Block Matrix Formulas

John A. Gubner
Department of Electrical and Computer Engineering
University of Wisconsin–Madison

Abstract

We derive a number of formulas for block matrices, including the block matrix inverse formulas, determinant formulas, psuedoinverse formulas, etc.

If you find this writeup useful, or if you find typos or mistakes, please let me know at John.Gubner@wisc.edu

1. Preliminary Observations

Given a block matrix

$$\Phi := \begin{bmatrix} A & B \\ C & D \end{bmatrix},$$

if D is invertible, then the **Schur complement** of D is

$$\Sigma := A - BD^{-1}C.$$

It is easy to check that

$$\Phi = \begin{bmatrix} I & BD^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} \Sigma & 0 \\ 0 & D \end{bmatrix} \begin{bmatrix} I & 0 \\ D^{-1}C & I \end{bmatrix}. \tag{1}$$

Then $\det\Phi = \det\Sigma \det D$, and we see that given D is invertible, Φ is invertible if and only if Σ is invertible. When Φ is invertible, taking the inverse of (1) yields

$$\begin{aligned} \Phi^{-1} &= \begin{bmatrix} I & 0 \\ -D^{-1}C & I \end{bmatrix} \begin{bmatrix} \Sigma^{-1} & 0 \\ 0 & D^{-1} \end{bmatrix} \begin{bmatrix} I & -BD^{-1} \\ 0 & I \end{bmatrix} \\ &= \left[\begin{array}{c|c} \Sigma^{-1} & -\Sigma^{-1}BD^{-1} \\ \hline -D^{-1}C\Sigma^{-1} & D^{-1}C\Sigma^{-1}BD^{-1} + D^{-1} \end{array}\right]. \end{aligned} \tag{2}$$

Instead of assuming D is invertible, suppose A is invertible. The Schur complement of A is

$$\Theta := D - CA^{-1}B,$$

and

$$\Phi = \begin{bmatrix} I & 0 \\ CA^{-1} & I \end{bmatrix} \begin{bmatrix} A & 0 \\ 0 & \Theta \end{bmatrix} \begin{bmatrix} I & A^{-1}B \\ 0 & I \end{bmatrix}. \tag{3}$$

Then $\det \Phi = \det A \det \Theta$, and we see that given A is invertible, Φ is invertible if and only if Θ is invertible. If Φ is invertible,

$$\Phi^{-1} = \begin{bmatrix} I & -A^{-1}B \\ 0 & I \end{bmatrix} \begin{bmatrix} A^{-1} & 0 \\ 0 & \Theta^{-1} \end{bmatrix} \begin{bmatrix} I & 0 \\ -CA^{-1} & I \end{bmatrix}$$

$$= \left[\begin{array}{c|c} A^{-1} + A^{-1}B\Theta^{-1}CA^{-1} & -A^{-1}B\Theta^{-1} \\ \hline -\Theta^{-1}CA^{-1} & \Theta^{-1} \end{array}\right]. \tag{4}$$

2. Results

Using (1)–(4), there are several identities that can easily be found.

2.1. Determinant Formulas

First, comparing (1) and (3) yields

$$\det \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \det(A)\det(D - CA^{-1}B) = \det(A - BD^{-1}C)\det(D). \tag{5}$$

In particular, taking $A = I_n$ and $D = I_k$ yields

$$\det(I_k - CB) = \det(I_n - BC). \tag{6}$$

2.2. Matrix Inversion Formulas

Next, comparing the upper-left blocks of (2) and (4), we see that

$$[A - BD^{-1}C]^{-1} = A^{-1} + A^{-1}B[D - CA^{-1}B]^{-1}CA^{-1}, \tag{7}$$

which is known as the **Sherman–Morrison–Woodbury formula** or sometimes just the **Woodbury formula**. The remaining corresponding blocks are also equal. For example, combining the top row of (2) with the bottom row of (4), we obtain

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} = \left[\begin{array}{c|c} [A - BD^{-1}C]^{-1} & -[A - BD^{-1}C]^{-1}BD^{-1} \\ \hline -[D - CA^{-1}B]^{-1}CA^{-1} & [D - CA^{-1}B]^{-1} \end{array}\right]. \tag{8}$$

2.3. Pseudoinverse Formulas

A nice application of (8) is to the pseudoinverse of a 2×2 block matrix. Recall that if M is full rank, then its pseudoinverse is $M^\dagger = (M^* M)^{-1} M^*$, where M^* is the complex conjugate transpose of M. Consider the case

$$M = \begin{bmatrix} G & H \end{bmatrix} \quad \text{and} \quad M^* = \begin{bmatrix} G^* \\ H^* \end{bmatrix}.$$

Then

$$M^* M = \begin{bmatrix} G^* G & G^* H \\ H^* G & H^* H \end{bmatrix},$$

and by (8),

$$(M^* M)^{-1} = \left[\begin{array}{c|c} (G^* P_H^\perp G)^{-1} & -(G^* P_H^\perp G)^{-1} G^* H (H^* H)^{-1} \\ \hline -(H^* P_G^\perp H)^{-1} H^* G (G^* G)^{-1} & (H^* P_G^\perp H)^{-1} \end{array}\right], \tag{9}$$

where $P_H^\perp := I - H(H^* H)^{-1} H^*$ and $P_G^\perp := I - G(G^* G)^{-1} G^*$. It now easily follows that

$$(M^* M)^{-1} M^* = \left[\begin{array}{c} (G^* P_H^\perp G)^{-1} G^* P_H^\perp \\ \hline (H^* P_G^\perp H)^{-1} H^* P_G^\perp \end{array}\right] = \left[\begin{array}{c} (P_H^\perp G)^\dagger \\ \hline (P_G^\perp H)^\dagger \end{array}\right]. \tag{10}$$

2.4. Positive Semidefinite Matrices

A matrix U is **positive semidefinite** if $U^* = U$ and $x^* U x \geq 0$ for all vectors x. In this case, we use the notation $U \geq 0$. If U and V are Hermitian, we write $U \geq V$ if $U - V$ is positive semidefinite. For real matrices, the condition $U = U^*$ is equivalent to $U = U^\mathsf{T}$, where U^T denotes the transpose of U.

Suppose that Φ is real with $\Phi^\mathsf{T} = \Phi$. Then $A = A^\mathsf{T}$, $D = D^\mathsf{T}$, and $C = B^\mathsf{T}$. Now rewrite (1) as

$$\begin{bmatrix} A - BD^{-1}B^\mathsf{T} & 0 \\ 0 & D \end{bmatrix} = \begin{bmatrix} I & -BD^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} A & B \\ B^\mathsf{T} & D \end{bmatrix} \begin{bmatrix} I & 0 \\ -D^{-1}B^\mathsf{T} & I \end{bmatrix}.$$

Suppose the center matrix on the right is known to be positive semidefinite. Then since the matrices on either side are transposes of each other, it is easy to see that the matrix on the left-hand side is also positive semidefinite. It then follows that $A - BD^{-1}B^\mathsf{T}$ is positive semidefinite; in symbols, $A \geq BD^{-1}B^\mathsf{T}$.

2.4.1. Properties of the Schur Complement

A matrix U is **positive definite** if in addition to being positive semidefinite, $x^*Ux > 0$ for all $x \neq 0$. Such a matrix is nonsingular and therefore invertible. We now assume that A and D are positive definite and that $C = B^*$. Then we can always write

$$\Sigma = A - BD^{-1}B^* = A - A^{1/2}KA^{1/2} = A^{1/2}(I-K)A^{1/2},$$

where

$$K = A^{-1/2}BD^{-1}B^*A^{-1/2}.$$

It follows that $\det\Sigma = \det A \det(I-K)$, and so Σ is singular if and only if $I-K$ is singular. The assumption $D > 0$ implies $D^{-1} > 0$, and so $K \geq 0$. Hence, Σ is positive semidefinite if and only if $I-K$ is positive semidefinite, in which case the eigenvalues of $I-K$ are nonnegative, which implies the eigenvalues of K are less than or equal to one. Since $K \geq 0$ as well, the eigenvalues of K lie between zero and one.

We next show that $K = 0$ if and only if $B = 0$. If $B = 0$, then obviously $K = 0$. Conversely, suppose $K = 0$. Write $K = JJ^*$, where $J := A^{-1/2}BD^{-1/2}$. Then $K = 0$ implies $0 = \operatorname{tr} K = \operatorname{tr}(JJ^*) = \sum_{i,j}|J_{i,j}|^2$, which implies $J = 0$, and it follows that $B = 0$.

We can also show that if K is singular, then so is B^*. Suppose $x \neq 0$ and $Kx = JJ^*x = 0$. Since $\ker JJ^* = \ker J^*$, we have $x \in \ker J^*$. So $D^{-1/2}B^*A^{-1/2}x = 0$. It follows that $B^*A^{-1/2}x = 0$, which implies $A^{-1/2}x \in \ker B^*$. Since $x \neq 0$, $A^{-1/2}x \neq 0$.

2.4.2. Upper-Left Block Submatrices

Given a 2×2 block matrix U, we denote its upper-left block by $\{U\}_{\mathsf{ULB}}$. For the matrix Φ defined earlier, $\{\Phi\}_{\mathsf{ULB}} = A$. We show below that if $\Phi = \Phi^*$, then

$$\{U^*\Phi^{-1}U\}_{\mathsf{ULB}} \geq \{U^*\}_{\mathsf{ULB}}\{\Phi\}_{\mathsf{ULB}}^{-1}\{U\}_{\mathsf{ULB}}.$$

More specifically, if

$$U := \begin{bmatrix} P & Q \\ R & S \end{bmatrix},$$

then

$$\{U^*\Phi^{-1}U\}_{\mathsf{ULB}} \geq P^*A^{-1}P.$$

To derive the result, we first use (4) to compute the left-hand block column of $\Phi^{-1}U$. We get

$$\left[\begin{array}{c} \{A^{-1} + A^{-1}B\Theta^{-1}CA^{-1}\}P - A^{-1}B\Theta^{-1}R \\ \hline -\Theta^{-1}CA^{-1}P + \Theta^{-1}R \end{array}\right].$$

It follows that

$$\begin{aligned}
\{U^*\Phi^{-1}U\}_{\mathsf{ULB}} &= \left[\, P^* \mid R^* \,\right] \left[\begin{array}{c} \{A^{-1} + A^{-1}B\Theta^{-1}CA^{-1}\}P - A^{-1}B\Theta^{-1}R \\ \hline -\Theta^{-1}CA^{-1}P + \Theta^{-1}R \end{array}\right] \\
&= P^*A^{-1}P + P^*A^{-1}B\Theta^{-1}CA^{-1}P \\
&\quad - P^*A^{-1}B\Theta^{-1}R - R\Theta^{-1}CA^{-1}P + R^*\Theta^{-1}R \\
&= P^*A^{-1}P + (B^*A^{-1}P - R)^*\Theta^{-1}(CA^{-1}P - R).
\end{aligned}$$

Since $\Phi = \Phi^*$, we have $C = B^*$, it follows that $\{U^*\Phi^{-1}U\}_{\mathsf{ULB}} \geq P^*A^{-1}P$.

Remarks. (i) The above display does not depend on Q or S.

(ii) The only requirement on U is that its size be such that the required matrix multiplications are defined. The number of columns in Q and S must be the same, but this number is arbitrary, and could be zero, in which case U is a 2×1 block matrix. In particular, there is no requirement that U be a square matrix.

References

[1] W. W. Hager, "Updating the inverse of a matrix," *SIAM Review*, vol. 31, no. 2, pp. 221–239, 1989.

[2] Wikipedia, "Schur complement — Wikipedia, The Free Encyclopedia," [Online]. Available: `https://en.wikipedia.org/w/index.php?title=Schur_complement&oldid=658902757`, accessed July 6, 2015.

[3] Wikipedia, "Woodbury matrix identity — Wikipedia, The Free Encyclopedia," [Online]. Available: `https://en.wikipedia.org/w/index.php?title=Woodbury_matrix_identity&oldid=661757721`, accessed July 6, 2015.