Acta Math. Hungar.
80 (1–2) (1998), 83–88.

A CLOSED FORMULA FOR THE MOORE–PENROSE GENERALIZED INVERSE OF A COMPLEX MATRIX OF GIVEN RANK

D. CONSTALES (Gent)

Abstract. For the Moore–Penrose generalized inverse of complex matrices, we establish closed forms valid on each set of matrices of given rank. These expressions are matrices of rational functions of the matrix coefficients and their complex conjugates, so that it is seen explicitly and constructively that taking the Moore–Penrose inverse is a real-analytic operation when restricted to the subsets of matrices of given rank.

1. Introduction

The Moore–Penrose generalized inverse (cf. [2]) $A^\dagger$ of a complex $m \times n$ matrix $A \in \mathbf{C}^{m\times n}$ is defined as the unique solution X to the system of equations

$$(1) \qquad AXA = A, \quad XAX = X, \quad (AX)^+ = AX, \quad (XA)^+ = XA,$$

in which $\cdot^+$ stands for Hermitian transposition (transposition combined with complex conjugation). If A has full rank $r = \min(m, n)$, there exist well-known (cf. [2]) closed-form expressions for its Moore–Penrose generalized inverse, viz $A^\dagger = A^+(AA^+)^{-1}$ if $r = m$ and $A^\dagger = (A^+A)^{-1}A^+$ if $r = n$.

In this note we shall prove an extension of these formulas to rank-deficient matrices A, i.e., those for which $r < \min(n, m)$, and establish appropriate formulas for each value of r. Each of these will express $A^\dagger$ as a matrix of rational functions of the components of A and of their complex conjugates.

When $r = 0$, i.e., A is zero, $A^\dagger = 0$ trivially provides such a closed-form expression. From now on, we concentrate on the cases where the rank satisfies $1 \leqq r < \min(n, m)$.

Let us first introduce a few notations. For any matrix $A \in \mathbf{C}^{m\times n}$ and for any $I = \{i_1, i_2, \dots, i_p\} \subseteqq \{1, 2, \dots, m\}$, $J = \{j_1, j_2, \dots, j_q\} \subseteqq \{1, 2, \dots, n\}$, the matrix $A[I, J]$ denotes the $p \times q$ submatrix of A formed by all the components of A that belong to the rows $i_1, i_2, \dots, i_p$ and to the columns $j_1, j_2, \dots, j_q$. For any square matrix $A = (a_{ij})$ of $\mathbf{C}^{n\times n}$ the matrix A^{adj} denotes the classical adjoint of A, i.e., $A^{\mathrm{adj}} = (A_{ji})$ where A_{ji} is the cofactor of a_{ji}.

2. Rank deficiency 1

Let $A \in \mathbf{C}^{m\times n}$, $m \leqq n$, be of rank $m-1$. Since AA^+ is a Hermitian matrix, there exists a unitary matrix U such that

$$AA^+ = U \operatorname{diag}(d_1, d_2, \ldots, d_{m-1}, 0)U^+;$$

the vanishing of the last diagonal component then implies that the last column of U^+A^+U is zero. Clearly,

$$(AA^+)^\dagger = U \operatorname{diag}(d_1^{-1}, d_2^{-1}, \ldots, d_{m-1}^{-1}, 0)U^+, \tag{2}$$

$$\big((AA^+)^{\mathrm{adj}}\big)^\dagger = U \operatorname{diag}\big(0, 0, \ldots, 0, (d_1 d_2 \ldots d_{m-1})^{-1}\big)\, U^+ \tag{3}$$

and since the last column of U^+A^+U is zero, we see that

$$A^+\big((AA^+)^{\mathrm{adj}}\big)^\dagger = 0.$$

Therefore the well-known expression (cf. [2])

$$A^\dagger = A^+(AA^+)^\dagger$$

can be transformed to

$$A^\dagger = A^+\big((AA^+)^\dagger + \big((AA^+)^{\mathrm{adj}}\big)^\dagger\big).$$

Using the general fact (cf. [2]) that for matrices $X, Y \in \mathbf{C}^{m\times m}$ such that $XY = YX = 0$, $(X+Y)^\dagger = X^\dagger + Y^\dagger$, we obtain with $X = AA^+$, $Y = (AA^+)^{\mathrm{adj}}$,

$$A^\dagger = A^+\big(AA^+ + (AA^+)^{\mathrm{adj}}\big)^\dagger,$$

where in view of (2) and (3) the generalized inverse is applied to a nonsingular matrix, so

$$A^\dagger = A^+\big(AA^+ + (AA^+)^{\mathrm{adj}}\big)^{-1}.$$

Hence we have shown that for any $m \times n$ matrix A, $m \leqq n$, of rank $m-1$ over $\mathbf{C}$,

$$A^\dagger = A^+ \frac{\big(AA^+ + (AA^+)^{\mathrm{adj}}\big)^{\mathrm{adj}}}{\det\big(AA^+ + (AA^+)^{\mathrm{adj}}\big)}. \tag{4}$$

If $m \geqq n$ and A has rank $n-1$, a similar formula can be obtained by applying (4) to A^+ and taking the Hermitian transpose of both sides, relying on the identity $A^\dagger = \big((A^+)^\dagger\big)^+$.

3. The general case for a singular matrix

Now we assume that the rank r of $A \in \mathbf{C}^{m\times n}$ satisfies $1 \leqq r \leqq \min(m,n) - 1$. The following elementary result will be used.

LEMMA 1. *If $X, Y \in \mathbf{C}^{n\times n}$ satisfy $XY = YX = 0$, then*

$$\operatorname{range}(X + Y^+) = \operatorname{range} X \oplus \operatorname{range} Y^+.$$

PROOF. Obviously $\operatorname{range}(X + Y^+) \subseteqq \operatorname{range} X + \operatorname{range} Y^+$. If $Xx + Y^+y \in \operatorname{range} X + \operatorname{range} Y^+$ there are x', y' such that $Xx + Y^+y = XX^+x' + Y^+Yy' = (X + Y^+)(X^+x' + Yy') \in \operatorname{range}(X + Y^+)$. The sum is direct because $Xx = Y^+y$ implies $X^+Xx = X^+Y^+y = 0$ so $Xx = 0$. This completes the proof of Lemma 1.

Again, we shall obtain a formula

$$A^\dagger = A^+(AA^+ + B^+B)^{-1}$$

for some suitable B, constructed as follows.

Let $M = \{1, 2, \dots, m\}$ and $N = \{1, 2, \dots, n\}$; if $I = \{i_1, i_2, \dots, i_{r+1}\} \subseteqq M$ and $J = \{j_1, j_2, \dots, j_{r+1}\} \subseteqq N$, define $B_{I,J}$ as the $n \times m$ matrix with entries

$$(B_{I,J})_{i,j} = \begin{cases} \left(A[I,J]^{\mathrm{adj}}\right)_{kl} & \text{if } i = j_k \text{ and } j = i_l \text{ for some } k \text{ and } l, \\ 0 & \text{otherwise.} \end{cases}$$

To clarify this construction, let $n = m = 4$, $r = 2$, $I = \{1, 2, 4\}$, $J = \{1, 3, 4\}$ and write

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix}.$$

If

$$\begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix} = A[I,J]^{\mathrm{adj}}, \qquad A[I,J] = \begin{pmatrix} a_{11} & a_{13} & a_{14} \\ a_{21} & a_{23} & a_{24} \\ a_{41} & a_{43} & a_{44} \end{pmatrix},$$

then

$$B_{I,J} = \begin{pmatrix} b_{11} & b_{12} & 0 & b_{13} \\ 0 & 0 & 0 & 0 \\ b_{21} & b_{22} & 0 & b_{23} \\ b_{31} & b_{32} & 0 & b_{33} \end{pmatrix}.$$

Combining the minors that make up $B_{I,J}$ with the components of A, it is readily seen that the components of $B_{I,J}A$ are either determinants of matrices with at least two equal columns, or minors of order $(r+1)$ of A, whose rank is only r: in both cases, they must vanish, so $B_{I,J}$ annihilates A from the left. Illustrations of these cases on the earlier example (remember that $r=2$ there) are

$$(B_{I,J}A)_{11} = b_{11}a_{11} + b_{12}a_{21} + b_{13}a_{41} = \begin{vmatrix} a_{11} & a_{13} & a_{14} \\ a_{21} & a_{23} & a_{24} \\ a_{41} & a_{43} & a_{44} \end{vmatrix} = 0,$$

$$(B_{I,J}A)_{12} = b_{11}a_{12} + b_{12}a_{22} + b_{13}a_{42} = \begin{vmatrix} a_{12} & a_{13} & a_{14} \\ a_{22} & a_{23} & a_{24} \\ a_{42} & a_{43} & a_{44} \end{vmatrix} = 0,$$

$$(B_{I,J}A)_{13} = b_{11}a_{13} + b_{12}a_{23} + b_{13}a_{43} = \begin{vmatrix} a_{13} & a_{13} & a_{14} \\ a_{23} & a_{23} & a_{24} \\ a_{43} & a_{43} & a_{44} \end{vmatrix} = 0.$$

We then define B by stacking these $B_{I,J}$ vertically, for all $\binom{m}{r+1}\binom{n}{r+1}$ pairs (I,J).

LEMMA 2. *The matrix B annihilates A from the left and has rank at least* $(m-r)$.

PROOF. Because each $B_{I,J}$ annihilates A from the left, so does B, since it is obtained by stacking them vertically. Because the matrix A has rank r, it contains a non-singular $r\times r$ submatrix. To put it more technically, we can find a set $I_0=\{i_1,i_2,\dots,i_r\}$ of row indices and a set $J_0=\{j_1,j_2,\dots,j_r\}$ of column indices such that the submatrix $A[I_0,J_0]$ has non-zero determinant.

The set I_0 having r elements, there clearly are $m-r$ subsets $I_1,I_2,\dots,I_{m-r}$ of M that have $r+1$ elements and contain I_0 as a subset. On the other hand, J_0 has r elements, so there are $n-r$ subsets of N having $r+1$ elements and containing J_0; we shall only need one of these, say $J=J_0\cup\{j\}$ for some $j\in N\setminus J_0$.

Relying on the definitions the j^{th} row of $B[I_p,J]$, where $p=1,2,\dots,m-r$ has 0 on the columns of $M\setminus I_p$ and $\pm\det A[I_0,J_0]$ on the column whose index is the sole element of $I_p\setminus I_0$. Considering the $m-r$ possible values of p we find a non-zero multiple of the $(m-r)\times(m-r)$ unit matrix in B, which must have rank at least $m-r$. This completes the proof of Lemma 2.

From

$$A^+(B^+B)^\dagger = A^+(B^+B)(B^+BB^+B)^\dagger = (BA)^+B(B^+BB^+B)^\dagger = 0$$

it follows, by a similar reasoning as in Section 2, that

$$A^\dagger = A^+(AA^+)^\dagger = A^+(AA^+ + B^+B)^\dagger.$$

Relying on the lemmas, we have that

$$\begin{aligned}\operatorname{rank}(AA^+ + B^+B) &= \operatorname{rank} AA^+ + \operatorname{rank} B^+B \\ &= \operatorname{rank} A + \operatorname{rank} B \geqq r + (m-r) = m.\end{aligned}$$

Hence in this case too the generalized inverse is applied to a non-singular square matrix (of dimension $m \times m$), and therefore the closed-form expression

$$A^\dagger = A^+ \frac{(AA^+ + B^+B)^{\mathrm{adj}}}{\det(AA^+ + B^+B)} \tag{5}$$

is obtained.

4. Concluding remarks

The matrix B being obtained by stacking the $B_{I,J}$ for all subsets I and J of cardinality $(r+1)$, the expression B^+B which occurs in the formula (5) can also be written as

$$B^+B = \sum_{I,J} B_{I,J}^+ B_{I,J},$$

but we preferred to use B because it is easier to prove that its rank is at least $m-r$. In fact, B can be replaced in the proof by any product CB, where C is a non-singular matrix, thus giving rise to many other closed-form expressions.

Thus we have proved the following theorem.

THEOREM 1. *Let $A \in \mathbf{C}^{m\times n}$, $m \leqq n$, be a matrix of* rank r. *Then its Moore–Penrose generalized inverse is given by the closed formula*

$$A^\dagger = A^+ \frac{(AA^+ + A')^{\mathrm{adj}}}{\det(AA^+ + A')},$$

where the matrix A' satisfies

(i) $A' = 0$ *if* $r = m$;

(ii) $A' = I_m$ *if* $r = 0$;

(iii) *otherwise, i.e., when* $1 \leqq r < m$:

$$A' = \sum_{I,J} B_{I,J}^+ B_{I,J},$$

where the sum ranges over all subsets $I \subseteqq \{1,\dots,m\}$ and $J \subseteqq \{1,\dots,n\}$ of cardinality $r+1$.

The case $m \geqq n$ is completely analogous and can be obtained by Hermitian transposition.

As an application, note that these formulas express each component of $A^\dagger$ as a rational function of the components of A and of their complex conjugates, so it follows at once that $A \to A^\dagger$ is a real-analytic mapping when restricted to a subset of matrices of fixed rank.

References

[1] A. Albert, *Regression and the Moore–Penrose pseudo-inverse*, Academic Press (New York, 1972).

[2] A. Ben-Israel and T. N. E. Greville, *Generalized Inverses: Theory and Applications,*, Wiley (New York, 1974).

[3] S. L. Campbell and C. D. Meyer Jr., *Generalized inverses of linear transformations*, Pitman (London, 1979).

(Received December 18, 1996; revised February 25, 1997)

UNIVERSITEIT GENT
VAKGROEP WISKUNDIGE ANALYSE
GALGLAAN 2
B-9000 GENT
BELGIUM
E-MAIL: DCONS@WORLD.STD.COM