

The Use of a Computer Algebra System in Modern Matrix Algebra

Karsten Schmidt

Faculty of Economics and Management Science,
University of Applied Sciences, Schmalkalden, Germany
e-mail: kschmidt@wi.fh-schmalkalden.de

Introduction

The concept of generalized inverses of matrices was not developed until the 20th century (cf. References). Whilst the inverse A^{-1} of a matrix A only exists if A is square and non-singular, the generalized inverse (g-inverse) A^{-} exists for all matrices A . Any matrix A^{-} satisfying the condition

$$AA^{-}A = A \quad (1)$$

is a generalized inverse of A . If A is square and nonsingular, we have $A^{-} = A^{-1}$, i.e. the generalized inverse is unique. Otherwise, the number of generalized inverses of a matrix is infinite.

Therefore, a special generalized inverse, the Moore-Penrose inverse (MP inverse) A^{+} , attracted greater attention. This matrix satisfies condition (1), i.e. $AA^{+}A = A$, and in addition

$$A^{+}AA^{+} = A^{+} \quad (2)$$

$$(A^{+}A)' = A^{+}A \quad (3)$$

$$(AA^{+})' = AA^{+} \quad (4)$$

assuring its uniqueness.

Until recently, these matrices did not play a principal role in first and second year foundations in mathematics and statistics in such areas as economics and management science. But with the availability of powerful computers in the classroom, it became possible to apply these modern concepts, for example to the solution of systems of linear equations or to the linear regression model.

In this paper we demonstrate how *DERIVE* can be used to teach the concepts of generalized inverses and the Moore-Penrose inverse.

In Section 2 we will introduce an algorithm for the computation of a generalized inverse of a matrix, and in Section 3 an algorithm for the computation of the unique Moore-Penrose inverse will be presented. Both algorithms will be illustrated by examples.

In Section 4 we will see how the g-inverse can be used to check if solutions for a system of linear equations $Ax = b$ exist and to provide the general solution. Some examples are given to demonstrate how this method works for different matrices A , both regular and singular.

Computation of a generalized inverse

We now introduce an algorithm for the computation of a generalized inverse A^- of any matrix A . This algorithm is based on the well known Gauss algorithm which is also frequently applied to calculate the inverse A^{-1} of a regular matrix A . It comprises four steps:

Step 1

We concatenate the identity matrix I to the right of A :

$$\begin{bmatrix} A & I \\ m \times n & m \times m \end{bmatrix}$$

Step 2

By successively performing elementary row operations to the matrix $[A \ I]$, i.e.

- by multiplying $[A \ I]$ from the left with matrices Z_i , where the Z_i are elementary matrices, in order to transform A into the Hermite normal form,

$$\begin{aligned} Z_1[A \ I] &= [Z_1A \ Z_1] \\ Z_2Z_1[A \ I] &= [Z_2Z_1A \ Z_2Z_1] \\ &\vdots \\ Z[A \ I] &= \begin{bmatrix} \underline{ZA} & Z \\ \underline{H} & \end{bmatrix} \end{aligned}$$

we get $Z = Z_k Z_{k-1} \cdots Z_1$.

Step 3

If the resulting matrix $H = ZA$ is not already of the form

$$R = \begin{bmatrix} I & K \\ 0 & 0 \end{bmatrix} \quad (5)$$

where $r = \text{rank}(H) = \text{rank}(A)$, we transform it into this form by interchanging the columns. This is equivalent to multiplying H from the right with a permutation matrix P which is equal to the identity matrix with interchanged columns (i.e. if H is already of the form (5) we have $P = I$):

$$\begin{aligned} [H \ Z]P &= [HP \ ZP] \\ &= [R \ ZP] \end{aligned}$$

Step 4

Having determined Z and P we can calculate a g-inverse of A by

$$\underset{n \times m}{A}^{-} = \underset{n \times n}{P} \underbrace{\begin{bmatrix} \underset{r \times r}{I} & \underset{r \times m}{0} \\ \underset{m \times r}{0} & \underset{m \times m}{0} \end{bmatrix}}_{n \times m} \underset{m \times m}{Z}$$

where $r = \text{rank}(A)$.

We illustrate this algorithm by calculating a g-Inverse of

$$A = \begin{pmatrix} 0 & 2 & 1 \\ 0 & 1 & \frac{1}{2} \end{pmatrix}$$

which is a matrix of rank 1. A copy of the corresponding expressions is given in Appendix A.

The *DERIVE*-function ROW_REDUCE performs the first two steps simultaneously. Obviously, the resulting H is not of the form (5). We therefore have to use a permutation matrix P different from I .

Finally, we compute a generalized inverse

$$\begin{aligned} \underset{3 \times 2}{A}^{-} &= \underset{3 \times 3}{P} \underbrace{\begin{bmatrix} \underset{1 \times 1}{I} & \underset{1 \times 2}{0} \\ \underset{2 \times 1}{0} & \underset{2 \times 2}{0} \end{bmatrix}}_{3 \times 2} \underset{2 \times 2}{Z} \\ &= \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & -2 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \end{aligned}$$

and show that condition 1 is indeed satisfied.

Computation of the Moore-Penrose inverse

In this section we introduce an algorithm for the computation of the Moore-Penrose inverse $\underset{n \times m}{A}^{+}$ of any matrix $\underset{m \times n}{A}$. This iterative algorithm, known as Greville algorithm, leads to the unique MP inverse in a finite number of iterations.

Since the Moore-Penrose inverse A^{+} is also a generalized inverse of A , this algorithm provides another method to calculate a generalized inverse A^{-} .

We start with a simple formula to calculate the MP inverse if $A = \underset{n \times 1}{a}$ is a vector:

$$\mathbf{a}^+ = \begin{cases} \frac{1}{\mathbf{a}'\mathbf{a}}\mathbf{a}' & \text{if } \mathbf{a} \neq \mathbf{0} \\ \mathbf{0}' & \text{if } \mathbf{a} = \mathbf{0} \end{cases} \quad (6)$$

We now consider the column notation of \mathbf{A} :

$$\mathbf{A}_{m \times n} = [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \cdots \quad \mathbf{a}_n]$$

and denote the submatrix, that comprises the first k columns of \mathbf{A} , by

$$\mathbf{A}_k = [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \cdots \quad \mathbf{a}_k]_{m \times k}$$

Hence

$$\mathbf{A}_k = [\mathbf{A}_{k-1} \quad \mathbf{a}_k]$$

Moreover, we define the following vectors for $j \geq 2$:

$$\begin{aligned} \mathbf{d}_j' &= \mathbf{a}_j' \mathbf{A}_{j-1}^+ \mathbf{A}_{j-1}^+ \\ \mathbf{c}_j &= (\mathbf{I} - \mathbf{A}_{j-1} \mathbf{A}_{j-1}^+) \mathbf{a}_j \\ \mathbf{b}_j' &= \mathbf{c}_j^+ + \frac{1 - \mathbf{c}_j^+ \mathbf{c}_j}{1 + \mathbf{d}_j' \mathbf{a}_j} \mathbf{d}_j' \end{aligned}$$

Note that \mathbf{d}_j' is a row vector, \mathbf{c}_j a column vector (and hence \mathbf{c}_j^+ a row vector) and \mathbf{b}_j' a row vector. Then we have

$$\mathbf{A}_j^+ = [\mathbf{A}_{j-1} \quad \mathbf{a}_j]^+ = \begin{bmatrix} \mathbf{A}_{j-1}^+ - \mathbf{A}_{j-1}^+ \mathbf{a}_j \mathbf{b}_j' \\ \mathbf{b}_j' \end{bmatrix} \quad (7)$$

Since $\mathbf{A}_1 = \mathbf{a}_1$ is a matrix which has only one column, its MP inverse is easily calculated by (6).

Using (7) we can then iteratively calculate $\mathbf{A}_2^+, \mathbf{A}_3^+, \dots, \mathbf{A}_n^+ = \mathbf{A}^+$.

This algorithm is easily implemented on a computer with a matrix programming language such as *GAUSS*. An example of a procedure for the calculation of the Moore-Penrose inverse can be found in Schmidt/Trenkler (1998, p. 123).

However, in this paper we provide a solution in *DERIVE*, where for the sake of simplicity we confine ourselves to matrices $\mathbf{A}_{m \times n}$ with $\min(m, n) \leq 2$, i.e. vectors, and matrices which have either only two rows or only two columns.

The set of functions on the following page could be used as a utility file. After being loaded, the function `MPI` calculates the MP inverse of the matrix passed as parameter, or terminates with an error message if $\min(m,n) > 2$.

We use the `MPI`-function to calculate the MP inverse of the matrix we already used in the previous section to illustrate the calculation of a g-inverse. A copy of the corresponding expressions, including the check showing that conditions 1 to 4 are satisfied, is given in Appendix B. Note that conditions (3) and (4) require both A^+A and AA^+ to be symmetric.

```

•
•
MPIV(a) := IF(a` • a = 0, 0-a`, -----)
•
•
ELEMENT(a` • a, 1, 1) f

A1(a) := DELETE_ELEMENT(a`, 2)`
A2(a) := DELETE_ELEMENT(a`, 1)`
D2T(a2, a1plus) := a2` • a1plus` • a1plus
C2(a1, a1plus, a2) := (IDENTITY_MATRIX(DIMENSION(a1)) - a1 • a1plus) • a2
B2T(c2, d2t, a2) := MPIV(c2) + -----d2t
1 + d2t • a2
A2PLUS(a1plus, a2, b2t) := APPEND(a1plus - a1plus • a2 • b2t, b2t)
MPI2(a) := A2PLUS(MPIV(A1(a)), A2(a), B2T(C2(A1(a), MPIV(A1(a)), A2(a)),
D2T(A2(a), MPIV(A1(a))), A2(a)))
MPI(a) := IF(MIN(DIMENSION(a), DIMENSION(a`)) > 2, "Error: MIN(m,n) > 2",
IF(MIN(DIMENSION(a), DIMENSION(a`)) = 1, IF(DIMENSION(a`) = 1,
MPIV(a), MPIV(a`)), IF(DIMENSION(a`) > 2, MPI2(a`), MPI2(a))))

```

Application to systems of linear equations

We consider a system of linear equations

$$\underset{m \times n}{A} \underset{n \times 1}{x} = \underset{m \times 1}{b} \quad (8)$$

The g-inverse of A can be applied to such a system

- to check if it is *consistent*, i.e. to investigate if it has solutions or not, and
- if it is consistent, to provide the general solution, which may consist of either one unique solution or an infinite number of solutions.

System (8) is consistent if and only if

$$AA^-b = b \quad (9)$$

for any generalized inverse A^- .

If $Ax = b$ is consistent, its general solution is given by

$$x = A^-b + \begin{pmatrix} I - A^-A \\ n \times n \end{pmatrix} z \quad (10)$$

where $z \in \mathbb{R}^n$ is an arbitrary vector.

Note that in applying (9) and (10), any g-inverse A^- is helpful. Therefore, we can use the Moore-Penrose inverse A^+ as well. Furthermore, since the vector $z \in \mathbb{R}^n$ in (10) is arbitrary, we can choose $z = 0$. Consequently, one (possibly unique) solution of $Ax = b$ is always given by

$$x = A^+b$$

The following function could be used as a utility file. After being loaded, the function SOLVESLE solves a system of linear equations $Ax = b$ where the matrix A and the vector b have been passed as parameters, or displays a message, if a solution does not exist.

```

,, z1 †
z := |  |
... z2 ‡

SOLVESLE(a, b) := IF(a • mpi • a • b = b,
    mpi • a • b + (IDENTITY_MATRIX(DIMENSION(a`)) - mpi • a • a) • z,
    "A solution does not exist!")

```

Finally, we analyze the consistency of three systems of linear equations, and calculate solutions, if possible. Copies of the corresponding expressions are given in Appendix C. We start with

$$A_{2 \times 2} = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}; b_{2 \times 1} = \begin{pmatrix} 2 \\ \frac{7}{2} \end{pmatrix} \quad (11)$$

By checking condition (9) using the MP inverse of A , we find that system (11) is consistent:

$$AA^+b = \begin{pmatrix} 2 \\ \frac{7}{2} \end{pmatrix} = b$$

Note that in this case A is a regular matrix. Hence $A^+ = A^{-1}$ and $AA^+ = I$.

The general solution is provided by (10); clearly, system (11) has a unique solution:

$$x = A^+b + (I - A^+A)z = \begin{pmatrix} 1 \\ \frac{1}{2} \end{pmatrix}$$

The second system of linear equations is described by

$$\underset{2 \times 2}{\mathbf{A}} = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}; \underset{2 \times 1}{\mathbf{b}} = \begin{pmatrix} 2 \\ 4 \end{pmatrix} \quad (12)$$

When we check condition (9) we find that system (12) is also consistent:

$$\mathbf{A}\mathbf{A}^+\mathbf{b} = \begin{pmatrix} 2 \\ 4 \end{pmatrix} = \mathbf{b}$$

Note that in this case \mathbf{A} is a singular matrix. Hence, \mathbf{A}^{-1} does not exist and $\mathbf{A}\mathbf{A}^+ \neq \mathbf{I}$. The general solution is provided by (10); obviously system (12) has an infinite number of solutions:

$$\mathbf{x} = \mathbf{A}^+\mathbf{b} + (\mathbf{I} - \mathbf{A}^+\mathbf{A})\mathbf{z} = \begin{pmatrix} \frac{2}{5} + \frac{4}{5}z_1 - \frac{2}{5}z_2 \\ \frac{4}{5} - \frac{2}{5}z_1 + \frac{1}{5}z_2 \end{pmatrix}$$

For example, by choosing $\mathbf{z} = \mathbf{0}$ we get the solution

$$\mathbf{x} = \begin{pmatrix} \frac{2}{5} \\ \frac{4}{5} \end{pmatrix}$$

The third system of linear equations is described by

$$\underset{2 \times 2}{\mathbf{A}} = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}; \underset{2 \times 1}{\mathbf{b}} = \begin{pmatrix} 2 \\ 3 \end{pmatrix} \quad (13)$$

By checking condition (9), we find that system (13) is inconsistent:

$$\mathbf{A}\mathbf{A}^+\mathbf{b} = \begin{pmatrix} \frac{8}{5} \\ \frac{16}{5} \end{pmatrix} \neq \mathbf{b}$$

i.e. a solution does not exist.

References

- Ben-Israel, A. & T.N.E. Greville (1974), *Generalized Inverses: Theory and Applications*, New York (Wiley).
- Greville, T.N.E. (1960), Some Applications of the Pseudoinverse of a Matrix, *SIAM Review* **2**, 15-22.
- Moore, E.H. (1920), On the Reciprocal of the General Algebraic Matrix (Abstract), *Bulletin of the American Mathematical Society* **26**, 394-395.
- Penrose, R. (1955), A Generalized Inverse for Matrices, *Proceedings of the Cambridge Philosophical Society* **51**, 406-413.

Rao, C.R. (1962), A Note on a Generalized Inverse of a Matrix with Applications to Problems in Mathematical Statistics, *Journal of the Royal Statistical Society B* **24**, 152-158.

Rao, C.R. & S.K. Mitra (1971), *Generalized Inverse of Matrices and Its Applications*, New York (Wiley).

Schmidt, K. & G. Trenkler (1998), *Moderne Matrix-Algebra*, Berlin (Springer).

Appendix A

$$a := \begin{pmatrix} 0 & 2 & 1 \\ 1 & 1 \\ 0 & 1 \\ \dots & 2 \end{pmatrix}$$

ROW_REDUCE(a, IDENTITY_MATRIX(DIMENSION(a)))

$$\begin{pmatrix} 1 \\ 0 & 1 \\ 2 \\ \dots & 0 & 0 & 0 & 1 & -2 \end{pmatrix}$$

$$\begin{bmatrix} H & Z \end{bmatrix}$$

$$h := \begin{pmatrix} 1 \\ 0 & 1 \\ 2 \\ \dots & 0 & 0 & 0 \end{pmatrix}$$

$$z := \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}$$

$$p := \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ \dots & 0 & 0 & 1 \end{pmatrix}$$

$$h \cdot p$$

$$\begin{pmatrix} 1 \\ 1 & 0 \\ 2 \\ \dots & 0 & 0 & 0 \end{pmatrix}$$

$$R$$

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \\ \dots & 0 & 0 \end{pmatrix}$$

$$\underbrace{\begin{bmatrix} I & 0 \\ 1 \times 1 & 0 \\ 0 & 0 \end{bmatrix}}_{3 \times 2}$$

$$p \cdot \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ \dots & 0 & 0 \end{pmatrix} \cdot z$$

$$\begin{pmatrix} 0 & 0 \\ 0 & 1 \\ \dots & 0 & 0 \end{pmatrix}$$

$$A^-$$

$$a \cdot \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \cdot a$$

$$\begin{array}{c}
 \begin{array}{c} | \quad | \\ \dots 0 \ 0 \ddagger \end{array} \\
 \begin{array}{c} \text{„} 0 \ 2 \ 1 \ \ddagger \\ | \quad | \\ | \quad 1 \ | \\ | 0 \ 1 \ \text{-----} | \\ \dots \quad 2 \ \ddagger \end{array}
 \end{array}
 \quad A$$

Appendix B

$$\begin{array}{c}
 \text{„} 0 \ 2 \ 1 \ \ddagger \\
 | \quad | \\
 a := | \quad 1 \ | \\
 | 0 \ 1 \ \text{-----} | \\
 \dots \quad 2 \ \ddagger
 \end{array}$$

MPI(a)

$$\begin{array}{c}
 \text{„} \ 0 \ 0 \ \ddagger \\
 | \quad | \\
 | \ 8 \ 4 \ | \\
 | \text{-----} | \\
 | 25 \ 25 \ | \\
 | \quad | \\
 | \ 4 \ 2 \ | \\
 | \text{-----} | \\
 \dots \ 25 \ 25 \ \ddagger
 \end{array}$$

$$\begin{array}{c}
 \text{„} \ 0 \ 0 \ \ddagger \\
 | \quad | \\
 | \ 8 \ 4 \ | \\
 | \text{-----} | \\
 aplus := | 25 \ 25 \ | \\
 | \quad | \\
 | \ 4 \ 2 \ | \\
 | \text{-----} | \\
 \dots \ 25 \ 25 \ \ddagger
 \end{array}$$

a • aplus • a

$$\begin{array}{c}
 \text{„} 0 \ 2 \ 1 \ \ddagger \\
 | \quad | \\
 | \quad 1 \ | \\
 | 0 \ 1 \ \text{-----} | \\
 \dots \quad 2 \ \ddagger
 \end{array}$$

aplus • a • aplus

$$\begin{array}{c}
 \text{„} \ 0 \ 0 \ \ddagger \\
 | \quad | \\
 | \ 8 \ 4 \ | \\
 | \text{-----} | \\
 | 25 \ 25 \ | \\
 | \quad | \\
 | \ 4 \ 2 \ | \\
 | \text{-----} | \\
 \dots \ 25 \ 25 \ \ddagger
 \end{array}$$

aplus • a

$$\begin{array}{c} \text{,, 0 0 0 } \dagger \\ \vdots \quad \vdots \\ \vdots \quad 4 \quad 2 \quad \vdots \\ \vdots \quad 0 \quad \frac{\quad}{5 \quad 5} \quad \frac{\quad}{\quad} \quad \vdots \\ \vdots \quad \vdots \\ \vdots \quad 2 \quad 1 \quad \vdots \\ \vdots \quad 0 \quad \frac{\quad}{5 \quad 5} \quad \frac{\quad}{\quad} \quad \vdots \\ \dots \quad 5 \quad 5 \quad \ddagger \end{array}$$

a • aplus

$$\begin{array}{c} \text{,, 4 2 } \dagger \\ \vdots \quad \vdots \\ \vdots \quad \frac{\quad}{5 \quad 5} \quad \frac{\quad}{\quad} \quad \vdots \\ \vdots \quad \vdots \\ \vdots \quad 2 \quad 1 \quad \vdots \\ \vdots \quad \frac{\quad}{5 \quad 5} \quad \frac{\quad}{\quad} \quad \vdots \\ \dots \quad 5 \quad 5 \quad \ddagger \end{array}$$

Appendix C

$$\begin{array}{c} \text{,, 1 2 } \dagger \\ \text{a} := \vdots \quad \vdots \\ \dots \quad 2 \quad 3 \quad \ddagger \end{array}$$

$$\begin{array}{c} \text{,, 2 } \dagger \\ \vdots \quad \vdots \\ \text{b} := \vdots \quad 7 \quad \vdots \\ \vdots \quad \frac{\quad}{\quad} \quad \vdots \\ \dots \quad 2 \quad \ddagger \end{array}$$

SOLVESLE(a, b)

$$\begin{array}{c} \text{,, 1 } \dagger \\ \vdots \quad \vdots \\ \vdots \quad 1 \quad \vdots \\ \vdots \quad \frac{\quad}{\quad} \quad \vdots \\ \dots \quad 2 \quad \ddagger \end{array}$$

$$\begin{array}{c} \text{,, 1 2 } \dagger \\ \text{a} := \vdots \quad \vdots \\ \dots \quad 2 \quad 4 \quad \ddagger \end{array}$$

$$\begin{array}{c} \text{,, 2 } \dagger \\ \text{b} := \vdots \quad \vdots \\ \dots \quad 4 \quad \ddagger \end{array}$$

SOLVESLE(a, b)

$$\begin{array}{c} \text{,, 4·z1 \quad 2·z2 \quad 2 } \dagger \\ \vdots \quad \frac{\quad}{5 \quad 5 \quad 5} \quad \frac{\quad}{\quad} \quad \frac{\quad}{\quad} \quad \vdots \\ \vdots \quad \vdots \\ \vdots \quad 2·z1 \quad z2 \quad 4 \quad \vdots \\ \vdots \quad \frac{\quad}{5 \quad 5 \quad 5} \quad \frac{\quad}{\quad} \quad \frac{\quad}{\quad} \quad \vdots \\ \dots \quad 5 \quad 5 \quad 5 \quad \ddagger \end{array}$$

$$\begin{array}{c} \text{,, 4·0 \quad 2·0 \quad 2 } \dagger \\ \vdots \quad \frac{\quad}{5 \quad 5 \quad 5} \quad \frac{\quad}{\quad} \quad \frac{\quad}{\quad} \quad \vdots \\ \vdots \quad \vdots \\ \vdots \quad \vdots \end{array}$$

$$\begin{array}{c} | \quad 2 \cdot 0 \quad 0 \quad 4 \quad | \\ | \quad \hline \dots \quad 5 \quad 5 \quad 5 \quad \ddagger \end{array} + \dots + \dots$$

$$\begin{array}{c} \text{,, } 2 \quad \ddagger \\ | \quad \hline | \quad 5 \quad | \\ | \quad | \\ | \quad 4 \quad | \\ | \quad \hline \dots \quad 5 \quad \ddagger \end{array}$$

$$\begin{array}{c} \text{,, } 2 \quad \ddagger \\ \mathbf{b} := | \quad | \\ \dots 3 \quad \ddagger \end{array}$$

SOLVESLE(a, b)

"A solution does not exist!"