THE MATRIX EQUATION $AX-YB=C$ AND RELATED PROBLEMS

Yu. A. Al'pin* and S. N. Il'in*

UDC 519.61

The main result of the paper is a theorem, using which a new proof of Roth's theorem is obtained, a new solvability criterion for the matrix equation $AX-YB=C$ is proved, a formula for a particular solution of the latter is derived, and the least of the orders of square nonsingular matrices containing a given rectangular matrix as a submatrix is determined. Bibliography: 5 titles.

1. INTRODUCTION AND AUXILIARY RESULTS

In 1952, Roth published the following theorem [1]: the matrix equation $AX-YB=C$ has a solution if and only if the matrices $\begin{pmatrix} A & C \\ 0 & B \end{pmatrix}$ and $\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$ are equivalent. Another proof of Roth’s theorem was given in [2] and reproduced in [3].

The main result of the present paper is Theorem 1. Based on this theorem, a new proof of Roth’s theorem is suggested, a new solvability criterion for the equation $AX-YB=C$ is found, a formula for a particular solution of the latter is derived, and the least order of a square nonsingular matrix that contains a given rectangular matrix as a submatrix is determined.

Our considerations are in terms of skeleton decompositions of matrices and pseudoinverse matrices. Recall the necessary facts related to these objects.

Let $M$ be a matrix of rank $n > 0$. A skeleton decomposition of the matrix $M$ is its arbitrary representation in the form of a product $M = LR$, where $n$ is the number of columns in $L$ and the number of rows in $R$. The ranks of the factors of a skeleton decomposition both necessarily equal $n = \text{rg} M$. We will use the following simple properties of skeleton decompositions.

**Proposition 1.** A matrix $L$ (R) is a left (right) factor of a skeleton decomposition of a matrix $M$ if and only if the columns of $L$ (the rows of $R$) are linearly independent, and their span coincides with the span of the columns (rows) of the matrix $M$. If the left (right) factor of a skeleton decomposition is known, then the right (left) factor is the unique solution of the matrix equation $LZ = M$ ($ZR = M$).

**Proof.** The validity of this assertion stems from the definition of a skeleton decomposition and the standard theory of simultaneous linear equations.

**Proposition 2.** Let a decomposition $M = LR$ whose factors satisfy the condition $\text{rg} L = \text{rg} R = n$ be given, where $n$ is the number of columns in $L$ and the number of rows in $R$. Then $\text{rg} M = n$, i.e., this decomposition is a skeleton decomposition of the matrix $M$.

**Proof.** In order to prove this assertion, we write the Sylvester inequality for the product $LR$ (see, e.g., [4]):

$$\text{rg} L + \text{rg} R - n \leq \text{rg} LR \leq \min(\text{rg} L, \text{rg} R).$$

On substituting the ranks of $L$ and $R$ into this inequality, we immediately obtain $n \leq \text{rg} LR \leq n$, i.e., $\text{rg} LR = n$.

The propositions below provide some properties of pseudoinverse matrices used in what follows.

**Proposition 3** (see, e.g., [3]). If the matrix equation $AX = C$ is solvable, then the matrix $A^+C$ is a solution. Similarly, if the equation $YB = C$ is solvable, then the matrix $CB^+$ is a solution.

**Proposition 4.** If $M = LR$ is a skeleton decomposition, then

$$M^+ = R^+L^+, \quad L = MR^+, \quad R = L^+M, \quad LL^+ = MM^+, \quad R^+R = M^+M. \quad (1)$$

**Proof.** Equality (1) can be found in [4], whereas (2) and (3) can readily be derived from elementary properties of pseudoinverses presented in that monograph.

*Kazan’ State University, Kazan’, Russia, e-mail: Yuri.Alpin@ksu.ru, Sergey.Ilyin@ksu.ru.

The main theorem

**Theorem 1.** Let a block matrix \( M = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \) satisfy the conditions

\[
\operatorname{rg}(A, B) = \operatorname{rg} A + \operatorname{rg} B, \tag{4}
\]

\[
\operatorname{rg} \begin{pmatrix} A \\ C \end{pmatrix} = \operatorname{rg} A + \operatorname{rg} C. \tag{5}
\]

Then the following assertions are equivalent:

- \( \operatorname{rg} M = \operatorname{rg} A + \operatorname{rg} B + \operatorname{rg} C \);
- the equation \( CX + YB = D \) has a solution.

**Proof.** The proof consists of three parts.

1. First consider the simplest case where \( B \) or \( C \) is a zero matrix. If the submatrix \( B \) is zero, then, with regard to (5), the assertion of the theorem takes the form

\[
\operatorname{rg} \begin{pmatrix} A \\ C \end{pmatrix} = \operatorname{rg} A + \operatorname{rg} C \iff \text{the equation } CX = D \text{ has a solution}
\]

and is obviously valid. Similarly, if the submatrix \( C \) is zero, then, with account for (4), the assertion of the theorem amounts to the equivalence

\[
\operatorname{rg} \begin{pmatrix} A & B \\ 0 & D \end{pmatrix} = \operatorname{rg}(A, B) \iff \text{the equation } YB = D \text{ has a solution},
\]

which is trivial.

2. Now let the submatrices \( B \) and \( C \) be nonzero and let \( A \) be zero. Fix arbitrary skeleton decompositions

\[
B = B_1B_2, \quad C = C_1C_2 \tag{6}
\]

and consider the matrices

\[
M = \begin{pmatrix} 0 & B \\ C & D \end{pmatrix}, \quad M_1 = \begin{pmatrix} 0 & B_1 \\ C_1 & DB_2^+ \end{pmatrix}.
\]

As is readily seen, \( \operatorname{rg} M_1 = \operatorname{rg} B_1 + \operatorname{rg} C_1 = \operatorname{rg} B + \operatorname{rg} C \). Since \( \begin{pmatrix} B_1 \\ DB_2^+ \end{pmatrix} = \begin{pmatrix} B \\ D \end{pmatrix} B_2^+ \), it is clear that the span of the columns of \( M_1 \) is contained in the span of the columns of \( M \).

Assume that the spans of the columns of \( M \) and \( M_1 \) coincide, i.e., the equality \( \operatorname{rg} M = \operatorname{rg} B + \operatorname{rg} C \) is valid. Then, by Proposition 1, the matrix \( M_1 \) is the left factor of a skeleton decomposition of the matrix \( M \), and the equation \( M_1Z = M \) has a unique solution. Partition the matrix \( Z \) into blocks in such a way that the blockwise matrix multiplication is feasible. Then the equation takes the form

\[
\begin{pmatrix} 0 & B_1 \\ C_1 & DB_2^+ \end{pmatrix} \begin{pmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{pmatrix} = \begin{pmatrix} 0 & B \\ C & D \end{pmatrix} \tag{7}
\]

This matrix equation amounts to the following four simultaneous equations:

\[
\begin{align*}
B_1Z_{21} &= 0, \\
B_1Z_{22} &= B, \\
C_1Z_{11} + DB_2^+ Z_{21} &= C, \\
C_1Z_{12} + DB_2^+ Z_{22} &= D.
\end{align*}
\]

This system is solved as follows. Since the columns of \( B_1 \) are linearly independent, we have \( Z_{21} = 0 \). Taking into account equalities (6) and Proposition 1, we see that \( Z_{22} = B_2 \) is a solution of the second equation, and