Consistency and efficient solution of the Sylvester equation for ∗-congruence:

\[ AX + X^*B = C \]

Fernando De Terán and Froilán M. Dopico

ICMAT and Departamento de Matemáticas, Universidad Carlos III de Madrid, Spain

Special thanks to Daniel Kressner

CEDYA 2011. Palma de Mallorca, Spain, September 5-9, 2011
Abstract

Given $A \in \mathbb{C}^{m \times n}$, $B \in \mathbb{C}^{n \times m}$, and $C \in \mathbb{C}^{m \times m}$, we study the equations

$$AX + X^* B = C,$$

where $X \in \mathbb{C}^{n \times m}$ is the unknown to be determined. More precisely:

3. Efficient and stable numerical algorithm for computing the unique solution (De Terán and D. 2011).
4. Very briefly, general solution and dimension of solution space of $AX + X^* B = 0$ (De Terán, D., Guillery, Montealegre, Reyes, 2011)

We establish parallelism/differences with well-known Sylvester equation

$$AX - XB = C,$$

$A \in \mathbb{C}^{m \times m}$, $B \in \mathbb{C}^{n \times n}$, $C \in \mathbb{C}^{m \times n}$. 
Abstract

Given $A \in \mathbb{C}^{m \times n}$, $B \in \mathbb{C}^{n \times m}$, and $C \in \mathbb{C}^{m \times m}$, we study the equations

$$AX + X^* B = C,$$

where $X \in \mathbb{C}^{n \times m}$ is the unknown to be determined. More precisely:

3. **Efficient and stable numerical algorithm for computing the unique solution** (De Terán and D. 2011).
4. **Very briefly, general solution and dimension of solution space of**

$$AX + X^* B = 0$$

(De Terán, D., Guillery, Montealegre, Reyes, 2011)

We establish parallelism/differences with **well-known Sylvester equation**

$$AX - XB = C,$$

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

F. M. Dopico (U. Carlos III, Madrid)
Abstract

Given $A \in \mathbb{C}^{m \times n}$, $B \in \mathbb{C}^{n \times m}$, and $C \in \mathbb{C}^{m \times m}$, we study the equations

$$AX + X^* B = C,$$

$$(X^* = X^T \text{ or } X^*),$$

where $X \in \mathbb{C}^{n \times m}$ is the unknown to be determined. More precisely:

3. **Efficient and stable numerical algorithm for computing the unique solution** (De Terán and D. 2011).
4. Very briefly, general solution and dimension of solution space of $AX + X^* B = 0$ (De Terán, D., Guillery, Montealegre, Reyes, 2011)

We establish parallelism/differences with **well-known Sylvester equation**

$$AX - XB = C,$$

$$A \in \mathbb{C}^{m \times m}, B \in \mathbb{C}^{n \times n}, C \in \mathbb{C}^{m \times n}. $$
Outline

1 Motivation

2 Consistency of the Sylvester equation for ⋆-congruence

3 Uniqueness of solutions

4 Efficient and stable algorithm to compute unique solutions

5 General “nonunique” solution of $AX + X^*B = C$

6 Conclusions
Outline

1. Motivation
2. Consistency of the Sylvester equation for $\star$-congruence
3. Uniqueness of solutions
4. Efficient and stable algorithm to compute unique solutions
5. General “nonunique” solution of $AX + X^*B = C$
6. Conclusions
Motivation for studying $AX + X^*B = C$ \hspace{1cm} (I)

It is well known that given a block upper triangular matrix (computed by the QR-algorithm for eigenvalues, when the matrix is real or several eigenvalues form a cluster), then

$$
\begin{bmatrix}
I & X \\
0 & I
\end{bmatrix}
\begin{bmatrix}
A & C \\
0 & B
\end{bmatrix}
\begin{bmatrix}
I & X \\
0 & I
\end{bmatrix}^{-1} =
\begin{bmatrix}
A & C - (AX - XB) \\
0 & B
\end{bmatrix}.
$$

Therefore, to find a solution of the Sylvester equation $AX - XB = C$ allows us to block-diagonalize block-triangular matrices via similarity

$$
\begin{bmatrix}
I & X \\
0 & I
\end{bmatrix}
\begin{bmatrix}
A & C \\
0 & B
\end{bmatrix}
\begin{bmatrix}
I & -X \\
0 & I
\end{bmatrix} =
\begin{bmatrix}
A & 0 \\
0 & B
\end{bmatrix}.
$$

This is indeed done in practice in numerical algorithms (LAPACK, MATLAB) to compute bases of invariant subspaces (eigenvectors) of matrices, via the classical Bartels-Stewart algorithm (Comm ACM, 1972) or level-3 BLAS variants of it Jonsson-Kågström (ACM TMS, 2002).
Motivation for studying $AX + X^* B = C$ (I)

It is well known that given a block upper triangular matrix (computed by the QR-algorithm for eigenvalues, when the matrix is real or several eigenvalues form a cluster), then

$$
\begin{bmatrix}
I & X \\
0 & I
\end{bmatrix}
\begin{bmatrix}
A & C \\
0 & B
\end{bmatrix}
\begin{bmatrix}
I & X \\
0 & I
\end{bmatrix}^{-1} = \begin{bmatrix}
A & C - (AX - XB) \\
0 & B
\end{bmatrix}.
$$

Therefore, to find a solution of the **Sylvester equation** $AX - XB = C$ allows us to block-diagonalize block-triangular matrices via similarity

$$
\begin{bmatrix}
I & X \\
0 & I
\end{bmatrix}
\begin{bmatrix}
A & C \\
0 & B
\end{bmatrix}
\begin{bmatrix}
I & -X \\
0 & I
\end{bmatrix} = \begin{bmatrix}
A & 0 \\
0 & B
\end{bmatrix}.
$$

This is indeed done in practice in numerical algorithms (LAPACK, MATLAB) to compute bases of invariant subspaces (eigenvectors) of matrices, via the classical Bartels-Stewart algorithm (Comm ACM, 1972) or level-3 BLAS variants of it Jonsson-Kågström (ACM TMS, 2002).
Motivation for studying $AX + X^*B = C$ (I)

It is well known that given a block upper triangular matrix (computed by the QR-algorithm for eigenvalues, when the matrix is real or several eigenvalues form a cluster), then

$$\begin{bmatrix} I & X \\ 0 & I \end{bmatrix} \begin{bmatrix} A & C \\ 0 & B \end{bmatrix} \begin{bmatrix} I & X \\ 0 & I \end{bmatrix}^{-1} = \begin{bmatrix} A & C - (AX - XB) \\ 0 & B \end{bmatrix}.$$

Therefore, to find a solution of the Sylvester equation $AX - XB = C$ allows us to block-diagonalize block-triangular matrices via similarity

$$\begin{bmatrix} I & X \\ 0 & I \end{bmatrix} \begin{bmatrix} A & C \\ 0 & B \end{bmatrix} \begin{bmatrix} I & -X \\ 0 & I \end{bmatrix} = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}.$$

This is indeed done in practice in numerical algorithms (LAPACK, MATLAB) to compute bases of invariant subspaces (eigenvectors) of matrices, via the classical Bartels-Stewart algorithm (Comm ACM, 1972) or level-3 BLAS variants of it Jonsson-Kågström (ACM TMS, 2002).
Motivation for studying $AX + X^* B = C$ (II)

Structured numerical algorithms for linear palindromic eigenproblems $(Z + \lambda Z^*)$ compute an anti-triangular Schur form via unitary $\star$-congruence:

**Theorem** (Kressner, Schröder, Watkins (Numer. Alg., 2009) and Mackey$^2$, Mehl, Mehrmann (NLAA, 2009))

Let $Z \in \mathbb{C}^{n \times n}$. Then there exists a unitary matrix $U \in \mathbb{C}^{n \times n}$ such that

$$M = U^* Z U = \begin{bmatrix}
* & \cdots & \cdots & * \\
\vdots & \ddots & \ddots & 0 \\
\vdots & \ddots & \ddots & \vdots \\
* & 0 & \cdots & 0
\end{bmatrix}$$

The matrix $M$ can be computed, for instance, through structure-preserving
- QR-type methods for matrices in anti-Hessenberg form (Kressner, Schröder, Watkins (Numer. Alg., 2009)),
- Jacobi-type methods (Mackey$^2$, Mehl, Mehrmann (NLAA, 2009)),

and compute eigenvalues of $Z + \lambda Z^*$ with exact pairing $\lambda, 1/\lambda^*$. 
Motivation for studying $AX + X^* B = C$ (II)

Structured numerical algorithms for linear palindromic eigenproblems $(Z + \lambda Z^*)$ compute an anti-triangular Schur form via unitary $\star$-congruence:

**Theorem (Kressner, Schröder, Watkins (Numer. Alg., 2009) and Mackey$^2$, Mehl, Mehrmann (NLAA, 2009))**

Let $Z \in \mathbb{C}^{n \times n}$. Then there exists a unitary matrix $U \in \mathbb{C}^{n \times n}$ such that

$$M = U^* Z U = \begin{bmatrix}
* & \cdots & \cdots & * \\
\vdots & \ddots & \ddots & 0 \\
\vdots & \ddots & \ddots & \vdots \\
* & 0 & \cdots & 0
\end{bmatrix}$$

The matrix $M$ can be computed, for instance, through structure-preserving

- QR-type methods for matrices in anti-Hessenberg form (Kressner, Schröder, Watkins (Numer. Alg., 2009)),
- Jacobi-type methods (Mackey$^2$, Mehl, Mehrmann (NLAA, 2009)),

and compute eigenvalues of $Z + \lambda Z^*$ with exact pairing $\lambda, 1/\lambda^*$. 
Motivation for studying $AX + X^* B = C$ (II)

Structured numerical algorithms for linear palindromic eigenproblems $(Z + \lambda Z^*)$ compute an **anti-triangular Schur form** via unitary $\star$-congruence:

**Theorem (Kressner, Schröder, Watkins (Numer. Alg., 2009) and Mackey^2, Mehl, Mehrmann (NLAA, 2009))**

Let $Z \in \mathbb{C}^{n \times n}$. Then there exists a unitary matrix $U \in \mathbb{C}^{n \times n}$ such that

$$M = U^* Z U = \begin{bmatrix}
* & \cdots & \cdots & * \\
\vdots & \ddots & \ddots & \vdots \\
\vdots & \ddots & 0 \\
* & 0 & \cdots & 0 
\end{bmatrix}$$

The matrix $M$ can be computed, for instance, through **structure-preserving**

- QR-type methods for matrices in anti-Hessenberg form (Kressner, Schröder, Watkins (Numer. Alg., 2009)),
- Jacobi-type methods (Mackey^2, Mehl, Mehrmann (NLAA, 2009)),

and compute eigenvalues of $Z + \lambda Z^*$ with **exact pairing** $\lambda, 1/\lambda^*$. 
Motivation for studying $AX + X^*B = C$ (III)

Given a block upper ANTI-triangular matrix (computed via structured algorithms for linear palindromic eigenproblems, when the matrix is real or several eigenvalues form a cluster), then

$$\begin{bmatrix} I & 0 \\ -X & I \end{bmatrix}^* \begin{bmatrix} C & A \\ B & 0 \end{bmatrix} \begin{bmatrix} I & 0 \\ -X & I \end{bmatrix} = \begin{bmatrix} C - (AX + X^*B) & A \\ B & 0 \end{bmatrix}. $$

Therefore, to find a solution of the Sylvester equation for $\star$-congruence allows us to block-ANTI-diagonalize block-ANTI-triangular matrices via $\star$-congruence

$$\begin{bmatrix} I & -X^* \\ 0 & I \end{bmatrix} \begin{bmatrix} C & A \\ B & 0 \end{bmatrix} \begin{bmatrix} I & 0 \\ -X & I \end{bmatrix} = \begin{bmatrix} 0 & A \\ B & 0 \end{bmatrix},$$

and to compute deflating subspaces of palindromic pencils.

**GOAL:** To understand Sylvester equations for $\star$-congruence and develop efficient and stable numerical algorithms for its solution in order to completely solve the linear palindromic eigenproblem numerically and to determine the conditioning of its deflating subspaces under structured perturbations.
Motivation for studying $AX + X^* B = C$ (III)

Given a block upper ANTI-triangular matrix (computed via structured algorithms for linear palindromic eigenproblems, when the matrix is real or several eigenvalues form a cluster), then

$$
\begin{bmatrix}
I & 0 \\
-X & I
\end{bmatrix}^* \begin{bmatrix}
C & A \\
B & 0
\end{bmatrix} \begin{bmatrix}
I & 0 \\
-X & I
\end{bmatrix} = \begin{bmatrix}
C - (AX + X^* B) & A \\
B & 0
\end{bmatrix}.
$$

Therefore, to find a solution of the Sylvester equation for $\star$-congruence allows us to block-ANTI-diagonalize block-ANTI-triangular matrices via $\star$-congruence

$$
\begin{bmatrix}
I & -X^* \\
0 & I
\end{bmatrix} \begin{bmatrix}
C & A \\
B & 0
\end{bmatrix} \begin{bmatrix}
I & 0 \\
-X & I
\end{bmatrix} = \begin{bmatrix}
0 & A \\
B & 0
\end{bmatrix},
$$

and to compute deflating subspaces of palindromic pencils.

**GOAL:** To understand Sylvester equations for $\star$-congruence and develop efficient and stable numerical algorithms for its solution in order to completely solve the linear palindromic eigenproblem numerically and to determine the conditioning of its deflating subspaces under structured perturbations.
Motivation for studying $AX + X^*B = C$ (III)

Given a block upper ANTI-triangular matrix (computed via structured algorithms for linear palindromic eigenproblems, when the matrix is real or several eigenvalues form a cluster), then

$$\begin{bmatrix} I & 0 \\ -X & I \end{bmatrix}^* \begin{bmatrix} C & A \\ B & 0 \end{bmatrix} \begin{bmatrix} I & 0 \\ -X & I \end{bmatrix} = \begin{bmatrix} C - (AX + X^*B) & A \\ B & 0 \end{bmatrix}.$$ 

Therefore, to find a solution of the Sylvester equation for $\star$-congruence allows us to block-ANTI-diagonalize block-ANTI-triangular matrices via $\star$-congruence

$$\begin{bmatrix} I & -X^* \\ 0 & I \end{bmatrix} \begin{bmatrix} C & A \\ B & 0 \end{bmatrix} \begin{bmatrix} I & 0 \\ -X & I \end{bmatrix} = \begin{bmatrix} 0 & A \\ B & 0 \end{bmatrix},$$

and to compute deflating subspaces of palindromic pencils.

**GOAL:** To understand Sylvester equations for $\star$-congruence and develop efficient and stable numerical algorithms for its solution in order to completely solve the linear palindromic eigenproblem numerically and to determine the conditioning of its deflating subspaces under structured perturbations.
1 Motivation

2 Consistency of the Sylvester equation for $\star$-congruence

3 Uniqueness of solutions

4 Efficient and stable algorithm to compute unique solutions

5 General “nonunique” solution of $AX + X^*B = C$

6 Conclusions
Consistency of $AX + X^*B = C$

**Theorem (Wimmer (LAA, 1994), De Terán and D. (ELA, 2011))**

Let $\mathbb{F}$ be a field of characteristic different from two and let $A \in \mathbb{F}^{m \times n}$, $B \in \mathbb{F}^{n \times m}$, $C \in \mathbb{F}^{m \times m}$ be given. There is some $X \in \mathbb{F}^{n \times m}$ such that

$$AX + X^*B = C$$

if and only if

$$\begin{bmatrix} C & A \\ B & 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 & A \\ B & 0 \end{bmatrix} \quad \text{are } \ast\text{-congruent.}$$

**Remarks:**

- The implication $\implies$ very easy: done in previous slide.
- The implication $\impliedby$ more challenging.
- Wimmer proved in 1994 the result, for $\mathbb{F} = \mathbb{C}$ and $\ast = \ast$, without any reference to palindromic eigenproblems.
- His motivation was the study of standard Sylvester equations with Hermitian solutions.
Consistency of $AX + X^* B = C$

**Theorem (Wimmer (LAA, 1994), De Terán and D. (ELA, 2011))**

Let $\mathbb{F}$ be a field of characteristic different from two and let $A \in \mathbb{F}^{m \times n}$, $B \in \mathbb{F}^{n \times m}$, $C \in \mathbb{F}^{m \times m}$ be given. There is some $X \in \mathbb{F}^{n \times m}$ such that

$$AX + X^* B = C$$

if and only if

$$\begin{bmatrix} C & A \\ B & 0 \end{bmatrix} \text{ and } \begin{bmatrix} 0 & A \\ B & 0 \end{bmatrix} \text{ are } \ast\text{-congruent.}$$

**Remarks:**

- The implication $\implies$ very easy: done in previous slide.
- The implication $\impliedby$ more challenging.
- Wimmer proved in 1994 the result, for $\mathbb{F} = \mathbb{C}$ and $\ast = \ast$, without any reference to palindromic eigenproblems.
- His motivation was the study of standard Sylvester equations with Hermitian solutions.
...to be compared with Roth’s criterion for standard Sylvester equation

**Theorem (Roth (Proc. AMS, 1952))**

Let $\mathbb{F}$ be any field and let $A \in \mathbb{F}^{m \times m}$, $B \in \mathbb{F}^{n \times n}$, $C \in \mathbb{F}^{m \times n}$ be given. There is some $X \in \mathbb{F}^{m \times n}$ such that

$$AX - XB = C$$

if and only if

$$\begin{bmatrix} A & C \\ 0 & B \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$$

are similar.
Outline

1 Motivation

2 Consistency of the Sylvester equation for ⋆-congruence

3 Uniqueness of solutions

4 Efficient and stable algorithm to compute unique solutions

5 General “nonunique” solution of $AX + X^*B = C$

6 Conclusions
Uniqueness of solutions of $AX + X^* B = C$  \hfill (I)

Remarks:

- If the matrices $A \in \mathbb{F}^{m \times n}$ and $B \in \mathbb{F}^{n \times m}$ are rectangular ($m \neq n$), then the equation does not have a unique solution for every right-hand side $C$,

- that is, the operator

\[
\mathbb{F}^{n \times m} \rightarrow \mathbb{F}^{m \times m}
\]

\[
X \mapsto AX + X^* B
\]

is never invertible.

- It is of course possible that $m > n$ and that for particular $A$, $B$ and $C$, a solution exists and is unique,

- but we restrict ourselves here to the square case $m = n$. 
Uniqueness of solutions of $AX + X^* B = C$ (I)

Remarks:

- If the matrices $A \in \mathbb{F}^{m \times n}$ and $B \in \mathbb{F}^{n \times m}$ are rectangular $(m \neq n)$, then the equation does not have a unique solution for every right-hand side $C$,

- that is, the operator

$$\mathbb{F}^{n \times m} \longrightarrow \mathbb{F}^{m \times m}$$

$$X \longmapsto AX + X^* B$$

is never invertible.

- It is of course possible that $m > n$ and that for particular $A$, $B$ and $C$, a solution exists and is unique,

- but we restrict ourselves here to the square case $m = n$. 

F. M. Dopico (U. Carlos III, Madrid)

Sylvester equation for congruence  
CEDYA 2011 12 / 24
Remarks:

- If the matrices $A \in \mathbb{F}^{m \times n}$ and $B \in \mathbb{F}^{n \times m}$ are rectangular ($m \neq n$), then the equation does not have a unique solution for every right-hand side $C$,

- that is, the operator

$$\begin{align*}
\mathbb{F}^{n \times m} & \longrightarrow \mathbb{F}^{m \times m} \\
X & \longmapsto AX + X^* B
\end{align*}$$

is never invertible.

- It is of course possible that $m > n$ and that for particular $A$, $B$ and $C$, a solution exists and is unique,

- but we restrict ourselves here to the square case $m = n$. 

F. M. Dopico (U. Carlos III, Madrid)  Sylvester equation for congruence  CEDYA 2011
Uniqueness of solutions of \( AX + X^* B = C \) \( \text{ (I)} \)

**Remarks:**

- If the matrices \( A \in \mathbb{F}^{m \times n} \) and \( B \in \mathbb{F}^{n \times m} \) are rectangular \((m \neq n)\), then the equation **does not have a unique solution for every right-hand side** \( C \),

- that is, **the operator**

\[
\begin{array}{c}
\mathbb{F}^{n \times m} \\
X
\end{array} \quad \longrightarrow \quad \begin{array}{c}
\mathbb{F}^{m \times m} \\
AX + X^* B
\end{array}
\]

**is never invertible.**

- It is of course possible that \( m > n \) and that for particular \( A, B \) and \( C \), a solution exists and is unique,

- but **we restrict ourselves here to the square case** \( m = n \).
Uniqueness of solutions of $AX + X^*B = C$ (II)

Definition: a set $\{\lambda_1, \ldots, \lambda_n\} \subset \mathbb{C}$ is $\star$-reciprocal free if $\lambda_i \neq 1/\lambda_j^*$ for any $1 \leq i, j \leq n$. We admit $0$ and/or $\infty$ as elements of $\{\lambda_1, \ldots, \lambda_n\}$.

Theorem (Byers, Kressner (SIMAX, 2006), Kressner, Schröder, Watkins, (Num. Alg., 2009))

Let $A, B \in \mathbb{C}^{n \times n}$ be given. Then:

- $AX + X^T B = C$ has a unique solution $X$ for every right-hand side $C \in \mathbb{C}^{n \times n}$ if and only if the following conditions hold:
  1) The pencil $A - \lambda B^T$ is regular, and
  2) the set of eigenvalues of $A - \lambda B^T \setminus \{1\}$ is $T$-reciprocal free and if $1$ is an eigenvalue of $A - \lambda B^T$, then it has algebraic multiplicity $1$.

- $AX + X^* B = C$ has a unique solution $X$ for every right-hand side $C \in \mathbb{C}^{n \times n}$ if and only if the following conditions hold:
  1) The pencil $A - \lambda B^*$ is regular, and
  2) the set of eigenvalues of $A - \lambda B^*$ is $\star$-reciprocal free.
Uniqueness of solutions of $AX + XB = C$ (II)

**Definition:** A set $\{\lambda_1, \ldots, \lambda_n\} \subset \mathbb{C}$ is $\star$-reciprocal free if $\lambda_i \neq 1/\lambda_j^\star$ for any $1 \leq i, j \leq n$. We admit 0 and/or $\infty$ as elements of $\{\lambda_1, \ldots, \lambda_n\}$.

**Theorem (Byers, Kressner (SIMAX, 2006), Kressner, Schröder, Watkins, (Num. Alg., 2009))**

Let $A, B \in \mathbb{C}^{n \times n}$ be given. Then:

- $AX + X^T B = C$ has a unique solution $X$ for every right-hand side $C \in \mathbb{C}^{n \times n}$ if and only if the following conditions hold:
  1) The pencil $A - \lambda B^T$ is regular, and
  2) the set of eigenvalues of $A - \lambda B^T \setminus \{1\}$ is $T$-reciprocal free and if 1 is an eigenvalue of $A - \lambda B^T$, then it has algebraic multiplicity 1.

- $AX + X^\star B = C$ has a unique solution $X$ for every right-hand side $C \in \mathbb{C}^{n \times n}$ if and only if the following conditions hold:
  1) The pencil $A - \lambda B^\star$ is regular, and
  2) the set of eigenvalues of $A - \lambda B^\star$ is $\star$-reciprocal free.
Uniqueness of solutions of $AX + X^* B = C$ (II)

**Definition:** a set $\{\lambda_1, \ldots, \lambda_n\} \subset \mathbb{C}$ is $\ast$-reciprocal free if $\lambda_i \neq 1/\lambda_j^\ast$ for any $1 \leq i, j \leq n$. We admit 0 and/or $\infty$ as elements of $\{\lambda_1, \ldots, \lambda_n\}$.

**Theorem (Byers, Kressner (SIMAX, 2006), Kressner, Schröder, Watkins, (Num. Alg., 2009))**

Let $A, B \in \mathbb{C}^{n \times n}$ be given. Then:

- $AX + X^T B = C$ has a unique solution $X$ for every right-hand side $C \in \mathbb{C}^{n \times n}$ if and only if the following conditions hold:
  1) The pencil $A - \lambda B^T$ is regular, and
  2) the set of eigenvalues of $A - \lambda B^T \setminus \{1\}$ is $T$-reciprocal free and if 1 is an eigenvalue of $A - \lambda B^T$, then it has algebraic multiplicity 1.

- $AX + X^* B = C$ has a unique solution $X$ for every right-hand side $C \in \mathbb{C}^{n \times n}$ if and only if the following conditions hold:
  1) The pencil $A - \lambda B^\ast$ is regular, and
  2) the set of eigenvalues of $A - \lambda B^\ast$ is $\ast$-reciprocal free.
...to be compared with uniqueness conditions for standard Sylvester eq

**Theorem**

Let $A \in \mathbb{C}^{m \times m}$ and $B \in \mathbb{C}^{n \times n}$ be given. Then:

- $AX - XB = C$ has a unique solution $X$ for every right-hand side $C \in \mathbb{C}^{m \times n}$ if and only if $A$ and $B$ have no eigenvalues in common.

**Remark:** Comparison of both results brings to our attention a key difference that appears always between solution methods for $AX + X^*B = C$ and $AX - XB = C$:

- In $AX + X^*B = C$, one starts dealing with the eigenproblem of $A - \lambda B^*$, that is, one deals from the very beginning simultaneously with $A$ and $B$.
- By contrast in $AX - XB = C$, one starts dealing independently with the eigenproblems of $A$ and $B$.  

F. M. Dopico (U. Carlos III, Madrid)  
Sylvester equation for congruence  
CEDYA 2011 14 / 24
Theorem

Let $A \in \mathbb{C}^{m \times m}$ and $B \in \mathbb{C}^{n \times n}$ be given. Then:

- $AX - XB = C$ has a unique solution $X$ for every right-hand side $C \in \mathbb{C}^{m \times n}$ if and only if $A$ and $B$ have no eigenvalues in common.

Remark: Comparison of both results brings to our attention a key difference that appears always between solution methods for $AX + X*B = C$ and $AX - XB = C$:

- In $AX + X*B = C$, one starts dealing with the eigenproblem of $A - \lambda B*$, that is, one deals from the very beginning simultaneously with $A$ and $B$.
- By contrast in $AX - XB = C$, one starts dealing independently with the eigenproblems of $A$ and $B$. 
Theorem

Let $A \in \mathbb{C}^{m \times m}$ and $B \in \mathbb{C}^{n \times n}$ be given. Then:

- $AX - XB = C$ has a unique solution $X$ for every right-hand side $C \in \mathbb{C}^{m \times n}$ if and only if $A$ and $B$ have no eigenvalues in common.

Remark: Comparison of both results brings to our attention a key difference that appears always between solution methods for $AX + X*B = C$ and $AX - XB = C$:

- In $AX + X*B = C$, one starts dealing with the eigenproblem of $A - \lambda B^*$, that is, one deals from the very beginning simultaneously with $A$ and $B$.
- By contrast in $AX - XB = C$, one starts dealing independently with the eigenproblems of $A$ and $B$.
Theorem

Let $A \in \mathbb{C}^{m \times m}$ and $B \in \mathbb{C}^{n \times n}$ be given. Then:

- $AX - XB = C$ has a unique solution $X$ for every right-hand side $C \in \mathbb{C}^{m \times n}$ if and only if $A$ and $B$ have no eigenvalues in common.

Remark: Comparison of both results brings to our attention a key difference that appears always between solution methods for $AX + X^*B = C$ and $AX - XB = C$:

- In $AX + X^*B = C$, one starts dealing with the eigenproblem of $A - \lambda B^*$, that is, one deals from the very beginning simultaneously with $A$ and $B$.

- By contrast in $AX - XB = C$, one starts dealing independently with the eigenproblems of $A$ and $B$. 

F. M. Dopico (U. Carlos III, Madrid)

Sylvester equation for congruence

CEDYA 2011
Outline

1. Motivation

2. Consistency of the Sylvester equation for $*$-congruence

3. Uniqueness of solutions

4. Efficient and stable algorithm to compute unique solutions

5. General “nonunique” solution of $AX + X^*B = C$

6. Conclusions
The fundamental transformation

In this section in $AX + X^* B = C$ all matrices are in $\mathbb{C}^{n \times n}$ and the solution is unique for every $C$.

$AX + X^* B = C$ is equivalent to a linear system for the $n^2$ entries of $X$ if $\star = T$ and to a linear system for the $2n^2$ entries of $(\text{Re} X, \text{Im} X)$ if $\star = \ast$. From now on, we say simply “linear system” for $X$.

Then, it is possible to use Gaussian elimination on the equivalent system (constructed via $\text{vec}(X)$, $\text{vec}(C)$, $\otimes$), but it costs $O(n^6)$ flops, which is prohibitive except for small $n$.

IDEA: transform $AX + X^* B = C$ into an equation of the same type but with much simpler coefficients instead of $A$ and $B$ and that can be easily solved to get a total cost of $O(n^3)$ flops.

To this purpose, use QZ algorithm to compute in $O(n^3)$ flops the generalized Schur decomposition of

$$A - \lambda B^* = U(R - \lambda S)V,$$

where $\left\{ \begin{array}{ll} R, S & \text{are upper triangular} \\ U, V & \text{are unitary matrices} \end{array} \right.$$

If $A, B$ real matrices: use real arithmetic to get quasi-triangular $R$. We do not consider this for brevity.
The fundamental transformation

- In this section in \( AX + X^* B = C \) all matrices are in \( \mathbb{C}^{n \times n} \) and the solution is unique for every \( C \).

- \( AX + X^* B = C \) is equivalent to a linear system for the \( n^2 \) entries of \( X \) if \( * = T \) and to a linear system for the \( 2n^2 \) entries of \( (\text{Re} X, \text{Im} X) \) if \( * = * \). From now on, we say simply “linear system” for \( X \).

- Then, it is possible to use Gaussian elimination on the equivalent system (constructed via \( \text{vec}(X), \text{vec}(C), \otimes \)), but it costs \( O(n^6) \) flops, which is prohibitive except for small \( n \).

- **IDEA:** transform \( AX + X^* B = C \) into an equation of the same type but with much simpler coefficients instead of \( A \) and \( B \) and that can be easily solved to get a total cost of \( O(n^3) \) flops.

- To this purpose, use **QZ algorithm** to compute in \( O(n^3) \) flops the generalized Schur decomposition of

\[
A - \lambda B^* = U(R - \lambda S)V, \quad \text{where} \quad \begin{cases} R, S \\ U, V \end{cases} \text{ are upper triangular} \quad \begin{cases} \text{are unitary matrices} \\
\end{cases}
\]

If \( A, B \) real matrices: use real arithmetic to get *quasi-triangular* \( R \). We do not consider this for brevity.
The fundamental transformation

- In this section in $AX + X^* B = C$ all matrices are in $\mathbb{C}^{n \times n}$ and the solution is unique for every $C$.

- $AX + X^* B = C$ is equivalent to a linear system for the $n^2$ entries of $X$ if $\star = T$ and to a linear system for the $2n^2$ entries of $(\Re X, \Im X)$ if $\star = \ast$. From now on, we say simply “linear system” for $X$.

- Then, it is possible to use Gaussian elimination on the equivalent system (constructed via $\text{vec}(X), \text{vec}(C), \otimes$), but it costs $O(n^6)$ flops, which is prohibitive except for small $n$.

- IDEA: transform $AX + X^* B = C$ into an equation of the same type but with much simpler coefficients instead of $A$ and $B$ and that can be easily solved to get a total cost of $O(n^3)$ flops.

- To this purpose, use QZ algorithm to compute in $O(n^3)$ flops the generalized Schur decomposition of

$$A - \lambda B^* = U(R - \lambda S)V,$$

where

$$\{ R, S \} \text{ are upper triangular}$$

$$U, V \text{ are unitary matrices}$$

If $A, B$ real matrices: use real arithmetic to get quasi-triangular $R$. We do not consider this for brevity.
In this section in $AX + X^* B = C$ all matrices are in $\mathbb{C}^{n \times n}$ and the solution is unique for every $C$.

$AX + X^* B = C$ is equivalent to a linear system for the $n^2$ entries of $X$ if $\star = T$ and to a linear system for the $2n^2$ entries of $(\text{Re} X, \text{Im} X)$ if $\star = \circ$. From now on, we say simply “linear system” for $X$.

Then, it is possible to use Gaussian elimination on the equivalent system (constructed via $\text{vec}(X), \text{vec}(C), \otimes$), but it costs $O(n^6)$ flops, which is prohibitive except for small $n$.

**IDEA**: transform $AX + X^* B = C$ into an equation of the same type but with much simpler coefficients instead of $A$ and $B$ and that can be easily solved to get a total cost of $O(n^3)$ flops.

To this purpose, use QZ algorithm to compute in $O(n^3)$ flops the generalized Schur decomposition of

$$A - \lambda B^* = U(R - \lambda S)V,$$

where

$$\begin{cases} 
R, S \\
U, V
\end{cases}$$

are upper triangular matrices.

If $A, B$ real matrices: use real arithmetic to get quasi-triangular $R$. We do not consider this for brevity.
The fundamental transformation

In this section in $AX + X^* B = C$ all matrices are in $\mathbb{C}^{n \times n}$ and the solution is unique for every $C$.

$AX + X^* B = C$ is equivalent to a linear system for the $n^2$ entries of $X$ if $\star = T$ and to a linear system for the $2n^2$ entries of $(\text{Re} \ X, \text{Im} \ X)$ if $\star = \ast$. From now on, we say simply “linear system” for $X$.

Then, it is possible to use Gaussian elimination on the equivalent system (constructed via $\text{vec}(X), \text{vec}(C), \otimes$), but it costs $O(n^6)$ flops, which is prohibitive except for small $n$.

**IDEA:** transform $AX + X^* B = C$ into an equation of the same type but with much simpler coefficients instead of $A$ and $B$ and that can be easily solved to get a total cost of $O(n^3)$ flops.

To this purpose, use **QZ algorithm** to compute in $O(n^3)$ flops the generalized Schur decomposition of

$$A - \lambda B^* = U(R - \lambda S)V,$$

where

\[
\begin{align*}
R, S & \text{ are upper triangular} \\
U, V & \text{ are unitary matrices}
\end{align*}
\]

If $A, B$ real matrices: use real arithmetic to get quasitriangular $R$. We do not consider this for brevity.
Algorithm to solve $AX + X^* B = C$ in $O(n^3)$ flops

**INPUT:** $A, B, C \in \mathbb{C}^{n \times n}$

**OUTPUT:** $X \in \mathbb{C}^{n \times n}$

**Step 1.** Compute via QZ algorithm $R, S, U$ and $V$ such that

$$A = U R V, \quad B^* = U S V,$$

where $\left\{ \begin{align*} R, S & \text{ are upper triangular} \\ U, V & \text{ are unitary matrices} \end{align*} \right.$$

**Step 2.** Compute $E = U^* C (U^*)^*$

**Step 3.** Solve for $W \in \mathbb{C}^{n \times n}$ the transformed equation

$$R W + W^* S^* = E$$

**Step 4.** Compute $X = V^* W U^*$
Algorithm to solve $AX + X^* B = C$ in $O(n^3)$ flops

**INPUT:** $A, B, C \in \mathbb{C}^{n \times n}$

**OUTPUT:** $X \in \mathbb{C}^{n \times n}$

**Step 1.** Compute via QZ algorithm $R, S, U$ and $V$ such that

$$A = URV, \quad B^* = USV,$$

where

$$\begin{cases} R, S \text{ are upper triangular} \\ U, V \text{ are unitary matrices} \end{cases}$$

**Step 2.** Compute $E = U^* C (U^*)^*$

**Step 3.** How to solve for $W \in \mathbb{C}^{n \times n}$ the transformed equation

$$RW + W^* S^* = E?$$

**Step 4.** Compute $X = V^* W U^*$
Algorithm to solve the transformed equation $RW + W^*S^* = E$ \hspace{1cm} (I)

We illustrate with $4 \times 4$ example for simplicity:

$$
\begin{bmatrix}
    r_{11} & r_{12} & r_{13} & r_{14} \\
    0 & r_{22} & r_{23} & r_{24} \\
    0 & 0 & r_{33} & r_{34} \\
    0 & 0 & 0 & r_{44}
\end{bmatrix}
+ 
\begin{bmatrix}
    w_{11} & w_{12} & w_{13} & w_{14} \\
    w_{21} & w_{22} & w_{23} & w_{24} \\
    w_{31} & w_{32} & w_{33} & w_{34} \\
    w_{41} & w_{42} & w_{43} & w_{44}
\end{bmatrix}
= 
\begin{bmatrix}
    e_{11} & e_{12} & e_{13} & e_{14} \\
    e_{21} & e_{22} & e_{23} & e_{24} \\
    e_{31} & e_{32} & e_{33} & e_{34} \\
    e_{41} & e_{42} & e_{43} & e_{44}
\end{bmatrix}
$$

If we equate the $(4,4)$-entry, then we get

$$r_{44} \quad w_{44} \quad + \quad w_{44}^* \quad s_{44}^* \quad = \quad e_{44} ,$$

a scalar equation that allows us to determine $w_{44}$. 

Algorithm to solve the transformed equation \( RW + W^* S^* = E \) (I)

We illustrate with \( 4 \times 4 \) example for simplicity:

\[
\begin{bmatrix}
  r_{11} & r_{12} & r_{13} & r_{14} \\
  0 & r_{22} & r_{23} & r_{24} \\
  0 & 0 & r_{33} & r_{34} \\
  0 & 0 & 0 & r_{44}
\end{bmatrix}
+ \begin{bmatrix}
  w_{11}^* & w_{21}^* & w_{31}^* & w_{41}^* \\
  w_{12}^* & w_{22}^* & w_{32}^* & w_{42}^* \\
  w_{13}^* & w_{23}^* & w_{33}^* & w_{43}^* \\
  w_{14}^* & w_{24}^* & w_{34}^* & w_{44}^*
\end{bmatrix}
\begin{bmatrix}
  s_{11}^* & 0 & 0 & 0 \\
  s_{12}^* & s_{22}^* & 0 & 0 \\
  s_{13}^* & s_{23}^* & s_{33}^* & 0 \\
  s_{14}^* & s_{24}^* & s_{34}^* & s_{44}^*
\end{bmatrix}
= \begin{bmatrix}
  e_{11} & e_{12} & e_{13} & e_{14} \\
  e_{21} & e_{22} & e_{23} & e_{24} \\
  e_{31} & e_{32} & e_{33} & e_{34} \\
  e_{41} & e_{42} & e_{43} & e_{44}
\end{bmatrix}
\]

If we equate the \((4,4)\)-entry, then we get

\[
 r_{44} \quad w_{44}^* + \quad w_{44}^* \quad s_{44}^* = e_{44} ,
\]

a scalar equation that allows us to determine \( w_{44} \).
Algorithm to solve the transformed equation $RW + W^* S^* = E$ (I)

We illustrate with $4 \times 4$ example for simplicity:

\[
\begin{bmatrix}
  r_{11} & r_{12} & r_{13} & r_{14} \\
  0 & r_{22} & r_{23} & r_{24} \\
  0 & 0 & r_{33} & r_{34} \\
  0 & 0 & 0 & r_{44}
\end{bmatrix}
+\begin{bmatrix}
  w_{11}^* & w_{21}^* & w_{31}^* & w_{41}^* \\
  w_{12}^* & w_{22}^* & w_{32}^* & w_{42}^* \\
  w_{13}^* & w_{23}^* & w_{33}^* & w_{43}^* \\
  w_{14}^* & w_{24}^* & w_{34}^* & w_{44}^*
\end{bmatrix}
\begin{bmatrix}
  s_{11}^* & 0 & 0 & 0 \\
  s_{12}^* & s_{22}^* & 0 & 0 \\
  s_{13}^* & s_{23}^* & s_{33}^* & 0 \\
  s_{14}^* & s_{24}^* & s_{34}^* & s_{44}^*
\end{bmatrix}
= \begin{bmatrix}
  e_{11} & e_{12} & e_{13} & e_{14} \\
  e_{21} & e_{22} & e_{23} & e_{24} \\
  e_{31} & e_{32} & e_{33} & e_{34} \\
  e_{41} & e_{42} & e_{43} & e_{44}
\end{bmatrix}
\]

If we equate the $(3,4)$ and $(4,3)$ entries, then we get

\[
s_{33} \ w_{34} + w_{43}^* \ r_{44}^* = e_{43} - s_{34} \ w_{44}^* \\
r_{33} \ w_{34} + w_{43}^* \ s_{44}^* = e_{34} - r_{34} \ w_{44}^*
\]

a $2 \times 2$ system of scalar equations that allows us to determine $w_{34}$ and $w_{43}$ simultaneously.
Algorithm to solve the transformed equation \( RW + W^* S^* = E \) (I)

We illustrate with a \( 4 \times 4 \) example for simplicity:

\[
\begin{bmatrix}
  r_{11} & r_{12} & r_{13} & r_{14} \\
  0 & r_{22} & r_{23} & r_{24} \\
  0 & 0 & r_{33} & r_{34} \\
  0 & 0 & 0 & r_{44}
\end{bmatrix}
+ \begin{bmatrix}
  w_{11}^* & w_{21}^* & w_{31}^* & w_{41}^* \\
  w_{12}^* & w_{22}^* & w_{32}^* & w_{42}^* \\
  w_{13}^* & w_{23}^* & w_{33}^* & w_{43}^* \\
  w_{14}^* & w_{24}^* & w_{34}^* & w_{44}^*
\end{bmatrix}
= \begin{bmatrix}
  s_{11}^* & 0 & 0 & 0 \\
  s_{12}^* & s_{22}^* & 0 & 0 \\
  s_{13}^* & s_{23}^* & s_{33}^* & 0 \\
  s_{14}^* & s_{24}^* & s_{34}^* & s_{44}^*
\end{bmatrix}
= \begin{bmatrix}
  e_{11} & e_{12} & e_{13} & e_{14} \\
  e_{21} & e_{22} & e_{23} & e_{24} \\
  e_{31} & e_{32} & e_{33} & e_{34} \\
  e_{41} & e_{42} & e_{43} & e_{44}
\end{bmatrix}
\]

If we equate the (3,4) and (4,3) entries, then we get

\[
s_{33} w_{34} + w_{43}^* r_{44} = e_{43}^* - s_{34} w_{44}
\]

\[
r_{33} w_{34} + w_{43}^* s_{44} = e_{34} - r_{34} w_{44}
\]

a \( 2 \times 2 \) system of scalar equations that allows us to determine \( w_{34} \) and \( w_{43} \) simultaneously.
Algorithm to solve the transformed equation $RW + W^* S^* = E$ (I)

We illustrate with $4 \times 4$ example for simplicity:

$$
\begin{bmatrix}
    r_{11} & r_{12} & r_{13} & r_{14} \\
    0 & r_{22} & r_{23} & r_{24} \\
    0 & 0 & r_{33} & r_{34} \\
    0 & 0 & 0 & r_{44}
\end{bmatrix}
+ 
\begin{bmatrix}
    w_{11}^* & w_{21}^* & w_{31}^* & w_{41}^* \\
    w_{12}^* & w_{22}^* & w_{32}^* & w_{42}^* \\
    w_{13}^* & w_{23}^* & w_{33}^* & w_{43}^* \\
    w_{14}^* & w_{24}^* & w_{34}^* & w_{44}^*
\end{bmatrix}
= 
\begin{bmatrix}
    w_{11} & w_{12} & w_{13} & w_{14} \\
    w_{21} & w_{22} & w_{23} & w_{24} \\
    w_{31} & w_{32} & w_{33} & w_{34} \\
    w_{41} & w_{42} & w_{43} & w_{44}
\end{bmatrix}
- 
\begin{bmatrix}
    s_{11}^* & 0 & 0 & 0 \\
    s_{12}^* & s_{22}^* & 0 & 0 \\
    s_{13}^* & s_{23}^* & s_{33}^* & 0 \\
    s_{14}^* & s_{24}^* & s_{34}^* & s_{44}^*
\end{bmatrix}
= 
\begin{bmatrix}
    e_{11} & e_{12} & e_{13} & e_{14} \\
    e_{21} & e_{22} & e_{23} & e_{24} \\
    e_{31} & e_{32} & e_{33} & e_{34} \\
    e_{41} & e_{42} & e_{43} & e_{44}
\end{bmatrix}
$$

If we equate the $(2,4)$ and $(4,2)$ entries, then we get

$$
\begin{aligned}
    s_{22} & w_{24} + w_{42}^* & r_{44}^* &= e_{42}^* - s_{23} & w_{34}^* - s_{24} & w_{44} \\
    r_{22} & w_{24} + w_{42}^* & s_{44}^* &= e_{24} - r_{23} & w_{34}^* - r_{24} & w_{44}
\end{aligned}
$$

a $2 \times 2$ system of scalar equations that allows us to determine $w_{24}$ and $w_{42}$ simultaneously.
Algorithm to solve the transformed equation $RW + W^* S^* = E$ (I)

We illustrate with $4 \times 4$ example for simplicity:

$$
\begin{bmatrix}
    r_{11} & r_{12} & r_{13} & r_{14} \\
    0 & r_{22} & r_{23} & r_{24} \\
    0 & 0 & r_{33} & r_{34} \\
    0 & 0 & 0 & r_{44}
\end{bmatrix}
+ 
\begin{bmatrix}
    w_{11}^* & w_{21}^* & w_{31}^* & w_{41}^* \\
    w_{12}^* & w_{22}^* & w_{32}^* & w_{42}^* \\
    w_{13}^* & w_{23}^* & w_{33}^* & w_{43}^* \\
    w_{14}^* & w_{24}^* & w_{34}^* & w_{44}^*
\end{bmatrix}
\begin{bmatrix}
    s_{11}^* & 0 & 0 & 0 \\
    s_{12}^* & s_{22}^* & 0 & 0 \\
    s_{13}^* & s_{23}^* & s_{33}^* & 0 \\
    s_{14}^* & s_{24}^* & s_{34}^* & s_{44}^*
\end{bmatrix}
= 
\begin{bmatrix}
    e_{11} & e_{12} & e_{13} & e_{14} \\
    e_{21} & e_{22} & e_{23} & e_{24} \\
    e_{31} & e_{32} & e_{33} & e_{34} \\
    e_{41} & e_{42} & e_{43} & e_{44}
\end{bmatrix}
$$

If we equate the (2,4) and (4,2) entries, then we get

$$
s_{22} w_{24} + w_{42}^* r_{44}^* = e_{42}^* - s_{23}^* w_{34} - s_{24}^* w_{44}^*,
$$

$$
r_{22} w_{24} + w_{42}^* s_{44}^* = e_{24}^* - r_{23}^* w_{34} - r_{24}^* w_{44}^*,
$$

a $2 \times 2$ system of scalar equations that allows us to determine $w_{24}$ and $w_{42}$ simultaneously.
Algorithm to solve the transformed equation \( RW + W^* S^* = E \) \( (I) \)

We illustrate with \( 4 \times 4 \) example for simplicity:

\[
\begin{bmatrix}
  r_{11} & r_{12} & r_{13} & r_{14} \\
  0 & r_{22} & r_{23} & r_{24} \\
  0 & 0 & r_{33} & r_{34} \\
  0 & 0 & 0 & r_{44}
\end{bmatrix}
\begin{bmatrix}
  w_{11} & w_{12} & w_{13} & w_{14} \\
  w_{21} & w_{22} & w_{23} & w_{24} \\
  w_{31} & w_{32} & w_{33} & w_{34} \\
  w_{41} & w_{42} & w_{43} & w_{44}
\end{bmatrix}
\begin{bmatrix}
  s^*_{11} & 0 & 0 & 0 \\
  s^*_{12} & s^*_{22} & 0 & 0 \\
  s^*_{13} & s^*_{23} & s^*_{33} & 0 \\
  s^*_{14} & s^*_{24} & s^*_{34} & s^*_{44}
\end{bmatrix}
= \begin{bmatrix}
  e_{11} & e_{12} & e_{13} & e_{14} \\
  e_{21} & e_{22} & e_{23} & e_{24} \\
  e_{31} & e_{32} & e_{33} & e_{34} \\
  e_{41} & e_{42} & e_{43} & e_{44}
\end{bmatrix}
\]

If we equate the (1,4) and (4,1) entries, then we get

\[
\begin{align*}
  s^*_{11} w_{14} + w^*_{41} & = e^*_{41} - s_{12} w_{24} - s_{13} w_{34} - s_{14} w_{44} \\
  r_{11} w_{14} + w^*_{41} & = e_{14} - r_{12} w_{24} - r_{13} w_{34} - r_{14} w_{44}
\end{align*}
\]

a \( 2 \times 2 \) system of scalar equations that allows us to determine \( w_{14} \) and \( w_{41} \) simultaneously.
Algorithm to solve the transformed equation $RW + W^* S^* = E$ (I)

We illustrate with a $4 \times 4$ example for simplicity:

$$
\begin{bmatrix}
  r_{11} & r_{12} & r_{13} & r_{14} \\
  0 & r_{22} & r_{23} & r_{24} \\
  0 & 0 & r_{33} & r_{34} \\
  0 & 0 & 0 & r_{44}
\end{bmatrix}
\begin{bmatrix}
  w_{11} & w_{12} & w_{13} & w_{14} \\
  w_{21} & w_{22} & w_{23} & w_{24} \\
  w_{31} & w_{32} & w_{33} & w_{34} \\
  w_{41} & w_{42} & w_{43} & w_{44}
\end{bmatrix}
+ 
\begin{bmatrix}
  w_{11}^* & w_{21}^* & w_{31}^* & w_{41}^* \\
  w_{12}^* & w_{22}^* & w_{32}^* & w_{42}^* \\
  w_{13}^* & w_{23}^* & w_{33}^* & w_{43}^* \\
  w_{14}^* & w_{24}^* & w_{34}^* & w_{44}^*
\end{bmatrix}
\begin{bmatrix}
  s_{11}^* & 0 & 0 & 0 \\
  s_{12}^* & s_{22}^* & 0 & 0 \\
  s_{13}^* & s_{23}^* & s_{33}^* & 0 \\
  s_{14}^* & s_{24}^* & s_{34}^* & s_{44}^*
\end{bmatrix}
= 
\begin{bmatrix}
  e_{11} & e_{12} & e_{13} & e_{14} \\
  e_{21} & e_{22} & e_{23} & e_{24} \\
  e_{31} & e_{32} & e_{33} & e_{34} \\
  e_{41} & e_{42} & e_{43} & e_{44}
\end{bmatrix}
$$

If we equate the (1,4) and (4,1) entries, then we get

\[
\begin{align*}
  s_{11} & \quad w_{14}^* + \quad w_{41}^* = e_{41}^* - s_{12} \\
  r_{11} & \quad w_{14}^* + w_{41}^* = e_{14} - r_{12}
\end{align*}
\]

a $2 \times 2$ system of scalar equations that allows us to determine $w_{14}$ and $w_{41}$ simultaneously.
Algorithm to solve the transformed equation \( RW + W^* S^* = E \)

We illustrate with a 4 \( \times \) 4 example for simplicity:

\[
\begin{bmatrix}
  r_{11} & r_{12} & r_{13} & r_{14} \\
  0 & r_{22} & r_{23} & r_{24} \\
  0 & 0 & r_{33} & r_{34} \\
  0 & 0 & 0 & r_{44}
\end{bmatrix}
+ \begin{bmatrix}
  w_{11}^* & w_{21}^* & w_{31}^* & w_{41}^* \\
  w_{12}^* & w_{22}^* & w_{32}^* & w_{42}^* \\
  w_{13}^* & w_{23}^* & w_{33}^* & w_{43}^* \\
  w_{14}^* & w_{24}^* & w_{34}^* & w_{44}^*
\end{bmatrix}
= \begin{bmatrix}
  s_{11}^* & 0 & 0 & 0 \\
  s_{12}^* & s_{22}^* & 0 & 0 \\
  s_{13}^* & s_{23}^* & s_{33}^* & 0 \\
  s_{14}^* & s_{24}^* & s_{34}^* & s_{44}^*
\end{bmatrix}
\]

If we equate the (1:3,1:3) submatrices, then we get

\[
\begin{bmatrix}
  r_{11} & r_{12} & r_{13} \\
  0 & r_{22} & r_{23} \\
  0 & 0 & r_{33}
\end{bmatrix}
= \begin{bmatrix}
  e_{11} & e_{12} & e_{13} \\
  e_{21} & e_{22} & e_{23} \\
  e_{31} & e_{32} & e_{33}
\end{bmatrix}
- \begin{bmatrix}
  r_{14} \\
  r_{24} \\
  r_{34}
\end{bmatrix}
\]

which is a 3 \( \times \) 3 matrix equation of the same type as the original one.
Algorithm to solve the transformed equation $RW + W^* S^* = E$ (I)

We illustrate with $4 \times 4$ example for simplicity:

$$
\begin{bmatrix}
    r_{11} & r_{12} & r_{13} & r_{14} \\
    0 & r_{22} & r_{23} & r_{24} \\
    0 & 0 & r_{33} & r_{34} \\
    0 & 0 & 0 & r_{44}
\end{bmatrix}
\begin{bmatrix}
    w_{11} & w_{12} & w_{13} & w_{14} \\
    w_{21} & w_{22} & w_{23} & w_{24} \\
    w_{31} & w_{32} & w_{33} & w_{34} \\
    w_{41} & w_{42} & w_{43} & w_{44}
\end{bmatrix}
\begin{bmatrix}
    s_{11} & 0 & 0 & 0 \\
    s_{12} & s_{22} & 0 & 0 \\
    s_{13} & s_{23} & s_{33} & 0 \\
    s_{14} & s_{24} & s_{34} & s_{44}
\end{bmatrix}
= 
\begin{bmatrix}
    e_{11} & e_{12} & e_{13} & e_{14} \\
    e_{21} & e_{22} & e_{23} & e_{24} \\
    e_{31} & e_{32} & e_{33} & e_{34} \\
    e_{41} & e_{42} & e_{43} & e_{44}
\end{bmatrix}
$$

If we equate the $(1:3,1:3)$ submatrices , then we get

$$
\begin{bmatrix}
    r_{11} & r_{12} & r_{13} \\
    0 & r_{22} & r_{23} \\
    0 & 0 & r_{33}
\end{bmatrix}
\begin{bmatrix}
    w_{11} & w_{12} & w_{13} \\
    w_{21} & w_{22} & w_{23} \\
    w_{31} & w_{32} & w_{33}
\end{bmatrix}
+ 
\begin{bmatrix}
    s_{11} & 0 & 0 & 0 \\
    s_{12} & s_{22} & 0 & 0 \\
    s_{13} & s_{23} & s_{33} & 0
\end{bmatrix}
= 
\begin{bmatrix}
    e_{11} & e_{12} & e_{13} \\
    e_{21} & e_{22} & e_{23} \\
    e_{31} & e_{32} & e_{33} \\
    e_{41} & e_{42} & e_{43} & e_{44}
\end{bmatrix}
$$

which is a $3 \times 3$ matrix equation of the same type as the original one.
Algorithm to solve the transformed equation \( RW + W^* S^* = E \)

**INPUT:** \( R, S, E \in \mathbb{C}^{n \times n} \), with \( R \) and \( S \) upper triangular

**OUTPUT:** \( W \in \mathbb{C}^{n \times n} \)

for \( j = p : -1 : 1 \)

solve \( r_{jj}w_{jj} + w_{jj}^* s_{jj}^* = e_{jj} \) to get \( w_{jj} \)

for \( i = j - 1 : -1 : 1 \)

solve \( \begin{cases} s_{ii}w_{ij} + w_{ji}^* r_{jj}^* &= e_{ji}^* - \sum_{k=i+1}^{j} s_{ik}w_{kj} \\ r_{ii}w_{ij} + w_{ji}^* s_{jj}^* &= e_{ij} - \sum_{k=i+1}^{j} r_{ik}w_{kj} \end{cases} \) to get \( w_{ij}, w_{ji} \)

end

\[
-(S(1 : j - 1, j)W(j, 1 : j - 1))^*
\]

end

**Cost**

- \( 2n^3 + O(n^2) \) flops for real input matrices and a total cost \( 76n^3 + O(n^2) \) flops for the whole algorithm for \( AX + X^* B = C \).
Algorithm to solve the transformed equation $RW + W^* S^* = E$ \hspace{1cm} (II)

**INPUT:** $R, S, E \in \mathbb{C}^{n \times n}$, with $R$ and $S$ upper triangular

**OUTPUT:** $W \in \mathbb{C}^{n \times n}$

for $j = p : -1 : 1$

solve $r_{jj}w_{jj} + w_{jj}^* s_{jj}^* = e_{jj}$ to get $w_{jj}$

for $i = j - 1 : -1 : 1$

solve \[
\begin{align*}
    s_{ii}w_{ij} + w_{ji}^* r_{jj}^* &= e_{ji} - \sum_{k=i+1}^{j} s_{ik}w_{kj} \\
    r_{ii}w_{ij} + w_{ji}^* s_{jj}^* &= e_{ij} - \sum_{k=i+1}^{j} r_{ik}w_{kj}
\end{align*}
\]

end


$-(S(1 : j - 1, j)W(j, 1 : j - 1))^*$

end

**Cost**

- $2n^3 + O(n^2)$ flops for real input matrices and a total cost $76n^3 + O(n^2)$ flops for the whole algorithm for $AX + X^* B = C$. 

F. M. Dopico (U. Carlos III, Madrid)  Sylvester equation for congruence  CEDYA 2011 19 / 24
Remarks on algorithm to solve $AX + X^* B = C$

- **Roundoff errors:** $\hat{X}$, computed solution of $AX + X^* B = C$, satisfies
  \[
  \|A\hat{X} + \hat{X}^* B - C\|_F \leq \alpha \, u \, n^{5/2} \left(\|A\|_F + \|B\|_F\right) \|\hat{X}\|_F,
  \]
  with $u$ unit roundoff and $\alpha$ small integer constant.

- The algorithm for solving $RW + W^* S^* = E$ is dominated by level-2 BLAS operations. In modern computers, a blocked-version dominated by level-3 BLAS operations would be more efficient (future work), but...

- for $AX + X^* B = C$, cost is dominated by the QZ-alg on $A - \lambda B^*$.

- The algorithm should be compared with Bartels-Stewart algorithm for Sylvester equation $AX - XB = C$:
  1. Compute independently triang. Schur forms $T_A$ and $T_B$ of $A$ and $B$.
  2. Solve $T_A Y - Y T_B = D$ for $Y$.
  3. Recover $X$ from $Y$.

- Same flavor, but also differences: $A - \lambda B^*$ and $T_A Y - Y T_B = D$ allows us to compute each entry of $Y$ in terms on previous ones (from left and from bottom) **without using $2 \times 2$ linear-systems**.
Remarks on algorithm to solve \( AX + X^* B = C \)

- **Roundoff errors**: \( \hat{X} \), computed solution of \( AX + X^* B = C \), satisfies

\[
\| A\hat{X} + \hat{X}^* B - C \|_F \leq \alpha u n^{5/2} (\| A \|_F + \| B \|_F) \| \hat{X} \|_F,
\]

with \( u \) unit roundoff and \( \alpha \) small integer constant.

- The algorithm for solving \( RW + W^* S^* = E \) is dominated by level-2 BLAS operations. In modern computers, a blocked-version dominated by level-3 BLAS operations would be more efficient (future work), but...

- for \( AX + X^* B = C \), cost is dominated by the QZ-alg on \( A - \lambda B^* \).

- The algorithm should be compared with Bartels-Stewart algorithm for Sylvester equation \( AX - XB = C \):

1. Compute independently triang. Schur forms \( T_A \) and \( T_B \) of \( A \) and \( B \).
2. Solve \( T_A Y - Y T_B = D \) for \( Y \).
3. Recover \( X \) from \( Y \).

- Same flavor, but also differences: \( A - \lambda B^* \) and \( T_A Y - Y T_B = D \) allows us to compute each entry of \( Y \) in terms on previous ones (from left and from bottom) without using \( 2 \times 2 \) linear-systems.
Remarks on algorithm to solve $A X + X^* B = C$

- **Roundoff errors:** $\hat{X}$, computed solution of $AX + X^*B = C$, satisfies

$$\|A\hat{X} + \hat{X}^* B - C\|_F \leq \alpha u n^{5/2} (\|A\|_F + \|B\|_F) \|\hat{X}\|_F,$$

with $u$ unit roundoff and $\alpha$ small integer constant.

- The algorithm for solving $RW + W^* S^* = E$ is dominated by level-2 BLAS operations. In modern computers, a blocked-version dominated by level-3 BLAS operations would be more efficient (future work), but...

- for $AX + X^*B = C$, cost is dominated by the QZ-alg on $A - \lambda B^*$.

- The algorithm should be compared with Bartels-Stewart algorithm for Sylvester equation $AX - XB = C$:

1. Compute independently triang. Schur forms $T_A$ and $T_B$ of $A$ and $B$.
2. Solve $T_A Y - Y T_B = D$ for $Y$.
3. Recover $X$ from $Y$.

- Same flavor, but also differences: $A - \lambda B^*$ and $T_A Y - Y T_B = D$ allows us to compute each entry of $Y$ in terms on previous ones (from left and from bottom) **without using $2 \times 2$ linear-systems.**

F. M. Dopico (U. Carlos III, Madrid) Sylvester equation for congruence CEDYA 2011 20 / 24
Remarks on algorithm to solve $AX + X^* B = C$

- **Roundoff errors:** $\hat{X}$, computed solution of $AX + X^* B = C$, satisfies
  \[ \|A\hat{X} + \hat{X}^* B - C\|_F \leq \alpha u n^{5/2} (\|A\|_F + \|B\|_F) \|\hat{X}\|_F, \]
  with $u$ unit roundoff and $\alpha$ small integer constant.

- The algorithm for solving $RW + W^* S^* = E$ is dominated by level-2 BLAS operations. In modern computers, a blocked-version dominated by level-3 BLAS operations would be more efficient (future work), but...

- for $AX + X^* B = C$, cost is dominated by the QZ-alg on $A - \lambda B^*$.

- The algorithm should be compared with Bartels-Stewart algorithm for Sylvester equation $AX - XB = C$:
  1. Compute **independently** triang. Schur forms $T_A$ and $T_B$ of $A$ and $B$.
  2. Solve $T_A Y - Y T_B = D$ for $Y$.
  3. Recover $X$ from $Y$.

- Same flavor, but also differences: $A - \lambda B^*$ and $T_A Y - Y T_B = D$ allows us to compute each entry of $Y$ in terms on previous ones (from left and from bottom) without using $2 \times 2$ linear-systems.
Remarks on algorithm to solve $AX + X^*B = C$

- **Roundoff errors:** $\hat{X}$, computed solution of $AX + X^*B = C$, satisfies
  \[
  \|A\hat{X} + \hat{X}^*B - C\|_F \leq \alpha \, u \, n^{5/2} \left( \|A\|_F + \|B\|_F \right) \|\hat{X}\|_F,
  \]
  with $u$ unit roundoff and $\alpha$ small integer constant.

- The algorithm for solving $RW + W^*S^* = E$ is dominated by level-2 BLAS operations. In modern computers, a blocked-version dominated by level-3 BLAS operations would be more efficient (future work), but...

- for $AX + X^*B = C$, cost is dominated by the QZ-alg on $A - \lambda B^*$.

- The algorithm should be compared with Bartels-Stewart algorithm for Sylvester equation $AX - XB = C$:
  1. Compute independently triang. Schur forms $T_A$ and $T_B$ of $A$ and $B$.
  2. Solve $T_A Y - Y T_B = D$ for $Y$.
  3. Recover $X$ from $Y$.

- Same flavor, but also differences: $A - \lambda B^*$ and $T_A Y - Y T_B = D$ allows us to compute each entry of $Y$ in terms on previous ones (from left and from bottom) **without using $2 \times 2$ linear-systems**.
1 Motivation

2 Consistency of the Sylvester equation for $\star$-congruence

3 Uniqueness of solutions

4 Efficient and stable algorithm to compute unique solutions

5 General "nonunique" solution of $AX + X^*B = C$

6 Conclusions
Theoretical method to solve $AX + X^* B = 0$

- In case of consistency, but “nonuniqueness”, general solution of $AX + X^* B = C$ is $X = X_p + X_h$, where
  1. $X_p$ is a particular solution and
  2. $X_h$ is the general solution of $AX + X^* B = 0$.

The latter found a few weeks ago by De Terán, D., Guillery, Montealegre, Reyes (REU program, U. of California at S. Barbara, M.I. Bueno).

**KEY IDEA:** If $E - \lambda F^*$ is the Kronecker Canonical form (KCF) of $A - \lambda B^*$, then $AX + X^* B = 0$ can be transformed into

$$EY + Y^* F = 0.$$ 

- If $E = E_1 \oplus \cdots \oplus E_d$, $F^* = F_1^* \oplus \cdots \oplus F_d^*$, and $Y = [Y_{ij}]$ is partitioned into blocks accordingly, then this equation decouples in

$$E_i Y_{ii} + Y_{ii}^* F_i = 0$$

and

$$\begin{cases} E_i Y_{ij} + Y_{jj}^* F_i = 0 \\ E_j Y_{ji} + Y_{ji}^* F_i = 0 \end{cases}, \quad (1 \leq i < j \leq d),$$

which produce 14 different types of matrix (systems) equations, whose explicit solutions have been found. Much more complicated general solution than standard Sylvester eq: $AX - XB = 0$.  

F. M. Dopico (U. Carlos III, Madrid)  
Sylvester equation for congruence  
CEDYA 2011 22 / 24
Theoretical method to solve $AX + X^* B = 0$

- In case of consistency, but “nonuniqueness”, general solution of $AX + X^* B = C$ is $X = X_p + X_h$, where
  1. $X_p$ is a particular solution and
  2. $X_h$ is the general solution of $AX + X^* B = 0$.

The latter found a few weeks ago by De Terán, D., Guillery, Montealegre, Reyes (REU program, U. of California at S. Barbara, M.I. Bueno).

**KEY IDEA:** If $E - \lambda F^*$ is the Kronecker Canonical form (KCF) of $A - \lambda B^*$, then $AX + X^* B = 0$ can be transformed into $EY + Y^* F = 0$.

- If $E = E_1 \oplus \cdots \oplus E_d$, $F^* = F_1^* \oplus \cdots \oplus F_d^*$, and $Y = [Y_{ij}]$ is partitioned into blocks accordingly, then this equation decouples in

  $E_i Y_{ii} + Y_{ii}^* F_i = 0$  \quad and \quad \begin{cases} E_i Y_{ij} + Y_{ij}^* F_j = 0 \\ E_j Y_{ji} + Y_{ji}^* F_i = 0 \end{cases}, \quad (1 \leq i < j \leq d),$

- which produce 14 different types of matrix (systems) equations, whose explicit solutions have been found. Much more complicated general solution than standard Sylvester eq: $AX - XB = 0$.
Theoretical method to solve $AX + X^* B = 0$

- In case of consistency, but “nonuniqueness”, general solution of $AX + X^* B = C$ is $X = X_p + X_h$, where
  1. $X_p$ is a particular solution and
  2. $X_h$ is the general solution of $AX + X^* B = 0$.

The latter found a few weeks ago by De Terán, D., Guillery, Montealegre, Reyes (REU program, U. of California at S. Barbara, M.I. Bueno).

**KEY IDEA:** If $E - \lambda F^*$ is the Kronecker Canonical form (KCF) of $A - \lambda B^*$, then $AX + X^* B = 0$ can be transformed into $EY + Y^* F = 0$.

- If $E = E_1 \oplus \cdots \oplus E_d$, $F^* = F_1^* \oplus \cdots \oplus F_d^*$, and $Y = [Y_{ij}]$ is partitioned into blocks accordingly, then this equation decouples in
  
  $$E_i Y_{ii} + Y_{ii}^* F_i = 0$$
  and
  $$\begin{cases} 
  E_i Y_{ij} + Y_{ij}^* F_j = 0 \\
  E_j Y_{ji} + Y_{ji}^* F_i = 0 
  \end{cases}, \quad (1 \leq i < j \leq d),$$

  which produce 14 different types of matrix (systems) equations, whose explicit solutions have been found. Much more complicated general solution than standard Sylvester eq: $AX - XB = 0$. 

F. M. Dopico (U. Carlos III, Madrid)  
Sylvester equation for congruence  
CEDYA 2011  
22 / 24
Theoretical method to solve $AX + X^* B = 0$

- In case of consistency, but “nonuniqueness”, general solution of $AX + X^* B = C$ is $X = X_p + X_h$, where
  1. $X_p$ is a particular solution and
  2. $X_h$ is the general solution of $AX + X^* B = 0$.

The latter found a few weeks ago by De Terán, D., Guillery, Montealegre, Reyes (REU program, U. of California at S. Barbara, M.I. Bueno).

**KEY IDEA:** If $E - \lambda F^*$ is the Kronecker Canonical form (KCF) of $A - \lambda B^*$, then $AX + X^* B = 0$ can be transformed into

$$EY + Y^* F = 0.$$ 

If $E = E_1 \oplus \cdots \oplus E_d$, $F^* = F_1^* \oplus \cdots \oplus F_d^*$, and $Y = [Y_{ij}]$ is partitioned into blocks accordingly, then this equation decouples in

$$E_i Y_{ii} + Y_{ii}^* F_i = 0$$

and

$$\begin{cases} 
E_i Y_{ij} + Y_{ji}^* F_j = 0 \\
E_j Y_{ji} + Y_{ij}^* F_i = 0
\end{cases}, \quad (1 \leq i < j \leq d),$$

which produce 14 different types of matrix (systems) equations, whose explicit solutions have been found. Much more complicated general solution than standard Sylvester eq: $AX - XB = 0$. 

F. M. Dopico (U. Carlos III, Madrid)
Outline

1 Motivation
2 Consistency of the Sylvester equation for \( \star \)-congruence
3 Uniqueness of solutions
4 Efficient and stable algorithm to compute unique solutions
5 General “nonunique” solution of \( AX + X^\star B = C \)
6 Conclusions
Conclusions

Several questions related to the Sylvester equation for \( \star \)-congruence
\[ AX + X^* B = C \] are well-understood:

1. Necessary and sufficient conds for consistency/uniqueness of sols.
2. Efficient and stable nume. algorithm for computing unique solution.
3. General solution of \[ AX + X^* B = 0. \]

Connections with std. Syl. eq \[ AX - XB = C \] but also relevant diffs:

1. Use of QZ-algor for pencil \( A - \lambda B^\star \) instead of QR-algor for matrices.
2. Use of KCF for pencil \( A - \lambda B^\star \) instead of JCF for matrices in general homogeneous solution: much more complicated solution.
3. The Canonical Form for Congruence only useful in the particular case \( AX + X^* A = 0 \).

Several problems still remain open. Among them:

1. Combine the algor for \( AX + X^* B = C \) with algors for computing the anti-triangular Schur form for completely solving the linear palindromic eigenproblem via congruence.
2. Numerical method for computing basis of the solution space of \( AX + X^* B = 0 \) via staircase-form of \( A - \lambda B^\star \).
3. Eigenvalues of the operator \( X \mapsto AX + X^* B \).
Conclusions

Several questions related to the Sylvester equation for \(-\)-congruence $AX + X^* B = C$ are well-understood:

1. Necessary and sufficient conds for consistency/uniqueness of sols.
2. Efficient and stable nume. algorithm for computing unique solution.
3. General solution of $AX + X^* B = 0$.

Connections with strd. Syl. eq $AX - XB = C$ but also relevant diffs:

1. Use of QZ-algor for pencil $A - \lambda B^*$ instead of QR-algor for matrices.
2. Use of KCF for pencil $A - \lambda B^*$ instead of JCF for matrices in general homogeneous solution: much more complicated solution.
3. The Canonical Form for Congruence only useful in the particular case $AX + X^* A = 0$.

Several problems still remain open. Among them:

1. Combine the algor for $AX + X^* B = C$ with algors for computing the anti-triangular Schur form for completely solving the linear palindromic eigenproblem via congruence.
2. Numerical method for computing basis of the solution space of $AX + X^* B = 0$ via staircase-form of $A - \lambda B^*$.
3. Eigenvalues of the operator $X \mapsto AX + X^* B$. 

F. M. Dopico (U. Carlos III, Madrid)
Conclusions

- Several questions related to the Sylvester equation for $\ast$-congruence $AX + X^*B = C$ are well-understood:
  1. Necessary and sufficient conditions for consistency/uniqueness of solutions.
  2. Efficient and stable numerical algorithm for computing unique solution.
  3. General solution of $AX + X^*B = 0$.

- Connections with standard Sylvester eq $AX - XB = C$ but also relevant differences:
  1. Use of QZ-algorithm for pencil $A - \lambda B^*$ instead of QR-algorithm for matrices.
  2. Use of KCF for pencil $A - \lambda B^*$ instead of JCF for matrices in general homogeneous solution: much more complicated solution.
  3. The Canonical Form for Congruence only useful in the particular case $AX + X^*A = 0$.

- Several problems still remain open. Among them:
  1. Combine the algorithm for $AX + X^*B = C$ with algorithms for computing the anti-triangular Schur form for completely solving the linear palindromic eigenproblem via congruence.
  2. Numerical method for computing basis of the solution space of $AX + X^*B = 0$ via staircase-form of $A - \lambda B^*$.
  3. Eigenvalues of the operator $X \mapsto AX + X^*B$. 

F. M. Dopico (U. Carlos III, Madrid)