MCIT Modeling, control & information technologies

On solvability of the matrix equation AX+YB=C over principal ideal domains

https://doi.org/10.31713/MCIT.2024.081

Volodymyr Prokip
IAPMM NAS of Ukraine,
v.prokip@gmail.com

Abstract **– In this communication we present conditions for solvability of the matrix equation AX + YB = C over the ring of integers $\mathbb{Z}$. The necessary and sufficient conditions for the solvability of this equation in terms of the Smith normal forms of the matrices A and B are given. The conditions under which this equation has a minimal solution are also given.**

Keywords **– ring of integers; field; Smith normal form; matrix equation; solution**

I. Introduction

Let $\mathbb{Z}$ be the ring of integers. We denote by (a,b) the greatest common divisor of nonzero elements $a,b \in \mathbb{Z}$. Further, let $\mathbb{Z}_{m,n}$ be the set of $m\times n$ matrices over $\mathbb{Z}$. Denote by I_n the identity matrix of dimension n and by $0_{m,n}$ the zero $m\times n$ matrix. We will denote by $GL(m,\mathbb{Z})$ the set of invertible matrices in $\mathbb{Z}_{m,m}$.

Let $D\in\mathbb{Z}_{m,n}$ and $rank\, D = r$. For matrix D there exist matrices $U,V\in GL(m,\mathbb{Z})$ such that

$$UDV = S_D = diag(d_1, d_2, \ldots, d_r, 0, \ldots, 0)$$

is the Smith normal form of matrix D, i.e. $d_i \in \mathbb{Z}$ and $d_i \mid d_{i+1}$ (divides) for all $i = 1,2,\ldots,r-1$. The matrix S_A we write in the form $S_A = diag\,(S(A),0,\ldots,0)$, where

$$S(A) = diag(a_1, a_2, \ldots, a_r).$$

Consider the matrix equation

$$AX + YB = C, \qquad (1)$$

where $A\in\mathbb{Z}_{m,m}$, $B\in\mathbb{Z}_{n,n}$, $C\in\mathbb{Z}_{m,n}$, X and Y are unknown $m\times n$ matrices over $\mathbb{Z}$.

The main problem of studying the linear matrix equation (1) is to give conditions for its solvability. In addition, it is also necessary to know the structure of a general solution of this equation. The equation (1) is one of the best known matrix equations in matrix theory and its applications. The problem of solvability of equation (1) has drawn the attention of many mathematicians.

Roth [16] showed that equation (1) is consistent over a field $\mathbb{F}$ if and only if matrices

$$M_C = \begin{bmatrix} A & C \\ 0_{n,m} & B \end{bmatrix} \text{ and } M_0 = \begin{bmatrix} A & 0_{m,n} \\ 0_{n,m} & B \end{bmatrix}$$

are equivalent. Roth's condition is equivalent to the following statement: Equation (1) is consistent over a field $\mathbb{F}$ if and only if

$$rank\, M_C = rank\, A + rank\, B.$$

Through the generalized inverses of matrices, Baksalary and Kala [1] showed that equation (1) over a field is consistent if and only if the following condition holds: $(I_m - AA^-)C(I_m - B^-B) = 0_{m,n}$ and gave the general solutions of (1), where A^- and B^- are generalized inverse of matrices A and B respectively. Many authors addressed the question when the equation (1) (over the real numbers R, the complex number C, the quaternion skew field H or commutative rings has a solution belonging to a special class of matrices (see [4], [6–13], [19] and references therein).

Let R be a commutative ring with identity. In [7] was prove the following statement: The matrix equation $AX+YB=C$ is solvable over R if and only if matrices M_C and M_0 are equivalent (see also [8]).

If matrices $A\in\mathbb{Z}_{m,m}$ and $B\in\mathbb{Z}_{n,n}$ are nonsingular with respectively prime determinants then equation (1) is solvable over $\mathbb{Z}$ for arbitrary matrix $C\in\mathbb{Z}_{m,n}$. In this report we give necessary and sufficient conditions of solvability of matrix equation (1) over $\mathbb{Z}$ in terms of the Smith normal forms of matrices A and B.

II. Main results

Theorem 1. *Let $A\in\mathbb{Z}_{m,m}$, $B\in\mathbb{Z}_{n,n}$, $C\in\mathbb{Z}_{m,n}$ and $rank\, A = p$, $rank\, B = q$. Further, let $U_1, V_1 \in GL(m,\mathbb{Z})$ and $U_2, V_2 \in GL(n,\mathbb{Z})$ such that*

$$U_1AV_1 = S_A = diag(a_1, a_2, \ldots, a_p, 0, \ldots, 0)$$

and

$$U_2BV_2 = S_B = diag(b_1, b_2, \ldots, b_q, 0, \ldots, 0).$$

Matrix equation $AX+YB=C$ is solvable over $\mathbb{Z}$ if and only if the following conditions are held

a) $U_1CV_2 = \left[\begin{array}{ccc|c} f_{11} & \cdots & f_{1q} & \\ \vdots & \ddots & \vdots & P \\ f_{p1} & \cdots & f_{pq} & \\ \hline & Q & & 0_{m-p,n-q} \end{array}\right]$,

where $f_{ij} \in \mathbb{Z}$, $P \in \mathbb{Z}_{p,n-p}$ *and* $Q \in \mathbb{Z}_{m-p,q}$;

b) $(a_i, b_j) \mid f_{ij}$ *for all* $i = 1,2,\ldots,p$; $j = 1,2,\ldots,q$ *and* $P = diag(a_1,a_2,\ldots,a_r)P_1$, $Q = Q_1 diag(b_1,b_2,\ldots,b_q)$.

***Proof*.** Let $m \times n$ matrices X_0 and Y_0 over $\mathbb{Z}$ be a solution of equation (1). Further, let $U_1, V_1 \in GL(m,\mathbb{Z})$ and $U_2, V_2 \in GL(n,\mathbb{Z})$ such that

$$U_1AV_1 = S_A = diag\,(S(A),0,\ldots,0)$$

and

$$U_2BV_2 = S_B diag\,(S(B),0,\ldots,0).$$

From equality $AX_0 + Y_0B = C$ we have

$$S_A\tilde{X}_0 + \tilde{Y}_0S_B = C_0, \quad (2)$$

Put

$$X_0 = \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix},\ Y_0 = \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix},\ C_0 = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix},$$

where

$$\tilde{X}_0 = V_1^{-1}X_0V_2 = \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix},\ X_{11} \in \mathbb{Z}_{p,q}\ \ X_{12} \in \mathbb{Z}_{p,n-p},$$

$$\tilde{Y}_0 = U_1Y_0U_2^{-1} = \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix},\ Y_{11} \in \mathbb{Z}_{p,q},\ Y_{21} \in \mathbb{Z}_{m-p,q},$$

$$C_0 = U_1CV_2 = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix},\ C_{11} \in \mathbb{Z}_{p,q},\ C_{12} \in \mathbb{Z}_{m-q,q},$$

$C_{21} \in \mathbb{Z}_{n-p,q}$.

From equality (2) follows

$$\begin{cases} S(A)X_{11} + Y_{11}S(B) &= F, \\ S(A)X_{12} &= P, \\ Y_{21}S(B) &= Q, \\ 0_{m-p,n-q} &= C_{22}. \end{cases}$$

Thus, from the first equality of this system of equalities, we obtain $(a_i, b_j) \mid f_{ij}$ for all $1 \le i \le p$ and $1 \le j \le q$.

Conversely, suppose $(a_i, b_j) \mid f_{ij}$ for all $1 \le i \le p$ and $1 \le j \le q$. Thus, there exist elements $p_{ik}, q_{kj} \in \mathbb{Z}$ such that $a_i p_{ik} + b_j q_{kj} = f_{ij}$. Hence, there exist matrices $P_{11} \in \mathbb{Z}_{p,q}$, $Q_{11} \in \mathbb{Z}_{p,q}$ such that $S(A)P_{11} + Q_{11}S(B) = F$. Considering equalities $S(A)X_{12} = P$ and $Y_{21}S(B) = Q$, we obtain $X_{12} = P_{12} \in \mathbb{Z}_{p,n-p}$ and $Y_{21} = Q_{21} \in \mathbb{Z}_{m-p,q}$. So, for any matrices $P_{22} \in \mathbb{Z}_{m-p,n-q}$ and $Q_{22} \in \mathbb{Z}_{m-p,n-q}$ the following equality holds

$$S_A\begin{bmatrix} P_{11} & P_{12} \\ 0 & P_{22} \end{bmatrix} + \begin{bmatrix} Q_{11} & 0 \\ Q_{21} & Q_{22} \end{bmatrix}S_B = \begin{bmatrix} P_{21} & P_{21} \\ P_{21} & 0 \end{bmatrix}.$$

From this equality it follows that equation (2) is solvable.

It is obvious that the pair of matrices $X_P = V^{-1}\begin{bmatrix} P_{11} & P_{12} \\ 0 & P_{22} \end{bmatrix}V_2$ and $Y_Q = U_1^{-1}\begin{bmatrix} Q_{11} & 0 \\ Q_{21} & Q_{22} \end{bmatrix}U_2^{-1}$ is a general solution of equation (1). The proof of Theorem 1 is complete.

From the proof of the sufficiency of Theorem 1, we obtain a method for constructing the general solution of equation (1).

The ring $\mathbb{Z}$ is an Euclidian domain. In the ring $\mathbb{Z}$ is defined a nonnegative integral valued function so satisfying the following conditions:

a) $\varphi(a) = 0$ *if* $a = 0$ *is the zero of* $\mathbb{Z}$;

b) $\varphi(a) = \varphi(b)$ if a divides b;

c) for any pair of elements $a, b \in \mathbb{Z}$ there exist elements $q, r \in \mathbb{Z}$ such that $a = bq + r$ and $\varphi(r) < \varphi(b)$.

In general, the existence of q and r is not assumed to be unique. We defined the integral function of the Euclidean norm in the domain $\mathbb{Z}$ as follows $\varphi(b) = |b|$ and $\varphi(r) \ge 0$, i.e. $\varphi(r) \in \{0,1,\ldots,|b|-1\}$. Thus, for arbitrary elements $a, b \in E$ there exist unique elements $q, r \in \mathbb{Z}$ such that $a = bq + r$ and $\varphi(r) < \varphi(b)$. We note that this property of the Euclidean norm φ holds for the ring of polynomials $F[\lambda]$ over a field F.

Consider the equation $ax + by = c$, where $a, b, c \in \mathbb{Z}$ and x, y are unknown elements in $\mathbb{Z}$. We say that a pair of elements $x_0, y_0 \in \mathbb{Z}$ is a minimal solution (with respect to y) of the equation $ax + by = c$ if

$$\varphi(y_0) < \varphi(a).$$

***Proposition 1*.** *Let* $a, b, c \in \mathbb{Z}$ *be nonzero elements. The equation* $ax + by = c$ *has a unique "minimal" solution* x_0, y_0 *over* $\mathbb{Z}$ *such that* $\varphi(y_0) < \varphi(a)$ *if and only if elements* a *and* b *are relatively prime, i.e.* $(a,b) = e$.

Denote by $\mathbb{Z}_+$ the set of nonnegative integers of the domain $\mathbb{Z}$. For nonsingular matrix $A \in \mathbb{Z}_{m,m}$ there exists a matrix $W \in GL(m,\mathbb{Z})$ such that

$$AW = H_A = \begin{bmatrix} h_1 & 0 & \ldots & \ldots & 0 \\ h_{21} & h_2 & 0 & \ldots & 0 \\ \ldots & & \ldots & & \ldots \\ h_{m1} & h_{m2} & \ldots & h_{m,m-1} & h_m \end{bmatrix}$$

is a lower triangular matrix in which $h_i \in \mathbb{Z}_+$ for all $1 \le i \le m$ 1 and $\varphi(h_{ij}) < \varphi(h_i)$ for all $1 \le j < i \le m$. The matrix H_A is called the Hermitian normal form of the matrix A. Taking into account Proposition 1 and article [15], we get the following statement.

***Theorem 2*.** *Let* $A \in \mathbb{Z}_{m,m}$ *and* $B \in \mathbb{Z}_{n,n}$ *be nonsingular matrices and* $C \in \mathbb{Z}_{m,n}$. *Further, let*

$$H_A = AW = \begin{bmatrix} h_1 & 0 & \dots & \dots & 0 \\ h_{21} & h_2 & 0 & \dots & 0 \\ \dots & & \dots & & \dots \\ h_{m1} & h_{m2} & \dots & h_{m,m-1} & h_m \end{bmatrix}$$

be the Hermitian normal form of the matrix A, where $W \in GL(m,\mathbb{Z})$. *The matrix equation*

$$AX + YB = C$$

has a unique "minimal" solution $X_0 \in \mathbb{Z}_{n,m}$ and

$$Y_0 = \begin{Vmatrix} y_{11} & y_{12} & \cdots & y_{1n-1} & y_{1n} \\ \dots & & \dots & & \dots \\ y_{k1} & y_{k2} & \cdots & y_{k,n-1} & y_{kn} \\ \dots & & \dots & & \dots \\ y_{m1} & y_{m2} & \cdots & y_{m,n-1} & y_{mn} \end{Vmatrix} \in \mathbb{Z}_{m,n}$$

such that for all $\varphi(y_{ik}) < \varphi(h_i)$ *for all* $1 \le i \le m$ *and* $1 \le k \le n$ *if and only if* $(\det A, \det B) = 1$.

Let $F[\lambda]$ be the polynomials ring over a field F. We defined the integral function of the Euclidean norm in the domain $F[\lambda]$ as follows $\varphi(a(\lambda)) = \deg a(\lambda)$.

The polynomial Diophantine equations are used in the theory of control systems as a special mathematical object whose application makes it possible to efficiently solve a broad range of the problems of synthesis of systems, including control systems. The entire classical automata theory is, in fact, constructed on the basis of the Diophantine equations (both in the scalar form and in the matrix form). A significant part of available methods used for the synthesis of linear systems can be described as methods aimed at finding solutions of the matrix polynomial equation

$$A(\lambda)X(\lambda) + Y(\lambda)B(\lambda) = C(\lambda) \quad (3)$$

(and, in particular, of the "minimal" solutions). In this case, the search and construction of all classes of solutions of linear polynomial equations depends on the number of steps of divisibility of polynomials with remainders and can be obtained in the explicit form as a result of proper substitutions.

In his investigations of the structures of polynomial matrices over the field, Barnett [2] showed that regular matrices $A(\lambda) \in F_{m,m}[\lambda]$ and $B(\lambda) \in F_{n,n}[\lambda]$ have coprime determinants if and only if for any $m \times n$ matrix $C(\lambda)$ such that

$$\deg C(\lambda) \le \deg A(\lambda) + \deg B(\lambda)$$

matrix equation (3) possesses a unique minimal solution such that $\deg X_0(\lambda) < \deg B(\lambda)$ and $\deg Y_0(\lambda) < \deg A(\lambda)$.

The conditions under which the minimal solution of equation (3) exists were weakened in [5] and [14, 18]. Thus, in [5], it was proved that the Barnett condition for the existence of the unique minimal solution can be weakened as follows: the unique minimal solution of equation (3) exists in the case where only one of the matrices $A(\lambda)$ or $B(\lambda)$ is regular. More recent investigations showed that the conditions of existence of the minimal solution of equation (3) established in [5] can also be weakened. Indeed, the following assertion was proved in [14]: The unique minimal solution of equation (3) exists if at least one of irregular and nonsingular matrices $A(\lambda)$ and $B(\lambda)$ with coprime determinants can be regularized ($A(\lambda)$ from the left or $B(\lambda)$ from the right). Thus, the conditions of [5] cover a much broader class of equations of the form (3) for which it is possible to indicate the minimal solution. The methods used for the construction of solutions of equation (3) under certain restrictions were presented in [15, 16].

We defined the Euclidean norm in the domain $F[\lambda]$ as follows $\varphi(a(\lambda)) = \deg a(\lambda)$.

Proposition 2. *Let* $a(\lambda), b(\lambda), c(\lambda) \in F[\lambda]$ *be nonzero elements. The equation*

$$a(\lambda)x(\lambda) + b(\lambda)y(\lambda) = c(\lambda)$$

has a unique "minimal" solution $x_0(\lambda), y_0(\lambda)$ *over* $F[\lambda]$ *such that* $\deg y_0(\lambda) < \deg a(\lambda)$ *if and only if elements* $a(\lambda)$ *and* $b(\lambda)$ *are relatively prime, i.e.* $(a(\lambda), b(\lambda)) = 1$.

Using Proposition 2 and [16], we formulate a condition under which equation (6) has a "minimal" solution under certain restrictions.

Theorem 3. *Let* $A(\lambda) \in F_{m,m}[\lambda]$ *and* $B(\lambda) \in F_{n,n}[\lambda]$ *be nonsingular matrices and* $C(\lambda) \in F_{m,n}[\lambda]$. *Further, let*

$$H_A(\lambda) = A(\lambda)W(\lambda) = \begin{bmatrix} h_1(\lambda) & 0 & \dots & \dots & 0 \\ h_{21}(\lambda) & h_2(\lambda) & 0 & \dots & 0 \\ \dots & & \dots & & \dots \\ h_{m1}(\lambda) & h_{m2}(\lambda) & \dots & h_{m,m-1}(\lambda) & h_m(\lambda) \end{bmatrix}$$

be the Hermitian normal form of the matrix $A(\lambda)$, *where* $W(\lambda) \in GL(m, F[\lambda])$, $h_i(\lambda) \in F[\lambda]$ *are monic polynomials and* $\deg h_{ij}(\lambda) < \deg h_i(\lambda)$ *for all* $1 \le i \le m$ *and* $1 \le j < i \le m$. *The matrix equation*

$$A(\lambda)X(\lambda) + Y(\lambda)B(\lambda) = C(\lambda)$$

has a unique "minimal" solution $X_0(\lambda) \in F_{n,m}(\lambda)$ and

$$Y_0(\lambda) = \begin{Vmatrix} y_{11}(\lambda) & \dots & y_{1l}(\lambda) & \dots & y_{1n}(\lambda) \\ \dots & \dots & \dots & \dots & \dots \\ y_{k1}(\lambda) & \dots & y_{kl}(\lambda) & \dots & y_{kn}(\lambda) \\ \dots & \dots & \dots & \dots & \dots \\ y_{m1}(\lambda) & \dots & y_{ml}(\lambda) & \dots & y_{mn}(\lambda) \end{Vmatrix} \in F_{m,n}[\lambda]$$

such that for all $\deg y_{ik}(\lambda) < \deg h_i(\lambda)$ *for all* $1 \le i \le m$ *and* $1 \le k \le n$ *if and only if* $(\det A(\lambda), \det B(\lambda)) = 1$.

REFERENCES

[1] J.K. Baksalary, R. Kala, "The matrix equation *AX-YB=C*", Linear Algebra and Applications, 1979, Vol. **25**, pp. 41–43.

[2] S. Barnett, "Regular polynomial matrices having relatively prime determinants", Math. Proc. Camb. Phil. Soc., 1969, Vol. **65**, No.3, pp.585–590.

[3] S. Chen., Y. Tian, "On solutions of generalized Sylvester equation in polynomial matrices", J. Franklin Inst., 2014, Vol. **351**, No.12, pp.5376–5385.

[4] E. Emre, L.M. Silverman, "The equation $XR + QY = \Phi$: a characterization of solutions", SIAM J. Control Optim., 1981, Vol. **19**, No.1, pp.33–38.

[5] J. Feinstein, Y. Bar-Ness, "On the uniqueness minimal solution to the matrix polynomial equation $A(\lambda)X(\lambda)+Y(\lambda)B(\lambda) = C(\lambda)$, J. Franklin Inst., 1980, Vol. **310**, No.2, pp.131–134.

[6] H. Flanders, H.K. Wimmer, "On the matrix equations AX-*XB=C* and *AX-YB=C*", SIAM Journal on Applied Mathematics, 1977, Vol. **32**, № 4, pp.707–710.

[7] W.H. Gustafson, "Roth's theorems over commutative rings", Linear Algebra and its Applications, 1979, Vol. **23**, pp.245–251.

[8] R.E. Harwig, "Roth's equivalence problem in unit regular ring", Proc. Amer. Math. Society, 1976, Vol. **76**, № 1, pp.39–44.

[9] J. (Jr.) Jones, "A Diophantine matrix equation", Amer. Math. Monthly, 1955, Vol. **62**, No.4, pp. 244–247.

[10] Y.H. Liu, "Ranks of solutions of the linear matrix equation *AX + YB= C*", Computers & Mathematics with Applications, 2006, Vol.**52**, No.6-7, pp.861–872.

[11] Y. Ke, C. Ma, "An alternating direction method for nonnegative solutions of the matrix equation *AX+ YB= C* ", Computational and Applied Mathematics, 2017, Vol.***36***, pp.359–365.

[12] I. Kyrchei, "Cramer's Rules for Sylvester Quaternion Matrix Equation and Its Special Cases", Adv. Appl. Clifford Algebras, 2018, Vol. **28**, pp.1–26.

[13] V. Olshevsky, "Similarity of block diagonal and block triangular matrices", Integral equations and operator theory, 1992, Vol.**15**, № 5, pp. 853–863.

[14] V.M. Prokip, "About the uniqueness solution of the matrix polynomial equation $A(\lambda)X(\lambda) - Y(\lambda)B(\lambda) = C(\lambda)$ ", Lobachevskii J. Math., 2008, Vol.**29**, No.3, pp.186–191.

[15] V.M. Prokip, "On the divisibility of matrices with remainder over the domain of principal ideals", J. Math. Sci., 2019, Vol. **243**, No.1, pp.45–55.

[16] V.M. Prokip, O.M. Mel'nyk and R.V. Kolyada, "On divisibility with a remainder of polynomial matrices over an arbitrary field", J. Math. Sci., 2024. Vol. **250**, No.1-2, pp.24–39.

[17] W. E. Roth, "The equation *AX – YB = C* and *AX – XB = C* in matrices", Proc. Amer. Math. Society, 1952, Vol.**97**, pp.392–396.

[18] Y. Tian, C. Xia, "On the low-degree solution of the Sylvester matrix polynomial equation", Hindawi J. Math., 2021, Vol.**2021**, Article ID 4612177.

[19] K. Zietak. "Properties of the l_1-solutions of the Linear Matrix Equation *AX+ YB= C*", IMA journal of numerical analysis, 1987, Vol. **7**, No.2, pp.223–233.