JOURNAL OF OPTIMIZATION THEORY AND APPLICATIONS: Vol. 94, No. 1, pp. 23-28, JULY 1997

An Alternative Proof of the Greville Formula

F. E. UDWADIA[1] AND R. E. KALABA[2]

Abstract. A simple proof of the Greville formula for the recursive computation of the Moore-Penrose (MP) inverse of a matrix is presented. The proof utilizes no more than the elementary properties of the MP inverse.

Key Words. Moore-Penrose inverse, simple proof for recursive relation, Greville formula.

1. Introduction

The recursive determination of the MP inverse of a matrix has found extensive application in the fields of statistical inference and estimation theory (Refs. 1 and 2), and more recently in the field of analytical dynamics (Ref. 3). The reason for its extensive applicability is that it provides a systematic method to generate updates, whenever a sequential addition of data or new information is made available and updated estimates which take into account this additional information are required.

The recursive scheme for the computation of the Moore-Penrose (MP) inverse of a matrix (Refs. 4 and 5) was ingeniously obtained in a famous paper by Greville in 1960 (Ref. 6). However, due to the complexity of the solution technique, the Greville proof is not quoted or outlined even in specialized texts which deal solely with generalized inverses of matrices (e.g., books like Refs. 2 and 7-9), though his result is invariably stated because of its wide applicability. In this paper, we present a simple proof of the Greville result based on nothing more than the elementary properties of the MP inverse of a matrix.

The Greville result (1960) amounts to the following (Ref. 6). Let B be an $m \times k$ matrix, and let it be partitioned as $B=[A, a]$, where A consists of

[1]Professor of Mechanical Engineering, Civil Engineering, and Business, University of Southern California, Los Angeles, California.

[2]Professor of Biomedical Engineering, Electrical Engineering, and Economics, University of Southern California, Los Angeles, California.

0022-3239/97/0700-0023$12.50/0

the first $k-1$ columns of B and a is its last column. Since the case where $a=0$ is trivial, we shall consider in what follows only the case where $a \neq 0$. Then, the Moore–Penrose inverse of B can be written, utilizing knowledge of A^+, as

$$B^+ = \begin{bmatrix} A^+ - A^+ac^+ \\ c^+ \end{bmatrix}, \tag{1}$$

where,

$$c = (I - AA^+)a, \qquad \text{for } a \neq AA^+a, \tag{2a}$$

$$c = [1 + a^T(AA^T)^+a](AA^T)^+a/[a^T(AA^T)^+(AA^T)^+a], \qquad \text{for } a = AA^+a,\ a \neq 0. \tag{2b}$$

Thoughout this paper, the superscript + will indicate the MP-inverse.

In applications, the column vector a comprises new or additional information, while the matrix A comprises accumulated past data. The generalized inverse B^+ of the updated matrix B is then sought, given that the generalized inverse A^+ of the matrix A corresponding to past accumulated data is available.

Since right multiplication of AA^+ by any m-vector in the column space of A leaves that vector unchanged, Eq. (2a) when $a \neq AA^+a$ deals with a vector a, or new data, which is not in the column space of A. When $a = AA^+a$, as in Eq. (2b), the vector a is in the column space of A.

2. Proof for the Recursive Determination of the Moore–Penrose Inverse of a Matrix

Consider the least-squares problem

$$(Bx-b)^T(Bx-b) = \min, \qquad \text{over all } x. \tag{3}$$

Let the $m \times k$ matrix B be partitioned as $[A, a]$, where A is an $m \times (k-1)$ matrix and a is an m-vector. Similarly, let the column vector x be partitioned as $\left[\begin{smallmatrix} z \\ s \end{smallmatrix}\right]$ where z is a $(k-1)$-vector and s is a scalar. Equation (3) can then be expressed in the following form:

$$(Az + as - b)^T(Az + as - b) = \min, \qquad \text{over all vectors } z \text{ and scalars } s. \tag{4}$$

The least-square minimum-length solution of (3), by the definition of the MP inverse, can be written as

$$\tilde{x} = \begin{bmatrix} \tilde{z} \\ \tilde{s} \end{bmatrix} = B^+b \tag{5}$$

for arbitrary b. The solution given by Eq. (5) can be interpreted as follows. We are looking for all those pairs (z, s) from among all possible $(k-1)$-vectors z and scalars s such that

$$J(z, s) = \|Az + as - b\|$$

is a minimum; from these pairs, we select that pair for which the $(k-1)$-vector $\tilde{z}$ and the scalar $\tilde{s}$ are such that $z^T z + s^2$ is a minimum.

First, we begin by setting $s = s_o$, where s_o is some fixed scalar. Thus, we have

$$J(z, s_o) = \|Az - (b - as_o)\|. \tag{6}$$

Minimizing $J(z, s_o)$ such that $\hat{z}^T\hat{z}$ is also a minimum for all $(k-1)$-vectors z, using the definition of the MP-inverse, we get

$$\hat{z}(s_o) = A^+(b - as_o). \tag{7}$$

Thus for a given s_o, the vector $\hat{z}$ is a function of s_o. Using Eq. (7) in Eq. (6), we can now find s_o such that

$$\begin{aligned} J(\hat{z}(s_o), s_o) &= \|AA^+(b - as_o) + as_o - b\|^2 \\ &= \|(I - AA^+)as_o - (I - AA^+)b\|^2 \end{aligned} \tag{8}$$

is a minimum. Depending on the vector $c = (I - AA^+)a$, we must now deal with two distinct cases; $c \neq 0$ and $c = 0$.

(i) For $c \neq 0$, the unique value of s_o which minimizes $J(\hat{z}(s_o), s_o)$ is given by

$$\tilde{s}_o = [(I - AA^+)a]^+(I - AA^+)b = c^+(I - AA^+)b. \tag{9}$$

But the MP inverse of a nonzero m-vector c is given by

$$c^+ = c^T / c^T c,$$

and so Eq. (9) becomes

$$\tilde{s}_o = c^T(I - AA^+)b / c^T c. \tag{11}$$

Moreover, since

$$c^T = a^T(I - AA^+)^T$$

and the matrix $I - AA^+$ is symmetric and idempotent, Eq. (11) reduces simply to

$$\tilde{s}_o = c^+ b. \tag{12}$$

Combining this last expression with Eq. (7), we can rewrite Eq. (5) as

$$\begin{bmatrix} \tilde{z} \\ \tilde{s} \end{bmatrix} = B^+ b = \begin{bmatrix} A^+ - A^+ac^+ \\ c^+ \end{bmatrix} b. \tag{13}$$

Noting that this is true for all m-vectors b, we obtain

$$B^+ = \begin{bmatrix} A^+ - A^+ac^+ \\ c^+ \end{bmatrix}, \qquad \text{for } c \neq 0. \tag{14}$$

(ii) For $c=0$, we observe that $J(\hat{z}(s_o), s_o)$ as given in Eq. (8) is not a function of s_o. We thus only need to minimize

$$J_1(s_o) = \hat{z}(s_o)^T \hat{z}(s_o) + s_o^2,$$

over all values of s_o, where $\hat{z}(s_o)$ is given by Eq. (7). For convenience, we shall write it as

$$J_1(s_o) = s_o^2 + s_o^2 v_1^T v_1 - 2 s_o v_1^T v_2 + v_2^T v_2, \tag{15}$$

where

$$v_1 = A^+ a, \qquad v_2 = A^+ b.$$

The value of s_o that minimizes J_1 is obtained from

$$(1/2)\, \partial J_1/\partial s_o = \tilde{s}_o(1 + v_1^T v_1) - v_1^T v_2 = 0, \tag{16}$$

yielding

$$\tilde{s}_o = v_1^T v_2/(1 + v_1^T v_1)\,. \tag{17}$$

Since

$$(1/2)\, \partial^2 J_1/\partial s_o^2 = (1 + v_1^T v_1)$$

is greater than zero, the $\tilde{s}_o$ obtained in Eq. (17) indeed gives a minimum for J_1. Substituting the values of v_1 and v_2 into Eq. (17), we then obtain

$$\tilde{s}_o = (A^+a)^T A^+ b/[1 + a^T (A^+)^T A^+ a], \tag{18}$$

which may be simplified to

$$\tilde{s}_o = a^T (AA^T)^+ b/[1 + a^T (AA^T)^+ a] = e^T b, \tag{19}$$

where

$$e = (AA^T)^+ a/[1 + a^T (AA^T)^+ a]. \tag{20}$$

It is easy to show that $e=0$ if and only if $a=0$. The if part of the statement is obvious. Since the denominator of (20) is never zero, $e=0$ implies

$(AA^T)^+a=0$. Taking the singular value decomposition of A to be $A=U\Lambda V^T$ (where Λ is nonsingular and square), $e=0$ then requires $U\Lambda^{-2}U^Ta=0$, which in turn requires that $U^Ta=0$. But our condition $c=0$ implies that $a=AA^+a$, which in turn requires that $a=UU^Ta$. Using the fact that $U^Ta=0$, the last equation implies $a=0$. Hence, $a\neq 0$ implies $e\neq 0$.

Using Eq. (7), Eq. (5) can now be written as

$$\begin{bmatrix} \tilde{z} \\ \tilde{s} \end{bmatrix} = B^+b = \begin{bmatrix} A^+ - A^+ae^T \\ e^T \end{bmatrix} b, \tag{21}$$

from which it follows, as before, that

$$B^+ = \begin{bmatrix} A^+ - A^+ae^T \\ e^T \end{bmatrix}, \qquad \text{for } c=0. \tag{22}$$

Equations (14) and (22) constitute the Greville result. Equation (14) is identical to Eq. (2a); when e^T is set equal to c^+, Eqs. (22) and (2b) become identical because $c=e/(e^Te)$. In the event that a, or equivalently e, is a null vector, then c is a null vector. □

It is perhaps worthwhile noting that the three properties of the MP inverse which we have mainly used in obtaining the recursive relation are: (i) that the MP inverse solves the least-square minimum-length problem; (ii) the MP inverse of a column vector is proportional to its transpose; and (iii) the matrix $(I-AA^+)$ is symmetric and idempotent.

References

1. 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, Series B, Vol. 24, pp. 152–158, 1962.
2. Graybill, F., *Matrices and Applications to Statistics*, 2nd Edition, Wadsworth, Belmont, California, 1983.
3. Udwadia, F. E., and Kalaba, R. E., *Analytical Dynamics: A New Approach*, Cambridge University Press, Cambridge, England, 1996.
4. Moore, E. H., *On the Reciprocal of the General Algebraic Matrix*, Bulletin of the American Mathematical Society, Vol. 26, pp. 394–395, 1920.
5. Penrose, R. A., *Generalized Inverse for Matrices*, Proceedings of the Cambridge Philosophical Society, Vol. 51, pp. 406–413, 1955.
6. Greville, T. N. E., *Some Applications of the Pseudoinverse of a Matrix*, SIAM Review, Vol. 2, pp. 15–22, 1960.

7. ALBERT, A., *Regression and the Moore–Penrose Pseudoinverse*, Academic Press, New York, New York, 1972.
8. PRINGLE, R. M., and RAYNER, A. A., *Generalized Inverse Matrices*, Hafner, New York, New York, 1971.
9. RAO, C. R., and MITRA, S. K., *Generalized Inverse of Matrices and Its Applications*, Wiley, New York, New York, 1971.