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

KURT S. RIEDEL

Abstract. Matrices of the form \( A + (V_1 + W_1)G(V_2 + W_2)^* \) are considered where \( A \) is a singular \( \ell \times \ell \) matrix and \( G \) is a nonsingular \( k \times k \) matrix, \( k \leq \ell \). Let the columns of \( V_1 \) be in the column space of \( A \) and the columns of \( W_1 \) be orthogonal to \( A \). Similarly, let the columns of \( V_2 \) be in the column space of \( A^* \) and the columns of \( W_2 \) be orthogonal to \( A^* \). An explicit expression for the inverse is given, provided that \( W_i^*W_i \) 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 wellknown Sherman-Morrison-Woodbury matrix identity [1]:

\[
(A + X_1G X_2^T)^{-1} = A^{-1} - A^{-1}X_1(G^{-1} + X_2^T A^{-1} X_1)^{-1} X_2^T A^{-1}
\]

is widely used. Several excellent review articles have appeared recently [2-4]. However (1) is only valid when \( A \) is nonsingular \(^1\). In this article, we consider matrix inverses of the form \( A + X_1G X_2^T \) where the rank of \( A + X_1G X_2^T \) is larger than the rank of \( A \).

We decompose the matrix \( X_1 \) into \( V_1 + W_1 \), where the columns of \( V_1 \) are contained in the column space of \( A \) and the columns of \( W_1 \) are orthogonal to it. Similarly, we decompose \( X_2 \) into \( V_2 + W_2 \), where the columns of \( V_2 \) are contained in the column space of \( A^* \) and the columns of \( W_2 \) are orthogonal to \( M(A^*) \). We denote the column space of \( A \) by \( M(A) \). The Moore-Penrose generalized inverse will be denoted by the superscript \(^+\). We denote the \( k \times k \) matrix \( W_i^*W_i \) by \( B_i \) and define \( C_i = W_i(W_i^*W_i)^{-1} \). We will require \( B_i \) to be nonsingular. However the rank of the perturbation, \( k \), can be significantly less than the size of the original matrix. We note that \( V_i^*V_i = 0 \) and \( W_i^*C_i = I_k \). Finally the projection operator onto the column space of \( W_i \) satisfies \( W_iB_i^{-1}W_i^* = W_iC_i = C_iW_i^* \).

Theorem 1. Let \( A \) be a \( \ell \times \ell \) matrix of rank \( \ell_1, \ell_1 < \ell \), \( V_i \) and \( W_i \) be \( \ell \times k \) matrices and \( G \) be a \( k \times k \) nonsingular matrix. Let the columns of \( V_1 \in M(A) \) and the columns of \( W_1 \) be orthogonal to \( M(A) \). Similarly, let the columns of \( V_2 \in M(A^*) \) and the columns of \( W_2 \) be orthogonal to \( M(A^*) \). Let \( B_i = W_i^*W_i \) have rank \( k \). The matrix,

\[
\Omega \equiv A + (V_1 + W_1)G(V_2 + W_2)^*
\]

has the following Moore-Penrose generalized inverse:

\[
\Omega^+ = A^+ - C_2V_2^*A^+ - A^+V_1^*C_1^+ + C_2(G^+ + V_2^*A^+ V_1)C_1^+.
\]

\(^1\) We denote the transpose of a matrix, \( A \) by \( A^T \) and the hermitian or conjugate transpose by \( A^* \).
Proof: We recall that the Moore-Penrose inverse is the unique generalized inverse which satisfies the following four conditions, (Ref. [5], p. 26):

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

The identity is verified by direct computation,

\[
\Omega \Omega^+ = A A^+ - A C_2 V_2^* A^+ - A A^+ V_1 C_1^+ + A C_2 (G^+ + V_2 A^+ V_1) C_1^\dagger
\]

\[
+ (V_1 + W_1) G (V_2 + W_2)^* A^+ - (V_1 + W_1) G (V_2 + W_2)^* C_2 V_2^* A^+
\]

\[
- (V_1 + W_1) G (V_2 + W_2)^* A^+ V_1 C_1^+
\]

\[
+(V_1 + W_1) G (V_2 + W_2)^* C_2 V_2^* A^+ V_1 C_1^+ + (V_1 + W_1) G W_2^* C_2 G^+ C_1^\dagger.
\]

Since \( W_2 \) is orthogonal to \( A^* \), we have \( A W_2 = 0 \), \( W_2^* A^+ = 0 \), and \( V_2 W_2 = 0 \), which simplifies the previous expression to

\[
\Omega \Omega^+ = A A^+ - A A^+ V_1 C_1^+ + (V_1 + W_1) G V_2^* A^+
\]

\[
- (V_1 + W_1) G W_2^* C_2 V_2^* A^+ - (V_1 + W_1) G V_2^* A^+ V_1 C_1^+
\]

\[
+(V_1 + W_1) G W_2^* C_2 V_2^* A^+ V_1 C_1^+ + (V_1 + W_1) G W_2^* C_2 G^+ C_1^\dagger.
\]

This expression may be simplified using \( G W_2^* C_2 G^+ C_1^\dagger = C_1^\dagger \), and \( G W_2^* C_2 V_2^* = G V_2^* \), and \( A A^+ V_1 = V_1 \) to

\[
\Omega \Omega^+ = A A^+ + W_1 C_1^\dagger,
\]

and clearly condition (c) is satisfied.

The corresponding identity for \( \Omega^+ \Omega = A^+ A + C_2 W_2^* \) requires the decomposition to satisfy \( A^+ W_1 = 0 \), \( W_1^* A = 0 \), \( V_1 W_1 = 0 \), and \( V_2 A^+ = V_2 \). In addition, the matrix \( G \) must satisfy \( C_2 G^+ C_1^\dagger W_1 G = C_2 \) and \( V_1 C_1^\dagger W_1 G = V_1 G \). These requirements guarantee that conditions (a), (b) and (d) are also satisfied. \[\]

Remark: The conditions that \( G \) and \( W_1^* W_1 \) have rank \( k \) may be replaced by the somewhat weaker but more complicated conditions that \( G W_2^* C_2 G^+ C_1^\dagger = C_1^\dagger \), \( G W_2^* C_2 V_2^* = G V_2^* \), \( C_2 G^+ C_1^\dagger W_1 G = C_2 \) and \( V_1 C_1^\dagger W_1 G = V_1 G \).

Note that the generalized inverse in (2) is singular and tends to infinity as \( W_i \) approaches to zero. Thus (2) does not reduce to the (1) as the perturbation tends to zero. When the perturbation of the column space of \( A \) is zero, i.e. \( V = 0 \), theorem 1 simplifies to

\[
\Omega^+ = A^+ + C_2 G^+ C_1.
\]

When \( A \) is a symmetric matrix, the column spaces of \( A \) and \( A^* \) are identical. Thus, for the case of symmetric \( A \) and \( \Omega \), Thm. 1 reduces to

**Theorem 2.** Let \( A \) be a symmetric \( \ell \times \ell \) matrix of rank \( \ell_1 \), \( \ell_1 < \ell \), \( V \) and \( W \) be \( \ell \times k \) matrices and \( G \) be a \( k \times k \) nonsingular matrix. Let \( V \in M(A) \) and the
columns of \( W \) be orthogonal to \( M(A) \). Let \( B \equiv W^* W \) have rank \( k \). The matrix,
\[
\Omega \equiv A + (V + W)G(V + W)^* ,
\]
has the following Moore-Penrose generalized inverse:
\[
\Omega^+ = A^+ - C V^* A^+ - A^+ V C^* + C(G^+ + V^* A^+ V)C^*. \tag{4}
\]

For concreteness, we specialise the preceding identities to the case of rank one perturbations. In this special case, \( k = 1 \), and \( V_i \) and \( W_i \) reduce to \( \ell \) 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 \( A \) and \( \ell \) vectors \( v_i \),
\[
(A + v_1 v_2^*)^{-1} = A^{-1} - \frac{(A^{-1}v_1)(v_2^* A^{-1})}{(1 + v_2^* A^{-1}v_1)}. \tag{5}
\]

In this case, theorem 1 reduces to the analogous result for an arbitrary singular matrix \( A \) with a rank one perturbation which contains a component perpendicular to the column space of \( A \). Noting that \( G \equiv 1 \) and \( C_i \equiv w_i/|w_i|^2 \), theorem 1 simplifies to the following result.

**Theorem 3.** Let \( A \) be a \( \ell \times \ell \) matrix of rank \( \ell_1 \), \( \ell_1 < \ell \), and \( v_i, w_i, i = 1,2 \) be \( \ell \) vectors. Let \( v_1 \in M(A) \) and \( w_1 \) be orthogonal to \( M(A) \), and \( v_2 \in M(A^*) \) and \( w_2 \) be orthogonal to \( M(A^*) \). Assume \( w_2 \) is parallel to \( w_1 \) and \( w_i \neq 0 \). Let
\[
\Omega \equiv A + (v_1 + w_1)(v_2 + w_2)^* ,
\]
The Moore-Penrose generalized inverse is
\[
\Omega^+ = A^+ - \frac{w_2 v_2^* A^+}{|w_2|^2} - \frac{A^+ v_1 w_1^*}{|w_1|^2} + (1 + v_2^* A^+ v) \frac{w_2 w_1^*}{|w_1|^2 |w_2|^2}. \tag{6}
\]
This generalized inverse is singular and tends to infinity as \( 1/|w_1||w_2| \), as \( w_i \) approaches to zero. Thus (6) does not reduce to Bartlett’s identity.

The projection operator onto the row space of \( \Omega \) is
\[
P_{X^*} = A^+ A^* + \frac{w_i w_i^*}{|w_i|^2}.
\]

The symmetric version of Thm. 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 squares and products (SSP) matrix needs to be inverted. We apply Thm. 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 \( i \)-th individual observation has the form
\[
x_i = \bar{x} + \tilde{x}_i \tag{8}.
\]
Let \( X \) denote the \( n \times \ell \) data matrix whose rows consist of \( x_1^* \) and \( \tilde{X} \) be the centered data matrix whose rows consist of \( \tilde{x}_1^* \).

We assume that some of the independent variables, \( x_k \), have not been varied. Thus \( X^* X \) is singular. The inverse of the uncentered sum of squares and crossproducts
matrix, $X^*X$ can now be expressed in terms of the Moore Penrose generalized inverse of the centered covariance matrix, $\tilde{X}^*\tilde{X}$.

We decompose a multiple of the mean value vector, $\sqrt{n}\bar{x}$, into $v + w$, where $v \in M(\tilde{X}^*\tilde{X})$ and $w \perp M(\tilde{X}^*\tilde{X})$.

The data matrix has the form

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

Thus we have rewritten $X^*X$ in a form appropriate to the application of theorem 3.

In conclusion, the application of these matrix identities requires the decomposition of $X_i$ into the orthogonal components, $V_i$ and $W_i$. Thus our theorems are most useful in situations where the decomposition is trivial.

Acknowledgments
The helpful comments of the referees are gratefully acknowledged.

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, p. 660, (1944).
2. H.V. Henderson and S.R. Searle, “On deriving the inverse of a sum of matrices”, SIAM Review, 23, p.53, (1981).
3. D.V. Ouellette, “Schur complements and statistics”, Journal of Linear Algebra, 36, p. 187, (1981).
4. W.W. Hager, “Updating the inverse of a matrix”, SIAM Review, 31, p.221, (1989).
5. C.R. Rao, Linear Statistical Inference and Its Applications, p. 26,33, J. Wiley and Sons, New York, 1973.
6. M.S. Bartlett, “An inverse matrix adjustment arising in discriminant analysis”, Annals of Mathematical Statistics, vol. 22, p107, (1951).
7. K.S. Riedel, New York University Report MF-118, National Technical Information Service document no. DOE/ER/53223-98 "Advanced Statistics for Tokamak Transport: Collinearity and Tokamak to Tokamak Variation" (1989).