On the $\star$-Sylvester Equation $AX \pm X^\star B^\star = C$

Chun-Yueh Chiang[*] Eric King-wah Chu[†] Wen-Wei Lin[‡]

October 14, 2008

Abstract

We consider the solution of the $\star$-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 stable numerical methods are considered, in terms of the (generalized and periodic) Schur and QR decompositions. We emphasize on the square cases where $m = n$ but the rectangular cases will be considered.

Keywords. error analysis, least squares, linear matrix equation, Lyapunov equation, palindromic eigenvalue problem, QR decomposition, generalized algebraic Riccati equation, Schur decomposition, singular value decomposition, solvability, Sylvester equation

AMS subject classifications. 15A03, 15A09, 15A24

1 Introduction

In [3], the Lyapunov-like linear matrix equation

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

with $(\cdot)^\star = (\cdot)^T$ was considered using generalized inverses. Applications occur in Hamiltonian mechanics. At the end of [3], the more general Sylvester-like equation

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

was proposed without solution. The equation (with $\star = T$) was studied, again using generalized inverses, in [10, 14]. However in [14], 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 III for the main result). In [10], some necessary or sufficient conditions for solvability were derived. A (seemingly wrong) formula for X in terms of generalized inverse was also proposed

[*]Department of Applied Mathematics, National Chiao Tung University, Hsinchu 300, Taiwan; d907201@oz.nthu.edu.tw

[†]School of Mathematical Sciences, Building 28, Monash University, VIC 3800, Australia; Phone: +61-3-99054480, FAX: +61-3-99054403; eric.chu@sci.monash.edu.au

[‡]Department of Applied Mathematics, National Chiao Tung University, Hsinchu 300, Taiwan; wwlin@am.nctu.edu.tw, This research work is partially supported by the National Science Council and the National Center for Theoretical Sciences in Taiwan.

(see Section 2.2 for more details on the approach taken in [10]). Consult also [4, 16], where solvability conditions for the $\star$-Sylvester equations with $m = n$ were obtained, without explicitly considering the numerical solution of the equations.

In this paper, the (numerical) solution of the $\star$-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, singular value and QR decompositions [11]. We are mainly interested in the square cases when $m = n$.

Our interest in the $\star$-Sylvester equation originates from the solution of the $\star$-Riccati equation

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

from an application related to the palindromic eigenvalue problem [4, 6, 16] (where eigenvalues appears in reciprocal pairs λ and $\lambda^{-\star}$). The solution of the $\star$-Riccati equation is difficult and the application of Newton's method is an obvious possibility. The solution of the $\star$-Sylvester equation is required in the Newton iterative process. Interestingly, the $\star$-Sylvester and $\star$-Lyapunov equations behave very differently from the ordinary Sylvester and Lyapunov equations. For example, from Theorem 2.1 below, the $\star$-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 detail of this application, see Appendix I.

Another application of the $\star$-Sylvester equation involves the generalized algebraic Riccati equations (GARE) in [5, 17], whose solutions by Newton's method require the solution of a coupled set of two T-Lyapunov equations, which is equivalent to a T-Sylvester equation, as described in Section 2.3. See Appendix II for more detail for this application.

The paper is organized as follows. After this introduction, Section 2 considers the $\star$-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 $\star$-Sylvester equation — $AXB^\star \pm X^\star = C$, $AXB^\star \pm CX^\star D^\star = E$ and the $\star$-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 [4, 16]) in the Appendices.

2 $\star$-Sylvester Equation

Consider the $\star$-Sylvester equation

$$AX \pm X^\star B^\star = C \ , \quad A, B, X \in \mathbb{C}^{n\times n} \ . \tag{1}$$

This includes the special cases of the T-Sylvester equation when $\star = T$ and the H-Sylvester equation when $\star = H$.

With the Kronecker product and for $\star = T$, (1) can be written as

$$\mathcal{P}\,\mathrm{vec}(X) = \mathrm{vec}(C) \ , \quad \mathcal{P} \equiv I \otimes A \pm (B \otimes I)E \tag{2}$$

where vec$X$ stacks the columns of X onto a column vector and E is the permutation matrix which maps vec(X) into vec(X^T). The matrix operator on the left-hand-side of (2) is $n^2 \times n^2$

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

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 $\star = H$ case, (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

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

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 above Kronecker product formulation for T-Sylvester equations can then be applied. Such a formulation will be less efficient for the numerical solution of (1), but may be useful as a theoretical tool.

Another approach will be to transform (1) by some unitary P and Q, so that (1) becomes:

$$PAQ \cdot \overline{Q}^T XP^T \pm PX^T\overline{Q} \cdot Q^TB^TP^T = PCP^T \tag{3}$$

or, for $\star = H$:

$$PAQ \cdot Q^H XP^H \pm PX^HQ \cdot Q^HB^HP^H = PCP^H\ . \tag{4}$$

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 [11]. The transformed equations in (3) and (4) then have the form

$$\begin{bmatrix} a_{11} & 0^T \\ a_{21} & A_{22} \end{bmatrix}\begin{bmatrix} x_{11} & x_{12}^\star \\ x_{21} & X_{22} \end{bmatrix} \pm \begin{bmatrix} x_{11}^\star & x_{21}^\star \\ x_{12} & X_{22}^\star \end{bmatrix}\begin{bmatrix} b_{11}^\star & b_{21}^\star \\ 0 & B_{22}^\star \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12}^\star \\ c_{21} & C_{22} \end{bmatrix}\ . \tag{5}$$

Multiply the matrices out, we have

$$a_{11}x_{11} \pm b_{11}^\star x_{11}^\star = c_{11}\ , \tag{6}$$

$$a_{11}x_{12}^\star \pm x_{21}^\star B_{22}^\star = c_{12}^\star \mp x_{11}^\star b_{21}^\star\ , \tag{7}$$

$$A_{22}x_{21} \pm b_{11}^\star x_{12} = c_{21} - x_{11}a_{21}\ , \tag{8}$$

$$A_{22}X_{22} \pm X_{22}^\star B_{22}^\star = \widetilde{C}_{22} \equiv C_{22} - a_{21}x_{12}^\star \mp x_{12}b_{21}^\star\ . \tag{9}$$

From (6) for $\star = T$, we have

$$(a_{11} \pm b_{11})x_{11} = c_{11} \tag{10}$$

Let $(a_{11}, b_{11}) \in \sigma(A, B)$. The solvability condition of the above equation is

$$a_{11} \pm b_{11} \neq 0 \Leftrightarrow \lambda_i = a_{11}/b_{11} \neq \mp 1 \tag{11}$$

Obviously, when $n = 1$, (11) is the only solvability condition.

From (6) when $\star = H$, we have

$$a_{11}x_{11} \pm \overline{b}_{11}\overline{x}_{11} = c_{11} \tag{12}$$

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

$$a_r x_r - a_i x_i \pm b_r x_r \mp b_i x_i = c_r \ , \quad a_r x_i + a_i x_r \mp b_r x_i \mp b_i x_r = c_i \ .$$

These imply

$$\begin{bmatrix} a_r \pm b_r & -a_i \mp b_i \\ a_i \mp b_i & a_r \mp b_r \end{bmatrix} \begin{bmatrix} x_r \\ x_i \end{bmatrix} = \begin{bmatrix} c_r \\ c_i \end{bmatrix} . \tag{13}$$

Let $\lambda = a_{11}/b_{11} \in \sigma(A, B)$. The determinant of the matrix operator in (13):

$$d = (a_r^2 - b_r^2) - (b_i^2 - a_i^2) = |a_{11}|^2 - |b_{11}|^2 \neq 0 \Leftrightarrow |\lambda| \neq 1 \tag{14}$$

requiring that no eigenvalue $\lambda \in \sigma(A, B)$ lies on the unit circle. Again, (14) is the solvability condition when $n = 1$.

Another way to solve (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 (7) and (8), we obtain

$$\begin{bmatrix} a_{11}^\star I & \pm B_{22} \\ \pm b_{11}^\star I & A_{22} \end{bmatrix} \begin{bmatrix} x_{12} \\ x_{21} \end{bmatrix} = \begin{bmatrix} \widetilde{c}_{12} \\ \widetilde{c}_{21} \end{bmatrix} \equiv \begin{bmatrix} c_{12} \\ c_{21} \end{bmatrix} + x_{11} \begin{bmatrix} \mp b_{21} \\ -a_{21} \end{bmatrix} \tag{15}$$

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 (19). If $a_{11} \neq 0$, (15) is then equivalent to

$$\begin{bmatrix} a_{11}^\star I & \pm B_{22} \\ 0 & A_{22} - \frac{b_{11}^\star}{a_{11}^\star} B_{22} \end{bmatrix} \begin{bmatrix} x_{12} \\ x_{21} \end{bmatrix} = \begin{bmatrix} \widetilde{c}_{12} \\ \widehat{c}_{21} \end{bmatrix} \equiv \begin{bmatrix} \widetilde{c}_{12} \\ \widetilde{c}_{21} \mp \frac{b_{11}^\star}{a_{11}^\star} \widetilde{c}_{12} \end{bmatrix} . \tag{16}$$

The solvability condition of (15) and (16) is

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

or, with $(a_{11}, b_{11}) \in \sigma(A, B)$, $(b_{11}^\star, a_{11}^\star)$ cannot be another eigenvalue of (A, B). Note that $\widetilde{A}_{22}$ is still lower-triangular, just like A_{22} or B_{22}.

If $b_{11} \neq 0$, (15) is equivalent to

$$\left[\begin{array}{cc} 0 & B_{22} - \frac{a_{11}^\star}{b_{11}^\star} A_{22} \\ b_{11}^\star I & \pm A_{22} \end{array}\right] \left[\begin{array}{c} x_{12} \\ x_{21} \end{array}\right] = \left[\begin{array}{c} \widehat{c}_{12} \\ \pm\widetilde{c}_{21} \end{array}\right] \equiv \left[\begin{array}{c} \pm\widetilde{c}_{12} - \frac{a_{11}^\star}{b_{11}^\star}\widetilde{c}_{21} \\ \pm\widetilde{c}_{21} \end{array}\right] \tag{17}$$

with an identical solvability condition (19).

Lastly, (9) is of the same form as (1) but of smaller size.

Remark 2.1 *Interestingly, for the ordinary Sylvester equation $AX - XB = C$, numerical solution will be possible when (A, B) is transformed into quasi-triangular/triangular form (not necessarily both of the same type) or the cheaper quasi-triangular/Hessenberg form. It is not the case for (1) and the $\star$ somehow alters the behaviour of the equation greatly.*

Remark 2.2 *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 TSylvester$_{\rm R}$ 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 (15) for x_{12} and x_{21}. The equation has the form, for some right-hand-side R_1:*

$$\left[\begin{array}{ccc|ccc} r_{11} & & & r_{12} & & \\ & s_{11} & & * & s_{12} & \\ & & \ddots & * & * & \ddots \\ \hline r_{21} & & & r_{22} & & \\ & s_{21} & & * & s_{22} & \\ & & \ddots & * & * & \ddots \end{array}\right] \left[\begin{array}{c} z_{r1} \\ z_{s1} \\ \vdots \\ \hline z_{r2} \\ z_{s2} \\ \vdots \end{array}\right] = R_1 \ . \tag{18}$$

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

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

where

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

Consequently, (18) is a quasi-lower-triangular linear system with at most 2×2 diagonal blocks. By implication, so is (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 (1) in the following theorem:

Theorem 2.1 *The $\star$-Sylvester equation (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:

$$a_{ii} a_{jj}^\star - b_{ii} b_{jj}^\star \neq 0 \quad (\forall i \neq j) \ ; \tag{19}$$

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

$$a_{ii} \pm b_{ii} \neq 0 \ \ (\text{for } \star = T) \ , \quad |\lambda_i| \neq 1 \ \ (\text{for } \star = H) \ . \tag{20}$$

Remark 2.3 *Condition (11) or (14) actually imply in (19). However, these conditions are only sufficient for (19) and have to be restated in (20) in Theorem 2.1. In terms of the eigenvalues* $\lambda_i \equiv a_{ii}/b_{ii}$*, (19) means that* $\lambda_j \neq \lambda_i^{-\star}$ $(i \neq j)$*, and (20) means that* $\lambda_i \neq \mp 1$ $(\star = T)$ *or* $|\lambda_i| \neq 1$ $(\star = H)$*. Consequently,* $\lambda = \pm 1$ *can be an eigenvalue of* (A, B) *but must be simple for the corresponding* $\star$*-Sylvester equation to be uniquely solvable.*

The solution process in this subsection is summarized below:

Algorithm SSylvester (for the unique solution of $AX \pm X^\star B^\star = C$; $A, B, C, X \in \mathbb{C}^{n\times n}$)
`if` (A, B) is not in lower-triangular generalized Schur form, `then`
  `compute` the lower-triangular generalized Schur form (PAQ, PBQ) using QZ algorithm [11]
`store` (PAQ, PBQ, PCP^*) in (A, B, C)
`solve` (10) for $\star = T$, or (13) for $\star = H$; `if` fail, `exit`
`if` $n = 1$ or $|a_{11}|^2 + |b_{11}|^2 \leq$ tolerance, `exit`
`if` $|a_{11}| \geq |b_{11}|$, `then`
  `if` $\widetilde{A}_{22} \equiv A_{22} - \frac{b_{11}^\star}{a_{11}^\star} B_{22}$ has any negligible diagonal elements, `then exit`
  `else compute` $x_{21} = \widetilde{A}_{22}^{-1}\widehat{c}_{21}$ by backsubstitution, $x_{12} = (\widetilde{c}_{12} \mp B_{22}x_{21})/a_{11}^\star$ (from (16))
`else if` $\widetilde{B}_{22} \equiv B_{22} - \frac{a_{11}^\star}{b_{11}^\star} A_{22}$ has any negligible diagonal elements, `then exit`
  `else compute` $x_{21} = \widetilde{B}_{22}^{-1}\widehat{c}_{12}$ by backsubstitution, $x_{12} = (\pm\widetilde{c}_{21} \mp A_{22}x_{21})/b_{11}^\star$ (from (17))
`apply` **Algorithm SSylvester** to $A_{22}X_{22} \pm X_{22}^\star B_{22}^\star = \widetilde{C}_{22}$ (from (9)), with $n \leftarrow n - 1$
`output` $X \leftarrow QX\overline{P}$ for $\star = T$, or $X \leftarrow QXP$ for $\star = H$
end of algorithm

Let the operation count of the Algorithm SSylvester, in addition to the $66n^3$ complex flops for the QZ procedure [11] for the generalized Schur decomposition of (A, B), be $f(n)$ complex flops, mainly involving the solution of (9) and (16) or (17). This involves forming and inverting $\widetilde{A}_{22}$ or $\widetilde{B}_{22}$ (n^2 flops), computing x_{12} ($\frac{1}{2}n^2$ flops) and forming $\widetilde{C}_{22}$ ($2n^2$). Thus $f(n) \approx f(n-1) + \frac{7}{2}n^2$, ignoring $O(n)$ terms. This implies that $f(n) \approx \frac{7}{6}n^3$ and the total operation count for Algorithm SSylvester is $67\frac{1}{6}n^3$ complex flops, ignoring $O(n^2)$ terms.

From the above analysis and Theorem 2.1, the condition of (1) will be bad if the separation $\lambda_i\lambda_j^\star - 1$ is narrow (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 detail.

2.1 The real case or divide-and-conquer

When A, B and C are all real, the solution X, judging from (2), will be real. To guarantee a real solution X, the generalized real Schur form [11] for (A, B) has to be used. The transformed equation in (3) or (4) has the form

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

or

$$
\begin{aligned}
A_{11}X_{11} \pm X_{11}^\star B_{11}^\star &= C_{11} \ , && (21)\\
A_{11}X_{12}^\star \pm X_{21}^\star B_{22}^\star &= \widetilde{C}_{12}^\star \equiv C_{12}^\star \mp X_{11}^\star B_{21}^\star \ , && (22)\\
A_{22}X_{21} \pm X_{12}B_{11}^\star &= \widetilde{C}_{21} \equiv C_{21} - A_{21}X_{11} \ , && (23)\\
A_{22}X_{22} \pm X_{22}^\star B_{22}^\star &= \widetilde{C}_{22} \equiv C_{22} - A_{21}X_{12}^\star \mp X_{12}B_{21}^\star \ ; && (24)
\end{aligned}
$$

where A_{11} and B_{11} may be 1×1 or 2×2. The former case will be trivial as in (6) and the latter can be handled using the Kronecker product. The theory leading to the conditions in (11) and (19) 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 (22) and (23). A better approach is to consider $(22)^\star$ and (23):

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

A linear combination of these equations will be

$$(\beta A_{22} \pm \alpha B_{22})X_{21} + X_{12}(\alpha A_{11}^\star \pm \beta B_{11}^\star) = \alpha\widetilde{C}_{12} + \beta\widetilde{C}_{21} \ . \qquad (25)$$

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

$$X_{12} = -(\beta A_{22} \pm \alpha B_{22})X_{21}(\alpha A_{11}^\star \pm \beta B_{11}^\star)^{-1} + \widehat{C}_{12} \ , \quad \widehat{C}_{12} \equiv (\alpha\widetilde{C}_{12} + \beta\widetilde{C}_{21})(\alpha A_{11}^\star \pm \beta B_{11}^\star)^{-1} \ . \quad (26)$$

Substitute X_{12} in (26) into (23), we have a generalized Sylvester equation [7] for X_{21}:

$$A_{22}X_{21} - (\alpha B_{22} \pm \beta A_{22})X_{21}(\alpha A_{11}^\star \pm \beta B_{11}^\star)^{-1}B_{11}^\star = \widetilde{C}_{21} \mp \widehat{C}_{12}^\star B_{11}^\star \ . \qquad (27)$$

From [7], (27) is uniquely solvable when there is no intersection of the spectra $\sigma(A_{22}, \alpha B_{22} \pm \beta A_{22})$ and $\sigma(B_{11}^\star, \alpha A_{11}^\star \pm \beta B_{11}^\star)$. 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 solvability condition is

$$\frac{\alpha_j}{\alpha\beta_j \pm \beta\alpha_j} \neq \frac{\beta_i^\star}{\alpha\alpha_i^\star \pm \beta\beta_i^\star} \Leftrightarrow \alpha_i^\star\alpha_j \neq \beta_i^\star\beta_j \ ,$$

exactly condition (19). The same conclusion is reached when $\alpha = 0$, which implies that B_{11} is invertible, and X_{12} in (26) should then be substituted into the $\star$ of (22) to produce a similar generalized Sylvester equation for X_{21}:

$$B_{22}X_{21} - A_{22}X_{21}B_{11}^{-\star}A_{11}^\star = \pm\widetilde{C}_{12} + \widehat{C}_{12}A_{11}^\star \ . \qquad (28)$$

Also X_{12} is retrievable from (26) in a numerical stable and efficient manner. Note that the matrix operator $\alpha A_{11}^\star \pm \beta B_{11}^\star$ in (26) is 2×2 with (α, β) controlling its condition. In (27), A_{22} and B_{22} are block-lower triangular with $(\alpha A_{11}^\star \pm \beta B_{11}^\star)^{-1}B_{11}^\star$ being at most 2×2, enabling X_{21} to be easily calculated as in the generalized Bartels-Stewart algorithm in [7]. (For illustration, let us consider (28). With B_{22} and A_{22} being lower-triangular and $B_{11}^{-\star}A_{11}^\star$ being 2×2, the first row and column of X_{21} can be computed easily, 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 (27) consecutively from the top, solving a 2×2 system for each row of X_{21}. Equation (28) can be solved analogously, also one row at a time.

We can then solve recursively (24), a smaller equation similar to (9).

Lastly, the procedure discussed in this subsection can be applied as a divide-and-conquer strategy, with A_{11} and B_{11} being $[\frac{n}{2}] \times [\frac{n}{2}]$. After transforming (1) using the (real) Schur form of (A, B), the resulting equation can be split up in the middle, with the sizes of A_{11} and A_{22} roughly equal. Subsequent systems in terms of (A_{ii}, B_{ii}) $(i = 1, 2)$ can then be treated recursively in the same divide-and-conquer fashion, yielding a more efficient version of our algorithm. We summarize the procedure in this subsection in the following algorithm, with the subscripts "R" for "Real".

Algorithm TSylvester$_{\mathrm{R}}$ (for the unique solution of $AX \pm X^T B^T = C$; $A, B, C, X \in \mathbb{R}^{n\times n}$)
`if` (A, B) is not in quasi-lower-triangular generalized real Schur form, `then`
  `compute` the quasi-lower-triangular generalized real Schur form (PAQ, PBQ) by QZ
`store` (PAQ, PBQ, PCP^T) in (A, B, C)
`solve` (21) for X_{11}; `if` fail, `exit`
`if` last block reached with $n = 1, 2$, `exit`
`if` A_{11} and B_{11} are scalar, `solve` (22) and (23) for X_{12} and X_{21} as in **Algorithm TSylvester**
  `if` fail, `exit`; `else solve` (27) or (28) with appropriate α, β for X_{21} row-wise,
    using Gaussian elimination on the 2×2 systems; `if` fail, `exit`
`retrieve` X_{12} from (26)
`apply` **Algorithm TSylvester**$_{\mathrm{R}}$ to $A_{22}X_{22} \pm X_{22}^T B_{22}^T = \widetilde{C}_{22}$ (c.f. (24)), $n \leftarrow n-1$ or $n-2$
`output` $X \leftarrow QXP$
end of algorithm

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

Remark 2.4 *Similar to Remark 2.2, Algorithms TSylvester*$_{\mathrm{R}}$ *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 (22) and (23) using the Kronecker product yields a linear system with matrix operator*

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

assuming without loss of generality that $\star = T$. The matrix has the same form as the one in (18), except the elements may be 2×2 blocks, producing a series of 4×4 linear systems. Similar arguments as those in Remark 2.2 thus follows.

2.2 Error analysis

We shall discuss the condition and error associated with Algorithms TSylvester and TSylvester$_{\mathrm{R}}$, following the development in [12, Chapter 16] and [13].

Condition

The condition of (1) is obviously identical to that of (2). However, E reshuffles the columns of $B \otimes I$, making the analysis of the matrix operator $\mathcal{P}$ difficult. We shall investigate the eigenvalues of $\mathcal{P}$, collaborating Theorem 2.1. First consider the trivial example when $n = 2$, $A = [a_{ij}]$ and $B = [b_{ij}]$, we have

$$\mathcal{P} = \left[\begin{array}{cccc} a_{11} \pm b_{11}^\star & & a_{12} \pm b_{12}^\star & \\ \pm b_{21}^\star & a_{11} & \pm b_{22}^\star & a_{12} \\ a_{21} & \pm b_{11}^\star & a_{22} & \pm b_{12}^\star \\ & a_{21} \pm b_{21}^\star & & a_{22} \pm b_{22}^\star \end{array} \right] .$$

To make things easier, let (A, B) be in lower-triangular generalized Schur form after some QZ procedure. We then have $a_{12}, b_{12} = 0$ and the eigenvalues of the corresponding $\widetilde{\mathcal{P}}$ are $a_{ii} \pm b_{ii}^\star$ $(i = 1, 2)$ and those of the middle block W_{12} where

$$W_{ij} \equiv \left[\begin{array}{cc} a_{ii} & \pm b_{jj}^\star \\ \pm b_{ii}^\star & a_{jj} \end{array} \right] .$$

The characteristic polynomial of W_{ij}, identical to that for W_{ji}, is $\lambda^2 - (a_{ii} + a_{jj})\lambda + \det W_{ij}$ with $\det W_{ij} = a_{ii}a_{jj} - b_{ii}^\star b_{jj}^\star$, and the eigenvalues are

$$\lambda_{W_{ij}} = \frac{1}{2} \left[a_{ii} + a_{jj} \pm \sqrt{(a_{ii} - a_{jj})^2 + 4 b_{ii}^\star b_{jj}^\star} \right] .$$

Note that some $\lambda_{W_{ij}}$ or $\det W_{ij} = 0$ if and only if (19) in Theorem 2.1 is violated. For larger values of n, the equivalence of our algorithms and quasi-triangular linear systems (after the initial QZ step), as mentioned in Remarks 2.2 and 2.4, means that $\widetilde{\mathcal{P}}$ in (2) is quasi-triangular with the correct permutation of the variables and equations. The ordering considers the first diagonal element x_{11}, then the first components in x_{12} and x_{21}, and then their second components etc. until exhaustion, and then recursively ordering X_{22} in the same fashion. From (18), the eigenvalues of $\widetilde{\mathcal{P}}$ thus consists of $a_{ii} \pm b_{ii}^\star$ (n of them, for $i = 1, \cdots, n$) and the eigenvalues $\lambda_{W_{ij}}$ (${}^nC_2 = n(n-1)/2$ of them, for $j > i$ and $i, j = 1, \cdots, n$). Notice that, as expected, there are exactly $2({}^nC_2) + n = n^2$ eigenvalues for $\widetilde{\mathcal{P}}$. It is conceptually simple but operationally tedious to reorder $\widetilde{\mathcal{P}}$ to show this result even for $n = 3$ and that will be left as an exercise.

Residual

As indicated in Remarks 2.2 and 2.4, Algorithms SSylvester and SSylvester$_\mathrm{R}$ can be arranged into quasi-triangular linear systems. We can then apply the error analysis for triangular linear systems in [12, Theorem 8.5] to obtain

$$\|R\|_F \equiv \|C - (A\widehat{X} \pm \widehat{X}^\star B^\star)\|_F \leq c_n u (\|A\|_F + \|B\|_F) \|\widehat{X}\|_F \tag{29}$$

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 (27), (28) and Remarks 2.2 and 2.4 are not bad. Note that the QZ transformation of (A, B) is backward stable, similar to the QR process in [12, Equation 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.

Backward error

Like for ordinary Sylvester equations, the numerical solution of (1) is not backward stable in general. Similar to [12, §16.2] and with "δ" indicating perturbation, we can define the normwise relative backward error of an approximate solution Y by

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

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) [11], the $Y^\star$ terms do not affect the analysis in [12, §16.2]. With $\Sigma = \mathrm{diag}\{\sigma_1, \cdots, \sigma_n\}$, it can be shown that

$$\eta(Y) \le \mu \frac{\|R\|_F}{(\alpha+\beta)\|Y\|_F + \gamma} \tag{30}$$

where

$$\mu \equiv \frac{(\alpha+\beta)\|Y\|_F + \gamma}{[(\alpha^2+\beta^2)\sigma_n^2 + \gamma^2]^{1/2}} \ , \quad R \equiv \delta A\, Y \pm Y^\star \delta B^\star - \delta C \ .$$

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 [12, §16.2], we do not know the conditions under which a $\star$-Sylvester equation has a well-conditioned solution.

Perturbation and practical error bounds

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

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

Application of norm gives rise to

$$\|\delta X\| \le \|S^{-1}\| \left\{\|\delta C\| + (\|\delta A\| + \|\delta B\|)(\|X\| + \|\delta X\|)\right\} \ .$$

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

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

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

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

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 [12, §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 (1) and its $\star$, producing

$$(A+B)X + X^\star(A+B)^\star = C + C^\star \ , \quad (A-B)X - X^\star(A-B)^\star = C - C^\star \ . \tag{31}$$

The pair of equations represent the symmetric/Hermitian and skew-symmetric/Hermitian parts of (1) and can be solved using the generalized Schur form of $(A+B, A-B)$. Identical solvability condition as (19) can be derived. In terms of the eigenvalues $\widetilde{\lambda}_i \in \sigma(A+B, A-B)$, (1) and (31) are uniquely solvability if and only if $\widetilde{\lambda}_i + \widetilde{\lambda}_j \neq 0$, with $\widetilde{\lambda}_i = (\lambda_i+1)/(\lambda_i-1)$ 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 [10], a formula for the solution X of (1) (for $\star = T$ and the "+" case) was derived using the first equation in (31) only, throwing away the information in the second equation. We cannot see how the formula can be correct using only half the information of (1) in the first half of (31). In the extreme case with $A = -B$, the first equation in (1) will be degenerate and the solution X will be totally free. Anyway, X is a solution of (1) if and only if it is also a solution of (31), but a solution of half of (31) in general does not satisfy (1).

3 Numerical Examples

In this section, we apply Algorithm SSylvester (denoted by ASS) and the Kronecker product approach in (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 3GB 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` *[15] commands, we have*

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

To guarantee condition (19), let $b = randn(n,1)$, $a = 2b$. In Table 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 SSA 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)). The results in Table 1 show that the advantage of ASS over KRP in CPU time grows rapidly as n increases, as predicted by the operation counts. Even with better management of sparsity or parallellism, 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) is solved by Gaussian elimination in an unstructured way. See the other examples for more comparison of the residuals of ASS and KRP.*

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 1: Results for Example 3.1

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 $\sigma(A, B) = \{\frac{\alpha+\varepsilon}{\beta}, \frac{\beta}{\alpha}\}$, and $|\lambda_1\lambda_2 - 1| = \frac{\varepsilon}{\alpha}$. Judging from (19), (1) has worsening condition when 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 2. The results show that if (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 2) reflects partially the condition of (1), as indicated in (29). 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 an ill-conditioned problem.*

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

Table 2: Results for Example 3.2

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*

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

and $C = AX_e + X_e^T B^T$. Solving the corresponding T-Sylvester equation by Algorithm SSylvester produces the results in Table 3, using symbols from Section 2.2. The column for the backward error $\eta(Y)$ (estimated using the bound (30)) 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)/$\|X\|$

m	Res(ASS)	RRes(ASS)	$\frac{\lvert X_{ASS}-X_e \rvert}{\lvert X_e \rvert}$	$O(\|X\|)$	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

Table 3: Results for Example 3.3

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 $\star$-Sylvester equation I

Consider the more general version of the $\star$-Sylvester equation (1):

$$AXB^\star \pm X^\star = C \tag{32}$$

with $A, B^\star, X^\star \in \mathbb{C}^{m\times n}$ and $m \neq n$. The generalized Kronecker canonical form [8, 9] 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 $\star$-Sylvester equation in Section 2 when either A or B is nonsingular. In general, consider the periodic Schur or PQZ decomposition [2] 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 $\star = H$:

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

or for $\star = T$:

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

The case when (A, B) are real and $\star = T$ with a real PQZ decomposition is similar but will be ignored here.

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

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

We then have

$$\begin{aligned} a_{11} b_{11}^\star x_{11} \pm x_{11}^\star &= c_{11}\ , && (33)\\ a_{11} x_{12}^\star B_{22}^\star \pm x_{21}^\star &= c_{12}^\star - a_{11} x_{11} b_{12}^\star\ , && (34)\\ b_{11}^\star A_{22} x_{21} \pm x_{12} &= c_{21} - b_{11}^\star x_{11} a_{21}\ , && (35)\\ A_{22} X_{22} B_{22}^\star \pm X_{22}^\star &= C_{22} - x_{11} a_{21} b_{12}^\star - A_{22} x_{21} b_{12}^\star - a_{21} x_{12}^\star B_{22}^\star\ . && (36) \end{aligned}$$

Inspection of (33)–(36) shows the solvability condition

$$a_{ii} b_{ii}^{\star} \neq \mp 1 \ , \quad a_{ii}^{\star} b_{ii}^{\star} a_{jj} b_{jj} \neq 1 \quad (\forall i \neq j) \ ; \tag{37}$$

analogous to (19) and (20) and a special case of (41). Algorithms can easily be constructed from (33)–(36) but will be ignored here.

4.2 Generalized $\star$-Sylvester equation II

Consider the more general version of the $\star$-Sylvester equation (1) and (32):

$$AXB^{\star} \pm CX^{\star}D^{\star} = E \tag{38}$$

with the complex matrices $A^{\star}$ and $C^{\star}$, $B^{\star}$ and $D^{\star}$, B and C, and A and D possessing the same number of columns. This is a more general equation than the rectangular $\star$-Sylvester equation in Section 2.4. It is also a special case of the equation in Section 4.5. We do not know how to tackle this equation.

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

$$PAR \cdot R^H XS \cdot S^H B^H Q \pm PCS \cdot S^H X^H R \cdot R^H D^H Q = PEQ \tag{39}$$

or, for $\star = T$:

$$PA\overline{R} \cdot R^T XS \cdot S^H B^T Q \pm PC\overline{S} \cdot S^T X^T R \cdot R^H D^T Q = PEQ \ . \tag{40}$$

These equation 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 [2] 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 $i \neq j = 1$:

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

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

$$\alpha_i \beta_i^{\star} \pm \gamma_i \delta_i^{\star} \neq 0 \ , \quad \alpha_i \alpha_j \beta_i^{\star} \beta_j^{\star} \neq \gamma_i \gamma_j \delta_i^{\star} \delta_j^{\star} \quad (\forall i \neq j) \ ; \tag{41}$$

conditions more general than but similar to (19) and (20), or (37).

4.3 $\star$-Lyapunov equation

Consider the $\star$-Lyapunov equation

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

With unitary P and Q, the equation can be transformed to, for $\star = T$:

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

or, for $\star = H$:

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

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

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

Multiply the matrices out, we have

$$a_{11}x_{11} \pm a_{11}^\star x_{11}^\star = c_{11} \ , \quad (42)$$

$$a_{11}x_{12}^\star \pm x_{21}^\star A_{22}^\star = \widetilde{c}_{12}^\star \equiv c_{12}^\star \mp x_{11}^\star a_{21}^\star \ , \quad (43)$$

$$A_{22}X_{22} \pm X_{22}^\star A_{22}^\star = \widetilde{C}_{22} \equiv C_{22} - a_{21}x_{12}^\star \mp x_{12}a_{21}^\star \ . \quad (44)$$

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, requiring $c_{11} = 0$ 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

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

which is underdetermined when $\widetilde{c}_{12}$ 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 $\star = T$:

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

or for $\star = H$:

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

we can impose the (anti-)symmetry constraint $X^\star = \pm X$. Equations (42)–(44) 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), we have

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

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 $c_{11} = 0$ automatically from the anti-symmetry of C. When $i = j$ and $\star = H$, no eigenvalue λ_i can be purely imaginary/real. Note that x_{11} is underdetermined and so are all the diagonal elements of X.

4.3.2 Rectangular A

The T-Lyapunov equation with rectangular A has been studied in [3] using generalized inverse (which can only be realized using the SVD). Please consult [3] 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 [11] is used instead to derive the same solution.

When A is rectangular, the SVD of A:

$$A = UDV^H = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1 & V_2 \end{bmatrix}^H \tag{45}$$

gives rise to the transformed T-Lyapunov equation:

$$UDV^H X \pm X^T \overline{V} D^T U^T = C \Leftrightarrow D(V^H X \overline{U}) \pm (U^H X^T \overline{V})D = U^H C \overline{U}$$

or the transformed H-Lyapunov equation:

$$UDV^H X \pm X^H V D^T U^H = C \Leftrightarrow D(V^H XU) \pm (U^H X^H V)D = U^H CU \ .$$

We then have

$$\begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix} \pm \begin{bmatrix} X_{11}^\star & X_{21}^\star \\ X_{12}^\star & X_{22}^\star \end{bmatrix} \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} \\ \pm C_{12}^\star & C_{22} \end{bmatrix} \tag{46}$$

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 least squares 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 consistency 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 [3] with A in SVD, we obtain

$$\widetilde{X} \equiv \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix} = \begin{bmatrix} \frac{1}{2}\Sigma^{-1} C_{11} + Z_{11}\Sigma & \Sigma^{-1} C_{12} \\ Y_{21} & Y_{22} \end{bmatrix} \tag{47}$$

where Y_{21} and Y_{22} are arbitrary and $Z_{11} = \mp Z_{11}^\star$. The solutions are identical except the (under-determined) calculations involving X_{11} is handled differently in [3] by the choice of Z_{11}. For a general A, we have to choose an arbitrary Z such that

$$(P_2^T Z P_2)^T = \mp P_2^T Z P_2 \tag{48}$$

where $P_2 = A^T G$ with G satisfying $A^T G A^T = A^T$. To choose Z using the SVD in (45), we have $P_2 = V_1 V_1^H$ and (48) becomes

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

implying the same condition for Z_{11} $(= \mp Z_{11}^\star)$ as in (47). Consequently, we might as well use the SVD of A to solve the T-Lyapuniov equation as in (46).

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 $\star = H$:

$$R(\Pi X Q) \pm (\Pi X Q)^H R^H = Q^H C Q .$$

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

$$\begin{aligned} R_{11}X_{11} \pm X_{11}^\star R_{11}^\star &= C_{11} - R_{12}X_{21} \mp X_{21}^\star R_{12}^\star , \\ R_{11}X_{12} &= C_{12} - R_{12}X_{22} . \end{aligned}$$

with X_{21} and X_{22} free. We can obtain X_{12} from the second equation and then retrieve X_{11} from the first. The first equation can be solved using the techniques in Section 4.2.

Alternatively, let $A^\star = QR\Pi$, then we have $R^\star(Q^\star X) \pm (X^\star Q)R = \Pi^T C \Pi$ and similar procedures follow. Minimum norm and minimum residual solutions are feasible from the above formulation.

Note that the solution of the $\star$-Lyapunov equation, with more symmetry, is easier than that of the Lyapunov equation, which requires the more expensive Schur decomposition.

5 Conclusions

We have considered the solution of the $\star$-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 $\star$-Lyapunov equation, have also been considered.

It is interesting and exciting that the $\star$ above the second X in (1) makes the equation behave very differently. The solvability condition in terms of non-intersecton of the spectra $\sigma(A)$ and $\sigma(B)$, for the ordinary Sylvester equation $AX \pm XB = C$, is shifted to (19) for the generalized spectrum $\sigma(A, B)$. In addition, (1) looks like a Sylvester equation associated with continuous-time but (19) is satisfied when $\sigma(A, B)$ in totally inside the unit circle, hinting at a discrete-time type of stability behaviour.

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

On future work, we are interested in the solution of the $\star$-Riccati equation (Appendix I) and the generalized algebraic Riccati equations (Appendix II). Preliminary numerical results by Newton's method are encouraging but also reveal several problems. The related work will be published elsewhere. Other possible future research problems include the estimation of practical error bounds and condition numbers, the conditions which guarantee good condition of X in (1), (32) and (38), more thorough numerical tests and a cheaper algorithm (preferably involving the Schur/Hessenberg decomposition of (A, B)) for (1), as well as the detailed analysis and numerical solution of the rectangular case of the $\star$-Sylvester equation and the other related equations in

Section 4 and Appendix II, and the behaviour of the alternative sep functions

$$\mathrm{sep}_2(A,B) = \min_{X\neq 0} \frac{\|AX - X^\star B^\star\|}{\|X\|} \;, \quad \mathrm{sep}_3(A,B) = \min_{X\neq 0} \frac{\|AXB^\star - X^\star\|}{\|X\|}$$

and

$$\mathrm{sep}_4\{(A,C);(B,D)\} = \min_{X\neq 0} \frac{\|AXB - CX^\star D\|}{\|X\|}$$

for $A, B, C, D \in \mathbb{C}^{n\times n}$.

Appendix I: Palindromic Linearization $\lambda Z + Z^\star$

An interesting application, for the $\star$-Sylvester equation (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 [6]

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

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 $\star$-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 $\lambda(XA + C) + (AX^\star + B)^\star$ obtained for free.

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

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

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 $\star$-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

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

which is a $\star$-Sylvester equation for δX_{k+1}.

Appendix II: Generalized Algebraic Riccati Equations

In [5], the numerical solution of the following generalized algebraic Riccati equation (GARE) was investigated:

$$A_a^T X_a + X_a^T A_a + (C_a^T J C_a - B_a J' B_a^T) - X_a^T B_a (J')^{-1} B_a^T X_a = 0 \quad \text{such that} \quad E_a^T X_a = X_a^T E_a$$

where

$$E_a = \begin{bmatrix} E & 0 \\ 0 & 0 \end{bmatrix}, \quad A_a = \begin{bmatrix} A & B \\ 0 & I \end{bmatrix}, \quad C_a = \begin{bmatrix} C & D \end{bmatrix}, \quad B_a = \begin{bmatrix} 0 \\ -I \end{bmatrix}$$

and some $E, A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{m\times n}$, $C \in \mathbb{R}^{p\times n}$, $D \in \mathbb{R}^{p\times m}$, $J \in \mathbb{R}^{p\times p}$, $J' \in \mathbb{R}^{m\times m}$, E is singular and J, J' are symmetric and nonsingular. Applying Newton's method, each iterative step will involve the solution of the coupled set of two T-Lyapunov equations

$$(\widetilde{A}\widetilde{X} + \widetilde{X}^T\widetilde{A}^T,\ \widetilde{E}\widetilde{X} - \widetilde{X}^T\widetilde{E}^T) = (\widetilde{B},\ \widetilde{C})$$

with some square $\widetilde{A}$, $\widetilde{B}$, $\widetilde{C}$, $\widetilde{E}$ and $\widetilde{X}$ with $\widetilde{B}$ (and $\widetilde{C}$) being (anti-)symmetric. This coupled set of equations is equivalent to a T-Sylvester equation, as described in Section 2.3. Numerical solution can be achieved through the equivalent T-Sylvester equation, or directly through the generalized Schur decomposition of $(\widetilde{A}, \widetilde{E})$.

There is also a similar GARE in [17]:

$$A^T X + X^T A + C^T C + X^T B B^T X = 0 \quad \text{such that} \quad E^T X = X^T E$$

in the H_∞ control of the descriptor system

$$E\dot{x} = Ax + Bu\ , \quad y = Cx$$

where $E, A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, and $C \in \mathbb{R}^{p\times n}$. This GARE may also be solved similarly.

Appendix III: Solution of T-Sylvester Equations using Generalized Inverses

In [14], the solution of

$$AX + X^T C = B\ ; \quad A, C \in \mathbb{C}^{m\times n} \quad (m \neq n) \tag{49}$$

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

Let $G \equiv A^{(1)}$, the 1-inverse which satisfies $AGA = A$, with $AP_1 = P_2 A = A$ and the projections $P_1 = GA$ and $P_2 = AG$. In addition, let $A_1 \equiv A^T + C$, $A_2 \equiv A^T - C$, $A_3 \equiv (I - P_{22})A_1$, $B_1 \equiv B + B^T$, $B_2 \equiv B - B^T$, $B_3 \equiv 2B - A_1^T P_{22}^T Z_2 P_{22} A_2 - A_2^T P_{22}^T Z_2 P_{22} A_1$, $G_1 \equiv A_1^{(1)}$, $G_2 \equiv A_2^{(1)}$, $G_3 \equiv [(I - P_{22})A_1]^{(1)}$, $P_{11} \equiv G_1 A_1$, $P_{12} \equiv A_1 G_1$, $P_{21} \equiv G_2 A_2$, $P_{22} \equiv A_2 G_2$, $P_{31} \equiv G_3 (I - P_{22}) A_1$, and $P_{32} \equiv [(I - P_{22})A_1]^{(1)}$. We have the following result for the solution of (49):

Theorem 5.1 *[14, Extension 2] The necessary and sufficient conditions for the solvability of (49) are:*

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

and

$$B_3 = B_2 - \left\{ \frac{1}{2} P_{11}^T B_1 P_{11} + A_2^T G_1^T B_1 (I - P_{11}) + P_{12}^T Z_1 P_{12} A_1 - \right.$$
$$\left. \left[\frac{1}{2} P_{11}^T B_1 G_1 + (I - P_{11}^T) B_1 G_1 - A_1^T (P_{12}^T Z_1 P_{12}) A_2 \right] \right\} \ , \tag{50}$$

where $Z_1^T = -Z_1$ and $Z_2^T = Z_3$.

When the above conditions are satisfied, the general solution to (49) is

$$X = \frac{1}{2} G_1^T B_1 P_{11} + G_1^T B_1 (I - P_{11}) + (P_{12}^T Z_1 P_{12}) A_1 +$$
$$(I - P_{12}^T) \left[\frac{1}{2} G_3^T B_3 P_{31} + G_3^T B_3 (I - P_{31}) + (I - P_{31}^T) Y + P_{32}^T Z P_{32} A_3 \right] \ ,$$

with Y and Z being arbitrary.

(The first G_1 inside the square brackets in (50) was mistaken to be an undefined G_{11} in [14].)

It is obvious that the above result is so complicated that it is virtually impossible to implement.

Acknowledgement

The authors wish to thank Professor Fan Hung-Yuan (National Taiwan Normal University) for many interesting and valuable suggestions on the manuscript.

References

[1] D.S. BERNSTEIN, *Matrix Mathematics: Theory, Facts, and Formulas with Applications to Linear Systems Theory*, Princeton University Press, Princeton and Oxford, 2005.

[2] A. BOJANCZYK, G. GOLUB AND P. VAN DOOREN, The periodic Schur decomposition: algorithm and applications, *Proc. SPIE Conference*, **1770** (1992) 31–42.

[3] H.W. BRADEN, The equations $A^TX \pm X^TA = B$, *SIAM J. Matrix Anal. Applic.*, **20** (1998) 295–302.

[4] R. BYERS AND D. KRESSNER, Structured condition numbers for invariant subspaces, *SIAM J. Matrix Anal. Applic.*, **28** (2006) 326–347.

[5] D. CHU, W.-W. LIN AND R.C.E. TAN, A numerical method for a generalized algebraic Riccati equation, *SIAM J. Control Optim.*, **45** (2006) 1222–1250.

[6] E.K.-W. CHU, T.-M. HWANG, W.-W. LIN AND C.-T. WU, Vibration of fast trains, palindromic eigenvalue problems and structure-preserving doubling algorithms, *J. Comput. Appl. Maths.* (to appear 2007; doi:10.1016/j.cam.2007.07.016).

[7] K.-W. E. CHU, The solution of the matrix equations $AXB-CXD=E$ and $(YA-DZ, YC-BZ)=(E,F)$, *Lin. Alg. Applic.*, **93** (1987) 93–105.

[8] J. DEMMEL AND 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) 160–174.

[9] J. DEMMEL AND 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.

[10] Z.-B. FAN AND C.-J. BU, The solution of the matrix equation $AX-X^TB=C$, *College Maths.*, **20** (2004) 100–102.

[11] G.H. GOLUB AND C.F. VAN LOAN, *Matrix Computations*, 2nd Edn., Johns Hopkins University Press, Baltimore, MD, 1989.

[12] N.J. HIGHAM, *Accuracy and Stability of Numerical Algorithms*, 2nd edn., SIAM, Philadelphia, 2002.

[13] A.R. GHAVIMI AND A.J. LAUB, Backward error, sensitivity, and refinement of computed solutions of algebraic Riccati equations, *Numer. Lin. Alg. with Applic.*, **2(1)** (1995) 29–49.

[14] X.-Z. MA, J.-H. HAN AND F.-X. PIAO, A study of solution existence for matrix equation $AX+X^TC=B$, *J. Shenyang Inst. Aero. Eng.*, **20** (2003) 64–66.

[15] MATHWORKS, *MATLAB User's Guide*, 2002.

[16] D. KRESSNER, C. SCHRÖDER AND D.S. WATKINS, Implicit QR algorithms for palindromic and even eigenvalue problems, DFG Research Centre Matheon Preprint **432**, 2008.

[17] H.-S. WANG, C.-F. YUNG AND F.-R. CHANG, Bounded real lemma and H_∞ control for descriptor systems, *IEE Proc. Control Theory Applic.*, **145** (1998) 316–322.