AISM (2007) 59: 807–820
DOI 10.1007/s10463-006-0075-3

Yoshio Takane · Yongge Tian · Haruo Yanai

On constrained generalized inverses of matrices and their properties

Received: 24 October 2005 / Revised: 28 March 2006 /
Published online: 8 August 2006

Abstract A matrix G is called a generalized inverse (g-inverse) of matrix A if $AGA = A$ and is denoted by $G = A^-$. Constrained g-inverses of A are defined through some matrix expressions like $E(AE)^-$, $(FA)^-F$ and $E(FAE)^-F$. In this paper, we derive a variety of properties of these constrained g-inverses by making use of the matrix rank method. As applications, we give some results on g-inverses of block matrices, and weighted least-squares estimators for the general linear model.

Keywords Linear matrix expression · Moore–Penrose inverse · Constrained generalized inverses · Matrix equation · Projector · Idempotent matrix · Rank equalities · General linear model · Weighted least-squares estimator

1 Introduction

Throughout this paper, the symbols $\mathbb{R}^{m\times n}$ and $\mathbb{C}^{m\times n}$ stand for the two sets of all $m \times n$ real and complex matrices, respectively; A', A^*, $r(A)$ and $\mathcal{R}(A)$ stand for the transpose, the conjugate transpose, the rank and the range (column space) of a complex matrix A, respectively.

Y. Takane (✉)
Department of Psychology, McGill University, Montréal, QC, Canada H3A 1B1
E-mail: takane@psych.mcgill.ca

Y. Tian
School of Economics, Shanghai University of Finance and Economics, Shanghai 200433, China
E-mail: yongge@mail.shufe.edu.cn

H. Yanai
St. Luke's College of Nursing, Chuo-ku, Tokyo, Japan
E-mail: hyanai@slcn.ac.jp

For an $A \in \mathbb{C}^{m\times n}$, the Moore-Penrose inverse of A, denoted by $A^{\dagger}$, is defined to be the unique matrix X satisfying the four Penrose equations

$$\text{(i) } AXA = A, \quad \text{(ii) } XAX = X, \quad \text{(iii) } (AX)^* = AX, \quad \text{(iv) } (XA)^* = XA.$$

A matrix $X \in \mathbb{C}^{n\times m}$ is called a generalized inverse (g-inverse) of A, denoted by $X = A^-$, if it satisfies (i), while the collection of all g-inverses is denoted by $\{A^-\}$. For convenience, let $P_A = I_m - AA^-$ and $Q_A = I_n - A^-A$. General properties of g-inverses can be found in Ben-Israel and Greville (2003), Campbell and Meyer (1991) and Rao and Mitra (1971).

In addition to the conventional g-inverse A^- of $A \in \mathbb{C}^{m\times n}$, there are many other types of g-inverse of A with constraints. For example, Rao and Mitra (1971, Sect. 4.11) investigated a set of matrices induced by A

$$G_1 = E(AE)^-, \quad G_2 = (FA)^-F, \quad G_3 = E(FAE)^-F, \tag{1}$$

where $E \in \mathbb{C}^{n\times p}$ and $F \in \mathbb{C}^{q\times m}$ are two weighted matrices. These three matrices are not necessarily g-inverses of A for given E and F. It has been shown that

$$\{E(AE)^-\} \subseteq \{A^-\}, \quad \{(FA)^-F\} \subseteq \{A^-\}, \quad \{E(FAE)^-F\} \subseteq \{A^-\} \tag{2}$$

hold iff

$$r(AE) = r(A), \quad r(FA) = r(A), \quad r(FAE) = r(A) \tag{3}$$

hold respectively; see Mitra (1968) and Yanai (1990). In these cases, the three matrices are often called weighted g-inverses of A.

Constrained g-inverses of matrices occur widely in various problems in mathematics and statistics on weighted least-squares solutions and weighted minimum norm solutions of linear matrix equations, weighted least-squares estimators (WLSEs) and best linear unbiased estimators (BLUEs) for general linear models. For example,

(a) It is easy to verify by the definition of g-inverse that

$$\{B^-(ABB^-)^-\} \subseteq \{(AB)^-\}, \quad \{(A^-AB)^-A^-\} \subseteq \{(AB)^-\},$$

$$\{B^-(A^-ABB^-)^-A^-\} \subseteq \{(AB)^-\},$$

$$\{B^*(ABB^*)^-\} \subseteq \{(AB)^-\}, \quad \{(A^*AB)^-A^*\} \subseteq \{(AB)^-\},$$

$$\{B^*(A^*ABB^*)^-A^*\} \subseteq \{(AB)^-\}.$$

Obviously, the matrices $B^-(ABB^-)^-$, $(A^-AB)^-A^-$, $B^-(A^-ABB^-)^-A^-$, $B^*(ABB^*)^-$, $(A^*AB)^-A^*$ and $B^*(A^*ABB^*)^-A^*$ are special cases of (1).

(b) The following set inclusions on g-inverses of block matrices hold

$$\left\{\begin{bmatrix} (P_BA)^-P_B \\ B^- - B^-A(P_BA)^-P_B \end{bmatrix}\right\} \subseteq \{[A,\ B]^-\},$$

$$\{[Q_C(AQ_C)^-,\ C^- - Q_C(AQ_C)^-AC^-]\} \subseteq \left\{\begin{bmatrix} A \\ C \end{bmatrix}^-\right\}.$$

Clearly, $(P_BA)^-P_B$ and $Q_C(AQ_C)^-$ are special cases of (1).

(c) A WLSE of $X\beta$ in the general linear model $y = X\beta + \varepsilon$ with respect to a nonnegative definite (nnd) matrix V can be written as

$$\mathrm{WLSE}(X\beta) = X(X'VX)^{-}X'Vy,$$

where the matrix $(X'VX)^{-}X'V$ is a special case of G_3, see, e.g., Baksalary and Puntanen (1989) and Rao and Mitra (1971).

(d) The weighted Moore–Penrose inverses of an $m \times n$ matrix A with respect to a pair of positive definite matrices M and N is defined to be the unique $n \times m$ matrix X of satisfying the following four matrix equations

$$\text{(i)}AXA = A, \text{ (ii)}XAX = X, \text{ (iii) } (MAX)^* = MAX, \text{ (iv)}(NXA)^* = NXA,$$

and is denoted as $X = A^{\dagger}_{M,N}$. $A^{\dagger}_{M,N}$ can be rewritten as

$$A^{\dagger}_{M,N} = N^{-\frac{1}{2}}(M^{\frac{1}{2}}AN^{-\frac{1}{2}})^{\dagger}M^{\frac{1}{2}},$$

where $M^{\frac{1}{2}}$ and $N^{\frac{1}{2}}$ are the positive definite square roots of M and N, respectively. This expression is a special case of G_3.

(e) Constrained g-inverses and the projectors associated play important roles in instrumental variables (IV) estimation (Takane and Yanai, 1999 and three stage least-squares estimation for regression models in econometrics, see, e.g., Amemiya, 1985; Davidson and MacKinnon, 2004).

Other applications of G_1, G_2 and G_3 in (1) in statistics can be found in Mitra and Rao (1974) and Yanai (1990).

Although constrained g-inverses have been studied by some authors, there are still many problems on constrained g-inverses worth further investigation. As is well known, the rank of matrix provides a powerful method for investigating various properties of g-inverses of matrices. Recall that $AXA = A$ if and only if $r(A - AXA) = 0$. This indicates that rank can be used to characterize g-inverses of matrices. In this paper, we use the matrix rank method for a new investigation of constrained g-inverses, such as uniqueness, maximal and minimal ranks, rank invariance and range invariance of constrained g-inverses. In addition, we give some applications of constrained g-inverses in partitioned matrices and general linear models.

Some basic rank formulas for partitioned matrices and g-inverses due to Marsaglia and Styan (1974) are given in the following lemma.

Lemma 1 *Let $A \in \mathbb{C}^{m\times n}$, $B \in \mathbb{C}^{m\times k}$ and $C \in \mathbb{C}^{l\times n}$. Then,*

$$r[A,\ B] = r(A) + r[(I_m - AA^{-})B] = r(B) + r[(I_m - BB^{-})A], \tag{4}$$

$$r\begin{bmatrix} A \\ C \end{bmatrix} = r(A) + r[C(I_n - A^{-}A)] = r(C) + r[A(I_n - C^{-}C)], \tag{5}$$

$$r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix} = r(B) + r(C) + r[(I_m - BB^{-})A(I_n - C^{-}C)], \tag{6}$$

where A^{-}, B^{-} and C^{-} are any g-inverses of A, B and C.

In addition, we need a set of formulas for extremal ranks of linear matrix expressions in Tian (2002a, b) and Tian and Cheng (2003).

Lemma 2 *Let $A \in \mathbb{C}^{m\times n}$, $B \in \mathbb{C}^{m\times k}$ and $C \in \mathbb{C}^{l\times n}$. Then,*

$$\max_X \, r(A - BX) = \min\{r[A,\ B],\ \ n\}, \tag{7}$$

$$\min_X \, r(A - BX) = r[A,\ B] - r(B), \tag{8}$$

$$\max_X \, r(A - BXC) = \min\left\{r[A,\ B],\ \ r\begin{bmatrix} A \\ C \end{bmatrix}\right\}, \tag{9}$$

$$\min_X \, r(A - BXC) = r[A,\ B] + r\begin{bmatrix} A \\ C \end{bmatrix} - r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix}. \tag{10}$$

Lemma 3 *The maximal and minimal ranks of the Schur complement $D - CA^{-}B$ are given by*

$$\max_{A^-} \, r(D - CA^{-}B) = \min\left\{r[C,\ D],\ \ r\begin{bmatrix} B \\ D \end{bmatrix},\ \ r\begin{bmatrix} A & B \\ C & D \end{bmatrix} - r(A)\right\}, \tag{11}$$

$$\begin{aligned} \min_{A^-} \, r(D - CA^{-}B) &= r(A) + r[C,\ D] + r\begin{bmatrix} B \\ D \end{bmatrix} \\ &\quad + r\begin{bmatrix} A & B \\ C & D \end{bmatrix} - r\begin{bmatrix} A & 0 & B \\ 0 & C & D \end{bmatrix} - r\begin{bmatrix} A & 0 \\ 0 & B \\ C & D \end{bmatrix}. \end{aligned} \tag{12}$$

In particular,

$$\max_{A^-} \, r(CA^{-}B) = \min\left\{r(B),\ \ r(C),\ \ r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix} - r(A)\right\}, \tag{13}$$

$$\min_{A^-} \, r(CA^{-}B) = r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix} - r\begin{bmatrix} A \\ C \end{bmatrix} - r[A,\ B] + r(A). \tag{14}$$

Equalities (11), (12), (13) and (14) are shown in Tian (2002a), see also Tian (2004). Many direct consequences can be derived from (11), (12), (13) and (14). For example, by equating (11) and (12), we can obtain a necessary and sufficient condition for the rank invariance of $D - CA^{-}B$. Equalities (11), (12), (13) and (14) are powerful tools for investigating various matrix expressions involving g-inverses. In Sect. 2, we use (11), (12), (13) and (14) to derive a group of properties of matrix expressions involving g-inverses.

2 Constrained g-inverses with forms $E(AE)^{-}$, $(FA)^{-}F$ and $E(FAE)^{-}F$

Recall that the general expression of A^{-} can be written as

$$A^{-} = A^{\dagger} + U_1(I_m - AA^{\dagger}) + (I_n - A^{\dagger}A)U_2,$$

where U_1 and U_2 are arbitrary. Applying this expressions to G_1, G_2 and G_3 in (1) gives the general expressions

$$G_1 = E(AE)^{\dagger} + EU_1[I_m - AE(AE)^{\dagger}] + [E - E(AE)^{\dagger}AE\,]U_2, \tag{15}$$
$$G_2 = (FA)^{\dagger}F + U_3[F - FA(FA)^{\dagger}F] + [I_n - (FA)^{\dagger}FA\,]U_4F, \tag{16}$$
$$G_3 = E(FAE)^{\dagger}F + EU_5[F - FAE(FAE)^{\dagger}F] + [E-(FAE)^{\dagger}FAE]U_6F, \tag{17}$$

where $U_1, \ldots, U_6$ are arbitrary. Clearly, G_1, G_2 and G_3 are linear matrix expressions that vary as functions of $U_1, \ldots, U_6$. Two other matrices associated with G_1 and G_2 in (1) are given by

$$E(AE)^- + V_1[I_m - AE(AE)^-], \quad (FA)^-F + [I_n - (FA)^-FA]V_2, \tag{18}$$

where V_1 and V_2 are arbitrary. The two matrices in (18) also occur in the expressions of weighted least-squares solutions and minimum norm solutions of linear matrix equations.

Rao and Mitra (1971, Sect. 4.11) and Yanai (1990) considered the constrained g-inverses of the forms given in (1) and their applications when they investigated least-squares solutions and minimum norm solutions of linear matrix equations. Recall that a vector x is called the V-weighted-least-squares solution of the linear matrix equation $Ax = b$ if it minimizes

$$\|b - Ax\|_V^2 = (b - Ax)^*V(b - Ax), \tag{19}$$

where V is an nnd matrix. As is well known, the normal equation corresponding to (19) is $A^*VAx = A^*Vb$. This equation is consistent and the general solution is

$$x_0 = (A^*VA)^-A^*Vb + [I_n - (A^*VA)^-(A^*VA)]v,$$

where $v \in \mathbb{C}^{n\times 1}$ is arbitrary. If $b \neq 0$, x_0 can be written as

$$x_0 = \{(A^*VA)^-A^*V + [I_n - (A^*VA)^-(A^*VA)]U\}b, \tag{20}$$

where U is arbitrary. The matrix $(A^*VA)^-A^*V$ on the right-hand side of (20) is a special case of G_2 in (1) with $F = A^*V$, while the matrix

$$M = (A^*VA)^-A^*V + [I_n - (A^*VA)^-(A^*VA)]U$$

is a special case of the second expression in (18) with $F = A^*V$. According to Mitra and Rao (1974), the product of A and M

$$P_{A:V} \overset{\text{def}}{=} AM = A(A^*VA)^-A^*V + [A - A(A^*VA)^-(A^*VA)]U \tag{21}$$

is called the projector into the range of A under the seminorm $\|\cdot\|_V$, where U is arbitrary. This matrix is not necessarily unique. Thus, one can choose $(A^*VA)^-$ and U such that $P_{A:V}$ has various possible prescribed properties. Some applications of $P_{A:V}$ in a linear linear model are given in Sect. 3.

Because G_1, G_2 and G_3 in (1) could be regarded as some special cases of the Schur complement $D - CA^-B$, we are able to use (11), (12), (13) and (14) to find their properties, as well as the relations among G_1, G_2, G_3 and A^-.

Some general results on the relations between a matrix X and A^-, AA^-, A^-A are given below.

Theorem 4 *Let* $A \in \mathbb{C}^{m\times n}$, $X_1 \in \mathbb{C}^{n\times m}$, $X_2 \in \mathbb{C}^{m\times m}$ *and* $X_3 \in \mathbb{C}^{n\times n}$. *Then*:

(a) $\min_{A^-} r(X_1 - A^-) = r(A - AX_1A)$.
(b$_1$) $\min_{A^-} r(X_2 - AA^-) = r(A - X_2A) + r[X_2,\ A] - r[X_2A,\ A]$.
(b$_2$) $X_2 \in \{AA^-\}$ *iff* $r(X_2) = r(A)$ *and* $X_2A = A$.
(c$_1$) $\min_{A^-} r(X_3 - A^-A) = r(A - AX_3) + r[X_3^*,\ A^*] - r[(AX_3)^*,\ A^*]$.
(c$_2$) $X_3 \in \{A^-A\}$ *iff* $r(X_3) = r(A)$ *and* $AX_3 = A$.

Proof Results (a), (b$_1$) and (c$_1$) are derived from (11) and (12). The details are omitted. Results (b$_2$) and (c$_2$) follow from (b$_1$) and (c$_1$). □

The following two theorems are derived from Lemma 3, see also Tian (2002a,b) for their proofs.

Theorem 5 *Let* $A,\ B \in \mathbb{C}^{m\times n}$. *Then*:

(a) $\min_{A^-,\,B^-} r(A^- - B^-) = r(A - B) - r[A,\ B] - r[A^*,\ B^*] + r(A) + r(B)$.
(b) $\max_{A^-} r(B - BA^-B) = \min\{r(B),\ \ r(A - B) - r(A) + r(B)\}$.
(c) *A and B have a common g-inverse iff* $r(A - B) = r[A,\ B] + r[A^*,\ B^*] - r(A) - r(B)$.
(d) $\{A^-\} \subseteq \{B^-\}$ *iff* $r(A - B) = r(A) - r(B)$.
(e) $\{A^-\} = \{B^-\}$ *iff* $A = B$.

Theorem 6 *Let* $A \in \mathbb{C}^{m\times n}$ *and* $B \in \mathbb{C}^{m\times p}$. *Then*:

(a) $\min_{A^-,\,B^-} r(AA^- - BB^-) = \max\{r[A,\ B] - r(A),\ \ r[A,\ B] - r(B)\}$.
(b) *There exist A^- and B^- such that $AA^- = BB^-$ iff $\mathcal{R}(A) = \mathcal{R}(B)$, in which case*, $\{AA^-\} = \{BB^-\}$.

More results on common g-inverses of two matrices can be found in Tian and Takane (2006). We have seen from the above results that the rank of matrix is quite useful for investigating the relations between two matrix expressions involving variable matrices. Applying the matrix rank method to the three matrices in (1) gives us the following results on the extremal ranks of G_1, G_2 and G_3.

Theorem 7 *Let* $A \in \mathbb{C}^{m\times n}$, $E \in \mathbb{C}^{n\times p}$ *and* $F \in \mathbb{C}^{q\times m}$. *Then*:

(a$_1$) $\max_{(AE)^-} r[E(AE)^-] = \min\{m,\ \ r(E)\}$.
(a$_2$) $\min_{(AE)^-} r[E(AE)^-] = r(AE)$.
(a$_3$) *The rank of $E(AE)^-$ is invariant iff $r(AE) = m$ or $r(AE) = r(E)$.*
(a$_4$) *$E(AE)^-$ is unique iff $E = 0$ or $r(AE) = r(E) = m$.*
(a$_5$) $\max_{(AE)^-} r[E(AE)^-A] = \min\{r(A),\ \ r(E)\}$.
(a$_6$) $\min_{(AE)^-} r[E(AE)^-A] = r(AE)$.
(a$_7$) *The rank of $E(AE)^-A$ is invariant iff $r(AE) = r(A)$ or $r(AE) = r(E)$.*
(a$_8$) *$E(AE)^-A$ is unique iff any one of* (i) $A = 0$, (ii) $E = 0$, (iii) $r(AE) = r(A) = r(E)$ *holds. In such a case*, $E(AE)^-A \in \{A^-A\}$ *and* $E(AE)^-A \in \{EE^-\}$.

(b$_1$) $\max\limits_{(FA)^-} r[(FA)^- F] = \min\{n,\ r(F)\}$.
(b$_2$) $\min\limits_{(FA)^-} r[(FA)^- F] = r(FA)$.
(b$_3$) *The rank of* $(FA)^- F$ *is invariant iff* $r(FA) = n$ *or* $r(FA) = r(F)$.
(b$_4$) $(FA)^- F$ *is unique iff* $F = 0$ *or* $r(FA) = r(F) = n$.
(b$_5$) $\max\limits_{(FA)^-} r[A(FA)^- F] = \min\{r(A),\ r(F)\}$.
(b$_6$) $\min\limits_{(FA)^-} r[A(FA)^- F] = r(FA)$.
(b$_7$) *The rank of* $A(FA)^- F$ *is invariant iff* $r(FA) = r(A)$ *or* $r(FA) = r(F)$.
(b$_8$) $A(FA)^- F$ *is unique iff any one of* (i) $A = 0$, (ii) $F = 0$, (iii) $r(FA) = r(A) = r(F)$ *holds. In such a case*, $A(FA)^- F \in \{AA^-\}$ *and* $A(FA)^- F \in \{F^- F\}$.
(c$_1$) $\max\limits_{(FAE)^-} r[E(FAE)^- F] = \min\{r(E),\ r(F)\}$.
(c$_2$) $\min\limits_{(FAE)^-} r[E(FAE)^- F] = r(FAE)$.
(c$_3$) *The rank of* $E(FAE)^- F$ *is invariant iff* $r(FAE) = r(E)$ *or* $r(FAE) = r(F)$.
(c$_4$) $E(FAE)^- F$ *is unique iff any one of* (i) $E = 0$, (ii) $F = 0$, (iii) $r(FAE) = r(E) = r(F)$ *holds*.

Proof The results related to the ranks of the matrix products are derived from (13) and (14). The uniqueness of the matrix products involving $E(AE)^-$, $(FA)^- F$ and $E(FAE)^- F$ is derived from the following simple result: $CA^- B$ is unique iff any one of (i) $B = 0$, (ii) $C = 0$, (iii) $\mathcal{R}(B) \subseteq \mathcal{R}(A)$ and $\mathcal{R}(C^*) \subseteq \mathcal{R}(A^*)$ holds. The details are omitted. □

The following theorem gives the extremal ranks of $A - AG_1A$, $A - AG_2A$ and $A - AG_3A$ and their consequences.

Theorem 8 *Let* $A \in \mathbb{C}^{m\times n}$, $E \in \mathbb{C}^{n\times p}$ *and* $F \in \mathbb{C}^{q\times m}$. *Then*:

(a$_1$) $r[A - AE(AE)^- A] = r(A) - r(AE)$ *for any* $(AE)^-$.
(a$_2$) $\min\limits_{A^-,\,(AE)^-} r[A^- A - E(AE)^- A] = \min\limits_{A^-,\,(AE)^-} r[AA^- - AE(AE)^-] = r(A) - r(AE)$.
(a$_3$) $\{E(AE)^-\} \subseteq \{A^-\} \Leftrightarrow$ there exist A^- *and* $(AE)^-$ *such that* $A^- A = E(AE)^- A \Leftrightarrow$ *there exist* A^- *and* $(AE)^-$ *such that* $AA^- = AE(AE)^- \Leftrightarrow r(AE) = r(A)$.
(b$_1$) $r[A - A(FA)^- FA] = r(A) - r(FA)$ *for any* $(FA)^-$.
(b$_2$) $\min\limits_{A^-,\,(FA)^-} r[AA^- - A(FA)^- F] = \min\limits_{A^-,\,(FA)^-} r[AA^- - (FA)^- FA] = r(A) - r(FA)$.
(b$_3$) $\{(FA)^- F\} \subseteq \{A^-\} \Leftrightarrow$ *there exist* A^- *and* $(FA)^-$ *such that* $AA^- = A(FA)^- F$ $\Leftrightarrow$ *there exist* A^- *and* $(FA)^-$ *such that* $A^- A = (FA)^- FA \Leftrightarrow r(FA) = r(A)$.
(c$_1$) $r[A - AE(FAE)^- FA] = r(A) - r(FAE)$ *for any* $(FAE)^-$.
(c$_2$) $\min\limits_{A^-,\,(FAE)^-} r[AA^- - AE(FAE)^- F] = \min\limits_{A^-,(FAE)^-} r[A^- A - E(FAE)^- FA]$ $= r(A) - r(FAE)$ *for any* A^- *and* $(FAE)^-$.
(c$_3$) $E(FAE)^- F \in \{A^-\} \Leftrightarrow$ *there exist* A^- *and* $(FAE)^-$ *such that* $AA^- = AE(FAE)^- F \Leftrightarrow$*there exist* A^- *and* $(FAE)^-$ *such that* $A^- A = E(FAE)^- FA$ $\Leftrightarrow r(FAE) = r(A)$. *In these cases*, $E(FAE)^- F$ *is unique*.

Proof From (4) and (5), we obtain

$$r[A - AE(AE)^{-}A] = r[A,\ AE] - r(AE) = r(A) - r(AE),$$
$$r[A - A(FA)^{-}FA] = r\begin{bmatrix} A \\ FA \end{bmatrix} - r(FA) = r(A) - r(FA).$$

Thus, we have (a_1) and (b_1). Results (a_2) and (b_2) follow from Theorem 4(b_1) and (c_1), and (a_1) and (b_1) of this theorem. Results (a_3) and (b_3) follow from (a_1), (a_2), (b_1), (b_2) and (c_1). Result (c_1) is derived from (11) and (12). Result (c_2) is derived from Theorem 4(b_1) and (c_1). Result (c_3) follows from (c_1) and (c_2). □

It is well known that the two products AA^{-} and $A^{-}A$ are idempotent for any A^{-}. However, the product $E(AE)^{-}A$ is necessarily idempotent for a pair of matrices A and E. From Theorem 7(a_7), we obtain the following result on the idempotency of $E(AE)^{-}A$.

Theorem 9 *Let $A \in \mathbb{C}^{m\times n}$ and $E \in \mathbb{C}^{n\times p}$. Then the following statements are equivalent*:

(a) *$E(AE)^{-}A$ is idempotent for any $(AE)^{-}$.*
(b) *$r(AE) = r(A)$ or $r(AE) = r(E)$, i.e., the rank of $E(AE)^{-}A$ is invariant.*

Proof Recall that if a square matrix M is idempotent, then $r(M) = \text{trace}(M)$. Hence $r(M) = \text{trace}(MM^{-}) = \text{trace}(M^{-}M)$. If $E(AE)^{-}A$ is idempotent for any $(AE)^{-}$, then

$$r[E(AE)^{-}A] = \text{trace}[E(AE)^{-}A] = \text{trace}[AE(AE)^{-}] = r(AE)$$

for any $(AE)^{-}$, which implies that the rank of $E(AE)^{-}$ is invariant. Conversely, if $r(AE) = r(A)$ or $r(AE) = r(E)$, then $\mathcal{R}(AE) = \mathcal{R}(A)$ or $\mathcal{R}[(AE)^{*}] = \mathcal{R}(E^{*})$. Hence $(AE)(AE)^{-}A = A$ or $E(AE)^{-}(AE) = E$, under which, $E(AE)^{-}A$ is idempotent. □

Replacing E and F in (1) with $E = WA^{*}$ and $F = A^{*}V$, where both W and V are square matrices, gives

$$G_4 = E(AE)^{-} = WA^{*}(AWA^{*})^{-}, \quad G_5 = (FA)^{-}F = (A^{*}VA)^{-}A^{*}V, \quad (22)$$

$$G_6 = E(FAE)^{-}F = WA^{*}(A^{*}VAWA^{*})^{-}A^{*}V. \quad (23)$$

When V is nnd, $A(A^{*}VA)^{-}A^{*}V$ is called the V-orthogonal projector in the literature. Yanai (1990) investigated the matrices $WA^{*}(AWA^{*})^{-}A$ and $A(A^{*}VA)^{-}A^{*}V$, as well as their relations with A^{-}. In particular, we find the following consequences if V and W are nnd from Theorems 7–9.

Corollary 10 *Let $A \in \mathbb{C}^{m\times n}$, and let $W \in \mathbb{C}^{n\times n}$ and $V \in \mathbb{C}^{m\times m}$ be nnd. Then*:

(a_1) $r[WA^{*}(AWA^{*})^{-}] = r[WA^{*}(AWA^{*})^{-}A] = r(AW)$ *for any* $(AWA^{*})^{-}$.
(a_2) *$WA^{*}(AWA^{*})^{-}A$ is unique iff $AW = 0$ or $r(AW) = r(A)$, in the latter case*,

$$(WA^{*})(AWA^{*})^{-}A \in \{A^{-}A\}, \quad WA^{*}(AWA^{*})^{-}A \in \{(WA^{*})(WA^{*})^{-}\}.$$

(b$_1$) $r[(A^*VA)^-A^*V] = r[A(A^*VA)^-A^*V] = r(VA)$ *for* $(A^*VA)^-$.
(b$_2$) $A(A^*VA)^-A^*V$ *is unique iff* $VA = 0$ *or* $r(A^*VA) = r(A^*V)$, *in the latter cases*,

$$A(A^*VA)^-A^*V \in \{AA^-\}, \quad A(A^*VA)^-A^*V \in \{(A^*V)^-(A^*V)\}.$$

Corollary 11 *Let* $A \in \mathbb{C}^{m\times n}$, *and let* $W \in \mathbb{C}^{n\times n}$ *and* $V \in \mathbb{C}^{m\times m}$ *be nnd. Then*:

(a$_1$) $WA^*(AWA^*)^-A$ *is idempotent for any* $(AWA^*)^-$.
(a$_2$) $r[A - AWA^*(AWA^*)^-A] = r(A) - r(AW)$ *for any* $(AWA^*)^-$.
(a$_3$) $\{WA^*(AWA^*)^-\} \subseteq \{A^-\} \Leftrightarrow r(AW) = r(A)$.
(b$_1$) $A(A^*VA)^-A^*V$ *is idempotent for any* $(A^*VA)^-$.
(b$_2$) $r[A - A(A^*VA)^-A^*VA] = r(A) - r(VA)$ *for any* $(A^*VA)^-$.
(b$_3$) $\{(A^*VA)^-A^*V\} \subseteq \{A^-\} \Leftrightarrow r(VA) = r(A)$.

3 Some applications

Many consequences and applications can be derived from the previous theorems. In this section, we first give some results on g-inverses of block matrices.

Theorem 12 *Let* $A \in \mathbb{C}^{m\times n}$, $B \in \mathbb{C}^{n\times k}$ *and* $C \in \mathbb{C}^{l\times n}$, *and let* $M = \begin{bmatrix} A & B \\ C & 0 \end{bmatrix}$. *Then*

$$r[A - A(P_BA)^-P_BA] = r[B - B(P_AB)^-P_AB] = r(A) + r(B) - r[A,\ B] \tag{24}$$

$$r[A - AQ_C(AQ_C)^-A] = r[C - CQ_A(CQ_A)^-C] = r(A) + r(C) - r[A^*,\ C^*] \tag{25}$$

$$r[A - AQ_C(P_BAQ_C)^-P_BA] = r(A) + r(B) + r(C) - r(M) \tag{26}$$

hold for any $(P_AB)^-$, $(P_BA)^-$, $(AQ_C)^-$, $(CQ_A)^-$ *and* $(P_BAQ_C)^-$. *Hence*,

(i) *The following statements are equivalent*:
  (a) $\{(P_BA)^-P_B\} \subseteq \{A^-\}$.
  (b) $\{(P_AB)^-P_A\} \subseteq \{B^-\}$.
  (c) $\left\{\begin{bmatrix} (P_BA)^-P_B \\ (P_AB)^-P_A \end{bmatrix}\right\} \subseteq \{[A,\ B]^-\}$.
  (d) $\mathcal{R}(A) \cap \mathcal{R}(B) = \{0\}$.
(ii) *The following statements are equivalent*:
  (a) $\{Q_C(AQ_C)^-\} \subseteq \{A^-\}$.
  (b) $\{Q_A(CQ_A)^-\} \subseteq \{C^-\}$.
  (c) $\{[Q_C(AQ_C)^-,\ Q_A(CQ_A)^-]\} \subseteq \left\{\begin{bmatrix} A \\ C \end{bmatrix}^-\right\}$.
  (d) $\mathcal{R}(A^*) \cap \mathcal{R}(C^*) = \{0\}$.
(iii) *The following statements are equivalent*:
  (a) $\{Q_C(P_BAQ_C)^-P_B\} \subseteq \{A^-\}$.

(b) $\left\{\begin{bmatrix} Q_C(P_BAQ_C)^-P_B & Q_A(CQ_A)^- \\ (P_AB)^-P_A & 0 \end{bmatrix}\right\} \subseteq \left\{\begin{bmatrix} A & B \\ C & 0 \end{bmatrix}^-\right\}$.

(c) $\mathcal{R}(A) \cap \mathcal{R}(B) = \{0\}$ and $\mathcal{R}(A^*) \cap \mathcal{R}(C^*) = \{0\}$.

Proof Equation (24) follows from Theorem 8(a_1) and (4); (25) follows from Theorem 8(b_1) and (5); (26) follows from Theorem 8(c_1) and (6). It is easy to verify that

$$\begin{aligned}
&[A,\, B] - [A,\, B]\begin{bmatrix} (P_BA)^-P_B \\ (P_AB)^-P_A \end{bmatrix}[A,\, B] \\
&= [A - A(P_BA)^-P_BA,\, B - B(P_AB)^-P_AB].
\end{aligned}$$

Hence

$$\begin{aligned}
[A,B]\begin{bmatrix} (P_BA)^-P_B \\ (P_AB)^-P_A \end{bmatrix}[A,B] = [A,B] \Leftrightarrow\ & A(P_BA)^-P_BA = A, \\
\text{and } & B(P_AB)^-P_AB = B.
\end{aligned}$$

The equivalences in (i) follow from this equivalence and (24). The equivalences in (ii) and (iii) can be shown similarly. □

Note from Sect. 1 that the two matrices $B^*(A^*ABB^*)^-A^*$ and $B^\dagger(A^\dagger ABB^\dagger)^-A^\dagger$ are special cases of G_3 in (1), and the set inclusions

$$\{B^*(A^*ABB^*)^-A^*\} \subseteq \{(AB)^-\} \ \text{ and } \ \{B^\dagger(A^\dagger ABB^\dagger)^-A^\dagger\} \subseteq \{(AB)^-\}$$

hold. In these cases, it is of interest to find the relation between the two matrix sets $\{B^*(A^*ABB^*)^-A^*\}$ and $\{B^\dagger(A^\dagger ABB^\dagger)^-A^\dagger\}$. A further investigation yields the following result.

Theorem 13 *Let $A \in \mathbb{C}^{m\times n}$ and $B \in \mathbb{C}^{m\times m}$ be given. Then the following set equality holds*

$$\{B^*(A^*ABB^*)^-A^*\} = \{B^\dagger(A^\dagger ABB^\dagger)^-A^\dagger\}. \tag{27}$$

Proof Equality (27) is obviously equivalent to

$$\min_{(A^\dagger ABB^\dagger)^-} r[B^*(A^*ABB^*)^-A^* - B^\dagger(A^\dagger ABB^\dagger)^-A^\dagger] = 0, \tag{28}$$

$$\min_{(A^*ABB^*)^-} r[B^*(A^*ABB^*)^-A^* - B^\dagger(A^\dagger ABB^\dagger)^-A^\dagger] = 0. \tag{29}$$

By (12), $\mathcal{R}(M^{\dagger}) = \mathcal{R}(M^{*})$ for any M and elementary block matrix operations

$$\begin{aligned}
&\min_{(A^{\dagger}ABB^{\dagger})^{-}} r[B^{*}(A^{*}ABB^{*})^{-}A^{*} - B^{\dagger}(A^{\dagger}ABB^{\dagger})^{-}A^{\dagger}] \\
&= r(A^{\dagger}ABB^{\dagger}) + r[B^{\dagger},\ B^{*}(A^{*}ABB^{*})^{-}A^{*}] + r\begin{bmatrix} A^{\dagger} \\ B^{*}(A^{*}ABB^{*})^{-}A^{*} \end{bmatrix} \\
&\quad + r\begin{bmatrix} A^{\dagger}ABB^{\dagger} & A^{\dagger} \\ B^{\dagger} & B^{*}(A^{*}ABB^{*})^{-}A^{*} \end{bmatrix} - r\begin{bmatrix} A^{\dagger}ABB^{\dagger} & 0 & A^{\dagger} \\ 0 & B^{\dagger} & B^{*}(A^{*}ABB^{*})^{-}A^{*} \end{bmatrix} \\
&\quad - r\begin{bmatrix} A^{\dagger}ABB^{\dagger} & 0 \\ 0 & A^{\dagger} \\ B^{\dagger} & B^{*}(A^{*}ABB^{*})^{-}A^{*} \end{bmatrix} \\
&= r(AB) + r(B) + r(A) + r\begin{bmatrix} 0 & A^{\dagger} - A^{\dagger}ABB^{*}(A^{*}ABB^{*})^{-}A^{*} \\ B^{\dagger} & 0 \end{bmatrix} - 2r(A) - 2r(B) \\
&= r(AB) + r[A^{*} - A^{*}ABB^{*}(A^{*}ABB^{*})^{-}A^{*}] - r(A) \\
&= r(AB) + r[A^{*},\ A^{*}ABB^{*}] - r(A^{*}ABB^{*}) - r(A) \ \text{ by (4)} \\
&= 0,
\end{aligned}$$

establishing (28). Equation (29) can be shown similarly. □

In the remaining of the section, we give some results on the WLSE of the parametric function $X\beta$ in the general linear (Gauss–Markov) model

$$y = X\beta + \varepsilon, \tag{30}$$

where $X \in \mathbb{R}^{n\times p}$ is a known matrix, $y \in \mathbb{R}^{n\times 1}$ is an observable random vector with the expectation vector $\mathrm{E}(y) = X\beta$ and the covariance matrix $\mathrm{Cov}(y) = \Sigma$, $\beta \in \mathbb{R}^{p\times 1}$ is a vector of unknown parameters, and $\Sigma \in \mathbb{R}^{n\times n}$ is an nnd matrix, known entirely except for a positive constant multiplier. The general expression of the WLSE of the parametric function $X\beta$ with respect to an nnd $V \in \mathbb{R}^{n\times n}$ is given by

$$\mathrm{WLSE}(X\beta) = P_{X:V}\, y, \tag{31}$$

where

$$P_{X:V} = X(X'VX)^{-}X'V + [X - X(X'VX)^{-}(X'VX)]U \tag{32}$$

with $U \in \mathbb{R}^{p\times n}$ arbitrary. The expectation and the covariance matrix of $\mathrm{WLSE}(X\beta)$ are given by

$$\mathrm{E}[\mathrm{WLSE}(X\beta)] = P_{X:V}X\beta \ \text{ and } \ \mathrm{Cov}[\mathrm{WLSE}(X\beta)] = \sigma^{2}P_{X:V}\Sigma P'_{X:V}. \tag{33}$$

Because the equality $P_{X:V}X = X$ does not necessarily hold for a given U in (32), $\mathrm{WLSE}(X\beta)$ in (31) is not unbiased for $X\beta$ in (30) for a given projector $P_{X:V}$.

Theorem 14 *Let the general linear model be as given in* (30), *and* $P_{X:V}$ *be as given in* (32). *Then*

$$\max\ r(P_{X:V}) = r(X), \tag{34}$$

$$\min\ r(P_{X:V}) = r(VX), \tag{35}$$

$$\max\ r\{\mathrm{Cov}[\mathrm{WLSE}(X\beta)]\} = \min\{\, r(\Sigma),\ \ r(X) + r(\Sigma VX) - r(VX)\,\}, \tag{36}$$

$$\min\ r\{\mathrm{Cov}[\mathrm{WLSE}(X\beta)]\} = r(\Sigma VX). \tag{37}$$

Hence,

(a) *The following statements are equivalent*:
  (i) *The rank of* $P_{X:V}$ *is invariant*.
  (ii) $P_{X:V}$ *is unique*.
  (iii) $r(VX) = r(X)$, *i.e.*, $\mathcal{R}(X'V) = \mathcal{R}(X')$.
  In these cases, the $\mathrm{WLSE}(X\beta)$ *can uniquely be written as*

$$\mathrm{WLSE}(X\beta) = X(X'VX)^{\dagger}X'Vy$$

  with $\mathrm{E}[\mathrm{WLSE}(X\beta)] = X\beta$ *and*

$$\mathrm{Cov}[\mathrm{WLSE}(X\beta)] = \sigma^2 X(X'VX)^{\dagger}X'V\Sigma VX(X'VX)^{\dagger}X'.$$

(b) *The following statements are equivalent*:
  (i) *The rank of* $\mathrm{Cov}[\mathrm{WLSE}(X\beta)]$ *is invariant*.
  (ii) $r(\Sigma VX) = r(\Sigma)$ *or* $r(VX) = r(X)$.

Proof Applying (7), (8) and elementary block matrix operations to $P_{X:V}$ gives

$$\begin{aligned}\max_{P_{X:V}}\ r(P_{X:V}) &= \max_{U} r\{X(X'VX)^{-}X'V - [X - X(X'VX)^{-}(X'VX)]U\}\\ &= r[X(X'VX)^{-}X'V,\ X - X(X'VX)^{-}(X'VX)],\end{aligned}$$

and

$$\begin{aligned}&\min_{P_{X:V}}\ r(P_{X:V})\\ &\quad= \min_{U}\ r\{X(X'VX)^{-}X'V - [X - X(X'VX)^{-}(X'VX)]U\}\\ &\quad= r[X(X'VX)^{-}X'V, X - X(X'VX)^{-}(X'VX)] - r[X - X(X'VX)^{-}(X'VX)].\end{aligned}$$

By (5), the ranks of the two matrix expressions reduce to

$$\begin{aligned}&r[X(X'VX)^{-}X'V,\ X - X(X'VX)^{-}(X'VX)]\\ &\quad= r\begin{bmatrix} X(X'VX)^{-}X'V & X\\ 0 & X'VX\end{bmatrix} - r(X'VX)\\ &\quad= r\begin{bmatrix} 0 & X\\ X'VX(X'VX)^{-}X'V & 0\end{bmatrix} - r(VX)\\ &\quad= r(X) + r[X'VX(X'VX)^{-}X'V] - r(VX) = r(X),\end{aligned}$$

and

$$r[X - X(X'VX)^{-}(X'VX)] = r\begin{bmatrix} X \\ X'VX \end{bmatrix} - r(X'VX) = r(X) - r(VX).$$

Hence we have (34) and (35). Because Σ is nnd, we obtain from (33) that

$$r\{\mathrm{Cov}[\mathrm{WLSE}(X\beta)]\} = r(\sigma^2 P_{X:V}\Sigma P'_{X:V}) = r(P_{X:V}\Sigma).$$

Applying (9), (10) and elementary block matrix operations to $P_{X:V}\Sigma$ gives

$$\begin{aligned} &\max_{P_{X:V}} r\{\mathrm{Cov}[\mathrm{WLSE}(X\beta)]\} \\ &= \max_{U} r\{X(X'VX)^{-}X'V\Sigma - [X - X(X'VX)^{-}(X'VX)]U\Sigma\} \\ &= \min\left\{ r[X(X'VX)^{-}X'V\Sigma, X{-}X(X'VX)^{-}(X'VX)],\ r\begin{bmatrix} X(X'VX)^{-}X'V\Sigma \\ \Sigma \end{bmatrix} \right\}, \end{aligned} \tag{38}$$

and

$$\begin{aligned} &\min_{P_{X:V}} r\{\mathrm{Cov}[\mathrm{WLSE}(X\beta)]\} \\ &= \min_{U} r\{X(X'VX)^{-}X'V\Sigma - [X - X(X'VX)^{-}(X'VX)]U\Sigma\} \\ &= r[X(X'VX)^{-}X'V\Sigma,\ X - X(X'VX)^{-}(X'VX)] + r\begin{bmatrix} X(X'VX)^{-}X'V\Sigma \\ \Sigma \end{bmatrix} \\ &\quad - r\begin{bmatrix} X(X'VX)^{-}X'V\Sigma & X - X(X'VX)^{-}(X'VX) \\ \Sigma & 0 \end{bmatrix}. \end{aligned} \tag{39}$$

By (4) and (5),

$$\begin{aligned} r[X(X'VX)^{-}X'V\Sigma,\ X - X(X'VX)^{-}(X'VX)] &= r(X) + r(\Sigma VX) - r(VX), \\ r\begin{bmatrix} X(X'VX)^{-}X'V\Sigma \\ \Sigma \end{bmatrix} &= r(\Sigma), \\ r\begin{bmatrix} X(X'VX)^{-}X'V\Sigma & X - X(X'VX)^{-}(X'VX) \\ \Sigma & 0 \end{bmatrix} &= r(\Sigma) + r(X) - r(VX). \end{aligned}$$

Substituting these three rank equalities into (38) and (39) gives (36) and (37). The equivalence of (i), (ii) and (iii) in (a) follows from equating (34) and (35). The equivalence of (i) and (ii) in (b) follows from equating (36) and (37). □

It can be seen from (32) and (33) that if $P_{X:V}$ is unique, then the product $P_{X:V}\Sigma P'_{X:V}$ in Cov[WLSE($X\beta$)] is unique, too. However, the contrary is not true. An open problem is to give necessary and sufficient conditions for $P_{X:V}\Sigma P'_{X:V}$ to be unique.

Acknowledgment We are grateful to an anonymous referee for his/her helpful comments and suggestions to an earlier version of this paper.

References

Amemiya, T. (1985). *Advanced econometrics*. Cambridge: Harvard University Press.
Ben-Israel, A. Greville, T. N. E. (2003). *Generalized inverses: theory and applications*, 2nd Ed. Berlin Heidelberg, New York: Springer .
Baksalary, J. K., Puntanen, S. (1989). Weighted-least-squares estimation in the general Gauss–Markov model. In Y. Dodge (Ed.), *Statistical data analysis and inference* (pp. 355–368). Amsterdam: Elsevier.
Campbell, S. L., Meyer, C. D. (1991). *Generalized inverses of linear transformations*. Corrected reprint of the 1979 original, Dover Publications, Inc., New York.
Davidson, R., MacKinnon, J. G. (2004). *Econometric theory and methods*. New York: Oxford University Press.
Marsaglia, G., Styan, G. P. H. (1974). Equalities and inequalities for ranks of matrices. *Linear and Multilinear Algebra*, *2*, 269–292.
Mitra, S. K. (1968). On a generalized inverse of a matrix and applications. *Sankhyā Ser A*, *30*, 107–114.
Mitra, S. K., Rao, C. R. (1974). Projections under seminorms and generalized Moore-Penrose inverses. *Linear Algebra and Its Applications*, *9*, 155–167.
Rao, C. R., Mitra, S. K. (1971). *Generalized inverse of matrices and its applications*. New York: Wiley.
Takane, Y., Yanai, H. (1999). On oblique projectors. *Linear Algebra and Its Applications*, *289*, 297–310.
Tian, Y. (2002a). Upper and lower bounds for ranks of matrix expressions using generalized inverses. *Linear Algebra and Its Applications*, *355*, 187–214.
Tian, Y. (2002b). The maximal and minimal ranks of some expressions of generalized inverses of matrices. *Southeast Asian Bulletin of Mathematics*, *25*, 745–755.
Tian, Y. (2004). More on maximal and minimal ranks of Schur complements with applications. *Applied Mathematics and Computation*, *152*, 675–692.
Tian, Y., Cheng, S. (2003). The maximal and minimal ranks of $A - BXC$ with applications. *New York Journal of Mathematics*, *9*, 345–362.
Tian, Y., Takane, Y. (2006). On common generalized inverses of a pair of matrices. *Linear and Multilinear Algebra*, *54*, 195–209.
Yanai, H. (1990). Some generalized forms of a least squares g-inverse, minimum norm g-inverse, and Moore-Penrose inverse matrices. *Computational Statistics and Data Analysis*, *10*, 251–260.