Global Journal of Pure and Applied Mathematics.
ISSN 0973-1768 Volume 12, Number 4 (2016), pp. 3401–3405

http://www.ripublication.com/gjpam.htm

A Note on Heuristic Analog of Cline's Formula

Anatoliy Shmyrin, Nikolay Mishachev and Eugene Trofimov

Department of Mathematics,
Lipetsk State Technical University,
Lipetsk, 398600, Russia.

Abstract

In this note, we consider the problem of block matrix pseudo-inversion. We propose a new heuristic formula for pseudo-inverse of partitioned matrix. For matrices of full column rank we give a justification and discuss the time complexity of the new formula. In this case the new formula is much simpler then well-known Cline's formula and in some issues it has an advantage in comparison with the usual formula $C^+ = (C^T C)^{-1} C^T$ for pseudo-inverse of full rank matrix.

AMS subject classification:
Keywords: Pseudo-inverse, partitioned matrix, Cline's formula.

1. Introduction

The problems of identification and control in the theory of linear neighborhood systems [1] often leads to the matrix equation of the form $Cz = Ax + By = D$, where the unknown vector z is a concatenation of vectors x and y. In this context, x and y can be the state and the control vectors (when we are dealing with control) or unknown coefficients on the state and control variables (when we are dealing with identification). If for the matrix $C = [A|B]$ the *pseudo-inverse* (or *Moore-Penrose inverse* [2]) $C = \begin{bmatrix} \widetilde{A} \\ \widetilde{B} \end{bmatrix}$ is known then the *pseudo-solution* to the system, i.e. the least-squares solution of smallest norm, can be written as

$$z = \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \widetilde{A}D \\ \widetilde{B}D \end{bmatrix}.$$

The Cline's formula [3] for block matrix pseudo-inverse

$$z = \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} C_A^+ + P_A K_A A^T (B^+)^T B^+ (I - AC_A^+) \\ C_B^+ + P_B K_B B^T (A^+)^T A^+ (I - AC_B^+) \end{bmatrix},$$

where

$$C_B = (I - AA^+)B, C_A = (I - BB^+)A,$$
$$P_B = I - C_B^+ C_B, P_A = I - C_A^+ C_A,$$
$$K_B = [I + P_B B^T (A^+)^T A^+ B P_B]^{-1}, K_A = [I + P_A A^T (B^+)^T B^+ A P_A]^{-1},$$

allows us, in particular, to find the matrix $\widetilde{A}$ (or $\widetilde{B}$) and, therefore, to find the unknown vector x (or y) without calculating the entire matrix C^+. In case when we are interested only in x (or y) it gives a hope for reducing the volume of computation. However, for the matrix $C = [A|B]$ of full column rank the time complexity of the Cline's formula is much greater than that of the usual formula $C^+ = (C^T C)^{-1} C^T = [(C^T C)^{-1} A^T, (C^T C)^{-1} B^T]$. In what follows, we suggest a heuristic computation for block matrix $C = [A|B]$ pseudo-inverse and then justify the obtained result for matrix of full column rank. Unlike of the Cline's formula, the new formula contains much less matrix multiplications and its time complexity is comparable with that of the formula $C^+ = (C^T C)^{-1} C$.

2. Heuristic pseudo-inverse

Consider the matrix equation

$$C_{m\times(n+k)} z = A_{m\times n} x_n + B_{m\times k} y_k = D_m \tag{1}$$

Let x_n and y_k are the pseudo-solutions of the equations

$$A_{m\times n} x_n = D_m - B_{m\times k} y_k \quad \text{and} \quad B_{m\times k} y_k = D_m - A_{m\times n} x_n \tag{2}$$

i.e.

$$x_n = A^+_{n\times m} D_m - A^+_{n\times m} B_{m\times k} y_k \tag{3}$$

and

$$y_k = B^+_{k\times m} D_m - B^+_{k\times m} A_{m\times n} x_n \tag{4}$$

From (3) and (4) we get

$$x_n = A^+_{n\times m} D_M - A^+_{n\times m} B_{m\times k} (B^+_{k\times m} D_m - B^+_{k\times m} A_{m\times n} x_n) \tag{5}$$

and hence

$$(I_{n\times n} - A^+_{n\times m} B_{m\times k} B^+_{k\times m} A_{m\times n}) x_n = (A^+_{n\times m} - A^+_{n\times m} B_{m\times k} B^+_{k\times m}) D_m \tag{6}$$

Now let x_n be the pseudo-solution of the matrix equation (6):

$$x_n = (I_{n\times n} - A^+_{n\times m} B_{m\times k} B^+_{k\times m} A_{m\times n})^+ (A^+_{n\times m} - A^+_{n\times m} B_{m\times k} B^+_{k\times m}) D_m \tag{7}$$

or, in short notation, $x_n = A^\tau_{n\times m} D_m$. Substituting this expression for x_n in (4), we get

$$y_k = B^+_{k\times m} (I_{m\times m} - A_{m\times n} A^\tau_{n\times m}) D_m \tag{8}$$

Therefore, the matrix

$$C^{\tau} = \begin{bmatrix} (I_{n\times n} - A^{+}_{n\times m} B_{m\times k} B^{+}_{k\times m} A_{m\times n})^{+} (A^{+}_{n\times m} - A^{+}_{n\times m} B_{m\times k} B^{+}_{k\times m}) \\ B^{+}_{k\times m} (I_{m\times m} - A_{m\times n} A^{\tau}_{n\times m}) \end{bmatrix}$$

or, in short notation,

$$C^{\tau} = \begin{bmatrix} A^{\tau}_{n\times m} \\ B^{\tau}_{k\times m} \end{bmatrix} \tag{9}$$

is a candidate for the pseudo-inverse C^{+}. Of course, our formal calculation leaves open the question, is the found vector $\begin{bmatrix} x_n \\ y_k \end{bmatrix}$ the genuine pseudo-solution of source system (1) and is the matrix C^{τ} the genuine pseudo-inverse for $C = [A|B]$.

3. Justification of the heuristic formula for the matrices of full column rank

In this case

$$A^{+}_{n\times m} A_{m\times n} = I_{n\times n} \quad \text{and} \quad B^{+}_{k\times m} B_{m\times k} = I_{k\times k}.$$

Besides, in this situation the matrix $(I_{n\times n} - A^{+}_{n\times m} B_{m\times k} B^{+}_{k\times m} A_{m\times n})$ is invertible. Indeed, the equality $A^{+}_{n\times m} B_{m\times k} B^{+}_{k\times m} A_{m\times n} z_n = z_n$ implies

$$A_{m\times n} A^{+}_{n\times m} B_{m\times k} B^{+}_{k\times m} (A_{m\times n} z_n) = (A_{m\times n} z_n). \tag{10}$$

On the other hand, $A_{m\times n} A^{+}_{n\times m}$ and $B_{m\times k} B^{+}_{k\times m}$ are projectors on linearly independent subspaces and hence (10) implies $A_{m\times n} z_n = 0$ and thus $z_n = 0$. Therefore, we can write down further the inverse

$$(I_{n\times n} - A^{+}_{n\times m} B_{m\times k} B^{+}_{k\times m} A_{m\times n})^{-1}$$

instead of the pseudo-inverse

$$(I_{n\times n} - A^{+}_{n\times m} B_{m\times k} B^{+}_{k\times m} A_{m\times n})^{+}.$$

Consider the product of hypothetical pseudo-inverse (9) and the source matrix:

$$\begin{bmatrix} A^{\tau}_{n\times m} \\ B^{\tau}_{k\times m} \end{bmatrix} \left[A_{m\times n} | B_{m\times k} \right] = \begin{bmatrix} A^{\tau}_{n\times m} A_{m\times n} & A^{\tau}_{n\times m} B_{m\times k} \\ B^{\tau}_{k\times m} A_{m\times n} & B^{\tau}_{k\times m} B_{m\times k} \end{bmatrix} = \begin{bmatrix} I & III \\ IV & II \end{bmatrix}.$$

We compute separately each of the blocks.

Block I:

$$\begin{aligned} &(I_{n\times n} - A^{+}_{n\times m} B_{m\times k} B^{+}_{k\times m} A_{m\times n})^{+} (I_{n\times n} - A^{+}_{n\times m} B_{m\times k} B^{+}_{k\times m} A_{m\times n}) A_{m\times n} \\ &\quad = (I_{n\times n} - A^{+}_{n\times m} B_{m\times k} B^{+}_{k\times m} A_{m\times n})^{-1} (I_{n\times n} - A^{+}_{n\times m} B_{m\times k} B^{+}_{k\times m} A_{m\times n}) = I_{n\times n} \end{aligned}$$

Block II:

$$B^{+}_{k\times m}(I_{m\times m}-A_{m\times n}A^{\tau}_{n\times m})B_{m\times k}$$
$$=B^{+}_{k\times m}[I_{m\times m}-A_{m\times n}(I_{n\times n}-A^{+}_{n\times m}B_{m\times k}B^{+}_{k\times m}A_{m\times n})^{+}$$
$$A^{+}_{n\times m}(I_{m\times m}-B_{m\times k}B^{+}_{k\times m})]B_{m\times k}$$
$$=B^{+}_{k\times m}B_{m\times k}-B^{+}_{k\times m}A_{m\times n}(I_{n\times n}-A^{+}_{n\times m}B_{m\times k}B^{+}_{k\times m}A_{m\times n})^{+}$$
$$(A^{+}_{n\times m}B_{m\times k}-A^{+}_{n\times m}B_{m\times k}B^{+}_{k\times m}B_{m\times k})$$
$$=I_{k\times k}-B^{+}_{k\times m}A_{m\times n}(I_{n\times n}-A^{+}_{n\times m}B_{m\times k}B^{+}_{k\times m}A_{m\times n})^{+}$$
$$(A^{+}_{n\times m}B_{m\times k}-A^{+}_{n\times m}B_{m\times k})=I_{k\times k}$$

Block III:

$$(I_{n\times n}-A^{+}_{n\times m}B_{m\times k}B^{+}_{k\times m}A_{m\times n})^{+}(A^{+}_{n\times m}-A^{+}_{n\times m}B_{m\times k}B^{+}_{k\times m})B_{m\times k}$$
$$=(I_{n\times n}-A^{+}_{n\times m}B_{m\times k}B^{+}_{k\times m}A_{m\times n})^{+}(A^{+}_{n\times m}B_{m\times k}-A^{+}_{n\times m}B_{m\times k}B^{+}_{k\times m}B_{m\times k})$$
$$=(I_{n\times n}-A^{+}_{n\times m}B_{m\times k}B^{+}_{k\times m}A_{m\times n})^{+}(A^{+}_{n\times m}B_{m\times k}-A^{+}_{n\times m}B_{m\times k})=0.$$

Block IV:

$$B^{+}_{k\times m}[I_{m\times m}-A_{m\times n}(I_{n\times n}-A^{+}_{n\times m}B_{m\times k}B^{+}_{k\times m}A_{m\times n})^{+}$$
$$A^{+}_{n\times m}(I_{m\times m}-B_{m\times k}B^{+}_{k\times m})]A_{m\times n}$$
$$=B^{+}_{k\times m}[A_{m\times n}-A_{m\times n}(I_{n\times n}-A^{+}_{n\times m}B_{m\times k}B^{+}_{k\times m}A_{m\times n})^{-1}$$
$$(A^{+}_{n\times m}A_{m\times n}-A^{+}_{n\times m}B_{m\times k}B^{+}_{k\times m}A_{m\times n})]$$
$$=B^{+}_{k\times m}(A_{m\times n}-A_{m\times n})=0.$$

Thus, $C^{\tau}_{(n+k)\times m}C_{m\times(n+k)}=I_{(n+k)\times(n+k)}$. Hence, the first three Moor-Penrose equations [2] are satisfied. It remains to verify the fourth equation $(CC^{\tau})^{T}=CC^{\tau}$. Note that

$$CC^{\tau}=AA^{\tau}+BB^{\tau}=AA^{\tau}+BB^{+}(I-AA^{\tau})=BB^{+}+(I-BB^{+})AA^{\tau}$$
$$=BB^{+}+(I-BB^{+})A[I-A^{+}BB^{+}A]^{-1}A^{+}(I-BB^{+}),$$
$$A^{+}=(A^{T}A)^{-1}A^{T},\ A=(A^{+})^{T}[A^{+}(A^{+})^{T}]^{-1},$$

therefore

$$(CC^{\tau})^{T}=(BB^{+})^{T}+(I-BB^{+})^{T}(A^{+})^{T}[I-A^{T}(BB^{+})^{T}(A^{+})^{T}]^{-1}A^{T}(I-BB^{+})^{T}$$
$$=BB^{+}+(I-BB^{+})A(A^{T}A)^{-1}[I-A^{T}BB^{+}(A^{+})^{T}]^{-1}[A^{+}(A^{+})^{T}]^{-1}A^{+}(I-BB^{+})$$
$$=BB^{+}+(I-BB^{+})A[(A^{+}(A^{+})^{T})(I-A^{T}BB^{+}(A^{+})^{T})(A^{T}A)]^{-1}A^{+}(I-BB^{+})$$
$$=BB^{+}+(I-BB^{+})A[A^{+}(A^{+})^{T}A^{T}A-A^{+}(A^{+})^{T}A^{T}BB^{+}(A^{+})^{T}A^{T}A]^{-1}A^{+}(I-BB^{+})$$
$$=BB^{+}+(I-BB^{+})A[A^{+}(AA^{+})^{T}A-A^{+}(AA^{+})^{T}BB^{+}(AA^{+})^{T}A]^{-1}A^{+}(I-BB^{+})$$
$$=BB^{+}+(I-BB^{+})A[A^{+}(AA^{+})A-A^{+}(AA^{+})BB^{+}(AA^{+})A]^{-1}A^{+}(I-BB^{+})$$
$$=BB^{+}+(I-BB^{+})A[A^{+}A-A^{+}BB^{+}A]^{-1}A^{+}(I-BB^{+})$$
$$=BB^{+}(I-BB^{+})A[I-A^{+}BB^{+}A]^{-1}A^{+}(I-BB^{+})=CC^{\tau}$$

Remark 3.1. For (non square) matrices of full row rank the numerical experiments with the formula (9) have shown that in all cases the first three Moore-Penrose equations are satisfied and the fourth failed.

4. Evaluation of the time complexity

As we already mentioned, the formula (9) requires considerably less computation than the Cline's formula. However, for matrices of full column rang we need to compare (9) with the formula

$$C^+ = (C^T C)^{-1} C^T \tag{11}$$

instead of the Clineâ's formula. Let, for example, $n = k \approx m/2$. The formula (11) contains the inversion and two matrix products, therefore the time complexity is approximately

$$(4/3)(2n)^3 + 2(2n)^3 = (80/3)n^3$$

(we are counting only the multiplications). The calculation of matrices A^+ and B^+ in (9) by the formula (11) requires $2\,[(4/3)n^3 + 2n^3 + 2n^3] = (32/3)n^3$ operations.The inversion $(I_{n\times n} - A^+_{n\times m} B_{m\times k} B^+_{k\times m} A_{m\times n})^{-1}$ requires $(4/3)n^3$ operations. Besides, we need in (9) four matrix products for the upper block ($8n^3$ operations) and two for the lower ($4n^3$ operations). So, we need in (9)

$$(32/3)n^3 + (4/3)n^3 + 12n^3 = (72/3)n^3$$

operations, i.e. approximately as much as in the formula (11). If we need to find only the upper block of the pseudo-inverse, the time complexity of (9) and (11) is reduced to $(68/3)n^3$ and $(60/3)n^3$. Besides, in cases when the pseudo-inverse A^+ and B^+ are already known due a specific situation, the time complexity of (9) is $(40/3)n^3$.

Acknowledgments

The work is supported by the Russian Fund for Basic Research (project ¹ 16-07-00854 a).

References

[1] S.L. Blyumin, A.M. Shmyrin, *Neighborhood systems*, Lipetsk: LEGI (2005).

[2] A. Ben-Israel, T.N.E. Greville, *Generalized inverses: theory and applications*, Springer-Verlag, New York (2003).

[3] R.E. Cline, *Representation for the generalized inverse of a partitioned matrix*, J. Soc. Industr. Appl. Math., 12 (1964), 588–600.