ZASTOSOWANIA MATEMATYKI
APPLICATIONES MATHEMATICAE
XVIII, 1 (1983), p. 49-54

G. S. ROGERS (Las Cruces, N. Mex.)

RANK CONDITIONS FOR ESTIMABILITY OF COVARIANCES

Let Y be a random vector with expectation $\mathrm{E}[Y] = X\beta$ and covariance matrix

$$\Sigma = \sum_{g=1}^{m} \sigma_g G_g.$$

(A precise description of the model is given in Section 2.) As usual, a linear function of $\sigma = (\sigma_1, \ldots, \sigma_m)'$, say $q'\sigma$, is said to be *estimable* if there is a quadratic form $Y'DY$ such that $\mathrm{E}[Y'DY] = q'\sigma$ for all σ and β. The main results of the paper (Section 4) are necessary and sufficient conditions for estimability of such linear functions in terms of the ranks of certain matrices. This is an extension of the work of Baksalary and Kala [1] with respect to linear functions of β which appears naturally as part of this exposition; they, in turn, extended the results of Milliken [5]. For completeness, the principle algebraic basis of all this work is given in the first part of Section 1; the second part contains material for dealing with the covariance matrix. Section 2 contains the statistical preliminaries for derivation of the results mentioned above.

1. Algebraic preliminaries. The following paragraph is standard (see, e.g., [2]):

Let F_1 and F_2 be finite-dimensional vector spaces and let T be a linear transformation from F_1 into F_2. Then $\mathscr{R}(T) = \{Tf\colon f \in F_1\}$ is a subspace of F_2 whose dimension $\dim\mathscr{R}(T)$ is equal to the rank $\varrho(T)$. For $\mathscr{N}(T) = \{f \in F_1\colon Tf = 0\}$,

$$\dim\mathscr{R}(T) + \dim\mathscr{N}(T) = \dim F_1.$$

Baksalary and Kala [1] used this to get a rank condition for the solution of a system of matrix equations:

Let the matrices X, being $M \times N$, and A, being $M \times K$, be given; let I_M denote an $M \times M$ identity matrix. Let $F_1 = \mathscr{R}(X)$ and $T = I_M - AA^-$, where A^- is any generalized inverse of A (see [7]). $\mathscr{N}(T) = \mathscr{R}(X) \cap \mathscr{R}(A)$ since $(I_M - AA^-)f = 0$ iff $AA^-f = f$ iff $f \in \mathscr{R}(A)$; also, $\mathscr{R}(T)$

$= \mathscr{R}\big((I_M - AA^-)X\big)$. Hence

$$\dim\mathscr{R}\big((I_M - AA^-)X\big) + \dim\big(\mathscr{R}(X)\cap\mathscr{R}(A)\big) = \dim\mathscr{R}(X)$$

or

$$\dim\mathscr{R}\big((I_M - AA^-)X\big) = \dim\mathscr{R}(X) - \dim\big(\mathscr{R}(X)\cap\mathscr{R}(A)\big).$$

Then

$$\varrho\big((I_M - AA^-)X\big) = \varrho(X) - \varrho(A)$$

iff

$$\varrho\big(\mathscr{R}(X)\cap\mathscr{R}(A)\big) = \varrho(A) \text{ iff } \mathscr{R}(A)\subset\mathscr{R}(X) \text{ iff } A = XX^-A$$

iff there is a D such that $A = XD$.

As in [5],

$$\begin{aligned}\varrho\big((I_M - AA^-)X\big) &= \varrho\Big((I_M - AA^-)X\big((I_M - AA^-)X\big)^-\Big)\\ &= \operatorname{tr}(I_M - AA^-)\,X\big((I_M - AA^-)X\big)^-\end{aligned}$$

since the matrix in the trace is idempotent. Obviously, the rank (trace) condition is easier to check than the other necessary and sufficient conditions. The general solution is then

$$D = X^-A + (I_N - X^-X)Z,$$

where Z is an arbitrary $(N\times K)$-matrix.

The following notation and concepts will be used later. For the $(t\times s)$-matrix C with columns $c_1, c_2, \ldots, c_s$, $\mathscr{P}(C) = (c_1', c_2', \ldots, c_s')'$ is an $(st\times 1)$-vector called the *pack* of C (see [4]). When the product ABC is defined, $\mathscr{P}(ABC) = (C'\otimes A)\mathscr{P}(B)$, where $\otimes$ denotes the Kronecker product. The trace of a product AB is $\operatorname{tr}AB = \mathscr{P}'(A')\mathscr{P}(B)$. For $\Sigma = \sum_{g=1}^{m}\sigma_g G_g$,

$$\mathscr{P}(\Sigma) = \sum_{g=1}^{m}\sigma_g\mathscr{P}(G_g) = G'\sigma,$$

where $G' = \big(\mathscr{P}(G_1), \ldots, \mathscr{P}(G_m)\big)$. Also, $\mathscr{P}(C') = I_{(s,t)}\mathscr{P}(C)$, where $I_{(s,t)}$ is the permuted identity matrix such that the ij-th submatrix is $s\times t$ with 1 at its ji-th position and zeroes elsewhere.

2. Statistical format. The $(n\times 1)$-vector Y has components which are real random variables with mean $\mathrm{E}[Y] = X\beta$, where X is a known $(n\times p)$-matrix and β is a $(p\times 1)$-vector of parameters in an open subset Ω_1 of the real Euclidean p-space R^p. The covariance matrix of Y is

$$\Sigma = \sum_{g=1}^{m}\sigma_g G_g,$$

where $G_1, \ldots, G_m$ are known linearly independent symmetric $(n\times n)$-matrices fixed for $\sigma = (\sigma_1, \ldots, \sigma_m)'$ in an open subset Ω_2 of R^m,

$m \leqslant n(n+1)/2$. (Σ is said to have a *linear structure* or to be *patterned*; Rogers and Young [8] give a number of references utilizing this concept.) Note that $m = n(n+1)/2$ is the case of an arbitrary Σ, non-negative definite.

Definition. If A is $k \times p$, $A\beta$ is *unbiasedly estimable* (u.e.) if there is a $(k \times n)$-matrix D such that $\mathrm{E}[DY] = A\beta$ for all β in Ω_1 and all σ in Ω_2. If Q' is $r \times m$, $Q'\sigma$ is *u. e.* if there is an $(r \times n^2)$-matrix C' such that $\mathrm{E}[C'\mathscr{P}(YY')] = Q'\sigma$ for all σ in Ω_2 and all β in Ω_1; $r \leqslant m$.

Now $A\beta$ is u. e. iff $DX\beta = A\beta$ for all β in Ω_1 iff $DX = A$ is consistent. By the results in Section 1, $A' = X'D'$ is consistent iff

$$\varrho\big((I_p - A'A'^-)X'\big) = \varrho(X') - \varrho(A')$$

or, equivalently,

$$\varrho\big(X(I_p - A^-A)\big) = \varrho(X) - \varrho(A).$$

In the classical cases, X has full rank $p \leqslant n$ and A has full rank $k \leqslant p$. For a single linear function, say $a'\beta$, $k = 1$ and $\varrho(A) = \varrho(a') = 1$.

In some problems, like hypothesis testing, it is also specified that $V\beta = b$ is consistent, where V is $s \times p$ and b is $s \times 1$, both given. Then the variable can be considered as $\begin{pmatrix} Y \\ b \end{pmatrix}$ with expectation $\begin{pmatrix} X \\ V \end{pmatrix}\beta$. The condition that $A\beta$ be u. e. is then

$$\varrho\big((I_p - A'A'^-)(X', V')\big) = \varrho\big((X', V')\big) - \varrho(A').$$

Since $\mathscr{R}(V') \subset \mathscr{R}(X', V')$, we have

$$\varrho\big((I_p - V'V'^-)(X', V')\big) = \varrho\big((X', V')\big) - \varrho(V')$$

or

$$\varrho\big((X', V')\big) = \varrho\big((I_p - V'V'^-)X'\big) + \varrho(V').$$

Thus $A\beta$ is u. e., given $V\beta = b$, iff

$$\varrho\big((I_p - A'A'^-)(X', V')\big) = \varrho\big((I_p - V'V'^-)X'\big) + \varrho(V') - \varrho(A')$$

or, equivalently,

$$\varrho\left(\begin{pmatrix} X \\ V \end{pmatrix}(I_p - A^-A)\right) = \varrho\big(X(I_p - V^-V)\big) + \varrho(V) - \varrho(A).$$

3. Conditions for $Q'\sigma$. First consider $Q'\sigma$ with $r = 1$, say $q_1'\sigma$. Then $c_1' = \mathscr{P}'(C_1)$ is $1 \times n^2$ with C_1 being $n \times n$ and

$$\begin{aligned} \mathrm{E}[c_1'\mathscr{P}(YY')] &= c_1'\mathscr{P}(\Sigma + X\beta\beta'X') = c_1'G'\sigma + \mathscr{P}'(C_1)\mathscr{P}(X\beta\beta'X') \\ &= c_1'G'\sigma + \beta'X_1'C_1X\beta \end{aligned}$$

(see Section 1). Hence $q_1'\sigma$ is u. e. iff $Gc_1 = q_1$ and $X'C_1X = 0$.

Now $\mathscr{P}(X'C_1X) = (X'\otimes X')c_1$; with $X_1 = X'\otimes X'$, the necessary and sufficient condition that $q_1'\sigma$ be u. e. is that

$$\begin{pmatrix} G \\ X_1 \end{pmatrix} c_1 = \begin{pmatrix} q_1 \\ 0 \end{pmatrix}$$

be consistent. It follows that $Q'\sigma$ is u. e. iff

$$\begin{pmatrix} G \\ X_1 \end{pmatrix} \cdot C = \begin{pmatrix} Q \\ 0 \end{pmatrix}$$

is consistent; here C is $n^2\times r$, Q is $m\times r$, and the zero matrix is $p^2\times r$.

Since

$$\begin{pmatrix} Q \\ 0 \end{pmatrix}^- = (Q^-, W'),$$

where W' is an arbitrary $(r\times p^2)$-matrix, we have

$$\begin{pmatrix} Q \\ 0 \end{pmatrix}\begin{pmatrix} Q \\ 0 \end{pmatrix}^- = \begin{pmatrix} QQ^- & QW' \\ 0 & 0 \end{pmatrix}$$

and some simplification is obtained by taking $W' = 0$. Then the rank condition for consistency is

$$\begin{aligned} \varrho\left(\begin{pmatrix} (I_m - QQ^-)G \\ X_1 \end{pmatrix}\right) &= \varrho\left(\begin{pmatrix} G \\ X_1 \end{pmatrix}\right) - \varrho\left(\begin{pmatrix} Q \\ 0 \end{pmatrix}\right) \\ &= \varrho\big((I_{n^2} - X_1'X'^{-})(G', X_1')\big) + \varrho(X_1') - \varrho(Q) \\ &= \varrho\big(G(I_{n^2} - X_1^- X_1)\big) + \varrho(X_1) - \varrho(Q). \end{aligned}$$

Note that σ itself is u. e. when $Q' = I_m$; thus

$$\varrho(X_1) = \varrho\big(G(I_{n^2} - X_1^- X_1)\big) + \varrho(X_1) - \varrho(I_m) \quad \text{or} \quad \varrho\big(G(I_{n^2} - X_1^- X_1)\big) = m.$$

If $\varrho\big((I_{n^2} - X_1^- X_1)\big) = n^2 - \varrho(X_1) < m$, then σ is not u. e.; here $\varrho(X_1) = \varrho(X'\otimes X') = \big(\varrho(X)\big)^2$. Also, σ is not u. e. if $\varrho(G) < m$, whence the condition of linear independence to begin with.

4. Symmetric forms. Up to this point, no a priori restriction has been put on the matrices C_i in the quadratic forms

$$c_i'\mathscr{P}(YY') = \mathscr{P}'(C_i)\mathscr{P}(YY') = \operatorname{tr} C_i' YY' = Y'C_iY, \quad i = 1(1)r.$$

Most often, these matrices are taken to be symmetric. Then

$$\mathscr{P}(C_i) = \mathscr{P}(C_i') = I_{(n,n)}\mathscr{P}(C_i) \quad \text{or} \quad (I_{n^2} - I_{(n,n)})c_i = 0.$$

The procedure above can then be followed through with

$$X_1 = \begin{pmatrix} X'\otimes X' \\ I_{n^2} - I_{(n,n)} \end{pmatrix}.$$

Note that $\pi = (I_{n^2} - I_{(n,n)})/2 = \pi' = \pi^2$; that is, π is an orthogonal projector.

5. Forms with variance free of β. If the covariance matrix of $P(YY')$ is assumed to be

$$2(\Sigma\otimes\Sigma+\Sigma\otimes X\beta\beta'X'+X\beta\beta'X'\otimes\Sigma)$$

(as it would be if Y were normally distributed), then the variance of $c_i'\mathscr{P}(YY')$ is

$$(1/2)\operatorname{tr}(C_i+C_i')\Sigma(C_i'+C_i)\Sigma+\beta'X'(C_i+C_i')\Sigma(C_i'+C_i)X\beta.$$

This will be free of β for all Σ (when there is at least one σ^0 in Ω_2 such that $\sum_{g=1}^{m}\sigma_g^0 G_g$ is positive definite) iff

$$(C_i+C_i')X = 0$$

iff

$$(X'\otimes I_n)(I_{n^2}+I_{(n,n)})c_i = 0$$

iff

$$(X'\otimes I_n)(I_{n^2}-\pi)c_i = 0.$$

Note that $(C_i+C_i')X = 0$ implies $X'C_iX = -X'C_i'X$, whence $\beta'X'C_iX\beta = 0$ for all β in Ω_1. That is, if this variance is free of β, so is the mean $\mathrm{E}[c_i'\mathscr{P}(YY')]$.

If C_i is symmetric, then $\pi c_i = 0$ and the above condition reduces to $(X'\otimes I_n)c_i = 0$ or $C_iX = 0$. Then

$$(Y-X\beta_0)'C_i(Y-X\beta_0) = Y'C_iY$$

for all β_0 in Ω_1; $Y'C_iY$ is said to be *translation invariant*. (Rao [6] and La Motte [3] use these various conditions on C_i in related contexts.) Proceeding as before, we infer that each $q_i'\sigma$, $i = 1(1)r$, is u. e. by a symmetric translation invariant form $c_i'\mathscr{P}(YY') = Y'C_iY$ iff

$$\varrho\left(\begin{pmatrix}(I_m-QQ^-)G\\ X_1\end{pmatrix}\right) = \varrho\big(G(I_{n^2}-X_1^-X_1)\big)+\varrho(X_1)-\varrho(Q),$$

$$\text{where } X_1 = \begin{pmatrix}\pi\\ X'\otimes I_n\end{pmatrix}.$$

6. Forms under a given constraint. As in the case of β, it is possible that some linear constraint is also put on σ, say $H'\sigma = h$, H' being $s\times m$ and h being $s\times 1$, both given. Then, for e_i being $s\times 1$,

$$\mathrm{E}\left[(c_i', e_i')\begin{pmatrix}\mathscr{P}(YY')\\ h\end{pmatrix}\right] = q_i'\sigma \quad\text{or}\quad (c_i', e_i')\begin{pmatrix}G'\sigma+\mathscr{P}(X\beta\beta'X')\\ H'\end{pmatrix} = q_i'\sigma$$

for all σ in Ω_2 and all β in Ω_1 iff $Gc_i+He_i = q_i$ and $(X'\otimes X')c_i = 0$ are consistent. More generally, $Q'\sigma$ is u. e. iff

$$\begin{pmatrix} G & H \\ X_1 & 0_1 \end{pmatrix}\begin{pmatrix} C \\ E \end{pmatrix} = \begin{pmatrix} Q \\ 0_2 \end{pmatrix},$$

where $X_1 = X'\otimes X'$, C is $n^2\times r$, E is $s\times r$, 0_1 is $p^2\times s$, and 0_2 is $p^2\times r$.

Conditions of symmetry or translation invariance can be imposed on $C_1, \ldots, C_r$ by the appropriate choice of X_1 as indicated above. The rank condition is

$$\varrho\left(\left(\begin{pmatrix} I_m - QQ^- \\ X_1 \end{pmatrix} G \begin{pmatrix} I_m - QQ^- \\ 0 \end{pmatrix} H\right)\right) = \varrho\big(G(I_{n^2} - X_1^- X_1), H\big) + \varrho(X) - \varrho(Q).$$

References

[1] J. K. Baksalary and R. Kala, *Extensions of Milliken's estimability criterion*, Ann. Statist. 4 (1976), p. 639-641.

[2] N. Jacobson, *Lectures in abstract algebra*, Vol. 2. *Linear algebra*, New York 1953.

[3] L. R. La Motte, *Quadratic estimation of variance components*, Biometrics 29 (1973), p. 311-330.

[4] E. C. MacRae, *Matrix derivatives with an application to an adaptive linear decision problem*, Ann. Statist. 2 (1974), p. 337-346.

[5] G. A. Milliken, *New criteria for estimability for linear models*, Ann. Math. Statist. 42 (1971), p. 1588-1594.

[6] C. R. Rao, *Estimation of variance and covariance components — MINQUE theory*, J. Multivariate Anal. 1 (1971), p. 257-275.

[7] — and S. K. Mitra, *Generalized inverse of matrices and its applications*, New York 1971.

[8] G. S. Rogers and D. L. Young, *Explicit maximum likelihood estimators for certain patterned covariance matrices*, Comm. Statist. A6 (1977), p. 121-133.

DEPARTMENT OF MATHEMATICAL SCIENCES
NEW MEXICO STATE UNIVERSITY
LAS CRUCES, N. MEX. 88003

Received on 8. 11. 1978