The Sylvester equation for congruence and some related equations

Fernando De Terán

Departamento de Matemáticas
Universidad Carlos III de Madrid
(Spain)

Joint work with: **F. M. Dopico** (UC3M-ICMAT)
N. Guillery
D. Montealegre
N. Z. Reyes

Leganés, June 29th, 2012

Outline

1. Definition. Goals. Related equations and some history.
2. Motivation
3. Necessary and sufficient conditions
4. The solution of $AX + X^{\star}B = 0$

Outline

Sylvester equation for congruence

$A \in \mathbb{F}^{m\times n}, B \in \mathbb{F}^{n\times m}, C \in \mathbb{F}^{m\times m}$ ($\mathbb{F}$ an arbitrary field)

$$AX + X^{\star}B = C$$

Sylvester equation for $\star$-congruence

$X \in \mathbb{F}^{n\times m}$, unknown

($\star = T$ or $*$)

(Other name in the literature: “Sylvester-transpose matrix equation")

Solution of Sylvester equation for congruence

$AX + X^\star B = C \qquad (\star = T \text{ or } *)$ Sylvester equation for congruence

GOALS:

- Find necessary and sufficient conditions for consistency.
- Find the dimension of the solution space.
- Find an expression for the solution.
- Find necessary and sufficient conditions for uniqueness of the solution.
- Find an (efficient) algorithm to compute the solution (when unique).

Related equations and history

$AX + X^\star B = C \qquad (\star = T \text{ or } *) \qquad$ Sylvester equation for congruence

(a) Sylvester equation: $AX + XB = C$ (A, B must be **square**!!)

- Solution know since (at least) the 1950's (Gantmacher).
- Characterization of consistency and uniqueness of solution already known (Roth, Gantmacher).
- Efficient algorithm for the unique solution already known (Bartels-Stewart).
- Mathscinet:
  - 83 references containing “Sylvester equation" in the title.
  - 44 references containing “Sylvester matrix equation" in the title.
  - 227 references containing “Sylvester equation" anywhere.
  - 91 references containing “Sylvester matrix equation" anywhere.

Related equations and history (II)

(b) $AX \pm X^\star A^\star = C, \qquad A \in \mathbb{F}^{m\times n}, C \in \mathbb{F}^{m\times m}$:

- Hodges (1957): Solution over finite fields.
- Taussky-Wielandt (1962): Eigenvalues of $g(X) = A^T X + X^T A$.
- Lancaster-Rozsa (1983), Braden (1999): Necessary and sufficient conditions for consistency. Closed-form formula for the solution (using projectors and generalized inverses) and dimension of the solution space.
- Djordjević (2007): Extends Lancaster-Rozsa to A, C, X bounded linear operators on Hilbert spaces (with closed rank).

(c) $AX + X^\star A = C, \qquad A, C \in \mathbb{C}^{n\times n}$:

- Ballantine (1969): $H = PA + AP^*$, with H hermitian and A, P with certain structure.
- DT-Dopico (2011): Complete solution for $C = 0$. Related to the theory of (congruence) orbits.

Related equations and history (III)

(d) The Sylvester equation for congruence: $AX + X^\star B = C$:

- Necessary and sufficient conditions for **consistency**: **Wimmer (1994)**, Piao-Zhang-Wang (2007, involved), DT-Dopico (2011, another proof of Wimmer's).
- Necessary and sufficient conditions for **unique solution**: Byers -Kressner (2006, $\star = T$), Kressner-Schröder-Watkins (2009, $\star = *$).
- **Formula** for the solution: Piao-Zhang-Wang (2007, involved), Cvetković-Ilić (2008, operators with certain restrictions), DT-Dopico-Guillery-Montealegre-Reyes (submitted, $C = 0$).
- **Algorithm** for the (unique) solution: DT-Dopico (2011, $O(n^3)$), Vorontsov-Ivanov (2011), Chiang-Chu-Lin (2012).

(e) $AXB + CX^\star D = E$:

- Numerical iterative methods to find the solution (when unique) or some structured solutions: Wang-Chen-Wei (2007), Hajarian-Mehghan (2010), Xie-Liu-Yang (2010), Song-Chen (2011).

Outline

Orbit theory

$XA + AX^{\star} = 0, \qquad A \in \mathbb{C}^{n\times n}$

Set:

$$\mathcal{O}(A) = \left\{ PAP^T \;:\; P \text{ nonsingular} \right\} \quad \textbf{Congruence orbit of } A$$
$$\mathcal{O}_s(A) = \left\{ PAP^{-1} \;:\; P \text{ nonsingular} \right\} \quad \text{Similarity orbit of } A$$

Then:

$$T_{\mathcal{O}(A)}(A) = \left\{ XA + AX^T \;:\; X \in \mathbb{C}^{n\times n} \right\} \quad \text{Tangent space of } \mathcal{O}(A) \text{ at } A$$
$$T_{\mathcal{O}_s(A)}(A) = \left\{ XA - AX \;:\; X \in \mathbb{C}^{n\times n} \right\} \quad \text{Tangent space of } \mathcal{O}_s(A) \text{ at } A$$

(a) $\operatorname{codim} \mathcal{O}(A) = \operatorname{codim} T_{\mathcal{O}(A)}(A) = \dim(\text{solution space of } XA + AX^T = 0)$

(b) $\operatorname{codim} \mathcal{O}_s(A) = \operatorname{codim} T_{\mathcal{O}_s(A)}(A) = \dim(\text{solution space of } XA - AX = 0)$

Orbit theory

$$XA + AX^{\star} = 0, \qquad A \in \mathbb{C}^{n\times n}$$

Set:

$$\mathcal{O}(A) = \left\{PAP^{T} \;:\; P \text{ nonsingular}\right\}$$ **Congruence** orbit of A

$$\mathcal{O}_s(A) = \left\{PAP^{-1} \;:\; P \text{ nonsingular}\right\}$$ Similarity orbit of A

Then:

$$T_{\mathcal{O}(A)}(A) = \left\{XA + AX^{T} \;:\; X \in \mathbb{C}^{n\times n}\right\}$$ Tangent space of $\mathcal{O}(A)$ at A

$$T_{\mathcal{O}_s(A)}(A) = \left\{XA - AX \;:\; X \in \mathbb{C}^{n\times n}\right\}$$ Tangent space of $\mathcal{O}_s(A)$ at A

(a) codim $\mathcal{O}(A)$ = codim $T_{\mathcal{O}(A)}(A)$ = dim(solution space of $XA + AX^{T} = 0$)

(b) codim $\mathcal{O}_s(A)$ = codim $T_{\mathcal{O}_s(A)}(A)$ = dim(solution space of $XA - AX = 0$)

Orbit theory

$XA + AX^\star = 0, \qquad A \in \mathbb{C}^{n\times n}$

Set:

$$\mathscr{O}(A) = \left\{PAP^T \ : \ P \text{ nonsingular}\right\} \quad \textbf{Congruence} \text{ orbit of } A$$
$$\mathscr{O}_s(A) = \left\{PAP^{-1} \ : \ P \text{ nonsingular}\right\} \quad \text{Similarity orbit of } A$$

Then:

$$T_{\mathscr{O}(A)}(A) = \left\{XA + AX^T \ : \ X \in \mathbb{C}^{n\times n}\right\} \quad \text{Tangent space of } \mathscr{O}(A) \text{ at } A$$
$$T_{\mathscr{O}_s(A)}(A) = \{XA - AX \ : \ X \in \mathbb{C}^{n\times n}\} \quad \text{Tangent space of } \mathscr{O}_s(A) \text{ at } A$$

(a) codim $\mathscr{O}(A) =$ codim $T_{\mathscr{O}(A)}(A) =$ dim(solution space of $XA + AX^T = 0$)

(b) codim $\mathscr{O}_s(A) =$ codim $T_{\mathscr{O}_s(A)}(A) =$ dim(solution space of $XA - AX = 0$)

Reduction by congruence to anti-triangular form

$$\overbrace{\begin{bmatrix} X^\star & I \\ I & 0 \end{bmatrix}}^{P} \begin{bmatrix} 0 & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \overbrace{\begin{bmatrix} X & I \\ I & 0 \end{bmatrix}}^{P^\star} = \begin{bmatrix} 0 & A_{12} \\ A_{21} & 0 \end{bmatrix}$$

$$\Leftrightarrow A_{21}X + X^\star A_{12} = -A_{22}.$$

Application: Anti-triangular form of palindromic pencils $A + \lambda A^\star$.

(Analogous to:

$$\overbrace{\begin{bmatrix} I & X \\ 0 & I \end{bmatrix}}^{P} \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \overbrace{\begin{bmatrix} I & -X \\ 0 & I \end{bmatrix}}^{P^{-1}} = \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix}$$

$\Leftrightarrow A_{11}X - XA_{22} = A_{12} \rightsquigarrow$ **Sylvester equation**)

Reduction by congruence to anti-triangular form

$$\overbrace{\begin{bmatrix} X^\star & I \\ I & 0 \end{bmatrix}}^{P} \begin{bmatrix} 0 & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \overbrace{\begin{bmatrix} X & I \\ I & 0 \end{bmatrix}}^{P^\star} = \begin{bmatrix} 0 & A_{12} \\ A_{21} & 0 \end{bmatrix}$$

$$\Leftrightarrow A_{21}X + X^\star A_{12} = -A_{22}.$$

Application: Anti-triangular form of palindromic pencils $A + \lambda A^\star$.

(Analogous to:

$$\overbrace{\begin{bmatrix} I & X \\ 0 & I \end{bmatrix}}^{P} \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \overbrace{\begin{bmatrix} I & -X \\ 0 & I \end{bmatrix}}^{P^{-1}} = \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix}$$

$\Leftrightarrow A_{11}X - XA_{22} = A_{12} \rightsquigarrow$ **Sylvester equation**)

Reduction by congruence to anti-triangular form

$$\overbrace{\begin{bmatrix} X^\star & I \\ I & 0 \end{bmatrix}}^{P} \begin{bmatrix} 0 & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \overbrace{\begin{bmatrix} X & I \\ I & 0 \end{bmatrix}}^{P^\star} = \begin{bmatrix} 0 & A_{12} \\ A_{21} & 0 \end{bmatrix}$$

$$\Leftrightarrow A_{21}X + X^\star A_{12} = -A_{22}.$$

Application: Anti-triangular form of palindromic pencils $A + \lambda A^\star$.

(Analogous to:

$$\overbrace{\begin{bmatrix} I & X \\ 0 & I \end{bmatrix}}^{P} \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \overbrace{\begin{bmatrix} I & -X \\ 0 & I \end{bmatrix}}^{P^{-1}} = \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix}$$

$\Leftrightarrow A_{11}X - XA_{22} = A_{12} \rightsquigarrow$ **Sylvester equation**)

Outline

Consistency

Theorem (Wimmer 1994, DT-Dopico 2011)

Let $\mathbb{F}$ be a field with $\operatorname{char}\mathbb{F} \neq 2$, $A \in \mathbb{F}^{m\times n}, B \in \mathbb{F}^{n\times m}, C \in \mathbb{F}^{m\times m}$. Then

$$AX + X^\star B = C \qquad \text{is consistent}$$

if and only if

$$P^\star \begin{bmatrix} C & A \\ B & 0 \end{bmatrix} P = \begin{bmatrix} 0 & A \\ B & 0 \end{bmatrix},$$

for some nonsingular P.

(Compare with Roth's criterion:
"$AX - XB = C$ is consistent if and only if

$$P^{-1} \begin{bmatrix} A & C \\ 0 & B \end{bmatrix} P = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix},$$

for some nonsingular P".)

Consistency

Theorem (Wimmer 1994, DT-Dopico 2011)

Let $\mathbb{F}$ be a field with $\operatorname{char}\mathbb{F} \neq 2$, $A \in \mathbb{F}^{m\times n}, B \in \mathbb{F}^{n\times m}, C \in \mathbb{F}^{m\times m}$. Then

$$AX + X^\star B = C \qquad \text{is consistent}$$

if and only if

$$P^\star \begin{bmatrix} C & A \\ B & 0 \end{bmatrix} P = \begin{bmatrix} 0 & A \\ B & 0 \end{bmatrix},$$

for some nonsingular P.

(Compare with Roth's criterion:
"$AX - XB = C$ is consistent if and only if

$$P^{-1} \begin{bmatrix} A & C \\ 0 & B \end{bmatrix} P = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix},$$

for some nonsingular P".)

Consistency: proof

Wimmer's proof: Dimensionality arguments.
DT-Dopico's proof: Based on:

Theorem (Wimmer 1994, Syrmos-Lewis 1994, Beitia-Gracia 1996)

$A_1, A_2 \in \mathbb{F}^{m\times n}$, $B_1, B_2 \in \mathbb{F}^{p\times k}$, $C_1, C_2 \in \mathbb{F}^{m\times k}$. Then

$$\begin{array}{c} A_1X + YB_1 = C_1 \\ A_2X + YB_2 = C_2 \end{array} \qquad \text{is consistent}$$

if and only if

$$P\begin{bmatrix} A_1 - \lambda A_2 & C_1 - \lambda C_2 \\ 0 & B_1 - \lambda B_2 \end{bmatrix} Q = \begin{bmatrix} A_1 - \lambda A_2 & 0 \\ 0 & B_1 - \lambda B_2 \end{bmatrix},$$

for some P, Q nonsingular.

Uniqueness of solution

Theorem (Byers-Kressner 2006, Kressner-Schröder-Watkins 2009)

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

$$AX + X^{\star}B = C \qquad \text{has a unique solution}$$

if and only if

(1) $A + \lambda B^{\star}$ is regular, and

(2) $\star = T$: If $\mu \in \mathrm{Spec}\,(A + \lambda B^{T}) \setminus \{-1\}$, then $1/\mu \notin \mathrm{Spec}\,(A + \lambda B^{T}) \setminus \{-1\}$ and, if $-1 \in \mathrm{Spec}\,(A + \lambda B^{T})$, then it has algebraic multiplicity one.
$\star = *$: If $\mu \in \mathrm{Spec}\,(A + \lambda B^{*})$, then $1/\overline{\mu} \notin \mathrm{Spec}\,(A + \lambda B^{*})$.

Uniqueness of solution: Algorithm

Using `vec` and Gaussian elimination: $M \cdot \mathrm{vec} X = \mathrm{vec} C \rightsquigarrow \mathbf{O(n^6)}$!!!!

Algorithm 1 (Solution of $AX + X^\star B = C$)

$A, B \in \mathbb{C}^{n\times n}$, $A + \lambda B^\star$ regular

Step 1. Compute the generalized Schur decomposition of $A + \lambda B^\star$ (with the QZ algorithm):

$$A = URV, \qquad B^\star = USV.$$

Step 2. Compute $E = U^* C (U^\star)^*$.
Step 3. Solve $RW + W^\star S^\star = E$.
Step 4. Compute $X = V^* W U^\star$.

Cost of `Algorithm 1` : $76n^3 + O(n^2)$

Outline

The pencil $A+\lambda B^\star$

Notation: $\mathscr{S}(A,B)=\{X : AX+X^\star B=0\}$

Lemma

If $P(A+\lambda B^\star)Q=\widetilde{A}+\lambda\widetilde{B}^\star$ then there is a one-to-one linear map:

$$\begin{array}{ccc} \mathscr{S}(A,B) & \to & \mathscr{S}(\widetilde{A},\widetilde{B}) \\ X & \mapsto & Y=Q^{-1}XP^\star. \end{array}$$

IDEA: Reduce $A+\lambda B^\star$ to its Kronecker Canonical Form (KCF), $K_1+\lambda K_2^\star$, and solve $K_1X+X^\star K_2=0$.

(Compare:

$AX-XB=0$: Depends on the Jordan canonical form of A,B

$AX+X^\star B=0$: Depends on the KCF of $A+\lambda B^\star$.)

The pencil $A+\lambda B^\star$

Notation: $\mathscr{S}(A,B)=\{X : AX+X^\star B=0\}$

Lemma

If $P(A+\lambda B^\star)Q=\widetilde{A}+\lambda\widetilde{B}^\star$ then there is a one-to-one linear map:

$$\begin{array}{ccc} \mathscr{S}(A,B) & \to & \mathscr{S}(\widetilde{A},\widetilde{B}) \\ X & \mapsto & Y=Q^{-1}XP^\star. \end{array}$$

IDEA: Reduce $A+\lambda B^\star$ to its Kronecker Canonical Form (KCF), $K_1+\lambda K_2^\star$, and solve $K_1X+X^\star K_2=0$.

(Compare:

$AX-XB=0$: Depends on the Jordan canonical form of A,B

$AX+X^\star B=0$: Depends on the KCF of $A+\lambda B^\star$.)

Partition into blocks

Lemma

Let $E = \text{diag}(E_1, \ldots, E_d)$ and $F^\star = \text{diag}(F_1^\star, \ldots, F_d^\star)$, and partition $X = [X_{ij}]_{i,j=1:d}$. Then

$$EX + X^\star F = 0$$

is equivalent to the set of equations

$$\begin{aligned} E_i X_{ij} + X_{ij}^\star F_j &= 0 \\ E_j X_{ji} + X_{ij}^\star F_i &= 0, \end{aligned}$$

for $i, j = 1, \ldots, d$.

Note that we have:

$$i = j \rightarrow E_i X_{ii} + X_{ii}^\star F_i = 0 \qquad \text{(1 equation)}$$

$$i \neq j \rightarrow \begin{cases} E_i X_{ij} + X_{ji}^\star F_j = 0 \\ E_j X_{ji} + X_{ij}^\star F_i = 0 \end{cases} \qquad \text{(system of 2 equations)}$$

Partition into blocks

Lemma

Let $E=\text{diag}(E_1,\ldots,E_d)$ and $F^\star=\text{diag}(F_1^\star,\ldots,F_d^\star)$, and partition $X=[X_{ij}]_{i,j=1:d}$. Then

$$EX+X^\star F=0$$

is equivalent to the set of equations

$$\begin{aligned} E_iX_{ij}+X_{ij}^\star F_j&=0\\ E_jX_{ji}+X_{ij}^\star F_i&=0, \end{aligned}$$

for $i,j=1,\ldots,d$.

Note that we have:

$$i=j\rightarrow E_iX_{ii}+X_{ii}^\star F_i=0 \qquad \text{(1 equation)}$$

$$i\neq j\rightarrow \left\{\begin{array}{l} E_iX_{ij}+X_{ji}^\star F_j=0\\ E_jX_{ji}+X_{ij}^\star F_i=0 \end{array}\right. \qquad \text{(system of 2 equations)}$$

Using the KCF

By particularizing to $F+\lambda F^{\star}$ as the KCF of $A+\lambda B^{\star}$, i.e.: direct sum of blocks:

Type 1: “finite blocks": $J_k(\lambda_i)+\lambda I_k$

Type 2: “infinite blocks": $\lambda J_m(0)+I_m$

Type 3: “right singular blocks": L_{ε}

Type 4: “left singular blocks": L_{η}

we have to solve:

(a) $EX+X^{\star}F=0$, with $E+\lambda F^{\star}$ of **type 1–4** $\rightsquigarrow$ **4** equations

(b) $\begin{array}{c} E_iX+Y^{\star}F_j=0 \\ E_jY+X^{\star}F_i=0 \end{array}$, with $E_i+\lambda F_i^{\star}$, $E_j+\lambda F_j^{\star}$ of **type 1–4** $\rightsquigarrow$ **10** systems

The KCF of $A+\lambda B^\star$

Let $A\in\mathbb{C}^{m\times n}$, $B\in\mathbb{C}^{n\times m}$, set $A+\lambda B^\star$ with Kronecker canonical form

$$\begin{aligned} K_1+\lambda K_2^\star = {} & L_{\varepsilon_1}\oplus L_{\varepsilon_2}\oplus\cdots\oplus L_{\varepsilon_a} \\ & \oplus L_{\eta_1}^T\oplus L_{\eta_2}^T\oplus\cdots\oplus L_{\eta_b}^T \\ & \oplus(\lambda J_{u_1}(0)+I_{u_1})\oplus(\lambda J_{u_2}(0)+I_{u_2})\oplus\cdots\oplus(\lambda J_{u_c}(0)+I_{u_c}) \\ & \oplus(J_{k_1}(\mu_1)+\lambda I_{k_1})\oplus(J_{k_2}(\mu_2)+\lambda I_{k_2})\oplus\cdots\oplus(J_{k_d}(\mu_d)+\lambda I_{k_d}), \end{aligned}$$

where $\varepsilon_1\le\varepsilon_2\le\cdots\le\varepsilon_a$, $\eta_1\le\eta_2\le\cdots\le\eta_b$, and $u_1\le u_2\le\cdots\le u_c$. Then the dimension of the solution space of the matrix equation

$$AX+X^\star B=0$$

depends only on $K_1+\lambda K_2^\star$.

Codimension count

Theorem

The **dimension** of the solution space of $AX + X^T B = 0$ is:

$$\dim \mathscr{S}(A,B) = \sum_{i=1}^{a} \varepsilon_i + \sum_{\mu_i=1} \lfloor k_i/2 \rfloor + \sum_{\mu_j=-1} \lceil k_j/2 \rceil +$$

$$\sum_{\substack{i,j=1 \\ i<j}}^{a} (\varepsilon_i + \varepsilon_j) + \sum_{\substack{i<j \\ \mu_i \mu_j = 1}} \min\{k_i, k_j\}$$

$$+ \sum_{i,j} \left(\eta_j - \varepsilon_i + 1\right) +$$

$$a \sum_{i=1}^{c} u_i + a \sum_{i=1}^{d} k_i + \sum_{\substack{i,j \\ \mu_j = 0}} \min\{u_i, k_j\}$$

Solution of $AX + X^\star B = 0$

- **Explicit formulas** available. Depend on P, Q, K_1, K_2, where

$$P(A + \lambda B^\star)Q = K_1 + \lambda K_2{}^\star ,$$

the *KCF* of $A + \lambda B^\star$.

- Solution (and codimension count) **over** $\mathbb{C}$.

Solution of $AX + X^\star B = 0$

- **Explicit formulas** available. Depend on P, Q, K_1, K_2, where

$$P(A + \lambda B^\star)Q = K_1 + \lambda K_2{}^\star,$$

  the *KCF* of $A + \lambda B^\star$.

- Solution (and codimension count) **over** $\mathbb{C}$.

- M. A. BEITIA, J.-M. GRACIA, *Sylvester matrix equation for matrix pencils*, LAA 232 (1996).
- H. W. BRADEN, *The equations $A^TX\pm X^TA=B$*, SIMAX 20 (1998).
- R. BYERS, D. KRESSNER, *Structured condition numbers for invariant subspaces*, SIMAX 28 (2) (2006).
- F. DE TERÁN, F. M. DOPICO, *The solution of the equation $XA+AX^T=0$ and ts application to the theory of orbits*, LAA 434 (2011).
- F. DE TERÁN, F. M. DOPICO, *The equation $AX+X^*B=0$ and the dimension of *congruence orbits*, ELA 22 (2011).
- F. DE TERÁN, F. M. DOPICO, *Consistency and efficient solution of the Sylvester equation for $\star-$congruence: $AX+X^\star B=C$*, ELA 22 (2011).
- F. DE TERÁN, F. M. DOPICO, N. GUILLERY, D. MONTEALEGRE, N. Z. REYES, The solution of the equation $AX+X^\star B=0$, submitted.
- D. KRESSNER, C. SCHRÖDER, D. S. WATKINS, *Implicit QR algorithms for palindromic and even eigenvalue problems*, NA 51(2) (2009).
- H. K. WIMMER, *Roth's Theorems for matrix equations with symmetry constraints*, LAA 199 (1994).