'s are distinct scalar complex variables. We may write $\tilde{X}=X_1+iX_2,\ i=\sqrt{(-1)},\ X_1,\ X_2$ being real $p\times q$ matrices. Let $\tilde{T}$ be a $p\times q$ parameter matrix and $\tilde{T}=T_1+iT_2,\ T_1,\ T_2$ being real $p\times q$ matrices. The conjugate transposes of $\tilde{X}$ and $\tilde{T}$ are denoted by $\tilde{X}^*$ and $\tilde{T}^*$, respectively. Then,

$$\begin{aligned}\text{tr}(\tilde{T}\tilde{X}^*)&=\text{tr}[(T_1+iT_2)(X_1'-iX_2')]\\&=\text{tr}[T_1X_1'+T_2X_2'+i(T_2X_1'-T_1X_2')]\\&=\text{tr}(T_1X_1')+\text{tr}(T_2X_2')+i\,\text{tr}(T_2X_1'-T_1X_2').\end{aligned}$$

If $T_1 = (t_{ij}^{(1)}),\ X_1 = (x_{ij}^{(1)}),\ X_2 = (x_{ij}^{(2)}),\ T_2 = (t_{ij}^{(2)})$, $\text{tr}(T_1 X_1') = \sum_{i=1}^{p}\sum_{j=1}^{q} t_{ij}^{(1)} x_{ij}^{(1)}$, $\text{tr}(T_2 X_2') = \sum_{i=1}^{p}\sum_{j=1}^{q} t_{ij}^{(2)} x_{ij}^{(2)}$. In other words, $\text{tr}(T_1X_1') + \text{tr}(T_2X_2')$ gives all the x_{ij}'s in the real and complex parts of $\tilde{X}$ multiplied by the corresponding t_{ij}'s in the real and complex parts of $\tilde{T}$. That is, $E[\mathrm{e}^{\text{tr}(T_1X_1')+\text{tr}(T_2X_2')}]$ gives a moment generating function (mgf) associated with the complex matrix-variate Gaussian density that is consistent with real multivariate mgf. However, $[\text{tr}(T_1X_1') + \text{tr}(T_2X_2')] = \Re(\text{tr}[\tilde{T}\tilde{X}^*])$, $\Re(\cdot)$ denoting the real part of $(\cdot)$. Thus, in the complex case, the mgf for any real-valued scalar function $g(\tilde{X})$ of the complex matrix argument $\tilde{X}$, where $g(\tilde{X})$ is a density, is defined as

$$\tilde{M}_{\tilde{X}}(\tilde{T}) = \int_{\tilde{X}} \mathrm{e}^{\Re[\text{tr}(\tilde{T}\tilde{X}^*)]} g(\tilde{X})\mathrm{d}\tilde{X} \tag{4.3a.1}$$

whenever the expected value exists. On replacing $\tilde{T}$ by $i\tilde{T},\ i = \sqrt{(-1)}$, we obtain the characteristic function of $\tilde{X}$ or that associated with $\tilde{f}$, denoted by $\phi_{\tilde{X}}(\tilde{T}) = \phi_{\tilde{f}}(\tilde{T})$. That is,

$$\phi_{\tilde{X}}(\tilde{T}) = \int_{\tilde{X}} \mathrm{e}^{\Re[\text{tr}(i\tilde{T}\tilde{X}^*)]} g(\tilde{X})\mathrm{d}\tilde{X}. \tag{4.3a.2}$$

Then, the mgf of the matrix-variate Gaussian density in the complex domain is available by paralleling the derivation in the real case and making use of Lemma 3.2*a*.1:

$$\begin{aligned}\tilde{M}_{\tilde{X}}(\tilde{T}) &= E[\mathrm{e}^{\Re[\text{tr}(\tilde{T}\tilde{X}^*)]}] \\ &= \mathrm{e}^{\Re[\text{tr}(\tilde{T}\tilde{M}^*)]+\frac{1}{4}\Re[\text{tr}(A^{-\frac{1}{2}}\tilde{T}B^{-1}\tilde{T}^*A^{-\frac{1}{2}})]}.\end{aligned} \tag{4.3a.3}$$

The corresponding characteristic function is given by

$$\phi_{\tilde{X}}(\tilde{T}) = \mathrm{e}^{\Re[\text{tr}(i\tilde{T}\tilde{M}^*)]-\frac{1}{4}\Re[\text{tr}(A^{-\frac{1}{2}}\tilde{T}B^{-1}\tilde{T}^*A^{-\frac{1}{2}})]}. \tag{4.3a.4}$$

Note that when $A = A^* > O$ and $B = B^* > O$ (Hermitian positive definite),

$$(A^{-\frac{1}{2}}\tilde{T}B^{-1}\tilde{T}^*A^{-\frac{1}{2}})^* = A^{-\frac{1}{2}}\tilde{T}B^{-1}\tilde{T}^*A^{-\frac{1}{2}},$$

that is, this matrix is Hermitian. Thus, letting $\tilde{U} = A^{-\frac{1}{2}}\tilde{T}B^{-1}\tilde{T}^*A^{-\frac{1}{2}} = U_1 + iU_2$ where U_1 and U_2 are real matrices, $U_1 = U_1'$ and $U_2 = -U_2'$, that is, U_1 and U_2 are respectively symmetric and skew symmetric real matrices. Accordingly, $\text{tr}(\tilde{U}) = \text{tr}(U_1) + i\,\text{tr}(U_2) = \text{tr}(U_1)$ as the trace of a real skew symmetric matrix is zero. Therefore, $\Re[\text{tr}(A^{-\frac{1}{2}}\tilde{T}B^{-1}\tilde{T}^*A^{-\frac{1}{2}})] = \text{tr}(A^{-\frac{1}{2}}\tilde{T}B^{-1}\tilde{T}^*A^{-\frac{1}{2}})$, the diagonal elements of a Hermitian matrix being real.

When $p = 1$, we have the usual q-variate complex multivariate normal density and taking the 1×1 matrix A to be 1, the mgf is as follows:

$$\tilde{M}_{\tilde{X}}(\tilde{T}) = \mathrm{e}^{\Re(\tilde{T}\tilde{M}^*)+\frac{1}{4}(\tilde{T}B^{-1}\tilde{T}^*)} \tag{4.3a.5}$$

where $\tilde{T},\ \tilde{M}$ are $1 \times q$ vectors and $B = B^* > O$ (Hermitian positive definite), the corresponding characteristic function being given by

$$\phi_{\tilde{X}}(\tilde{T}) = \mathrm{e}^{\Re(i\tilde{T}\tilde{M}^*)-\frac{1}{4}(\tilde{T}B^{-1}\tilde{T}^*)}. \tag{4.3a.6}$$

Example 4.3a.1. Consider a 2×2 matrix $\tilde{X}$ in the complex domain having a complex matrix-variate density with the following parameters:

$$\tilde{X} = \begin{bmatrix} \tilde{x}_{11} & \tilde{x}_{12} \\ \tilde{x}_{21} & \tilde{x}_{22} \end{bmatrix},\ E[\tilde{X}] = \tilde{M} = \begin{bmatrix} 1+i & i \\ 2-i & 1 \end{bmatrix},$$
$$A = \begin{bmatrix} 2 & i \\ -i & 3 \end{bmatrix},\ B = \begin{bmatrix} 2 & -i \\ i & 1 \end{bmatrix}.$$

Determine whether A and B are Hermitian positive definite; then, obtain the mgf of this distribution and provide the exponential part explicitly.

Solution 4.3a.1. Clearly, $A > O$ and $B > O$. We first determine $A^{-1},\ B^{-1},\ A^{-1}\tilde{T}$, $B^{-1}\tilde{T}^*$:

$$A^{-1} = \frac{1}{5}\begin{bmatrix} 3 & -i \\ i & 2 \end{bmatrix},\ A^{-1}\tilde{T} = \frac{1}{5}\begin{bmatrix} 3 & -i \\ i & 2 \end{bmatrix}\begin{bmatrix} \tilde{t}_{11} & \tilde{t}_{12} \\ \tilde{t}_{21} & \tilde{t}_{22} \end{bmatrix} = \frac{1}{5}\begin{bmatrix} 3\tilde{t}_{11} - i\tilde{t}_{21} & 3\tilde{t}_{12} - i\tilde{t}_{22} \\ i\tilde{t}_{11} + 2\tilde{t}_{21} & i\tilde{t}_{12} + 2\tilde{t}_{22} \end{bmatrix},$$
$$B^{-1} = \begin{bmatrix} 1 & i \\ -i & 2 \end{bmatrix},\ B^{-1}\tilde{T}^* = \begin{bmatrix} \tilde{t}_{11}^* + i\tilde{t}_{12}^* & \tilde{t}_{21}^* + i\tilde{t}_{22}^* \\ -i\tilde{t}_{11}^* + 2\tilde{t}_{12}^* & -i\tilde{t}_{21}^* + 2\tilde{t}_{22}^* \end{bmatrix}.$$

Letting $\delta = \frac{1}{2}\mathrm{tr}(A^{-1}\tilde{T}B^{-1}\tilde{T}^*)$,

$$\begin{aligned} 10\delta = \{&(3\tilde{t}_{11} - i\tilde{t}_{21})(\tilde{t}_{11}^* + i\tilde{t}_{12}^*) + (3\tilde{t}_{12} - i\tilde{t}_{22})(-i\tilde{t}_{11}^* + 2\tilde{t}_{12}^*) \\ &+ (i\tilde{t}_{11} + 2\tilde{t}_{21})(\tilde{t}_{21}^* + i\tilde{t}_{22}^*) + (i\tilde{t}_{12} + 2\tilde{t}_{22})(-i\tilde{t}_{21}^* + 2\tilde{t}_{22}^*)\}, \end{aligned}$$

$$\begin{aligned} 10\delta = \{&3\tilde{t}_{11}\tilde{t}_{11}^* + 3i\tilde{t}_{11}\tilde{t}_{12}^* - i\tilde{t}_{21}\tilde{t}_{11}^* + \tilde{t}_{21}\tilde{t}_{12}^* \\ &+ 6\tilde{t}_{12}\tilde{t}_{12}^* - \tilde{t}_{22}\tilde{t}_{11}^* - 2i\tilde{t}_{22}\tilde{t}_{12}^* - 3i\tilde{t}_{12}\tilde{t}_{11}^* \\ &+ i\tilde{t}_{11}\tilde{t}_{21}^* - \tilde{t}_{11}\tilde{t}_{22}^* + 2\tilde{t}_{21}\tilde{t}_{21}^* + 2i\tilde{t}_{21}\tilde{t}_{22}^* \\ &+ \tilde{t}_{12}\tilde{t}_{21}^* + 2i\tilde{t}_{12}\tilde{t}_{22}^* - 2i\tilde{t}_{22}\tilde{t}_{21}^* + 4\tilde{t}_{22}\tilde{t}_{22}^*\}, \end{aligned}$$

$$
\begin{aligned}
10\delta = {} & 3\tilde{t}_{11}\tilde{t}_{11}^{*} + 6\tilde{t}_{12}\tilde{t}_{12}^{*} + 2\tilde{t}_{21}\tilde{t}_{21}^{*} + 4\tilde{t}_{22}\tilde{t}_{22}^{*} \\
& + 3i[\tilde{t}_{11}\tilde{t}_{12}^{*} - \tilde{t}_{12}\tilde{t}_{11}^{*}] - [\tilde{t}_{22}\tilde{t}_{11}^{*} + \tilde{t}_{11}\tilde{t}_{22}^{*}] \\
& + i[\tilde{t}_{11}\tilde{t}_{21}^{*} - \tilde{t}_{11}^{*}\tilde{t}_{21}] + [\tilde{t}_{12}\tilde{t}_{21}^{*} + \tilde{t}_{12}^{*}\tilde{t}_{21}] \\
& + 2i[\tilde{t}_{21}\tilde{t}_{22}^{*} - \tilde{t}_{21}^{*}\tilde{t}_{22}] + 2i[\tilde{t}_{12}\tilde{t}_{22}^{*} - \tilde{t}_{12}^{*}\tilde{t}_{22}].
\end{aligned}
$$

Letting $\tilde{t}_{rs} = t_{rs1} + it_{rs2},\ i = \sqrt{(-1)},\ t_{rs1},\ t_{rs2}$ being real, for all r and s, then δ, the exponent in the mgf, can be expressed as follows:

$$
\begin{aligned}
\delta = \frac{1}{10}\{ & 3(t_{111}^2 + t_{112}^2) + 6(t_{121}^2 + t_{122}^2) + 2(t_{211}^2 + t_{212}^2) + 4(t_{221}^2 + t_{222}^2) \\
& - 6(t_{112}t_{121} - t_{111}t_{122}) - 2(t_{111}t_{221} - t_{112}t_{222}) - 2(t_{112}t_{211} - t_{111}t_{212}) \\
& + 2(t_{121}t_{211} + t_{122}t_{212}) - 4(t_{212}t_{221} - t_{211}t_{222}) - 4(t_{122}t_{221} - t_{121}t_{222})\}.
\end{aligned}
$$

This completes the computations.

4.3.1. Distribution of the exponent, real case

Let us determine the distribution of the exponent in the $p \times q$ real matrix-variate Gaussian density. Letting $u = \mathrm{tr}(AXBX')$, its density can be obtained by evaluating its associated mgf. Then, taking t as its scalar parameter since u is scalar, we have

$$M_u(t) = E[\mathrm{e}^{tu}] = E[\mathrm{e}^{t\,\mathrm{tr}(AXBX')}].$$

Since this expected value depends on X, we can integrate out over the density of X:

$$
\begin{aligned}
M_u(t) &= C\int_X \mathrm{e}^{t\,\mathrm{tr}(AXBX') - \frac{1}{2}\mathrm{tr}(AXBX')}\mathrm{d}X \\
&= C\int_X \mathrm{e}^{-\frac{1}{2}(1-2t)(\mathrm{tr}(AXBX'))}\mathrm{d}X \quad \text{for } 1-2t > 0 \qquad (i)
\end{aligned}
$$

where

$$C = \frac{|A|^{\frac{q}{2}}|B|^{\frac{p}{2}}}{(2\pi)^{\frac{pq}{2}}}.$$

The integral in *(i)* is convergent only when $1-2t > 0$. Then distributing $\sqrt{(1-2t)}$ to each element in X and X', and denoting the new matrix by X_t, we have $X_t = \sqrt{(1-2t)}X \Rightarrow \mathrm{d}X_t = (\sqrt{(1-2t)})^{pq}\mathrm{d}X = (1-2t)^{\frac{pq}{2}}\mathrm{d}X$. Integral over X_t, together with C, yields 1 and hence

$$M_u(t) = (1-2t)^{-\frac{pq}{2}},\ \text{provided } 1-2t > 0. \qquad (4.3.6)$$

The corresponding density is a real chi-square having pq degrees of freedom or a real gamma density with parameters $\alpha = \frac{pq}{2}$ and $\beta = 2$. Thus, the resulting density, denoted by $f_{u_1}(u_1)$, is given by

$$f_{u_1}(u_1) = [2^{\frac{pq}{2}}\Gamma(pq/2)]^{-1}u_1^{\frac{pq}{2}-1}\mathrm{e}^{-\frac{u_1}{2}},\ 0 \le u_1 < \infty,\ p,q = 1,2,\ldots, \tag{4.3.7}$$

and $f_{u_1}(u_1) = 0$ elsewhere.

4.3a.1. Distribution of the exponent, complex case

In the complex case, letting $\tilde{u} = \mathrm{tr}(A^{\frac{1}{2}}\tilde{X}B\tilde{X}^*A^{\frac{1}{2}})$, we note that $\tilde{u} = \tilde{u}^*$ and $\tilde{u}$ is a scalar, so that $\tilde{u}$ is real. Hence, the mgf of $\tilde{u}$, with real parameter t, is given by

$$M_{\tilde{u}}(t) = E[\mathrm{e}^{t\,\mathrm{tr}(A^{\frac{1}{2}}\tilde{X}B\tilde{X}^*A^{\frac{1}{2}})}] = C_1\int_{\tilde{X}}\mathrm{e}^{-(1-t)\mathrm{tr}(A^{\frac{1}{2}}\tilde{X}B\tilde{X}^*A^{\frac{1}{2}})}\mathrm{d}\tilde{X},\ \ 1-t>0,\ \ \text{with}$$

$$C_1 = \frac{|\mathrm{det}(A)|^q|\mathrm{det}(B)|^p}{\pi^{pq}}.$$

On making the transformation $\tilde{Y} = A^{\frac{1}{2}}\tilde{X}B^{\frac{1}{2}}$, we have

$$M_{\tilde{u}}(t) = \frac{1}{\pi^{pq}}\int_{\tilde{Y}}\mathrm{e}^{-(1-t)\mathrm{tr}(\tilde{Y}\tilde{Y}^*)}\mathrm{d}\tilde{Y}.$$

However,

$$\mathrm{tr}(\tilde{Y}\tilde{Y}^*) = \sum_{r=1}^{p}\sum_{s=1}^{q}|\tilde{y}_{rs}|^2 = \sum_{r=1}^{p}\sum_{s=1}^{q}(y_{rs1}^2 + y_{rs2}^2)$$

where $\tilde{y}_{rs} = y_{rs1} + iy_{rs2},\ i = \sqrt{(-1)},\ y_{rs1},\ y_{rs2}$ being real. Hence

$$\frac{1}{\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\mathrm{e}^{-(1-t)(y_{rs1}^2+y_{rs2}^2)}\mathrm{d}y_{rs1}\wedge\mathrm{d}y_{rs2} = \frac{1}{1-t},\ \ 1-t>0.$$

Therefore,

$$M_{\tilde{u}}(t) = (1-t)^{-pq},\ \ 1-t>0, \tag{4.3a.7}$$

and $\tilde{u} = v$ has a real gamma density with parameters $\alpha = pq,\ \beta = 1$, or a chi-square density in the complex domain with pq degrees of freedom, that is,

$$f_v(v) = \frac{1}{\Gamma(pq)}v^{pq-1}\mathrm{e}^{-v},\ \ 0 \le v < \infty, \tag{4.3a.8}$$

and $f_v(v) = 0$ elsewhere.

4.3.2. Linear functions in the real case

Let the $p\times q$ real matrix $X=(x_{ij})$ of the real scalar random variables x_{ij}'s have the density in (4.2.2), namely

$$f_{p,q}(X)=\frac{|A|^{\frac{q}{2}}|B|^{\frac{p}{2}}}{(2\pi)^{\frac{pq}{2}}}\mathrm{e}^{-\frac{1}{2}\mathrm{tr}(A(X-M)B(X-M)')} \tag{4.3.8}$$

for $A>O, B>O$, where M is a $p\times q$ location parameter matrix. Let L_1 be a $p\times 1$ vector of constants. Consider the linear function $Z_1=L_1'X$ where Z_1 is $1\times q$. Let T be a $1\times q$ parameter vector. Then the mgf of the $1\times q$ vector Z_1 is

$$\begin{aligned}M_{Z_1}(T)&=E[\mathrm{e}^{(TZ_1')}]=E[\mathrm{e}^{(TX'L_1)}]=E[\mathrm{e}^{\mathrm{tr}(TX'L_1)}]\\&=E[\mathrm{e}^{\mathrm{tr}((L_1T)X')}].\end{aligned} \tag{i}$$

This can be evaluated by replacing T by L_1T in (4.3.4). Then

$$\begin{aligned}M_{Z_1}(T)&=\mathrm{e}^{\mathrm{tr}((L_1T)M')+\frac{1}{2}\mathrm{tr}(A^{-1}L_1TB^{-1}(L_1T)')}\\&=\mathrm{e}^{\mathrm{tr}(TM'L_1)+\frac{1}{2}\mathrm{tr}[(L_1'A^{-1}L_1)TB^{-1}T']}.\end{aligned} \tag{ii}$$

Since $L_1'A^{-1}L_1$ is a scalar,

$$(L_1'A^{-1}L_1)TB^{-1}T'=T(L_1'A^{-1}L_1)B^{-1}T'.$$

On comparing the resulting expression with the mgf of a q-variate real normal distribution, we observe that Z_1 is a q-variate real Gaussian vector with mean value vector $L_1'M$ and covariance matrix $[L_1'A^{-1}L_1]B^{-1}$. Hence the following result:

Theorem 4.3.1. *Let the real $p\times q$ matrix X have the density specified in (4.3.8) and L_1 be a $p\times 1$ constant vector. Let Z_1 be the linear function of X, $Z_1=L_1'X$. Then Z_1, which is $1\times q$, has the mgf given in (ii) and thereby Z_1 has a q-variate real Gaussian density with the mean value vector $L_1'M$ and covariance matrix $[L_1'A^{-1}L_1]B^{-1}$.*

Theorem 4.3.2. *Let L_2 be a $q\times 1$ constant vector. Consider the linear function $Z_2=XL_2$ where the $p\times q$ real matrix X has the density specified in (4.3.8). Then Z_2, which is $p\times 1$, is a p-variate real Gaussian vector with mean value vector ML_2 and covariance matrix $[L_2'B^{-1}L_2]A^{-1}$.*

The proof of Theorem 4.3.2 is parallel to the derivation of that of Theorem 4.3.1. Theorems 4.3.1 and 4.3.2 establish that when the $p\times q$ matrix X has a $p\times q$-variate real Gaussian density with parameters $M,\ A>O,\ B>O$, then all linear functions of the form $L_1'X$ where L_1 is $p \times 1$ are q-variate real Gaussian and all linear functions of the type XL_2 where L_2 is $q \times 1$ are p-variate real Gaussian, the parameters in these Gaussian densities being given in Theorems 4.3.1 and 4.3.2.

By retracing the steps, we can obtain characterizations of the density of the $p \times q$ real matrix X through linear transformations. Consider all possible $p \times 1$ constant vectors L_1 or, equivalently, let L_1 be arbitrary. Let T be a $1 \times q$ parameter vector. Then the $p \times q$ matrix L_1T, denoted by $T_{(1)}$, contains pq free parameters. In this case the mgf in *(ii)* can be written as

$$M(T_{(1)}) = \mathrm{e}^{\mathrm{tr}(T_{(1)}M')+\frac{1}{2}\mathrm{tr}(A^{-1}T_{(1)}B^{-1}T_{(1)}')}, \qquad (iii)$$

which has the same structure of the mgf of a $p \times q$ real matrix-variate Gaussian density as given in (4.3.8), whose the mean value matrix is M and parameter matrices are $A > O$ and $B > O$. Hence, the following result can be obtained:

Theorem 4.3.3. *Let L_1 be a constant $p \times 1$ vector, X be a $p \times q$ matrix whose elements are real scalar variables and $A > O$ be $p \times p$ and $B > O$ be $q \times q$ constant real positive definite matrices. If for an arbitrary vector L_1, $L_1'X$ is a q-variate real Gaussian vector as specified in Theorem 4.3.1, then X has a $p \times q$ real matrix-variate Gaussian density as given in (4.3.8).*

As well, a result parallel to this one follows from Theorem 4.3.2:

Theorem 4.3.4. *Let L_2 be a $q \times 1$ constant vector, X be a $p \times q$ matrix whose elements are real scalar variables and $A > O$ be $p \times p$ and $B > O$ be $q \times q$ constant real positive definite matrices. If for an arbitrary constant vector L_2, XL_2 is a p-variate real Gaussian vector as specified in Theorem 4.3.2, then X is $p \times q$ real matrix-variate Gaussian distributed as in (4.3.8).*

Example 4.3.2. Consider a 2×2 matrix-variate real Gaussian density with the parameters

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix},\ B = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix},\ M = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} = E[X],\ X = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix}.$$

Letting $U_1 = L_1'X,\ U_2 = XL_2,\ U_3 = L_1'XL_2$, evaluate the densities of $U_1,\ U_2,\ U_3$ by applying Theorems 4.3.1 and 4.3.2 where $L_1' = [1, 1],\ L_2' = [1, -1]$; as well, obtain those densities without resorting to these theorems.

Solution 4.3.2. Let us first compute the following quantities:

$$A^{-1},\ B^{-1},\ L_1'A^{-1}L_1,\ L_2'B^{-1}L_2,\ L_1'M,\ ML_2,\ L_1'ML_2.$$

They are

$$A^{-1}=\begin{bmatrix}1&-1\\-1&2\end{bmatrix},\ B^{-1}=\frac{1}{3}\begin{bmatrix}2&-1\\-1&2\end{bmatrix},$$
$$L_1'M=[1,1]\begin{bmatrix}1&-1\\0&1\end{bmatrix}=[1,0],\ ML_2=\begin{bmatrix}2\\-1\end{bmatrix},$$
$$L_1'A^{-1}L_1=[1,1]\begin{bmatrix}1&-1\\-1&2\end{bmatrix}\begin{bmatrix}1\\1\end{bmatrix}=1,\ L_2'B^{-1}L_2=\frac{1}{3}[1,-1]\begin{bmatrix}2&-1\\-1&2\end{bmatrix}\begin{bmatrix}1\\-1\end{bmatrix}=2,$$
$$L_1'ML_2=[1,0]\begin{bmatrix}2\\-1\end{bmatrix}=2.$$

Let $U_1=L_1'X, U_2=XL_2, U_3=L_1'XL_2$. Then by making use of Theorems 4.3.1 and 4.3.2 and then, results from Chap. 2 on q-variate real Gaussian vectors, we have the following:

$$U_1\sim N_2((1,0),(1)B^{-1}),\ U_2\sim N_2(ML_2,2A^{-1}),\ U_3\sim N_1(1,(1)(2))=N_1(1,2).$$

Let us evaluate the densities without resorting to these theorems. Note that $U_1=[x_{11}+x_{21},x_{12}+x_{22}]$. Then U_1 has a bivariate real distribution. Let us compute the mgf of U_1. Letting t_1 and t_2 be real parameters, the mgf of U_1 is

$$M_{U_1}(t_1,t_2)=E[\mathrm{e}^{t_1(x_{11}+x_{21})+t_2(x_{12}+x_{22})}]=E[\mathrm{e}^{t_1x_{11}+t_1x_{21}+t_2x_{12}+t_2x_{22}}],$$

which is available from the mgf of X by letting $t_{11}=t_1,\ t_{21}=t_1,\ t_{12}=t_2,\ t_{22}=t_2$. Thus,

$$A^{-1}T=\begin{bmatrix}1&-1\\-1&2\end{bmatrix}\begin{bmatrix}t_1&t_2\\t_1&t_2\end{bmatrix}=\begin{bmatrix}0&0\\t_1&t_2\end{bmatrix}$$
$$B^{-1}T'=\frac{1}{3}\begin{bmatrix}2&-1\\-1&2\end{bmatrix}\begin{bmatrix}t_1&t_1\\t_2&t_2\end{bmatrix}=\frac{1}{3}\begin{bmatrix}2t_1-t_2&2t_1-t_2\\-t_1+2t_2&-t_1+2t_2\end{bmatrix},$$

so that

$$\frac{1}{2}\mathrm{tr}(A^{-1}TB^{-1}T')=\frac{1}{2}\Big\{\frac{1}{3}[2t_1^2+2t_2^2-2t_1t_2]\Big\}=\frac{1}{2}[t_1,t_2]B^{-1}\begin{bmatrix}t_1\\t_2\end{bmatrix}. \qquad (i)$$

Since

$$U_2=XL_2=\begin{bmatrix}x_{11}-x_{12}\\x_{21}-x_{22}\end{bmatrix},$$

we let $t_{11} = t_1,\ t_{12} = -t_1,\ t_{21} = t_2, t_{22} = -t_2$. With these substitutions, we have the following:

$$A^{-1}T = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} t_1 & -t_1 \\ t_2 & -t_2 \end{bmatrix} = \begin{bmatrix} t_1 - t_2 & -t_1 + t_2 \\ -t_1 + 2t_2 & t_1 - 2t_2 \end{bmatrix}$$
$$B^{-1}T' = \frac{1}{3}\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} t_1 & t_2 \\ -t_1 & -t_2 \end{bmatrix} = \begin{bmatrix} t_1 & t_2 \\ -t_1 & -t_2 \end{bmatrix}.$$

Hence,

$$\begin{aligned}\mathrm{tr}(A^{-1}TB^{-1}T') &= t_1(t_1 - t_2) - t_1(-t_1 + t_2) + t_2(-t_1 + 2t_2) - t_2(t_1 - 2t_2)\\ &= 2[t_1, t_2]\begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} t_1 \\ t_2 \end{bmatrix}.\end{aligned}$$

Therefore, U_2 is a 2-variate real Gaussian with covariance matrix $2A^{-1}$ and mean value vector $\begin{bmatrix} 2 \\ -1 \end{bmatrix}$. That is, $U_2 \sim N_2(ML_2, 2A^{-1})$. For determining the distribution of U_3, observe that $L_1'XL_2 = L_1'U_2$. Then, $L_1'U_2$ is univariate real Gaussian with mean value $E[L_1'U_2] = L_1'ML_2 = [1, 1]\begin{bmatrix} 2 \\ -1 \end{bmatrix} = 1$ and variance $L_1'\mathrm{Cov}(U_2)L_1 = L_1'2A^{-1}L_1 = 2$. That is, $U_3 = u_3 \sim N_1(1, 2)$. This completes the solution.

The results stated in Theorems 4.3.1 and 4.3.2 are now generalized by taking sets of linear functions of X:

Theorem 4.3.5. *Let C' be a $r \times p$, $r \le p$, real constant matrix of full rank r and G be a $q \times s$ matrix, $s \le q$, of rank s. Let $Z = C'X$ and $W = XG$ where X has the density specified in (4.3.8). Then, Z has a $r \times q$ real matrix-variate Gaussian density with M replaced by $C'M$ and A^{-1} replaced by $C'A^{-1}C$, B^{-1} remaining unchanged, and $W = XG$ has a $p \times s$ real matrix-variate Gaussian distribution with B^{-1} replaced by $G'B^{-1}G$ and M replaced by MG, A^{-1} remaining unchanged.*

Example 4.3.3. Let the 2×2 real $X = (x_{ij})$ have a real matrix-variate Gaussian distribution with the parameters M, A and B. Consider the set of linear functions $U = C'X$ where

$$M = \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix},\ A = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix},\ B = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix},\ C' = \begin{bmatrix} \sqrt{2} & -\frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} \end{bmatrix}.$$

Show that the rows of U are independently distributed real q-variate Gaussian vectors with common covariance matrix B^{-1} and the rows of M as the mean value vectors.

Solution 4.3.3. Let us compute A^{-1} and $C'A^{-1}C$:

$$A^{-1} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$$
$$C'A^{-1}C = \begin{bmatrix} \sqrt{2} & -\frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} \sqrt{2} & 0 \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2.$$

In the mgf of $U = C'X$, A^{-1} is replaced by $C'A^{-1}C = I_2$ and B^{-1} remains the same. Then, the exponent in the mgf of U, excluding $\mathrm{tr}(TM')$ is $\frac{1}{2}\mathrm{tr}(TB^{-1}T') = \frac{1}{2}\sum_{j=1}^{p} T_j B^{-1} T_j'$ where T_j is the j-th row of T. Hence the p rows of U are independently distributed q-variate real Gaussian with the common covariance matrix B^{-1}. This completes the computations.

The previous example entails a general result that now is stated as a corollary.

Corollary 4.3.1. *Let X be a $p \times q$-variate real Gaussian matrix with the usual parameters M, A and B, whose density is as given in (4.3.8). Consider the set of linear functions $U = C'X$ where C is a $p \times p$ constant matrix of full rank p and C is such that $A = CC'$. Then $C'A^{-1}C = C'(CC')^{-1}C = C'(C')^{-1}C^{-1}C = I_p$. Consequently, the rows of U, denoted by $U_1, \ldots, U_p$, are independently distributed as real q-variate Gaussian vectors having the common covariance matrix B^{-1}.*

It is easy to construct such a C. Since $A = (a_{ij})$ is real positive definite, set it as $A = CC'$ where C is a lower triangular matrix with positive diagonal elements. The first row, first column element in $C = (c_{ij})$ is $c_{11} = +\sqrt{a_{11}}$. Note that since $A > O$, all the diagonal elements are real positive. The first column of C is readily available from the first column of A and c_{11}. Now, given a_{22} and the first column in C, c_{22} can be determined, and so on.

Theorem 4.3.6. *Let C, G and X be as defined in Theorem 4.3.5. Consider the $r \times s$ real matrix $Z = C'XG$. Then, when X has the distribution specified in (4.3.8), Z has an $r \times s$ real matrix-variate Gaussian density with M replaced by $C'MG$, A^{-1} replaced by $C'A^{-1}C$ and B^{-1} replaced by $G'B^{-1}G$.*

Example 4.3.4. Let the 2×2 matrix $X = (x_{ij})$ have a real matrix-variate Gaussian density with the parameters M, A and B, and consider the set of linear functions $Z = C'XG$ where C' is a $p \times p$ constant matrix of rank p and G is a $q \times q$ constant matrix of rank q, where

$$M = \begin{bmatrix} 2 & -1 \\ 1 & 5 \end{bmatrix},\ A = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix},\ B = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix},$$
$$C' = \begin{bmatrix} \sqrt{2} & -\frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} \end{bmatrix},\ G = \begin{bmatrix} \sqrt{2} & 0 \\ \frac{1}{\sqrt{2}} & \sqrt{\frac{5}{2}} \end{bmatrix}.$$

Show that all the elements z_{ij}'s in $Z = (z_{ij})$ are mutually independently distributed real scalar standard Gaussian random variables when $M = O$.

Solution 4.3.4. We have already shown in Example 4.3.3 that $C'A^{-1}C = I$. Let us verify that $GG' = B$ and compute $G'B^{-1}G$:

$$GG' = \begin{bmatrix} \sqrt{2} & 0 \\ \frac{1}{\sqrt{2}} & \sqrt{\frac{5}{2}} \end{bmatrix} \begin{bmatrix} \sqrt{2} & \frac{1}{\sqrt{2}} \\ 0 & \sqrt{\frac{5}{2}} \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix} = B;$$

$$B^{-1} = \frac{1}{5}\begin{bmatrix} 3 & -1 \\ -1 & 2 \end{bmatrix},$$
$$G'B^{-1}G = \frac{1}{5}\begin{bmatrix} \sqrt{2} & \frac{1}{\sqrt{2}} \\ 0 & \sqrt{\frac{5}{2}} \end{bmatrix}\begin{bmatrix} 3 & -1 \\ -1 & 2 \end{bmatrix} G = \frac{1}{5}\begin{bmatrix} 3\sqrt{2} - \frac{1}{\sqrt{2}} & 0 \\ -\sqrt{\frac{5}{2}} & 2\sqrt{\frac{5}{2}} \end{bmatrix} G$$
$$= \frac{1}{5}\begin{bmatrix} 3\sqrt{2} - \frac{1}{\sqrt{2}} & 0 \\ -\sqrt{\frac{5}{2}} & 2\sqrt{\frac{5}{2}} \end{bmatrix}\begin{bmatrix} \sqrt{2} & 0 \\ \frac{1}{\sqrt{2}} & \sqrt{\frac{5}{2}} \end{bmatrix} = \frac{1}{5}\begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix} = I_2.$$

Thus, A^{-1} is replaced by $C'A^{-1}C = I_2$ and B^{-1} is replaced by $G'B^{-1}G = I_2$ in the mgf of Z, so that the exponent in the mgf, excluding $\text{tr}(TM')$, is $\frac{1}{2}\text{tr}(TT')$. It follows that all the elements in $Z = C'XG$ are mutually independently distributed real scalar standard normal variables. This completes the computations.

The previous example also suggests the following results which are stated as corollaries:

Corollary 4.3.2. *Let the $p \times q$ real matrix $X = (x_{ij})$ have a real matrix-variate Gaussian density with parameters M, A and B, as given in (4.3.8). Consider the set of linear functions $Y = XG$ where G is a $q \times q$ constant matrix of full rank q, and let $B = GG'$. Then, the columns of Y, denoted by $Y_{(1)}, \ldots, Y_{(q)}$, are independently distributed p-variate real Gaussian vectors with common covariance matrix A^{-1} and mean value $(MG)_{(j)}$, $j = 1, \ldots, q$, where $(MG)_{(j)}$ is the j-th column of MG.*

Corollary 4.3.3. *Let $Z = C'XG$ where C is a $p \times p$ constant matrix of rank p, G is a $q \times q$ constant matrix of rank q and X is a real $p \times q$ Gaussian matrix whose parameters are $M = O$, A and B, the constant matrices C and G being such that $A = CC'$ and $B = GG'$. Then, all the elements z_{ij} in $Z = (z_{ij})$ are mutually independently distributed real scalar standard Gaussian random variables.*

4.3a.2. Linear functions in the complex case

We can similarly obtain results parallel to Theorems 4.3.1–4.3.6 in the complex case. Let $\tilde{X}$ be $p \times q$ matrix in the complex domain, whose elements are scalar complex variables. Assume that $\tilde{X}$ has a complex $p \times q$ matrix-variate density as specified in (4.2*a*.9) whose associated moment generating function is as given in (4.3*a*.3). Let $\tilde{C}_1$ be a $p \times 1$ constant vector, $\tilde{C}_2$ be a $q \times 1$ constant vector, $\tilde{C}$ be a $r \times p$, $r \leq p$, constant matrix of rank r and $\tilde{G}$ be a $q \times s,\ s \leq q$, a constant matrix of rank s. Then, we have the following results:

Theorem 4.3a.1. *Let $\tilde{C}_1$ be a $p \times 1$ constant vector as defined above and let the $p \times q$ matrix $\tilde{X}$ have the density given in (4.2a.9) whose associated mgf is as specified in (4.3a.3). Let $\tilde{U}$ be the linear function of $\tilde{X}$, $\tilde{U} = \tilde{C}_1^*\tilde{X}$. Then $\tilde{U}$ has a q-variate complex Gaussian density with the mean value vector $\tilde{C}_1^*\tilde{M}$ and covariance matrix $[\tilde{C}_1^*A^{-1}\tilde{C}_1]B^{-1}$.*

Theorem 4.3a.2. *Let $\tilde{C}_2$ be a $q \times 1$ constant vector. Consider the linear function $\tilde{Y} = \tilde{X}\tilde{C}_2$ where the $p \times q$ complex matrix $\tilde{X}$ has the density (4.2a.9). Then $\tilde{Y}$ is a p-variate complex Gaussian random vector with the mean value vector $\tilde{M}\tilde{C}_2$ and the covariance matrix $[\tilde{C}_2^*B^{-1}\tilde{C}_2]A^{-1}$.*

Note 4.3a.1. Consider the mgf's of $\tilde{U}$ and $\tilde{Y}$ in Theorems 4.3a.1 and 4.3a.2, namely $M_{\tilde{U}}(\tilde{T}) = E[\mathrm{e}^{\Re(\tilde{T}\tilde{U}^*)}]$ and $M_{\tilde{Y}}(\tilde{T}) = E[\mathrm{e}^{\Re(\tilde{Y}^*\tilde{T})}]$ with the conjugate transpose of the variable part in the linear form in the exponent; then $\tilde{T}$ in $M_{\tilde{U}}(\tilde{T})$ has to be $1 \times q$ and $\tilde{T}$ in $M_{\tilde{Y}}(\tilde{T})$ has to be $p \times 1$. Thus, the exponent in Theorem 4.3a.1 will be of the form $[\tilde{C}_1^*A^{-1}\tilde{C}_1]\tilde{T}B^{-1}\tilde{T}^*$ whereas the corresponding exponent in Theorem 4.3a.2 will be $[\tilde{C}_2^*B^{-1}\tilde{C}_2]\tilde{T}^*A^{-1}\tilde{T}$. Note that in one case, we have $\tilde{T}B^{-1}\tilde{T}^*$ and in the other case, $\tilde{T}$ and $\tilde{T}^*$ are interchanged as are A and B. This has to be kept in mind when applying these theorems.

Example 4.3a.2. Consider a 2×2 matrix $\tilde{X}$ having a complex 2×2 matrix-variate Gaussian density with the parameters $M = O$, A and B, as well as the 2×1 vectors L_1 and L_2 and the linear functions $L_1^*\tilde{X}$ and $\tilde{X}L_2$ where

$$A=\begin{bmatrix}2 & -i\\ i & 1\end{bmatrix},\ B=\begin{bmatrix}1 & i\\ -i & 2\end{bmatrix},\ L_1=\begin{bmatrix}-2i\\ 3i\end{bmatrix},\ L_2=\begin{bmatrix}-i\\ -2i\end{bmatrix}\text{ and }\tilde{X}=\begin{bmatrix}\tilde{x}_{11} & \tilde{x}_{12}\\ \tilde{x}_{21} & \tilde{x}_{22}\end{bmatrix}.$$

Evaluate the densities of $\tilde{U}=L_1^*\tilde{X}$ and $\tilde{Y}=\tilde{X}L_2$ by applying Theorems 4.3a.1 and 4.3a.2, as well as independently.

Solution 4.3a.2. First, we compute the following quantities:

$$A^{-1}=\begin{bmatrix}1 & i\\ -i & 2\end{bmatrix},\ B^{-1}=\begin{bmatrix}2 & -i\\ i & 1\end{bmatrix},$$

$$L_1^*=[2i,-3i],\ L_2^*=[i,2i],$$

so that

$$L_1^*A^{-1}L_1=[2i,-3i]\begin{bmatrix}1 & i\\ -i & 2\end{bmatrix}\begin{bmatrix}-2i\\ 3i\end{bmatrix}=22$$
$$L_2^*B^{-1}L_2=[i,2i]\begin{bmatrix}2 & -i\\ i & 1\end{bmatrix}\begin{bmatrix}-i\\ -2i\end{bmatrix}=6.$$

Then, as per Theorems 4.3a.1 and 4.3a.2, $\tilde{U}$ is a q-variate complex Gaussian vector whose covariance matrix is $22\,B^{-1}$ and $\tilde{Y}$ is a p-variate complex Gaussian vector whose covariance matrix is $6\,A^{-1}$, that is, $\tilde{U}\sim\tilde{N}_2(O,22\,B^{-1})$, $\tilde{Y}\sim\tilde{N}_2(O,6\,A^{-1})$. Now, let us determine the densities of $\tilde{U}$ and $\tilde{Y}$ without resorting to these theorems. Consider the mgf of $\tilde{U}$ by taking the parameter vector $\tilde{T}$ as $\tilde{T}=[\tilde{t}_1,\tilde{t}_2]$. Note that

$$\tilde{T}\tilde{U}^*=t_1(-2i\tilde{x}_{11}^*+3i\tilde{x}_{21}^*)+t_2(-2i\tilde{x}_{12}^*+3i\tilde{x}_{22}^*). \qquad (i)$$

Then, in comparison with the corresponding part in the mgf of $\tilde{X}$ whose associated general parameter matrix is $\tilde{T}=(\tilde{t}_{ij})$, we have

$$\tilde{t}_{11}=-2i\tilde{t}_1,\ \tilde{t}_{12}=-2i\tilde{t}_2,\ \tilde{t}_{21}=3i\tilde{t}_1,\ \tilde{t}_{22}=3i\tilde{t}_2. \qquad (ii)$$

We now substitute the values of *(ii)* in the general mgf of $\tilde{X}$ to obtain the mgf of $\tilde{U}$. Thus,

$$A^{-1}\tilde{T}=\begin{bmatrix}1 & i\\ -i & 2\end{bmatrix}\begin{bmatrix}-2i\tilde{t}_1 & -2i\tilde{t}_2\\ 3i\tilde{t}_1 & 3i\tilde{t}_2\end{bmatrix}=\begin{bmatrix}(-3-2i)\tilde{t}_1 & (-3-2i)\tilde{t}_2\\ (-2+6i)\tilde{t}_1 & (-2+6i)\tilde{t}_2\end{bmatrix}$$
$$B^{-1}\tilde{T}^*=\begin{bmatrix}2 & -i\\ i & 1\end{bmatrix}\begin{bmatrix}2i\tilde{t}_1^* & -3i\tilde{t}_1^*\\ 2i\tilde{t}_2^* & -3i\tilde{t}_2^*\end{bmatrix}=\begin{bmatrix}4i\tilde{t}_1^*+2\tilde{t}_2^* & -6i\tilde{t}_1^*-3\tilde{t}_2^*\\ -2\tilde{t}_1^*+2i\tilde{t}_2^* & 3\tilde{t}_1^*-3i\tilde{t}_2^*\end{bmatrix}.$$

Here an asterisk only denotes the conjugate as the quantities are scalar.

$$\begin{aligned}\mathrm{tr}[A^{-1}\tilde{T}B^{-1}\tilde{T}^*] &= [-3-2i]\tilde{t}_1[4i\tilde{t}_1^*+2\tilde{t}_2^*]+[(-3-2i)\tilde{t}_2[-2\tilde{t}_1^*+2i\tilde{t}_2^*]\\ &\quad +[(-2+6i)\tilde{t}_1][(-6i\tilde{t}_1^*-3\tilde{t}_2^*]+[(-2+6i)\tilde{t}_2[3\tilde{t}_1^*-3i\tilde{t}_2^*]\\ &= 22\,[2\tilde{t}_1\tilde{t}_1^*-i\tilde{t}_1\tilde{t}_2^*+i\tilde{t}_2\tilde{t}_1^*+\tilde{t}_2\tilde{t}_2^*]\\ &= 22\,[\tilde{t}_1,\tilde{t}_2]\begin{bmatrix}2 & -i\\ i & 1\end{bmatrix}\begin{bmatrix}\tilde{t}_1^*\\ \tilde{t}_2^*\end{bmatrix}=22\,\tilde{T}B^{-1}\tilde{T}^*,\ \tilde{T}=[\tilde{t}_1,\tilde{t}_2].\end{aligned}$$

This shows that $\tilde{U}\sim\tilde{N}_2(O,22B^{-1})$. Now, consider

$$\tilde{Y}=\tilde{X}L_2=\begin{bmatrix}\tilde{x}_{11} & \tilde{x}_{12}\\ \tilde{x}_{21} & \tilde{x}_{22}\end{bmatrix}\begin{bmatrix}-i\\ -2i\end{bmatrix}=\begin{bmatrix}-i\tilde{x}_{11}-2i\tilde{x}_{12}\\ -i\tilde{x}_{21}-2i\tilde{x}_{22}\end{bmatrix}.$$

Then, on comparing the mgf of $\tilde{Y}$ with that of $\tilde{X}$ whose general parameter matrix is $\tilde{T}=(\tilde{t}_{ij})$, we have

$$\tilde{t}_{11}=i\tilde{t}_1,\ \tilde{t}_{12}=2i\tilde{t}_1,\ \tilde{t}_{21}=i\tilde{t}_2,\ \tilde{t}_{22}=2i\tilde{t}_2.$$

On substituting these values in the mgf of $\tilde{X}$, we have

$$\begin{aligned}A^{-1}\tilde{T}&=\begin{bmatrix}1 & i\\ -i & 2\end{bmatrix}\begin{bmatrix}i\tilde{t}_1 & 2i\tilde{t}_1\\ i\tilde{t}_2 & 2i\tilde{t}_2\end{bmatrix}=\begin{bmatrix}i\tilde{t}_1-\tilde{t}_2 & 2i\tilde{t}_1-2\tilde{t}_2\\ \tilde{t}_1+2i\tilde{t}_2 & 2\tilde{t}_1+4i\tilde{t}_2\end{bmatrix}\\ B^{-1}\tilde{T}^*&=\begin{bmatrix}2 & -i\\ i & 1\end{bmatrix}\begin{bmatrix}-i\tilde{t}_1^* & -i\tilde{t}_2^*\\ -2i\tilde{t}_1^* & -2i\tilde{t}_2^*\end{bmatrix}=\begin{bmatrix}(-2-2i)\tilde{t}_1^* & (-2-2i)\tilde{t}_2^*\\ (1-2i)\tilde{t}_1^* & (1-2i)\tilde{t}_2^*\end{bmatrix},\end{aligned}$$

so that

$$\begin{aligned}\mathrm{tr}[A^{-1}\tilde{T}B^{-1}\tilde{T}^*] &= [(\tilde{t}_1-\tilde{t}_2)][(-2-2i)\tilde{t}_1^*]+[2i\tilde{t}_1-2\tilde{t}_2][(1-2i)\tilde{t}_1^*]\\ &\quad +[\tilde{t}_1+2i\tilde{t}_2][(-2-2i)\tilde{t}_2^*]+[2\tilde{t}_1+4i\tilde{t}_2][(1-2i)\tilde{t}_2^*]\\ &= 6\,[\tilde{t}_1\tilde{t}_1^*-i\tilde{t}_1\tilde{t}_2^*+i\tilde{t}_2\tilde{t}_1^*+2\tilde{t}_2\tilde{t}_2^*]\\ &= 6\,[\tilde{t}_1^*,\tilde{t}_2^*]\begin{bmatrix}1 & i\\ -i & 2\end{bmatrix}\begin{bmatrix}\tilde{t}_1\\ \tilde{t}_2\end{bmatrix}=6\,\tilde{T}^*A^{-1}\tilde{T};\end{aligned}$$

refer to Note 4.3a.1 regarding the representation of the quadratic forms in the two cases above. This shows that $\tilde{Y}\sim\tilde{N}_2(O,6A^{-1})$. This completes the computations.

Theorem 4.3a.3. *Let $\tilde{C}_1$ be a constant $p\times 1$ vector, $\tilde{X}$ be a $p\times q$ matrix whose elements are complex scalar variables and let $A=A^*>O$ be $p\times p$ and $B=B^*>O$ be $q\times q$ constant Hermitian positive definite matrices, where an asterisk denotes the conjugate* *transpose. If, for an arbitrary $p \times 1$ constant vector $\tilde{C}_1$, $\tilde{C}_1^* \tilde{X}$ is a q-variate complex Gaussian vector as specified in Theorem 4.3a.1, then $\tilde{X}$ has the $p \times q$ complex matrix-variate Gaussian density given in (4.2a.9).*

As well, a result parallel to this one follows from Theorem 4.3a.2:

Theorem 4.3a.4. *Let $\tilde{C}_2$ be a $q \times 1$ constant vector, $\tilde{X}$ be a $p \times q$ matrix whose elements are complex scalar variables and let $A > O$ be $p \times p$ and $B > O$ be $q \times q$ Hermitian positive definite constant matrices. If, for an arbitrary constant vector $\tilde{C}_2$, $\tilde{X}\tilde{C}_2$ is a p-variate complex Gaussian vector as specified in Theorem 4.3a.2, then $\tilde{X}$ is $p \times q$ complex matrix-variate Gaussian matrix which is distributed as in (4.2a.9).*

Theorem 4.3a.5. *Let $\tilde{C}^*$ be a $r \times p$, $r \le p$, complex constant matrix of full rank r and $\tilde{G}$ be a $q \times s$, $s \le q$, constant complex matrix of full rank s. Let $\tilde{U} = \tilde{C}^* \tilde{X}$ and $\tilde{W} = \tilde{X}\tilde{G}$ where $\tilde{X}$ has the density given in (4.2a.9). Then, $\tilde{U}$ has a $r \times q$ complex matrix-variate density with $\tilde{M}$ replaced by $\tilde{C}^*\tilde{M}$, A^{-1} replaced by $\tilde{C}^* A^{-1} \tilde{C}$ and B^{-1} remaining the same, and $\tilde{W}$ has a $p \times s$ complex matrix-variate distribution with B^{-1} replaced by $\tilde{G}^* B^{-1} \tilde{G}$, $\tilde{M}$ replaced by $\tilde{M}\tilde{G}$ and A^{-1} remaining the same.*

Theorem 4.3a.6. *Let $\tilde{C}^*$, $\tilde{G}$ and $\tilde{X}$ be as defined in Theorem 4.3a.5. Consider the $r \times s$ complex matrix $\tilde{Z} = \tilde{C}^* \tilde{X} \tilde{G}$. Then when $\tilde{X}$ has the distribution specified by (4.2a.9), $\tilde{Z}$ has an $r \times s$ complex matrix-variate density with $\tilde{M}$ replaced by $\tilde{C}^* \tilde{M} \tilde{G}$, A^{-1} replaced by $\tilde{C}^* A^{-1} \tilde{C}$ and B^{-1} replaced by $\tilde{G}^* B^{-1} \tilde{G}$.*

Example 4.3a.3. Consider a 2×3 matrix $\tilde{X}$ having a complex matrix-variate Gaussian distribution with the parameter matrices $\tilde{M} = O$, A and B where

$$A = \begin{bmatrix} 2 & i \\ -i & 1 \end{bmatrix},\ B = \begin{bmatrix} 3 & i & 0 \\ -i & 1 & i \\ 0 & -i & 2 \end{bmatrix},\ \tilde{X} = \begin{bmatrix} \tilde{x}_{11} & \tilde{x}_{12} & \tilde{x}_{13} \\ \tilde{x}_{21} & \tilde{x}_{22} & \tilde{x}_{23} \end{bmatrix}.$$

Consider the linear forms

$$C^*\tilde{X} = \begin{bmatrix} i\tilde{x}_{11} - i\tilde{x}_{21} & i\tilde{x}_{12} - i\tilde{x}_{22} & i\tilde{x}_{13} - i\tilde{x}_{23} \\ \tilde{x}_{11} + 2\tilde{x}_{21} & \tilde{x}_{12} + 2\tilde{x}_{22} & \tilde{x}_{13} + 2\tilde{x}_{23} \end{bmatrix}$$

$$\tilde{X}G = \begin{bmatrix} \tilde{x}_{11} + i\tilde{x}_{12} + 2\tilde{x}_{13} & \tilde{x}_{12} & i\tilde{x}_{11} - i\tilde{x}_{12} + i\tilde{x}_{13} \\ \tilde{x}_{21} + i\tilde{x}_{22} + 2\tilde{x}_{23} & \tilde{x}_{22} & i\tilde{x}_{21} - i\tilde{x}_{22} + i\tilde{x}_{23} \end{bmatrix}.$$

(1): Compute the distribution of $\tilde{Z} = C^*\tilde{X}G$; (2): Compute the distribution of $\tilde{Z} = C^*\tilde{X}G$ if A remains the same and G is equal to

$$\begin{bmatrix} \sqrt{3} & 0 & 0 \\ -\frac{i}{\sqrt{3}} & \sqrt{\frac{2}{3}} & 0 \\ 0 & -i\sqrt{\frac{3}{2}} & \frac{1}{\sqrt{2}} \end{bmatrix},$$

and study the properties of this distribution.

Solution 4.3a.3. Note that $A = A^*$ and $B = B^*$ and hence both A and B are Hermitian. Moreover, $|A| = 1$ and $|B| = 1$ and since all the leading minors of A and B are positive, A and B are both Hermitian positive definite. Then, the inverses of A and B in terms of the cofactors of their elements are

$$A^{-1} = \begin{bmatrix} 1 & -i \\ i & 2 \end{bmatrix},\ [\mathrm{Cof}(B)]' = \begin{bmatrix} 1 & -2i & -1 \\ 2i & 6 & -3i \\ -1 & 3i & 2 \end{bmatrix} = B^{-1}.$$

The linear forms provided above in connection with part (1) of this exercise can be respectively expressed in terms of the following matrices:

$$C^* = \begin{bmatrix} i & -i \\ 1 & 2 \end{bmatrix},\ G = \begin{bmatrix} 1 & 0 & i \\ i & 1 & -i \\ 2 & 0 & i \end{bmatrix}.$$

Let us now compute $C^*A^{-1}C$ and $G^*B^{-1}G$:

$$C^*A^{-1}C = \begin{bmatrix} i & -i \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 1 & -i \\ i & 2 \end{bmatrix}\begin{bmatrix} -i & 1 \\ i & 2 \end{bmatrix} = 3\begin{bmatrix} 1 & 1-i \\ 1+i & 3 \end{bmatrix}$$

$$G^*B^{-1}G = \begin{bmatrix} 1 & -i & 2 \\ 0 & 1 & 0 \\ -i & i & -i \end{bmatrix}\begin{bmatrix} 1 & -2i & -1 \\ 2i & 6 & -3i \\ -1 & 3i & 2 \end{bmatrix}\begin{bmatrix} 1 & 0 & i \\ i & 1 & -i \\ 2 & 0 & i \end{bmatrix}$$
$$= \begin{bmatrix} 3 & -2i & -2+i \\ 2i & 6 & 1-6i \\ -2-i & 1+6i & 7 \end{bmatrix}.$$

Then in (1), $C^*\tilde{X}G$ is a 2×3 complex matrix-variate Gaussian with A^{-1} replaced by $C^*A^{-1}C$ and B^{-1} replaced by $G^*B^{-1}G$ where $C^*A^{-1}C$ and $G^*B^{-1}G$ are given above. For answering (2), let us evaluate $G^*B^{-1}G$:

$$G^*B^{-1}G = \begin{bmatrix} \sqrt{3} & \frac{i}{\sqrt{3}} & 0 \\ 0 & \sqrt{\frac{2}{3}} & i\sqrt{\frac{3}{2}} \\ 0 & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 1 & -2i & -1 \\ 2i & 6 & -3i \\ -1 & 3i & 2 \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 & 0 \\ -\frac{i}{\sqrt{3}} & \sqrt{\frac{2}{3}} & 0 \\ 0 & -i\sqrt{\frac{3}{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$
$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I_3.$$

Observe that this $q \times q$ matrix G which is such that $GG^* = B$, is nonsingular; thus, $G^*B^{-1}G = G^*(GG^*)^{-1}G = G^*(G^*)^{-1}G^{-1}G = I$. Letting $\tilde{Y} = \tilde{X}G$, $\tilde{X} = \tilde{Y}G^{-1}$, and the exponent in the density of $\tilde{X}$ becomes

$$\text{tr}(A^{-1}\tilde{X}B\tilde{X}^*) = \text{tr}(A^{-1}\tilde{Y}G^{-1}B(G^*)^{-1}\tilde{Y}^*) = \text{tr}(Y^*A\tilde{Y}) = \sum_{j=1}^{p} \tilde{Y}_{(j)}^* A\tilde{Y}_{(j)}$$

where the $\tilde{Y}_{(j)}$'s are the columns of $\tilde{Y}$, which are independently distributed complex p-variate Gaussian vectors with common covariance matrix A^{-1}. This completes the computations.

The conclusions obtained in the solution to Example 4.3a.1 suggest the corollaries that follow.

Corollary 4.3a.1. *Let the $p \times q$ matrix $\tilde{X}$ have a matrix-variate complex Gaussian distribution with the parameters $M = O$, $A > O$ and $B > O$. Consider the transformation $\tilde{U} = C^*\tilde{X}$ where C is a $p \times p$ nonsingular matrix such that $CC^* = A$ so that $C^*A^{-1}C = I$. Then the rows of $\tilde{U}$, namely $\tilde{U}_1, \ldots, \tilde{U}_p$, are mutually independently distributed q-variate complex Gaussian vectors with common covariance matrix B^{-1}.*

Corollary 4.3a.2. *Let the $p \times q$ matrix $\tilde{X}$ have a matrix-variate complex Gaussian distribution with the parameters $M = O$, $A > O$ and $B > O$. Consider the transformation $\tilde{Y} = \tilde{X}G$ where G is a $q \times q$ nonsingular matrix such that $GG^* = B$ so that $G^*B^{-1}G = I$. Then the columns of $\tilde{Y}$, namely $\tilde{Y}_{(1)}, \ldots, \tilde{Y}_{(q)}$, are independently distributed p-variate complex Gaussian vectors with common covariance matrix A^{-1}.*

Corollary 4.3a.3. *Let the $p \times q$ matrix $\tilde{X}$ have a matrix-variate complex Gaussian distribution with the parameters $M = O$, $A > O$ and $B > O$. Consider the transformation $\tilde{Z} = C^*\tilde{X}G$ where C is a $p \times p$ nonsingular matrix such that $CC^* = A$ and G is a $q \times q$ nonsingular matrix such that $GG^* = B$. Then, the elements $\tilde{z}_{ij}$'s of $\tilde{Z} = (\tilde{z}_{ij})$ are mutually independently distributed complex standard Gaussian variables.*

4.3.3. Partitioning of the parameter matrix

Suppose that in the $p \times q$ real matrix-variate case, we partition T as $\begin{bmatrix} T_1 \\ T_2 \end{bmatrix}$ where T_1 is $p_1 \times q$ and T_2 is $p_2 \times q$, so that $p_1 + p_2 = p$. Let $T_2 = O$ (a null matrix). Then,

$$TB^{-1}T' = \begin{pmatrix} T_1 \\ O \end{pmatrix} B^{-1} \begin{pmatrix} T_1' & O \end{pmatrix} = \begin{bmatrix} T_1B^{-1}T_1' & O_1 \\ O_2 & O_3 \end{bmatrix}$$

where $T_1B^{-1}T_1'$ is a $p_1 \times p_1$ matrix, O_1 is a $p_1 \times p_2$ null matrix, O_2 is a $p_2 \times p_1$ null matrix and O_3 is a $p_2 \times p_2$ null matrix. Let us similarly partition A^{-1} into sub-matrices:

$$A^{-1} = \begin{bmatrix} A^{11} & A^{12} \\ A^{21} & A^{22} \end{bmatrix},$$

where A^{11} is $p_1 \times p_1$ and A^{22} is $p_2 \times p_2$. Then,

$$\mathrm{tr}(A^{-1}TB^{-1}T') = \mathrm{tr} \begin{bmatrix} A^{11}T_1B^{-1}T_1' & O \\ O & O \end{bmatrix} = \mathrm{tr}(A^{11}T_1B^{-1}T_1').$$

If A is partitioned as

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

where A_{11} is $p_1 \times p_1$ and A_{22} is $p_2 \times p_2$, then, as established in Sect. 1.3, we have

$$A^{11} = (A_{11} - A_{12}A_{22}^{-1}A_{21})^{-1}.$$

Therefore, under this special case of T, the mgf is given by

$$E[\mathrm{e}^{\mathrm{tr}(T_1X_1)}] = \mathrm{e}^{\frac{1}{2}\mathrm{tr}((A_{11}-A_{12}A_{22}^{-1}A_{21})^{-1}T_1B^{-1}T_1')}, \tag{4.3.9}$$

which is also the mgf of the $p_1 \times q$ sub-matrix of X. Note that the mgf's in (4.3.9) and (4.3.1) share an identical structure. Hence, due to the uniqueness of the mgf, X_1 has a real $p_1 \times q$ matrix-variate Gaussian density wherein the parameter B remains unchanged and A is replaced by $A_{11} - A_{12}A_{22}^{-1}A_{21}$, the A_{ij}'s denoting the sub-matrices of A as described earlier.

4.3.4. Distributions of quadratic and bilinear forms

Consider the real $p \times q$ Gaussian matrix U defined in (4.2.17) whose mean value matrix is $E[X] = M = O$ and let $U = XB^{\frac{1}{2}}$. Then,

$$U'AU = \begin{bmatrix} U_1'AU_1 & U_1'AU_2 & \dots & U_1'AU_q \\ U_2'AU_1 & U_2'AU_2 & \dots & U_2'AU_q \\ \vdots & \vdots & \ddots & \vdots \\ U_q'AU_1 & U_q'AU_2 & \dots & U_q'AU_q \end{bmatrix} \tag{4.3.10}$$

where the $p \times 1$ column vectors of U, namely, $U_1, \dots, U_q$, are independently distributed as $N_p(O, A^{-1})$ vectors, that is, the U_j's are independently distributed real p-variate Gaussian (normal) vectors whose covariance matrix is $A^{-1} = E[UU']$, with density

$$f_{U_j}(U_j) = \frac{|A|^{\frac{1}{2}}}{(2\pi)^{\frac{p}{2}}} \mathrm{e}^{-\frac{1}{2}(U_j'AU_j)},\ A > O. \tag{4.3.11}$$

What is then the distribution of $U_j'AU_j$ for any particular j and what are the distributions of $U_i'AU_j,\ i \neq j = 1, \dots, q$? Let $z_j = U_j'AU_j$ and $z_{ij} = U_i'AU_j$, $i \neq j$. Letting t be a scalar parameter, consider the mgf of z_j:

$$\begin{aligned} M_{z_j}(t) = E[\mathrm{e}^{tz_j}] &= \int_{U_j} \mathrm{e}^{tU_j'AU_j} f_{U_j}(U_j)\mathrm{d}U_j \\ &= \frac{|A|^{\frac{1}{2}}}{(2\pi)^{\frac{p}{2}}} \int_{U_j} \mathrm{e}^{-\frac{1}{2}(1-2t)U_j'AU_j}\mathrm{d}U_j \\ &= (1-2t)^{-\frac{p}{2}} \quad \text{for } 1-2t > 0, \end{aligned}$$

which is the mgf of a real gamma random variable with parameters $\alpha = \frac{p}{2},\ \beta = 2$ or a real chi-square random variable with p degrees of freedom for $p = 1, 2, \dots$. That is,

$$U_j'AU_j \sim \chi_p^2 \ \text{(a real chi-square random variable having } p \text{ degrees of freedom).} \tag{4.3.12}$$

In the complex case, observe that $\tilde{U}_j^*A\tilde{U}_j$ is real when $A = A^* > O$ and hence, the parameter in the mgf is real. On making the transformation $A^{\frac{1}{2}}\tilde{U}_j = \tilde{V}_j$, $|\det(A)|$ is canceled. Then, the exponent can be expressed in terms of

$$-(1-t)\tilde{Y}^*\tilde{Y} = -(1-t)\sum_{j=1}^{p}|\tilde{y}_j|^2 = -(1-t)\sum_{j=1}^{p}(y_{j1}^2 + y_{j2}^2),$$

where $\tilde{y}_j = y_{j1} + iy_{j2},\ i = \sqrt{(-1)}$. The integral gives $(1-t)^{-p}$ for $1-t > 0$. Hence, $\tilde{V}_j = \tilde{U}_j^* A\tilde{U}_j$ has a real gamma distribution with the parameters $\alpha = p,\ \beta = 1$, that is, a chi-square distribution with p degrees of freedom in the complex domain. Thus, $2\tilde{V}_j$ is a real chi-square random variable with $2p$ degrees of freedom, that is,

$$2\tilde{V}_j = 2\,\tilde{U}_j^* A\tilde{U}_j \sim \chi^2_{2p}. \tag{4.3a.9}$$

What is then the distribution of $U_i' A U_j,\ i \neq j$? Let us evaluate the mgf of $U_i' A U_j = z_{ij}$. As z_{ij} is a function of U_i and U_j, we can integrate out over the joint density of U_i and U_j where U_i and U_j are independently distributed p-variate real Gaussian random variables:

$$\begin{aligned} M_{z_{ij}}(t) = E[\mathrm{e}^{tz_{ij}}] &= \int_{U_i}\int_{U_j} \mathrm{e}^{t\,U_i' A U_j} f_{U_i}(U_i) f_{U_j}(U_j)\mathrm{d}U_i \wedge \mathrm{d}U_j \\ &= \frac{|A|}{(2\pi)^p}\int\int \mathrm{e}^{t\ U_i' A U_j - \frac{1}{2}U_j' A U_j - \frac{1}{2}U_i' A U_i}\mathrm{d}U_i \wedge \mathrm{d}U_j. \end{aligned}$$

Let us first integrate out U_j. The relevant terms in the exponent are

$$-\frac{1}{2}(U_j' A\,U_j) + \frac{1}{2}(2t)(U_i' A\,U_j) = -\frac{1}{2}(U_j - C)' A\,(U_j - C) + \frac{1}{2}t^2 U_i' A\,U_i,\ C = t\,U_i.$$

But the integral over U_j which is the integral over $U_j - C$ will result in the following representation:

$$\begin{aligned} M_{z_{ij}}(t) &= \frac{|A|^{\frac{1}{2}}}{(2\pi)^{\frac{p}{2}}}\int_{U_i} \mathrm{e}^{\frac{t^2}{2}U_i' A U_i - \frac{1}{2}U_i' A U_i}\mathrm{d}U_i \\ &= (1-t^2)^{-\frac{p}{2}} \quad \text{for } 1 - t^2 > 0. \end{aligned} \tag{4.3.13}$$

What is the density corresponding to the mgf $(1-t^2)^{-\frac{p}{2}}$? This is the mgf of a real scalar random variable u of the form $u = x - y$ where x and y are independently distributed real scalar chi-square random variables. For $p = 2$, x and y will be exponential variables so that u will have a double exponential distribution or a real Laplace distribution. In the general case, the density of u can also be worked out when x and y are independently distributed real gamma random variables with different parameters whereas chi-squares with equal degrees of freedom constitutes a special case. For the exact distribution of covariance structures such as the z_{ij}'s, see Mathai and Sebastian (2022).

Exercises 4.3

4.3.1. In the moment generating function (mgf) (4.3.3), partition the $p \times q$ parameter matrix T into column sub-matrices such that $T = (T_1, T_2)$ where T_1 is $p \times q_1$ and T_2 is $p \times q_2$ with $q_1 + q_2 = q$. Take $T_2 = O$ (the null matrix). Simplify and show that if X is similarly partitioned as $X = (Y_1, Y_2)$, then Y_1 has a real $p \times q_1$ matrix-variate Gaussian density. As well, show that Y_2 has a real $p \times q_2$ matrix-variate Gaussian density.

4.3.2. Referring to Exercises 4.3.1, write down the densities of Y_1 and Y_2.

4.3.3. If T is the parameter matrix in (4.3.3), then what type of partitioning of T is required so that the densities of (1): the first row of X, (2): the first column of X can be determined, and write down these densities explicitly.

4.3.4. Repeat Exercises 4.3.1–4.3.3 by taking the mgf in (4.3*a*.3) for the corresponding complex case.

4.3.5. Write down the mgf explicitly for $p = 2$ and $q = 2$ corresponding to (4.3.3) and (4.3*a*.3), assuming general $A > O$ and $B > O$.

4.3.6. Partition the mgf in the complex $p \times q$ matrix-variate Gaussian case, corresponding to the partition in Sect. 4.3.1 and write down the complex matrix-variate density corresponding to $\tilde{T}_1$ in the complex case.

4.3.7. In the real $p \times q$ matrix-variate Gaussian case, partition the mgf parameter matrix into $T = (T_{(1)}, T_{(2)})$ where $T_{(1)}$ is $p \times q_1$ and $T_{(2)}$ is $p \times q_2$ with $q_1 + q_2 = q$. Obtain the density corresponding to $T_{(1)}$ by letting $T_{(2)} = O$.

4.3.8. Repeat Exercise 4.3.7 for the complex $p \times q$ matrix-variate Gaussian case.

4.3.9. Consider $v = \tilde{U}_j^* A \tilde{U}_j$. Provide the details of the steps for obtaining (4.3*a*.9).

4.3.10. Derive the mgf of $\tilde{U}_i^* A \tilde{U}_j$, $i \neq j$, in the complex $p \times q$ matrix-variate Gaussian case, corresponding to (4.3.13).

4.4. Marginal Densities in the Real Matrix-variate Gaussian Case

On partitioning the real $p \times q$ Gaussian matrix into X_1 of order $p_1 \times q$ and X_2 of order $p_2 \times q$ so that $p_1 + p_2 = p$, it was determined by applying the mgf technique that X_1 has a $p_1 \times q$ matrix-variate Gaussian distribution with the parameter matrices B remaining unchanged while A was replaced by $A_{11} - A_{12}A_{22}^{-1}A_{21}$ where the A_{ij}'s are the sub-matrices of A. This density is then the marginal density of the sub-matrix X_1 with respect to the joint density of X. Let us see whether the same result is available by direct integration of the remaining variables, namely by integrating out X_2. We first consider the real case. Note that

$$\begin{aligned}\mathrm{tr}(AXBX') &= \mathrm{tr}\left[A\begin{pmatrix}X_1\\X_2\end{pmatrix}B(X_1'\ \ X_2')\right]\\ &= \mathrm{tr}\left[A\begin{pmatrix}X_1BX_1' & X_1BX_2'\\ X_2BX_1' & X_2BX_2'\end{pmatrix}\right].\end{aligned}$$

Now, letting A be similarly partitioned, we have

$$\begin{aligned}\mathrm{tr}(AXBX') &= \mathrm{tr}\left[\begin{pmatrix}A_{11} & A_{12}\\ A_{21} & A_{22}\end{pmatrix}\begin{pmatrix}X_1BX_1' & X_1BX_2'\\ X_2BX_1' & X_2BX_2'\end{pmatrix}\right]\\ &= \mathrm{tr}(A_{11}X_1BX_1') + \mathrm{tr}(A_{12}X_2BX_1')\\ &\quad + \mathrm{tr}(A_{21}X_1BX_2') + \mathrm{tr}(A_{22}X_2BX_2'),\end{aligned}$$

as the remaining terms do not appear in the trace. However, $(A_{12}X_2BX_1')' = X_1BX_2'A_{21}$, and since $\mathrm{tr}(PQ) = \mathrm{tr}(QP)$ and $\mathrm{tr}(S) = \mathrm{tr}(S')$ whenever $S,\ PQ$ and QP are square matrices, we have

$$\mathrm{tr}(AXBX') = \mathrm{tr}(A_{11}X_1BX_1') + 2\mathrm{tr}(A_{21}X_1BX_2') + \mathrm{tr}(A_{22}X_2BX_2').$$

We may now complete the quadratic form in $\mathrm{tr}(A_{22}X_2BX_2') + 2\mathrm{tr}(A_{21}X_1BX_2')$ by taking a matrix $C = A_{22}^{-1}A_{21}X_1$ and replacing X_2 by $X_2 + C$. Note that when $A > O$, $A_{11} > O$ and $A_{22} > O$. Thus,

$$\begin{aligned}\mathrm{tr}(AXBX') &= \mathrm{tr}(A_{22}(X_2+C)B(X_2+C)') + \mathrm{tr}(A_{11}X_1BX_1') - \mathrm{tr}(A_{12}A_{22}^{-1}A_{21}X_1BX_1')\\ &= \mathrm{tr}(A_{22}(X_2+C)B(X_2+C)') + \mathrm{tr}((A_{11} - A_{12}A_{22}^{-1}A_{21})X_1BX_1').\end{aligned}$$

On applying a result on partitioned matrices from Sect. 1.3, we have

$$|A| = |A_{22}|\,|A_{11} - A_{12}A_{22}^{-1}A_{21}|,$$

and clearly, $(2\pi)^{\frac{pq}{2}} = (2\pi)^{\frac{p_1q}{2}}(2\pi)^{\frac{p_2q}{2}}$. When integrating out X_2, $|A_{22}|^{\frac{q}{2}}$ and $(2\pi)^{\frac{p_2q}{2}}$ are getting canceled. Hence, the marginal density of X_1, the $p_1 \times q$ sub-matrix of X, denoted by $f_{p_1,q}(X_1)$, is given by

$$f_{p_1,q}(X_1) = \frac{|B|^{\frac{p_1}{2}}|A_{11} - A_{12}A_{22}^{-1}A_{21}|^{\frac{q}{2}}}{(2\pi)^{\frac{p_1q}{2}}}\mathrm{e}^{-\frac{1}{2}\mathrm{tr}((A_{11}-A_{12}A_{22}^{-1}A_{21})X_1BX_1')}. \tag{4.4.1}$$

When $p_1 = 1,\ p_2 = 0$ and $p = 1$, we have the usual multivariate Gaussian density. When $p = 1$, the 1×1 matrix A will be taken as 1 without any loss of generality.

Then, from (4.4.1), the multivariate (q-variate) Gaussian density corresponding to (4.2.3) is given by

$$f_{1,q}(X_1) = \frac{(1/2)^{\frac{q}{2}}|B|^{\frac{1}{2}}}{\pi^{\frac{q}{2}}}\mathrm{e}^{-\frac{1}{2}\mathrm{tr}(X_1BX_1')} = \frac{|B|^{\frac{1}{2}}}{(2\pi)^{\frac{q}{2}}}\mathrm{e}^{-\frac{1}{2}X_1BX_1'}$$

since the 1×1 quadratic form X_1BX_1' is equal to its trace. It is usually expressed in terms of $B = V^{-1},\ V > O$. When $q = 1$, X is reducing to a $p\times 1$ vector, say Y. Thus, for a $p\times 1$ column vector Y with a location parameter μ, the density, denoted by $f_{p,1}(Y)$, is the following:

$$f_{p,1}(Y) = \frac{1}{|V|^{\frac{1}{2}}(2\pi)^{\frac{p}{2}}}\mathrm{e}^{-\frac{1}{2}(Y-\mu)'V^{-1}(Y-\mu)}, \tag{4.4.2}$$

where $Y' = (y_1,\ldots,y_p),\ \mu' = (\mu_1,\ldots,\mu_p),\ -\infty < y_j < \infty,\ -\infty < \mu_j < \infty,\ j = 1,\ldots,p,\ V > O$. Observe that when Y is $p\times 1$, $\mathrm{tr}(Y-\mu)'V^{-1}(Y-\mu) = (Y-\mu)'V^{-1}(Y-\mu)$. From symmetry, we can write down the density of the sub-matrix X_2 of X from the density given in (4.4.1). Let us denote the density of X_2 by $f_{p_2,q}(X_2)$. Then,

$$f_{p_2,q}(X_2) = \frac{|B|^{\frac{p_2}{2}}|A_{22}-A_{21}A_{11}^{-1}A_{12}|^{\frac{q}{2}}}{(2\pi)^{\frac{p_2q}{2}}}\mathrm{e}^{-\frac{1}{2}\mathrm{tr}((A_{22}-A_{21}A_{11}^{-1}A_{12})X_2BX_2')}. \tag{4.4.3}$$

Note that $A_{22}-A_{21}A_{11}^{-1}A_{12} > O$ as $A > O$, our intial assumptions being that $A > O$ and $B > O$.

Theorem 4.4.1. *Let the $p\times q$ real matrix X have a real matrix-variate Gaussian density with the parameter matrices $A > O$ and $B > O$ where A is $p\times p$ and B is $q\times q$. Let X be partitioned into sub-matrices as $X = \begin{pmatrix}X_1\\X_2\end{pmatrix}$ where X_1 is $p_1\times q$ and X_2 is $p_2\times q$, with $p_1+p_2 = p$. Let A be partitioned into sub-matrices as $A = \begin{bmatrix}A_{11} & A_{12}\\A_{21} & A_{22}\end{bmatrix}$ where A_{11} is $p_1\times p_1$. Then X_1 has a $p_1\times q$ real matrix-variate Gaussian density with the parameter matrices $A_{11}-A_{12}A_{22}^{-1}A_{21} > O$ and $B > O$, as given in (4.4.1) and X_2 has a $p_2\times q$ real matrix-variate Gaussian density with the parameter matrices $A_{22}-A_{21}A_{11}^{-1}A_{12} > O$ and $B > O$, as given in (4.4.3).*

Observe that the p_1 rows taken in X_1 need not be the first p_1 rows. They can be any set of p_1 rows. In that instance, it suffices to make the corresponding permutations in the rows and columns of A and B so that the new set of p_1 rows can be taken as the first p_1 rows, and similarly for X_2.

Can a similar result be obtained in connection with a matrix-variate Gaussian distribution if we take a set of column vectors and form a sub-matrix of X? Let us partition the

$p\times q$ matrix X into sub-matrices of columns as $X=(Y_1\ Y_2)$ where Y_1 is $p\times q_1$ and Y_2 is $p\times q_2$ such that $q_1+q_2=q$. The variables $Y_1,\ Y_2$ are used in order to avoid any confusion with $X_1,\ X_2$ utilized in the discussions so far. Let us partition B as follows:

$$B=\begin{bmatrix}B_{11} & B_{12}\\ B_{21} & B_{22}\end{bmatrix},\ B_{11}\text{ being }q_1\times q_1,\ B_{22}\text{ being }q_2\times q_2.$$

Then,

$$\begin{aligned}\mathrm{tr}(AXBX')&=\mathrm{tr}[A(Y_1\ Y_2)\begin{bmatrix}B_{11} & B_{12}\\ B_{21} & B_{22}\end{bmatrix}\begin{pmatrix}Y_1'\\ Y_2'\end{pmatrix}]\\ &=\mathrm{tr}(AY_1B_{11}Y_1')+\mathrm{tr}(AY_2B_{21}Y_1')+\mathrm{tr}(AY_1B_{12}Y_2')+\mathrm{tr}(AY_2B_{22}Y_2')\\ &=\mathrm{tr}(AY_1B_{11}Y_1')+2\mathrm{tr}(AY_1B_{12}Y_2')+\mathrm{tr}(AY_2B_{22}Y_2').\end{aligned}$$

As in the previous case of row sub-matrices, we complete the quadratic form:

$$\begin{aligned}\mathrm{tr}(AXBX')&=\mathrm{tr}(AY_1B_{11}Y_1')-\mathrm{tr}(AY_1(B_{12}B_{22}^{-1}B_{21}Y_1')+\mathrm{tr}(A(Y_2+C)B_{22}(Y_2+C)')\\ &=\mathrm{tr}(AY_1(B_{11}-B_{12}B_{22}^{-1}B_{21})Y_1')+\mathrm{tr}(A(Y_2+C)B_{22}(Y_2+C)').\end{aligned}$$

Now, by integrating out Y_2, we have the result, observing that $A>O,\ B>O,\ B_{11}-B_{12}B_{22}^{-1}B_{21}>O$ and $|B|=|B_{22}|\,|B_{11}-B_{12}B_{22}^{-1}B_{21}|$. A similar result follows for the marginal density of Y_2. These results will be stated as the next theorem.

Theorem 4.4.2. *Let the $p\times q$ real matrix X have a real matrix-variate Gaussian density with the parameter matrices $M=O,\ A>O$ and $B>O$ where A is $p\times p$ and B is $q\times q$. Let X be partitioned into column sub-matrices as $X=(Y_1\ Y_2)$ where Y_1 is $p\times q_1$ and Y_2 is $p\times q_2$ with $q_1+q_2=q$. Then Y_1 has a $p\times q_1$ real matrix-variate Gaussian density with the parameter matrices $A>O$ and $B_{11}-B_{12}B_{22}^{-1}B_{21}>O$, denoted by $f_{p,q_1}(Y_1)$, and Y_2 has a $p\times q_2$ real matrix-variate Gaussian density denoted by $f_{p,q_2}(Y_2)$ where*

$$f_{p,q_1}(Y_1)=\frac{|A|^{\frac{q_1}{2}}|B_{11}-B_{12}B_{22}^{-1}B_{21}|^{\frac{p}{2}}}{(2\pi)^{\frac{pq_1}{2}}}\mathrm{e}^{-\frac{1}{2}\mathrm{tr}[AY_1(B_{11}-B_{12}B_{22}^{-1}B_{21})Y_1']}\qquad(4.4.4)$$

$$f_{p,q_2}(Y_2)=\frac{|A|^{\frac{q_2}{2}}|B_{22}-B_{21}B_{11}^{-1}B_{12}|^{\frac{p}{2}}}{(2\pi)^{\frac{pq_2}{2}}}\mathrm{e}^{-\frac{1}{2}\mathrm{tr}[AY_2(B_{22}-B_{21}B_{11}^{-1}B_{12})Y_2']}.\qquad(4.4.5)$$

If $q=1$ and $q_2=0$ in (4.4.4), $q_1=1$. When $q=1$, the 1×1 matrix B is taken to be 1. Then Y_1 in (4.4.4) is $p\times 1$ or a column vector of p real scalar variables. Let it still be denoted by Y_1. Then the corresponding density, which is a real p-variate Gaussian (normal) density, available from (4.4.4) or from the basic matrix-variate density, is the following:

$$f_{p,1}(Y_1) = \frac{|A|^{\frac{1}{2}}(1/2)^{\frac{p}{2}}}{\pi^{\frac{p}{2}}} \mathrm{e}^{-\frac{1}{2}\mathrm{tr}(AY_1Y_1')}$$
$$= \frac{|A|^{\frac{1}{2}}}{(2\pi)^{\frac{p}{2}}} \mathrm{e}^{-\frac{1}{2}\mathrm{tr}(Y_1'AY_1)} = \frac{|A|^{\frac{1}{2}}}{(2\pi)^{\frac{p}{2}}} \mathrm{e}^{-\frac{1}{2}Y_1'AY_1}, \tag{4.4.6}$$

observing that $\mathrm{tr}(Y_1'AY_1) = Y_1'AY_1$ since Y_1 is $p \times 1$ and then, $Y_1'AY_1$ is 1×1. In the usual representation of a multivariate Gaussian density, A replaced by $A = V^{-1}$, V being positive definite.

Example 4.4.1. Let the 2×3 matrix $X = (x_{ij})$ have a real matrix-variate distribution with the parameter matrices $M = O,\ A > O,\ B > O$ where

$$X = \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \end{bmatrix},\ A = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix},\ B = \begin{bmatrix} 3 & -1 & 0 \\ -1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix}.$$

Let us partition $X,\ A$ and B as follows:

$$X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = [Y_1, Y_2],\ A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},\ B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

where $A_{11} = (2),\ A_{12} = (1),\ A_{21} = (1),\ A_{22} = (1),\ X_1 = [x_{11}, x_{12}, x_{13}],\ X_2 = [x_{21}, x_{22}, x_{23}]$,

$$Y_1 = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix},\ Y_2 = \begin{bmatrix} x_{13} \\ x_{23} \end{bmatrix},\ B_{11} = \begin{bmatrix} 3 & -1 \\ -1 & 1 \end{bmatrix},\ B_{12} = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

$B_{21} = [0, 1]$, $B_{22} = (2)$. Compute the densities of X_1, X_2, Y_1 and Y_2.

Solution 4.4.1. We need the following quantities: $A_{11} - A_{12}A_{22}^{-1}A_{21} = 2 - 1 = 1$, $A_{22} - A_{21}A_{11}^{-1}A_{12} = 1 - \frac{1}{2} = \frac{1}{2}$, $|B| = 1$,

$$B_{22} - B_{21}B_{11}^{-1}B_{12} = 2 - [0, 1]\Big(\frac{1}{2}\Big)\begin{bmatrix} 1 & 1 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = 2 - \frac{3}{2} = \frac{1}{2}$$
$$B_{11} - B_{12}B_{22}^{-1}B_{21} = \begin{bmatrix} 3 & -1 \\ -1 & 1 \end{bmatrix} - \begin{bmatrix} 0 \\ 1 \end{bmatrix}\Big(\frac{1}{2}\Big)[0, 1] = \begin{bmatrix} 3 & -1 \\ -1 & \frac{1}{2} \end{bmatrix}.$$

Let us compute the constant parts or normalizing constants in the various densities. With our usual notations, the normalizing constants in $f_{p_1,q}(X_1)$ and $f_{p_2,q}(X_2)$ are

$$\frac{|B|^{\frac{p_1}{2}}|A_{11}-A_{12}A_{22}^{-1}A_{21}|^{\frac{q}{2}}}{(2\pi)^{\frac{p_1q}{2}}}=\frac{|B|^{\frac{1}{2}}(1)^{\frac{3}{2}}}{(2\pi)^{\frac{3}{2}}}$$
$$\frac{|B|^{\frac{p_2}{2}}|A_{22}-A_{21}A_{11}^{-1}A_{12}|^{\frac{3}{2}}}{(2\pi)^{\frac{p_2q}{2}}}=\frac{|B|^{\frac{1}{2}}(\frac{1}{2})^{\frac{3}{2}}}{(2\pi)^{\frac{3}{2}}}.$$

Hence, the corresponding densities of X_1 and X_2 are the following:

$$f_{1,3}(X_1)=\frac{|B|^{\frac{1}{2}}}{(2\pi)^{\frac{3}{2}}}\mathrm{e}^{-\frac{1}{2}X_1BX_1'},\ -\infty<x_{1j}<\infty,\ j=1,2,3,$$
$$f_{1,3}(X_2)=\frac{|B|^{\frac{1}{2}}}{2^{\frac{3}{2}}(2\pi)^{\frac{3}{2}}}\mathrm{e}^{-\frac{1}{4}(X_2BX_2')},\ -\infty<x_{2j}<\infty,\ j=1,2,3.$$

Let us now evaluate the normalizing constants in the densities $f_{p,q_1}(Y_1)$, $f_{p,q_2}(Y_2)$:

$$\frac{|A|^{\frac{q_1}{2}}|B_{11}-B_{12}B_{22}^{-1}B_{21}|^{\frac{p}{2}}}{(2\pi)^{\frac{pq_1}{2}}}=\frac{|A|^{\frac{1}{2}}(\frac{1}{2})^1}{4\pi^2}=\frac{1}{8\pi^2},$$
$$\frac{|A|^{\frac{q_2}{2}}|B_{22}-B_{21}B_{11}^{-1}B_{12}|^{\frac{p}{2}}}{(2\pi)^{\frac{pq_2}{2}}}=\frac{|A|^{\frac{1}{2}}(\frac{1}{2})^1}{2\pi}=\frac{1}{4\pi}.$$

Thus, the density of Y_1 is

$$\begin{aligned}f_{2,2}(Y_1)&=\frac{1}{8\pi^2}\mathrm{e}^{-\frac{1}{2}\mathrm{tr}\{AY_1(B_{11}-B_{12}B_{22}^{-1}B_{21})Y_1'\}}\\&=\frac{1}{8\pi^2}\mathrm{e}^{-\frac{1}{2}Q},\ -\infty<x_{ij}<\infty,\ i,j=1,2,\end{aligned}$$

where

$$\begin{aligned}Q&=\mathrm{tr}\left\{\begin{bmatrix}2&1\\1&1\end{bmatrix}\begin{bmatrix}x_{11}&x_{12}\\x_{21}&x_{22}\end{bmatrix}\begin{bmatrix}3&-1\\-1&\frac{1}{2}\end{bmatrix}\begin{bmatrix}x_{11}&x_{21}\\x_{12}&x_{22}\end{bmatrix}\right\}\\&=6x_{11}^2+x_{12}^2+3x_{21}^2+\frac{1}{2}x_{22}^2-4x_{11}x_{12}-2x_{11}x_{22}\\&\quad+3x_{11}x_{21}+x_{12}x_{22}-2x_{12}x_{21}-x_{22}x_{21},\end{aligned}$$

the density of Y_2 being

$$f_{2,1}(Y_2) = \frac{1}{4\pi}e^{-\frac{1}{2}\mathrm{tr}\{AY_2(B_{22}-B_{21}B_{11}^{-1}B_{12})Y_1'\}}$$
$$= \frac{1}{4\pi}e^{-\frac{1}{4}[2x_{13}^2+x_{23}^2+2x_{13}x_{23}]},\ -\infty < x_{i3} < \infty,\ i = 1, 2.$$

This completes the computations.

4.4a. Marginal Densities in the Complex Matrix-variate Gaussian Case

The derivations of the results are parallel to those provided in the real case. Accordingly, we will state the corresponding results.

Theorem 4.4a.1. *Let the $p\times q$ matrix $\tilde{X}$ have a complex matrix-variate Gaussian density with the parameter matrices $M = O,\ A > O,\ B > O$ where A is $p \times p$ and B is $q \times q$. Consider a row partitioning of $\tilde{X}$ into sub-matrices $\tilde{X}_1$ and $\tilde{X}_2$ where $\tilde{X}_1$ is $p_1 \times q$ and $\tilde{X}_2$ is $p_2 \times q$, with $p_1 + p_2 = p$. Then, $\tilde{X}_1$ and $\tilde{X}_2$ have $p_1 \times q$ complex matrix-variate and $p_2 \times q$ complex matrix-variate Gaussian densities with parameter matrices $A_{11}-A_{12}A_{22}^{-1}A_{21}$ and B, and $A_{22}-A_{21}A_{11}^{-1}A_{12}$ and B, respectively, denoted by $\tilde{f}_{p_1,q}(\tilde{X}_1)$ and $\tilde{f}_{p_2,q}(\tilde{X}_2)$. The density of $\tilde{X}_1$ is given by*

$$\tilde{f}_{p_1,q}(\tilde{X}_1) = \frac{|\det(B)|^{p_1}|\det(A_{11} - A_{12}A_{22}^{-1}A_{21})|^q}{\pi^{p_1q}}e^{-\mathrm{tr}((A_{11}-A_{12}A_{22}^{-1}A_{21})\tilde{X}_1B\tilde{X}_1^*)}, \quad (4.4a.1)$$

the corresponding vector case for $p = 1$ being available from (4.4a.1) for $p_1 = 1,\ p_2 = 0$ and $p = 1$; in this case, the density is

$$\tilde{f}_{1,q}(\tilde{X}_1) = \frac{|\det(B)|}{\pi^q}e^{-(\tilde{X}_1-\mu)B(\tilde{X}_1-\mu)^*} \quad (4.4a.2)$$

where $\tilde{X}_1$ and μ are $1 \times q$ and μ is a location parameter vector. The density of $\tilde{X}_2$ is the following:

$$\tilde{f}_{p_2,q}(\tilde{X}_2) = \frac{|\det(B)|^{p_2}|\det(A_{22} - A_{21}A_{11}^{-1}A_{12})|^q}{\pi^{p_2q}}e^{-\mathrm{tr}((A_{22}-A_{21}A_{11}^{-1}A_{12})\tilde{X}_2B\tilde{X}_2^*)}. \quad (4.4a.3)$$

Theorem 4.4a.2. *Let $\tilde{X}$, A and B be as defined in Theorem 4.2a.1 and let $\tilde{X}$ be partitioned into column sub-matrices $\tilde{X} = (\tilde{Y}_1\ \ \tilde{Y}_2)$ where $\tilde{Y}_1$ is $p \times q_1$ and $\tilde{Y}_2$ is $p \times q_2$, so that $q_1 + q_2 = q$. Then $\tilde{Y}_1$ and $\tilde{Y}_2$ have $p \times q_1$ complex matrix-variate and $p \times q_2$ complex matrix-variate Gaussian densities given by*

$$\tilde{f}_{p,q_1}(\tilde{Y}_1) = \frac{|\det(A)|^{q_1}|\det(B_{11} - B_{12}B_{22}^{-1}B_{21})|^p}{\pi^{pq_1}} \times \mathrm{e}^{-\mathrm{tr}(A\tilde{Y}_1(B_{11}-B_{12}B_{22}^{-1}B_{21})\tilde{Y}_1^*)} \quad (4.4a.4)$$

$$\tilde{f}_{p,q_2}(\tilde{Y}_2) = \frac{|\det(A)|^{q_2}|\det(B_{22} - B_{21}B_{11}^{-1}B_{21})|^p}{\pi^{pq_2}} \times \mathrm{e}^{-\mathrm{tr}(A\tilde{Y}_2(B_{22}-B_{21}B_{11}^{-1}B_{12})\tilde{Y}_2^*)}. \quad (4.4a.5)$$

When $q = 1$, we have the usual complex multivariate case. In this case, it will be a p-variate complex Gaussian density. This is available from (4.4a.4) by taking $q_1 = 1,\ q_2 = 0$ and $q = 1$. Now, $\tilde{Y}_1$ is a $p \times 1$ column vector. Let μ be a $p \times 1$ location parameter vector. Then the density is

$$\tilde{f}_{p,1}(\tilde{Y}_1) = \frac{|\det(A)|}{\pi^p}\, \mathrm{e}^{-(\tilde{Y}_1-\mu)^* A(\tilde{Y}_1-\mu)} \quad (4.4a.6)$$

where $A > O$ (Hermitian positive definite), $\tilde{Y}_1 - \mu$ is $p \times 1$ and its $1 \times p$ conjugate transpose is $(\tilde{Y}_1 - \mu)^*$.

Example 4.4a.1. Consider a 2×3 complex matrix-variate Gaussian distribution with the parameters $M = O,\ A > O,\ B > O$ where

$$\tilde{X} = \begin{bmatrix} \tilde{x}_{11} & \tilde{x}_{12} & \tilde{x}_{13} \\ \tilde{x}_{21} & \tilde{x}_{22} & \tilde{x}_{23} \end{bmatrix},\ A = \begin{bmatrix} 2 & i \\ -i & 2 \end{bmatrix},\ B = \begin{bmatrix} 2 & -i & i \\ i & 2 & -i \\ -i & i & 2 \end{bmatrix}.$$

Consider the partitioning

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},\ B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix},\ \tilde{X} = \begin{bmatrix} \tilde{X}_1 \\ \tilde{X}_2 \end{bmatrix} = [\tilde{Y}_1, \tilde{Y}_2]$$

where

$$\tilde{Y}_1 = \begin{bmatrix} \tilde{x}_{11} \\ \tilde{x}_{21} \end{bmatrix},\ \tilde{Y}_2 = \begin{bmatrix} \tilde{x}_{12} & \tilde{x}_{13} \\ \tilde{x}_{22} & \tilde{x}_{23} \end{bmatrix},\ B_{22} = \begin{bmatrix} 2 & -i \\ i & 2 \end{bmatrix},\ B_{21} = \begin{bmatrix} i \\ -i \end{bmatrix},$$

$\tilde{X}_1 = [\tilde{x}_{11}, \tilde{x}_{12}, \tilde{x}_{13}]$, $\tilde{X}_2 = [\tilde{x}_{21}, \tilde{x}_{22}, \tilde{x}_{23}]$, $A_{11} = (2)$, $A_{12} = (i)$, $A_{21} = (-i)$, $A_{22} = 2$; $B_{11} = (2)$, $B_{12} = [-i, i]$. Compute the densities of $\tilde{X}_1, \tilde{X}_2, \tilde{Y}_1, \tilde{Y}_2$.

Solution 4.4a.1. It is easy to ascertain that $A = A^*$ and $B = B^*$; hence both matrices are Hermitian. As well, all the leading minors of A and B are positive so that $A > O$ and $B > O$. We need the following numerical results: $|A| = 3$, $|B| = 2$,

$$A_{11} - A_{12}A_{22}^{-1}A_{21} = 2 - (i)(1/2)(-i) = 2 - \frac{1}{2} = \frac{3}{2}$$
$$A_{22} - A_{21}A_{11}^{-1}A_{12} = 2 - (-i)(1/2)(i) = 2 - \frac{1}{2} = \frac{3}{2}$$

$$B_{11}-B_{12}B_{22}^{-1}B_{21}=2-[-i,i]\Big(\frac{1}{3}\Big)\begin{bmatrix}2 & i\\ -i & 2\end{bmatrix}\begin{bmatrix}i\\ -i\end{bmatrix}=\frac{2}{3}$$

$$B_{22}-B_{21}B_{11}^{-1}B_{12}=\begin{bmatrix}2 & -i\\ i & 2\end{bmatrix}-\begin{bmatrix}i\\ -i\end{bmatrix}\Big(\frac{1}{2}\Big)[-i,i]=\begin{bmatrix}2 & -i\\ i & 2\end{bmatrix}-\frac{1}{2}\begin{bmatrix}1 & -1\\ -1 & 1\end{bmatrix}$$
$$=\frac{1}{2}\begin{bmatrix}3 & 1-2i\\ 1+2i & 3\end{bmatrix}.$$

With these preliminary calculations, we can obtain the required densities with our usual notations:

$$\tilde{f}_{p_1,q}(\tilde{X}_1)=\frac{|\det(B)|^{p_1}|\det(A_{11}-A_{12}A_{22}^{-1}A_{21})|^{q}}{\pi^{p_1q}}$$
$$\times\, \mathrm{e}^{-\mathrm{tr}[(A_{11}-A_{12}A_{22}^{-1}A_{21})\tilde{X}_1B\tilde{X}_1^*]},\ \text{that is,}$$
$$\tilde{f}_{1,3}(\tilde{X}_1)=\frac{2(3/2)^3}{\pi^3}\mathrm{e}^{-\frac{3}{2}\tilde{X}_1B\tilde{X}_1^*}$$

where

$$Q_1=\tilde{X}_1B\tilde{X}_1^*=[\tilde{x}_{11},\tilde{x}_{12},\tilde{x}_{13}]\begin{bmatrix}2 & -i & i\\ i & 2 & -i\\ -i & i & 2\end{bmatrix}\begin{bmatrix}\tilde{x}_{11}^*\\ \tilde{x}_{12}^*\\ \tilde{x}_{13}^*\end{bmatrix}$$
$$=2\tilde{x}_{11}\tilde{x}_{11}^*+2\tilde{x}_{12}\tilde{x}_{12}^*+2\tilde{x}_{13}\tilde{x}_{13}^*-i\tilde{x}_{11}\tilde{x}_{12}^*+i\tilde{x}_{11}\tilde{x}_{13}^*$$
$$+i\tilde{x}_{12}\tilde{x}_{11}^*-i\tilde{x}_{12}\tilde{x}_{13}^*-i\tilde{x}_{13}\tilde{x}_{11}^*+i\tilde{x}_{13}\tilde{x}_{12}^*;$$

$$\tilde{f}_{p_2,q}(\tilde{X}_2)=\frac{|\det(A)|^{p_2}|\det(A_{22}-A_{21}A_{11}^{-1}A_{12})|^{q}}{\pi^{p_2q}}$$
$$\times\, \mathrm{e}^{-\mathrm{tr}[(A_{22}-A_{21}A_{11}^{-1}A_{12})\tilde{X}_2B\tilde{X}_2^*]},\ \text{that is,}$$
$$\tilde{f}_{1,3}(\tilde{X}_2)=\frac{2(3/2)^3}{\pi^3}\mathrm{e}^{-\frac{3}{2}\tilde{X}_2B\tilde{X}_2^*}$$

where let $Q_2=\tilde{X}_2B\tilde{X}_2^*$, Q_2 being obtained by replacing $\tilde{X}_1$ in Q_1 by $\tilde{X}_2$;

$$\tilde{f}_{p,q_1}(\tilde{Y}_1)=\frac{|\det(A)|^{q_1}|\det(B_{11}-B_{12}B_{22}^{-1}B_{21})|^{p}}{\pi^{pq_1}}$$
$$\times\, \mathrm{e}^{-\mathrm{tr}[A\tilde{Y}_1(B_{11}-B_{12}B_{22}^{-1}B_{21})\tilde{Y}_1^*]},\ \text{that is,}$$
$$\tilde{f}_{2,1}(\tilde{Y}_1)=\frac{3(2/3)^2}{\pi^2}\mathrm{e}^{-\mathrm{tr}(\frac{2}{3}Q_3)}$$

where

$$\begin{aligned}Q_3=\tilde{Y}_1^*A\tilde{Y}_1&=[\tilde{x}_{11}^*,\tilde{x}_{21}^*]\begin{bmatrix}2&i\\-i&2\end{bmatrix}\begin{bmatrix}\tilde{x}_{11}\\\tilde{x}_{21}\end{bmatrix}\\&=2\tilde{x}_{11}\tilde{x}_{11}^*+2\tilde{x}_{21}\tilde{x}_{21}^*+i\tilde{x}_{21}\tilde{x}_{11}^*-i\tilde{x}_{21}^*\tilde{x}_{11};\end{aligned}$$

$$\begin{aligned}\tilde{f}_{p,q_2}(\tilde{Y}_2)&=\frac{|\det(A)|^{q_2}|\det(B_{22}-B_{21}B_{11}^{-1}B_{12})|^p}{\pi^{pq_2}}\\&\quad\times \mathrm{e}^{-\mathrm{tr}[A\tilde{Y}_2(B_{22}-B_{21}B_{11}^{-1}B_{12})\tilde{Y}_2^*]},\ \text{that is,}\\\tilde{f}_{2,2}(\tilde{Y}_2)&=\frac{3^2}{\pi^4}\mathrm{e}^{-\frac{1}{2}Q}\end{aligned}$$

where

$$\begin{aligned}2Q&=\mathrm{tr}\begin{bmatrix}2&i\\-i&2\end{bmatrix}\begin{bmatrix}\tilde{x}_{12}&\tilde{x}_{13}\\\tilde{x}_{22}&\tilde{x}_{23}\end{bmatrix}\begin{bmatrix}3&1-2i\\1+2i&3\end{bmatrix}\begin{bmatrix}\tilde{x}_{12}^*&\tilde{x}_{22}^*\\\tilde{x}_{13}^*&\tilde{x}_{23}^*\end{bmatrix}\\&=6\tilde{x}_{12}\tilde{x}_{12}^*+6\tilde{x}_{13}\tilde{x}_{13}^*+6\tilde{x}_{22}\tilde{x}_{22}^*+6\tilde{x}_{23}\tilde{x}_{23}^*\\&\quad+[2(1-2i)(\tilde{x}_{12}\tilde{x}_{13}^*+\tilde{x}_{22}\tilde{x}_{23}^*)+2(1+2i)(\tilde{x}_{23}\tilde{x}_{22}^*+\tilde{x}_{13}\tilde{x}_{12}^*)]\\&\quad+[i(1-2i)(\tilde{x}_{22}\tilde{x}_{13}^*-\tilde{x}_{12}\tilde{x}_{23}^*)-i(1+2i)(\tilde{x}_{13}\tilde{x}_{22}^*-\tilde{x}_{23}\tilde{x}_{12}^*)]\\&\quad+3i[\tilde{x}_{22}\tilde{x}_{12}^*+\tilde{x}_{23}\tilde{x}_{13}^*-\tilde{x}_{12}\tilde{x}_{22}^*-\tilde{x}_{13}\tilde{x}_{23}^*].\end{aligned}$$

This completes the computations.

Exercises 4.4

4.4.1. Write down explicitly the density of a $p\times q$ matrix-variate Gaussian for $p=3, q=3$. Then by integrating out the other variables, obtain the density for the case (1): $p=2, q=2$; (2): $p=2, q=1$; (3): $p=1, q=2$; (4): $p=1, q=1$. Take the location matrix $M=O$. Let A and B to be general positive definite parameter matrices.

4.4.2. Repeat Exercise 4.4.1 for the complex case.

4.4.3. Write down the densities obtained in Exercises 4.4.1 and 4.4.2. Then evaluate the marginal densities for $p=2, q=2$ in both the real and complex domains by partitioning matrices and integrating out by using matrix methods.

4.4.4. Let the 2×2 real matrix $A>O$ where the first row is $(1, 1)$. Let the real $B>O$ be 3×3 where the first row is $(1, -1, 1)$. Complete A and B with numbers of your choosing so that $A > O$, $B > O$. Consider a real 2×3 matrix-variate Gaussian density with these A and B as the parameter matrices. Take your own non-null location matrix. Write down the matrix-variate Gaussian density explicitly. Then by integrating out the other variables, either directly or by matrix methods, obtain (1): the 1×3 matrix-variate Gaussian density; (2): the 2×2 matrix-variate Gaussian density from your 2×3 matrix-variate Gaussian density.

4.4.5. Repeat Exercise 4.4.4 for the complex case if the first row of A is $(1, 1+i)$ and the first row of B is $(2, 1+i, 1-i)$ where $A = A^* > O$ and $B = B^* > O$.

4.5. Conditional Densities in the Real Matrix-variate Gaussian Case

Consider a real $p \times q$ matrix-variate Gaussian density with the parameters $M = O$, $A > O$, $B > O$. Let us consider the partition of the $p \times q$ real Gaussian matrix X into row sub-matrices as $X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$ where X_1 is $p_1 \times q$ and X_2 is $p_2 \times q$ with $p_1 + p_2 = p$. We have already established that the marginal density of X_2 is

$$f_{p_2,q}(X_2) = \frac{|A_{22} - A_{21}A_{11}^{-1}A_{12}|^{\frac{q}{2}}|B|^{\frac{p_2}{2}}}{(2\pi)^{\frac{p_2 q}{2}}} \,\mathrm{e}^{-\frac{1}{2}\mathrm{tr}[(A_{22}-A_{21}A_{11}^{-1}A_{12})X_2BX_2']}.$$

Thus, the conditional density of X_1 given X_2 is obtained as

$$f_{p_1,q}(X_1|X_2) = \frac{f_{p,q}(X)}{f_{p_2,q}(X_2)} = \frac{|A|^{\frac{q}{2}}|B|^{\frac{p}{2}}}{|A_{22} - A_{21}A_{11}^{-1}A_{12}|^{\frac{q}{2}}|B|^{\frac{p_2}{2}}} \frac{(2\pi)^{\frac{p_2 q}{2}}}{(2\pi)^{\frac{pq}{2}}}$$
$$\times\, \mathrm{e}^{-\frac{1}{2}[\mathrm{tr}(AXBX')]+\mathrm{tr}[(A_{22}-A_{21}A_{11}^{-1}A_{12})X_2BX_2']}.$$

Note that

$$AXBX' = A\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} B(X_1' \ \ X_2') = A\begin{bmatrix} X_1BX_1' & X_1BX_2' \\ X_2BX_1' & X_2BX_2' \end{bmatrix}$$
$$= \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}\begin{bmatrix} X_1BX_1' & X_1BX_2' \\ X_2BX_1' & X_2BX_2' \end{bmatrix} = \begin{bmatrix} \alpha & * \\ * & \beta \end{bmatrix}$$

where $\alpha = A_{11}X_1BX_1' + A_{12}X_2BX_1'$, $\beta = A_{21}X_1BX_2' + A_{22}X_2BX_2'$ and the asterisks designate elements that are not utilized in the determination of the trace. Then

$$\mathrm{tr}(AXBX') = \mathrm{tr}(A_{11}X_1BX_1' + A_{12}X_2BX_1') + \mathrm{tr}(A_{21}X_1BX_2' + A_{22}X_2BX_2').$$

Thus the exponent in the conditional density simplifies to

$$\begin{aligned}
\mathrm{tr}(A_{11}X_1BX_1') &+ 2\mathrm{tr}(A_{12}X_2BX_1') + \mathrm{tr}(A_{22}X_2BX_2') - \mathrm{tr}[(A_{22} - A_{21}A_{11}^{-1}A_{12})X_2BX_2')] \\
&= \mathrm{tr}(A_{11}X_1BX_1') + 2\mathrm{tr}(A_{12}X_2BX_1') + \mathrm{tr}[A_{21}A_{11}^{-1}A_{12}X_2BX_2'] \\
&= \mathrm{tr}[A_{11}(X_1 + C)B(X_1 + C)'],\ C = A_{11}^{-1}A_{12}X_2.
\end{aligned}$$

We note that $E(X_1|X_2) = -C = -A_{11}^{-1}A_{12}X_2$: the regression of X_1 on X_2, the constant part being $|A_{11}|^{\frac{q}{2}}|B|^{\frac{p_1}{2}}/(2\pi)^{\frac{p_1q}{2}}$. Hence the following result:

Theorem 4.5.1. *If the $p \times q$ matrix X has a real matrix-variate Gaussian density with the parameter matrices $M = O,\ A > O$ and $B > O$ where A is $p \times p$ and B is $q \times q$ and if X is partitioned into row sub-matrices $X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$ where X_1 is $p_1 \times q$ and X_2 is $p_2 \times q$, so that $p_1 + p_2 = p$, then the conditional density of X_1 given X_2, denoted by $f_{p_1,q}(X_1|X_2)$, is given by*

$$f_{p_1,q}(X_1|X_2) = \frac{|A_{11}|^{\frac{q}{2}}|B|^{\frac{p_1}{2}}}{(2\pi)^{\frac{p_1q}{2}}}\mathrm{e}^{-\frac{1}{2}\mathrm{tr}[A_{11}(X_1+C)B(X_1+C)']} \tag{4.5.1}$$

where $C = A_{11}^{-1}A_{12}X_2$ if the location parameter is a null matrix; otherwise $C = -M_1 + A_{11}^{-1}A_{12}(X_2 - M_2)$ with M partitioned into row sub-matrices M_1 and M_2, M_1 being $p_1 \times q$ and M_2, $p_2 \times q$.

Corollary 4.5.1. *Let $X,\ X_1,\ X_2,\ M,\ M_1$ and M_2 be as defined in Theorem 4.5.1; then, in the real Gaussian case, the conditional expectation of X_1 given X_2, denoted by $E(X_1|X_2)$, is*

$$E(X_1|X_2) = M_1 - A_{11}^{-1}A_{12}(X_2 - M_2). \tag{4.5.2}$$

We may adopt the following general notation to represent a real matrix-variate Gaussian (or normal) density:

$$X \sim N_{p,q}(M, A, B),\ A > O,\ B > O, \tag{4.5.3}$$

which signifies that the $p \times q$ matrix X has a real matrix-variate Gaussian distribution with location parameter matrix M and parameter matrices $A > O$ and $B > O$ where A is $p \times p$ and B is $q \times q$. Accordingly, the usual q-variate multivariate normal density will be denoted as follows:

$$X_1 \sim N_{1,q}(\mu, B),\ B > O \Rightarrow X_1' \sim N_q(\mu', B^{-1}),\ B > O, \tag{4.5.4}$$

where μ is the location parameter vector, which is the first row of M, and X_1 is a $1\times q$ row vector consisting of the first row of the matrix X. Note that $B^{-1} = \mathrm{Cov}(X_1)$ and the covariance matrix usually appears as the second parameter in the standard notation $N_p(\cdot,\cdot)$. In this case, the 1×1 matrix A will be taken as 1 to be consistent with the usual notation in the real multivariate normal case. The corresponding column case will be denoted as follows:

$$Y_1 \sim N_{p,1}(\mu_{(1)}, A),\ A > O \Rightarrow Y_1 \sim N_p(\mu_{(1)}, A^{-1}),\ A > O,\ A^{-1} = \mathrm{Cov}(Y_1) \quad (4.5.5)$$

where Y_1 is a $p\times 1$ vector consisting of the first column of X and $\mu_{(1)}$ is the first column of M. With this partitioning of X, we have the following result:

Theorem 4.5.2. *Let the real matrices $X,\ M,\ A$ and B be as defined in Theorem 4.5.1 and X be partitioned into column sub-matrices as $X = (Y_1\ \ Y_2)$ where Y_1 is $p\times q_1$ and Y_2 is $p\times q_2$ with $q_1+q_2 = q$. Let the density of X, the marginal densities of Y_1 and Y_2 and the conditional density of Y_1 given Y_2, be respectively denoted by $f_{p,q}(X),\ f_{p,q_1}(Y_1),\ f_{p,q_2}(Y_2)$ and $f_{p,q_1}(Y_1|Y_2)$. Then, the conditional density of Y_1 given Y_2 is*

$$f_{p,q_1}(Y_1|Y_2) = \frac{|A|^{\frac{q_1}{2}}|B_{11}|^{\frac{p}{2}}}{(2\pi)^{\frac{pq_1}{2}}}\,\mathrm{e}^{-\frac{1}{2}\mathrm{tr}[A(Y_1-M_{(1)}+C_1)B_{11}(Y_1-M_{(1)}+C_1)']} \quad (4.5.6)$$

where $A > O,\ B_{11} > O$ and $C_1 = (Y_2 - M_2)B_{21}B_{11}^{-1}$, so that the conditional expectation of Y_1 given Y_2, or the regression of Y_1 on Y_2, is obtained as

$$E(Y_1|Y_2) = M_{(1)} - (Y_2 - M_{(2)})B_{21}B_{11}^{-1},\ M = (M_{(1)}\ \ M_{(2)}), \quad (4.5.7)$$

where $M_{(1)}$ is $p\times q_1$ and $M_{(2)}$ is $p\times q_2$ with $q_1+q_2 = q$. As well, the conditional density of Y_2 given Y_1 is the following:

$$f_{p,q_2}(Y_2|Y_1) = \frac{|A|^{\frac{q_2}{2}}|B_{22}|^{\frac{p}{2}}}{(2\pi)^{\frac{pq_2}{2}}}\,\mathrm{e}^{-\frac{1}{2}\mathrm{tr}[A(Y_2-M_{(2)}+C_2)B_{22}(Y_2-M_{(2)}+C_2)']} \quad (4.5.8)$$

where

$$M_{(2)} - C_2 = M_{(2)} - (Y_1 - M_{(1)})B_{12}B_{22}^{-1} = E[Y_2|Y_1]. \quad (4.5.9)$$

Example 4.5.1. Consider a 2×3 real matrix $X = (x_{ij})$ having a real matrix-variate Gaussian distribution with the parameters M, $A > O$ and $B > O$ where

$$M = \begin{bmatrix} 1 & -1 & 1\\ 2 & 0 & -1\end{bmatrix},\ A = \begin{bmatrix} 2 & 1\\ 1 & 3\end{bmatrix},\ B = \begin{bmatrix} 2 & -1 & 1\\ -1 & 3 & 0\\ 1 & 0 & 1\end{bmatrix}.$$

Let X be partitioned as $\begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$=$[Y_1, Y_2]$ where $X_1 = [x_{11}, x_{12}, x_{13}]$, $X_2 = [x_{21}, x_{22}, x_{23}]$, $Y_1 = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix}$ and $Y_2 = \begin{bmatrix} x_{13} \\ x_{23} \end{bmatrix}$. Determine the conditional densities of X_1 given X_2, X_2 given X_1, Y_1 given Y_2, Y_2 given Y_1 and the conditional expectations $E[X_1|X_2]$, $E[X_2|X_1]$, $E[Y_1|Y_2]$ and $E[Y_2|Y_1]$.

Solution 4.5.1. Given the specified partitions of X, A and B are partitioned accordingly as follows:

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},\ B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix},\ B_{11} = \begin{bmatrix} 2 & -1 \\ -1 & 3 \end{bmatrix},\ B_{12} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$B_{21} = [1, 0],\ B_{22} = (1),\ A_{11} = (2),\ A_{12} = (1),\ A_{21} = (1),\ A_{22} = (3).$$

The following numerical results are needed:

$$A_{11} - A_{12}A_{22}^{-1}A_{21} = 2 - (1)(1/3)(1) = \frac{5}{3}$$
$$A_{22} - A_{21}A_{11}^{-1}A_{12} = 3 - (1)(1/2)(1) = \frac{5}{2}$$
$$B_{11} - B_{12}B_{22}^{-1}B_{21} = \begin{bmatrix} 2 & -1 \\ -1 & 3 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \end{bmatrix}(1)[1, 0] = \begin{bmatrix} 1 & -1 \\ -1 & 3 \end{bmatrix}$$
$$B_{22} - B_{21}B_{11}^{-1}B_{12} = 1 - [1, 0](1/5)\begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{2}{5};$$

$$|A_{11}| = 2,\ |A_{22}| = 3,\ |A| = 5,\ |A_{11} - A_{12}A_{22}^{-1}A_{21}| = \frac{5}{3},\ |A_{22} - A_{21}A_{11}^{-1}A_{12}| = \frac{5}{2},$$
$$|B_{11}| = 5,\ |B_{22}| = 1,\ |B_{11} - B_{12}B_{22}^{-1}B_{21}| = 2,\ |B_{22} - B_{21}B_{11}^{-1}B_{12}| = \frac{2}{5},\ |B| = 2;$$

$$A_{11}^{-1}A_{12} = \frac{1}{2},\ A_{22}^{-1}A_{21} = \frac{1}{3},\ B_{22}^{-1}B_{21} = [1, 0]$$
$$B_{11}^{-1}B_{12} = \frac{1}{5}\begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{1}{5}\begin{bmatrix} 3 \\ 1 \end{bmatrix},\ M_1 = [1, -1, 1],\ M_2 = [2, 0, -1]$$
$$M_{(1)} = \begin{bmatrix} 1 & -1 \\ 2 & 0 \end{bmatrix},\ M_{(2)} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

All the conditional expectations can now be determined. They are

$$E[X_1|X_2] = M_1 - A_{11}^{-1}A_{12}(X_2 - M_2) = [1, -1, 1] - \frac{1}{2}(x_{21} - 2, x_{22}, x_{23} + 1)$$
$$= [1 - \frac{1}{2}(x_{21} - 2), -1 - \frac{1}{2}x_{22}, 1 - \frac{1}{2}(x_{23} + 1)] \qquad (i)$$
$$E[X_2|X_1] = M_2 - A_{22}^{-1}A_{21}(X_1 - M_1) = [2, 0, -1] - \frac{1}{3}[x_{11} - 1, x_{12} + 1, x_{13} - 1]$$
$$= [2 - \frac{1}{3}(x_{11} - 1), -\frac{1}{3}(x_{12} + 1), -1 - \frac{1}{3}(x_{13} - 1)]; \qquad (ii)$$

$$E[Y_1|Y_2] = M_{(1)} - (Y_2 - M_{(2)})B_{21}B_{11}^{-1} = \begin{bmatrix} 1 & -1 \\ 2 & 0 \end{bmatrix} - \frac{1}{5}\begin{bmatrix} x_{13} - 1 \\ x_{23} + 1 \end{bmatrix}[1, 0]\begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix}$$
$$= \begin{bmatrix} 1 - \frac{3}{5}(x_{13} - 1) & -1 - \frac{1}{5}(x_{13} - 1) \\ 2 - \frac{3}{5}(x_{23} + 1) & -\frac{1}{5}(x_{23} + 1) \end{bmatrix} \qquad (iii)$$
$$E[Y_2|Y_1] = M_{(2)} - (Y_1 - M_{(1)})B_{12}B_{22}^{-1}$$
$$= \begin{bmatrix} 1 \\ -1 \end{bmatrix} - \begin{bmatrix} x_{11} - 1 & x_{12} + 1 \\ x_{21} - 2 & x_{22} \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix}[(1)] = \begin{bmatrix} 2 - x_{11} \\ 1 - x_{21} \end{bmatrix}. \qquad (iv)$$

The conditional densities can now be obtained. That of X_1 given X_2 is

$$f_{p_1,q}(X_1|X_2) = \frac{|A_{11}|^{\frac{q}{2}}|B|^{\frac{p_1}{2}}}{(2\pi)^{\frac{p_1 q}{2}}}\,\mathrm{e}^{-\frac{1}{2}\mathrm{tr}[A_{11}(X_1 - M_1 + C)B(X_1 - M_1 + C)']}$$

for the matrices $A > O$ and $B > O$ previously specified; that is,

$$f_{1,3}(X_1|X_2) = \frac{4}{(2\pi)^{\frac{3}{2}}}\,\mathrm{e}^{-\frac{2}{2}(X_1 - M_1 + C)B(X_1 - M_1 + C)'}$$

where $M_1 - C = E[X_1|X_2]$ is given in *(i)*. The conditional density of $X_2|X_1$ is the following:

$$f_{p_2,q}(X_2|X_1) = \frac{|A_{22}|^{\frac{q}{2}}|B|^{\frac{p_2}{2}}}{(2\pi)^{\frac{p_2 q}{2}}}\,\mathrm{e}^{-\frac{1}{2}\mathrm{tr}[A_{22}(X_2 - M_2 + C_1)B(X_2 - M_2 + C_1)']},$$

that is,

$$f_{1,3}(X_2|X_1) = \frac{(3^{\frac{3}{2}})(2^{\frac{1}{2}})}{(2\pi)^{\frac{3}{2}}}\,\mathrm{e}^{-\frac{3}{2}(X_2 - M_2 + C_1)B(X_2 - M_2 + C_2)'}$$

where $M_2 - C_2 = E[X_2|X_1]$ is given in *(ii)*. The conditional density of Y_1 given Y_2 is

$$f_{p,q_1}(Y_1|Y_2) = \frac{|A|^{\frac{q_1}{2}}|B_{11}|^{\frac{p}{2}}}{(2\pi)^{\frac{pq_1}{2}}} \mathrm{e}^{-\frac{1}{2}\mathrm{tr}[A(Y_1-M_{(1)}+C_2)B_{11}(Y_1-M_{(1)}+C_2)']};$$

that is,

$$f_{2,2}(Y_1|Y_2) = \frac{25}{(2\pi)^2} \mathrm{e}^{-\frac{1}{2}\mathrm{tr}[A(Y_1-M_{(1)}+C_2)B_{11}(Y_1-M_{(1)}+C_2)']}$$

where $M_{(1)} - C_1 = E[Y_1|Y_2]$ is specified in *(iii)*. Finally, the conditional density of $Y_2|Y_1$ is the following:

$$f_{p,q_2}(Y_2|Y_1) = \frac{|A|^{\frac{q_2}{2}}|B_{22}|^{\frac{p}{2}}}{(2\pi)^{\frac{pq_2}{2}}} \mathrm{e}^{-\frac{1}{2}\mathrm{tr}[A(Y_2-M_{(2)}+C_3)B_{22}(Y_2-M_{(2)}+C_3)']};$$

that is,

$$f_{2,1}(Y_2|Y_1) = \frac{\sqrt{5}}{(2\pi)} \mathrm{e}^{-\mathrm{tr}[A(Y_2-M_{(2)}+C_3)B_{22}(Y_2-M_{(2)}+C_3)']}$$

where $M_{(2)} - C_3 = E[Y_2|Y_1]$ is given in *(iv)*. This completes the computations.

4.5a. Conditional Densities in the Matrix-variate Complex Gaussian Case

The corresponding distributions in the complex case closely parallel those obtained for the real case. A tilde will be utilized to distinguish them from the real distributions. Thus,

$$\tilde{X} \sim \tilde{N}_{p,q}(\tilde{M}, A, B),\ A = A^* > O,\ B = B^* > O$$

will denote a complex $p \times q$ matrix $\tilde{X}$ having a $p \times q$ matrix-variate complex Gaussian density. For the $1 \times q$ case, that is, the q-variate multivariate normal distribution in the complex case, which is obtained from the marginal distribution of the first row of $\tilde{X}$, we have

$$\tilde{X}_1 \sim \tilde{N}_{1,q}(\mu, B),\ B > O,\ \tilde{X}_1 \sim \tilde{N}_q(\mu, B^{-1}),\ B^{-1} = \mathrm{Cov}(\tilde{X}_1),$$

where $\tilde{X}_1$ is $1 \times q$ vector having a q-variate complex normal density with $E(\tilde{X}_1) = \mu$. The case $q = 1$ corresponds to a column vector in $\tilde{X}$, which constitutes a $p \times 1$ column vector in the complex domain. Letting it be denoted as $\tilde{Y}_1$, we have

$$\tilde{Y}_1 \sim \tilde{N}_{p,1}(\mu_{(1)}, A),\ A > O,\ \ \text{that is,}\ \tilde{Y}_1 \sim \tilde{N}_p(\mu_{(1)}, A^{-1}),\ A^{-1} = \mathrm{Cov}(\tilde{Y}_1),$$

where $\mu_{(1)}$ is the first column of M.

Theorem 4.5a.1. *Let $\tilde{X}$ be $p \times q$ matrix in the complex domain having a $p \times q$ matrix-variate complex Gaussian density denoted by $\tilde{f}_{p,q}(\tilde{X})$. Let $\tilde{X} = \begin{pmatrix} \tilde{X}_1 \\ \tilde{X}_2 \end{pmatrix}$ be a row partitioning of $\tilde{X}$ into sub-matrices where $\tilde{X}_1$ is $p_1 \times q$ and $\tilde{X}_2$ is $p_2 \times q$, with $p_1 + p_2 = p$. Then the conditional density of $\tilde{X}_1$ given $\tilde{X}_2$ denoted by $\tilde{f}_{p_1,q}(\tilde{X}_1|\tilde{X}_2)$, is given by*

$$\tilde{f}_{p_1,q}(\tilde{X}_1|\tilde{X}_2) = \frac{|\det(A_{11})|^q|\det(B)|^{p_1}}{\pi^{p_1 q}}\,\mathrm{e}^{-\mathrm{tr}[A_{11}(\tilde{X}_1-M_1+\tilde{C})B(\tilde{X}_1-M_1+\tilde{C})^*]} \tag{4.5a.1}$$

where $\tilde{C} = A_{11}^{-1}A_{12}(\tilde{X}_2 - M_2)$, $E[\tilde{X}] = M = \begin{bmatrix} M_1 \\ M_2 \end{bmatrix}$, and the regression of $\tilde{X}_1$ on $\tilde{X}_2$ is as follows:

$$E(\tilde{X}_1|\tilde{X}_2) = \begin{cases} M_1 - A_{11}^{-1}A_{12}(\tilde{X}_2 - M_2) \text{ if } M = \begin{pmatrix} M_1 \\ M_2 \end{pmatrix} \\ -A_{11}^{-1}A_{12}\tilde{X}_2 \text{ if } M = O. \end{cases} \tag{4.5a.2}$$

Analogously, the conditional density of $\tilde{X}_2$ given $\tilde{X}_1$ is

$$\tilde{f}_{p_2,q}(\tilde{X}_2|\tilde{X}_1) = \frac{|\det(A_{22})|^q|\det(B)|^{p_2}}{\pi^{p_2 q}}\,\mathrm{e}^{-\mathrm{tr}[A_{22}(\tilde{X}_2-M_2+C_1)B(\tilde{X}_2-M_2+C_1)^*]} \tag{4.5a.3}$$

where $C_1 = A_{22}^{-1}A_{21}(\tilde{X}_1 - M_1)$, so that the conditional expectation of $\tilde{X}_2$ given $\tilde{X}_1$ or the regression of $\tilde{X}_2$ on $\tilde{X}_1$ is given by

$$E[\tilde{X}_2|\tilde{X}_1] = M_2 - A_{22}^{-1}A_{21}(\tilde{X}_1 - M_1). \tag{4.5a.4}$$

Theorem 4.5a.2. *Let $\tilde{X}$ be as defined in Theorem 4.5a.1. Let $\tilde{X}$ be partitioned into column submatrices, that is, $\tilde{X} = (\tilde{Y}_1\ \ \tilde{Y}_2)$ where $\tilde{Y}_1$ is $p \times q_1$ and $\tilde{Y}_2$ is $p \times q_2$, with $q_1 + q_2 = q$. Then the conditional density of $\tilde{Y}_1$ given $\tilde{Y}_2$, denoted by $\tilde{f}_{p,q_1}(\tilde{Y}_1|\tilde{Y}_2)$ is given by*

$$\tilde{f}_{p,q_1}(\tilde{Y}_1|\tilde{Y}_2) = \frac{|\det(A)|^{q_1}|\det(B_{11})|^p}{\pi^{pq_1}}\,\mathrm{e}^{-\mathrm{tr}[A(\tilde{Y}_1-M_{(1)}+\tilde{C}_{(1)})B_{11}(\tilde{Y}_1-M_{(1)}+\tilde{C}_{(1)})]} \tag{4.5a.5}$$

where $\tilde{C}_{(1)} = (\tilde{Y}_2 - M_{(2)})B_{21}B_{11}^{-1}$, and the regression of $\tilde{Y}_1$ on $\tilde{Y}_2$ or the conditional expectation of $\tilde{Y}_1$ given $\tilde{Y}_2$ is given by

$$E(\tilde{Y}_1|\tilde{Y}_2) = M_{(1)} - (\tilde{Y}_2 - M_{(2)})B_{21}B_{11}^{-1} \tag{4.5a.6}$$

with $E[\tilde{X}] = M = [M_{(1)}\ M_{(2)}] = E[\tilde{Y}_1\ \tilde{Y}_2]$. *As well the conditional density of* $\tilde{Y}_2$ *given* $\tilde{Y}_1$ *is the following:*

$$\tilde{f}_{p,q_2}(\tilde{Y}_2|\tilde{Y}_1) = \frac{|\det(A)|^{q_2}|\det(B_{22})|^p}{\pi^{pq_2}} \,\mathrm{e}^{-\mathrm{tr}[A(\tilde{Y}_2 - M_{(2)} + C_{(2)})B_{22}(\tilde{Y}_2 - M_{(2)} + C_{(2)})^*]} \tag{4.5a.7}$$

where $C_{(2)} = (\tilde{Y}_1 - M_{(1)})B_{12}B_{22}^{-1}$ *and the conditional expectation of* $\tilde{Y}_2$ *given* $\tilde{Y}_1$ *is then*

$$E[\tilde{Y}_2|\tilde{Y}_1] = M_{(2)} - (\tilde{Y}_1 - M_{(1)})B_{12}B_{22}^{-1}. \tag{4.5a.8}$$

Example 4.5a.1. Consider a 2×3 matrix-variate complex Gaussian distribution with the parameters

$$A = \begin{bmatrix} 2 & i \\ -i & 1 \end{bmatrix},\ B = \begin{bmatrix} 3 & -i & 0 \\ i & 2 & i \\ 0 & -i & 1 \end{bmatrix},\ M = E[\tilde{X}] = \begin{bmatrix} 1+i & i & -i \\ i & 2+i & 1-i \end{bmatrix}.$$

Consider the partitioning of $\tilde{X} = \begin{bmatrix} \tilde{X}_1 \\ \tilde{X}_2 \end{bmatrix} = [\tilde{Y}_1\ \tilde{Y}_2]$ where $\tilde{X}_1 = [\tilde{x}_{11}, \tilde{x}_{12}, \tilde{x}_{13}]$, $\tilde{X}_2 = [\tilde{x}_{21}, \tilde{x}_{22}, \tilde{x}_{23}]$, $\tilde{Y}_1 = \begin{bmatrix} \tilde{x}_{11} \\ \tilde{x}_{21} \end{bmatrix}$ and $\tilde{Y}_2 = \begin{bmatrix} \tilde{x}_{12} & \tilde{x}_{13} \\ \tilde{x}_{22} & \tilde{x}_{23} \end{bmatrix}$. Determine the conditional densities of $\tilde{X}_1|\tilde{X}_2,\ \tilde{X}_2|\tilde{X}_1,\ \tilde{Y}_1|\tilde{Y}_2$ and $\tilde{Y}_2|\tilde{Y}_1$ and the corresponding conditional expectations.

Solution 4.5a.1. As per the partitioning of $\tilde{X}$, we have the following partitions of $A,\ B$ and M:

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}, B_{22} = \begin{bmatrix} 2 & i \\ -i & 1 \end{bmatrix}, B_{22}^{-1} = \begin{bmatrix} 1 & -i \\ i & 2 \end{bmatrix}, B_{21} = \begin{bmatrix} i \\ 0 \end{bmatrix},$$

$$A_{11} = (2), A_{12} = (i), A_{21} = (-i), A_{22} = (1), B_{12} = [-i, 0], A_{11}^{-1} = \frac{1}{2}, A_{22}^{-1} = 1, B_{11}^{-1} = \frac{1}{3},$$

$$A_{11} - A_{12}A_{22}^{-1}A_{21} = 2 - (i)(1)(-i) = 1,\ |A_{11} - A_{12}A_{22}^{-1}A_{21}| = 1,$$

$$A_{22} - A_{21}A_{11}^{-1}A_{12} = \frac{1}{2},\ |A_{22} - A_{21}A_{11}^{-1}A_{12}| = \frac{1}{2},\ |A| = 1,\ |B| = 2,$$

$$B_{11} - B_{12}B_{22}^{-1}B_{21} = 3 - [-i\ 0]\begin{bmatrix} 1 & -i \\ i & 2 \end{bmatrix}\begin{bmatrix} i \\ 0 \end{bmatrix} = 2,\ |B_{11} - B_{12}B_{22}^{-1}B_{21}| = 2,$$

$$B_{22} - B_{21}B_{11}^{-1}B_{12} = \begin{bmatrix} 2 & i \\ -i & 1 \end{bmatrix} - \begin{bmatrix} i \\ 0 \end{bmatrix}(1/3)[-i\ 0] = \begin{bmatrix} \frac{5}{3} & i \\ -i & 1 \end{bmatrix},\ |B_{22} - B_{21}B_{11}^{-1}B_{12}| = \frac{2}{3}.$$

$$M_1 = [1+i,\ i,\ -i], M_2 = [i,\ 2+i,\ 1-i], M_{(1)} = \begin{bmatrix} 1+i \\ i \end{bmatrix}, M_{(2)} = \begin{bmatrix} i & -i \\ 2+i & 1-i \end{bmatrix}.$$

All the conditional expectations can now be determined. They are

$$E[\tilde{X}_1|\tilde{X}_2] = M_1 - A_{11}^{-1}A_{12}(\tilde{X}_2 - M_2) = [1+i,\ i,\ \ -i] - \frac{i}{2}(\tilde{X}_2 - M_2) \qquad (i)$$

$$E[\tilde{X}_2|\tilde{X}_1] = M_2 - A_{22}^{-1}A_{21}(\tilde{X}_1 - M_1) = [i,\ 2+i,\ 1-i] + i(\tilde{X}_1 - M_1) \qquad (ii)$$

$$E[\tilde{Y}_1|\tilde{Y}_2] = M_{(1)} - (\tilde{Y}_2 - M_{(2)})B_{21}B_{11}^{-1} = \frac{1}{3}\begin{bmatrix} 2+i\tilde{x}_{12} \\ -1+5i-i\tilde{x}_{22}\end{bmatrix} \qquad (iii)$$

$$\begin{aligned} E[\tilde{Y}_2|\tilde{Y}_1] &= M_{(2)} - (\tilde{Y}_1 - M_{(1)})B_{12}B_{22}^{-1} \\ &= \begin{bmatrix} i & -i \\ 2+i & 1-i\end{bmatrix} - \begin{bmatrix} \tilde{x}_{11} - (1+i) \\ \tilde{x}_{21} - i\end{bmatrix}\begin{bmatrix} -i & 0\end{bmatrix}\begin{bmatrix} 1 & -i \\ -i & 2\end{bmatrix} \\ &= \begin{bmatrix} i\tilde{x}_{11} + 1 & \tilde{x}_{11} - 1 - 2i \\ -i\tilde{x}_{21} + i + 3 & \tilde{x}_{21} + 1 - 2i\end{bmatrix}. \end{aligned} \qquad (iv)$$

Now, on substituting the above quantities in equations (4.5*a*.1), (4.5*a*.3), (4.5*a*.5) and (4.5*a*.7), the following densities are obtained:

$$\tilde{f}_{1,3}(\tilde{X}_1|\tilde{X}_2) = \frac{2^4}{\pi^3}\,\mathrm{e}^{-2(\tilde{X}_1 - E_1)B(\tilde{X}_1 - E_1)^*}$$

where $E_1 = E[\tilde{X}_1|\tilde{X}_2]$ given in *(i)*;

$$\tilde{f}_{1,3}(\tilde{X}_2|\tilde{X}_1) = \frac{2}{\pi^3}\,\mathrm{e}^{-(\tilde{X}_2 - E_2)B(\tilde{X}_2 - E_2)^*}$$

where $E_2 = E[\tilde{X}_2|\tilde{X}_1]$ given in *(ii)*;

$$\tilde{f}_{2,1}(\tilde{Y}_1|\tilde{Y}_2) = \frac{3^2}{\pi^2}\,\mathrm{e}^{-3\mathrm{tr}[A(\tilde{Y}_1 - E_3)(\tilde{Y}_1 - E_3)^*]}$$

where $E_3 = E[\tilde{Y}_1|\tilde{Y}_2]$ given in *(iii)*;

$$\tilde{f}_{2,2}(\tilde{Y}_2|\tilde{Y}_1) = \frac{1}{\pi^4}\,\mathrm{e}^{-\mathrm{tr}[A(\tilde{Y}_2 - E_4)B_{22}(\tilde{Y}_2 - E_4)^*]}$$

where $E_4 = E[\tilde{Y}_2|\tilde{Y}_1]$ given in *(iv)*. The exponent in the density of $\tilde{Y}_1|\tilde{Y}_2$ can be simplified as follows:

$$\begin{aligned} &-\mathrm{tr}[A(\tilde{Y}_1 - M_{(1)})B_{11}(\tilde{Y}_1 - M_{(1)})^*] = -3(\tilde{Y}_1 - M_{(1)})^* A(\tilde{Y}_1 - M_{(1)}) \\ &= -3[(\tilde{x}_{11} - (1+i))^*\ (\tilde{x}_{21} - i)^*]\begin{bmatrix} 2 & i \\ -i & 1\end{bmatrix}\begin{bmatrix} (\tilde{x}_{11} - (1+i)) \\ (\tilde{x}_{21} - i)\end{bmatrix} \\ &= -6\{(x_{111}^2 + x_{112}^2) + \tfrac{1}{2}(x_{211}^2 + x_{212}^2) + (x_{112}x_{211} - x_{111}x_{212}) - 2x_{112} - x_{111} - x_{211} + \tfrac{3}{2}\} \end{aligned}$$

by writing $\tilde{x}_{k1} = x_{k11} + ix_{k12},\ k = 1, 2,\ i = \sqrt{(-1)}$. This completes the computations.

4.5.1. Re-examination of the case $q = 1$

When $q = 1$, we have a $p \times 1$ vector-variate or the usual p-variate Gaussian density of the form in (4.5.5). Let us consider the real case first. Let the $p \times 1$ vector be denoted by Y_1 with

$$Y_1 = \begin{bmatrix} y_1 \\ \vdots \\ y_p \end{bmatrix} = \begin{bmatrix} Y_{(1)} \\ Y_{(2)} \end{bmatrix}, \ Y_{(1)} = \begin{bmatrix} y_1 \\ \vdots \\ y_{p_1} \end{bmatrix}, \ Y_{(2)} = \begin{bmatrix} y_{p_1+1} \\ \vdots \\ y_p \end{bmatrix};$$

$$M_{(1)} = \begin{bmatrix} M_{(1)}^{(p_1)} \\ M_{(2)}^{(p_2)} \end{bmatrix}, \ M_{(1)}^{(p_1)} = \begin{bmatrix} m_1 \\ \vdots \\ m_{p_1} \end{bmatrix}, \ M_{(2)}^{(p_2)} = \begin{bmatrix} m_{p_1+1} \\ \vdots \\ m_p \end{bmatrix}, \ E[Y_1] = M_{(1)}, \, p_1 + p_2 = p.$$

Then, from (4.5.2) wherein $q = 1$, we have

$$E[Y_{(1)}|Y_{(2)}] = M_{(1)}^{(p_1)} - A_{11}^{-1} A_{12}(Y_{(2)} - M_{(2)}^{(p_2)}), \tag{4.5.10}$$

with $A = \Sigma^{-1}$, Σ being the covariance matrix of Y_1, that is, $\text{Cov}(Y_1) = E[(Y_1 - E(Y_1))(Y_1 - E(Y_1))']$. Let

$$A^{-1} = \Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}, \text{ where } \Sigma_{11} \text{ is } p_1 \times p_1 \text{ and } \Sigma_{22} \text{ is } p_2 \times p_2.$$

From the partitioning of matrices presented in Sect. 1.3, we have

$$-A_{11}^{-1} A_{12} = A^{12}(A^{22})^{-1} = \Sigma_{12}\Sigma_{22}^{-1}. \tag{4.5.11}$$

Accordingly, we may rewrite (4.5.10) in terms of the sub-matrices of the covariance matrix as

$$E[Y_{(1)}|Y_{(2)}] = M_{(1)}^{(p_1)} + \Sigma_{12}\Sigma_{22}^{-1}(Y_{(2)} - M_{(2)}^{(p_2)}). \tag{4.5.12}$$

If $p_1 = 1$, then $Y_{(2)}$ will contain $p - 1$ elements, denoted by $Y'_{(2)} = (y_2, \ldots, y_p)$. Letting $E[y_1] = m_1$, we have

$$E[y_1|Y_{(2)}] = m_1 + \Sigma_{12}\Sigma_{22}^{-1}(Y_{(2)} - M_{(2)}^{(p_2)}), \ p_2 = p - 1. \tag{4.5.13}$$

The conditional expectation (4.5.13) is *the best predictor of* y_1 *at the preassigned values of* $y_2, \ldots, y_p$, where $m_1 = E[y_1]$. It will now be shown that $\Sigma_{12}\Sigma_{22}^{-1}$ can be expressed in terms of variances and correlations. Let $\sigma_j^2 = \sigma_{jj} = \text{Var}(y_j)$ where $\text{Var}(\cdot)$ denotes the variance of $(\cdot)$. Note that $\sigma_{ij} = \mathrm{Cov}(y_i, y_j)$ or the covariance between y_1 and y_j. Letting ρ_{ij} be the correlation between y_i and y_j, we have

$$\begin{aligned}\Sigma_{12} &= [\mathrm{Cov}(y_1, y_2), \ldots, \mathrm{Cov}(y_1, y_p)]\\ &= [\sigma_1\sigma_2\rho_{12}, \ldots, \sigma_1\sigma_p\rho_{1p}].\end{aligned}$$

Then

$$\Sigma = \begin{bmatrix} \sigma_1\sigma_1 & \sigma_1\sigma_2\rho_{12} & \cdots & \sigma_1\sigma_p\rho_{1p} \\ \sigma_2\sigma_1\rho_{21} & \sigma_2\sigma_2 & \cdots & \sigma_2\sigma_p\rho_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_p\sigma_1\rho_{p1} & \sigma_p\sigma_2\rho_{p2} & \cdots & \sigma_p\sigma_p \end{bmatrix}, \ \rho_{ij} = \rho_{ji}, \ \rho_{jj} = 1,$$

for all j. Let $D = \mathrm{diag}(\sigma_1, \ldots, \sigma_p)$ be a diagonal matrix whose diagonal elements are $\sigma_1, \ldots, \sigma_p$, the standard deviations of $y_1, \ldots, y_p$, respectively. Letting $R = (\rho_{ij})$ = denote the correlation matrix wherein ρ_{ij} is the correlation between y_i and y_j, we can express Σ as DRD, that is,

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_{pp} \end{bmatrix} = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_p \end{bmatrix} \begin{bmatrix} 1 & \rho_{12} & \cdots & \rho_{1p} \\ \rho_{21} & 1 & \cdots & \rho_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{p1} & \rho_{p2} & \cdots & 1 \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_p \end{bmatrix}$$

so that

$$\Sigma^{-1} = D^{-1}R^{-1}D^{-1}, \ p = 2, 3, \ldots \tag{4.5.14}$$

We can then re-express (4.5.13) in terms of variances and correlations since

$$\Sigma_{12}\Sigma_{22}^{-1} = \sigma_1 R_{12} D_{(2)} D_{(2)}^{-1} R_{22}^{-1} D_{(2)}^{-1} = \sigma_1 R_{12} R_{22}^{-1} D_{(2)}^{-1}$$

where $D_{(2)} = \mathrm{diag}(\sigma_2, \ldots, \sigma_p)$ and R is partitioned accordingly. Thus,

$$E[y_1|Y_{(2)}] = m_1 + \sigma_1 R_{12} R_{22}^{-1} D_{(2)}^{-1}(Y_{(2)} - M_{(2)}^{(p_2)}). \tag{4.5.15}$$

An interesting particular case occurs when $p = 2$, as there are then only two real scalar variables y_1 and y_2, and

$$E[y_1|y_2] = m_1 + \frac{\sigma_1}{\sigma_2}\rho_{12}(y_2 - m_2), \tag{4.5.16}$$

which is the *regression of* y_1 *on* y_2 *or the best predictor of* y_1 *at a given value of* y_2.

4.6. Sampling from a Real Matrix-variate Gaussian Density

Let the $p\times q$ matrix $X_\alpha=(x_{ij\alpha})$ have a $p\times q$ real matrix-variate Gaussian density with parameter matrices $M,\ A>O$ and $B>O$. When n independently and identically distributed (iid) matrix random variables that are distributed as X_α are available, we say that we have *a simple random sample of size n from* X_α *or from the population distributed as* X_α. We will consider simple random samples from a $p\times q$ matrix-variate Gaussian population in the real and complex domains. Since the procedures are parallel to those utilized in the vector variable case, we will recall the particulars in connection with that particular case. Some of the following materials are re-examinations of those already presented Chap. 3. For $q=1$, we have a p-vector which will be denoted by Y_1. In our previous notations, Y_1 is the same Y_1 for $q_1=1,\ q_2=0$ and $q=1$. Consider a sample of size n from a population distributed as Y_1 and let the $p\times n$ sample matrix be denoted by $\mathbf{Y}$. Then,

$$\mathbf{Y}=[Y_1,\ldots,Y_n]=\begin{bmatrix}y_{11}&y_{12}&\cdots&y_{1n}\\y_{21}&y_{22}&\cdots&y_{2n}\\\vdots&\vdots&\ddots&\vdots\\y_{p1}&y_{p2}&\cdots&y_{pn}\end{bmatrix},\ Y_1=\begin{bmatrix}y_{11}\\\vdots\\y_{p1}\end{bmatrix}.$$

In this case, the columns of $\mathbf{Y}$, that is, $Y_j,\ j=1,\ldots,n$, are iid variables, distributed as Y_1. Let an $n\times 1$ column vector whose components are all equal to 1 be denoted by J and consider

$$\bar{Y}=\frac{1}{n}\mathbf{Y}J=\frac{1}{n}\begin{bmatrix}y_{11}&\cdots&y_{1n}\\y_{21}&\cdots&y_{2n}\\\vdots&\ddots&\vdots\\y_{p1}&\cdots&y_{pn}\end{bmatrix}\begin{bmatrix}1\\\vdots\\1\end{bmatrix}=\begin{bmatrix}\bar{y}_1\\\bar{y}_2\\\vdots\\\bar{y}_p\end{bmatrix}$$

where $\bar{y}_j=\frac{\sum_{k=1}^n y_{jk}}{n}$ denotes the average of the variables, distributed as y_j. Let

$$S=(\mathbf{Y}-\bar{\mathbf{Y}})(\mathbf{Y}-\bar{\mathbf{Y}})' \text{ where the bold-faced } \bar{\mathbf{Y}}=\begin{bmatrix}\bar{y}_1&\cdots&\bar{y}_1\\\bar{y}_2&\cdots&\bar{y}_2\\\vdots&\ddots&\vdots\\\bar{y}_p&\cdots&\bar{y}_p\end{bmatrix}=[\bar{Y},\ldots,\bar{Y}].$$

Then,

$$S=(s_{ij}),\ s_{ij}=\sum_{k=1}^n(y_{ik}-\bar{y}_i)(y_{jk}-\bar{y}_j) \text{ for all } i \text{ and } j. \tag{4.6.1}$$

This matrix S is known as the *sample sum of products matrix or corrected sample sum of products matrix*. Here "corrected" indicates that the deviations are taken from the respective averages $\bar{y}_1, \ldots, \bar{y}_p$. Note that $\frac{1}{n}s_{ij}$ is equal to the sample covariance between y_i and y_j and when $i = j$, it is the sample variance of y_i. Observing that

$$J = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \Rightarrow JJ' = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix} \text{ and } J'J = n,$$

we have

$$\mathbf{Y}\big(\frac{1}{n}JJ'\big) = \bar{Y} \Rightarrow \mathbf{Y} - \bar{\mathbf{Y}} = \mathbf{Y}[I - \frac{1}{n}JJ'].$$

Hence

$$S = (\mathbf{Y} - \bar{\mathbf{Y}})(\mathbf{Y} - \bar{\mathbf{Y}})' = \mathbf{Y}[I - \frac{1}{n}JJ'][I - \frac{1}{n}JJ']'\mathbf{Y}'.$$

However,

$$\begin{aligned}[I - \frac{1}{n}JJ'][I - \frac{1}{n}JJ']' &= I - \frac{1}{n}JJ' - \frac{1}{n}JJ' + \frac{1}{n^2}JJ'JJ' \\ &= I - \frac{1}{n}JJ' \text{ since } J'J = n.\end{aligned}$$

Thus,

$$S = \mathbf{Y}[I - \frac{1}{n}JJ']\mathbf{Y}'. \tag{4.6.2}$$

Letting $C_1 = (I - \frac{1}{n}JJ')$, we note that $C_1^2 = C_1$ and that the rank of C_1 is $n - 1$. Accordingly, C_1 is an idempotent matrix having $n-1$ eigenvalues equal to 1, the remaining one being equal to zero. Now, letting $C_2 = \frac{1}{n}JJ'$, it is easy to verify that $C_2^2 = C_2$ and that the rank of C_2 is one; thus, C_2 is idempotent with $n - 1$ eigenvalues equal to zero, the remaining one being equal to 1. Further, since $C_1C_2 = O$, that is, C_1 and C_2 are orthogonal to each other, $\mathbf{Y} - \bar{\mathbf{Y}} = \mathbf{Y}C_1$ and $\bar{\mathbf{Y}} = \mathbf{Y}C_2$ are independently distributed, so that $\mathbf{Y} - \bar{\mathbf{Y}}$ and $\bar{Y}$ are independently distributed. Consequently, $S = (\mathbf{Y} - \bar{\mathbf{Y}})(\mathbf{Y} - \bar{\mathbf{Y}})'$ and $\bar{Y}$ are independently distributed as well. This will be stated as the next result.

Theorem 4.6.1, 4.6a.1. *Let $Y_1, \ldots, Y_n$ be a simple random sample of size n from a p-variate real Gaussian population having a $N_p(\mu, \Sigma)$, $\Sigma > O$, distribution. Let $\bar{Y}$ be the sample average and S be the sample sum of products matrix; then, $\bar{Y}$ and S are statistically independently distributed. In the complex domain, let the $\tilde{Y}_j$'s be iid $N_p(\tilde{\mu}, \tilde{\Sigma})$, $\tilde{\Sigma} =$*

$\tilde{\Sigma}^* > O$, and $\bar{\tilde{Y}}$ and $\tilde{S}$ denote the sample average and sample sum of products matrix; then, $\bar{\tilde{Y}}$ and $\tilde{S}$ are independently distributed.

4.6.1. The distribution of the sample sum of products matrix, real case

Reprising the notations of Sect. 4.6, let the $p \times n$ matrix $\mathbf{Y}$ denote a sample matrix whose columns $Y_1, \ldots, Y_n$ are iid as $N_p(\mu, \Sigma),\ \Sigma > O$, Gaussian vectors. Let the sample mean be $\bar{Y} = \frac{1}{n}(Y_1 + \cdots + Y_n) = \frac{1}{n}\mathbf{Y}J$ where $J' = (1, \ldots, 1)$. Let the bold-faced matrix $\bar{\mathbf{Y}} = [\bar{Y}, \ldots, \bar{Y}] = \mathbf{Y}C_1$ where $C_1 = I_n - \frac{1}{n}JJ'$. Note that $C_1 = I_n - C_2 = C_1^2$ and $C_2 = \frac{1}{n}JJ' = C_2^2$, that is, C_1 and C_2 are idempotent matrices whose respective ranks are $n-1$ and 1. Since $C_1 = C_1'$, there exists an $n \times n$ orthonormal matrix P, $PP' = I_n,\ P'P = I_n$, such that $P'C_1P = D$ where

$$D = \begin{bmatrix} I_{n-1} & O \\ O & 0 \end{bmatrix} = P'C_1P.$$

Let $\mathbf{Y} = ZP'$ where Z is $p \times n$. Then, $\mathbf{Y} = ZP' \Rightarrow \mathbf{Y}C_1 = ZP'C_1 = ZP'PDP' = ZDP'$, so that

$$S = (\mathbf{Y}C_1)(\mathbf{Y}C_1)' = \mathbf{Y}C_1C_1'\mathbf{Y}' = Z\begin{bmatrix} I_{n-1} & O \\ O & 0 \end{bmatrix}\begin{bmatrix} I_{n-1} & O \\ O & 0 \end{bmatrix}Z'$$

$$= (Z_{n-1}, O)(Z_{n-1}, O)' = Z_{n-1}Z_{n-1}' \tag{4.6.3}$$

where Z_{n-1} is a $p \times (n-1)$ matrix obtained by deleting the last column of the $p \times n$ matrix Z. Thus, $S = Z_{n-1}Z_{n-1}'$ where Z_{n-1} contains $p(n-1)$ distinct real variables. Accordingly, Theorems 4.2.1, 4.2.2, 4.2.3, and the analogous results in the complex domain, are applicable to Z_{n-1} as well as to the corresponding quantity $\tilde{Z}_{n-1}$ in the complex case. Observe that when $Y_1 \sim N_p(\mu, \Sigma)$, $\mathbf{Y}-\bar{\mathbf{Y}}$ has expected value $\mathbf{M}-\mathbf{M} = O, \mathbf{M} = (\mu, \ldots, \mu)$. Hence, $\mathbf{Y} - \bar{\mathbf{Y}} = (\mathbf{Y} - \mathbf{M}) - (\bar{\mathbf{Y}} - \mathbf{M})$ and therefore, without any loss of generality, we can assume Y_1 to be coming from a $N_p(O, \Sigma),\ \Sigma > O$, vector random variable whenever $\mathbf{Y} - \bar{\mathbf{Y}}$ is involved.

Theorem 4.6.2. *Let $\mathbf{Y},\ \bar{Y},\ \bar{\mathbf{Y}},\ J,\ C_1$ and C_2 be as defined in this section. Then, the $p \times n$ matrix $(\mathbf{Y} - \bar{\mathbf{Y}})J = O$, which implies that there exist linear relationships among the columns of $\mathbf{Y}$. However, all the elements of Z_{n-1} as defined in (4.6.3) are distinct real variables. Thus, Theorems 4.2.1, 4.2.2 and 4.2.3 are applicable to Z_{n-1}.*

Note that the corresponding result for the complex Gaussian case also holds.

4.6.2. Linear functions of sample vectors

Let $Y_j \overset{iid}{\sim} N_p(\mu, \Sigma),\ \Sigma > O,\ j = 1, \ldots, n$, or equivalently, let the Y_j's constitutes a simple random sample of size n from this p-variate real Gaussian population. Then, the density of the $p \times n$ sample matrix $\mathbf{Y}$, denoted by $L(\mathbf{Y})$, is the following:

$$L(\mathbf{Y}) = \frac{1}{(2\pi)^{\frac{np}{2}}|\Sigma|^{\frac{n}{2}}}\,\mathrm{e}^{-\frac{1}{2}\mathrm{tr}[\Sigma^{-1}(\mathbf{Y}-\mathbf{M})(\mathbf{Y}-\mathbf{M})']},$$

where $\mathbf{M} = (\mu, \ldots, \mu)$ is $p \times n$ whose columns are all equal to the $p \times 1$ parameter vector μ. Consider a linear function of the sample values $Y_1, \ldots, Y_n$. Let the linear function be $U = \mathbf{Y}A$ where A is an $n \times q$ constant matrix of rank q, $q \le p \le n$, so that U is $p \times q$. Let us consider the mgf of U. Since U is $p \times q$, we employ a $q \times p$ parameter matrix T so that $\mathrm{tr}(TU)$ will contain all the elements in U multiplied by the corresponding parameters. The mgf of U is then

$$\begin{aligned}M_U(T) &= E[\mathrm{e}^{\mathrm{tr}(TU)}] = E[\mathrm{e}^{\mathrm{tr}(T\mathbf{Y}A)}] = E[\mathrm{e}^{\mathrm{tr}(AT\mathbf{Y})}]\\ &= \mathrm{e}^{\mathrm{tr}(AT\mathbf{M})}E[\mathrm{e}^{\mathrm{tr}(AT(\mathbf{Y}-\mathbf{M}))}]\end{aligned}$$

where $\mathbf{M} = (\mu, \ldots, \mu)$. Letting $W = \Sigma^{-\frac{1}{2}}(\mathbf{Y} - \mathbf{M})$, $\mathrm{d}\mathbf{Y} = |\Sigma|^{\frac{n}{2}}\mathrm{d}W$ and

$$\begin{aligned}M_U(T) &= \mathrm{e}^{\mathrm{tr}(AT\mathbf{M})}|\Sigma|^{\frac{n}{2}}E[\mathrm{e}^{\mathrm{tr}(AT\Sigma^{\frac{1}{2}}W)}]\\ &= \frac{\mathrm{e}^{\mathrm{tr}(AT\mathbf{M})}}{(2\pi)^{\frac{np}{2}}}\int_W \mathrm{e}^{\mathrm{tr}(AT\Sigma^{\frac{1}{2}}W)-\frac{1}{2}\mathrm{tr}(WW')}\mathrm{d}W.\end{aligned}$$

Now, expanding

$$\mathrm{tr}[(W - C)(W - C)'] = \mathrm{tr}(WW') - 2\mathrm{tr}(WC') + \mathrm{tr}(CC').$$

and comparing the resulting expression with the exponent in the integrand, which excluding $-\frac{1}{2}$, is $\mathrm{tr}(WW') - 2\mathrm{tr}(AT\Sigma^{\frac{1}{2}}W)$, we may let $C' = AT\Sigma^{\frac{1}{2}}$ so that $\mathrm{tr}(CC') = \mathrm{tr}(AT\Sigma T'A') = \mathrm{tr}(T\Sigma T'A'A)$. Since $\mathrm{tr}(AT\mathbf{M}) = \mathrm{tr}(T\mathbf{M}A)$ and

$$\frac{1}{(2\pi)^{\frac{np}{2}}}\int_W \mathrm{e}^{-\frac{1}{2}((W-C)(W-C)')}\mathrm{d}W = 1,$$

we have

$$M_U(T) = M_{\mathbf{Y}A}(T) = \mathrm{e}^{\mathrm{tr}(T\mathbf{M}A)+\frac{1}{2}\mathrm{tr}(T\Sigma T'A'A)}$$

where $\mathbf{M}A = E[\mathbf{Y}A]$, $\Sigma > O$, $A'A > O$, A being a full rank matrix, and $T\Sigma T'A'A$ is a $q \times q$ positive definite matrix. Hence, the $p \times q$ matrix $U = \mathbf{Y}A$ has a matrix-variate real Gaussian density with the parameters $\mathbf{M}A = E[\mathbf{Y}A]$ and $A'A > O,\ \Sigma > O$. Thus, the following result:

Theorem 4.6.3, 4.6a.2. *Let $Y_j \overset{iid}{\sim} N_p(\mu, \Sigma),\ \Sigma > O,\ j = 1, \ldots, n$, or equivalently, let the Y_j's constitutes a simple random sample of size n from this p-variate real Gaussian population. Consider a set of linear functions of $Y_1, \ldots, Y_n$, $U = \mathbf{Y}A$ where $\mathbf{Y} = (Y_1, \ldots, Y_n)$ is a $p \times n$ sample matrix and A is an $n \times q$ constant matrix of rank q, $q \le p \le n$. Then, U has a nonsingular $p \times q$ matrix-variate real Gaussian distribution with the parameters $\mathbf{M}A = E[\mathbf{Y}A],\ A'A > O$, and $\Sigma > O$. Analogously, in the complex domain, $\tilde{U} = \tilde{\mathbf{Y}}A$ is a $p \times q$-variate complex Gaussian distribution with the corresponding parameters $E[\tilde{\mathbf{Y}}A],\ A^*A > O$, and $\tilde{\Sigma} > O$, A^* denoting the conjugate transpose of A. In the usual format of a $p \times q$ matrix-variate $N_{p,q}(M, A, B)$ real Gaussian density, M is replaced by MA, A, by $A'A$ and B, by Σ, in the real case, with corresponding changes for the complex case.*

A certain particular case turns out to be of interest. Observe that $\mathbf{M}A = \mu(J'A),\ J' = (1, \ldots, 1)$, and that when $q = 1$, we are considering only one linear combination of $Y_1, \ldots, Y_n$ in the form $U_1 = a_1Y_1 + \cdots + a_nY_n$, where $a_1, \ldots, a_n$ are real scalar constants. Then $J'A = \sum_{j=1}^{n} a_j,\ A'A = \sum_{j=1}^{n} a_j^2$, and the $p \times 1$ vector U_1 has a p-variate real nonsingular Gaussian distribution with the parameters $(\sum_{j=1}^{n} a_j)\mu$ and $(\sum_{j=1}^{n} a_j^2)\Sigma$. This result was stated in Theorem 3.5.4.

Corollary 4.6.1, 4.6a.1. *Let A as defined in Theorem 4.6.3 be $n \times 1$, in which case A is a column vector whose components are $a_1, \ldots, a_n$, and the resulting single linear function of $Y_1, \ldots, Y_n$ is $U_1 = a_1Y_1 + \cdots + a_nY_n$. Let the population be p-variate real Gaussian with the parameters μ and $\Sigma > O$. Then U_1 has a p-variate nonsingular real normal distribution with the parameters $(\sum_{j=1}^{n} a_j)\mu$ and $(\sum_{j=1}^{n} a_j^2)\Sigma$. Analogously, in the complex Gaussian population case, $\tilde{U}_1 = a_1\tilde{Y}_1 + \cdots + a_n\tilde{Y}_n$ is distributed as a complex Gaussian with mean value $(\sum_{j=1}^{n} a_j)\tilde{\mu}$ and covariance matrix $(\sum_{j=1}^{n} a_j^* a_j)\tilde{\Sigma}$. Taking $a_1 = \cdots = a_n = \frac{1}{n}$, $U_1 = \frac{1}{n}(Y_1 + \cdots + Y_n) = \bar{Y}$, the sample average, which has a p-variate real Gaussian density with the parameters μ and $\frac{1}{n}\Sigma$. Correspondingly, in the complex Gaussian case, the sample average $\bar{\tilde{Y}}$ is a p-variate complex Gaussian vector with the parameters $\tilde{\mu}$ and $\frac{1}{n}\tilde{\Sigma},\ \tilde{\Sigma} = \tilde{\Sigma}^* > O$.*

4.6.3. The general real matrix-variate case

In order to avoid a multiplicity of symbols, we will denote the $p \times q$ real matrix-variate random variable by $X_\alpha = (x_{ij\alpha})$ and the corresponding complex matrix by $\tilde{X}_\alpha = (\tilde{x}_{ij\alpha})$. Consider a simple random sample of size n from the population represented by the real

$p \times q$ matrix $X_\alpha = (x_{ij\alpha})$. Let $X_\alpha = (x_{ij\alpha})$ be the α-th sample value, so that the X_α's, $\alpha = 1, \ldots, n$, are iid as X_1. Let the $p \times nq$ sample matrix be denoted by the bold-faced $\mathbf{X} = [X_1, X_2, \ldots, X_n]$ where each X_j is $p \times q$. Let the sample average be denoted by $\bar{X} = (\bar{x}_{ij}),\ \bar{x}_{ij} = \frac{1}{n}\sum_{\alpha=1}^{n} x_{ij\alpha}$. Let $\mathbf{X_d}$ be the sample deviation matrix which is the $p \times qn$ matrix

$$\mathbf{X_d} = [X_1 - \bar{X}, X_2 - \bar{X}, \ldots, X_n - \bar{X}],\ X_\alpha - \bar{X} = (x_{ij\alpha} - \bar{x}_{ij}), \tag{4.6.4}$$

wherein the corresponding sample average is subtracted from each element. For example,

$$\begin{aligned} X_\alpha - \bar{X} &= \begin{bmatrix} x_{11\alpha} - \bar{x}_{11} & x_{12\alpha} - \bar{x}_{12} & \cdots & x_{1q\alpha} - \bar{x}_{1q} \\ x_{21\alpha} - \bar{x}_{21} & x_{22\alpha} - \bar{x}_{22} & \cdots & x_{2q\alpha} - \bar{x}_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ x_{p1\alpha} - \bar{x}_{p1} & x_{p2\alpha} - \bar{x}_{p2} & \cdots & x_{pq\alpha} - \bar{x}_{pq} \end{bmatrix} \\ &= \begin{bmatrix} C_{1\alpha} & C_{2\alpha} & \cdots & C_{q\alpha} \end{bmatrix} \end{aligned} \tag{i}$$

where $C_{j\alpha}$ is the j-th column in the α-th sample deviation matrix $X_\alpha - \bar{X}$. In this notation, the $p \times qn$ sample deviation matrix can be expressed as follows:

$$\mathbf{X_d} = [C_{11}, C_{21}, \ldots, C_{q1}, C_{12}, C_{22}, \ldots, C_{q2}, \ldots, C_{1n}, C_{2n}, \ldots, C_{qn}] \tag{ii}$$

where, for example, $C_{\gamma\alpha}$ denotes the γ-th column in the α-th $p \times q$ matrix, $X_\alpha - \bar{X}$, that is,

$$C_{\gamma\alpha} = \begin{bmatrix} x_{1\gamma\alpha} - \bar{x}_{1\gamma} \\ x_{2\gamma\alpha} - \bar{x}_{2\gamma} \\ \vdots \\ x_{p\gamma\alpha} - \bar{x}_{p\gamma} \end{bmatrix}. \tag{iii}$$

Then, the sample sum of products matrix, denoted by S, is given by

$$\begin{aligned} S = \mathbf{X_d}\mathbf{X_d}' = &\ C_{11}C_{11}' + C_{21}C_{21}' + \cdots + C_{q1}C_{q1}' \\ &+ C_{12}C_{12}' + C_{22}C_{22}' + \cdots + C_{q2}C_{q2}' \\ &\ \vdots \\ &+ C_{1n}C_{1n}' + C_{2n}C_{2n}' + \cdots + C_{qn}C_{qn}'. \end{aligned} \tag{iv}$$

Let us rearrange these matrices by collecting the terms relevant to each column of $\mathbf{X}$ which are

$$\begin{bmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_{p1} \end{bmatrix}, \begin{bmatrix} x_{12} \\ x_{22} \\ \vdots \\ x_{p2} \end{bmatrix}, \ldots, \begin{bmatrix} x_{1q} \\ x_{2q} \\ \vdots \\ x_{pq} \end{bmatrix}.$$

Then, the terms relevant to these columns are the following:

$$\begin{aligned} S = \mathbf{X_d}\mathbf{X_d}' &= C_{11}C_{11}' + C_{21}C_{21}' + \cdots + C_{q1}C_{q1}' \\ &\quad + C_{12}C_{12}' + C_{22}C_{22}' + \cdots + C_{q2}C_{q2}' \\ &\quad \vdots \\ &\quad + C_{1n}C_{1n}' + C_{2n}C_{2n}' + \cdots + C_{qn}C_{qn}' \\ &\equiv S_1 + S_2 + \cdots + S_q \qquad (v) \end{aligned}$$

where S_1 denotes the $p \times p$ sample sum of products matrix in the first column of $\mathbf{X}$, S_2, the $p \times p$ sample sum of products matrix corresponding to the second column of $\mathbf{X}$, and so on, S_q being equal to the $p \times p$ sample sum of products matrix corresponding to the q-th column of $\mathbf{X}$.

Theorem 4.6.4. *Let $X_\alpha = (x_{ij\alpha})$ be a real $p \times q$ matrix of distinct real scalar variables $x_{ij\alpha}$'s. Letting $X_\alpha,\ \bar{X},\ \mathbf{X},\ \mathbf{X_d},\ S$, and $S_1, \ldots, S_q$ be as previously defined, the sample sum of products matrix in the $p \times nq$ sample matrix $\mathbf{X}$, denoted by S, is given by*

$$S = S_1 + \cdots + S_q. \tag{4.6.5}$$

Example 4.6.1. Consider a 2×2 real matrix-variate $N_{2,2}(O, A, B)$ distribution with the parameters

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \text{ and } B = \begin{bmatrix} 3 & -1 \\ -1 & 2 \end{bmatrix}.$$

Let $X_\alpha,\ \alpha = 1, \ldots, 5$, be a simple random sample of size 5 from this real Gaussian population. Suppose that the following observations on $X_\alpha,\ \alpha = 1, \ldots, 5$, were obtained:

$$X_1 = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix},\ X_2 = \begin{bmatrix} -1 & 1 \\ -2 & 1 \end{bmatrix},\ X_3 = \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix}$$
$$X_4 = \begin{bmatrix} -1 & 1 \\ 1 & 2 \end{bmatrix},\ X_5 = \begin{bmatrix} -4 & 1 \\ -1 & -2 \end{bmatrix}.$$

Compute the sample matrix, the sample average, the sample deviation matrix and the sample sum of products matrix.

Solution 4.6.1. The sample average is available as

$$\bar{X} = \frac{1}{5}[X_1 + \cdots + X_5]$$
$$= \frac{1}{5}\begin{bmatrix} 1+(-1)+0+(-1)+(-4) & 1+1+1+1+1 \\ 1+(-2)+1+1+(-1) & 2+1+2+2+(-2) \end{bmatrix} = \begin{bmatrix} -1 & 1 \\ 0 & 1 \end{bmatrix}.$$

The deviations are then

$$X_{1d} = X_1 - \bar{X} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} - \begin{bmatrix} -1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix},\ X_{2d} = \begin{bmatrix} 0 & 0 \\ -2 & 0 \end{bmatrix}$$
$$X_{3d} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix},\ X_{4d} = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix},\ X_{5d} = \begin{bmatrix} -3 & 0 \\ -1 & -3 \end{bmatrix}.$$

Thus, the sample matrix, the sample average matrix and the sample deviation matrix, denoted by bold-faced letters, are the following:

$$\mathbf{X} = [X_1, X_2, X_3, X_4, X_5],\ \bar{\mathbf{X}} = [\bar{X}, \ldots, \bar{X}] \text{ and } \mathbf{X_d} = [X_{1d}, X_{2d}, X_{3d}, X_{4d}, X_{5d}].$$

The sample sum of products matrix is then

$$S = [\mathbf{X} - \bar{\mathbf{X}}][\mathbf{X} - \bar{\mathbf{X}}]' = [\mathbf{X_d}][\mathbf{X_d}]' = S_1 + S_2$$

where S_1 is obtained from the first columns of each of $X_{\alpha d},\ \alpha = 1, \ldots, 5$, and S_2 is evaluated from the second columns of $X_{\alpha d},\ \alpha = 1, \ldots, 5$. That is,

$$S_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}[2\ 1] + \begin{bmatrix} 0 \\ -2 \end{bmatrix}[0\ -2] + \begin{bmatrix} 1 \\ 1 \end{bmatrix}[1\ 1] + \begin{bmatrix} 0 \\ 1 \end{bmatrix}[0\ 1] + \begin{bmatrix} -3 \\ -1 \end{bmatrix}[-3\ -1]$$
$$= \begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 4 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 9 & 3 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} 14 & 6 \\ 6 & 8 \end{bmatrix};$$
$$S_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}[0\ 1] + \begin{bmatrix} 0 \\ 0 \end{bmatrix}[0\ 0] + \begin{bmatrix} 0 \\ 1 \end{bmatrix}[0\ 1] + \begin{bmatrix} 0 \\ 1 \end{bmatrix}[0\ 1] + \begin{bmatrix} 0 \\ -3 \end{bmatrix}[0\ -3]$$
$$= \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} + O + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 9 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 12 \end{bmatrix};$$
$$S = S_1 + S_2 = \begin{bmatrix} 14 & 6 \\ 6 & 20 \end{bmatrix}.$$

This S can be directly verified by taking $[\mathbf{X} - \bar{\mathbf{X}}][\mathbf{X} - \bar{\mathbf{X}}]' = [\mathbf{X_d}][\mathbf{X_d}]' =$ where

$$\mathbf{X} - \bar{\mathbf{X}} = \mathbf{X_d} = \begin{bmatrix} 2 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & -3 & 0 \\ 1 & 1 & -2 & 0 & 1 & 1 & 1 & 1 & -1 & -3 \end{bmatrix},\ S = \mathbf{X}_d\mathbf{X}_d'.$$

4.6a. The General Complex Matrix-variate Case

The preceding analysis has its counterpart for the complex case. Let $\tilde{X}_\alpha = (\tilde{x}_{ij\alpha})$ be a $p \times q$ matrix in the complex domain with the $\tilde{x}_{ij\alpha}$'s being distinct complex scalar variables. Consider a simple random sample of size n from this population designated by $\tilde{X}_1$. Let the α-th sample matrix be $\tilde{X}_\alpha,\ \alpha = 1,\ldots,n$, the $\tilde{X}_\alpha$'s being iid as $\tilde{X}_1$, and the $p \times nq$ sample matrix be denoted by the bold-faced $\tilde{\mathbf{X}} = [\tilde{X}_1,\ldots,\tilde{X}_n]$. Let the sample average be denoted by $\bar{\tilde{X}} = (\bar{\tilde{x}}_{ij})$, $\bar{\tilde{x}}_{ij} = \frac{1}{n}\sum_{\alpha=1}^{n}\tilde{x}_{ij\alpha}$, and $\tilde{\mathbf{X}}_{\mathbf{d}}$ be the sample deviation matrix:

$$\tilde{\mathbf{X}}_{\mathbf{d}} = [\tilde{X}_1 - \bar{\tilde{X}},\ldots,\tilde{X}_n - \bar{\tilde{X}}].$$

Let $\tilde{S}$ be the sample sum of products matrix, namely, $\tilde{S} = \tilde{\mathbf{X}}_{\mathbf{d}}\tilde{\mathbf{X}}_{\mathbf{d}}^*$ where an asterisk denotes the complex conjugate transpose and let $\tilde{S}_j$ be the sample sum of products matrix corresponding to the j-th column of $\tilde{\mathbf{X}}$. Then we have the following result:

Theorem 4.6a.3. *Let $\tilde{\mathbf{X}},\ \bar{\tilde{X}},\ \tilde{\mathbf{X}}_{\mathbf{d}},\ \tilde{S}$ and $\tilde{S}_j$ be as previously defined. Then,*

$$\tilde{S} = \tilde{S}_1 + \cdots + \tilde{S}_q = \tilde{\mathbf{X}}_d\tilde{\mathbf{X}}_d^*. \qquad (4.6a.1)$$

Example 4.6a.1. Consider a 2×2 complex matrix-variate $\tilde{N}_{2,2}(O, A, B)$ distribution where

$$A = \begin{bmatrix} 2 & 1+i \\ 1-i & 3 \end{bmatrix} \ \text{and}\ B = \begin{bmatrix} 2 & i \\ -1 & 2 \end{bmatrix}.$$

A simple random sample of size 4 from this population is available, that is, $\tilde{X}_\alpha \overset{iid}{\sim} \tilde{N}_{2,2}(O, A, B),\ \alpha = 1, 2, 3, 4$. The following are one set of observations on these sample values:

$$\tilde{X}_1 = \begin{bmatrix} 2 & i \\ -i & 1 \end{bmatrix},\ \tilde{X}_2 = \begin{bmatrix} 3 & -i \\ i & 1 \end{bmatrix},\ \tilde{X}_3 = \begin{bmatrix} 1 & 1-i \\ 1+i & 3 \end{bmatrix},\ \tilde{X}_4 = \begin{bmatrix} 2 & 3+i \\ 3-i & 7 \end{bmatrix}.$$

Determine the observed sample average, the sample matrix, the sample deviation matrix and the sample sum of products matrix.

Solution 4.6a.1. The sample average is

$$\begin{aligned}\bar{\tilde{X}} &= \frac{1}{4}[\tilde{X}_1 + \tilde{X}_2 + \tilde{X}_3 + \tilde{X}_4] \\ &= \frac{1}{4}\left\{\begin{bmatrix} 2 & i \\ -i & 1 \end{bmatrix} + \begin{bmatrix} 3 & -i \\ i & 1 \end{bmatrix} + \begin{bmatrix} 1 & 1-i \\ 1+i & 3 \end{bmatrix} + \begin{bmatrix} 2 & 3+i \\ 3-i & 7 \end{bmatrix}\right\} = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}\end{aligned}$$

and the deviations are as follows:

$$\tilde{X}_{1d} = \tilde{X}_1 - \bar{\tilde{X}} = \begin{bmatrix} 0 & -1+i \\ -1-i & -2 \end{bmatrix},\ \tilde{X}_{2d} = \begin{bmatrix} 1 & -1-i \\ -1+i & -2 \end{bmatrix},$$
$$\tilde{X}_{3d} = \begin{bmatrix} -1 & -i \\ i & 0 \end{bmatrix},\ \tilde{X}_{4d} = \begin{bmatrix} 0 & 2+i \\ 2-i & 4 \end{bmatrix}.$$

The sample deviation matrix is then $\tilde{\mathbf{X}}_{\mathbf{d}} = [\tilde{X}_{1d}, \tilde{X}_{2d}, \tilde{X}_{3d}, \tilde{X}_{4d}]$. If $V_{\alpha 1}$ denotes the first column of $\tilde{X}_{\alpha d}$, then with our usual notation, $\tilde{S}_1 = \sum_{j=1}^{4} V_{\alpha 1} V_{\alpha 1}^*$ and similarly, if $V_{\alpha 2}$ is the second column of $\tilde{\mathbf{X}}_{\alpha\mathbf{d}}$, then $\tilde{S}_2 = \sum_{\alpha=1}^{4} V_{\alpha 2} V_{\alpha 2}^*$, the sample sum of products matrix being $\tilde{S} = \tilde{S}_1 + \tilde{S}_2$. Let us evaluate these quantities:

$$\tilde{S}_1 = \begin{bmatrix} 0 \\ -1-i \end{bmatrix} [0\ \ -1+i] + \begin{bmatrix} 1 \\ -1+i \end{bmatrix} [1\ \ -1-i] + \begin{bmatrix} -1 \\ i \end{bmatrix} [-1\ \ -i] + \begin{bmatrix} 0 \\ 2-i \end{bmatrix} [0\ 2+i]$$
$$= \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix} + \begin{bmatrix} 1 & -1-i \\ -1+i & 2 \end{bmatrix} + \begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 5 \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ -1 & 10 \end{bmatrix},$$

$$\tilde{S}_2 = \begin{bmatrix} -1+i \\ -2 \end{bmatrix} [-1-i\ \ -2] + \begin{bmatrix} -1-i \\ -2 \end{bmatrix} [-1+i\ \ -2] + \begin{bmatrix} -i \\ 0 \end{bmatrix} [i\ 0] + \begin{bmatrix} 2+i \\ 4 \end{bmatrix} [2-i\ 4]$$
$$= \begin{bmatrix} 2 & 2-2i \\ 2+2i & 4 \end{bmatrix} + \begin{bmatrix} 2 & 2+2i \\ 2-2i & 4 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 5 & 8+4i \\ 8-4i & 16 \end{bmatrix}$$
$$= \begin{bmatrix} 10 & 12+4i \\ 12-4i & 24 \end{bmatrix},$$

and then,

$$\tilde{S} = \tilde{S}_1 + \tilde{S}_2 = \begin{bmatrix} 2 & -1 \\ -1 & 10 \end{bmatrix} + \begin{bmatrix} 10 & 12+4i \\ 12-4i & 24 \end{bmatrix} = \begin{bmatrix} 12 & 11+4i \\ 11-4i & 34 \end{bmatrix}.$$

This can also be verified directly as $\tilde{S} = [\tilde{\mathbf{X}}_{\mathbf{d}}][\tilde{\mathbf{X}}_{\mathbf{d}}]^*$ where the deviation matrix is

$$\tilde{\mathbf{X}}_{\mathbf{d}} = \begin{bmatrix} 0 & -1+i & 1 & -1-i & -1 & -i & 0 & 2+i \\ -i-1 & -2 & -1+i & -2 & i & 0 & 2-i & 4 \end{bmatrix}.$$

As expected,

$$[\tilde{\mathbf{X}}_{\mathbf{d}}][\tilde{\mathbf{X}}_{\mathbf{d}}]^* = \begin{bmatrix} 12 & 11+4i \\ 11-4i & 34 \end{bmatrix}.$$

This completes the calculations.

Exercises 4.6

4.6.1. Let A be a 2×2 matrix whose first row is $(1, 1)$ and B be 3×3 matrix whose first row is $(1, -1, 1)$. Select your own real numbers to complete the matrices A and B so that $A > O$ and $B > O$. Then consider a 2×3 matrix X having a real matrix-variate Gaussian density with the location parameter $M = O$ and the foregoing parameter matrices A and B. Let the first row of X be X_1 and its second row be X_2. Determine the marginal densities of X_1 and X_2, the conditional density of X_1 given X_2, the conditional density of X_2 given X_1, the conditional expectation of X_1 given $X_2 = (1, 0, 1)$ and the conditional expectation of X_2 given $X_1 = (1, 2, 3)$.

4.6.2. Consider the matrix X utilized in Exercise 4.6.1. Let its first two columns be Y_1 and its last one be Y_2. Then, obtain the marginal densities of Y_1 and Y_2, and the conditional densities of Y_1 given Y_2 and Y_2 given Y_1, and evaluate the conditional expectation of Y_1 given $Y_2' = (1, -1)$ as well as the conditional expectation of Y_2 given $Y_1 = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$.

4.6.3. Let $A > O$ and $B > O$ be 2×2 and 3×3 matrices whose first rows are $(1, 1-i)$ and $(2, i, 1+i)$, respectively. Select your own complex numbers to complete the matrices $A = A^* > O$ and $B = B^* > O$. Now, consider a 2×3 matrix $\tilde{X}$ having a complex matrix-variate Gaussian density with the aforementioned matrices A and B as parameter matrices. Assume that the location parameter is a null matrix. Letting the row partitioning of $\tilde{X}$, denoted by $\tilde{X}_1, \tilde{X}_2$, be as specified in Exercise 4.6.1, answer all the questions posed in that exercise.

4.6.4. Let A, B and $\tilde{X}$ be as given in Exercise 4.6.3. Consider the column partitioning specified in Exercise 4.6.2. Then answer all the questions posed in Exercise 4.6.2.

4.6.5. Repeat Exercise 4.6.4 with the non-null location parameter

$$\tilde{M} = \begin{bmatrix} 2 & 1-i & i \\ 1+i & 2+i & -3i \end{bmatrix}.$$

4.7. The Singular Matrix-variate Gaussian Distribution

Consider the moment generating function specified in (4.3.3) for the real case, namely,

$$M_X(T) = M_f(T) = \mathrm{e}^{\mathrm{tr}(TM')+\frac{1}{2}\mathrm{tr}(\Sigma_1 T\Sigma_2 T')} \qquad (4.7.1)$$

where $\Sigma_1 = A^{-1} > O$ and $\Sigma_2 = B^{-1} > O$. In the complex case, the moment generating function is of the form

$$\tilde{M}_{\tilde{X}}(\tilde{T}) = \mathrm{e}^{\Re[\mathrm{tr}(\tilde{T}\tilde{M}^*)]+\frac{1}{4}\mathrm{tr}(\Sigma_1\tilde{T}\Sigma_2\tilde{T}^*)}. \qquad (4.7a.1)$$

The properties of the singular matrix-variate Gaussian distribution can be studied by making use of (4.7.1) and (4.7a.1). Suppose that we restrict Σ_1 and Σ_2 to be positive semi-definite matrices, that is, $\Sigma_1 \geq O$ and $\Sigma_2 \geq O$. In this case, one can also study many properties of the distributions represented by the mgf's given in (4.7.1) and (4.7a.1); however, the corresponding densities will not exist unless the matrices Σ_1 and Σ_2 are both strictly positive definite. The $p \times q$ real or complex matrix-variate density does not exist if at least one of A or B is singular. When either or both Σ_1 and Σ_2 are only positive semi-definite, the distributions corresponding to the mgf's specified by (4.7.1) and (4.7a.1) are respectively referred to as *real matrix-variate singular Gaussian and complex matrix-variate singular Gaussian.*

For instance, let

$$\Sigma_1 = \begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix} \text{ and } \Sigma_2 = \begin{bmatrix} 3 & -1 & 0 \\ -1 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

in the mgf of a 2×3 real matrix-variate Gaussian distribution. Note that $\Sigma_1 = \Sigma_1'$ and $\Sigma_2 = \Sigma_2'$. Since the leading minors of Σ_1 are $|(4)| = 4 > 0$ and $|\Sigma_1| = 0$ and those of Σ_2 are $|(3)| = 3 > 0$, $\begin{vmatrix} 3 & -1 \\ -1 & 2 \end{vmatrix} = 5 > 0$ and $|\Sigma_2| = 2 > 0$, Σ_1 is positive semi-definite and Σ_2 is positive definite. Accordingly, the resulting Gaussian distribution does not possess a density. Fortunately, its distributional properties can nevertheless be investigated via its associated moment generating function.

References

A.M. Mathai (1997): *Jacobians of Matrix Transformations and Functions of Matrix Argument*, World Scientific Publishing, New York.

A.M. Mathai and Nicy Sebastian (2022): On distributions of covariance structures, *Communications in Statistics*, *Theory and Methods*, http://doi.org/10.1080/03610926.2022.2045022.