Extremal quadratic properties of least-squares solutions of a linear matrix equation with statistical applications

Bo Jiang\textsuperscript{a}, Yongge Tian\textsuperscript{b,∗}

\textsuperscript{a}College of Mathematics and Information Science, Shandong Institute of Business and Technology, Yantai, China
\textsuperscript{b}China Economics and Management Academy, Central University of Finance and Economics, Beijing, China

Abstract. Assume that a quadratic matrix-valued function \( \psi(X) = Q - X'PX \) is given and let \( S = \{ X \in \mathbb{R}^{m \times n} | \text{trace}(AX - B)'(AX - B) \} \) be the set of all least-squares solutions of the linear matrix equation \( AX = B \). In this paper, we first establish explicit formulas for calculating the maximum and minimum ranks and inertias of \( \psi(X) \) subject to \( X \in S \), and then derive from the formulas the analytic solutions of the two optimization problems

\[ \psi(X) = \max \quad \text{and} \quad \psi(X) = \min \quad \text{subject to} \quad X \in S \]


As a statistical application, we present some results on equalities and inequalities of the ordinary least squares estimator (OLSE) of the unknown parameter vector in a general linear model.

\begin{align*}
\text{Mathematics Subject Classifications:} & \quad 15A03; 15A09; 62H12; 62J05 \\
\text{Key Words:} & \quad \text{Matrix equation; least-squares solution; quadratic matrix-valued function; rank; inertia; Löwner partial ordering; linear model}
\end{align*}

1 Introduction

Throughout this paper, \( \mathbb{R}^{m \times n} \) stands for the set of all \( m \times n \) real matrices. \( A' \), \( r(A) \) and \( \mathcal{R}(A) \) stand for the transpose, rank, and range (column space) of a matrix \( A \in \mathbb{R}^{m \times n} \), respectively. \( I_{m} \) denotes the identity matrix of order \( m \). \([A, B]\) denotes a row block matrix consisting of \( A \) and \( B \). The Moore–Penrose inverse of \( A \in \mathbb{R}^{m \times n} \), denoted by \( A^+ \), is defined to be the unique solution \( X \) satisfying the four matrix equations \( AXA = A \), \( XAX = X \), \( (AX)' = AX \), and \( (XA)' = XA \). \( E_{m}^{A} \) and \( F_{A} \) stand for \( E_{m}^{A} = \begin{bmatrix} I_{m} & -A^{+} \end{bmatrix} \) and \( F_{A} = I_{n} - A'A \) with \( r(E_{A}) = m - r(A) \) and \( r(F_{A}) = n - r(A) \). The Frobenius norm of a matrix \( A \in \mathbb{R}^{m \times n} \) is defined to be \( ||A||_{F} = \sqrt{\text{trace}(AA')} \). The symbols \( i_{+}(A) \) and \( i_{-}(A) \) for \( A = A' \in \mathbb{R}^{m \times n} \), called the positive and negative inertias of \( A \), respectively, denote the number of the positive and negative eigenvalues of \( A \) counted with multiplicities, respectively, both of which satisfy \( r(A) = i_{+}(A) + i_{-}(A) \). For brief, we use \( i_{+}(A) \) to denote the both numbers.

For a symmetric matrix \( A = A' \in \mathbb{R}^{m \times m} \), the notations \( A \succ 0 \), \( A \succcurlyeq 0 \), \( A \prec 0 \), and \( A \preccurlyeq 0 \) mean that \( A \) is positive definite, positive semi-definite, negative definite, and negative semi-definite, respectively. Two symmetric matrices \( A \) and \( B \) of the same size are said to satisfy the inequalities \( A \succcurlyeq B \), \( A \succ B \), \( A \prec B \), and \( A \preccurlyeq B \) in the Löwner partial ordering if \( A - B \) is positive definite, positive semi-definite, negative definite, and negative semi-definite respectively. It is well known that the Löwner partial ordering is a surprisingly strong and useful relation between two complex Hermitian (real symmetric) matrices. For more issues about connections between the inertias and the Löwner partial ordering of complex Hermitian (real symmetric) matrices, as well as applications of the matrix inertia and the Löwner partial ordering in statistical analysis, see, e.g., [6].

Matrix equations are one of the prominent topics in matrix theory and applications, and the study on

\begin{equation}
AX = B \tag{1.1}
\end{equation}

is the very beginning of the theory on matrix equations, where \( A \in \mathbb{R}^{n \times n} \) and \( B \in \mathbb{R}^{p \times m} \) are two given matrices. As is well known, the least-squares solution (LSS for short) of (1.1) is defined to be a matrix \( X \in \mathbb{R}^{n \times m} \) that satisfies the following minimization problem

\begin{equation}
\| AX - B \|_{F}^{2} = \min \tag{1.2}
\end{equation}

In this paper, we consider some extremal quadratic properties of LSSs of (1.1) and present some applications of the quadratic properties in statistical analysis of linear regression models.

Eq. (1.1) is one of the most important objects of study in matrix analysis due to its fundamental form. A seminal work on this equation was given in [5], while many results on the equation and its applications occur everywhere in current theory of mathematics and other disciplines. Under the assumptions that \( A \) is singular, the LSS of \( A \) is not unique, and general LSS of (1.2) can be written as certain linear matrix expression that involve one arbitrary matrix. In such a case, the performances of the LSS depend on the choices of the arbitrary matrix. Recall that ranks and inertias of real symmetric (complex Hermitian) matrices are conceptual foundation in elementary linear algebra, which are the most significant finite integers in reflecting

\begin{flushright}
∗Corresponding author. E-mail Addresses: yongge.tian@gmail.com, jiangbo@email.cufe.edu.cn
\end{flushright}
intrinsic properties of matrices. In a recent paper [12], we established a group of formulas for calculating maximum and minimum ranks and inertias of the quadratic matrix-valued function

$$\phi(X) = Q - XPX'$$

subject to the LSSs of (1.1), and derive many quadratic properties of the LSSs of (1.1) from the rank and inertia formulas. In order to find more quadratic properties of the LSSs of (1.1), we construct a new quadratic matrix-valued function as follows

$$\psi(X) = Q - X'PX,$$

where $$P = P' \in \mathbb{R}^{n \times n}$$ and $$Q = Q' \in \mathbb{R}^{m \times m}$$ are given matrices, and give analytical solutions of the following matrix rank and inertia optimization problems and their applications.

**Problem 1.1.** Let

$$S = \{ X \in \mathbb{R}^{n \times m} \mid \| AX - B \|_F^2 = \min \}.$$

Then establish exact algebraic formulas for calculating the following six maximum and minimum ranks and inertias

$$\max_{X \in S} r(Q - XPX'), \quad \min_{X \in S} r(Q - XPX'),$$

$$\max_{X \in S} i_+(Q - XPX'), \quad \min_{X \in S} i_+(Q - XPX').$$

**Problem 1.2.** Establish necessary and sufficient conditions for the following two partial ordering optimization problems

$$\max \{ Q - X'PX \mid X \in S \}, \quad \min \{ Q - X'PX \mid X \in S \}$$

to have solutions, respectively, and give exact algebraic expressions of the solutions.

**Problem 1.3.** Establish necessary and sufficient conditions for the following constrained quadratic matrix equation

$$X'PX = Q \quad \text{s.t.} \quad X \in S$$

(1.9)

to have a solution, as well as necessary and sufficient conditions for the following four constrained quadratic matrix inequalities

$$X'PX \succ Q \ (\succeq Q, \prec Q, \preceq Q)$$

(1.10)

to hold for a $$X \in S$$ (all matrices $$X \in S$$), respectively.

These problems may occur in the analysis of constrained quadratic matrix-valued functions. In particular, many problems in the covariance matrix analysis of estimators of unknown parameters under linear regression models can be summarized as the special cases of Problems 1.1–1.3, while the solutions to these problems, as algebraic tools, can directly be applied to solve many optimization problems occurred in statistical analysis.

Concerning the general solution of (1.1), the following result is well known.

**Lemma 1.4** ([5]). The matrix equation in (1.1) has a solution if and only if $$AA' = B$$. In this case, the general solution can be written in the following parametric form

$$X = A' B + F_A V,$$

(1.11)

where $$V \in \mathbb{R}^{n \times m}$$ is arbitrary. The solution of (1.1) is unique if and only if $$r(A) = n$$. If (1.1) is inconsistent, then the normal equation of (1.1) is

$$A'AX = A'B,$$

(1.12)

and the general expression of the LSSs of (1.1) can be written as

$$X = A' B + F_A V,$$

(1.13)

where $$V \in \mathbb{R}^{n \times m}$$ is arbitrary. The LSSs of (1.1) is unique if and only if $$r(A) = n$$.

In order to solve the previous problems, we need to use the following lemma on the connections of matrix equalities/inequalities and matrix ranks/inertias.
Lemma 1.5 ([9]). Let $S$ be a set consisting of matrices over $\mathbb{R}^{m \times n}$ and let $\mathcal{T}$ be a set consisting of symmetric matrices of order $m$. Then, the following results hold.

(a) Under $m = n$, there exists a nonsingular matrix $X \in S$ if and only if $\max_{X \in \mathcal{S}} r(X) = m$.
(b) Under $m = n$, all $X \in S$ are nonsingular if and only if $\min_{X \in \mathcal{S}} r(X) = m$.
(c) $0 \in S$ if and only if $\min_{X \in \mathcal{S}} r(X) = 0$.
(d) $S = \{0\}$ if and only if $\max_{X \in \mathcal{S}} r(X) = 0$.
(e) All $X \in S$ have the same rank if and only if $\max_{X \in \mathcal{S}} r(X) = \min_{X \in \mathcal{S}} r(X)$.
(f) $\mathcal{T}$ has a matrix $X > 0 \ (X > 0)$ if and only if $\max_{X \in \mathcal{T}} i_+(X) = m \ (\max_{X \in \mathcal{T}} i_-(X) = m)$.
(g) All $X \in \mathcal{T}$ satisfy $X > 0 \ (X < 0)$ if and only if $\min_{X \in \mathcal{T}} i_+(X) = m \ (\min_{X \in \mathcal{T}} i_-(X) = m)$.
(h) $\mathcal{T}$ has a matrix $X \succ 0 \ (X \prec 0)$ if and only if $\max_{X \in \mathcal{T}} i_+(X) = 0 \ (\max_{X \in \mathcal{T}} i_-(X) = 0)$.
(i) All $X \in \mathcal{T}$ satisfy $X \succ 0 \ (X \prec 0)$ if and only if $\max_{X \in \mathcal{T}} i_+(X) = 0 \ (\max_{X \in \mathcal{T}} i_-(X) = 0)$.
(j) All $X \in \mathcal{T}$ have the same positive inertia if and only if $\max_{X \in \mathcal{T}} i_+(X) = \min_{X \in \mathcal{T}} i_+(X)$.
(k) All $X \in \mathcal{T}$ have the same negative inertia if and only if $\max_{X \in \mathcal{T}} i_-(X) = \min_{X \in \mathcal{T}} i_-(X)$.

The assertions in Lemma 1.5 directly follow from the definitions of rank/inertia, definiteness, and semi-definiteness of (symmetric) matrices, which provide a highly flexible framework for characterizing equalities and inequalities of matrices via matrix ranks and inertias. Under the guidance of the principles in Lemma 1.5, a huge amount of exact formulas for calculating the ranks/inertias of matrices are established, and many matrix equalities/inequalities are proved via the rank/inertia formulas. We next present some basic matrix rank/inertia formulas used in this paper.

Lemma 1.6 ([7]). Let $A \in \mathbb{C}^{m \times n}$, $B \in \mathbb{C}^{m \times k}$, and $C \in \mathbb{C}^{l \times n}$. Then,

\[ r[ A, B ] = r(A) + r(EBA) = r(B) + r(AB), \]
\[ r \begin{bmatrix} A & C \\ B & 0 \end{bmatrix} = r(C) + r(AF) \]
\[ r \begin{bmatrix} A & B \\ C & 0 \end{bmatrix} = r(C) + r(AB). \]

Lemma 1.7 ([9]). Let $A = A' \in \mathbb{R}^{m \times m}$, $B = B' \in \mathbb{R}^{n \times n}$ and $Q \in \mathbb{R}^{n \times n}$, and assume that $P \in \mathbb{R}^{m \times m}$ is nonsingular. Then,

\[ i_\pm(PAP^t) = i_\pm(A), \quad (\text{Sylvester's law of inertia}), \]
\[ di_\pm(\lambda A) = \begin{cases} i_\pm(A) & \text{if } \lambda > 0 \\ i_\mp(A) & \text{if } \lambda < 0 \end{cases}, \]
\[ i_\pm \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} = i_\pm(A) + i_\pm(B), \]
\[ i_\pm \begin{bmatrix} 0 & Q \\ Q' & 0 \end{bmatrix} = i_\mp \begin{bmatrix} 0 & Q \\ Q' & 0 \end{bmatrix} = r(Q). \]

Lemma 1.8 ([9]). Let $A = A' \in \mathbb{R}^{m \times m}$ and $B \in \mathbb{R}^{m \times n}$. Then, the following expansion formulas hold

\[ i_\pm \begin{bmatrix} A & B \\ 0 & B' \end{bmatrix} = r(B) + i_\pm(E_BAE_B), \quad r \begin{bmatrix} A & B \\ 0 & B' \end{bmatrix} = 2r(B) + r(E_BAE_B). \]

Lemma 1.9 ([11]). Let $\psi(X)$ be as given in (1.4) and assume that (1.1) is consistent. Also, let

\[ T_1 = \begin{bmatrix} Q & B' \\ B & 0 \end{bmatrix}, \quad T_2 = \begin{bmatrix} B & A' \\ 0 & A \end{bmatrix}. \]

Then,

\[ \max_{AX=B} r(Q - X'PX) = \min\{ m, r(T_1) - 2r(A) \}, \]
\[ \min_{AX=B} r(Q - X'PX) = \max\{ 0, r(T_1) - 2r(T_2) + 2r(A), i_+(T_1) - r(T_2) + r(A), i_-(T_1) - r(T_2) + r(A) \}, \]
\[ \max_{AX=B} i_+(Q - X'PX) = \min\{ m, i_+(T_1) - r(A) \}, \]
\[ \min_{AX=B} i_+(Q - X'PX) = \max\{ 0, i_+(T_1) - r(T_2) + r(A) \}. \]
It is a little bit surprised in mathematics that people now can establish a huge amount of algebraic
formulas like those in Lemmas 1.6–1.8 and 1.9 for calculating the (maximum and minimum possible) ranks
and inertias of matrices and their operations. But it takes really a long march for people to realize this fact,
make enough preparations and use them in mathematics and other fields in the past several decades.

2 Rank and inertia formulas of \( Q - X'PX \) for the LSSs of \( AX = B \)

Based on the preparations in Section 1, we are ready to solve Problems 1.1–1.3 in Section 1. We first give a
group of formulas for calculating the maximum and minimum ranks and inertias in (1.6) and (1.7).

**Theorem 2.1.** Let \( \psi(X) \) and \( S \) be as given in (1.4) and (1.5), respectively. Also, denote

\[
T_1 = \begin{bmatrix} Q & B'A & 0 \\ A'B & 0 & A'A \\ 0 & A'A & -P \end{bmatrix}, \quad T_2 = \begin{bmatrix} A'B & 0 & A'A \\ 0 & A'A & P \end{bmatrix}.
\]  

Then, the global maximum and minimum ranks and inertias of \( Q - X'PX \) subject to \( X \in S \) are given by

\[
\begin{align*}
\max_{X \in S} r(Q - X'PX) &= \max_{A'AX = A'B} r(Q - X'PX), \\
\min_{X \in S} r(Q - X'PX) &= \min_{A'AX = A'B} r(Q - X'PX), \\
\max_{X \in S} i_{\pm}(Q - X'PX) &= \max_{A'AX = A'B} i_{\pm}(Q - X'PX), \\
\min_{X \in S} i_{\pm}(Q - X'PX) &= \min_{A'AX = A'B} i_{\pm}(Q - X'PX).
\end{align*}
\]

Proof. It can be seen from (1.5) and (1.12) that

In these cases, replacing \( A \) with \( A' \) and \( B \) with \( B' \) in (1.23)–(1.26) and simplifying, we obtain (2.2)–

(2.5).

In the light of the exact formulas (2.2)–(2.5) and Lemma 1.5, we obtain the following consequences. Their
derivations are simple and direct, and therefore are omitted.

**Corollary 2.2.** Let \( AX = B \) be given in (1.1), and let \( \psi(X) \), \( T_1 \), and \( T_2 \) be of the forms in (1.4) and (2.1). Then, the following results hold.

(a) \( AX = B \) has an LSS such that \( Q - X'PX \) is nonsingular if and only if \( r(T_1) \geq 2r(A) + m \).

(b) \( Q - X'PX \) is nonsingular for all LSSs of \( AX = B \) if and only if \( r(T_1) - 2r(T_2) + 2r(A) = m \), or

\( i_+(T_1) - r(T_2) + r(A) = m \), or \( i_-(T_1) - r(T_2) + r(A) = m \).

(c) \( AX = B \) has an LSS satisfying \( X'PX = Q \) if and only if \( i_+(T_1) \leq r(T_2) - r(A) \) and \( i_-(T_1) \leq r(T_2) - r(A) \).

(d) All LSSs of \( AX = B \) satisfy \( X'PX = Q \) if and only if \( r(T_1) = 2r(A) \).

(e) \( AX = B \) has an LSS such that \( X'PX < Q \) if and only if \( i_+(T_1) \geq m + r(A) \).

(f) All LSSs of \( AX = B \) satisfy \( X'PX < Q \) if and only if \( i_+(T_1) - r(T_2) + r(A) = m \).

(g) \( AX = B \) has an LSS such that \( X'PX > Q \) if and only if \( i_-(T_1) \geq m + r(A) \).

(h) All LSSs of \( AX = B \) satisfy \( X'PX > Q \) if and only if \( i_-(T_1) - r(T_2) + r(A) = m \).

(i) \( AX = B \) has an LSS such that \( X'PX \leq Q \) if and only if \( i_-(T_1) \leq r(T_2) - r(A) \).

(j) All LSSs of \( AX = B \) satisfy \( X'PX \leq Q \) if and only if \( i_-(T_1) = r(A) \).
(k) $AX = B$ has an LSS such that $X'PX \succ Q$ if and only if $i_+(T_1) \leq r(T_2) - r(A)$.

(l) All LSSs of $AX = B$ satisfy $X'PX \succ Q$ if and only if $i_+(T_1) = r(A)$.

Under the conditions that $P$ and $Q$ are both positive definite, or especially, $P = I_n$ and/or $Q = I_m$, simplifying the results in Theorem 2.1 and Corollary 2.2 via Lemmas 1.6–1.8, we obtain the following three corollaries, and their proofs are also omitted.

**Corollary 2.3.** Let $\psi(X)$ be as given in (1.4) with $P \succ 0$ and $Q \succ 0$, and let $S$ be as given in (1.5). Then,

$$\max_{X \in S} r(Q - X'PX) = \min \{ m, \ m + n + r(A'AP^{-1}A'A - A'BP^{-1}B'A) - 2r(A) \},$$
\[ (2.6) \]

$$\min_{X \in S} r(Q - X'PX) = \max \{ m - n + r(A'AP^{-1}A'A - A'BP^{-1}B'A), \ i_+(A'AP^{-1}A'A - A'BP^{-1}B'A) \},$$
\[ (2.7) \]

$$\max_{X \in S} i_+(Q - X'PX) = m + i_+(A'AP^{-1}A'A - A'BP^{-1}B'A) - r(A),$$
\[ (2.8) \]

$$\max_{X \in S} i_+(Q - X'PX) = \min \{ m, \ m + i_+(A'AP^{-1}A'A - A'BP^{-1}B'A) - r(A) \},$$
\[ (2.9) \]

$$\min_{X \in S} i_+(Q - X'PX) = \max \{ 0, \ m - n + i_+(A'AP^{-1}A'A - A'BP^{-1}B'A) \},$$
\[ (2.10) \]

$$\min_{X \in S} i_+(Q - X'PX) = i_+(A'AP^{-1}A'A - A'BP^{-1}B'A).$$
\[ (2.11) \]

In consequence, the following results hold.

(a) $AX = B$ has an LSS such that $Q - X'PX$ is nonsingular if and only if $r(A'AP^{-1}A'A - A'BP^{-1}B'A) \geq 2r(A) - n$.

(b) $Q - X'PX$ is nonsingular for all LSSs of $AX = B$ if and only if $r(A'AP^{-1}A'A - A'BP^{-1}B'A) = n$ or $i_+(A'AP^{-1}A'A - A'BP^{-1}B'A) = m$.

(c) $AX = B$ has an LSS such that $X'PX = Q$ if and only if $r(A'AP^{-1}A'A - A'BP^{-1}B'A) \leq n - m$ and $A'AP^{-1}A'A \succ A'BP^{-1}B'A$.

(d) All LSSs of $AX = B$ satisfy $X'PX = Q$ if and only if $r(A'AP^{-1}A'A - A'BP^{-1}B'A) = 2r(A) - m - n$.

(e) $AX = B$ has an LSS such that $X'PX \prec Q$ if and only if $i_+(A'AP^{-1}A'A - A'BP^{-1}B'A) = r(A)$.

(f) All LSSs of $AX = B$ satisfy $X'PX \prec Q$ if and only if $A'AP^{-1}A'A \succ A'BP^{-1}B'A$.

(g) $AX = B$ has an LSS such that $X'PX \succ Q$ if and only if $i_+(A'AP^{-1}A'A - A'BP^{-1}B'A) \geq m - n + r(A)$.

(h) All LSSs of $AX = B$ satisfy $X'PX \succ Q$ if and only if $i_-(A'AP^{-1}A'A - A'BP^{-1}B'A) = m$.

(i) $AX = B$ has an LSS such that $X'PX \preceq Q$ if and only if $A'AP^{-1}A'A \succeq A'BP^{-1}B'A$.

(j) All LSSs of $AX = B$ satisfy $X'PX \preceq Q$ if and only if $A'AP^{-1}A'A \succeq A'BP^{-1}B'A$ and $r(A) = n$.

(k) $AX = B$ has an LSS such that $X'PX \succ Q$ if and only if $i_+(A'AP^{-1}A'A - A'BP^{-1}B'A) \leq n - m$.

(l) All LSSs of $AX = B$ satisfy $X'PX \succ Q$ if and only if $i_+(A'AP^{-1}A'A - A'BP^{-1}B'A) = r(A) - m$.

**Corollary 2.4.** Let $\psi(X)$ and $S$ be as given in (1.4) and (1.5), respectively, and let $T = \begin{bmatrix} Q & B' \ A \\ A' B & (A'A)^2 \end{bmatrix}$. Then,

$$\max_{X \in S} r(Q - X'X) = \min \{ m, \ n + r(T) - 2r(A) \},$$
\[ (2.12) \]

$$\min_{X \in S} r(Q - X'X) = \max \{ r(T) - n, \ i_+(T) \},$$
\[ (2.13) \]

$$\max_{X \in S} i_+(Q - X'X) = \min \{ m, \ i_+(T) - r(A) + i_+(I_n) \},$$
\[ (2.14) \]

$$\min_{X \in S} i_+(Q - X'X) = \max \{ 0, \ i_+(T) - i_+(I_n) \}.$$
\[ (2.15) \]

In consequence, the following results hold.

(a) $AX = B$ has an LSS such that $Q - X'X$ is nonsingular if and only if $r(T) \geq 2r(A) + m - n$.

(b) $Q - X'X$ is nonsingular for all LSSs of $AX = B$ if and only if $r(T) = m + n$ or $i_+(T) = m$.  

5
Corollary 2.5. Let $AX = B$ has an LSS such that $X'X = Q$ if and only if $T \succ 0$ and $r(T) \leq n$.

(d) All LSSs of $AX = B$ satisfy $X'X = Q$ if and only if $r(T) = 2r(A) - n$.

(e) $AX = B$ has an LSS such that $X'X \prec Q$ if and only if $i_+(T) \geq r(A) + m$.

(f) All LSSs of $AX = B$ satisfy $X'X \prec Q$ if and only if $i_+(T) = m + n$.

(g) $AX = B$ has an LSS such that $X'X \succ Q$ if and only if $i_-(T) \geq r(A) + m - n$.

(h) All LSSs of $AX = B$ satisfy $X'X \succ Q$ if and only if $i_-(T) = m$.

(i) $AX = B$ has an LSS such that $X'X \preceq Q$ if and only if $T \succ 0$.

(j) All LSSs of $AX = B$ satisfy $X'X \preceq Q$ if and only if $r(A) = n$ and $T \preceq 0$.

(k) $AX = B$ has an LSS such that $X'X \succ Q$ if and only if $i_+(T) \leq n$.

(l) All LSSs of $AX = B$ satisfy $X'X \succ Q$ if and only if $i_+(T) = r(A)$.

Corollary 2.5. Let $S$ be as given in (1.5). Then,

$$\max_{X \in S} r(A) = \min \{ m, \ m + n + r[(A'A)^2 - A'BB'A] - 2r(A) \}$$

$$\min_{X \in S} r(A) = \max \{ m - n + [(A'A)^2 - A'BB'A], \ i_-(A'A)^2 - A'BB'A] \}$$

$$\max_{X \in S} i_+(A'A)^2 - A'BB'A] = m + i_+[(A'A)^2 - A'BB'A] - r(A)$$

$$\max_{X \in S} i_-[(A'A)^2 - A'BB'A] = \min \{ m, \ n + i_-[(A'A)^2 - A'BB'A] - r(A) \}$$

$$\min_{X \in S} i_+(A'A)^2 - A'BB'A] = \max \{ 0, \ m - n + i_+[(A'A)^2 - A'BB'A] \}$$

$$\min_{X \in S} i_-[(A'A)^2 - A'BB'A] = \max \{ 0, \ m - n + i_-[(A'A)^2 - A'BB'A] \}$$

In consequence, the following results hold.

(a) $AX = B$ has an LSS such that $I_m - X'X$ is nonsingular if and only if $r[(A'A)^2 - A'BB'A] \geq 2r(A) - n$.

(b) $I_m - X'X$ is nonsingular for all LSSs of $AX = B$ if and only if $r[(A'A)^2 - A'BB'A] \geq 0$ or $i_-[(A'A)^2 - A'BB'A] = 0$.

(c) $AX = B$ has an LSS such that $X'X = I_m$ if and only if $r[(A'A)^2 - A'BB'A] \leq n - m$ and $(A'A)^2 \succ A'BB'A$.

(d) All LSSs of $AX = B$ satisfy $X'X = I_m$ if and only if $r[(A'A)^2 - A'BB'A] = 2r(A) - m - n$.

(e) $AX = B$ has an LSS such that $X'X \preceq I_m$ if and only if $i_+[(A'A)^2 - A'BB'A] = r(A)$.

(f) All LSSs of $AX = B$ satisfy $X'X \prec I_m$ if and only if $(A'A)^2 \succ A'BB'A$.

(g) $AX = B$ has an LSS such that $X'X \succ I_m$ if and only if $i_-[(A'A)^2 - A'BB'A] \geq m - n + r(A)$.

(h) All LSSs of $AX = B$ satisfy $X'X \succ I_m$ if and only if $i_-[(A'A)^2 - A'BB'A] = m$.

(i) $AX = B$ has an LSS such that $X'X \preceq I_m$ if and only if $(A'A)^2 \succ A'BB'A$.

(j) All LSSs of $AX = B$ satisfy $X'X \preceq I_m$ if and only if $(A'A)^2 \succ A'BB'A$ and $r(A) = n$.

(k) $AX = B$ has an LSS such that $X'X \succ I_m$ if and only if $i_+[(A'A)^2 - A'BB'A] \leq n - m$.

(l) All LSSs of $AX = B$ satisfy $X'X \succ I_m$ if and only if $i_+[(A'A)^2 - A'BB'A] = r(A) - m$.

In mathematics, the collection of all matrices $X$ that satisfy $X'X = I_m$ is called a complex Stiefel manifold; see, e.g., [2, 3], while the collections of all matrices $X$ that satisfy $X'X \succ I_m$ ($\succ I_m, \prec I_m, \prec I_m$) are called generalized complex Stiefel manifolds. The results in Corollary 2.5 characterize some basic relations between the manifold $S$ and these Stiefel manifolds, and can be applied to reveal many deep and profound properties of LSSs of linear matrix equations.

In the remaining part of this section, we derive analytical solutions to the two optimization problems in (1.8), i.e., to find $X, \hat{X}$ such that

$$\psi(X) \leq \psi(\hat{X}) \text{ for all } X \in S,$$  \hspace{1cm} (2.22)

$$\psi(X) \geq \psi(\hat{X}) \text{ for all } X \in S$$  \hspace{1cm} (2.23)

hold, respectively.
Theorem 2.6. Let $S$ be as given in (1.5). Then, the following results hold.

(a) There exists an $\tilde{X} \in S$ such that (2.22) holds if and only if

$$F_A P F_A \succeq 0 \quad \text{and} \quad \mathcal{R} \left[ \begin{array}{cc} 0 & A' A \\ A' B & 0 \end{array} \right] \subseteq \mathcal{R} \left[ \begin{array}{cc} P & A' A \\ A' A & 0 \end{array} \right].$$

In this case, the global maximizer $\tilde{X}$ satisfying (2.22) is determined by the consistent matrix equation

$$\begin{bmatrix} A' A \\ F_A P \end{bmatrix} \tilde{X} = \begin{bmatrix} A' B \\ 0 \end{bmatrix},$$

and the general expression of the global maximizer is given by

$$\argmax \{ Q - X'PX | X \in S \} = \begin{bmatrix} A' A \\ F_A P \end{bmatrix} \begin{bmatrix} A' B \\ 0 \end{bmatrix} + \left( I_n - \begin{bmatrix} A' A \\ F_A P \end{bmatrix} \right) U,$$

where $U \in \mathbb{R}^{n \times m}$ is arbitrary. In particular, the global maximizer is unique if and only if $r(F_A P F_A) = r(F_A)$.

(b) There exists an $\tilde{X} \in S$ such that (2.23) holds if and only if

$$F_A P F_A \preceq 0 \quad \text{and} \quad \mathcal{R} \left[ \begin{array}{cc} 0 & A' A \\ A' B & 0 \end{array} \right] \subseteq \mathcal{R} \left[ \begin{array}{cc} P & A' A \\ A' A & 0 \end{array} \right].$$

In this case, the global maximizer $\tilde{X}$ satisfying (2.23) is determined by the consistent matrix equation

$$\begin{bmatrix} A' A \\ F_A P \end{bmatrix} \tilde{X} = \begin{bmatrix} A' B \\ 0 \end{bmatrix},$$

and the global minimizer is given by

$$\argmin \{ Q - X'PX | X \in S \} = \begin{bmatrix} A' A \\ F_A P \end{bmatrix} \begin{bmatrix} A' B \\ 0 \end{bmatrix} + \left( I_n - \begin{bmatrix} A' A \\ F_A P \end{bmatrix} \right) U,$$

where $U \in \mathbb{R}^{n \times m}$ is arbitrary. The global minimizer is unique if and only if $r(F_A P F_A) = r(F_A)$.

Proof. Let

$$\psi_M(X) = \psi(X) - \psi(\tilde{X}) = \tilde{X}'P\tilde{X} - X'PX, \quad \psi_m(X) = \psi(X) - \psi(\tilde{X}) = \tilde{X}'P\tilde{X} - X'PX.$$

Then, (2.22) and (2.23) are equivalent to

$$\psi_M(X) \preceq 0 \quad \text{for all} \quad X \in S,$$

$$\psi_m(X) \succeq 0 \quad \text{for all} \quad X \in S,$$  

respectively. From Corollary 2.2(j) and (l), (2.30) and (2.31) are equivalent to

$$i_+ \begin{bmatrix} \tilde{X}'P\tilde{X} & B'A & 0 \\ A'B & 0 & A'A \\ 0 & A'A & -P \end{bmatrix} = r(A), \quad \tilde{X} \in S,$$

$$i_- \begin{bmatrix} \tilde{X}'P\tilde{X} & B'A & 0 \\ A'B & 0 & A'A \\ 0 & A'A & -P \end{bmatrix} = r(A), \quad \tilde{X} \in S,$$

respectively. It is easy to derive from congruence transformations for symmetric matrices and (1.21) that

$$i_+ \begin{bmatrix} \tilde{X}'P\tilde{X} & B'A & 0 \\ A'B & 0 & A'A \\ 0 & A'A & -P \end{bmatrix} = i_+ \begin{bmatrix} 0 & 0 & \tilde{X}'P \\ A'B & 0 & A'A \\ 0 & A'A & -P \end{bmatrix} = r(A) + i_+ \begin{bmatrix} 0 & F_A P \tilde{X} & -F_A P F_A \\ \tilde{X}'P \tilde{X} & A'A & -P \end{bmatrix} \succeq r(A),$$

$$i_- \begin{bmatrix} \tilde{X}'P\tilde{X} & B'A & 0 \\ A'B & 0 & A'A \\ 0 & A'A & -P \end{bmatrix} = i_- \begin{bmatrix} 0 & 0 & \tilde{X}'P \\ A'B & 0 & A'A \\ 0 & A'A & -P \end{bmatrix} = r(A) + i_- \begin{bmatrix} 0 & F_A P \tilde{X} & -F_A P F_A \\ \tilde{X}'P \tilde{X} & A'A & -P \end{bmatrix} \preceq r(A).$$
In consequence, (2.32) and (2.33) are equivalent to

\[
F_A PF_A \geq 0, \quad \begin{bmatrix} A'A \\ F_A P \end{bmatrix} \tilde{X} = \begin{bmatrix} A'B \\ 0 \end{bmatrix}, \tag{2.36}
\]

\[
F_A PF_A \leq 0, \quad \begin{bmatrix} A'A \\ F_A P \end{bmatrix} \tilde{X} = \begin{bmatrix} A'B \\ 0 \end{bmatrix}, \tag{2.37}
\]

respectively. From Lemma 1.4, the matrix equation in (2.36) has a solution if and only if

\[
r \begin{bmatrix} A'A & A'B \\ F_A P & 0 \end{bmatrix} = r \begin{bmatrix} A'A \\ F_A P \end{bmatrix}, \ i.e., \ r \begin{bmatrix} P & A'A \\ A'A & 0 \end{bmatrix} = r \begin{bmatrix} P & A'A \\ A'A & 0 \end{bmatrix}. \tag{2.38}
\]

In this case, the general solution of (2.36) can be written as (2.26). The results in (b) can be shown similarly.

Optimization problems of quadratic matrix-valued functions subject to linear matrix equation restrictions are classic objects of study in mathematics and other fields. It is always mathematicians’ desire to give analytical solutions of optimization problems from the theoretical and applied points of view. Theorem 2.6 provides a group of exact algebraic solutions to the two optimization problems in (2.22) and (2.23) for a special quadratic matrix-valued function subject to linear matrix equation restriction, which we believe will serve as a standard tool in treating various theoretical and applied problems related to (1.1) and (1.4).

3 An application to least-squares estimators under a general linear model

Statistical methods in many areas of application require mathematical computations with vectors and matrices, while various formulas and algebraic tricks for handling matrices in linear algebra and matrix theory play important roles in the derivations and characterizations of estimators and their properties under linear statistical analysis. Let us consider a general linear model defined by

\[
y = X\beta + e, \quad E(e) = 0, \quad \text{Cov}(e) = \sigma^2\Sigma, \tag{3.1}
\]

where \( y \) is an \( n \times 1 \) observable random vector, \( X \) is an \( n \times p \) known matrix of arbitrary rank, \( \beta \) is a \( p \times 1 \) fixed but unknown parameter vector, \( e \) is a random error vector, \( \sigma^2 \) is an unknown positive number, and \( \Sigma \) is an \( n \times n \) known positive semi-definite matrix of arbitrary rank.

Linear regression models were the first type of models to be studied rigorously in regression analysis, which were regarded without doubt as a noble and magnificent part in current statistical theory. As demonstrated in most statistical textbooks, a best-known estimator of the unknown parameter vector in (3.1) is the Ordinary Least-Squares Estimator (OLSE for short). Recall that the well-known OLSE of the unknown parameter vector \( \beta \) in (3.1) is defined to be

\[
\hat{\beta} = \arg\min_{\beta \in \mathbb{R}^p} (y - X\beta)'(y - X\beta), \tag{3.2}
\]

while the OLSE of the parametric vector \( K\beta \) under (3.1) is defined to be \( K\hat{\beta} \). A direct decomposition of the norm \( (y - X\beta)'(y - X\beta) \) in (3.2) is

\[
(y - X\beta)'(y - X\beta) = (y - XX'\gamma)'(y - XX'\gamma) + (XX'\gamma - X\beta)'(XX'\gamma - X\beta) = y'Exy + (Pxy - X\beta)'(Pxy - X\beta),
\]

where the two terms on the right-hand side satisfy \( y'Exy \geq 0 \) and \( (Pxy - X\beta)'(Pxy - X\beta) \geq 0 \). Hence, \( \min_{\beta \in \mathbb{R}^p} (y - X\beta)'(y - X\beta) = y'Exy + \min_{\beta \in \mathbb{R}^p} (Pxy - X\beta)'(Pxy - X\beta) = y'Exy \), where the equation \( X\beta = Pxy \), which is equivalent to the so-called normal equation \( X'X\beta = X'y \) by premultiplying \( X' \), is always consistent; see, e.g., [1, p.114] and [8, pp. 164–165]. An alternative definition of the OLSE of \( \beta \) in (3.1) is given by

\[
\hat{\beta} = \tilde{L}y \quad \text{and} \quad \tilde{L} = \arg\min_{L \in \mathbb{R}^{n \times n}} (y - XLy)'(y - XLy). \tag{3.3}
\]

Also notice that \( (y - XLy)'(y - XLy) \) in (3.3) can be decomposed as

\[
(y - XLy)'(y - XLy) = y'Exy + y'(P - XL)'(P - XL)y,
\]
where \( y'Xy \geq 0 \) and \( y'(P_X - XL)(P_X - XL)y \geq 0 \). Hence,

\[
\min_{L \in \mathbb{R}^{p \times n}} (y - XLy)'(y - XLy) = y'Exy + \min_{L \in \mathbb{R}^{p \times n}} y'(P_X - XL)'(P_X - XL)y = y'Exy,
\]

where the matrix equation \( XL = P_X \) is always solvable for \( L \), say, \( L = X' \); see [12]. Solving the equation \( X\beta = P_Xy \) by Lemma 1.4 yields the following well-known results.

**Lemma 3.1.** Assume that \( P_Xy \neq 0 \). Then, the OLSEs of \( \beta \) and \( K\beta \) under (3.1) can be written as

\[
\hat{\beta} = (X' + F_XU) P_Xy = (X' + F_XU) P_Xy, \tag{3.4}
\]

\[
K\hat{\beta} = (KX' + KF_XU) P_Xy = (KX' + KF_XU) P_Xy, \tag{3.5}
\]

where \( U \in \mathbb{R}^{p \times n} \) is arbitrary.

The results in Section 3 can be used to derive some new algebraic properties of OLSEs under linear models. It is obvious that the OLSE of the parametric vector \( K\beta \) under (3.1) is not unique, and thus we are interested in establishing certain possible equalities and inequalities for the length of OLSE \( K\beta \). In particular, we directly obtain from Corollary 2.4 a group of results on the length of \( \beta \) as follows.

**Corollary 2.2.** Let \( \hat{\beta} \) be given as in (3.4), and \( s > 0 \). Also, let \( T = \begin{bmatrix} s & y'X (X'X)^{-1} \end{bmatrix} \). Then,

\[
\max_{\hat{\beta}} r(s - ||\hat{\beta}||^2) = \min \{ 1, \ p + r(T) - 2r(X) \},
\]

\[
\min_{\hat{\beta}} r(s - ||\hat{\beta}||^2) = \max \{ r(T) - p, \ i_-(T) \},
\]

\[
\max_{\hat{\beta}} i_\pm(s - ||\hat{\beta}||^2) = \min \{ 1, \ i_\pm(T) - r(X) + i_+(I_p) \},
\]

\[
\min_{\hat{\beta}} i_\pm(s - ||\hat{\beta}||^2) = \max \{ 0, \ i_\pm(T) - i_+(I_p) \}.
\]

In conclusion, the following results hold.

(a) There exists \( \hat{\beta} \) such that \( ||\hat{\beta}||^2 \neq s \) if and only if \( r(T) \geq 2r(X) + 1 - p \).

(b) \( ||\hat{\beta}||^2 \neq s \) for all OLSEs if and only if \( r(T) = 1 + p \) or \( i_-(T) = 1 \).

(c) There exists \( \hat{\beta} \) such that \( ||\hat{\beta}||^2 = s \) if and only if \( T \succ 0 \) and \( r(T) \leq p \).

(d) All OLSEs satisfy \( ||\hat{\beta}||^2 = s \) if and only if \( r(T) = 2r(X) - p \).

(e) There exists \( \hat{\beta} \) such that \( ||\hat{\beta}||^2 < s \) if and only if \( i_+(T) \geq r(X) + 1 \).

(f) All OLSEs satisfy \( ||\hat{\beta}||^2 < s \) if and only if \( i_+(T) = 1 + p \).

(g) There exists \( \hat{\beta} \) such that \( ||\hat{\beta}||^2 > s \) if and only if \( i_-(T) \geq r(X) + 1 - p \).

(h) All OLSEs satisfy \( ||\hat{\beta}||^2 > s \) if and only if \( i_-(T) = 1 \).

(i) There exists \( \hat{\beta} \) such that \( ||\hat{\beta}||^2 \leq s \) if and only if \( T \succ 0 \).

(j) All OLSEs satisfy \( ||\hat{\beta}||^2 \leq s \) if and only if \( r(X) = p \) and \( T \preceq 0 \).

(k) There exists \( \hat{\beta} \) such that \( ||\hat{\beta}||^2 \geq s \) if and only if \( i_+(T) \leq p \).

(l) All OLSEs satisfy \( ||\hat{\beta}||^2 \geq s \) if and only if \( i_+(T) = r(X) \).

It can be figured out that equalities and inequalities of \( ||K\hat{\beta}||^2 \) in (3.5) can be established as well from Corollary 2.4. More problems on extremal quadratic properties of OLSEs, in particular, many new properties of covariance matrices of OLSEs, as well as other types of estimator of unknown parameters under linear regression models be can be proposed and solved from the results in Section 3.

All the results in the this paper once again show some essential applications of matrix rank and inertia formulas in characterizing matrix equalities and inequalities, and deriving exact algebraic solutions of matrix-valued function optimization problems in the Löwner partial ordering. In fact, thousands of formulas for calculating ranks/inertias of matrices like those in this paper were obtained by many authors since 1980s, which showed bright insight into matrix rank/inertia theory, and provided at the same time significant
advances to general algebraical methodology in linear algebra and matrix theory. Many matrix problems solved by matrix rank/inertia formulas, as demonstrated above, are so simple and explicit that they are easy to understand, and thus are easy to accept and use in matrix theory and applications. It is really necessary for mathematicians to see the great truth hidden behind the matrix rank/inertia formulas, and to recognize that the materials in textbooks on matrix rank/inertia formulas and their applications really need to update for linear algebra learners. It is no doubt that matrix rank/inertia theory will definitely re-dominate, as its orthodox origination, the linear algebra and matrix theory in the coming future. Before stepping into the world of advanced mathematics, it is really better for all mathematicians to master the basic skills of establishing matrix equalities/inequalities by the matrix rank/inertia methods.

References