12-1990

Characterization of Matrix Variate Normal Distributions

A.K. Gupta
Bowling Green State University - Main Campus

T. Varga
Rose-Hulman Institute of Technology

MSTR 90-10

CHARACTERIZATION OF MATRIX VARIATE NORMAL DISTRIBUTIONS

A.K. Gupta and T. Varga

MS TR 90-10

December 1990

Department of Mathematics
Rose-Hulman Institute of Technology
Terre Haute, IN 47803

FAX(812) 877-3198 Phone: (812) 877-8391

CHARACTERIZATION OF MATRIX VARIATE NORMAL DISTRIBUTIONS

A.K. Gupta and T. Varga

MS TR 90-10

December 1990

Department of Mathematics
Rose-Hulman Institute of Technology
Terre Haute, IN 47803

FAX(812) 877-3198

Phone: (812) 877-8391

CHARACTERIZATION OF MATRIX VARIATE NORMAL DISTRIBUTIONS

by

A. K. Gupta and T. Varga*
Bowling Green State University

ABSTRACT

In this paper, it is shown that two random matrices have a joint matrix variate distribution if conditioning each one on the other the resulting distributions satisfy certain conditions. A general result involving more than two matrices is also proved.

*Now at the Rose-Hulman Institute of Technology.

AMS 1985 subject classification: Primary 62H10; Secondary 62H05

Key words and phrases: random matrices; conditional distributions; linearly independent; normal distribution.

1. INTRODUCTION

The characterization of multivariate normal distribution through conditional distributions has been studied by many authors in recent years. Results on bivariate normal distribution were obtained by Brucker (1979), and Fraser and Streit (1980). Khatri (1979) gave characterizations of multivariate normality through regression. In the present paper, their results are generalized for the matrix variate normal distribution.

2. THE MAIN RESULT

In order to derive the results of this paper the following lemma will be useful. It shows how the joint density of two random matrices can be obtained from the conditional densities.

LEMMA 1.1. *Let* $X \in \mathbb{R}^{p\times n}$ *and* $Y \in \mathbb{R}^{q\times m}$ *be random matrices with joint probability density function* $f(X,Y)$. *Let* $g_1(X)$ *and* $g_2(Y)$ *denote the marginal densities, and* $h_1(X \mid Y)$ *and* $h_2(Y \mid X)$ *be the conditional densities. Assume* $f(X,Y), g_1(X), g_2(Y), h_1(X \mid Y)$, *and* $h_2(Y \mid X)$ *are defined for all* $X \in \mathbb{R}^{p\times n}, Y \in \mathbb{R}^{q\times m}$. *Suppose there exists* $Y_0 \in \mathbb{R}^{q\times m}$ *such that* $h_2(Y_0 \mid X) \neq 0$ *for all* $X \in \mathbb{R}^{p\times n}$. *Then*

$$f(X,Y) = k\,\frac{h_2(Y \mid X)\, h_1(X \mid Y_0)}{h_2(Y_0 \mid X)},$$

where k *is a constant.*

Proof. Let $k = g_2(Y_0)$. Then

$$k\,\frac{h_2(Y \mid X)\, h_1(X \mid Y_0)}{h_2(Y_0 \mid X)} = g_2(Y_0)\,\frac{\dfrac{f(X,Y)}{g_1(X)}\,\dfrac{f(X,Y_0)}{g_2(Y_0)}}{\dfrac{f(X,Y_0)}{g_1(X)}} = f(X,Y). \blacksquare$$

Now we can derive the main results.

THEOREM 1.1. *Let* X: $p \times n$, Y: $q \times n$ *be random matrices and suppose that* $Y \mid X \sim N_{q,n}(C + DX, \Sigma_2 \otimes \Phi)$, $X \mid Y = Y_0 \sim N_{p,n}(M, \Sigma_1 \otimes \Phi)$, *where* C: $q \times n$, D: $q \times p$, Σ_2: $q \times q$, Φ: $n \times n$, M: $p \times n$, Σ_1: $p \times p$, $\Sigma_1 > 0$, $\Sigma_2 > 0$, $\Phi > 0$, *and* Y_0 *is a fixed* $q \times n$ *matrix. Define* $B = \Sigma_1 D' \Sigma_2^{-1}$, $A = M - \Sigma_1 D' \Sigma_2^{-1} Y_0$. *Then*

$$Z = \begin{pmatrix} X \\ Y \end{pmatrix} \sim N_{p+q,n}\left(\begin{pmatrix} (I_p - BD)^{-1}(A + BC) \\ (I_q - DB)^{-1}(C + DA) \end{pmatrix}, \begin{pmatrix} (I_p - BD)^{-1}\Sigma_1 & (I_p - BD)^{-1}B\Sigma_2 \\ D(I_p - BD)^{-1}\Sigma_1 & (I_q - DB)^{-1}\Sigma_2 \end{pmatrix} \otimes \Phi\right) \qquad (2.1)$$

Proof. Let $f(X,Y)$ be the joint probability density function of X and Y, $g_1(X)$ the marginal density of X, $g_2(Y)$ the marginal density of Y, and $h_1(X \mid Y_0)$, $h_2(Y \mid X)$ the conditional densities. Throughout the proof k will denote constants which need not be equal. Using Lemma 1.1, we obtain,

$$f(X,Y) = k \frac{\operatorname{etr}\left\{-\frac{1}{2}(Y - (C + DX))'\Sigma_2^{-1}(Y - (C + DX))\Phi\right\} \operatorname{etr}\left\{-\frac{1}{2}(X - M)'\Sigma_1^{-1}(X - M)\Phi^{-1}\right\}}{\operatorname{etr}\left\{-\frac{1}{2}(Y_0 - (C + DX))'\Sigma_2^{-1}(Y_0 - (C + DX))\Phi^{-1}\right\}}$$

$$= k \operatorname{etr}\left\{-\frac{1}{2}(X'\Sigma_1^{-1}X - 2X'\Sigma_1^{-1}Y + Y'\Sigma_2^{-1}Y - 2Y'\Sigma_2^{-1}C - 2X'D'\Sigma_2^{-1}Y)\Phi\right\}$$

$$= k \operatorname{etr}\left\{-\frac{1}{2}(Z'\Omega Z + Z'R)\Phi^{-1}\right\}$$

where $\Omega = \begin{pmatrix} \Sigma_1^{-1} & -D'\Sigma_2^{-1} \\ -\Sigma_2^{-1}D & \Sigma_2^{-1} \end{pmatrix}$ and $R = \begin{pmatrix} -2\Sigma_1^{-1}A \\ -2\Sigma_2^{-1}C \end{pmatrix}$. Since $f(X,Y)$ is a probability density function, the symmetric matrix Ω must be positive definite and hence is non-singular. Thus

$$f(Z) = k\,\mathrm{etr}\left\{-\frac{1}{2}\left(Z-\left(-\frac{1}{2}\Omega^{-1}R\right)\right)'\Omega\left(Z-\left(-\frac{1}{2}\Omega^{-1}R\right)\right)\Phi\right\}.$$

Now using the fact that $D(I_p - BD)^{-1} = (I_q - DB)^{-1}D$ and $B(I_q - DB)^{-1} = (I_p - BD)^{-1}B$ is is easy to see that

$$\Omega^{-1} = \begin{pmatrix} (I_p - BD)^{-1}\Sigma_1 & (I_p - BD)^{-1}B\Sigma_2 \\ D(I_p - BD)^{-1}\Sigma_1 & (I_q - DB)^{-1}\Sigma_2 \end{pmatrix}$$

Then

$$-\frac{1}{2}\Omega^{-1}R = \begin{pmatrix} (I_p - BD)^{-1}(A + BC) \\ (I_q - DB)^{-1}(C + DA) \end{pmatrix}$$

which completes the proof. ■

Using Theorem 1.1 we can derive the following result.

<u>COROLLARY 1.1.1</u>. *Let* X: $p \times n$, Y: $q \times n$ *be random matrices and suppose that*

$$X \mid Y \sim N_{p,n}(A + BY, \Sigma_1 \otimes \Phi), \tag{2.2}$$

$$Y \mid X \sim N_{q,n}(C + DX, \Sigma_2 \otimes \Phi), \tag{2.3}$$

where A: $p \times n$, B: $p \times q$, Σ_1: $p \times p$, Φ: $n \times n$, C: $q \times n$, D: $q \times p$, Σ_2: $q \times q$, $\Sigma_1 > 0$, $\Sigma_2 > 0$, $\Phi > 0$. *Then*

$$\Sigma_2 B' = D\Sigma_1 \tag{2.4}$$

and

$$Z = \begin{pmatrix} X \\ Y \end{pmatrix} \sim N_{p+q,n}\left(\begin{pmatrix} (I_p - BD)^{-1}(A + BC) \\ (I_q - DB)^{-1}(C + DA) \end{pmatrix}, \begin{pmatrix} (I_p - BD)^{-1}\Sigma_1 & (I_p - BD)^{-1}B\Sigma_2 \\ D(I_q - BD)^{-1}\Sigma_1 & (I_q - DB)^{-1}\Sigma_2 \end{pmatrix} \otimes \Phi\right) \tag{2.5}$$

Proof. Let $Y_0 = 0$. Then $X \mid Y = Y_0 \sim N_{p,n}(A, \Sigma_1 \otimes \Phi)$ Using Theorem 1.1 we conclude that Z has the distribution (2.1) if we replace B by $B^* = \Sigma_1 D' \Sigma_2^{-1}$ in formula (2.1). It follows from (2.1) that

$$E(X \mid Y) = B^*Y + A.$$

Comparing this with (2.2) we get $B^* = B$ which proves (2.5). Hence (2.5) is established. ■

Another type of characterization is given in the following theorem.

THEOREM 1.2. *Let* X: $p \times n$ *and* Y: $q \times n$ *be random matrices. Suppose that* $Y \mid X \sim N_{q,n}(C + DX, \Sigma_2 \otimes \Phi)$ *and* $X \sim N_{p,n}(F, \Sigma_1 \otimes \Phi)$, *where* C: $q \times n$, D: $q \times p$, Σ_2: $q \times q$, Φ: $n \times n$, F: $p \times n$, Σ_1: $p \times p$, $\Sigma_1 > 0$, $\Sigma_2 > 0$, $\Phi > 0$. *Then*

$$Z = \begin{pmatrix} X \\ Y \end{pmatrix} \sim N_{p+q,n}\left(\begin{pmatrix} F \\ DF + C \end{pmatrix}, \begin{pmatrix} \Sigma_1 & \Sigma_1 D' \\ D\Sigma_1 & \Sigma_2 + D\Sigma_1 D' \end{pmatrix} \otimes \Phi\right) \tag{2.6}$$

Proof. Let ϕ be the joint characteristic function of X and Y. For S: $p \times n$, T: $q \times n$, we have

$$\phi(S,T) = E(\mathrm{etr}(iS'X + iT'Y)) = E(\mathrm{etr}(iS'X)E_{Y \mid X}(\mathrm{etr}(iT'Y)))$$

$$= E(\mathrm{etr}(i(S' + T'D)X))\ \mathrm{etr}\left(iCT' - \frac{1}{2}\Sigma_2 T\Phi T'\right)$$

$$= \mathrm{etr}(i(S'F + T'(C + DF))$$

$$-\frac{1}{2}((T'(\Sigma_2 + D\Sigma_1 D')T + S'\Sigma_1 S + 2S'\Sigma_1 D'T)\Phi))$$

which proves (2.6). ■

The next theorem is an extension of a multivariate result of Khatri (1979) to matrices.

THEOREM 1.3. *Let* $X_1,\ldots,X_k$ *be random matrices of dimension* $p \times n$. *Assume that the random variables which are the elements of the matrix* $X = (X_1',\ldots,X_k')$ *are not linearly dependent. Suppose that the following conditions are satisfied:*

a) $E(X_i \mid X_1,\ldots,X_{i-1},X_{i+1},\ldots,X_k) = A_{i1}X_1 + \ldots + A_{i,i-1}X_{i-1} + A_{i,i+1}X_{i+1} + \ldots + A_{ik}X_k + B_i$, $i = 1,\ldots,k$, *where* A_{ij}, $i = 1,\ldots,k$, $j = 1,\ldots,k$, $j \neq i$, *are* $p \times p$ *and* B_i, $i = 1,\ldots,k$, *are* $p \times n$ *constant matrices,*

b) $X_1 \mid X_2,\ldots,X_k$ *depends on* $X_2,\ldots,X_k$ *only through* $E(X_1 \mid X_2,\ldots,X_k)$,

c) $\mathrm{Var}(\mathrm{vec}(X_2') \mid X_1,X_3,\ldots,X_k) = \Sigma \otimes \Phi$, *where* Σ: $p \times p$, Φ: $n \times n$, $\Sigma > 0$, $\Phi > 0$,

d) A_{12} *and* A_{21} *are nonsingular,*

e)
$$\begin{pmatrix} -I & A_{12} & A_{13} & \cdots & A_{1i} \\ A_{21} & -I & A_{23} & \cdots & A_{2i} \\ \vdots & & & & \\ A_{i1} & A_{i2} & A_{i3} & \cdots & -I \end{pmatrix} = A_{(i)} \text{ is nonsingular for } i = 1,2,\ldots,k,$$

f)
$$(A_{i1},A_{i2},\ldots,A_{i,i-1})A_{(i-1)}^{-1}\begin{pmatrix} A_{1i} \\ A_{2i} \\ \vdots \\ A_{i-1,i} \end{pmatrix} \text{ is nonsingular for } i = 2,3,\ldots,k,$$

g) *with the notation* $A_{ii} = -I_p$, $i = 1,\ldots,k$,

$$\begin{pmatrix} A_{11} & A_{13} & \cdots & A_{1k} \\ A_{21} & A_{23} & & A_{2k} \\ \vdots & & & \\ A_{i-1,1} & A_{i-1,3} & & A_{i-1,k} \\ A_{i+1,1} & A_{i+1,3} & & A_{i+1,k} \\ \vdots & & & \\ A_{k,1} & A_{k,3} & & A_{k,k} \end{pmatrix} = A^{(i)} \textit{ is nonsingular, } i = 1,2,\ldots,k,$$

h) $$I_p + (A_{21} \;\; A_{23} \;\; A_{24} \; \ldots \; A_{2k}) \cdot A^{(2)^{-1}} \begin{pmatrix} A_{12} \\ A_{32} \\ A_{42} \\ \vdots \\ A_{k2} \end{pmatrix} \textit{ is nonsingular.}$$

Then $X = (X_1' \;\; X_2' \;\; \ldots \;\; X_k')'$ *has a matrix variate normal distribution.*

<u>Proof</u>. Define

$$\tilde{X}_i = \mathrm{vec}(X_i'), \;\; i = 1,\ldots,k,$$

$$\tilde{U}_i = \mathrm{vec}(B_i'), \;\; i = 1,\ldots,k,$$

$$\tilde{A}_{ij} = \begin{cases} I_p \otimes I_n & \text{if } i = j \\ -A_{ij} \otimes I_n & \text{if } i \neq j, \end{cases}$$

$$\tilde{\Sigma} = \Sigma \otimes \Phi,$$

$$\tilde{Z}_i = \sum_{j=1}^{k} \tilde{A}_{ij}\tilde{X}_j + \tilde{U}_i. \qquad (2.7)$$

Then it is easy to see that $\tilde{X}_i$, $\tilde{U}_i$, $\tilde{Z}_i$, $\tilde{A}_{ij}$, $\tilde{\Sigma}$ satisfy the conditions of Theorem 2 of Khatri (1979). From that theorem we conclude that $\tilde{Z} = (\tilde{Z}_1' \;\; \tilde{Z}_2' \;\; \ldots \;\; \tilde{Z}_k')'$ has a multivariate normal distribution.

Let $\tilde{X} = (\tilde{X}_1' \ \ \tilde{X}_2' \ \ ... \ \ \tilde{X}_k')'$. Since $\tilde{X} = -(A_{(k)}^{-1} \otimes I_n)(\tilde{Z} - \tilde{U})$, $\tilde{X}$ is also multivariate normal and

$$\tilde{X} \sim N_{kpn}(\mu, \Sigma^*) \tag{2.8}$$

where $\mu = (\mu_1' \ \ \mu_2' \ \ ... \ \ \mu_k')'$. The only remaining thing to prove is that $\Sigma^* = S \otimes \Phi$ where S: $(kp) \times (kp)$, Φ: $n \times n$ because from this the theorem follows immediately.

Let us partition Σ into

$$\Sigma^* = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} & \cdots & \Sigma_{1k} \\ \Sigma_{21} & \Sigma_{22} & \cdots & \Sigma_{2k} \\ \vdots & & & \\ \Sigma_{k1} & \Sigma_{k2} & \cdots & \Sigma_{kk} \end{pmatrix}$$

where Σ_{ij}: $(pn) \times (pn)$, $i,j = 1,...,k$. From condition (a) we get

$$E(\tilde{X}_i \mid \tilde{X}_1,...,\tilde{X}_{i-1},\tilde{X}_{i+1},...,\tilde{X}_k) = B_i \otimes I_n + (A_{i1},...,A_{i,i-1},A_{i,i+1},...,A_{ik}) \otimes I_n) \begin{pmatrix} \tilde{X}_1 \\ \vdots \\ \tilde{X}_{i-1} \\ \tilde{X}_{i-1} \\ \vdots \\ \tilde{X}_k \end{pmatrix},$$

$$i = 1,2,...,k. \tag{2.9}$$

From (2.8) we get

$$E(\tilde{X}_i \mid \tilde{X}_1,...,\tilde{X}_{i-1},\tilde{X}_{i+1},...,\tilde{X}_k) = \mu_i^* + (\Sigma_{i1},...,\Sigma_{i,i-1},\Sigma_{i,i+1},...,\Sigma_{ik})$$

$$\cdot \begin{pmatrix} \Sigma_{11},\ldots,\Sigma_{1,i-1},\Sigma_{1,i+1},\ldots,\Sigma_{1k} \\ \vdots \\ \Sigma_{i-1,1},\ldots,\Sigma_{i-1,i-1},\Sigma_{i-1,i+1},\ldots,\Sigma_{i-1,k} \\ \Sigma_{i+1,1},\ldots,\Sigma_{i+1,i-1},\Sigma_{i+1,i+1},\ldots,\Sigma_{i+1,k} \\ \vdots \\ \Sigma_{k1},\ldots,\Sigma_{k,i-1},\Sigma_{k,i+1},\ldots,\Sigma_{kk} \end{pmatrix}^{-1} \begin{pmatrix} \tilde{X}_1 \\ \vdots \\ \tilde{X}_{i-1} \\ \tilde{X}_{i+1} \\ \vdots \\ \tilde{X}_k \end{pmatrix}, \quad i = 1,2,\ldots,k. \tag{2.10}$$

Comparing the coefficient matrix of $(\tilde{X}_1' \ldots \tilde{X}_{i-1}' \tilde{X}_{i+1}' \ldots \tilde{X}_k')'$ in (2.9) and (2.10) we find $\Sigma_{ij} = \sum_{\substack{\ell=1 \\ \ell \neq i}}^{k} (A_{i\ell} \otimes I_n)\Sigma_{\ell j}$, $i = 1,2,\ldots,k$, $j \neq i$ which can be rewritten using the definition $A_{ii} = -I_p$ as $0 = \sum_{\ell=1}^{k} (A_{i\ell} \otimes I_n)\Sigma_{\ell j}$, $i = 1,2,\ldots,k$, $j \neq i$. Hence

$$(-A_{i2} \otimes I_n)\Sigma_{2j} = \sum_{\substack{\ell=1 \\ \ell \neq 2}}^{k} (A_{i\ell} \otimes I_n)\Sigma_{\ell j}, \quad i = 1,2,\ldots,k, \quad j \neq i. \tag{2.11}$$

Let us fix j, then from (2.11) we get

$$(-p^{(j)} \otimes I_n)\Sigma_{2j} = (A^{(j)} \otimes I_n) \begin{pmatrix} \Sigma_{1j} \\ \Sigma_{3j} \\ \vdots \\ \Sigma_{kj} \end{pmatrix}$$

where $p^{(j)} = \begin{pmatrix} A_{12} \\ A_{22} \\ \vdots \\ A_{j-1,2} \\ A_{j+1,2} \\ \vdots \\ A_{k,2} \end{pmatrix}$. Thus

$$\begin{pmatrix} \Sigma_{1j} \\ \Sigma_{3j} \\ \vdots \\ \Sigma_{kj} \end{pmatrix} = ((-A^{(j)^{-1}} p^{(j)}) \otimes I_n)\Sigma_{2j}, \ j = 1,2,\ldots,k. \qquad (2.12)$$

Let $j = 2$, then we have

$$\begin{pmatrix} \Sigma_{12} \\ \Sigma_{32} \\ \vdots \\ \Sigma_{k2} \end{pmatrix} = ((-A^{(2)^{-1}} p^{(2)}) \otimes I_n)\Sigma_{22}. \qquad (2.13)$$

Let $\tilde{\Sigma}_{21} = (\Sigma_{21} \ \Sigma_{23} \ \Sigma_{23} \ \ldots \ \Sigma_{2k})$, $\tilde{\Sigma}_{12} = \tilde{\Sigma}_{23}'$ and

$$\tilde{\Sigma}_{11} = \begin{pmatrix} \Sigma_{11} & \Sigma_{13} & \Sigma_{14} & \ldots & \Sigma_{1k} \\ \Sigma_{31} & \Sigma_{33} & \Sigma_{34} & \ldots & \Sigma_{3k} \\ \Sigma_{41} & \Sigma_{43} & \Sigma_{44} & \ldots & \Sigma_{4k} \\ \vdots & & & & \\ \Sigma_{k1} & \Sigma_{k3} & \Sigma_{k4} & \ldots & \Sigma_{kk} \end{pmatrix}.$$

With this notation (2.13) can be written as

$$\tilde{\Sigma}_{12} = ((-A^{(2)^{-1}} p^{(2)}) \otimes I_n)\Sigma_{22}. \qquad (2.14)$$

If we take $i = 2$ in (2.9) and (2.10) and equate the coefficient matrices of

$(\tilde{X}_1' \;\; \tilde{X}_3' \; \dots \; \tilde{X}_k')$ in the two equations we get

$$\tilde{\Sigma}_{21}\tilde{\Sigma}_{11}^{-1} = Q_2 \otimes I_n, \tag{2.15}$$

where $Q_2 = (A_{21} \;\; A_{23} \;\; A_{24} \dots \;\; A_{2k})$.

Now using (2.14) and (2.15) we get

$$\begin{aligned} \mathrm{Var}(\tilde{X}_2 \mid \tilde{X}_1, \tilde{X}_3, \dots, \tilde{X}_k) &= \Sigma_{22} - \tilde{\Sigma}_{21}\tilde{\Sigma}_{11}^{-1}\tilde{\Sigma}_{12} \\ &= \Sigma_{22} - (Q_2 \otimes I_n)((-A^{(2)^{-1}} p^{(2)}) \otimes I_n)\Sigma_{22} \\ &= ((I_p + Q_2 A^{(2)^{-1}} p^{(2)}) \otimes I_n)\Sigma_{22}. \end{aligned} \tag{2.16}$$

Using condition (c) and (2.16), we get

$$\Sigma_{22} = ((I_p + Q_2 A^{(2)^{-1}} p^{(2)})^{-1} \Sigma) \otimes \Phi.$$

Therefore Σ_{22} can be written as

$$\Sigma_{22} = S_{22} \otimes \Phi. \tag{2.17}$$

From (2.13) and (2.17) we get

$$\begin{pmatrix} \Sigma_{12} \\ \Sigma_{32} \\ \vdots \\ \Sigma_{k2} \end{pmatrix} = ((-A^{(2)^{-1}} p^{(2)}) \otimes I_n)(S_{22} \otimes \Phi) = (-A^{(2)^{-1}} p^{(2)} S_{22}) \otimes \Phi.$$

Therefore

$$\Sigma_{i2} = S_{i2} \otimes \Phi, \;\; i = 1,2,\dots,k. \tag{2.18}$$

Hence

$$\Sigma_{2i} = S'_{i2} \otimes \Phi, \ i = 1,2,\ldots,k. \tag{2.19}$$

From (2.12) and (2.19) we obtain

$$\begin{pmatrix} \Sigma_{1j} \\ \Sigma_{3j} \\ \vdots \\ \Sigma_{kj} \end{pmatrix} = ((-A^{(j)^{-1}} p^{(j)}) \otimes I_n)(S'_{i2} \otimes \Phi) = (-A^{(j)^{-1}} p^{(j)} S_{22}) \otimes \Phi, \ j = 1,3,\ldots,k.$$

Therefore

$$\Sigma_{ij} = S_{ij} \otimes \Phi, \ i = 1,2,\ldots,k, \ j = 1,2,\ldots,k.$$

Hence $\Sigma^* = S \otimes \Phi$ which completes the proof. ■

If in Theorem 1.3 we consider only two matrices, the conditions of the theorem can be weakened as the following result shows.

__THEOREM 1.4__. *Let* X *and* Y *be random matrices of dimension* $p \times n$. *Assume that the random variables which are the elements of the matrix* $Z = \begin{pmatrix} X \\ Y \end{pmatrix}$ *are not linearly dependent. Suppose that the following conditions are satisfied:*

a) $$E(X \mid Y) = A + BY, \tag{2.20}$$

$$E(Y \mid X) = C + DX, \tag{2.21}$$

where B, D *are* $p \times p$ *and* A, C *are* $p \times n$ *matrices,*

b) $X \mid Y$ *depends on* Y *only through* $E(X \mid Y)$,

c) $$\mathrm{Var}(\mathrm{vec}(Y') \mid X) = \Sigma \otimes \Phi \tag{2.22}$$

where Σ: $p \times p$, Φ: $n \times n$, $\Sigma > 0$, $\Phi > 0$

d) B *and* D *are nonsingular*

e) $\begin{pmatrix} -I_p & B \\ D & -I_p \end{pmatrix}$ *is nonsingular.*

Let us define $\Sigma_1 = B\Sigma D'^{-1}$ *and* $\Sigma_2 = \Sigma$. *Then* $\Sigma_1 > 0$ *and*

$$Z = \begin{pmatrix} X \\ Y \end{pmatrix} \sim N_{2p,n}\left(\begin{pmatrix} (I_p - BD)^{-1}(A + BC) \\ (I_q - DB)^{-1}(C + DA) \end{pmatrix}, \begin{pmatrix} (I_p - BD)^{-1}\Sigma_1 & (I_p - BD)^{-1}B\Sigma_2 \\ D(I_p - BD)^{-1}\Sigma_1 & (I_q - DB)^{-1}\Sigma_2 \end{pmatrix} \otimes \Phi\right). \tag{2.23}$$

Proof: First we use Theorem 1.3 with $k = 2$, $X_1 = X$, $X_2 = Y$, $A_{12} = B$, $A_{21} = D$, $B_1 = A$, $B_2 = C$. It is easy to see that the conditions (a)-(h) of that theorem are satisfied. Therefore

$$Z = \begin{pmatrix} X \\ Y \end{pmatrix} \sim N_{2p,n}\left(\begin{pmatrix} M_1 \\ M_2 \end{pmatrix}, \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \otimes \Phi\right) \tag{2.24}$$

where $S_{11}, S_{12}, S_{21}, S_{22}$ are $p \times p$ and M_1, M_2 are $p \times n$ matrices. Then from (2.24) we get

$$E(X \mid Y) = M_1 - S_{12}S_{22}^{-1}M_2 + S_{12}S_{22}^{-1}Y$$

and

$$E(Y \mid X) = M_2 - S_{21}S_{11}^{-1}M_1 + S_{21}S_{11}^{-1}X.$$

Then using (2.20) and (2.21) we get $B = S_{12}S_{22}^{-1}$ and $D = S_{21}S_{11}^{-1}$. Now

$$Var(vec(X') \mid Y) = S_{11\cdot 2} \otimes \Phi$$

and

$$Var(vec(Y') \mid X) = S_{22\cdot 1} \otimes \Phi.$$

Then using (2.22) we get $S_{22 \cdot 1} = \Sigma$. But

$$S_{11 \cdot 2}(S_{11}^{-1}S_{12}) = (S_{12}S_{22}^{-1})S_{22 \cdot 1},$$

therefore $S_{11 \cdot 2} = B \cdot \Sigma \cdot D'^{-1}$. Hence

$$X \mid Y \sim N_{p,n}(A + BY, \Sigma_1 \otimes \Phi)$$

$$Y \mid X \sim N_{p,n}(C + DX, \Sigma_2 \otimes \Phi).$$

An application of Corollary 1.1.1 completes the proof. ■

REFERENCES

Brucker, J. (1979). A note on the bivariate normal distribution. *Communications in Statistics-Theor. Meth.*, A8(2), 175-177.

Fraser, D. A. S. and Streit, F. (1980). A further note on the bivariate normal distribution. *Communications in Statistics - Theor. Meth.*, A9(10), 1097-1099.

Khatri, C. G. and Rao, C. Radhakrishna (1976). Characterizations of Multivariate Normality. I. Through Independence of Some Statistics. *Journal of Multivariate Analysis*, **6**, 81-94.

Khatri, C. G. (1979). Characterizations of Multivariate Normality. II. Through Linear Regressions. *Journal of Multivariate Analysis*, **9**, 589-598.