Applied Mathematics and Computation 218 (2012) 8393–8407

Contents lists available at SciVerse ScienceDirect

Applied Mathematics and Computation

journal homepage: www.elsevier.com/locate/amc

On the ★-Sylvester equation $AX \pm X^{\star} B^{\star} = C$

Chun-Yueh Chiang [a,*], Eric King-Wah Chu [b], Wen-Wei Lin [c]

[a] Center for General Education, National Formosa University, Huwei 632, Taiwan
[b] School of Mathematical Sciences, Building 28, Monash University, VIC 3800, Australia
[c] Department of Applied Mathematics, National Chiao Tung University, Hsinchu 300, Taiwan

ARTICLE INFO

Keywords:
Linear matrix equation
Lyapunov equation
Palindromic eigenvalue problem
QR decomposition
Generalized algebraic Riccati equation
Schur decomposition
Singular value decomposition
Solvability

ABSTRACT

We consider the solution of the ★-Sylvester equations $AX \pm X^{\star}B^{\star} = C$, for $\star = T$, H and $A, B, \in \mathbb{C}^{n\times n}$, and the related linear matrix equations $AXB^{\star} \pm X^{\star} = C$, $AXB^{\star} \pm CX^{\star}D^{\star} = E$ and $AX \pm X^{\star}A^{\star} = C$. Solvability conditions and numerical methods are considered, in terms of the (generalized and periodic) Schur and QR decompositions. We emphasize the square cases where $m = n$ but the rectangular cases will be considered.

1. Introduction

In [4], the Lyapunov-like linear matrix equation

$$A^{\star}X + X^{\star}A = B, \quad A, X \in \mathbb{C}^{m\times n}\ (m \neq n)$$

with $(\cdot)^{\star} = (\cdot)^T$ was considered using generalized inverses. Applications occur in Hamiltonian mechanics. At the end of [4], the more general Sylvester-like equation

$$A^{\star}X + X^{\star}C = B, \quad A, C, X \in \mathbb{C}^{m\times n}\ (m \neq n)$$

was proposed without solution. The equation (with $\star = T$) was studied, again using generalized inverses, in [11,16]. However in [16], the necessary and sufficient conditions for solvability may be too complicated for most applications. The formula for X for the special case, assuming $m = n$, $B^T = B$ and the invertibility of $A \pm C^T$, may not be numerically stable or efficient (see Appendix B for the main result). In [11], some necessary or sufficient conditions for solvability were derived. A (seemingly wrong) formula for X in terms of generalized inverse was also proposed (see Section 2.2 for more details on the approach taken in [11]). Consult also [5, Lemma 5.10] and [18, Lemma 7], where solvability conditions for the ★-Sylvester equations with $m = n$ were obtained, without considering the details of the solution process. In recent years, an extensive amount of iterative methods based on the conjugate gradient method were studied and developed for solving the generalized T-Sylvester equation

$$AXB + CX^TD = E, \quad A, B, C, D, E, X \in \mathbb{R}^{n\times n}.$$

See, e.g., [15,21–23] and the references cited therein.

In this paper, the (numerical) solution of the ★-Sylvester equation (with $\star = T, H$; the latter indicating the complex conjugate transpose), as well as some related equations, will be studied. Our tools include the (generalized and periodic) Schur,

* Corresponding author.
E-mail addresses: chiang@nfu.edu.tw (C.-Y. Chiang), eric.chu@monash.edu (E.K.-W. Chu), wwlin@am.nctu.edu.tw (W.-W. Lin).

0096-3003/$ - see front matter
doi:10.1016/j.amc.2012.01.065

singular value and QR decompositions [13]. We are mainly interested in the square cases when $m = n$. Other relative work can be found in [1,19,20].

Our interest in the ★-Sylvester equation originates from the solution of the ★-Riccati equation

$$XAX^\star + XB + CX^\star + D = 0$$

from an application related to the palindromic eigenvalue problem [5–7,18] (where eigenvalues appears in reciprocal pairs λ and $\lambda^{-\star}$). The solution of the ★-Riccati equation is difficult and the application of Newton's method is an obvious possibility. The solution of the ★-Sylvester equation is required in the Newton iterative process. Interestingly, the ★-Sylvester and ★-Lyapunov equations behave very differently from the ordinary Sylvester and Lyapunov equations. For example, from Theorem 2.1 below, the ★-Sylvester equation is uniquely solvable only if the generalized spectrum $\sigma(A, B)$ (the set of ordered pairs $\{(a_i, b_i)\}$ representing the eigenvalues of the matrix pencil $A - \lambda B$ or matrix pair (A, B) by $\lambda_i = a_i/b_i$) does not contain λ and $\lambda^{-\star}$ simultaneously, some sort of *apalindromic*[1] requirement. For more details of this application, see Appendix A.

The paper is organized as follows. After this introduction, Section 2 considers the ★-Sylvester equation, in terms of its solvability, the proposed algorithms and the associated error analysis. Section 3 contains several small illustrative examples. Section 4 considers some generalizations of the ★-Sylvester equation—$AXB^\star \pm X^\star = C$, $AXB^\star \pm CX^\star D^\star = E$ and the ★-Lyapunov equation $AX \pm X^\star A^\star = C$. (Similar equations like $AX \pm BX^\star = C$ can be treated similarly and will not be pursued here.) We conclude in Section 5 before describing two applications (in addition to those in [5,6,18]) and a solution formula in terms of generalized inverse from [16] in the Appendices.

2. ★-Sylvester equation

Consider the ★-Sylvester equation

$$AX \pm X^\star B^\star = C, \quad A, B, X \in \mathbb{C}^{n\times n}. \tag{2.1}$$

This includes the special cases of the T-Sylvester equation when ★ = T and the H-Sylvester equation when ★ = H. Justified by associated applications and for efficient exposition, we shall consider ★ = H, T simultaneously, as far as possible.

Remark 2.1. Although it is seemingly simpler to consider only the "+" case in (2.1) and replace B by it negative for the "−" case, this will not be applicable for ★-Lyapunov equations. Also, note that some solvability conditions are dependent on the sign while others are not, thus our results will be more revealing with ± in (2.1). While all these features make our results more general, the (small) price to pay will be the occasional confusing symbols to unfamiliar eyes. If necessary, please concentrate on one of the four cases, e.g. the ★ = T and "−" case, which interests you most.

With the Kronecker product and ★ = T, (2.1) can be written as

$$\mathcal{P}\text{vec}(X) = \text{vec}(C), \mathcal{P} \equiv I \otimes A \pm (B \otimes I)E, \tag{2.2}$$

where vec$X$ stacks the columns of X into a column vector and E is the permutation matrix which maps vec (X) into vec (X^T) [2]; i.e., $E = \sum_{1\leqslant i,j\leqslant n} e_j e_i^T \otimes e_i e_j^T$, where e_i denotes the i-th column of the $n \times n$ identity matrix I_n. The matrix operator on the left-hand-side of (2.2) is $n^2 \times n^2$ and the application of Gaussian elimination and the like will be inefficient. In addition, the approach ignores the structure of the original problem, introducing errors to the solution process unnecessarily.

For the ★ = H case, (2.1) can be rewritten as an expanded T-Sylvester equation:

$$\mathcal{A}\mathcal{X} \pm \mathcal{X}^T\mathcal{B}^T = \mathcal{C}, \quad \mathcal{A}, \mathcal{B}, \mathcal{X} \in \mathbb{R}^{2n\times 2n},$$

where

$$\mathcal{A} \equiv \begin{bmatrix} A_r & A_i \\ -A_i & A_r \end{bmatrix}, \quad \mathcal{B} \equiv \begin{bmatrix} B_r & B_i \\ -B_i & B_r \end{bmatrix}, \quad \mathcal{C} \equiv \begin{bmatrix} C_r & C_i \\ -C_i & C_r \end{bmatrix}, \quad \mathcal{X} \equiv \begin{bmatrix} X_r & X_i \\ -X_i & X_r \end{bmatrix};$$

with the original matrices written in their real and imaginary parts:

$$A = A_r + iA_i, \quad B = B_r + iB_i, \quad C = C_r + iC_i, \quad X = X_r + iX_i.$$

The Kronecker product formulation for T-Sylvester equations can then be applied. Such a formulation will be less efficient for the numerical solution of (2.1), but may be useful as a theoretical tool.

A more efficient approach will be to transform (2.1) by some unitary P and Q, so that (2.1) becomes:

$$PAQ \cdot \overline{Q}^T XP^T \pm PX^T\overline{Q} \cdot Q^T B^T P^T = PCP^T \tag{2.3}$$

or, for ★ = H:

$$PAQ \cdot Q^H XP^H \pm PX^H Q \cdot Q^H B^H P^H = PCP^H. \tag{2.4}$$

[1] Not being palindromic, with "anti-palindromic" already describes something different.

Note that minimum residual and minimum norm solutions are possible with the unitary P and Q. Let $(Q^HA^HP^H, Q^HB^HP^H)$ be in (upper-triangular) generalized Schur form [13]. The transformed equations in (2.3) and (2.4) then have the form

$$\begin{bmatrix} a_{11} & 0^T \\ a_{21} & A_{22} \end{bmatrix}\begin{bmatrix} x_{11} & x_{12}^\star \\ x_{21} & X_{22} \end{bmatrix} \pm \begin{bmatrix} x_{11}^\star & x_{21}^\star \\ x_{12} & X_{22}^\star \end{bmatrix}\begin{bmatrix} b_{11}^\star & b_{21}^\star \\ 0 & B_{22}^\star \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12}^\star \\ c_{21} & C_{22} \end{bmatrix}. \tag{2.5}$$

Multiplying the matrices out, we have

$$a_{11}x_{11} \pm b_{11}^\star x_{11}^\star = c_{11}, \tag{2.6}$$
$$a_{11}x_{12}^\star \pm x_{21}^\star B_{22}^\star = c_{12}^\star \mp x_{11}^\star b_{21}^\star, \tag{2.7}$$
$$A_{22}x_{21} \pm b_{11}^\star x_{12} = c_{21} - x_{11}a_{21} \tag{2.8}$$
$$A_{22}X_{22} \pm X_{22}^\star B_{22}^\star = \widetilde{C}_{22} \equiv C_{22} - a_{21}x_{12}^\star \mp x_{12}b_{21}^\star. \tag{2.9}$$

From (2.6) for ★ = T, we have

$$(a_{11} \pm b_{11})x_{11} = c_{11}. \tag{2.10}$$

With $(a_{11}, b_{11}) \in \sigma(A, B)$, the above equation is uniquely solvable if and only if

$$a_{11} \pm b_{11} \neq 0 \iff \lambda = a_{11}/b_{11} \neq \mp 1. \tag{2.11}$$

Obviously, when $n = 1$, (2.11) is the only condition for the equation to be uniquely solvable.

From (2.6) when ★ = H, we have

$$a_{11}x_{11} \pm \bar{b}_{11}\bar{x}_{11} = c_{11}. \tag{2.12}$$

Let $x_{11} \equiv x_r + ix_i$, $a_{11} \equiv a_r + ia_i$, $b_{11} \equiv b_r + ib_i$ and $c_{11} \equiv c_r + ic_i$. The above equation becomes

$$(a_r + ia_i)(x_r + ix_i) \pm (b_r - ib_i)(x_r - ix_i) = c_r + ic_i$$

or

$$a_rx_r - a_ix_i \pm b_rx_r \mp b_ix_i = c_r, \quad a_rx_i + a_ix_r \mp b_rx_i \mp b_ix_r = c_i.$$

These imply

$$\begin{bmatrix} a_r \pm b_r & -a_i \mp b_i \\ a_i \mp b_i & a_r \mp b_r \end{bmatrix}\begin{bmatrix} x_r \\ x_i \end{bmatrix} = \begin{bmatrix} c_r \\ c_i \end{bmatrix}. \tag{2.13}$$

Denote alternatively $\lambda = a_{11}/b_{11} \in \sigma(A, B)$. The determinant of the matrix operator in (2.13):

$$d = \left(a_r^2 - b_r^2\right) - \left(b_i^2 - a_i^2\right) = |a_{11}|^2 - |b_{11}|^2 \neq 0 \iff |\lambda| \neq 1, \tag{2.14}$$

requiring that no eigenvalue $\lambda \in \sigma(A, B)$ lies on the unit circle. Again, (2.14) is the condition for the equation to be uniquely solvable when $n = 1$.

Another way to solve (2.12) is to write it together with its complex conjugate in the composite form

$$\begin{bmatrix} a_{11} & \pm b_{11}^\star \\ \pm b_{11} & a_{11}^\star \end{bmatrix}\begin{bmatrix} x_{11} \\ x_{11}^\star \end{bmatrix} = \begin{bmatrix} c_{11} \\ c_{11}^\star \end{bmatrix}$$

which produces the equivalent formula

$$x_{11} = \frac{a_{11}^\star c_{11} \mp b_{11}^\star c_{11}^\star}{|a_{11}|^2 - |b_{11}|^2}.$$

From (2.7) and (2.8), we obtain

$$\begin{bmatrix} a_{11}^\star I & \pm B_{22} \\ \pm b_{11}^\star I & A_{22} \end{bmatrix}\begin{bmatrix} x_{12} \\ x_{21} \end{bmatrix} = \begin{bmatrix} \tilde{c}_{12} \\ \tilde{c}_{21} \end{bmatrix} \equiv \begin{bmatrix} c_{12} \\ c_{21} \end{bmatrix} + x_{11}\begin{bmatrix} \mp b_{21} \\ -a_{21} \end{bmatrix}. \tag{2.15}$$

With $a_{11} = b_{11} = 0$, x_{11} will be undetermined. However, (A, B) then forms a singular pencil, $\sigma(A,B) = \mathbb{C}$ and this case will be excluded by (2.22) in Theorem 2.1. If $a_{11} \neq 0$, (2.15) is then equivalent to

$$\begin{bmatrix} a_{11}^\star I & \pm B_{22} \\ 0 & A_{22} - \frac{b_{11}^\star}{a_{11}^\star}B_{22} \end{bmatrix}\begin{bmatrix} x_{12} \\ x_{21} \end{bmatrix} = \begin{bmatrix} \tilde{c}_{12} \\ \hat{c}_{21} \end{bmatrix} \equiv \begin{bmatrix} \tilde{c}_{12} \\ \tilde{c}_{21} \mp \frac{b_{11}^\star}{a_{11}^\star}\tilde{c}_{12} \end{bmatrix}, \tag{2.16}$$

which is uniquely solvable if and only if

$$\det \widetilde{A}_{22} \neq 0, \quad \widetilde{A}_{22} \equiv A_{22} - \frac{b_{11}^\star}{a_{11}^\star}B_{22}, \tag{2.17}$$

or

$$\left(b_{11}^{\star}, a_{11}^{\star}\right) \notin \sigma(A,B) \tag{2.18}$$

with $(a_{11}, b_{11}) \in \sigma(A, B)$. Note that $\widetilde{A}_{22}$ is still lower-triangular, just like A_{22} or B_{22}.

Alternatively, if $b_{11} \neq 0$, (2.15) is equivalent to

$$\begin{bmatrix} 0 & B_{22} - \frac{a_{11}^{\star}}{b_{11}^{\star}} A_{22} \\ b_{11}^{\star} I & \pm A_{22} \end{bmatrix} \begin{bmatrix} x_{12} \\ x_{21} \end{bmatrix} = \begin{bmatrix} \hat{c}_{12} \\ \pm\tilde{c}_{21} \end{bmatrix} \equiv \begin{bmatrix} \pm\tilde{c}_{12} - \frac{a_{11}^{\star}}{b_{11}^{\star}}\tilde{c}_{21} \\ \pm\tilde{c}_{21} \end{bmatrix}, \tag{2.19}$$

which is uniquely solvable if and only if

$$\det \widetilde{B}_{22} \neq 0, \ \widetilde{B}_{22} \equiv B_{22} - \frac{a_{11}^{\star}}{b_{11}^{\star}} A_{22}, \tag{2.20}$$

or (2.18) again. Note that (2.17) and (2.20) are "symmetric", in the sense that when interchanging A and B as well as a_{11} and b_{11} in one equation reproduces the other, and $\widetilde{B}_{22}$ is still lower-triangular as A_{22} and B_{22}. Lastly, (2.9) is of the same form as (2.1) but of smaller size, leading to a recursive algorithm.

Remark 2.2. Interestingly, for the ordinary Sylvester equation $AX - XB = C$, numerical solution will be possible when (A, B) is transformed into the quasi-triangular/triangular form (not necessarily both of the same type) or the cheaper quasi-triangular/Hessenberg form [13]. It is not the case for (2.1) and the ★ somehow alters the behavior of the equation greatly.

Remark 2.3. We can arrange the above solution process into a large quasi-triangular linear system. This enables us to apply the error analysis of triangular linear systems to proposed Algorithms SSylvester and $\text{TSylvester}_\text{R}$ in Section 2.2. Because x_{11} can be solved via a scalar or 2×2 system and X_{22} can be treated recursively, we only need to consider the solution of (2.15) for x_{12} and x_{21}. The equation has the form, for some right-hand-side R_1:

$$\left[\begin{array}{ccc|ccc} r_{11} & & & r_{12} & & \\ & s_{11} & & * & s_{12} & \\ & & \ddots & * & * & \ddots \\ \hline r_{21} & & & r_{22} & & \\ & s_{21} & & * & s_{22} & \\ & & \ddots & * & * & \ddots \end{array}\right] \begin{bmatrix} z_{r1} \\ z_{s1} \\ \vdots \\ z_{r2} \\ z_{s2} \\ \vdots \end{bmatrix} = R_1. \tag{2.21}$$

This is equivalent to a series of 2×2 systems, for known right-hand-sides $R_r, R_s, \ldots$:

$$M_r z_r = R_r, \quad M_s z_s = R_s, \ldots$$

where

$$M_r \equiv [r_{ij}], M_s \equiv [s_{ij}], \ldots; \quad z_r \equiv [z_{r1}, z_{r2}]^T, z_s \equiv [z_{s1}, z_{s2}]^T, \ldots.$$

Consequently, (2.21) is a quasi-lower-triangular linear system with at most 2×2 diagonal blocks. By implication, so is (2.5). This comment still holds when a_{11} and b_{11} are replaced by 2×2 blocks, as in Section 2.1. In that case, the diagonal blocks in the corresponding quasi-triangular matrix will be at most 4×4.

We summarize the solvability condition for (2.1) in the following theorem:

Theorem 2.1. *The ★-Sylvester equation (2.1):*

$$AX \pm X^{\star}B^{\star} = C, \quad A, B \in \mathbb{C}^{n\times n}$$

is uniquely solvable if and only if, for $\{(a_{ii}, b_{ii})\} = \sigma(A, B)$, *the following conditions are satisfied:*

$$a_{ii}a_{jj}^{\star} - b_{ii}b_{jj}^{\star} \neq 0 \quad (\forall i \neq j); \tag{2.22}$$

and, for $\lambda_i \equiv a_{ii}/b_{ii}$ *and all i,*

$$a_{ii} \pm b_{ii} \neq 0 \ (\text{for } \star = T), \ |\lambda_i| \neq 1 \ (\text{for } \star = H). \tag{2.23}$$

Remark 2.4. Condition (2.11) or (2.14) are actually contained in (2.22) when $i = j = 1$. The corresponding conditions have to be restated in (2.23) in Theorem 2.1 for an arbitrary ordering of the eigenvalues $\{(a_{ii}, b_{ii})\}$ or $\{\lambda_i\}$. In terms of the traditional representation of the eigenvalues $\lambda_i \equiv a_{ii}/b_{ii}$, (2.22) means that $\lambda_j \neq \lambda_i^{-\star}$ $(i \neq j)$, and (2.23) means that $\lambda_i \neq \mp 1$ (★ = T) or $|\lambda_i| \neq 1$ (★ = H). Consequently, $\lambda = \pm 1$ can be an eigenvalue of (A, B) but must be simple for the corresponding T-Sylvester equation to be uniquely solvable. As expected, (2.22) or (2.23) imply that (A, B) has to be regular for (2.1) to be uniquely solvable. When ★ = T, it is worth to mention that conditions (2.22) and (2.23) are equivalent as Lemma 5.10 in [5].

The solution process in this subsection is summarized below, using MATLAB-like notation $X(i{:}j, k)$ (or $X(k, i{:}j)$) for the row (column), consisting of the intersection of row i to j (or k) and column k (or i to j):

Algorithm SSylvester

Input: Given the matrix $A, B, C \in \mathbb{C}^{n\times n}$, τ (a small tolerance).
Output: The unique solution X of $AX \pm X^{\star} B^{\star} = C$.

```
Compute the lower-triangular Schur form (PAQ, PBQ) (if necessary)
  (A, B, C) ← (PAQ, PBQ, PCP★)
for i = 1,2,...,n − 1
  x_ii = c_11/(a_11 + b_11)
  if dim(A) = 1 or |a_11|^2 + |b_11|^2 ⩽ τ, exit
  if |a_11| ⩾ |b_11|, then
    if λ ∈ diag(Ã_22) and |λ| < τ, exit
      else compute x_21 = Ã_22^{-1} ĉ_21, x_12 = (c̃_12 ∓ B_22 x_21)/a★_11 (from (2.16))
    else if λ ∈ diag(B̃_22) and |λ| < τ, then exit
      else compute x_21 = B̃_22^{-1} ĉ_12, x_12 = (±c̃_21 ∓ A_22 x_21)/b★_11 (from (2.19))
    X(i+1 : n, i) = x_21, X(i, i+1 : n) = x_12^T
    C ← C(2 : n−i+1, 2 : n−i+1) − A_21 X(i, i+1 : n) − X(i, i+1 : n)★ B★_21
    A ← A(2:n − i + 1,2:n − i + 1), B ← B(2:n − i + 1,2:n − i + 1)
end for
x_nn = c_11/(a_11 + b_11)
X ← Q X P̄ (★ = T), or X ← Q X P (★ = H)
```

Let the operation count of the Algorithm SSylvester be $f(n)$ complex flops, on top of the $66n^3$ complex flops for the QZ procedure [13] for the generalized Schur decomposition of (A, B). The count in $f(n)$ involves the solution of (2.9) (for $f(n-1)$ complex flops) and (2.16) or (2.19). This involves forming and inverting A_{22} or B_{22} (n^2 flops), computing x_{12} ($\frac{1}{2}n^2$ flops) and forming C_{22} ($2n^2$). Thus $f(n) = f(n-1) + \frac{7}{2}n^2 + O(n)$. This implies that $f(n) = \frac{7}{6}n^3 + O(n^2)$ and the total operation count for Algorithm SSylvester is $67\frac{1}{6}n^3 + O(n^2)$ complex flops.

From the above analysis and Theorem 2.1, the condition of (2.1) will be bad if the separation $\lambda_i\lambda_j^{\star} - 1$ is close to zero (or when the assumption for unique solvability is nearly violated). The same conclusion can also be drawn from the analogous analysis in Section 2.1 below. For error analysis, see Section 2.2 for more details.

2.1. The real case

When A, B and C are all real, the solution X, judging from (2.2), will be real. To guarantee a real solution X, the generalized real Schur form [13] for (A, B) has to be used. The transformed equation in (2.3) or (2.4) has the form

$$\begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix}\begin{bmatrix} X_{11} & X_{12}^T \\ X_{21} & X_{22} \end{bmatrix} \pm \begin{bmatrix} X_{11}^T & X_{21}^T \\ X_{12} & X_{22}^T \end{bmatrix}\begin{bmatrix} B_{11}^T & B_{21}^T \\ 0 & B_{22}^T \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12}^T \\ C_{21} & C_{22} \end{bmatrix}$$

or

$$A_{11}X_{11} \pm X_{11}^T B_{11}^T = C_{11}, \tag{2.24}$$

$$A_{11}X_{12}^T \pm X_{21}^T B_{22}^T = \widetilde{C}_{12}^T \equiv C_{12}^T \mp X_{11}^T B_{21}^T, \tag{2.25}$$

$$A_{22}X_{21} \pm X_{12}B_{11}^T = \widetilde{C}_{21} \equiv C_{21} - A_{21}X_{11}, \tag{2.26}$$

$$A_{22}X_{22} \pm X_{22}^T B_{22}^T = \widetilde{C}_{22} \equiv C_{22} - A_{21}X_{12}^T \mp X_{12}B_{21}^T, \tag{2.27}$$

where A_{11} and B_{11} may be 1×1 or 2×2 real matrix. The 1×1 case will be trivial as in (2.6) and the 2×2 case can be solved using Kronecker products. The theory leading to the conditions in (2.11) and (2.22) from the complex Schur form still holds. We shall assume that A_{11} and B_{11} are not scalar in the rest of this subsection.

Again, the Kronecker product can be applied to (2.25) and (2.26). A better approach is to consider $(2.25)^T$ and (2.26):

$$\pm B_{22}X_{21} + X_{12}A_{11}^T = \widetilde{C}_{12}, \quad A_{22}X_{21} \pm X_{12}B_{11}^T = \widetilde{C}_{21}.$$

A linear combination of these equations will be

$$(\beta A_{22} \pm \alpha B_{22})X_{21} + X_{12}\left(\alpha A_{11}^T \pm \beta B_{11}^T\right) = \alpha\widetilde{C}_{12} + \beta\widetilde{C}_{21}. \tag{2.28}$$

Assuming regularity of the pencil (A, B), there exists real α and β such that $\alpha A_{11}^T \pm \beta B_{11}^T$ is nonsingular (or well-conditioned). We then have

$$X_{12} = -(\beta A_{22} \pm \alpha B_{22})X_{21}\left(\alpha A_{11}^T \pm \beta B_{11}^T\right)^{-1} + \widehat{C}_{12}X_{12} \tag{2.29}$$

with $\widehat{C}_{12} \equiv (\alpha\widetilde{C}_{12} + \beta\widetilde{C}_{21})\left(\alpha A_{11}^T \pm \beta B_{11}^T\right)^{-1}$. Substitute X_{12} in (2.29) into (2.26), we have a generalized Sylvester equation [8] for X_{21}:

$$A_{22}X_{21} - (\alpha B_{22} \pm \beta A_{22})X_{21}(\alpha A_{11}^T \pm \beta B_{11}^T)^{-1}B_{11}^T = \widetilde{C}_{21} \mp \widehat{C}_{12}^T B_{11}^T. \tag{2.30}$$

Remark 2.5. Alternatively from (2.28), one can choose α and β to optimize the condition of $\beta A_{22} \pm \alpha B_{22}$, express X_{21} in terms of X_{12} and obtain a generalized Sylvester equation in X_{12} analogous to (2.30).

From [8], (2.30) is uniquely solvable when there is no intersection of the spectra $\sigma(A_{22}, \alpha B_{22} \pm \beta A_{22})$ and $\sigma\left(B_{11}^T, \alpha A_{11}^T \pm \beta B_{11}^T\right)$. Let (A_{11}, B_{11}) and (A_{22}, B_{22}) be transformed into generalized real Schur forms with diagonal elements (α_i, β_i) and (α_j, β_j) respectively. For $\alpha \neq 0$, the condition for the equation to be uniquely solvable is

$$\frac{\alpha_j}{\alpha\beta_j \pm \beta\alpha_j} \neq \frac{\beta_i}{\alpha\alpha_i \pm \beta\beta_i} \iff \alpha_i\alpha_j \neq \beta_i\beta_j,$$

exactly condition (2.22). The same conclusion is reached when $\alpha = 0$, which implies that B_{11} is invertible, and X_{12} in (2.29) should then be substituted into $(2.25)^T$ to produce a similar generalized Sylvester equation for X_{21}:

$$X_{21} - A_{22}X_{21}B_{11}^{-T}A_{11}^T = \pm\widetilde{C}_{12} + \widehat{C}_{12}A_{11}^T.B_{22} \tag{2.31}$$

Also X_{12} is retrievable from (2.29) in a numerical stable and efficient manner. Note that the matrix operator $\alpha A_{11}^T \pm \beta B_{11}^T$ in (2.29) is 2×2 with (α, β) controlling its condition. In (2.30), A_{22} and B_{22} are quasi-lower-triangular with $(\alpha A_{11}^T \pm \beta B_{11}^T)^{-1}B_{11}^T$ being at most 2×2, enabling X_{21} to be calculated as in the generalized Bartels–Stewart algorithm in [8]. (For illustration, let us consider (2.31). With B_{22} and A_{22} being quasi-lower-triangular and $B_{11}^{-T}A_{11}^T$ being 2×2, the first row of X_{21} can be computed, leaving a smaller but similar system. This can then be solved recursively and similarly.) A slightly more efficient alternative will be to consider the rows of (2.30) consecutively from the top, solving a 2×2 system for each row of X_{21}. Eq. (2.31) can be solved analogously, also one row at a time.

We can then solve recursively (2.27), a smaller equation similar to (2.9). We summarize the procedure in this subsection in the following algorithm, with the subscripts "R" for "Real".

Algorithm TSylvester$_\mathbf{R}$

Input: Given the matrix $A, B, C \in \mathbb{R}^{n\times n}$, τ (a small tolerance).
Output: The unique solution X of $AX \pm X^T B^T = C$.
`Compute` the real Schur form (PAQ, PBQ) (if necessary)
  $(A, B, C) \leftarrow (PAQ, PBQ, PCP^T)$
`Solve` (2.24) for X_{11}; `if` fail, `exit`
`if` last block reached with $n = 1, 2$, `exit`
`if` A_{11} and B_{11} are scalar, `then`
  `Solve` (2.25) and (2.26) for X_{12} and X_{21} as in **Algorithm SSylvester**
  `if` fail, `exit`
  `else Solve` (2.30) or (2.31) with appropriate α, β for X_{21} row-wise,
    using Gaussian elimination on the 2×2 systems; `if` fail, `exit`
    `Retrieve` X_{12} from (2.29)
`Apply` **Algorithm TSylvester$_\mathbf{R}$** to
    $A_{22}X_{22} \pm X_{22}^T B_{22}^T = C_{22}$ (c.f. (2.27)), $n \leftarrow n-1$ or $n-2$
$X \leftarrow QXP$

The operation count of Algorithm TSylvester$_R$ is approximately equal to $67\frac{1}{6}n^3$ flops, similar to Algorithm SSylvester and overwhelmed by the initial QZ process.

Remark 2.6. We have not written Algorithm TSylvester$_R$ as in Algorithm SSylvester, without calling itself. As A_{11} and B_{11} can be 1×1 or 2×2, the alternative algorithm will be repetitive for the different cases.

Remark 2.7. Similar to Remark 2.3, Algorithms TSylvester$_R$ is equivalent to solving quasi-lower-triangular linear systems after the initial QZ step. The equations for the scalar (or 2×2) X_{11} can be written as a 2×2 (or 8×8) linear system for the real and imaginary parts of the elements of X_{11}. For X_{12} and X_{21}, expanding (2.25) and (2.26) using the Kronecker product yields a linear system with matrix operator

$$\begin{bmatrix} I_{n-2}\otimes A_{11} & B_{22}\otimes I_2 \\ I_{n-2}\otimes B_{11} & A_{22}\otimes I_2 \end{bmatrix}.$$

This has the same form as the one in (2.21), except the elements may be 2×2 blocks, producing a series of 4×4 linear systems. Similar arguments as those in Remark 2.3 thus follows.

2.2. Error analysis

We shall discuss the error associated with Algorithms TSylvester and TSylvester$_R$, following the development in [14, Chapter 16] and [12].

2.2.1. Residual

As indicated in Remarks 2.3 and 2.7, Algorithms SSylvester and SSylvester$_R$ can be arranged into quasi-triangular linear systems. We can then apply the error analysis for triangular linear systems in [14, Theorem 8.5] to obtain

$$\|R\|_F \equiv \|C - (A\widehat{X} \pm \widehat{X}^\star B^\star)\|_F \leqslant c_n u(\|A\|_F + \|B\|_F)\|\widehat{X}\|_F \tag{2.32}$$

for a computed solution $\widehat{X}$ from our algorithms, where c_n is a constant dependent on n and u is the unit round-off (typically $O(10^{-16})$), when the condition of the 2×2, 4×4 or 8×8 linear systems in (2.30), (2.31) and Remarks 2.3 and 2.7 are not bad. Note that the QZ transformation of (A, B) is backward stable, similar to the QR process in [14, Eq. (16.9)]. Consequently, the relative residual is bounded by a modest multiple of the unit round-off u. See the collaborating numerical examples in Section 3.

2.2.2. Backward error

Like for ordinary Sylvester equations, the numerical solution of (2.1) is not backward stable in general. Similar to [14, §16.2] and with "δ" indicating perturbation, we can define the normwise relative backward error of an approximate solution Y by

$$\eta(Y) \equiv \min\{\epsilon : (A+\delta A)Y \pm Y^\star(B+\delta B)^\star = C+\delta C,$$
$$\|\delta A\|_F \leqslant \epsilon\alpha, \quad \|\delta B\|_F \leqslant \epsilon\beta, \quad \|\delta C\|_F \leqslant \epsilon\gamma\}, \tag{2.33}$$

where $\alpha \equiv \|A\|_F$, $\beta \equiv \|B\|_F$ and $\gamma \equiv \|C\|_F$. With $Y = U\Sigma V^H$ in singular value decomposition (SVD) [13], the $Y^\star$ terms do not affect the analysis in [14, §16.2]. With $\Sigma = \mathrm{diag}\{\sigma_1,\ldots,\sigma_n\}$, it can be shown that

$$\eta(Y) \leqslant \mu\frac{\|R\|_F}{(\alpha+\beta)\|Y\|_F+\gamma}, \tag{2.34}$$

where the residual $R \equiv C - (AY \pm Y^\star B)$ and

$$\mu \equiv \frac{(\alpha+\beta)\|Y\|_F+\gamma}{\left[(\alpha^2+\beta^2)\|X^{-1}\|_2^{-2}+\gamma^2\right]^{1/2}}.$$

Consequently, $\eta(Y)$ can be large when Y is ill-conditioned, and a small residual R does not always imply a small backward error $\eta(Y)$. This phenomenon has been observed in Example 3.3, where Y is ill-conditioned. However, from our experience, severely backward unstable $\star$-Sylvester equations are rare and have to be artificially constructed. This suggests that our algorithms may well be conditionally backward stable. Similar to the Sylvester equation [14, Section 16.2], we do not know the conditions under which a $\star$-Sylvester equation has a well-conditioned solution.

2.2.3. Perturbation and practical error bounds

For perturbation, the usual results for linear systems apply. In terms of the $\star$-Sylvester equation (2.1), consider the perturbed equation

$$(A+\delta A)(X+\delta X) \pm (X+\delta X)^\star(B+\delta B)^\star = C+\delta C.$$

Define the $\star$-Sylvester operator

$$S(X) \equiv AX \pm X^\star B^\star,$$

we then obtain

$$S(\delta X) = \delta C - \delta AX \mp X^\star\delta B^\star - \delta A\delta X \mp \delta X^\star\delta B^\star.$$

Application of norm gives rise to

$$\|\delta X\|_F \leqslant \|S^{-1}\|_F\{\|\delta C\|_F + (\|\delta A\|_F + \|\delta B\|_F)(\|X\|_F + \|\delta X\|_F)\}.$$

When $\|\delta S\|_F \equiv \|\delta A\|_F + \|\delta B\|_F$ is small enough so that $1 \geqslant \|S^{-1}\|_F \cdot \|\delta S\|_F$, we can rearrange the above result to

$$\frac{\|\delta X\|_F}{\|X\|_F} \leqslant \frac{\|S^{-1}\|_F}{1-\|S^{-1}\|_F \cdot \|\delta S\|_F}\left(\frac{\|\delta C\|_F}{\|X\|_F} + \|\delta S\|_F\right).$$

With $\|C\|_F = \|S(X)\|_F \leqslant \|S\|_F \cdot \|X\|_F$ and the condition number $\kappa(S) \equiv \|S\|_F \cdot \|S^{-1}\|_F$, we arrive at the standard perturbation result

$$\frac{\|\delta X\|_F}{\|X\|_F} \leqslant \frac{\kappa(S)}{1-\kappa(S)\cdot \|\delta S\|_F/\|S\|_F}\left(\frac{\|\delta C\|_F}{\|C\|_F} + \frac{\|\delta S\|_F}{\|S\|_F}\right).$$

Thus the relative error in X is controlled by those in A, B and C, magnified by the condition number $\kappa(S)$.

As indicated in [14, Section 16.4], practical error bounds can be estimated, just like for other linear matrix equations. Several applications of the solution algorithm will be required. More work has to be done along this direction.

2.3. An alternative formulation

We can consider the sum/difference of (2.1) and its $\star$, producing

$$(A+B)X + X^\star(A+B)^\star = C + C^\star, \quad (A-B)X - X^\star(A-B)^\star = C - C^\star. \tag{2.35}$$

The pair of equations represent the symmetric/Hermitian and skew-symmetric/Hermitian parts of (2.1) and can be solved using the generalized Schur form of $(A + B, A - B)$. Identical solvability condition as (2.22) can be derived. In terms of the eigenvalues $\tilde{\lambda}_i \in \sigma(A+B, A-B)$, (2.1) and (2.35) are uniquely solvability if and only if $\tilde{\lambda}_i + \tilde{\lambda}_j \neq 0$ or $\lambda_i\lambda_j \neq 1$, with $\tilde{\lambda}_i = (\lambda_i + 1)/(\lambda_i - 1)$ for some $\lambda_i \in \sigma(A, B)$. It is easy to see that mapping between (A, B) and $(A + B, A - B)$ corresponds to some (inverse) Cayley transformations.

In [11], a formula for the solution X of (2.1) (for $\star = T$ and the "+" case) was derived using the first equation in (2.35) only, throwing away the information in the second equation. We cannot see how the formula can be correct using only half the information of (2.1) in the first half of (2.35). In the extreme case with $A = -B$, the first equation in (2.1) will be degenerate and the solution X will be totally free. Anyway, X is a solution of (2.1) if and only if it is also a solution of (2.35), but a solution of half of (2.35) in general does not satisfy (2.1).

3. Numerical examples

In this section, we apply Algorithm SSylvester (denoted by ASS) and, for the lack of alternative algorithms, the Kronecker product approach in (2.2) (denoted by KRP) to some examples for illustrative and comparative purposes. All computations were performed in MATLAB/version 7.5 on a PC with an Intel Pentium-IV 4.3GHZ processor and 3 GB main memory, using IEEE double-precision.

Example 3.1. We choose $\widehat{A}, \widehat{B} \in \mathbb{R}^{n\times n}$ to be real lower-triangular matrices with given diagonal elements (specified by $a, b \in \mathbb{R}^n$) and random strictly lower-triangular elements. They are then reshuffled by the orthogonal matrices $Q, Z \in \mathbb{R}^{n\times n}$ to form $(A, B) = (Q\widehat{A}Z, Q\widehat{B}Z)$. In MATLAB [17] commands, we have

$$\widehat{A} = tril(randn(n), -1) + diag(a), \quad \widehat{B} = tril(randn(n), -1) + diag(b), \quad C = randn(n).$$

To guarantee condition (2.22), let $b = randn(n, 1)$, $a = 2b$. In Table 3.1, we list the CPU time ratios of the ASS and the KRP approaches as well as the corresponding residuals and their ratios, with increasing dimensions n = 16, 20, 25, 30, 35, 40. Note that the operation counts for the ASS and KRP methods are approximately $67n^3$ and $\frac{2}{3}n^6$ flops respectively (the latter for the LU decomposition of the $n^2 \times n^2$ matrix in (2.2)). The results in Table 3.1 show that the advantage of ASS over KRP in CPU time grows rapidly as n increases, as predicted by. Even with better management of sparsity or parallelism, the $O(n^6)$ operation count makes the KRP approach uncompetitive even for moderate size n. The residuals from ASS is also better than that from KRP, as (2.2) is solved by Gaussian elimination in an unstructured way. See the other examples for more comparison of the residuals of ASS and KRP.

Example 3.2. With the same construction as in Example 3.1 and $n = 2$, let $a = [\alpha + \epsilon, \beta]^T$, $b = [\beta, \alpha]^T$. Here α, β are two randomly numbers greater than 1, with the spectral set $\sigma(A, B) = \left\{\frac{\alpha+\varepsilon}{\beta}, \frac{\beta}{\alpha}\right\}$, and $|\lambda_1\lambda_2 - 1| = \frac{\varepsilon}{\alpha}$. Judging from (2.22) and (2.1) has worsening condition as ε decreases. We report a comparison of absolute residuals for the ASS and KRP approaches for $\varepsilon = 10^{-1}, 10^{-3}, 10^{-5}, 10^{-7}$ and 10^{-9} in Table 3.2. The results show that if (2.2) is solved by Gaussian elimination, its residual will be larger than that for ASS especially for smaller ε. Note that the size of X (the last column in Table 3.2) reflects partially the condition of (2.1), as indicated in (2.32). The residuals will be worsen for large values of $\|X\|_F$, with the quotient of res (ASS) and $\|X\|$ approximately equal to the unit round-off u. The KRP approach copes less well than the ASS approach for ill-conditioned problems.

Example 3.3. With $n = 2$ and let $Q \in \mathbb{R}^{n\times n}$ be orthogonal and the exact solution be X_e, where

Table 3.1
Results for Example 3.1.

n	$\frac{t_{KRP}}{t_{ASS}}$	Res (ASS)	Res (KRP)	$\frac{\text{Res(KRP)}}{\text{Res(ASS)}}$
16	1.00e+00	1.8527e−17	2.1490e−17	1.16
25	1.31e+01	2.3065e−17	2.8686e−17	1.24
30	2.61e+01	3.1126e−18	5.7367e−18	2.20
35	6.48e+01	7.0992e−18	1.2392e−17	1.75
40	1.05e+02	1.7654e−18	6.4930e−18	3.68

Table 3.2
Results for Example 3.2.

ϵ	Res (ASS)	Res (KRP)	$\frac{\text{Res(KRP)}}{\text{Res(ASS)}}$	$O(\|X\|)$
1.0e−1	2.0673e−15	2.4547e−15	1.19	10^1
1.0e−3	8.6726e−13	4.3279e−13	0.50	10^3
1.0e−5	2.3447e−12	2.4063e−12	1.03	10^3
1.0e−7	5.9628e−10	1.1786e−09	1.98	10^6
1.0e−9	5.8632e−08	3.4069e−07	5.81	10^8

$$X_e \equiv Q^T \begin{bmatrix} 10^{-m} & 0 \\ 0 & 10^m \end{bmatrix} Q, \quad A = \begin{bmatrix} randn & 0 \\ randn & 10^{-m} \end{bmatrix} Q, \quad B = \begin{bmatrix} randn & 0 \\ randn & 2 * 10^{-m} \end{bmatrix} Q$$

and $C = AX_e + X_e^T B^T$. Solving the corresponding T-Sylvester equation by Algorithm SSylvester produces the results in Table 3.3, using symbols from Section 2.2. The column for the backward error $\eta(Y)$ (estimated using the bound (2.34)) confirms that our algorithm is not numerically backward stable. The problem is increasingly ill-conditioned for increasing values of m and the large values of μ worsen the backward errors $\eta(Y)$, although the relative residuals RRes (ASS)=Res (ASS)/$\|X\|$ are of machine accuracy. On the other hand, from our experience, badly behaved examples are rare and have to be artificially constructed.

4. Related equations

4.1. Generalized ★-Sylvester equation I

Consider the more general version of the ★-Sylvester equation (2.1):

$$AXB^\star \pm X^\star = C \tag{4.1}$$

with $A, B^\star, X^\star \in \mathbb{C}^{m\times n}$ and $m \neq n$. The generalized Kronecker canonical form [9,10] for $(A, B^\star)$ may be used to analyze and solve the equation. We shall not pursuit this line of attack further.

For $A, B, C \in \mathbb{C}^{n\times n}$, the equation is equivalent to the ★-Sylvester equation in Section 2 when either A or B is nonsingular. In general, consider the periodic Schur or PQZ decomposition [3] for $B^H A^H$ so that $(Q^H A^H P^H, PB^H Q)$ is in upper triangular form.

Consider the transformed equation, for ★ = H:

$$PAQ \cdot Q^H X P^H \cdot PB^H Q \pm PX^H Q = PCQ$$

or for ★ = T:

$$PAQ \cdot Q^H X P^T \cdot \overline{P} B^T \overline{Q} \pm PX^T \overline{Q} = PC\overline{Q}.$$

The case when (A, B) are real and ★ = T with a real PQZ decomposition is similar but will be skipped over.

With $(Q^H A^H P^H, PB^H Q)$ or $(Q^H A^H P^H, \overline{P} B^T \overline{Q})$ being upper-triangular, the transformed equations look like

Table 3.3
Results for Example 3.3.

m	Res (ASS)	RRes (ASS)	$\frac{\|X_{ASS}-X_e\|}{\|X_e\|}$	$O(\|X\|)$	μ	$\eta(X_{ASS})$
0	1.0129e−16	10^{-16}	2.6624e−16	10^0	3.2440e+00	2.7169e−16
2	1.5268e−14	10^{-16}	2.0519e−15	10^2	9.7188e+01	5.8991e−15
4	2.4170e−12	10^{-16}	5.0599e−13	10^4	7.3715e+03	1.0410e−12
6	1.6955e−10	10^{-16}	2.4933e−11	10^6	9.0423e+05	6.8488e−11
8	3.7545e−09	10^{-17}	2.7786e−09	10^8	8.4485e+07	1.2658e−09

$$\begin{bmatrix} a_{11} & 0^T \\ a_{21} & A_{22} \end{bmatrix}\begin{bmatrix} x_{11} & x_{12}^\star \\ x_{21} & X_{22} \end{bmatrix}\begin{bmatrix} b_{11}^\star & b_{21}^\star \\ 0 & B_{22}^\star \end{bmatrix} \pm \begin{bmatrix} x_{11}^\star & x_{21}^\star \\ x_{12} & X_{22}^\star \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12}^\star \\ c_{21} & C_{22} \end{bmatrix}.$$

We then have

$$a_{11}b_{11}^\star x_{11} \pm x_{11}^\star = c_{11}, \tag{4.2}$$

$$a_{11}x_{12}^\star B_{22}^\star \pm x_{21}^\star = c_{12}^\star - a_{11}x_{11}b_{21}^\star, \tag{4.3}$$

$$b_{11}^\star A_{22}x_{21} \pm x_{12} = c_{21} - b_{11}^\star x_{11}a_{21}, \tag{4.4}$$

$$A_{22}X_{22}B_{22}^\star \pm X_{22}^\star = C_{22} - x_{11}a_{21}b_{21}^\star - A_{22}x_{21}b_{21}^\star - a_{21}x_{12}^\star B_{22}^\star. \tag{4.5}$$

Inspection of (4.2)–(4.5) shows the equation is uniquely solvable if and only if

$$a_{ii}b_{ii}^\star \neq \mp 1, \quad a_{ii}^\star b_{ii}^\star a_{jj}b_{jj} \neq 1 \ (\forall i \neq j); \tag{4.6}$$

analogous to (2.22) and (2.23) and a special case of (4.10). Algorithms can easily be constructed from (4.2)–(4.5) but will be ignored here.

4.2. Generalized ★-Sylvester equation II

Consider the more general version of the ★-Sylvester equations (2.1) and (4.1):

$$AXB^\star \pm CX^\star D^\star = E. \tag{4.7}$$

For $A, B, C, D, E \in \mathbb{C}^{n\times n}$, the equation is equivalent to the ★-Sylvester equation in Section 2, when A and D (or B and C) are nonsingular. In general, we can transform the equation to, for ★ = H:

$$PAR \cdot R^H XS \cdot S^H B^H Q \pm PCS \cdot S^H X^H R \cdot R^H D^H Q = PEQ \tag{4.8}$$

or, for ★ = T:

$$PA\overline{R} \cdot R^T XS \cdot S^H B^T Q \pm PC\overline{S} \cdot S^T X^T R \cdot R^H D^T Q = PEQ. \tag{4.9}$$

These equations have the form

$$\widetilde{A}\widetilde{X}\widetilde{B}^\star \pm \widetilde{C}\widetilde{X}^\star \widetilde{D}^\star = \widetilde{E}.$$

The transformation can be realized using the periodic Schur or PQZ decomposition [3] for $B^{-1}DA^{-1}C$ (or other similar formations), where P, Q, R and S are unitary, and $\widetilde{A}$, $\widetilde{B}$, $\widetilde{C}$ and $\widetilde{D}$ are (quasi-) lower-triangular (with diagonal elements α_i, β_i, γ_i and δ_i, respectively). Consequently, similar solution procedure as in Section 2 applies, with both minimum norm and minimum residual solutions feasible. The transformed equations give rise to equations in the form, for $i \neq j$:

$$(\alpha_i\beta_i^\star \pm \gamma_i\delta_i^\star)x_{ii} = \tilde{e}_{ii}, \quad \begin{bmatrix} \alpha_i\beta_j^\star & \pm\gamma_i\delta_j^\star \\ \pm\gamma_j\delta_i^\star & \alpha_j\beta_i^\star \end{bmatrix}\begin{bmatrix} x_{ij} \\ x_{ji} \end{bmatrix} = \begin{bmatrix} \tilde{e}_{ij} \\ \tilde{e}_{ji} \end{bmatrix}$$

for some known $\tilde{e}_{ii}$, $\tilde{e}_{ij}$ and $\tilde{e}_{ji}$, with x_{ii} and x_{ij} solved in the correct order. The equation will then be uniquely solvable if and only if

$$\alpha_i\beta_i^\star \pm \gamma_i\delta_i^\star \neq 0, \quad \alpha_i\alpha_j\beta_i^\star\beta_j^\star \neq \gamma_i\gamma_j\delta_i^\star\delta_j^\star \ (\forall i \neq j); \tag{4.10}$$

conditions more general than but similar to (2.22) and (2.23), or (4.6).

4.3. ★-Lyapunov equation

Consider the ★-Lyapunov equation

$$AX \pm X^\star A^\star = C, \quad A \in \mathbb{C}^{n\times n}.$$

(The more general $AXB^\star \pm B^\star X^\star A^\star = C$ can be treated similarly.) Here we assumed that this system is consistent. With unitary P and Q, the equation can be transformed to, for ★ = T:

$$PAQ \cdot Q^H XP^T \pm PX^T\overline{Q} \cdot Q^T A^T P^T = PCP^T$$

or, for ★ = H:

$$PAQ \cdot Q^H XP^H \pm PX^H Q \cdot Q^H A^H P^H = PCP^H.$$

Obviously, ★-Lyapunov equations are special cases of ★-Sylvester equations, with $B = A$, and associated solvability conditions in Theorem 2.1 (with $b_{kk} = a_{kk}$). We recognize that ★-Lyapunov equations are not uniquely solvable, with (2.22) being violated. However, we can still solve these equations in the traditional, least squares or minimum norm sense, as seen below.

Note that the unitary transformation of A allows for minimum norm or residual solutions of the equations. We can choose P and Q from the SVD of A. This is more suited to the case when A is rectangular and this line of attack will be pursued later. For a square A, we can choose $Q = P^H$ using the Schur decomposition of A, solving the equation in a similar fashion as in Section 2. The transformed equation has the form

$$\begin{bmatrix} a_{11} & 0^T \\ a_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_{11} & x_{12}^\star \\ x_{21} & X_{22} \end{bmatrix} \pm \begin{bmatrix} x_{11}^\star & x_{21}^\star \\ x_{12} & X_{22}^\star \end{bmatrix} \begin{bmatrix} a_{11}^\star & a_{21}^\star \\ 0 & A_{22}^\star \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12}^\star \\ \pm c_{12} & C_{22} \end{bmatrix}.$$

Note that the form of the right-hand-side came from the obvious consistency condition $C = \pm C^\star$, implying that the diagonal elements $c_{ii} = 0$ in C when the $\star = T$ and "$-$" case.

Multiply the matrices out, we have

$$a_{11}x_{11} \pm a_{11}^\star x_{11}^\star = c_{11}, \tag{4.11}$$

$$a_{11}x_{12}^\star \pm x_{21}^\star A_{22}^\star = \tilde{c}_{12}^\star \equiv c_{12}^\star \mp x_{11}^\star a_{21}^\star, \tag{4.12}$$

$$A_{22}X_{22} \pm X_{22}^\star A_{22}^\star = \widetilde{C}_{22} \equiv C_{22} - a_{21}x_{12}^\star \mp x_{12}a_{21}^\star. \tag{4.13}$$

Because of the (anti-) symmetry of the $\star$-Lyapunov equation, we only need to consider the above three equations, with the fourth containing redundant information.

For $\star = T$, x_{11} is free for the "$-$" case, with $c_{11} = 0$ for consistency. For the "+" case, $x_{11} = \frac{c_{11}}{2a_{11}}$ when the eigenvalue $\lambda_1 = a_{11} \in \sigma(A)$ is nonzero. For $\star = H$, we have the underdetermined equation $\Re e(a_{11}x_{11}) = c_{11}$ (for the "+" case) or $\Im m(a_{11}x_{11}) = 0$ (for the "$-$" case). For x_{12} and x_{21}, we have the equation

$$\begin{bmatrix} a_{11}^\star I & A_{22} \end{bmatrix} \begin{bmatrix} x_{12} \\ x_{21} \end{bmatrix} = \tilde{c}_{12}$$

which is underdetermined when $\tilde{c}_{12}$ is in the span of $[a_{11}^\star I, A_{22}]$ (always holds if A is nonsingular).

The equation for X_{22} is smaller but similar to the original $\star$-Lyapunov equation.

4.3.1. Symmetric/hermitian solution

With the transformed equations, for $\star = T$:

$$PAP^H \cdot PXP^T \pm PX^TP^T \cdot \overline{P}A^TP^T = PCP^T$$

or for $\star = H$:

$$PAP^H \cdot PXP^H \pm PX^HP^H \cdot PA^HP^H = PCP^H,$$

we can impose the (anti-) symmetry constraint $X^\star = \pm X$. Eqs. (4.11)–(4.13) then imply similar equations for x_{11} and X_{22} as in the non-symmetric/Hermitian case. For $x_{12} = x_{21}$ (and similarly for the anti-symmetric/Hermitian case), (4.12) becomes

$$(a_{11}^\star I \pm A_{22})x_{12} = \tilde{c}_{12},$$

retrieving the solvability condition for the ordinary Sylvester/Lyapunov equation. This requires the eigenvalues λ_j and λ_j of A to satisfy $\lambda_i^\star \pm \lambda_j \neq 0$. When $i = j$ and $\star = T$, this indicates that we cannot have zero eigenvalues for the "+" case and the "$-$" case gives rise to an undetermined x_{11}, with $c_{11} = 0$ from the anti-symmetric C. When $i = j$ and $\star = H$, no eigenvalue λ_i can be purely imaginary/real.

4.3.2. Rectangular A

The T-Lyapunov equation with rectangular A has been studied in [4] using generalized inverse (which can only be realized using the SVD). Please consult [4] for solvability conditions and the formula for the general solution. Here we construct the solution, and implicitly derive the solvability conditions, using the SVD. In the next subsection, the cheaper QR decomposition [13] is used instead to derive the same solution.

When A is rectangular, the SVD of A:

$$A = UDV^H = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1 & V_2 \end{bmatrix}^H \tag{4.14}$$

gives rise to the transformed T-Lyapunov equation:

$$UDV^HX \pm X^T\overline{V}D^TU^T = C \iff D(V^HX\overline{U}) \pm (U^HX^T\overline{V})D = U^HC\overline{U}$$

or the transformed H-Lyapunov equation:

$$UDV^HX \pm X^HVD^TU^H = C \iff D(V^HXU) \pm (U^HX^HV)D = U^HCU.$$

We then have

$$\begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix} \pm \begin{bmatrix} X_{11}^\star & X_{21}^\star \\ X_{12}^\star & X_{22}^\star \end{bmatrix}\begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} \\ \pm C_{12}^\star & C_{22} \end{bmatrix} \tag{4.15}$$

or

$$\begin{aligned} &\Sigma X_{11} \pm X_{11}^\star \Sigma = C_{11}, \\ &\Sigma X_{12} = C_{12}, \\ &X_{21}, X_{22} = \text{free}; \end{aligned}$$

requiring $C_{22} = 0$ for consistency. With σ_k being the singular values of A, the first equation has the form

$$\sigma_i x_{ij} \pm \sigma_j x_{ji}^\star = c_{ij}.$$

For $i \neq j$, we can solve these equations in the minimum norm solution sense:

$$\begin{bmatrix} x_{ij} \\ x_{ji}^\star \end{bmatrix} = \frac{c_{ij}}{\sigma_i^2 + \sigma_j^2}\begin{bmatrix} \sigma_i \\ \pm\sigma_j \end{bmatrix},$$

or let x_{ji} $(j > i)$ be free and express x_{ij} in terms of x_{ji}:

$$x_{ij} = \frac{c_{ij} \mp \sigma_j x_{ji}^\star}{\sigma_i}.$$

For $i = j$, we have

$$\sigma_i(x_{ii} \pm x_{ii}^\star) = c_{ii}.$$

When $\star = T$, $x_{ii} = \frac{c_{ii}}{2\sigma_i}$ for the "+" case, or x_{ii} is free requiring $c_{ii} = 0$ (from the anti-symmetry of C) for the "$-$" case. When $\star = H$, $\Re e(x_{ii}) = \frac{c_{ii}}{2\sigma_i}$ with $\Im m(x_{ii})$ free for the "+" case, or $\Im m(x_{ii}) = \frac{c_{ii}}{2\sigma_i}$ with $\Re e(x_{ii})$ free for the " $-$ " case.

Note that minimum norm and minimum residual solutions are feasible from the above formulation.

Applying the formula in [4] with A in SVD, we obtain

$$\widetilde{X} \equiv \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix} = \begin{bmatrix} \frac{1}{2}\Sigma^{-1}C_{11} + Z_{11}\Sigma & \Sigma^{-1}C_{12} \\ Y_{21} & Y_{22} \end{bmatrix}, \tag{4.16}$$

where Y_{21} and Y_{22} are arbitrary and $Z_{11} = \mp Z_{11}^\star$. The solutions are identical except the (underdetermined) calculations involving X_{11} is handled differently in [4] by the choice of Z_{11}. To apply the formula in [4] to a general A, we need to choose an arbitrary Z such that

$$\left(P_2^T Z P_2\right)^T = \mp P_2^T Z P_2, \tag{4.17}$$

where $P_2 = A^T G$ with G satisfying $A^T G A^T = A^T$. To choose Z using the SVD in (4.14), we have $P_2 = V_1 V_1^H$ and (4.17) becomes

$$\overline{V}_1 V_1^T (Z^T \pm Z) V_1 V_1^H = 0 \Longleftrightarrow V_1^T \overline{V}(\widetilde{Z}^T \pm \widetilde{Z}) V^H V_1 = 0, \quad \widetilde{Z} \equiv V^T Z V = \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{bmatrix};$$

implying the same condition for Z_{11} $(= \mp Z_{11}^\star)$ as in (4.16). Consequently, we might as well use the SVD of A to solve the T-Lyapuniov equation as in (4.15).

4.3.3. QR

The SVD in Section 4.3.2 can be replaced by the cheaper but equally effective QR decomposition. Let

$$A = QR\Pi = Q\begin{bmatrix} R_{11} & R_{12} \\ 0 & 0 \end{bmatrix}\Pi$$

for some nonsingular R_{11} and permutation matrix Π. The transformed equation is, for $\star = T$:

$$R(\Pi X\overline{Q}) \pm (\Pi X\overline{Q})^T R^T = Q^H C\overline{Q}$$

or, for $\star = H$:

$$R(\Pi XQ) \pm (\Pi XQ)^H R^H = Q^H CQ.$$

These have the form

$$\begin{bmatrix} R_{11} & R_{12} \\ 0 & 0 \end{bmatrix}\begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix} \pm \begin{bmatrix} X_{11}^\star & X_{21}^\star \\ X_{12}^\star & X_{22}^\star \end{bmatrix}\begin{bmatrix} R_{11}^\star & 0 \\ R_{12}^\star & 0 \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} \\ \pm C_{12}^\star & C_{22} \end{bmatrix}.$$

Then we have

$$R_{11}X_{11} \pm X_{11}^{\star}R_{11}^{\star} = C_{11} - R_{12}X_{21} \mp X_{21}^{\star}R_{12}^{\star},$$
$$R_{11}X_{12} = C_{12} - R_{12}X_{22}$$

with X_{21} and X_{22} free. We can obtain X_{11} and X_{12} from the decoupled equations, with the first equation solved using the techniques in Section 4.2.

5. Conclusions

We have considered the solution of the ★-Sylvester equation which has not been fully investigated before. For the square case, solvability conditions have been derived and algorithms have been proposed. Preliminary numerical results shows that the algorithms behave promisingly. The rectangular case and some related equations, especially the ★-Lyapunov equation, have also been considered.

It is interesting and exciting that the ★ above the second X in (2.1) makes the equation behave very differently. The solvability condition in terms of non-intersection of the spectra $\sigma(A)$ and $\sigma(B)$, for the ordinary Sylvester equation $AX \pm XB = C$, is shifted to (2.22) for the generalized spectrum $\sigma(A, B)$. In addition, (2.1) looks like a Sylvester equation associated with continuous-time but (2.22) is satisfied when $\sigma(A, B)$ in totally inside the unit circle, hinting at a discrete-time type of stability behavior.

For numerical solution, the varying levels of difficulty and complexity for various equations are also intriguing. In terms of increasing complexity, the ★-Lyapunov, Lyapunov, Sylvester, ★-Sylvester and generalized ★-Sylvester equations require, respectively, the QR, Schur, Schur-Hessenberg, generalized Schur and periodic Schur decompositions. The ★ makes the Lyapunov equation easier (by creating more symmetry) yet forces the Sylvester equation the opposite direction.

Acknowledgements

The authors thank Professor Fan Hung-Yuan (National Taiwan Normal University) for many interesting and valuable suggestions on the manuscript. This research work is partially supported by the National Science Council and the National Center for Theoretical Sciences in Taiwan.

Appendix A. Palindromic linearization $\lambda Z + Z^{\star}$

An interesting application, for the ★-Sylvester equation (2.1)

$$AX \pm X^{\star}B^{\star} = C, \quad A, B, X \in \mathbb{C}^{n\times n}$$

arises from the eigensolution of the palindromic linearization [7]

$$(\lambda Z + Z^{\star})x = 0, \quad Z = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \in \mathbb{C}^{2n\times 2n}.$$

Applying congruence, we have

$$\begin{bmatrix} I_n & 0 \\ X & I_n \end{bmatrix}(\lambda Z + Z^{\star})\begin{bmatrix} I_n & X^{\star} \\ 0 & I_n \end{bmatrix} =$$

$$\begin{bmatrix} \lambda A + A^{\star} & \lambda(AX^{\star} + B) + (XA + C)^{\star} \\ \lambda(XA + C) + (AX^{\star} + B)^{\star} & \lambda\mathcal{R}(X) + \mathcal{R}(X)^{*} \end{bmatrix}$$

with

$$\mathcal{R}(X) \equiv XAX^{\star} + XB + CX^{\star} + D.$$

If we can solve the ★-Riccati equation

$$\mathcal{R}(X) = 0,$$

the palindromic linearization can then be "square-rooted". We then have to solve the generalized eigenvalue problem for the pencil $\lambda(AX^{\star} + B) + (XA + C)^{\star}$, with the reciprocal eigenvalues in $\lambda(XA + C) + (A\,X^{\star} + B)^{\star}$ obtained for free.

It is easy to show from the ★-Riccati equation that its solution corresponds to the (stabilizing) deflating subspaces of $\lambda Z + Z^{\star}$ spanned by

$$(S_1, S_2) \equiv \left(\begin{bmatrix} X^{\star} \\ I \end{bmatrix}, \quad \begin{bmatrix} I \\ -X \end{bmatrix}\right).$$

It turns out that the palindromic symmetry in the problem leads to the orthogonality property $S_1^{\star}S_2 = 0$, allowing the above congruence to annihilate the lower-right corner of the transformed pencil, thus square-rooting the problem.

Solving the ★-Riccati equation is of course as difficult as the original eigenvalue problem of $\lambda Z + Z^{\star}$. The usual invariance/deflating subspace approach for Riccati equations leads back to the original difficult eigenvalue problem. The obvious application of Newton's method lead to the iterative process

$$\delta X_{k+1}\left(AX_k^{\star} + B\right) + (X_kA + C)\delta X_{k+1}^{\star} = -\mathcal{R}(X_k),$$

which is a ★-Sylvester equation for δX_{k+1}.

Appendix B. Solution of *T*-Sylvester equations using generalized inverses [16]

We aim to show the result is tedious and difficult to implement.

The solution of

$$AX + X^TC = B; \quad A, C \in \mathbb{C}^{m\times n}\ (m \neq n) \tag{5.1}$$

was investigated using generalized inverses. We shall only quote the main result, ignoring some special cases.

Let $G \equiv A^{(1)}$ [2], the 1-inverse which satisfies $AGA = A$, with the projections $P_{11} \equiv G_1A_1$, $P_{12} \equiv A_1G_1$, $P_{21} \equiv G_2A_2$ and $P_{22} \equiv A_2G_2$, and

$$A_1 \equiv A^T + C, \quad B_1 \equiv B + B^T, \quad G_1 \equiv A_1^{(1)};$$
$$A_2 \equiv A^T - C, \quad B_2 \equiv B - B^T, \quad G_2 \equiv A_2^{(1)};$$
$$A_3 \equiv (I - P_{22})A_1, \quad G_3 \equiv [(I - P_{22})A_1]^{(1)}.$$

We have the following result for the solution of (5.1):

Theorem 5.1 (*17, Extension 2*). *The necessary and sufficient conditions for the solvability of (5.1) are:*

$$\left(I - P_{11}^T\right)B_1(I - P_{11}) = 0, \quad \left(I - P_{21}^T\right)B_2(I - P_{21}) = 0, \quad \left(I - P_{31}^T\right)B_3(I - P_{31}) = 0$$

where $B_3 \equiv 2B - A_1^TP_{22}^TZ_2P_{22}A_2 - A_2^TP_{22}^TZ_2P_{22}A_1$, *which is satisfying*

$$B_3 = B_2 - \left\{\frac{1}{2}P_{11}^TB_1P_{11} + A_2^TG_1^TB_1(I - P_{11}) + P_{12}^TZ_1P_{12}A_1 - \left[\frac{1}{2}P_{11}^TB_1G_1 + \left(I - P_{11}^T\right)B_1G_1 - A_1^T\left(P_{12}^TZ_1P_{12}\right)A_2\right]\right\}, \tag{5.2}$$

where $Z_1^T = -Z_1$ *and* $Z_2^T = Z_2$.

When the above conditions are satisfied, the general solution to (5.1) is

$$X = \frac{1}{2}G_1^TB_1P_{11} + G_1^TB_1(I - P_{11}) + \left(P_{12}^TZ_1P_{12}\right)A_1 + \left(I - P_{12}^T\right)\left[\frac{1}{2}G_3^TB_3P_{31} + G_3^TB_3(I - P_{31}) + (I - P_{31}^T)Y + P_{32}^TZP_{32}A_3\right],$$

with Y and Z being arbitrary, where $P_{31} \equiv G_3\,(I - P_{22})A_1$ *and* $P_{32} \equiv [(I - P_{22})A_1]^{(1)}$.

(The first G_1 inside the square brackets in (5.2) was mistaken to be an undefined G_{11} in [16].).

References

[1] C.S. Ballantine, A note on the matrix equation $H = AP + PA^{\star}$, Linear Algebra Appl. 2 (1969) 37–47.
[2] D.S. bernstein, Matrix Mathematics: Theory, Facts, and Formulas with Applications to Linear Systems Theory, Princeton University Press, Princeton and Oxford, 2005.
[3] A. Bojanczyk, G. Golub, P. Van dooren, The periodic Schur decomposition: algorithm and applications, in: Proc. SPIE Conference, vol. 1770, 1992, pp. 31–42.
[4] H.W. Braden, The equations $A^TX \pm X^TA = B$, SIAM J. Matrix Anal. Appl. 20 (1998) 295–302.
[5] R. Byers, D. Kressner, Structured condition numbers for invariant subspaces, SIAM J. Matrix Anal. Appl. 28 (2006) 326–347.
[6] C.Y. Chiang, E.K.-W. Chu, W.-W. Lin, Structured factorization of palindromic matrix polynomials of odd degree, Technical Report, National Chiao Tung University, Hsinchu, Taiwan, 2009.
[7] E.K.-W. Chu, T.-M. Hwang, W.-W. Lin, C.-T. Wu, Vibration of fast trains, palindromic eigenvalue problems and structure-preserving doubling algorithms, J. Comput. Appl. Math. 219 (2008) 237–252, doi:10.1016/j.cam.2007.07.016.
[8] K.-W.E. chu, The solution of the matrix equations $AXB - CXD = E$ and $(YA - DZ, YC - BZ) = (E,F)$, Linear Algebra Appl. 93 (1987) 93–105.
[9] J. Demmel, B. Kagström, The generalized Schur decomposition of an arbitrary pencil $A - \lambda B$: robust software with error bounds and applications, Part I: Theory and algorithms, ACM Trans. Math. Software 19 (1993) 60–174.
[10] J. Demmel, B. Kagström, The generalized Schur decomposition of an arbitrary pencil $A - \lambda B$: robust software with error bounds and applications, Part II: Software and applications, ACM Trans. Math. Software 19 (1993) 175–201.
[11] Z.-B. Fan, C.-J. Bu, The solution of the matrix equation $AX - X^TB = C$, College Math. 20 (2004) 100–102.
[12] A.R. Ghavimi, A.J. laub, Backward error, sensitivity, and refinement of computed solutions of algebraic Riccati equations, Numer. Linear Algebra Appl. 2 (1995) 29–49.
[13] G.H. golub, C.F. Van loan, Matrix Computations, second ed., Johns Hopkins University Press, Baltimore, MD, 1989.
[14] N.J. Higham, Accuracy and Stability of Numerical Algorithms, second ed., SIAM, Philadelphia, 2002.
[15] K. Liang, J. Liu, Iterative algorithms for the minimum-norm solution and the least-squares solution of the linear matrix equations $A_1XB_1 + C_1X^TD_1 = M_1$, $A_2XB_2 + C_2X^TD_2 = M_2$, Appl. Math. Comput. 218 (7) (2011) 3166–3175.
[16] X.-Z. Ma, J.-H. Han, F.-X. Piao, A study of solution existence for matrix equation $AX + X^TC = B$, J. Shenyang Inst. Aero. Eng. 20 (2003) 64–66.
[17] Mathworks, MATLAB User's Guide, 2002.

[18] D. Kressner, C. Schröder, D.S. Watkins, Implicit QR algorithms for palindromic and even eigenvalue problems, DFG Research Centre Matheon, Preprint 432, 2008.
[19] F. de Teran, F. Dopico, Consistency and efficient solution of the Sylvester equation for ★-congruence: $AX + X^{\star}B = C$, Electron. J. Linear Alg. 22 (2011) 849–863.
[20] O. Taussky, H. Wielandt, On the matrix function $AX + X^{\star}A_0^{\star}$, Arch. Ration. Mech. Anal. 9 (1962) 93–96.
[21] Y. Su, G. Chen, Iterative methods for solving linear matrix equation and linear matrix system, Int. J. Comput. Math. 87 (4) (2010) 763–774.
[22] M. Wang, X. Cheng, M. Wei, Iterative algorithms for solving the matrix equation $AXB + CX^TD = E$, Appl. Math. Comput. 187 (2) (2007) 622–629.
[23] L. Xie, Y. Liu, H. Yuan, Gradient based and least squares based iterative algorithms for matrix equations $AXB + CX^TD = F$, Appl. Math. Comput. 217 (5) (2010) 2191–2199.