Block Matrix Inverse and Woodbury Formula

Patrick Yin

May 28, 2020

1 Introduction

In this note, I will prove the block matrix inverse and the Woodbury formula. Despite being somewhat tedious and unintuitive, I have found that these formulas are really useful tools when proving ideas related to multivariate gaussians. For this reason, I want to document and share an explanation behind these two formulas.

2 Block Matrix Inverse

Any matrix $M \in \mathbb{R}^{m,n}$ can be partitioned as

$$M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}. \tag{1}$$

Theorem 1 (Block Matrix Inverse)**.** *Assuming A and D are invertible square matrices, we can express M^{-1} in terms of blocks A, B, C, D, A^{-1}, and D^{-1}:*

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} = \begin{bmatrix} S_1^{-1} & -A^{-1}BS_2^{-1} \\ -D^{-1}CS_1^{-1} & S_2^{-1} \end{bmatrix} \tag{2}$$

or equivalently,

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} = \begin{bmatrix} S_1^{-1} & -S_1^{-1}BD^{-1} \\ -S_2^{-1}CA^{-1} & S_2^{-1} \end{bmatrix} \tag{3}$$

where $S_1 = A - BD^{-1}C$ and $S_2 = D - CA^{-1}B$.

Proof. We want to find a M^{-1} such that $MM^{-1} = I$ and $M^{-1}M = I$. Let us write M^{-1} as

$$M^{-1} = \begin{bmatrix} W & X \\ Y & Z \end{bmatrix}. \tag{4}$$

Using $MM^{-1} = I$, we find that:

$$\begin{cases} AX + BZ = 0 \\ CW + DY = 0 \\ AW + BY = I \\ CX + DZ = I \end{cases} \tag{5}$$

Rearranging the first two equations, we get:

$$\begin{cases} X = -A^{-1}BZ \\ Y = -D^{-1}CW \end{cases} \tag{6}$$

Rearranging the next two equations and plugging in our new values for X and Y, we get:

$$\begin{cases} A^{-1} = (I - A^{-1}BD^{-1}C)W \\ D^{-1} = (I - D^{-1}CA^{-1}B)Z \end{cases} \tag{7}$$

Since $(AB)^{-1} = B^{-1}A^{-1}$, we can rearrange equation 7 to get:

$$\begin{cases} W = (A - BD^{-1}C)^{-1} \\ Z = (D - CA^{-1}B)^{-1} \end{cases} \tag{8}$$

After plugging in the values of X, Y, Z, and W that we have found in equation 6 and 8 into equation 4, we find that:

$$M^{-1} = \begin{bmatrix} (A - BD^{-1}C)^{-1} & -A^{-1}B(D - CA^{-1}B)^{-1} \\ -D^{-1}C(A - BD^{-1}C)^{-1} & (D - CA^{-1}B)^{-1} \end{bmatrix}. \tag{9}$$

We have proved equation 2. In order to show equation 3, we simply start with $M^{-1}M = I$ instead. Going along this route, we end up with an equivalent result for W and Z. Y and X are different (as excepted):

$$\begin{cases} X = -WBD^{-1} \\ Y = -ZCA^{-1} \end{cases} \tag{10}$$

□

Remark. *Why are S_1 and S_2 invertible? It seems that we didn't make such an assumption in our theorem. In fact, we can definitely construct a counterexample where A and D are invertible but S_1 and S_2 are not invertible. However, note that we are also making the assumption that M is invertible. Since inverses are unique, S_1 and S_2 must be invertible. This is not a rigorous proof but an intuitive explanation. I may come back and try to rigorously proof this in the future.*

3 Woodbury Formula

The block matrix inverse is nice, but can we break down S_1^{-1} and S_2^{-1} further? This is where Woodbury Formula comes in.

Theorem 2 (Woodbury Formula)**.** *Let E and G be square invertible matrices of dimension n-by-n and k-by-k respectively. Let F and H be matrices of size n-by-k and k-by-n respectively. Then,*

$$(E + FGH)^{-1} = E^{-1} - E^{-1}F(G^{-1} + HE^{-1}F)^{-1}HE^{-1} \tag{11}$$

Remark. *Notice that* S_1^{-1} *and* S_2^{-1} *fit the form of* $(E+FGH)^{-1}$*, so Woodbury Formula helps us break down these two terms.*

Proof. Equating equations 6 and 10 for X gives $-WBD^{-1} = -A^{-1}BZ$. Plugging in the values of Z and W from equation 8, we get:

$$(A - BD^{-1}C)^{-1}BD^{-1} = A^{-1}B(D - CA^{-1}B)^{-1} \tag{12}$$

Post-multiplying both sides by CA^{-1}, we get

$$(A - BD^{-1}C)^{-1}BD^{-1}CA^{-1} = A^{-1}B(D - CA^{-1}B)^{-1}CA^{-1} \tag{13}$$

Already, we see that the right side of equation 13 is identical that the second term on the right side of equation 11. Similarly, the left side of equation 13 is identical to the left side of equation 11 other than than the fact that there is a $BD^{-1}CA^{-1}$ attached. Thus, we shall try to remove $BD^{-1}CA^{-1}$ from the left side of equation 13.

Notice that

$$(A - BD^{-1}C)(-A^{-1}) = BD^{-1}CA^{-1} - I \tag{14}$$

so

$$(A - BD^{-1}C)(-A^{-1}) + I = BD^{-1}CA^{-1} \tag{15}$$

Plugging this into the $BD^{-1}CA^{-1}$ term of equation 13 gives us

$$A^{-1} + A^{-1}B(D - CA^{-1}B)^{-1}CA^{-1} = (A - BD^{-1}C)^{-1} \tag{16}$$

This equation is in fact already Woolbury Formula. All we need to do now is plug in E for A, $-B$ for F, D^{-1} for G, and C for H to achieve the same form as equation 11. □