Matrix rank/inertia formulas for least-squares solutions with statistical applications

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

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

Abstract. Least-Squares Solution (LSS) of a linear matrix equation and Ordinary Least-Squares Estimator (OLSE) of unknown parameters in a general linear model are two standard algebraical methods in computational mathematics and regression analysis. Assume that a symmetric quadratic matrix-valued function \( \phi(Z) = Q - ZPZ' \) is given, where \( Z \) is taken as the LSS of the linear matrix equation \( AZ = B \). In this paper, we establish a group of formulas for calculating maximum and minimum ranks and inertias of \( \phi(Z) \) subject to the LSS of \( AZ = B \), and derive many quadratic matrix equalities and inequalities for LSSs from the rank and inertia formulas. This work is motivated by some inference problems on OLSEs under general linear models, while the results obtained can be applied to characterize many algebraical and statistical properties of the OLSEs.

Keywords: Linear model, matrix equation, LSS, OLSE, quadratic matrix-valued function, rank, inertia

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 \( (XA)' = XA \). \( E_A \) and \( F_A \) stand for \( E_A = I_m - AA^+ \) and \( F_A = I_n - A^+A \) with \( r(E_A) = m - r(A) \) and \( r(F_A) = n - r(A) \). The symbols \( i_+(A) \) and \( i_-(A) \) for \( A = A' \in \mathbb{R}^{m \times n} \), called the partial inertia of \( A \), 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 \( i_{\pm}(A) \) to denote the both numbers. The Frobenius norm of a matrix \( A \) is defined to be \( \|A\|_F = \sqrt{\text{trace}(AA')} \). \( A \succ 0, A \succeq 0, A \prec 0, \) and \( A \preceq 0 \) mean that \( A \) is symmetric positive definite, positive semi-definite, negative definite, and negative semi-definite, respectively. Two symmetric matrices \( A \) and \( B \) of same size are said to satisfy the inequalities \( A \succ B, A \succeq B, A \prec B, \) and \( A \preceq B \) 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 Hermitian matrices. For more issues about connections between the inertias and the Löwner partial ordering of Hermitian matrices, as well as applications of the matrix inertia and Löwner partial ordering in statistics, see, e.g., [17, 19].

Consider a general linear model defined by

\[
y = X\beta + e, \quad E(e) = 0, \quad \text{Cov}(e) = \sigma^2\Sigma, \tag{1.1}
\]

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, \( \sigma^2 \) is an unknown positive number, \( \Sigma \) is an \( n \times n \) known positive semi-definite matrix of arbitrary rank.

Recall that the well-known Ordinary Least Squares Estimator (OLSE) of the unknown parameter vector \( \beta \) in (1.1) is defined to be

\[
\hat{\beta} = \arg\min_{\beta} (y - X\beta)'(y - X\beta), \tag{1.2}
\]

while the OLSE of the parametric vector \( K\beta \) under (1.1) is defined to be \( K\hat{\beta} \). A direct decomposition of the norm \( (y - X\beta)'(y - X\beta) \) in (1.2) is

\[
(y - X\beta)'(y - X\beta) = (y - XX'\beta)'(y - XX'\beta) + (XX'\beta - X\beta)'(XX'\beta - X\beta) = y'Exy + (P_{XY} - X\beta)'(P_{XY} - X\beta),
\]

where \( y'Exy \geq 0 \) and \( (P_{XY} - X\beta)'(P_{XY} - X\beta) \geq 0 \). Hence,

\[
\min_{\beta} (y - X\beta)'(y - X\beta) = y'Exy + \min_{\beta} (P_{XY} - X\beta)'(P_{XY} - X\beta) = y'Exy.
\]

E-mail Address: yongge.tian@gmail.com, jiangbo@cafe.edu.cn
where the equation \( X \beta = P_{X}y \), which is equivalent to the so-called normal equation \( X'X \beta = X'y \) by premultiplying \( X' \), is always consistent; see, e.g., [7, p.114] and [18, pp.164–165]. An alternative definition of OLSE of \( \beta \) in (1.1) is given by

\[
\hat{\beta} = \hat{L}g, \quad \hat{L} = \text{argmin}_{L}(y - XLy)'(y - XLy).
\] (1.3)

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

\[
(y - XLy)'(y - XLy) = y'Exy + y'(P_{X} - XL)'(P_{X} - XL)y,
\]

where \(y'Exy \geq 0\) and \(y'(P_{X} - XL)'(P_{X} - XL)y \geq 0\). Hence,

\[
\min_{L}(y - XLy)'(y - XLy) = y'Exy + \min_{L}y'(P_{X} - XL)'(P_{X} - XL)y = y'Exy,
\]

where the matrix equation \( XL = P_{X} \) is always consistent for \( L \), say, \( L = X^{+} \). Solving the equation \( X \beta = P_{X}y \) by Lemma 1.5 below yields the following well-known results.

**Lemma 1.1.** Assume that \( P_{XY} \neq 0 \). Then, the OLSEs of \( \beta \) and \( K \beta \) under (1.1) can be written as

\[
\hat{\beta} = (X'^{+} + F_{X}U)P_{X}y = (X'^{+} + F_{X}UP_{X})y,
\]

\[
K \hat{\beta} = (KX'^{+} + KF_{X}U)P_{X}y = (KX'^{+} + KF_{X}UP_{X})y,
\]

where \( U \in \mathbb{R}^{p \times n} \) is arbitrary. In this setting, the following results hold.

(a) The expectations of \( \hat{\beta} \) and \( K \hat{\beta} \) are

\[
E(\hat{\beta}) = (X'^{+} + F_{X}UX)\beta, \quad E(K \hat{\beta}) = (KX'^{+} + KF_{X}UX)\beta.
\]

(b) The dispersion matrices of \( \hat{\beta} \) and \( K \hat{\beta} \) are

\[
\text{Cov}(\hat{\beta}) = \sigma^{2}(X'^{+} + F_{X}U)P_{X}\Sigma P_{X}(X'^{+} + F_{X}U)',
\]

\[
\text{Cov}(K \hat{\beta}) = \sigma^{2}(KX'^{+} + KF_{X}U)P_{X}\Sigma P_{X}(KX'^{+} + KF_{X}U}'.
\]

(c) The matrix mean square errors (MMSEs) of \( \hat{\beta} \) and \( K \hat{\beta} \) are

\[
\text{MMSE}(\hat{\beta}) = E[(\hat{\beta} - \beta)(\hat{\beta} - \beta)'] = \text{Cov}(\hat{\beta}) + \text{Bias}(\hat{\beta})\text{Bias}(\hat{\beta})',
\]

\[
\text{MMSE}(K \hat{\beta}) = E[(K \hat{\beta} - K \beta)(K \hat{\beta} - K \beta)'] = \text{Cov}(K \hat{\beta}) + \text{Bias}(K \hat{\beta})\text{Bias}(K \hat{\beta})'.
\]

(d) There exists \( U \in \mathbb{R}^{p \times n} \) such that

\[
E(K \hat{\beta}) = (KX'^{+} + KF_{X}UX)\beta = K \beta
\]

holds for all \( \beta \) if and only if \( \mathcal{R}(K') \subseteq \mathcal{R}(X') \), namely, \( K \beta \) is estimable. In this case,

\[
K \hat{\beta} = KX'^{+}y, \quad E(K \hat{\beta}) = K \beta, \quad \text{MMSE}(K \hat{\beta}) = \text{Cov}(K \hat{\beta}) = \sigma^{2}KX'^{+}\Sigma(KX'^{+})';
\]

in particular,

\[
X \hat{\beta} = P_{X}y, \quad E(X \hat{\beta}) = X \beta, \quad \text{MMSE}(X \hat{\beta}) = \text{Cov}(X \hat{\beta}) = \sigma^{2}P_{X}\Sigma P_{X}.
\]

In the inference theory of linear models, there has been considerable interest in establishing estimators of the unknown parameters by certain linear functions of the observed response vectors in the models. The OLSE of unknown parameters in a linear model, as described above, is a simplest linear estimator with extensive applications in regression analysis, while many results on computational and algebraic properties of OLSEs were established in the statistical literature. Once an estimator is defined and derived, it is always desirable to know more behaviors of the estimators under the models. In particular, equalities and inequalities for the covariance matrices of given estimators and the corresponding matrix risk functions, such as

\[
\text{Cov}(K \hat{\beta}) = Q \quad (\succ Q \succ Q, \prec Q, \preceq Q), \quad \text{MMSE}(K \hat{\beta}) = Q \quad (\succ Q \succ Q, \prec Q, \preceq Q)
\]

for a given symmetric matrix \( Q \), as well as the matrix minimization problems

\[
\text{Cov}(K \hat{\beta}) \overset{1}{=} \min, \quad \text{MMSE}(K \hat{\beta}) \overset{1}{=} \min
\]
in the Löwner partial ordering, can be used to characterize mathematical and statistical properties of the estimators. Note from Lemma 1.1 that $\text{Cov}(K\hat{\beta})$ and $\text{MMSE}(K\hat{\beta})$ are in fact symmetric quadratic matrix-valued functions with arbitrary matrix $U$. Hence, equalities and inequalities for $\text{Cov}(K\hat{\beta})$ and $\text{MMSE}(K\hat{\beta})$ depend on the choices of $U$. This work can be reformulated as some general problems on establishing equalities and inequalities for a symmetric quadratic matrix-valued function subject to the Least-Squares Solution (LSS) of a linear matrix equation as follows.

Problem 1.2. Let
\[
AZ = B
\]  
(1.4)

be a linear matrix equation, where $A \in \mathbb{R}^{p \times n}$ and $B \in \mathbb{R}^{p \times m}$ are two given matrices. Then, the LSS of (1.4) is defined to be the matrix
\[
Z_0 = \arg\min_{Z \in \mathbb{R}^{n \times m}} \| AZ - B \|_F^2 = \arg\min_{Z \in \mathbb{R}^{n \times m}} \text{tr}[(AZ - B)'(AZ - B)].
\]
The solution is not necessarily unique, and let $S$ be the collection of all LSSs of (1.4):
\[
S = \{ Z \in \mathbb{R}^{n \times m} \mid \text{tr}[(AZ - B)'(AZ - B)] = \min \}.\]

Further, let $\phi(Z) = Q - ZPZ'$ be a quadratic matrix-valued function, where $P = P' \in \mathbb{R}^{m \times m}$ and $Q = Q' \in \mathbb{R}^{n \times n}$ are given matrices. In this setting, establish exact algebraic formulas for calculating the following six maximum and minimum ranks and inertias
\[
\max_{Z \in S} r(Q - ZPZ'), \quad \min_{Z \in S} r(Q - ZPZ'),
\]
\[
\max_{Z \in S} i_\pm(Q - ZPZ'), \quad \min_{Z \in S} i_\pm(Q - ZPZ').
\]

Problem 1.3. Under the assumptions in Problem 1.2, establish necessary and sufficient conditions for the following two partial ordering optimization problems
\[
\{ Q - ZPZ' \mid Z \in S \} \overset{1}{\overset{\leq}{\equiv}} \max, \quad \{ Q - ZPZ' \mid Z \in S \} \overset{1}{\overset{\leq}{\equiv}} \min
\]
to have solutions, respectively, and give exact algebraic expressions of the solutions.

Problem 1.4. Under the assumptions in Problem 1.2, establish necessary and sufficient conditions for the following constrained quadratic matrix equation
\[
ZPZ' = Q \quad \text{s.t. } Z \in S
\]
to have a solution, as well as necessary and sufficient conditions for the following four constrained quadratic matrix inequalities
\[
ZPZ' \succeq Q, \quad ZPZ' \preceq Q, \quad ZPZ' \equiv Q
\]
to hold for a $Z \in S$ (all matrices $Z \in S$), respectively.

Concerning the LSS of (1.4), we have the following direct derivation. Decompose $(B - AZ)'(B - AZ)$ as
\[
(B - AZ)'(B - AZ) = (B - AA^TB)'(B - AA^TB) + (AA^TB - AZ)'(AA^TB - AZ)
\]
\[
= B'E_AB + (P_AB - AZ)'(P_AB - AZ).
\]
Hence,
\[
\text{tr}[(B - AZ)'(B - AZ)] = \text{tr}(B'E_AB) + \text{tr}[(P_AB - AZ)'(P_AB - AZ)],
\]
\[
\min_{Z \in \mathbb{R}^{n \times m}} \text{tr}[(B - AZ)'(B - AZ)] = \text{tr}(B'E_AB) + \min_{Z \in \mathbb{R}^{n \times m}} \text{tr}[(P_AB - AZ)'(P_AB - AZ)]
\]
\[
= \text{tr}(B'E_AB),
\]
where the equation $AZ = P_AB$, which is equivalent to $A'AZ = A'B$, is always consistent. A seminal result due to [16] on (1.4) is given below.

Lemma 1.5. The matrix equation in (1.4) has a solution if and only if $AA^TB = B$. In this case, the general solution can be written in the following parametric form $Z = A^TB + F_AV$, where $V \in \mathbb{R}^{n \times m}$ is arbitrary. The solution of (1.4) is unique if and only if $r(A) = n.$ If (1.4) is inconsistent, then the normal equation of (1.4) is $A'AZ = A'B$, and the general expression of the LSSs of (1.4) can be written as $Z = A^TB + F_AV$, where $V \in \mathbb{R}^{n \times m}$ is arbitrary. The LSSs of (1.4) is unique if and only if $r(A) = n.$
Lemma 1.6 ([15]). The pair of matrix equations $AZ = B$ and $ZC = D$ have a common solution if and only if $AA^tB = B$, $DC^tC = D$ and $AD = BC$. In this case, the general common solution can be written in the following parametric form $Z = A^tB + DC^t - A^tADC + F_{AV}E_{CC}$, where $V \in \mathbb{R}^{n \times m}$ is arbitrary.

Lemma 1.7 ([21]). Let $\phi(Z)$ be as given in Problem 1.2, and assume that (1.4) is consistent. Then,

$$
\max_{Z \in S} r(Q - ZPZ') = \min \{ 2n + r(AQA' - BPB') - 2r(A), n + r[AQ, BP] - r(A), r(Q) + r(P) \},
$$

$$
\min_{Z \in S} r(Q - ZPZ') = \max \{ s_1, s_2, s_3, s_4 \},
$$

$$
\max_{Z \in S} i_\pm(Q - ZPZ') = \min \{ n + i_\pm(AQA' - BPB'), r(A), i_\pm(Q) + i_\mp(P) \},
$$

$$
\min_{Z \in S} i_\pm(Q - ZPZ') = \max \{ u_\pm, v_\pm \},
$$

where

$$
s_1 = r(AQA' - BPB') + 2r[ AQ, BP] - 2r[AQA', BP],
$$

$$
s_2 = 2r[ AQ, BP] + r(Q) - r(P) - 2r(AQ),
$$

$$
s_3 = 2r[ AQ, BP] + i_\pm(AQA' - BPB') - r[AQA', BP] + i_\pm(Q) - i_\mp(P) - r(AQ),
$$

$$
s_4 = 2r[ AQ, BP] + i_\pm(AQA' - BPB') - r[AQA', BP] + i_\mp(Q) - i_\mp(P) - r(AQ),
$$

$$
u_\pm = i_\pm(AQA' - BPB') + r[ AQ, BP] - r[AQA', BP],
$$

$$
v_\pm = r[ AQ, BP] + i_\pm(Q) - i_\mp(P) - r(AQ).
$$

There do exist matrices $Z$ satisfying the six formulas in 1.7, while the constructions of these matrices were described or formulated in [21] or other papers.

Lemma 1.8 ([19]). Let $S$ be a set consisting of matrices over $\mathbb{R}^{m \times n}$, and let $H$ be a set consisting of symmetric matrices over $\mathbb{R}^{m \times m}$. Then, the following assertions hold.

(a) Under $m = n$, there exists a nonsingular matrix $Z \in S$ if and only if $\max_{Z \in S} r(Z) = m$.

(b) Under $m = n$, all $Z \in S$ are nonsingular if and only if $\min_{Z \in S} r(Z) = m$.

(c) $0 \in S$ if and only if $\min_{Z \in S} r(Z) = 0$.

(d) $S = \{ 0 \}$ if and only if $\max_{Z \in S} r(Z) = 0$.

(e) All $Z \in S$ have the same rank if and only if $\max_{Z \in S} r(Z) = \min_{Z \in S} r(Z)$.

(f) $H$ has a matrix $Z \succ 0$ ($Z \prec 0$) if and only if $\max_{Z \in H} i_+(Z) = m$ ($\max_{Z \in H} i_-(Z) = m$).

(g) All $Z \in H$ satisfy $Z \succ 0$ ($Z \prec 0$) if and only if $\min_{Z \in H} i_+(Z) = m$ ($\min_{Z \in H} i_-(Z) = m$).

(h) $H$ has a matrix $Z \succ 0$ ($Z \prec 0$) if and only if $\min_{Z \in H} i_+(Z) = 0$ ($\min_{Z \in H} i_+(Z) = 0$).

(i) All $Z \in H$ satisfy $Z \succ 0$ ($Z \prec 0$) if and only if $\max_{Z \in H} i_-(Z) = 0$ ($\max_{Z \in H} i_+(Z) = 0$).

(j) All $Z \in H$ have the same positive index of inertia if and only if $\max_{Z \in H} i_+(Z) = \min_{Z \in H} i_+(Z)$.

(k) All $Z \in H$ have the same negative index of inertia if and only if $\max_{Z \in H} i_-(Z) = \min_{Z \in H} i_-(Z)$.

The assertions in Lemma 1.8 directly follow from the definitions of rank/inertia, definiteness, and semi-definiteness of (symmetric) matrices. These assertions show that if certain expansion formulas for calculating ranks/inertias of differences of (symmetric) matrices are established, we can use them to characterize the corresponding matrix equalities and inequalities. This fact reflects without doubt the most exciting values of ranks/inertias in matrix analysis and applications, and thus it is really necessary to produce thousands of matrix rank/inertia formulas from the theoretical and applied points of view.

2 Main results

Theorem 2.1. Let $\phi(Z)$ and $S$ be as given in Problem 1.2. Then, the following results hold.

(a) The maximum rank of $\phi(Z)$ subject to $Z \in S$ is

$$
\max_{Z \in S} r(Q - ZPZ') = \min \{ 2n + r(A'AQA' - A'BPB'A) - 2r(A), n + r[A'AQ, A'BP] - r(A), r(Q) + r(P) \}.
$$
(b) The minimum rank of $\phi(Z)$ subject to $Z \in S$ is
\[
\min_{Z \in S} r(Q - ZPZ') = \max \{ s_1, s_2, s_3, s_4 \},
\]
where
\[
s_1 = r(A'QA'A - A'BPB'B) + 2r[A'AQ, A'BP] - 2r[A'QA', A'BP],
\]
\[
s_2 = 2r[A'AQ, A'BP] + r(Q) - r(P) - 2r(AQ),
\]
\[
s_3 = 2r[A'AQ, A'BP] + i_+(A'Q'AA - A'BPB'B) - r[A'QA', A'BP]
\]
\[
+ i_-(Q) - i_-(P) - r(AQ),
\]
\[
s_4 = 2r[A'AQ, A'BP] + i_-(A'Q'AA - A'BPB'B) - r[A'QA', A'BP]
\]
\[
+ i_+(Q) - i_+(P) - r(AQ).
\]

(c) The maximum partial inertia of $\phi(Z)$ subject to $Z \in S$ are
\[
\max_{Z \in S} i_\pm(Q - ZPZ') = \min \{ n + i_\pm(A'Q'AA - A'BPB'B) - r(A), i\pm(Q) + i\pm(P) \}.
\]

(d) The minimum partial inertia of $\phi(Z)$ subject to $Z \in S$ are
\[
\min_{Z \in S} i_\pm(Q - ZPZ') = \max \{ u_\pm, v_\pm \},
\]
where
\[
u_\pm = r[A'AQ, A'BP] + i_\pm(Q) - i_\pm(P) - r(AQ),
\]
\[
u_\pm = r[A'AQ, A'BP] + i_\pm(Q) - i_\pm(P) - r(AQ).
\]

Proof. It can be seen from Lemma 1.5 that
\[
\max_{Z \in S} r(Q - ZPZ') = \max_{A'AZ = A'B} r(Q - ZPZ'), \quad \min_{Z \in S} r(Q - ZPZ') = \min_{A'AZ = A'B} r(Q - ZPZ'),
\]
\[
\max_{Z \in S} i_\pm(Q - ZPZ') = \max_{A'AZ = A'B} i_\pm(Q - ZPZ'), \quad \min_{Z \in S} i_\pm(Q - ZPZ') = \min_{A'AZ = A'B} i_\pm(Q - ZPZ').
\]

In these cases, replacing $A$ with $A'A$ and $B$ with $A'B$ in Lemma 1.7 and simplifying, we obtain the formulas in (a)–(d).

Applying Lemma 1.8 to Theorem 2.1, we obtain the following consequences. Their proofs are omitted.

**Corollary 2.2.** Let $\phi(Z)$ and $S$ be as given in Problem 1.2, $s_1, \ldots, s_4$ be as given in Lemma 1.7. Then, the following results hold.

(a) $AZ = B$ has a LSS such that $Q - ZPZ'$ is nonsingular if and only if
\[
r(A'QA'A - A'BPB'B) \geq 2r(A) - n, \quad r[A'AQ, A'BP] = r(A), \quad r(Q) + r(P) \geq n.
\]

(b) $Q - ZPZ'$ is nonsingular for all LSSs of $AZ = B$ if and only if one of $s_i = n, i = 1, \ldots, 4$ holds.

(c) $AZ = B$ has an LSS such that $ZPZ' = Q$ if and only if
\[
A'Q'AA = A'BPB'B, \quad \mathcal{P}(A'AQ) \subseteq \mathcal{P}(A'BP),
\]
\[
i_\pm(Q) \leq i_\pm(P), \quad r(A'BP) + i_\pm(Q) \leq i_\pm(P) + r(AQ).
\]

(d) All LSSs of $AZ = B$ satisfy $ZPZ' = Q$ if and only if $r(A) = n$ and $A'QA'A = A'BPB'B$, or $Q = 0$ and $P = 0$.

(e) $AZ = B$ has a LSS such that $Q - ZPZ' > 0$ if and only if
\[
i_\pm(A'Q'AA - A'BPB'B) = r(A) \text{ and } i_\pm(Q) + i_\pm(P) \geq n.
\]

(f) All LSSs of $AZ = B$ satisfy $Q - ZPZ' > 0$ if and only if
\[
i_\pm(A'Q'AA - A'BPB'B) + r[A'AQ, A'BP] = n + r[A'QA', A'BP]
\]
or
\[
r[A'AQ, A'BP] + i_\pm(Q) = n + i_\pm(P) + r(AQ).
\]

5
(g) $AZ = B$ has a LSS such that $Q - ZPZ' < 0$ if and only if
$$i_-(A'AQA'A - A'BPB'A) = r(A) \text{ and } i_-(Q) + i_+(P) \geq n.$$ 

(h) All LSSs of $AZ = B$ satisfy $Q - ZPZ' < 0$ if and only if
$$i_-(A'AQA'A - A'BPB'A) + r[A'AQ, A'BP] = n + r[A'AQA', A'BP]$$
or
$$r[A'AQ, A'BP] + i_-(Q) = n + i_-(P) + r(AQ).$$

(i) $AZ = B$ has a LSS such that $Q - ZPZ' \succ 0$ if and only if
$$A'AQA'A \succ A'BPB'A, \quad r[A'AQA', A'BP] = r[A'AQ, A'BP] \leq i_-(P) - i_-(Q) + r(AQ).$$

(j) All LSSs of $AZ = B$ satisfy $Q - ZPZ' \succ 0$ if and only if $r(A) = n$ and $A'AQA'A \succ A'BPB'A$, or $Q \succ 0$ and $P \preceq 0$;

(k) $AZ = B$ has a LSS such that $Q - ZPZ' \preceq 0$ if and only if
$$A'AQA'A \preceq A'BPB'A, \quad r[A'AQA', A'BP] = r[A'AQ, A'BP] \leq i_+(P) - i_+(Q) + r(AQ).$$

(l) All LSSs of $AZ = B$ satisfy $Q - ZPZ' \preceq 0$ if and only if $r(A) = n$ and $A'AQA'A \preceq A'BPB'A$, or $Q \preceq 0$ and $P \succeq 0$.

The following two corollaries follow directly from Theorem 2.1.

**Corollary 2.3.** Let $\phi(Z)$ and $S$ be as given in Problem 1.2 with $P \succ 0$ and $Q \succ 0$. Then,

$$\max_{Z \in S} r(Q - ZPZ') = \min \{n, \quad 2n + r(A'AQA'A - A'BPB'A) - 2r(A)\},$$

$$\min_{Z \in S} r(Q - ZPZ') = \max \{r(A'AQA'A - A'BPB'A), \quad i_-(A'AQA'A - A'BPB'A) + n - m\},$$

$$\max_{Z \in S} i_+(Q - ZPZ') = \min \{n + i_+(A'AQA'A - A'BPB'A) - r(A), \quad i_+(I_n) + i_+(I_m)\},$$

$$\min_{Z \in S} i_+(Q - ZPZ') = \max \{i_+(A'AQA'A - A'BPB'A), \quad i_+(I_n) - i_+(I_m)\}.$$

In consequence, the following results hold.

(a) $AZ = B$ has a LSS such that $Q - ZPZ'$ is nonsingular if and only if
$$r(A'AQA'A - A'BPB'A) \geq 2r(A) - n.$$ 

(b) $Q - ZPZ'$ is nonsingular for all LSSs of $AZ = B$ if and only if
$$r(A'AQA'A - A'BPB'A) = n \quad \text{or} \quad i_-(A'AQA'A - A'BPB'A) = m.$$ 

(c) $AZ = B$ has a least-squares solution such that $ZPZ' = Q$ if and only if $A'AQA'A = A'BPB'A$ and $m \geq n$.

(d) All LSSs of $AZ = B$ satisfy $ZPZ' = Q$ if and only if $A'AQA'A = A'BPB'A$ and $r(A) = n$.

(e) $AZ = B$ has a LSS such that $Q \succeq ZPZ'$ if and only if $i_+(A'AQA'A - A'BPB'A) = r(A)$.

(f) All LSSs of $AZ = B$ satisfy $Q \succeq ZPZ'$ if and only if $i_+(A'AQA'A - A'BPB'A) = n$.

(g) $AZ = B$ has a LSS such that $Q \preceq ZPZ'$ if and only if $i_-(A'AQA'A - A'BPB'A) = r(A)$ and $m \geq n$.

(h) All LSSs of $AZ = B$ satisfy $Q \preceq ZPZ'$ if and only if $i_-(A'AQA'A - A'BPB'A) = n$.

(i) $AZ = B$ has a LSS such that $Q \preceq ZPZ'$ if and only if $A'AQA'A \succ A'BPB'A$.

(j) All LSSs of $AZ = B$ satisfy $Q \preceq ZPZ'$ if and only if $A'AQA'A \succ A'BPB'A$ and $r(A) = n$.

(k) $AZ = B$ has a LSS such that $Q \preceq ZPZ'$ if and only if $A'AQA'A \preceq A'BPB'A$ and $n \leq m$.

(l) All LSSs of $AZ = B$ satisfy $Q \preceq ZPZ'$ if and only if $A'AQA'A \preceq A'BPB'A$ and $r(A) = n$. 


Corollary 2.4. Let \( S \) be as given in Problem 1.2. Then,

\[
\begin{align*}
\max_{Z \in S} r(I_n - ZZ') &= \min\{n, 2n + r(A'AA' - A'BB'A) - 2r(A)\}, \\
\min_{Z \in S} r(I_n - ZZ') &= \max\{r(A'AA' - A'BB'A), i_-(A'AA' - A'BB'A) + n - m\}, \\
\max_{I_\pm(I_n - ZZ')} &= \{n + i_\pm(A'AA' - A'BB'A) - r(A), i_\pm(I_n) + i_\mp(I_m)\}, \\
\min_{I_\pm(I_n - ZZ')} &= \max\{i_\pm(A'AA' - A'BB'A), i_\pm(I_n) - i_\pm(I_m)\}.
\end{align*}
\]

In consequence, the following results hold.

(a) \( AZ = B \) has a LSS such that \( I_n - ZZ' \) is nonsingular if and only if \( r(A'AA' - A'BB'A) \geq 2r(A) - n \).

(b) \( I_n - ZZ' \) is nonsingular for all LSSs of \( AZ = B \) if and only if \( r(A'AA' - A'BB'A) = n \) or \( i_-(A'AA' - A'BB'A) = m \).

(c) \( AZ = B \) has a LSS such that \( ZZ' = I_n \), i.e., the rows of a solution of \( AZ = B \) are orthogonal, if and only if \( A'AA' = A'BB'A \) and \( m \geq n \).

(d) All LSSs of \( AZ = B \) satisfy \( ZZ' = I_n \) if and only if \( A'AA' = A'BB'A \) and \( r(A) = n \).

(e) \( AZ = B \) has a LSS such that \( ZZ' < I_n \) if and only if \( i_+(A'AA' - A'BB'A) = r(A) \).

(f) All LSSs of \( AZ = B \) satisfy \( ZZ' < I_n \) if and only if \( i_+(A'AA' - A'BB'A) = n \).

(g) \( AZ = B \) has a LSS such that \( ZZ' > I_n \) if and only if \( i_-(A'AA' - A'BB'A) = r(A) \) and \( m \geq n \).

(h) All LSSs of \( AZ = B \) satisfy \( ZZ' > I_n \) if and only if \( i_-(A'AA' - A'BB'A) = n \).

(i) \( AZ = B \) has a LSS such that \( ZZ' \leq I_n \) if and only if \( A'AA' \succ A'BB'A \).

(j) All LSSs of \( AZ = B \) satisfy \( ZZ' \leq I_n \) if and only if \( i_-(A'AA' - A'BB'A) = n - r(A) \).

(k) \( AZ = B \) has a LSS such that \( ZZ' \geq I_n \) if and only if \( A'AA' \preceq A'BB'A \) and \( m \geq n \).

(l) All LSSs of \( AZ = B \) satisfy \( ZZ' \geq I_n \) if and only if \( i_+(A'AA' - A'BB'A) = n - r(A) \).

In mathematics, the collection of all matrices \( Z \) that satisfy \( ZZ' = I_n \) is called a complex Stiefel manifold; see, e.g., [8, 11], while the collections of all matrices \( Z \) that satisfy \( ZZ' \succeq I_n \) (\( \geq I_n \), \( > I_n \), \( \preceq I_n \)) are called generalized complex Stiefel manifolds. The results in Corollary 2.4 characterize some basic relations between the manifold \( S \) and these Stiefel manifolds.

In the remaining part of this section, we solve Problem 1.3, i.e., to find \( \tilde{Z} \), \( \tilde{Z} \in S \) such that

\[
\begin{align*}
\phi(Z) &\preceq \phi(\tilde{Z}) \quad \text{s.t.} \quad Z \in S, \quad (2.1) \\
\phi(Z) &\succeq \phi(\tilde{Z}) \quad \text{s.t.} \quad Z \in S \quad (2.2)
\end{align*}
\]

hold, respectively.

Theorem 2.5. Let \( \phi(Z) \) and \( S \) be as given in Problem 1.2 with \( r(A) < n \). Then, the following results hold.

(a) There exists a \( \tilde{Z} \in S \) such that (2.1) holds if and only if

\[
P \succeq 0 \quad \text{and} \quad A'BP = 0.
\]

In this case, the maximizer and the corresponding maximum matrix are given by

\[
\text{argmax}_{\succ} \{ \phi(Z) \mid Z \in S \} = A^1B + F_aUF_p, \quad \max_{\succ} \{ \phi(Z) \mid Z \in S \} = Q,
\]

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

(b) There exists a \( \tilde{Z} \in S \) such that (2.2) holds if and only if

\[
P \preceq 0 \quad \text{and} \quad A'BP = 0.
\]

In this case, the minimizer and the corresponding minimum matrix are given by

\[
\text{argmin}_{\preceq} \{ \phi(Z) \mid Z \in S \} = A^1B + F_aUF_p, \quad \min_{\preceq} \{ \phi(Z) \mid Z \in S \} = Q,
\]

where \( U \in \mathbb{R}^{n \times m} \) is arbitrary.
Proof. Under \( r(A) = n \), the LSS of \( AZ = B \) is unique, so that (2.1) and (2.2) become trivial. Let 
\[
\phi_M(Z) = \phi(Z) - \phi(\hat{Z}) = \hat{Z}P\hat{Z}' - ZP', \quad \phi_m(Z) = \phi(Z) - \phi(\hat{Z}) = \hat{Z}P\hat{Z}' - ZP'.
\]
Then, (2.1) and (2.2) are equivalent to 
\[
\phi_M(Z) \preceq 0, \quad Z, \hat{Z} \in \mathcal{S},
\]
\[
\phi_m(Z) \succeq 0, \quad Z, \hat{Z} \in \mathcal{S},
\]
respectively. It can be seen from Corollary 2.2(j) and (l) that they are further equivalent to 
\[
\hat{Z}P\hat{Z}' \preceq 0, \quad A'\hat{A}\hat{Z} = A'B, \quad P \succ 0,
\]
\[
\hat{Z}P\hat{Z}' \succeq 0, \quad A'\hat{A}\hat{Z} = A'B, \quad P \preceq 0,
\]
respectively. The two inequalities \( \hat{Z}P\hat{Z}' \preceq 0 \) and \( \hat{Z}P\hat{Z}' \succeq 0 \) are equivalent to \( \hat{Z}P = 0 \) and \( \hat{Z}P = 0 \) when \( P \) is definite. Hence, the above two groups of equalities and inequalities reduce to 
\[
\hat{Z}P = 0, \quad A'\hat{A}\hat{Z} = A'B, \quad P \succ 0,
\]
\[
\hat{Z}P = 0, \quad A'\hat{A}\hat{Z} = A'B, \quad P \preceq 0,
\]
respectively. From Lemma 1.6, the above two equations for \( \hat{Z} \) have a common solution if and only if \( A'BP = 0 \). In this case, the general common solution \( \hat{Z} \) is \( \hat{Z} = \hat{A}B + F_\Lambda UP \). Substituting this solution into \( \phi(Z) \) gives \( \phi(\hat{Z}) = Q \), establishing (a). Result (b) can be shown similarly. \( \square \)

The results and techniques in this section can be applied to establish many rank/inertia formulas, equalities, and inequalities for the covariance matrices and MMSEs in Lemma 1.1. For instance, let \( X \in \mathbb{R}^{n \times p} \) and denote 
\[
S = \{ Z \mid tr([XZ - I_n]'(XZ - I_n)] = \min \} = \{ Z \mid X'XZ = X' \}.
\]
Then, we obtain from Lemma 1.1(b) and (c) that 
\[
Q - \text{MMSE}(\hat{\beta}) = Q - \sigma^2(X' + F_XU)'P_X\Sigma P_X(X' + F_XU)' - \text{Bias}(\hat{\beta})\text{Bias}(\hat{\beta})',
\]
\[
= Q - \text{Bias}(\hat{\beta})\text{Bias}(\hat{\beta})' - \sigma^2 ZP_X\Sigma P_XZ', \quad Z \in \mathcal{S},
\]
which are special forms of \( S \) and \( \phi(Z) \) in Problem 1.2. Hence, many new and valuable features of the MMSE, as exercises in linear algebra, can be derived from the previous algebraic methods.

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 regression models. In particular, ranks and inertias of real symmetric (complex Hermitian) matrices are conceptual foundation in elementary linear algebra, which are the most significant finite integer quantities in reflecting intrinsic properties of matrices. Dislike the quantities with continuous properties of determents, norms, traces of matrices, rank/inertia are unique quantities to demonstrate finite-dimensional properties of matrix algebra, and are unreplaceable in role and cannot directly be extended to infinite-dimensional matrices and operators. There were many classic approaches on rank/inertia theory of real symmetric (complex Hermitian) matrices and their applications in the mathematical literature; see, e.g., [1, 2, 4, 5, 6, 9, 10, 12]. Motivated by many requirements of establishing matrix equalities and inequalities, the present author reconsidered matrix rank/inertia and established in [13, 14, 19, 20, 21, 22, 23, 27, 28] a variety of formulas for calculating maximum/minimum ranks/inertias of linear/nonlinear real symmetric (complex Hermitian) matrix-valued functions. In addition, the present author’s contributions on matrix rank/inertia formulas also attracted much attention in the field of matrix analysis, and many results on combinations and extensions of the previous work on minmaxity of ranks/inertias of matrix-valued functions by other authors can be found in the literature. It has been realized that matrix rank/inertia formulas can be utilized to characterize many features and performances of real symmetric (complex Hermitian) matrix-valued functions, such as, establishing and simplifying various complicated matrix expressions, deriving matrix equalities/inequalities that involve generalized inverses of matrices, characterizing definiteness/semi-definiteness of real symmetric (complex Hermitian) matrix-valued functions, and deriving exact algebraic solutions to the corresponding real symmetric (complex Hermitian) matrix-valued function optimization problems in the L"{o}wner partial ordering. In fact, thousands of formulas for calculating ranks/inertias of matrices like those in this section were
obtained by the present author 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