PREPRINT

MAR 1972

3

Morris L. Eaton

A Note on the Gauss-Markov Theorem

UNIVERSITY OF COPENHAGEN
INSTITUTE OF
MATHEMATICAL STATISTICS

Morris L. Eaton[1]

A Note on the Gauss-Markov Theorem

Preprint 1972 No. 3

INSTITUTE OF MATHEMATICAL STATISTICS
UNIVERSITY OF COPENHAGEN

March 1972

[1]Currently on leave from the University of Chicago.

§1: Introduction and Notation:

The purpose of this note is to give a necessary and sufficient condition for the existence of a Gauss-Markov (G-M.) estimator of a mean vector for a general linear model to be defined below. A careful examination of Theorem 3 in Kruskal (1968) suggests that the proper context for G-M estimation is within finite dimensional vector spaces without inner products, even though Kruskal's proof does involve an inner product. Using results due to Kruskal (1968), Eaton (1970) gave necessary and sufficient conditions for the existence of a G-M estimator when the covariance operators in the linear model were non-singular.

The conditions given in the current paper do not require non-singularity of the covariance operators. In fact, the arguments presented here show that the singularity or non-singular of the covariance operators in the linear model is really irrelevant, and it is the introduction of an inner product which leads one to worry about the singularity of covariance operators.

It is assumed that the reader is familiar with finite dimensional vector space theory as presented in Halmos(1958) or other comparable texts. Throughout, V will be a finite dimensional real vector space, V' will denote the dual space of V, and the value of the linear functional $\xi \in V'$ at $x \in V$ is $[\xi,x]$. The usual canonical identification of V'' with V is assumed so the value of $x \in V(= V'')$ at $\xi \in V'$ is $[x,\xi] \equiv [\xi,x]$. Let $Y \in V$ be a random vector with $\mathcal{E}[\xi,Y]^2 < +\infty$ for all $\xi \in V'$. Then the mean vector of Y, say $\mu \equiv \mathcal{E}Y$, exists and the covariance operator of Y, say $\Sigma \equiv \mathrm{Cov}(Y)$ exists.

By definition $[\xi,\mu] = E[\xi,Y]$ for $\xi \in V'$, and Covariance$\{[\xi_1,Y],[\xi_2,Y]\} = [\xi_1,\Sigma\xi_2] \equiv \mathrm{Cov}\{[\xi_1,Y],[\xi_2,Y]\}$. Thus Σ is a linear transforamtion (l.t.) on V' to V and $\Sigma \geqq 0$; that is, Σ is self adjoint and positive semi-definite. If $A : V \to W$ is a linear transformation, A' is the adjoint of A, $R(A)$ is the range of A, and $N(A)$ is the null space of A. For a linear manifold $M \subseteq V$, $M^o \subseteq V'$ denotes the annihilator of M. In this notation, a result which is often used in this paper is $(R(A))^o = N(A')$ (Halmos (1958)). Suppose $\Sigma : V' \to V$ and $\Sigma \geqq 0$. Let M be a linear manifold in V.

Then

$$\Sigma(M^o) \cap M = \{0\}\ . \tag{1.1}$$

To see this, first note that $N(\Sigma) = \{\xi \mid [\xi,\Sigma\xi] = 0\}$.

Now, $u \in M$ iff $[\xi,u] = 0$ for all $\xi \in M^o$; and $u \in \Sigma(M^o)$ implies $u = \Sigma\xi_1$ for $\xi_1 \in M^o$. Hence $u = \Sigma\xi_1 \in \Sigma(M^o) \cap M$ implies $[\xi_1,\Sigma\xi_1] = 0$ so $u = \Sigma\xi_1 = 0$.

§ 2: The Simple Linear Model:

By a simple linear model we mean a random vector $Y \in V$ such that: (a) the mean vector of Y, μ, ranges over a fixed linear manifold M and (b) $\mathrm{Cov}(Y) = \Sigma_1$ where $\Sigma_1 \geqq 0$ is a fixed known linear transformation.

Given the simple linear model (V, M, Σ_1), define the set $\acute{A}$ of linear transformations on V to V by

$$\acute{A} = \{A \mid A : V \to V, \quad Ax = x, \quad x \in M\}. \tag{2.1}$$

Clearly, if $B : V \to V$, then BY is an unbiased estimator of μ iff $B \in \acute{A}$.

<u>Definition 2.1</u>: $A_1 \in \acute{A}$ is a <u>Gauss-Markov Operator</u> (G-M.O.) iff

$$[\xi, A_1 \Sigma_1 A'_1 \xi] \leqq [\xi, A \Sigma_1 A \xi] \tag{2.2}$$

for all $\xi \in V'$ and $A \in \acute{A}$.

Since $[\xi, A\Sigma_1 A' \xi] = \text{Variance } ([\xi, AY])$, $A_1 \in \acute{A}$ is a G-M.O. if A_1 minimizes $\mathrm{Var}\,[\xi, AY]$ (over $\acute{A}$) for all $\xi \in V'$. This is the usual interpretation of a G-M estimator.

<u>Theorem 2.1</u>: Let $N \subseteqq V$ be a linear manifold complimentary to M such that $N \supseteqq \Sigma_1(M^o)$ and let P_1 denote the projection on M along N.

Then P_1 is a G-M.O.

<u>Proof</u>: By assumption, $\acute{N}(P_1) \supseteqq \Sigma_1(M^o)$. Since $R((I-P_1)') = [\acute{N}(I-P_1)]^o = M^o$, $P_1 \Sigma_1 (I - P_1)' = 0$ iff $P_1 x = 0$ for all $x \in \Sigma_1(M^o)$. Thus $\acute{N}(P_1) \supseteqq \Sigma_1(M^o)$ is equivalent to

$$P_1 \Sigma_1 (I - P_1)' = 0. \tag{2.3}$$

For $A \in \acute{A}$, $AP_1 = P_1$ so $A = AP_1 + A(I - P_1) = P_1 + A(I - P_1)$. Thus

$$\mathrm{Var}\ [\xi, AY] = \mathrm{Var}\ [\xi, P_1 Y + A(I - P_1)Y] = \mathrm{Var}\ [\xi, P_1 Y] +$$

$$2\,[\xi, P_1 \Sigma_1 (I - P_1)' A' \xi] + \mathrm{Var}\ [\xi, A(I - P_1)Y] \quad . \tag{2.4}$$

From (2.3), we have

$$\mathrm{Var}\ [\xi, AY] \geqq \mathrm{Var}\ [\xi, P_1 Y] \tag{2.5}$$

with equality iff $\mathrm{Var}\ [\xi, A(I - P_1)Y] = 0$ for all $\xi \in V'$ .//

Theorem 2.2: $A_1 \in \acute{A}$ is a G-M.O. iff $\acute{N}(A_1) \supseteqq \Sigma_1(M^o)$.

Proof: From the condition for equality in (2.5), A_1 is a G-M.O. iff $\mathrm{Var}\,[\xi, A_1(I - P_1)Y] = 0$ iff $A(I - P_1)\Sigma_1(I - P_1)'A' = 0$. Since $\Sigma_1 \geqq 0$, A_1 is a G-M.O.

$$A_1(I - P_1)\Sigma_1 = 0.$$

But (2.6) holds iff $\acute{N}(A_1) \supseteqq R((I - P_1)\Sigma_1)$. However, $\acute{R}((I - P_1)\Sigma_1) = \Sigma_1(M^o)$ since $\acute{R}(P_1) = M$ and $\acute{N}(P_1) \supseteqq \Sigma_1(M^o)$.//

Corollary 2.1: A G-M.O. is unique iff $M + \Sigma_1(M^o) = V$.

Corollary 2.2: $A_1 \in \acute{A}$ is a G-M.O. iff $A_1 Y$ and $(I - P_1)Y$ are uncorrelated.

Proof: $A_1 Y$ and $(I - P_1)Y$ are uncorrelated iff $A_1 \Sigma (I - P_1)' = o$ iff

$$\acute{N}(A_1) \supseteqq \acute{R}(\Sigma_1(I - P_1)') = \Sigma_1(M^o).$$

The conclusion follows from Theorem 2.2.//

<u>Corollary 2.3</u>: If $\Sigma_1 : V' \to V$ is non-singular, then the G-M.O, A_1, is unique. Further A_1 is the projection on M along $\Sigma_1(M^o) = [\Sigma_1^{-1}(M)]^o$.

<u>Proof</u>. From the previous results, we must only establish that $\Sigma_1(M^o) = [\Sigma_1^{-1}(M)]^o$ and to do this, it will be shown that $(\Sigma_1(M^o))^o = \Sigma_1^{-1}(M)$. Now $u \in (\Sigma_1(M^o))^o$ iff $[u,\Sigma_1\eta] = 0$ for all $\eta \in M^o$ iff $[\Sigma_1 u,\eta] = 0$ for all $\eta \in M^o$. But $[\Sigma_1 u,\eta] = 0$ for all $\eta \in M^o$ iff $u \in \Sigma_1^{-1}(M)$.//

<u>Remark</u>: An alternative interpretation of a G-M.O. is the following. Let $C_o : V \to V'$ be any positive definite self-adjoint linear transformation. Consider the simple linear model (V,M,Σ_1). Define a bilinear function T on pairs of linear transformations A,B (mapping V to V) by

$$T(A,B) = \acute{E}_\mu [C_o A(Y - \mu), B(Y - \mu)] \ . \qquad (2.7)$$

It is easy to show that if AY and BY are incorrelated, then $T(A,B) = 0$. Further, $T(A,A) = 0$ iff $A\Sigma_1 A' = 0$. The following result is not hard to prove.

<u>Theorem 2.3</u>: $A_o \in \acute{A}$ is a G-M.O. iff $T(A_o,A_o) \leqq T(A,A)$ for all $A \in \acute{A}$.

§ 3: The General Linear Model:

By a general linear model we mean a random vector $Y \in V$ such that: (a) $E(Y) = \mu$ ranges over a fixed linear manifold M, and (b) $\mathrm{Cov}(Y) = \Sigma$ ranges over a fixed set Γ of positive semi-definite linear transformations on V' to V.

Definition 3.1: $A_o \in A$ is a G-M.O. for the model (V,M,Γ) iff A_o is a G-M.O. for each of the simple linear models (V,M,Σ), $\Sigma \in \Gamma$.

With this definition, the following results are obvious.

Theorem 3.1: A G-M.O. A_o exists for the model (V,M,Γ) iff there exists a manifold $N \subseteqq V$ complementary to M such that $\Sigma(M^o) \subseteqq N$ for all $\Sigma \in \Gamma$.

Theorem 3.2: Consider the linear model (V,M,Γ) and suppose each $\Sigma \in \Gamma$ is non-singular. A G-M.O. exists iff $\Sigma_1^{-1}M = \Sigma_2^{-1}M$ for all $\Sigma_1, \Sigma_2 \in \Gamma$.

Note that Theorem 3.2 was given by Eaton (1970) in the context of inner product spaces. Also, if one is working with an inner product space, then M^o becomes $M^\perp$ - the orthogonal complement of M, and the conditions in the above theorems are then in terms of $\Sigma(M^\perp)$.

In particular, when the identity operator on V to V is in Γ, then application of Theorem 3.1 shows that a G-M.O. exists iff $\Sigma(M^\perp) \subseteqq M^\perp$ for all $\Sigma \in \Gamma$. But $\Sigma(M^\perp) \subseteqq M^\perp$ is equvivalent to $\Sigma(M) \subseteqq M$ since each Σ is self adjoint. Hence we have

Theorem 3.3: Consider the linear model (V,M,Γ) where V is an inner product space with inner product $(\cdot,\cdot)$ and $\Sigma : V \to V$, $\Sigma \in \Gamma$, is self adjoint and positive semi-definite with respect to $(\cdot,\cdot)$. If $I \in \Gamma$, then a G-M.O. exists iff

$$\Sigma(M) \subseteqq M \quad \text{for all } \Sigma \in \Gamma. \tag{3.1}$$

Applications of the results in this section include those given by Eaton (1970) and Kruskal (1968). Also, the references in these two papers contain many examples to which the above results are directly applicable.

References.

Eaton, M.L. (1970). Gauss-Markov estimation for multivariate linear models: A coordinate free approach. *Ann. Math. Statist.* *41*, 528-538.

Halmos, P. (1958). *Finite Dimensional Vector Spaces.* Van Nostrand, Princeton.

Kruskal, W.K. (1968). When are Gauss-Markov and least squares estimators the same? A coordinate free approach. *Ann. Math. Statist.* *39*, 70-75.