Eigenvalue Condition Numbers: Structured versus Traditional

S. Noschese and L. Pasquini

Università degli Studi di Roma "La Sapienza"
Dipartimento di Matematica "Guido Castelnuovo"

E-mail:
noschese@mat.uniroma1.it,
lionello.pasquini@uniroma1.it

This work has been performed within the activities of the National GNCS-INDAM Research Project: "Algebra Lineare Numerica per matrici con struttura" and supported by the related funds.

OUTLINE

1. Wilkinson Eigenvalue Perturbation Theory
2. Structured Conditioning. Motivations
3. Zero structures and symmetry patterns
4. Structured Conditioning. Definitions
5. Some numerical tests
6. Zero-structured condition number
   - 6.1 Existing theory
   - 6.2 The new definition
   - 6.3 Zero-structured analogue of the Wilkinson perturbation
7. Toeplitz-patterned condition number
   - 7.1 A first formulation
   - 7.2 The new definition
   - 7.3 Toeplitz-patterned analogue of the Wilkinson perturbation
8. Further numerical tests
   - 8.1 Frank matrix case
   - 8.2 Bessel matrix case

References

1. Wilkinson Eigenvalue Perturbation Theory ([14] , [2])

$$A \in \mathbb{C}^{n\times n}, \qquad \lambda \in \sigma(A), \quad \lambda \text{ simple.}$$

We can write ([14] , [2] , [1])

$$(A + \varepsilon\, E)\, x(\varepsilon) = \lambda(\varepsilon)\, x(\varepsilon), \qquad \|E\|_F = 1,$$

$$\lambda(\varepsilon) = \lambda + \frac{y^H E\, x}{y^H x}\, \varepsilon + O(\varepsilon^2),$$

where x and y respectively are the right and left eigenvector associated to λ, with $\|x\|_2 = \|y\|_2 = 1$.

$$\left| \frac{y^H E\, x}{y^H x} \right| \leq \frac{1}{|y^H x|}. \tag{1}$$

The upper-bound in (1) is attained when $E = y\, x^H$.

$W_\lambda := y\, x^H$, **Wilkinson perturbation**,

$\kappa(\lambda) := \dfrac{1}{|y^H x|}$, **traditional condition number**.

2. Structured Conditioning. Motivations

- However, many perturbations E result in slight modifications in λ and some perturbations even annihilate the ratio
$$\frac{y^H E\, x}{y^H x}$$
(see [10] , [9]).
- In addition, very often it does not make any sense to take into consideration any unit norm matrix perturbation E, but one has only to consider perturbations satisfying specific requirements. As an example, this happens when A is a zero structured matrix or a symmetry patterned matrix and the machine perturbations are considered. For instance, if A is a tridiagonal Toeplitz matrix, one could be interested in considering only matrix perturbations E which are tridiagonal Toeplitz matrices themselves.

3. Zero structures and symmetry patterns

Examples of **matrix zero structures** are: diagonal, bidiagonal, tridiagonal, pentadiagonal, banded, triangular, Hessenberg and even sparse. We can treat any kind of zero structured matrices [7] .

Examples of **matrix symmetry patterns** are: symmetric, antisymmetric, persymmetric, circulant, Toeplitz, and Hankel. We considered Toeplitz and nonreal Hankel matrices [8] since they are nonnormal (we focus on Toeplitz matrices in this talk).

4. Structured Conditioning. Definitions

We consider a subspace $\mathcal{S} \subseteq \mathbb{C}^{n\times n}$ of zero structured matrix perturbations and, bearing Wilkinson's formulation in mind, we start with the following definition of the individual $\mathcal{S}-$**structured condition number** $s\kappa(\lambda)$ of a simple eigenvalue $\lambda \in \sigma(A)$

$$s\kappa(\lambda) := \max\left\{\left|\frac{y^H E\, x}{y^H x}\right|, \|E\|_F = 1,\ E \in \mathcal{S}\right\}.$$

Analogously, if, for instance, A is a Toeplitz matrix, we consider a subspace $\mathcal{T} \subseteq \mathbb{C}^{n\times n}$ of Toeplitz matrices and give the following definition of the individual $\mathcal{T}-$**patterned condition number** $\tau\kappa(\lambda)$ of a simple eigenvalue $\lambda \in \sigma(A)$

$$\tau\kappa(\lambda) := \max\left\{\left|\frac{y^H E\, x}{y^H x}\right|, \|E\|_F = 1,\ E \in \mathcal{T}\right\}.$$

• *How to choose subspaces $\mathcal{S}$ and $\mathcal{T}$*

5. Some numerical tests

- A rough heuristic study leads us to
  - n large enough
  - band narrow enough
  - worst conditioned eigenvalues
- An algorithm we defined [8] looks for tridiagonal complex Toeplitz matrices yielding significant gains

n	$\kappa(\lambda)$	$s\kappa(\lambda)$	$\tau\kappa(\lambda)$	$\tau\kappa/s\kappa$
50	1.0503+17	1.0799+01	5.1114-01	4.7331-02
100	1.1755+24	2.6176+02	8.8262-01	3.3718-03
150	4.4678+32	2.4046+05	3.2860+02	1.3665-03
200	4.0869+32	1.7864+05	1.3457+02	7.5330-04
250	4.8907+33	5.5122+06	6.0402+03	1.0958-03
300	1.0592+29	3.0394+05	3.3856+02	1.1139-03
350	4.6276+32	5.8314+07	1.1386+04	1.9526-04
400	1.0705+19	1.9215+02	1.3907-01	7.2374-04
450	1.4637+27	6.0350+05	4.0743+02	6.7511-04
500	2.2154+29	7.5651+07	9.5991+03	1.2689-04

6. Zero-structured condition number

6.1 Existing theory

We use existing theory ([4] , [5] , [12] , [6]) to rewrite the above definition of the $\mathcal{S}-$structured condition number $s\kappa(\lambda)$ which becomes

$$s\kappa(\lambda) = \frac{\left\| y^H (x^T \otimes I)\, B \right\|_F}{|y^H x|}.$$

Here $\otimes$ denotes the Kronecker product and $B = B(\mathcal{S})$ is a suitable matrix which depends on $\mathcal{S}$.

Steps to define B

1. Mark the structure-positions and denote by m their number.

2. Consider the column vector $\xi = \xi(E) \in \mathbb{C}^m,\ E \in \mathcal{S}$, defined by arranging by columns the m entries that are in the m marked positions in $E \in \mathcal{S}$.

3. Consider the column vector $\mathrm{vec}(E) \in \mathbb{C}^{n^2}$, $E \in \mathcal{S}$.

- B is the unique matrix in $\mathbb{C}^{n^2 \times m}$ such that: $\mathrm{vec}(E) = B\,\xi$, for each $E \in \mathcal{S}$.

6.2 The new definition

▼ The components of the n^2-length row vector $y^H\,(x^T\otimes I)$ are the entries of $\overline{W}_\lambda$ arranged by columns: $(\overline{y}_i\,x_j,\ i=1:n),\ j=1:n$.

▼ Post-multiplying $y^H\,(x^T\otimes I)$ by B results in deleting the components $\overline{y}_i\,x_j$ of $y^H\,(x^T\otimes I)$ whose indices locate positions out of the structure of the matrices in $\mathcal{S}$.

▼ We express the norm of the $m-$length row vector $y^H\,(x^T\otimes I)\,B$ in a form which is more significant and cheaper. This form is as follows

$$\left\|y^H\,(x^T\otimes I)\,B\right\|_F=\left\|\mathrm{vec}(\overline{W}_\lambda\mid_{\mathcal{S}})\right\|_F,$$

where $\overline{W}_\lambda\mid_{\mathcal{S}}$ denotes the restriction of $\overline{W}_\lambda$ to $\mathcal{S}$.

▼ Then we observe that

$$\left\|y^H\,(x^T\otimes I)\,B\right\|_F=\left\|\mathrm{vec}(\overline{W}_\lambda\mid_{\mathcal{S}})\right\|_F=$$
$$=\left\|\overline{W}_\lambda\mid_{\mathcal{S}}\right\|_F=\left\|W_\lambda\mid_{\mathcal{S}}\right\|_F.$$

► As a consequence [7] :

$\mathcal{S}$-structured condition number

$$s\kappa(\lambda)=\frac{\left\|W_\lambda\mid_{\mathcal{S}}\right\|_F}{|y^Hx|}.$$

6.3 Zero-structured analogue of the Wilkinson perturbation

▼ The role played by $W_\lambda\,|_{\mathcal{S}}$ is fully highlighted by noting that

$$y^H\,W_\lambda\,|_{\mathcal{S}}\;x=\sum_{i=1}^{n}\overline{y}_i\sum_{j=1}^{n}(W_\lambda\,|_{\mathcal{S}})_{ij}\,x_j=$$

$$=\sum_{\mathcal{S}}\overline{y}_i\,y_i\,\overline{x}_j\,x_j=\sum_{\mathcal{S}}|y_i|^2\,|x_j|^2=$$

$$=\sum_{\mathcal{S}}|y_i\,x_j|^2=\|W_\lambda\,|_{\mathcal{S}}\|_F^2\,.$$

▼ Consequently

$$y^H\,\frac{W_\lambda\,|_{\mathcal{S}}}{\|W_\lambda\,|_{\mathcal{S}}\|_F}\,x=\|W_\lambda\,|_{\mathcal{S}}\|_F\,,$$

so that we can conclude that one of the "worst" matrices E is $W_\lambda\,|_{\mathcal{S}}/\,\|W_\lambda\,|_{\mathcal{S}}\|_F$.

► This leads to define [7] :

$\mathcal{S}$-structured analogue of the Wilkinson perturbation W_λ

$$\frac{W_\lambda\,|_{\mathcal{S}}}{\|W_\lambda\,|_{\mathcal{S}}\|_F}.$$

7. Toeplitz-patterned condition number

7.1 A first formulation

As before, existing results lead us to write:

$$\tau\kappa(\lambda) = \frac{\left\| y^H (x^T \otimes I)\, B\, D^{-1} \right\|_F}{|y^H x|}.$$

B and D are suitable matrices depending on $\mathcal{T}$

Steps to define B

1. Mark the structure-diagonals and denote by k their number.

2. From the $i-$th, $i = 1 : k$, marked diagonal of E, $E \in \mathcal{T}$, take its common entry ξ_i forming a column vector $\xi = \xi(E) \in \mathbb{C}^k$.

• B is the unique matrix in $\mathbb{C}^{n^2 \times k}$ such that: $\mathrm{vec}(E) = B\,\xi$, for each $E \in \mathcal{T}$.

Definition of D

Its diagonal entry d_i, $i = 1 : k$, is the square root of the length of the $i-$th marked diagonal.

• D is the unique diagonal matrix in $\mathbb{C}^{k \times k}$ such that $\|E\|_F = \|D\,\xi\|_2$, for each $E \in \mathcal{T}$.

7.2 The new definition

▼ We proved that $y^H(x^T \otimes I) = \left(\mathrm{vec}(\overline{W}_\lambda)\right)^T$.

▼ Post-multiplying $y^H(x^T \otimes I)$ by B results in a row vector of length k. Its $i-$th component, $i = 1 : k$, is the sum of the $\overline{y}_p\, x_q$'s whose indices p, q locate the entries in the $i-$th marked diagonal in $\overline{W}_\lambda$.

▼ Thus, $y^H(x^T \otimes I)\, B\, D^{-1}$ is a row vector of length k whose $i-$th component, $i = 1 : k$, is that of $y^H(x^T \otimes I)\, B$ divided by the square root of the length of the $i-$th marked diagonal.

▼ We express the norm of the $y^H(x^T \otimes I)\, B\, D^{-1}$ in a more significant and cheaper form:

$$\left\| y^H(x^T \otimes I)\, B\, D^{-1} \right\|_F = \left\| \overline{W_\lambda \downarrow_{\mathcal{T}}} \right\|_F ,$$

where $W_\lambda \downarrow_{\mathcal{T}}$ denotes the matrix obtained from $W_\lambda \mid_{\mathcal{S}}$ by replacing all the entries in each diagonal with their mean.

▶ As a consequence [8] :

$\mathcal{T}$-patterned condition number

$$\tau\kappa(\lambda) = \frac{\|W_\lambda \downarrow_{\mathcal{T}}\|_F}{|y^H x|}.$$

7.3 Toeplitz-patterned analogue of the Wilkinson perturbation

▼ The role played by $W_\lambda \downarrow_{\mathcal{T}}$ is analogous to that of $W_\lambda \mid_{\mathcal{S}}$. In fact, an argument very similar to the one used in the zero-structured case but quite complicate can be used to prove the equality

$$y^H W_\lambda \downarrow_{\mathcal{T}} x = \|W_\lambda \downarrow_{\mathcal{T}}\|_F^2 ,$$

which is perfectly analogous to the one obtained above.

▼ Consequently

$$y^H \frac{W_\lambda \downarrow_{\mathcal{T}}}{\|W_\lambda \downarrow_{\mathcal{T}}\|_F} x = \|W_\lambda \downarrow_{\mathcal{T}}\|_F ,$$

so that we can conclude that one of the "worst" matrices E is just $W_\lambda \downarrow_{\mathcal{T}} / \|W_\lambda \downarrow_{\mathcal{T}}\|_F$.

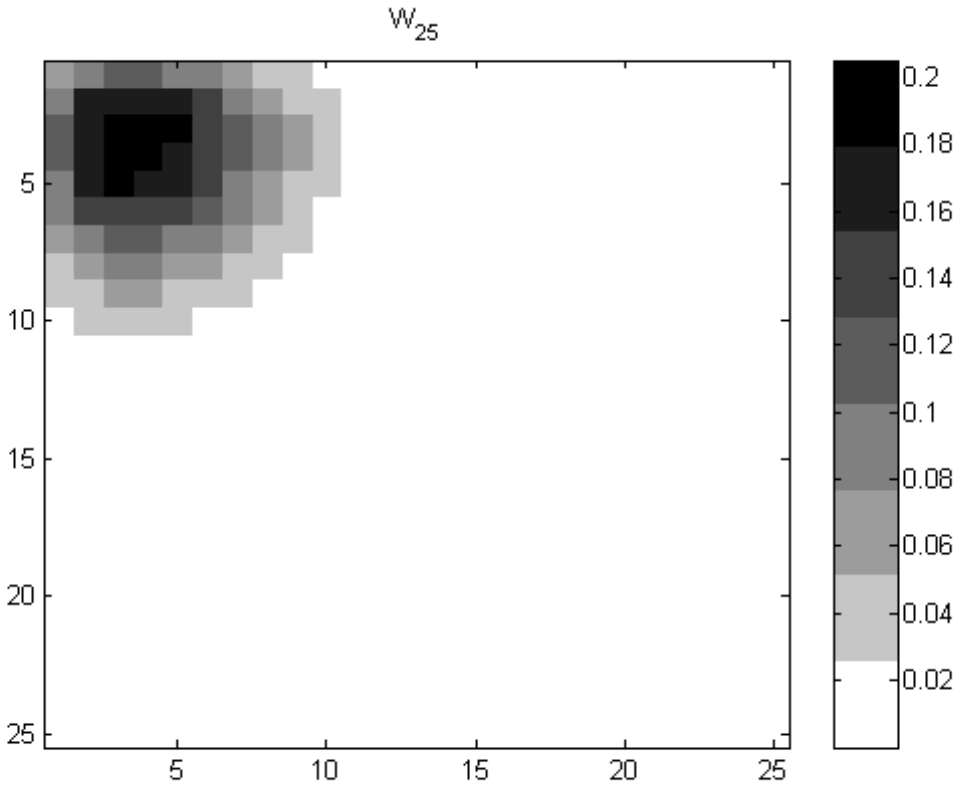
▶ This leads to define [8] :

$\mathcal{T}$-patterned analogue of the Wilkinson perturbation W_λ

$$\frac{W_\lambda \downarrow_{\mathcal{T}}}{\|W_\lambda \downarrow_{\mathcal{T}}\|_F}.$$

8. Further numerical tests

8.1 Bessel matrix case ([3] , [11])

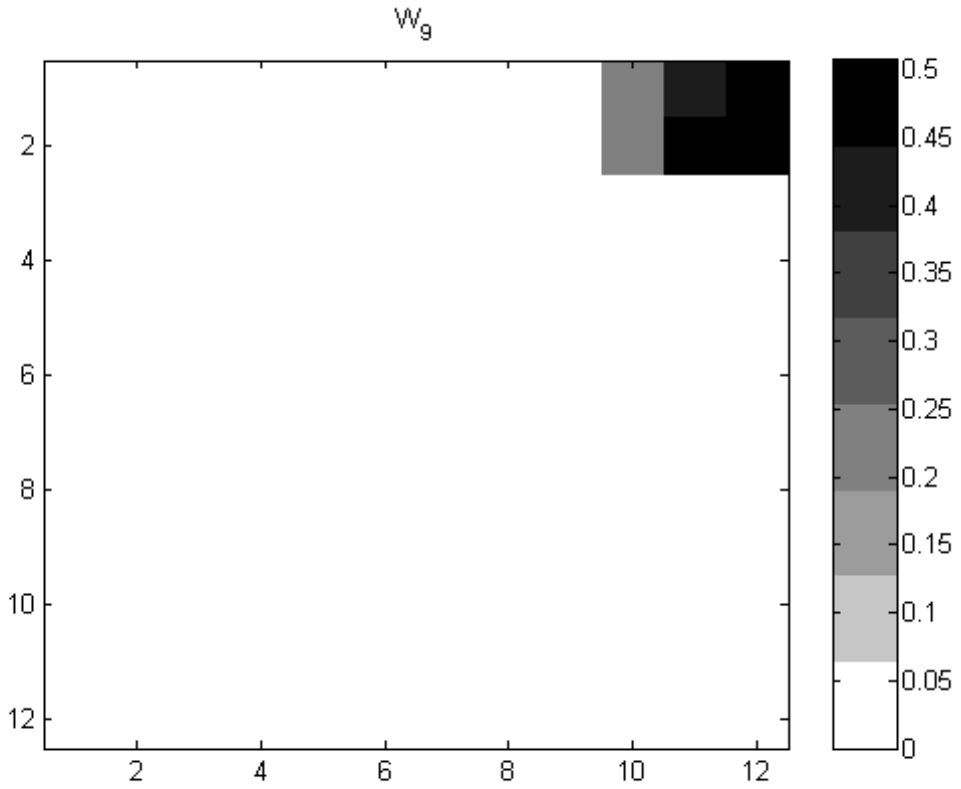


λ	$k(\lambda)$	$sk(\lambda)$
25	3.9408e+12	2.0839e+12

λ	$k(\lambda)$	$sk(\lambda)$
25	3.9408e+12	2.0800e+12

λ	$k(\lambda)$	$sk(\lambda)$
25	3.9408e+12	1.5999e+05

8.2 Frank matrix case ([13] , [2] , [15])



λ	$k(\lambda)$	$sk(\lambda)$
8	1.8283e+7	1.8283e+7
9	3.8774e+7	3.8774e+7
10	2.6646e+7	2.6646e+7

λ	$k(\lambda)$	$sk(\lambda)$
8	1.8283e+7	6.6138e+0
9	3.8774e+7	3.6994e+0
10	2.6646e+7	2.6031e+0

References

[1] G. H. Golub and C. F. Van Loan. *Matrix Computations*. The Johns Hopkins University Press, Baltimore and London, third edition, 1996.

[2] G. H. Golub and J. H. Wilkinson. Ill–conditioned eigensystems and the computation of the Jordan canonical form. *SIAM Review*, 18:578–619, 1976.

[3] E. Grosswald. *Bessel Polynomials*, volume 698 of *Lecture Notes in Mathematics*. Springer–Verlag, New York, 1978.

[4] D. J. Higham and N. J. Higham. Backward Error and Condition of Structured Linear Systems. *SIAM J. Matrix Anal. Appl.*, 13(1):162–175, 1992.

[5] D. J. Higham and N. J. Higham. Structured Backward Error and Condition of Generalized Eigenvalue Problems. *SIAM J. Matrix Anal. Appl.*, 20(2):493–512, 1998.

[6] P. Lancaster and M. Tismenetsky. *The Theory of Matrices*. Academic Press, New York, second edition, 1985.

[7] S. Noschese and L. Pasquini. Eigenvalue Condition Numbers: Zero-Structured versus Traditional. *Submitted*.

[8] S. Noschese and L. Pasquini. Eigenvalue patterned condition numbers: Toeplitz and Hankel cases. *To be submitted*.

[9] S. Noschese and L. Pasquini. Two algorithms to construct matrix modifications keeping essentially unaltered sets of eigenvalues and applications. *Submitted*.

[10] S. Noschese and L. Pasquini. How to Find Matrix Modifications Keeping Essentially Unaltered a Selected Set of Eigenvalues. In *Linear Algebra Proceedings*, SIAM Conference on Applied Linear Algebra, Williamsburg, VA, July, 15-19 2003. CP5 in http://www.siam.org/meetings/la03/proceedings/.

[11] L. Pasquini. Accurate Computation of the Zeros of the Generalized Bessel Polynomials. *Numer. Math.*, 86(3):507–538, 2000.

[12] F. Tisseur. A Chart of Backward Errors for Singly and Doubly Structured Eigenvalue Prob-

lems. *SIAM J. Matrix Anal. Appl.*, 24(3):877–897, 2003.

[13] J. H. Wilkinson. Rounding errors in algebraic processes. *Her Majesty's Stationery Office*, 1963.

[14] J. H. Wilkinson. *The algebraic eigenvalue problem*. Clarendon Press, Oxford, England, 1965.

[15] J. H. Wilkinson. Sensitivity of eigenvalues II. *Utilitas Mathematica*, 30:243–286, 1986.