Algorithm for Finding a Positive Definite Solution to the Sylvester Matrix Equation

Oleksii Bychkov [1], Ganna Marzafey [2] and George Dimitrov [3]

[1] *Taras Shevcenko National University of Kyjv, Volodymyrska str. 60, Kyjv, Ukraine, 01601*
[2] *University of Library Studies and Information Technologies, Blvd. "Tsarigradsko Shose" 119, 7-Kilometar, Sofia, Bulgaria, 1784*

Abstract

The article examines the problem of constructing an algorithm for finding positive definite matrices that are the solution to Sylvester's three-term matrix equation. The problem is that, unlike the Lyapunov equation, such a condition cannot be written in terms of eigenvalues. The condition for the existence of a solution to the Sylvester equation is based on the principle of contraction mappings. The article also proposes an iterative procedure and algorithm for finding a solution.

Keywords[1]

algorithm, stability, iteration procedure, Silvestr, solution

1. Introduction

The Sylvester matrix equation $AX + XB = C$, also sometimes called the continuous Sylvester equation, and the Stein matrix equation $X + AXB = C$, in turn, sometimes called the discrete Sylvester equation, as well as their special cases - the Lyapunov equations $A^T X + XA = C$ and $X - AXA^T = C$, are well-studied and frequently encountered (for example, in the theory of differential equations) classes of matrix equations [1-6]. The conditions for the unique solvability of these equations have long been known; there are numerical algorithms for solving them, for example, the Bartels-Stewart and Golub-Nash-Van Loan algorithms.

An analogue of Sylvester's equation $AX + XB = C$ began to attract the attention of researchers relatively recently [7-9]. Although this equation is superficially very similar to Sylvester's equation, their natures are profoundly different. Let's give a simple example to illustrate this difference. If all matrices are square and $A = B$, then Sylvester's equation has a unique solution X - for any right-hand side C. For the same coefficients A and B, the equation $AX + XB = C$ has a solution only if the matrix C is symmetric. If this condition is met and X satisfies this equation, then $X + K$, where K is an arbitrary skew-symmetric matrix, is also a solution.

Equation $A^T X + XA = C$, as well as the equations $AX + XB = C$ we will generally call two member equations of Sylvester type. The relevance of studying this kind of equations is beyond doubt. Let us give several examples showing why the study of equations of Sylvester type is justified. Equation $AX + XB = C$ was first encountered by us in article [10-12], where it was studied under the additional assumption $C = C^T$.

Solvability conditions and a description of the general solution were given in terms of generalized inverses for matrices A and B and their associated projectors. These conditions are not entirely

Dynamical System Modeling and Stability Investigation (DSMSI-2023), December 19-21, 2023, Kyiv, Ukraine
EMAIL: oleksiibychkov@knu.ua (O. Bychkov); annamartsafei@knu.ua (G, Marzafey); geo.p.dimitrov@gmail.com (G. Dimitrov)
ORCID: 0000-0002-9378-9535 (O. Bychkov); 0000-0001-5064-3168 (G. Dimitrov)

constructive and are difficult to verify. Homogeneous equation $AX + XB = C$ was studied in article [2] in the special case $B = A$. The authors were motivated by the fact that the set $AX + XB \in C^{n\times n}$ is the tangent space (calculated at point A) of the orbit of matrix A under the action of congruences. The codimension of this orbit is exactly the dimension of the space of solutions to the equation $AX + AX = 0$. The main result of work [2,3] was the establishment of the canonical structure of matrices in general position with respect to congruences. In a similar way, the same authors in [4] studied the equation $AX + AX = 0$. In the publication [4] the equations $AX + XB = C$ arise when constructing an algorithm for palindromic eigenvalue problems $Ax = \lambda A^T x$.

In the process of reducing matrix A to antitriangular form, the need arises to solve these equations numerically. In article [5], the conditions for unique solvability and a numerical algorithm for solving the equation $AX + XB = C$ are formulated.

By analogy with equations of Sylvester type, equations of Stein type will be called equations $X + AXB = C$. The first of them was partially studied in publication [6]. The question naturally arises about the completion of this study. The topic of solvability of the other two equations, on the contrary, is explored exhaustively in this publication.

No less interesting is the problem of obtaining sufficient solvability conditions for the Sylvester matrix equation

$$A^T H + HA + B^T HB = -C,$$

where C is a positive definite matrix. The need to solve such equations is emphasized in [13-14].

It is noted in [9] that there are now several approaches to solving the Sylvester equation. The first is to reduce the matrix equation to a vector (linear algebraic equation of increased dimension) and then the condition for the solvability of this equation is expressed through the non-degeneracy of the corresponding matrix. The second approach uses the small parameter method. It is also possible to obtain a spectral sparsity criterion for an equation with mutually commutable matrices. An analytical solution of this equation is possible only for the case of the two-term Sylvester equation.

At the same time, the question of the existence of a positive definite solution in the general case remains open. The procedure for finding this solution is also of interest for another investigations [15-16]. The purpose of the article is to obtain sufficient conditions for the existence of a solution to the Sylvester matrix equation on a set of positive definite matrices. The article will also present an iterative procedure and algorithm for finding these solutions.

2. Main result

The approach is based on a modification of the contraction mapping method. If on a complete metric space M the following operator is specified $F[x], x \in M$, that maps points $x \in M$to points of the same space $F[x] \in M$ and satisfies the contraction condition $\rho(F[x], F[y]) \le \alpha\rho(x, y), 0 < \alpha < 1$, where $\rho(x, y)$ is the metric of the space M, then the operator equation

$$x = F[x]$$

has a unique solution $x^* \in M$, and it can be found by the method of successive iterations

$$x = \lim_{n\to\infty} x_n, x_n = F[x_{n-1}], n = 1,2, \dots, x_0 = x^0.$$

Let x_0 - be an arbitrary point in space, H is some complete metric space with metric $\rho(H_1, H_2)$. Let us fix the following operators:

$$F, G: H \to H.$$

Consider the operator equation

$$F[H] = G[H] \tag{1}$$

Let us show that to solve this equation we can use an iterative procedure based on the following implicit scheme

$$F[H_{s+1}] = G[H_s], H_0 = H^0, s = 1,2,..,;\ H^0 \in \mathrm{H} \quad (2)$$

Definition. We call method (2) converging to the solution of equation (1) if the sequence $\{y_s\}, y_s = F[H_s]$, at $s \to \infty$ converges. Let us prove a theorem whose conditions ensure the convergence of method (2) to the solution of equation (1).

Let us assume that the solution to equation (1) lies in the set $H(A, r)$:

$$H(A, r) = \{H \in H\colon\ \rho\ (F[H], F[A])\langle r\}.$$

Theorem 1. Let us fix the set $H(A, r)$, i.e. the values of r and A are determined. If for arbitrary $H_1, H_2 \in H(A, r)$ operators $F, G\colon H \ \to\ H$ satisfy the contraction condition

$$\rho(G[H_1], G[H_2]) < q\ \rho(F[H_1], F[H_2]), 0 < q < 1, \quad (3)$$

and

$$\rho(G[A], F[A]) < (1-q)r\ , \quad (4)$$

then operator equation (1) has a unique solution in the set $H(a, r)$ and the sequence $\{y_s\}, y_s = F[H_s]$ constructed according to scheme (2) converges to $y^* = F[H^*]$ for any starting point $H_0 \epsilon H(A, r)$. For the error of method (2), the following estimate is valid:

$$\rho(F[H_k], F[H^*]) < \frac{q^k}{1-q}\rho(G[H_0], F[H_0]). \quad (5)$$

Proof. Let H_0 - be an arbitraryelement of *H(A, r)* . Let us prove that the sequence H_k built according to scheme (2) will not leave the set *H(A, r)*. By condition $H_0 \in H(A, r)$. Let us assume that $H_k \in H(A, r)$for some fixed k. Let us show that then $H_{k+1} \in H(A, r)$. Consider the equality:

$$F[H_{s+1}] = G[H_s]$$

Let us subtract the $F[A]$ from both sides of the above mentioned equation, we get

$$F[H_{k+1}] - F[A] = G[H_k] - F[A] = (G[H_k] - G[A]) + (G[A] - F[A])$$

Then the inequality holds

$$\rho(F[H_{k+1}], F[A]) < \rho(G[H_k], G[A]) + \rho(G[A], F[A]).$$

Using the induction hypothesis, we have

$$\rho(G[H_k], G[A]) < q\rho(F[H_k], F[A]) < \mathrm{qr}$$

and since by the conditions of the theorem $\rho(G[A], F[A]) < (1-q)r$ we get that

$$\rho(F[H_{k+1}], F[A]) < \mathrm{qr} + (1-q)r = r.$$

Thus $H_{k+1} \in H(A, r)$.

Now we will show that the following sequence $\{y_s\}, y_s = F[H_s]$ is Cauchy. Let's consider the difference $F[H_{k+1}] - F[H_k]$:

$$F[H_{k+1}] - F[H_k] = G[H_k] - G[H_{k-1}]$$

Because the $H_k \in H(A, r)$, then using condition (3) we obtain

$$\rho(F[H_{k+1}], F[H_k]) = \rho(G[H_k], G[H_{k-1}]) < q\rho(F[H_k], F[H_{k-1}])$$

and therefore

$$\rho(F[H_{k+1}], F[H_k]) < q^k\rho(F[H_1], F[H_0]) \quad (6)$$

Let $p \in N$. Then it's fair

$$F[H_{k+p}] - F[H_k] = \sum_{j=1}^{p}\left(F[H_{k+j}] - F[H_{k+j-1}]\right).$$

And according to (6) we get:

$$\rho\left(F\left[H_{k+p}\right],F\left[H_k\right]\right)<\frac{q^k}{1-q}\rho(F[H_1],F[H_0]). \qquad (7)$$

Because $\lim_{k\to ¥}\frac{q^k}{1-q}=0$ and does not depend on p, then the sequence $\{y_s\}$ - Cauchy. Therefore there is a limit

$$\lim_{k\to\infty} y_s=y^*, y^*=F[H^*], H^*\in H(A,r)$$

Let us pass to the limit in (2) at $k\to\infty$, then we get that $F[H^*]=G[H^*]$. Hence H^*- is a solution of equation (1). Let us prove the uniqueness of this solution. Let H_* is a some another solution (1) and $H_*\in H(A,r)$. Then

$$F[H_*]-F[H^*]=G[H_*]-G[H^*]$$

and

$$\rho(F[H_*],F[H^*])<q\rho(F[H_*],F[H^*]).$$

Because $0<q<1$, That $H_*=H^*$.

Let us prove the estimate of the error of method (2). Let k be fixed and $p\to\infty$. Then from (7) we obtain

$$\rho(F[H_k],F[H_*])<\frac{q^k}{1-q}\rho(F[H_1],F[H_0]), k=1,2,K.$$

Therefore the theorem is proven.

2.1. Existance of solution

Denote as H the space of symmetric matrices with the metric -$\rho(H_1,H_2)=|H_1-H_2|$.

Let us write the Lagrange formula for the operators F and G. We have:

$$F[H_1]-F[H_2]=H_2(x)(H_1-H_2), G[H_1]-G[H_2]=H_1(x)(H_1-H_2),$$

Where $H_1(.), H_2(.)$ are the Gâteaux derivatives of the operators F and G at some midpoint.

Assume that there is an inverse operator H_2^{-1}. Then we get:

$$G[H_1]-G[H_2]=H_1(x)H_2^{-1}(x)(F[H_1]-F[H_2]).$$

Thus we have

$$|\ G[H_1]-G[H_2]|\le\left|H_1(x)H_2^{-1}(x)\right||F[H_1]-F[H_2]| \qquad (8)$$

It follows that condition (3) of Theorem 1 will be satisfied if

$$\left|H_1(x)H_2^{-1}(x)\right|\le q<1. \qquad (9)$$

This condition is not always convenient. Using Theorem 1 and inequality (9), we obtain the conditions for the solvability of the Sylvester matrix equation.

Let us define the operators F and G as follows:

$$F[H]=-A^T H-\mathrm{HA}, G[H]=C+B^T\mathrm{HB},$$

where A,B are some matrices, C is a positive definite matrix. Then the Sylvester matrix equation can be rewritten as:

$$F[H]=G[H]. \qquad (10)$$

Let us obtain constructive solvability conditions for this equation.

Lemma 1. A necessary condition for the solvability of the Sylvester matrix equation on the set of positive definite matrices is that the matrix A is Hurwitz.

Proof. Let H_0 is a positive definite solution to Sylvester's equation. Consider the expression

$$-A^T H_0-H_0A=C+B^T H_0B.$$

Let's denote $C_1=C+B^T H_0 B$. Because C and H_0 are positive definite, then C_1 will be a positive definite. Then

$$-A^T H_0 - H_0 A = C_1$$

is a matrix Lyapunov equation and since H_0 is his solution, then A is a Hurwitz matrix.

2.2. Iteration procedure

Lemma 2. Let matrices A, B satisfy the following conditions:

$$1)\ A\text{- is Hurwitz matrix,} \tag{11}$$

$$2)|(B^T \times \mathrm{B})(-A^T \times \mathrm{I} - \mathrm{I} \times \mathrm{A})^{-1}| \le q, \tag{12}$$

where $\times$ denotes the Kronecker product, then there is a unique positive definite solution to the Sylvester matrix equation and it can be found using the iterative procedure

$$-A^T H_{k+1} - H_{k+1} A = C + B^T H_k B,$$

where C is an arbitrary positive definite matrix.

Proof. Let us apply inequality (9) to the operators defining the Sylvester equation. For operator G we have:

$$H_1(\xi) = B^T \times \mathrm{B}.$$

Since A is Hurwitz, then the operator $H_2^{-1}(\xi)$ exists and can be written

$$H_2^{-1}(\xi) = (-A^T \mathrm{xI} - \mathrm{IxA})^{-1}.$$

Consequently, we find that inequality (12) ensures that condition (3) of Theorem 1 is satisfied. Let us take as the working set the entire space of positive definite matrices, then $r=\infty$. Thus, we obtain conditions for the existence and uniqueness of a solution to the Sylvester matrix equation.

Let us show that the solution obtained by the iterative procedure

$$-A^T H_{k+1} - H_{k+1} A = C + B^T H_k B$$

will be a positive definite matrix. Let us prove this by mathematical induction. Matrix H_0 positive definite - due to the choice of the starting point. Let H_k is a positive definite matrix. Let's show positive definiteness H_{k+1}.

Consider the equation

$$-A^T H_{k+1} - H_{k+1} A = C_k,$$

where $C_k = C + B^T H_k B, k = 1, \dots, n, \dots$

Since A is Hurwitz and C_k- is always positive definite, thenH_{k+1}, as a solution to the Lyapunov equation will be a positive definite matrix. The lemma is proven.

2.3. Application examples

Let us consider examples of the fulfillment of the conditions of Lemma 2. Let in the first example

$$A = \begin{pmatrix} 3 & 5 & 0 \\ 2 & 1 & 0 \\ 0 & 3 & 5 \end{pmatrix} \quad B = \begin{pmatrix} 0.1 & 0.5 & 0 \\ 4 & 0.3 & 0.9 \\ 0.1 & 0 & 0.5 \end{pmatrix}$$

where matrix A is Hurwitz. We substitute the values of matrices A and B into condition (12), carry out calculations and find that all elements in the resulting matrix are less than 1 in absolute value, which indicates that the condition is met.

In the second example

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 3 & 4 & 0 \\ 0 & 1 & 2 \end{pmatrix} \quad B = \begin{pmatrix} 0.1 & 0.2 & 0.1 \\ 0.3 & 0.1 & 0.5 \\ 0.2 & 0.4 & 0.1 \end{pmatrix}$$

we carry out similar calculations and obtain the fulfillment of conditions (11) and (12).

Let us give an example of non-fulfillment of the conditions of Lemma 2. Let the values of the matrices be as follows:

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 3 & 4 & 0 \\ 0 & 1 & 2 \end{pmatrix} \qquad B = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 4 & 5 \\ 2 & 4 & 1 \end{pmatrix}$$

In this case, when calculating condition (12), we obtain a matrix where some elements have absolute values greater than 1, and this indicates that the condition is not met.

2.4. Algoritm for solving Silvest’r equation

Let us consider an algorithm for finding positive definite solutions. For finding a solution of the Sylvester equation, we apply the following algorithm:

1. Check the Hurwitz property of matrix A.
2. Check the condition $|(B^T \times \mathrm{B})(-A^T \times \mathrm{I} - \mathrm{I} \times \mathrm{A})^{-1}| \leq q$.
3. If 1.2 are true, then a solution exists and perform step 4. Otherwise, step 9.
4. We let $k = 0, H_0 = I, I -$ identity matrix.
5. At the kth step we calculate $C_1 = C + B^T H_k B$.
6. Solve the Lyapunov equation $-A^T H_{k+1} - H_{k+1} A = C_1$.
7. Check the condition for ending the iteration procedure. If it is not fulfilled, then $k = k + 1$ and go to step 5, otherwise go to step 8.
8. Resulting matrix H_k - is a solution to the Sylvester matrix solution.
9. The end.

3. Acknowledgements

We thanks professor Khusainov Denys and Erasmus+ for supporting our research.

4. Conclusions and discussion of results

The article studies the solvability of Sylvester's three-term matrix equation. Using the principle of contraction mappings, a sufficient condition for the existence of a positive definite solution is obtained. It should be noted that such conditions have not been published before. In other works authors solve similar equations by expanding the equation into a system of linear algebraic equations with constant coefficients and then solving it, for example, using the Gauss method. This approach does not allow us to obtain the condition for the existence of the necessary solution. In this article, in addition to the sufficient condition for the existence of a positive definite solution to the three-term matrix Sylvester equation, an algorithm for finding it is proposed and the convergence of this algorithm is proven.

The method presented in this article will allow us to more effectively solve problems of stability and controllability, which are presented in [17-22]. The solvability condition for Sylvester's three-term matrix equation is easy to verify. It is constructive. The proposed algorithm effectively finds a symmetric positive definite solution.

5. References

[1]. Xing Lili, Li Weiguo, Bao Wendi, Some results for Kaczmarz method to solve Sylvester matrix equations, Journal of the Franklin Institute, Volume 360, Issue 11, 2023, pp. 7457-7461.

[2]. Piao F., Zhang Q., Wang Z. The solution to matrix equation AX + XTC = B // J. Franklin Inst. 2007. V. 344. N 8. P. 1056-1062.
[3]. Zebin Chen, Xuesong Chen, Modification on the convergence results of the Sylvester matrix equation AX+XB=C, Journal of the Franklin Institute, Volume 359, Issue 7, 2022, pp. 3126-3147
[4]. Teran F., Bopico F. M. The solution of the equation XA+AXT = 0 and its application to the theory of orbits // Linear Algebra Appl. 2011. V. 434. N 1. P. 44-67.
[5]. Zhaolu Tian, Yudong Wang, Yinghui Dong, Xuefeng Duan, The shifted inner–outer iteration methods for solving Sylvester matrix equations, Journal of the Franklin Institute, Volume 361, Issue 5, 2024, https://www.sciencedirect.com/science/article/pii/S0016003224000954
[6]. Shihai Li, Changfeng Ma, Factor gradient iterative algorithm for solving a class of discrete periodic Sylvester matrix equations, Journal of the Franklin Institute, Volume 359, Issue 17, n2022, pp. 9952-9970.
[7]. Soheila Ghoroghi Shafiei, Masoud Hajarian, Developing Kaczmarz method for solving Sylvester matrix equations, Journal of the Franklin Institute, Volume 359, Issue 16, 2022, pp. 8991-9005.
[8]. Lakhlifa Sadek, Ahmad Sami Bataineh, Osman Rasit Isik, Hamad Talibi Alaoui, Ishak Hashim, A numerical approach based on Bernstein collocation method: Application to differential Lyapunov and Sylvester matrix equations, Mathematics and Computers in Simulation, Volume 212, 2023, pp. 475-488.
[9]. Kressner V., Schroder C., Watkins BS Implicit QR algorithms for palindromic and even eigenvalue problems // Numerical Algorithms. 2009. V. 51. N 2. P. 209-238.
[10]. George A., Ikramov Kh. D., Matushkina E.V., Tang W.-P. On a QR-Like Algorithm for Some Structured Eigenvalue Problems // SIAM J. Matrix Anal. Appl. 1995. V. 16. N. 4. P. 1107-1126.
[11]. Zhou W., Lam J., Buan G.-R. Toward solution of matrix equation X = Ai(X)B + C // Linear Algebra Appl. 2011. V. 435. N 6. P. 1370-1398.
[12]. Wenli Wang, Caiqin Song, Iterative algorithms for discrete-time periodic Sylvester matrix equations and its application in antilinear periodic system, Applied Numerical Mathematics, Volume 168, 2021, pp. 251-273.
[13]. Zhaolu Tian, Yudong Wang, Yinghui Dong, Shiyu Wang, New results of the IO iteration algorithm for solving Sylvester matrix equation, Journal of the Franklin Institute, Volume 359, Issue 15, 2022, pp. 8201-8217.
[14]. Soheila Ghoroghi Shafiei, Masoud Hajarian, An iterative method based on ADMM for solving generalized Sylvester matrix equations, Journal of the Franklin Institute, Volume 359, Issue 15, 2022, pp. 8155-8170.
[15]. Korenivsky D.G. Destabilizing effect of parametric white noise in continuous and discrete dynamic systems. – Proceedings of the Institute of Mathematics of the National Academy of Sciences of Ukraine, vol. 59, 2006. – 110 p.
[16] Biswa Datta Numerical methods for linear control systems: design and analysis. Elsevier Academic Press, Amsterdam ; Boston. ISBN 978-0-12-203590-6.
[17]. Shatyrko A., Khusainov D. On the Interval Stability of Weak-Nonlinear Control Systems with Aftereffect // Open Source Journal. The Scientific World Journal, vol. 2016, Article ID 6490826, 8 pages, 2016. doi:10.1155/2016/6490826
[18]. Khusainov D.Ya., Bychkov A.S. Estimates of exponential convergence of difference systems with delay //Differential equations, vol. 38, no. 6, 2002, pp. 1803-1806
[19]. Khusainov D.Ya., Shatyrko A.V. Absolute stability of multidelay regulation systems. Journal of Automation and Information Sciences, 1995, 27(3-4). P.33-42
[20]. Bychkov A.S., Merkuryev M.G. Stability of continuous hybrid automata //Cybernetics and systems analysis. – 2007. - No. 2. - With. 123-128
[21]. K. Merkulova and Y. Zhabska, "Input Data Requirements for Person Identification Information Technology", Proceedings of the 1st International Workshop on Computer Information Technologies in Industry 4.0 (CITI 2023), vol. 3468, 2023, pp. 24-37
[22] O. Kalivoshko, V. Kraevsky, K. Burdeha, I. Lyutyy and N. Kiktev, "The Role of Innovation in Economic Growth: Information and Analytical Aspect," *2021 IEEE 8th International Conference on Problems of Infocommunications, Science and Technology (PIC S&T)*, Kharkiv, Ukraine, 2021, pp. 120-124, doi: 10.1109/PICST54195.2021.9772201