On Solvability of the Matrix Equation $AX - XB = C$ over Integer Rings

https://doi.org/10.31713/MCIT.2019.04

Volodymyr Prokip
dept. of algebra
Pidstryhach Institute for Applied Problems of Mechanics and Mathematics NAS of Ukraine
L'viv, Ukraine
v.prokip@gmail.com

Abstract — **In this communication we present conditions of solvability of Sylvester matrix equation AX – XB = C over integer domains. The necessary and sufficient conditions of solvability of Sylvester equation in term of columns equivalence of matrices constructed in a certain way by using the coefficients of this equation are proposed.**

Keywords — matrix, Sylvester equation, integer domain.

I. Introduction

Let R denote an integer domain with identity $1 \neq 0$. Further, let $R_{m,n}$ be the set of $m \times n$ matrices over R. Denote by I_n the identity matrix of dimension n and $0_{m,n}$ the zero $m \times n$ matrix. For any $A \in R_{m,n}$ $rank\, A$ and A^T denote the rank and the transpose of matrix A respectively. In what follows we shall denote by '$a(\lambda)$' the characteristic polynomial of a matrix $A \in R_{m,m}$, i.e. $a(\lambda) = \det(I_m \lambda - A)$.

Theorem 1. (Sylvester's Theorem) Let $R = F$ be an algebraically closed field. Further, let $A \in F_{m,m}$ and $B \in F_{n,n}$. If A and B have no eigenvalues in common, then for each matrix $C \in F_{m,n}$, the equation

$$AX - XB = C, \qquad (1)$$

has a unique solution $X \in R_{m,n}$.

Sylvester discovered the result in 1884 [35]. The theorem 1 is well-known but deserves to be even better known, if for no other reason than its proof highlights the power of switching back and forth between matrices and linear transformations.

Note that $AX_0 - X_0 B = C$, then matrices $M_C = \begin{bmatrix} A & C \\ O & B \end{bmatrix}$ and $M_O = \begin{bmatrix} A & O \\ O & B \end{bmatrix}$ (of dimension $m+n$) are similar with the similarity transformation by the matrix $T = \begin{bmatrix} I_m & X_0 \\ O & I_n \end{bmatrix}$. Here the $O's$ are zero matrices.

W.E. Roth [32] proved that the converse is also true, i.e. if the matrices M_C and M_O are similar, and then Sylvester's matrix equation (1) has a solution. This result is called the similarity theorem of Roth. We note that the relationship of the Sylvester matrix equation with the block-diagonalization of block triangular matrices is evident. Generalizations of Sylvester's Theorem to bounded operators, as well as to some classes of unbounded operators, are due to Dalecki [4] and Rosenblum [31].

Another proof of Roth's theorem was given by Flanders and Wimmer in [7] (see also [8], [16], [20, 21]). In [18], Jameson studied the matrix equation $AX - XB = C$ by the method of characteristic polynomial. In [25] it is shown that Roth's theorems do not admit a generalization for infinite dimensional case. Analogous results for other linear matrix equations such as Stein equation $X - AXB = C$ were established by Wimmer [38] and other authors ([3], [9]).

The problem of solvability of equation (1) draws the attention of many mathematicians. Reviews of theoretical results of this problem are given in [10, 11, 20, 22, 30], and computational algorithms for solving equation (1) are described in [5, 18, 33]. It is well know that Sylvester equation is one of the important matrix equations in linear algebra and its applications (the control theory, the system theory and other). There is a vast literature in linear algebra and matrix theory on Sylvester equations.

Roth's similarity theorem have also been proved for matrices over commutative rings [10, 13, 17] and unit regular rings [15], but most proofs are existence proofs not furnishing any solution. These results were discussed in [9] for matrices over arbitrary rings. W.H. Gustafson [13] extended the result of Roth [32] over fields and proved the following theorem:

Theorem 2. (Gustafson's theorem). Let $A \in R_{m,m}$, $B \in R_{n,n}$ and $C \in R_{m,n}$ be non zero matrices over a commutative ring R with identity. The equation

$$AX - XB = C, \qquad (2)$$

has a solution over R if and only if matrices

$M_C = \begin{bmatrix} A & C \\ O & B \end{bmatrix}$ and $M_O = \begin{bmatrix} A & O \\ O & B \end{bmatrix}$ are similar over R.

This implies that solvability of matrix equation (2) contains the problem of similarity of matrices over commutative rings. For matrices over a commutative ring R with identity, including a residue ring Z/p^nZ, there is no simple criterion of similarity of matrices, or analogue of canonical forms of a polynomial matrix over a field is known. However, no intelligible results on the structure of similarity classes over R in the general case have been obtained yet. Apparently, all attempts at a complete classification of similarity classes of matrices over R are doomed to failure, because this problem is a "wild" [6]. In spite of the difficulties, in certain partial cases the description of similarity classes and the estimate of the number of classes are possible (see [1, 2, 12, 14, 19, 23, 26, 27, 29, 34] and references therein).

In this communication, we establish conditions for solvability of Sylvester matrix equation (2) in terms of columns equivalence and the Hermite normal forms of matrices constructed in a certain way by using the coefficients of this equation.

II. Main results

In what follows the notation R means an integer domain with identity $1 \neq 0$. The operator vec for any matrix $C = \lfloor c_{ij} \rfloor \in R_{p,q}$ is defined in the following way (see [8, 16, 21])

$$\bar{C} = \begin{bmatrix} c_{11} \dots c_{1q} & c_{21} \dots c_{2q} & \dots & c_{p1} \dots c_{pq} \end{bmatrix}^T,$$

i.e. the entries of C are stacked columnwise forming a vector of length mn. And we define the right Kronecker product by

$$A \otimes B = \begin{bmatrix} a_{11} B & \dots & a_{1n} B \\ \vdots & \ddots & \vdots \\ a_{m1} B & \dots & a_{m,n} B \end{bmatrix} \in R_{mp,nq},$$

where $A = \lfloor a_{ij} \rfloor \in R_{m,n}$ and $B \in R_{p,q}$.

Let $A \in R_{m,m}$, $B \in R_{n,n}$ and $C \in M_{m,n}$. Put

$$G = A \otimes I_n - I_m \otimes B^T.$$

Theorem 3. Let $A \in R_{m,m}$, $B \in R_{n,n}$ and $C \in R_{m,n}$. The following statements are equivalent:

1) equation $AX - XB = C$ has a solution $X_0 \in R_{m,n}$;

2) the system of linear equations $G\vec{X} = \vec{C}$ has a solution $\vec{X}_0 \in R_{mn,1}$;

3) matrices $\lfloor G \quad 0_{mn,1} \rfloor$ and $\begin{bmatrix} G & \bar{C} \end{bmatrix}$ are right equivalence over R.

Proof. (1) $\to$ (2). Let $X_0 \in R_{m,n}$ be a solution of the equation $AX - XB = C$, i.e., $AX_0 - X_0 B = C$. From this we have $G\bar{X}_0 = \bar{C}$ (see [21]). Thus, the system of equations $G\bar{X} = \bar{C}$ is solvable.

(2) $\to$ (3). Let $G\bar{X}_0 = \bar{C}$, where $\bar{X}_0 \in R_{mn,1}$. For the matrix $T = \begin{bmatrix} I_{mn} & \bar{X}_0 \\ 0_{1,mn} & -1 \end{bmatrix} \in GL(mn+1, R)$ we have

$$\begin{bmatrix} G & \bar{C} \end{bmatrix} T = \begin{bmatrix} G & 0_{mn,1} \end{bmatrix}.$$

Thus, matrices $\lfloor G \quad 0_{mn,1} \rfloor$ and $\begin{bmatrix} G & \bar{C} \end{bmatrix}$ are right equivalence.

(3)$\to$(1). Let

$$\begin{bmatrix} G & \bar{C} \end{bmatrix} = \begin{bmatrix} G & 0_{mn,1} \end{bmatrix} U. \tag{3}$$

where $U \in GL(mn+1, R)$. The matrix U we write in the form as $U = \begin{bmatrix} U_{11} & U_{12} \\ U_{2,1} & U_{22} \end{bmatrix}$, where $U_{11} \in R_{mn,mn}$ and $U_{12} \in R_{mn,1}$.

From (3) it follows $GU_{12} = \bar{C}$, where

$$U_{12} = \begin{bmatrix} \alpha_{11} \dots \alpha_{1n} & \alpha_{21} \dots \alpha_{2n} & \dots & \alpha_{m1} \dots c_{mn} \end{bmatrix}^T.$$

Thus, the matrix $X_0 = \begin{bmatrix} \alpha_{11} & \dots & \alpha_{1n} \\ \vdots & \ddots & \vdots \\ \alpha_{m1} & \dots & \alpha_{mn} \end{bmatrix} \in R_{m,m}$ is the solution of the equation $AX - XB = C$. The proof of Theorem 3 is complete.

Corollary 1. Let $A \in R_{m,m}$ and $B \in R_{n,n}$. The following statements are equivalent:

1) for each matrix $C \in R_{m,n}$ equation $AX - XB = C$ has a solution;
2) $a(B) \in GL(n, R)$;
3) $b(A) \in GL(m, R)$.

Corollary 2. Let $A \in R_{m,m}$ and $B \in R_{n,n}$ be matrices with characteristic polynomials $a(\lambda)$ and $b(\lambda)$ respectively. The following statements are equivalent:

1) the equation $AX = XB$ has a non-zero solution $X_0 \in R_{m,n}$;
2) matrices $a(B)$ and $b(A)$ are singular.

Remark. If a non-zero matrix $X_0 \in R_{m,n}$ is a solution of the equation $AX = XB$, then
$rank\, X_0 \le \min\{ m - rank\, a(B), \quad n - rank\, b(A) \}$.

Naturally, the following question arose: Let $A \in R_{m,m}$ be given. What shall a matrix $C \in R_{m,m}$ such that the matrix equation $AX - XA = C$ has a solution? In other words, what are matrices that are commutators? This equation was studied in many papers (see [30, 36, 37] and references therein). The answer to this question is as follows:

Theorem 4. Let matrices $A, C \in R_{m,m}$ be given. The matrix equation $AX - XA = C$ has a solution $X_0 \in R_{m,m}$ if and only if a trace of $C \in R_{m,m}$ is equal to zero and matrices

$\begin{bmatrix} A\otimes I_m - I_m A^T & 0_{mn,1} \end{bmatrix}$ and $\begin{bmatrix} A\otimes I_m - I_m A^T & \bar{C} \end{bmatrix}$ are right equivalent.

III. Applications

Presented results are true for matrices over domains of elementary divisors and Bezout domains. They can be generalized to equations over commutative rings of a more general algebraic nature.

A1. Let $R = D$ be a Bezout domain (see Chapter 1, [8]) and $A \in D_{m,n}$ be a non-zero matrix with $rank\, A = r$ in which the first k rows are zero, i.e., $A = \begin{bmatrix} 0_{k,n} \\ A_1 \end{bmatrix}$ and the first row of the matrix A_1 is non-zero, then, for A, there exists a matrix $W \in GL(n, D)$ such that

$$AW = H_A = \begin{bmatrix} & 0_{k,n} \\ H_1 & 0_{m_1, n-1} \\ H_2 & 0_{m_2, n-1} \\ \vdots & \vdots \\ H_r & 0_{m_r, n-r} \end{bmatrix},$$

where $H_1 = \begin{bmatrix} a_1 \\ * \end{bmatrix} \in D_{m_1,1}$, $H_2 = \begin{bmatrix} h_{21} & a_1 \\ & * \end{bmatrix} \in D_{m_2,2}, \cdots,$

$H_r = \begin{bmatrix} h_{r1} \ldots h_{r,r-1} & a_r \\ * & \end{bmatrix} \in D_{m_r,r}$, $k + m_1 + \ldots + m_r = m$.

The lower block-triangular matrix H_A is called a (right) Hermitian form of the matrix A and it is uniquely defined for A (see [8, 24]). Thus, matrices $A \in D_{m,n}$ and $B \in D_{m,n} \frac{1}{2}$ are right equivalent over a Bezout domain D if and only if the Hermitian normal forms of A and B coincide. Thus, we have arrived to the following result (see also [28]).

Theorem 4. The system of linear nonhomogeneous equations $A\bar{x} = \bar{b}$, where $A \in D_{m,n}$ and $\bar{b} \in D_{m,1}$, is solvable over a Bezout domain D if and only if the Hermitian normal forms of the matrices $\begin{bmatrix} A & 0_{m,1} \end{bmatrix}$ and $\begin{bmatrix} A & \bar{b} \end{bmatrix}$ coincide.

From the proof of Theorem 3 and Theorem 4 we obtain the algorithm for solving of Sylvester matrix equation over Bezout domains and domains of principal ideals in particular.

Corollary 3. Let $A \in D_{m,m}$, $B \in D_{n,n}$ and $C \in D_{m,n}$. Equation $AX - XB = C$ has a solution over D if and only if the Hermitian forms of the matrices $\begin{bmatrix} G & 0_{mn,1} \end{bmatrix}$ and $\begin{bmatrix} G & \bar{C} \end{bmatrix}$ coincide.

A2. Let Z denote the set of integers. Further, let $A \in Z_{m,m}$, $B \in Z_{n,n}$ and $Z \in R_{m,n}$. From corollary 3 it follows the following statement. Matrix equation $AX - XB = C$ has a solution $X_0 \in Z_{m,n}$ if and only if the Hermitian forms of the matrices $\begin{bmatrix} G & 0_{mn,1} \end{bmatrix}$ and $\begin{bmatrix} G & \bar{C} \end{bmatrix}$ coincide.

The nonzero integer $d \in Z$ is said to be *square-free* if it has no square factor, i.e. $a^2 \mid d$ if and only if $a = \pm 1$. Let $d \in Z$ be a square-free integer not equal to 0 or 1. We now consider a ring of quadratic integers

$$K = Z + Z\sqrt{d} = \{a + b\sqrt{d},\ where\ a, b \in Z\}.$$

Put $x = x_1 + x_2\sqrt{d}$ and $y = y_1 + y_2\sqrt{d}$. It is clear that K is an integral domain, i.e. K is a commutative ring with unit $1 \neq 0$ such that if $x \cdot y = 0$, then either $x = 0$ or $y = 0$ and with the usual addition $x + y = x_1 + y_1 + (x_2 + y_2)\sqrt{d}$ and multiplication $x \cdot y = x_1 y_1 + d x_2 y_2 + (x_1 y_2 + x_2 y_1)\sqrt{d}$. The study of rings of quadratic integers is basic for many questions of algebraic number theory.

Let $A \in K_{m,m}$, $B \in K_{n,n}$ and $C \in K_{m,n}$. Consider the equation $AX - XB = C$. These matrices we rewrite in the forms $A = A_1 + \sqrt{d} A_2$, $B = B_1 + \sqrt{d} B_2$, $C = C_1 + \sqrt{d} C_2$, where A_i, B_i, C_i are matrices with entries from the ring Z and put $X = X_1 + \sqrt{d} X_2$, where X_i are unknown $m \times n$ matrices over Z. From this equation it follows

$$\begin{cases} (A_1 X_1 - X_1 B_1) & + d(A_2 X_2 - X_2 B_2) & = C_1, \\ (A_2 X_1 - X_1 B_2) & + (A_1 X_2 - X_2 B_1) & = C_2. \end{cases} \quad (4)$$

The system of equations (4) we rewrite in the form

$$\begin{bmatrix} A_1 \otimes I_n - I_m \otimes B_1^T & d(A_2 \otimes I_n - I_m \otimes B_2^T) \\ A_2 \otimes I_n - I_m \otimes B_2^T & (A_1 \otimes I_n - I_m \otimes B_1^T) \end{bmatrix} \begin{bmatrix} \bar{X}_1 \\ \bar{X}_2 \end{bmatrix} = \begin{bmatrix} \bar{C}_1 \\ \bar{C}_2 \end{bmatrix}.$$

Corollary 4. Let $A \in K_{m,m}$, $B \in K_{n,n}$ and $C \in K_{m,n}$. The equation $AX - XB = C$ is solvable over the ring K if and only if the Hermitian forms of the matrices

$$\begin{bmatrix} (A_1 \otimes I_n - I_m \otimes B_1^T) & d(A_2 \otimes I_n - I_m \otimes B_2^T) & 0_{mn,1} \\ (A_2 \otimes I_n - I_m \otimes B_2^T) & (A_1 \otimes I_n - I_m \otimes B_1^T) & 0_{mn,1} \end{bmatrix}$$

and

$$\begin{bmatrix} (A_1 \otimes I_n - I_m \otimes B_1^T) & d(A_2 \otimes I_n - I_m \otimes B_2^T) & \bar{C}_1 \\ (A_2 \otimes I_n - I_m \otimes B_2^T) & (A_1 \otimes I_n - I_m \otimes B_1^T) & \bar{C}_2 \end{bmatrix}$$

coincide.

Let F be a field. Put $R = F[\lambda]$ the polynomial ring over F. Further, let $A \in F_{m,m}[\lambda], B \in F_{n,n}[\lambda]$ and $C \in F_{m,n}[\lambda]$. Suppose that the Hermitian forms of matrices $\begin{bmatrix} G[\lambda] & 0_{mn,1} \end{bmatrix}$ and $\begin{bmatrix} G[\lambda] & \bar{C}[\lambda] \end{bmatrix}$ coincide. Then for every solution $X_0[\lambda]$ of equation $A[\lambda]X[\lambda] - X[\lambda]B[\lambda] = C[\lambda]$ we have

$$\deg X_0[\lambda] \leq \deg A[\lambda] + \deg B[\lambda] - \deg C[\lambda].$$

References

[1] N. Avni, U. Onn, A. Prasad, L. Vaserstein, "Similarity classes of 3×3 matrices over a local principal ideal ring", Communications in Algebra, 2009, Vol. **37**, № 8, pp.2601–2615.

[2] V. M. Bondarenko, "On the similarity of matrices over a ring of the classes of residues", in: Mathematical Collection [in Russian], Naukova Dumka, Kiev, 1976, pp.275–277.

[3] G. Bourgeois, “How to solve the matrix equation XA – AX=f(X)”, Linear Algebra and Appl., 2011, Vol. **434**, №. 3, pp.657–668.

[4] Ju. L. Dalecki, On the asymptotic solution of a vector differential equation, Dokl. Akad.Nauk SSSR, 1953, **92,** №5, pp.881–884 (in Russian).

[5] A. Dmytryshyn, B. Kågström, “Orbit closure hierarchies of skew-symmetric matrix pencils”, SIAM Journal on Matrix Analysis and Applications. – 2014. – T. 35, № 4, pp.1429-1443.

[6] Yu. A. Drozd, “Tame and wild matrix problems,” in: Representations and Quadratic Forms [in Russian], Institute of Mathematics, Ukrainian Academy of Sciences, Kiev, pp. 39–74, 1979.

[7] 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.

[8] S. Friedland, “Matrices: Algebra, Analysis and Applications”, Singapore: World Scientific Publishing Co., 2015.

[9] M.C. Gouveia, “On the solution of Sylvester, Lyapunov and Stein equations over arbitrary rings”, International Journal of Pure and Applied Mathematics, 2005, Vol. **24,** №. 1, pp.131–137.

[10] R. M. Guralnick, “Roth's theorems and decomposition of modules”, Linear Algebra and Its Applications, 1981, Vol. **39**, pp.155–165.

[11] R. M. Guralnick, “Roth's theorems for sets of matrices”, Linear algebra and its applications, 1985, Vol. **71**, pp.113–117.

[12] R. M. Guralnick, “Similarity of matrices over commutative rings”, Linear Algebra Appl., 1991, Vol. **157**, pp. 55–68.

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

[14] W.H. Gustafson, “On matrix similarity over commutative rings”, Linear and Multilinear Algebra, 1981, Vol. **10**, № 3, pp.249–252.

[15] R.E. Harwig, “Roth’s equivalence problem in unit regular ring”, Proc. Amer. Math. Society, 1976, Vool. 76, № 1, pp.39–44.

[16] R. Horn, Ch. R. Johnson, “Topics in matrix analysis”, Cambridge University Presss, Cambridge, 1991.

[17] L. Huang, J. Liu, “The extension of Roth's theorem for matrix equations over a ring”, Linear algebra and its applications, 1997, Vol. **259**, pp.229–235.

[18] A. Jameson, “Solution of the equation AX + XB=C by inversion of an M×M or N×N matrix”, SIAM J. Applied Mathematics, 1968, Vol.**16**, pp.1020–1023.

[19] V.L. Kurakin, “Similarity invariants for matrices over an Artinian ring”, Mathematical Notes, 2006, **Vol**. 80, pp. 387–395. Translated from Matematicheskie Zametki, Vol. 80, № 5, pp. 763–770.

[20] P. Lancaster, “Explicit solution of linear matrix equations”, SIAM Review, 1970, Vol.**12**, pp.544–566.

[21] P. Lancaster, M. Tismenetsky, “The theory of matrices: with applications”, Academic Press, 1985.

[22] S.G. Lee, Q.P. Vu, “Simultaneous solutions of matrix equations and simultaneous equivalence of matrices”, Linear Algebra and its Applications, 2012, V. 437, № 9, pp. 2325–2339.

[23] A. A. Nechaev, “On the similarity of matrices over a commutative Artinian ring”, Journal of Mathematical Sciences, 1986, Vol. **33**, № 5, pp.1221–1237. Translated from Trudy Seminara imeni I. G. Petrovskogo, 1983, No. 9, pp. 81–101

[24] M. Newman, “Integral Matrices,” Acad.emic Press, New York and London, 1972.

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

[26] V. Prokip, “On similarity of matrices over commutative rings,” Linear algebra and its applications, 2005, Vol. **399**, pp. 225–233.

[27] V. M. Prokip, “Diagonalizability of matrices over a principal ideal domain, matrices over a principal ideal domain,” Ukrainian Mathematical Journal, 2012, Vol.**64**, № 2, pp.316–323. Translated from Ukrains’kyi Mat. Zhurnal, Vol. 64, No. 2, pp. 283–288.

[28] V. M. Prokip, “On solvability On the solvability of a system of linear equations over the domain of principal ideals”, Ukrainian Mathematical Journal, 2014, Vol. **66**, № 4, pp.566–570. Translated from Ukrains’kyi Matematychnyi Zhurnal, Vol. 66, No.4, pp. 566–570,

[29] V. M. Prokip, “On the structure of matrices over a domain of principal ideals with respect to similarity transformations” (in Ukrainian), Proc. of the International Geometry Center, 2019, Vol.12, № 1, pp.56–69.

[30] E. L. Rabkin, “Full investigation of the matrix equation AX − XB=C and specifically of the equation AX − XA=C”, St. Peterburg Math. J., 2015, Vol.**26**, № 1, pp.117–130. Translated from Algebra i Analiz, 2014, Vol.**26,** № 1, pp.165–184.

[31] M. Rosenblum, “On the operator equation BX − XA = Q”, Duke Math. J., 1956, Vol. **23**, pp.263–270.

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

[33] I.V. Shapochka, N.V. Yurchenko, “On the Sylvester matrix equation over the domain of principal ideals”, Scientific Bulletin of Uzhhorod National University. Series Mathematics and Informatics [in Ukrainian], 2015, Issue **26** (1), pp.170–174.

[34] S. V. Sidorov, “On the Similarity of Certain Integer Matrices with Single Eigenvalue over the Ring of Integers”, Mathematical Notes, 2019, Vol. **105**, Issue 5-6, pp.756-762. Translated from Matematicheskie Zametki, 2019, Vol. **105**, № 5, pp.763–770.

[35] J. Sylvester, “Sur l’ȇquations en matrices px=xq”, C.R. Acad. Sci. Paris, 1884, V.**99** , pp. 67–71.

[36] A. Stasinski, "Similarity and commutators of matrices over principal ideal rings." Trans. Amer. Math. Society, 2016, Vol.**368**, № 4, pp.2333–2354.

[37] A. Stasinski, “Commutators of trace zero matrices over principal ideal rings”, Israel Journal of Mathematics, 2018, Vol.228, № 1, pp. 211–227.

[38] H. K. Wimmer, “The matrix equation X − AXB = C and analogue of Roth’s theorem”, Linear Algebra and its Applications, 1988, Vol. 109, pp.145–147.