The covariance matrix of a multivariate locally best unbiased estimator
Sander, P.C.

Published: 01/01/1979

Document Version
Publisher’s PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:
• A submitted manuscript is the author’s version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher’s website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

Citation for published version (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal?

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
The covariance matrix of a multivariate locally best unbiased estimator

by Peter C. Sander

Eindhoven University of Technology
Netherlands
April 1979

BDK/ORS-79/3
SUMMARY

The covariance matrix of a multivariate locally best unbiased estimator

Barankin (1949) has given the variance of a univariate locally best unbiased estimator. In this note the covariance matrix is given of a multivariate locally best unbiased estimator. The Cramér-Rao and the Chapman-Robbins-Kiefer bounds appear as special cases.
1. INTRODUCTION

Let $X$ be a random variable defined on a measure space $(X, \mathcal{A})$. Let $P = \{p(\cdot | \theta) | \theta \in \Theta\}$ be a family of probability densities of $X$ with respect to the sigma finite measure $\mu$; $\Theta$ is an index set. Define $g(\theta) := (g_1(\theta), \ldots, g_m(\theta))^t$ with $g_i : \Theta \rightarrow \mathbb{R}$ ($i = 1, \ldots, m$).

Estimates of $g(\theta)$ are denoted by $\delta(x) = (\delta_1(x), \ldots, \delta_m(x))^t$ with $\delta_i : X \rightarrow \mathbb{R}$ ($i=1, \ldots, m$) $\mu$-measurable; $\delta(X)$ is the corresponding estimator.

The class of unbiased estimators $\delta(X)$ of $g(\theta)$ is denoted by $Z$, while $Z(\theta^0), \theta^0 \in \Theta$, denotes the subclass of $Z$ with the property $\operatorname{E}(\delta_i(X)|\theta^0) < \infty$ ($i = 1, \ldots, m$).

**Definition 1.** $\delta^0(X)$ is a locally best unbiased estimator (LBUE) of $g(\theta)$ in $\theta^0$, if the following is true for every $a = (a_1, \ldots, a_m)^t \in \mathbb{R}^m$

i) $\delta^0(X) \in Z(\theta^0)$

and

ii) $\operatorname{var}(\sum_{i=1}^{m} a_i \delta_i(X)|\theta^0) \leq \operatorname{var}(\sum_{i=1}^{m} a_i \delta_i(X)|\theta^0)$ for all $\delta \in Z$. \hfill $\Box$

Let the covariance matrix of $\delta(X) \in Z$ with respect to $p(\cdot | \theta)$ be given by $\Sigma^\delta(\theta) = \left((\sigma^2_{ij}(\theta))\right)$. If $\delta^0(X) \in Z(\theta^0)$, then $\delta^0(X)$ is a LBUE iff

$$
(1) \quad \sum_{i,j=1}^{m} a_i a_j \sigma^2_{ij}(\theta^0) \leq \sum_{i,j=1}^{m} a_i a_j \sigma^2_{ij}(\theta^0) \quad \text{for all } a \in \mathbb{R}^m, \delta \in Z.
$$

Formula (1) means $\Sigma^\delta(\theta^0) \leq \Sigma^\delta(\theta^0)$ is positive semi-definite, i.e. $\Sigma^\delta(\theta^0) \preceq \Sigma^\delta(\theta^0)$.

The following results are of immediate use (cf. Rao (1973) pp. 317-318).

---


Key words and phrases. Locally best unbiased estimator, covariance matrix, Cramer-Rao bound.
Theorem 1. Assume $m = 1$ and $\delta^0(X) \in Z(\theta^0)$. Then $\delta^0(X)$ is a LBUE of $g(\theta)$ in $\theta^0$ iff 
$\text{cov}(\delta^0(X), \delta(X) | \theta^0) = 0$ for all $\delta(X) \in Z(\theta^0)$.

Corollary 1

i) $\delta(X)$ is a LBUE of $g(\theta)$ in $\theta^0$ iff $\delta_i(X)$ is a LBUE of $g_i(\theta)$ in $\theta^0$ 
$(i = 1, \ldots, m)$

ii) if there are two LBUEs of $g(\theta)$ in $\theta^0$, they are the same except for 
a set of $p(\cdot | \theta^0)$ measure zero.

From this corollary it follows that LBUEs can be constructed coordinate wise, following Barankin (1949).

2. LINEAR TRANSFORMATIONS

Let $L^2(X, \nu)$ be the family of all square integrable real valued functions 
with respect to a $\sigma$-finite measure $\nu$ on $X$, with inner product

$$(f, g) := \int_X f g \, d\nu$$

and norm $||f|| = (f, f)^{\frac{1}{2}}$.

By Riesz' representation theorem with every bounded linear functional 
$H_0 : L^2(X, \nu) \to \mathbb{R}$ there corresponds an adjoint function $h_0 \in L^2(X, \nu)$, 
such that

$$(2) \quad H_0(g) = \int_X g h_0 \, d\nu, \text{ for all } g \in L^2(X, \nu).$$

Let $H(g) := (H_1(g), \ldots, H_m(g))^t$ with $H_i (i = 1, \ldots, m)$ a bounded linear functional with adjoint $h_i \in L^2(X, \nu)$. Then we will call $H : L^2(X, \nu) \to \mathbb{R}^m$ 
a bounded linear functional as well, with adjoint $H^t := (h_1, \ldots, h_m)$. 
Lemma 1. Let \( H : L^2(X, \nu) \to \mathbb{R}^m \) be a bounded linear functional with adjoint \( H^* = (h_1, \ldots, h_m) \).

Define

\[
S := \left( \left\{ \int_X h_i h_j \, d\nu \right\} \right)_{i,j=1}^m.
\]

Then

i) \( S \geq H(x)[H(x)]^{t/2} ||x||^2 \), for all \( x \in L^2(X, \nu), x \neq 0 \)

ii) if there is an \( m \times m \) matrix \( A \) with

\[
A \geq H(x)[H(x)]^{t/2} ||x||^2 \), for all \( x \in L^2(X, \nu), x \neq 0, \)

then

\( A \geq S. \)

Proof

i) Let \( a \in \mathbb{R}^m \). By Schwarz' inequality

\[
a^t S a \int_X ||x||^2 \, d\nu \geq \left[ \sum_i a_i \int_X x h_i \, d\nu \right]^2 \text{ for all } x \in L^2(X, \nu).
\]

Hence

\[
a^t S a \geq a^t H(x)[H(x)]^{t/2} a ||x||^2 \text{ for all } x \in L^2(X, \nu), x \neq 0.
\]

ii) If (3) then

\[
a^t A a - a^t S a + a^t S a - a^t H(x)[H(x)]^{t/2} a ||x||^2 \geq 0
\]

for all \( x \in L^2(X, \nu), x \neq 0 \) and all \( a \in \mathbb{R}^m \).

From (4) it follows that for an arbitrary but fixed \( a \in \mathbb{R}^m \) and

\[
x = \sum_{i=1}^m a_i h_i
\]
\[ |x|^{2} a^t S a = a^t H(x)[H(x)]^t a. \]

Hence, for this choice of \( x \),

\[ a^t A a - a^t S a \geq 0. \]  

Since (6) is true for all \( a \in \mathbb{R}^m \) the proof is complete. \( \square \)

**Notes**

1. Since \( S = (\int_{\mathcal{X}} h_i h_j dv)_{i,j=1}^{m} \), it is logical to denote \( S \) by \( HH^t \). This will be done in what remains.

2. If a matrix \( S \) satisfies i) and ii) of lemma 1, the following notation will be used

\[ S = \sup_{x \in L^2(\mathcal{X}, \nu)} H(x)[H(x)]^t / |x|^2. \]

Let \( K \) be a linear subspace of \( L^2(\mathcal{X}, \nu) \) and \( K^\perp \) the orthogonal complement of \( K \) in \( L^2(\mathcal{X}, \nu) \). Let \( F : K \to \mathbb{R}^m \) be a bounded linear functional.

Define the bounded linear extension \( F_e \) of \( F \) by

\[ F_e(x_1 + x_2) := F(x_1) \text{ for all } x_1 \in K, x_2 \in K^\perp. \]

**Lemma 2.** Let \( K \) be a linear subspace of \( L^2(\mathcal{X}, \nu) \) and let \( F : K \to \mathbb{R}^m \) be a bounded linear functional with extension \( F_e \).

Then

\[ F_e F_e^t = \sup_{x \in K} F(x)[F(x)]^t / |x|^2. \]

**Proof**

By Lemma 1

\[ F_e F_e^t \geq F_e(x)[F_e(x)]^t / |x|^2 \text{ for all } x \in L^2(\mathcal{X}, \nu), x \neq 0. \]
Especially

\[ F_e^t F_e^t \geq F(x_1)[F(x_1)]^t / ||x_1||^2 \text{ for all } x_1 \in K, x_1 \neq 0. \]

Assume there is an \( m \times m \) matrix \( T \) such that

\[ T \geq F(x_1)[F(x_1)]^t / ||x_1||^2 \text{ for all } x_1 \in K, x_1 \neq 0. \]

Then

\[ T \geq F_e(x_1)[F_e(x_1)]^t / ||x_1||^2 \text{ for all } x_1 \in K, x_1 \neq 0. \]

and hence, since \( ||x|| = ||x_1 + x_2|| \geq ||x_1|| \) (\( x_1 \in K, x_2 \in K^t \)),

\[ T \geq F_e(x)[F_e(x)]^t / ||x||^2 \text{ for all } x \in L^2(X, \nu), x \neq 0. \]

The result follows from lemma 1.

\[ \square \]

3. THE COVARIANCE MATRIX OF A LBUE

Define the measure \( \nu \) by

\[ \nu(A) = \int_A p(.|\theta^0) d\mu \text{ for all } A \in A. \]

Assume from now on:

I. \( p(x|\theta)/p(x|\theta^0) \) is defined \( \mu - a.e. \) for all \( \theta \in \Omega \) and this quotient is denoted by \( \pi(x|\theta) \)

II. \( \pi(.|\theta) \in L^2(X, \nu) \) for all \( \theta \in \Omega \).

Define \( K \) as the linear subspace of \( L^2(X, \nu) \) spanned by finite linear combinations of the \( \pi(.|\theta), \theta \in \Omega \).
Lemma 3. Let $F : \{\pi(., \theta) | \theta \in \Omega\} \rightarrow \mathbb{R}^m$ be defined by $F(\pi(., \theta)) = g(\theta)$. Then the propositions i) and ii) are equivalent:

i) there is a bounded linear transformation $A : L^2(\chi, \nu) \rightarrow \mathbb{R}^m$ such that $A(\pi(., \theta)) = F(\pi(., \theta))$.

ii) there is a positive semidefinite matrix $M$ such that

$$\sum_{i=1}^{n} a_i F(\pi(., \theta^i)) \sum_{i=1}^{n} a_i F(\pi(., \theta^i))^t \leq M \left| \sum_{i=1}^{n} a_i \pi(., \theta^i) \right|^2$$

for all $n \in \mathbb{N}$, $a = (a_1, \ldots, a_n) \in \mathbb{R}^n$ and $\theta^1, \ldots, \theta^n \in \Omega$.

Furthermore, let $z = \sum_{i=1}^{n} a_i \pi(., \theta^i)$ and define $\phi : K \rightarrow \mathbb{R}^m$ by

$$\phi(z) := \sum_{i=1}^{n} a_i F(\pi(., \theta^i)).$$

iii) If ii), then $\theta$ is well-defined. Furthermore $M_0 := \phi_e \phi_e^t$ satisfies (8) and $M \succeq M_0$.

Proof

i) $\Rightarrow$ ii). Choose $M = A A^t$.

iii) Assume $\hat{z} = \sum_{i=1}^{n} a_i \pi(., \theta^i)$ and $z = \hat{z}$. Then by (8)

$$[\phi(z) - \phi(\hat{z})][\phi(z) - \phi(\hat{z})]^t \leq M \left| z - \hat{z} \right| = 0,$$

hence $\phi$ is well-defined. Since $\phi$ is a bounded linear transformation the results follow from lemma 2.

ii) $\Rightarrow$ i). $\phi_e$ meets the requirements of i).

We are now able to give the covariance matrix of a multivariate locally best unbiased estimator.
Theorem 2. Let $\phi$, $\phi_0$, and $M_0$ be defined as in lemma 3. Then

i) $Z(\phi_0) \neq \emptyset$ iff

$$
(10) \quad (\sum_{i=1}^{n} a_i g(\xi^i)) (\sum_{i=1}^{n} a_i g(\xi^i))^t \leq M_0 ||\sum_{i=1}^{n} a_i \pi(\cdot | \xi^i)||^2
$$

for all $n \in \mathbb{N}$, $a = (a_1, \ldots, a_n) \in \mathbb{R}^n$, $\xi^1, \ldots, \xi^n \in \Omega$.

ii) If $\delta(x) \in Z(\phi_0)$ then $\Sigma^{\delta}(\phi_0) \geq M_0$.

iii) $Z(\phi_0) \neq \emptyset$ iff there is one and only one $(p(\cdot | \phi_0), q, \xi) : \delta_0(x) \in Z(\phi_0)$ with $\Sigma^{\delta}(\phi_0) = M_0$.

Proof:

i) Assume $\delta(x) \in Z(\phi_0)$ (the "only if" part). Then, in an obvious notation,

$$
\sum_{i=1}^{n} a_i g(\xi^i) = \int_X \sum_{i=1}^{n} a_i \pi(\cdot | \xi^i) dv.
$$

Hence

$$
(11) \quad u^t (\sum_{i=1}^{n} a_i g(\xi^i)) (\sum_{i=1}^{n} a_i g(\xi^i))^t u
$$

$$
= u^t \left( \int_X \sum_{i=1}^{n} a_i \pi(\cdot | \xi^i) dv \right) \left( \int_X \sum_{i=1}^{n} a_i \pi(\cdot | \xi^i) dv \right)^t u
$$

$$
= \left( \int_X (\sum_{i=1}^{m} b_i \xi_j) (\sum_{i=1}^{n} a_i \pi(\cdot | \xi^i) dv) \right)^2
$$

$$
\leq \int_X (\sum_{i=1}^{m} b_i \xi_j)^2 dv \int_X (\sum_{i=1}^{n} a_i \pi(\cdot | \xi^i))^2 dv
$$

$$
= u^t \Sigma^{\delta}(\phi_0) u ||\sum_{i=1}^{n} a_i \pi(\cdot | \xi^i)||^2
$$

for all $n \in \mathbb{N}$, $a = (a_1, \ldots, a_n) \in \mathbb{R}^n$, $\xi^1, \ldots, \xi^n \in \Omega$. 
From lemma 3 it follows that (10) is true for $M = M_0$.

Now the "if" part. By Riesz' representation theorem there is a 
$
\delta^0 := (\delta^0_1, \ldots, \delta^0_m)^t \text{ with } \delta^0_i \in L^2(\mathcal{X}, \nu) \ (i = 1, \ldots, m) \text{ such that }
$

\begin{equation}
\phi(\pi(\cdot|\theta)) = g(\theta) \text{ from (12) it follows that } \delta^0(\pi) \in Z(\theta^0).
\end{equation}

ii) This follows immediately from (11) and lemma 3.

iii) If $Z(\theta^0) \neq \emptyset$ then $\delta^0$ defined by (12) meets the requirements.

The unicity is a result of corollary 1.

4. THE CRAMÉR-RAO BOUND

The Cramér-Rao bound is a special case of theorem 2. This can be seen from what follows.

Theorem 3. Assume

i) $Z(\theta^0) \neq \emptyset$

ii) $\Omega \subseteq \mathbb{R}^k$, $\Omega$ is open

iii) $p(x|\theta) \neq 0$ $\mu$-ae for all $\theta \in \Omega$

iv) if $\theta = (\theta_1, \ldots, \theta_k)^t \in \Omega$, $\theta(i,h):= (\theta_1, \ldots, \theta_{i-1}, \theta_{i+h}, \theta_{i+1}, \ldots, \theta_k)^t$

and $h, a_1, \ldots, a_k \in \mathbb{R}$, then

$$
\lim_{h \to 0} \int_\mathcal{X} \sum_{i=1}^k \frac{\left\{ \sum_{i=1}^k (\pi(\cdot|\theta) - \pi(\cdot|\theta(i, h)))/h \right\}^2 \, d\nu}{h}
$$

$$
= \int_\mathcal{X} \sum_{i=1}^k \frac{\left\{ \lim_{h \to 0} \sum_{i=1}^k (\pi(\cdot|\theta) - \pi(\cdot|\theta(i, h)))/h \right\}^2 \, d\nu}{h}
$$

v) $M_0$ is positive definite

vi) $F := \left( \int_\mathcal{X} (\partial^2 \pi(\cdot|\theta)/\partial \theta_i \partial \theta_j) (\partial \pi(\cdot|\theta)/\partial \theta_i) \, d\nu \right)^k_{i, j=1; \theta = \theta_0}$
and

\[ G := \begin{pmatrix} \frac{\partial g_1(\theta)}{\partial \theta_1} & \cdots & \frac{\partial g_1(\theta)}{\partial \theta_k} \\ \vdots \\ \frac{\partial g_m(\theta)}{\partial \theta_1} & \cdots & \frac{\partial g_m(\theta)}{\partial \theta_k} \end{pmatrix} \]

are well-defined and \( F \) is nonsingular.

Then

\[ M_0 \succeq GF^{-1}C^t. \]

**Proof.** By theorem 2

\[ u^t (\sum_{i=1}^n a_{ij} g(\theta^i)) (\sum_{i=1}^n a_{ij} g(\theta^i))^t u \leq u^t M_0 u \left\| \sum_{i=1}^n a_{ij} \pi(\cdot, \theta^i) \right\|^2 \]

for all \( n \in \mathbb{N}, \ a = (a_1, \ldots, a_n) \in \mathbb{R}^n, \ \theta^1, \ldots, \ \theta^n \in \Omega, \ u \in \mathbb{R}^m. \)

Especially for \( a \) and \( \theta^1, \ldots, \ \theta^n \) fixed

(13) \[ \sup_{u \in \mathbb{R}^m} u^t (\sum_{i=1}^n a_{ij} g(\theta^i)) (\sum_{i=1}^n a_{ij} g(\theta^i))^t u / (u^t M_0 u) \leq \left\| \sum_{i=1}^n a_{ij} \pi(\cdot, \theta^i) \right\|^2. \]

Note that there is an \( m \times m \) matrix \( B \) such that \( M_0 = B^t B. \)

Define

\[ v := \sum_{i=1}^n a_{ij} g(\theta^i) \]
\[ p := (B^{-1})^t v \]
\[ g := Bu. \]
From
\[(p^t q)^2 \leq (p^t p)(q^t q)\]
it follows that
\[(v^t u)^2 \leq (v^t (B^t B)^{-1} v)(u^t B^t B u),\]
hence
\[(14) \sup_{u} u^t v v^t u/(u^t M_0 u) = v^t M_0^{-1} v.\]

From (13) and (14) it follows that
\[n \sum_{i=1}^{n} a_i \pi(\cdot \mid \theta^i) \leq \left(\sum_{i=1}^{n} a_i \pi(\cdot \mid \theta^i)\right)^2 \leq \left(\sum_{i=1}^{n} a_i \pi(\cdot \mid \theta^i)\right)^2 \]
for all \(a = (a_1, \ldots, a_n) \in \mathbb{R}^n\) and \(\theta^1, \ldots, \theta^n \in \Omega\).

In particular
\[\left(\sum_{i=1}^{k} a_i \pi(\cdot \mid \theta^0) - \pi(\cdot \mid \theta^0(i,h))\right)M_0^{-1} \left(\sum_{i=1}^{k} a_i \pi(\cdot \mid \theta^0) - \pi(\cdot \mid \theta^0(i,h))\right)/h \]
\[\leq \left(\sum_{i=1}^{k} a_i \pi(\cdot \mid \theta^0) - \pi(\cdot \mid \theta^0(i,h))\right)h.\]

Hence with iv)
\[a^t G M_0^{-1} G a \leq \int_{X} \left(\sum_{i=1}^{k} a_i \partial \pi(\cdot \mid \theta)/\partial \theta_i\right)_{\theta^0}^2 dv \]
\[= a^t F a \quad \text{for all } a \in \mathbb{R}^k.\]

Let \(I_n\) be the identity matrix of order \(n\), and \(B\) an arbitrary \(m \times k\) matrix. Note that
\[B^t B \leq I_k \iff B B^t \leq I_m.\]
From (15) it follows that
\[ G^t M_0^{-1} G \leq F, \]
hence
\[ F^{-\frac{1}{2}} G^t M_0^{-\frac{1}{2}} M_0^{-\frac{1}{2}} G F^{-\frac{1}{2}} \leq I_k. \]
The result now follows from (16).

5. **FINAL REMARKS**

In Lemma 3 \( M_0 \) is defined by \( M_0 = \phi e^t \), in other words
\[ M_0 = \sup_{n, a, \theta} \frac{n}{\left( \sum a_i g(\theta_i)(\sum a_i g(\theta_i))^t \right)} \frac{n}{\left( \sum a_i n(\theta_i) \right)^2}. \]
It is easily seen that (17) is equivalent to
\[ M_0 = \sup_{\omega_1 \neq \omega_2} \frac{[E_1 g(V) - E_2 g(V)][E_1 g(V) - E_2 g(V)]^t}{\int_X \{ \int_{\Omega} \frac{p(x|v) \mu(\omega_1(v) - \omega_2(v))^2}{p(x|\theta^0) \mu(x)} d\mu(x) \}} \]
with
\[ E_1 g(V) = \int_{\Omega} g(v) \omega_1(v) \]
and in which the supremum is taken over all measures \( \omega_1 \) and \( \omega_2 \) with \( \omega_1 \neq \omega_2 \) and such that the integrand in the denominator is defined \( \mu - a.e. \).

In particular for \( \omega_1 \) concentrated in \( \theta^0 \):
This last result is a logical generalization of the well-known one-dimensional Chapman-Robbins-Kiefer theorem (cf. Chapman and Robbins (1951) and Kiefer (1952)).
REFERENCES


Department of Industrial Engineering
Eindhoven University of Technology
5600 MB Eindhoven
Netherlands