Two Explicit Characterizations of the General Nonnegative-Definite Covariance Matrix Structure for Equality of BLUEs, WLSEs, and LSEs

Phil D. Young\textsuperscript{1}, Joshua D. Patrick\textsuperscript{2} & Dean M. Young\textsuperscript{2}

\textsuperscript{1} Department of Information Systems and Business Analytics, Baylor University, Waco, TX
\textsuperscript{2} Department of Statistical Science, Baylor University, Waco, TX

Correspondence: Phil D. Young, Department of Information Systems and Business Analytics, Baylor University, Waco, TX 76798, USA. Tel: 1-254-710-7394. E-mail: philip_young@baylor.edu

Received: June 30, 2020 Accepted: October 11, 2020 Online Published: October 21, 2020

\textbf{Abstract}

We provide a new, concise derivation of necessary and sufficient conditions for the explicit characterization of the general nonnegative-definite covariance structure $V$ of a general Gauss-Markov model with $E(y)$ and $Var(y)$ such that the best linear unbiased estimator, the weighted least squares estimator, and the least squares estimator of $X\beta$ are identical. In addition, we derive a representation of the general nonnegative-definite covariance structure $V$ defined above in terms of its Moore-Penrose pseudo-inverse.

\textbf{Keywords}: matrix equations, orthogonal-projection matrices, matrix column space, matrix rank, Moore-Penrose pseudo-inverse

\textbf{1. Introduction}

We consider the general Gauss-Markov model

$$y = X\beta + \epsilon,$$

(1)

where $y$ is an $n \times 1$ vector of observations, $X$ is an $n \times p$ known fixed, non-null model (design) matrix such that $\text{rank}(X) = p$, $\beta$ is a $p \times 1$ vector of unknown model parameters, and $\epsilon$ is an $n \times 1$ vector of random perturbations such that $E(\epsilon) = 0_{n \times 1}$ and $\text{Var}(\epsilon) = V$, where $V$ is a known $n \times n$ non-null, symmetric nonnegative-definite ($n.n.d.$) matrix. We denote the Gauss-Markov model defined above by $[y, X\beta, V]$, and we assume $y \in C(X : V)$, where $C(X : V)$ represents the column space of the partitioned matrix $(X : V)$.

Throughout the remainder of this paper, the notation $\mathbb{R}_{n \times m}$ represents the vector space of all $m \times n$ matrices over the real field $\mathbb{R}$. $\mathbb{R}_n$ denotes the set of all $n \times n$ real symmetric matrices, $\mathbb{R}_n^{+}$ represents the cone of all symmetric $n.n.d.$ matrices in $\mathbb{R}_{n \times n}$, and $\mathbb{R}_n^{*}$ denotes the interior of $\mathbb{R}_n^{+}$, which is the set of all symmetric positive-definite ($p.d.$) matrices. We use the notation $K'$ to denote the transpose of the real matrix $K \in \mathbb{R}_{m \times n}$. Furthermore, we let $K^+ \in \mathbb{R}_{m \times n}$ and $K^- \in \mathbb{R}_{m \times n}$ represent the Moore-Penrose pseudo-inverse and a generalized inverse of $K$, respectively. Also, for $K \in \mathbb{R}_{m \times n}$, we use the notation $P_K$ and $P_K^*$ to denote the orthogonal projection matrix onto $C(K)$ and $C(K)^{+}$, respectively.

Given $X$, we define the ordinary least squares (LS) estimator of $X\beta$ as

$$X\hat{\beta}_{LS} = X(X'X)^{-}X'y.$$

Puntanen, Styan, and Isotalo (2011) have defined the best linear unbiased (BLU) estimator of $X\beta$ as

$$X\hat{\beta}_{BLU} = X(X'T'X)^{-}X'T'y,$$

(2)

where $T = V + XU'X$ and $U \in \mathbb{R}_n$ is any $n \times n$ matrix such that $C(T) = C(X : V)$. Puntanen et al. (2011) have defined the weighted least squares (WLS) estimator as

$$X\hat{\beta}_{WLS} = X(X'V'X)^{-}X'V'y.$$

In this paper we give two characterizations of the general $n.n.d.$ error covariance structure $V$ in the Gauss-Markov model $[y, X\beta, V]$ for which $X\hat{\beta}_{BLU} = X\hat{\beta}_{WLS} = X\hat{\beta}_{LS}$ where $y \in C(X : V)$. We define these covariance matrices to be $BLU$-$WLS$-$LS$ estimator-equivalent (e.e.) covariance matrices. Specifically, in the first characterization we give a derivation of the
explicit general n.n.d. BLU-WLS-LS e.e. covariance structure that is considerably more concise and straightforward than the derivation given in Young, Odell and Hahn (2000). In the second characterization, we demonstrate that the Moore-Penrose pseudo-inverse of the covariance matrices contained in the set of n.n.d. BLU-WLS-LS e.e. covariance structures are themselves elements of the set.

A large majority of previous work has focused on implicitly and explicitly characterizing the general covariance matrix $V$ such that the BLU and LS estimators are equal. Puntanen and Styan (1989), Alalouf and Styan (1984), Tian and Wiens (2006), Proposition 10.1 in Puntanen et al. (2011), and numerous additional journal articles have presented many of these implicit characterizations.

However, we have found fewer results on explicit n.n.d. WLS-LS e.e. covariance structure characterizations. Plackett (1960), McElroy (1967), and Williams (1967) have derived sufficient (p.d.) WLS-LS e.e. covariance matrices. Additionally, for certain model matrices $X$, Herzberg and Aleong (1985) have presented a sufficient p.d. WLS-LS e.e. covariance matrix, and Zyskind and Martin (1969), Searle (1994), and Tian and Wiens (2006) have presented several implicit WLS-LS e.e. covariance-structure characterizations. Results on both implicit and explicit characterizations of the general n.n.d. BLU-WLS-LS e.e. covariance structure for the Gauss-Markov model $\{y, X\beta, V\}$ appear to be more sparse. Herzberg and Aleong (1985) have presented two sufficient WLS-BLU-LS e.e. covariance matrices. Moreover, Baksalary and Kala (1983) have given an implicit characterization of the general n.n.d. e.e. covariance structure for $V$, and Young et al. (2000) have explicitly characterized the general n.n.d. BLU-WLS-LS e.e. covariance structure.

We have organized the remainder of the paper as follows. In Section 2 we state two lemmas that we use to derive the first of our two theorems. In Section 3 we present a new concise derivation our general n.n.d. BLU-WLS-LS e.e. dependency structure characterization for $V$. We also demonstrate that the Moore-Penrose pseudo-inverse of elements contained in the set of n.n.d. BLU-WLS-LS e.e. covariance structures are themselves elements of this set. Last, in Section 4 we briefly summarize the two characterization results proven here.

2. Preliminary Lemmas

We next present two lemmas that we use in the proof of our first e.e.-covariance-structure characterization. The first lemma gives conditions for $V$ such that $X\hat{\beta}_{BLU} = X\hat{\beta}_{WLS} = X\hat{\beta}_{LS}$. A proof of part a) is in the lemma in Zyskind (1967), a proof of b) is in Zyskind and Martin (1969), and a proof of part c) is in Theorem 2.2 of Baksalary and Kala (1983).

**Lemma 1.** For the Gauss-Markov model $\{y, X\beta, V\}$, we have

- a) $X\hat{\beta}_{BLU} = X\hat{\beta}_{WLS}$ if and only if $C(X) \subset C(V)$,
- b) $X\hat{\beta}_{BLU} = X\hat{\beta}_{LS}$ if and only if $C(VX) \subset C(X)$, and
- c) $X\hat{\beta}_{WLS} = X\hat{\beta}_{LS}$ if and only if $X\hat{\beta}_{BLU} = X\hat{\beta}_{WLS} = X\hat{\beta}_{LS}$.

In the second lemma, we state the general symmetric n.n.d. solution matrix to a particular homogeneous matrix that contains the column space of a specified matrix.

**Lemma 2.** Let $A \in \mathbb{R}_{n \times q}$ such that $\text{rank}(A) = k$, where $k \leq q < n$, and let $\mathcal{W} := \{U \in \mathbb{R}_{n}^{q} : C(A) \subset C(U)\}$. Then, a representation of the general n.n.d. solution to $P_{A}ZP_{A}^{\perp} = 0$ such that $C(A) \subset C(Z)$ is

$$Z = P_{A}U_{1}P_{A} + P_{A}^{\perp}U_{2}P_{A}^{\perp},$$

where $U_{1} \in \mathcal{W}$ and $U_{2} \in \mathbb{R}_{n}^{q}$ is arbitrary.

**Proof.** The proof is similar to the proof of Lemma 6 in Young et al. (2000).

3. Main Results

We now present a concise proof of the explicit characterization of the general n.n.d. BLU-WLS-LS e.e. covariance structure. The proof immediately below is considerably shorter and more direct than a previous proof given in Young et al. (2000).

**Theorem 1.** For the general Gauss-Markov model $\{y, X\beta, V\}$, we have $X\hat{\beta}_{BLU} = X\hat{\beta}_{WLS} = X\hat{\beta}_{LS}$ if and only if $V \in \mathcal{V}$, where

$$\mathcal{V} := \{V \in \mathbb{R}_{n}^{q} : V = P_{X}W_{1}P_{X} + P_{X}^{\perp}W_{2}P_{X}^{\perp}\}$$

with

$$W_{1} \in \{W \in \mathbb{R}_{n}^{q} : C(X) \subset C(W)\}$$

and $W_{2} \in \mathbb{R}_{n}^{q}$ is arbitrary.
Last, again using the fact that
\[ X\hat{\beta}_{BLU} = X\hat{\beta}_{WLS} = X\hat{\beta}_{LS} \]
the proof proceeds as follows:

- **Proof.** From Lemmas 1 and 2, we have that
  \[ X\hat{\beta}_{BLU} = X\hat{\beta}_{WLS} = X\hat{\beta}_{LS} \]

Next, for the general Gauss-Markov model \( \{y, X\beta, V\} \), we characterize the n.n.d. e.e. covariance matrices \( V \in \mathcal{V} \), defined in (3), by showing that for \( V \in \mathcal{V} \), the Moore-Penrose inverse \( V^* \) has a particular form.

**Theorem 2.** For the general Gauss-Markov model \( \{y, X\beta, V\} \), consider the covariance matrices \( V \in \mathcal{V} \) defined in (3). Then, \( V \in \mathcal{V} \) if and only if \( V^* \in \mathcal{V} \).

**Proof.** We first prove the necessity portion of Theorem 2. Let \( V \in \mathcal{V} \) be defined as in (3). In addition, let

\[ V^* = P_XW_i^TP_X + P_XW_i^TP_X^T \]

Then, using the definition of a Moore-Penrose pseudo-inverse and the facts that for \( W_i \in \mathbb{R}^n \), \( P_XW_i^TP_X = 0 \), and \( P_XW_i = W_iP_X = W_i \), \( i = 1, 2 \), we have

\[ VV^*V = (P_XW_i^TP_X + P_XW_i^TP_X^T)(P_XW_i^TP_X + P_XW_i^TP_X^T)(P_XW_i^TP_X + P_XW_i^TP_X^T) \]

\[ = (P_XW_i^TP_X)(P_XW_i^TP_X) + (P_XW_i^TP_X)(P_XW_i^TP_X^T) \]

\[ = (P_XW_i^TP_X) + (P_XW_i^TP_X^T) \]

\[ = V. \]

Next, we have

\[ V^*VV^* = (P_XW_i^TP_X + P_XW_i^TP_X^T)(P_XW_i^TP_X + P_XW_i^TP_X^T)(P_XW_i^TP_X + P_XW_i^TP_X^T) \]

\[ = (P_XW_i^TP_X)(P_XW_i^TP_X) + (P_XW_i^TP_X)(P_XW_i^TP_X^T) \]

\[ = (P_XW_i^TP_X) + (P_XW_i^TP_X^T) \]

\[ = V^*. \]

Third, let \( W_i \) be defined as in (4). Then, using the fact that \( W_iW_i^T = \text{big}(W_iW_i^T) = W_i^TW_i = W_i^T W_i \), \( i = 1, 2 \), we have

\[ [VV^*] = \left[(P_XW_i^TP_X + P_XW_i^TP_X^T)(P_XW_i^TP_X + P_XW_i^TP_X^T)\right] \]

\[ = (P_XW_i^TP_X)(P_XW_i^TP_X) + (P_XW_i^TP_X)(P_XW_i^TP_X^T) \]

\[ = (P_XW_i^TP_X) + (P_XW_i^TP_X^T) \]

\[ = VV^*. \]

Last, again using the fact that \( W_iW_i^T = W_i^T W_i \), \( i = 1, 2 \), we have that

\[ [VV^*] = \left[(P_XW_i^TP_X + P_XW_i^TP_X^T)(P_XW_i^TP_X + P_XW_i^TP_X^T)\right] \]

\[ = (P_XW_i^TP_X)(P_XW_i^TP_X) + (P_XW_i^TP_X)(P_XW_i^TP_X^T) \]

\[ = (P_XW_i^TP_X) + (P_XW_i^TP_X^T) \]

\[ = VV^*. \]
Hence, $V^* = V^+$. The sufficiency portion of the proof is similar to the necessity portion because of the facts that $[V^+]^T = V$ and $[W_i^+]^T = W_i, i = 1, 2$.

The following corollary, which follows directly from Theorems 1 and 2, gives several implicit characterizations of the general n.n.d. BLU-WLS-LS e.e. dependency matrix.

**Corollary.** Let $V$ be defined as in (3). Then, $V \in \mathcal{V}$ if and only if $C(X) \subset C(V)$, and

\begin{align*}
a) \quad & P_X^\perp V P_X^\perp = P_X^\perp V \\
b) \quad & P_X^\perp V^+ P_X^\perp = P_X^\perp V^+ \\
c) \quad & P_X V P_X = P_X V \\
d) \quad & P_X V^+ P_X = P_X V^+ \\
e) \quad & P_X V = V P_X \\
f) \quad & P_X V^+ = V^+ P_X \\
g) \quad & P_X^\perp V = V P_X^\perp \\
h) \quad & P_X^\perp V^+ = V^+ P_X^\perp.
\end{align*}

4. **Summary**

We have derived two explicit characterizations of the general n.n.d. e.e. covariance structure such that $X\hat{\beta}_{BLU} = X\hat{\beta}_{WLS} = X\hat{\beta}_{LS}$. Theorem 1 provides a brief derivation of the explicit general n.n.d. BLU-WLS-LS e.e. dependency structure that considerably shortens a proof given in Young et al. (2000). Theorem 2 presents a second characterization of the general n.n.d. BLU-WLS-LS e.e. covariance matrix $V$ in which we prove that $V$ and $V^+$ have the same general structure. Last, we give some implicit characterizations of the general n.n.d. e.e. covariance matrices such that $X\hat{\beta}_{BLU} = X\hat{\beta}_{WLS} = X\hat{\beta}_{LS}$.

**Acknowledgements**

We wish to thank Joy L. Young for her help in the writing of this paper.

**References**


**Copyrights**

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/4.0/).