SIAM J. MATRIX ANAL. APPL.
Vol. 13, No. 2, pp. 659–662, April 1992

013

A SHERMAN–MORRISON–WOODBURY IDENTITY FOR RANK AUGMENTING MATRICES WITH APPLICATION TO CENTERING*

KURT S. RIEDEL†

Abstract. Matrices of the form $\mathbf{A} + (\mathbf{V}_1 + \mathbf{W}_1)\mathbf{G}(\mathbf{V}_2 + \mathbf{W}_2)^*$ are considered where $\mathbf{A}$ is a *singular* $\ell \times \ell$ matrix and $\mathbf{G}$ is a nonsingular $k \times k$ matrix, $k \leq \ell$. Let the columns of $\mathbf{V}_1$ be in the column space of $\mathbf{A}$ and the columns of $\mathbf{W}_1$ be orthogonal to $\mathbf{A}$. Similarly, let the columns of $\mathbf{V}_2$ be in the column space of $\mathbf{A}^*$ and the columns of $\mathbf{W}_2$ be orthogonal to $\mathbf{A}^*$. An explicit expression for the inverse is given, provided that $\mathbf{W}_i^*\mathbf{W}_i$ has rank k. An application to centering covariance matrices about the mean is given.

Key words. linear algebra, Schur matrices, generalized inverses

AMS(MOS) subject classifications. 65R10, 33A65, 35K05, 62G20, 65P05

The well-known Sherman–Morrison–Woodbury matrix identity [1]:

$$(1) \qquad (\mathbf{A} + \mathbf{X}_1\mathbf{G}\mathbf{X}_2^T)^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{X}_1(\mathbf{G}^{-1} + \mathbf{X}_2^T\mathbf{A}^{-1}\mathbf{X}_1)^{-1}\mathbf{X}_2^T\mathbf{A}^{-1}$$

is widely used.[1] Several excellent review articles have appeared recently [2]–[4]. However, (1) is only valid when $\mathbf{A}$ is nonsingular. In this article, we consider matrix inverses of the form $\mathbf{A} + \mathbf{X}_1\mathbf{G}\mathbf{X}_2^T$ where the rank of $\mathbf{A} + \mathbf{X}_1\mathbf{G}\mathbf{X}_2^T$ is larger than the rank of $\mathbf{A}$.

We decompose the matrix $\mathbf{X}_1$ into $\mathbf{V}_1 + \mathbf{W}_1$, where the columns of $\mathbf{V}_1$ are contained in the column space of $\mathbf{A}$ and the columns of $\mathbf{W}_1$ are orthogonal to it. We denote the column space of $\mathbf{A}$ by $M(\mathbf{A})$. Similarly, we decompose $\mathbf{X}_2$ into $\mathbf{V}_2+\mathbf{W}_2$, where the columns of $\mathbf{V}_2$ are contained in the column space of $\mathbf{A}^*$ and the columns of $\mathbf{W}_2$ are orthogonal to $M(\mathbf{A}^*)$. The Moore–Penrose generalized inverse will be denoted by the superscript $^+$. We denote the $k \times k$ matrix $\mathbf{W}_i^*\mathbf{W}_i$ by $\mathbf{B}_i$ and define $\mathbf{C}_i \equiv \mathbf{W}_i(\mathbf{W}_i^*\mathbf{W}_i)^{-1}$. We will require $\mathbf{B}_i$ to be nonsingular. However, the rank of the perturbation k can be significantly less than the size of the original matrix. We note that $\mathbf{V}_i^*\mathbf{W}_i = 0$ and $\mathbf{W}_i^*\mathbf{C}_i = \mathbf{I}_k$. Finally, the projection operator onto the column space of $\mathbf{W}$ satisfies $\mathbf{W}_i\mathbf{B}_i^{-1}\mathbf{W}_i^* = \mathbf{W}_1\mathbf{C}_1^* = \mathbf{C}_2\mathbf{W}_2^*$.

THEOREM 1. *Let $\mathbf{A}$ be an $\ell \times \ell$ matrix of rank ℓ_1, $\ell_1 < \ell$, $\mathbf{V}_i$ and $\mathbf{W}_i$ be $\ell \times k$ matrices and $\mathbf{G}$ be a $k \times k$ nonsingular matrix. Let the columns of $\mathbf{V}_1 \in M(\mathbf{A})$ and the columns of $\mathbf{W}_1$ be orthogonal to $M(\mathbf{A})$. Similarly, let the columns of $\mathbf{V}_2 \in M(\mathbf{A}^*)$ and the columns of $\mathbf{W}_2$ be orthogonal to $M(\mathbf{A}^*)$. Let $\mathbf{B}_i \equiv \mathbf{W}_i^*\mathbf{W}_i$ have rank k. $M(\mathbf{W}_1) = M(\mathbf{W}_2)$. The matrix,*

$$\Omega \equiv \mathbf{A} + (\mathbf{V}_1 + \mathbf{W}_1)\mathbf{G}(\mathbf{V}_2 + \mathbf{W}_2)^*$$

has the following Moore–Penrose generalized inverse:

$$(2) \qquad \Omega^+ = \mathbf{A}^+ - \mathbf{C}_2\mathbf{V}_2^*\mathbf{A}^+ - \mathbf{A}^+\mathbf{V}_1\mathbf{C}_1^* + \mathbf{C}_2(\mathbf{G}^+ + \mathbf{V}_2^*\mathbf{A}^+\mathbf{V}_1)\mathbf{C}_1^*.$$

* Received by the editors July 12, 1990; accepted for publication (in revised form) February 19, 1991. This work was supported by U.S. Department of Energy grant DE-FG02-86ER53223.

† Courant Institute of Mathematical Sciences, New York University, New York, New York 10012 (reidel@mfdd4.nyu.edu).

[1] We denote the transpose of a matrix $\mathbf{A}$ by $\mathbf{A}^T$ and the hermitian or conjugate transpose by $\mathbf{A}^*$.

Proof. We recall that the Moore–Penrose inverse is the unique generalized inverse which satisfies the following four conditions [5, p. 26]:

(a) $\Omega\Omega^+\Omega = \Omega$,
(b) $\Omega^+\Omega\Omega^+ = \Omega^+$,
(c) $(\Omega\Omega^+)^* = \Omega\Omega^+$,
(d) $(\Omega^+\Omega)^* = \Omega^+\Omega$.

The identity is verified by direct computation,

$$\begin{aligned}\Omega\Omega^+ \equiv \mathbf{A}\mathbf{A}^+ &- \mathbf{A}\mathbf{C}_2\mathbf{V}_2^*\mathbf{A}^+ - \mathbf{A}\mathbf{A}^+\mathbf{V}_1\mathbf{C}_1^* + \mathbf{A}\mathbf{C}_2(\mathbf{G}^+ + \mathbf{V}_2^*\mathbf{A}^+\mathbf{V}_1)\mathbf{C}_1^* \\ &+(\mathbf{V}_1+\mathbf{W}_1)\mathbf{G}(\mathbf{V}_2+\mathbf{W}_2)^*\mathbf{A}^+ - (\mathbf{V}_1+\mathbf{W}_1)\mathbf{G}(\mathbf{V}_2+\mathbf{W}_2)^*\mathbf{C}_2\mathbf{V}_2^*\mathbf{A}^+ \\ &-(\mathbf{V}_1+\mathbf{W}_1)\mathbf{G}\ (\mathbf{V}_2+\mathbf{W}_2)^*\mathbf{A}^+\mathbf{V}_1\mathbf{C}_1^* \\ &+(\mathbf{V}_1+\mathbf{W}_1)\mathbf{G}(\mathbf{V}_2+\mathbf{W}_2)^*\mathbf{C}_2(\mathbf{V}_2^*\mathbf{A}^+\mathbf{V}_1)\mathbf{C}_1^* \\ &+(\mathbf{V}_1+\mathbf{W}_1)\mathbf{G}(\mathbf{V}_2+\mathbf{W}_2)^*\mathbf{C}_2\mathbf{G}^+\mathbf{C}_1^*.\end{aligned}$$

Since $\mathbf{W}_2$ is orthogonal to $\mathbf{A}^*$, we have $\mathbf{A}\mathbf{W}_2 = 0$, $\mathbf{W}_2^*\mathbf{A}^+ = 0$, and $\mathbf{V}_2^*\mathbf{W}_2 = 0$, which simplifies the previous expression to

$$\begin{aligned}\Omega\Omega^+ \equiv \mathbf{A}\mathbf{A}^+ &- \mathbf{A}\mathbf{A}^+\mathbf{V}_1\mathbf{C}_1^* + +(\mathbf{V}_1+\mathbf{W}_1)\mathbf{G}\mathbf{V}_2^*\mathbf{A}^+ \\ &- (\mathbf{V}_1+\mathbf{W}_1)\mathbf{G}\mathbf{W}_2^*\mathbf{C}_2\mathbf{V}_2^*\mathbf{A}^+ - (\mathbf{V}_1+\mathbf{W}_1)\mathbf{G}\ \ \mathbf{V}_2^*\mathbf{A}^+\mathbf{V}_1\mathbf{C}_1^* \\ &+(\mathbf{V}_1+\mathbf{W}_1)\mathbf{G}\mathbf{W}_2^*\mathbf{C}_2\mathbf{V}_2^*\mathbf{A}^+\mathbf{V}_1\mathbf{C}_1^* + (\mathbf{V}_1+\mathbf{W}_1)\mathbf{G}\mathbf{W}_2^*\mathbf{C}_2\mathbf{G}^+\mathbf{C}_1^*.\end{aligned}$$

This expression may be simplified using $\mathbf{G}\mathbf{W}_2^*\mathbf{C}_2\mathbf{G}^+\mathbf{C}_1^* = \mathbf{C}_1^*$, $\mathbf{G}\mathbf{W}_2^*\mathbf{C}_2\mathbf{V}_2^* = \mathbf{G}\mathbf{V}_2^*$, and $\mathbf{A}\mathbf{A}^+\mathbf{V}_1 = \mathbf{V}_1$ to

$$\Omega\Omega^+ \equiv \mathbf{A}\mathbf{A}^+ + \mathbf{W}_1\mathbf{C}_1^*,$$

and clearly condition (c) is satisfied.

The corresponding identity for $\Omega^+\Omega \equiv \mathbf{A}^+\mathbf{A} + \mathbf{C}_2\mathbf{W}_2^*$ requires the decomposition to satisfy $\mathbf{A}^+\mathbf{W}_1 = 0$, $\mathbf{W}_1^*\mathbf{A} = 0$, $\mathbf{V}_1^*\mathbf{W}_1 = 0$, and $\mathbf{V}_2\mathbf{A}^+\mathbf{A} = \mathbf{V}_2$. In addition, the matrix $\mathbf{G}$ must satisfy $\mathbf{C}_2\mathbf{G}^+\mathbf{C}_1^*\mathbf{W}_1\mathbf{G} = \mathbf{C}_2$ and $\mathbf{V}_1\mathbf{C}_1^*\mathbf{W}_1\mathbf{G} = \mathbf{V}_1\mathbf{G}$. These requirements guarantee that conditions (a), (b), and (d) are also satisfied. □

Remark. The conditions that $\mathbf{G}$ and $\mathbf{W}_i^*\mathbf{W}_i$ have rank k may be replaced by the somewhat weaker but more complicated conditions that $\mathbf{G}\mathbf{W}_2^*\mathbf{C}_2\mathbf{G}^+\mathbf{C}_1^* = \mathbf{C}_1^*$, $\mathbf{G}\mathbf{W}_2^*\mathbf{C}_2\mathbf{V}_2^* = \mathbf{G}\mathbf{V}_2^*$, $\mathbf{C}_2\mathbf{G}^+\mathbf{C}_1^*\mathbf{W}_1\mathbf{G} = \mathbf{C}_2$ and $\mathbf{V}_1\mathbf{C}_1^*\mathbf{W}_1\mathbf{G} = \mathbf{V}_1\mathbf{G}$.

Note that the generalized inverse in (2) is singular and tends to infinity as $\mathbf{W}_i$ approaches zero. Thus (2) does not reduce to the (1) as the perturbation tends to zero. When the perturbation of the column space of $\mathbf{A}$ is zero, i.e., $\mathbf{V} \equiv 0$, Theorem 1 simplifies to

$$\Omega^+ = \mathbf{A}^+ + \mathbf{C}_2\mathbf{G}^+\mathbf{C}_1. \tag{3}$$

When $\mathbf{A}$ is a symmetric matrix, the column spaces of $\mathbf{A}$ and $\mathbf{A}^*$ are identical. Thus, for the case of symmetric $\mathbf{A}$ and Ω, Theorem 1 reduces to Theorem 2.

THEOREM 2. *Let $\mathbf{A}$ be a symmetric $\ell \times \ell$ matrix of rank ℓ_1, $\ell_1 < \ell$, $\mathbf{V}$ and $\mathbf{W}$ be $\ell \times k$ matrices, and $\mathbf{G}$ be a $k \times k$ nonsingular matrix. Let $\mathbf{V} \in M(\mathbf{A})$ and the columns of $\mathbf{W}$ be orthogonal to $M(\mathbf{A})$. Let $\mathbf{B} \equiv \mathbf{W}^*\mathbf{W}$ have rank k. The matrix*

$$\Omega \equiv \mathbf{A} + (\mathbf{V} + \mathbf{W})\mathbf{G}(\mathbf{V} + \mathbf{W})^*$$

has the following Moore–Penrose generalized inverse:

$$\Omega^+ = \mathbf{A}^+ - \mathbf{C}\mathbf{V}^*\mathbf{A}^+ - \mathbf{A}^+\mathbf{V}\mathbf{C}^* + \mathbf{C}(\mathbf{G}^+ + \mathbf{V}^*\mathbf{A}^+\mathbf{V})\mathbf{C}^*. \tag{4}$$

For concreteness, we specialize the preceding identities to the case of rank one perturbations. In this special case, $k \equiv 1$, and $\mathbf{V}_i$ and $\mathbf{W}_i$ reduce to ℓ vectors v_i and w_i. In the nonsingular case, (1) reduces to Bartlett's identity [6]. It states for an arbitrary nonsingular $\ell \times \ell$ matrix $\mathbf{A}$ and ℓ vectors v_i,

$$(\mathbf{A} + v_1 v_2^*)^{-1} = \mathbf{A}^{-1} - \frac{(\mathbf{A}^{-1}v_1)(v_2^*\mathbf{A}^{-1})}{(1 + v_2^*\mathbf{A}^{-1}v_1)}. \tag{5}$$

In this case, Theorem 1 reduces to the analogous result for an arbitrary singular matrix $\mathbf{A}$ with a rank one perturbation which contains a component perpendicular to the column space of $\mathbf{A}$. Noting that $\mathbf{G} \equiv 1$ and $\mathbf{C}_i \equiv w_i/|w_i|^2$, Theorem 1 simplifies to the following result.

THEOREM 3. *Let $\mathbf{A}$ be an $\ell \times \ell$ matrix of rank ℓ_1, $\ell_1 < \ell$, and v_i, w_i, $i = 1, 2$ be ℓ vectors. Let $v_1 \in M(\mathbf{A})$ and w_1 be orthogonal to $M(\mathbf{A})$, and $v_2 \in M(\mathbf{A}^*)$ and w_2 be orthogonal to $M(\mathbf{A}^*)$. Assume w_2 is parallel to w_1 and $w_i \neq 0$. Let*

$$\Omega \equiv \mathbf{A} + (v_1 + w_1)(v_2 + w_2)^*.$$

The Moore–Penrose generalized inverse is

$$\Omega^+ = \mathbf{A}^+ - \frac{w_2 v_2^* \mathbf{A}^+}{|w_2|^2} - \frac{\mathbf{A}^+ v_1 w_1^*}{|w_1|^2} + (1 + v_2^*\mathbf{A}^+ v_1)\frac{w_2 w_1^*}{|w_1|^2|w_2|^2}. \tag{6}$$

This generalized inverse is singular and tends to infinity as $1/|w_1||w_2|$, as w_i approaches zero. Thus (6) does not reduce to Bartlett's identity.

The projection operator onto the row space of Ω is

$$P_{X_T} = \mathbf{A}^+\mathbf{A} + \frac{w_i w_i^*}{|w_i|^2}.$$

The symmetric version of Theorem 3 was originally developed and applied by the author in his statistical analysis of magnetic fusion data [7]. To estimate the regression parameters in ordinary least squares regression, the sum of the squares and products (SSP) matrix needs to be inverted. We apply Theorem 3 to determine the inverse of the SSP matrix in terms of the inverse of the covariance matrix of the covariates.

We decompose the independent variable vector x into a mean value vector $\bar{x}$ and a fluctuating part $\tilde{x}$. Thus the ith individual observation has the form

$$x_i = \bar{x} + \tilde{x}_i.$$

Let $\mathbf{X}$ denote the $n \times \ell$ data matrix whose rows consist of x_i^T and let $\tilde{\mathbf{X}}$ be the centered data matrix whose rows consist of $\tilde{x}_i^T$.

We assume that some of the independent variables x_k have not been varied. Thus $\tilde{\mathbf{X}}^*\tilde{\mathbf{X}}$ is singular.

The inverse of the uncentered sum of squares and crossproducts matrix $\mathbf{X}^*\mathbf{X}$ can now be expressed in terms of the Moore–Penrose generalized inverse of the centered covariance matrix $\tilde{\mathbf{X}}^*\tilde{\mathbf{X}}$. We decompose a multiple of the mean value vector $\sqrt{n}\bar{x}$ into $v + w$, where $v \in M(\tilde{\mathbf{X}}^*\tilde{\mathbf{X}})$ and $w \perp M(\tilde{\mathbf{X}}^*\tilde{\mathbf{X}})$. The data matrix has the form

$$\mathbf{X}^*\mathbf{X} = \tilde{\mathbf{X}}^*\tilde{\mathbf{X}} + n\bar{x}\bar{x}^T = \tilde{\mathbf{X}}^*\tilde{\mathbf{X}} + (v + w)(v + w)^*.$$

Thus we have rewritten $\mathbf{X}^*\mathbf{X}$ in a form appropriate to the application of Theorem 3.

In conclusion, the application of these matrix identities requires the decomposition of $\mathbf{X}_i$ into the orthogonal components $\mathbf{V}_i$ and $\mathbf{W}_i$. Thus our theorems are most useful in situations where the decomposition is trivial.

Acknowledgments. The helpful comments of the referees are gratefully acknowledged.

Note added in proof. A first order approximation to the matrix identity given in Theorem 1 in the limit of small perturbing matrices is given in equation (3.24) of [8].

REFERENCES

[1] W. J. DUNCAN, *Some devices for the solution of large sets of simultaneous equations (with an appendix on the reciprocation of partitioned matrices)*, The London, Edinburgh and Dublin Philosophical Magazine and Journal of Science, Seventh Series, 35 (1944), pp. 660–670.

[2] H. V. HENDERSON AND S. R. SEARLE, *On deriving the inverse of a sum of matrices*, SIAM Rev., 23 (1981), pp. 53–60.

[3] D. V. OUELLETE, *Schur complements and statistics*, J. Linear Algebra, 36 (1981), pp. 187–295.

[4] W. W. HAGER, *Updating the inverse of a matrix*, SIAM Rev., 31 (1989), pp. 221–239.

[5] C. R. RAO, *Linear Statistical Inference and Its Applications*, John Wiley and Sons, New York, 1973.

[6] M. S. BARTLETT, *An inverse matrix adjustment arising in discriminant analysis*, Ann. Math. Statist., 22 (1951), pp. 107-111.

[7] K. S. RIEDEL, *Advanced statistics for Tokamak transport: Collinearity and Tokamak to Tokamak variation*, New York University Report MF-118, National Technical Information Service Document no. DOE/ER/53223-98, 1989.

[8] G. W. STEWART, *On the perturbation of pseudo-inverses, projections and linear least squares*, SIAM Rev., 19 (1977), pp. 634–663.