APPENDIX B

Linear Equations and Generalized Inverses

B.1 INTRODUCTION

In various sections of the book, and particularly in Section 4.6, we have discussed how one solves a set of simultaneous linear equations where the matrix of coefficients has a regular inverse. Moreover, the pivotal method has been described as an illustrative computational procedure for obtaining the desired inverse.

In this appendix interest centers on a matrix of coefficients $\mathbf{A}$ for which no regular inverse $\mathbf{A}^{-1}$ exists. That is, $\mathbf{A}$ may be rectangular or, even if square, it may be singular.[1] The *generalized inverse* is a concept that provides a way to solve a set of consistent linear equations in which a regular inverse does not exist. As we shall note later, several different types of generalized inverse have been defined, although we concentrate here on only two variations, the Moore–Penrose inverse (Penrose, 1955) and the g inverse (Rao, 1962).

Before discussing generalized inverses, we provide a review of the types of solutions: (a) none, (b) one, or (c) infinitely many, that one can obtain in attempting to solve a set of simultaneous linear equations. Aspects of homogeneous equations and nonhomogeneous equations are described and illustrated numerically. Finally, a general method is evolved for solving sets of equations.

We then introduce the topic of generalized inverse in terms of a set of properties that such inverses are designed to satisfy. Following this, the Moore–Penrose type of inverse is defined and related to the concept of basic structure (or singular value decomposition) discussed in Chapter 5.

The second type of inverse, called the g inverse, is then introduced. This inverse is required to satisfy only one of the Penrose properties and, in practice, is easier to compute. Illustrations of its computation are presented, and this type of generalized inverse is related to procedures for solving linear equations. In so doing, a general procedure for computing inverses—regular or generalized—is described and related to earlier material involving the solution of simultaneous equations.

[1] That is, its determinant is zero.

B.2 SIMULTANEOUS LINEAR EQUATIONS

In the examples considered in the book we often had occasion to solve the system of equations

$$\boxed{\mathbf{Ax} = \mathbf{b}}$$

by means of matrix inversion. Recall that $\mathbf{A}$ is the matrix of coefficients, $\mathbf{x}$ is the vector of unknowns, and $\mathbf{b}$ is the vector of constants. In this case $\mathbf{A}$ was $n \times n$ and $r(\mathbf{A}) = n$. That is, $\mathbf{A}$ was square and nonsingular. The pivotal method was employed as a general solution technique.

However, suppose $\mathbf{A}$ is either rectangular or square singular so that a regular inverse does not exist. What happens then? Before launching into this topic, let us review some of the basic results related to solving a system of simultaneous linear equations. First, let us consider the set of equations

$$3x_1 + x_2 = 5$$

$$5x_1 + 2x_2 = 9$$

This set of equations, as could be easily verified, has the solution $x_1 = 1$, $x_2 = 2$. Furthermore, the solution is unique—only that specific set of values satisfies the set of equations.

Let us next consider the simultaneous equations

$$3x_1 + x_2 = 5$$

$$6x_1 + 2x_2 = 11$$

If we try to eliminate x_1 by taking twice the first equation and subtracting it from the second, we get the result

$$0 = 1$$

and, of course, something is wrong. This is most easily observed by noting that insofar as the left-hand side of the equations is concerned, the second equation is twice that of the first, but this relationship is not true for the right-hand side. The equations in this case are said to be *inconsistent,* and no solution exists.

Next, let us take the three equations

$$x_1 + x_2 + 3x_3 = 8$$

$$x_1 + 2x_2 + 6x_3 = 14$$

$$x_2 + 3x_3 = 6$$

We first eliminate x_1 from the second equation by subtracting the first from the second to get the pair of equations

$$x_2 + 3x_3 = 6$$

$$x_2 + 3x_3 = 6$$

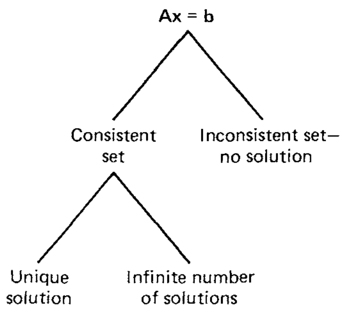

Fig. B.1 Tree diagram of types of solutions to a set of linear equations.

Note that these are identical. Thus, we have

$$x_2 = 6-3x_3$$
$$x_1 = 8-(6-3x_3)-3x_3 = 2$$

In this case, then, more than a single solution exists. For example, we have

$$x_1 = 2; \qquad x_2 = 3; \qquad x_3 = 1$$
$$x_1 = 2; \qquad x_2 = 0; \qquad x_3 = 2$$
$$x_1 = 2; \qquad x_2 = -3; \qquad x_3 = 3$$

and so on.

The tree diagram in Fig. B.1 shows the three cases of interest. We first want to examine whether the set of equations is consistent or not. If inconsistent, no solutions exist. If consistent, either a single (and unique) solution exists or an infinite number of solutions exist.

The three theorems[2] of interest in determining which condition prevails are:

1. A set of linear equations is consistent if and only if the rank of the augmented matrix (found by appending the **b** vector to the matrix of coefficients **A**) is equal to the rank of the original coefficients matrix.

2. A set of consistent linear equations has a unique solution if and only if the rank of the coefficients matrix **A** equals its order; that is, if and only if $r(\mathbf{A}) = n$, where **A** is of order $n \times n$ and n unknowns are present.

3. A consistent set of linear equations, where **A** is of rank k, can be solved for k unknowns in terms of the remaining $n - k$ unknowns if and only if the submatrix of coefficients (obtained from **A**) is of rank k.

With these theorems to guide us, let us return to the pair of equations

$$3x_1 + x_2 = 5$$
$$6x_1 + 2x_2 = 11$$

[2] Proofs can be bound in Graybill (1969).

and now form the matrix of coefficients **A** and the augmented matrix **M**:

$$\mathbf{A} = \begin{bmatrix} 3 & 1 \\ 6 & 2 \end{bmatrix}; \qquad \mathbf{M} = \begin{bmatrix} 3 & 1 & \vdots & 5 \\ 6 & 2 & \vdots & 11 \end{bmatrix}$$

By inspection we see that the second row of **A** is twice the first; hence $r(\mathbf{A}) = 1$. However, if we apply the reduction to echelon form procedure (from Section 4.7) to **M** we get the echelon matrix $\mathbf{H_M}$:

$$\mathbf{H_M} = \begin{bmatrix} 1 & 1/3 & \vdots & 5/3 \\ 0 & 0 & \vdots & 1 \end{bmatrix}$$

We note that both rows of $\mathbf{H_M}$ have at least one nonzero entry. Hence, $r(\mathbf{H_M}) = r(\mathbf{M}) = 2$ while $r(\mathbf{A}) = 1$; the set of equations is not consistent, and no solution exists.

Next, taking the equations

$$3x_1 + x_2 = 5$$

$$5x_1 + 2x_2 = 9$$

we have the matrices

$$\mathbf{A} = \begin{bmatrix} 3 & 1 \\ 5 & 2 \end{bmatrix}; \qquad \mathbf{M} = \begin{bmatrix} 3 & 1 & \vdots & 5 \\ 5 & 2 & \vdots & 9 \end{bmatrix}$$

After reduction to echelon form we obtain

$$\mathbf{H_A} = \begin{bmatrix} 1 & 1/3 \\ 0 & 1 \end{bmatrix}; \qquad \mathbf{H_M} = \begin{bmatrix} 1 & 1/3 & \vdots & 5/3 \\ 0 & 1 & \vdots & 2 \end{bmatrix}$$

and note, then, that $r(\mathbf{A}) = r(\mathbf{M}) = 2$, which is also equal to the order of **A**. In this case the equations are consistent, and a unique solution exists.

Finally, if we take the set of the three equations

$$x_1 + x_2 + 3x_3 = 8$$

$$x_1 + 2x_2 + 6x_3 = 14$$

$$x_2 + 3x_3 = 6$$

we have

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 3 \\ 1 & 2 & 6 \\ 0 & 1 & 3 \end{bmatrix}; \qquad \mathbf{M} = \begin{bmatrix} 1 & 1 & 3 & \vdots & 8 \\ 1 & 2 & 6 & \vdots & 14 \\ 0 & 1 & 3 & \vdots & 6 \end{bmatrix}$$

with associated echelon forms

$$\mathbf{H_A} = \begin{bmatrix} 1 & 1 & 3 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix}; \qquad \mathbf{H_M} = \begin{bmatrix} 1 & 1 & 3 & \vdots & 8 \\ 0 & 1 & 3 & \vdots & 6 \\ 0 & 0 & 0 & \vdots & 0 \end{bmatrix}$$

and, we note that the rank in each case is 2, while the order of **A** is 3. Thus, we can solve for $k = 2$ unknowns in terms of $n - k = 3 - 2 = 1$ remaining unknown. The results suggest a general approach to solving sets of simultaneous linear equations.

B.2.1 A General Procedure for Solving Linear Equations

As might be surmised at this point, in solving sets of linear equations we must determine whether a solution exists and, if so, whether the solution is unique or whether an infinity of solutions exists. The reduction of the matrix to echelon form via elementary row (or column) operations provides a practical way to find the rank of the coefficients matrix **A** and the rank of the augmented matrix **M**. As it turns out, however, reduction of a matrix to echelon form, followed by a few additional operations, provides us with a very general method for solving sets of simultaneous equations. As recalled, elementary row (column) operations permit

1. the interchange of two rows (columns);
2. the multiplication of each entry in a row (column) by any scalar $\lambda \neq 0$;
3. the addition, to each entry of some row (column), of λ times the corresponding element of some other row (column).

Each of these operations can be carried out on the rows of **A** by means of premultiplying **A** by a matrix that, in turn, can be obtained by performing the given elementary row operation on the *identity* matrix.[3] For example, let

$$\mathbf{A} = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$$

and assume that we wish to

1. interchange rows 1 and 2;
2. multiply row 2 by the scalar 4;
3. add twice row 3 to row 1.

If these three operations are separately performed on the identity matrix **I**, of order 3 x 3, we have, respectively,

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}; \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{bmatrix}; \quad \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

The reader can convince himself that premultiplication of **A** by each of the three matrices above will effect the desired row operation. Similarly, elementary column operations can be carried out by performing the indicated operation on the columns of a 2 x 2 identity matrix and *postmultiplying* **A** by the appropriate matrix. Successive operations are represented, of course, by a set of matrices whose sequence is determined by the desired sequence in which the elementary operations are to be performed.

[3] In the case of elementary column operations, the matrix **A** is *postmultiplied* by the specified elementary column operation on the identity matrix.

To illustrate the notion of a sequence of elementary row operations, let us simultaneously transform **A** and **I** by the three operations noted above, in the order given:

Interchange rows 1 and 2:

$$\mathbf{A}_1 = \begin{bmatrix} 2 & 5 \\ 1 & 4 \\ 3 & 6 \end{bmatrix}; \qquad \mathbf{B}_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Multiply row 2 by the scalar 4:

$$\mathbf{A}_2 = \begin{bmatrix} 2 & 5 \\ 4 & 16 \\ 3 & 6 \end{bmatrix}; \qquad \mathbf{B}_2 = \begin{bmatrix} 0 & 1 & 0 \\ 4 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Add twice row 3 to row 1:

$$\mathbf{A}_3 = \begin{bmatrix} 8 & 17 \\ 4 & 16 \\ 3 & 6 \end{bmatrix}; \qquad \mathbf{B}_3 = \begin{bmatrix} 0 & 1 & 2 \\ 4 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Finally, we note that $\mathbf{A}_3$ can be obtained from the combined row operations–in the *indicated order*–by

$$\mathbf{A}_3 = \overset{\mathbf{B}_3}{\begin{bmatrix} 0 & 1 & 2 \\ 4 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}} \overset{\mathbf{A}}{\begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}} = \begin{bmatrix} 8 & 17 \\ 4 & 16 \\ 3 & 6 \end{bmatrix}$$

Moreover, we also recall from Chapter 4 that $\mathbf{B}_1, \mathbf{B}_2, \ldots,$ is each nonsingular and that the rank of **A** is unaffected by elementary row (column) operations.

With this review information out of the way, the formal method for solving sets of simultaneous equations can be stated. First, we start with the augmented matrix **M**. Then we carry out elementary operations to reduce **M** to echelon form $\mathbf{H}_M$. As a final step we carry out additional elementary row operations on $\mathbf{H}_M$ so as to obtain an identity matrix in the subset of columns corresponding to **A**, the matrix of coefficients. To illustrate, let us take the matrix **M**, as used earlier for the set of two simultaneous equations for which the unique solution was $x_1 = 1, x_2 = 2$. We then apply the echelon reduction procedure to get $\mathbf{H}_M$. That is,

$$\begin{aligned} 3x_1 + x_2 &= 5 \\ 5x_1 + 2x_2 &= 9 \end{aligned}; \qquad \mathbf{M} = \left[\begin{array}{cc:c} 3 & 1 & 5 \\ 5 & 2 & 9 \end{array}\right]; \qquad \mathbf{H}_M = \left[\begin{array}{cc:c} 1 & 1/3 & 5/3 \\ 0 & 1 & 2 \end{array}\right]$$

Next, subtract 1/3 of the second row of $\mathbf{H_M}$ from the first row to get

$$\mathbf{N} = \left[\begin{array}{cc:c} 1 & 0 & 1 \\ 0 & 1 & 2 \end{array}\right]$$

Now, let us consider $\mathbf{N}$ in its original context of two linear equations

$$1x_1 + 0x_2 = 1$$

$$0x_1 + 1x_2 = 2$$

with the desired solution $x_1 = 1, x_2 = 2$.

Next, let us take the case of the two inconsistent equations:

$$\begin{aligned} 3x_1 + x_2 &= 5 \\ 6x_1 + 2x_2 &= 11 \end{aligned}; \qquad \mathbf{M} = \left[\begin{array}{cc:c} 3 & 1 & 5 \\ 6 & 2 & 11 \end{array}\right]; \qquad \mathbf{H_M} = \left[\begin{array}{cc:c} 1 & 1/3 & 5/3 \\ 0 & 0 & 1 \end{array}\right]$$

In this case we need go no further since in the second row of $\mathbf{H_M}$ we note the inconsistency:

$$0x_1 + 0x_2 = 1$$

which, of course, shows that this set of equations has no solution. Further evidence for this is found by examining the left-hand submatrix of $\mathbf{H_M}$; it is of rank 1 while $\mathbf{H_M}$ itself is of rank 2.

Finally, let us take the third example:

$$\begin{aligned} x_1 + x_2 + 3x_3 &= 8 \\ x_1 + 2x_2 + 6x_3 &= 14; \\ x_2 + 3x_3 &= 6 \end{aligned} \qquad \mathbf{M} = \left[\begin{array}{ccc:c} 1 & 1 & 3 & 8 \\ 1 & 2 & 6 & 14 \\ 0 & 1 & 3 & 6 \end{array}\right]; \qquad \mathbf{H_M} = \left[\begin{array}{ccc:c} 1 & 1 & 3 & 8 \\ 0 & 1 & 3 & 6 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

If we subtract the second row of $\mathbf{H_M}$ from the first, we get

$$\mathbf{N} = \left[\begin{array}{ccc:c} 1 & 0 & 0 & 2 \\ 0 & 1 & 3 & 6 \\ 0 & 0 & 0 & 0 \end{array}\right]; \qquad \begin{aligned} 1x_1 + 0x_2 + 0x_3 &= 2 \\ 0x_1 + 1x_2 + 3x_3 &= 6 \\ 0x_1 + 0x_2 + 0x_3 &= 0 \end{aligned}$$

In this case the best we can do is obtain a 2 x 2 identity matrix for the *first two* rows and columns of $\mathbf{N}$. As illustrated earlier, we can then transfer x_3 to the right-hand side, giving us

$$x_1 = 2; \qquad x_2 = 6 - 3x_3$$

If we then treat x_3 as a parameter, by setting it equal to (say) γ_3, we have

$$x_1 = 2; \qquad x_2 = 6 - 3\gamma_3; \qquad x_3 = \gamma_3$$

or, in vector form,

$$\mathbf{x} = \begin{bmatrix} 2 \\ 6 - 3\gamma_3 \\ \gamma_3 \end{bmatrix}$$

and an infinity of solutions exists depending upon what value we choose for γ_3.

B.2.2 Other Cases

In the cases examined so far we dealt with square matrices **A** of order $n \times n$ and vectors **x** and **b**, each of order $n \times 1$.

In the more general case, **A** can be of order $m \times n$. First, let us assume that $r(\mathbf{A}) = r(\mathbf{M}) = k$. If so, then at least one solution must exist. Next, let us assume that $k = n$, the number of unknowns. Since k cannot exceed m, the number of rows of **A** (or **M**), and $k = n$, then either $m = n$ or $m > n$. If $m = n = k$, then we know that **A** is square and nonsingular, and the solution is unique. However, if $m > n$, the echelon matrix $\mathbf{H_M}$ will have $m - k$ rows of zeros, and we can say that $m - k$ equations are redundant. If so, we proceed as before and solve k equations in $n = k$ variables. The submatrix **N**, of order $k \times k$, is still nonsingular, and the solution is still unique.

Next, suppose that $k < n$. If this case exists, then either $k = m$ or $k < m$. (We know that k cannot exceed m.) If we assume that $k = m < n$, we shall have an infinite number of solutions, as illustrated earlier. That is, $n - k$ of the unknowns can be treated as parameters.

Finally, assume that $k < m$ (and $k < n$). In this case, we have not only unknowns to spare but redundant equations as well. Not surprisingly, by following the formal method outlined earlier, we shall end up with an infinity of solutions *and* redundant equations in the bargain. To illustrate,

$$
\begin{aligned}
x_1 + x_2 + 3x_3 &= 8 \\
x_1 + 2x_2 + 6x_3 &= 14 \\
x_2 + 3x_3 &= 6 \\
2x_1 + 3x_2 + 9x_3 &= 22
\end{aligned}
\quad ; \quad
\mathbf{M} = \left[\begin{array}{ccc|c} 1 & 1 & 3 & 8 \\ 1 & 2 & 6 & 14 \\ 0 & 1 & 3 & 6 \\ 2 & 3 & 9 & 22 \end{array}\right]
$$

If we then reduce **M** to echelon form, we get

$$
\mathbf{H_M} = \left[\begin{array}{ccc|c} 1 & 1 & 3 & 8 \\ 0 & 1 & 3 & 6 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]
$$

As can be noted from $\mathbf{H_M}$, $r(\mathbf{A}) = r(\mathbf{M}) = 2$, and at least one solution exists. The next step is to find an identity submatrix by further elementary row operations on $\mathbf{H_M}$, giving us

$$
\mathbf{N} = \left[\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 3 & 6 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]
$$

with the solution

$$\mathbf{x} = \begin{bmatrix} 2 \\ 6 - 3\gamma_3 \\ \gamma_3 \end{bmatrix}$$

as found earlier. We see in this case that the fourth equation is redundant with the others. As a matter of fact, it is simply the sum of the first two equations (whereas the third equation represents their difference).

Thus, if we have m equations in n unknowns (of the form $\mathbf{Ax} = \mathbf{b}$) in which $r(\mathbf{A}) = r(\mathbf{M}) = k$, while $k < n$ and $k < m$, we have an infinite number of solutions in which $n - k$ variables can be treated as parameters and $m - k$ equations are redundant. In effect, then, the relationship between k and n deals with the question of a single versus infinite number of solutions, while the relationship between k and m concerns whether some of the equations (viz., $m - k$) are redundant.

B.2.3 Homogeneous Equations

Up to this point we have been discussing the case of $\mathbf{Ax} = \mathbf{b}$, involving nonhomogeneous equations. Sometimes the multivariate analyst will encounter sets of linear equations of the form

$$\boxed{\mathbf{Ax} = \mathbf{0}}$$

These are called homogeneous equations. First of all, we note that one possible solution is to let $\mathbf{x} = \mathbf{0}$. That is, if we assign 0 to each unknown, the equation is satisfied, since $\mathbf{A0} = \mathbf{0}$. This is called the *trivial* solution.

Viewed another way, if we append the zero vector to $\mathbf{A}$ to get the augmented matrix $\mathbf{M}$, then $r(\mathbf{A})$ will always equal $r(\mathbf{M})$, and the equations will always be consistent. Hence, we shall always have at least one solution, namely, the trivial solution.

The basic question, then, becomes one of determining the conditions under which solutions other than the trivial one exist. As with the case for nonhomogeneous equations, the answer depends on the relationship between $r(\mathbf{A}) = r(\mathbf{M}) = k$ and n, the number of unknowns. Since k cannot exceed n, we are left with the two cases: (a) $k = n$ and (b) $k < n$. The results in each case are contained in the following assertions:

1. Given a set of homogeneous linear equations $\mathbf{Ax} = \mathbf{0}$, involving m equations in n unknowns, a unique (and trivial) solution $\mathbf{x} = \mathbf{0}$ exists if $r(\mathbf{A}) = k = n$.
2. Given a set of homogeneous linear equations $\mathbf{Ax} = \mathbf{0}$, involving m equations in n unknowns, an infinite number of solutions exist if $r(\mathbf{A}) = k < n$.

We can illustrate these cases by the following set of $m = 3$ equations in $n = 2$ unknowns:

$$\begin{aligned} 2x_1 + x_2 &= 0 \\ 3x_1 + 2x_2 &= 0; \\ 7x_1 + 4x_2 &= 0 \end{aligned} \qquad \mathbf{A} = \begin{bmatrix} 2 & 1 \\ 3 & 2 \\ 7 & 4 \end{bmatrix}$$

With homogeneous equations, there is no point in obtaining **M**, the augmented matrix, since the appended column would be the zero vector. Rather, we can reduce **A** itself to echelon form, so as to get

$$\mathbf{H_A} = \begin{bmatrix} 1 & 1/2 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$

From $\mathbf{H_A}$ we see that $r(\mathbf{A}) = k = n = 2$. Next, if we subtract $\frac{1}{2}$ times row 2 from row 1, we get

$$\mathbf{N} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}; \qquad \begin{aligned} 1x_1 + 0x_2 &= 0 \\ 0x_1 + 1x_2 &= 0 \\ 0x_1 + 0x_2 &= 0 \end{aligned}$$

with the trivial solution $x_1 = 0, x_2 = 0$. Moreover, the third original equation is redundant and, as a matter of fact, equals twice the first equation plus the second.

An illustration of the second case, $r(\mathbf{A}) = k < n$, is the following set of $m = 2$ equations in $n = 3$ unknowns:

$$2x_1 + x_2 + 3x_3 = 0$$
$$x_1 + x_2 - x_3 = 0$$

As before we find the echelon form of **A** as

$$\mathbf{H_A} = \begin{bmatrix} 1 & 1/2 & 3/2 \\ 0 & 1 & -5 \end{bmatrix}.$$

and note that $r(\mathbf{H_A}) = r(\mathbf{A}) = 2$ and, hence, $k < n$. Next, we find the identity submatrix for the first two rows and two columns by subtracting $\frac{1}{2}$ of row 2 from row 1:

$$\mathbf{N} = \begin{bmatrix} 1 & 0 & 4 \\ 0 & 1 & -5 \end{bmatrix}.$$

The equations can now be written[4] as

$$1x_1 + 0x_2 = -4x_3$$
$$0x_1 + 1x_2 = 5x_3$$

and if we let $x_3 = \gamma_3$, we have the general solution, in vector form, as

$$\mathbf{x} = \begin{bmatrix} -4\gamma_3 \\ 5\gamma_3 \\ \gamma_3 \end{bmatrix} = \gamma_3 \begin{bmatrix} -4 \\ 5 \\ 1 \end{bmatrix}$$

[4] Note here that the implied vector $\begin{bmatrix} 4x_3 \\ -5x_3 \end{bmatrix}$ in the third column of the preceding matrix has simply been transposed to the right-hand side of the equation.

Note, here, that $m < n$; hence, $k < n$. If the number of equations is less than the number of unknowns, we must have an infinite number of solutions.[5]

One point of major difference between the present case and the counterpart case involving nonhomogeneous equations concerns the solution vector in situations involving $k < n$. In the case of homogeneous equations, we find that

$$\mathbf{x} = \begin{bmatrix} -4\gamma_3 \\ 5\gamma_3 \\ \gamma_3 \end{bmatrix}$$

and observe that *each* entry of x involves the arbitrary parameter γ_3. Thus, if we set $\gamma_3 = 0$, then $\mathbf{x} = \mathbf{0}$, and the trivial solution is included. Also, in the present illustration, where we have only one parameter γ_3, each solution is a scalar multiple of each other solution. For example, if we let $\gamma_3 = 1$ and then let $\gamma_3 = 2$, we have

$$\mathbf{x}_1 = \begin{bmatrix} -4 \\ 5 \\ 1 \end{bmatrix}; \qquad \mathbf{x}_2 = \begin{bmatrix} -8 \\ 10 \\ 2 \end{bmatrix} = 2\mathbf{x}_1$$

In contrast, if we reproduce the solution

$$\mathbf{x} = \begin{bmatrix} 2 \\ 6 - 3\gamma_3 \\ \gamma_3 \end{bmatrix}$$

in Section B.2.2 dealing with nonhomogeneous equations, we see that the first entry (2) does *not* involve γ_3.

To sum up, if some $\mathbf{x}^\circ \neq \mathbf{0}$ is a solution, then $\lambda\mathbf{x}^\circ$ (where λ is an arbitrary scalar) is also a solution in the case of homogeneous equations (a fact that was noted in Chapter 5 in the context of matrix eigenstructures), provided that $n - k = 1$. If *more than one* free parameter is found (i.e., $n - k > 1$), then it no longer follows that all solutions are scalar multiples of each other. However, if $\mathbf{x}^\circ \neq \mathbf{0}$ is a solution, it still follows that $\lambda\mathbf{x}^\circ$ is also a solution. This can be easily seen by noting that

$$\lambda\mathbf{A}\mathbf{x}^\circ = \mathbf{A}(\lambda\mathbf{x}^\circ) = \mathbf{0}$$

Again, as recalled from the discussion of eigenstructures in Chapter 5, if $\mathbf{A}$ is nonsingular and $r(\mathbf{A}) = k = n$, then nontrivial solutions cannot exist.

In summary, we can recapitulate the general method for solving either case:

1. nonhomogeneous equations of the form $\mathbf{A}\mathbf{x} = \mathbf{b}$;
2. homogeneous equations of the form $\mathbf{A}\mathbf{x} = \mathbf{0}$.

[5] As observed, in the case of homogeneous equations, either one solution exists (i.e., the trivial solution) or an infinity of solutions exists, depending upon the relationship between matrix rank and number of unknowns.

In the nonhomogeneous equations case the augmented matrix **M** is reduced to echelon form and then, via additional elementary row operations, to identity matrix form for the appropriate submatrix.[6] In the homogeneous equations case the analogous operations are carried out on the coefficients matrix **A**.

Figure B.2 recapitulates the various outcomes in tree diagram form. If we examine the case of nonhomogeneous equations first, we see that the primary outcomes are (a) none, (b) one, and (c) infinitely many solutions. We check $r(\mathbf{A})$ versus $r(\mathbf{M})$ to ascertain which condition prevails.

Assuming that $r(\mathbf{A}) = r(\mathbf{M}) = k$—and, hence, at least one solution exists—we check to see whether $k = n$, the number of unknowns. If so, then a unique solution exists. Next, if $k = m$, all equations are independent, while if $k < m$, some equations ($m - k$ of them) are redundant. Part (a) designates the first case in the tree diagram, while Part (b) designates the second.

If $k < n$, an infinite number of solutions exist. Again, we check to see if $k = m$ or $k < m$ so as to see if the equations are either all independent or not. Similar remarks pertain to the case of homogeneous equations with the exception, of course, that just one (the trivial solution) or infinitely many solutions exist in this instance; that is, if the system is homogeneous, it is consistent.

B.3 INTRODUCTORY ASPECTS OF GENERALIZED INVERSES

In Section B.2 a general method, utilizing reduction to echelon form followed by additional elementary row operations for finding an appropriate identity submatrix, was illustrated for solving sets of simultaneous equations. As was noted, provided that the equations are consistent, a solution—indeed an infinite number of solutions—can be found if $\mathbf{A}^{-1}$, the regular inverse of the coefficients matrix, does not exist.

At this point our interest centers on cases in which $\mathbf{A}^{-1}$, in the usual sense, does not exist, and yet we would still like to solve the set of equations of interest. This is the type of problem that the concept of *generalized inverse* has been developed to solve.

The literature on generalized inverses is of relatively recent origin, and its nomenclature and mathematical notation are not standard across authors. Two basic types of generalized inverse are discussed here:

1. the Moore–Penrose inverse (sometimes referred to as the pseudoinverse), written as $\mathbf{A}^{+}$;
2. the g inverse (sometimes referred to as the conditional inverse), written as $\mathbf{A}^{-}$.

However, the reader should be made aware of the fact that many types of generalized inverses exist—each obeying a particular set of properties.[7] Also, different ways have been developed to define these inverses. Our discussion in this appendix merely scratches the surface of an already broad and still expanding topic.

[6] The general procedure for converting the echelon reduced form to an identity submatrix is formalized in Section B.4. As noted there, the complete procedure involves reduction of **M** to what is known as Hermite form (in the case of a square coefficients matrix).

[7] The names and symbols for the Moore–Penrose and the g inverse follow those of Good (1969); other authors use different names, symbols, or both.

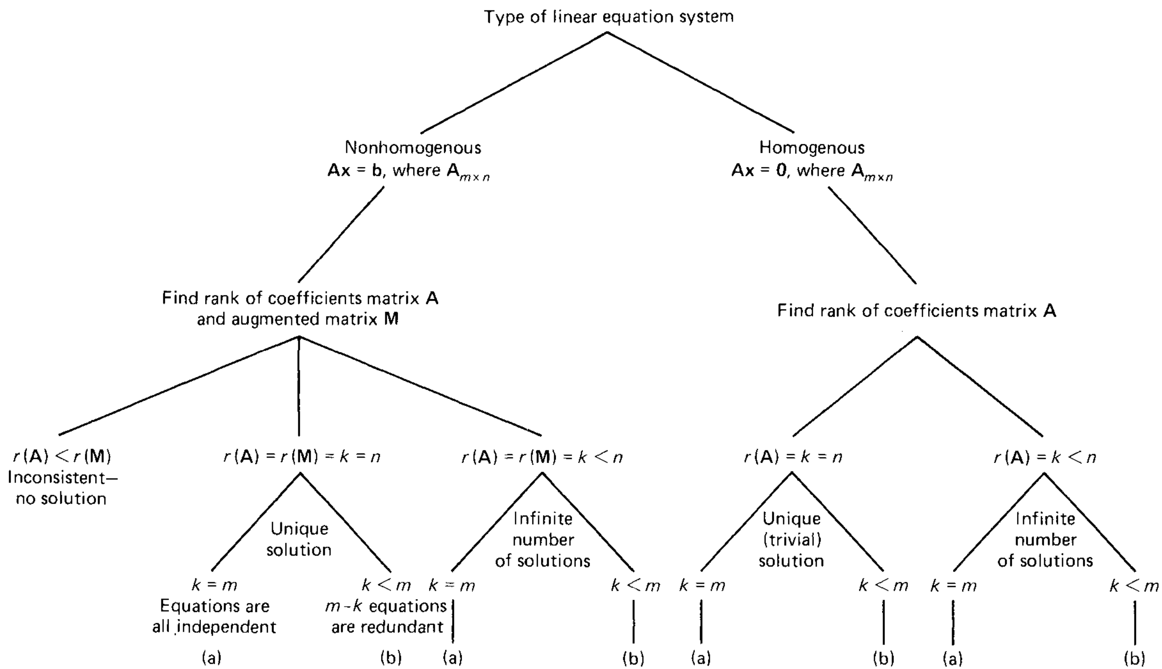

Fig. B.2 A classification of results obtained by solving simultaneous equations.

B.3.1 The Penrose Conditions

Research on generalized inverses goes back at least to 1920 with the work of Moore. Working independently, Penrose (1955) later defined the concept of a *unique* generalized inverse (now often called the Moore–Penrose inverse, denoted by $\mathbf{A}^+$) as a matrix that, in conjunction with the matrix $\mathbf{A}$ from which it is derived, satisfies four conditions:

(i) $\mathbf{AA}^+\mathbf{A} = \mathbf{A}$ (ii) $\mathbf{A}^+\mathbf{AA}^+ = \mathbf{A}^+$

(iii) $(\mathbf{AA}^+)' = \mathbf{AA}^+$ (symmetry) (iv) $(\mathbf{A}^+\mathbf{A})' = \mathbf{A}^+\mathbf{A}$ (symmetry)

There are alternative contexts in which to discuss $\mathbf{A}^+$. One context concerns the familiar case of solving a set of nonhomogeneous linear equations:

$$\boxed{\mathbf{Ax} = \mathbf{b}}$$

As we know, if $\mathbf{A}$ is nonsingular, a unique solution exists and is given by

$$\boxed{\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}}$$

Moreover, the reader can easily observe that $\mathbf{A}^{-1}$, the *regular* inverse, satisfies the four Penrose conditions. Still, cases might exist where $\mathbf{A}$ is either rectangular or else square and singular so that $\mathbf{A}^{-1}$ does not exist.

Suppose, then, that we define a matrix $\mathbf{A}$ of order $m \times n$ where $r(\mathbf{A}) = k \leqslant \min(m, n)$. If a Moore–Penrose inverse $\mathbf{A}^+$ exists, it will be of order $n \times m$; this must be so because $\mathbf{AA}^+$ is symmetric and, hence, square. It can be proved that for any matrix $\mathbf{A}$, there exists a unique matrix $\mathbf{A}^+$ that satisfies the four Penrose conditions.[8]

However, in solving a set of simultaneous equations, it is not always necessary that the solution be unique, as pointed out in Section B.2 in the context of echelon matrices. Moreover, it might be of interest to consider generalized inverses that obey only one (or more) of the four Penrose conditions. For example, if

$$\boxed{\mathbf{x} = \mathbf{A}^-\mathbf{b}}$$

is a solution to a set of consistent equations, $\mathbf{A}^-$ (also of order $n \times m$ if $\mathbf{A}$ is $m \times n$) need not be unique.

Indeed, in order for $\mathbf{A}^-$, called the g inverse, to exist, only the *first* of Penrose's four conditions, namely,

$$\boxed{\mathbf{AA}^-\mathbf{A} = \mathbf{A}}$$

need be satisfied. Thus, if our interest centers on solving simultaneous equations, and we do not require the generalized inverse to be unique, it may be easier to compute $\mathbf{A}^-$—and usually it is—than to compute $\mathbf{A}^+$, the Moore–Penrose generalized inverse.

Accordingly, we shall wish to examine both the "stronger" (Moore–Penrose inverse) case and the "weaker" (g inverse) case. We say "stronger" since all Moore–Penrose generalized inverses are g inverses but not the converse.

[8] A proof of this assertion can be found in Graybill (1969, p. 97).

We start with the Moore–Penrose generalized inverse by showing its relevance to basic structure (or singular value decomposition), a concept already discussed in Chapter 5. Not only does the concept of basic structure provide one way to define $\mathbf{A}^+$, but the present discussion should also help illuminate earlier remarks on matrix decomposition into its basic structure.

Then we turn to a discussion of $\mathbf{A}^-$, the g inverse. Our interest in this type of (nonunique) generalized inverse stems from the relative ease with which it can be computed and its close connection with solution methods for simultaneous equations that have already been discussed in Section B.2.

B.3.2 Left and Right General Inverses

In the discussion (Section 5.7) of the basic structure of an arbitrary matrix $\mathbf{A}$, we recall that $\mathbf{A}$, of order $m \times n$, can be decomposed into the triple product

$$\boxed{\mathbf{A} = \mathbf{P\Delta Q}'}$$

where $\mathbf{P}'\mathbf{P} = \mathbf{Q}'\mathbf{Q} = \mathbf{I}$ and Δ is diagonal of order $k \times k$, with $k \leqslant \min(m, n)$ positive entries that can be arranged in decreasing order of magnitude.

For the moment, let us place no restrictions on $\mathbf{A}$—it need be neither square nor basic and, hence, the rank of $\mathbf{A}$ may be less than its smaller order. Next, let us consider the following matrix $\mathbf{A}^+$, defined as being of rank k and of order $n \times m$:[9]

$$\boxed{\mathbf{A}^+ = \mathbf{Q\Delta}^{-1}\mathbf{P}'}$$

We obtain $\mathbf{A}^+$ from $\mathbf{A} = \mathbf{P\Delta Q}'$ by taking the reciprocals of the diagonal entries of Δ and then transposing the triple product $\mathbf{P\Delta}^{-1}\mathbf{Q}'$ into $\mathbf{Q\Delta}^{-1}\mathbf{P}'$.

Let us see what happens if we then premultiply $\mathbf{A}$ by $\mathbf{A}^+$:

$$\begin{aligned}\mathbf{A}^+\mathbf{A} &= (\mathbf{Q\Delta}^{-1}\mathbf{P}')(\mathbf{P\Delta Q}') = \mathbf{Q\Delta}^{-1}(\mathbf{P}'\mathbf{P})\mathbf{\Delta Q}' = \mathbf{Q\Delta}^{-1}(\mathbf{I}_{k\times k})\mathbf{\Delta Q}'\\ &= \mathbf{Q}(\mathbf{\Delta}^{-1}\mathbf{\Delta})\mathbf{Q}' = \mathbf{Q}(\mathbf{I}_{k\times k})\mathbf{Q}' = \mathbf{QQ}'\end{aligned}$$

What is found here is the major product moment of the right orthonormal section of $\mathbf{A}$. As we know $\mathbf{QQ}'$ is symmetric, since it is a product moment matrix.

Suppose next that $\mathbf{A}$ is postmultiplied by $\mathbf{A}^+$. Without going through the algebra, the result is

$$\mathbf{AA}^+ = \mathbf{PP}'$$

where $\mathbf{P}$ is the left orthonormal section of $\mathbf{A}$. Again, $\mathbf{PP}'$ is symmetric since it is a (major) product-moment matrix.

The reader will observe that $\mathbf{A}^+\mathbf{A}$ and $\mathbf{AA}^+$ are each symmetric. Moreover, we also have

$$\begin{array}{ll}\text{(i)}\ \ \mathbf{AA}^+\mathbf{A} = \mathbf{A} & \text{(ii)}\ \ \mathbf{A}^+\mathbf{AA}^+ = \mathbf{A}^+\\ \quad\ \ \mathbf{PP}'\mathbf{P\Delta Q}' = \mathbf{P\Delta Q}' & \quad\ \ \mathbf{QQ}'\mathbf{Q\Delta}^{-1}\mathbf{P}' = \mathbf{Q\Delta}^{-1}\mathbf{P}'\\ \quad\ \ = \mathbf{A} & \quad\ \ = \mathbf{A}^+\end{array}$$

[9] As recalled from Chapter 5, decomposition to basic structure is unique; also $\mathbf{A}^+$ is unique, given that $\mathbf{Q\Delta}^{-1}\mathbf{P}'$ is also of rank k.

and, hence, conclude that all four of the Penrose conditions are met. Notice that $r(\mathbf{A}^+) = r(\mathbf{A}) = k$ where $k \leqslant \min(m, n)$.

Next, let us suppose that $\mathbf{A}_{m \times n}$ (with $m > n$) is *basic* (as described in Chapter 5). If so, $k = n$ and $\mathbf{Q}'$ in the triple product $\mathbf{P\Delta Q}'$ will be square, of order $n \times n$, resulting in

$$\boxed{\mathbf{A}^+\mathbf{A} = \mathbf{QQ}' = \mathbf{Q}'\mathbf{Q} = \mathbf{I}_{n \times n}}$$

If this relation is met, we let $\mathbf{A}^+ = \mathbf{L}$; the matrix $\mathbf{L}$ is sometimes referred to as the *left pseudoinverse* of $\mathbf{A}$.

By the same token, if $\mathbf{A}$ is "horizontal," of order $m \times n$ ($m < n$), and basic, then $k = m$ and $\mathbf{P}$ will be square, of order $m \times m$, and we shall have

$$\boxed{\mathbf{AA}^+ = \mathbf{PP}' = \mathbf{P}'\mathbf{P} = \mathbf{I}_{m \times m}}$$

If this relation is met, we let $\mathbf{A}^+ = \mathbf{R}$; the matrix $\mathbf{R}$ is sometimes referred to as the *right pseudoinverse* of $\mathbf{A}$. Finally, it should be clear that if and only if $\mathbf{A}$ is both square and basic (i.e., nonsingular) will it possess *both* a left and right pseudoinverse and these inverses *will be the same.*[10]

However, suppose we return to the first case in which $\mathbf{A}$ is nonbasic:

$$\boxed{r(\mathbf{A}) = k < \min(m, n)}$$

It is possible, of course, to find a generalized inverse of $\mathbf{A}$ that meets only the first Penrose condition

$$\boxed{\mathbf{AA}^-\mathbf{A} = \mathbf{A}}$$

by defining $\mathbf{A}^-$ in terms of the square roots of *all* of the eigenvalues of either $\mathbf{AA}'$, if $m \leqslant n$, or $\mathbf{A}'\mathbf{A}$, if $m \geqslant n$, *including those eigenvalues that turn out to be zero.*

If so, we refer to this case as a g inverse of $\mathbf{A}$ and continue to denote it as $\mathbf{A}^-$. In this version of the generalized inverse the basic diagonal $\mathbf{\Delta}$ of $\mathbf{A}(=\mathbf{P\Delta Q}')$ has $\min(m, n) - k$ zeros, and $\mathbf{P}$ and $\mathbf{Q}'$ are no longer unique. It turns out, however, that

$$\boxed{\mathbf{x} = \mathbf{A}^-\mathbf{b}}$$

is still a solution to the set of consistent equations

$$\boxed{\mathbf{Ax} = \mathbf{b}}$$

and, in this sense, $\mathbf{A}^-$ is still a generalized inverse, specifically a g inverse.

[10] Still, it should be pointed out that although $\mathbf{A}^+$ is unique, the matrix $\mathbf{A}$—if basic but singular—will, in general, have an infinity of other matrices that satisfy $\mathbf{LA} = \mathbf{I}$ or $\mathbf{AR} = \mathbf{I}$, as the case may be. However, only one of this infinity of matrices will be the Moore–Penrose inverse. Furthermore, if $\mathbf{A}$ is nonsingular, only one matrix $\mathbf{A}^{-1}$ ($=\mathbf{A}^+$) exists. Thus, $\mathbf{A}$, if nonsingular, has exactly one inverse, $\mathbf{A}^{-1}$. If $\mathbf{A}$ is singular, it has an infinity of generalized inverses, one of which is the (uniquely specified) Moore-Penrose inverse, $\mathbf{A}^+$.

Why should we ever want to find a version of $\mathbf{Q}\boldsymbol{\Delta}^{-1}\mathbf{P}'$ whose diagonal is of larger order than $k \times k$, where $r(\mathbf{A}) = k$? Again, the motivation may be pragmatic in that it may be easier to compute $\mathbf{A}^-$ (as defined above) even though it is no longer unique.

We now turn to computational methods for finding the Moore–Penrose inverse, after which the g inverse $\mathbf{A}^-$ is discussed.

B.3.3 Some Numerical Procedures for Computing $\mathbf{A}^+$

To illustrate the computation of $\mathbf{A}^+$, let us consider the 3×2 matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 3 \\ 1 & 2 \\ 2 & 1 \end{bmatrix}$$

If $r(\mathbf{A}) = n = 2$, we should be able to find a left general inverse such that $\mathbf{LA} = \mathbf{I}_{2\times 2}$.

Using the procedure described in Section 5.7.3, we first find the product-moment matrix with the smaller order, in this case the minor product moment

$$\mathbf{A}'\mathbf{A} = \begin{bmatrix} 6 & 7 \\ 7 & 14 \end{bmatrix}$$

and solve for its eigenstructure

$$\mathbf{D} = \begin{bmatrix} 18.062 & 0 \\ 0 & 1.938 \end{bmatrix}; \qquad \mathbf{Q} = \begin{bmatrix} -0.502 & 0.865 \\ -0.865 & -0.502 \end{bmatrix}$$

We then compute the basic diagonal and its inverse

$$\boldsymbol{\Delta} = \mathbf{D}^{1/2} = \begin{bmatrix} 4.250 & 0 \\ 0 & 1.392 \end{bmatrix}; \qquad \boldsymbol{\Delta}^{-1} = \begin{bmatrix} 0.235 & 0 \\ 0 & 0.718 \end{bmatrix}$$

and then solve for $\mathbf{P}$:

$$\mathbf{P} = \mathbf{A}\mathbf{Q}\boldsymbol{\Delta}^{-1} = \begin{bmatrix} 1 & 3 \\ 1 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} -0.502 & 0.865 \\ -0.865 & -0.502 \end{bmatrix} \begin{bmatrix} 0.235 & 0 \\ 0 & 0.718 \end{bmatrix}$$

$$\mathbf{P} = \begin{bmatrix} -0.728 & -0.460 \\ -0.525 & -0.099 \\ -0.439 & -0.882 \end{bmatrix}$$

The matrix $\mathbf{A}$ is now expressed in terms of basic structure as $\mathbf{A} = \mathbf{P}\boldsymbol{\Delta}\mathbf{Q}'$. The left general inverse is then

$$\mathbf{A}^+ = \mathbf{L} = \mathbf{Q}\boldsymbol{\Delta}^{-1}\mathbf{P}' = \begin{bmatrix} -0.2 & 0 & 0.6 \\ 0.313 & 0.142 & -0.228 \end{bmatrix}$$

and we have the desired result:

$$\mathbf{LA} = \begin{bmatrix} -0.2 & 0 & 0.6 \\ 0.313 & 0.142 & -0.228 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 1 & 2 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

However, now let us take the case where **A** is nonbasic. To illustrate, consider

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 1 & 2 \\ 2 & 4 \end{bmatrix}$$

In this case we know that **A** is not basic since the second column is twice that of the first column and $r(\mathbf{A}) = 1$. We first find

$$\mathbf{A}'\mathbf{A} = \begin{bmatrix} 6 & 12 \\ 12 & 24 \end{bmatrix} = 6 \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$$

By procedures identical to those just illustrated, the basic structure of **A** is then found to be

$$\mathbf{A} = \overset{\mathbf{P}}{\begin{bmatrix} -0.408 \\ -0.408 \\ -0.816 \end{bmatrix}} \overset{\Delta}{[5.477]} \overset{\mathbf{Q}'}{[-0.477 \quad -0.894]} = \begin{bmatrix} 1 & 2 \\ 1 & 2 \\ 2 & 4 \end{bmatrix}$$

and $\mathbf{A}^+$, of rank $k(\mathbf{A}) = 1$, is

$$\mathbf{A}^+ = \mathbf{Q}\Delta^{-1}\mathbf{P}' = \begin{bmatrix} 0.033 & 0.033 & 0.067 \\ 0.067 & 0.067 & 0.133 \end{bmatrix}$$

The reader can verify that the four Penrose conditions are met, although now it is no longer true that $\mathbf{A}^+\mathbf{A} = \mathbf{I}$.

Computing $\mathbf{A}^+$ after first solving for the basic structure of a matrix is only one of many solution methods. By way of contrast, let us consider another method, due to Penrose himself (1956), that can also be used to find $\mathbf{A}^+$. This method involves implementation of a fairly simple algorithm that entails the following steps:

1. Compute $\mathbf{B} = \mathbf{A}'\mathbf{A}$.
2. Let $\mathbf{C}_1 = \mathbf{I}$, the identity matrix.
3. Compute $\mathbf{C}_{i+1} = \mathbf{I}(1/i)\mathrm{tr}(\mathbf{C}_i\mathbf{B}) - \mathbf{C}_i\mathbf{B}$ for $i = 1, 2, \ldots, k-1$.[11]
4. Compute $k\mathbf{C}_k\mathbf{A}'/\mathrm{tr}(\mathbf{C}_k\mathbf{B})$, to get $\mathbf{A}^+$.
5. Also, it will be found that $\mathbf{C}_{k+1}\mathbf{B} = \phi$; $\mathrm{tr}(\mathbf{C}_k\mathbf{B}) \neq 0$, so that $r(\mathbf{B}) = r(\mathbf{A}) = k$.

[11] The reader should recall that the trace (tr) of a square matrix $\mathbf{A}_{n \times n}$ is equal to the sum of its main diagonal elements:

$$\mathrm{tr}(\mathbf{A}) = \sum_{i=1}^{n} a_{ii}$$

Applying the procedure to the last problem (where **A** is nonbasic) gives us

1. $\mathbf{B} = \mathbf{A}'\mathbf{A} = \begin{bmatrix} 6 & 12 \\ 12 & 24 \end{bmatrix}$

2. $\mathbf{C}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}; \quad \mathbf{C}_1\mathbf{B} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 6 & 12 \\ 12 & 24 \end{bmatrix} = \begin{bmatrix} 6 & 12 \\ 12 & 24 \end{bmatrix}; \quad \text{tr}\,\mathbf{C}_1\mathbf{B} = 30$

3. $\mathbf{C}_2 = \mathbf{I}\,\text{tr}(\mathbf{C}_1\mathbf{B}) - \mathbf{C}_1\mathbf{B} = \begin{bmatrix} 30 & 0 \\ 0 & 30 \end{bmatrix} - \begin{bmatrix} 6 & 12 \\ 12 & 24 \end{bmatrix} = \begin{bmatrix} 24 & -12 \\ -12 & 6 \end{bmatrix}; \quad \mathbf{C}_2\mathbf{B} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

Since $\mathbf{C}_2\mathbf{B} = \phi$ and $\text{tr}(\mathbf{C}_1\mathbf{B}) \neq 0$, we know that $r(\mathbf{B}) = r(\mathbf{A}) = 1$; we can then go on to find $\mathbf{A}^+$.

4. $\mathbf{A}^+ = \dfrac{1\mathbf{C}_1\mathbf{A}'}{\text{tr}(\mathbf{C}_1\mathbf{B})} = \dfrac{1}{30}\begin{bmatrix} 1 & 1 & 2 \\ 2 & 2 & 4 \end{bmatrix} = \begin{bmatrix} 0.033 & 0.033 & 0.067 \\ 0.067 & 0.067 & 0.133 \end{bmatrix}$

We find, of course, the same solution for $\mathbf{A}^+$ as found earlier. The Penrose procedure, like the basic structure approach, can be used to find $\mathbf{A}^+$, whether or not **A** is basic. To complete the discussion, let us apply the Penrose computational procedure to the first case, where **A** is basic:

1. $\mathbf{A} = \begin{bmatrix} 1 & 3 \\ 1 & 2 \\ 2 & 1 \end{bmatrix}; \quad \mathbf{B} = \mathbf{A}'\mathbf{A} = \begin{bmatrix} 6 & 7 \\ 7 & 14 \end{bmatrix}$

2. $\mathbf{C}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}; \quad \mathbf{C}_1\mathbf{B} = \begin{bmatrix} 6 & 7 \\ 7 & 14 \end{bmatrix}; \quad \text{tr}(\mathbf{C}_1\mathbf{B}) = 20$

3. $\mathbf{C}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} 20 - \begin{bmatrix} 6 & 7 \\ 7 & 14 \end{bmatrix} = \begin{bmatrix} 14 & -7 \\ -7 & 6 \end{bmatrix}; \quad \mathbf{C}_2\mathbf{B} = \begin{bmatrix} 35 & 0 \\ 0 & 35 \end{bmatrix}$

4. $\mathbf{A}^+ = \dfrac{2\mathbf{C}_2\mathbf{A}'}{\text{tr}(\mathbf{C}_2\mathbf{B})}, \quad \text{where} \quad \text{tr}(\mathbf{C}_2\mathbf{B}) = 70$

$$= \frac{2}{70}\begin{bmatrix} 14 & -7 \\ -7 & 6 \end{bmatrix}\begin{bmatrix} 1 & 1 & 2 \\ 3 & 2 & 1 \end{bmatrix} = \begin{bmatrix} -0.2 & 0 & 0.6 \\ 0.313 & 0.142 & -0.228 \end{bmatrix}$$

Note that the Penrose computational procedure involves less computation—particularly, no need to find eigenstructures—than the method based on matrix decomposition via basic structure.[12]

[12] Note also that

$$\mathbf{C}_3 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}(70/2) - \begin{bmatrix} 35 & 0 \\ 0 & 35 \end{bmatrix} = \phi$$

and, $\text{tr}(\mathbf{C}_2 B) \neq 0$; hence, $r(\mathbf{A}) = 2$.

B.3.4 Some Properties of the Moore–Penrose Inverse

In many respects the Moore–Penrose inverse A^+ acts like a regular inverse A^{-1}. (Indeed, A^+ equals A^{-1} if A is nonsingular.) However, even in other cases, A^+ possesses a number of properties, many of which are similar to those displayed by the regular inverse. Some of the more important of these properties are listed below:

1. The Moore–Penrose inverse of the transpose of A is the transpose of the Moore–Penrose inverse of A: $(A')^+ = (A^+)'$.
2. The Moore–Penrose inverse of A^+ is equal to A: $(A^+)^+ = A$.
3. The rank of the Moore–Penrose inverse of A is equal to the rank of A: $r(A^+) = r(A)$.
4. For any matrix A, $(A'A)^+ = A^+(A')^+$.
5. For any matrix A, $(AA^+)^+ = AA^+$; $(A^+A)^+ = A^+A$.
6. If $A = A'$, then $A^+ = (A^+)'$.
7. If $A = A'$, then $AA^+ = A^+A$.
8. If A is nonsingular, then $A^{-1} = A^+$.
9. If A is an $m \times n$ matrix of rank m, then $A^+ = A'(AA')^{-1}$ and $AA^+ = I$ (as related to Section B.3.2).
10. If A is an $m \times n$ matrix of rank n, then $A^+ = (A'A)^{-1}A'$ and $A^+A = I$ (as related to Section B.3.2).

In addition to the properties listed above, the Moore–Penrose inverse figures prominently in the solution of sets of linear equations. More specifically, given the set of nonhomogeneous equations

$$\boxed{Ax = b}$$

where A is of order $m \times n$ and b is an $m \times 1$ vector of constants, the system of equations is consistent if and only if

$$\boxed{AA^+b = b}$$

Second, given that the system is consistent (and, hence, has at least one solution), then for each $n \times 1$ vector γ, the $n \times 1$ vector x is a solution where

$$\boxed{x = A^+b + (I - A^+A)\gamma}$$

and every solution to the system can be so written for some $n \times 1$ vector γ.

As just indicated, the Moore–Penrose generalized inverse, assuming it can be found easily, provides a way to solve sets of linear equations. While in principle, we could always compute A^+ in solving sets of simultaneous equations, it is usually the case that we do not need the stronger properties of the Moore–Penrose inverse to get the job done.

However, as recalled, a g inverse A^- also provides a solution to a set of consistent equations, albeit one that is not unique but, on the other hand, one that is relatively easy to compute. Accordingly, we now turn to a discussion of the g inverse A^- and its role in solving sets of simultaneous equations.

B.4 THE g INVERSE

If we let $\mathbf{A}$ be an $m \times n$ matrix, a matrix $\mathbf{A}^-$, of order $n \times m$, is defined to be a g inverse of $\mathbf{A}$ if and only if it satisfies the first of the Penrose conditions

$$\boxed{\mathbf{A}\mathbf{A}^-\mathbf{A} = \mathbf{A}}$$

As pointed out earlier, the Moore–Penrose inverse of $\mathbf{A}$ is also a g inverse of $\mathbf{A}$, but the converse does not hold in general. Moreover, in general $\mathbf{A}^-$ is not unique for a given $\mathbf{A}$.

A g inverse is particularly useful in the practical setting of solving sets of simultaneous equations. Fully analogous to the Moore–Penrose inverse, the system of equations $\mathbf{Ax} = \mathbf{b}$ has a solution if and only if

$$\boxed{\mathbf{A}\mathbf{A}^-\mathbf{b} = \mathbf{b}}$$

Second, given that the system is consistent, then for each $n \times 1$ vector γ, the $n \times 1$ vector $\mathbf{x}$ is a solution where $\mathbf{x}$ is

$$\boxed{\mathbf{x} = \mathbf{A}^-\mathbf{b} + (\mathbf{I} - \mathbf{A}^-\mathbf{A})\gamma}$$

Finally, every solution to the system can be so written for some $n \times 1$ vector γ.

The value of a g inverse relates to its relative ease of calculation, particularly by means of the echelon form of a matrix, as considered in Section B.2. However, before discussing a general approach to computing $\mathbf{A}^-$ (in the context of solving sets of equations), we consider the concept of Hermite form. The Hermite form of a matrix provides the key concept for obtaining $\mathbf{A}^-$.

B.4.1 The Hermite Form of a Square Matrix

A square ($n \times n$) matrix $\mathbf{J}$ is defined to be in (upper) Hermite form if and only if it satisfies the following conditions:

1. $\mathbf{J}$ is upper triangular.
2. Only zeros and ones are on its main diagonal.
3. If a row has a zero on the diagonal, then every entry in the row is zero.
4. If a row has a one on the diagonal, then every other entry is zero in the column in which the one appears.

If $\mathbf{J}$ is of Hermite form, it is also the case that

$$\boxed{\mathbf{J} = \mathbf{J}^2}$$

and $\mathbf{J}$ is said to be *idempotent*. Moreover, for any $n \times n$ matrix $\mathbf{A}$, there exists a nonsingular matrix $\mathbf{G}$ such that

$$\boxed{\mathbf{GA} = \mathbf{J}_\mathbf{A}}$$

and so $\mathbf{A}$ can *always* be reduced to Hermite form via $\mathbf{G}$. (However, $\mathbf{G}$ is nonunique, in general, although $\mathbf{J_A}$ will be.)

Just as was the case in reducing $\mathbf{A}$ to echelon form $\mathbf{H_A}$, we can use elementary row operations to reduce $\mathbf{A}_{n \times n}$ to Hermite form $\mathbf{J_A}$ [13] The matrix $\mathbf{G}$ is the nonsingular matrix that brings about the reduction of $\mathbf{A}_{n \times n}$ to Hermite form.

Moreover, once this is done it turns out that the job is finished since $\mathbf{A}^-$ can be defined as

$$\boxed{\mathbf{A}^- = \mathbf{G}}$$

That is, $\mathbf{G}$ is the g inverse of $\mathbf{A}$, and all we have to do to find $\mathbf{G}$ is to reduce $\mathbf{A}$ to Hermite form via elementary row operations while performing companion operations on $\mathbf{I}$.

However, since the Hermite form does not exist for rectangular matrices, a slight modification is required to find $\mathbf{A}^- = \mathbf{G}$ when $\mathbf{A}$ is rectangular. If $\mathbf{A}$ is vertical ($m \times n$, with $m > n$), we can append a set of $\mathbf{0}$ column vectors to make $\mathbf{A}$ square. That is,

$$\mathbf{A}_0 = \left[\mathbf{A} \;\vdots\; \phi\right]$$

where $\mathbf{A}_0$ is $m \times m$. Then, if $\mathbf{G}$ is a nonsingular matrix, such that $\mathbf{G}_0\mathbf{A}_0 = \mathbf{J}_{\mathbf{A}_0}$, where $\mathbf{J}_{\mathbf{A}_0}$ is the Hermite form of $\mathbf{A}_0$, we have

$$\mathbf{G}_0 = \begin{bmatrix} \mathbf{G} \\ \cdots \\ \mathbf{G}_1 \end{bmatrix}$$

where $\mathbf{G}$ is the upper $n \times m$ submatrix of $\mathbf{G}_0$. This is the g inverse of $\mathbf{A}$. Similarly, if $\mathbf{A}$ is horizontal ($m \times n$, with $n > m$), we can append a set of $\mathbf{0}'$ row vectors to make $\mathbf{A}$ square and proceed to find $\mathbf{G}_{n \times m}$, the left-hand submatrix of $\mathbf{G}_0$.[14]

The strategy should now be clear. In the rectangular case, we make $\mathbf{A}$ square by adding columns or rows of zeros, as the case may be. We then find a nonsingular matrix that reduces $\mathbf{A}_0$ to Hermite form. The matrix $\mathbf{G}$ is the g inverse of $\mathbf{A}$.

However, one more facet of the problem has to be introduced before proceeding to find $\mathbf{A}^-$. In Section B.2 a general procedure was introduced for solving sets of simultaneous equations via reduction of either the coefficients matrix $\mathbf{A}$ or the augmented matrix $\mathbf{M}$ to echelon form. As might be surmised, if $\mathbf{A}$ is already square, or made square by appending columns (or rows) of zeros, the Hermite form $\mathbf{J_A}$ of $\mathbf{A}$ can be obtained from its echelon form $\mathbf{H_A}$. This is done by transforming rows of $\mathbf{H}$, via additional elementary row operations, until $\mathbf{J_A}$ is found. The matrix $\mathbf{G}$ that summarizes the *full set of elementary row operations* used in reducing $\mathbf{A}$ to $\mathbf{H_A}$ and then $\mathbf{H_A}$ to $\mathbf{J_A}$ is

[13] As will be shown, if $\mathbf{A}$ is square to begin with (or can be made square by procedures to be described later), we can compute $\mathbf{J_A}$ via additional elementary operations on $\mathbf{H_A}$.

[14] The matrix would appear as

$$\mathbf{G}_0 = \left[\mathbf{G} \;\vdots\; \mathbf{G}_1\right].$$

The next section shows some numerical examples of the general procedure, including a case in which $\mathbf{A}$ is rectangular.

A^-, the desired g inverse. In general, A^- will not be unique.[15] However, the matrix J_A in Hermite form *is* unique for a given matrix A.

Before proceeding with the computation of A^-, three additional properties of matrices in Hermite form are of interest to note:

1. The Hermite form J_A of A has the same rank as A.
2. If A, of order $n \times n$, is nonsingular, then J_A is the $n \times n$ identity matrix I.
3. The rank of A is equal to the number of diagonal elements of J_A that are equal to unity.

With the foregoing comments as background, a general procedure can now be stated for finding A^-:

1. If A is rectangular, make it square by appending columns (or rows) of zeros.
2. Via elementary row operations reduce A to echelon form and then to Hermite form. At the same time, perform the same operations on I, the associated $n \times n$ identity matrix.
3. If A is nonsingular, then its Hermite form is I and $A^- = A^{-1}$.
4. If A is singular, then I will be transformed to $G = A^-$ as A is being reduced to J_A, its Hermite form. And A^- will be the g inverse of interest.

B.4.2 Some Numerical Examples

Let us now consider some illustrations of finding A^- by means of the method presented above. First, let us take a nonsingular matrix A and its companion identity matrix:

$$\mathbf{A} = \begin{bmatrix} 1 & 4 & 2 \\ 2 & 5 & 3 \\ 3 & 6 & 5 \end{bmatrix}; \qquad \mathbf{I} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

As outlined earlier, the task is to reduce A to echelon form H_A and then into Hermite form J_A via a series of elementary row operations. Each elementary row operation that is performed on A is also performed concurrently on the associated starting identity matrix I. As A is reduced to Hermite form J_A, I is transformed to A^-, the desired g inverse.

The reader should note the similarity of this procedure to that followed in Section B.2. In the present case J_A, the Hermite form of A, takes on the role of the identity submatrix computed from the echelon matrix in Section 3.2.

We can now start the row operations, bearing in mind that these, in general, are not unique. We first subtract twice row 1 from row 2 and subtract 3 times row 1 from row 3:

$$\begin{bmatrix} 1 & 4 & 2 \\ 0 & -3 & -1 \\ 0 & -6 & -1 \end{bmatrix}; \qquad \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix}$$

[15] The reason why A^- (=G) is not unique is simply because, in general, there are different sets of elementary row operations (summarized in G) that can lead to J_A—as a matter of fact, an infinity of such sets.

Subtract twice row 2 from row 3:

$$\begin{bmatrix} 1 & 4 & 2 \\ 0 & -3 & -1 \\ 0 & 0 & 1 \end{bmatrix}; \qquad \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & -2 & 1 \end{bmatrix}$$

Multiply row 2 by $-1/3$:

$$\mathbf{H_A} = \begin{bmatrix} 1 & 4 & 2 \\ 0 & 1 & 1/3 \\ 0 & 0 & 1 \end{bmatrix}; \qquad \begin{bmatrix} 1 & 0 & 0 \\ 2/3 & -1/3 & 0 \\ 1 & -2 & 1 \end{bmatrix}$$

At this point we note that **A** is in echelon form $\mathbf{H_A}$ and that $r(\mathbf{A}) = 3$. The next task is to reduce $\mathbf{H_A}$ to Hermite form $\mathbf{J_A}$. To do this, subtract 4 times row 2 from row 1:

$$\begin{bmatrix} 1 & 0 & 2/3 \\ 0 & 1 & 1/3 \\ 0 & 0 & 1 \end{bmatrix}; \qquad \begin{bmatrix} -5/3 & 4/3 & 0 \\ 2/3 & -1/3 & 0 \\ 1 & -2 & 1 \end{bmatrix}$$

Subtract 2/3 of row 3 from row 1 and subtract 1/3 of row 3 from row 2:

$$\mathbf{J_A} = \mathbf{I} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}; \qquad \begin{bmatrix} -7/3 & 8/3 & -2/3 \\ 1/3 & 1/3 & -1/3 \\ 1 & -2 & 1 \end{bmatrix} = \mathbf{G} = \mathbf{A}^- = \mathbf{A}^{-1}$$

The reader can then check to see that

$$\mathbf{AA}^- = \mathbf{AA}^{-1} = \mathbf{I}$$

Note also that the Hermite form $\mathbf{J_A}$ of a nonsingular matrix **A** is an identity matrix of the same order.

Next, we consider the case where **A** is square but not of full rank. In Section B.2.1 we encountered a matrix of this type:

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 3 \\ 1 & 2 & 6 \\ 0 & 1 & 3 \end{bmatrix}; \qquad \mathbf{I} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

If we subtract the first row from the second, we get

$$\begin{bmatrix} 1 & 1 & 3 \\ 0 & 1 & 3 \\ 0 & 1 & 3 \end{bmatrix}; \qquad \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Next we subtract the second row from the third:

$$\mathbf{H_A} = \begin{bmatrix} 1 & 1 & 3 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix}; \qquad \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & -1 & 1 \end{bmatrix}$$

and note that we have the echelon form of $\mathbf{A}$, $\mathbf{H_A}$, in which $r(\mathbf{A}) = 2$. Next, we subtract the second row from the first:

$$\mathbf{J_A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix}; \qquad \begin{bmatrix} 2 & -1 & 0 \\ -1 & 1 & 0 \\ 1 & -1 & 1 \end{bmatrix} = \mathbf{A}^-$$

to obtain $\mathbf{J_A}$, the Hermite form of $\mathbf{A}$. We find $\mathbf{A}^-$ as well and note that

$$\mathbf{J_A} = \mathbf{J_A}^2$$

as it should, since $\mathbf{J_A}$ is idempotent.

We then solve for $\mathbf{x}$ as

$$\mathbf{x} = \mathbf{A}^-\mathbf{b}$$

where, from Section B.2.1, $\mathbf{b}' = (8, 14, 6)$, so that

$$\mathbf{x} = \overset{\mathbf{A}^-}{\begin{bmatrix} 2 & -1 & 0 \\ -1 & 1 & 0 \\ 1 & -1 & 1 \end{bmatrix}} \overset{\mathbf{b}}{\begin{bmatrix} 8 \\ 14 \\ 6 \end{bmatrix}} = \begin{bmatrix} 2 \\ 6 \\ 0 \end{bmatrix}$$

However, as noted earlier, when the rank of $\mathbf{A}$ is less than the number of unknowns, there is an infinite number of solutions. Hence, we express $\mathbf{x}$ in the general form:

$$\mathbf{x} = \mathbf{A}^-\mathbf{b} + (\mathbf{I} - \mathbf{A}^-\mathbf{A})\boldsymbol{\gamma}$$

where $\boldsymbol{\gamma}$ is an arbitrary vector. Specifically, we have

$$\mathbf{x} = \overset{\mathbf{A}^-\mathbf{b}}{\begin{bmatrix} 2 \\ 6 \\ 0 \end{bmatrix}} + \left\{ \overset{\mathbf{I}}{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}} - \overset{\mathbf{A}^-\mathbf{A}}{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix}} \right\} \overset{\boldsymbol{\gamma}}{\begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \end{bmatrix}}$$

$$= \begin{bmatrix} 2 \\ 6 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ -3\gamma_3 \\ \gamma_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 6 - 3\gamma_3 \\ \gamma_3 \end{bmatrix}$$

where γ_3 is considered as an arbitrary parameter. Note that this is the same result as found in Section B.2.1.

Finally, let us consider the case where $\mathbf{A}$ is rectangular. As an illustration, we again return to Section B.2.2 and consider the 4 x 3 vertical matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 3 \\ 1 & 2 & 6 \\ 0 & 1 & 3 \\ 2 & 3 & 9 \end{bmatrix}$$

After appending a column vector of zeros to make $\mathbf{A}$ square, we have

$$\mathbf{A}_0 = \left[\begin{array}{ccc:c} 1 & 1 & 3 & 0 \\ 1 & 2 & 6 & 0 \\ 0 & 1 & 3 & 0 \\ 2 & 3 & 9 & 0 \end{array}\right]; \qquad \mathbf{I}_0 = \left[\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \hdashline 0 & 0 & 0 & 1 \end{array}\right]$$

Note that the associated identity matrix $\mathbf{I}_0$ has an extra row added. Reduction of $\mathbf{A}_0$ to echelon form (with concurrent elementary row operations on $\mathbf{I}_0$) gives us first the echelon form:

$$\mathbf{H}_{\mathbf{A}_0} = \left[\begin{array}{ccc:c} 1 & 1 & 3 & 0 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]; \qquad \left[\begin{array}{cccc} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 1 & -1 & 1 & 0 \\ \hdashline -1 & -1 & 0 & 1 \end{array}\right]$$

and a further elementary row operation that subtracts row 2 from row 1 leads to

$$\mathbf{J}_{\mathbf{A}_0} = \left[\begin{array}{ccc:c} 1 & 0 & 0 & 0 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]; \qquad \left[\begin{array}{cccc} 2 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 1 & -1 & 1 & 0 \\ \hdashline -1 & -1 & 0 & 1 \end{array}\right] = \mathbf{G}_0$$

We then write $\mathbf{A}^-$ by dropping the fourth row of the transformed identity matrix to get the 3 x 4 matrix

$$\mathbf{G} = \mathbf{A}^- = \begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 1 & -1 & 1 & 0 \end{bmatrix}$$

Having found $\mathbf{A}^-$, the g inverse, we can go on to solve for x in terms of the vector $\mathbf{b}' = (8, 14, 6, 22)$, as from Section B.2.2:

$$\mathbf{x} = \overset{\mathbf{A}^-}{\begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 1 & -1 & 1 & 0 \end{bmatrix}} \overset{\mathbf{b}}{\begin{bmatrix} 8 \\ 14 \\ 6 \\ 22 \end{bmatrix}} = \begin{bmatrix} 2 \\ 6 \\ 0 \end{bmatrix}$$

Since the rank of A is only two, we have an infinite number of solutions, written in general form as

$$\mathbf{x} = \overset{\mathbf{A}^-\mathbf{b}}{\begin{bmatrix} 2 \\ 6 \\ 0 \end{bmatrix}} + \left\{ \overset{\mathbf{I}}{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}} - \overset{\mathbf{A}^-\mathbf{A}}{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix}} \right\} \overset{\boldsymbol{\gamma}}{\begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \end{bmatrix}}$$

$$= \begin{bmatrix} 2 \\ 6 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ -3\gamma_3 \\ \gamma_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 6-\gamma_3 \\ \gamma_3 \end{bmatrix}$$

as was also obtained in Section B.2.2.

We observe, in passing, that the fourth row in A is redundant with the others, since it is the sum of the first two rows.

The procedure just outlined is quite general and can be applied to square or rectangular coefficients matrices, using the modification (to make A square) that was just illustrated. Of course, if one or more rows of zeros are appended to make A square, then the resulting additional columns of the transformed identity matrix are dropped in finding $\mathbf{A}^-$. Otherwise, the method is the same. In brief, the present method of computing g inverses is fully consistent with the echelon procedure of Section B.2. Thus, the echelon procedure (and consequent reduction of A to an identity submatrix) of Section B.2 is a general method for solving a specific set of linear equations.

The present method of solving for $\mathbf{A}^-$ is also fully general for finding a g inverse and then solving the system

$$\boxed{\mathbf{x} = \mathbf{A}^-\mathbf{b} + (\mathbf{I} - \mathbf{A}^-\mathbf{A})\boldsymbol{\gamma}}$$

for *any* desired b of constants. If A is nonsingular, then $\mathbf{A}^- = \mathbf{A}^{-1}$, and the second term on the right drops out. If A is singular, then $\mathbf{A}^-$ will still exist (as long as the equations are consistent, a test that can be made via the procedure of Section B.2).

Generalized inverses, of various types, are playing increasingly important roles in multivariate analysis. For more extensive discussion of the topic, the reader is referred to books by Pringle and Rayner (1971) and Rao and Mitra (1971).

B.5 SUMMARY

The role of generalized inverses in solving sets of linear equations has been the main subject of this appendix. Generalized inverses provide a counterpart role to regular inverses in cases where the coefficients matrix is singular. We started the discussion by reviewing a general procedure, involving reduction of either a coefficients or an augmented matrix to echelon form, for determining whether a set of simultaneous equations had (a) none, (b) exactly one, or (c) infinitely many solutions.

A general solution procedure, employing elementary row operations, was described and illustrated numerically. After reducing the augmented matrix **M** or the coefficients matrix **A** to echelon form, $\mathbf{H}_\mathrm{M}$ or $\mathbf{H}_\mathrm{A}$, additional elementary row operations were employed to reduce $\mathbf{H}_\mathrm{M}$ (or $\mathbf{H}_\mathrm{A}$) to an identity submatrix. Illustrations were provided for both nonhomogeneous and homogeneous sets of equations.

We then turned to a discussion of generalized inverses. The Penrose conditions were introduced, and the Moore–Penrose inverse $\mathbf{A}^+$ was described in the context of basic structure. The related concepts of left and right pseudoinverses were also illustrated, and properties of the (unique) Moore–Penrose were listed.

The appendix was concluded with a companion discussion of the *g* inverse $\mathbf{A}^-$. This (nonunique) generalized inverse need satisfy only the first Penrose condition. In general, $\mathbf{A}^-$ is easier to compute than $\mathbf{A}^+$ and, furthermore, plays a central role in solving sets of simultaneous equations. Several numerical illustrations of one procedure, involving reduction of a square coefficients matrix to Hermite form, were presented and tied in with the general procedure (of Section B.2) that was based on matrix reduction to echelon form.

REVIEW QUESTIONS

1. By means of reduction to echelon form, find the rank of

a. $\mathbf{A} = \begin{bmatrix} 1 & 1 & 2 & 4 \\ 3 & 3 & 6 & 12 \end{bmatrix}$ b. $\mathbf{B} = \begin{bmatrix} 2 & 1 & 3 \\ 1 & 2 & 0 \end{bmatrix}$

c. $\mathbf{C} = \begin{bmatrix} 1 & 4 & 1 \\ 3 & 2 & 7 \end{bmatrix}$ d. $\mathbf{D} = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 5 \\ 2 & 4 & 8 \end{bmatrix}$

2. Reduce the following matrices to identity submatrices:

a. $\mathbf{A} = \begin{bmatrix} 1 & 2 & 0 & -1 \\ 3 & 4 & 1 & 2 \\ -2 & 3 & 2 & 5 \end{bmatrix}$ b. $\mathbf{B} = \begin{bmatrix} 0 & 2 & 3 & 4 \\ 2 & 3 & 5 & 4 \\ 4 & 8 & 13 & 12 \end{bmatrix}$

3. Using elementary row operations on **A** and **I**, simultaneously, find the inverse of

a. $$\mathbf{A} = \begin{bmatrix} 2 & 3 \\ 1 & 4 \end{bmatrix}$$ b. $$\mathbf{A} = \begin{bmatrix} 1 & 3 & 3 \\ 1 & 4 & 3 \\ 1 & 3 & 4 \end{bmatrix}$$

4. Find a left pseudoinverse of

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 \\ 3 & 4 & 3 \\ 2 & 1 & 5 \\ 3 & 3 & 4 \end{bmatrix}$$

5. Consider the system of equations

$$\begin{aligned} x_1 + 3x_2 - 2x_3 - x_4 + 2x_5 &= 1 \\ 2x_1 + 6x_2 - 4x_3 - 2x_4 + 4x_5 &= 2 \\ x_1 + 3x_2 - 2x_3 + x_4 &= -1 \\ 2x_1 + 6x_2 + x_3 - x_4 &= 4 \end{aligned}$$

Reduce the augmented matrix **M** to echelon form and find **x**.

6. Consider the matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 3 \\ 2 & 1 \\ 3 & 2 \end{bmatrix}$$

(a) Find the Moore–Penrose inverse $\mathbf{A}^+$.

(b) Find a *g* inverse $\mathbf{A}^-$ via reduction of **A** to Hermite form.