A RECURSIVE ALGORITHM FOR COMPUTING THE WEIGHTED MOORE-PENROSE INVERSE A^+_{MN}*

WANG GUO-RONG (王国荣)
(Shanghai Normal University, Shanghai, China)

CHEN YONG-LIN (陈永林)
(Nanjing Normal University, Nanjing, China)

Abstract

In this paper, we give a recursive algorithm for computing the weighted Moore-Penrose inverse A^+_{MN}. This method is a generalization of Greville's method for computing Moore-Penrose inverse A^+, and the technique of its proof is new. This method suits the weighted least-squares problem.

§ 1. Introduction

Throughout this paper, let M and N be positive definite matrices of order m and n respectively. Let $A\in C^{m\times n}$. Then there is a unique matrix $X\in C^{n\times m}$ satisfying

$$AXA=A,\ XAX=X,\ (MAX)^*=MAX,\ (NXA)^*=NXA. \tag{1.1}$$

This X is called the weighted $M-P$ inverse of A, and is denoted by $X=A^+_{MN}$. Especially, when $M=I_m$ and $N=I_n$, the matrix X satisfying (1.1) is called the $M-P$ inverse of A, and is denoted by $X=A^+$, i.e., $A^+=A^+_{I_mI_n}$.

In 1960, A famous recursive method for computing the $M-P$ inverse of A was given by Greville[1].

Let $A_k\in C^{m\times k}$ be the submatrix of $A\in C^{m\times n}$ consisting of its first k columns. For $k=2, \cdots, n$ the matrix A_k is partitioned as

$$A_k=[A_{k-1}\ a_k],$$

where a_k is the k-th column of A. For $k=2, \cdots, n$ the vectors d_k and c_k are defined by $d_k=A^+_{k-1}a_k$ and $c_k=a_k-A_{k-1}d_k=(I-A_{k-1}A^+_{k-1})a_k$. Then, the $M-P$ inverse of A_k is

$$A^+_k=\begin{pmatrix} A^+_{k-1}-d_kb^*_k \\ b^*_k \end{pmatrix},$$

where

$$b^*_k=\begin{cases}(c^*_kc_k)^{-1}c^*_k & \text{if } c_k\neq 0,\\ (1+d^*_kd_k)^{-1}d^*_kA^+_{k-1} & \text{if } c_k=0.\end{cases}$$

In [2, 3, 4] three different proofs for Greville's method were presented. Greville's method is natural in some applications, for example, the least-squares polynomial approximation problem, regression analysis, etc[5].

There are many formulas for computing the weighted $M-P$ inverse A^+_{NM}[6], but they are very complex. In this paper, we will give a recursive algorithm for computing A^+_{MN}. This method is a generalization of Greville's method, and the technique of its proof is new. This method suits the weighted least-squares problem.

* Received February 26, 1985.

§ 2. Preliminaries

In this section we will give three lemmas.

Lemma 1. *Let* $A \in C_r^{m\times n}$, $X = A_{MN}^+$; then

(i) $R(X) = N^{-1}R(A^*)$, $N(X) = M^{-1}N(A^*)$,
$R(X^*) = MR(A)$, $N(X^*) = MN(A)$,
$AX = P_{R(A),M^{-1}N(A^*)}$, $XA = P_{N^{-1}R(A^*),N(A)}$.

(ii) $AX = A(A^*MA)^+A^*M$, $XA = N^{-1}A^*(AN^{-1}A^*)^+A$.

(iii) $(A_{MN}^+)^* = (A^*)^+_{N^{-1}M^{-1}}$.

(iv) *Let* $U \in C_{m-r}^{m\times(m-r)}$ *and* $V \in C_{n-r}^{n\times(n-r)}$ *such that* $A^*U = 0$ *and* $AV = 0$; *then*
$I - XA = V(V^*NV)^{-1}V^*N$, $I - AX = M^{-1}U(U^*M^{-1}U)^{-1}U^*$.

Proof. (i) See [1, chap. 3].

(ii) and (iii) See [2, chap. 3].

(iv) By hypothesis, $R(U) = N(A^*)$ and $R(V) = N(A)$. Since V^*NV is p.d., inverse $(V^*NV)^{-1}$ exists and is also p.d. Set $V(V^*NV)^{-1}V^*N = E$. Then E is idempotent, and $R(E) = R(V) = N(A)$ and $N(E) = N(V^*N) = N^{-1}N(V^*) = N^{-1}R(A^*)$. Hence $V(V^*NV)^{-1}V^*N = P_{N(A),N^{-1}R(A^*)} = I - P_{N^{-1}R(A^*),N(A)} = I - XA$. A similar argument shows $I - AX = M^{-1}U(U^*M^{-1}U)^{-1}U^*$.

Lemma 2. *Let* $A \in C_r^{m\times n}$, $U \in C_{m-r}^{m\times(m-r)}$ *and* $V \in C_{n-r}^{n\times(n-r)}$ *such that*

$$A^*U = 0 \text{ and } AV = 0. \tag{2.1}$$

Then

(i)
$$\begin{pmatrix} A & M^{-1}U \\ V^*N & 0 \end{pmatrix} \text{ is nonsingular.} \tag{2.2}$$

(ii)
$$\begin{pmatrix} A & M^{-1}U \\ V^*N & 0 \end{pmatrix}^{-1} = \begin{pmatrix} A_{MN}^+ & V(V^*NV)^{-1} \\ (U^*M^{-1}U)^{-1}U^* & 0 \end{pmatrix}. \tag{2.3}$$

Proof. Set $X = A_{MN}^+$. From Lemma 1, we have

$$AX + M^{-1}U(U^*M^{-1}U)^{-1}U^* = AX + (I - AX) = I \tag{2.4}$$

and from (2.1)

$$AV(V^*NV)^{-1} = 0. \tag{2.5}$$

Since $V^*NXA = V^*(NXA)^* = V^*A^*X^*N = 0$,

$$V^*NX = V^*NXAX = 0 \tag{2.6}$$

and, obviously

$$V^*NV(V^*NV)^{-1} = I. \tag{2.7}$$

Using (2.4)—(2.7), we may obtain (2.2) and (2.3) immediately.

Lemma 3. *Let* $\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$ *be a partitioned matrix which is nonsingular, and let the submatrix* A_{11} *also be nonsingular. Then*

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}^{-1} = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}, \tag{2.8}$$

where

$$B_{11} = A_{11}^{-1} + A_{11}^{-1}A_{12}B_{22}A_{21}A_{11}^{-1}, \tag{2.9}$$

$$B_{12} = -A_{11}^{-1}A_{12}B_{22}, \tag{2.10}$$

$$B_{21} = -B_{22}A_{21}A_{11}^{-1}, \tag{2.11}$$

$$B_{22} = (A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1}. \tag{2.12}$$

§ 3. Main Results

In this section we will give the recursive algorithm for computing A^+_{MN}.

Theorem 1. *Let $A \in C^{m\times n}$ and A_k be the submatrix of A consisting of its first k columns. For $k=2, \cdots, n$, the matrix A_k is partitioned as*

$$A_k = [A_{k-1}\ a_k], \tag{3.1}$$

where a_k is the k-th column of A, and the matrix $N_k \in C^{k\times k}$ is the leading principal submatrix of N, and N_k is partitioned as

$$N_k = \begin{pmatrix} N_{k-1} & l_k \\ l_k^* & n_{kk} \end{pmatrix}. \tag{3.2}$$

Let the matrices X_{k-1} and X_k be defined by

$$X_{k-1} = (A_{k-1})^+_{MN_{k-1}}, \quad X_k = (A_k)^+_{MN_k}, \tag{3.3}$$

the vectors d_k and c_k be defined by

$$d_k = X_{k-1}a_k, \tag{3.4}$$

$$c_k = a_k - A_{k-1}d_k \tag{3.5}$$

$$= (I - A_{k-1}X_{k-1})a_k. \tag{3.6}$$

Then

$$X_k = \begin{pmatrix} X_{k-1} - (d_k + (I - X_{k-1}A_{k-1})N_{k-1}^{-1}l_k)b_k^* \\ b_k^* \end{pmatrix}, \tag{3.7}$$

where

$$b_k^* = \begin{cases} (c_k^* M c_k)^{-1} c_k^* M & \text{if } c_k \neq 0, \\ \delta_k^{-1}(d_k^* N_{k-1} - l_k^*) X_{k-1} & \text{if } c_k = 0, \end{cases} \tag{3.8}$$

and

$$\delta_k = n_{kk} + d_k^* N_k d_k - (d_k^* l_k + l_k^* d_k) - l_k^* (I - X_{k-1}A_{k-1}) N_{k-1}^{-1} l_k \tag{3.9}$$

is a positive real number.

Proof. It is obvious that

$$X_1 = \begin{cases} 0 & \text{if } a_1 = 0, \\ (a_1^* M a_1)^{-1} a_1^* M & \text{if } a_1 \neq 0. \end{cases} \tag{3.10}$$

Now we distinguish between two cases according as c_k is or is not 0.

Case I ($c_k = 0$).

Let the columns of U_{k-1} and V_{k-1} be bases for $N(A_{k-1}^*)$ and $N(A_{k-1})$ respectively; then

$$A_{k-1}^* U_{k-1} = 0, \tag{3.11}$$

$$A_{k-1} V_{k-1} = 0. \tag{3.12}$$

By Lemma 2, we have

$$\begin{pmatrix} A_{k-1} & M^{-1}U_{k-1} \\ V_{k-1}^* N_{k-1} & 0 \end{pmatrix}^{-1} = \begin{pmatrix} X_{k-1} & V_{k-1}(V_{k-1}^* N_{k-1} V_{k-1})^{-1} \\ (U_{k-1}^* M^{-1} U_{k-1})^{-1} U_{k-1}^* & 0 \end{pmatrix}. \tag{3.13}$$

From $c_k=0$, we have

$$A_{k-1}d_k=a_k, \tag{3.14}$$

$$R(A_{k-1})=R(A_k) \tag{3.15}$$

and

$$N(A_{k-1}^*)=N(A_k^*). \tag{3.16}$$

From (3.12) and (3.14), we have

$$A_k^*U_{k-1}=0. \tag{3.17}$$

Set

$$V_k=\begin{pmatrix} V_{k-1} & -d_k \\ 0 & 1 \end{pmatrix}. \tag{3.18}$$

Then

$$A_kV_k=0. \tag{3.19}$$

This shows the columns of U_{k-1} and V_k are bases for $N(A_k^*)$ and $N(A_k)$ respectively. By Lemma 2 again, we have

$$\begin{pmatrix} A_k & M^{-1}U_{k-1} \\ V_k^*N_k & 0 \end{pmatrix}^{-1}=\begin{pmatrix} X_k & V_k(V_k^*N_kV_k)^{-1} \\ (U_{k-1}^*M^{-1}U_{k-1})^{-1}U_{k-1}^* & 0 \end{pmatrix}. \tag{3.20}$$

Since

$$V_k^*N_k=\begin{pmatrix} V_{k-1}^*N_{k-1} & V_{k-1}^*l_k \\ l_k^*-d_k^*N_{k-1} & n_{kk}-d_k^*l_k \end{pmatrix}, \tag{3.21}$$

$$\left(\begin{array}{cc|c} A_{k-1} & a_k & M^{-1}U_{k-1} \\ \hline V_{k-1}^*N_{k-1} & V_{k-1}^*l_k & 0 \\ l_k^*-d_k^*N_{k-1} & n_{kk}-d_k^*l_k & 0 \end{array}\right)^{-1}=\left(\begin{array}{c|cc} X_k & * & * \\ & * & * \\ \hline * & 0 & 0 \end{array}\right). \tag{3.22}$$

And set

$$K=\left(\begin{array}{cc|c} A_{k-1} & M^{-1}U_{k-1} & a_k \\ V_k^*N_{k-1} & 0 & V_{k-1}^*l_k \\ \hline l_k^*-d_k^*N_{k-1} & 0 & n_{kk}-d_k^*l_k \end{array}\right) \tag{3.23}$$

and

$$P=\begin{pmatrix} I & 0 & 0 \\ 0 & 0 & 1 \\ 0 & I & 0 \end{pmatrix}, \tag{3.24}$$

where the two submatrices I are identity matrices of a suitable order, and the partition of P is conformable with that of K. From (3.20)—(3.24), we have

$$P^TK^{-1}=(KP)^{-1}=\left(\begin{array}{c|cc} X_k & * & * \\ & * & * \\ \hline * & 0 & 0 \end{array}\right). \tag{3.25}$$

Thus, if the matrix K^{-1} is partitioned as of 3×3 block form, then X_k is a 2×1 block submatrix that lies in the upper left-hand corner of P^TK^{-1}, that is, X_k consists of the first and third block in the first block column of K^{-1}. Now, computing the related blocks of K^{-1} by Lemma 3 and (3.20), we may obtain

$$(K^{-1})_{11}=\left(\begin{array}{c|c} X_{k-1}+\delta_k^{-1}d_k(l_k^*-d_k^*N_{k-1})X_{k-1} & * \\ +\delta_k^{-1}V_{k-1}(V_{k-1}^*N_{k-1}V_{k-1})^{-1}V_{k-1}^*l_k(l_k^*-d_k^*N_{k-1})X_{k-1} & \\ \hline * & * \end{array}\right), \tag{3.26}$$

$$(K^{-1})_{21}=[-\delta_k^{-1}(l_k^*-d_k^*N_{k-1})X_{k-1}\,|\,*], \tag{3.27}$$

where

$$\delta_k=1/(K^{-1})_{22}=r_{kk}-d_k^*l_k-(l_k^*-d_k^*N_{k-1})(X_{k-1}a_k+V_{k-1}(V_{k-1}^*N_{k-1}V_{k-1})^{-1}V_{k-1}^*l_k). \tag{3.28}$$

By Lemma 1,

$$V_{k-1}(V_{k-1}^*N_{k-1}V_{k-1})^{-1}V_{k-1}^*=(I-X_{k-1}A_{k-1})N_{k-1}^{-1} \tag{3.29}$$

and since

$$d_k^*N_{k-1}X_{k-1}A_{k-1}=d_k^*A_{k-1}^*X_{k-1}^*N_{k-1}=a_k^*X_{k-1}^*N_{k-1}=d_k^*N_{k-1}, \tag{3.30}$$

$$d_k^*N_{k-1}(I-X_{k-1}A_{k-1})=0. \tag{3.31}$$

Thus, the expression of δ_k may be simplified to (2.9), and

$$(K^{-1})_{11}=\left(\begin{array}{c|c} X_{k-1}+(d_k+(I-X_{k-1}A_{k-1})N_{k-1}^{-1}l_k)\delta_k^{-1}(l_k^*-d_k^*N_{k-1})X_{k-1} & * \\ \hline * & * \end{array}\right). \tag{3.32}$$

From (3.32) and (3.27), we may obtain

$$X_k=\begin{pmatrix} X_{k-1}+(d_k+(I-X_{k-1}A_{k-1})N_{k-1}^{-1}l_k)\delta_k^{-1}(l_k^*-d_k^*N_{k-1})X_{k-1} \\ -\delta_k^{-1}(l_k^*-d_k^*N_{k-1})X_{k-1} \end{pmatrix}. \tag{3.33}$$

Now we prove δ_k is a positive real number. Since

$$V_k^*N_kV_k=\begin{pmatrix} V_{k-1}^*N_{k-1}V_{k-1} & V_{k-1}^*(l_k-N_{k-1}d_k) \\ (l_k^*-d_k^*N_{k-1})V_{k-1} & t_k \end{pmatrix}, \tag{3.34}$$

where $t_k=n_{kk}+d_k^*N_{k-1}d_k-(d_k^*l_k+l_k^*d_k)$, is p.d., $(V_k^*N_kV_k)^{-1}$ is also p.d. If the final diagonal element of $(V_k^*N_kV_k)^{-1}$ is r_k, then $r_k>0$. But, by Lemma 3, (3.29) and (3.31), and noticing the following equality

$$(I-X_{k-1}A_{k-1})d_k=0 \tag{3.35}$$

we have

$$\begin{aligned} r_k^{-1}&=t_k-l_k^*(I-X_{k-1}A_{k-1})N_{k-1}^{-1}l_k \\ &=n_{kk}+d_k^*N_{k-1}d_k-(d_k^*l_k+l_k^*d_k)-l_k^*(I-X_{k-1}A_{k-1})N_{k-1}^{-1}l_k. \end{aligned} \tag{3.36}$$

Obviously $\delta_k=r_k^{-1}>0$.

Case II ($c_k\neq 0$).

Let the columns of U_{k-1} and V_{k-1} be bases for $N(A_{k-1}^*)$ and $N(A_{k-1})$ respectively; then

$$\begin{cases} A_{k-1}^*U_{k-1}=0, \\ A_{k-1}V_{k-1}=0. \end{cases} \tag{3.37}$$

By Lemma 2, we have

$$\begin{pmatrix} A_{k-1} & M^{-1}U_{k-1} \\ V_{k-1}^*N_{k-1} & 0 \end{pmatrix}^{-1}=\begin{pmatrix} X_{k-1} & V_{k-1}(V_{k-1}^*N_{k-1}V_{k-1})^{-1} \\ (U_{k-1}^*M^{-1}U_{k-1})^{-1}U_{k-1}^* & 0 \end{pmatrix}. \tag{3.38}$$

Since $c_k \neq 0$ is equivalent to $a_k \notin R(A_{k-1})$, we have

$$\operatorname{rank}(A_k) = \operatorname{rank}(A_{k-1}) + 1, \tag{3.39}$$

$$\dim N(A_{k-1}^*) = \dim N(A_k^*) + 1, \tag{3.40}$$

and

$$\dim N(A_k) = \dim N(A_{k-1}). \tag{3.41}$$

Set

$$V_k = \begin{pmatrix} V_{k-1} \\ 0 \end{pmatrix}. \tag{3.42}$$

Then

$$A_k V_k = 0. \tag{3.43}$$

This shows the columns of V_k are a basis for $N(A_k)$. Since

$$c_k^* M = c_k^*(I - A_{k-1}X_{k-1})^* M = a_k^* M (I - A_{k-1}X_{k-1})$$

we have

$$c_k^* M A_{k-1} = 0 \tag{3.44}$$

and

$$c_k^* M a_k = a_k^* M (I - A_{k-1}X_{k-1})(I - A_{k-1}X_{k-1}) a_k = c_k^* M c_k > 0. \tag{3.45}$$

Thus

$$A_k^* M c_k = \begin{pmatrix} A_{k-1}^* \\ a_k^* \end{pmatrix} M c_k = \begin{pmatrix} 0 \\ a_k^* M c_k \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{3.46}$$

This means

$$M c_k \overline{\in} N(A_k^*) \tag{3.46$'$}$$

and Mc_k can be taken as a vector in a basis of $N(A_{k-1})$. Let the columns of U_k be a basis for $N(A_k^*)$ such that

$$U_{k-1} = [U_k \mid M c_k]. \tag{3.47}$$

Then

$$A_k^* U_k = 0 \tag{3.48}$$

i.e.,

$$A_{k-1}^* U_k = 0 \text{ and } a_k^* U_k = 0. \tag{3.49}$$

Hence

$$U_k^* c_k = U_k^* a_k - U_k^* A_{k-1} X_{k-1} a_k = 0. \tag{3.50}$$

From (3.43), (3.48) and Lemma 2, we have

$$\begin{pmatrix} A_k & M^{-1}U_k \\ V_k^* N_k & 0 \end{pmatrix}^{-1} = \begin{pmatrix} X_k & V_k(V_k^* N_k V_k)^{-1} \\ (U_k^* M^{-1} U_k)^{-1} U_k^* & 0 \end{pmatrix}. \tag{3.51}$$

By

$$M^{-1}U_{k-1} = [M^{-1}U_k \mid c_k] \text{ and } V_k^* N_k = [V_{k-1}^* N_{k-1} \mid V_{k-1}^* l_k] \tag{3.52}$$

(3.38) and (3.51) become

$$\left(\begin{array}{c|cc} A_{k-1} & M^{-1}U_k & c_k \\ \hline V_{k-1}^* N_{k-1} & 0 & 0 \end{array}\right)^{-1} = \left(\begin{array}{c|c} X_{k-1} & * \\ \hline * & 0 \\ * & 0 \end{array}\right) \tag{3.53}$$

and

$$\left(\begin{array}{cc|c} A_{k-1} & a_k & M^{-1}U_k \\ \hline V_{k-1}^* N_{k-1} & V_{k-1}^* l_k & 0 \end{array}\right)^{-1} = \left(\begin{array}{c|c} X_k & * \\ & * \\ \hline * & 0 \end{array}\right). \tag{3.54}$$

Set

$$G=\left(\begin{array}{ccc|c} A_{k-1} & a_k & M^{-1}U_k & c_k \\ V_{k-1}^* N_{k-1} & V_{k-1}^* l_k & 0 & 0 \\ \hline 0 & 0 & 0 & 1 \end{array}\right). \tag{3.55}$$

Then G is a nonsingular block triangular matrix, and

$$G^{-1}=\left(\begin{array}{c|c|c} X_k & \begin{matrix} * \\ * \end{matrix} & \\ \cline{1-2} * & 0 & * \\ \hline \multicolumn{2}{c|}{0} & 1 \end{array}\right). \tag{3.56}$$

Again set

$$F=\left(\begin{array}{ccc|c} A_{k-1} & M^{-1}U_k & c_k & a_k \\ V_{k-1}^* N_{k-1} & 0 & 0 & V_{k-1}^* l_k \\ \hline 0 & 0 & 1 & 0 \end{array}\right) \tag{3.57}$$

and

$$P_1=\begin{pmatrix} I & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & I & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},\quad P_2=\begin{pmatrix} I & 0 & 0 & 0 \\ 0 & I & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \tag{3.58}$$

where the submatrices I are identity matrices of a suitable order, such that

$$GP_1P_2=F. \tag{3.59}$$

Thus

$$P_1P_2F^{-1}=G^{-1}. \tag{3.60}$$

If F^{-1} is partitioned as of 4×3 block form

$$F^{-1}=\left(\begin{array}{cc|c} X_{k-1} & * & * \\ * & 0 & * \\ * & 0 & * \\ \hline * & * & * \end{array}\right), \tag{3.61}$$

then, from (3.60), (3.56) and (3.61), the first and forth block in the first block column of F^{-1} form X_k. By Lemma 3, (3.53) and (3.57), we may obtain

$$(F^{-1})_{11}=\left(\begin{array}{c|c} \begin{matrix} X_{k-1}+\beta^{-1}d_k e^T(U_{k-1}^* M^{-1}U_{k-1})^{-1}U_{k-1}^* \\ +\beta^{-1}V_{k-1}(V_{k-1}^* N_{k-1}V_{k-1})^{-1}V_{k-1}^* l_k e^T(U_{k-1}^* M^{-1}U_{k-1})^{-1}U_{k-1}^* \end{matrix} & * \\ \hline * & 0 \\ * & 0 \end{array}\right), \tag{3.62}$$

$$(F^{-1})_{21}=-\beta^{-1}[e^T(U_{k-1}^* M^{-1}U_{k-1})^{-1}U_{k-1}^* | 0], \tag{3.63}$$

where $e^T=[0, \cdots, 0, 1]$ whose dimensions equal the number of the columns of U_{k-1}, and

$$\beta=1/(F^{-1})_{22}=-e^T(U_{k-1}^* M^{-1}U_{k-1})^{-1}U_{k-1}^* a_k. \tag{3.64}$$

From (3.47) and (3.50)

$$(U_{k-1}^*M^{-1}U_{k-1})^{-1}=\begin{pmatrix}(U_k^*M^{-1}U_k)^{-1} & 0\\ 0 & (c_k^*Mc_k)^{-1}\end{pmatrix}. \tag{3.65}$$

Thus

$$e^T(U_{k-1}^*M^{-1}U_{k-1})^{-1}U_{k-1}^*=(c_kMc_k)^{-1}c_k^*M \tag{3.66}$$

and

$$\beta=-(c_k^*Mc_k)^{-1}c_k^*Ma_k=-1. \tag{3.67}$$

By (3.66), (3.67) and (3.29), we have

$$X_k=\begin{pmatrix}X_{k-1}-(d_k+(I-X_{k-1}A_{k-1})N_{k-1}^{-1}l_k)(c_k^*Mc_k)^{-1}c_k^*M\\ (c_k^*Mc_k)^{-1}c_k^*M\end{pmatrix}. \tag{3.68}$$

This completes the proof.

Note. If $M=I_m$ and $N=I_n$, the result of this theorem agrees with Greville's method. Thus, we have provided a new proof for Greville's method.

The practical methods for computing A_{MN}^+ and N_k^{-1} are as follows:

Algorithm 1. Let $A\in C^{m\times n}$, M and N be p.d. matrices of order m and n respectively. This algorithm computes the weighted $M-P$ inverse A_{MN}^+.

(1) $A_1=a_1$.

(2) If $a_1=0$, then $X_1=(a_1^*Ma_1)^{-1}a_1^*M$; else $X_1=0$.

(3) For $k=2, \cdots, n$

1) $d_k=X_{k-1}a_k$,

2) $c_k=a_k-A_{k-1}d_k$,

3) If $c_k\neq 0$ then $b_k^*=(c_k^*Mc_k)^{-1}c_k^*M$, goto (3.6),

4) $\delta_k=n_{kk}+d_k^*N_{k-1}d_k-(d_k^*l_k+l_k^*d_k)-l_k^*(I-X_{k-1}A_{k-1})N_{k-1}^{-1}l_k$,

5) $b_k^*=\delta_k^{-1}(d_k^*N_{k-1}-l_k^*)X_{k-1}$,

6) $X_k=\begin{pmatrix}X_{k-1}-(d_k+(I-X_{k-1}A_{k-1})N_{k-1}^{-1}l_k)b_k^*\\ b_k^*\end{pmatrix}$.

(4) $A_{MN}^+=X_n$.

Algorithm 2. Let $N_k=\begin{pmatrix}N_{k-1} & l_k\\ l_k^* & n_{kk}\end{pmatrix}\in C^{k\times k}$ be the leading principal submatrix of p.d. matrix N. This algorithm computes the inverse matrix $N_k^{-1}=\begin{pmatrix}E_{k-1} & f_k\\ f_k^* & g_{kk}\end{pmatrix}$.

(1) $N_1^{-1}=n_{11}^{-1}$.

(2) For $k=2, \cdots, n$

1) $g_{kk}=(n_{kk}-l_k^*N_{k-1}^{-1}l_k)^{-1}$,

2) $f_k=-g_{kk}N_{k-1}^{-1}l_k$,

3) $E_{k-1}=N_{k-1}^{-1}+g_{kk}^{-1}f_kf_k^*$,

4) $N_k^{-1}=\begin{pmatrix}E_{k-1} & f_k\\ f_k^* & g_{kk}\end{pmatrix}$.

§ 4. Applications

Let $A\in C^{m\times n}$, $y\in C^m$, $m>n$. If the linear system

$$Ax=y \tag{4.1}$$

is inconsistent, then the minimum-N-norm least-M-squares solution of (4.1) is

$$\hat{x}=A_{MN}^+y. \tag{4.2}$$

Our algorithm for computing A^+_{MN} may be adapted for the computation of $A^+_{MN}y$, for any $y\in C^m$, without computing A^+_{MN}. This is done as follows:

Let

$$\widetilde{A}=[A\,|\,y]\in C^{m\times(n+1)}. \tag{4.3}$$

Then (3.7)—(3.8) give

$$X_k\widetilde{A}=\begin{pmatrix} X_{k-1}\widetilde{A}-(d_k+(I-X_{k-1}A_{k-1})N^{-1}_{k-1}l_k)b^*_k\widetilde{A} \\ b^*_k\widetilde{A}\end{pmatrix}, \tag{4.4}$$

where

$$b^*_k\widetilde{A}=\begin{cases}(c^*_kM^{-1}c_k)^{-1}c^*_kM\widetilde{A} & \text{if } c_k\neq 0, \qquad (4.5)\\ \delta_k^{-1}(d^*_kN_{k-1}-l^*_k)X_{k-1}\widetilde{A} & \text{if } c_k=0, \qquad (4.6)\end{cases}$$

and

$$\delta_k=n_{kk}+d^*_kN_{k-1}d_k-(d^*_kl_k+l^*_kd_k)-l^*_k(I-X_{k-1}A_{k-1})N^{-1}_{k-1}l_k. \tag{4.7}$$

As the $(n+1)$-th column of $X_n\widetilde{A}$ equals $A^+_{MN}y$, we may obtain a recursive formula for computing $X_k\widetilde{A}$, from (4.4)—(4.6).

Firstly, if $c_k\neq 0$, then $c^*_kMc_k=c^*_kMa_k$ from (3.45). Hence $c^*_kMc_k$ is the k-th element of the row vector $c^*_kM\widetilde{A}$. Secondly, $X_{k-1}A_{k-1}$ is the submatrix of $X_{k-1}\widetilde{A}$ consisting of its first $k-1$ columns, i.e., $X_{k-1}A_{k-1}=X_{k-1}\widetilde{A}\begin{pmatrix}I_{k-1}\\0\end{pmatrix}$. Finally, d_k is the k-th column of $X_{k-1}\widetilde{A}$.

Algorithm 3. Let $A\in C^{m\times n}$. For any $y\in C^m$, this algorithm computes $A^+_{MN}y$ without computing A^+_{MN}.

(1) If $a_1\neq 0$, then $X_1\widetilde{A}=(a^*_1Ma_1)^{-1}a^*_1M\widetilde{A}$; else $X_1\widetilde{A}=0$.

(2) For $k=2,\cdots,n$

1) $L_{k-1}=X_{k-1}\widetilde{A}\begin{pmatrix}I_{k-1}\\0\end{pmatrix}$;

2) $c_k=a_k-L_{k-1}a_k$;

3) If $c_k\neq 0$ then 1) $r_k=c^*_kMa_k$, 2) $u^*_k=r_k^{-1}(c^*_kM\widetilde{A})$

else 1) $w_k=(I-L_{k-1})N^{-1}_{k-1}l_k$, 2) $d_k=(X_{k-1}\widetilde{A})e_k$,

3) $\delta_k=n_{kk}+d^*_kN_{k-1}d_k-(d^*_kl_k+l^*_kd_k)-l^*_kw_k$,

4) $b^*_k=\delta_k^{-1}(d^*_kN_{k-1}-l^*_k)X_{k-1}\widetilde{A}$;

4) $X_k\widetilde{A}=\begin{pmatrix}X_{k-1}\widetilde{A}-(d_k+w_k)u^*_k\\ u^*_k\end{pmatrix}$.

(3) $\hat{x}=(X_n\widetilde{A})e_{n+1}$.

There are applications on the other hand which call for inverting the matrix by adjoining additional rows. This problem of computing the $M-P$ inverse of A was considered in [1, chap. 5]. However, it is more difficult to compute the weighted $M-P$ inverse of A. Now we need some new notation.

Let $A\in C^{m\times n}$ and $A_{(k)}\in C^{k\times n}$ be the submatrix of A consisting of its first k rows. For $k=2,\cdots,m$, $A_{(k)}$ is partitioned as

$$A_{(k)}=\begin{pmatrix}A_{(k-1)}\\ a^*_{(k)}\end{pmatrix}, \tag{4.8}$$

where $a^*_{(k)}$ is the k-th row of A.

Set $M^{-1}=H$. Let $H_k\in C^{k\times k}$ be the leading principal submatrix of H and H_k be partitioned as

$$H_k=\begin{pmatrix} H_{k-1} & h_k \\ h_k^* & h_{kk} \end{pmatrix}. \tag{4.9}$$

Then $H_m=H$. Let the matrices Y_{k-1}, Y_k, Z_{k-1} and Z_k be defined by

$$Y_{k-1}=(A^*_{(k-1)})^+_{N^{-1}H_{k-1}}, \quad Y_k=(A^*_{(k)})^+_{N^{-1}H_k} \tag{4.10}$$

and

$$Z_{k-1}=Y^*_{k-1}, \quad Z_k=Y^*_k \tag{4.11}$$

the vectors $d_{(k)}$ and $c_{(k)}$ be defined by

$$d_{(k)}=Y_{k-1}a_{(k)}, \tag{4.12}$$

$$c_{(k)}=a_{(k)}-A^*_{(k-1)}d_{(k)} \tag{4.13}$$

$$=(I-A^*_{(k-1)}Y_{(k-1)})a_{(k)}. \tag{4.14}$$

Then, the relation between Z_{k-1} and Z_k is as follows:

Theorem 2. *Under the above notation, for $k=2, \cdots, m$,*

$$Z_k=[Z_{k-1}-b_{(k)}(d^*_{(k)}+h^*_kH^{-1}_{k-1}(I-A_{(k-1)}Z_{k-1}))\mid b_{(k)}], \tag{4.15}$$

where

$$b_{(k)}=\begin{cases}(c^*_{(k)}N^{-1}c_{(k)})^{-1}N^{-1}c_{(k)} & \text{if } c_{(k)}\neq 0, \quad (4.16)\\ \delta^{-1}_{(k)}Z_{k-1}(H_{k-1}d_{(k)}-h_k) & \text{if } c_{(k)}=0 \quad (4.17)\end{cases}$$

and

$$d^*_{(k)}=a^*_{(k)}Z_{k-1}, \tag{4.18}$$

$$c^*_{(k)}=a^*_{(k)}(I-Z_{k-1}A_{(k-1)}) \tag{4.19}$$

$$=a^*_{(k)}-d^*_{(k)}A_{(k-1)}, \tag{4.20}$$

$$\delta_{(k)}=h_{kk}+d^*_{(k)}H_{k-1}d_{(k)}-(d^*_{(k)}h_k+h^*_kd_{(k)})-h^*_kH^{-1}_{k-1}(I-A_{(k-1)}Z_{k-1})h_k, \tag{4.21}$$

$$Z_m=A^+_{MN}. \tag{4.22}$$

Proof. Applying Theorem 1 to the conjugate transpose of matrix (4.8) with p.d. matrices N^{-1} and H, we may obtain

$$Y_k=\begin{pmatrix} Y_{k-1}-(d_{(k)}+(I-Y_{k-1}A^*_{(k-1)})H^{-1}_{k-1}h_k)b^*_{(k)} \\ b^*_{(k)} \end{pmatrix}, \tag{4.23}$$

where

$$b^*_{(k)}=\begin{cases}(c^*_{(k)}N^{-1}c_{(k)})^{-1}c^*_{(k)}N^{-1} & \text{if } c_{(k)}\neq 0, \quad (4.24)\\ \delta^{-1}_{(k)}(d^*_{(k)}H_{k-1}-h^*_k)Y_{k-1} & \text{if } c_{(k)}=0 \quad (4.25)\end{cases}$$

and

$$\delta_{(k)}=h_{kk}+d^*_{(k)}H_{k-1}d_{(k)}-(d^*_{(k)}h_k+h^*_kd_{(k)})-h^*_k(I-Y_{k-1}A^*_{(k-1)})H^{-1}_{k-1}h_k. \tag{4.26}$$

Taking the conjugate transpose of (4.23)—(4.26) respectively, and using Lemma 1 (iii), we may obtain (4.16)—(4.22) immediately.

Obviously,

$$Z_1=\begin{cases}0 & \text{if } a_{(1)}=0,\\ (a^*_{(1)}N^{-1}a_{(1)})^{-1}N^{-1}a_{(1)} & \text{if } a_{(1)}\neq 0.\end{cases} \tag{4.27}$$

The practical method for computing A^+_{MN} by Theorem 2 is similar to Algorithm 1.

§ 5. Numerical Example

Let

$$A=\begin{pmatrix}1&2&0\\0&0&1\\1&2&0\end{pmatrix}, \quad M=\begin{pmatrix}1&0&1\\0&2&0\\1&0&3\end{pmatrix}, \quad N=\begin{pmatrix}1&1&1\\1&2&0\\1&0&3\end{pmatrix}.$$

We compute A^{+}_{MN} by Theorems 1 and 2 respectively, as follows:

(i) By Theorem 1:

$$A_1=a_1=\begin{pmatrix}1\\0\\1\end{pmatrix},\quad a_2=\begin{pmatrix}2\\0\\2\end{pmatrix},\quad X_1=(a_1^*Ma_1)^{-1}a_1^*M=\frac{1}{3}[1\ 0\ 2];$$

$$d_2=X_1a_2=2,\ c_2=a_2-A_1d_2=0,\ N_1=1,\ N_1^{-1}=1,\ l_2=1,\ (I-X_1A_1)N_1^{-1}l_2=0,$$

$$\delta_2=n_{22}+d_2^*N_1d_2-(d_2^*l_2+l_2^*d_2)-l_2^*(I-X_1A_1)N_1^{-1}l_2=2,$$

$$b_2^*=\delta_2^{-1}(d_2^*N_1-l_2^*)X_1=[1\ 0\ 2],$$

$$X_2=\begin{pmatrix}X_1-(d_2+(I-X_1A_1)N_1^{-1}l_2)b_2^*\\ b_2^*\end{pmatrix}=\frac{1}{6}\begin{pmatrix}0&0&0\\1&0&2\end{pmatrix};$$

$$A_2=\begin{pmatrix}1&2\\0&0\\1&2\end{pmatrix},\quad a_3=\begin{pmatrix}0\\1\\0\end{pmatrix},\quad d_3=X_2a_3=\begin{pmatrix}0\\0\end{pmatrix},$$

$$c_3=a_3=\begin{pmatrix}0\\1\\0\end{pmatrix},\quad N_2=\begin{pmatrix}1&1\\1&2\end{pmatrix},\quad N_2^{-1}=\begin{pmatrix}2&-1\\-1&2\end{pmatrix},$$

$$l_3=\begin{pmatrix}1\\0\end{pmatrix},\quad (I-X_2A_2)N_2^{-1}l_3=\begin{pmatrix}2\\-1\end{pmatrix},\quad b_3^*=(c_3^*Mc_3)^{-1}c_3^*M=[0\ 1\ 0],$$

$$X_3=\begin{pmatrix}X_2-(d_3+(I-X_2A_2)N_2^{-1}l_3)b_3^*\\ b_3^*\end{pmatrix}=\begin{pmatrix}0&-2&0\\1/6&1&1/3\\0&1&0\end{pmatrix}=A^{+}_{MN}.$$

(ii) By Theorem 2:

$$N^{-1}=\begin{pmatrix}6&-3&-2\\-3&2&1\\-2&1&1\end{pmatrix},\quad M^{-1}=H=\frac{1}{2}\begin{pmatrix}3&0&-1\\0&1&0\\-1&0&1\end{pmatrix}.$$

$$A_{(1)}=a_{(1)}^*=[1\ 2\ 0],\ Z_1=(a_{(1)}^*N^{-1}a_{(1)})^{-1}N^{-1}a_{(1)}=\begin{pmatrix}0\\1/2\\0\end{pmatrix};\ a_{(2)}^*=[0\ 0\ 1],$$

$$d_{(2)}^*=a_{(2)}^*Z_1=0,\ c_{(2)}^*=a_{(2)}^*-d_{(2)}^*A_{(1)}=[0\ 0\ 1],\ H_1=3/2,\ H_1^{-1}=2/3,\ h_2=0,$$

$$h_2^*H_1^{-1}(I-A_{(1)}Z_1)=0,\ b_{(2)}=(c_{(1)}^*N^{-1}c_{(1)})^{-1}N^{-1}c_{(1)}=\begin{pmatrix}-2\\1\\1\end{pmatrix},$$

$$Z_2=[Z_1-b_{(2)}(d_{(2)}^*+h_2^*H_1^{-1}(I-A_{(1)}Z_1))\,|\,b_{(2)}]=\begin{pmatrix}0&-2\\1/2&1\\0&1\end{pmatrix};$$

$$A_{(2)}=\begin{pmatrix}1&2&0\\0&0&1\end{pmatrix},\ a_{(3)}^*=[1\ 2\ 0],\ d_{(3)}^*=a_{(3)}^*Z_2=[1\ 0],\ c_{(3)}^*=a_{(3)}^*-d_{(3)}^*A_{(2)}=0,$$

$$H_2=\frac{1}{2}\begin{pmatrix}3&0\\0&1\end{pmatrix},\quad H_2^{-1}=\frac{2}{3}\begin{pmatrix}1&0\\0&3\end{pmatrix},\quad h_3=\begin{pmatrix}-1\\0\end{pmatrix},\quad h_3^*H_2^{-1}(I-A_{(2)}Z_2)=0,$$

$$\delta_{(3)}=h_{33}+d_{(3)}^*H_2d_{(3)}-(d_{(3)}^*h_3+h_3^*d_{(3)})-h_3^*H_2^{-1}(I-A_{(2)}Z_2)h_3=3,$$

$$b_{(3)}=\delta_{(3)}^{-1}Z_2(H_2d_{(3)}-h_3)=\begin{pmatrix}0\\1/3\\0\end{pmatrix},$$

$$Z_3=[Z_2-b_{(3)}(d_{(3)}^*+h_3^*H_2^{-1}(I-A_{(2)}Z_2))\mid b_{(3)}]=\begin{pmatrix}0&-2&0\\1/6&1&1/3\\0&1&0\end{pmatrix}=A_{MN}^+.$$

References

[1] T. N. E. Greville, Some applications of the pseudoinverse of a matrix, *SIAM Review*, **2** (1960), 15—22.

[2] S. L. Campbell, C. D. Meyer, Jr., Generalized Inverses of Linear Transformations, London, Pitman, 1979.

[3] Dong Song-shi, The Theory of Matrices and Its Applications in Geodesy and Cartography (in Chinese), Beijing, 1980.

[4] Wang Guo-rong, A New Proof of Greville's Method for Computing the $M-P$ Inverse, *Journal of Shanghai Normal University* (Natural Science Edition), **3**, 1985.

[5] A. Ben-Israel, T. N. E. Greville, Generalized Inverse: Theory and Applications, John Wiley, New York, 1974.

[6] Chen Yong-lin, On the weighted projector and weighted generalized inverse matrices, *Acta Mathematicae Applicatae Sinica*, **6**: 3 (1983), 282—291.