Structured condition numbers for multiple eigenvalues

Julio Moro

Departamento de Matemáticas
Universidad Carlos III de Madrid
(Spain)

Joint work with

Daniel Kressner & **María J. Peláez**

Structured Numerical Linear Algebra Problems:
Analysis, Algorithms and Applications

Cortona
September 15 – 19, 2008

An example to begin with

Consider the complex skew-symmetric matrix

$$A = \begin{bmatrix} 0 & 1 & 0 & 1 \\ -1 & 0 & -i & 0 \\ 0 & i & 0 & i \\ -1 & 0 & -i & 0 \end{bmatrix},$$

which has one single eigenvalue $\lambda_0 = 0$ and Jordan form

$$J_3(0) \oplus J_1(0).$$

An example to begin with

Consider the complex skew-symmetric matrix

$$A = \begin{bmatrix} 0 & 1 & 0 & 1 \\ -1 & 0 & -i & 0 \\ 0 & i & 0 & i \\ -1 & 0 & -i & 0 \end{bmatrix},$$

which has one single eigenvalue $\lambda_0 = 0$ and Jordan form

$$J_3(0) \oplus J_1(0).$$

- If A is subject to a **small** perturbation

$$A(\varepsilon) = A + \varepsilon E, \qquad \varepsilon << 1,$$

with E an arbitrary 4×4 complex matrix, then $A(\varepsilon)$ has generically three eigenvalues of order $O(\varepsilon^{1/3})$, and one of order $O(\varepsilon)$.

An example to begin with

Consider the complex skew-symmetric matrix

$$A = \begin{bmatrix} 0 & 1 & 0 & 1 \\ -1 & 0 & -i & 0 \\ 0 & i & 0 & i \\ -1 & 0 & -i & 0 \end{bmatrix},$$

which has one single eigenvalue $\lambda_0 = 0$ and Jordan form

$$J_3(0) \oplus J_1(0).$$

- If A is subject to a **small** perturbation

  $$A(\varepsilon) = A + \varepsilon E, \qquad \varepsilon << 1,$$

  with E an arbitrary 4×4 complex matrix, then $A(\varepsilon)$ has generically three eigenvalues of order $O(\varepsilon^{1/3})$, and one of order $O(\varepsilon)$.

- However, if E is restricted to be complex skew-symmetric, then $A(\varepsilon)$ has generically **four eigenvalues of order** $O(\varepsilon^{1/2})$.

Consider the complex skew-symmetric matrix

$$A = \begin{bmatrix} 0 & 1 & 0 & 1 \\ -1 & 0 & -i & 0 \\ 0 & i & 0 & i \\ -1 & 0 & -i & 0 \end{bmatrix},$$

which has one single eigenvalue $\lambda_0 = 0$ and Jordan form

$$J_3(0) \oplus J_1(0).$$

- If A is subject to a **small** perturbation

  $$A(\varepsilon) = A + \varepsilon E, \qquad \varepsilon << 1,$$

  with E an arbitrary 4×4 complex matrix, then $A(\varepsilon)$ has generically three eigenvalues of order $O(\varepsilon^{1/3})$, and one of order $O(\varepsilon)$.

- However, if E is restricted to be complex skew-symmetric, then $A(\varepsilon)$ has generically **four eigenvalues of order** $O(\varepsilon^{1/2})$.

Sensitivity under structured perturbation **qualitatively** different than under unstructured perturbation

Outline

1) Hölder condition numbers (structured & unstructured) for multiple eigenvalues

2) Comparing structured and unstructured condition numbers for

- **2.1)** **generic** structured perturbations
- **2.2)** **nongeneric** structured perturbations

3) Concluding remarks

- Let λ_0 be a simple eigenvalue of $A \in \mathbb{C}^{n\times n}$, $\|\cdot\|$ matrix 2-norm.

Definition

The condition number of λ_0 is

$$\kappa(\lambda_0) = \lim_{\varepsilon \to 0} \sup \left\{ \frac{|\Delta\lambda|}{\varepsilon} : E \in \mathbb{C}^{n\times n}, \|E\| \leq 1, \lambda_0 + \Delta\lambda \in \mathrm{sp}(A + \varepsilon E) \right\}$$

Condition numbers for simple eigenvalues

- Let λ_0 be a simple eigenvalue of $A \in \mathbb{C}^{n\times n}$, $\|\cdot\|$ matrix 2-norm.

Definition

The condition number of λ_0 is

$$\kappa(\lambda_0) = \lim_{\varepsilon\to 0} \sup \left\{ \frac{|\Delta\lambda|}{\varepsilon} : E \in \mathbb{C}^{n\times n}, \|E\| \le 1, \lambda_0 + \Delta\lambda \in \mathrm{sp}(A+\varepsilon E) \right\}$$

- If x (resp. y) right (resp. left) e-vector corresp. to λ_0 with $y^H x = 1$, then

$$\kappa(\lambda_0) = \|x\|\,\|y\|$$

Condition numbers for simple eigenvalues

- Let λ_0 be a simple eigenvalue of $A \in \mathbb{C}^{n\times n}$, $\|\cdot\|$ matrix 2-norm.

Definition

The condition number of λ_0 is

$$\kappa(\lambda_0) = \lim_{\varepsilon\to 0} \sup \left\{ \frac{|\Delta\lambda|}{\varepsilon} \; : \; E \in \mathbb{C}^{n\times n}, \|E\| \leq 1\,, \lambda_0 + \Delta\lambda \in \mathrm{sp}(A+\varepsilon E) \right\}$$

- If x (resp. y) right (resp. left) e-vector corresp. to λ_0 with $y^H x = 1$, then

$$\kappa(\lambda_0) = \|x\|\,\|y\|$$

But: if λ_0 is defective, then generically

$$\frac{\Delta\lambda}{\varepsilon} \to \infty \quad \text{as} \quad \varepsilon \to 0 \quad \Longrightarrow \quad \text{Need} \neq \text{definition for } \kappa(\lambda_0)$$

If the eigenproblem has some special structure (symmetric, skew-symmetric, Toeplitz, Hankel, zero patterns, symplectic, Hamiltonian,...)

- look for numerical algorithms that preserve the structure and spectral properties of the problem
- may lead to significantly faster and/or more accurate solutions,

If the eigenproblem has some special structure (symmetric, skew-symmetric, Toeplitz, Hankel, zero patterns, symplectic, Hamiltonian,...)

- look for numerical algorithms that preserve the structure and spectral properties of the problem
- may lead to significantly faster and/or more accurate solutions,

Hence, if $A \in \mathbb{S}$, class of structured matrices, $\Longrightarrow$ measure sensitivity with respect to perturbations $E \in \mathbb{S}$.

Leads to structured condition numbers.

$$\kappa(\lambda_0, \mathbb{S}) = \lim_{\varepsilon \to 0} \sup \left\{ \frac{|\Delta\lambda|}{\varepsilon} : E \in \mathbb{S}, \|E\| \leq 1, \lambda_0 + \Delta\lambda \in \mathrm{sp}(A + \varepsilon E) \right\}.$$

If the eigenproblem has some special structure (symmetric, skew-symmetric, Toeplitz, Hankel, zero patterns, symplectic, Hamiltonian,...)

- look for numerical algorithms that preserve the structure and spectral properties of the problem
- may lead to significantly faster and/or more accurate solutions,

Hence, if $A \in \mathbb{S}$, class of structured matrices, $\Longrightarrow$ measure sensitivity with respect to perturbations $E \in \mathbb{S}$.

Leads to structured condition numbers.

$$\kappa(\lambda_0, \mathbb{S}) = \lim_{\varepsilon \to 0} \sup \left\{ \frac{|\Delta\lambda|}{\varepsilon} : E \in \mathbb{S}, \|E\| \leq 1 \,, \lambda_0 + \Delta\lambda \in \mathrm{sp}(A + \varepsilon E) \right\}.$$

Question:

Is $\kappa(\lambda_0, \mathbb{S})$ much smaller than $\kappa(\lambda_0)$?

Many relevant contributions on structured condition numbers of simple eigenvalues

[Tisseur '03]

[Byers & Kressner '03]

[Noschese & Pasquini '06, '07]

[Karow, Kressner & Tisseur '06]

[Tisseur & Graillat '06]

[Bora '06]

...

as well as for structured pseudospectra

[Rump '06]

[Karow '06]

...

Many relevant contributions on structured condition numbers of simple eigenvalues

[Tisseur '03]

[Byers & Kressner '03]

[Noschese & Pasquini '06, '07]

[Karow, Kressner & Tisseur '06]

[Tisseur & Graillat '06]

[Bora '06]

...

as well as for structured pseudospectra

[Rump '06]

[Karow '06]

...

How about condition numbers for multiple, defective eigenvalues?

- Let λ_0 be a **multiple** e-value of $A \in \mathbb{C}^{n \times n}$, and

  let $n_1 \equiv$ size of largest Jordan block corresp. to λ_0.

  Then, the worst-case behaviour of λ_0 under small perturbations $A + \varepsilon E$ corresponds to

$$\widehat{\lambda}(\varepsilon) = \lambda_0 + (\xi_k)^{1/n_1} \varepsilon^{1/n_1} + o(\varepsilon^{1/n_1}),$$

  where ξ_k are the eigenvalues of a product $Y^H E X$

[Lidskii '66]

First order perturbation theory

- Let λ_0 be a **multiple** e-value of $A \in \mathbb{C}^{n\times n}$, and
  let $n_1 \equiv$ size of largest Jordan block corresp. to λ_0.
  Then, the worst-case behaviour of λ_0 under small perturbations $A+\varepsilon E$ corresponds to

$$\widehat{\lambda}(\varepsilon) = \lambda_0 + (\xi_k)^{1/n_1}\varepsilon^{1/n_1} + o(\varepsilon^{1/n_1}),$$

  where ξ_k are the eigenvalues of a product $Y^H EX$

  [Lidskii '66]

- Let

$$P^{-1}AP = \begin{bmatrix} J_0 & 0 \\ 0 & * \end{bmatrix}$$

  be a Jordan form of A, where J_0 gathers all r_1 Jordan blocks of size n_1 corresp. to λ_0. Then

$$Y^H E X \in \mathbb{C}^{r_1\times r_1}, \qquad \text{where}$$

- X contains those columns of P which are right e-vectors of A corresp. to blocks of largest size n_1 in J_0.
- Y^H contains those rows of P^{-1} which are left e-vectors of A corresp. to blocks of largest size n_1 in J_0.

Lidskii's theory shows that worst-case behavior is

$$|\Delta\lambda| = |\lambda(\varepsilon) - \lambda_0| \leq |\xi_k|^{1/n_1} \varepsilon^{1/n_1} + \dots$$

Lidskii's theory shows that worst-case behavior is

$$|\Delta\lambda| = |\lambda(\varepsilon) - \lambda_0| \leq |\xi_k|^{1/n_1} \varepsilon^{1/n_1} + \ldots \leq \rho(Y^H E X)^{1/n_1} \varepsilon^{1/n_1} + \ldots$$

Hölder condition numbers for multiple eigenvalues

Lidskii's theory shows that worst-case behavior is

$$|\Delta\lambda| = |\lambda(\varepsilon) - \lambda_0| \leq |\xi_k|^{1/n_1}\varepsilon^{1/n_1} + \ldots \leq \rho(Y^H EX)^{1/n_1}\varepsilon^{1/n_1} + \ldots$$

Definition [M., Burke & Overton '97]

λ_0 multiple e-value of $A \in \mathbb{C}^{n\times n}$,

$n_1 \equiv$ largest size of Jordan blocks J_{λ_0} in Jordan form of A

$Y \equiv$ left e-vectors taken from all $n_1 \times n_1$ Jordan blocks J_{λ_0}.

$X \equiv$ right e-vectors taken from all $n_1 \times n_1$ Jordan blocks J_{λ_0}.

The **Hölder condition number** of λ_0 is the pair $\boxed{\kappa(\lambda_0) = (n_1, \alpha),}$ where

$$\alpha = \sup_{\substack{E\in\mathbb{C}^{n\times n}\\ ||E||\leq 1}} \rho(Y^H EX), \qquad \rho(\cdot) \equiv \text{ spectral radius}$$

Lidskii's theory shows that worst-case behavior is

$$|\Delta\lambda| = |\lambda(\varepsilon) - \lambda_0| \leq |\xi_k|^{1/n_1}\varepsilon^{1/n_1} + \ldots \leq \rho(Y^H E X)^{1/n_1}\varepsilon^{1/n_1} + \ldots$$

Definition [M., Burke & Overton '97]

λ_0 multiple e-value of $A \in \mathbb{C}^{n\times n}$,

$n_1 \equiv$ largest size of Jordan blocks J_{λ_0} in Jordan form of A

$Y \equiv$ left e-vectors taken from all $n_1 \times n_1$ Jordan blocks J_{λ_0}.

$X \equiv$ right e-vectors taken from all $n_1 \times n_1$ Jordan blocks J_{λ_0}.

The **Hölder condition number** of λ_0 is the pair $\boxed{\kappa(\lambda_0) = (n_1, \alpha),}$ where

$$\alpha = \sup_{\substack{E\in\mathbb{C}^{n\times n}\\ \|E\|\leq 1}} \rho(Y^H E X), \qquad \rho(\cdot) \equiv \text{ spectral radius}$$

One can prove that for any unitarily invariant matrix norm,

$$\alpha = \|XY^H\|_2$$

Again, if $\lambda_0 \in \mathrm{sp}(A)$ for $A \in \mathbb{S}$, class of structured matrices, define structured Hölder condition number as a pair

$$\kappa(\lambda_0, \mathbb{S}) = (n_{\mathbb{S}}, \alpha_{\mathbb{S}}),$$

where

Again, if $\lambda_0 \in \mathrm{sp}(A)$ for $A \in \mathbb{S}$, class of structured matrices, define structured Hölder condition number as a pair

$$\kappa(\lambda_0, \mathbb{S}) = (n_{\mathbb{S}}, \alpha_{\mathbb{S}}),$$

where

- $n_{\mathbb{S}} \equiv$ reciprocal of smallest possible leading exponent in asymptotic expansions of $\widehat{\lambda}(\varepsilon) - \lambda_0$ among all $E \in \mathbb{S}$.

Again, if $\lambda_0 \in \mathrm{sp}(A)$ for $A \in \mathbb{S}$, class of structured matrices, define structured Hölder condition number as a pair

$$\kappa(\lambda_0, \mathbb{S}) = (n_{\mathbb{S}}, \alpha_{\mathbb{S}}),$$

where

- $n_{\mathbb{S}} \equiv$ reciprocal of smallest possible leading exponent in asymptotic expansions of $\widehat{\lambda}(\varepsilon) - \lambda_0$ among all $E \in \mathbb{S}$.
- $\alpha_{\mathbb{S}} \equiv$ maximal value for leading coefficient of asymptotic expansions among all $E \in \mathbb{S}$ giving rise to $O(\varepsilon^{1/n_{\mathbb{S}}})$ expansions.

Again, if $\lambda_0 \in \mathrm{sp}(A)$ for $A \in \mathbb{S}$, class of structured matrices, define structured Hölder condition number as a pair

$$\kappa(\lambda_0, \mathbb{S}) = (n_{\mathbb{S}}, \alpha_{\mathbb{S}}),$$

where

- $n_{\mathbb{S}} \equiv$ reciprocal of smallest possible leading exponent in asymptotic expansions of $\widehat{\lambda}(\varepsilon) - \lambda_0$ among all $E \in \mathbb{S}$.
- $\alpha_{\mathbb{S}} \equiv$ maximal value for leading coefficient of asymptotic expansions among all $E \in \mathbb{S}$ giving rise to $O(\varepsilon^{1/n_{\mathbb{S}}})$ expansions.

Our goal: Determine and compare the **structured** and **unstructured** condition numbers of defective e-values for particular classes $\mathbb{S}$ of matrices, e.g.,

complex symmetric, skew-symmetric, persymmetric, skew-Hermitian, Toeplitz, symmetric Toeplitz, Hankel, zero-structured, Hamiltonian, skew-Hamiltonian, symplectic,...

Again, if $\lambda_0 \in \mathrm{sp}(A)$ for $A \in \mathbb{S}$, class of structured matrices, define structured Hölder condition number as a pair

$$\kappa(\lambda_0, \mathbb{S}) = (n_{\mathbb{S}}, \alpha_{\mathbb{S}}),$$

where

- $n_{\mathbb{S}} \equiv$ reciprocal of smallest possible leading exponent in asymptotic expansions of $\widehat{\lambda}(\varepsilon) - \lambda_0$ among all $E \in \mathbb{S}$.
- $\alpha_{\mathbb{S}} \equiv$ maximal value for leading coefficient of asymptotic expansions among all $E \in \mathbb{S}$ giving rise to $O(\varepsilon^{1/n_{\mathbb{S}}})$ expansions.

Important:

$n_{\mathbb{S}}$ may be strictly smaller than n_1,

e.g. as in the initial 4×4 complex skew-symmetric example,

where $n_1 = 3$, $n_{\mathbb{S}} = 2$, and

$\mathbb{S} \equiv$ skew-symmetric matrices

$$\kappa(\lambda_0) = (n_1, \alpha), \qquad \kappa(\lambda_0, \mathbb{S}) = (n_{\mathbb{S}}, \alpha_{\mathbb{S}}).$$

First, consider the generic situation when

$$n_{\mathbb{S}} = n_1,$$

i.e., there is some $E \in \mathbb{S}$ giving rise to a perturbation expansion of order $O(\varepsilon^{1/n_1})$.

$$\kappa(\lambda_0) = (n_1, \alpha), \qquad \kappa(\lambda_0, \mathbb{S}) = (n_{\mathbb{S}}, \alpha_{\mathbb{S}}).$$

First, consider the generic situation when

$$n_{\mathbb{S}} = n_1,$$

i.e., there is some $E \in \mathbb{S}$ giving rise to a perturbation expansion of order $O(\varepsilon^{1/n_1})$. In that case,

$$\alpha = \sup_{\substack{E \in \mathbb{C}^{n\times n} \\ \|E\|_2 \le 1}} \rho(Y^H E X) = \|XY^H\|_2, \qquad \alpha_{\mathbb{S}} = \sup_{\substack{E \in \mathbb{S} \\ \|E\|_2 \le 1}} \rho(Y^H E X)$$

and we want to know whether $\boxed{\alpha_{\mathbb{S}} << \alpha}$ or not.

$$\kappa(\lambda_0) = (n_1, \alpha), \qquad \kappa(\lambda_0, \mathbb{S}) = (n_{\mathbb{S}}, \alpha_{\mathbb{S}}).$$

First, consider the generic situation when

$$n_{\mathbb{S}} = n_1,$$

i.e., there is some $E \in \mathbb{S}$ giving rise to a perturbation expansion of order $O(\varepsilon^{1/n_1})$. In that case,

$$\alpha = \sup_{\substack{E \in \mathbb{C}^{n\times n} \\ \|E\|_2 \le 1}} \rho(Y^H E X) = \|XY^H\|_2, \qquad \alpha_{\mathbb{S}} = \sup_{\substack{E \in \mathbb{S} \\ \|E\|_2 \le 1}} \rho(Y^H E X)$$

and we want to know whether $\boxed{\alpha_{\mathbb{S}} << \alpha}$ or not.

Usually, look first for some $E_{\mathbb{S}} \in \mathbb{S}$ such that

$$\rho(Y^H E_{\mathbb{S}} X) \approx \|XY^H\|.$$

Then,

$$\kappa(\lambda_0, \mathbb{S}) \approx \kappa(\lambda_0).$$

$\boxed{\text{Structured Jordan form for } A \in \mathbb{S}.}$

(see [Thompson'91], [Mehl'06], ...)

$\Downarrow \quad \Downarrow \quad \Downarrow$

$\boxed{\text{Induced structure in } XY^H.}$

$\Downarrow \quad \Downarrow \quad \Downarrow$

$\boxed{\text{Mapping theorems: } E_0 u = \beta v \ \text{ with } \ |\beta| = 1, \ \|E_0\|_2 \approx 1, \ E_0 \in \mathbb{S}.}$

(see [Rump'03], [Mackey,Mackey&Tisseur'06], ...)

$\Downarrow \quad \Downarrow \quad \Downarrow$

$$\boxed{\kappa(\lambda_0) \quad \approx \quad \kappa(\lambda_0, \mathbb{S})}$$

$\boxed{\text{Structured Jordan form for } A \in \mathbb{S}.}$

(see [Thompson'91], [Mehl'06], ...)

$\Downarrow \quad \Downarrow \quad \Downarrow$

$\boxed{\text{Induced structure in } XY^H.}$

$\Downarrow \quad \Downarrow \quad \Downarrow$

$\boxed{\text{Mapping theorems: } E_0 u = \beta v \ \text{ with } \ |\beta| = 1, \ \|E_0\|_2 \approx 1, \ E_0 \in \mathbb{S}.}$

(see [Rump'03], [Mackey,Mackey&Tisseur'06], ...)

$\Downarrow \quad \Downarrow \quad \Downarrow$

$$\boxed{\kappa(\lambda_0) \quad \approx \quad \kappa(\lambda_0, \mathbb{S})}$$

Take, for instance, $\mathbb{S} \equiv$ complex symm. matrices. Then

$$Y = \overline{X}, \quad \text{i.e.,} \quad \alpha = \|XY^H\|_2 = \|XX^T\|_2.$$

Let $XX^T = U\Sigma U^T$ be a Takagi factorization (i.e., an SVD), and let $u_1 = Ue_1$
Set $E_0 = \overline{u_1} u_1^H \in \mathbb{S}$. Then,

$$\rho(E_0 XX^T) = \rho(\overline{u_1} u_1^H XX^T) = \rho(u_1^H XX^T \overline{u_1}) = \sigma_{\max}(XX^T) = \|XX^T\|_2 = \alpha$$

$\square$

Mapping: $E_0 u_1 = \overline{u_1} u_1^H u_1 = \overline{u_1}$.

$\boxed{\text{Structured Jordan form for } A \in \mathbb{S}.}$

(see [Thompson'91], [Mehl'06], ...)

$\Downarrow \quad \Downarrow \quad \Downarrow$

$\boxed{\text{Induced structure in } XY^H.}$

$\Downarrow \quad \Downarrow \quad \Downarrow$

$\boxed{\text{Mapping theorems: } E_0 u = \beta v \ \text{ with } \ |\beta| = 1, \ \|E_0\|_2 \approx 1, \ E_0 \in \mathbb{S}.}$

(see [Rump'03], [Mackey,Mackey&Tisseur'06], ...)

$\Downarrow \quad \Downarrow \quad \Downarrow$

$$\boxed{\kappa(\lambda_0) \quad \approx \quad \kappa(\lambda_0, \mathbb{S})}$$

Works well for *complex* Toeplitz, Hankel, persymmetric, Hermitian, symmetric, *real* Hamiltonian, skew-Hamiltonian,...

Does **not** work for complex skew-symmetric $\longrightarrow$ can still be done using 'ad hoc' techniques

$$\boxed{\text{Structured Jordan form for } A \in \mathbb{S}.}$$

(see [Thompson'91], [Mehl'06], ...)

$$\Downarrow \qquad \Downarrow \qquad \Downarrow$$

$$\boxed{\text{Induced structure in } XY^H.}$$

$$\Downarrow \qquad \Downarrow \qquad \Downarrow$$

$$\boxed{\text{Mapping theorems: } E_0 u = \beta v \ \text{ with } \ |\beta| = 1, \ \ \|E_0\|_2 \approx 1, \ \ E_0 \in \mathbb{S}.}$$

(see [Rump'03], [Mackey,Mackey&Tisseur'06], ...)

$$\Downarrow \qquad \Downarrow \qquad \Downarrow$$

$$\boxed{\kappa(\lambda_0) \quad \approx \quad \kappa(\lambda_0, \mathbb{S})}$$

Works well for *complex* Toeplitz, Hankel, persymmetric, Hermitian, symmetric, *real* Hamiltonian, skew-Hamiltonian,...

Does **not** work for complex skew-symmetric $\longrightarrow$ can still be done using 'ad hoc' techniques

For more details (including matrix pencils & matrix polynomials) see

D. Kressner, M. J. Peláez, and J. Moro. Structured Hölder condition numbers for multiple eigenvalues, preprint, 2006.

$\boxed{\text{Structured Jordan form for } A \in \mathbb{S}.}$

(see [Thompson'91], [Mehl'06], ...)

$\Downarrow \quad \Downarrow \quad \Downarrow$

$\boxed{\text{Induced structure in } XY^H.}$

$\Downarrow \quad \Downarrow \quad \Downarrow$

$\boxed{\text{Mapping theorems: } E_0 u = \beta v \ \text{ with } \ |\beta| = 1, \ \|E_0\|_2 \approx 1, \ E_0 \in \mathbb{S}.}$

(see [Rump'03], [Mackey,Mackey&Tisseur'06], ...)

$\Downarrow \quad \Downarrow \quad \Downarrow$

$$\boxed{\kappa(\lambda_0) \quad \approx \quad \kappa(\lambda_0, \mathbb{S})}$$

Works well for *complex* Toeplitz, Hankel, persymmetric, Hermitian, symmetric, *real* Hamiltonian, skew-Hamiltonian,...

Does **not** work for complex skew-symmetric $\longrightarrow$ can still be done using 'ad hoc' techniques

What about **nongeneric** perturbations,
like skew-symmetric ones in the initial example?

Def: Given λ_0 e-value of A, and Y, X matrices of left and right e-vectors as before, a class $\mathbb{S}$ of structured matrices is **nongeneric** if $n_{\mathbb{S}} < n_1$ or, equivalently, if

$$\sup_{\substack{E \in \mathbb{S} \\ \|E\|_2 \leq 1}} \rho(Y^H E X) = 0$$

Def: Given λ_0 e-value of A, and Y, X matrices of left and right e-vectors as before, a class $\mathbb{S}$ of structured matrices is **nongeneric** if $n_{\mathbb{S}} < n_1$ or, equivalently, if

$$\sup_{\substack{E \in \mathbb{S} \\ \|E\|_2 \leq 1}} \rho(Y^H E X) = 0$$

For instance, in our initial 4 by 4 example, both $Y = y$ and $X = x$ are vectors, since there is one single largest Jordan block of size 3.

Moreover, $y = \overline{x}$ since A is complex skew-symmetric, so

$$Y^H E X = y^H E x = x^T E x = 0 \quad \text{for any skew-symmetric } E \in \mathbb{S}.$$

Def: Given λ_0 e-value of A, and Y, X matrices of left and right e-vectors as before, a class $\mathbb{S}$ of structured matrices is **nongeneric** if $n_{\mathbb{S}} < n_1$ or, equivalently, if

$$\sup_{\substack{E \in \mathbb{S} \\ \|E\|_2 \leq 1}} \rho(Y^H E X) = 0$$

For instance, in our initial 4 by 4 example, both $Y = y$ and $X = x$ are vectors, since there is one single largest Jordan block of size 3.

Moreover, $y = \overline{x}$ since A is complex skew-symmetric, so

$$Y^H E X = y^H E x = x^T E x = 0 \quad \text{for any skew-symmetric } E \in \mathbb{S}.$$

Definition

Given λ_0 e-value of A, and Y, X matrices of left and right e-vectors as before, a class $\mathbb{S}$ of structured matrices is said to be **fully** nongeneric if

$$Y^H E X = 0 \qquad \text{for any } E \in \mathbb{S}.$$

Def: Given λ_0 e-value of A, and Y, X matrices of left and right e-vectors as before, a class $\mathbb{S}$ of structured matrices is **nongeneric** if $n_{\mathbb{S}} < n_1$ or, equivalently, if

$$\sup_{\substack{E \in \mathbb{S} \\ \|E\|_2 \leq 1}} \rho(Y^H E X) = 0$$

For instance, in our initial 4 by 4 example, both $Y = y$ and $X = x$ are vectors, since there is one single largest Jordan block of size 3.

Moreover, $y = \overline{x}$ since A is complex skew-symmetric, so

$$Y^H E X = y^H E x = x^T E x = 0 \quad \text{for any skew-symmetric } E \in \mathbb{S}.$$

Definition

Given λ_0 e-value of A, and Y, X matrices of left and right e-vectors as before, a class $\mathbb{S}$ of structured matrices is said to be **fully** nongeneric if

$$Y^H E X = 0 \qquad \text{for any } E \in \mathbb{S}.$$

Can we somehow characterize **fully nongeneric** structures?

Let $\mathbb{S}$ be a **linear structure**, i.e. $\mathbb{S}$ is a linear subspace of $\mathbb{C}^{n\times n}$.

Def: Let $\mathbb{S}$ be a **linear subspace** of $\mathbb{C}^{n\times n}$. Then, the **skew-structure** associated with $\mathbb{S}$ is defined as

$$Skew(\mathbb{S}) = \{B \in \mathbb{C}^{n\times n} : vec(B)^H vec(A) = 0 \quad \forall A \in \mathbb{S}\},$$

where vec$\equiv$ stacking operator.

Theorem [Peláez & M. '08]

Let λ_0 be an e-value of A with e-vector matrices X and Y. Let y_i, x_j be, respectively, the columns of Y and X, and let $\mathbb{S}$ be a linear structure. Then $\mathbb{S}$ is fully nongeneric for λ_0 if and only if

$$y_i^H x_j \in Skew(\mathbb{S}) \qquad \text{for every} \quad i,j$$

The *Skew* operator produces all the 'customary´ skew-families:

Some linear structures and their skew-structures

$\mathbb{S}$	$Skew(\mathbb{S})$
Symmetric	Skewsymmetric
Pseudo Symmetric	Pseudo Skewsymmetric
Persymmetric	Perskewsymmetric
Hamiltonian	Skew-Hamiltonian
Hermitian	Skew-Hermitian
Pseudo-Hermitian	Pseudo-Skew-Hermitian
Toeplitz	zero d-sums
Hankel	zero ad-sums
Circulant	zero ed-sums
Cocirculant	zero ead-sums

$\mathbb{S}$	$Skew(\mathbb{S})$
Symmetric	Skewsymmetric
Pseudo Symmetric	Pseudo Skewsymmetric
Persymmetric	Perskewsymmetric
Hamiltonian	Skew-Hamiltonian
Hermitian	Skew-Hermitian
Pseudo-Hermitian	Pseudo-Skew-Hermitian
Toeplitz	zero d-sums
Hankel	zero ad-sums
Circulant	zero ed-sums
Cocirculant	zero ead-sums

Fully nongeneric structures: the leading exponents

For fully nongeneric structures,

$$\kappa(\lambda_0) = (n_1, \alpha), \qquad \kappa(\lambda_0, \mathbb{S}) = (n_{\mathbb{S}}, \alpha_{\mathbb{S}}) \qquad \text{with } n_{\mathbb{S}} < n_1,$$

where

- n_1 is the size of the largest λ_0-Jordan block

Fully nongeneric structures: the leading exponents

For fully nongeneric structures,

$$\kappa(\lambda_0) = (n_1, \alpha), \qquad \kappa(\lambda_0, \mathbb{S}) = (n_{\mathbb{S}}, \alpha_{\mathbb{S}}) \qquad \text{with } n_{\mathbb{S}} < n_1,$$

where

- n_1 is the size of the largest λ_0-Jordan block
- How can we find $n_{\mathbb{S}}$?

Fully nongeneric structures: the leading exponents

For fully nongeneric structures,

$$\kappa(\lambda_0) = (n_1, \alpha), \qquad \kappa(\lambda_0, \mathbb{S}) = (n_{\mathbb{S}}, \alpha_{\mathbb{S}}) \qquad \text{with } n_{\mathbb{S}} < n_1,$$

where

- n_1 is the size of the largest λ_0-Jordan block
- How can we find $n_{\mathbb{S}}$?

Use the **Newton diagram**, a geometric construction which gives both the leading exponents and the leading coefficients in the asymptotic expansions of the *roots of a polynomial in two variables.*

Fully nongeneric structures: the leading exponents

For fully nongeneric structures,

$$\kappa(\lambda_0) = (n_1, \alpha), \qquad \kappa(\lambda_0, \mathbb{S}) = (n_{\mathbb{S}}, \alpha_{\mathbb{S}}) \qquad \text{with } n_{\mathbb{S}} < n_1,$$

where

- n_1 is the size of the largest λ_0-Jordan block
- How can we find $n_{\mathbb{S}}$?

Use the **Newton diagram**, a geometric construction which gives both the leading exponents and the leading coefficients in the asymptotic expansions of the *roots of a polynomial in two variables.*

In our case,

$$p(\lambda, \varepsilon) = \det(A + \varepsilon E - \lambda I),$$

the characteristic polynomial of $A(\varepsilon) = A + \varepsilon E$.

For fully nongeneric structures,

$$\kappa(\lambda_0) = (n_1, \alpha), \qquad \kappa(\lambda_0, \mathbb{S}) = (n_{\mathbb{S}}, \alpha_{\mathbb{S}}) \qquad \text{with } n_{\mathbb{S}} < n_1,$$

where

- n_1 is the size of the largest λ_0-Jordan block
- How can we find $n_{\mathbb{S}}$?

Use the **Newton diagram**, a geometric construction which gives both the leading exponents and the leading coefficients in the asymptotic expansions of the *roots of a polynomial in two variables.*

In our case,

$$p(\lambda, \varepsilon) = \det(A + \varepsilon E - \lambda I),$$

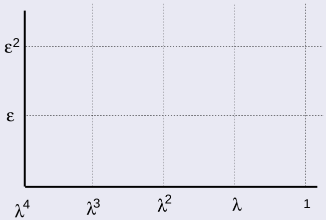
the characteristic polynomial of $A(\varepsilon) = A + \varepsilon E$.

How does the Newton diagram work?

Write the characteristic polynomial $p(\lambda, \varepsilon) = \det(A + \varepsilon E - \lambda I)$ of the perturbed matrix as a polynomial in λ with ε-dependent coefficients, e.g.,

$$p(\lambda, \varepsilon) = \lambda^4 + (2\varepsilon - \varepsilon^2)\lambda^3 + \varepsilon^2\lambda^2 + (\varepsilon - \varepsilon^3)\lambda + \varepsilon^2$$

Draw a cartesian grid and label the axes with λ, ε



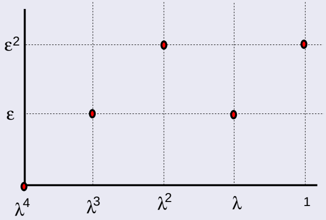
How does the ND work?

Write the characteristic polynomial $p(\lambda,\varepsilon)=\det(A+\varepsilon E-\lambda I)$ of the perturbed matrix as a polynomial in λ with ε-dependent coefficients, e.g.,

$$p(\lambda,\varepsilon)=\lambda^4+(2\varepsilon-\varepsilon^2)\lambda^3+\varepsilon^2\lambda^2+(\varepsilon-\varepsilon^3)\lambda+\varepsilon^2$$

Consider only the dominant terms

Step 1: plot a point for each dominant $\varepsilon^p\lambda^q$ terms

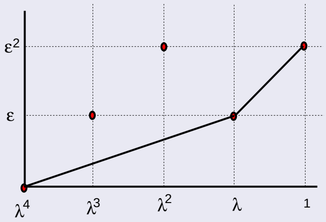


Write the characteristic polynomial $p(\lambda,\varepsilon)=\det(A+\varepsilon E-\lambda I)$ of the perturbed matrix as a polynomial in λ with ε-dependent coefficients, e.g.,

$$p(\lambda,\varepsilon)=\lambda^4+(2\varepsilon-\varepsilon^2)\lambda^3+\varepsilon^2\lambda^2+(\varepsilon-\varepsilon^3)\lambda+\varepsilon^2$$

Consider only the dominant terms

Step 2: draw the lower boundary of the convex hull: that's the ND

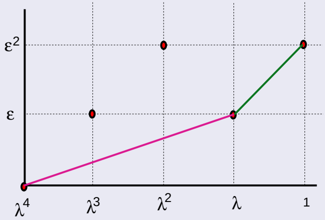


Write the characteristic polynomial $p(\lambda,\varepsilon) = \det(A + \varepsilon E - \lambda I)$ of the perturbed matrix as a polynomial in λ with ε-dependent coefficients, e.g.,

$$p(\lambda,\varepsilon) = \lambda^4 + (2\varepsilon - \varepsilon^2)\lambda^3 + \varepsilon^2\lambda^2 + (\varepsilon - \varepsilon^3)\lambda + \varepsilon^2$$

Consider only the dominant terms

Step 1: each slope is a leading power in the Puiseux ε-expansion



three e-vals of $O(\varepsilon^{1/3})$, one e-val of $O(\varepsilon)$

Unstructured perturbation

$$\begin{bmatrix} 0 & 1 & 0 & 1 \\ -1 & 0 & -i & 0 \\ 0 & i & 0 & i \\ -1 & 0 & -i & 0 \end{bmatrix} + \varepsilon E, \qquad E = \begin{bmatrix} e_{11} & e_{12} & e_{13} & e_{14} \\ e_{21} & e_{22} & e_{23} & e_{24} \\ e_{31} & e_{32} & e_{33} & e_{34} \\ e_{41} & e_{42} & e_{43} & e_{44} \end{bmatrix} \in \mathbb{C}^{4\times 4}$$

If $P^{-1}AP = J$, then $P^{-1}(A+\varepsilon E)P = J + \widetilde{E}$, where $J = J_3(0) \oplus J_1(0)$ and

$$\widetilde{E} = P^{-1}EP = \left[\begin{array}{ccc|c} * & * & * & * \\ * & * & * & * \\ \Phi & * & * & * \\ \hline * & * & * & * \end{array}\right], \qquad \text{with} \quad \Phi = y^H E x,$$

The 4×4 example: leading exponents via the ND

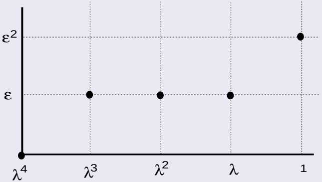
Unstructured perturbation

$$\begin{bmatrix} 0 & 1 & 0 & 1 \\ -1 & 0 & -i & 0 \\ 0 & i & 0 & i \\ -1 & 0 & -i & 0 \end{bmatrix} + \varepsilon E, \qquad E = \begin{bmatrix} e_{11} & e_{12} & e_{13} & e_{14} \\ e_{21} & e_{22} & e_{23} & e_{24} \\ e_{31} & e_{32} & e_{33} & e_{34} \\ e_{41} & e_{42} & e_{43} & e_{44} \end{bmatrix} \in \mathbb{C}^{4\times 4}$$

If $P^{-1}AP = J$, then $P^{-1}(A+\varepsilon E)P = J + \widetilde{E}$, where $J = J_3(0) \oplus J_1(0)$ and

$$\widetilde{E} = P^{-1}EP = \left[\begin{array}{ccc|c} * & * & * & * \\ * & * & * & * \\ \Phi & * & * & * \\ \hline * & * & * & * \end{array}\right], \qquad \text{with} \quad \Phi = y^H E x,$$

Step 1: plot the points



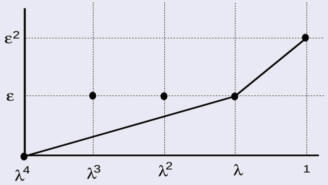
Unstructured perturbation

$$\begin{bmatrix} 0 & 1 & 0 & 1 \\ -1 & 0 & -i & 0 \\ 0 & i & 0 & i \\ -1 & 0 & -i & 0 \end{bmatrix} + \varepsilon E, \qquad E = \begin{bmatrix} e_{11} & e_{12} & e_{13} & e_{14} \\ e_{21} & e_{22} & e_{23} & e_{24} \\ e_{31} & e_{32} & e_{33} & e_{34} \\ e_{41} & e_{42} & e_{43} & e_{44} \end{bmatrix} \in \mathbb{C}^{4\times 4}$$

If $P^{-1}AP = J$, then $P^{-1}(A+\varepsilon E)P = J + \widetilde{E}$, where $J = J_3(0) \oplus J_1(0)$ and

$$\widetilde{E} = P^{-1}EP = \left[\begin{array}{ccc|c} * & * & * & * \\ * & * & * & * \\ \Phi & * & * & * \\ \hline * & * & * & * \end{array}\right], \qquad \text{with} \quad \Phi = y^H E x,$$

Step 2: draw lower boundary of convex hull



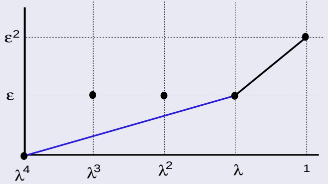
Unstructured perturbation

$$\begin{bmatrix} 0 & 1 & 0 & 1 \\ -1 & 0 & -i & 0 \\ 0 & i & 0 & i \\ -1 & 0 & -i & 0 \end{bmatrix} + \varepsilon E, \qquad E = \begin{bmatrix} e_{11} & e_{12} & e_{13} & e_{14} \\ e_{21} & e_{22} & e_{23} & e_{24} \\ e_{31} & e_{32} & e_{33} & e_{34} \\ e_{41} & e_{42} & e_{43} & e_{44} \end{bmatrix} \in \mathbb{C}^{4\times 4}$$

If $P^{-1}AP = J$, then $P^{-1}(A+\varepsilon E)P = J + \widetilde{E}$, where $J = J_3(0) \oplus J_1(0)$ and

$$\widetilde{E} = P^{-1}EP = \left[\begin{array}{ccc|c} * & * & * & * \\ * & * & * & * \\ \Phi & * & * & * \\ \hline * & * & * & * \end{array}\right], \qquad \text{with} \quad \Phi = y^H E x,$$

Step 2: draw lower boundary of convex hull



Smallest slope $1/3 \quad \Rightarrow n_1 = 3$

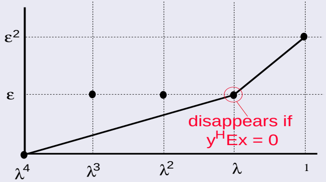
Unstructured perturbation

$$\begin{bmatrix} 0 & 1 & 0 & 1 \\ -1 & 0 & -i & 0 \\ 0 & i & 0 & i \\ -1 & 0 & -i & 0 \end{bmatrix} + \varepsilon E, \qquad E = \begin{bmatrix} e_{11} & e_{12} & e_{13} & e_{14} \\ e_{21} & e_{22} & e_{23} & e_{24} \\ e_{31} & e_{32} & e_{33} & e_{34} \\ e_{41} & e_{42} & e_{43} & e_{44} \end{bmatrix} \in \mathbb{C}^{4\times 4}$$

If $P^{-1}AP = J$, then $P^{-1}(A+\varepsilon E)P = J + \widetilde{E}$, where $J = J_3(0) \oplus J_1(0)$ and

$$\widetilde{E} = P^{-1}EP = \left[\begin{array}{ccc|c} * & * & * & * \\ * & * & * & * \\ \Phi & * & * & * \\ \hline * & * & * & * \end{array}\right], \qquad \text{with} \quad \Phi = y^H E x,$$

The term in $\lambda\varepsilon$ is present only if $\Phi = y^H E x \neq 0$



Smallest slope $1/3 \quad \Rightarrow n_1 = 3$

Structured perturbation

$$\begin{bmatrix} 0 & 1 & 0 & 1 \\ -1 & 0 & -i & 0 \\ 0 & i & 0 & i \\ -1 & 0 & -i & 0 \end{bmatrix} + \varepsilon E, \qquad E = \begin{bmatrix} 0 & e_{12} & e_{13} & e_{14} \\ -e_{12} & 0 & e_{23} & e_{24} \\ -e_{13} & -e_{23} & 0 & e_{34} \\ -e_{14} & -e_{24} & -e_{34} & 0 \end{bmatrix} \in \mathbb{S}$$

If $P^{-1}AP = J$, then $P^{-1}(A+\varepsilon E)P = J + \widetilde{E}$, where $J = J_3(0) \oplus J_1(0)$ and

$$\widetilde{E} = P^{-1}EP = \left[\begin{array}{ccc|c} * & * & * & * \\ * & * & * & * \\ 0 & * & * & * \\ \hline * & * & * & * \end{array}\right], \qquad \text{since} \quad \Phi = y^H E x = 0.$$

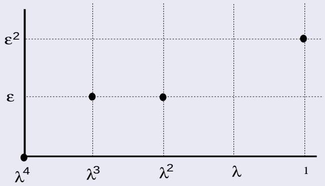
Structured perturbation

$$\begin{bmatrix} 0 & 1 & 0 & 1 \\ -1 & 0 & -i & 0 \\ 0 & i & 0 & i \\ -1 & 0 & -i & 0 \end{bmatrix} + \varepsilon E, \qquad E = \begin{bmatrix} 0 & e_{12} & e_{13} & e_{14} \\ -e_{12} & 0 & e_{23} & e_{24} \\ -e_{13} & -e_{23} & 0 & e_{34} \\ -e_{14} & -e_{24} & -e_{34} & 0 \end{bmatrix} \in \mathbb{S}$$

If $P^{-1}AP = J$, then $P^{-1}(A+\varepsilon E)P = J + \widetilde{E}$, where $J = J_3(0) \oplus J_1(0)$ and

$$\widetilde{E} = P^{-1}EP = \left[\begin{array}{ccc|c} * & * & * & * \\ * & * & * & * \\ 0 & * & * & * \\ \hline * & * & * & * \end{array}\right], \qquad \text{since} \quad \Phi = y^H E x = 0.$$

Step 1: plot the points



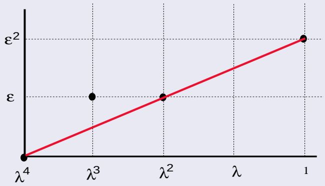
Structured perturbation

$$\begin{bmatrix} 0 & 1 & 0 & 1 \\ -1 & 0 & -i & 0 \\ 0 & i & 0 & i \\ -1 & 0 & -i & 0 \end{bmatrix} + \varepsilon E, \qquad E = \begin{bmatrix} 0 & e_{12} & e_{13} & e_{14} \\ -e_{12} & 0 & e_{23} & e_{24} \\ -e_{13} & -e_{23} & 0 & e_{34} \\ -e_{14} & -e_{24} & -e_{34} & 0 \end{bmatrix} \in \mathbb{S}$$

If $P^{-1}AP = J$, then $P^{-1}(A+\varepsilon E)P = J + \widetilde{E}$, where $J = J_3(0) \oplus J_1(0)$ and

$$\widetilde{E} = P^{-1}EP = \left[\begin{array}{ccc|c} * & * & * & * \\ * & * & * & * \\ 0 & * & * & * \\ \hline * & * & * & * \end{array}\right], \qquad \text{since} \quad \Phi = y^H E x = 0.$$

Step 2: draw lower boundary of convex hull



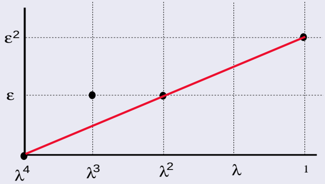
Structured perturbation

$$\begin{bmatrix} 0 & 1 & 0 & 1 \\ -1 & 0 & -i & 0 \\ 0 & i & 0 & i \\ -1 & 0 & -i & 0 \end{bmatrix} + \varepsilon E, \qquad E = \begin{bmatrix} 0 & e_{12} & e_{13} & e_{14} \\ -e_{12} & 0 & e_{23} & e_{24} \\ -e_{13} & -e_{23} & 0 & e_{34} \\ -e_{14} & -e_{24} & -e_{34} & 0 \end{bmatrix} \in \mathbb{S}$$

If $P^{-1}AP = J$, then $P^{-1}(A+\varepsilon E)P = J + \widetilde{E}$, where $J = J_3(0) \oplus J_1(0)$ and

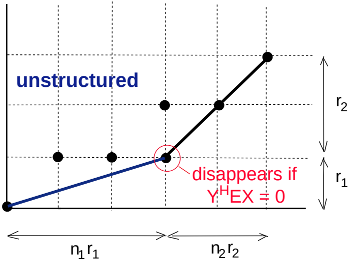
$$\widetilde{E} = P^{-1}EP = \left[\begin{array}{ccc|c} * & * & * & * \\ * & * & * & * \\ 0 & * & * & * \\ \hline * & * & * & * \end{array}\right], \qquad \text{since} \quad \Phi = y^H E x = 0.$$

Step 2: draw lower boundary of convex hull



Only slope $1/2$ $\quad \Rightarrow n_1 = 2$

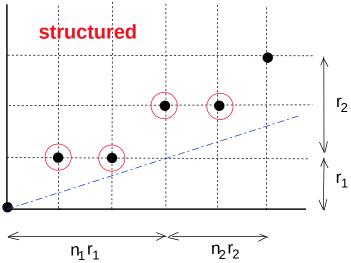
Same can be done in general for arbitrary **fully nongeneric** structure:



Bending point disappears from ND as soon as perturbations become fully nongeneric, i.e., when

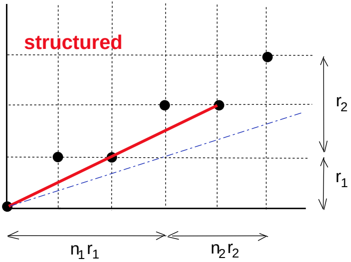
$$Y^H E X = 0 \quad \text{for all } E \in \mathbb{S}.$$

Same can be done in general for arbitrary **fully nongeneric** structure:



- For perturbations in $\mathbb{S}$, identify most likely points to lie on lowest segment in the ND
- Depend on sizes n_1, n_2 and numbers r_1, r_2 of λ_0-Jordan blocks. Possibly, also on the structure $\mathbb{S}$

Same can be done in general for arbitrary **fully nongeneric** structure:

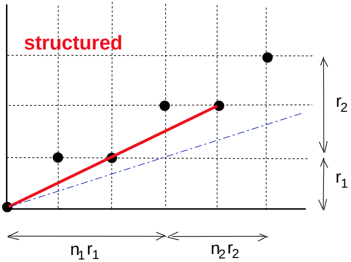


Finally, determine the reciprocal

$$n_{\mathbb{S}} < n_1$$

of the lowest possible slope

Same can be done in general for arbitrary **fully nongeneric** structure:



Finally, determine the reciprocal

$$n_{\mathbb{S}} < n_1$$

of the lowest possible slope

In this way, explicit formulas can be found for $n_{\mathbb{S}}$, depending on the quantities n_1, n_2, r_1, r_2

- Entry-wise information on the structure is important, to assess which points are present on the grid when the perturbations are structured.
- Otherwise, the formulas give just upper bounds on $n_{\mathbb{S}}$

Concluding remarks

- Structured and unstructured condition numbers for multiple eigenvalues can be defined and compared for several classes of structured matrices.

- Structured and unstructured condition numbers for multiple eigenvalues can be defined and compared for several classes of structured matrices.

- So far, not many significant differences between structured and unstructured condition number for **generic structures** except for a small number of special cases

  Main tools: structured canonical forms + mapping theorems

Concluding remarks

- Structured and unstructured condition numbers for multiple eigenvalues can be defined and compared for several classes of structured matrices.

- So far, not many significant differences between structured and unstructured condition number for **generic structures** except for a small number of special cases

  Main tools: structured canonical forms + mapping theorems

- First component $n_{\mathbb{S}}$ of structured condition number can be obtained for **fully nongeneric structures**

  Main tool: Newton diagram.

- Structured and unstructured condition numbers for multiple eigenvalues can be defined and compared for several classes of structured matrices.

- So far, not many significant differences between structured and unstructured condition number for **generic structures** except for a small number of special cases

  Main tools: structured canonical forms + mapping theorems

- First component $n_{\mathbb{S}}$ of structured condition number can be obtained for **fully nongeneric structures**

  Main tool: Newton diagram.

- Still several relevant classes to be explored (e.g., zero structures).