The closest normal structured matrix

Erna Begović Kovač

University of Zagreb

ebegovic@fkit.hr

Joint work with Heike Faßbender and Philip Saltenberger
(TU Braunschweig)

AppIMath18
17 September 2018

This work has been supported in part by Croatian Science Foundation under the project 3670.

OUTLINE

- Introduction
- Structured matrices
  and structure-preserving transformations
- Jacobi-type algorithm
  for the reduction to the canonical form
- Finding the closest normal matrix with a given structure
- Numerical examples

INTRODUCTION

- Set of normal matrices: $\mathcal{N} = \{X \;:\; XX^* = X^*X\}$
- X is normal if and only if there is unitary U such that

$$U^*XU = \begin{bmatrix} \diagdown \end{bmatrix}.$$

- A. Ruhe: *Closest normal matrix finally found!*
  BIT 27 (4) (1987) 585–598.

  <u>Does NOT preserve given matrix structure.</u>

INTRODUCTION

- Set of normal matrices: $\mathcal{N} = \{X \ : \ XX^* = X^*X\}$
- X is normal if and only if there is unitary U such that

$$U^*XU = \begin{bmatrix} \ddots & & \\ & & \\ & & \end{bmatrix}.$$

- A. Ruhe: *Closest normal matrix finally found!*
  BIT 27 (4) (1987) 585–598.

Does NOT preserve given matrix structure.

Suppose that A has a structure $\mathcal{S}$, $A \in \mathcal{S}$.

Minimization problem:

$$\min \left\{ \|A - X\|_F^2 \ : \ X \in \mathcal{N} \cap \mathcal{S} \right\}$$

MAXIMIZATION PROBLEM

Theorem (Causey 1964, Gabriel 1979)

Let $A \in \mathbb{C}^{n\times n}$ and let $X = ZDZ^*$, where Z is unitary and D is diagonal. Then X is a nearest normal matrix to A in the Frobenius norm if and only if

(a) $\|\mathrm{diag}(Z^*AZ)\|_F = \max_{QQ^*=I} \|\mathrm{diag}(Q^*AQ)\|_F$, and

(b) $D = \mathrm{diag}(Z^*AZ)$.

→ Finding the closest normal matrix is equivalent to finding an unitary transformation that maximizes Frobenius norm of the diagonal.

→ This theorem has to be modified to fulfill structure-preserving requirements.

- N. J. Higham: *Matrix nearness problem and applications*. In Applications of Matrix theory 22 (1989) 1–27.

STRUCTURED MATRICES

- **Hamiltonian** A (J-Hermitian):

$$(JA)^* = JA, \quad \text{that is } A^* = JAJ, \quad \text{where} \quad J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}.$$

We can write it as

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & -A_{11}^* \end{bmatrix}, \qquad A_{12}^* = A_{12},\; A_{21}^* = A_{21}.$$

STRUCTURED MATRICES

- **Hamiltonian** A (J-Hermitian):

$$(JA)^* = JA, \quad \text{that is } A^* = JAJ, \quad \text{where} \quad J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}.$$

  We can write it as

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & -A_{11}^* \end{bmatrix}, \qquad A_{12}^* = A_{12},\ A_{21}^* = A_{21}.$$

- **Skew-Hamiltonian** A (J-skew-Hermitian):

$$(JA)^* = -JA, \quad \text{that is } A^* = -JAJ.$$

  We can write it as

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{11}^* \end{bmatrix}, \qquad A_{12}^* = -A_{12},\ A_{21}^* = -A_{21}.$$

- For every skew-Hamiltonian W there is Hamiltonian H (and viceversa) such that $W = \imath H$.

STRUCTURED MATRICES–cont.

- **Per-Hermitian** A (R-Hermitian):

$$(RA)^* = RA, \quad \text{that is } A^* = RAR,$$

$$\text{where} \quad R = \begin{bmatrix} 0 & \cdots & 0 & 1 \\ \vdots & & \iddots & 0 \\ 0 & \iddots & & \vdots \\ 1 & 0 & \cdots & 0 \end{bmatrix}.$$

$\rightarrow$ Hermitian about its anti-diagonal

- **Perskew-Hermitian** A (R-skew-Hermitian):

$$(RA)^* = -RA, \quad \text{that is } A^* = -RAR.$$

$\rightarrow$ Skew-Hermitian about its anti-diagonal

- For every perskew-Hermitian K there is per-Hermitian M (and viceversa) such that $K = \imath M$.

STRUCTURE-PRESERVING TRANSFORMATIONS

- For Hamiltonian and skew-Hamiltonian

  $\rightarrow$ J-unitary

- For per-Hermitian and perskew-Hermitian

  $\rightarrow$ R-unitary

STRUCTURE-PRESERVING TRANSFORMATIONS

- For Hamiltonian and skew-Hamiltonian

$$M \text{ is } \textbf{symplectic} \text{ if } M^*JM = J.$$

- For per-Hermitian and perskew-Hermitian

$$M \text{ is } \textbf{perplectic} \text{ if } M^*RM = R.$$

STRUCTURE-PRESERVING TRANSFORMATIONS

- For Hamiltonian and skew-Hamiltonian

$$M \text{ is } \textbf{symplectic} \text{ if } M^*JM = J.$$

- For per-Hermitian and perskew-Hermitian

$$M \text{ is } \textbf{perplectic} \text{ if } M^*RM = R.$$

manifold	tangent subspace at I	orthogonal subspace at I
symplectic	Hamiltonian	skew-Hamiltonian
perplectic	perskew-Hermitian	per-Hermitian
Lie group	Lie algebra	Jordan algebra

Table: Geometric and algebraic setting for the structured matrices

CANONICAL FORM — HAMILTONIAN

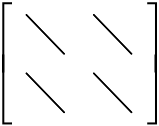
Theorem (BK, Faßbender, Saltenberger)

For every normal Hamiltonian $A \in \mathbb{C}^{2n\times 2n}$ there is unitary symplectic U such that

$$U^*AU = \begin{bmatrix} D_1 & 0 & 0 & 0 \\ 0 & D_2 & 0 & D_3 \\ 0 & 0 & -D_1^* & 0 \\ 0 & -D_3 & 0 & D_2 \end{bmatrix},$$

where D_j, $j = 1, 2, 3$ diagonal matrices,
$D_1 \in \mathbb{C}^{n_1\times n_1}$, $D_2 \in \imath\mathbb{R}^{n_2\times n_2}$, $D_3 \in \mathbb{R}^{n_2\times n_2}$, $n_1 + n_2 = n$.

$$U^*AU = \begin{bmatrix} \Lambda_1 & \Lambda_2 \\ -\Lambda_2 & -\Lambda_1^* \end{bmatrix} =$$



$$=: \Lambda_{\mathcal{H}}$$

CANONICAL FORM — PER-HERMITIAN

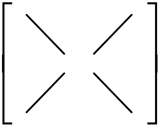
Theorem (BK, Faßbender, Saltenberger)

For every normal per-Hermitian $A \in \mathbb{C}^{2n \times 2n}$ there is unitary perplectic U such that

$$U^* A U = \begin{bmatrix} D_1 & 0 & 0 & 0 \\ 0 & D_2 & D_3 & 0 \\ 0 & RD_3R & RD_2R & 0 \\ 0 & 0 & 0 & RD_1R \end{bmatrix},$$

where D_1 i D_2 are diagonal, and D_3 is antidiagonal matrix, $D_1 \in \mathbb{C}^{n_1 \times n_1}$, $D_2 \in \mathbb{R}^{n_2 \times n_2}$, $D_3 \in \mathbb{R}^{n_2 \times n_2}$, $n_1 + n_2 = n$.

$$U^* A U = \begin{bmatrix} \Lambda_1 & \Lambda_2 \\ R\Lambda_2 R & R\Lambda_1^* R \end{bmatrix} =$$



$$=: \Lambda_{\mathcal{P}}$$

MAXIMIZATION ALGORITHM

$$\max_{ZZ^*=I,\ Z\in Sp_{2n}(\mathbb{C})} \left\{ f_{\mathcal{H}}(Z) := \|\mathrm{diag}(Z^*AZ)\|_F^2 + \|\mathrm{diag}(JZ^*AZ)\|_F^2 \right\}$$

- Iterative algorithm of the form

$$A^{(k+1)} = R_k^* A^{(k)} R_k, \quad k \geq 0.$$

- Transformations R_k are structure-preserving rotations obtained by embedding **two Jacobi rotations**

$$\begin{bmatrix} c & -s \\ s & c \end{bmatrix} := \begin{bmatrix} \cos\phi & -e^{\imath\alpha}\sin\phi \\ e^{-\imath\alpha}\sin\phi & \cos\phi \end{bmatrix} \quad \text{in } I_{2n}.$$

They are chosen to maximize

$$\|\mathrm{diag}(A^{(k+1)})\|_F^2 + \|\mathrm{diag}(JA^{(k+1)})\|_F^2.$$

- D. S. Mackey, N. Mackey, F. Tisseur: *Structured tools for structured matrices*. Electron. J. Linear Al. 10 (2003) 106–145.

MAXIMIZATION ALGORITHM

$$\max_{ZZ^*=I,\ Z\in Pp_{2n}(\mathbb{C})} \left\{ f_{\mathcal{P}}(Z) := \|\text{diag}(Z^*AZ)\|_F^2 + \|\text{diag}(RZ^*AZ)\|_F^2 \right\}$$

- Iterative algorithm of the form

$$A^{(k+1)} = R_k^* A^{(k)} R_k, \quad k \geq 0.$$

- Transformations R_k are structure-preserving rotations obtained by embedding **two Jacobi rotations**

$$\begin{bmatrix} c & -s \\ s & c \end{bmatrix} := \begin{bmatrix} \cos\phi & -e^{\imath\alpha}\sin\phi \\ e^{-\imath\alpha}\sin\phi & \cos\phi \end{bmatrix} \qquad \text{in } I_{2n}.$$

  They are chosen to maximize

$$\|\text{diag}(A^{(k+1)})\|_F^2 + \|\text{diag}(RA^{(k+1)})\|_F^2.$$

- D. S. Mackey, N. Mackey, F. Tisseur: *Structured tools for structured matrices*. Electron. J. Linear Al. 10 (2003) 106–145.

SYMPLECTIC ROTATIONS

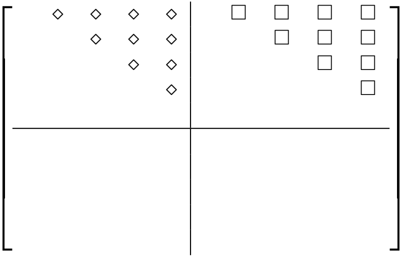
$$R(i,j,\phi,\alpha) = \left[\begin{array}{cc|cc} c & -s & & \\ \bar{s} & c & & \\ \hline & & c & -s \\ & & \bar{s} & c \end{array}\right] \begin{array}{c} i \\ j \\ n+i \\ n+j \end{array}$$

$$R(i,j,\phi,\alpha) = \left[\begin{array}{cc|cc} c & & & -s \\ & c & -\bar{s} & \\ \hline & s & c & \\ \bar{s} & & & c \end{array}\right] \begin{array}{c} i \\ j-n \\ n+i \\ j \end{array}$$

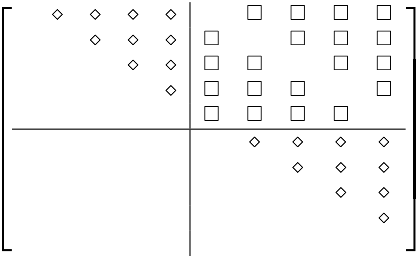
PIVOT POSITIONS (SYMPLECTIC)

$$
\left[\begin{array}{cccc|cccc}
\diamond & \diamond & \diamond & \diamond & \square & \square & \square & \square \\
 & \diamond & \diamond & \diamond & & \square & \square & \square \\
 & & \diamond & \diamond & & & \square & \square \\
 & & & \diamond & & & & \square \\
\hline
 & & & & & & & \\
 & & & & & & & \\
 & & & & & & & \\
 & & & & & & &
\end{array}\right]
$$

PIVOT POSITIONS (SYMPLECTIC)



considering double rotations
$\rightsquigarrow$



PERPLECTIC ROTATIONS

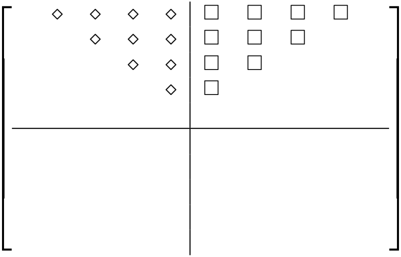
$$R(i,j,\phi,\alpha) = \left[\begin{array}{cc|cc} c & -s & & \\ \bar{s} & c & & \\ \hline & & c & \bar{s} \\ & & -s & c \end{array}\right] \begin{array}{c} i \\ j \\ 2n-j+1 \\ 2n-i+1 \end{array}$$

$$R(i,j,\phi,\alpha) = \left[\begin{array}{cc|cc} c & & -s & \\ & c & & \bar{s} \\ \hline \bar{s} & & c & \\ & -s & & c \end{array}\right] \begin{array}{c} i \\ 2n-j+1 \\ j \\ 2n-i+1 \end{array}$$

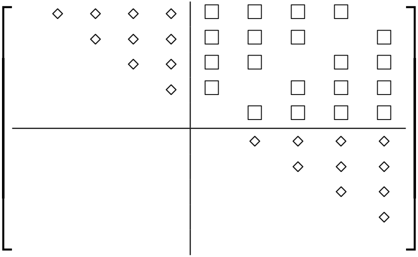
PIVOT POSITIONS (PERPLECTIC)

$$\left[\begin{array}{cccc|cccc}
\diamond & \diamond & \diamond & \diamond & \square & \square & \square & \square \\
 & \diamond & \diamond & \diamond & \square & \square & \square & \\
 & & \diamond & \diamond & \square & \square & & \\
 & & & \diamond & \square & & & \\
\hline
 & & & & & & & \\
 & & & & & & & \\
 & & & & & & & \\
 & & & & & & &
\end{array}\right]$$

PIVOT POSITIONS (PERPLECTIC)



considering double rotations $\rightsquigarrow$



REDUCTION TO CANONICAL FORM

Jacobi algorithm

Input: $A \in \mathbb{C}^{2n\times 2n} \in \mathcal{S}$, $Z_0 = I$
Output: structure-preserving unitary Z
REPEAT
  Select (i_k, j_k).
  Find ϕ_k and α_k.
  Form rotation matrix $R(i_k, j_k, \phi_k, \alpha_k)$.
  $A^{(k+1)} = R_k^* A^{(k)} R_k$
  $Z_{k+1} = Z_k R_k$
UNTIL convergence

REDUCTION TO CANONICAL FORM

Jacobi algorithm

Input: $A \in \mathbb{C}^{2n \times 2n} \in \mathcal{S}$, $Z_0 = I$
Output: structure-preserving unitary Z
REPEAT
  Select (i_k, j_k).
  Find ϕ_k and α_k.
  Form rotation matrix $R(i_k, j_k, \phi_k, \alpha_k)$.
  $A^{(k+1)} = R_k^* A^{(k)} R_k$
  $Z_{k+1} = Z_k R_k$
UNTIL convergence

- Cyclic pivot strategy
- Convergence condition:

$$|\langle \text{grad} f(Z), Z\dot{R}(i_k, j_k, 0, \alpha_k)\rangle| \geq \eta \|\text{grad} f(Z)\|_F,$$

where $\dot{R}(i,j,\phi,\alpha) = \frac{\partial}{\partial \phi} R(i,j,\phi,\alpha)$ and $f = f_{\mathcal{H}}$ or $f = f_{\mathcal{P}}$.

THE CLOSEST NORMAL MATRIX

- Let A be Hamiltonian. Analogy with unstructured case:
  (i) Find Z that maximizes
  $f_{\mathcal{H}}(Z) = \|\mathrm{diag}(Z^*AZ)\|_F^2 + \|\mathrm{diag}(JZ^*AZ)\|_F^2$,
  (ii) Extract the canonical form,
  (iii) Solution is given by $X = Z \begin{bmatrix} \diagdown & \diagdown \\ \diagdown & \diagdown \end{bmatrix} Z^*$.

  $\rightarrow$ But this can produce a matrix that is not normal!

THE CLOSEST NORMAL MATRIX

- Let A be Hamiltonian. Analogy with unstructured case:
  (i) Find Z that maximizes
  $f_{\mathcal{H}}(Z) = \|\text{diag}(Z^*AZ)\|_F^2 + \|\text{diag}(JZ^*AZ)\|_F^2$,
  (ii) Extract the canonical form,
  (iii) Solution is given by $X = Z \begin{bmatrix} \ddots & \ddots \\ \ddots & \ddots \end{bmatrix} Z^*$.

  $\rightarrow$ But this can produce a matrix that is not normal!

- All that can be done is to take

$$\text{diag}(Z^*AZ) = \begin{bmatrix} \Lambda_1 & \\ & -\Lambda_1^* \end{bmatrix} = \begin{bmatrix} \ddots & \\ & \ddots \end{bmatrix} =: \mathcal{D}.$$

THE CLOSEST NORMAL MATRIX–cont.

- We set

$$f_{\mathcal{D}}(Z) = \|\mathrm{diag}(Z^* A Z)\|_F^2.$$

(i) Find Z that maximizes $f_{\mathcal{D}}$.
(ii) Extract the diagonal.
(iii) Solution is given by $X = Z \begin{bmatrix} \diagdown & \\ & \diagdown \end{bmatrix} Z^*$.

- To find Z that maximizes $f_{\mathcal{D}}$ we add new rotations to the Jacobi algorithm.

ADDITIONAL ROTATIONS

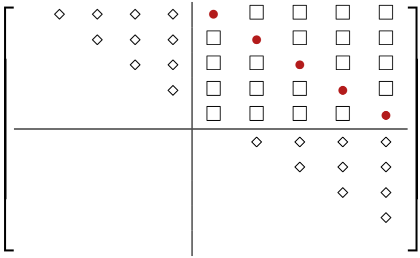
- Symplectic rotations

$$R(i, n+i, \phi, 0) = \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} \begin{matrix} i \\ n+i \end{matrix}$$

ADDITIONAL ROTATIONS

- Symplectic rotations

$\rightsquigarrow$



ADDITIONAL ROTATIONS

- Symplectic rotations

$$R(i, n+i, \phi, 0) = \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} \begin{matrix} i \\ n+i \end{matrix}$$

- Perplectic rotations

$$R(i, 2n-i+1, \phi, -\frac{\pi}{2}) = \begin{bmatrix} \cos\phi & \imath\sin\phi \\ \imath\sin\phi & \cos\phi \end{bmatrix} \begin{matrix} i \\ 2n-i+1 \end{matrix}$$

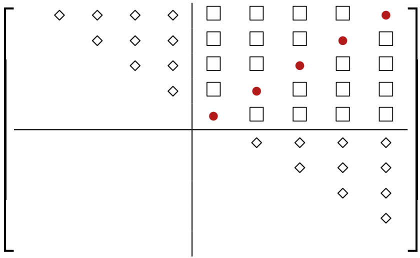
ADDITIONAL ROTATIONS

- Symplectic rotations

$$R(i, n+i, \phi, 0) = \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} \begin{matrix} i \\ n+i \end{matrix}$$

- Perplectic rotations

$\rightsquigarrow$



CONVERGENCE

Theorem (BK, Faßbender, Saltenberger)

Let A be Hamiltonian and let $(Z_k)_k$ be a sequence of unitary symplectic matrices generated by the Jacobi algorithm. Every accumulation point of $(Z_k)_k$ is a stationary point of function $f_{\mathcal{H}}$.

Theorem (BK, Faßbender, Saltenberger)

Let A be Hamiltonian and let $(Z_k)_k$ be a sequence of unitary symplectic matrices generated by the Jacobi algorithm
with additional rotations. Every accumulation point of $(Z_k)_k$ is a stationary point of function $f_{\mathcal{D}}$.

CONVERGENCE

Theorem (BK, Faßbender, Saltenberger)

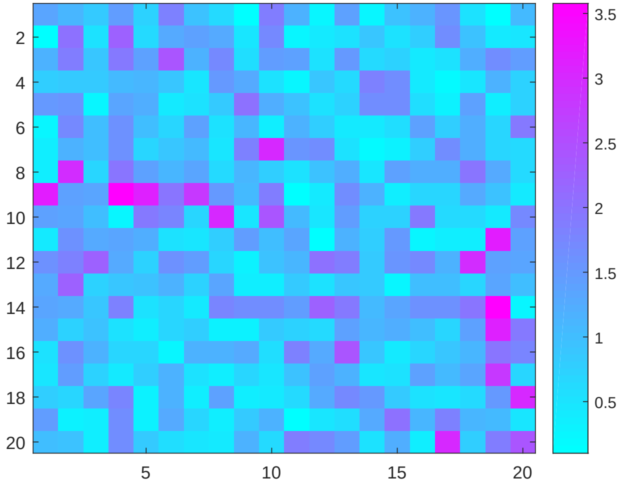
Let A be Hamiltonian and let $(Z_k)_k$ be a sequence of unitary symplectic matrices generated by the Jacobi algorithm. Every accumulation point of $(Z_k)_k$ is a stationary point of function $f_{\mathcal{H}}$.

Theorem (BK, Faßbender, Saltenberger)

Let A be Hamiltonian and let $(Z_k)_k$ be a sequence of unitary symplectic matrices generated by the Jacobi algorithm with additional rotations. Every accumulation point of $(Z_k)_k$ is a stationary point of function $f_{\mathcal{D}}$.

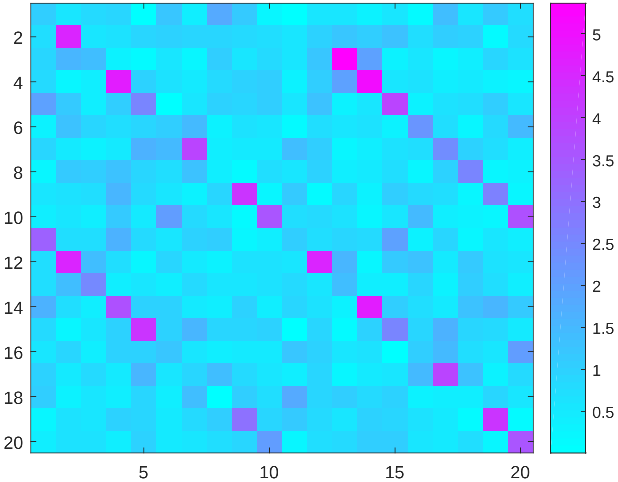
- M. Ishteva, P.-A. Absil, P. Van Dooren: *Jacobi algorithm for the best low multilinear rank approximation of symmetric tensors*. SIAM J. Matrix Anal. Appl. 34(2) (2013) 651–672.
- E. Begović Kovač, D. Kressner: *Structure-preserving low multilinear rank approximation of antisymmetric tensors*. SIAM. J. Matrix Anal. Appl. 38(3) (2017) 967–983.

NUMERICAL EXAMPLES — Canonical form



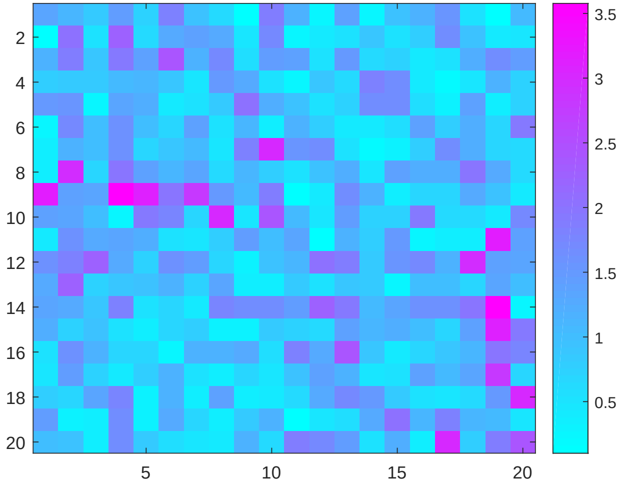
Random Hamiltonian 20×20 matrix.

NUMERICAL EXAMPLES — Canonical form



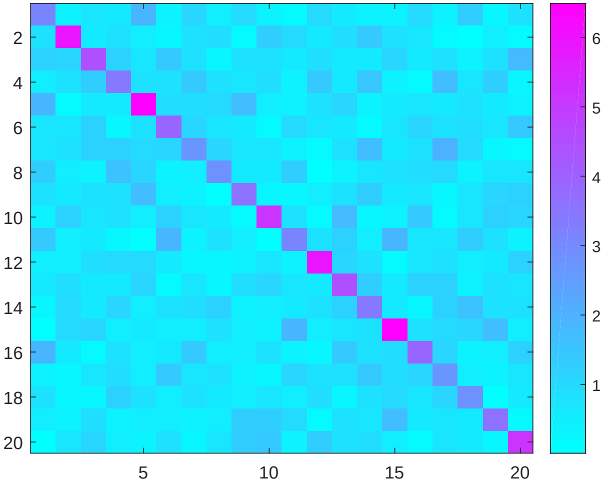
Reduction to the canonical form after 10 cycles.

NUMERICAL EXAMPLES — Diagonalization



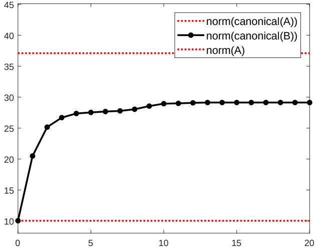
Random Hamiltonian 20×20 matrix.

NUMERICAL EXAMPLES — Diagonalization

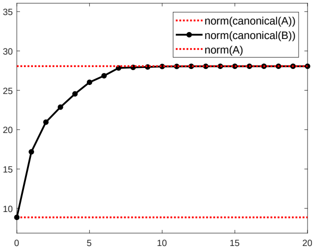


Diagonalization after 10 cycles.

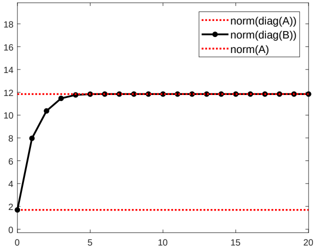
NUMERICAL EXAMPLES — Convergence



(a) Random Hamiltonian 30×30



(b) Normal Hamiltonian 30×30



Normal Hamiltonian 30×30 with no purely imaginary eigenvalue

THANK YOU!