JOURNAL OF MULTIVARIATE ANALYSIS **60**, 148–153 (1997)
ARTICLE NO. MV961649

A Note on Matrix Variate Normal Distribution

Truc T. Nguyen

Bowling Green State University

A characterization of the matrix variate normal distribution having identically distributed row vectors based on conditional normality is given. © 1997 Academic Press

1. INTRODUCTION AND BASIC RESULTS

Let X_1 and X_2 be two identically distributed random variables. Suppose that $X_1 \mid X_2 = x_2$ has a $N(ax_2 + b, \sigma^2)$ distribution for all $x_2 \in R$, where $-\infty < a, b < \infty$, $\sigma^2 > 0$. Ahsanullah (1985) showed that these conditions imply $|a| < 1$, X_1, and X_2 have a common normal distribution with mean $b/(1-a)$ and variance $\sigma^2/(1-a^2)$, and the joint distribution of X_1 and X_2 is a bivariate normal distribution with covariance $a\sigma^2/(1-a^2)$. Castillo and Galambos (1989) presented a unified extension of the characterizations of the bivariate normal distribution given by Brucker (1979) and Ahsanulla (1985). In his paper, Ahsanullah also proposed a conjecture on a multi-dimensional version of his result. Hamedani (1988), then Arnold and Pourahmadi (1988), gave counterexamples to this conjecture, and they also gave different characterizations for multivariate normal distribution based on conditional normality. For the details of these results, see the survey paper of Hamedani (1992). In this note we give other multidimensional versions of the result of Ahsanullah (1985), then apply these results to the characterization of a matrix variate normal distribution with identically distributed row vectors. Our technique is similar to the technique used by Ahsanullah in the bivariate case, combining it with some well known results on matrices and linear transformations on a real Euclidean space R^n.

The basic results on matrices and linear transformations used in the proof of Theorem 2.1 can be found in Halmos (1974) or Young and

Received July 14, 1993; revised September 12, 1996.

AMS 1992 subject classification: 62H05.

Key words and phrases: characteristic function, identically distributed, conditional distribution, norm of a matrix, linear transformation, vec notation.

0047-259X/97 $25.00

Gregory (1972). The vector norm we use here is the Euclidean norm, $\|x\| = (\sum_{i=1}^{n} x_i^2)^{1/2}$, and the matrix norm

$$\|A\| = \sup\{\|A\mathbf{x}\| : x \in R, \|x\| = 1\} = \sqrt{\lambda_1}, \tag{1.1}$$

where $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n \geqslant 0$ are the eigenvalues of $A'A$. In R^n under the topology corresponding to the metric $d(\mathbf{x}, y) = \|x - y\|$ the set $S = \{\mathbf{x} : \mathbf{x} \in R^n, \|x\| = 1\}$ is a compact set. Consequently, for every $n \times n$ matrix A, $\|A\mathbf{x}\|$ is a real valued continuous function on S. Hence, if A is a nonsingular matrix, there exist $\mathbf{s}_0$ and $\mathbf{s}_1$ of S such that

$$\begin{aligned} 0 < c_0 = \|A\mathbf{s}_0\| &= \inf\{\|A\mathbf{s}\| : \mathbf{s} \in S\} \leqslant \|A\mathbf{s}\| \leqslant \sup\{\|A\mathbf{s}\| : \mathbf{s} \in S\} \\ &= \|A\| = \|As_1\| = \sqrt{\lambda_1}, \qquad \forall s \in S. \end{aligned} \tag{1.2}$$

2. A CHARACTERIZATION OF MULTIVARIATE NORMAL DISTRIBUTION

Arnold and Pourahmadi (1988) and Ahsanullah and Sinha (1986) independently presented the extension of the result of Ahsanullah (1985) in the bivariate case to the multivariate case as follows.

Suppose that the random vector $\mathbf{X} = (X_1, ..., X_k)$ possesses a nonsingular distribution and suppose that the components of $\mathbf{X}$ have an exchangeable distribution and that $X_1 \mid X_2 = x_2, ..., X_k = x_k$ has a $N(\alpha_0 + \sum_{j=1}^{k} \alpha_j x_j, \sigma_0^2)$ distribution. Then X has a k-variate normal distribution.

In this section we give another characterization of multivariate normal distribution based on conditional normality. It is also a multivariate version of the result of Ahsanullah (1985). This result will be used to characterize a matrix variate normal distribution in Section 3.

THEOREM 2.1. *Let* $\mathbf{X}$ *and* $\mathbf{Y}$ *be two* $n \times 1$ *identically distributed random vectors. Suppose that* $\mathbf{X} \mid \mathbf{Y} = \mathbf{y} \sim N(A'\mathbf{y} + \mathbf{b}, \Sigma_0)$ *for all* $\mathbf{y} \in R^n$, *where* A *is an* $n \times n$ *matrix,* $\mathbf{b} \in R^n$, Σ_0 *is a* $n \times n$ *positive definite covariance matrix, and* A, $\mathbf{b}$, Σ_0 *do not depend on* $\mathbf{y}$. *Then*

(a) $\rho(A) < 1$ *where* $\rho(A)$ *is the spectral radius of* A.

(b) $\mathbf{X}$ *and* $\mathbf{Y}$ *have a common n-variate normal distribution with mean vector* $\mu = \lim_{k \to \infty} \sum_{j=0}^{k-1} (A^j)' \mathbf{b}$ *and covariance matrix* $\Sigma_{11} = \lim_{k \to \infty} \sum_{j=0}^{k-1} (A^j)' \Sigma_0 A^j$, *and the joint distribution of* X *and* Y *is a* $2n$*-variate normal distribution with mean vector* $(\mu', \mu')'$ *and covariance matrix*

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{11} \end{bmatrix},$$

where $\Sigma_{12} = \Sigma_{21}' = A'\Sigma_{11}$.

Proof. From $\mathbf{X} \mid \mathbf{Y} = \mathbf{y} \sim N(A'\mathbf{y} + \mathbf{b}, \Sigma_0)$,

$$\phi_{\mathbf{X},\,\mathbf{Y}}(\mathbf{s}, \mathbf{t}) = \phi(\mathbf{t} + A\mathbf{s}) \exp\{i\mathbf{b}'\mathbf{s} - \tfrac{1}{2}\mathbf{s}'\Sigma_0\mathbf{s}\}, \tag{2.1}$$

$\forall \mathbf{s}, \mathbf{t} \in R^n$, where ϕ is the common characteristic function of X and Y. Hence,

$$\phi(\mathbf{s}) = \phi(A\mathbf{s}) \exp\{i\mathbf{b}'\mathbf{s} - \tfrac{1}{2}s'\Sigma_0\mathbf{s}\}, \qquad \forall \mathbf{s} \in R^n. \tag{2.2}$$

Iterating (2.2) k times, $k = 0, 1, 2, \ldots$, we obtain

$$\phi(s) = \phi(A^k\mathbf{s}) \exp\left\{i\mathbf{b}' \sum_{j=0}^{k-1} A^j\mathbf{s} - \tfrac{1}{2}\mathbf{s}'\left(\sum_{j=0}^{k-1} (A^j)' \,\Sigma_0 A^j\right)\mathbf{s}\right\}. \tag{2.3}$$

Let $k \to \infty$. The limit of the right side of (2.3) always exists and is the characteristic function $\phi(\mathbf{s})$. A necessary and sufficient condition that $\lim_{k \to \infty} s' \,(\sum_{j=0}^{k-1} (A^j)' \,\Sigma_0 A^j)\, s$ exists for all $s \in R^n$ is that $\rho(A) < 1$. In this case,

$$\lim_{k \to \infty} \phi(A^k\mathbf{s}) = \phi(0) = 1,$$

$$\lim_{k \to \infty} \sum_{j=0}^{k-1} (A^j)' \,\mathbf{b} = ((I - A')^{-1})' \,\mathbf{b} = \mu, \tag{2.4}$$

$$\lim \sum_{j=0}^{k-1} (A^j)' \,\Sigma_0 A^j = \Sigma_{11}, \tag{2.5}$$

and

$$\phi(\mathbf{s}) = \exp\{i\mu'\mathbf{s} - \tfrac{1}{2}\mathbf{s}'\Sigma_{11}\mathbf{s}\}, \qquad \forall \mathbf{s} \in R^n. \tag{2.6}$$

Hence, X and Y have an n-variate normal distribution with mean vector and covariance matrix given by (2.4) and (2.5), respectively. The joint characteristic function of $\mathbf{X}$ and $\mathbf{Y}$ is obtained by substituting $\phi(\mathbf{s})$ of (2.6) to (2.1). From the fact that

$$\mu = \lim_{k \to \infty} \sum_{j=0}^{k-1} (A^j)' \,\mathbf{b} = \lim_{k \to \infty} \left(\mathbf{b} + A' \sum_{j=0}^{k-1} (A^j)' \,\mathbf{b}\right) = \mathbf{b} + A'\mu,$$

and

$$\begin{aligned}\Sigma_{11} &= \lim_{k \to \infty} \sum_{j=0}^{k-1} (A^j)' \,\Sigma_0 A^j \\ &= \lim_{k \to \infty} \left(\Sigma_0 + A' \left(\sum_{j=0}^{k-1} (A^j)' \,\Sigma_0 A^j\right) A\right) = \Sigma_0 + A'\Sigma_{11} A,\end{aligned}$$

then

$$\phi_{X,\,Y}(\mathbf{s},\mathbf{t})=\exp\{i\mu'\mathbf{s}+i\mu'\mathbf{t}-\tfrac{1}{2}[\mathbf{s}'\Sigma_{11}\mathbf{s}+\mathbf{t}'\Sigma_{11}\mathbf{t}+\mathbf{s}'A'\Sigma_{11}\mathbf{t}+\mathbf{t}'\Sigma_{11}A\mathbf{s}]\}, \tag{2.7}$$

$\forall \mathbf{s}, \mathbf{t}\in R^n$. Then $\mathbf{X}$ and $\mathbf{Y}$ have a joint $2n$-variate normal distribution with covariance matrix given by

$$\Sigma=\begin{bmatrix}\Sigma_{11} & \Sigma_{12}\\ \Sigma_{21} & \Sigma_{11}\end{bmatrix},$$

where $\Sigma_{12}=\Sigma'_{21}=A'\Sigma_{11}$.

3. A CHARACTERIZATION OF MATRIX VARIATE NORMAL DISTRIBUTION

Let M be a random matrix of size $k\times n$, where the row vectors are $\mathbf{X}'_i=(X_{i1},...,X_{in})$, $i=1,...,k$. The notation $\text{vec}(M')=(\mathbf{X}'_1,...,\mathbf{X}'_k)'$ is a $kn\times 1$ random vector. M is defined to have a $k\times n$ matrix variate normal distribution if $\text{vec}(M')$ has a kn variate normal distribution with mean vector $\mu=(\mu'_1,...,\mu'_k)'$, where $\mu_i=E[X_i]$, $i=1,...,k$, and the covariance matrix is partitioned into $k^2 n\times n$ matrices $(\Sigma_{i,j})$, where $\Sigma_{i,j}$ is the covariance matrix between $\mathbf{X}_i$ and $\mathbf{X}_j$, $i,j=1,...,k$.

The following theorem is a characterization of a matrix variate normal distribution having identically distributed row vectors.

THEOREM 3.1. *Let M be a random matrix with k rows and n columns, where $\mathbf{X}'_1,...,\mathbf{X}'_k$ are the row vectors of M and let $M^{(1)}$ be the submatrix of size $1\times n$ containing the first row and $M^{(2)}$ be the submatrix of size $(k-1)\times n$ containing the last $(k-1)$ rows of M. Suppose that $\mathbf{X}_1,...,\mathbf{X}_k$ are identically distributed. If* $\text{vec}(M^{(2)'})\mid\text{vec}(M^{(1)'})=\mathbf{x}_1$ *has a $(k-1)\,n$ variate normal distribution with mean vector $A'\mathbf{x}_1+\mathbf{b}$ and covariance matrix a constant positive definite matrix $\Sigma_0=(\Sigma^0_{ij})$, $i,j=2,...,k$, where A' is a $(k-1)\,n\times n$ matrix of constants and $\mathbf{b}=(\mathbf{b}'_2,...,\mathbf{b}'_k)'$ is a $(k-1)\,n\times 1$ constant vector, then*

(a) $\max\{\rho(A_i), i=2,...,k\}<1$, *where $A'_2,...,A'_k$ are submatrices of A containing the first n rows, the second n rows, ..., the last n rows of A', respectively.*

(b) $\mu=\lim_{k\to\infty}\sum_{j=0}^{k-1}(A_i^j)'\,\mathbf{b}_i$ *all are equal for $i=2,...,k$, and,*

$$\Sigma_{11}=\lim_{k\to\infty}\sum_{j=0}^{k-1}(A_i^j)'\,\Sigma^0_{ii}A_i^j \text{ all are equal for } i=2,...,k,$$

and the distribution of each vector $\mathbf{X}_i$, $i=1,...,k$, is an n-variate normal distribution with covariance matrix Σ_{11} and with mean vector μ.

(c) *The random matrix M has a $k \times n$ matrix variate normal distribution with* $\operatorname{vec}(M')$ *having mean vector* $(\mu', ..., \mu')'$ *and covariance matrix*

$$\Sigma = \begin{pmatrix} \Sigma_{(1)} & \Sigma_{(1,2)} \\ \Sigma_{(2,1)} & \Sigma_{(2)} \end{pmatrix},$$

where $\Sigma_{(1)} = \Sigma_{11}$, $\Sigma_{(2)} = A'\Sigma_{11}A + \Sigma_0$.

(d) $\operatorname{cov}(\mathbf{X}_1, \mathbf{X}_i) = A_i'\Sigma_{11}$, $\operatorname{cov}(\mathbf{X}_i\mathbf{X}_j) = A_i'\Sigma_{11}A_j + \Sigma_{ij}^0$, $i, j = 2, ..., n$.

Proof. The proof is similar to the proof of Theorem 2.1, using the result of Theorem 2.2. The joint characteristic function of $\operatorname{vec}(M^{(1)'})$ and $\operatorname{vec}(M^{(2)'})$ is

$$\phi_{\operatorname{vec}(M')}(\mathbf{s}, \mathbf{t}) = \phi(\mathbf{s} + A\mathbf{t}) \exp(i\mathbf{b}'\mathbf{t} - \tfrac{1}{2}\mathbf{t}'\Sigma_0\mathbf{t}), \tag{3.1}$$

$\forall \mathbf{s} \in R^n$, $\mathbf{t} \in R^{(k-1)n}$. For every $i = 2, ..., n$, $\mathbf{X}_i \mid \mathbf{X}_{(1)} = \mathbf{x}_1 \sim N(A_i'\mathbf{x}_1 + \mathbf{b}_i, \Sigma_{ii}^0)$, then by Theorem 2.2, $\|\rho(A_i)\| < 1$, $\mathbf{X}_1$, and $\mathbf{X}_i$ have a n-variate normal distribution with mean vector $\mu = \mu_i' = \lim_{k \to \infty} \sum_{j=0}^{k} A_i^j \mathbf{b}_i$ and covariance matrix $\Sigma_{11} = \lim_{k \to \infty} \sum_{j=0}^{k-1} (A_i^j)' \Sigma_{ii}^0 A_i^j$, and the covariance matrix between $\mathbf{X}_1$ and $\mathbf{X}_i$ is $\Sigma_{1i} = \Sigma_{i1}' = A_i'\Sigma_{11}$. Then (a) and (b) are obtained. The characteristic function of each $\mathbf{X}_1, ..., \mathbf{X}_k$ is $\phi(\mathbf{s}) = \exp\{i\mu'\mathbf{s} - \tfrac{1}{2}\mathbf{s}'\Sigma_{11}\mathbf{s}\}$. Substitute this characteristic function in (3.1),

$$\begin{aligned}\phi_{\operatorname{vec}(M')}(\mathbf{s}, \mathbf{t}) = \exp\{&i\mu'\mathbf{s} + i(A'\mu + \mathbf{b})'\,\mathbf{t} - \tfrac{1}{2}\mathbf{s}'\Sigma_{11}\mathbf{s} + \mathbf{t}'(A'\Sigma_{11}A + \Sigma_0)\,\mathbf{t}] \\ &+ \mathbf{s}'\Sigma_{11}A\mathbf{t} + \mathbf{t}'A'\Sigma_{11}\mathbf{s}\}, \qquad \forall \mathbf{s} \in R^n, \quad \mathbf{t} \in R^{(k-1)n}.\end{aligned} \tag{3.2}$$

Then M has a $k \times n$ matrix variate normal distribution with covariance matrix

$$\Sigma = \begin{pmatrix} \Sigma_{(1)} & \Sigma_{(12)} \\ \Sigma_{(21)} & \Sigma_{(2)} \end{pmatrix}$$

where

$$\Sigma_{(1)} = \Sigma_{11}, \Sigma_{(2)} = A'\Sigma_{11}A + \Sigma_0, \Sigma_{(12)} = \Sigma_{(21)}' = \Sigma_{11}A, \Sigma_{(12)}$$

is the covariance matrix between $\operatorname{vec}(M^{(1)'})$ and $\operatorname{vec}(M^{(2)'})$, $\Sigma_{(2)}$ is the covariance matrix of $\operatorname{vec}(M^{(2)'})$.

From (3.2), $\operatorname{cov}(X_i, X_j) = A_i'\Sigma_{11}A_j + \Sigma_{1j}^0$, so that part (c) of Theorem 3.1 is then proved.

Note that from (a), and also from (b), we have $\Sigma_{11} = A_i'\Sigma_{11}A_i + \Sigma_{ii}^0$ for all $i = 2, ..., k$.

The following result for characterizing a k-variate normal distribution, is a special case of Theorem 3.1 when $n = 1$.

COROLLARY 3.1. *Let* $\mathbf{X}=(X_1, ..., X_k)'$ *be a random vector having identically distributed components, and let* $\mathbf{X}^{(1)}=X_1$, $\mathbf{X}^{(2)}=(X_2, ..., X_k)'$. *Suppose that* $\mathbf{X}^{(2)} \mid \mathbf{X}^{(1)}=x_1$ *has a* $(k-1)$ *variate normal distribution with mean vector* $\mathbf{a}x_1+\mathbf{b}$ *and covariance matrix* Σ_0 *for all* $x_1 \in R$, *where* $\mathbf{a}=(a_2, ..., a_k)'$, $\mathbf{b}=(b_2, ..., b_k)' \in R^{k-1}$, $\Sigma_0=(\sigma^0_{ij})$, $i, j=2, ..., k$ *is a* $(k-1)\times(k-1)$ *positive definite covariance matrix of constants. Then*

(a) $\max\{|a_i|, i=2, ..., k\}<1, \qquad \mu_0=b_2/(1-a_2)=\cdots=b_k/(1-a_k),$
and

$$\sigma_0^2=\sigma^0_{2,2}/(1-a_2^2=\cdots=\sigma^0_{k,k}/(1-a_k^2).$$

(b) *The distribution of* $X_1, ..., X_k$ *is a normal distribution with mean* μ_0 *and variance* σ_0^2, *and the distribution of* $\mathbf{X}$ *is a k-variate normal distribution with covariance matrix*

$$\Sigma=\begin{pmatrix}\Sigma_{11} & \Sigma_{12}\\ \Sigma_{21} & \Sigma_{22}\end{pmatrix},$$

where $\Sigma_{11}=\sigma^2_{0'}$, $\Sigma_{22}=\sigma_0^2\mathbf{aa}'+\Sigma_0$, $\Sigma'_{12}=\Sigma_{21}=\sigma_0^2\mathbf{a}$.

ACKNOWLEDGMENT

I am grateful to the referees for their valuable comments which led to a substantial improvement in the presentation of the paper.

REFERENCES

[1] Ahsanullah, M. (1985). Some characterizations of the bivariate normal distribution. *Metrika* **32** 215–218.

[2] Ahsanullah, M., and Sinha, B. K. (1986). On normality via conditional normality. *Calcutta Statist. Assoc. Bull* **35** 139–140.

[3] Arnold, B. C, and Pourahmadi, M. (1988). Conditional characterizations of multivariate distributions. *Metrika* **35** 99–108.

[4] Brucker, J. (1979). A note on the bivariate normal distribution. *Comm. Statist. A* **8** 175–177.

[5] Castillo, E., and Galambos, J. (1989). Conditional distributions and the bivariate normal distribution. *Metrika* **36** 209–214.

[6] Halmos, P. R. (1974). *Finite Dimensional Vector Spaces*. Springer-Verlag, New York.

[7] Hamedani, G. G. (1988). On two recent characterizations of multivariate normal distribution. *Metrika* **35** 41–47.

[8] Hamedani, G. G. (1989). Bivariate and multivariate normal characterizations: A brief survey. *Comm. Statist. A* **21** 2665–2688.

[9] Young, D. M, and Gregory, R. T. (1973). *A Survey of Numerical Mathematics*. Addison–Wesley, Reading, MA.