Journal of Multivariate Analysis **76**, 153–157 (2001)
doi:10.1006/jmva.2000.1900, available online at http://www.idealibrary.com on IDEAL®

NOTE

On a Matrix Representation Lemma Useful in Determining Maximal Invariance Groups

H. Finner

Deutsches-Diabetes-Forschungsinstitut Düsseldorf, Düsseldorf, Germany

and

M. Roters

Universität Potsdam, Potsdam, Germany

Received December 9, 1997; published online November 7, 2000

testing problem invariant. His main result concerning the derivation of the maximal invariance group is heavily based on a matrix representation lemma which can be considered interesting in its own right. Unfortunately, the proof of this lemma is erroneous and there seems to be no trivial way to correct it. The aim of this note is to show the validity of the assertion. © 2000 Academic Press

AMS 1991 subject classifications: 62A05, 62H15.

Key words and phrases: invariance group, multivariate normal models, affine transformations.

The quest for maximal invariance groups is of some interest in hypothesis testing problems, especially in multivariate statistics, e.g., in MANOVA or GMANOVA models. As mentioned by Kariya (1985, p. 49), the problem of maximality has never been explicitly questioned even in a specific problem until Banken (1984), where a group leaving the GMANOVA problem invariant is shown to be maximal in the general affine linear group. In contrast to Banken (1984, 1986), Kariya (1985), as well as Nabeya and Kariya (1986), considered more general groups than the group of all affine transformations. However, the derivation of maximal invariance groups within the group of all affine transformations in Banken (1986) is heavily based on a matrix representation lemma which can be considered interesting in its own right. Unfortunately, the proof of this lemma contains a mistake which does not seem to be repairable in an easy manner. The aim of this note is to prove the correctness of the assertion.

We use notation similar to that in Banken (1986). Let $\mathbf{R}^{n\times p}$ denote the set of all real $n\times p$ matrices, $\mathrm{GL}(p)$ the set of all nonsingular real $p\times p$ matrices, $O(p)$ the set of all orthogonal real $p\times p$ matrices, $S_+(p)$ the set of all positive definite symmetric $p\times p$ matrices, and $S_{(+)}(p)$ the set of all

0047-259X/00 $35.00

positive semidefinite symmetric $p \times p$ matrices. The Kronecker product of matrices will be indicated by the sign $\otimes$ and the rank of a matrix A by $\operatorname{rk}(A)$. Moreover, for $n \times p$ matrices A and B we write $A \propto B$ if $A = \gamma B$ for some $\gamma \in \mathbf{R}$. The ith unit vector in $\mathbf{R}^p$ will be denoted by e_i and the $p \times p$ identity matrix by I_p.

Now the matrix representation lemma under consideration reads as follows.

Lemma 1. *For $C \in O(np)$ the following statements are equivalent:*

(i) $\forall \Sigma \in S_+(p) : \exists \Psi \in S_+(p) : C(I_n \otimes \Sigma)\, C' = I_n \otimes \Psi.$

(ii) $\exists P \in O(n) : \exists F \in O(p) : C = P \otimes F.$

The proof given in Banken (1986) fails on p. 159, lines 4–7. The linear mapping g defined there is not from $\mathbf{R}^p \to \mathbf{R}^p$ but from $\mathbf{R}^p \to \mathbf{R}^{p \times n}$. Therefore, the matrix representation of g indicated in line 7, which is essentially used in the rest of the proof, is invalid.

The following lemmas will be helpful during the proof of Lemma 1.

Lemma 2. *For all $A, B \in \mathbf{R}^{n \times p}$ it holds that $\forall x \in \mathbf{R}^p$; $Ax \propto Bx \Leftrightarrow A \propto B$.*

Proof. Only the sufficiency of the left-hand side for the right-hand side of the equivalence sign has to be proved. It follows from the assumption that for all columns $a_i = Ae_i$ of A and $b_i = Be_i$ of B there exist real numbers γ_i such that $a_i = \gamma_i b_i$ for $i = 1, ..., p$. First, if $b_i = 0$ then $a_i = 0$ and the proportionality relation between these vectors holds for each $\gamma_i \in \mathbf{R}$. Second, if b_i and b_j are non-zero linearly dependent columns of B, then there exists a $\gamma \neq 0$ such that $b_j = \gamma b_i$. Then for $x = -\gamma e_i + e_j$ we obtain $-\gamma(\gamma_i - \gamma_j)\, b_i = Ax \propto Bx = 0$, hence $\gamma_i = \gamma_j$. Third, consider the case that b_i and b_j are linearly independent. By assumption there exists a $\gamma \in \mathbf{R}$ such that $a_i + a_j = \gamma(b_i + b_j)$, hence $\gamma_i b_i + \gamma_j b_j = \gamma(b_i + b_j)$, or equivalently $(\gamma_i - \gamma)\, b_i + (\gamma_j - \gamma)\, b_j = 0$, which yields $\gamma_i = \gamma_j$ by the linear independence of b_i and b_j. Altogether, we have shown that $A \propto B$, as required. ∎

Lemma 3. *Let $U \in \mathrm{GL}(p)$ such that $\Sigma = U\Sigma U'$ for all $\Sigma \in S_{(+)}(p)$. Then $U = I_p$ or $U = -I_p$.*

Proof. Choosing $\Sigma = e_i e_i'$ yields $Ue_i = \delta_i e_i$ with $\delta_i \in \{-1, 1\}$, $i = 1, ..., p$. Applying then the assumption to $\Sigma = (e_i + e_j)(e_i + e_j)'$, we get $1 = e_i' \Sigma e_j = e_i' U \Sigma U' e_j = \delta_i \delta_j e_i' \Sigma e_j = \delta_i \delta_j$ for all $i \neq j$, hence $\delta_1 = \cdots = \delta_p$, which entails $U = I_p$ or $U = -I_p$. ∎

Proof of Lemma 1. Only the sufficiency of (i) for (ii) has to be proved. We first note that (i) is equivalent to

$$\forall \Sigma \in S_{(+)}(p) : \exists \Psi \in S_{(+)}(p) : C(I_n \otimes \Sigma)\, C' = I_n \otimes \Psi, \tag{1}$$

since, on the one hand, every $\Sigma \in S_{(+)}(p)$ can be represented as the difference of the two positive definite matrices $\Sigma + I_p$ and I_p, which have the same eigenvalues as their (obviously unique) corresponding Ψ-matrices, i.e., Ψ_1 (say) and I_p, so that the difference matrix $\Psi_1 - I_p$ is again an element of $S_{(+)}(p)$. On the other hand, if $\Sigma \in S_+(p)$ the corresponding matrix $\Psi \in S_{(+)}(p)$ must have full rank and hence is in $S_+(p)$.

It will be shown that (1) implies (ii). To this end let $g_i : \mathbf{R}^p \to \mathbf{R}^{p \times n}$ be defined by

$$g_i(x) = [0_{p \times (i-1)\, p}, I_p, 0_{p \times (n-i)\, p}]\; C(I_n \otimes x) = [C_{i1} x, ..., C_{in} x],$$

where $C = [C_{ij}]_{1 \leqslant i,\, j \leqslant n}$, $C_{ij} \in \mathbf{R}^{p \times p}$. Let $i \in \{1, ..., n\}$ be fixed in the following. We first show that $\text{rk}(g_i(x)) = 1$ for all $x \neq 0$. By assumption (1), for each $x \neq 0$ there exists a $\Phi_x \in S_{(+)}(p)$ such that $C(I_n \otimes xx')\, C' = I_n \otimes \Phi_x$. Since $n = \text{rk}(C(I_n \otimes xx')\, C') = \text{rk}(I_n \otimes \Phi_x)$, we have $\text{rk}(\Phi_x) = 1$. Moreover, it follows that

$$\begin{aligned} g_i(x)\, g_i(x)' &= [0_{p \times (i-1)\, p}, I_p, 0_{p \times (n-i)\, p}]\; C(I_n \otimes xx') \\ &\quad \times C'[0_{p \times (i-1)\, p}, I_p, 0_{p \times (n-i)\, p}]' \\ &= [0_{p \times (i-1)\, p}, I_p, 0_{p \times (n-i)\, p}](I_n \otimes \Phi_x) \\ &\quad \times [0_{p \times (i-1)\, p}, I_p, 0_{p \times (n-i)\, p}]' \\ &= \Phi_x, \end{aligned}$$

hence $\text{rk}(g_i(x)) = 1$ or equivalently

$$\forall x \neq 0 : \text{rk}([C_{i1} x, ..., C_{in} x]) = 1. \tag{2}$$

This implies that

$$\forall j \in \{1, ..., p\} : \exists r_j \in \{1, ..., n\} : [C_{ir_j} e_j \neq 0 \text{ and } \forall s \neq r_j : C_{is} e_j \propto C_{ir_j} e_j]. \tag{3}$$

Since $C \in O(np)$ it is $\text{rk}([C_{i1}, ..., C_{in}]) = p$. Then (3) implies that $C_{ir_1} e_1, ..., C_{ir_p} e_p$ are linearly independent. We now show $\text{rk}(C_{ir_1}) = p$. This follows by (3) if $C_{ir_1} e_j \neq 0$ for all $j \in \{2, ..., p\}$. Suppose there exists a $j \in \{2, ..., p\}$ such that $C_{ir_1}e_j=0$. For $x=e_1+e_j$ property (2) entails $\mathrm{rk}([C_{ir_1}(e_1+e_j), C_{ir_j}(e_1+e_j)])\leqslant 1$. But since $C_{ir_j}e_1=\gamma C_{ir_1}e_1$ for some $\gamma\in\mathbf{R}$ and

$$C_{ir_1}(e_1+e_j)=C_{ir_1}e_1+C_{ir_1}e_j=C_{ir_1}e_1,$$

$$C_{ir_j}(e_1+e_j)=C_{ir_j}e_1+C_{ir_j}e_j=\gamma C_{ir_1}e_1+C_{ir_j}e_j,$$

it follows that

$$\begin{aligned}\mathrm{rk}([C_{ir_1}(e_1+e_j), C_{ir_j}(e_1+e_j)])&=\mathrm{rk}([C_{ir_1}e_1, \gamma C_{ir_1}e_1+C_{ir_j}e_j])\\&=\mathrm{rk}([C_{ir_1}e_1, C_{ir_j}e_j])=2,\end{aligned}$$

which yields a contradiction. As a consequence, $C_{ir_1}x\neq 0$ for all $x\neq 0$, and by (2)

$$\forall j\in\{1, ..., n\} : \forall x\in\mathbf{R}^p : C_{ij}x \propto C_{ir_1}x,$$

hence by applying Lemma 2

$$\forall j\in\{1, ..., n\} : C_{ij} \propto C_{ir_1},$$

and finally

$$[C_{i1}, ..., C_{in}]=w_i'\otimes R_i$$

for some $w_i\in\mathbf{R}^n$ satisfying $w_i'w_i=1$ and some $R_i\in\mathrm{GL}(p)$. Obviously, we get $R_i\in O(p)$ from $C\in O(np)$. Moreover, $I_{np}=CC'=[w_i'w_jR_iR_j']_{1\leqslant i, j\leqslant n}$ entails $w_i'w_j=0$ for all $i\neq j$. On the other hand, (1) implies that for all $\Sigma\in S_{(+)}(p)$ there exists a $\Psi\in S_{(+)}(p)$ such that

$$\begin{aligned}I_n\otimes\Psi=C(I_n\otimes\Sigma)\,C'&=[w_i'w_jR_i\Sigma R_j']_{1\leqslant i, j\leqslant n}\\&=\mathrm{diag}(R_1\Sigma R_1', ..., R_n\Sigma R_n'),\end{aligned}$$

hence for all $i=2, ..., n$ and all $\Sigma\in S_{(+)}(p)$,

$$R_1\Sigma R_1'=R_i\Sigma R_i', \qquad \text{or equivalently} \quad \Sigma=R_1'R_i\Sigma R_i'R_1.$$

With Lemma 3 we get $R_i=R_1$ or $R_i=-R_1$ for all $i=2, ..., n$. Setting $F=R_1\in O(p)$ and $P=[p_1, ..., p_n]'$, where $p_i=w_i$ if $R_i=R_1$ and $p_i=-w_i$ if $R_i=-R_1$, $i=1, ..., n$, the properties of the w_i's derived before imply $P\in O(n)$. Thus we obtain the desired representation $C=P\otimes F$. ▮

REFERENCES

L. Banken, On the reduction of the general MANOVA model, Technical Report, Universität Trier, Trier, 1984.

L. Banken, The greatest invariance-group of multivariate models, *J. Multivariate Anal.* **19** (1986), 156–161.

T. Kariya, "Testing in the Multivariate General Linear Model," Kinokuniya, New York, 1985.

S. Nabeya and T. Kariya, Transformations preserving normality and Wishart-ness, *J. Multivariate Anal.* **20** (1986), 251–264.

Printed in Belgium