A CANCELLATION PROPERTY OF THE MOORE–PENROSE INVERSE OF TRIPLE PRODUCTS

TOBIAS DAMM and HARALD K. WIMMER

(Received 9 February 2007; accepted 17 October 2007)

Communicated by J. J. Koliha

Abstract

We study the matrix equation $C(BXC)\dagger B = X\dagger$, where $X\dagger$ denotes the Moore–Penrose inverse. We derive conditions for the consistency of the equation and express all its solutions using singular vectors of $B$ and $C$. Applications to compliance matrices in molecular dynamics, to mixed reverse-order laws of generalized inverses and to weighted Moore–Penrose inverses are given.

Keywords and phrases: Moore–Penrose generalized inverse, compliance matrix, reverse-order property, weighted generalized inverse, Wedderburn–Guttman theorem, matrix equations.

1. Introduction

Let $B$, $C$, $X$, be complex matrices of size $s \times n$, $m \times t$, $n \times m$, respectively, and let $X\dagger$ denote the Moore–Penrose generalized inverse of $X$. The purpose of this paper is to characterize all triples $(B, C, X)$ which satisfy

$$C(BXC)\dagger B = X\dagger.$$  \hfill (1.1)

We say that $(B, C, X)$ has the cancellation property if (1.1) holds. If $B$, $C$, $X$ are nonsingular $n \times n$ matrices then it is obvious that (1.1) holds. In that case we have $C(BXC)^{-1}B = X^{-1}$.

Our investigation is motivated by recent applications of compliance matrices in molecular dynamics (for instance, [2, 4–6, 15]). According to [15] the compliance matrix can be defined as the inverse of a force-constant matrix. While the force-constant matrix describes the forces between different parts of a molecule (acting in different directions), the compliance matrix shows how the molecule complies with certain external forces acting on it. In particular, centrifugal distortion constants
and high-temperature mean-square amplitudes depend directly on compliances rather than on force constants. Very often it is advantageous to model a molecule in a redundant coordinate system, which means using more variables than there are degrees of freedom in the molecule. This is, for instance, the case if all bond lengths and interbond angles are taken as coordinates. In a redundant coordinate system the compliance matrix \( N_r \) is defined as the Moore–Penrose inverse of the symmetric force-constant matrix \( F_r \) in redundant coordinates. Thus, if \( F_s \) is the force-constant matrix in a nonredundant coordinate system, then \( F_r \) is related to \( F_s \) via
\[
J^\dagger F_r (J^\dagger)^T = F_s,
\]
where \( J \) is a linearized coordinate transformation. In general \( F_r \) is much larger than \( F_s \), and the question arises whether \( N_r = F_r^\dagger \) can be obtained from \( N_s = F_s^\dagger \). This is exactly the case if the triple \((J, J^T, F_r)\) possesses the cancellation property.

The main result of our paper is Theorem 3.4 in Section 3 with necessary and sufficient conditions for the consistency of (1.1). We shall prove that (1.1) holds if and only if
\[
\text{Im } B^* B X = \text{Im } X \quad \text{and} \quad \text{Ker } X C C^* = \text{Ker } X. \tag{1.2}
\]
In Section 4 we study (1.1) as a matrix equation. If \( X \) is a solution of (1.1) then (1.2) implies that \( \text{Im } X \) and \( \text{Im } X^* \) are invariant under \( B^* B \) and \( C C^* \), respectively. This observation will be used to construct all solutions of (1.1). In Section 5 we consider topics which involve products of the form \( C (B X C) \dagger B \). In particular, we reexamine the issue of compliance matrices and apply Theorem 3.4 to mixed-type reverse-order laws and to weighted generalized inverses.

### 2. Notation, basic facts, auxiliary results

Let us first summarize the main issues related to the definition of the Moore–Penrose inverse. Consider a matrix \( A \in \mathbb{C}^{n \times m} \) and the corresponding linear mapping \( A : \mathbb{C}^m \to \mathbb{C}^n \). Let \( \text{Ker } A \) and \( \text{Im } A \) denote the kernel and the image of \( A \), respectively. The restriction \( A|_{\text{Im } A^*} : \text{Im } A^* \to \text{Im } A \) is invertible. Then \( A^\dagger \) is defined by \( A^\dagger x = (A|_{\text{Im } A^*})^{-1} x \) if \( x \in \text{Im } A \), and \( A^\dagger x = 0 \) if \( x \in (\text{Im } A)^\perp = \text{Ker } A^* \). This functional definition (see [3, p. 8]) can be illustrated in a diagram:

\[
\begin{align*}
\mathbb{C}^m &\quad (\text{Im } A^* = \text{Im } A^\dagger) \oplus \quad (\text{Im } A^*)^\perp = \text{Ker } A \\
\downarrow A &\quad \uparrow A^\dagger \quad A = 0 \quad \uparrow A^\dagger = 0 \\
\mathbb{C}^n &\quad \text{Im } A \oplus (\text{Im } A)^\perp = \text{Ker } A^* = \text{Ker } A^\dagger.
\end{align*}
\]

It follows that
\[
P_A = A A^\dagger : \mathbb{C}^n \to \mathbb{C}^n \quad \text{and} \quad P_{A^*} = A^\dagger A : \mathbb{C}^m \to \mathbb{C}^m
\]
are the orthogonal projections on \( \text{Im } A \) and \( \text{Im } A^* \), respectively. These properties characterize \( A^\dagger \) uniquely, so that \( W = A^\dagger \) is the unique solution of the two Moore equations [3, p. 9; 1, p. 370]
\[
A W = P_A \quad \text{and} \quad W A = P_{A^*}. \tag{2.1}
\]
It is also clear that $W = A^\dagger$ satisfies the four Penrose equations \[3, \text{p. 9}; \ 1, \text{p. 40}\]

\begin{align*}
AWA &= A \quad (1) \\
WAW &= A \quad (2) \\
(AW)^* &= AW \quad (3) \\
(WA)^* &= WA, \quad (4)
\end{align*}

and in fact these equations determine $A^\dagger$ uniquely. The sets of conditions (2.1) and (2.2) are equivalent such that $A^\dagger$ is rightly named the Moore–Penrose inverse of $A$. We shall exploit the equivalence of the three definitions where it is convenient.

Sometimes we will only consider a subset of the Penrose conditions (2.2). In accordance with \[1, \text{p. 40}\], let $A\{i, j, \ldots, p\}$ denote the set of matrices $W = A\{i, j, \ldots, p\} \in \mathbb{C}^{n \times m}$ which satisfy equations $(i), (j), \ldots, (p)$ from (2.2). Thus $\{A^\dagger\} = A\{1, 2, 3, 4\}$.

The following lemma gathers together some auxiliary results on kernel and image inclusions, matrix products and Moore–Penrose inverses.

\textbf{Lemme 2.1.} Let $X \in \mathbb{C}^{m \times n}$, $B \in \mathbb{C}^{k \times m}$ and $C \in \mathbb{C}^{n \times p}$.

(i) $\text{Im } X \subseteq \text{Im } B^*$ $\iff$ $X = B^\dagger BX$ $\iff$ $X^\dagger = X^\dagger B^\dagger B$.

(ii) $\text{Ker } C^* \subseteq \text{Ker } X$ $\iff$ $X = XCC^\dagger$ $\iff$ $X^\dagger = CC^\dagger X^\dagger$.

(iii) If $X = B^\dagger BX = XCC^\dagger$ then

\[X[C(BXC)^\dagger B] = B^\dagger P_{BX} B \quad \text{and} \quad [C(BXC)^\dagger B]X = CP_{(XC)^*}C^\dagger.\]

\textbf{Proof.}

(i) The assertion follows from $B^\dagger B = P_{B^*}$ and $(X^\dagger)^* = (X^*)^\dagger$, together with $\text{Im } X^\dagger = \text{Im } X^*$.

(ii) It suffices to note that Ker $C^* \subseteq \text{Ker } X$ is equivalent to $\text{Im } X^* \subseteq \text{Im } C$.

(iii) Note that $X = XCC^\dagger$ implies $\text{Im } BXC = \text{Im } BX$. Hence we have $P_{BXC} = P_{BX}$ and

\[X[C(BXC)^\dagger B] = B^\dagger [BXC(BXC)^\dagger]B = B^\dagger P_{BX} B = B^\dagger P_{BX} B. \quad \Box\]

\section{3. Main results}

In this section we characterize those triples $(B, C, X)$ which possess property (1.1). In particular, we aim for criteria which do not involve pseudoinverses and describe the cancellation property in terms of image and kernel inclusions. Our first criterion, presented in the following lemma, is rather technical and serves as an intermediate step in the derivation of the main Theorem 3.4.

\textbf{Lemma 3.1.} The following statements are equivalent.

(i) We have

\[C(BXC)^\dagger B = X^\dagger. \quad (3.1)\]

(ii) The matrices $K = B^\dagger P_{BX} B$ and $L = CP_{XC^*}C^\dagger$ are Hermitian, and

\[X = B^\dagger BX \quad \text{and} \quad X = XCC^\dagger. \quad (3.2)\]
Consider the special cases of Lemma 2.1

Put \( W = C(BXC)^\dagger B \).

We show that (i) implies (ii). From (3.1) it follows that \( X^\dagger = X^\dagger B^\dagger B = C C^\dagger X^\dagger \).
By Lemma 2.1(i) the preceding identity is equivalent to (3.2). Then Lemma 2.1(iii) implies that \( XW = K \) and \( WX = L \). By \( W = X^\dagger \) we have \( K = K^* \) and \( L = L^* \).

We show that (ii) implies (i). Using (3.2) we obtain \( W \in X \{1\} \) from

\[
X W X = [B^\dagger B X C C^\dagger] C (B X C)^\dagger B [B^\dagger B X C C^\dagger] = B^\dagger [(B X C) (B X C)^\dagger (B X C)] C^\dagger = B^\dagger B X C C^\dagger = X.
\]

The identity \( W X W = W \) is obvious. Hence \( W \in X \{1\} \). We know that (3.2) implies both \( K = X W \) and \( L = W X \). Since \( K \) and \( L \) are Hermitian, \( W \in X \{3, 4\} \). Therefore \( W = X^\dagger \), which is (3.1).

We remark that \( W \in X \{1\} \) can be deduced from a more general result. Note that (3.2) implies that rank \( X = \text{rank } B X = \text{rank } XC \). According to [9] the preceding rank condition holds if and only if \( C (B X C)^\dagger B \in X \{1\} \) for each \( (B X C)^\dagger B \in (B X C)^\dagger \{1\} \).

The following example shows that, in general, condition (3.2) on its own is not sufficiently strong to imply (3.1).

**Example 1.** Take

\[
B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \quad C = B^* \quad \text{and} \quad X = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.
\]

Then (3.2) holds and \( X^\dagger = X \). For

\[
B X C = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix},
\]
we find \( (B X C)^\dagger = \frac{1}{4} B X C \) and thus

\[
W = C (B X C)^\dagger B = \frac{1}{4} \begin{pmatrix} 4 & 2 \\ 2 & 1 \end{pmatrix} \neq X^\dagger.
\]

In fact we have \( W \in X \{1\} \), but \( W \) does not satisfy any of the conditions (2)–(4) in (2.2).

**Corollary 3.2.** \( X^\dagger = C(B X C)^\dagger B \) if and only if both

\[
X^\dagger = (B X)^\dagger B \quad \text{and} \quad X^\dagger = C (X C)^\dagger \tag{3.3}
\]

hold.

**Proof.** Consider the special cases of Lemma 3.1 with \( C = I \) or \( B = I \). Then \( X^\dagger = (B X)^\dagger B \) if and only if \( B^\dagger P_B X B \) is Hermitian and \( X = B^\dagger B X \). Similarly, \( X^\dagger = C (X C)^\dagger \) is valid if and only if \( C P_{(X C)^*} C^\dagger \) is Hermitian and \( X^\dagger = C (X C)^\dagger \). \( \square \)
The following statements are equivalent.

(i) \((BX)^\dagger B = X^\dagger\).

(ii) \(\text{Im } B^*B X = \text{Im } X\).

(iii) \(\text{Im } B^*B X \subseteq \text{Im } X \subseteq \text{Im } B^*\).

(iv) \((BX)^\dagger = X^\dagger B^\dagger \text{ and } X = B^\dagger B X\).

(v) \(\text{Im } X = \text{Im } (B^*B)^\dagger X\).

**Proof.** We prove that (i) implies (ii), that (ii) implies (iv), and that (iv) implies (i), and also that (ii) implies (iii), that (iii) implies (v), and that (v) implies (ii).

We show that (i) implies (ii). From \(\text{Ker } X^\dagger = [\text{Im } (X^\dagger)^*]^\perp = (\text{Im } X)^\perp\) and

\[\text{Ker } (BX)^\dagger B = [\text{Im } B^*((BX)^\dagger)^*]^\perp = [B^*\text{Im } ((BX)^\dagger)^*]^\perp = [\text{Im } B^*B X]^\perp,\]

we obtain \(\text{Im } B^*B X = \text{Im } X\).

We show that (ii) implies (iv). Clearly \(\text{Im } B^*B X = \text{Im } X\) implies \(\text{Im } B^* \supseteq \text{Im } X\), and thus \(X = B^\dagger B X\). We derive the reverse-order identity \((BX)^\dagger = X^\dagger B^\dagger\) from \(\text{Im } B^*B X = \text{Im } X\) as follows. We first prove that \(\text{Im } (BX)^\dagger = \text{Im } (X^\dagger B^\dagger)\) and \(\text{Ker } (BX)^\dagger = \text{Ker } (X^\dagger B^\dagger)\) and then show that \((BX)^\dagger z = X^\dagger B^\dagger z\) for all \(z \in \text{Im } B X\). We observe that

\[\text{Im } (BX)^\dagger = \text{Im } (BX)^* = \text{Im } X^*(B^*B X) = \text{Im } X^*X = \text{Im } X^*\]

and

\[\text{Im } X^\dagger B^\dagger = \text{Im } X^\dagger B^* = \text{Im } X^\dagger B^*B(X^\dagger)^*\]

\[= \text{Im } X^\dagger (B^*B X) = \text{Im } X^\dagger X = \text{Im } P_X X = \text{Im } X^*\].

From \(\text{Ker } (BX)^\dagger = \text{Ker } (BX)^*\) it follows that \((\text{Ker } (BX)^\dagger)^\perp = \text{Im } B X\). On the other hand,

\[\left(\text{Ker } X^\dagger B^\dagger\right)^\perp = \left(\text{Ker } X^\dagger B^\dagger\right)^\perp\]

\[= \text{Im } (B^\dagger)^*X = \text{Im } (B^\dagger)^*B^*B X = \text{Im } (BB^\dagger)^*B X\]

\[= P_B \text{Im } B X = \text{Im } B X.\]

If \(z \in \text{Im } B X\) then there is a unique \(u \in \text{Im } (BX)^* = \text{Im } X^*\) such that \(z = BX u\). Hence \((BX)^\dagger z = u\) and

\[X^\dagger B^\dagger z = X^\dagger (B^\dagger B X) u = X^\dagger X u = P_X u = u.\]

We show that (iv) implies (i). According to Lemma 2.1(i) the identity \(B^\dagger B X = X\) is equivalent to \(X^\dagger B^\dagger B = X^\dagger\). Hence we obtain \(X^\dagger B^\dagger B = (BX)^\dagger B = X^\dagger\).

It is obvious that (ii) implies (iii).

We show that (iii) implies (v). Set \(\beta = B^*B\mid_{\text{Im } B^*}\). Then \(\beta : \text{Im } B^* \to \text{Im } B^*\) is invertible and \(\beta^{-1} = (B^*B)^\dagger\mid_{\text{Im } B^*}\). Since \(\text{Im } X \subseteq \text{Im } B^*\) the inclusion \(\text{Im } B^*B X \subseteq \text{Im } X\) can be written as \(\beta (\text{Im } X) \subseteq \text{Im } X\). Hence \(\beta (\text{Im } X) = \text{Im } X\), and we obtain \(\text{Im } X = \beta^{-1} (\text{Im } X) = \text{Im } (B^*B)^\dagger X\).
We show that (v) implies (ii). We know that \( \text{Im} B^* B X = \text{Im} X \) implies that \( \text{Im} (B^* B) \dagger X = \text{Im} X \). The converse implication follows from \( (B^* B) \dagger = B^\dagger (B^\dagger)^* \). □

According to [1, p. 160, Example 22] or [3, p. 23, Theorem 1.4], the reverse-order property \( (B X) \dagger = X^\dagger B^\dagger \) is equivalent to

\[
\text{Im} X X^* B^* \subseteq \text{Im} B^* \quad \text{and} \quad \text{Im} B^* B X \subseteq \text{Im} X. \tag{3.4}
\]

We did not take advantage of this result in order to make the proof of Theorem 3.3 self-contained. Using (3.4) we could have deduced (iv) from (iii) as follows. Since \( \text{Im} X \subseteq \text{Im} B^* \) implies \( \text{Im} X X^* B^* \subseteq \text{Im} B^* \), both conditions of (3.4) are satisfied. Hence \( (B X) \dagger = X^\dagger B^\dagger \).

Using Corollary 3.2 we can combine Theorem 3.3 with the analogous results for \( (C X C^\dagger) = X \) to obtain equivalence of the first six statements in the following theorem.

**Theorem 3.4.** The following statements are equivalent.

(i) \( C(B X C)^\dagger B = X^\dagger \).

(ii) \((B X)^\dagger B = X^\dagger \) and \((C X C)^\dagger = X^\dagger \).

(iii) \( \text{Im} B^* B X = \text{Im} X \) and \( \text{Ker} X C C^* = \text{Ker} X \).

(iv) \( \text{Im} B^* B X \subseteq \text{Im} X \subseteq \text{Im} B^* \) and \( \text{Ker} C^* \subseteq \text{Ker} X \subseteq \text{Ker} X C C^* \).

(v) \( (B X)^\dagger = X^\dagger B^\dagger \), \( (C X)^\dagger = C^\dagger X^\dagger \), \( X = B^\dagger B X \), and \( X = X C C^\dagger \).

(vi) \( \text{Im}(B^* B)^\dagger X = \text{Im} X \) and \( \text{Ker}(C C^*)^\dagger = \text{Ker} X \).

(vii) \( (B X C)^\dagger = C^\dagger X^\dagger B^\dagger \) and \( X = B^\dagger B X C C^\dagger \).

**Proof.** It remains to include (vii) in the graph of equivalences. We will exploit the equivalence

\[
X = B^\dagger B X C C^\dagger \iff X = B^\dagger B X \quad \text{and} \quad X = X C C^\dagger. \tag{3.5}
\]

We show that (vii) implies (i). Put \( W = C(B X C)^\dagger B \). Then \( W = C C^\dagger X^\dagger B^\dagger B \).

Using (3.5), we find \( X W X = X \), \( W X W = W \), \( W X = C C^\dagger X^\dagger X C C^\dagger = (W X)^* \), and \( X W = B^\dagger B X X^\dagger B^\dagger B = (X W)^* \), which means that \( W = X^\dagger \).

We show that (v) implies (vii). By (3.5), \( X = B^\dagger B X C C^\dagger \). Note that \( X = X C C^\dagger \) implies that \( \text{Im} B X C = \text{Im} B X \).

Hence \( P_{B X} = P_{B X C} \). Analogously, \( P_{(X C)^*} = P_{(B X C)^*} \).

Put \( A = B X C \) and \( W = C^\dagger X^\dagger B^\dagger \). Then the Moore equations are easily verified as follows:

\[
A W = B(X C C^\dagger) (X^\dagger B^\dagger) = B X (B X)^\dagger = P_{B X} = P_{B X C} = P_A,
\]

\[
W A = (C^\dagger X^\dagger) (B^\dagger B X C) = (X C)^\dagger X C = P_{(X C)^*} = P_{(B X C)^*} = P_{A^*}. \quad \square
\]

Now let us assume that \( X \) is Hermitian and \( C = B^* \). Theorem 3.3(iii) yields a sufficient condition for the cancellation property, which will be applied to compliance matrices in Section 5.

**Corollary 3.5.** If \( X = X^* \) and \( \text{Im} B^* = \text{Im} X \) then

\[
X^\dagger = B^*(B X B^*)^\dagger B.
\]
4. The matrix equation $C(BXC)^\dagger B = X^\dagger$

In this section we consider the matrix equation

$$C(BXC)^\dagger B = X^\dagger,$$  \hspace{1cm} (4.1)

where $B \in \mathbb{C}^{s \times n}$ and $C \in \mathbb{C}^{m \times t}$ are given and $X \in \mathbb{C}^{n \times m}$ is unknown. We wish to determine all the solutions of (4.1). Set $\eta = \min\{s, t, m, n\}$. Then (4.1) implies that rank $X \leq \eta$. We know from Theorem 3.4 that $X$ satisfies (4.1) if and only if there exists a $B^*B$-invariant subspace $S$ of $\mathbb{C}^n$ and a $CC^*$-invariant subspace $T$ of $\mathbb{C}^m$ such that $\text{Im} \ X = S$ and $\text{Im} \ X^* = T$. Therefore the spectral decompositions of $B^*B$ and $CC^*$ should play a role in the following results. Suppose that rank $B = r$ and rank $C = q$. Let $\beta_i$, $i = 1, \ldots, k$, be the different nonzero eigenvalues of $B^*B$, and let $\nu_i$ be the multiplicity of $\beta_i$. Correspondingly, let $\gamma_j$, $j = 1, \ldots, \ell$, be the distinct nonzero eigenvalues of $CC^*$, and let $\mu_j$ be their multiplicities. Then there exists a matrix $U \in \mathbb{C}^{n \times r}$ with

$$U = (U_1, \ldots, U_k), \quad U_i \in \mathbb{C}^{n \times \nu_i}, \quad i = 1, \ldots, k, \quad U^*U = I_r,$$ \hspace{1cm} (4.2)

such that

$$B^*B = \sum_{i=1}^{k} \beta_i U_i U_i^*, \hspace{1cm} (4.3)$$

and a matrix $V \in \mathbb{C}^{q \times m}$ with

$$V = (V_1^*, \ldots, V_\ell^*)^*, \quad V_j \in \mathbb{C}^{\mu_j \times m}, \quad j = 1, \ldots, \ell, \quad VV^* = I_q,$$ \hspace{1cm} (4.4)

such that

$$CC^* = \sum_{j=1}^{\ell} \gamma_j V_j^* V_j.$$ \hspace{1cm} (4.5)

Suppose that the columns of a matrix $G \in \mathbb{C}^{n \times p}$ are an orthogonal basis of a subspace $S \subseteq \text{Im} \ B^*$. Then $S$ is invariant under $B^*B$ if and only if, for some permutation matrix $P$,

$$GP = (U_1 M_1, \ldots, U_k M_k) = U \text{diag}(M_1, \ldots, M_k),$$ \hspace{1cm} (4.6)

with

$$M_i \in \mathbb{C}^{\nu_i \times \tau_i}, \quad 0 \leq \tau_i \leq \nu_i, \quad M_i^* M_i = I_{\tau_i}, \quad i = 1, \ldots, k, \hspace{1cm} (4.7)$$

and $\text{Im} \ X = S$ if and only if

$$X = GL, \quad \text{for some } L \in \mathbb{C}^{p \times m} \text{ with rank } L = p.$$ \hspace{1cm} (4.8)

Similarly, if $T \subseteq \text{Im} \ C$ is a subspace of $\mathbb{C}^m$ with an orthogonal basis given by the columns of a matrix $H \in \mathbb{C}^{m \times p}$, then $T$ is invariant under $CC^*$ if and only if, for some permutation matrix $Q$,

$$QH = (V_1^* N_1^*, \ldots, V_\ell^* N_\ell^*)^* = \text{diag}(N_1, \ldots, N_\ell) V,$$ \hspace{1cm} (4.9)
with
\[ N_j \in \mathbb{C}^{\mu_j \times \omega_j}, \quad 0 \leq \omega_j \leq \mu_j, \quad N_j N_j^\dagger = I_{\omega_j}, \quad j = 1, \ldots, \ell. \quad (4.10) \]
Moreover, \( \text{Im } X^* = T \) if and only if
\[ X = KH \quad \text{for some } K \in \mathbb{C}^{n \times p} \text{ with rank } K = p. \quad (4.11) \]

**Theorem 4.1.** Let \( U_i, i = 1, \ldots, k, \) and \( V_j, j = 1, \ldots, \ell, \) be given as in \((4.2), (4.3), \) and \((4.4), (4.5), \) respectively. Suppose that \( p \leq \eta. \) Then \( X \) is a solution of \((4.1)\) and rank \( X = p \) if and only if
\[ X = U \text{diag}(M_1, \ldots, M_k)Z \text{diag}(N_1, \ldots, N_\ell)V \]
\[ = (U_1 M_1, \ldots, U_k M_k)Z \begin{pmatrix} N_1 V_1 \\ \vdots \\ N_\ell V_\ell \end{pmatrix}, \quad (4.12) \]
where \( M_i \in \mathbb{C}^{\nu_i \times \rho_i}, \ i = 1, \ldots, k, \) and \( N_j \in \mathbb{C}^{\omega_j \times \mu_j}, \ j = 1, \ldots, \ell, \) are as in \((4.7)\) and \((4.10)\), and \( \sum \rho_i = \sum \omega_j = p, \) and \( Z \in \mathbb{C}^{p \times p} \) is nonsingular.

**Proof.** Suppose that \( X \) is given as in \((4.12)\). Then \( L = Z \text{diag}(N_1, \ldots, N_\ell)V \in \mathbb{C}^{p \times m}, \) has full row rank. This leads to \((4.8)\), and therefore \( \text{Im } B^*B X = \text{Im } X. \) Similarly,
\[ K = U \text{diag}(M_1, \ldots, M_k)Z \in \mathbb{C}^{n \times p}, \]
has full column rank. Then \((4.11)\) yields \( \text{Im } CC^* X^* = \text{Im } X^*. \) Hence \( X \) satisfies \((4.1)\).

On the other hand, if \( X \) is a solution of \((4.1)\) then we have seen that \( X = GL = KH \) with \( G \) and \( H \) given by \((4.6)\) and \((4.9)\). Thus
\[ \text{rank } L = \text{rank } K = \text{rank } X = p \quad (4.13) \]
and
\[ X = XX^\dagger X = GL(GL)^\dagger KH = G[LL^\dagger G^\dagger K]H = GZH. \]
It follows from \((4.13)\) that \( Z = LL^\dagger G^\dagger K \in \mathbb{C}^{p \times p} \) has full rank. \( \square \)

5. Applications

We first discuss an issue related to compliance matrices. Then we consider reverse-order laws and weighted generalized inverses.

5.1. Compliance matrices

In Section 1 a compliance matrix \( N_r \) was introduced. Recall that \( N_r \) is the Moore–Penrose inverse of a symmetric matrix \( F_r, \) which is related to a Hessian matrix \( F_s \) by \( F_s = J^\dagger F_r (J^\dagger)^T. \) With regard to \([2] \) or \([8] \) it is important to retrieve information on \( F_r^\dagger \) from the matrix \( F_s^\dagger. \) In a chemical set-up Brandhorst \([2] \) makes the assumption that
\[ \theta = \text{rank } F_s = \text{rank } F_r, \quad (5.1) \]
where $\theta$ represents the maximal degree of freedom in a molecule. Let us show that the additional assumption

$$\text{Im } J \subseteq \text{Im } F_r$$

is sufficient to recover $F_r^\dagger$ completely from $F_s^\dagger$. Note that (5.1) implies that rank $J \geq$ rank $F_r$. Thus (5.2) implies $\text{Im } J = \text{Im } F_r$. Hence the following observation is an immediate consequence of Corollary 3.5.

**Proposition 5.1.** Assume that (5.1) and (5.2) hold. Then $F_r^\dagger = (J^\dagger)^T F_s^\dagger J^\dagger$.

5.2. **Reverse-order laws** Let $R \in \mathbb{C}^{m \times n}$ and $S \in \mathbb{C}^{n \times p}$. According to [13, p. 3110], the pair $(R, S)$ fulfills the reverse-order law $(RS)^\dagger = S^\dagger R^\dagger$ if and only if both mixed-type reverse-order laws

$$\begin{align*}
(RS)^\dagger &= S^\dagger (R^\dagger RSS^\dagger)^\dagger R^\dagger, \\
(RS)^\dagger &= S^\dagger (R^\dagger RSS^\dagger)^\dagger R^\dagger
\end{align*}$$

(5.3)

are satisfied. Both equations in (5.3) are of the form $X^\dagger = C(BXC)^\dagger B$ if we set $X = RS$ and $(B, C) = (R^\dagger, S^\dagger)$ or $(B, C) = (R^*, S^*)$. Hence we can use Theorem 3.4 to obtain the results on mixed-type reverse-order laws.

**Theorem 5.2** [12, Theorem 1]. The following statements are equivalent.

(i) $(RS)^\dagger = S^\dagger (R^\dagger RSS^\dagger)^\dagger R^\dagger$.

(ii) $\text{Im } (R^\dagger)^* R^\dagger RS = \text{Im } RS$ and $\text{Im } S^\dagger (S^\dagger)^*(RS)^* = \text{Im } (RS)^*$.

(iii) $(R^\dagger RS)^\dagger R^\dagger = S^\dagger (RSS^\dagger)^\dagger$.

**Proof.** To show that (i) implies (ii), apply part (iii) of Theorem 3.4.

We show that (iii) implies (i). Set $A = RS$ and $W = (R^\dagger RS)^\dagger R^\dagger$.

From $\text{Im } RSS^\dagger = \text{Im } RS$, it follows that $P_{RSS^\dagger} = P_{RS}$. If (iii) holds then

$$AW = RSS^\dagger (RSS^\dagger)^\dagger = P_{RSS^\dagger} = P_{RS}.$$ 

Similarly,

$$WA = (R^\dagger RS)^\dagger R^\dagger RS = P_{(R^\dagger RS)^*} = P_{(RS)^*}.$$

Thus $W = (RS)^\dagger$. Setting $(B, C, X) = (R^\dagger, S^\dagger, RS)$, we find that (iii) is equivalent to (3.3), and, by Corollary 3.2, also to (i).

The rank formula in (5.4) below is due to Mazko [7]. In the case where $BXC$ is an invertible square matrix the result is known as Wedderburn–Guttman theorem (see [10, 11]).

**Theorem 5.3.** Let $X \in \mathbb{C}^{n \times m}$, $B \in \mathbb{C}^{s \times n}$ and $C \in \mathbb{C}^{m \times t}$ so that rank $X = r$ and rank $BXC = h$. If $(BXC)^- \in (BXC)\{1\}$ then

$$\text{rank } [X - XC(BXC)^- BX] = r - h.$$  

(5.4)
In our context the special case with $r = h$ and $(BXC)^\dagger = (BXC)^\dagger$ is of interest. Note that (5.5) in the following proposition can be regarded as another cancellation property of the triple $(B, C, X)$. As we have seen in Example 1, it is weaker than $X^\dagger = C(BXC)^\dagger B$.

**Proposition 5.4.** The equality

$$XC(BXC)^\dagger BX = X$$

holds if and only if

$$\text{rank } X = \text{rank } BXC.$$  

**Proof.** The left-hand side of (5.5) can be written as

$$XC(BXC)^\dagger BX = XC(BXX^\dagger XC)^\dagger BX.$$ 

Thus we are in the setting of Theorem 3.4 with $X = X^\dagger$, $B = BX$ and $C = XC$. Because of

$$\text{Im } X^*B^*BXX^\dagger = \text{Im } X^*B^*BX = \text{Im } X^*B^*,$$

and $X^\dagger = X^*$, the condition $\text{Im } B^*BX = \text{Im } X$ can be expressed as $\text{Im } X^*B^* = \text{Im } X^*$, which implies rank $BX = \text{rank } X$. Similarly, $\text{Im } CC^*X^* = \text{Im } X^*$ is equivalent to rank $XC = \text{rank } X$. Since

$$\text{rank } X = \text{rank } BX = \text{rank } XC$$

is equivalent to (5.6), the proof is complete. 

5.3. Weighted Moore–Penrose inverses

Our third application deals with a generalization of the Moore–Penrose inverse. Let $M \in \mathbb{C}^{n\times n}$ and $N \in \mathbb{C}^{m\times m}$ be positive definite Hermitian matrices. If $A \in \mathbb{C}^{n\times m}$ then

$$A_{M,N}^\dagger = N^{-1/2}(M^{1/2}AN^{-1/2})^\dagger M^{1/2}$$

is the weighted Moore–Penrose inverse of $A$ with respect to $M$ and $N$. The following observation, which is contained in [1, p. 121, Example 42], is an immediate consequence of Theorem 3.4.

**Proposition 5.5.** The equality $A_{M,N}^\dagger = A^\dagger$ holds if and only if

$$\text{Im } MA = \text{Im } A \quad \text{and} \quad \text{Ker } AN = \text{Ker } A.$$ 

We indicate without proof a condition for the cancellation property in the case of a weighted generalized inverse.

**Theorem 5.6.** Let $B, C$ and $X$ be of size $n \times n$, $m \times m$ and $n \times m$, respectively. Then

$$C(BXC)_{M,N}^\dagger B = X_{M,N}^\dagger$$

if and only if

$$\text{Im } M^{-1}B^*MBX = \text{Im } X \quad \text{and} \quad \text{Ker } XCN^{-1}C^*N = \text{Ker } X.$$
6. Open problems

Consider the matrix

\[ M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}, \tag{6.1} \]

and let

\[ S_1 = A - BD^\dagger C \quad \text{and} \quad S_2 = D - CA^\dagger B \tag{6.2} \]

be the associated Schur complements. If \( A = BXC \) and \( D = X^\dagger \) then \( S_1 = BXC - B(X^\dagger)^\dagger C = 0 \) and \( S_2 = X^\dagger - C(BXC)^\dagger B \). Thus the cancellation property of \((B, C, X)\) is equivalent to \( S_2 = 0 \). The following problem arises. Let \( M \) be the matrix in (6.1) with Schur complements (6.2). When does \( S_1 = 0 \) imply \( S_2 = 0 \)?

A more detailed investigation of cancellation properties would require an understanding of relations of the form

\[ C(BXC)^{(i,\ldots,j)} B = X^{(i,\ldots,j)}. \]

We remark that a comprehensive study of triple matrix products and mixed-type reverse-order properties of the form

\[ (BXC)^{(i,\ldots,j)} = (XC)^{(i,\ldots,j)} X(BX)^{(i,\ldots,j)} \]

can be found in [14].

References


[12] Y. Tian, ‘The reverse-order law \((AB)^\dagger = B^\dagger (A^\dagger AB B^\dagger)^\dagger A^\dagger\) and its equivalent equalities’, *J. Math. Kyoto Univ.* **45** (2005), 841–850.


TOBIAS DAMM, Fachbereich Mathematik, TU Kaiserslautern, D-67663 Kaiserslautern, Germany
e-mail: damm@mathematik.uni-kl.de

HARALD K. WIMMER, Mathematisches Institut, Universität Würzburg, D-97074 Würzburg, Germany
e-mail: wimmer@mathematik.uni-wuerzburg.de