The Annals of Mathematical Statistics
1970, Vol. 41, No. 2, 528–538

GAUSS–MARKOV ESTIMATION FOR MULTIVARIATE LINEAR MODELS: A COORDINATE FREE APPROACH[1]

BY MORRIS L. EATON

University of Chicago

0. Introduction and summary. The coordinate free (geometric) approach to univariate linear models has added both insight and understanding to the problems of Gauss Markov (GM) estimation and hypothesis testing. One of the initial papers emphasizing the geometric aspects of univariate linear models is Kruskal's (1961). The coordinate free approach is used in this paper to treat GM estimation in a multivariate analysis context. In contrast to the univariate situation, a central question for multivariate linear models is the existence of GM estimates. Of course, it is the more complicated covariance structure in the multivariate case that creates the concern over the existence of GM estimates. As the emphasis is on GM estimation, first and second moment assumptions (as opposed to distributional assumptions) play the key role.

Classical results for the univariate linear model are outlined in Section 1. In addition, a recent theorem due to Kruskal (1968) concerning the equality of GM and Least Squares (LS) estimates is discussed. A minor modification of Kruskal's result gives a very useful, necessary and sufficient condition for the existence of GM estimators for arbitrary covariance structures and a fixed regression manifold.

In Section 2, the outer product of two vectors and the Kronecker product of linear transformations is discussed and applied to describe the covariance structure of a random matrix. This application includes the case of a random sample from a multivariate population with covariance matrix $\Sigma > 0$ ("$\Sigma > 0$" means that Σ is positive definite).

The question of GM estimation in the standard multivariate linear model is taken up in Section 3. This model is described as follows: a random matrix $Y: n \times p$, whose rows are uncorrelated and each row has a covariance matrix $\Sigma > 0$, is observed. The mean matrix of Y, μ, is assumed to have the form $\mu = ZB$ where $Z: n \times q$ is known and of rank q, and $B: q \times p$ is a matrix of regression coefficients. For this model, GM estimators for μ and B exist and are well known (see Anderson (1958) chapter 8). The main result in Section 3 establishes a converse to this classical result. Explicitly, let Y have the covariance structure as above and assume Ω is a fixed regression manifold. It is shown that if a GM estimator for $\mu \in \Omega$ exists, then

Received April 17, 1969; revised October 2, 1969.

[1] This research was supported in part by National Science Foundation Grant NSF GP 6681 while the author was visiting Stanford University during the summer of 1968. Also, this research was supported in part by Research Grant No. NSF GP 8026 from the Division of Mathematical, Physical and Engineering Sciences of the National Science Foundation, and in part by the Department of Statistics, University of Chicago, under partial sponsorship of the Statistics Branch, Office of Naval Research. Reproduction in whole or in part is permitted for any purpose of the United States Government.

each element $\mu \in \Omega$ can be written as $\mu = ZB$ where $Z: n \times q$ is fixed and $B: q \times p$ ranges over all $q \times p$ real matrices.

The results in Section 4 and Section 5 are similar to the main result of Section 3. A complete description of all regression manifolds for which GM estimators exist is given for two different kinds of covariance assumptions concerning Σ (Σ as above). In Section 4, it is assumed that Σ has a block diagonal form with two blocks. Section 5 is concerned with the case when Σ has the so-called intra-class correlation form.

1. The univariate linear model. Let $(V, (\cdot,\cdot))$ be a real n-dimensional inner product space and let Y be a random vector taking values in V. Throughout, assume that Y has a mean vector $\mu \equiv E(Y)$ and a covariance operator $\Sigma \equiv \mathrm{Cov}(Y)$ (see Kruskal (1961) for definitions and properties of μ and Σ).

The univariate linear model may be described as follows. Assume that μ lies in a known linear subspace, $\Omega \subseteq V$ and

(1.1) $$\mathrm{Cov}(Y) = \sigma^2 \Sigma_0$$

where $\Sigma_0 > 0$ is known and $\sigma^2 > 0$ is known or unknown. When (1.1) holds, the GM estimator of μ is $\hat{\mu} \equiv P_{\Sigma_0} Y$ where P_{Σ_0} is the orthogonal projection onto Ω relative to the inner product $(\cdot,\cdot)_{\Sigma_0}$ defined by

(1.2) $$(x, y)_{\Sigma_0} \equiv (x, \Sigma_0^{-1} y).$$

Also, the LS estimate of μ is defined to be $\mu^* \equiv P_I Y$ where I is the identity transformation.

THEOREM 1.1. (Kruskal (1968)). *The two estimators $\hat{\mu}$ and μ^* are equal if and only if* (iff) *Ω is invariant under Σ_0.*

As will be seen later, the covariance structure for many multivariate linear models does not satisfy (1.1). In view of this, assume

(1.3) $$\mathrm{Cov}(Y) \in \Theta$$

where Θ is a given set of positive definite operators in V. However, there is now the possibility that a GM estimator $\mu \in \Omega$ does not exist. In order to define a GM estimator precisely in the present situation, let $\mathrm{Hom}(V)$ be the set of all linear transformations on V to V and let

(1.4) $$\mathscr{A} = \{A \mid A \in \mathrm{Hom}(V), Ax = x \quad \text{for all} \quad x \in \Omega\}.$$

Note that BY is an unbiased estimate for μ if and only if $B \in \mathscr{A}$.

DEFINITION 1.1. The estimator $A_0 Y$ is GM for $\mu \in \Omega$ iff $A_0 \in \mathscr{A}$ and

$$\mathrm{Var}_\Sigma\{(x, A_0 Y)\} \leqq \mathrm{Var}_\Sigma\{(x, AY)\} \quad \text{for all} \quad A \in \mathscr{A}, \Sigma \in \Theta, x \in V.$$

Var_Σ denotes variance when $\mathrm{Cov}(Y) = \Sigma$. When (1.1) holds, the classical Gauss–Markov Theorem asserts that a GM estimator exists, is unique, and $A_0 Y = P_{\Sigma_0} Y$ is the GM estimator. In view of Theorem 1.1, we have

THEOREM 1.2. *A* GM *estimator exists* iff $\Sigma_1^{-1}\Omega = \Sigma_2^{-1}\Omega$ *for all* $\Sigma_1, \Sigma_2 \in \Theta$.

PROOF. Theorem 1.1 shows that $P_{\Sigma_0} = P_I$ iff $\Sigma_0\Omega = \Omega$ and a trivial modification of Kruskal's (1968) argument shows that $P_{\Sigma_1} = P_{\Sigma_2}$ iff $\Sigma_1^{-1}\Omega = \Sigma_2^{-1}\Omega$. Now, fix Σ_1 and $\Sigma_2 \in \Theta$. Applying the uniqueness assertion of the Gauss–Markov Theorem, a GM estimator exists iff $P_{\Sigma_1} = P_{\Sigma_2}$. This completes the proof.

Note that if $I \in \Theta$, then a GM estimator exists iff

$$\Sigma\Omega = \Omega \quad \text{for all} \quad \Sigma \in \Theta. \tag{1.5}$$

A central question in the remainder of this paper is: given a particular set Θ, describe all the manifolds Ω for which a GM estimator exists.

2. Covariance structure in the multivariate linear model. In a discussion of multivariate linear models, the notions of the outer product of vectors and the Kronecker product of linear transformations arise naturally. Let $(V_i, (\cdot,\cdot)_i)$ be p_i-dimensional real inner product spaces for $i = 1, 2$ and let $\mathscr{L}_{p_1,p_2}$ be the real vector space of linear transformations on V_1 to V_2. If $y \in V_2$ and $x \in V_1$, the outer product of y and x, $y \otimes x \in \mathscr{L}_{p_1,p_2}$ is defined by

$$(y \otimes x)z = (x, z)_1 y \quad \text{for} \quad z \in V_1. \tag{2.1}$$

Some basic facts are (see Halmos (1958) page 40):

(2.2) $y \otimes x$ is linear in both x and y

(2.3) If $A \in \mathscr{L}_{p_1,p_2}$ has rank one, then $A = y \otimes x$ for some $x \in V_1$ and $y \in V_2$

(2.4) If $\{x_1, \cdots, x_{p_1}\}$ ($\{y_1, \cdots, y_{p_2}\}$) is a basis for V_1 (V_2, respectively), then $\{y_j \otimes x_i \mid i = 1, \cdots, p_1, j = 1, \cdots, p_2\}$ is a basis for $\mathscr{L}_{p_1,p_2}$.

Given the two inner products $(\cdot,\cdot)_1$ and $(\cdot,\cdot)_2$, there is a natural inner product on $\mathscr{L}_{p_1,p_2}$. Let $\{v_1, \cdots, v_{p_1}\}$ ($\{u_1, \cdots, u_{p_2}\}$) be an orthonormal (ON) basis for V_1 (V_2, respectively). For $A, B \in \mathscr{L}_{p_1,p_2}$, write $A = \sum\sum_{ij} a_{ij} u_i \otimes v_j$ and $B = \sum\sum_{ij} b_{ij} u_i \otimes v_j$, and define $\langle\cdot,\cdot\rangle$ on $\mathscr{L}_{p_1,p_2} \times \mathscr{L}_{p_1,p_2}$ by

$$\langle A, B\rangle \equiv \textstyle\sum\sum_{ij} a_{ij} b_{ij}. \tag{2.5}$$

It is not hard to verify that $\langle\cdot,\cdot\rangle$ is the unique inner product on $\mathscr{L}_{p_1,p_2}$ which satisfies the basic relationship

$$\langle y_1 \otimes x_1, y_2 \otimes x_2\rangle = (x_1, x_2)_1 (y_1, y_2)_2 \tag{2.6}$$

for all $y_1, y_2 \in V_2$ and $x_1, x_2 \in V_1$.

Note that if $\{x_1, \cdots, x_{p_1}\}$ ($\{y_1, \cdots, y_{p_2}\}$) is an ON basis for V_1 (V_2, respectively), then $\{y_j \otimes x_i \mid i = 1, \cdots, p_1, j = 1, \cdots, p_2\}$ is an ON basis for $\mathscr{L}_{p_1,p_2}$ with respect to $\langle\cdot,\cdot\rangle$. Also, for $A, B \in \mathscr{L}_{p_1,p_2}$,

$$\langle A, B\rangle = \operatorname{tr}[A][B]' \tag{2.7}$$

where $[A]$ ($[B]$) is the matrix of A (B, respectively) with respect to the two ON bases. Here, tr denotes trace and $[B]'$ is the transpose of $[B]$.

For $P \in \mathscr{L}_{p_2,p_2}$, $Q \in \mathscr{L}_{p_1,p_1}$, the Kronecker product of P and Q, $P \otimes Q$, is that linear transformation on $\mathscr{L}_{p_1,p_2}$ to $\mathscr{L}_{p_1,p_2}$ whose value at $A \in \mathscr{L}_{p_1,p_2}$ is PAQ' where Q' is the adjoint of Q relative to $(\cdot,\cdot)_1$. Clearly

$$P_1 \otimes Q_1 \circ P_2 \otimes Q_2 = (P_1 P_2) \otimes (Q_1 Q_2)$$

where $\circ$ denotes composition. Also, the adjoint of $P \otimes Q$ with respect to $\langle \cdot,\cdot \rangle$ is $P' \otimes Q'$. For further properties of $P \otimes Q$ the reader may consult Bellman ((1960) page 227) or Halmos ((1958) page 90).

EXAMPLE 2.1. Let $Y: n \times p$ be a random matrix whose rows, y_i, are uncorrelated and each row has a common covariance matrix Σ. Take $V_1 = R^p$ and $V_2 = R^n$ with the standard inner products. Using the standard ON bases, the inner product on $\mathscr{L}_{p,n}$ is $\langle A, B \rangle = \operatorname{tr} AB'$. As is well known, $\operatorname{Cov}(Y) = I_n \otimes \Sigma$, where I_n is the $n \times n$ identity matrix. To show this directly, let $A, B \in \mathscr{L}_{p,n}$ have rows a_i and b_i. Then

$$\begin{aligned} \text{(2.8)} \quad \operatorname{Cov}\{\langle A, Y\rangle, \langle B, Y\rangle\} &= \textstyle\sum\sum_{ij} \operatorname{Cov}\{(a_i, y_i)_1, (b_j, y_j)_1\} \\ &= \textstyle\sum_i (a_i, \Sigma b_i)_1 = \operatorname{tr} A \Sigma B' = \langle A, (I_n \otimes \Sigma) B\rangle. \end{aligned}$$

This concludes Example 2.1.

Suppose $Y \in \mathscr{L}_{p_1,p_2}$ is such that $\operatorname{Cov}(Y) = I_{p_2} \otimes \Sigma$. Then

$$\text{(2.9)} \qquad \operatorname{Cov}((\Gamma \otimes I_{p_1})Y) = \operatorname{Cov}(Y)$$

for all orthogonal $\Gamma \in \mathscr{L}_{p_2,p_2}$. Note that (2.9) characterizes $\operatorname{Cov}(Y)$—that is, if (2.9) holds for all orthogonal $\Gamma \in \mathscr{L}_{p_2,p_2}$, then there exists a positive semi-definite $\Sigma \in \mathscr{L}_{p_1,p_1}$ such that $\operatorname{Cov}(Y) = I_{p_2} \otimes \Sigma$. The proof of this assertion is not hard and is omitted.

3. GM estimation in the standard multivariate linear model. In order to motivate the main result of this section, we begin with a discussion of GM estimation in the standard multivariate linear model.

EXAMPLE 3.1. As in Example 2.1, let $Y: n \times p$ be a random matrix with $\operatorname{Cov}(Y) = I_n \otimes \Sigma$ where Σ ranges over $\mathscr{S}_p^+$—the set of all $p \times p$ positive definite matrices. Each element of the regression manifold Ω is assumed to have the form $\mu = ZB$ where $Z: n \times q$ is a fixed known matrix of rank q and $B: q \times p$ is the matrix of unknown regression parameters. The matrix B is free to vary over all $q \times p$ real matrices. Since

$$\text{(3.1)} \qquad (I_n \otimes \Sigma)(ZB) = Z(B\Sigma) \in \Omega$$

for each B, we see that $(I_n \otimes \Sigma)\Omega = \Omega$ for all $\Sigma \in \mathscr{S}_p^+$. Thus, from Theorem 1.2, a GM estimator for μ exists and from Section 1, $\hat{\mu}$ is the orthogonal projection (relative to $\langle \cdot,\cdot \rangle$) of Y onto Ω. However, it is straightforward to check that $Z(Z'Z)^{-1}Z' \otimes I_p$ is the orthogonal projection onto Ω. Hence $\hat{\mu} = (Z(Z'Z)^{-1}Z' \otimes I_p)Y = Z(Z'Z)^{-1}Z'Y$. Since $B = (Z'Z)^{-1}Z'\mu$, the GM estimator for B is $\hat{B} = (Z'Z)^{-1}Z'\hat{\mu} = (Z'Z)^{-1}Z'Y = [(Z'Z)^{-1}Z' \otimes I_p]Y$. Thus, Cov

$(\hat{B}) = \text{Cov}[((Z'Z)^{-1}Z' \otimes I_p)Y] = [(Z'Z)^{-1}Z' \otimes I_p](I_n \otimes \Sigma)[(Z'Z)^{-1}Z' \otimes I_p]' = (Z'Z)^{-1} \otimes \Sigma$. Of course, all of the above results are well known (see Anderson (1958) chapter 8).

Now, consider the following problem. Suppose $Y \in \mathscr{L}_{p_1,p_2}$ has $\text{Cov}(Y) = I_{p_2} \otimes \Sigma$ where Σ varies over $\mathscr{S}^+_{p_1}$. Let Ω be a fixed regression manifold and assume a GM estimator for μ exists. The problem is to characterize Ω—that is, for what Ω's do we have GM estimators. The remainder of this section is devoted to showing that

$$\Omega = \{\mu \mid \mu = ZB,\ Z \in \mathscr{L}_{q,p_2} \text{ fixed},\ B \in \mathscr{L}_{p_1,q}\} \tag{3.2}$$

when a GM estimator exists. Of course, this result shows that GM estimators exist only for the standard multivariate linear model when the covariance structure is $I_{p_2} \otimes \Sigma$, $\Sigma \in \mathscr{S}^+_{p_1}$.

Now, let $\Omega \neq \{0\}$ be a linear manifold in $\mathscr{L}_{p_1,p_2}$ and assume that

$$(I_{p_2} \otimes \Sigma)\Omega = \Omega \quad \text{for all} \quad \Sigma \in \mathscr{S}^+_{p_1}. \tag{3.3}$$

By Theorem 1.2, (3.3) is equivalent to the existence of a GM estimator since $I_{p_2} \otimes I_{p_1}$ is a possible covariance operator.

LEMMA 3.1. *If* (3.3) *holds, then* $(I_{p_2} \otimes B)\Omega \subseteq \Omega$ *for all* $B \in \mathscr{L}_{p_1,p_2}$.

PROOF. For $A \in \Omega$, (3.3) is equivalent to $A\Sigma \in \Omega$ for all $A \in \Omega$ and $\Sigma \in \mathscr{S}^+_{p_1}$. Since Ω is a manifold, we see that $A\Sigma_1 - A\Sigma_2 = A(\Sigma_1 - \Sigma_2) \in \Omega$ for all $A \in \Omega$ and Σ_1, $\Sigma_2 \in \mathscr{S}^+_{p_1}$. However every symmetric operator S can be written in the form $\Sigma_1 - \Sigma_2$ for $\Sigma_1, \Sigma_2 \in \mathscr{S}^+_{p_1}$ ($S = S + \lambda I - \lambda I$ and for sufficiently large $\lambda > 0$, $S + \lambda I \in \mathscr{S}^+_{p_1}$). Thus (3.3) is equivalent to

$$AS \in \Omega \quad \text{for all} \quad A \in \Omega, S \in \mathscr{S}_{p_1} \tag{3.4}$$

where $\mathscr{S}_{p_1}$ is all symmetric operators on V_1 to V_1. Equation (3.4) yields $AS_1 S_2 \in \Omega$ for $A \in \Omega$, S_1, $S_2 \in \mathscr{S}_{p_1}$ and by induction we have

$$AS_1 S_2 \cdots S_k \in \Omega \quad \text{for all} \quad A \in \Omega, S_i \in \mathscr{S}_{p_1}. \tag{3.5}$$

But every element $B \in \mathscr{L}_{p_1,p_1}$ can be written in the form $S_1 \cdots S_k$ for some k (see Olkin (1966)). This completes the proof.

LEMMA 3.2. *If* (3.3) *holds, then there exists a rank one element in* Ω. *Further, if* $z_1 \otimes x_1 \in \Omega$ *for* $x_1 \neq 0$, *then* $z_1 \otimes x \in \Omega$ *for all* $x \in V_1$.

PROOF. For the first assertion, let $A \neq 0$ be in Ω. Then there exists $u_1 \in V_1$ such that $Au_1 \neq 0$. From Lemma 3.1, $A(u_1 \otimes u_1) \in \Omega$. However, $A(u_1 \otimes u_1) = (Au_1) \otimes u_1$ which is of rank one.

For the second assertion, if $z_1 \otimes x_1 \in \Omega$, then $(z_1 \otimes x_1)B = z_1 \otimes (B'x_1) \in \Omega$ for all $B \in \mathscr{L}_{p_1,p_1}$. As B varies over $\mathscr{L}_{p_1,p_1}$, $B'x_1$ can be any vector in V_1. The proof is completed.

Now, define $M \subseteq V_2$ by

$$M = \{z \mid z \in V_2, z \otimes x \in \Omega \quad \text{for all} \quad x \in V_1\}. \tag{3.6}$$

LEMMA 3.3. *The set M is a linear manifold and* $\dim(M) \geqq 1$.

PROOF. It is trivial to check that M is a manifold. From Lemma 3.2 there is a rank one element, say $z_1 \otimes x_1$, in Ω. Hence $z_1 \neq 0$ and $x_1 \neq 0$ so that $z_1 \otimes x \in \Omega$ for all $x \in V_1$. Thus $z_1 \in M$ and the conclusion follows.

Setting $q = \dim(M)$, let $\{z_1, \cdots, z_q, z_{q+1}, \cdots, z_{p_2}\}$ be an ON basis for V_2 such that $\{z_1, \cdots, z_q\}$ is an ON basis for M. Also, let $\{x_1, \cdots, x_{p_1}\}$ be an ON basis for V_1. As pointed out before, $\{z_1 \otimes x_j \mid i = 1, \cdots, p_2; j = 1, \cdots, p_1\}$ is an ON basis for $\mathscr{L}_{p_1, p_2}$ with inner product $\langle \cdot, \cdot \rangle$. Consider the manifold Ω_0 given by

$$\Omega_0 = \text{span}\{z_i \otimes x_j \mid i = 1, \cdots, q; j = 1, \cdots, p_1\}. \tag{3.7}$$

Since $z_i \otimes x_j \in \Omega$ for $i = 1, \cdots, q; j = 1, \cdots, p_1$, we see that $\Omega_0 \subseteq \Omega$.

THEOREM 3.1. *If* (3.3) *holds, then* $\Omega_0 = \Omega$.

PROOF. Let P_0 be the orthogonal projection onto Ω_0. The theorem will be established if we can show that $(I - P_0)A = 0$ for all $A \in \Omega$. If $A_1 \in \Omega$, then $(I - P_0)A_1 \equiv C_1 \in \Omega$ since $C_1 = A_1 - P_0 A_1$ and $P_0 A_1 \in \Omega_0 \subseteq \Omega$. Noting that $\{z_i \otimes x_j \mid i = q+1, \cdots, p_2; j = 1, \cdots, p_1\}$ is an ON basis for the orthogonal complement of Ω_0, we can write

$$C_1 = \textstyle\sum_{i=q+1}^{p_2} \sum_{j=1}^{p_1} c_{ij} z_i \otimes x_j \in \Omega. \tag{3.8}$$

Now, fix j_0 and choose $B \in \mathscr{L}_{p_1, p_1}$ such that $Bx_j = 0$ for $j \neq j_0$ and $Bx_{j_0} = x_{j_0}$. Then

$$C_1 B = \textstyle\sum_{i=q+1}^{p_2} c_{i,j_0} z_i \otimes x_{j_0} \equiv \tilde{z} \otimes x_{j_0} \in \Omega. \tag{3.9}$$

Since $x_{j_0} \neq 0$, $\tilde{z} \in M$ by Lemma 3.2. However, $\tilde{z}$ is orthogonal to M by the definition of $\tilde{z}$ so that $\tilde{z} = 0$. Hence $c_{i,j_0} = 0$ for $i = q+1, \cdots, p_2$. Since j_0 was arbitrary, $c_{ij} = 0$ for $i = q+1, \cdots, p_2$ and $j = 1, \cdots, p_1$ and thus $C_1 = 0$. This completes the proof.

From Theorem 3.1, we obtain

COROLLARY 3.1. *If* (3.3) *holds, then Ω has a coordinate representation*

$$\Omega = \{\mu \mid \mu = \left[\begin{smallmatrix} B \\ 0 \end{smallmatrix}\right]\}, \qquad \begin{array}{ll} B \ \ a \ \ q \times p_1 & \textit{real matrix} \\ 0 \ \ a \ \ (p_2 - q) \times p_1 & \textit{zero matrix.} \end{array} \tag{3.10}$$

PROOF. From Theorem 3.1, each $\mu \in \Omega$ has the form

$$\mu = \textstyle\sum_{i=1}^{q} \sum_{j=1}^{p_1} b_{ij} z_i \otimes x_j. \tag{3.11}$$

By choosing $\{z_i \otimes x_j \mid i = 1, \cdots, p_2, j = 1, \cdots p_1\}$ as an ON basis for $\mathscr{L}_{p_1, p_2}$, then the coordinate representation for $\mu \in \Omega$ is (3.10) with $B = \{b_{ij}\}$. This completes the proof.

We conclude this section by remarking that (3.10) is commonly referred to as the "canonical" form for the standard multivariate linear model. It is well known that any regression manifold $\Omega = \{\mu\}$, where $\mu = ZB$, $Z: p_2 \times q$ fixed and $B: q \times p_1$ can be reduced (via a change of coordinates and relabeling) to the form (3.10)—

see Anderson (1958), chapter 8. Thus, our original assertion that the existence of GM estimators when $\operatorname{Cov}(Y) = I_{p_2} \otimes \Sigma$, $\Sigma \in \mathscr{S}_{p_1}^+$, implies a standard regression model is established.

4. GM estimation for an alternative regression model. We begin this section by giving a necessary and sufficient condition that a GM estimator exists for a non-standard regression model. Also the GM estimator is given when it exists.

Consider Euclidean spaces, R^n, R^q, R^r, and R^p where $q \leqq n$, $r \leqq p \leqq n$, equipped with the standard inner product. Assume we observe $Y \in \mathscr{L}_{p,n}$ and

$$\operatorname{Cov} Y = I_n \otimes \Sigma \quad \text{for} \quad \Sigma \in \Theta \subseteq \mathscr{S}_p^{\,+} \tag{4.1}$$

where Θ is unspecified. However, we assume that $I_p \in \Theta$. Consider a regression manifold Ω described as follows: $\mu \in \Omega$ if and only if

$$\mu = Z_1 B Z_2 \tag{4.2}$$

where $Z_1 \in \mathscr{L}_{q,n}$, $Z_2 \in \mathscr{L}_{p,r}$ are both of full rank and known, and $B \in \mathscr{L}_{r,q}$. For a general discussion and bibliography concerning this model, see Gleser and Olkin (1966).

In order to discuss the existence of GM estimators in the present context, let $\mathscr{X}_2$ be the row space of Z_2. That is, $\mathscr{X}_2$ is the image of R^r in R^p under the mapping Z_2'. Of course, $\mathscr{X}_2$ is an r-dimensional manifold in R^p. We now have

PROPOSITION 4.1. *In order that a* GM *estimator of μ in* (4.2) *exists, it is necessary and sufficient that $\mathscr{X}_2$ be invariant under Σ for all $\Sigma \in \Theta$.*

PROOF. By Theorem 1.2, a GM estimate exists iff $(I_n \otimes \Sigma)\Omega = \Omega$ for all $\Sigma \in \Theta$ and this is equivalent to

$$\mu\Sigma \in \Omega \quad \text{for all} \quad \Sigma \in \Theta. \tag{4.2}$$

However, (4.2) is the same as saying, given $B \in \mathscr{L}_{r,q}$ and $\Sigma \in \Theta$, there exists a $\tilde{B}(\Sigma) \in \mathscr{L}_{r,q}$ such that

$$Z_1 B Z_2 \Sigma = Z_1 \tilde{B}(\Sigma) Z_2. \tag{4.3}$$

Solving for $\tilde{B}(\Sigma)$, we see that (4.3) is equivalent to

$$Z_1 B Z_2 \Sigma = Z_1 B Z_2 \Sigma Z_2'(Z_2 Z_2')^{-1} Z_2 \tag{4.4}$$

for all $B \in \mathscr{L}_{r,q}$ and $\Sigma \in \Theta$. But (4.4) is clearly equivalent to

$$Z_2 \Sigma = Z_2 \Sigma Z_2'(Z_2 Z_2')^{-1} Z_2 \quad \text{for all} \quad \Sigma \in \Theta. \tag{4.5}$$

Since $Z_2'(Z_2 Z_2')^{-1} Z_2$ is the orthogonal projection onto $\mathscr{X}_2$, (4.5) is the same as

$$\Sigma Z_2' x \in \mathscr{X}_2 \quad \text{for all} \quad x \in R^r, \quad \text{and for all} \quad \Sigma \in \Theta. \tag{4.6}$$

However, (4.6) is exactly the assertion that $\mathscr{X}_2$ is invariant under Σ for all $\Sigma \in \Theta$. This completes the proof.

Now assume that a GM estimator exists. Then it is given by the orthogonal projection of Y onto Ω. When Ω is given by (4.2), it is readily verified that the projection onto Ω, P, is

$$P = [Z_1(Z_1'Z_1)^{-1}Z_1'] \otimes [Z_2'(Z_2 Z_2')^{-1}Z_2] \tag{4.7}$$

so that

$$\hat{\mu} = Z_1(Z_1'Z_1)^{-1}Z_1'YZ_2'(Z_2 Z_2')^{-1}Z_2. \tag{4.8}$$

Since

$$B = (Z_1'Z_1)^{-1}Z_1'\mu Z_2'(Z_2 Z_2')^{-1} \tag{4.9}$$

we have

$$\hat{B} = (Z_1'Z_1)^{-1}Z_1'\hat{\mu}Z_2'(Z_2 Z_2')^{-1} = (Z_1'Z_1)^{-1}Z_1'YZ_2'(Z_2 Z_2')^{-1}. \tag{4.10}$$

From (4.10), it follows quickly that

$$\operatorname{Cov}(\hat{B}) = (Z_1'Z_1)^{-1} \otimes [(Z_2 Z_2')^{-1}Z_2 \Sigma Z_2'(Z_2 Z_2')^{-1}]. \tag{4.11}$$

Now, we turn the situation around. Consider $Y \in \mathscr{L}_{p_1,p_2}$ and let N be a fixed, non-trivial, linear manifold in V_1. Assume $\operatorname{Cov}(Y) = I_{p_2} \otimes \Sigma$ where $\Sigma \in \Theta$ and Θ is given by

$$\Theta = \{\Sigma \mid \Sigma \in \mathscr{S}_{p_1}^+, \Sigma N = N\}. \tag{4.12}$$

Since $\Sigma \in \mathscr{S}_{p_1}^+$ is symmetric, $\Sigma N = N$ iff $\Sigma N^\perp = N^\perp$ where $N^\perp$ is the orthogonal complement of N in V_1. Given the above covariance structure, we would like to describe the regression manifolds Ω for which GM estimates exist. For the remainder of this section, proofs will be rather abbreviated as the arguments parallel those of corresponding statements in Section 3.

Let $\Omega \neq \{0\}$ be a linear manifold in $\mathscr{L}_{p_1,p_2}$ and assume

$$(I_{p_2} \otimes \Sigma)\Omega = \Omega \quad \text{for all} \quad \Sigma \in \Theta. \tag{4.13}$$

As before, (4.13) is equivalent to the existence of a GM estimator. Define $\mathscr{L}_{p_1}(N)$ by

$$\mathscr{L}_{p_1}(N) = \{B \mid B \in \mathscr{L}_{p_1,p_1}, BN \subseteq N \quad \text{and} \quad BN^\perp \subseteq N^\perp\}. \tag{4.14}$$

LEMMA 4.1. *If* (4.13) *holds, then*

$$(I_{p_2} \otimes B)\Omega \subseteq \Omega \quad \textit{for all} \quad B \in \mathscr{L}_{p_1}(N). \tag{4.15}$$

PROOF. The proof is similar to that of Lemma 3.1.

LEMMA 4.2. *Suppose* (4.13) *holds. Then*

(i) *there exists a rank one element in* Ω.
(ii) *if* $x_0 \in N$, $x_0 \neq 0$ *and* $z \otimes x_0 \in \Omega$, *then* $z \otimes x \in \Omega$ *for all* $x \in N$
(iii) *if* $x_1 \in N^\perp$, $x_1 \neq 0$ *and* $z \otimes x_1 \in \Omega$, *then* $z \otimes x \in \Omega$ *for all* $x \in N^\perp$.

PROOF. See the proof of Lemma 3.2.

Now, consider

$$W_1 = \{z \mid z \otimes x \in \Omega \quad \text{for all} \quad x \in N\} \tag{4.16}$$

and

$$W_2 = \{z \mid z \otimes x \in \Omega \quad \text{for all} \quad x \in N^\perp\}. \tag{4.17}$$

Clearly, W_1 and W_2 are linear manifolds and at least one of these manifolds is non-trivial. Let $M_0 = W_1 \cap W_2$, $M_1 = W_1 \cap M_0^{\perp}$, and $M_2 = W_2 \cap M_0^{\perp}$. Then M_0, M_1, and M_2 are mutually orthogonal manifolds in V_2 and at least one is non-trivial.

Let $\{z_1, \cdots, z_{r_0}, z_{r_0+1}, \cdots, z_{r_1}, z_{r_1+1} \cdots z_{r_2}, z_{r_2+1}, \cdots z_{p_2}\}$ be an ON basis for V_2 such that $M_0 = \operatorname{span}\{z_1, \cdots, z_{r_0}\}$, $M_1 = \operatorname{span}\{z_{r_0+1}, \cdots, z_{r_1}\}$, and $M_2 = \operatorname{span}\{z_{r_1+1}, \cdots, z_{r_2}\}$. Also, let $\{x_1, \cdots, x_{p_0}, x_{p_0+1}, \cdots, x_{p_1}\}$ be an ON basis for V_1 such that $N = \operatorname{span}\{x_1, \cdots, x_{p_0}\}$. Now, consider

$$S_0 = \operatorname{span}\{z_i \otimes x_j \mid i = 1, \cdots, r_0; j = 1, \cdots, p_1\},$$

$$S_1 = \operatorname{span}\{z_i \otimes x_j \mid i = r_0+1, \cdots, r_1; j = 1, \cdots, p_0\},$$

and

$$S_2 = \operatorname{span}\{z_i \otimes x_j \mid i = r_1+1, \cdots, r_2; j = p_0+1, \cdots, p_1\}.$$

Clearly S_0, S_1, S_2 are mutually orthogonal in $\mathscr{L}_{p_1,p_2}$ and $S_i \subseteq \Omega$, $i = 0, 1, 2$. Setting $\Omega_0 = S_0 \oplus S_1 \oplus S_2$, we now have

THEOREM 4.1. *If* (4.13) *holds, then* $\Omega = \Omega_0$.

PROOF. The proof is essentially the same as the proof of Theorem 3.1 and is omitted.

In coordinate form, Theorem 4.1 can be stated as

COROLLARY 4.1. *If* (4.13) *holds, then each element* $\mu \in \Omega$ *has the coordinate representation*

$$\mu = \begin{bmatrix} & B_0 & \\ \hline B_1 & \vdots & 0 \\ \hline 0 & \vdots & B_2 \\ \hline & 0 & \end{bmatrix} \tag{4.18}$$

where B_0 *is* $r_0 \times p_1$, B_1 *is* $r_1 \times p_0$, B_2 *is* $r_2 \times (p_1 - p_0)$ *and the bottom block of zeros is* $(p_2 - r_0 - r_1 - r_2) \times p_1$. *As* μ *ranges over all of* Ω, B_0, B_1, *and* B_2 *range over all matrices of their respective dimensions.*

PROOF. The representation (4.18) is simply the coordinate form for μ in the basis specified above. The second assertion is immediate from Theorem 4.1. This completes the proof.

As with (3.10) of Section 3, it is appropriate to refer to (4.18) as the "canonical" form for the multivariate linear model of this section. The justification for this is that when Θ is given by (4.12) and when a GM estimator exists, then one can choose a coordinate system so that μ has the form (4.18).

5. The intra-class correlation model. As in Example 2.1, consider a random matrix $Y: n \times p$ such that $\operatorname{Cov}(Y) = I_n \otimes \Sigma$. In some experimental situations, it is reasonable to assume that Σ has the form

$$\Sigma = \sigma^2[(1-\rho)I_p + \rho\, e \otimes e] \tag{5.1}$$

where $\sigma^2 > 0$, $-1/(p-1) < \rho < 1$, and e is the vector of 1's. A covariance matrix having the form (5.1) is called an intra-class correlation matrix. The purpose of this section is to present a result which describes all the regression manifolds for which GM estimators exist when Σ has the form (5.1).

Again consider $Y \in \mathscr{L}_{p_1,p_2}$ where $\operatorname{Cov}(Y) = I_n \otimes \Sigma$. Let $u_0 \neq 0$ be a fixed vector in V_1 and define the set Θ by

$$(5.2) \qquad \Theta = [\Sigma \mid \Sigma \in \mathscr{S}^+_{p_1}, \Sigma = c_1 I + c_2 u_0 \otimes u_0]$$

where $c_1 > 0$ and $c_1 + (u_0, u_0)_1 c_2 > 0$. The conditions on c_1 and c_2 are necessary and sufficient that $c_1 I + c_2 u_0 \otimes u_0$ be positive definite. Now, we assume that Σ varies over Θ where $\operatorname{Cov}(Y) = I_n \otimes \Sigma$. Let $\Omega \neq \{0\}$ be a linear manifold in $\mathscr{L}_{p_1,p_2}$ and assume

$$(5.3) \qquad (I_n \otimes \Sigma)\Omega = \Omega \quad \text{for all} \quad \Sigma \in \Theta,$$

so that a GM estimator exists for the regression manifold Ω.

In order to describe the implications of (5.3), let

$$(5.4) \qquad R_1 = \{B \mid B \in \mathscr{L}_{p_1,p_2}, Bu_0 = 0\} \qquad \text{and}$$

$$(5.5) \qquad R_2 = \{B \mid B \in \mathscr{L}_{p_1,p_2}, B = y \otimes u_0 \text{ for some } y \in V_2\}.$$

It is straightforward to show that R_1 and R_2 are orthogonal manifolds in $\mathscr{L}_{p_1,p_2}$ and $R_1 \oplus R_2 = \mathscr{L}_{p_1,p_2}$. Also, $\dim(R_1) = p_2(p_1-1)$ and $\dim(R_2) = p_2$. Setting $\Omega_i = R_i \cap \Omega$, $i = 1, 2$, note that $\Omega_1 \oplus \Omega_2 \subseteq \Omega$. Further, let $\tilde{\Omega}_i$ be the orthogonal projection of the manifold Ω onto the manifold R_i, $i = 1, 2$ and note that $\Omega \subseteq \tilde{\Omega}_1 \oplus \tilde{\Omega}_2$.

LEMMA 5.1. *If* $\tilde{\Omega}_2 \subseteq \Omega$, *then* $\Omega = \Omega_1 \oplus \Omega_2 = \tilde{\Omega}_1 \oplus \tilde{\Omega}_2$.

PROOF. Let P_2 be the orthogonal projection onto R_2, and assume $\tilde{\Omega}_2 \subseteq \Omega$. Then $\tilde{\Omega}_2 \subseteq \Omega \cap R_2 = \Omega_2$. Also, $P_2 A \in \Omega$ for all $A \in \Omega$ when $\tilde{\Omega}_2 \subseteq \Omega$ so that $(I-P_2)A \in \Omega$ for all $A \in \Omega$. Hence $\tilde{\Omega}_1 \subseteq \Omega$ since $I - P_2$ is the orthogonal projection onto Ω_1. Hence $\tilde{\Omega}_1 \subseteq \Omega \cap R_1 = \Omega_1$. Combining the above we have $\tilde{\Omega}_I \oplus \tilde{\Omega}_2 \subseteq \Omega_1 \oplus \Omega_2 \subseteq \Omega \subseteq \tilde{\Omega}_1 \oplus \tilde{\Omega}_2$ and the result follows.

LEMMA 5.2. *For* $A \in \mathscr{L}_{p_1,p_2}$,

$$(5.6) \qquad P_2 A = [(Au_0) \otimes u_0]/(u_0, u_0)_1.$$

PROOF. Let $\tilde{u}_0 = u_0(u_0, u_0)_1^{-\frac{1}{2}}$, and let $z_1, \cdots, z_{p_2}$ be an ON basis for V_2. Then $\{z_i \otimes \tilde{u}_0 \mid i = 1, \cdots, p_2\}$ is an ON basis for R_2. Thus

$$(5.7) \qquad \begin{aligned} P_2 A &= \textstyle\sum_{i=1}^{p_2} \langle A, z_i \otimes \tilde{u}_0 \rangle z_i \otimes \tilde{u}_0 = \sum_{i=1}^{p_2} (z_i, A\tilde{u}_0)_2 (z_i \otimes \tilde{u}_0) \\ &= \{\textstyle\sum_{i=1}^{p_2} (z_i, A\tilde{u}_0)_2 z_i\} \otimes \tilde{u}_0 = (A\tilde{u}_0) \otimes \tilde{u}_0. \end{aligned}$$

The result follows.

THEOREM 5.1. *If* (5.3) *holds, then*

$$(5.8) \qquad \Omega = \Omega_1 \oplus \Omega_2.$$

PROOF. We will show that $\tilde{\Omega}_2 \subseteq \Omega$ and apply Lemma 5.1. Let

$$M = \{z \otimes u_0 \,|\, Au_0 = z \quad \text{for some} \quad A \in \Omega\}. \tag{5.9}$$

Clearly $M \subseteq R_2$ is a manifold. Since $(Au_0) \otimes u_0 = A(u_0 \otimes u_0)$, $M \subseteq \Omega$ by assumption (5.3). However, it follows immediately from Lemma 5.2 that $M = \tilde{\Omega}_2$. This completes the proof.

We note that the converse of Theorem 5.1 is immediate. That is, if $\Omega_1{}^* \subseteq R_1$ and $\Omega_2{}^* \subseteq R_2$ are manifolds, then $\Omega^* \equiv \Omega_1{}^* \oplus \Omega_2{}^*$ satisfies (5.3). The application of Theorem 5.1 to the intra-class correlation model is obvious.

6. Discussion. With the aid of Theorem 1.2, the results in Section 3, Section 4 and Section 5 describe the form of the regression manifolds which admit GM estimators for three different covariance structures. The covariance structures considered here represent only a small portion of such structures which arise in practice. However, the methods used above should carry over without difficulty to other cases and yield useful descriptions of regression manifolds which admit GM estimators. Other covariance structures of interest include the block diagonal form with more than two blocks and the circular covariance matrix as discussed by Olkin and Press (1969).

REFERENCES

[1] ANDERSON, T. W. (1958). *Introduction to Multivariate Statistical Analysis.* Wiley, New York.

[2] BELLMAN, R. (1960). *Introduction to Matrix Analysis.* McGraw Hill, New York.

[3] GLESER, L. J. and OLKIN, I. (1966). Linear models in multivariate analysis. Stanford Univ. Technical Report No. 21.

[4] HALMOS, P. (1958). *Finite Dimensional Vector Spaces.* Van Nostrand, Princeton.

[5] KRUSKAL, W. (1961). The coordinate-free approach to Gauss–Markov estimation, and its application to missing and extra observations. *Fourth Berkeley Symp. Math. Statist. Prob.* **1** 435–451.

[6] KRUSKAL, W. (1968). When are Gauss–Markov and least squares estimators the same? A coordinate free approach. *Ann. Math. Statist.* **39** 70–75.

[7] MCELROY, F. W. (1967). A necessary and sufficient condition that ordinary least squares estimators be best linear unbiased. *J. Amer. Statist. Assoc.* **62** 1302–1305.

[8] OLKIN, I. (1966). Special topics in matrix theory and inequalities (mimeographed notes, recorded by M. L. Eaton). Department of Statistics, Stanford Univ.

[9] OLKIN, I. and PRESS, S. J. (1969). Testing and estimation for a circular stationary model. *Ann. Math. Statist.* **40** 1358–1373.