Sylvester equations and the matrix symmetrizer problem

Froilán M. Dopico and Frank Uhlig

Departamento de Matemáticas
Universidad Carlos III de Madrid, Spain
and
Department of Mathematics and Statistics
Auburn University, AL, USA

19th International Linear Algebra Society Conference
Sungkyunkwan University, Seoul Campus
Seoul, Korea. August 6-9, 2014

Outline

Outline

The starting point: Frobenius 1910

Theorem

Let $A \in \mathbb{C}^{n\times n}$. *There are matrices* $B, C \in \mathbb{C}^{n\times n}$ *such that* $B = B^T$, $C = C^T$, *and*

$$A = B\,C.$$

Either B *or* C *can be chosen to be nonsingular.*

Remarks

- The result remains valid for any field $\mathbb{F}$.
- It can be found in standard books (Horn & Johnson, *Matrix Analysis*).
- Olga Taussky paid attention to this result in *"The role of symmetric matrices in the study of general matrices", LAA, 1972.*

The starting point: Frobenius 1910

Theorem

Let $A \in \mathbb{C}^{n\times n}$. There are matrices $B, C \in \mathbb{C}^{n\times n}$ such that $B = B^T$, $C = C^T$, and

$$A = B\,C.$$

Either B or C can be chosen to be nonsingular.

Remarks

- The result remains valid for any field $\mathbb{F}$.
- It can be found in standard books (Horn & Johnson, *Matrix Analysis*).
- Olga Taussky paid attention to this result in *"The role of symmetric matrices in the study of general matrices", LAA, 1972.*

The starting point: Frobenius 1910

Theorem

Let $A \in \mathbb{C}^{n \times n}$*. There are matrices* $B, C \in \mathbb{C}^{n \times n}$ *such that* $B = B^T$*,* $C = C^T$*, and*

$$A = B\,C.$$

Either B *or* C *can be chosen to be nonsingular.*

Remarks

- The result remains valid for any field $\mathbb{F}$.
- It can be found in standard books (Horn & Johnson, *Matrix Analysis*).
- Olga Taussky paid attention to this result in *"The role of symmetric matrices in the study of general matrices", LAA, 1972.*

Theorem

Let $A \in \mathbb{C}^{n\times n}$. There are matrices $B, C \in \mathbb{C}^{n\times n}$ such that $B = B^T$, $C = C^T$, and

$$A = B\,C.$$

Either B or C can be chosen to be nonsingular.

Remarks

- The result remains valid for any field $\mathbb{F}$.
- It can be found in standard books (Horn & Johnson, *Matrix Analysis*).
- Olga Taussky paid attention to this result in *"The role of symmetric matrices in the study of general matrices", LAA, 1972.*

Frobenius's result calls itself for extensions in several directions

Theorem (Frobenius, 1910)

Let $A \in \mathbb{C}^{n \times n}$. There are matrices $B, C \in \mathbb{C}^{n \times n}$ such that $B = B^T$, $C = C^T$, and

$$A = B\,C.$$

Either B or C can be chosen to be nonsingular.

Some relevant potential extensions:

- How many factorizations exist?: characterization of all possible symmetric factors B and C.
- How to compute factors B and C?
- *"Either B or C can be chosen to be nonsingular"* is a bit vague for present times. More informative would be *"Either B or C can be chosen to have condition number as small as..."*. So, *given A*,

  What is the smallest possible cond. number of a nonsingular C (or B)?

Frobenius's result calls itself for extensions in several directions

Theorem (Frobenius, 1910)

Let $A \in \mathbb{C}^{n\times n}$. There are matrices $B, C \in \mathbb{C}^{n\times n}$ such that $B = B^T$, $C = C^T$, and

$$A = B\,C.$$

Either B or C can be chosen to be nonsingular.

Some relevant potential extensions:

- How many factorizations exist?: characterization of all possible symmetric factors B and C.
- How to compute factors B and C?
- *"Either B or C can be chosen to be nonsingular"* is a bit vague for present times. More informative would be *"Either B or C can be chosen to have condition number as small as..."*. So, *given* A,

  What is the smallest possible cond. number of a nonsingular C (or B)?

Frobenius's result calls itself for extensions in several directions

Theorem (Frobenius, 1910)

Let $A \in \mathbb{C}^{n\times n}$. *There are matrices* $B, C \in \mathbb{C}^{n\times n}$ *such that* $B = B^T$, $C = C^T$, *and*

$$A = B\,C.$$

Either B *or* C *can be chosen to be nonsingular.*

Some relevant potential extensions:

- How many factorizations exist?: characterization of all possible symmetric factors B and C.
- How to compute factors B and C?
- *"Either* B *or* C *can be chosen to be nonsingular"* is a bit vague for present times. More informative would be *"Either* B *or* C *can be chosen to have condition number as small as..."*. So, *given* A,

  What is the smallest possible cond. number of a nonsingular C (or B)?

Frobenius's result calls itself for extensions in several directions

Theorem (Frobenius, 1910)

Let $A \in \mathbb{C}^{n\times n}$. There are matrices $B, C \in \mathbb{C}^{n\times n}$ such that $B = B^T$, $C = C^T$, and

$$A = B\,C.$$

Either B or C can be chosen to be nonsingular.

Some relevant potential extensions:

- How many factorizations exist?: characterization of all possible symmetric factors B and C.
- How to compute factors B and C?
- *"Either B or C can be chosen to be nonsingular"* is a bit vague for present times. More informative would be *"Either B or C can be chosen to have condition number as small as..."*. So, *given* A,

  What is the smallest possible cond. number of a nonsingular C (or B)?

- **First, note that one only needs to determine the non-singular factor**: take $C^T = C \in \mathbb{C}^{n\times n}$ in $A = BC$ to be nonsingular, then

$$A = BC \iff AC^{-1} = B.$$

- **Second, note that it is easier to determine** $X = C^{-1}$ since it is a nonsingular solution of a particular **Sylvester** equation with restrictions

$$\begin{aligned} AX - X^T A^T &= 0 \\ X &= X^T \end{aligned}$$

(equivalent to $B = B^T$)
(equivalent to $C = C^T$)

- Solutions X of these equations are called **right symmetrizers** of A.
- There are also **left symmetrizers**, which are right symmetrizers of A^T.
- Baksalary & Kala (LAA-1981) defined symmetrizers $X \neq X^T$ in a more general setting and characterized them via projectors and pseudoinverses (not via matrix equations).

The answers to these questions are easier by defining "symmetrizers"

- **First, note that one only needs to determine the non-singular factor**: take $C^T = C \in \mathbb{C}^{n\times n}$ in $A = BC$ to be nonsingular, then

$$A = BC \iff AC^{-1} = B.$$

- **Second, note that it is easier to determine** $X = C^{-1}$ since it is a nonsingular solution of a particular **Sylvester** equation with restrictions

$$\begin{aligned} AX - X^T A^T &= 0 \\ X &= X^T \end{aligned}$$

(equivalent to $B = B^T$)
(equivalent to $C = C^T$)

- Solutions X of these equations are called **right symmetrizers** of A.
- There are also **left symmetrizers**, which are right symmetrizers of A^T.
- Baksalary & Kala (LAA-1981) defined symmetrizers $X \neq X^T$ in a more general setting and characterized them via projectors and pseudoinverses (not via matrix equations).

The answers to these questions are easier by defining "symmetrizers"

- **First, note that one only needs to determine the non-singular factor**: take $C^T = C \in \mathbb{C}^{n\times n}$ in $A = BC$ to be nonsingular, then

$$A = BC \iff AC^{-1} = B.$$

- **Second, note that it is easier to determine** $X = C^{-1}$ since it is a nonsingular solution of a particular **Sylvester** equation with restrictions

$$\begin{aligned} AX - X^T A^T &= 0 \\ X &= X^T \end{aligned}$$

(equivalent to $B = B^T$)
(equivalent to $C = C^T$)

- Solutions X of these equations are called **right symmetrizers** of A.
- There are also **left symmetrizers**, which are right symmetrizers of A^T.
- Baksalary & Kala (LAA-1981) defined symmetrizers $X \neq X^T$ in a more general setting and characterized them via projectors and pseudoinverses (not via matrix equations).

- **First, note that one only needs to determine the non-singular factor**: take $C^T = C \in \mathbb{C}^{n\times n}$ in $A = BC$ to be nonsingular, then

$$A = BC \iff AC^{-1} = B.$$

- **Second, note that it is easier to determine** $X = C^{-1}$ since it is a nonsingular solution of a particular **Sylvester** equation with restrictions

$$\begin{aligned} AX - X^T A^T &= 0 \\ X &= X^T \end{aligned}$$

(equivalent to $B = B^T$)
(equivalent to $C = C^T$)

- Solutions X of these equations are called **right symmetrizers** of A.
- There are also **left symmetrizers**, which are right symmetrizers of A^T.
- Baksalary & Kala (LAA-1981) defined symmetrizers $X \neq X^T$ in a more general setting and characterized them via projectors and pseudoinverses (not via matrix equations).

- **First, note that one only needs to determine the non-singular factor**: take $C^T = C \in \mathbb{C}^{n\times n}$ in $A = BC$ to be nonsingular, then

$$A = BC \iff AC^{-1} = B.$$

- **Second, note that it is easier to determine** $X = C^{-1}$ since it is a nonsingular solution of a particular **Sylvester** equation with restrictions

$$\begin{aligned} AX - X^T A^T &= 0 \\ X &= X^T \end{aligned}$$

(equivalent to $B = B^T$)
(equivalent to $C = C^T$)

- Solutions X of these equations are called **right symmetrizers** of A.
- There are also **left symmetrizers**, which are right symmetrizers of A^T.
- Baksalary & Kala (LAA-1981) defined symmetrizers $X \neq X^T$ in a more general setting and characterized them via projectors and pseudoinverses (not via matrix equations).

The answers to these questions are easier by defining "symmetrizers"

- **First, note that one only needs to determine the non-singular factor**: take $C^T = C \in \mathbb{C}^{n\times n}$ in $A = BC$ to be nonsingular, then

$$A = BC \Longleftrightarrow AC^{-1} = B.$$

- **Second, note that it is easier to determine** $X = C^{-1}$ since it is a nonsingular solution of a particular **Sylvester** equation with restrictions

$$\begin{aligned} AX - X^T A^T &= 0 \\ X &= X^T \end{aligned}$$

(equivalent to $B = B^T$)
(equivalent to $C = C^T$)

- Solutions X of these equations are called **right symmetrizers** of A.
- There are also **left symmetrizers**, which are right symmetrizers of A^T.
- Baksalary & Kala (LAA-1981) defined symmetrizers $X \neq X^T$ in a more general setting and characterized them via projectors and pseudoinverses (not via matrix equations).

My goals in this talk

Given $A \in \mathbb{C}^{n\times n}$ and the Sylvester equation with restrictions

$$\begin{aligned} AX - X^T A^T &= 0 \\ X &= X^T \end{aligned}$$

- I will describe theoretically **the set of all solutions** using standard techniques. Easy problem but I have not found it in the literature.
- I will briefly mention the most **recent algorithms for computing solutions (symmetrizers) in an stable way**. Very difficult problem with relevant recent advances by Frank Uhlig.
- Comments on the **smallest possible condition number** of solutions. Very difficult completely open problem. Recent advances in joint work with Paul Van Dooren.

My goals in this talk

Given $A \in \mathbb{C}^{n\times n}$ and the Sylvester equation with restrictions

$$\begin{aligned} AX - X^T A^T &= 0 \\ X &= X^T \end{aligned}$$

- I will describe theoretically **the set of all solutions** using standard techniques. Easy problem but I have not found it in the literature.
- I will briefly mention the most **recent algorithms for computing solutions (symmetrizers) in an stable way**. Very difficult problem with relevant recent advances by Frank Uhlig.
- Comments on the **smallest possible condition number** of solutions. Very difficult completely open problem. Recent advances in joint work with Paul Van Dooren.

My goals in this talk

Given $A \in \mathbb{C}^{n\times n}$ and the Sylvester equation with restrictions

$$\begin{aligned} AX - X^T A^T &= 0 \\ X &= X^T \end{aligned}$$

- I will describe theoretically **the set of all solutions** using standard techniques. Easy problem but I have not found it in the literature.
- I will briefly mention the most **recent algorithms for computing solutions (symmetrizers) in an stable way**. Very difficult problem with relevant recent advances by Frank Uhlig.
- Comments on the **smallest possible condition number** of solutions. Very difficult completely open problem. Recent advances in joint work with Paul Van Dooren.

Given $A \in \mathbb{C}^{n\times n}$ and the Sylvester equation with restrictions

$$\begin{aligned} AX - X^T A^T &= 0 \\ X &= X^T \end{aligned}$$

- I will describe theoretically **the set of all solutions** using standard techniques. Easy problem but I have not found it in the literature.
- I will briefly mention the most **recent algorithms for computing solutions (symmetrizers) in an stable way**. Very difficult problem with relevant recent advances by Frank Uhlig.
- Comments on the **smallest possible condition number** of solutions. Very difficult completely open problem. Recent advances in joint work with Paul Van Dooren.

Applications of symmetrizers??

- To be honest, nowadays, I do not see any application for them.
- In the past (see Datta (1973), Venkaiah & Sen (JCAM, 1988),...), it was said that symmetrizers allow us to transform a nonsymmetric standard eigenproblem into a generalized symmetric eigenproblem.
- Nonsingular left symmetrizers Y are more convenient for this task

$$A\mathbf{v} = \lambda\mathbf{v} \Longleftrightarrow (YA)\mathbf{v} = \lambda Y\mathbf{v},$$

- But with present eigensolvers, I do not think that this transformation has any advantage.
- But "symmetrizers" are fun, classical, and a highly nontrivial problem!!

- To be honest, nowadays, I do not see any application for them.
- In the past (see Datta (1973), Venkaiah & Sen (JCAM, 1988),...), it was said that symmetrizers allow us to transform a nonsymmetric standard eigenproblem into a generalized symmetric eigenproblem.
- Nonsingular left symmetrizers Y are more convenient for this task

$$A\mathbf{v} = \lambda\mathbf{v} \Longleftrightarrow (YA)\mathbf{v} = \lambda Y\mathbf{v},$$

- But with present eigensolvers, I do not think that this transformation has any advantage.
- But "symmetrizers" are fun, classical, and a highly nontrivial problem!!

Applications of symmetrizers??

- To be honest, nowadays, I do not see any application for them.
- In the past (see Datta (1973), Venkaiah & Sen (JCAM, 1988),...), it was said that symmetrizers allow us to transform a nonsymmetric standard eigenproblem into a generalized symmetric eigenproblem.
- Nonsingular left symmetrizers Y are more convenient for this task

$$A\mathbf{v} = \lambda\mathbf{v} \iff (YA)\mathbf{v} = \lambda Y\mathbf{v},$$

- But with present eigensolvers, I do not think that this transformation has any advantage.
- But "symmetrizers" are fun, classical, and a highly nontrivial problem!!

Applications of symmetrizers??

- To be honest, nowadays, I do not see any application for them.
- In the past (see Datta (1973), Venkaiah & Sen (JCAM, 1988),...), it was said that symmetrizers allow us to transform a nonsymmetric standard eigenproblem into a generalized symmetric eigenproblem.
- Nonsingular left symmetrizers Y are more convenient for this task

$$A\mathbf{v} = \lambda\mathbf{v} \Longleftrightarrow (YA)\mathbf{v} = \lambda Y\mathbf{v},$$

- But with present eigensolvers, I do not think that this transformation has any advantage.
- But "symmetrizers" are fun, classical, and a highly nontrivial problem!!

- To be honest, nowadays, I do not see any application for them.
- In the past (see Datta (1973), Venkaiah & Sen (JCAM, 1988),...), it was said that symmetrizers allow us to transform a nonsymmetric standard eigenproblem into a generalized symmetric eigenproblem.
- Nonsingular left symmetrizers Y are more convenient for this task

$$A\mathbf{v} = \lambda \mathbf{v} \Longleftrightarrow (YA)\mathbf{v} = \lambda Y\mathbf{v},$$

- But with present eigensolvers, I do not think that this transformation has any advantage.
- But "symmetrizers" are fun, classical, and a highly nontrivial problem!!

Outline

Given $A \in \mathbb{C}^{n\times n}$ and the Sylvester equation with restrictions

$$\begin{aligned} AX - X^T A^T &= 0 \\ X &= X^T \end{aligned}$$

- Frobenius in 1910 noted that there are at least n linearly independent solutions (just counting equations and unknowns).
- Taussky and Zassenhaus (Pacific. J. Math., 1959) proved that there are exactly n linearly independent solutions if and only if A is nonderogatory.
- Uhlig (LAA, 1974) describes implicitly the set of solutions for $A \in \mathbb{R}^{n\times n}$ and describes all possible inertias attained by solutions.

Some antecedents that I have found in the literature...

Given $A \in \mathbb{C}^{n\times n}$ and the Sylvester equation with restrictions

$$\begin{aligned} AX - X^T A^T &= 0 \\ X &= X^T \end{aligned}$$

- Frobenius in 1910 noted that there are at least n linearly independent solutions (just counting equations and unknowns).
- Taussky and Zassenhaus (Pacific. J. Math., 1959) proved that there are exactly n linearly independent solutions if and only if A is nonderogatory.
- Uhlig (LAA, 1974) describes implicitly the set of solutions for $A \in \mathbb{R}^{n\times n}$ and describes all possible inertias attained by solutions.

Some antecedents that I have found in the literature...

Given $A \in \mathbb{C}^{n\times n}$ and the Sylvester equation with restrictions

$$\begin{aligned} AX - X^T A^T &= 0 \\ X &= X^T \end{aligned}$$

- Frobenius in 1910 noted that there are at least n linearly independent solutions (just counting equations and unknowns).
- Taussky and Zassenhaus (Pacific. J. Math., 1959) proved that there are exactly n linearly independent solutions if and only if A is nonderogatory.
- Uhlig (LAA, 1974) describes implicitly the set of solutions for $A \in \mathbb{R}^{n\times n}$ and describes all possible inertias attained by solutions.

Given $A \in \mathbb{C}^{n\times n}$ and the Sylvester equation with restrictions

$$\begin{aligned} AX - X^T A^T &= 0 \\ X &= X^T \end{aligned}$$

- Frobenius in 1910 noted that there are at least n linearly independent solutions (just counting equations and unknowns).
- Taussky and Zassenhaus (Pacific. J. Math., 1959) proved that there are exactly n linearly independent solutions if and only if A is nonderogatory.
- Uhlig (LAA, 1974) describes implicitly the set of solutions for $A \in \mathbb{R}^{n\times n}$ and describes all possible inertias attained by solutions.

- Standard technique for Sylvester eqs. (Gantmacher, Horn & Johnson)
- Let

$$A = P\,J\,P^{-1} \quad \text{with}$$
$$J = \text{diag}\left(J_{n_1}(\lambda_1), \ldots, J_{n_q}(\lambda_q)\right) \quad \text{Jordan canonical form (JCF) of } A,$$

where e-values are not necessarily pairwise different. Then

$$\begin{array}{rcl} AX - X^T A^T &=& 0 \\ X &=& X^T \end{array} \iff \begin{array}{rcl} (PJP^{-1})\,X - X^T\,(P^{-T}J^T P^T) &=& 0 \\ X &=& X^T \end{array}$$

$$\iff \begin{array}{rcl} J(P^{-1}XP^{-T}) - (P^{-1}X^T P^{-T})J^T &=& 0 \\ (P^{-1}XP^{-T}) &=& (P^{-1}X^T P^{-T}) \end{array}$$

Define $Y := (P^{-1}XP^{-T})$ and get $\begin{array}{rcl} JY - Y^T J^T &=& 0 \\ Y &=& Y^T \end{array}$

General solution via the Jordan canonical form of A

- Standard technique for Sylvester eqs. (Gantmacher, Horn & Johnson)
- Let

$$A = P\,J\,P^{-1} \qquad \text{with}$$
$$J = \text{diag}\left(J_{n_1}(\lambda_1), \ldots, J_{n_q}(\lambda_q)\right) \qquad \text{Jordan canonical form (JCF) of } A,$$

where e-values are not necessarily pairwise different. Then

$$\begin{array}{rcl} AX - X^T A^T & = & 0 \\ X & = & X^T \end{array} \iff \begin{array}{rcl} (PJP^{-1})\,X - X^T\,(P^{-T}J^TP^T) & = & 0 \\ X & = & X^T \end{array}$$

$$\iff \begin{array}{rcl} J(P^{-1}XP^{-T}) - (P^{-1}X^TP^{-T})J^T & = & 0 \\ (P^{-1}XP^{-T}) & = & (P^{-1}X^TP^{-T}) \end{array}$$

Define $Y := (P^{-1}XP^{-T})$ and get $\begin{array}{rcl} JY - Y^TJ^T & = & 0 \\ Y & = & Y^T \end{array}$

- Standard technique for Sylvester eqs. (Gantmacher, Horn & Johnson)
- Let

$$A = P\,J\,P^{-1} \qquad \text{with}$$
$$J = \text{diag}\left(J_{n_1}(\lambda_1), \ldots, J_{n_q}(\lambda_q)\right) \qquad \text{Jordan canonical form (JCF) of } A,$$

where e-values are not necessarily pairwise different. Then

$$\begin{array}{rcl} AX - X^T A^T & = & 0 \\ X & = & X^T \end{array} \iff \begin{array}{rcl} (PJP^{-1})\,X - X^T\,(P^{-T}J^T P^T) & = & 0 \\ X & = & X^T \end{array}$$

$$\iff \begin{array}{rcl} J(P^{-1}XP^{-T}) - (P^{-1}X^T P^{-T})J^T & = & 0 \\ (P^{-1}XP^{-T}) & = & (P^{-1}X^T P^{-T}) \end{array}$$

Define $Y := (P^{-1}XP^{-T})$ and get $\begin{array}{rcl} JY - Y^T J^T & = & 0 \\ Y & = & Y^T \end{array}$

- Standard technique for Sylvester eqs. (Gantmacher, Horn & Johnson)
- Let

$A = P\,J\,P^{-1}$ with

$J = \text{diag}\left(J_{n_1}(\lambda_1), \ldots, J_{n_q}(\lambda_q) \right)$ Jordan canonical form (JCF) of A,

where e-values are not necessarily pairwise different. Then

$$\begin{array}{rcl} AX - X^T A^T & = & 0 \\ X & = & X^T \end{array} \iff \begin{array}{rcl} (PJP^{-1})\,X - X^T\,(P^{-T} J^T P^T) & = & 0 \\ X & = & X^T \end{array}$$

$$\iff \begin{array}{rcl} J(P^{-1} X P^{-T}) - (P^{-1} X^T P^{-T}) J^T & = & 0 \\ (P^{-1} X P^{-T}) & = & (P^{-1} X^T P^{-T}) \end{array}$$

Define $Y := (P^{-1} X P^{-T})$ and get

$$\begin{array}{rcl} JY - Y^T J^T & = & 0 \\ Y & = & Y^T \end{array}$$

- Standard technique for Sylvester eqs. (Gantmacher, Horn & Johnson)
- Let

$$A = P\,J\,P^{-1} \qquad \text{with}$$

$$J = \text{diag}\left(J_{n_1}(\lambda_1), \ldots, J_{n_q}(\lambda_q)\right) \qquad \text{Jordan canonical form (JCF) of } A,$$

where e-values are not necessarily pairwise different. Then

$$\begin{array}{rcl} AX - X^T A^T & = & 0 \\ X & = & X^T \end{array} \iff \begin{array}{rcl} (PJP^{-1})\,X - X^T\,(P^{-T}J^TP^T) & = & 0 \\ X & = & X^T \end{array}$$

$$\iff \begin{array}{rcl} J(P^{-1}XP^{-T}) - (P^{-1}X^TP^{-T})J^T & = & 0 \\ (P^{-1}XP^{-T}) & = & (P^{-1}X^TP^{-T}) \end{array}$$

Define $Y := (P^{-1}XP^{-T})$ and get $\begin{array}{rcl} JY - Y^TJ^T & = & 0 \\ Y & = & Y^T \end{array}$

Theorem (General solution)

Let $A = P\, \text{diag}\left(J_{n_1}(\lambda_1), \ldots, J_{n_q}(\lambda_q) \right) P^{-1}$ *be the JCF of* $A \in \mathbb{C}^{n\times n}$*. Then the general solution of*

$$\begin{aligned} AX - X^T A^T &= 0 \\ X &= X^T \end{aligned}$$

is given by the formula $X = P\, Y\, P^T$, *where*

$$Y = (Y_{ij})_{i,j=1}^{q}$$

with

(a) $Y_{ij} \in \mathbb{C}^{n_i \times n_j}$,

(b) $Y_{ij} = Y_{ji}^T$,

(c) $Y_{ij} = 0$ *if* $\lambda_i \neq \lambda_j$,

(d) Y_{ij} **is an upper antitriangular Hankel matrix if** $\lambda_i = \lambda_j$

Example for a "set of all symmetrizers"

Consider

$$A = \text{diag}\,(\,J_3(1)\,,\;J_2(1)\,,\;J_3(2)\,)\,.$$

Then the set of all (right) symmetrizers of A consists of the matrices of the form

$$X = \left[\begin{array}{ccc|cc|ccc} \mathbf{a} & \mathbf{b} & \mathbf{c} & \mathbf{e} & \mathbf{f} & 0 & 0 & 0 \\ \mathbf{b} & \mathbf{c} & 0 & \mathbf{f} & 0 & 0 & 0 & 0 \\ \mathbf{c} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline \mathbf{e} & \mathbf{f} & 0 & \mathbf{g} & \mathbf{h} & 0 & 0 & 0 \\ \mathbf{f} & 0 & 0 & \mathbf{h} & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & \mathbf{r} & \mathbf{s} & \mathbf{t} \\ 0 & 0 & 0 & 0 & 0 & \mathbf{s} & \mathbf{t} & 0 \\ 0 & 0 & 0 & 0 & 0 & \mathbf{t} & 0 & 0 \end{array}\right],$$

where all the letters are free parameters.

Theorem (Dimension of the solution space)

Let $A = P\, \text{diag}\left(J_{n_1}(\lambda_1), \ldots, J_{n_q}(\lambda_q) \right) P^{-1}$ *be the JCF of* $A \in \mathbb{C}^{n\times n}$*. Then the space* $\mathcal{S}$ *of solutions of*

$$\begin{array}{rcl} AX - X^T A^T & = & 0 \\ X & = & X^T \end{array}$$

has dimension

$$\dim \mathcal{S} = \sum_{1\leq i\leq j\leq q} \nu_{ij}\,, \qquad \text{with} \quad \nu_{ij} = \begin{cases} 0, & \text{if } \lambda_i \neq \lambda_j, \\ \min\{n_i, n_j\} & \text{if } \lambda_i = \lambda_j. \end{cases}$$

This dimension can be also expressed as follows: let $\Lambda(A)$ *be the set of* **distinct** *eigenvalues of* A *and for each* $\lambda \in \Lambda(A)$*, let*

$$n_1(\lambda) \geq n_2(\lambda) \geq \cdots \qquad \text{be the sizes of its Jordan blocks,}$$

then

$$\dim \mathcal{S} = \sum_{\lambda\in\Lambda(A)} (n_1(\lambda) + 2\, n_2(\lambda) + 3\, n_3(\lambda) + \cdots)$$

Theorem (Dimension of the solution space)

Let $A = P\, \text{diag}\left(J_{n_1}(\lambda_1), \ldots, J_{n_q}(\lambda_q) \right) P^{-1}$ *be the JCF of* $A \in \mathbb{C}^{n\times n}$. *Then the space* $\mathcal{S}$ *of solutions of*

$$\begin{aligned} AX - X^T A^T &= 0 \\ X &= X^T \end{aligned}$$

has dimension

$$\dim \mathcal{S} = \sum_{1\le i\le j\le q} \nu_{ij}\,, \qquad \text{with} \quad \nu_{ij} = \begin{cases} 0, & \text{if } \lambda_i \neq \lambda_j, \\ \min\{n_i, n_j\} & \text{if } \lambda_i = \lambda_j. \end{cases}$$

This dimension can be also expressed as follows: let $\Lambda(A)$ *be the set of* **distinct** *eigenvalues of* A *and for each* $\lambda \in \Lambda(A)$, *let*

$$n_1(\lambda) \geq n_2(\lambda) \geq \cdots \qquad \text{be the sizes of its Jordan blocks,}$$

then

$$\dim \mathcal{S} = \sum_{\lambda\in\Lambda(A)} (n_1(\lambda) + 2\, n_2(\lambda) + 3\, n_3(\lambda) + \cdots)$$

Corollary (Diagonalizable matrices)

Let $A = V\, diag\,(\lambda_1, \dots, \lambda_n)\, V^{-1} \in \mathbb{C}^{n\times n}$ *be diagonalizable. Then, all matrices of the form*

$$X = V\,D\,V^T, \qquad \text{with } D \text{ any arbitrary diagonal matrix}$$

are solutions of

$$\begin{aligned} AX - X^T A^T &= 0 \\ X &= X^T \end{aligned}$$

Moreover, if $\lambda_i \neq \lambda_j$ *whenever* $i \neq j$, *then all solutions have this form.*

Remark

- Since **the columns of V are eigenvectors of A and almost all matrices are diagonalizable, this result opens the possibility of computing symmetrizers just by taking $D = I_n$,**
- **...but other D may be more convenient for good conditioning.**

Corollary (Diagonalizable matrices)

Let $A = V\, \mathit{diag}(\lambda_1, \ldots, \lambda_n)\, V^{-1} \in \mathbb{C}^{n\times n}$ be diagonalizable. Then, all matrices of the form

$$X = V\, D\, V^T, \qquad \text{with } D \text{ any arbitrary diagonal matrix}$$

are solutions of
$$\begin{aligned} AX - X^T A^T &= 0 \\ X &= X^T \end{aligned}$$

Moreover, if $\lambda_i \neq \lambda_j$ whenever $i \neq j$, then all solutions have this form.

Remark

- Since **the columns of V are eigenvectors of A and almost all matrices are diagonalizable, this result opens the possibility of computing symmetrizers just by taking $D = I_n$,**
- **...but other D may be more convenient for good conditioning.**

Corollary (Diagonalizable matrices)

Let $A = V \, diag(\lambda_1, \ldots, \lambda_n) \, V^{-1} \in \mathbb{C}^{n \times n}$ *be diagonalizable. Then, all matrices of the form*

$$X = V \, D \, V^T, \qquad \text{with } D \text{ any arbitrary diagonal matrix}$$

are solutions of

$$\begin{array}{rcl} AX - X^T A^T & = & 0 \\ X & = & X^T \end{array}$$

Moreover, if $\lambda_i \neq \lambda_j$ *whenever* $i \neq j$*, then all solutions have this form.*

Remark

- Since **the columns of V are eigenvectors of A and almost all matrices are diagonalizable, this result opens the possibility of computing symmetrizers just by taking $D = I_n$,**
- **...but other D may be more convenient for good conditioning.**

Corollary (Diagonalizable matrices)

Let $A = V\, \text{diag}(\lambda_1, \ldots, \lambda_n)\, V^{-1} \in \mathbb{C}^{n\times n}$ *be diagonalizable. Then, all matrices of the form*

$$X = V\, D\, V^T, \qquad \text{with } D \text{ any arbitrary diagonal matrix}$$

are solutions of

$$\begin{aligned} AX - X^T A^T &= 0 \\ X &= X^T \end{aligned}$$

Moreover, if $\lambda_i \neq \lambda_j$ *whenever* $i \neq j$*, then all solutions have this form.*

Remark

- Since **the columns of V are eigenvectors of A and almost all matrices are diagonalizable, this result opens the possibility of computing symmetrizers just by taking $D = I_n$,**
- **...but other D may be more convenient for good conditioning.**

Corollary (The simplest symmetrizer)

Let $A = P\, \text{diag}\left(J_{n_1}(\lambda_1), \ldots, J_{n_q}(\lambda_q) \right) P^{-1}$ *be the JCF of* $A \in \mathbb{C}^{n\times n}$ *and let*

$$E_i = \begin{bmatrix} 0 & & & 0 & 1 \\ & \iddots & \iddots & & 0 \\ 0 & \iddots & \iddots & & \\ 1 & 0 & & & 0 \end{bmatrix} \in \mathbb{C}^{n_i \times n_i}$$

be the reverse identity matrix. Then

$$X = P \begin{bmatrix} E_1 & & & \\ & E_2 & & \\ & & \ddots & \\ & & & E_q \end{bmatrix} P^T$$

is a solution of

$$\begin{aligned} AX - X^T A^T &= 0 \\ X &= X^T \end{aligned}$$

Outline

- There were attempts from early 1960s through 1970s to compute symmetrizers by several researchers
  - J. Howland & F. Farrel (1963),
  - J. Howland (1971),
  - B. N. Datta (1973),
  - L. Trapp (1975), ...
- These attempts used Hessenberg reduction and **they were unstable**.
- The problem of computing symmetrizers lay dormant and has been reconsidered again recently by Frank Uhlig (LAMA, 2012) and in D. & Frank Uhlig (Submitted, 2014).

A brief history on computing symmetrizers

- There were attempts from early 1960s through 1970s to compute symmetrizers by several researchers
  - J. Howland & F. Farrel (1963),
  - J. Howland (1971),
  - B. N. Datta (1973),
  - L. Trapp (1975), ...
- These attempts used Hessenberg reduction and **they were unstable**.
- The problem of computing symmetrizers lay dormant and has been reconsidered again recently by Frank Uhlig (LAMA, 2012) and in D. & Frank Uhlig (Submitted, 2014).

A brief history on computing symmetrizers

- There were attempts from early 1960s through 1970s to compute symmetrizers by several researchers
  - J. Howland & F. Farrel (1963),
  - J. Howland (1971),
  - B. N. Datta (1973),
  - L. Trapp (1975), ...
- These attempts used Hessenberg reduction and **they were unstable**.
- The problem of computing symmetrizers lay dormant and has been reconsidered again recently by Frank Uhlig (LAMA, 2012) and in D. & Frank Uhlig (Submitted, 2014).

The most important message on computing symmetrizers is...

- **It is easy to compute symmetrizers in an stable and efficient way ($O(n^3)$ cost)**,
- but, in difficult situations, it is not clear **how to compute full-rank well-conditioned symmetrizers**,
- but there are advances.

Remark

Compute in difficult situations "sufficiently" well-conditioned symmetrizers **is very hard** at present and one reason is that there are no yet theoretical results on the smallest condition number of the symmetrizers of a given matrix $A \in \mathbb{C}^{n \times n}$.

The most important message on computing symmetrizers is...

- **It is easy to compute symmetrizers in an stable and efficient way ($O(n^3)$ cost)**,
- but, in difficult situations, it is not clear **how to compute full-rank well-conditioned symmetrizers**,
- but there are advances.

Remark

Compute in difficult situations "sufficiently" well-conditioned symmetrizers **is very hard** at present and one reason is that there are no yet theoretical results on the smallest condition number of the symmetrizers of a given matrix $A \in \mathbb{C}^{n \times n}$.

The most important message on computing symmetrizers is...

- **It is easy to compute symmetrizers in an stable and efficient way ($O(n^3)$ cost)**,
- but, in difficult situations, it is not clear **how to compute full-rank well-conditioned symmetrizers**,
- but there are advances.

Remark

Compute in difficult situations "sufficiently" well-conditioned symmetrizers **is very hard** at present and one reason is that there are no yet theoretical results on the smallest condition number of the symmetrizers of a given matrix $A \in \mathbb{C}^{n \times n}$.

- **It is easy to compute symmetrizers in an stable and efficient way ($O(n^3)$ cost)**,
- but, in difficult situations, it is not clear **how to compute full-rank well-conditioned symmetrizers**,
- but there are advances.

Remark

Compute in difficult situations “sufficiently” well-conditioned symmetrizers **is very hard** at present and one reason is that there are no yet theoretical results on the smallest condition number of the symmetrizers of a given matrix $A \in \mathbb{C}^{n\times n}$.

Recall: if $A = V\,\Lambda\,V^{-1}$ $\in \mathbb{C}^{n\times n}$ is **diagonalizable**, then, all matrices of the form $X = V\,D\,V^T$, with D arbitrary diagonal matrix are symmetrizers of A.

MATLAB BASIC ALGORITHM

```
[V,Lamb] = eig(A)
X = V*V'
```

- `eig` command is Francis QR + computing eigenvectors, so $O(n^3)$ **cost**.
- **This works almost always very well**, but it was not used in the 1960s, 1970s,...: why waiting until D. & Uhlig, 2014?
- We are taking $D = I$ (MATLAB gives columns of norm 1) but other options may be more convenient for good conditioning...
- because if V is very ill-conditioned $X = (V\,V^T)$ **can be also very ill conditioned or even singular in floating point arithmetic.**
- **The residual errors can be proved to be very good.**

Recall: if $A = V \Lambda V^{-1}$ $\in \mathbb{C}^{n\times n}$ is **diagonalizable**, then, all matrices of the form $X = V D V^T$, with D arbitrary diagonal matrix are symmetrizers of A.

MATLAB BASIC ALGORITHM

```
[V,Lamb] = eig(A)
X = V*V'
```

- eig command is Francis QR + computing eigenvectors, so $O(n^3)$ **cost**.
- **This works almost always very well**, but it was not used in the 1960s, 1970s,...: why waiting until D. & Uhlig, 2014?
- We are taking $D = I$ (MATLAB gives columns of norm 1) but other options may be more convenient for good conditioning...
- because if V is very ill-conditioned $X = (V\,V^T)$ **can be also very ill conditioned or even singular in floating point arithmetic.**
- **The residual errors can be proved to be very good.**

Computing Symmetrizers via eigenanalysis (I): basic approach

Recall: if $A = V \Lambda V^{-1} \in \mathbb{C}^{n\times n}$ is **diagonalizable**, then, all matrices of the form $X = V D V^T$, with D arbitrary diagonal matrix are symmetrizers of A.

MATLAB BASIC ALGORITHM

```
[V,Lamb] = eig(A)
X = V*V'
```

- `eig` command is Francis QR + computing eigenvectors, so $O(n^3)$ **cost**.
- **This works almost always very well**, but it was not used in the 1960s, 1970s,...: why waiting until D. & Uhlig, 2014?
- We are taking $D = I$ (MATLAB gives columns of norm 1) but other options may be more convenient for good conditioning...
- because if V is very ill-conditioned $X = (V\,V^T)$ **can be also very ill conditioned or even singular in floating point arithmetic.**
- **The residual errors can be proved to be very good.**

Computing Symmetrizers via eigenanalysis (I): basic approach

Recall: if $A = V\,\Lambda\,V^{-1}$ $\in \mathbb{C}^{n\times n}$ is **diagonalizable**, then, all matrices of the form $X = V\,D\,V^T$, with D arbitrary diagonal matrix are symmetrizers of A.

MATLAB BASIC ALGORITHM

```
[V,Lamb] = eig(A)
X = V*V'
```

- `eig` command is Francis QR + computing eigenvectors, so $O(n^3)$ **cost**.
- **This works almost always very well**, but it was not used in the 1960s, 1970s,...: why waiting until D. & Uhlig, 2014?
- We are taking $D = I$ (MATLAB gives columns of norm 1) but other options may be more convenient for good conditioning...
- because if V is very ill-conditioned $X = (V\,V^T)$ **can be also very ill conditioned or even singular in floating point arithmetic.**
- **The residual errors can be proved to be very good.**

Recall: if $A = V\,\Lambda\,V^{-1} \in \mathbb{C}^{n\times n}$ is **diagonalizable**, then, all matrices of the form $X = V\,D\,V^T$, with D arbitrary diagonal matrix are symmetrizers of A.

MATLAB BASIC ALGORITHM

```
[V,Lamb] = eig(A)
X = V*V'
```

- `eig` command is Francis QR + computing eigenvectors, so $O(n^3)$ **cost**.
- **This works almost always very well**, but it was not used in the 1960s, 1970s,...: why waiting until D. & Uhlig, 2014?
- We are taking $D = I$ (MATLAB gives columns of norm 1) but other options may be more convenient for good conditioning...
- because if V is very ill-conditioned $X = (V\,V^T)$ **can be also very ill conditioned or even singular in floating point arithmetic.**
- **The residual errors can be proved to be very good.**

Computing Symmetrizers via eigenanalysis (I): basic approach

Recall: if $A = V \Lambda V^{-1} \in \mathbb{C}^{n \times n}$ is **diagonalizable**, then, all matrices of the form $X = V D V^T$, with D arbitrary diagonal matrix are symmetrizers of A.

MATLAB BASIC ALGORITHM

```
[V,Lamb] = eig(A)
X = V*V'
```

- `eig` command is Francis QR + computing eigenvectors, so $O(n^3)$ **cost**.
- **This works almost always very well**, but it was not used in the 1960s, 1970s,...: why waiting until D. & Uhlig, 2014?
- We are taking $D = I$ (MATLAB gives columns of norm 1) but other options may be more convenient for good conditioning...
- because if V is very ill-conditioned $X = (V V^T)$ **can be also very ill conditioned or even singular in floating point arithmetic.**
- **The residual errors can be proved to be very good.**

Recall: if $A = V \Lambda V^{-1}$ $\in \mathbb{C}^{n\times n}$ is **diagonalizable**, then, all matrices of the form $X = V D V^T$, with D arbitrary diagonal matrix are symmetrizers of A.

MATLAB BASIC ALGORITHM

```
[V,Lamb] = eig(A)
X = V*V'
```

- `eig` command is Francis QR + computing eigenvectors, so $O(n^3)$ **cost**.
- **This works almost always very well**, but it was not used in the 1960s, 1970s,...: why waiting until D. & Uhlig, 2014?
- We are taking $D = I$ (MATLAB gives columns of norm 1) but other options may be more convenient for good conditioning...
- because if V is very ill-conditioned $X = (V V^T)$ **can be also very ill conditioned or even singular in floating point arithmetic.**
- **The residual errors can be proved to be very good**.

Computing Symmetrizers via eigenanalysis (II): errors in basic approach

Theorem

Let $\widehat{V}$ be the eigenvector matrix of $A \in \mathbb{C}^{n\times n}$ computed by MATLAB and let $\widehat{X} = \mathsf{fl}(\widehat{V}\, D\, \widehat{V}^T)$ be the computed symmetrizer of A for a diagonal matrix D, then

$$\frac{\|A\widehat{X} - \widehat{X}A^T\|_2}{\|A\|_2\,\|\widehat{X}\|_2} \leq p(n)\,\mathbf{u}\,\frac{\|\widehat{V}\,D^{1/2}\|_2^2}{\|\widehat{X}\|_2},$$

with $p(n)$ a low-degree polynomial and $\mathbf{u}$ the unit-roundoff of the computer.

Remarks:

- The expected error in Sylvester equations is

$$\frac{\|A\widehat{X} - \widehat{X}A^T\|_2}{\|A\|_2\,\|\widehat{X}\|_2} \leq p(n)\,\mathbf{u}.$$

- The extra factor is unavoidable in this approach, since it cames from computing $\widehat{X} = fl(\widehat{V}\,D\,\widehat{V}^T)$,
- but we have not observed it in the many many tests we have performed.

Computing Symmetrizers via eigenanalysis (II): errors in basic approach

Theorem

Let $\widehat{V}$ be the eigenvector matrix of $A \in \mathbb{C}^{n\times n}$ computed by MATLAB and let $\widehat{X} = \mathrm{fl}(\widehat{V}\, D\, \widehat{V}^T)$ be the computed symmetrizer of A for a diagonal matrix D, then

$$\frac{\|A\widehat{X} - \widehat{X}A^T\|_2}{\|A\|_2\,\|\widehat{X}\|_2} \leq p(n)\,\mathbf{u}\,\frac{\|\widehat{V}\,D^{1/2}\|_2^2}{\|\widehat{X}\|_2},$$

with $p(n)$ a low-degree polynomial and $\mathbf{u}$ the unit-roundoff of the computer.

Remarks:

- The expected error in Sylvester equations is

$$\frac{\|A\widehat{X} - \widehat{X}A^T\|_2}{\|A\|_2\,\|\widehat{X}\|_2} \leq p(n)\,\mathbf{u}.$$

- The extra factor is unavoidable in this approach, since it cames from computing $\widehat{X} = fl(\widehat{V}\,D\,\widehat{V}^T)$,
- but we have not observed it in the many many tests we have performed.

Theorem

Let $\widehat{V}$ *be the eigenvector matrix of* $A \in \mathbb{C}^{n\times n}$ *computed by MATLAB and let* $\widehat{X} = \mathrm{fl}(\widehat{V}\, D\, \widehat{V}^T)$ *be the computed symmetrizer of* A *for a diagonal matrix* D*, then*

$$\frac{\|A\widehat{X} - \widehat{X}A^T\|_2}{\|A\|_2\,\|\widehat{X}\|_2} \leq p(n)\,\mathbf{u}\,\frac{\|\widehat{V}\,D^{1/2}\|_2^2}{\|\widehat{X}\|_2},$$

with $p(n)$ *a low-degree polynomial and* $\mathbf{u}$ *the unit-roundoff of the computer.*

Remarks:

- The expected error in Sylvester equations is

$$\frac{\|A\widehat{X} - \widehat{X}A^T\|_2}{\|A\|_2\,\|\widehat{X}\|_2} \leq p(n)\,\mathbf{u}.$$

- The extra factor is unavoidable in this approach, since it cames from computing $\widehat{X} = fl(\widehat{V}\, D\, \widehat{V}^T)$,
- but we have not observed it in the many many tests we have performed.

Theorem

Let $\widehat{V}$ be the eigenvector matrix of $A \in \mathbb{C}^{n\times n}$ computed by MATLAB and let $\widehat{X} = \mathrm{fl}(\widehat{V}\, D\, \widehat{V}^T)$ be the computed symmetrizer of A for a diagonal matrix D, then

$$\frac{\|A\widehat{X} - \widehat{X}A^T\|_2}{\|A\|_2\,\|\widehat{X}\|_2} \leq p(n)\,\mathbf{u}\,\frac{\|\widehat{V}\,D^{1/2}\|_2^2}{\|\widehat{X}\|_2},$$

with $p(n)$ a low-degree polynomial and $\mathbf{u}$ the unit-roundoff of the computer.

Remarks:

- The expected error in Sylvester equations is

$$\frac{\|A\widehat{X} - \widehat{X}A^T\|_2}{\|A\|_2\,\|\widehat{X}\|_2} \leq p(n)\,\mathbf{u}.$$

- The extra factor is unavoidable in this approach, since it cames from computing $\widehat{X} = fl(\widehat{V}\, D\, \widehat{V}^T)$,
- but we have not observed it in the many many tests we have performed.

- **Choose another $D \neq I_n$:** We do not know yet how to choose D to really get small as possible $\kappa_2(VDV^T)$.
- **I discuss next very very briefly two available approaches:**
  - Still use **eigenanalysis** via orthonormal bases of **principal subspaces** of A (subspaces generated by Jordan chains).
  - Completely different approach via an **iterative method** developed by **Uhlig** (LAMA, 2013) based on **Huang-Nong** method (LAA, 2010) for solving finite dimensional linear operator equations.

- **Choose another $D \neq I_n$:** We do not know yet how to choose D to really get small as possible $\kappa_2(VDV^T)$.
- **I discuss next very very briefly two available approaches:**
  - Still use **eigenanalysis** via orthonormal bases of **principal subspaces** of A (subspaces generated by Jordan chains).
  - Completely different approach via an **iterative method** developed by **Uhlig** (LAMA, 2013) based on **Huang-Nong** method (LAA, 2010) for solving finite dimensional linear operator equations.

- If the (scaled) eigenvector matrix V of A is ill-conditioned, then at least one eigenvalue λ of A is ill-conditioned (Stewart & Sun, 1990), and
- A is very close to a matrix with λ as multiple evalue (Wilkinson, 1972).
- If **one can identify well** a "cluster" of computed multiple eigenvalues (of $A+E$) **and from them a "unique" multiple approx evalue of a** A, then
- methods by Golub-Wilkinson (1976) + intricate developments allow us to compute reliably orthonormal bases of principal subspaces of A and
- **the Hessenberg matrices that represent A in these bases.**
- **From here it is easy, to compute symmetrizers.**
- **Drawback 1:** The method cost $O(n^4)$ for complicated Jordan structures.
- **Drawback 2:** It only improves conditioning of symmetrizers if "multiple evalues" are well identified. Old open problem in Numerical Lin. Algebra.

- If the (scaled) eigenvector matrix V of A is ill-conditioned, then at least one eigenvalue λ of A is ill-conditioned (Stewart & Sun, 1990), and
- A is very close to a matrix with λ as multiple evalue (Wilkinson, 1972).
- If **one can identify well** a "cluster" of computed multiple eigenvalues (of $A+E$) **and from them a "unique" multiple approx evalue of a** A, then
- methods by Golub-Wilkinson (1976) + intricate developments allow us to compute reliably orthonormal bases of principal subspaces of A and
- **the Hessenberg matrices that represent A in these bases.**
- **From here it is easy, to compute symmetrizers.**
- **Drawback 1:** The method cost $O(n^4)$ for complicated Jordan structures.
- **Drawback 2:** It only improves conditioning of symmetrizers if "multiple evalues" are well identified. Old open problem in Numerical Lin. Algebra.

Method 1: Orthonormal bases of principal subspaces of A

- If the (scaled) eigenvector matrix V of A is ill-conditioned, then at least one eigenvalue λ of A is ill-conditioned (Stewart & Sun, 1990), and
- A is very close to a matrix with λ as multiple evalue (Wilkinson, 1972).
- If **one can identify well** a "cluster" of computed multiple eigenvalues (of $A + E$) **and from them a "unique" multiple approx evalue of a A**, then
- methods by Golub-Wilkinson (1976) + intricate developments allow us to compute reliably orthonormal bases of principal subspaces of A and
- **the Hessenberg matrices that represent A in these bases.**
- **From here it is easy, to compute symmetrizers.**
- **Drawback 1:** The method cost $O(n^4)$ for complicated Jordan structures.
- **Drawback 2:** It only improves conditioning of symmetrizers if "multiple evalues" are well identified. Old open problem in Numerical Lin. Algebra.

Method 1: Orthonormal bases of principal subspaces of A

- If the (scaled) eigenvector matrix V of A is ill-conditioned, then at least one eigenvalue λ of A is ill-conditioned (Stewart & Sun, 1990), and
- A is very close to a matrix with λ as multiple evalue (Wilkinson, 1972).
- If **one can identify well** a "cluster" of computed multiple eigenvalues (of $A+E$) **and from them a "unique" multiple approx evalue of a A**, then
- methods by Golub-Wilkinson (1976) + intricate developments allow us to compute reliably orthonormal bases of principal subspaces of A and
- **the Hessenberg matrices that represent A in these bases.**
- **From here it is easy, to compute symmetrizers.**
- **Drawback 1:** The method cost $O(n^4)$ for complicated Jordan structures.
- **Drawback 2:** It only improves conditioning of symmetrizers if "multiple evalues" are well identified. Old open problem in Numerical Lin. Algebra.

Method 1: Orthonormal bases of principal subspaces of A

- If the (scaled) eigenvector matrix V of A is ill-conditioned, then at least one eigenvalue λ of A is ill-conditioned (Stewart & Sun, 1990), and
- A is very close to a matrix with λ as multiple evalue (Wilkinson, 1972).
- If **one can identify well** a "cluster" of computed multiple eigenvalues (of $A+E$) **and from them a "unique" multiple approx evalue of a A**, then
- methods by Golub-Wilkinson (1976) + intricate developments allow us to compute reliably orthonormal bases of principal subspaces of A and
- **the Hessenberg matrices that represent A in these bases.**
- **From here it is easy, to compute symmetrizers.**
- **Drawback 1:** The method cost $O(n^4)$ for complicated Jordan structures.
- **Drawback 2:** It only improves conditioning of symmetrizers if "multiple evalues" are well identified. Old open problem in Numerical Lin. Algebra.

- If the (scaled) eigenvector matrix V of A is ill-conditioned, then at least one eigenvalue λ of A is ill-conditioned (Stewart & Sun, 1990), and
- A is very close to a matrix with λ as multiple evalue (Wilkinson, 1972).
- If **one can identify well** a "cluster" of computed multiple eigenvalues (of $A + E$) **and from them a "unique" multiple approx evalue of a A**, then
- methods by Golub-Wilkinson (1976) + intricate developments allow us to compute reliably orthonormal bases of principal subspaces of A and
- **the Hessenberg matrices that represent A in these bases.**
- **From here it is easy, to compute symmetrizers.**
- **Drawback 1:** The method cost $O(n^4)$ for complicated Jordan structures.
- **Drawback 2:** It only improves conditioning of symmetrizers if "multiple evalues" are well identified. Old open problem in Numerical Lin. Algebra.

Method 1: Orthonormal bases of principal subspaces of A

- If the (scaled) eigenvector matrix V of A is ill-conditioned, then at least one eigenvalue λ of A is ill-conditioned (Stewart & Sun, 1990), and
- A is very close to a matrix with λ as multiple evalue (Wilkinson, 1972).
- If **one can identify well** a "cluster" of computed multiple eigenvalues (of $A + E$) **and from them a "unique" multiple approx evalue of a** A, then
- methods by Golub-Wilkinson (1976) + intricate developments allow us to compute reliably orthonormal bases of principal subspaces of A and
- **the Hessenberg matrices that represent A in these bases.**
- **From here it is easy, to compute symmetrizers.**
- **Drawback 1:** The method cost $O(n^4)$ for complicated Jordan structures.
- **Drawback 2:** It only improves conditioning of symmetrizers if "multiple evalues" are well identified. Old open problem in Numerical Lin. Algebra.

Method 1: Orthonormal bases of principal subspaces of A

- If the (scaled) eigenvector matrix V of A is ill-conditioned, then at least one eigenvalue λ of A is ill-conditioned (Stewart & Sun, 1990), and
- A is very close to a matrix with λ as multiple evalue (Wilkinson, 1972).
- If **one can identify well** a "cluster" of computed multiple eigenvalues (of $A + E$) **and from them a "unique" multiple approx evalue of a A**, then
- methods by Golub-Wilkinson (1976) + intricate developments allow us to compute reliably orthonormal bases of principal subspaces of A and
- **the Hessenberg matrices that represent A in these bases.**
- **From here it is easy, to compute symmetrizers.**
- **Drawback 1:** The method cost $O(n^4)$ for complicated Jordan structures.
- **Drawback 2:** It only improves conditioning of symmetrizers if "multiple evalues" are well identified. Old open problem in Numerical Lin. Algebra.

- Given $T : H_1 \longrightarrow H_2$ linear operator on finite dimensional inner-product spaces, Huang & Nong (LAA, 2010) developed an iterative algorithm for solving

$$T(x) = f$$

- Some features of this algorithm are
  - It resembles BICG.
  - It converges always in a finite number of steps.
  - It works when there are infinite solutions.
- Uhlig (LAMA, 2012) adapts Huang & Nong's algor in a non-trivial way to the symmetrizer problem by taking

$$T(X) = \begin{bmatrix} AX - X^T A^T \\ X - X^T \end{bmatrix}, \quad f = 0, \quad H_1 = \mathbb{C}^{n\times n}, \quad H_2 = \mathbb{C}^{2n\times n}$$

- **Positive property:** It finds **better conditioned symmetrizers in difficult situations** than methods based on eigenanalysis. (Why????)
- **Negative property:** It is **very slow** $O(n^5)$ up to $O(n^7)$ cost.

Method 2: Iterative method

- Given $T : H_1 \longrightarrow H_2$ linear operator on finite dimensional inner-product spaces, Huang & Nong (LAA, 2010) developed an iterative algorithm for solving
$$T(x) = f$$
- Some features of this algorithm are
  - It resembles BICG.
  - It converges always in a finite number of steps.
  - It works when there are infinite solutions.
- Uhlig (LAMA, 2012) adapts Huang & Nong's algor in a non-trivial way to the symmetrizer problem by taking
$$T(X) = \begin{bmatrix} AX - X^T A^T \\ X - X^T \end{bmatrix}, \quad f = 0, \quad H_1 = \mathbb{C}^{n\times n}, \quad H_2 = \mathbb{C}^{2n\times n}$$
- **Positive property:** It finds **better conditioned symmetrizers in difficult situations** than methods based on eigenanalysis. (Why????)
- **Negative property:** It is **very slow** $O(n^5)$ up to $O(n^7)$ cost.

Method 2: Iterative method

- Given $T : H_1 \longrightarrow H_2$ linear operator on finite dimensional inner-product spaces, Huang & Nong (LAA, 2010) developed an iterative algorithm for solving
$$T(x) = f$$
- Some features of this algorithm are
    - It resembles BICG.
    - It converges always in a finite number of steps.
    - It works when there are infinite solutions.
- Uhlig (LAMA, 2012) adapts Huang & Nong's algor in a non-trivial way to the symmetrizer problem by taking
$$T(X) = \begin{bmatrix} AX - X^T A^T \\ X - X^T \end{bmatrix}, \quad f = 0, \quad H_1 = \mathbb{C}^{n\times n}, \quad H_2 = \mathbb{C}^{2n\times n}$$
- **Positive property:** It finds **better conditioned symmetrizers in difficult situations** than methods based on eigenanalysis. (Why????)
- **Negative property:** It is **very slow** $O(n^5)$ up to $O(n^7)$ cost.

- Given $T : H_1 \longrightarrow H_2$ linear operator on finite dimensional inner-product spaces, Huang & Nong (LAA, 2010) developed an iterative algorithm for solving

$$T(x) = f$$

- Some features of this algorithm are
  - It resembles BICG.
  - It converges always in a finite number of steps.
  - It works when there are infinite solutions.
- Uhlig (LAMA, 2012) adapts Huang & Nong's algor in a non-trivial way to the symmetrizer problem by taking

$$T(X) = \begin{bmatrix} AX - X^T A^T \\ X - X^T \end{bmatrix}, \quad f = 0, \quad H_1 = \mathbb{C}^{n\times n}, \quad H_2 = \mathbb{C}^{2n\times n}$$

- **Positive property:** It finds **better conditioned symmetrizers in difficult situations** than methods based on eigenanalysis. (Why????)
- **Negative property:** It is **very slow** $O(n^5)$ up to $O(n^7)$ cost.

- Given $T : H_1 \longrightarrow H_2$ linear operator on finite dimensional inner-product spaces, Huang & Nong (LAA, 2010) developed an iterative algorithm for solving

$$T(x) = f$$

- Some features of this algorithm are
  - It resembles BICG.
  - It converges always in a finite number of steps.
  - It works when there are infinite solutions.
- Uhlig (LAMA, 2012) adapts Huang & Nong's algor in a non-trivial way to the symmetrizer problem by taking

$$T(X) = \begin{bmatrix} AX - X^T A^T \\ X - X^T \end{bmatrix}, \quad f = 0, \quad H_1 = \mathbb{C}^{n\times n}, \quad H_2 = \mathbb{C}^{2n\times n}$$

- **Positive property:** It finds **better conditioned symmetrizers in difficult situations** than methods based on eigenanalysis. (Why????)
- **Negative property:** It is **very slow** $O(n^5)$ up to $O(n^7)$ cost.

Numerical test 1: random matrices

Averages on 10 random 100×100 and 10 random 200×200 matrices.

A	max eig cond no	res. error	cond(X)	runtime average
100 by 100				
iter. method		2.2e-9	7.2e+03	13.75
eigenv. method	32.6	9.2e-15	9e+03	0.028
200 by 200				
iter. method		2.2e-7	2.8e+04	152.9
eigenv. method	49	1.5e-14	2.97e+04	0.116

Numerical test 2: 35×35 Frank matrix

Well-known type of matrices with ill-conditioned eigenvalues.

F 35 by 35	max eig cond no	res. error	cond(X)	rank(X)	runtime average
iter. method		8e-14	7e+13	35	0.8
eigenv. method	4.6e+8	2e-10	1.6e+17	32	0.012

Outline

Recall: $A = V \Lambda V^{-1}$ **diagonalizable**, symmetrizers are $X = V D V^T$, with D arbitrary diagonal.

Consider the real diagonalization

$$A = \begin{bmatrix} 1 & 1 \\ 0 & \delta \end{bmatrix} \begin{bmatrix} \lambda_1 & \\ & \lambda_2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & \delta \end{bmatrix}^{-1}$$

For $|\delta| \ll 1$ eigenvector matrix very ill-conditioned. Any symmetrizer of A has the form

$$X(d_1, d_2) = \begin{bmatrix} 1 & 1 \\ 0 & \delta \end{bmatrix} \begin{bmatrix} d_1 & \\ & d_2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & \delta \end{bmatrix}^T = \begin{bmatrix} (d_1 + d_2) & d_2\delta \\ d_2\delta & d_2\delta^2 \end{bmatrix}$$

and

$$\kappa_F(X(d_1, d_2)) = \frac{|d_1 + d_2|^2}{|d_1|\,|d_2|} \frac{1}{\delta^2} + (2 + \delta^2) \frac{|d_2|}{|d_1|}.$$

$$\min_{d_1, d_2} \kappa_F(X(d_1, d_2)) = 2,$$ attained with $d_1 = -d_2\,(1 + \delta^2)$

A 2×2 example

Recall: $A = V \Lambda V^{-1}$ **diagonalizable**, symmetrizers are $X = V D V^T$, with D arbitrary diagonal.

Consider the real diagonalization

$$A = \begin{bmatrix} 1 & 1 \\ 0 & \delta \end{bmatrix} \begin{bmatrix} \lambda_1 & \\ & \lambda_2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & \delta \end{bmatrix}^{-1}$$

For $|\delta| \ll 1$ eigenvector matrix very ill-conditioned. Any symmetrizer of A has the form

$$X(d_1, d_2) = \begin{bmatrix} 1 & 1 \\ 0 & \delta \end{bmatrix} \begin{bmatrix} d_1 & \\ & d_2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & \delta \end{bmatrix}^T = \begin{bmatrix} (d_1 + d_2) & d_2\delta \\ d_2\delta & d_2\delta^2 \end{bmatrix}$$

and

$$\kappa_F(X(d_1, d_2)) = \frac{|d_1 + d_2|^2}{|d_1|\,|d_2|} \frac{1}{\delta^2} + (2 + \delta^2) \frac{|d_2|}{|d_1|}.$$

$$\min_{d_1, d_2} \kappa_F(X(d_1, d_2)) = 2, \qquad \text{attained with } d_1 = -d_2\,(1 + \delta^2)$$

A 2×2 example

Recall: $A = V\,\Lambda\,V^{-1}$ **diagonalizable**, symmetrizers are $X = V\,D\,V^T$, with D arbitrary diagonal.

Consider the real diagonalization

$$A = \begin{bmatrix} 1 & 1 \\ 0 & \delta \end{bmatrix} \begin{bmatrix} \lambda_1 & \\ & \lambda_2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & \delta \end{bmatrix}^{-1}$$

For $|\delta| \ll 1$ eigenvector matrix very ill-conditioned. Any symmetrizer of A has the form

$$X(d_1, d_2) = \begin{bmatrix} 1 & 1 \\ 0 & \delta \end{bmatrix} \begin{bmatrix} d_1 & \\ & d_2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & \delta \end{bmatrix}^T = \begin{bmatrix} (d_1 + d_2) & d_2\delta \\ d_2\delta & d_2\delta^2 \end{bmatrix}$$

and

$$\kappa_F(X(d_1, d_2)) = \frac{|d_1 + d_2|^2}{|d_1|\,|d_2|}\frac{1}{\delta^2} + (2 + \delta^2)\frac{|d_2|}{|d_1|}.$$

$$\min_{d_1, d_2} \kappa_F(X(d_1, d_2)) = 2,$$ attained with $d_1 = -d_2\,(1 + \delta^2)$

A 2×2 example

Recall: $A = V \Lambda V^{-1}$ **diagonalizable**, symmetrizers are $X = V\, D\, V^T$, with D arbitrary diagonal.

Consider the real diagonalization

$$A = \begin{bmatrix} 1 & 1 \\ 0 & \delta \end{bmatrix} \begin{bmatrix} \lambda_1 & \\ & \lambda_2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & \delta \end{bmatrix}^{-1}$$

For $|\delta| \ll 1$ eigenvector matrix very ill-conditioned. Any symmetrizer of A has the form

$$X(d_1, d_2) = \begin{bmatrix} 1 & 1 \\ 0 & \delta \end{bmatrix} \begin{bmatrix} d_1 & \\ & d_2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & \delta \end{bmatrix}^T = \begin{bmatrix} (d_1 + d_2) & d_2 \delta \\ d_2 \delta & d_2 \delta^2 \end{bmatrix}$$

and

$$\kappa_F(X(d_1, d_2)) = \frac{|d_1 + d_2|^2}{|d_1|\,|d_2|} \frac{1}{\delta^2} + (2 + \delta^2) \frac{|d_2|}{|d_1|}.$$

$$\min_{d_1, d_2} \kappa_F(X(d_1, d_2)) = 2,$$ attained with $d_1 = -d_2\,(1 + \delta^2)$

A 2×2 example

Recall: $A = V \Lambda V^{-1}$ **diagonalizable**, symmetrizers are $X = V D V^T$, with D arbitrary diagonal.

Consider the real diagonalization

$$A = \begin{bmatrix} 1 & 1 \\ 0 & \delta \end{bmatrix} \begin{bmatrix} \lambda_1 & \\ & \lambda_2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & \delta \end{bmatrix}^{-1}$$

For $|\delta| \ll 1$ eigenvector matrix very ill-conditioned. Any symmetrizer of A has the form

$$X(d_1, d_2) = \begin{bmatrix} 1 & 1 \\ 0 & \delta \end{bmatrix} \begin{bmatrix} d_1 & \\ & d_2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & \delta \end{bmatrix}^T = \begin{bmatrix} (d_1 + d_2) & d_2\delta \\ d_2\delta & d_2\delta^2 \end{bmatrix}$$

and

$$\kappa_F(X(d_1, d_2)) = \frac{|d_1 + d_2|^2}{|d_1|\,|d_2|} \frac{1}{\delta^2} + (2 + \delta^2) \frac{|d_2|}{|d_1|}.$$

$$\min_{d_1, d_2} \kappa_F(X(d_1, d_2)) = 2,$$ attained with $d_1 = -d_2\,(1 + \delta^2)$

- **Matrices with very ill-conditioned (scaled) eigenvector matrix may have perfectly conditioned symmetrizers.**
- The minimun condition number attained in previous example is attained for any 2×2 diagonalizable matrix.
- For $n \times n$ matrices, we have examples of very-well conditioned symmetrizers of matrices with very ill-conditioned eigenvector matrices,
- but also lower bounds that guarantee that this is not always the case.
- How to get $\min_{D\ diag} \kappa_F(V\, D\, V^T)$?

Outline

Conclusions

- We know how to compute symmetrizers in an stable and very efficient way (via Francis QR),
- most of the times they are well-conditioned,
- but for “difficult” matrices they are not, and we do not know yet which is the lowest condition number of a symmetrizer neither how to compute the corresponding symmetrizers,
- although iterative Frank Uhlig’s method does a fair (slow) job (why??).