CHAPTER 1

Existence and Construction of Generalized Inverses

1. The Penrose equations

In 1955 Penrose [**637**] showed that, for every finite matrix A (square or rectangular) of real or complex elements, there is a unique matrix X satisfying the four equations (that we call the *Penrose equations*)

$$AXA = A\ , \tag{1}$$

$$XAX = X\ , \tag{2}$$

$$(AX)^* = AX\ , \tag{3}$$

$$(XA)^* = XA\ , \tag{4}$$

where A^* denotes the conjugate transpose of A. Because this unique generalized inverse had previously been studied (though defined in a different way) by E. H. Moore ([**571**], [**572**]), it is commonly known as the *Moore–Penrose inverse*, and is often denoted by $A^\dagger$.

If A is nonsingular, it is clear that $X = A^{-1}$ trivially satisfies the four equations. Since the Moore–Penrose inverse is known to be unique (as we shall prove shortly) it follows that the Moore–Penrose inverse of a nonsingular matrix is the same as the ordinary inverse.

Throughout this book we shall be much concerned with generalized inverses that satisfy some, but not all, of the four Penrose equations. As we shall wish to deal with a number of different subsets of the set of four equations, we need a convenient notation for a generalized inverse satisfying certain specified equations. Let $\mathbb{C}^{m\times n}$ $[\mathbb{R}^{m\times n}]$ denote the class of $m \times n$ complex [real] matrices.

Definition 1. For any $A \in \mathbb{C}^{m\times n}$, let $A\{i, j, \dots , k\}$ denote the set of matrices $X \in \mathbb{C}^{n\times m}$ which satisfy equations $(i), (j), \cdots , (k)$ from among the equations (1)–(4). A matrix $X \in A\{i, j, \dots , k\}$ is called[1] an $\{i, j, \dots , k\}$*–inverse* of A, and also denoted by $A^{(i,j,\dots,k)}$.

In Chapter 4 we shall extend the scope of this notation by enlarging the set of four matrix equations to include several further equations, applicable only to square matrices, that will play an essential role in the study of generalized inverses having spectral properties.

Exercises.

Ex. 1. If $A\{1, 2, 3, 4\}$ is nonempty, then it consists of a single element (Penrose [**637**]).

Proof. Let $X, Y \in A\{1, 2, 3, 4\}$. Then

$$\begin{aligned} X &= X(AX)^* = XX^*A^* = X(AX)^*(AY)^* \\ &= XAY = (XA)^*(YA)^*Y = A^*Y^*Y \\ &= (YA)^*Y = Y\ . \end{aligned}$$

□

Ex. 2. By means of a (trivial) example, show that $A\{2, 3, 4\}$ is nonempty.

[1] Some writers have adopted descriptive names to designate various classes of generalized inverses. However there is a notable lack of uniformity and consistency in the use of these terms by different writers. Thus, $X \in A\{1\}$ is called a *generalized inverse* (Rao [**674**]), *pseudoinverse* (Sheffield [**754**]), *inverse* (Bjerhammar [**100**]). $X \in A\{1, 2\}$ is called a *semi–inverse* (Frame [**279**]), *reciprocal inverse* (Bjerhammar), *reflexive generalized inverse* (Rohde [**711**]). $X \in A\{1, 2, 3\}$ is called a weak generalized inverse (Goldman and Zelen [**295**]). $X \in A\{1, 2, 3, 4\}$ is called the *general reciprocal* (Moore [**571, 572**]), *generalized inverse* (Penrose [**637**]), *pseudoinverse* (Greville [**318**]), the *Moore–Penrose inverse* (Ben-Israel and Charnes [**77**]). In view of this diversity of terminology, the unambiguous notation adopted here is considered preferable. This notation also emphasizes the lack of uniqueness of many of the generalized inverses considered.

2. Existence and construction of {1}–inverses

It is easy to construct a {1}–inverse of the matrix $R \in \mathbb{C}_r^{m\times n}$ given by

$$R = \begin{bmatrix} I_r & K \\ O & O \end{bmatrix} \tag{0.70}$$

For any $L \in \mathbb{C}^{(n-r)\times(m-r)}$, the $n \times m$ matrix

$$S = \begin{bmatrix} I_r & O \\ O & L \end{bmatrix}$$

is a {1}–inverse of (0.70). If R is of full column [row] rank, the two lower [right–hand] submatrices are interpreted as absent.

The construction of {1}–inverses for an arbitrary $A \in \mathbb{C}^{m\times n}$ is simplified by transforming A into a Hermite normal form, as shown in the following theorem.

THEOREM 1. Let $A \in \mathbb{C}_r^{m\times n}$, and let $E \in \mathbb{C}_m^{m\times m}$ and $P \in \mathbb{C}_n^{n\times n}$ be such that

$$EAP = \begin{bmatrix} I_r & K \\ O & O \end{bmatrix} . \tag{0.71}$$

Then for any $L \in \mathbb{C}^{(n-r)\times(m-r)}$, the $n \times m$ matrix

$$X = P \begin{bmatrix} I_r & O \\ O & L \end{bmatrix} E \tag{5}$$

is a {1}–inverse of A. The partitioned matrices in (0.71) and (5) must be suitably interpreted in case $r = m$ or $r = n$.

PROOF. Rewriting (0.71) as

$$A = E^{-1} \begin{bmatrix} I_r & K \\ O & O \end{bmatrix} P^{-1} , \tag{0.73}$$

it is easily verified that any X given by (5) satisfies $AXA = A$. □

In the trivial case of $r = 0$, when A is therefore the $m\times n$ null matrix, *any* $n\times m$ matrix is a {1}–inverse.

We note that since P and E are both nonsingular, the rank of X is the rank of the partitioned matrix in RHS(5). In view of the form of the latter matrix,

$$\operatorname{rank} X = r + \operatorname{rank} L . \tag{6}$$

Since L is arbitrary, it follows that a {1}–inverse of A exists having any rank between r and $\min\{m, n\}$, inclusive (see also Fisher [**268**]).

Theorem 1 shows that every finite matrix with elements in the complex field has a {1}–inverse, and suggests how such an inverse can be constructed.

Exercises.

EX. 3. What is the Hermite normal form of a nonsingular matrix A? In this case, whet is the matrix E, and what is its relationship to A? What is the permutation matrix P? What is the matrix X given by (5)?

EX. 4. An $m \times n$ matrix A has all its elements equal to 0 except for the (i, j)th element, which is 1. What is the Hermite normal form? Show that E can be taken as a permutation matrix. What are the simplest choices of E and P? (By "simplest" we mean having the smallest number of elements different from the corresponding elements of the unit matrix of the same order.) Using these choices of E and P, but regarding L as entirely arbitrary, what is the form of the resulting matrix X given by (5)? Is this X the most general {1}–inverse of A? (See Exercise 6, Introduction, and Ex. 11 below.)

EX. 5. Show that every square matrix has a nonsingular {1}–inverse.

EX. 6. *Computing a* $\{1\}$*–inverse.* This is demonstrated for the matrix A of (0.79), using (5) with E as computed in (0.81), and an arbitrary $L \in \mathbb{C}^{(n-r)\times(m-r)}$. Using the permutation matrix P selected in (0.82), and the corresponding submatrix K, we write

$$EAP = \begin{bmatrix} 1 & 0 & \vdots & 0 & \frac{1}{2} & 1-2i & -\frac{1}{2}i \\ 0 & 1 & \vdots & 0 & 0 & 2 & 1+i \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \vdots & 0 & 0 & 0 & 0 \end{bmatrix}, \quad \text{and take} \quad L = \begin{bmatrix} \alpha \\ \beta \\ \gamma \\ \delta \end{bmatrix} \in \mathbb{C}^{4\times 1}$$

since $m = 3$, $n = 6$, $r = 2$. A $\{1\}$–inverse of A is

$$X = P \begin{bmatrix} I_r & O \\ O & L \end{bmatrix} E \tag{5}$$

$$= \begin{bmatrix} 0 & 0 & \vdots & 1 & 0 & 0 & 0 \\ 1 & 0 & \vdots & 0 & 0 & 0 & 0 \\ 0 & 0 & \vdots & 0 & 1 & 0 & 0 \\ 0 & 1 & \vdots & 0 & 0 & 0 & 0 \\ 0 & 0 & \vdots & 0 & 0 & 1 & 0 \\ 0 & 0 & \vdots & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & \vdots & 0 \\ 0 & 1 & \vdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \vdots & \alpha \\ 0 & 0 & \vdots & \beta \\ 0 & 0 & \vdots & \gamma \\ 0 & 0 & \vdots & \delta \end{bmatrix} \begin{bmatrix} -\frac{1}{2}i & 0 & 0 \\ 0 & -\frac{1}{3} & 0 \\ i & \frac{1}{3} & 1 \end{bmatrix} = \begin{bmatrix} i\alpha & \frac{1}{3}\alpha & \alpha \\ -\frac{1}{2}i & 0 & 0 \\ i\beta & \frac{1}{3}\beta & \beta \\ 0 & -\frac{1}{3} & 0 \\ i\gamma & \frac{1}{3}\gamma & \gamma \\ i\delta & \frac{1}{3}\delta & \delta \end{bmatrix}. \tag{7}$$

Note that, in general, the scalars $i\alpha, i\beta, i\gamma, i\delta$ are not pure imaginaries since $\alpha, \beta, \gamma, \delta$ are complex.

3. Properties of $\{1\}$–inverses

Certain properties of $\{1\}$–inverses are given in Lemma 1. For a given matrix A, we denote any $\{1\}$–inverse by $A^{(1)}$. Note that, in general, $A^{(1)}$ is not a uniquely defined matrix (see Ex. 8 below). For any scalar λ we define $\lambda^\dagger$ by

$$\lambda^\dagger = \begin{cases} \lambda^{-1} & (\lambda \neq 0) \\ 0 & (\lambda = 0) \end{cases}. \tag{8}$$

It will be recalled that a square matrix E is called *idempotent* if $E^2 = E$. Idempotent matrices are intimately related to generalized inverses, and their properties are considered in some detail in Chapter 2.

LEMMA 1. Let $A \in \mathbb{C}_r^{m\times n}$, $\lambda \in \mathbb{C}$. Then,
(a) $(A^{(1)})^* \in A^*\{1\}$.
(b) If A is nonsingular, $A^{(1)} = A^{-1}$ uniquely (see also Ex. 7 below).
(c) $\lambda^\dagger A^{(1)} \in (\lambda A)\{1\}$.
(d) $\operatorname{rank} A^{(1)} \geq \operatorname{rank} A$.
(e) If S and T are nonsingular, $T^{-1}A^{(1)}S^{-1} \in SAT\{1\}$.
(f) $AA^{(1)}$ and $A^{(1)}A$ are idempotent and have the same rank as A.

PROOF. These are immediate consequences of the defining relation (1); (d) and the latter part of (f) depend on the fact that the rank of a product of matrices does not exceed the rank of any factor. □

If an $m \times n$ matrix A is of full column rank, its $\{1\}$–inverses are its left inverses. If it is of full row rank, its $\{1\}$–inverses are its right inverses.

LEMMA 2. Let $A \in \mathbb{C}_r^{m\times n}$. Then,
(a) $A^{(1)}A = I_n$ if and only if $r = n$.
(b) $AA^{(1)} = I_m$ if and only if $r = m$.

PROOF. (a) *If*: Let $A \in \mathbb{C}_r^{m\times n}$. Then the $n \times n$ matrix $A^{(1)}A$ is, by Lemma 1(f), idempotent and nonsingular. Multiplying $(A^{(1)}A)^2 = A^{(1)}A$ by $(A^{(1)}A)^{-1}$ gives $A^{(1)}A = I_n$.
Only if: $A^{(1)}A = I_n \implies \operatorname{rank} A^{(1)}A = n \implies \operatorname{rank} A = n$, by Lemma 1(f).
(b) Similarly proved. □

Exercises and examples.

EX. 7. Let $A = FHG$ where F is of full column rank and G is of full row rank. Then $\operatorname{rank} A = \operatorname{rank} H$. (*Hint*: Use Lemma 2.)

EX. 8. Show that A is nonsingular if and only if it has a unique $\{1\}$–inverse, which then coincides with A^{-1}.

PROOF. For any $\mathbf{x} \in N(A)$ $[\mathbf{y} \in N(A^*)]$, adding $\mathbf{x}$ $[\mathbf{y}^*]$ to any column [row] of an $X \in A\{1\}$ gives another $\{1\}$–inverse of A. The uniqueness of the $\{1\}$–inverse is therefore equivalent to

$$N(A) = \{\mathbf{0}\}\ ,\ N(A^*) = \{\mathbf{0}\}\ ,$$

i.e., to the nonsingularity of A. □

EX. 9. Show that if $A^{(1)} \in A\{1\}$, then $R(AA^{(1)}) = R(A)$, $N(AA^{(1)}) = N(A)$, and $R((A^{(1)}A)^*) = R(A^*)$.

PROOF. We have

$$R(A) \supset R(AA^{(1)}) \supset R(AA^{(1)}A) = R(A)\ ,$$

from which the first result follows.
Similarly,

$$N(A) \subset N(A^{(1)}A) \subset N(AA^{(1)}A) = N(A)$$

yields the second equation.
Finally, by Lemma 1(a),

$$R(A^*) \supset R(A^*(A^{(1)})^*) = R((A^{(1)}A)^*) \supset R(A^*(A^{(1)})^*A^*) = R(A^*)\ .$$

□

EX. 10. More generally, show that $R(AB) = R(A)$ if and only if $\operatorname{rank} AB = \operatorname{rank} A$, and $N(AB) = N(B)$ if and only if $\operatorname{rank} AB = \operatorname{rank} B$.

PROOF. Evidently, $R(A) \supset R(AB)$, and these two subspaces are identical if and only if they have the same dimension. But, the rank of any matrix is the dimension of its range.
Similarly, $N(B) \subset N(AB)$. Now, the nullity of any matrix is the dimension of its null space, and also the number of columns minus the rank. Thus, $N(B) = N(AB)$ if and only if B and AB have the same nullity, which is equivalent, in this case, to having the same rank, since the two matrices have the same number of columns. □

EX. 11. The answer to the last question in Ex. 4 indicates that, for particular choices of E and P, one does not get all the $\{1\}$–inverses of A merely by varying L in (5). Note, however, that Theorem 1 does not require P to be a permutation matrix. Could one get all the $\{1\}$–inverses by considering all nonsingular P and Q such that

$$QAP = \begin{bmatrix} I_r & O \\ O & O \end{bmatrix}\ ? \tag{9}$$

Given $A \in \mathbb{C}_r^{m\times n}$, show that $X \in A\{1\}$ if and only if

$$X = P \begin{bmatrix} I_r & O \\ O & L \end{bmatrix} Q \tag{10}$$

for some L and for some nonsingular P and Q satisfying (9).

PROOF. If (9) and (10) hold, X is a $\{1\}$–inverse of A by Theorem 1.
On the other hand, let $AXA = A$. Then, both AX and XA are idempotent and of rank r, by Lemma 1(f). Since any idempotent matrix E satisfies $E(E - I) = O$, its only eigenvalues are 0 and 1. Thus, the Jordan canonical forms of both AX and XA are of the form

$$\begin{bmatrix} I_r & O \\ O & O \end{bmatrix}\ ,$$

being of orders m and n, respectively. Therefore, there exist nonsingular P and R such that

$$R^{-1}AXR = \begin{bmatrix} I_r & O \\ O & O \end{bmatrix} , \; P^{-1}XAP = \begin{bmatrix} I_r & O \\ O & O \end{bmatrix} .$$

Thus,

$$\begin{aligned} R^{-1}AP &= R^{-1}AXAXAP = (R^{-1}AXR)R^{-1}AP(P^{-1}XAP) \\ &= \begin{bmatrix} I_r & O \\ O & O \end{bmatrix} R^{-1}AP \begin{bmatrix} I_r & O \\ O & O \end{bmatrix} . \end{aligned}$$

It follows that $R^{-1}AP$ is of the form

$$R^{-1}AP = \begin{bmatrix} H & O \\ O & O \end{bmatrix} ,$$

where $H \in \mathbb{C}_r^{r\times r}$, i.e., nonsingular. Let

$$Q = \begin{bmatrix} H^{-1} & O \\ O & I_{m-r} \end{bmatrix} R^{-1} .$$

Then (9) is satisfied. Consider the matrix $P^{-1}XQ^{-1}$. We have,

$$\begin{aligned} \begin{bmatrix} I_r & O \\ O & O \end{bmatrix} (P^{-1}XQ^{-1}) &= (QAP)\,(P^{-1}XQ^{-1}) = QAXQ^{-1} \\ &= \begin{bmatrix} H^{-1} & O \\ O & I_{m-r} \end{bmatrix} \begin{bmatrix} I_r & O \\ O & O \end{bmatrix} \begin{bmatrix} H & O \\ O & I_{m-r} \end{bmatrix} = \begin{bmatrix} I_r & O \\ O & O \end{bmatrix} \end{aligned}$$

and

$$\begin{aligned} (P^{-1}XQ^{-1}) \begin{bmatrix} I_r & O \\ O & O \end{bmatrix} &= (P^{-1}XQ^{-1})(QAP) = P^{-1}XAP \\ &= \begin{bmatrix} I_r & O \\ O & O \end{bmatrix} . \end{aligned}$$

From the latter two equations it follows that

$$P^{-1}XQ^{-1} = \begin{bmatrix} I_r & O \\ O & L \end{bmatrix}$$

for some L. But this is equivalent to (10). □

4. Existence and construction of $\{1,2\}$–inverses

It was first noted by Bjerhammar [**100**] that the existence of a $\{1\}$–inverse of a matrix A implies the existence of a $\{1,2\}$–inverse. This easily verified observation is stated as a lemma for convenience of reference.

LEMMA 3. Let $Y, Z \in A\{1\}$, and let

$$X = YAZ .$$

Then $X \in A\{1,2\}$.

Since the matrices A and X occur symmetrically in (1) and (2), $X \in A\{1,2\}$ and $A \in X\{1,2\}$ are equivalent statements, and in either case we can say that A and X are $\{1,2\}$–inverses *of each other.*

From (1) and (2) and the fact that the rank of a product of matrices does not exceed the rank of any factor, it follows at once that if A and X are $\{1,2\}$–inverses of each other, they have the same rank. Less obvious is the fact, first noted by Bjerhammar [**100**], that if X is a $\{1\}$–inverse of A and of the same rank as A, it is a $\{1,2\}$–inverse of A.

THEOREM 2. (Bjerhammar) Given A and $X \in A\{1\}$, $X \in A\{1,2\}$ if and only if $\operatorname{rank} X = \operatorname{rank} A$.

PROOF. *If*: Clearly $R(XA) \subset R(X)$. But $\operatorname{rank} XA = \operatorname{rank} A$ by Lemma 1(f), and so, if $\operatorname{rank} X = \operatorname{rank} A$, $R(XA) = R(X)$ by Ex. 10. Thus,

$$XAY = X$$

for some Y. Premultiplication by A gives

$$AX = AXAY = AY ,$$

and therefore

$$XAX = X .$$

Only if: This follows at once from (1) and (2). □

An equivalent statement is the following:

COROLLARY 1. Any two of the following three statement imply the third:

$$\begin{aligned} X &\in A\{1\} , \\ X &\in A\{2\} , \\ \operatorname{rank} X &= \operatorname{rank} A . \end{aligned}$$ □

In view of Theorem 2, (6) shows that the $\{1\}$–inverse obtained from the Hermite normal form is a $\{1, 2\}$–inverse if we take $L = O$. In other words,

$$X = P \begin{bmatrix} I_r & O \\ O & O \end{bmatrix} E \tag{11}$$

is a $\{1, 2\}$–inverse of A where P and E are nonsingular and satisfy (0.71).

Exercises.

EX. 12. Show that (5) gives a $\{1, 2\}$–inverse of A if and only if $L = O$.

EX. 13. Let $A = [a_{ij}] \in \mathbb{C}^{m\times n}$ be nonzero and upper triangular, i.e., $a_{ij} = 0$ if $i > j$. Find a $\{1, 2\}$–inverse of A.

SOLUTION. Let P, Q be permutation matrices such that

$$QAP = \begin{bmatrix} T & K \\ O & O \end{bmatrix}$$

where T is upper triangular and nonsingular (the K block, or the zero blocks, are absent if A is of full–rank.) Then

$$X = P \begin{bmatrix} T^{-1} & O \\ O & O \end{bmatrix} Q$$

is a $\{1, 2\}$–inverse of A (again, some zero blocks are absent if A is of full–rank.) Note that the inverse T^{-1} is obtained from T by back substitution. □

5. Existence and construction of $\{1, 2, 3\}$–, $\{1, 2, 4\}$– and $\{1, 2, 3, 4\}$–inverses

Just as Bjerhammar [**100**] showed that the existence of a $\{1\}$–inverse implies the existence of a $\{1.2\}$–inverse, Urquhart [**832**] has shown that the existence of a $\{1\}$–inverse of every finite matrix with elements in $\mathbb{C}$ implies the existence of a $\{1, 2, 3\}$–inverse and a $\{1, 2, 4\}$–inverse of every such matrix. However, in order to show the nonemptiness of $A\{1, 2, 3\}$ and $A\{1, 2, 4\}$ for any given A, we shall utilize the $\{1\}$–inverse not of A itself but of a related matrix. For that purpose we shall need the following lemma.

LEMMA 4. For any finite matrix A,

$$\operatorname{rank} AA^* = \operatorname{rank} A = \operatorname{rank} A^*A .$$

PROOF. If $A \in \mathbb{C}^{m\times n}$, both A and AA^* have m rows. Now, the rank of any m–rowed matrix is equal to m minus the number of independent linear relations among its rows. To show that $\operatorname{rank} AA^* = \operatorname{rank} A$, it is sufficient, therefore, to show that every linear relation among the rows of A holds for the corresponding rows of AA^*, and vice versa. Any nontrivial linear relation among the rows of a matrix H is equivalent to the existence of a nonzero row vector $\mathbf{x}^*$ such that $\mathbf{x}^*H = \mathbf{0}$. Now evidently,

$$\mathbf{x}^*A = \mathbf{0} \implies \mathbf{x}^*AA^* = \mathbf{0} \ ,$$

and, conversely,

$$\begin{aligned} \mathbf{x}^*AA^* = \mathbf{0} &\implies \mathbf{0} = \mathbf{x}^*AA^*\mathbf{x} = (A^*\mathbf{x})^*A^*\mathbf{x} \\ &\implies A^*\mathbf{x} = \mathbf{0} \implies \mathbf{0} = (A^*\mathbf{x})^* = \mathbf{x}^*A \ . \end{aligned}$$

Here we have used the fact that, for any column vector $\mathbf{y}$ of complex elements $\mathbf{y}^*\mathbf{y}$ is the sum of squares of the absolute values of the elements, and this sum vanishes only if every element is zero.

Finally, applying this result to the matrix A^* gives $\operatorname{rank} A^*A = \operatorname{rank} A^*$, and, of course, $\operatorname{rank} A^* = \operatorname{rank} A$. □

COROLLARY 2. For any finite matrix A, $R(AA^*) = R(A)$ and $N(AA^*) = N(A)$.

PROOF. This follows from Lemma 4 and Ex. 10. □

Using the preceding lemma, we can now prove the following theorem.

THEOREM 3. (Urquhart [**832**]) For every finite matrix A with complex elements,

$$Y = (A^*A)^{(1)}A^* \in A\{1,2,3\} \tag{12a}$$

and

$$Z = A^*(AA^*)^{(1)} \in A\{1,2,4\} \ . \tag{12b}$$

PROOF. Applying Corollary 2 to A^* gives

$$R(A^*A) = R(A^*) \ ,$$

and so,

$$A^* = A^*AU \tag{13}$$

for some U. Taking conjugate transpose gives

$$A = U^*A^*A \ . \tag{14}$$

Consequently,

$$AYA = U^*A^*A(A^*A)^{(1)}A^*A = U^*A^*A = A \ .$$

Thus, $Y \in A\{1\}$. But $\operatorname{rank} Y \geq \operatorname{rank} A$ by Lemma 1(d), and $\operatorname{rank} Y \leq \operatorname{rank} A^* = \operatorname{rank} A$ by the definition of Y. Therefore

$$\operatorname{rank} Y = \operatorname{rank} A \ ,$$

and, by Theorem 2, $Y \in A\{1,2\}$. Finally, (13) and (14) give

$$AY = U^*A^*A(A^*A)^{(1)}A^*AU = U^*A^*AU \ ,$$

which is clearly Hermitian. Thus, (12a) is established.

Relation (12b) is similarly proved. □

A $\{1,2\}$–inverse of a matrix A is, of course, a $\{2\}$–inverse, and similarly, a $\{1,2,3\}$–inverse is also a $\{1,3\}$–inverse and a $\{2,3\}$–inverse. Thus, if we can establish the existence of a $\{1,2,3,4\}$–inverse, we will have demonstrated the existence of an $\{i,j,\dots,k\}$–inverse for all possible choices of one, two or three integers $i,j,\dots,k$ from the set $\{1,2,3,4\}$. It was shown in Ex. 1 that if a $\{1,2,3,4\}$–inverse exists, it is unique. We know, as a matter of fact, that it does exist, because it is the well–known Moore–Penrose inverse, $A^\dagger$. However, we have not yet proved this. This is done in the next theorem.

THEOREM 4. (Urquhart [**832**]) For any finite matrix A of complex elements,

$$A^{(1,4)}AA^{(1,3)} = A^\dagger \tag{15}$$

PROOF. Let X denote LHS(15). It follows at once from Lemma 3 that $X \in A\{1,2\}$. Moreover, (15) gives

$$AX = AA^{(1,3)}\ ,\ XA = A^{(1,4)}A\ .$$

But, both $AA^{(1,3)}$ and $A^{(1,4)}A$ are Hermitian, by the definition of $A^{(1,3)}$ and $A^{(1,4)}$. Thus

$$X \in A\{1,2,3,4\}\ .$$

However, by Ex. 1, $A\{1,2,3,4\}$ contains at most a single element. Therefore, it contains exactly one element, which we denote by $A^\dagger$, and $X = A^\dagger$. □

6. Explicit formula for $A^\dagger$

C. C. MacDuffee apparently was the first to point out, in private communications about 1959, that a full–rank factorization of a matrix A leads to an explicit formula for its Moore–Penrose inverse, $A^\dagger$.

THEOREM 5. (MacDuffee). If $A \in \mathbb{C}_r^{m\times n}$, $r > 0$, has a full–rank factorization

$$A = FG\ , \tag{16}$$

then

$$A^\dagger = G^*(F^*AG^*)^{-1}F^*\ . \tag{17}$$

PROOF. First, we must show that F^*AG^* is nonsingular. By (16),

$$F^*AG^* = (F^*F)(GG^*)\ , \tag{18}$$

and both factors of the right member are $r \times r$ matrices. Also, by Lemma 4, both are of rank r. Thus, F^*AG^* is the product of two nonsingular matrices, and therefore nonsingular. Moreover, (18) gives

$$(F^*AG^*)^{-1} = (G^*G)^{-1}(F^*F)^{-1}\ .$$

Denoting by X the right member of (17), we now have

$$X = G^*(GG^*)^{-1}(F^*F)^{-1}F^*\ , \tag{19}$$

and it is easily verified that this expression for X satisfies the Penrose equations (1)–(4). As $A^\dagger$ is the sole element of $A\{1,2,3,4\}$, (17) is therefore established. □

Exercises.

EX. 14. Theorem 5 provides an alternative proof of the existence of the $\{1,2,3,4\}$–inverse (previously established by Theorem 4). However, Theorem 5 excludes the case $r = 0$. Complete the alternative existence proof by showing that if $r = 0$, (2) has a unique solution for X, and this X satisfies (1), (3) and (4).

EX. 15. Compute $A^\dagger$ for the matrix A of (0.79).

EX. 16. What is the most general $\{1,2\}$–inverse of the special matrix A of Ex. 4? What is its Moore–Penrose inverse?

EX. 17. Show that if $A = FG$ is a rank factorization, then

$$A^\dagger = G^\dagger F^\dagger\ . \tag{20}$$

EX. 18. Show that for every matrix A,

(a) $(A^\dagger)^\dagger = A$ (b) $(A^*)^\dagger = (A^\dagger)^*$
(c) $(A^T)^\dagger = (A^\dagger)^T$ (d) $A^\dagger = (A^*A)^\dagger A^* = A^*(AA^*)^\dagger$

Ex. 19. If $\mathbf{a}$ and $\mathbf{b}$ are column vectors, then

(a) $\mathbf{a}^\dagger = (\mathbf{a}^*\mathbf{a})^\dagger\mathbf{a}^*$ (b) $(\mathbf{a}\mathbf{b}^*)^\dagger = (\mathbf{a}^*\mathbf{a})^\dagger(\mathbf{b}^*\mathbf{b})^\dagger\mathbf{b}\mathbf{a}^*$.

Ex. 20. Show that if H is Hermitian and idempotent, $H^\dagger = H$.

Ex. 21. Show that $H^\dagger = H$ if and only if H^2 is Hermitian and idempotent and $\operatorname{rank} H^2 = \operatorname{rank} H$.

Ex. 22. If $D = \operatorname{diag}(d_1, d_2, \dots, d_n)$, show that $D^\dagger = \operatorname{diag}\left(d_1^\dagger, d_2^\dagger, \dots, d_n^\dagger\right)$.

Ex. 23. Let $J_k(0)$ be a Jordan block corresponding to the eigenvalue zero. Then $(J_k(0))^\dagger = (J_k(0))^T$, showing that for a square matrix A, $A^\dagger$ is in general not a polynomial in A (if it were, then A and $A^\dagger$ would commute).

Ex. 24. Let $A, B \in \mathbb{C}^{n\times n}$ be similar, i.e. $B = S^{-1}AS$ for some nonsingular S. Then, in general, $B^\dagger \neq S^{-1}A^\dagger S$.

Ex. 25. If U and V are unitary matrices, show that

$$(UAV)^\dagger = V^*A^\dagger U^*$$

for any matrix A for which the product UAV is defined.

In particular, if $A, B \in \mathbb{C}^{n\times n}$ are unitarily similar, i.e. $B = U^{-1}AU$ for some unitary matrix U, then $B^\dagger = U^{-1}A^\dagger U$.

7. Construction of {2}–inverses of prescribed rank

Following the proof of Theorem 1, we described A.G. Fisher's construction of a {1}–inverse of a given $A \in \mathbb{C}_r^{m\times n}$ having any prescribed rank between r and $\min(m, n)$, inclusive. From (2) it is easily deduced that

$$\operatorname{rank} A^{(2)} \leq r .$$

We note also that the $n \times m$ null matrix is a {2}–inverse of rank 0, and any $A^{(1,2)}$ is a {2}–inverse of rank r, by Theorem 2. For $r > 1$, is there a construction analogous to Fisher's for a {2}–inverse of rank s for arbitrary s between 0 and r? Using full–rank factorization, we can readily answer the question in the affirmative.

Let $X_0 \in A\{1, 2\}$ have a rank factorization

$$X_0 = YZ .$$

Then, $Y \in \mathbb{C}_r^{m\times r}$ and $Z \in \mathbb{C}_r^{r\times n}$, and (2) becomes

$$YZAYZ = YZ .$$

In view of Lemma 2, multiplication on the left by $Y^{(1)}$ and on the right by $Z^{(1)}$ gives (see Stewart [**787**])

$$ZAY = I_r . \tag{21}$$

Let Y_s denote the submatrix of Y consisting of the first s columns, and Z_s the submatrix of Z consisting of the first s rows. Then, both Y_s and Z_s are of full rank s, and it follows from (21) that

$$Z_sAY_s = I_s . \tag{22}$$

Now, let

$$X_s = Y_sZ_s .$$

Then, $\operatorname{rank} X_s = s$, by Ex. 7 and (22) gives

$$X_sAX_s = X_s .$$

Exercises.

EX. 26. For

$$A = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

find elements of $A\{2\}$ of ranks 1, 2, and 3, respectively.

EX. 27. With A as in Ex. 26, find a $\{2\}$–inverse of rank 2 having zero elements in the last two rows and the last two columns.

EX. 28. Show that there is at most one matrix X satisfying the three equations $AX = B$, $XA = D$, $XAX = X$ (Cline; see Cline and Greville [**199**]).

EX. 29. Let $A = FG$ be a rank factorization of $A \in \mathbb{C}_r^{m\times n}$, i.e., $F \in \mathbb{C}_r^{m\times r}$, $G \in \mathbb{C}_r^{r\times n}$. Then
(a) $G^{(i)}F^{(1)} \in A\{i\}$, $(i = 1, 2, 4)$, (b) $G^{(1)}F^{(j)} \in A\{j\}$, $(j = 1, 2, 3)$.

PROOF.
(a) $i = 1$:

$$FGG^{(1)}F^{(1)}FG = FG ,$$

since

$$F^{(1)}F = GG^{(1)} = I_r$$

by Lemma 2.
$i = 2$:

$$G^{(2)}F^{(1)}FGG^{(2)}F^{(1)} = G^{(2)}F^{(1)}$$

since

$$F^{(1)}F = I_r \ , \ G^{(2)}GG^{(2)} = G^{(2)} \ .$$

$i = 4$:

$$G^{(4)}F^{(1)}FG = G^{(4)}G = (G^{(4)}G)^* \ .$$

(b) Similarly proved, with the roles of F and G interchanged. □

EX. 30. Let A, F, G be as in Ex. 29. Then

$$A^\dagger = G^\dagger F^{(1,3)} = G^{(1,4)}F^\dagger \ .$$

Suggested further reading

Section 1. Urquhart [**833**].
Section 2. Rao [**674**], Sheffield [**754**].
Section 3. Rao ([**673**], [**676**]).
Section 4. Deutsch [**224**], Frame [**279**], Greville [**324**], Hartwig [**372**], Przeworska–Rolewicz and Rolewicz [**655**].
Section 5. Hearon and Evans [**404**], Rao [**676**], Sibuya [**762**].
Section 6. Sakallioğlu and Akdeniz [**727**].