JOURNAL OF OPTIMIZATION THEORY AND APPLICATIONS: Vol. 127, No. 3, pp. 639–663, December 2005 (© 2005)
DOI: 10.1007/s10957-005-7508-7

Recursive Determination of the Generalized Moore–Penrose *M*-Inverse of a Matrix

F. E. UDWADIA[1] AND P. PHOHOMSIRI[2]

Abstract. In this paper, we obtain recursive relations for the determination of the generalized Moore–Penrose M-inverse of a matrix. We develop separate relations for situations when a rectangular matrix is augmented by a row vector and when such a matrix is augmented by a column vector.

Key Words. Generalized inverse, Moore–Penrose M-inverse, recursive formulas, least squares problems.

1. Introduction

The concept of generalized inverses was introduced first by Moore (Ref. 1) in 1920 and independently rediscovered by Penrose (Ref. 2) in 1955. In 1960, the recursive scheme of computing the Moore–Penrose (MP) inverse of a matrix was obtained by Greville (Ref. 3). Because of its extensive applicability, the Greville result appears in almost every book on the subject of generalized inverses. Nevertheless, due to the complexity of the proof, the Greville paper is seldom quoted or outlined even in specialized textbooks which concentrate only on generalized inverses like Refs. 4–6. Recently in 1997, Udwadia and Kalaba (Ref. 7) provided an alternative, simple constructive proof of the Greville formulas. They obtained also the recursive determination of different types of generalized inverses of a matrix which includes the generalized inverse, the least-squares generalized inverse, the minimum-norm generalized inverse, and the Moore–Penrose inverse of a matrix (Refs. 8–9).

[1]Professor of Aerospace and Mechanical Engineering, Civil Engineering, Mathematics, and Information and Operation Management, University of Southern California, Los Angeles, California.
[2]Postdoctoral Research Associate, Aerospace and Mechanical Engineering, University of Southern California, Los Angeles, California.

Generalized inverses have various applications in the areas of statistical inference (Ref. 10), filtering theory, estimation theory (Ref. 11), system identification (Ref. 12), optimization and control (Ref. 11), and lately analytical dynamics (Ref. 13). The reason for the extensive applicability is that it provides a systematic method to generate updates whenever sequential addition of data or new information becomes available and updated estimates which take into account this additional information are required.

In the seventies, the concept of the Moore–Penrose (MP) inverse was expanded to the generalized Moore–Penrose M-inverse of a matrix (Ref. 6). This MP M-inverse is used in many areas of application, especially statistics, and recently has found also applicability in the field of analytical dynamics. It appears explicitly (Ref. 14) in the general equations of motion describing constrained mechanical systems.

In this paper, we obtain recursive formulae for the computation of the generalized Moore–Penrose (MP) M-inverse of a matrix. Since the generalized MP M-inverse of a matrix is not as well known as the regular Moore–Penrose inverse, we provide its properties. For a given m by n matrix B, the MP M-inverse of the matrix B_M^+ is the unique matrix that satisfies the following four properties (Ref. 6):

$$BB_M^+B = B, \tag{1}$$

$$B_M^+BB_M^+ = B_M^+, \tag{2}$$

$$(BB_M^+)^T = BB_M^+, \tag{3}$$

$$(B_M^+B)^T = MB_M^+BM^{-1}. \tag{4}$$

Throughout this paper, the superscript + represents the MP inverse and the subscript M denotes the generalized M-inverse. The matrix M in Eq. (4) is a symmetric positive-definite n by n matrix.

Consider a set of linear equations

$$Bx = b, \tag{5}$$

where B is an m by n matrix, b is an m-vector (m by 1 matrix), and x is an n-vector. The minimum M-norm least squares solution of Eq. (5) is given by

$$x = B_M^+b.$$

The Moore–Penrose (MP) M-inverse of the matrix B (B_M^+ and not B^+) is then obtained by finding that n-vector x for which

$$G = \|Bx - b\|^2, \tag{6}$$

$$H = x^T M x = \left\| M^{1/2} x \right\|^2 = \|x\|_M^2 \tag{7}$$

are both minimized, where M is an n by n symmetric positive-definite matrix. Note that, when M becomes αI_n, where α is a given constant which is greater than zero and I_n is the n by n identity matrix, we have

$$H = \alpha x^T x = \alpha \|x\|^2;$$

thus, the MP M-inverse turns out to be the regular MP inverse ($B_M^+ = B^+$).

Equation (5) can be written as

$$Bx = \begin{bmatrix} \hat{A} \\ \hat{a} \end{bmatrix} x = \begin{bmatrix} \hat{b} \\ s \end{bmatrix} = b, \tag{8}$$

or

$$Bx = \begin{bmatrix} A | a \end{bmatrix} x = b, \tag{9}$$

where $\hat{A}$ is an $(m-1)$ by n matrix, $\hat{b}$ is a column vector of $(m-1)$ components, $\hat{a}$ is a row vector of n components, s is a scalar, A is an m by $(n-1)$ matrix, and a is a column vector of m components. We have thus defined the row-wise partitioned matrix $B = \begin{bmatrix} \hat{A} \\ \hat{a} \end{bmatrix}$ in Eq. (8) and the column-wise partitioned matrix $B = \begin{bmatrix} A | a \end{bmatrix}$ in Eq. (9).

Our goal is to obtain the generalized M-inverse of B in terms of the generalized M-inverse of $\hat{A}$ or the generalized M_--inverse of A, where M_- is the $(n-1)$ by $(n-1)$ principal minor of the matrix M, thereby obtaining recursive relations. We develop different recursive formulae for the generalized M-inverse of a matrix applicable for the two cases when a row is added to a matrix and when a column is added to it. The reason we develop both the row-wise and column-wise formulae is because it does not appear to be straightforward to obtain one set of formulae from the other. While for the (regular) MP inverse, we can use the property $\left(B^+\right)^T = \left(B^T\right)^+$ to obtain the row-wise formulae from the column-wise formulae (and viceversa, as was done in Ref. 13), we do not have such a property to draw upon for the MP M-inverse of a matrix. Furthermore, the recursive generalized M-inverse depends upon whether the added column a in Eq. (9) (added row $\hat{a}$ in Eq. (8)) is linearly independent or linearly dependent on the columns (rows) of A ($\hat{A}$). For each of these cases, we develop the explicit recursive relations.

2. Generalized M-Inverse of a Row-Wise Partitioned Matrix

In this section, we develop recursive relations for the generalized M-inverse of an m by n matrix B obtained by adding an n component row vector $\hat{a}$, to an $(m-1)$ by n matrix $\hat{A}$. We assume that M is a given n by n positive-definite matrix.

Result 2.1. Given the row-wise partitioned m by n matrix $B=\begin{bmatrix}\hat{A}\\ \hat{a}\end{bmatrix}$, its generalized MP M-inverse is given by

$$B_M^+=\begin{bmatrix}\hat{A}\\ \hat{a}\end{bmatrix}_M^+=\left[\hat{A}_M^+-c_M^+\hat{a}\hat{A}_M^+ \;\middle|\; c_M^+\right], \qquad c\neq 0. \tag{10}$$

$$=\left[\hat{A}_M^+-\hat{A}_M^+u^Tu/(1+uu^T) \;\middle|\; \hat{A}_M^+u^T/(1+uu^T)\right], \qquad c=0. \tag{11}$$

where

$$c=\hat{a}(I-\hat{A}_M^+\hat{A}) \quad \text{and} \quad u=\hat{a}\hat{A}_M^+.$$

Proof. We first consider a set of linear equations

$$\hat{A}\hat{x}=\hat{b}, \tag{12}$$

where $\hat{A}$ is an $(m-1)$ by n matrix, $\hat{b}$ is an $(m-1)$-vector, and $\hat{x}$ is an n-vector. Suppose that the generalized MP M-inverse of $\hat{A}$, which we denote by $\hat{A}_M^+$, is given, where the matrix M is an n by n symmetric positive-definite matrix. The matrix $\hat{A}_M^+$ is n by $(m-1)$. Let us then add another equation,

$$\hat{a}\hat{x}=s, \tag{13}$$

to the equation set (12) as the last row, where $\hat{a}$ is a row vector of n components and s is a scalar.

Consequently, the new set of equations becomes

$$Bx=\begin{bmatrix}\hat{A}\\ \hat{a}\end{bmatrix}x=\begin{bmatrix}\hat{b}\\ s\end{bmatrix}=b, \tag{14}$$

where x is the n-vector, which needs to be updated due to the addition of the row vector $\hat{a}$.

Our aim here is to obtain the solution x of Eq. (14) so that

$$G = \|Bx - b\|^2 = \left\| \begin{bmatrix} \hat{A} \\ \hat{a} \end{bmatrix} x - \begin{bmatrix} \hat{b} \\ s \end{bmatrix} \right\|^2 = \left\| \hat{A}x - \hat{b} \right\|^2 + \left\| \hat{a}x - s \right\|^2, \tag{15}$$

$$H = \|x\|_M^2 = \left\| M^{1/2} x \right\|^2 = x^T M x \tag{16}$$

are both minimized.

Let us start by noting that any n-vector x can be expressed in the form (see Appendix, Property 5.1)

$$x = \hat{A}_M^+ (\hat{b} + y) + (I - \hat{A}_M^+ \hat{A}) w, \tag{17}$$

where y is a suitable $(m-1)$-vector and w is a suitable n-vector.

Substituting Eq. (17) in Eq. (15) and noting that

$$\hat{A}(I - \hat{A}_M^+ \hat{A}) = 0,$$

we obtain

$$G(y, w) = \left\| \hat{A}\hat{A}_M^+ (\hat{b} + y) - \hat{b} \right\|^2 + \left\| \hat{a}\hat{A}_M^+ (\hat{b} + y) + cw - s \right\|^2, \tag{18}$$

where we have defined the 1 by n row vector

$$c = \hat{a}(I - \hat{A}_M^+ \hat{A}).$$

We see from Eq. (18) that the scalar cw plays an important role in finding the minimum of $G(y, cw)$: when $c = 0, G$ is no longer a function of w. Therefore, we shall consider the two cases separately: when $c = 0$ and when $c \neq 0$. When $c = 0$, the row vector $\hat{a}$ is a linear combination of the rows of the matrix $\hat{A}$; when $c \neq 0$, the row vector $\hat{a}$ is not a linear combination of the rows of the matrix $\hat{A}$; see Property 5.2 in the Appendix.

(a) Case $c \neq 0$. Let us first set $y = y_0$, where y_0 is some fixed vector. By using the definition of the MP M-inverse, the value of w that minimizes G, which is given by Eq. (18), is

$$w = c_M^+ (s - \hat{a}\hat{A}_M^+ (\hat{b} + y_0)) + (I - c_M^+ c) l_1, \tag{19}$$

where l_1 is any arbitrary column vector of n components. We note that c_M^+ is the least squares generalized M-inverse of c (see Appendix, Property 5.3).

Substituting Eq. (19) in Eq. (18), we obtain

$$G(y_0, w(y_0)) = \left\| \hat{A}\hat{A}_M^+(\hat{b} + y_0) - \hat{b} \right\|^2 + \left\| \hat{a}\hat{A}_M^+(\hat{b} + y_0) + cc_M^+(s - \hat{a}\hat{A}_M^+(\hat{b} + y_0)) - s \right\|^2, \tag{20}$$

which can be written as

$$G(y_0, w(y_0)) = \left\| \hat{A}\hat{A}_M^+ y_0 - (I - \hat{A}\hat{A}_M^+)\hat{b} \right\|^2 + \left\| (1 - cc_M^+)\hat{a}\hat{A}_M^+(\hat{b} + y_0) - (1 - cc_M^+)s \right\|^2. \tag{21}$$

Since $cc_M^+ = 1$ (see Appendix, Property 5.4), Equation (21) becomes

$$G(y_0, w(y_0)) = \left\| \hat{A}\hat{A}_M^+ y_0 - (I - \hat{A}\hat{A}_M^+)\hat{b} \right\|^2. \tag{22}$$

Because

$$\begin{aligned}(\hat{A}\hat{A}_M^+ y_0)^T (I - \hat{A}\hat{A}_M^+)b &= y_0^T (\hat{A}\hat{A}_M^+)^T (I - \hat{A}\hat{A}_M^+)b \\ &= y_0^T \hat{A}\hat{A}_M^+ (I - \hat{A}\hat{A}_M^+)b \\ &= 0,\end{aligned}$$

Equation (24) can be expressed as

$$G(y_0, w(y_0)) = \left\| \hat{A}\hat{A}_M^+ y_0 \right\|^2 + \left\| (I - \hat{A}\hat{A}_M^+)\hat{b} \right\|^2. \tag{23}$$

For G to be minimized, we must choose y_0 so that

$$\hat{A}\hat{A}_M^+ y_0 = 0. \tag{24}$$

Premultiplying Eq. (24) by $\hat{A}_M^+$, we obtain

$$\hat{A}_M^+ y_0 = 0. \tag{25}$$

Substitution of Eqs. (19) and (25) in Eq. (17) now gives us

$$x = \hat{A}_M^+\hat{b} + (I - \hat{A}_M^+\hat{A})c_M^+(s - \hat{a}\hat{A}_M^+\hat{b}) + (I - \hat{A}_M^+\hat{A})(I - c_M^+ c)l_1, \tag{26}$$

which now minimizes G. Here, l_1 is any arbitrary column vector of n components. In the Appendix (Property 5.5), we show that $\hat{A}c_M^+ = 0$. Denoting

$$K = (I - \hat{A}_M^+\hat{A})(I - c_M^+ c),$$

Equation (26) then simplifies to

$$x=\hat{A}_M^+\hat{b}+c_M^+(s-\hat{a}\hat{A}_M^+\hat{b})+Kl_1. \tag{27}$$

We next choose l_1 so as to minimize

$$\begin{aligned} H &= \|x\|_M^2 \\ &= \left\| \hat{A}_M^+\hat{b}+c_M^+(s-\hat{a}\hat{A}_M^+\hat{b})+Kl_1 \right\|_M^2, \\ &= \left\| \hat{A}_M^+\hat{b}+c_M^+(s-\hat{a}\hat{A}_M^+\hat{b} \right\|_M^2 + \|Kl_1\|_M^2 \\ &\quad +2l_1^T K^T M\left(\hat{A}_M^+\hat{b}+c_M^+(s-\hat{a}\hat{A}_M^+\hat{b})\right). \end{aligned} \tag{28}$$

From the Appendix (Property 5.6), the last term of Eq. (28) is zero. Hence, to minimize H, we require Kl_1 to be zero. Thus, Equation (27) becomes

$$x=\hat{A}_M^+b+c_M^+(s-\hat{a}\hat{A}_M^+\hat{b}), \tag{29}$$

which can be expressed as

$$x=B_M^+b=\begin{bmatrix}\hat{A}\\ \hat{a}\end{bmatrix}_M^+\begin{bmatrix}\hat{b}\\ s\end{bmatrix}=\left[\hat{A}_M^+-c_M^+\hat{a}\hat{A}_M^+\middle| c_M^+\right]\begin{bmatrix}\hat{b}\\ s\end{bmatrix}. \tag{30}$$

Since B_M^+ is unique and the result is valid for any given $(m-1)$-vector $\hat{b}$ and scalar s, we have

$$B_M^+=\begin{bmatrix}\hat{A}\\ \hat{a}\end{bmatrix}_M^+=\left[\hat{A}_M^+-c_M^+\hat{a}\hat{A}_M^+\middle| c_M^+\right], \quad c=\hat{a}(I-\hat{A}_M^+\hat{A})\neq 0. \qquad \square$$

(b) Case $c=0$. When $c=0$, Equation (18) becomes

$$G=\left\|\hat{A}\hat{A}_M^+(\hat{b}+y)-\hat{b}\right\|^2+\left\|\hat{a}\hat{A}_M^+(\hat{b}+y)-s\right\|^2, \tag{31}$$

which can be rewritten as

$$G=\left\|\hat{A}\hat{A}_M^+\hat{b}+k-\hat{b}\right\|^2+\left\|\hat{a}\hat{A}_M^+(\hat{b}+k)-s\right\|^2, \tag{32}$$

where we have defined

$$k=\hat{A}\hat{A}_M^+y$$

and have used

$$\hat{A}_M^+ y = \hat{A}_M^+ k.$$

Defining $u = \hat{a}\hat{A}_M^+$ and minimizing G with respect to k, we obtain

$$\partial G/\partial k = 2\hat{A}\hat{A}_M^+\hat{b} + 2k - 2\hat{b} + 2u^T\left(u(\hat{b}+k) - s\right) = 0. \tag{33}$$

The last equality of Eq. (33) gives

$$(I + u^T u)k = u^T(s - u\hat{b}) + (I - \hat{A}\hat{A}_M^+)\hat{b}. \tag{34}$$

Since $I + u^T u$ is a positive-definite matrix, we obtain

$$k = (I + u^T u)^{-1} u^T (s - u\hat{b}) + (I + u^T u)^{-1}(I - \hat{A}\hat{A}_M^+)\hat{b}, \tag{35}$$

which minimizes G. Using the fact that

$$(I + u^T u)^{-1} = I - u^T u/(1 + uu^T)$$

(see Property 5.7 in the Appendix) and

$$\begin{aligned}(I + u^T u)^{-1} u^T &= (I - u^T u/(1 + uu^T))u^T \\ &= u^T/(1 + uu^T),\end{aligned}$$

we get

$$\begin{aligned} k &= \left[u^T/(1 + uu^T)\right](s - u\hat{b}) + \left[I - u^T u/(1 + uu^T)\right](I - \hat{A}\hat{A}_M^+)\hat{b}, \\ &= \left[u^T/(1 + uu^T)\right](s - u\hat{b}) + (I - \hat{A}\hat{A}_M^+)\hat{b} - \left[u^T u/(1 + uu^T)\right] \\ &\quad \times (I - \hat{A}\hat{A}_M^+)\hat{b}. \end{aligned} \tag{36}$$

Since

$$u(I - \hat{A}\hat{A}_M^+)\hat{b} = (a\hat{A}_M^+)(I - \hat{A}\hat{A}_M^+)\hat{b} = 0,$$

the last term on the right-hand side of Eq. (36) vanishes. Premultiplying Eq. (36) by $\hat{A}_M^+$, we have

$$\begin{aligned}\hat{A}_M^+ k &= \hat{A}_M^+(\hat{A}\hat{A}_M^+ y) \\ &= \hat{A}_M^+ y = \left[\hat{A}_M^+ u^T/(1 + uu^T)\right](s - u\hat{b}).\end{aligned} \tag{37}$$

Let us then substitute the last equality of Eq. (37) in Eq. (17) to obtain

$$x = \hat{A}_M^+ \left\{ \hat{b} + \left[u^T/(1+uu^T) \right] (s - u\hat{b}) \right\} + (I - \hat{A}_M^+ \hat{A}) w, \tag{38}$$

which now minimizes G.

Since

$$(I - \hat{A}_M^+ \hat{A})^T M \hat{A}_M^+ = (I - M\hat{A}_M^+ \hat{A} M^{-1}) M \hat{A}_M^+ = 0,$$

the first and second terms on the right-hand side of Eq. (38) are M-orthogonal. So, we get

$$\begin{aligned} H = \|x\|_M^2 &= \left\| \hat{A}_M^+ \left\{ \hat{b} + \left[u^T/(1+uu^T) \right] (s - u\hat{b}) \right\} + (I - \hat{A}_M^+ \hat{A}) w \right\|_M^2 \\ &= \left\| \hat{A}_M^+ \left\{ \hat{b} + \left[u^T/(1+uu^T) \right] (s - u\hat{b}) \right\} \right\|_M^2 + \left\| (I - \hat{A}_M^+ \hat{A}) w \right\|_M^2 . \end{aligned} \tag{39}$$

For H to be minimized, $M^{1/2}(I - \hat{A}_M^+ \hat{A}) w$ must then be zero. Because M is nonsingular, this gives

$$(I - \hat{A}_M^+ \hat{A}) w = 0.$$

Hence, from (38) we have

$$\begin{aligned} x = B_M^+ b &= \begin{bmatrix} \hat{A} \\ \hat{a} \end{bmatrix}_M^+ \begin{bmatrix} \hat{b} \\ s \end{bmatrix} \\ &= \hat{A}_M^+ \left\{ \hat{b} + \left[u^T/(1+uu^T) \right] (s - u\hat{b}) \right\} \\ &= \left[\hat{A}_M^+ - \hat{A}_M^+ u^T u/(1+uu^T) \middle| \hat{A}_M^+ u^T/(1+uu^T) \right] \begin{bmatrix} \hat{b} \\ s \end{bmatrix}. \end{aligned} \tag{40}$$

For

$$c = \hat{a}(I - \hat{A}_M^+ \hat{A}) = 0,$$

we thus have, arguing as before, that

$$B_M^+ = \begin{bmatrix} \hat{A} \\ \hat{a} \end{bmatrix}_M^+ = \left[\hat{A}_M^+ - \hat{A}_M^+ u^T u/(1+uu^T) \middle| \hat{A}_M^+ u^T/(1+uu^T) \right]. \qquad \square$$

Corollary 2.1. When $M = \alpha I_n, \alpha > 0$, the generalized M-inverse of a matrix becomes the (regular) Moore–Penrose inverse and we obtain the recursive relations

$$\begin{bmatrix} \hat{A} \\ a \end{bmatrix}^{+}_{M=\alpha I_n} = \begin{bmatrix} \hat{A} \\ a \end{bmatrix}^{+} = \left[\hat{A}^{+} - c^{+}\hat{a}\hat{A}^{+} \middle| c^{+} \right], \qquad c \neq 0,$$
$$\begin{bmatrix} \hat{A} \\ a \end{bmatrix}^{+}_{M=\alpha I_n} = \begin{bmatrix} \hat{A} \\ \hat{a} \end{bmatrix}^{+} = \left[\hat{A}^{+} - \hat{A}^{+}u^T u/(1+uu^T) \middle| \hat{A}^{+}u^T/(1+uu^T) \right], \; c = 0,$$

where

$$c = \hat{a}(I - \hat{A}^{+}\hat{A}) \quad \text{and} \quad u = \hat{a}\hat{A}^{+}.$$

Proof. When $M = \alpha I_n$, we obtain

$$\hat{A}^{+}_{M} = \hat{A}^{+} \quad \text{and} \quad c^{+}_{M} = c^{+}.$$

Consequently, we have

$$\begin{aligned} \begin{bmatrix} \hat{A} \\ a \end{bmatrix}^{+}_{M=\alpha I_n} &= \left[\hat{A}^{+}_{M} - c^{+}_{M}\hat{a}\hat{A}^{+}_{M} \middle| c^{+}_{M} \right]_{M=\alpha I_n} \\ &= \left[\hat{A}^{+} - c^{+}\hat{a}\hat{A}^{+} \middle| c^{+} \right], \quad c \neq 0, \end{aligned}$$

$$\begin{aligned} \begin{bmatrix} \hat{A} \\ a \end{bmatrix}^{+}_{M=\alpha I_n} &= \left[\hat{A}^{+}_{M} - \hat{A}^{+}_{M}u^T u/(1+uu^T) \middle| \hat{A}^{+}_{M}u^T/(1+uu^T) \right]_{M=\alpha I_n} \\ &= \left[\hat{A}^{+} - \hat{A}^{+}u^T u/(1+uu^T) \middle| \hat{A}^{+}u^T/(1+uu^T) \right], \quad c = 0, \end{aligned}$$

where

$$c = \hat{a}(I - \hat{A}^{+}\hat{A}) \quad \text{and} \quad u = \hat{a}\hat{A}^{+}. \qquad \square$$

Thus, we note that the recursive relations for the (regular) generalized MP inverse (Refs. 3 and 13) are obtained from the recursive relations for the generalized M-inverse by simply suppressing the subscript M in Result 2.1.

3. Generalized M-Inverse of a Column-Wise Partitioned Matrix

In this section, we provide recursive relations for the generalized M-inverse of an m by n matrix B obtained by the addition of a column m-vector a to an m by $(n-1)$ matrix A. We assume that M is a given n by n positive-definite matrix.

Result 3.1. Given the column-wise partitioned m by n matrix $B = [A\,|a]$, its generalized MP M-inverse is given by

$$B_M^+ = \left[A\,\middle|\,a\right]_M^+ = \begin{bmatrix} A_{M_-}^+ - A_{M_-}^+ ad^+ - pd^+ \\ d^+ \end{bmatrix}, \qquad d \neq 0, \tag{41}$$

$$= \begin{bmatrix} A_{M_-}^+ - A_{M_-}^+ ah - ph \\ h \end{bmatrix}, \qquad d = 0, \tag{42}$$

where

$$d = (I - AA_{M_-}^+)a, \quad p = (I - A_{M_-}^+ A)M_-^{-1}\tilde{m}, \quad h = (q^T/q^T Mq)MU,$$

$$U = \begin{bmatrix} A_{M_-}^+ \\ 0_{1\times m} \end{bmatrix}, \quad q = \begin{bmatrix} A_{M_-}^+ a + p \\ -1 \end{bmatrix}, \quad M = \begin{bmatrix} M_- & \tilde{m} \\ \tilde{m}^T & \bar{m} \end{bmatrix}.$$

Proof. Consider a system of linear equations

$$A\tilde{x} = b, \tag{43}$$

where A is an m by $(n-1)$ matrix, b is an m-vector, and $\tilde{x}$ is an $(n-1)$-vector. We assume that the generalized M_--inverse of A, which we denote by the $(n-1)$ by m matrix $A_{M_-}^+$, is given. Here, M_- is the symmetric positive-definite $(n-1)$ by $(n-1)$ principal minor of the matrix M.

Let us next add an extra column vector a of m components to the matrix A in the equation set above to get

$$Bx = [A\,|a]x = b, \tag{44}$$

where x is now an n-vector and needs to be determined.

We aim to find the solution x of Eq. (44) so that

$$\begin{aligned} G(z,r) &= \|Bx - b\|^2 \\ &= \left\| \left[A\,\middle|\,a\right]\begin{bmatrix} z \\ r \end{bmatrix} - b \right\|^2 \\ &= \|Az + ar - b\|^2, \end{aligned} \tag{45}$$

$$H = \left\| M^{1/2}x \right\|^2 \tag{46}$$

are both minimized. In Eq. (45), we have denoted

$$x = \begin{bmatrix} z \\ r \end{bmatrix},$$

where z is an $(n-1)$-vector and r is a scalar. It is also noted that

$$M=\begin{bmatrix} M_- & \tilde{m} \\ \tilde{m}^T & \bar{m} \end{bmatrix}, \tag{47}$$

where M is a positive definite n by n matrix obtained by augmenting the positive definite matrix M_- by one last column $\tilde{m}$ of $(n-1)$ components and then one last row $[\tilde{m}^T|\bar{m}]$ of n components, where $\bar{m}$ is a scalar.

We begin with minimizing $G(z,r)$ for a fixed value of r, say $r=r_0$. The value of $z(r_0)$ that minimizes $G(z(r_0),r_0)$ is then given by

$$z(r_0)=A^+_{M_-}(b-ar_0)+(I-A^+_{M_-}A)t_1, \tag{48}$$

where t_1 is an arbitrary $(n-1)$-vector, since $A^+_{M_-}$ is the generalized M_- inverse of A.

Substituting Eq. (48) in Eq. (45) gives

$$\begin{aligned} G(z(r_0),r_0) &= \left\| AA^+_{M_-}(b-ar_0)+ar_0-b \right\|^2 \\ &= \left\| dr_0-(I-AA^+_{M_-})b \right\|^2, \end{aligned} \tag{49}$$

where

$$d=(I-AA^+_{M_-})a.$$

Next, we need to find r_0 so that $G(z(r_0),r_0)$ is a minimum. From Eq. (49), we see that the value of d becomes important for this minimization, because when $d=0$, $G(z(r_0),r_0)$ is not a function of r_0. Therefore, two separate cases, when $d\neq 0$ and when $d=0$, need to be considered. Note that $d=0$ means that the column vector a is a linear combination of the columns of the matrix A (see Property 5.8 in the Appendix).

(a) Case $d\neq 0$. We shall determine first r_0 that minimizes G and then find $(I-A^+_{M_-}A)t_1$ that minimizes H.

Using the definition of the MP inverse, the scalar r_0 that minimizes $G(z(r_0),r_0)$ is given by

$$r_0=d^+(I-AA^+_{M_-})b+(1-d^+d)t_2, \tag{50}$$

where t_2 is an arbitrary scalar. Since

$$\begin{aligned} d^+A &= (d^Td)^{-1}[(I-AA^+_{M_-})a]^TA \\ &= (d^Td)^{-1}a^T(I-AA^+_{M_-})^TA \\ &= 0 \end{aligned}$$

and

$$d^+d = ((d^Td)^{-1}d^T)d = 1,$$

we have

$$r_0 = d^+b. \tag{51}$$

By Eqs. (48) and (51), G is minimized by the vector

$$\begin{aligned} x &= \begin{bmatrix} z(r_0) \\ r_0 \end{bmatrix} \\ &= \begin{bmatrix} A^+_{M_-}(b - ad^+b) + (I - A^+_{M_-}A)t_1 \\ d^+b \end{bmatrix}, \end{aligned} \tag{52}$$

where the vector t_1 is arbitrary. For convenience, we write Eq. (52) as

$$x = f + E(I - A^+_{M_-}A)e. \tag{53}$$

In Eq. (53) we have denoted

$$f = \begin{bmatrix} A^+_{M_-}(I - ad^+) \\ d^+ \end{bmatrix} b, \quad E = \begin{bmatrix} I_{n-1} \\ 0_{1\times(n-1)} \end{bmatrix}, \quad e = (I - A^+_{M_-}A)t_1,$$

where I_{n-1} is the $(n-1)$ by $(n-1)$ identity matrix and $0_{1\times(n-1)}$ is the zero row vector of $(n-1)$ components. Note that $I - A^+_{M_-}A$ is idempotent.

Now, we determine e so that $H = \left\| M^{1/2}x \right\|^2$ is minimized. Taking the partial derivative of H with respect to e, we have

$$\begin{aligned} \partial H/\partial e &= \partial \left\| M^{1/2}x \right\|^2 /\partial e \\ &= \partial \left\| M^{1/2}(f + E(I - A^+_{M_-}A)e) \right\|^2 /\partial e \\ &= 2(I - A^+_{M_-}A)^T E^T M(f + E(I - A^+_{M_-}A)e) = 0. \end{aligned} \tag{54}$$

Since

$$(I - A^+_{M_-}A)^T = M_-(I - A^+_{M_-}A)M_-^{-1}$$

and

$$E^T M E = M_-$$

(see Property 5.9 in the Appendix), from the last equality of Eq. (54) we have

$$\begin{aligned}&[M_-(I-A^+_{M_-}A)M_-^{-1}]M_-(I-A^+_{M_-}A)e\\&=-[M_-(I-A^+_{M_-}A)M_-^{-1}]E^TMf,\end{aligned} \tag{55}$$

which yields

$$M_-(I-A^+_{M_-}A)e=-[M_-(I-A^+_{M_-}A)M_-^{-1}]E^TMf. \tag{56}$$

It should be noted that, since

$$\begin{aligned}(I-A^+_{M_-}A)e&=(I-A^+_{M_-}A)[(I-A^+_{M_-}A)t_1]\\&=(I-A^+_{M_-}A)t_1=e,\end{aligned}$$

the left-hand side of Eq. (56) can be rewritten as

$$M_-e=-[M_-(I-A^+_{M_-}A)M_-^{-1}]E^TMf. \tag{57}$$

Since

$$\begin{aligned}E^TMf&=\begin{bmatrix}I_{n-1}, & 0^T_{n-1}\end{bmatrix}\begin{bmatrix}M_- & \tilde{m}\\ \tilde{m}^T & \bar{m}\end{bmatrix}\begin{bmatrix}A^+_{M_-}(I-ad^+)\\ d^+\end{bmatrix}b\\&=M_-A^+_{M_-}(I-ad^+)b+\tilde{m}d^+b,\end{aligned}$$

we obtain the $(n-1)$-vector e that minimizes H as

$$\begin{aligned}e&=-(I-A^+_{M_-}A)M_-^{-1}E^TMf\\&=-(I-A^+_{M_-}A)M_-^{-1}\tilde{m}d^+b\\&=-pd^+b\end{aligned} \tag{58}$$

where

$$p=(I-A^+_{M_-}A)M_-^{-1}\tilde{m}.$$

We note again that the matrix M_- is positive definite. Thus, we have

$$\begin{aligned}x&=B^+_Mb\\&=\begin{bmatrix}z(r_0)\\ r_0\end{bmatrix}\\&=\begin{bmatrix}A^+_{M_-}(b-ad^+b)+e\\ d^+b\end{bmatrix}\\&=\begin{bmatrix}A^+_{M_-}-A^+_{M_-}ad^+-pd^+\\ d^+\end{bmatrix}b,\end{aligned} \tag{59}$$

which gives

$$B_M^+ = [A \mid a]_M^+ \\ = \begin{bmatrix} A_{M_-}^+ - A_{M_-}^+ ad^+ - pd^+ \\ d^+ \end{bmatrix},$$

when $d = (I - AA_{M_-}^+)a \neq 0$. □

(b) Case $d = 0$. When $d = 0$, Equation (49) becomes

$$G = \|(I - AA_{M_-}^+)b\|^2,$$

which does not depend on r_0. As a result, by the relation (48), the vector x that minimizes G is given by

$$x = \begin{bmatrix} z(r_0) \\ r_0 \end{bmatrix} \\ = \begin{bmatrix} A_{M_-}^+(b - ar_0) + (I - A_{M_-}^+ A)t_1 \\ r_0 \end{bmatrix}, \tag{60}$$

which can be expressed as

$$x = \begin{bmatrix} A_{M_-}^+(b - ar_0) \\ r_0 \end{bmatrix} + \begin{bmatrix} I_{n-1} \\ 0_{1\times(n-1)} \end{bmatrix} (I - A_{M_-}^+ A)t_1, \tag{61}$$

or alternatively as

$$x = g + E(I - A_{M_-}^+ A)j, \tag{62}$$

where we have defined

$$g = \begin{bmatrix} A_{M_-}^+(b - ar_0) \\ r_0 \end{bmatrix}, \quad E = \begin{bmatrix} I_{n-1} \\ 0_{1\times(n-1)} \end{bmatrix}, \quad j = (I - A_{M_-}^+ A)t_1.$$

For a fixed r_0, we next find $j(r_0)$ so that

$$H(j, r_0) := \|x\|_M^2$$

is minimized. Minimizing with respect to j, we obtain

$$\partial H/\partial j = \partial \left\| M^{1/2}x \right\|^2 / \partial j \\ = \partial \left\| M^{1/2}(g + E(I - A_{M_-}^+ A)j) \right\|^2 / \partial j, \\ = 2(I - A_M^+ A)^T E^T M(g + E(I - A_{M_-}^+ A)j) = 0. \tag{63}$$

Since

$$(I - A^+_{M_-} A)^T = M_-(I - A^+_{M_-} A)M_-^{-1} \text{ and } E^T M E = M_-,$$

see Property 5.9 in the Appendix, from the last equality of Eq. (63) we have

$$\begin{aligned}&[M_-(I - A^+_M A)M_-^{-1}]M_-(I - A^+_{M_-} A)j\\ &= -[M_-(I - A^+_M A)M_-^{-1}]E^T Mg,\end{aligned} \tag{64}$$

which yields

$$M_-(I - A^+_{M_-} A)j = -M_-(I - A^+_{M_-} A)M_-^{-1} E^T Mg. \tag{65}$$

Noting that

$$\begin{aligned}(I - A^+_{M_-} A)j &= (I - A^+_{M_-} A)[(I - A^+_{M_-} A)t_1]\\ &= (I - A^+_{M_-} A)t_1\\ &= j,\end{aligned}$$

Eq. (65) can be written as

$$M_- j = -M_-(I - A^+_{M_-} A)M_-^{-1} E^T Mg. \tag{66}$$

Since

$$\begin{aligned}E^T Mg &= \begin{bmatrix} I_{n-1} & 0^T_{1\times(n-1)} \end{bmatrix} \begin{bmatrix} M_- & \tilde{m} \\ \tilde{m}^T & \bar{m} \end{bmatrix} \begin{bmatrix} A^+_{M_-}(b - ar_0) \\ r_0 \end{bmatrix}\\ &= M_- A^+_{M_-}(b - ar_0) + \tilde{m} r_0,\end{aligned}$$

Equation (66) gives

$$\begin{aligned}j &= -(I - A^+_{M_-} A)M_-^{-1}[M_- A^+_{M_-}(b - ar_0) + \tilde{m} r_0]\\ &= -[(I - A^+_{M_-} A)A^+_{M_-}](b - ar_0) - (I - A^+_{M_-} A)M_-^{-1}\tilde{m} r_0.\end{aligned}$$

Thus, we obtain

$$\begin{aligned}j(r_0) &= -(I - A^+_{M_-} A)M_-^{-1}\tilde{m} r_0\\ &= -pr_0,\end{aligned} \tag{67}$$

where we have denoted

$$p = (I - A^+_{M_-} A)M_-^{-1}\tilde{m}.$$

Now, we find the value of r_0 such that $H(j(r_0), r_0)$ is minimized. Substituting Eq. (67) in Eq. (60), we get

$$\begin{aligned} H(j(r_0), r_0) &= \left\| M^{1/2} x \right\|^2 \\ &= \left\| M^{1/2} \begin{bmatrix} A^+_{M\_}(b - ar_0) - pr_0 \\ r_0 \end{bmatrix} \right\|^2 \\ &= \left\| M^{1/2}(Ub - qr_0) \right\|^2, \end{aligned} \tag{68}$$

where

$$U = \begin{bmatrix} A^+_{M\_} \\ 0_{1\times m} \end{bmatrix}, \quad v = A^+_{M\_} a, \quad q = \begin{bmatrix} v + p \\ -1 \end{bmatrix},$$

and $0_{1\times m}$ is the zero row vector with m components.

Minimizing H with respect to r_0, we obtain

$$\partial H/\partial r_0 = 2q^T M(Ub - qr_0) = 0, \tag{69}$$

which gives, nothing that $q^T Mq$ is a scalar greater than zero,

$$\begin{aligned} r_0 &= \left(q^T / q^T Mq\right) MUb \\ &= hb, \end{aligned} \tag{70}$$

where

$$h = \left(q^T / q^T Mq\right) MU.$$

Since

$$j = (I - A^+_{M\_} A)t_1,$$

using Eqs. (67) and (70) in Eq. (60), we have

$$\begin{aligned} x &= [A \mid a]^+_M b \\ &= \begin{bmatrix} A^+_{M\_} b - A^+_{M\_} ar_0 + j \\ r_0 \end{bmatrix} \\ &= \begin{bmatrix} A^+_{M\_} b - A^+_{M\_} ahb - phb \\ hb \end{bmatrix}, \end{aligned} \tag{71}$$

which gives, arguing as we have done before,

$$\begin{aligned} B_M^+ &= [A|a]_M^+ \\ &= \begin{bmatrix} A_{M_-}^+ - A_{M_-}^+ ah - ph \\ h \end{bmatrix}, \qquad d = (I - AA_{M_-}^+)a = 0, \end{aligned}$$

where

$$p = (I - A_{M_-}^+ A)M_-^{-1}\tilde{m}, \quad h = \left(q^T/q^T Mq\right) MU, \quad U = \begin{bmatrix} A_{M_-}^+ \\ 0_{1\times m} \end{bmatrix},$$

$$q = \begin{bmatrix} v + p \\ -1 \end{bmatrix}, \quad v = A_{M_-}^+ a. \qquad \square$$

Corollary 3.2. When $M = \alpha I_n, \alpha > 0$, the generalized M-inverse of a matrix becomes the Moore–Penrose inverse and the recursive relations in Result 3.1 reduce to

$$\begin{aligned} [A|\,a]_{M=\alpha I_n}^+ &= [A|\,a]^+ \\ &= \begin{bmatrix} A^+ - A^+ ad^+ \\ d^+ \end{bmatrix}, \qquad d \neq 0, \end{aligned} \tag{72}$$

$$\begin{aligned} [A|a]_{M=\alpha I_n}^+ &= [A|a]^+ \\ &= \begin{bmatrix} A^+ - vv^T A^+/(1 + v^T v) \\ v^T A^+/(1 + v^T v) \end{bmatrix}, \qquad d = 0, \end{aligned} \tag{73}$$

where

$$d = (I - AA^+)a, \quad v = A^+ a.$$

Proof. When $M = \alpha I_n$, $\alpha > 0$, we have

$$\tilde{m} = 0, \quad p = (I - A_{M_-}^+ A)M_-^{-1}\tilde{m} = 0, \quad A_{M_-}^+ = A^+,$$

$$U = \begin{bmatrix} A^+ \\ 0_{1\times m} \end{bmatrix}, \quad v = A^+ a, \quad q = \begin{bmatrix} v \\ -1 \end{bmatrix}, \quad q^T q = 1 + v^T v,$$

$$\begin{aligned} h &= q^T U/q^T q \\ &= [1/(1 + v^T v)][v^T, -1]\begin{bmatrix} A^+ \\ 0_{1\times m} \end{bmatrix} \\ &= v^T A^+/(1 + v^T v). \end{aligned}$$

Thus, we have

$$[A|a]^+_{M=\alpha I_n} = \begin{bmatrix} A^+_{M_-} - A^+_{M_-} ad^+ - pd^+ \\ d^+ \end{bmatrix}_{M=\alpha I_n}$$
$$= \begin{bmatrix} A^+ - A^+ ad^+ \\ d^+ \end{bmatrix}, \qquad d \neq 0,$$

$$[A, a]^+_{M=\alpha I_n} = \begin{bmatrix} A^+_{M_-} - A^+_{M_-} ah - ph \\ h \end{bmatrix}_{M=\alpha I_n}$$
$$= \begin{bmatrix} A^+ - A^+ ah \\ h \end{bmatrix}$$
$$= \begin{bmatrix} A^+ - vv^T A^+/(1+v^T v) \\ v^T A^+/(1+v^T v) \end{bmatrix}, \qquad d = 0,$$

where

$$d = (I - AA^+)a \quad \text{and} \quad v = A^+ a.$$

These relations are identical to those given in Ref. 7.

4. Conclusions

In this paper, we present explicit recursive relations for the Moore–Penrose (MP) M-inverse of any matrix B. To the best of our knowledge, this is first time that such recursive relations for generalized M-inverses of matrices have been obtained. The value of these recursive relations is seen from their ready applicability to fields as diverse as analytical mechanics, statistics, filtering theory, signal processing, optimization, and controls, where additional information, which often takes the form of augmenting a matrix by a row or a column, is required to be processed in a recursive manner.

The formulae for the MP M-inverse of a matrix B obtained by augmenting a given matrix by a row and by augmenting it by a column are different. Due to the reason that these row-wise and column-wise MP M-inverse formulae may not be easily obtained from each other as in the case of the regular MP inverse, we provide herein both recursive row-wise and column-wise formulae. In Section 2, the MP M-inverse formulae of the partitioned matrix $\begin{bmatrix} \hat{A} \\ \hat{a} \end{bmatrix}$ are derived. The analysis presented here points out that two separate cases require to be distinguished: when the additional row $\hat{a}$ is a linear combination of the rows of the matrix $\hat{A}$, and when it is not. Accordingly, different recursive relations for each of these cases are obtained. In Section 3, we derive the column-wise formulae for the MP M-inverse of the partitioned matrix $[A\,|\,a]$. Similar to Section 2, two distinct cases are considered: the case when the column vector a is not a linear combination $(d\neq 0)$ of the columns of A, and when it is $(d=0)$.

5. Appendix

This appendix provides several important properties, mainly of generalized M-inverses, which are used in the main proofs in the paper.

Property 5.1. Any n-vector x can be expressed as

$$x=\hat{A}_M^{+}(\hat{b}+y)+(I-\hat{A}_M^{+}\hat{A})w.$$

Proof. Since M is positive definite, the above equation can be rewritten as

$$\begin{aligned}x&=M^{-1/2}\left(M^{1/2}\hat{A}_M^{+}(\hat{b}+y)+M^{1/2}(I-\hat{A}_M^{+}\hat{A})w\right)\\&=M^{-1/2}(S+T),\end{aligned}$$

where

$$S=M^{1/2}\hat{A}_M^{+}(\hat{b}+y),\quad T=M^{1/2}(I-\hat{A}_M^{+}\hat{A})w.$$

Let us then define

$$E=M^{1/2}(I-\hat{A}_M^{+}\hat{A}),$$

so that

$$\begin{aligned}E^TS&=\left[M^{1/2}(I-\hat{A}_M^{+}\hat{A})\right]^T M^{1/2}\hat{A}_M^{+}(\hat{b}+y)\\&=(I-\hat{A}_M^{+}\hat{A})^T M^{1/2}M^{1/2}\hat{A}_M^{+}(\hat{b}+y)\\&=\left[M(I-\hat{A}_M^{+}\hat{A})M^{-1}\right]M\hat{A}_M^{+}(\hat{b}+y)=0.\end{aligned}$$

Since T belongs to the range space of E and from the above equation S belongs to the null space of E^T, $S+T$ spans the whole space R^n. Because M is a positive definite matrix, $M^{-1/2}(S+T)$ also spans the whole space R^n. Thus, any n-vector x can always be expressed as

$$x=M^{-1/2}(S+T)=\hat{A}_M^{+}(\hat{b}+y)+(I-\hat{A}_M^{+}\hat{A})w. \qquad \square$$

Property 5.2. The row vector $c=0$ if and only if the row vector $\hat{a}$ is a linear combination of rows of the matrix $\hat{A}$.

Proof. Let us assume that ρ is a row vector of $(m-1)$ components. We need to show that

$$c=0 \Leftrightarrow \hat{a}=\rho\hat{A}.$$

When

$$c=\hat{a}(I-\hat{A}_M^+\hat{A})=0 \quad \text{or} \quad \hat{a}=\hat{a}\hat{A}_M^+\hat{A},$$

we have $\hat{a}=\rho\hat{A}$, where $\rho=\hat{a}\hat{A}_M^+$. If $\hat{a}=\rho\hat{A}$, then postmultiplying $\hat{a}=\rho\hat{A}$ by $I-\hat{A}_M^+\hat{A}$, we have $\hat{a}(I-\hat{A}_M^+\hat{A})=c=0$. Hence, the result. □

Property 5.3. $c_M^+=M^{-1}c^T/(cM^{-1}c^T)$.

Proof. If $M^{-1}c^T/(cM^{-1}c^T)$ is the MP M-inverse of c, it must satisfy Eqs. (1)–(4) as follows:

$$\begin{aligned}
cc_M^+c &= c\ M^{-1}c^T/(cM^{-1}c^T)c=c,\\
c_M^+cc_M^+ &= \left[M^{-1}c^T/\left(cM^{-1}c^T\right)\right]c\left[M^{-1}c^T/\left(cM^{-1}c^T\right)\right]\\
&= \left[M^{-1}c^T/\left(cM^{-1}c^T\right)\right]\left[cM^{-1}c^T/\left(cM^{-1}c^T\right)\right]\\
&= M^{-1}c^T/(cM^{-1}c^T)\\
&= c_M^+,
\end{aligned}$$

$$\begin{aligned}
(cc_M^+)^T &= \left[c\ M^{-1}c^T/\left(cM^{-1}c^T\right)\right]^T\\
&= c\ M^{-1}c^T/(cM^{-1}c^T)\\
&= cc_M^+,
\end{aligned}$$

$$\begin{aligned}
(c_M^+c)^T &= \left[M^{-1}c^T/\left(cM^{-1}c^T\right)c\right]^T\\
&= c^TcM^{-1}/(cM^{-1}c^T)\\
&= M\left[M^{-1}c^T/\left(cM^{-1}c^T\right)c\right]M^{-1}\\
&= Mc_M^+cM^{-1}.
\end{aligned}$$

Thus, $M^{-1}c^T/(cM^{-1}c^T)$ is the MP M-inverse of c. □

Property 5.4. $cc_M^+ = 1.$

Proof. Using Property 5.3 above, we have

$$cc_M^+ = c\left[M^{-1}c^T/cM^{-1}c^T\right] = 1. \qquad \square$$

Property 5.5. $\hat{A}c_M^+ = 0.$

Proof. Since

$$c_M^+ = M^{-1}c^T/(cM^{-1}c^T) \quad \text{and} \quad c = \hat{a}(I - \hat{A}_M^+\hat{A}),$$

we have

$$\begin{aligned}\hat{A}c_M^+ &= \hat{A}\left[M^{-1}c^T/\left(cM^{-1}c^T\right)\right]\\ &= \hat{A}M^{-1}[\hat{a}(I - \hat{A}_M^+\hat{A})]^T/\left(cM^{-1}c^T\right)\\ &= \hat{A}M^{-1}(I - \hat{A}_M^+\hat{A})^T\hat{a}^T/\left(cM^{-1}c^T\right),\\ &= \hat{A}M^{-1}[M(I - \hat{A}_M^+\hat{A})M^{-1}]\hat{a}^T/\left(cM^{-1}c^T\right)\\ &= [\hat{A}(I - \hat{A}_M^+\hat{A})]M^{-1}\hat{a}^T/\left(cM^{-1}c^T\right) = 0. \qquad \square\end{aligned}$$

Property 5.6. $l_1^T K^T M\left(\hat{A}_M^+\hat{b} + c_M^+(s - \hat{a}\hat{A}_M^+\hat{b})\right) = 0$, where $K = (I - \hat{A}_M^+\hat{A})(I - c_M^+c)$.

Proof. Since

$$\begin{aligned}K^T M &= (I - c_M^+c)^T(I - \hat{A}_M^+\hat{A})^T M\\ &= [M(I - c_M^+c)M^{-1}][M(I - \hat{A}_M^+\hat{A})M^{-1}]M\\ &= M(I - c_M^+c)(I - \hat{A}_M^+\hat{A}),\end{aligned}$$

we have

$$K^T M\hat{A}_M^+ = M(I - c_M^+c)(I - \hat{A}_M^+\hat{A})\hat{A}_M^+ = 0.$$

Using, $\hat{A}c_M^+ = 0$, (see Property 5.5 above) we have

$$\begin{aligned}K^T Mc_M^+ &= M(I - c_M^+c)(I - \hat{A}_M^+\hat{A})c_M^+\\ &= M(I - c_M^+c)c_M^+\\ &= 0.\end{aligned}$$

Thus,

$$\begin{aligned} & l_1^T K^T M\left(\hat{A}_M^+\hat{b}+c_M^+(s-\hat{a}\hat{A}_M^+\hat{b})\right) \\ & \quad = l_1^T\left(K^T M\hat{A}_M^+\right)\hat{b}+l_1^T\left(K^T Mc_M^+\right)(s-\hat{a}\hat{A}_M^+\hat{b})=0. \end{aligned}$$ □

Property 5.7. $(I+u^T u)^{-1}=I-u^T u/(1+uu^T)$, where u is a suitably dimensioned row vector.

Proof. If $I-u^T u/(1+uu^T)$ is the inverse of the matrix $I+u^T u$, we must have

$$\begin{aligned} (I+u^T u)^{-1}(I+u^T u) &= (I+u^T u)(I+u^T u)^{-1} \\ &= I. \end{aligned}$$

Since

$$\begin{aligned} (I+u^T u)^{-1}(I+u^T u) &= (I+u^T u)(I+u^T u)^{-1} \\ &= I+u^T u-u^T u/(1+uu^T)-[u^T u/(1+uu^T)]u^T u, \\ &= I+u^T u\left[1-1/(1+uu^T)-uu^T/(1+uu^T)\right], \\ &= I+u^T u\left[(1+uu^T-1-uu^T)/(1+uu^T)\right] \\ &= I, \end{aligned}$$

the result follows. □

Property 5.8. The column vector $d=0$ if and only if the column vector a is a linear combination of columns of the matrix A.

Proof. Let us assume that γ is a column vector of $n-1$ components. We need to show that

$$d=0 \Leftrightarrow a=A\gamma.$$

When $d=(I-AA_{M_-}^+)a=0$, $a=AA_{M_-}^+a$, and so we have $a=A\gamma$ where $\gamma=A_{M_-}^+a$. When $a=A\gamma$, then premultiplying $a=A\gamma$ by $I-AA_{M_-}^+$ we have $(I-AA_{M_-}^+)a=d=0$. Hence, the result. □

Property 5.9. $E^T ME=M_-$.

Proof. Using

$$E=\begin{bmatrix} I_{n-1} \\ 0_{1\times(n-1)} \end{bmatrix} \quad \text{and} \quad M=\begin{bmatrix} M_- & \widetilde{m} \\ \widetilde{m}^T & \overline{m} \end{bmatrix},$$

we have

$$\begin{aligned} E^T M E &= \left[I_{n-1} \,|\, 0^T_{1\times(n-1)} \right] \begin{bmatrix} M_- & \widetilde{m} \\ \widetilde{m}^T & \overline{m} \end{bmatrix} \begin{bmatrix} I_{n-1} \\ 0_{1\times(n-1)} \end{bmatrix} \\ &= [M_- \;\; \widetilde{m}] \begin{bmatrix} I_{n-1} \\ 0_{1\times(n-1)} \end{bmatrix} \\ &= M_-. \end{aligned}$$

□

References

1. MOORE, E. H., *On the Reciprocal of the General Algebraic Matrix*, Bulletin of the American Mathematical Society, Vol. 26, pp. 394–395, 1920.
2. PENROSE, R. A., *Generalized Inverses for Matrices*, Proceedings of the Cambridge Philosophical Society, Vol. 51, pp. 406–413, 1955.
3. GREVILLE, T. N. E., *Some Applications of the Pseudoinverse of a Matrix*, SIAM Review, Vol. 2, pp. 15–22, 1960.
4. ALBERT, A., *Regression and the Moore-Penrose Pseudoinverse*, Academic Press, New York, NY, 1972.
5. PRINGLE, R. M., and RAYNER, A. A., *Generalized Inverse Matrices*, Hafner, New York, NY, 1971.
6. RAO, C. R., and MITRA, S. K., *Generalized Inverses of Matrices and Their Applications*, Wiley, New York, NY, 1971.
7. UDWADIA, F. E., and KALABA, R. E., *An Alternative Proof of the Greville Formula*, Journal of Optimization Theory and Applications, Vol. 94, pp. 23–28, 1997.
8. UDWADIA, F. E., and KALABA, R. E., *A Unified Approach for the Recursive Determination of Generalized Inverses*, Computers and Mathematics with Applications, Vol. 37, pp. 125–130, 1999.
9. UDWADIA, F. E., and KALABA, R. E., *Sequential Determination of the {1, 4}-Inverse of a Matrix*, Journal of Optimization Theory and Applications, Vol. 117, pp. 1–7, 2003.
10. RAO, C. R., *A Note on a Generalized Inverse of a Matrix with Applications to Problems in Mathematical Statistics*, Journal of the Royal Statistical Society, Vol. 24 B, pp. 152–158, 1962.
11. GRAYBILL, F., *Matrices and Applications to Statistics*, 2nd Edition, Wadsworth, Belmont, California, 1983.
12. KALABA, R. E., and UDWADIA, F. E., *Associative Memory Approach to the Identification of Structural and Mechanical Systems*, Journal of Optimization Theory and Applications, Vol. 76, pp. 207–223, 1993.

13. UDWADIA, F. E., and KALABA, R. E., *Analytical Dynamics: A New Approach*, Cambridge University Press, Cambridge, England, 1996.
14. UDWADIA, F. E., *On Constrained Motion*, Applied Mathematics and Computations, Vol. 164, pp. 313–320, 2005.