JOURNAL OF
COMPUTATIONAL AND
APPLIED MATHEMATICS

Journal of Computational and Applied Mathematics 137 (2001) 317–329

www.elsevier.com/locate/cam

On the perturbation and subproper splittings for the generalized inverse $A^{(2)}_{T,S}$ of rectangular matrix A ☆

Yimin Wei[a,*], Hebing Wu[b]

[a]Department of Mathematics and Laboratory of Mathematics for Nonlinear Science, Fudan University, Shanghai 200433, People's Republic of China

[b]Institute of Mathematics, Fudan University, Shanghai 200433, People's Republic of China

Received 8 December 1999; received in revised form 9 November 2000

Abstract

In this paper, the perturbation and subproper splittings for the generalized inverse $A^{(2)}_{T,S}$, the unique matrix X such that $XAX = X$, $R(X) = T$ and $N(X) = S$, are considered. We present lower and upper bounds for the perturbation of $A^{(2)}_{T,S}$. Convergence of subproper splittings for computing the special solution $A^{(2)}_{T,S}b$ of restricted rectangular linear system $Ax=b$, $x \in T$, are studied. For the solution $A^{(2)}_{T,S}b$ we develop a characterization. Therefore, we give a unified treatment of the related problems considered in literature by Ben-Israel, Berman, Hanke, Neumann, Plemmons, etc. © 2001 Published by Elsevier Science B.V.

MSC: 15A09; 65F15; 65F20

Keywords: Generalized inverse; Perturbation theory; Subproper splitting; Proper splitting, convergence; Characterization

1. Introduction and preliminary results

In this paper, we adopt the same notations on generalized inverses used in Ben-Israel and Greville [3] and Berman and Plemmons [6]. It is well known that the commonly important six kinds of generalized inverses: the Moore–Penrose inverse A^+, the weighted Moore–Penrose inverse A^+_{MN}, the Drazin inverse A^{D}, the group inverse A_{g}, the Bott–Duffin inverse $A^{(-1)}_{(L)}$ and the generalized

☆ Project 19901006 supported by National Natural Science Foundation of China and Science Foundation of Laboratory of Computational Physics. Partial work was finished when the first author visited Harvard University and MIT. The second author is also supported by the State Major Key Project for Basic Researches in China.

* Corresponding author.

E-mail address: ymwei@fudan.edu.cn (Y. Wei).

0377-0427/01/$ - see front matter
PII: S0377-0427(00)00710-X

Bott–Duffin inverse $A_{(L)}^{(+)}$ can all be viewed as special $\{2\}$-inverses $A_{T,S}^{(2)}$, having the prescribed range T and null space S.

The $\{2\}$-inverse has many applications, for example, the application in the iterative methods for solving nonlinear equations [3,18] and the applications to statistics [11,13]. In particular, $\{2\}$-inverse plays an important role in stable approximations of ill-posed problems and in linear and nonlinear problems involving rank-deficient generalized inverse [17,27]. There are many numerical methods for computing $A_{T,S}^{(2)}$ (see [8,30] and references therein).

In 1966, Ben-Israel [1] gave the perturbation for the Moore–Penrose inverse A^+, and recently this result has been extended to the Drazin inverse A^{D} and the group inverse A_{g}, see [28,31]. In this paper, our first goal is to make a unified treatment for the perturbation of the generalized inverse $A_{T,S}^{(2)}$.

In the past decades, many authors studied various kinds of linear equations. Ben-Israel [2], Verghese [24], and Wang [25] considered the Cramer rule for minimum norm solution or least-squares solutions of consistent and inconsistent linear equations

$$Ax = b,$$

where $A \in \mathbb{C}^{m\times n}$, and $b \in \mathbb{C}^m$. Werner [33] presented the Cramer rule for restricted linear equations:

$$Ax = b, \quad x \in K,$$

where $A \in \mathbb{C}^{m\times n}$, $b \in \mathbb{C}^m$, and K is a complementary subspace of $N(A)$. Wang [26] discussed the special singular equations

$$Ax = b, \quad x \in R(A^k),$$

where $A \in \mathbb{C}^{n\times n}$, $k = \mathrm{index}(A)$, and $b \in R(A^k)$. Summarizing these equations, Chen [7,9] discussed a general restricted linear equation:

$$Ax = b, \quad x \in T, \tag{1.1}$$

where $A \in \mathbb{C}^{m\times n}$, and T is a subspace of $\mathbb{C}^n$. In [7] Chen showed that the restricted system (1.1) has a solution if and only if $b \in AT$, and has a unique solution if and only if $b \in AT$ and $T \cap N(A) = \{0\}$. Let subspace $S \in \mathbb{C}^m$ be such that $AT \oplus S = \mathbb{C}^m$, then one sees that $A_{T,S}^{(2)}b$ is a solution or the unique solution of system (1.1) assuming it is consistent. Our second goal is to establish a characterization of this special solution despite system (1.1) is solvable or not.

To obtain a solution of (1.1) and its special cases, one often splits A into the form

$$A = U - V. \tag{1.2}$$

Then there results the natural iteration

$$x_{i+1} = U^- V x_i + U^- b, \tag{1.3}$$

where U^- is some generalized inverse of A. In literature, various choices of U are considered in order to compute the corresponding solution of (1.1), see [4,5,14,15,19,29] and Section 4. Our last contribution to this problem is to consider two kinds of subproper splittings (1.2) of A for computing the solution $A_{T,S}^{(2)}b$ of (1.1), which, to our best knowledge, include and extend all kinds of existing splittings for solution of (1.1).

The following lemmata are needed in what follows.

Lemma 1.1 (Ben-Israel and Greville [3]). *Let $A \in \mathbb{C}^{m\times n}$, $\operatorname{rank}(A)=r$, and let T and S be subspaces of $\mathbb{C}^n$ and $\mathbb{C}^m$, respectively, $\dim T=\dim S^{\perp}=t\leqslant r$. Then A has a $\{2\}$-inverse X such that $R(X)=T$ and $N(X)=S$ if and only if*

$$AT \oplus S = \mathbb{C}^m, \tag{1.4}$$

in which case X is unique and denoted by $X = A^{(2)}_{T,S}$.

Note 1 (Ben-Israel and Greville [3]). When $t=r$ in Lemma 1.1, condition (1.4) is equivalent to

$$R(A) \oplus S = \mathbb{C}^m \quad \text{and} \quad T \oplus N(A) = \mathbb{C}^n, \tag{1.5}$$

in which case $A^{(2)}_{T,S}$ is indeed $A^{(1,2)}_{T,S}$, the unique $\{1,2\}$-inverse of A having range T and null space S.

Note 2 (Ben-Israel and Greville [3]). $AA^{(2)}_{T,S}=P_{AT,S}$, $A^{(2)}_{T,S}A=P_{T,(A^*S^{\perp})^{\perp}}$, $AA^{(1,2)}_{T,S}=P_{R(A),S}$ and $A^{(1,2)}_{T,S}A=P_{T,N(A)}$.

Lemma 1.2 (Chen [8]). (1) *Let $A\in\mathbb{C}^{m\times n}$. Then, for the Moore–Penrose inverse A^+ and the weighted Moore–Penrose inverse A^+_{MN}, one has*

(i) $A^+ = A^{(1,2)}_{R(A^*),N(A^*)} = A^{(2)}_{R(A^*),N(A^*)}$.
(ii) $A^+_{\mathrm{M,N}} = A^{(1,2)}_{R(A^\#),N(A^\#)} = A^{(2)}_{R(A^\#),N(A^\#)}$, *where $A^\# = N^{-1}A^*M$. N and M are Hermitian positive-definite matrices of order n and m, respectively.*
(2) *Let $A \in \mathbb{C}^{n\times n}$. Then, for the Drazin inverse A^{D}, the group inverse A_{g}, the Bott–Duffin inverse $A^{(-1)}_{(L)}$ and the generalized Bott–Duffin inverse $A^{(+)}_{(L)}$, one has*
(iii) $A^{\mathrm{D}} = A^{(2)}_{R(A^k),N(A^k)}$ *where* $k = \operatorname{index}(A)$.
(iv) *In particular, when* $\operatorname{index}(A)=1$, $A_{\mathrm{g}} = A^{(2)}_{R(A),N(A)}$.
(v) $A^{(-1)}_{(L)} = A^{(2)}_{L,L^{\perp}} = P_L(AP_L+P_{L^{\perp}})^{-1}$, *where L is a subspace of $\mathbb{C}^n$ such that $AL\oplus L^{\perp}=\mathbb{C}^n$ and P_L is the orthogonal projector on L;*
(vi) $A^{(+)}_{(L)} = A^{(-1)}_{(S)} = A^{(2)}_{S,S^{\perp}}$ *where* $S=R(P_LA)$.

2. Perturbation theory

Let $B=A+E\in\mathbb{C}^{m\times n}$. Let subspaces $T,\tilde{T}\subseteq\mathbb{C}^n$ and subspaces $S,\tilde{S}\in\mathbb{C}^m$ be such that $AT\oplus S=\mathbb{C}^m$ and $B\tilde{T}\oplus\tilde{S}=\mathbb{C}^m$. To present perturbation bounds for generalized inverse $A^{(2)}_{T,S}$ we give first a simple expression of $B^{(2)}_{\tilde{T},\tilde{S}}$ in terms of $A^{(2)}_{T,S}$ and E.

Theorem 2.1. *Let $B=A+E\in\mathbb{C}^{m\times n}$. Let subspaces $T,\tilde{T}$ of $\mathbb{C}^n$ and subspaces $S,\tilde{S}$ of $\mathbb{C}^m$ be such that $AT\oplus S=\mathbb{C}^m$ and $B\tilde{T}\oplus\tilde{S}=\mathbb{C}^m$. Then*

$$B^{(2)}_{\tilde{T},\tilde{S}} = (I+A^{(2)}_{T,S}E)^{-1}A^{(2)}_{T,S} = A^{(2)}_{T,S}(I+EA^{(2)}_{T,S})^{-1} \tag{2.1}$$

if and only if

$$\tilde{T}=T \quad \textit{and} \quad \tilde{S}=S. \tag{2.2}$$

Proof. The necessary part is obvious. Assume (2.2) holds. We first show that $I+A^{(2)}_{T,S}E$ is invertible. If not, then there is a nonzero vector x such that

$$(I + A^{(2)}_{T,S}E)x = 0,$$

i.e.,

$$x = -A^{(2)}_{T,S}Ex \in T = \tilde{T}. \tag{2.3}$$

Multiplying $A^{(2)}_{T,S}A$ with (2.3) yields

$$A^{(2)}_{T,S}Bx = 0,$$

which implies $Bx \in S=\tilde{S}$. But from (2.3) we see $Bx \in B\tilde{T}$. Hence $Bx=0$. Therefore, $x=B^{(2)}_{\tilde{T},\tilde{S}}Bx=0$ by $x \in \tilde{T}$ and Note 2, a contradiction.

By direct computation, we have

$$\begin{aligned}(I + A^{(2)}_{T,S}E)B^{(2)}_{T,S} &= B^{(2)}_{T,S} + A^{(2)}_{T,S}BB^{(2)}_{T,S} - A^{(2)}_{T,S}AB^{(2)}_{T,S}\\ &= B^{(2)}_{T,S} + A^{(2)}_{T,S} - B^{(2)}_{T,S}\\ &= A^{(2)}_{T,S}.\end{aligned}$$

Thus $B^{(2)}_{T,S} = (I + A^{(2)}_{T,S}E)^{-1}A^{(2)}_{T,S}$.

Since $(I+A^{(2)}_{T,S}E)A^{(2)}_{T,S}=A^{(2)}_{T,S}(I+EA^{(2)}_{T,S})$, so $B^{(2)}_{T,S}=A^{(2)}_{T,S}(I+EA^{(2)}_{T,S})^{-1}$ and the proof is complete. □

Corollary 2.2 (Wei [32]). *Let $B=A+E\in\mathbb{C}^{m\times n}$, and let T,S be subspaces of $\mathbb{C}^n$ and $\mathbb{C}^m$, respectively, such that $AT \oplus S = \mathbb{C}^m$. Suppose $R(E)\subseteq AT$ and $R(E^*)\subseteq A^*S^{\perp}$. If $\|EA^{(2)}_{T,S}\| < 1$, then $B^{(2)}_{T,S}$ exists and*

$$B^{(2)}_{T,S} = (I + A^{(2)}_{T,S}E)^{-1}A^{(2)}_{T,S} = A^{(2)}_{T,S}(I + EA^{(2)}_{T,S})^{-1}.$$

Proof. The hypothesis $\|EA^{(2)}_{T,S}\| < 1$ implies that $I + A^{(2)}_{T,S}E$ and $I + EA^{(2)}_{T,S}$ are nonsingular. By the assumption and Note 2,

$$B = A(I + A^{(2)}_{T,S}E) = (I + EA^{(2)}_{T,S})A \tag{2.4}$$

holds and $BT \subseteq AT$. Let P_T be the orthogonal projector on T, then $BT=R(BP_T)$ and $AT=R(AP_T)$. It follows from (2.4) that

$$BP_T = (I + EA^{(2)}_{T,S})AP_T.$$

Thus $\operatorname{rank}(BP_T)=\operatorname{rank}(AP_T)$ and $BT=AT$.

Notice that $AT\oplus S=\mathbb{C}^m$, then $BT\oplus S=\mathbb{C}^m$ and $B^{(2)}_{T,S}$ exist. We conclude the proof of the corollary applying Theorem 2.1. □

From Theorem 2.1 and Corollary 2.2, we immediately obtain the following results.

Corollary 2.3. *Let $B=A+E$. Let subspaces $T\subseteq\mathbb{C}^n$ and $S\subseteq\mathbb{C}^m$ be such that $T\oplus N(A)=\mathbb{C}^n$ and $R(A)\oplus S=\mathbb{C}^m$. Suppose $R(E)\subseteq R(A)$ and $R(E^*)\subseteq R(A^*)$. If $\|EA^{(1,2)}_{T,S}\| < 1$, then $B^{(1,2)}_{T,S}$ exists and*

$$B^{(1,2)}_{T,S} = (I + A^{(1,2)}_{T,S}E)^{-1}A^{(1,2)}_{T,S} = A^{(1,2)}_{T,S}(I + EA^{(1,2)}_{T,S})^{-1}. \tag{2.5}$$

Corollary 2.4 (Ben-Israel [1]). *Let $B = A + E$. Suppose $R(E) \subseteq R(A)$ and $R(E^*) \subseteq R(A^*)$. If $\|EA^+\| < 1$, then*

$$B^+ = (I + A^+E)^{-1}A^+ = A^+(I + EA^+)^{-1}. \tag{2.6}$$

Corollary 2.5 (Wei [29]). *Let $B = A + E \in \mathbb{C}^{n\times n}$ with $\mathrm{index}(A) = k$. Suppose $R(E) \subseteq R(A^k)$ and $R(E^*) \subseteq R(A^{k^*})$. If $\|EA^{\mathrm{D}}\| < 1$, then $\mathrm{index}(B) = k$ and*

$$B^{\mathrm{D}} = (I + A^{\mathrm{D}}E)^{-1}A^{\mathrm{D}} = A^{\mathrm{D}}(I + EA^{\mathrm{D}})^{-1}. \tag{2.7}$$

We are in a position to present the main result of this section.

Theorem 2.6. *Let $B = A + E \in \mathbb{C}^{m\times n}$. Let subspaces $T \subseteq \mathbb{C}^n$ and $S \subseteq \mathbb{C}^m$ be such that $AT \oplus S = \mathbb{C}^m$. Suppose $R(E) \subseteq AT$ and $R(E^*) \subseteq A^*S^\perp$. If $\triangle = \|EA^{(2)}_{T,S}\| < 1$, then*

$$\frac{\triangle}{(1+\triangle)\mathscr{K}(A)} \leqslant \frac{\|B^{(2)}_{T,S} - A^{(2)}_{T,S}\|}{\|A^{(2)}_{T,S}\|} \leqslant \frac{\triangle}{1-\triangle}, \tag{2.8}$$

where $\mathscr{K}(A) = \|A\|\,\|A^{(2)}_{T,S}\|$ is the condition number of generalized inverse $A^{(2)}_{T,S}$.

Proof. From Theorem 2.1 and Corollary 2.2 we have

$$B^{(2)}_{T,S} - A^{(2)}_{T,S} = -B^{(2)}_{T,S}EA^{(2)}_{T,S}. \tag{2.9}$$

Taking norm of (2.9) leads to

$$\|B^{(2)}_{T,S} - A^{(2)}_{T,S}\| \leqslant \triangle\|B^{(2)}_{T,S}\|. \tag{2.10}$$

It follows from (2.1) that

$$\frac{\|A^{(2)}_{T,S}\|}{1+\triangle} \leqslant \|B^{(2)}_{T,S}\| \leqslant \frac{\|A^{(2)}_{T,S}\|}{1-\triangle}. \tag{2.11}$$

Substituting (2.11) into (2.10), one has

$$\frac{\|B^{(2)}_{T,S} - A^{(2)}_{T,S}\|}{\|A^{(2)}_{T,S}\|} \leqslant \frac{\triangle}{1-\triangle}. \tag{2.12}$$

On the other hand, we note that $BB^{(2)}_{T,S} = P_{BT,S} = P_{AT,S} = AA^{(2)}_{T,S}$ since $AT = BT$, so we have

$$EA^{(2)}_{T,S} = B(A^{(2)}_{T,S} - B^{(2)}_{T,S}). \tag{2.13}$$

Substituting (2.4) into (2.13) we have

$$EA^{(2)}_{T,S} = (I + EA^{(2)}_{T,S})A(A^{(2)}_{T,S} - B^{(2)}_{T,S}). \tag{2.14}$$

It follows from (2.14) that

$$\frac{\|B^{(2)}_{T,S} - A^{(2)}_{T,S}\|}{\|A^{(2)}_{T,S}\|} \geqslant \frac{\triangle}{(1+\triangle)\mathscr{K}(A)} \tag{2.15}$$

and, considering (2.12) and (2.15), the proof is complete. □

Corollary 2.7 (Ben-Israel [1]). *Let $B = A + E \in \mathbb{C}^{m\times n}$. Suppose $R(E) \subseteq R(A)$ and $R(E^*) \subseteq R(A^*)$. If $\triangle = \|EA^+\| < 1$, then*

$$\frac{\triangle}{(1+\triangle)\mathscr{K}(A)} \leqslant \frac{\|B^+ - A^+\|}{\|A^+\|} \leqslant \frac{\triangle}{1-\triangle}, \tag{2.16}$$

where $\mathscr{K}(A) = \|A\|\,\|A^+\|$ is the condition number of Moore–Penrose inverse A^+.

Corollary 2.8 (Wei [29]). *Let $B = A + E \in \mathbb{C}^{n\times n}$ with $\operatorname{index}(A) = k$. Suppose $R(E) \subseteq R(A^k)$ and $R(E^*) \subseteq R(A^{k^*})$. If $\triangle = \|EA^{\mathrm{D}}\| < 1$, then*

$$\frac{\triangle}{(1+\triangle)\mathscr{K}(A)} \leqslant \frac{\|B^{\mathrm{D}} - A^{\mathrm{D}}\|}{\|A^{\mathrm{D}}\|} \leqslant \frac{\triangle}{1-\triangle}, \tag{2.17}$$

where $\mathscr{K}(A) = \|A\|\,\|A^{\mathrm{D}}\|$ is the condition number of Drazin inverse A^{D}.

3. Characterization of $A^{(2)}_{T,S}b$

In this section, we shall establish a characteristic for the special solution $A^{(2)}_{T,S}b$ of the restricted linear equations (1.1) despite $b \in AT$ or not. Our result gives better characterization of $A^{(2)}_{T,S}b$ than that of Chen [7], and extends the results of Hanke and Neumann [12] for $A^{(1,2)}_{T,S}b$ and that of Ben-Israel [3] for A^+b.

Recall that $AA^{(2)}_{T,S} = P_{AT,S}$ and $A^{(2)}_{T,S}A = P_{T,(A^*S^\perp)^\perp}$. As done in [10,12], we define the norms

$$\|x\|^2_{AT,S} = \|P_{AT,S}x\|_2^2 + \|(I - P_{AT,S})x\|_2^2 \quad \forall x \in \mathbb{C}^m \tag{3.1}$$

and

$$\|y\|^2_{T,(A^*S^\perp)^\perp} = \|P_{T,(A^*S^\perp)^\perp}y\|_2^2 + \|(I - P_{T,(A^*S^\perp)^\perp})y\|_2^2 \quad \forall y \in \mathbb{C}^n. \tag{3.2}$$

Theorem 3.1. *The solution $A^{(2)}_{T,S}b$ is a least-squares solution of $Ax = b$ in $N(A) + T$ with respect to the norm $\|.\|_{AT,S}$, i.e.,*

$$\|b - AA^{(2)}_{T,S}b\|_{AT,S} = \min_{x\in N(A)+T} \|b - Ax\|_{AT,S}. \tag{3.3}$$

Moreover, $A^{(2)}_{T,S}b$ is the unique minimum $\|.\|_{T,(A^*S^\perp)^\perp}$-norm solution to $A^{(2)}_{T,S}Ax = A^{(2)}_{T,S}b$.

Proof. For any $x \in N(A) + T$, we have

$$\begin{aligned}\|b - Ax\|^2_{AT,S} &= \|P_{AT,S}(b - Ax)\|_2^2 + \|(I - P_{AT,S})(b - Ax)\|_2^2 \\ &= \|P_{AT,S}b - Ax\|_2^2 + \|(I - P_{AT,S})b\|_2^2 \\ &\geqslant \|(I - P_{AT,S})b\|_2^2 = \|b - AA^{(2)}_{T,S}b\|^2_{AT,S},\end{aligned} \tag{3.4}$$

which implies $A^{(2)}_{T,S}b$ is a least-squares solution $Au = b$ in $N(A) + T$ with respect to the norm $\|.\|_{AT,S}$.

It is evident that $A^{(2)}_{T,S}Ax = A^{(2)}_{T,S}b$ is consistent and the general solution is $x = A^{(2)}_{T,S}b + z$, $\forall z \in (A^*S^\perp)^\perp$. Thus

$$\begin{aligned}
\|x\|^2_{T,(A^*S^\perp)^\perp} &= \|P_{T,(A^*S^\perp)^\perp}(A^{(2)}_{T,S}b+z)\|_2^2 + \|(I-P_{T,(A^*S^\perp)^\perp})(A^{(2)}_{T,S}b+z)\|_2^2 \\
&= \|A^{(2)}_{T,S}b\|_2^2 + \|z\|_2^2 \\
&\geqslant \|A^{(2)}_{T,S}b\|_2^2 = \|A^{(2)}_{T,S}b\|^2_{T,(A^*S^\perp)^\perp}.
\end{aligned} \tag{3.5}$$

The equality holds in (3.5) if and only if $z=0$ which implies $A^{(2)}_{T,S}b$ is the minimum $\|.\|_{T,(A^*S^\perp)^\perp}$-norm solution of $A^{(2)}_{T,S}Ax = A^{(2)}_{T,S}b$. $\square$

Corollary 3.2 (Hanke and Neumann [12]). *Let $A \in \mathbb{C}^{m\times n}$, and let subspaces $T \subseteq \mathbb{C}^n$ and $S \subseteq \mathbb{C}^m$ be such that $T \oplus N(A) = \mathbb{C}^n$ and $R(A) \oplus S = \mathbb{C}^m$. Then $A^{(1,2)}_{T,S}b$ is a least-squares solution of $Ax = b$ with respect to the norm $\|.\|_{R(A),S}$; moreover, $A^{(1,2)}_{T,S}b$ is the unique minimum $\|.\|_{T,N(A)}$-norm solution of $A^{(1,2)}_{T,S}Ax = A^{(1,2)}_{T,S}b$.*

For $T = R(A^*)$ and $S = N(A^*)$, we obtain

Corollary 3.3 (Ben-Israel and Greville [3]). *Let $A \in \mathbb{C}^{m\times n}$. Then A^+b is the minimum norm least-squares solution of $Ax = b$.*

4. Subproper splittings

Iterative methods for solution of system (1.1) of the form

$$x_{i+1} = Bx_i + c, \quad i = 0,1,2,\ldots, \tag{4.1}$$

where B is the nth-order complex matrix, are often employed. For this reason B is commonly called the iteration matrix. As mentioned in Section 1, B arises in a large number of cases from a splitting (1.2) of A.

The purpose of this section is to unify and extend well-known results concerning the convergence of iterative scheme (4.1) to a special solution of (1.1), that is, the solution $A^{(2)}_{T,S}b$.

In our study, we find it instructive to classify the known results into the following categories:

(i) $m = n$, A and U in (1.2) are nonsingular, and the iteration matrix of (4.1) is $B = U^{-1}V$ (see [23,34,20,16,22]).
(ii) $m = n$, in (1.2) A is a singular matrix and U is a nonsingular matrix and the iteration matrix is $B = U^{-1}V$ (see [15,21]).
(iii) A and U in (1.2) are rectangular matrices and $R(U) = R(A)$ and $N(U) = N(A)$. This splitting is called proper splitting by Berman and Plemmons. In this case, the iteration matrix is $B = U^+V$ (see [4,5]).
(iv) A and U in (1.2) are rectangular matrices, $m \geqslant n$, U has full column rank, and $UU^+A = A$, and the iteration matrix is $B = U^+V$ (see [14]).

(v) $m=n$, A and U in (1.2) satisfy $R(U)=R(A^k)$ and $N(U)=N(A^k)$ where $k=\text{index}(A)$. This splitting is called index splitting by Wei. In this case the iteration matrix is U_gV (see [29]).
(vi) A and U in (1.2) are rectangular matrices. Neumann call splitting (1.2) is subproper if $R(A) \subseteq R(U)$ and $R(A^*) \subseteq R(U^*)$. In this case the iteration matrix of (4.1) is $B=U^+V$ (see [19]).

To unify all kinds of splittings above we give

Definition 4.1. Let $A=U-V \in \mathbb{C}^{m\times n}$. Let subspaces $T, \tilde{T} \subseteq \mathbb{C}^n$ and $S, \tilde{S} \subseteq \mathbb{C}^m$ be such that $AT \oplus S = \mathbb{C}^m$ and $U\tilde{T} \oplus \tilde{S} = \mathbb{C}^m$. The splitting (1.2) is called subproper if $T \subseteq \tilde{T}$ and $\tilde{S} \subseteq S$, and is called proper if $T=\tilde{T}$ and $S=\tilde{S}$.

It is easy to see that the splittings from (i)–(vi) but (v) are all subproper splittings according to Definition 4.1. We shall remark that the concept of subproper splitting in Definition 4.1 is more general than that of Neumann.

Now we give

Definition 4.2. Let $A, U \in \mathbb{C}^{m\times n}$. Let subspaces $T, \tilde{T} \subseteq \mathbb{C}^n$ and $S, \tilde{S} \subseteq \mathbb{C}^m$ be such that $AT \oplus S = \mathbb{C}^m$ and $U\tilde{T} \oplus \tilde{S} = \mathbb{C}^m$. Splitting (1.2) is called subproper if $\tilde{T} \subseteq T$ and $S \subseteq \tilde{S}$, and is called proper if $T=\tilde{T}$ and $S=\tilde{S}$.

The index splitting (v) is a subproper splitting in Definition 4.2. To make a distinction between the subproper splittings in Definitions 4.1 and 4.2, we shall call them in the sequel types I and II, respectively. In both cases the iterative scheme (4.1) becomes

$$x_{i+1} = U^{(2)}_{\tilde{T},\tilde{S}}Vx_i + U^{(2)}_{\tilde{T},\tilde{S}}b, \quad i=0,1,2,\ldots . \tag{4.2}$$

Theorem 4.3. *Let $A=U-V \in \mathbb{C}^{m\times n}$, and* let *$T, \tilde{T} \subseteq \mathbb{C}^n$ and $S, \tilde{S} \subseteq \mathbb{C}^m$ be such that $AT \oplus S = \mathbb{C}^m$ and $U\tilde{T} \oplus S = \mathbb{C}^m$. Let splitting* (1.2) *be of type* I *or type* II *subproper splitting. Then* (i)

$$A^{(2)}_{T,S} = (I - U^{(2)}_{\tilde{T},\tilde{S}}V)^{-1}U^{(2)}_{\tilde{T},\tilde{S}} = U^{(2)}_{\tilde{T},\tilde{S}}(I - VU^{(2)}_{\tilde{T},\tilde{S}})^{-1} \tag{4.3}$$

if and only if $T=\tilde{T}$ and $S=\tilde{S}$, i.e., the splitting is proper.

(ii) *The sequence of the iterations of* (4.2) *converges to $A^{(2)}_{T,S}b$ for every $x_0 \in \mathbb{C}^n$ and for every $b \in \mathbb{C}^m$ if and only if $\rho(U^{(2)}_{\tilde{T},\tilde{S}}V) < 1$ and splitting* (1.2) *is proper.*

Proof. Part (i) follows directly from Theorem 2.1, we have only to prove (ii). Assume that $\rho(U^{(2)}_{\tilde{T},\tilde{S}}V) < 1$ and $T=\tilde{T}$, $S=\tilde{S}$. Then by (i)

$$(I - U^{(2)}_{\tilde{T},\tilde{S}}V)^{-1}U^{(2)}_{\tilde{T},\tilde{S}} = A^{(2)}_{T,S}.$$

From (4.2) it is easily proven by induction that

$$x_i = (U^{(2)}_{\tilde{T},\tilde{S}}V)^i x_0 + \sum_{j=0}^{i-1} (U^{(2)}_{\tilde{T},\tilde{S}}V)^j U^{(2)}_{\tilde{T},\tilde{S}}b. \tag{4.4}$$

Since $\rho(U^{(2)}_{\tilde{T},\tilde{S}}V) < 1$, it follows from (4.4) that

$$\lim_{i\to\infty} x_i = (I - U^{(2)}_{\tilde{T},\tilde{S}}V)^{-1}U^{(2)}_{\tilde{T},\tilde{S}}b = A^{(2)}_{T,S}b.$$

Conversely, assume that the sequence of $\{x_i\}$ with respect to (4.2) converges to $A^{(2)}_{T,S}b$ independent of the initial guess x_0 and b, we must have $\rho(U^{(2)}_{\tilde{T},\tilde{S}}V) < 1$ and

$$A^{(2)}_{T,S} = (I - U^{(2)}_{\tilde{T},\tilde{S}}V)^{-1}U^{(2)}_{\tilde{T},\tilde{S}} = U^{(2)}_{\tilde{T},\tilde{S}}(I - VU^{(2)}_{\tilde{T},\tilde{S}})^{-1}.$$

By Theorem 2.1 we have $T = \tilde{T}$ and $S = \tilde{S}$ and the proof is complete. □

Corollary 4.4 (Berman and Plemmons [5]). *Let $A = U - V \in \mathbb{C}^{m\times n}$ be such that $R(U) = R(A)$ and $N(U) = N(A)$. Then*

(i) $A^+ = (I - U^+V)^{-1}U^+ = U^+(I - VU^+)^{-1}$.
(ii) *The sequence of iterations*

$$x_{i+1} = U^+Vx_i + U^+b$$

converges to A^+b for every $x_0 \in \mathbb{C}^n$ and every $b \in \mathbb{C}^m$ if and only if $\rho(U^+V) < 1$.

Corollary 4.5 (Wei [29]). *Let $A \in \mathbb{C}^{n\times n}$ with* index(A)=k. *Let $A=U-V$ be such that $R(U)=R(A^k)$ and $N(U) = N(A^k)$. Then*

(i) $A^{\mathrm{D}} = (I - U_gV)^{-1}U_g = U_g(I - VU_g)^{-1}$.
(ii) *The sequence of iterations*

$$x_{i+1} = U_gVx_i + U_gb$$

converges to $A^{\mathrm{D}}b$ for every $x_0 \in \mathbb{C}^n$ and every $b \in \mathbb{C}^n$ if and only if $\rho(U_gV) < 1$.

To present convergence theorem for type I subproper splitting, we assume that $b \in AT$ for the moment.

Lemma 4.6. *Let subspaces $T, \tilde{T} \subseteq \mathbb{C}^n$ and $S, \tilde{S} \subseteq \mathbb{C}^m$ be such that $AT \oplus S = \mathbb{C}^m$ and $U\tilde{T} \oplus \tilde{S} = \mathbb{C}^m$. Let $b \in AT$ and splitting* (1.2) *be type* I *subproper splitting of A. Then every solution $x \in \tilde{T}$ (of which at least one exists, e.g., $A^{(2)}_{T,S}b$) satisfies $x = U^{(2)}_{\tilde{T},\tilde{S}}Vx + U^{(2)}_{\tilde{T},\tilde{S}}b$.*

Proof. The thesis follows immediately on applying $U^{(2)}_{\tilde{T},\tilde{S}}$ to $Ux = Vx + b$. □

Let $x \in \tilde{T}$ be any fixed solution of $Ax = b$ and define the "errors"

$$e_i = x_i - x. \tag{4.5}$$

Then from (4.2) and Lemma 4.6, the recursion satisfied by these errors is

$$e_i = Be_{i-1} = \cdots = B^ie_0, \tag{4.6}$$

where $B = U^{(2)}_{\tilde{T},\tilde{S}}V$ is the iteration matrix. From this relation we immediately have the following theorem.

Theorem 4.7. *Under the conditions of Lemma* 4.6, *the following three statements are equivalent*:

(i) *For any* x_0 *the sequence* $\{x_i\}$ *of* (4.2) *converges to a solution of* $Ax=b$ *in* $\tilde{T}$.
(ii) *For every* e_0 *the error sequence* $\{e_i\}$ *of* (4.6) *converges to a vector in* $N(A)\cap\tilde{T}$.
(iii) *The iteration matrix* $B=U^{(2)}_{\tilde{T},\tilde{S}}V$ *is semi-convergent, i.e.*,

$$\rho(B)\leqslant 1, \tag{4.7}$$

$$\lambda\in\sigma(B)\quad \textit{and}\quad |\lambda|=1\Rightarrow\lambda=1, \tag{4.8}$$

$$\mathrm{index}(I-B)\leqslant 1. \tag{4.9}$$

Proof. The equivalence of (i) and (ii) and (iii) are evident. □

When splitting (1.2) is a type I subproper splitting and the iteration matrix B is semi-convergent, then iteration (4.2) will converge to

$$(I-B)_g U^{(2)}_{\tilde{T},\tilde{S}}b+[I-(I-B)(I-B)_g]x_0. \tag{4.10}$$

For the convergence of type II subproper splitting, we choose the initial vector $x_0\in\tilde{T}$, then the iterative scheme (4.2) is equivalent to

$$x_{i+1}=x_i+U^{(2)}_{\tilde{T},\tilde{S}}(b-Ax_i) \tag{4.11}$$

for $i=0,1,2,\ldots$.

We now introduce a relaxation parameter β and consider

$$x_{i+1}=x_i+\beta U^{(2)}_{\tilde{T},\tilde{S}}(b-Ax_i). \tag{4.12}$$

Theorem 4.8. *Let* $A=U-V\in\mathbb{C}^{m\times n}$. *Let subspaces* $T,\tilde{T}\subseteq\mathbb{C}^n$ *and* $S,\tilde{S}\subseteq\mathbb{C}^m$ *be such that* $AT\oplus S=\mathbb{C}^m$ *and* $U\tilde{T}\oplus\tilde{S}=\mathbb{C}^m$. *Let splitting* (1.2) *be a type* II *subproper splitting of* A *and* $\beta>0$. *Then*

(i) *For any initial guess* $x_0\in T$, *the sequence* $\{x_i\}$ *of* (4.12) *converges to* $A^{(2)}_{T,S}b$ *if and only if*

$$\rho(P_T-\beta U^{(2)}_{\tilde{T},\tilde{S}}A)<1. \tag{4.13}$$

(ii) *Let* (4.13) *holds and* $x_i\to x$ *as* $i\to\infty$. *Then* $x=A^{(2)}_{T,S}b$ *has the representation*

$$x=\beta(P_{T^\perp}+\beta U^{(2)}_{\tilde{T},\tilde{S}}A)^{-1}U^{(2)}_{\tilde{T},\tilde{S}}b. \tag{4.14}$$

Moreover, if $\|P_T-\beta U^{(2)}_{\tilde{T},\tilde{S}}A\|=q<1$, *then* x_i *has the error estimation*

$$\|x_i-A^{(2)}_{T,S}b\|\leqslant q^i(\|x_0\|+\beta(1-q)^{-1}\|U^{(2)}_{\tilde{T},\tilde{S}}b\|). \tag{4.15}$$

(iii) *If all eigenvalues of* $U^{(2)}_{\tilde{T},\tilde{S}}A$ *are nonnegative, then* (4.13) *holds if and only if*

$$\mathrm{rank}(U^{(2)}_{\tilde{T},\tilde{S}}A)=\dim T\quad \textit{and}\quad 0<\beta<\frac{2}{\lambda_{\max}(U^{(2)}_{\tilde{T},\tilde{S}}A)}, \tag{4.16}$$

where $\lambda_{\max}(U^{(2)}_{\tilde{T},\tilde{S}}A)$ *is the largest eigenvalue of* $U^{(2)}_{\tilde{T},\tilde{S}}A$.

Proof. Denote $A^{(2)}_{T,S}$ by G for simplicity. It is easy to observe that $x=Gb$ satisfies $x=x+\beta U^{(2)}_{\tilde{T},\tilde{S}}(b-Ax)$. Hence from (4.12) we have

$$\begin{aligned} x_i - Gb &= x_{i-1} - Gb - \beta U^{(2)}_{\tilde{T},\tilde{S}}A(x_{i-1} - Gb) \\ &= (P_T - \beta U^{(2)}_{\tilde{T},\tilde{S}}A)(x_{i-1} - Gb) \quad (\text{by } x_i \in T). \end{aligned}$$

Using this relation repeatedly, it follows that

$$x_i - Gb = (P_T - \beta U^{(2)}_{\tilde{T},\tilde{S}}A)^i(x_0 - Gb). \tag{4.17}$$

(i) If (4.13) holds, then from (4.17) we have obviously $x_i \to Gb$ as $i \to \infty$ for any $x_0 \in T$. Conversely, suppose that $x_i \to Gb$ for every $x_0 \in T$. This x_0 is expressible as $x_0 = P_T z$, z arbitrary, and now (4.17) can be written as

$$\begin{aligned} x_i - Gb &= (P_T - \beta U^{(2)}_{\tilde{T},\tilde{S}}A)^i P_T(z - Gb) \\ &= (P_T - \beta U^{(2)}_{\tilde{T},\tilde{S}}AP_T)^i(z - Gb). \end{aligned} \tag{4.18}$$

Let the unitary matrix $Q = [Q_1, Q_2]$ be such that $T = R(Q_1)$ and $T^\perp = R(Q_2)$. Then $P_T = Q_1Q_1^*$. Therefore,

$$Q^*(P_T - \beta U^{(2)}_{\tilde{T},\tilde{S}}A)Q = \begin{bmatrix} I - \beta Q_1^* U^{(2)}_{\tilde{T},\tilde{S}}AQ_1 & * \\ 0 & 0 \end{bmatrix}, \tag{4.19}$$

$$Q^*(P_T - \beta U^{(2)}_{\tilde{T},\tilde{S}}AP_T)Q = \begin{bmatrix} I - \beta Q_1^* U^{(2)}_{\tilde{T},\tilde{S}}AQ_1 & 0 \\ 0 & 0 \end{bmatrix} \tag{4.20}$$

and

$$\begin{aligned} Q^*(x_i - Gb) &= (Q^*(P_T - \beta U^{(2)}_{\tilde{T},\tilde{S}}A)Q)^i Q^*(z - Gb) \\ &= \begin{bmatrix} (I - \beta Q_1^* U^{(2)}_{\tilde{T},\tilde{S}}AQ_1)^i \tilde{z} \\ 0 \end{bmatrix}, \quad \tilde{z} \text{ arbitrary.} \end{aligned} \tag{4.21}$$

Since $x_i \to Gb$ as $i \to \infty$, it follows from (4.21) that $\rho(I - \beta U^{(2)}_{\tilde{T},\tilde{S}}A) < 1$. In view of (4.19), we have $\rho(P_T - \beta U^{(2)}_{\tilde{T},\tilde{S}}A) < 1$.

(ii) Note that

$$\begin{aligned} x_i &= (P_T - \beta U^{(2)}_{\tilde{T},\tilde{S}}A)x_{i-1} + \beta U^{(2)}_{\tilde{T},\tilde{S}}b \\ &= [I + (P_T - \beta U^{(2)}_{\tilde{T},\tilde{S}}A) + (P_T - \beta U^{(2)}_{\tilde{T},\tilde{S}}A)^2 + \cdots + (P_T - \beta U^{(2)}_{\tilde{T},\tilde{S}}A)^{i-1}]\beta U^{(2)}_{\tilde{T},\tilde{S}}b \\ &\quad + (P_T - \beta U^{(2)}_{\tilde{T},\tilde{S}}A)^i x_0. \end{aligned} \tag{4.22}$$

If $\rho(P_T - \beta U^{(2)}_{\tilde{T},\tilde{S}}A) < 1$, then one has

$$x_i = \beta[I - (P_T - \beta U^{(2)}_{\tilde{T},\tilde{S}}A)]^{-1}U^{(2)}_{\tilde{T},\tilde{S}}b = \beta(P_{T^\perp} + \beta U^{(2)}_{\tilde{T},\tilde{S}}A)^{-1}U^{(2)}_{\tilde{T},\tilde{S}}b.$$

Furthermore, by (4.22) and (4.14), we get

$$x_i - Gb = -(P_T - \beta U^{(2)}_{\tilde{T},\tilde{S}}A)^i[I + (P_T - \beta U^{(2)}_{\tilde{T},\tilde{S}}A) + (P_T - \beta U^{(2)}_{\tilde{T},\tilde{S}}A)^2 + \cdots]\beta U^{(2)}_{\tilde{T},\tilde{S}}b$$
$$+ (P_T - \beta U^{(2)}_{\tilde{T},\tilde{S}}A)^i x_0$$

giving (4.15) if $\|P_T - \beta U^{(2)}_{\tilde{T},\tilde{S}}A\| = q < 1$.

(iii) For the proof of (iii) see [8, Theorem 2.1]. □

We should note that Chen proved an analogous result for the iterative scheme similar to (4.12) in the case of $b \in AT$ in [9]. However, in Theorem 4.7, we do not assume that (1.1) is consistent.

The following convergence theorem for type II subproper splitting is the direct result of Theorem 4.8.

Theorem 4.9. *Let $A = U - V \in \mathbb{C}^{m\times n}$. Suppose subspaces $T, \tilde{T} \subseteq \mathbb{C}^n$ and $S, \tilde{S} \subseteq \mathbb{C}^m$ be such that $AT \oplus S = \mathbb{C}^m$ and $U\tilde{T} \oplus \tilde{S} = \mathbb{C}^m$. Let* (1.2) *be a type* II *subproper splitting of A. Then*

(i) *If $\rho(P_T - U^{(2)}_{\tilde{T},\tilde{S}}A) < 1$, then the sequence $\{x_i\}$ of* (4.2) *converges to $A^{(2)}_{T,S}b$ provided $x_0 \in \tilde{T}$. Moreover, if $\|P_T - U^{(2)}_{\tilde{T},\tilde{S}}A\| = q < 1$, then x_i has the error estimation*

$$\|x_i - A^{(2)}_{T,S}b\| \leqslant q^i(\|x_0\| + (1-q)^{-1}\|U^{(2)}_{\tilde{T},\tilde{S}}b\|). \tag{4.23}$$

(ii) *Suppose all the eigenvalues of $U^{(2)}_{\tilde{T},\tilde{S}}A$ are nonnegative and $0 < \lambda_{\max}(U^{(2)}_{\tilde{T},\tilde{S}}A) < 2$. If splitting* (1.2) *is proper, then $\rho(P_T - U^{(2)}_{\tilde{T},\tilde{S}}A) < 1$ and thus the sequence of iterations of* (4.2) *converges to $A^{(2)}_{T,S}b$ for any $x_0 \in T$.*

Acknowledgements

The authors would like to thank the referee for his/her comments on the presentation of this paper.

References

[1] A. Ben-Israel, On error bounds for generalized inverse, SIAM J. Numer. Anal. 3 (1966) 585–592.
[2] A. Ben-Israel, A cramer rule for least-norm solution of consistent linear equations, Linear Algebra Appl. 43 (1982) 223–226.
[3] A. Ben-Israel, T.N.E. Greville, Generalized Inverses Theory and Applications, Wiley, New York, 1974.
[4] A. Berman, M. Neumann, Proper splittings of rectangular matrices, SIAM J. Appl. Math. 31 (1976) 307–312.
[5] A. Berman, R.J. Plemmons, Cones and iterative methods for best least squares solutions of linear systems, SIAM J. Numer. Anal. 11 (1974) 145–154.
[6] A. Berman, R.J. Plemmons, Nonnegative Matrices in Mathematical Sciences, Academic Press, New York, 1979.
[7] Y. Chen, A cramer rule for solution of the general restricted linear equations, Linear Multilinear Algebra 34 (1993) 177–186.
[8] Y. Chen, Iterative methods for computing the generalized inverse $A^{(2)}_{T,S}$ of a matrix A, Appl. Math. Comput. 75 (1996) 207–222.
[9] Y. Chen, Iterative methods for solving restricted linear equations, Appl. Math. Comput. 86 (1997) 171–184.
[10] M. Eiermann, I. Marek, W. Niethammer, On the solution of singular systems of algebraic equations by semiiterative methods, Numer. Math. 53 (1988) 265–283.

[11] A.J. Getson, F.C. Hsuan, {2}-Inverses and Their Statistical Applications, Lecture Notes in Statistics, Springer, Berlin, 1988.
[12] M. Hanke, M. Neumann, Preconditionings and splittings for rectangular systems, Numer. Math. 57 (1990) 85–95.
[13] F. Husen, P. Langenberg, A.J. Getson, The {2}-inverse with applications to statistics, Linear Algebra Appl. 70 (1985) 241–248.
[14] V.N. Joshi, A note on the solution of rectangular linear systems by iterations, SIAM Rev. 12 (1970) 463–466.
[15] H.B. Keller, On the solution of singular and semidefinite linear systems by iterations, SIAM J. Numer Anal. 2 (1965) 281–290.
[16] O.L. Mangasarian, A convergent splitting for matrices, Numer. Math. 15 (1970) 351–353.
[17] M.Z. Nashed, Generalized Inverses and Applications, Academic Press, New York, 1976.
[18] M.Z. Nashed, X. Chen, Convergence of Newton-like methods for singular operator equations using outer inverse, Numer. Math. 66 (1993) 235–257.
[19] M. Neumann, Subproper splitting for rectangular matrices, Linear Algebra Appl. 14 (1976) 41–51.
[20] J.M. Ortega, W.C. Rheinboldt, Monotone iterations for nonlinear equations with applications to Gauss–Seidel methods, SIAM J. Numer. Anal. 4 (1967) 171–190.
[21] R.J. Plemmons, Regular splittings and discrete Neumann problem, Numer. Math. 25 (1976) 153–161.
[22] J.S. Vandergraft, Applications of partial orderings to the study of positive definiteness, monotonicity, and convergence of iterative methods for linear systems, SIAM J. Numer. Anal. 4 (1972) 97–104.
[23] R.S. Varga, Matrix Iterative Analysis, Prentice-Hall, Englewood Cliffs, NJ, 1971.
[24] G.C. Verghese, A cramer rule for least-norm least-squares-error solution of inconsistent linear equations, Linear Algebra Appl. 48 (1982) 315–316.
[25] G. Wang, A cramer rule for minimum-norm (T) least-squares (S) solution of inconsistent linear equations, Linear Algebra Appl. 74 (1986) 213–218.
[26] G. Wang, A cramer rule for finding the solution of restricted linear equations, Linear Algebra Appl. 116 (1989) 27–34.
[27] P.-Å. Wedin, Perturbation results and condition numbers for outer inverses and especially for projections, Technical Report UMINF 124.85, S-901 87, Inst. of Info. Proc., University of Umeå, Umeå, Sweden, 1985.
[28] Y. Wei, G. Wang, The perturbation theory for the Drazin inverse and its application, Linear Algebra Appl. 258 (1997) 179–186.
[29] Y. Wei, Index splitting for the Drazin inverse and the singular linear systems, Appl. Math. Comput. 95 (1998) 115–124.
[30] Y. Wei, A characterization and representation for the generalized inverse $A_{T,S}^{(2)}$ and its applications, Linear Algebra Appl. 280 (1998) 87–96.
[31] Y. Wei, On the perturbation of the group inverse and oblique projection, Appl. Math. Comput. 98 (1999) 29–42.
[32] Y. Wei, Solving singular linear systems and generalized inverses, Ph.D. Thesis, Institute of Mathematics, Fudan University, Shanghai, China, March 1997.
[33] H.J. Werner, On extension of Cramer's rule for solution of restricted linear systems, Linear Multilinear Algebra 15 (1984) 319–330.
[34] Z. Woźnicki, Nonnegative splitting theory, Jpn J. Indust. Appl. Math. 11 (1994) 289–342.