MATRIX VARIATE θ-GENERALIZED NORMAL DISTRIBUTION

A. K. GUPTA AND T. VARGA

ABSTRACT. In this paper, the matrix variate θ-generalized normal distribution is introduced. Then its properties are studied. In particular, it is proved that this distribution has maximal entropy in a certain class of distributions.

1. INTRODUCTION

Goodman and Kotz (1973) introduced the family of the multivariate θ-generalized normal distributions and derived its basic properties. In the present paper, a definition of the matrix variate θ-generalized normal distribution is proposed. These distributions are very useful in robustness studies. It is shown that for matrices with one column, this definition is equivalent to the definition of the multivariate θ-generalized normal distribution. On the other hand, the matrix variate normal distribution is also a special case of this family of distributions. Properties of the matrix variate θ-generalized normal distribution are also studied.

In this paper, the following notations will be used. Let A be a $p \times n$ matrix; then A_{ij} denotes the (i, j)th element of A, $i = 1, \ldots, p$, $j = 1, \ldots, n$. Let B be a $p \times p$ nonsingular matrix; then B_{ij}^{-1} denotes the (i, j)th element of B^{-1}, $i, j = 1, \ldots, p$. Let the columns of matrix A $(p \times n)$ be denoted by $A^{(i)}$, $i = 1, \ldots, n$; then (see Gupta and Varga, 1993)

$$\text{vec}(A) = \begin{pmatrix} A^{(1)} \\ A^{(2)} \\ \vdots \\ A^{(n)} \end{pmatrix}.$$

2. BASIC RESULTS

Definition 2.1. For $\theta > 0$, $X = (X_{ij})$, $i, j = 1, \ldots, n$, has a matrix variate standard θ-generalized normal distribution if the X_{ij}'s are independent identically distributed random variables with probability density function

$$f(X_{ij}) = \left\{ 2\Gamma \left(1 + \frac{1}{\theta} \right) \right\}^{-1} e^{-|X_{ij}|^{\theta}}, \qquad X_{ij} \in \mathbb{R}, \ i, j = 1, \ldots, n.$$

Definition 2.2. For $\theta > 0$, $Y : p \times n$ is said to have a matrix variate θ-generalized normal distribution if Y can be written as $Y = AXB + M$ where X is a standard θ-generalized normal random matrix, $A : p \times p$, $B : n \times n$, and $M : p \times n$ are constant matrices, with A and B being nonsingular. The distribution of Y is denoted by $N_{p,n}(M, A, B, \theta)$.

For $n = 1$, we get the multivariate θ-generalized normal distribution as defined by Goodman and Kotz (1973). Furthermore the case $n = p = 1$ reduces to the Laplace density for $\theta = 1$, and the normal density for $\theta = 2$. It approaches the uniform density as $\theta \to \infty$, and an improper uniform one over the real line as $\theta \to 0$. The probability density function of a matrix variate θ-generalized normal distribution is given in the following theorem.

Theorem 2.1. Let $Y \sim N_{p,n}(M, A, B, \theta)$. Then the probability density function of Y is

$$(2.1) \quad f(Y) = k_{p,n}(A, B, \theta) \exp\left\{ -\sum_{i=1}^{p}\sum_{j=1}^{n}\left|\sum_{k=1}^{p}\sum_{\ell=1}^{n} A_{ik}^{-1}(Y_{k\ell} - M_{k\ell})B_{\ell j}^{-1}\right|^{\theta} \right\},$$

and $k_{p,n}(A, B, \theta) = (2\Gamma(1 + \frac{1}{\theta}))^{-pn}|A|^{-n}|B|^{-p}$.

Proof. If $X \sim N_{p,n}(0, I_p, I_n, \theta)$ then the p.d.f. of X is

$$g(X) = \left(2\Gamma\left(1 + \frac{1}{\theta}\right)\right)^{-pn} \exp\left\{ -\sum_{i=1}^{p}\sum_{j=1}^{n} |X_{ij}|^{\theta} \right\},$$

and from $Y = AXB + M$ it follows that (2.1) is true. $\square$

Linear transformations of matrices with matrix variate θ-generalized normal distribution also have matrix variate θ-generalized normal distribution. This is proved in the next theorem.

Theorem 2.2. Let $Y \sim N_{p,n}(M, A, B, \theta)$. Let C $(p \times p)$, D $(n \times n)$ be nonsingular matrices, L be a $p \times n$ matrix, and define $Z = CYD + L$. Then

$$(2.2) \quad Z \sim N_{p,n}(CMD + L, CA, BD, \theta).$$

Proof. Let $X \sim N_{p,n}(0, I_p, I_n, \theta)$ and $Y = AXB + M$. Then $Z = (CA) \times (BD) + (CMD + L)$, where CA and BD are nonsingular. From this (2.2) follows. $\square$

In multivariate analysis, the normal distribution plays a central role. By definition, the $p \times n$ random matrix Y is said to have a matrix variate normal distribution if its probability density function is

$$f(Y) = (2\pi)^{-\frac{pn}{2}}|\Sigma|^{-\frac{n}{2}}|\Phi|^{-\frac{p}{2}} \exp\left(\text{tr}\left(-\frac{1}{2}(Y - M)'\Sigma^{-1}(Y - M)\Phi^{-1}\right)\right),$$

where $\Sigma : p \times p$ and $\Phi : n \times n$ are positive definite matrices, and M is a $p \times n$ matrix. We denote this by $Y \sim N_{p,n}(M, \Sigma \otimes \Phi)$, e.g. see Gupta and Varga (1992). It is an important fact that the class of matrix variate θ-generalized normal distributions contains the matrix variate normal distribution as a special case. We can state this result more precisely in the following remark.

Remark 2.1. $N_{p,n}(M, A, B, 2) = N_{p,n}(M, \frac{1}{2}(AA') \otimes (B'B))$.

Indeed, let $Y \sim N_{p,n}(M, A, B, 2)$. Then $Y = AXB + M$, where $\sqrt{2}X \sim N_{p,n}(0, I_p \otimes I_n)$ from which the statement follows.

The relationship between matrix variate θ-generalized normal distributions and multivariate θ-generalized normal distributions is pointed out in the next theorem.

Theorem 2.3. $Y \sim N_{p,n}(M, A, B, \theta)$ *if and only if*

$$\widetilde{Y} = \text{vec}(Y') \sim N_{pn,1}(\text{vec}(M'), A \otimes B', 1, \theta).$$

Proof. $Y = AXB + M$ is equivalent to $\text{vec}(Y') = (A \otimes B')\text{vec}(X') + \text{vec}(M')$, from which the statement of the theorem follows. $\square$

The next theorem shows that the parameters of a matrix variate θ-generalized normal distribution are not uniquely determined.

Theorem 2.4. $N_{p,n}(M, A, B, \theta)$ *and* $N_{p,n}(M^*, A^*, B^*, \theta)$ *define the same distribution if and only if* $M = M^*$ *and*

(a) *in the case of* $\theta = 2$, *there exist* $G : p \times p$ *and* $H : n \times n$ *orthogonal matrices and* $c > 0$ *such that* $A^* = cAG$ *and* $B^* = \frac{1}{c}HB$;

(b) *in the case of* $\theta \neq 2$, *there exist* $P : p \times p$ *and* $Q : n \times n$ *signed permutation matrices and* $c > 0$ *such that* $A^* = cAP$ *and* $B^* = \frac{1}{c}QB$.

Proof. The sufficiency of the conditions is obvious. To prove the necessity of the conditions assume that

$$N_{pn,1}(\text{vec}(M'), A \otimes B', 1, \theta) \quad \text{and} \quad N_{pn,1}(\text{vec}(M^{*'}), A^* \otimes B^{*'}, 1, \theta)$$

define the same distribution. Since the first distribution is symmetric about $\text{vec}(M')$ and the second one about $\text{vec}(M^{*'})$, we must have $M = M^*$.

(a) If $\theta = 2$, we get

$$N_{pn}\left(\text{vec}(M'), \frac{1}{2}(AA') \otimes (B'B)\right) = N_{pn}\left(\text{vec}(M'), \frac{1}{2}(A^* A^{*'}) \otimes (B^{*'}B^*)\right).$$

Hence there exists $c^2 > 0$ such that $A^* A^{*'} = c^2 AA'$ and $B^{*'}B^* = \frac{1}{c^2}B'B$. But then we can find $G : p \times p$ and $H : n \times n$ orthogonal matrices such that $A^* = cAG$ and $B^* = \frac{1}{c}HB$ (see Muirhead (1982)).

(b) If $\theta \neq 2$ the result follows from Theorem 3 of Goodman and Kotz (1973). $\square$

The first four moments of a matrix variate θ-generalized normal distribution are derived next.

Theorem 2.5. *Let* $X \sim N_{p,n}(0, I_p, I_n, \theta)$. *Then* $\mathscr{E}(X_{ij}) = 0$,

$$\mathscr{E}(X_{i_1 j_1} X_{i_2 j_2}) = \frac{\Gamma(3/\theta)}{\Gamma(1/\theta)}\delta_{i_1,i_2}\delta_{j_1,j_2}, \qquad \mathscr{E}(X_{i_1 j_1} X_{i_2 j_2} X_{i_3 j_3}) = 0,$$

and

$$\mathscr{E}(X_{i_1 j_1} X_{i_2 j_2} X_{i_3 j_3} X_{i_4 j_4}) = \left(\frac{\Gamma(5/\theta)}{\Gamma(1/\theta)} - 3\frac{\Gamma^2(3/\theta)}{\Gamma^2(1/\theta)}\right)\Sigma_{i_1 i_2 i_3 i_4}\Sigma_{j_1 j_2 j_3 j_4}$$

$$+ \frac{\Gamma^2(3/\theta)}{\Gamma^2(1/\theta)}(\delta_{i_1 i_2}\delta_{j_1 j_2}\delta_{i_3 i_4}\delta_{j_3 j_4} + \delta_{i_1 i_3}\delta_{j_1 j_3}\delta_{i_2 i_4}\delta_{j_2 j_4} + \delta_{i_1 i_4}\delta_{j_1 j_4}\delta_{i_2 i_3}\delta_{j_2 j_3}),$$

where

$$\delta_{ab} = \begin{cases} 1 & if\ a = b\,, \\ 0 & if\ a \neq b \end{cases} \quad and \quad \Sigma_{abcd} = \begin{cases} 1 & if\ a = b = c = d\,, \\ 0 & otherwise. \end{cases}$$

Proof. If X_{ij} has the pdf

$$f(X_{ij}) = \left\{ 2\Gamma\left(1 + \frac{1}{\theta}\right)\right\}^{-1} e^{-|X_{ij}|^{\theta}}$$

and $k > 1$, then

$$\int_0^{\infty} X_{ij}^k f(X_{ij})\,dX_{ij} = \frac{\Gamma((k+1)/\theta)}{2\Gamma(1/\theta)}.$$

Then if k is a nonnegative integer, we have

(2.3) $$\mathscr{E}(X_{ij}^k) = \frac{1 + (-1)^k}{2} \cdot \frac{\Gamma((k+1)/\theta)}{\Gamma(1/\theta)}.$$

Using (2.3) and the fact that the elements of X are independent of one another we obtain the results of the theorem. $\quad\square$

Corollary 2.5.1. *Let* $Y \sim N_{p,n}(0, A, B, \theta)$. *Then* $\mathscr{E}(Y_{ij}) = 0$,

$$\mathscr{E}(Y_{i_1j_1} Y_{i_2j_2}) = \frac{\Gamma(3/\theta)}{\Gamma(1/\theta)} g_{i_1i_2} h_{j_1j_2}\,, \quad \mathscr{E}(Y_{i_1j_1} Y_{i_2j_2} Y_{i_3j_3}) = 0,$$

and

$$\mathscr{E}(Y_{i_1j_1} Y_{i_2j_2} Y_{i_3j_3} Y_{i_4j_4}) = \left(\frac{\Gamma(5/\theta)}{\Gamma(1/\theta)} - 3\frac{\Gamma^2(3/\theta)}{\Gamma^2(1/\theta)}\right) r_{i_1i_2i_3i_4} q_{j_1j_2j_3j_4}$$
$$+ \frac{\Gamma^2(3/\theta)}{\Gamma^2(1/\theta)} (g_{i_1i_2} h_{j_1j_2} g_{i_3i_4} h_{j_3j_4} + g_{i_1i_3} h_{j_1j_3} g_{i_2i_4} h_{j_2j_4} + g_{i_1i_4} h_{j_1j_4} g_{i_2i_3} h_{j_2j_3}),$$

where

$$g_{uv} = \sum_{k=1}^p A_{uk} A_{vk}\,, \quad h_{uv} = \sum_{\ell=1}^n B_{\ell u} B_{\ell v}\,, \quad r_{uvwt} = \sum_{k=1}^p A_{uk} A_{vk} A_{wk} A_{tk}\,,$$

and

$$q_{uvwt} = \sum_{\ell=1}^n B_{\ell u} B_{\ell v} B_{\ell w} B_{\ell t}.$$

Corollary 2.5.2. *Let* $Y \sim N_{p,n}(M, A, B, \theta)$. *Then* $\mathscr{E}(Y_{ij}) = M_{ij}$,

$$\mathscr{E}(Y_{i_1j_1} Y_{i_2j_2}) = \frac{\Gamma(3/\theta)}{\Gamma(1/\theta)} g_{i_1i_2} h_{j_1j_2} + M_{i_1j_1} M_{i_2j_2}\,,$$

$$\mathscr{E}(Y_{i_1j_1} Y_{i_2j_2} Y_{i_3j_3}) = \frac{\Gamma(3/\theta)}{\Gamma(1/\theta)} [g_{i_1i_2} h_{j_1j_2} M_{i_3j_3} + g_{i_1i_3} h_{j_1j_3} M_{i_2j_2} + g_{i_2i_3} h_{j_2j_3} M_{i_1j_1}]$$
$$+ M_{i_1j_1} M_{i_2j_2} M_{i_3j_3}\,,$$

and

$$\mathscr{E}(Y_{i_1 j_1} Y_{i_2 j_2} Y_{i_3 j_3} Y_{i_4 j_4}) = \left(\frac{\Gamma(5/\theta)}{\Gamma(1/\theta)} - 3\frac{\Gamma^2(3/\theta)}{\Gamma^2(1/\theta)}\right) r_{i_1 i_2 i_3 i_4} q_{j_1 j_2 j_3 j_4}$$

$$+ \frac{\Gamma^2(3/\theta)}{\Gamma^2(1/\theta)}(g_{i_1 i_2} h_{j_1 j_2} g_{i_3 i_4} h_{j_3 j_4} + g_{i_1 i_3} h_{j_1 j_3} g_{i_2 i_4} h_{j_2 j_4} + g_{i_1 i_4} h_{j_1 j_4} g_{i_2 i_3} h_{j_2 j_3})$$

$$+ \frac{\Gamma(3/\theta)}{\Gamma(1/\theta)}(M_{i_1 j_1} M_{i_2 j_2} g_{i_3 i_4} h_{j_3 j_4} + M_{i_1 j_1} M_{i_3 j_3} g_{i_2 i_4} h_{j_2 j_4} + M_{i_1 j_1} M_{i_4 j_4} g_{i_2 i_3} h_{j_2 j_4}$$

$$+ M_{i_2 j_2} M_{i_3 j_3} g_{i_1 i_4} h_{j_1 j_4} + M_{i_2 j_2} M_{i_4 j_4} g_{i_1 i_3} h_{j_1 j_3} + M_{i_3 j_3} M_{i_4 j_4} g_{i_1 i_2} h_{j_1 j_2})$$

$$+ M_{i_1 j_1} M_{i_2 j_2} M_{i_3 j_3} M_{i_4 j_4} ,$$

where the functions g, h, r, and q are defined in Corollary 2.5.1.

Corollary 2.5.3. *Let $X \sim N_{p,n}(M, A, B, \theta)$. Then $\mathscr{E}(X) = M$ and*

$$(2.4) \qquad \operatorname{Cov}(\operatorname{vec}(X')) = \frac{\Gamma(3/\theta)}{\Gamma(1/\theta)}(AA') \otimes (B'B).$$

Corollary 2.5.4. *Let $Y \sim N_{p,n}(M, A, B, \theta)$. Then*

$$\operatorname{Corr}(Y_{i_1 j_1}, Y_{i_2 j_2}) = \frac{g_{i_1 i_2} h_{j_1 j_2}}{\sqrt{g_{i_1 i_1} g_{i_2 i_2} h_{j_1 j_1} h_{j_2 j_2}}},$$

and hence the $\operatorname{Corr}(Y_{i_1 j_1}, Y_{i_2 j_2})$ does not depend on θ.

Using the expressions for the moments the following result can be derived.

Theorem 2.6. *Let $X \sim N_{p,n}(M, A, B, \theta)$ and $E : r \times p$, $C : n \times k$, $F : q \times p$, and $D : n \times \ell$ be constant matrices. Then EXC and FXD are uncorrelated if and only if either $C'B'BD = 0$ or $EAA'F' = 0$. Specifically, XC and XD are uncorrelated iff $C'B'BD = 0$, and EX and FX are uncorrelated iff $EAA'F' = 0$.*

Proof. Using (2.4) we get

$$\operatorname{Cov}(\operatorname{vec}(EXC)', \operatorname{vec}(FXD)') = \operatorname{Cov}((E \otimes C')\operatorname{vec}(X'), (F \otimes D')\operatorname{vec}(X'))$$

$$= (E \otimes C')\frac{\Gamma(3/\theta)}{\Gamma(1/\theta)}[(AA') \otimes (B'B)](F' \otimes D)$$

$$= \frac{\Gamma(3/\theta)}{\Gamma(1/\theta)}(EAA'F') \otimes (C'B'BD),$$

and the last expression equals zero iff $EAA'F' = 0$ or $C'B'BD = 0$. $\square$

3. CHARACTERIZATION RESULTS

The first theorem in this section shows that matrix variate θ-generalized normal distribution has maximal entropy in a certain class of distributions. It is an extension of Theorem 10 of Goodman and Kotz (1973).

Theorem 3.1. *Let $X : p \times n$ be a random matrix with p.d.f. f such that*

$$(3.1) \qquad \mathscr{E}\|AXB + M\|_\theta = c$$

where $A : p \times p$, $B : n \times n$ are nonsingular matrices, M is $p \times n$ matrix, c is a given scalar, and for a $p \times n$ matrix Y we define $\|Y\|_\theta = \sum_{i=1}^p \sum_{j=1}^n |y_{ij}|^\theta$.

Then the entropy of X, that is, $\mathscr{E}_X(-\log f(X))$, is maximized iff $X = Y$ a.e. where

(3.2) $$Y \sim N_{p,n}\left(-A^{-1}MB^{-1}, \left(\frac{\theta c}{pn}\right)^{\frac{1}{\theta}} A^{-1}, B^{-1}, \theta\right).$$

The maximal entropy is

(3.3) $$\frac{pn}{\theta}\left(1 + \log\left(\frac{\theta c}{pn}\right)\right) - \log k_{p,n}(A^{-1}, B^{-1}, \theta).$$

Proof. Let $g(Y)$ be the p.d.f. of Y defined by (3.2). Then

$$\mathscr{E}\|AYB + M\|_\theta = \int_{\mathbb{R}^{p\times n}} \|AYB + M\|_\theta k_{p,n}\left(\left(\frac{\theta c}{pn}\right)^{\frac{1}{\theta}} A^{-1}, B^{-1}, \theta\right)$$

$$\cdot \exp\left\{-\left\|\left(\frac{pn}{\theta c}\right)^{\frac{1}{\theta}}(AYB + M)\right\|_\theta\right\} dY.$$

Let $Y = (\theta c/pn)^{\frac{1}{\theta}} A^{-1}ZB^{-1} - M$. Then

$$\mathscr{E}\|AYB + M\|_\theta = \frac{\theta c}{pn}\int_{\mathbb{R}^{p\times n}}\left(2\Gamma\left(1 + \frac{1}{\theta}\right)\right)^{-pn}\|Z\|_\theta \exp\{-\|Z\|_\theta\} dZ$$

$$= \frac{\theta c}{pn}pn\mathscr{E}(|Z_1|^\theta) = c.$$

Next assume that X satisfies (3.1). Then

(3.4) $$\mathscr{E}_X(-\log f(X)) = \int_{\mathbb{R}^{p\times n}} -f(X)\log f(X)\,dX \le -\int_{\mathbb{R}^{p\times n}} f(X)\log g(X)\,dX$$

where equality holds iff $f(X) = g(X)$ almost everywhere in $\mathbb{R}^{p\times n}$. Therefore

$$\mathscr{E}_X(-\log f(X)) \le -\int_{\mathbb{R}^{p\times n}} f(X)\log g(X)\,dX$$

$$= -\int_{\mathbb{R}^{p\times n}} f(X)\left[\log k_{p,n}\left(\left(\frac{\theta c}{pn}\right)^{\frac{1}{\theta}} A^{-1}, B^{-1}, \theta\right)\right.$$

(3.5)

$$\left. - \frac{pn}{\theta c}\|AXB + M\|_\theta\right] dX$$

$$= \frac{pn}{\theta}\left(1 + \log\frac{\theta c}{pn}\right) - \log k_{p,n}(A^{-1}, B^{-1}, \theta),$$

and taking $f(X) = g(X)$ in (3.4), (3.5) shows that the maximum entropy is given by (3.3). $\square$

For characterizing matrix variate θ-generalized normal distributions which are invariant under certain transformations, we need the following lemma. Before we state the lemma we define the sets

$$\mathscr{P} = \{P : P \text{ is a } p \times p \text{ permutation matrix}\},$$
$$\mathscr{R} = \{P : P \text{ is a } p \times p \text{ signed permutation matrix}\},$$
$$\mathscr{G} = \{G : G \text{ is a } p \times p \text{ orthogonal matrix}\}.$$

Lemma 3.1. (a) *Let A be a $p \times p$ matrix $(p > 1)$. Then for every $P \in \mathcal{P}$ there exists a $Q \in \mathcal{R}$ such that $PA = AQ$ iff $A = SR$ where $R \in \mathcal{R}$ and S has one of the following structures.*

(3.6)   (i) $S = (S_{ij})$, *where* $S_{ij} = \begin{cases} a & \text{if } i \neq j, \\ b & \text{if } i = j, \end{cases}$

                (ii) $S = (S_{ij})$, *where* $S_{ij} = a_j$,

(3.7)   (iii) $S = \begin{pmatrix} a & a & a & b \\ a & -a & -a & b \\ -a & a & -a & b \\ -a & -a & a & b \end{pmatrix}$, *if $p = 4$,*

                (iv) $S = \begin{pmatrix} a & 0 & -a \\ -a & a & 0 \\ 0 & -a & a \end{pmatrix}$, *if $p = 3$,*

(3.8)   (v) $S = \begin{pmatrix} a & b \\ -a & b \end{pmatrix}$, *if $p = 2$,*

                (vi) $S = \begin{pmatrix} a & b \\ -a & -b \end{pmatrix}$, *if $p = 2$.*

   (b) *Let A be a $p \times p$ nonsingular matrix $(p > 1)$. Then for every $P \in \mathcal{P}$ there exists a $Q \in \mathcal{R}$ such that $PA = AQ$ iff $A = SR$ where $R \in \mathcal{R}$ and S has either the structure* (i) *with $a \neq b$ and $a \neq \frac{b}{1-p}$, or* (iii), *with $ab \neq 0$, or* (v) *with $ab \neq 0$.*

Proof. (a) It is easy to see that if the $p \times p$ matrix S has the structure described in (i)–(vi) and $P \in \mathcal{P}$, then there exists $T \in \mathcal{R}$ such that $PS = ST$. Hence if $A = SR$ where $R \in \mathcal{R}$, then $PA = PSR = STR = SRR'TR = AQ$ with $Q = R'TR \in \mathcal{R}$; therefore, A satisfies the condition of the theorem.

   Conversely, let A be a $p \times p$ $(p > 1)$ matrix so that for every $P \in \mathcal{P}$ there exists a $Q \in \mathcal{R}$ such that $PA = AQ$. This must be satisfied for each $P \in \mathcal{P}$ of the form

$$P_{ij} = \begin{cases} 1 & \text{if } (i, j) = (k, \ell) \text{ or } (i, j) = (\ell, k), \\ 1 & \text{if } i = j \text{ but } i \neq k, i \neq \ell, \\ 0 & \text{elsewhere,} \end{cases}$$

where $1 \leq k \neq \ell \leq p$. Let us denote this P by $P(k, \ell)$. If the elements in each column of A are equal, we get the case (ii). If there exists a column of A which contains at least two different elements, then without loss of generality we can assume that this is the first column $A^{(1)}$. Assume that the elements of $A^{(1)}$ have k different values, and p_i of them are equal to a_i $(i = 1, \ldots, k)$.

   Let $\mathcal{U} = \{P(k, \ell) : A_{k1} \neq A_{\ell 1}\}$. Let $P(k, \ell) \in \mathcal{U}$. Then there exists a $Q \in \mathcal{R}$ such that $P(k, \ell)A = AQ$ or equivalently $P(k, \ell)AQ' = A$. Since $Q' \in \mathcal{R}$, that means $P(k, \ell)A$ must contain a column which differs from the first column of A by at most a (-1) multiplier. Therefore A must contain a column $A(k, \ell)$ or $-A(k, \ell)$, where $A(k, \ell) = P(k, \ell)A^{(1)}$.

   1. Assume there exists $P(k, \ell) \in \mathcal{U}$ such that $A(k, \ell) = \pm A^{(1)}$. Fix k and ℓ. $A(k, \ell) = A^{(1)}$ is not possible since $A_{k1} \neq A_{\ell 1}$. If $A(k, \ell) = -A^{(1)}$, then

$A_{k1} = -A_{\ell 1}$, and $A_{i1} = 0$ if $i \neq k$ and $i \neq \ell$. If $p = 2$, we get (v) and (vi). If $p = 3$, we get (iv). If $p \geq 4$, then comparing $A(i, j)$'s, $i \in \{k, \ell\}$; $j \notin \{k, \ell\}$ we see that $\pm A(i, j)$'s are all different from one another and also different from $A^{(1)}$. So A must contain at least $2(p - 2) + 1$ different columns. But $2(p - 2) + 1 > p$ if $p \geq 4$ and A has only p columns which is a contradiction.

2. Assume that if $P(k, \ell) \in \mathscr{U}$, then $A(k, \ell) \neq \pm A^{(1)}$ but on the other hand there exist $P(k, \ell), P(r, s) \in \mathscr{U}$, $P(k, \ell) \neq P(r, s)$, such that $A(k, \ell) = \pm A(r, s)$. Fix k, ℓ, r, and s. Then it is easy to see that $A(k, \ell) \neq A(r, s)$. If $A(k, \ell) = -A(r, s)$, then either $\{k, \ell\} \cap \{r, s\} = \varnothing$ or $\{k, \ell\} \cap \{r, s\}$ consists of one element. If $\{k, \ell\} \cap \{r, s\} \neq \varnothing$, e.g. $k = r$, then $A(k, \ell) = -A(r, s)$ implies that $A^{(1)} = 0$ which is impossible since $A^{(1)}$ should contain at least two different elements.

If $\{k, \ell\} \cap \{r, s\} = \varnothing$, then $A(k, \ell) = -A(r, s)$ implies that $A_{i1} = 0$ if $i \notin \{k, \ell, r, s\}$. Moreover, if $p = 4$, then $|A_{k1}| = |A_{r1}|$ must hold because otherwise considering $A(i, j)$ where $i \in \{k, \ell\}$, $j \in \{r, s\}$ we would find 4 additional columns of A besides $A^{(1)}$, which is impossible. If $|A_{k1}| = |A_{r1}|$, we obtain (iii). If $p \geq 5$, then considering $A(i, j)$ where $i \in \{k, \ell, r, s\}$ and $j \notin \{k, \ell, r, s\}$ we see that A must contain at least $4(p - 4)$ additional columns. But $1 + 1 + 4(p - 4) > p$ if $p \geq 5$ and A has p columns which is a contradiction.

3. Assume that if $P(k, \ell) \in \mathscr{U}$, then $A(k, \ell) \neq \pm A^{(1)}$; moreover, if $P(r, s) \in \mathscr{U}$ also holds and $P(k, \ell) \neq P(r, s)$, then $A(k, \ell) \neq \pm A(r, s)$. But that means that each $P(k, \ell) \in \mathscr{U}$ will correspond to a different column of A: $A(k, \ell)$.

Now $\mathscr{U}$ has $\ell = \frac{p_1(p-p_1)+p_2(p-p_2)+\cdots+p_k(p-p_k)}{2}$ elements and it can be seen that $1 + \ell \geq p$ where equality holds if and only if $k = 2$ and $p_1 = 1$ or $p_k = 1$. So $A^{(1)}$ must contain $p - 1$ equal elements and one which is different from the others, say A_{k1}. Then considering $A(k, i)$, $i \neq k$, shows that A must have the structure (i).

(b) The result for A nonsingular follows if we consider that the determinant of the matrix S is 0 in (ii), (iv), and (vi), $(b - a)^{p-1}(b + (p - 1)a)$, in (i), $-16a^3b$ in (iii) and $2ab$ in (v). □

Now, we can prove the characterization result.

Theorem 3.2. *Let* $X \sim N_{p,n}(M, A, B, \theta)$, $p \geq 2$, $\theta \neq 2$. *Then*

(a) *for every* $P \in \mathscr{P}$, $PX \approx X$; *that is*, PX *and* X *are identically distributed if the rows of* M *are identical and* $A = SU$ *where* $U \in \mathscr{R}$ *and* S *has the structure described by* (3.6), (3.7), *or* (3.8);

(b) *for every* $R \in \mathscr{R}$, $RX \approx X$ *if* $M = 0$ *and* $A = SU$ *where*

(3.9) $$S = (S_{ij}), \quad where \ S_{ij} = \begin{cases} b \neq 0 & if \ i = j, \\ 0 & otherwise, \end{cases}$$

or

(3.10) $$S = \begin{pmatrix} a & a \\ -a & a \end{pmatrix} \quad where \ a \neq 0 \ if \ p = 2;$$

(c) *for every* $G \in \mathscr{G}$, $GX \approx X$ *is impossible; that is*, X *cannot be left-invariant.*

Proof. If $X \sim N_{p,n}(M, A, B, \theta)$ and $C : p \times p$ is nonsingular, then $CX \sim N_{p,n}(CM, CA, B, \theta)$. So $X \approx CX$ iff

(3.11) $M = CM$ and there exists $V \in \mathcal{R}$ such that $CA = AV$.

Note that $\mathcal{P} \subset \mathcal{R} \subset \mathcal{G}$.

(a) $PM = M$ for every $P \in \mathcal{P}$ iff the rows of M are identical. Using Lemma 3.1 we obtain the structure of A.

(b) Here we have a subset of the distributions in (a). It is easy to see that $M = 0$ and $A = SU$ where S is given by (3.9) or (3.10) satisfies (3.11). On the other hand, let $R = -I_p$. This gives $-I_p M = M$ and so $M = 0$. Now let

$$
R = \begin{pmatrix} -1 & & & 0 \\ & 1 & & \\ & & \ddots & \\ 0 & & & 1 \end{pmatrix}.
$$

Then there must be a $Q \in \mathcal{R}$ such that $RA = AQ$. Hence $RS = SuQu'$ where $uQu' \in \mathcal{R}$. However, this is only possible if $a = 0$ in (3.6), and (3.7), and $|a| = |b|$ in (3.8). Taking $a = 0$ in (3.7) makes $|A| = 0$ which is not allowed. So we are left with (3.9) and (3.10).

(c) Here we must have a subset of the distributions in (b). Let $G \in \mathcal{G}$ such that the first row of G is $(\frac{1}{\sqrt{p}}, \frac{1}{\sqrt{p}}, \ldots, \frac{1}{\sqrt{p}})$. Then we can see that there is no $R \in \mathcal{R}$ for which $GS = SR$. $\square$

BIBLIOGRAPHY

T. R. Goodman and S. Kotz, *Multivariate θ-generalized normal distributions*, J. Multivariate Anal. **3** (1973), 204–219.

A. K. Gupta and T. Varga, *Characterization of matrix variate normal distributions*, J. Multivariate Anal. **41** (1992), 80–88.

———, *Elliptically contoured models in statistics*, Kluwer Academic, Dordrecht, 1993.

R. J. Muirhead, *Aspects of multivariate statistical theory*, Wiley, New York, 1982.

DEPARTMENT OF MATHEMATICS AND STATISTICS, BOWLING GREEN STATE UNIVERSITY, BOWLING GREEN, OHIO 43403
E-mail address: gupta@math.bgsu.edu