Multiple Eigenvalues with Prespecified Multiplicites and Pseudospectra

Emre Mengi

Department of Mathematics
Koç University
emengi@ku.edu.tr
http://home.ku.edu.tr/∼emengi

SIAM Conference on Applied Linear Algebra 2009

Problem

- ϵ-pseudospectrum

$$\begin{aligned}\Lambda_\epsilon(A) &= \{\lambda \in \mathbb{C} : \exists E \text{ s.t. } \|E\|_2 \le \epsilon \text{ and } \det(A + E - \lambda I) = 0\} \\ &= \{\lambda \in \mathbb{C} : \sigma_n(A - \lambda I) \le \epsilon\}.\end{aligned}$$

Problem

- ϵ-pseudospectrum

$$\begin{aligned}\Lambda_\epsilon(A) &= \{\lambda \in \mathbb{C} : \exists E \text{ s.t. } \|E\|_2 \leq \epsilon \text{ and } \det(A + E - \lambda I) = 0\} \\ &= \{\lambda \in \mathbb{C} : \sigma_n(A - \lambda I) \leq \epsilon\}.\end{aligned}$$

- Backward error of an eigenvalue λ

$$\inf\{\|\Delta A\| : \det(A + \Delta A - \lambda I) = 0\} = \sigma_n(A - \lambda I)$$

Problem

- ϵ-pseudospectrum

$$\begin{aligned}\Lambda_\epsilon(A) &= \{\lambda \in \mathbb{C} : \exists E \text{ s.t. } \|E\|_2 \leq \epsilon \text{ and } \det(A + E - \lambda I) = 0\} \\ &= \{\lambda \in \mathbb{C} : \sigma_n(A - \lambda I) \leq \epsilon\}.\end{aligned}$$

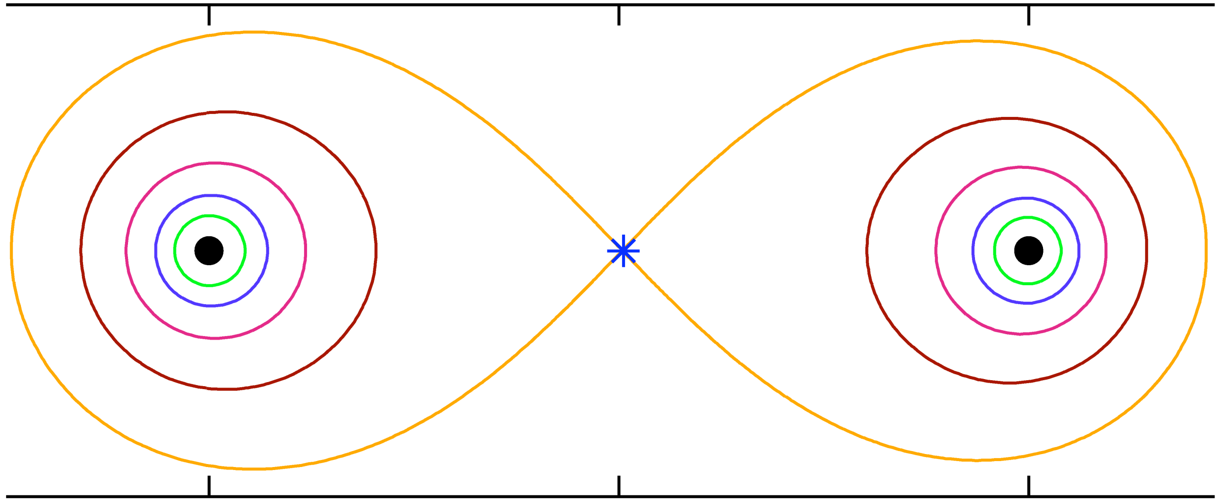
- Backward error of an eigenvalue λ

$$\inf\{\|\Delta A\| : \det(A + \Delta A - \lambda I) = 0\} = \sigma_n(A - \lambda I)$$

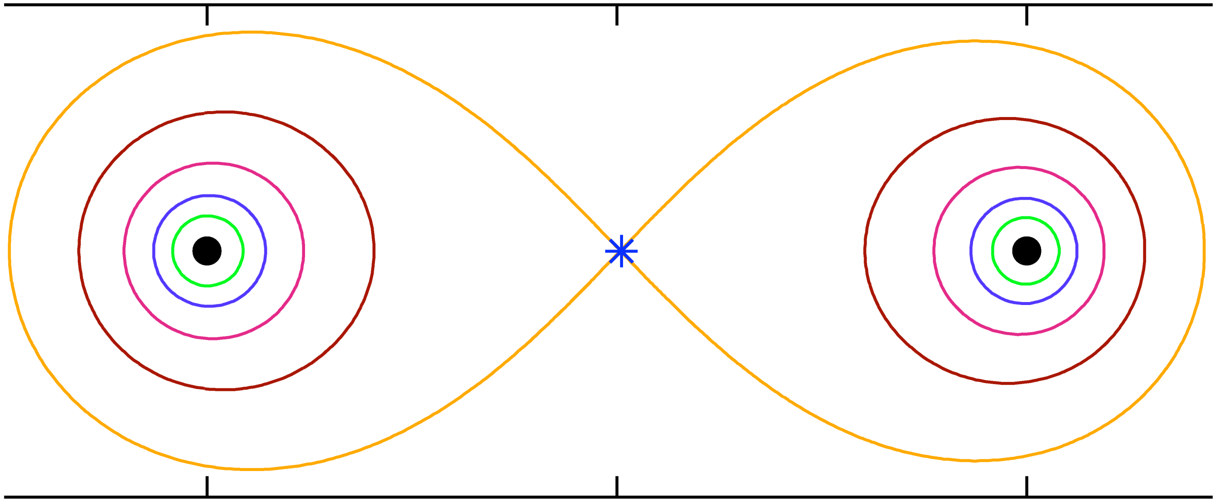
- If λ is a point on the boundary of $\Lambda_\epsilon(A)$, then

  $\exists \Delta A$ such that $\|\Delta A\| = \epsilon$ and $(A + \Delta A)$ has λ as an eigenvalue.

Problem



Problem

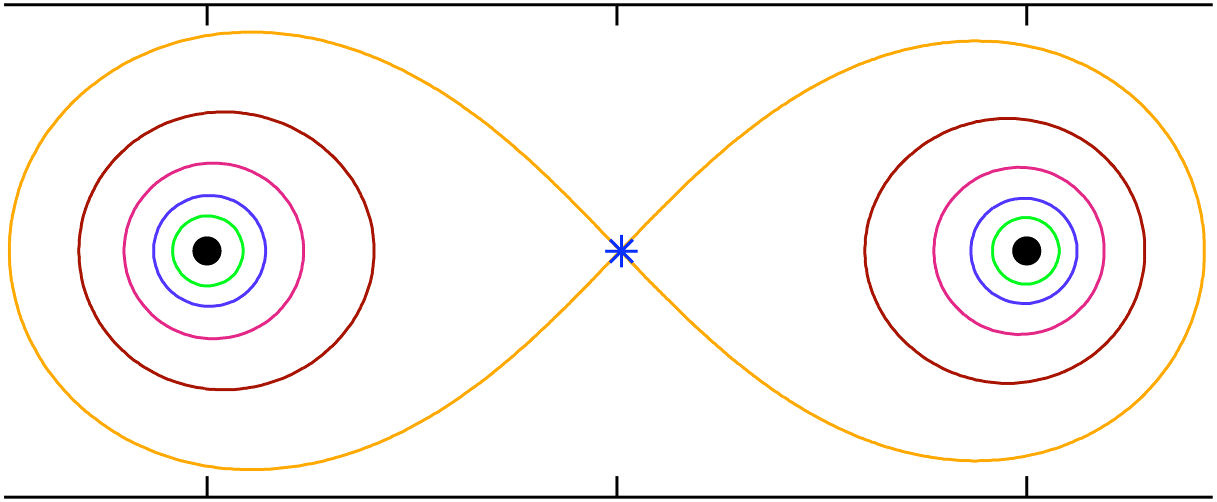


1. If λ is a point where two components of $\Lambda_\epsilon(A)$ coalesce, then

   $\exists \Delta A$ s.t. $\|\Delta A\| = \epsilon$ and $(A + \Delta A)$ has λ as an eval with mult $\geq r = 2$.

2. Smallest ϵ s.t. two components of $\Lambda_\epsilon(A)$ coalesce is the distance to the nearest matrix with an eigenvalue of multiplicity $\geq r = 2$.

Problem



1. If λ is a point where two components of $\Lambda_\epsilon(A)$ coalesce, then

   $\exists \Delta A$ s.t. $\|\Delta A\| = \epsilon$ and $(A + \Delta A)$ has λ as an eval with mult $\geq r = 2$.

2. Smallest ϵ s.t. two components of $\Lambda_\epsilon(A)$ coalesce is the distance to the nearest matrix with an eigenvalue of multiplicity $\geq r = 2$.

How do 1. and 2. generalize for an arbitrary r?

Motivation and Definition

- Absolute condition number of an eigenvalue λ

$$\kappa(\lambda) = \lim_{\delta \to 0} \sup_{\|\Delta A\| \leq \delta} \frac{|\delta\lambda|}{\|\Delta A\|} = \frac{1}{y^* x}$$

where

$\delta\lambda = \lambda(\Delta A) - \lambda$

$y, x \in \mathbb{C}^n$: unit left and right eigenvectors associated with λ

Motivation and Definition

- Absolute condition number of an eigenvalue λ

$$\kappa(\lambda) = \lim_{\delta \to 0} \sup_{\|\Delta A\| \leq \delta} \frac{|\delta\lambda|}{\|\Delta A\|} = \frac{1}{y^* x}$$

where

$\delta\lambda = \lambda(\Delta A) - \lambda$

$y, x \in \mathbb{C}^n$: unit left and right eigenvectors associated with λ

(*i.e.* $Ax = \lambda x$ and $y^* A = \lambda y^*$)

Motivation and Definition

- Absolute condition number of an eigenvalue λ

$$\kappa(\lambda) = \lim_{\delta \to 0} \sup_{\|\Delta A\| \leq \delta} \frac{|\delta\lambda|}{\|\Delta A\|} = \frac{1}{y^* x}$$

where

$\delta\lambda = \lambda(\Delta A) - \lambda$

$y, x \in \mathbb{C}^n$: unit left and right eigenvectors associated with λ

(*i.e.* $Ax = \lambda x$ and $y^* A = \lambda y^*$)

- $y^* x = 0$ for a (defective) eigenvalue associated with a Jordan block of size two.

Motivation and Definition

The eigenvalues corresponding to non-linear elementary divisors must, in general, be regarded as ill-conditioned ... However we must not be misled into thinking that this is the main form of ill-conditioning. Even if the eigenvalues are distinct and well separated they may still be very ill-conditioned.

J.H. WILKINSON, THE ALGEBRAIC EIGENVALUE PROBLEM

Motivation and Definition

The eigenvalues corresponding to non-linear elementary divisors must, in general, be regarded as ill-conditioned ... However we must not be misled into thinking that this is the main form of ill-conditioning. Even if the eigenvalues are distinct and well separated they may still be very ill-conditioned.

J.H. WILKINSON, THE ALGEBRAIC EIGENVALUE PROBLEM

$$W = \begin{bmatrix} 20 & 20 & & & & \\ & 19 & 20 & & & \\ & & 18 & 20 & & \\ & & & \ddots & \ddots & \\ & & & & 2 & 20 \\ & & & & & 1 \end{bmatrix}, \quad \kappa(\lambda) = \frac{20^{19}}{10!\,9!} \text{ for } \lambda = 10, 11$$

Motivation and Definition

$$W(\epsilon) = \begin{bmatrix} 20 & 20 & & & & \\ & 19 & 20 & & & \\ & & 18 & 20 & & \\ & & & \ddots & \ddots & \\ & & & & 2 & 20 \\ \epsilon & & & & & 1 \end{bmatrix}$$

$W(\epsilon)$ has $\lambda = 10.5$ as a defective multiple eigenvalue for $\epsilon \approx 7.8 \times 10^{-14}$.

Motivation and Definition

$$W(\epsilon) = \begin{bmatrix} 20 & 20 & & & & \\ & 19 & 20 & & & \\ & & 18 & 20 & & \\ & & & \ddots & \ddots & \\ & & & & 2 & 20 \\ \epsilon & & & & & 1 \end{bmatrix}$$

$W(\epsilon)$ has $\lambda = 10.5$ as a defective multiple eigenvalue for $\epsilon \approx 7.8 \times 10^{-14}$.

Definition (Wilkinson Distance):

The distance in 2-norm from A to the nearest defective matrix

$$\begin{aligned} \mathcal{W}(A) &= \inf\{\|\Delta A\|_2 : \exists\lambda \ \ (A + \Delta A) \text{ has } \lambda \text{ as a defect eigval}\} \\ &= \inf\{\|\Delta A\|_2 : \exists\lambda \ \ (A + \Delta A) \text{ has } \lambda \text{ as a mult eigval}\} \end{aligned}$$

is called the *Wilkinson distance* of A.

Motivation and Definition

- Wilkinson distance measures the sensitivity of the worst-conditioned eigenvalue to perturbations.
  - Any matrix close to defectiveness has an ill-conditioned eigenvalue.

Motivation and Definition

- Wilkinson distance measures the sensitivity of the worst-conditioned eigenvalue to perturbations.
  - Any matrix close to defectiveness has an ill-conditioned eigenvalue.
  - Conversely, any matrix with an ill-conditioned eigenvalue is close to defectiveness (Ruhe, 1970 and Wilkinson, 1971).

  Wilkinson's bound

$$\mathcal{W}(A) \leq \|A\|_2 / \sqrt{\kappa(\lambda)^2 - 1}$$

Wilkinson Distance and Pseudospectra

- The ϵ-pseudospectrum of a matrix A (the set of eigenvalues of matrices within an ϵ-neighborhood of A)

Wilkinson Distance and Pseudospectra

- The ϵ-pseudospectrum of a matrix A (the set of eigenvalues of matrices within an ϵ-neighborhood of A)

$$\Lambda_\epsilon(A) \quad = \quad \{\lambda \in \mathbb{C} : \exists E \text{ s.t. } \|E\|_2 \leq \epsilon \text{ and } \det(A + E - \lambda I) = 0\}.$$

Wilkinson Distance and Pseudospectra

- The ϵ-pseudospectrum of a matrix A (the set of eigenvalues of matrices within an ϵ-neighborhood of A)

$$\begin{aligned}\Lambda_\epsilon(A) &= \{\lambda \in \mathbb{C} : \exists E \text{ s.t. } \|E\|_2 \le \epsilon \text{ and } \det(A + E - \lambda I) = 0\} \\ &= \{\lambda \in \mathbb{C} : \sigma_n(A - \lambda I) \le \epsilon\}.\end{aligned}$$

Wilkinson Distance and Pseudospectra

- The ϵ-pseudospectrum of a matrix A (the set of eigenvalues of matrices within an ϵ-neighborhood of A)

$$\begin{aligned}\Lambda_\epsilon(A) &= \{\lambda \in \mathbb{C} : \exists E \text{ s.t. } \|E\|_2 \leq \epsilon \text{ and } \det(A + E - \lambda I) = 0\} \\ &= \{\lambda \in \mathbb{C} : \sigma_n(A - \lambda I) \leq \epsilon\}.\end{aligned}$$

- Define

$$\mathcal{C}(A) = \inf\{\epsilon : \#comp\,(\Lambda_\epsilon(A)) \leq n - 1\}$$

Wilkinson Distance and Pseudospectra

- The ϵ-pseudospectrum of a matrix A (the set of eigenvalues of matrices within an ϵ-neighborhood of A)

$$\begin{aligned}\Lambda_\epsilon(A) &= \{\lambda \in \mathbb{C} : \exists E \text{ s.t. } \|E\|_2 \leq \epsilon \text{ and } \det(A + E - \lambda I) = 0\} \\ &= \{\lambda \in \mathbb{C} : \sigma_n(A - \lambda I) \leq \epsilon\}.\end{aligned}$$

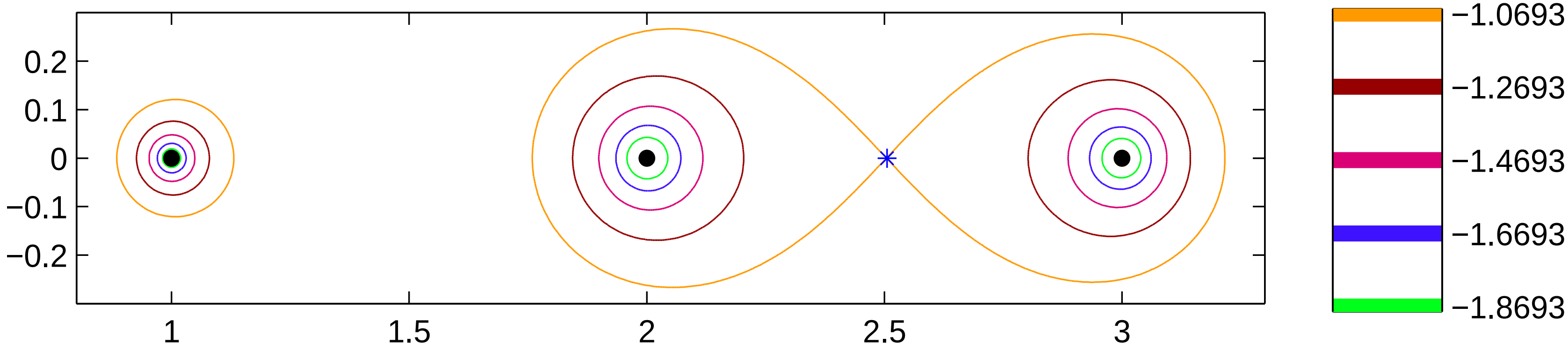
- Define

$$\mathcal{C}(A) = \inf\{\epsilon : \#comp\,(\Lambda_\epsilon(A)) \leq n - 1\}$$

- It was conjectured by Demmel (1983), and later proven by Alam and Bora (2005) that
  - $\mathcal{W}(A) = \mathcal{C}(A)$,
  - Two components of $\Lambda_\epsilon(A)$ for $\epsilon = \mathcal{C}(A)$ coalesce at λ_* iff a nearest matrix at a distance of $\mathcal{W}(A)$ has λ_* as a multiple eigenvalue.

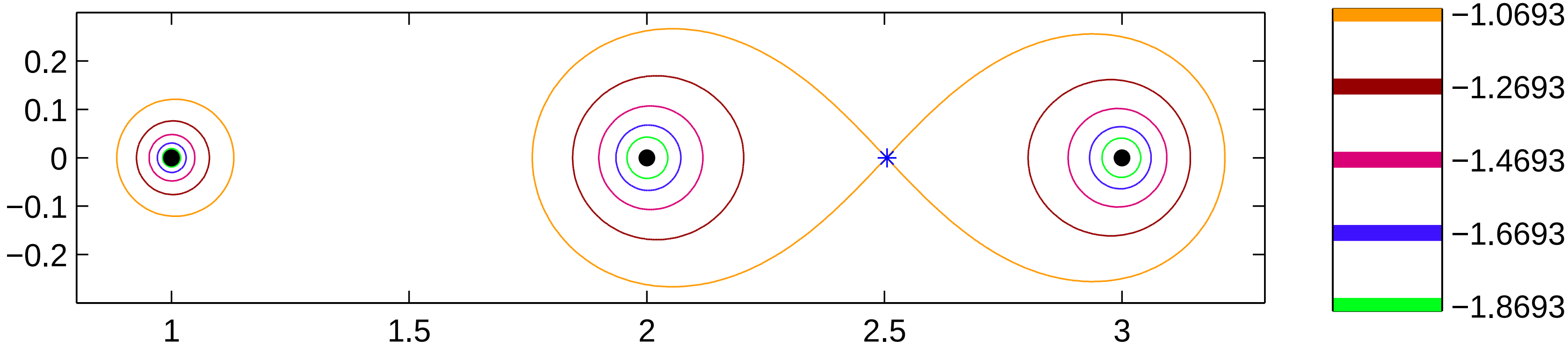
Wilkinson Distance and Pseudospectra

$$A = \begin{bmatrix} 1 & -1 & 2 \\ 0 & 2 & -2 \\ 0 & 0 & 3 \end{bmatrix}$$



Wilkinson Distance and Pseudospectra

$$A = \begin{bmatrix} 1 & -1 & 2 \\ 0 & 2 & -2 \\ 0 & 0 & 3 \end{bmatrix}$$



- $\mathcal{W}(A) = \mathcal{C}(A) = 10^{-1.0693}$
- $\lambda_* = 2.5057$ (point of coalescence marked with asterisk) is the multiple eigenvalue of a nearest matrix.

Wilkinson Distance and Pseudospectra

Proof of $\mathcal{W}(A) = \mathcal{C}(A)$

Wilkinson Distance and Pseudospectra

Proof of $\mathcal{W}(A) = \mathcal{C}(A)$

- Let $(A + \Delta A_*)$ be the nearest matrix such that $\|\Delta A_*\| = \mathcal{W}(A)$ and with λ_* as a multiple eigenvalue.

Wilkinson Distance and Pseudospectra

Proof of $\mathcal{W}(A) = \mathcal{C}(A)$

- Let $(A + \Delta A_*)$ be the nearest matrix such that $\|\Delta A_*\| = \mathcal{W}(A)$ and with λ_* as a multiple eigenvalue.
  - There exist two continuous curves $\lambda_1, \lambda_2 : [0, 1] \to \mathbb{C}$ s.t.
    1. $\lambda_1(t), \lambda_2(t)$ are distinct eigenvalues of $A + t\Delta A_*$ for $t < 1$.
    2. $\lambda_1(1) = \lambda_2(1) = \lambda_*$.

Wilkinson Distance and Pseudospectra

Proof of $\mathcal{W}(A) = \mathcal{C}(A)$

- Let $(A + \Delta A_*)$ be the nearest matrix such that $\|\Delta A_*\| = \mathcal{W}(A)$ and with λ_* as a multiple eigenvalue.
  - There exist two continuous curves $\lambda_1, \lambda_2 : [0,1] \to \mathbb{C}$ s.t.
    1. $\lambda_1(t), \lambda_2(t)$ are distinct eigenvalues of $A + t\Delta A_*$ for $t < 1$.
    2. $\lambda_1(1) = \lambda_2(1) = \lambda_*$.
  - $\lambda_1(t), \lambda_2(t) \in \Lambda_\epsilon(A)$ for all $t \in [0,1]$ and $\epsilon = \mathcal{W}(A)$.

Wilkinson Distance and Pseudospectra

Proof of $\mathcal{W}(A) = \mathcal{C}(A)$

- Let $(A + \Delta A_*)$ be the nearest matrix such that $\|\Delta A_*\| = \mathcal{W}(A)$ and with λ_* as a multiple eigenvalue.
  - There exist two continuous curves $\lambda_1, \lambda_2 : [0, 1] \to \mathbb{C}$ s.t.
    1. $\lambda_1(t), \lambda_2(t)$ are distinct eigenvalues of $A + t\Delta A_*$ for $t < 1$.
    2. $\lambda_1(1) = \lambda_2(1) = \lambda_*$.
  - $\lambda_1(t), \lambda_2(t) \in \Lambda_\epsilon(A)$ for all $t \in [0, 1]$ and $\epsilon = \mathcal{W}(A)$.
  - $\#comp\,(\Lambda_\epsilon(A)) < n$ for $\epsilon = \mathcal{W}(A)$.

Wilkinson Distance and Pseudospectra

Proof of $\mathcal{W}(A) = \mathcal{C}(A)$

- Let $(A + \Delta A_*)$ be the nearest matrix such that $\|\Delta A_*\| = \mathcal{W}(A)$ and with λ_* as a multiple eigenvalue.
  - There exist two continuous curves $\lambda_1, \lambda_2 : [0, 1] \to \mathbb{C}$ s.t.
    1. $\lambda_1(t), \lambda_2(t)$ are distinct eigenvalues of $A + t\Delta A_*$ for $t < 1$.
    2. $\lambda_1(1) = \lambda_2(1) = \lambda_*$.
  - $\lambda_1(t), \lambda_2(t) \in \Lambda_\epsilon(A)$ for all $t \in [0, 1]$ and $\epsilon = \mathcal{W}(A)$.
  - $\#comp\,(\Lambda_\epsilon(A)) < n$ for $\epsilon = \mathcal{W}(A)$.
- Therefore $\mathcal{W}(A) \geq \mathcal{C}(A)$.

Wilkinson Distance and Pseudospectra

Proof of $\mathcal{W}(A) = \mathcal{C}(A)$

- For the other direction let λ_* be a critical point of $g(\lambda) = \sigma_n(A - \lambda I)$.

Wilkinson Distance and Pseudospectra

Proof of $\mathcal{W}(A) = \mathcal{C}(A)$

- For the other direction let λ_* be a critical point of $g(\lambda) = \sigma_n(A - \lambda I)$.

  Theorem (Rellich):

  Let $f : \mathbb{R} \to \mathbb{R}$ be defined as $f(\omega) = \sigma_j(\mathcal{A}(\omega))$ (σ_j denoting the jth largest singular value). If the multiplicity of $\sigma_j(\mathcal{A}(\tilde{\omega}))$ is one and $\sigma_j(\mathcal{A}(\tilde{\omega})) \neq 0$, then $f(\omega)$ is real analytic at $\tilde{\omega}$ with the derivative

  $$f'(\tilde{\omega}) = \mathrm{Real}\left(u^* \frac{d\mathcal{A}(\tilde{\omega})}{d\omega} v\right)$$

  where u and v consist of a consistent pair of a unit left and a right singular vector associated with $\sigma_j(\mathcal{A}(\tilde{\omega}))$.

Wilkinson Distance and Pseudospectra

Proof of $\mathcal{W}(A) = \mathcal{C}(A)$

- For the other direction let λ_* be a critical point of $g(\lambda) = \sigma_n(A - \lambda I)$.

  Theorem (Rellich):

  Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be defined as $f(\omega) = \sigma_j(\mathcal{A}(\omega))$ (σ_j denoting the jth largest singular value). If the multiplicity of $\sigma_j(\mathcal{A}(\tilde{\omega}))$ is one and $\sigma_j(\mathcal{A}(\tilde{\omega})) \neq 0$, then $f(\omega)$ is real analytic at $\tilde{\omega}$ with the derivative

  $$f'(\tilde{\omega}) = \text{Real}\left(u^* \frac{d\mathcal{A}(\tilde{\omega})}{d\omega} v\right)$$

  where u and v consist of a consistent pair of a unit left and a right singular vector associated with $\sigma_j(\mathcal{A}(\tilde{\omega}))$.

- Viewing $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ yields

$$\text{Real}(u^* v) = \text{Real}(iu^* v) = 0 \Longrightarrow u^* v = 0$$

Wilkinson Distance and Pseudospectra

Proof of $\mathcal{W}(A) = \mathcal{C}(A)$

- Let $g(\lambda_*) = \sigma_n(A - \lambda I) = \epsilon$. Then defining $\Delta A = -\epsilon u v^*$ we have

$$\begin{aligned} (A + \Delta A - \lambda I)v = &\quad (A - \lambda I)v + \Delta A v = &\quad \epsilon u - \epsilon u = &\quad 0 \\ u^*(A + \Delta A - \lambda I) = &\quad u^*(A - \lambda I) + u^* \Delta A = &\quad \epsilon v^* - \epsilon v^* = &\quad 0 \end{aligned}$$

Wilkinson Distance and Pseudospectra

Proof of $\mathcal{W}(A) = \mathcal{C}(A)$

- Let $g(\lambda_*) = \sigma_n(A - \lambda I) = \epsilon$. Then defining $\Delta A = -\epsilon u v^*$ we have

$$\begin{aligned} (A + \Delta A - \lambda I)v = &\quad (A - \lambda I)v + \Delta A v = &\quad \epsilon u - \epsilon u = &\quad 0 \\ u^*(A + \Delta A - \lambda I) = &\quad u^*(A - \lambda I) + u^* \Delta A = &\quad \epsilon v^* - \epsilon v^* = &\quad 0 \end{aligned}$$

- Therefore u, v are left, right eigenvectors of $A + \Delta A$ associated with λ_* such that $u^* v = 0$ implying λ_* is a multiple eigenvalue of $A + \Delta A$.

Wilkinson Distance and Pseudospectra

Proof of $\mathcal{W}(A) = \mathcal{C}(A)$

- Let $g(\lambda_*) = \sigma_n(A - \lambda I) = \epsilon$. Then defining $\Delta A = -\epsilon u v^*$ we have

$$\begin{aligned}(A + \Delta A - \lambda I)v &= (A - \lambda I)v + \Delta A v = \epsilon u - \epsilon u = 0\\ u^*(A + \Delta A - \lambda I) &= u^*(A - \lambda I) + u^*\Delta A = \epsilon v^* - \epsilon v^* = 0\end{aligned}$$

- Therefore u, v are left, right eigenvectors of $A + \Delta A$ associated with λ_* such that $u^* v = 0$ implying λ_* is a multiple eigenvalue of $A + \Delta A$.

> Theorem (Alam and Bora):
> Let $\lambda_* \in \mathbb{C}$ be a critical point of $g(\lambda) = \sigma_n(A - \lambda I)$ such that $g(\lambda_*) = \epsilon > 0$. There exists a perturbation ΔA with norm ϵ such that $A + \Delta A$ has λ_* as a multiple eigenvalue.

Wilkinson Distance and Pseudospectra

Proof of $\mathcal{W}(A) = \mathcal{C}(A)$

- Now suppose λ_* is a point of coalescence of two components of $\Lambda_\epsilon(A)$ for $\epsilon = \mathcal{C}(A)$.

Wilkinson Distance and Pseudospectra

Proof of $\mathcal{W}(A) = \mathcal{C}(A)$

- Now suppose λ_* is a point of coalescence of two components of $\Lambda_\epsilon(A)$ for $\epsilon = \mathcal{C}(A)$.
  - Suppose the multiplicity of $\sigma_n(A - \lambda_* I) = \mathcal{C}(A)$ is two or greater with the left singular vectors u_1, u_2 and right singular vectors v_1, v_2. Define $\Delta A = -\mathcal{C}(A)[u_1 \;\; u_2][v_1 \;\; v_2]^*$ and note

$$(A + \Delta A - \lambda_* I)[v_1 \;\; v_2] = 0.$$

Wilkinson Distance and Pseudospectra

Proof of $\mathcal{W}(A) = \mathcal{C}(A)$

- Now suppose λ_* is a point of coalescence of two components of $\Lambda_\epsilon(A)$ for $\epsilon = \mathcal{C}(A)$.
  - Suppose the multiplicity of $\sigma_n(A - \lambda_* I) = \mathcal{C}(A)$ is two or greater with the left singular vectors u_1, u_2 and right singular vectors v_1, v_2. Define $\Delta A = -\mathcal{C}(A)[u_1 \ \ u_2][v_1 \ \ v_2]^*$ and note

$$(A + \Delta A - \lambda_* I)[v_1 \ \ v_2] = 0.$$

  - If the multiplicity of $\sigma_n(A - \lambda_* I) = \mathcal{C}(A)$ is one, λ_* must be a critical point. From previous theorem we have a perturbation ΔA of norm $\mathcal{C}(A)$ such that $A + \Delta A$ has λ_* as a multiple eigenvalue.

Wilkinson Distance and Pseudospectra

Proof of $\mathcal{W}(A) = \mathcal{C}(A)$

- Now suppose λ_* is a point of coalescence of two components of $\Lambda_\epsilon(A)$ for $\epsilon = \mathcal{C}(A)$.
  - Suppose the multiplicity of $\sigma_n(A - \lambda_* I) = \mathcal{C}(A)$ is two or greater with the left singular vectors u_1, u_2 and right singular vectors v_1, v_2. Define $\Delta A = -\mathcal{C}(A)[u_1 \;\; u_2][v_1 \;\; v_2]^*$ and note

$$(A + \Delta A - \lambda_* I)[v_1 \;\; v_2] = 0.$$

  - If the multiplicity of $\sigma_n(A - \lambda_* I) = \mathcal{C}(A)$ is one, λ_* must be a critical point. From previous theorem we have a perturbation ΔA of norm $\mathcal{C}(A)$ such that $A + \Delta A$ has λ_* as a multiple eigenvalue.

- Therefore $\mathcal{C}(A) \geq \mathcal{W}(A)$.

Wilkinson Distance and Pseudospectra

The following characterizations are useful for computational purposes.

- The smallest saddle point characterization

$$\mathcal{W}(A) = \inf \{\sigma_n(A - \lambda I) : \lambda \in \mathbb{C} \text{ is a saddle point of } \sigma_n(A - \lambda I)\}$$

Wilkinson Distance and Pseudospectra

The following characterizations are useful for computational purposes.

- The smallest saddle point characterization

$$\mathcal{W}(A) = \inf \{\sigma_n(A - \lambda I) : \lambda \in \mathbb{C} \text{ is a saddle point of } \sigma_n(A - \lambda I)\}$$

- The singular value characterization (Malyshev, 1999)

$$\mathcal{W}(A) = \inf_{\lambda \in \mathbb{C}} \sup_{\gamma \in (0,1]} \sigma_{2n-1}\left(\begin{bmatrix} A - \lambda I & \gamma I \\ 0 & A - \lambda I \end{bmatrix}\right)$$

Generalized Wilkinson Distance and Pseudospectra

Definition (Generalized Wilkinson Distance):

The distance from A to the nearest matrix with an eigenvalue of multiplicity r or greater

$\mathcal{W}_r(A) = \inf\{\|\Delta A\|_2 : \exists\lambda \ (A + \Delta A) \text{ has } \lambda \text{ as an eigenvalue of mult} \geq r\}$

is called the *generalized Wilkinson distance* of A.

Generalized Wilkinson Distance and Pseudospectra

Definition (Generalized Wilkinson Distance):

The distance from A to the nearest matrix with an eigenvalue of multiplicity r or greater

$\mathcal{W}_r(A) = \inf\{\|\Delta A\|_2 : \exists\lambda \;\; (A + \Delta A) \text{ has } \lambda \text{ as an eigenvalue of mult} \geq r\}$

is called the *generalized Wilkinson distance* of A.

- The singular value characterization (M. 2008)

$$\mathcal{W}_r(A) = \inf_{\lambda\in\mathbb{C}} \sup_{\gamma} \sigma_{nr-r+1} \left(\begin{bmatrix} A-\lambda I & \gamma_{1,2}I & \gamma_{1,3}I & \dots & \gamma_{1,r}I \\ 0 & A-\lambda I & \gamma_{2,3}I & & \vdots \\ 0 & 0 & \ddots & & \\ & & & A-\lambda I & \gamma_{r-1,r}I \\ & & & 0 & A-\lambda I \end{bmatrix} \right)$$

where $\gamma = [\gamma_{1,2} \; \gamma_{1,3} \;\; \dots \;\; \gamma_{r-1,r}]^T \in \mathbb{C}^{(r-1)r/2}$.

Generalized Wilkinson Distance and Pseudospectra

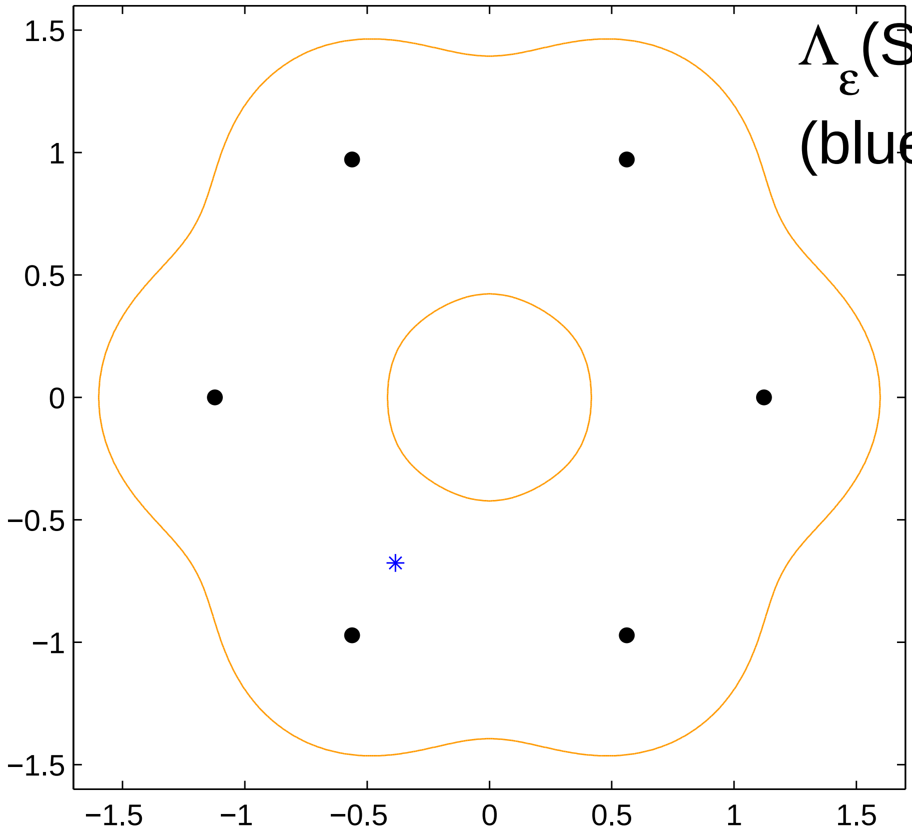
For the 6×6 smoke matrix S

- $\mathcal{W}_3(S) = 0.3270$ and the nearest matrix has $\lambda_* = -0.3841 - 0.6767i$ as the eigenvalue with multiplicity 3.

Generalized Wilkinson Distance and Pseudospectra

For the 6×6 smoke matrix S

- $\mathcal{W}_3(S) = 0.3270$ and the nearest matrix has $\lambda_* = -0.3841 - 0.6767i$ as the eigenvalue with multiplicity 3.



$\Lambda_\varepsilon(S)$ for $\varepsilon = W_3(S)$

(blue asterisks marks λ_*)

Generalized Wilkinson Distance and Pseudospectra

- $\mathcal{W}_r(A)$ does not appear to be relevant to $\Lambda_\epsilon(A)$.

Generalized Wilkinson Distance and Pseudospectra

- $\mathcal{W}_r(A)$ does not appear to be relevant to $\Lambda_\epsilon(A)$.
- The generalized (ϵ, r)-pseudospectrum (the eigenvalues with multiplicity r or greater of the matrices within an ϵ neighborhood)

$$\Lambda_{\epsilon,r}(A) = \{\lambda \in \mathbb{C} : \exists E \text{ s.t. } \|E\|_2 \leq \epsilon \text{ and } \operatorname{rank}(A + E - \lambda I)^r \leq n - r\}.$$

Generalized Wilkinson Distance and Pseudospectra

- $\mathcal{W}_r(A)$ does not appear to be relevant to $\Lambda_\epsilon(A)$.
- The generalized (ϵ, r)-pseudospectrum (the eigenvalues with multiplicity r or greater of the matrices within an ϵ neighborhood)

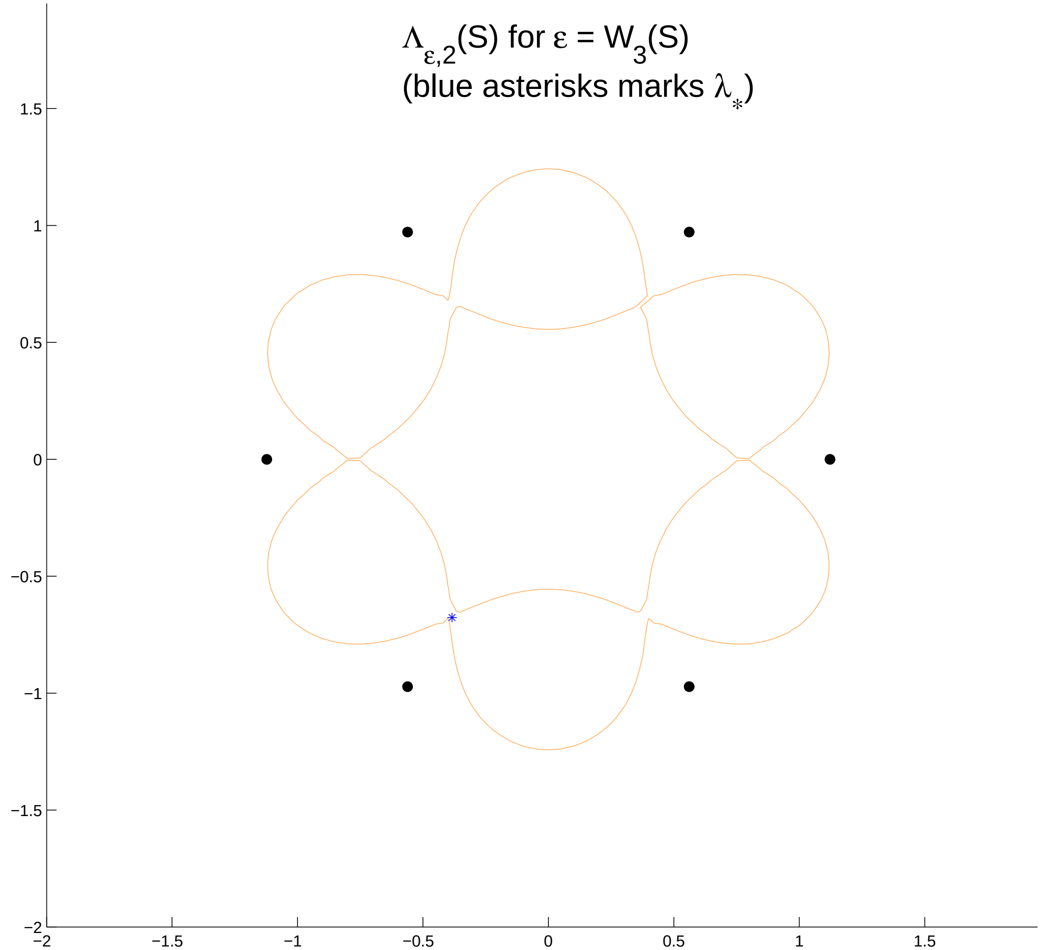
$$\Lambda_{\epsilon,r}(A) = \{\lambda \in \mathbb{C} : \exists E \text{ s.t. } \|E\|_2 \leq \epsilon \text{ and } \operatorname{rank}(A + E - \lambda I)^r \leq n - r\}.$$

- For instance $\Lambda_{\epsilon,2}$ consists of the multiple eigenvalues of the matrices within an ϵ neighborhood.

$$\begin{aligned} \Lambda_{\epsilon,2}(A) &= \{\lambda \in \mathbb{C} : \exists E \text{ s.t. } \|E\|_2 \leq \epsilon \text{ and } \operatorname{rank}(A + E - \lambda I)^2 \leq n - 2\} \\ &= \left\{\lambda \in \mathbb{C} : \sup_{\gamma \in (0,1]} \sigma_{2n-1}\left(\begin{bmatrix} A - \lambda I & \gamma I \\ 0 & A - \lambda I \end{bmatrix}\right) \leq \epsilon\right\} \end{aligned}$$

Generalized Wilkinson Distance and Pseudospectra

For the 6×6 smoke matrix S



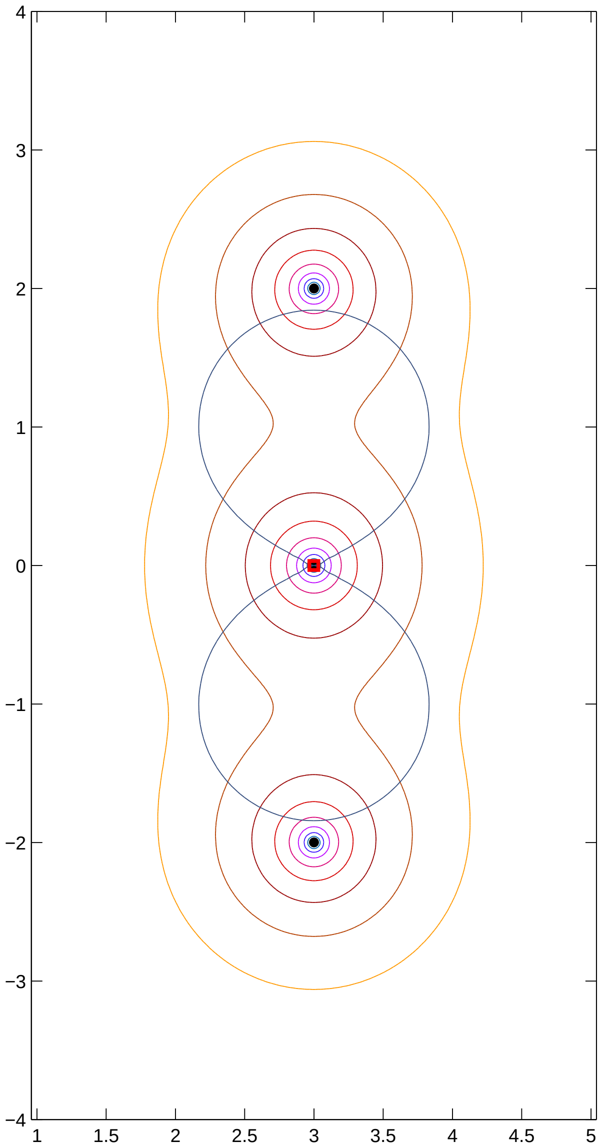
Generalized Wilkinson Distance and Pseudospectra

Blue curve
$\Lambda_{\varepsilon,2}$ for $\varepsilon = W_3(T)$

Orange curve
Λ_{ε} for $\varepsilon = W_3(T)$

Red Square (λ_*)

eigenvalue of multiplicity three of the nearest matrix



Generalized Wilkinson Distance and Pseudospectra

- It appears that
  - often two components of $\Lambda_{\epsilon,r}(A)$ coalesce for $\epsilon = \mathcal{W}_{r+1}(A)$,

Generalized Wilkinson Distance and Pseudospectra

- It appears that
  - often two components of $\Lambda_{\epsilon,r}(A)$ coalesce for $\epsilon = \mathcal{W}_{r+1}(A)$,
  - a point of coalescence is λ_* whenever a nearest matrix at a distance of $\mathcal{W}_{r+1}(A)$ has λ_* as an eigenvalue of multiplicity $r+1$.

Generalized Wilkinson Distance and Pseudospectra

- It appears that
  - often two components of $\Lambda_{\epsilon,r}(A)$ coalesce for $\epsilon = \mathcal{W}_{r+1}(A)$,
  - a point of coalescence is λ_* whenever a nearest matrix at a distance of $\mathcal{W}_{r+1}(A)$ has λ_* as an eigenvalue of multiplicity $r+1$.
- Alam and Bora showed that

$$\sigma_n(A - \lambda_* I) = \epsilon > 0 \;\text{ and }\; (\lambda_* \text{ is a critical point of } \sigma_n(A - \lambda I))$$

$$\Longrightarrow$$

$$\exists \Delta A \text{ s.t. } \|\Delta A\| = \epsilon \text{ and } (\lambda_* \text{ is a multiple eigenvalue of } A + \Delta A)$$

Generalized Wilkinson Distance and Pseudospectra

Proposition

$$g(\lambda_*) = \epsilon > 0 \;\text{ and }\; (\lambda_* \text{ is a critical point of } g(\lambda))$$

$$\Longrightarrow$$

$$\exists \Delta A \text{ s.t. } \|\Delta A\| = \epsilon \text{ and } (\lambda_* \text{ is an eigenvalue of } A + \Delta A \text{ with multiplicity } \geq r+1)$$

Generalized Wilkinson Distance and Pseudospectra

Proposition

$$g(\lambda_*) = \epsilon > 0 \;\text{ and }\; (\lambda_* \text{ is a critical point of } g(\lambda))$$

$$\Longrightarrow$$

$$\exists \Delta A \text{ s.t. } \|\Delta A\| = \epsilon \text{ and } (\lambda_* \text{ is an eigenvalue of } A + \Delta A \text{ with multiplicity } \geq r+1)$$

where

$$g(\lambda) = \sup_{\gamma} \sigma_{nr-r+1} \left(\underbrace{\begin{bmatrix} A-\lambda I & \gamma_{1,2} I & \gamma_{1,3} I & \dots & \gamma_{1,r} I \\ 0 & A-\lambda I & \gamma_{2,3} I & & \vdots \\ 0 & 0 & \ddots & & \\ & & & A-\lambda I & \gamma_{r-1,r} I \\ & & & 0 & A-\lambda I \end{bmatrix}}_{\mathcal{A}(\lambda,\gamma)} \right)$$

Generalized Wilkinson Distance and Pseudospectra

Outline of the proof of the proposition

Assumptions

- Suppose $g(\lambda_*) = \sigma_{nr-r+1}\left(\mathcal{A}(\lambda_*, \gamma_*)\right)$. Then the multiplicity of $\sigma_{nr-r+1}\left(\mathcal{A}(\lambda_*, \gamma_*)\right)$ is one.

Generalized Wilkinson Distance and Pseudospectra

Outline of the proof of the proposition

Assumptions

- Suppose $g(\lambda_*) = \sigma_{nr-r+1}\left(\mathcal{A}(\lambda_*, \gamma_*)\right)$. Then the multiplicity of $\sigma_{nr-r+1}\left(\mathcal{A}(\lambda_*, \gamma_*)\right)$ is one.

- Let $\mathcal{U} = [u_1^T \; u_2^T \; \ldots \; u_r^T]$ and $\mathcal{V} = [v_1^T \; v_2^T \; \ldots \; v_r^T]$ be a consistent pair of unit left and right singular vectors associated with $\sigma_{nr-r+1}\left(\mathcal{A}(\lambda_*, \gamma_*)\right)$ where $u_1, \ldots, u_r, v_1, \ldots, v_r \in \mathbb{C}^n$. The sets

$$\{u_1, \; u_2, \; \ldots \; u_r\} \quad \text{and} \quad \{v_1, \; v_2, \; \ldots \; v_r\}$$

are linearly independent.

Generalized Wilkinson Distance and Pseudospectra

Outline of the proof of the proposition

- A nearest matrix with λ_* as an eigenvalue of multiplicity r or greater is given by

$$A - \epsilon[u_1 \; u_2 \; \ldots \; u_r][v_1 \; v_2 \; \ldots \; v_r]^\dagger$$

Generalized Wilkinson Distance and Pseudospectra

Outline of the proof of the proposition

- A nearest matrix with λ_* as an eigenvalue of multiplicity r or greater is given by

$$A - \epsilon[u_1 \ u_2 \ \ldots \ u_r][v_1 \ v_2 \ \ldots \ v_r]^\dagger$$

- The fact that $g(\lambda_*) = \sup_\gamma \sigma_{nr-r+1}\left(\mathcal{A}(\lambda_*, \gamma)\right)$ implies

$$u_j^* v_k = 0 \ \text{ for all } j, k \text{ such that } j < k.$$

Generalized Wilkinson Distance and Pseudospectra

Outline of the proof of the proposition

- Define for $i = 1, \ldots, n$

$$g_i(\lambda) = \sup_{\gamma} \sigma_{nr-r+1} \left(\begin{bmatrix} A - \lambda_* I & \gamma_{1,2} I & \gamma_{1,3} I & \cdots & \gamma_{1,r} I \\ 0 & \ddots & \gamma_{2,3} I & & \vdots \\ 0 & 0 & \underbrace{A - \lambda I}_{ith \text{ block row}} & & \\ & & & \ddots & \gamma_{r-1,r} I \\ & & & 0 & A - \lambda_* I \end{bmatrix} \right)$$

Generalized Wilkinson Distance and Pseudospectra

Outline of the proof of the proposition

- Define for $i = 1, \ldots, n$

$$g_i(\lambda) = \sup_{\gamma} \sigma_{nr-r+1} \left(\begin{bmatrix} A - \lambda_* I & \gamma_{1,2} I & \gamma_{1,3} I & \ldots & \gamma_{1,r} I \\ 0 & \ddots & \gamma_{2,3} I & & \vdots \\ 0 & 0 & \underbrace{A - \lambda I}_{ith \text{ block row}} & & \\ & & & \ddots & \gamma_{r-1,r} I \\ & & & 0 & A - \lambda_* I \end{bmatrix} \right)$$

- $g_i(\lambda)$ is the distance to the nearest matrix with $r - 1$ of the eigenvalues equal to λ_* and the remaining equal to λ. Therefore

$$g_i(\lambda) = g_j(\lambda) \ \text{ for all } i, j.$$

Generalized Wilkinson Distance and Pseudospectra

Outline of the proof of the proposition

- Take the derivatives of g_i for all i

$$g_i'(\lambda_*) = u_i^* v_i = u_j^* v_j = g_j'(\lambda_*)$$

Outline of the proof of the proposition

- Take the derivatives of g_i for all i

$$g_i'(\lambda_*) = u_i^* v_i = u_j^* v_j = g_j'(\lambda_*)$$

- Furthermore λ_* is a critical point of $g(\lambda)$ implying

$$g'(\lambda_*) = \sum_{i=1}^{n} u_i^* v_i = 0 \quad \Longrightarrow \quad u_i^* v_i = 0 \text{ for all } i$$

Generalized Wilkinson Distance and Pseudospectra

Outline of the proof of the proposition

Theorem:

Let u_1 be a left eigenvector and $\{v_1, v_2, \ldots, v_r\}$ be a linearly independent set of right eigenvectors of A associated with λ. If $u_1^* v_j = 0$ for $j = 1, \ldots, r$, then λ is an eigenvalue of A with multiplicity $r + 1$ or greater.

Generalized Wilkinson Distance and Pseudospectra

Outline of the proof of the proposition

Theorem:
Let u_1 be a left eigenvector and $\{v_1, v_2, \ldots, v_r\}$ be a linearly independent set of right eigenvectors of A associated with λ. If $u_1^* v_j = 0$ for $j = 1, \ldots, r$, then λ is an eigenvalue of A with multiplicity $r + 1$ or greater.

- For the matrix

$$A_p = A - \epsilon[u_1 \ u_2 \ \ldots \ u_r][v_1 \ v_2 \ \ldots \ v_r]^\dagger$$

$u_1, u_2, \ldots u_r$ are the left eigenvectors and $v_1, v_2, \ldots, v_r$ are the right eigenvectors associated with λ_*.

Generalized Wilkinson Distance and Pseudospectra

Outline of the proof of the proposition

- Since $u_1^* v_j = 0$ for $j = 1, \ldots, r$, it follows that λ_* is an eigenvalue of A_p with multiplicity $r + 1$ or greater.

Generalized Wilkinson Distance and Pseudospectra

Outline of the proof of the proposition

- Since $u_1^* v_j = 0$ for $j = 1, \ldots, r$, it follows that λ_* is an eigenvalue of A_p with multiplicity $r+1$ or greater.
- It can also be shown that

$$\|A_p - A\| = \| -\epsilon[u_1 \; u_2 \; \ldots \; u_r][v_1 \; v_2 \; \ldots \; v_r]^\dagger \| = \epsilon$$

deducing the proposition.

Generalized Wilkinson Distance and Pseudospectra

Corollary of the Proposition

- Define

$$\mathcal{C}_r(A) = \inf\{\epsilon : \text{ two components of } \Lambda_{\epsilon,r-1}(A) \text{ coalesce}\}$$

Generalized Wilkinson Distance and Pseudospectra

Corollary of the Proposition

- Define

$$\mathcal{C}_r(A) = \inf\{\epsilon : \text{ two components of } \Lambda_{\epsilon,r-1}(A) \text{ coalesce}\}$$

- If at a point of coalescence of $\Lambda_{\epsilon,r-1}(A)$ for $\epsilon = \mathcal{C}_r(A)$ the multiplicity and linear independence assumptions hold, then

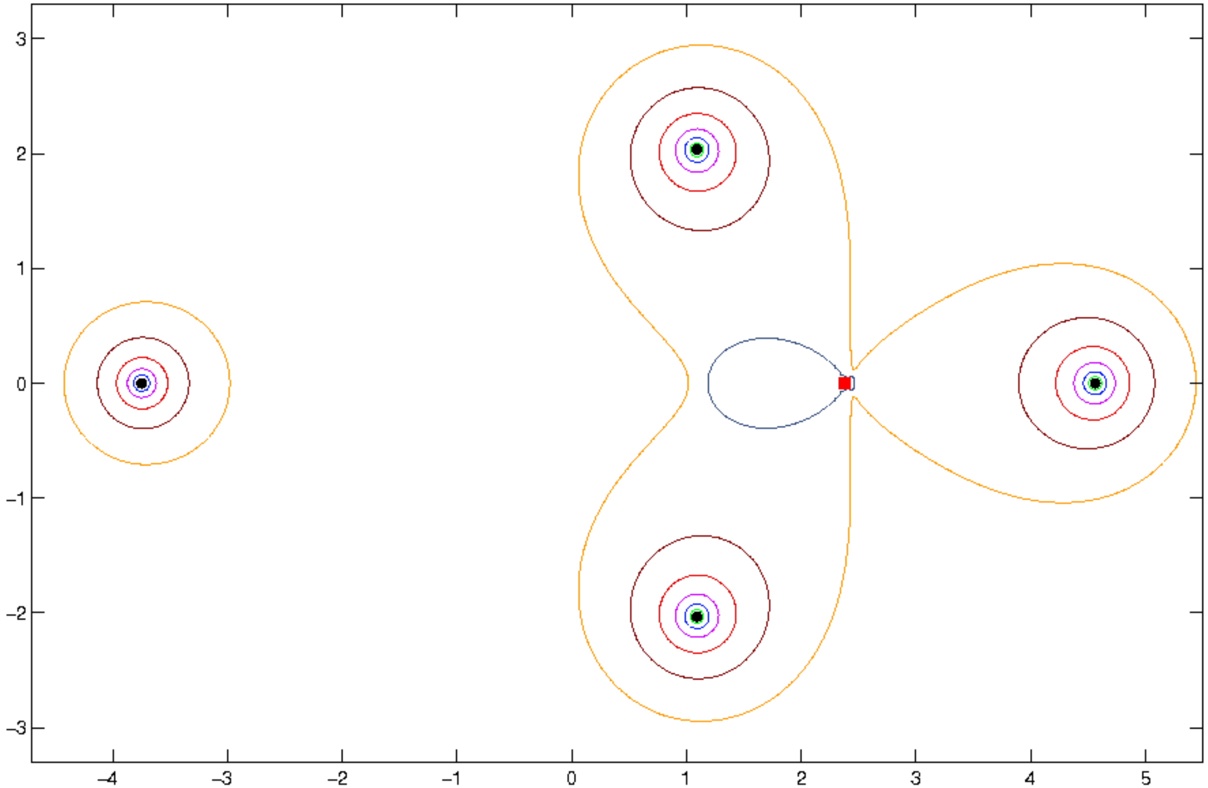
$$\mathcal{W}_r(A) \leq \mathcal{C}_r(A)$$

Generalized Wilkinson Distance and Pseudospectra

- The other direction $\mathcal{W}_r(A) \geq \mathcal{C}_r(A)$ seems to be usually true, but not always.

Generalized Wilkinson Distance and Pseudospectra

- The other direction $\mathcal{W}_r(A) \geq \mathcal{C}_r(A)$ seems to be usually true, but not always.



Blue Curve – $\Lambda_{\varepsilon,2}(H)$ for $\varepsilon = W_3(H)$

Orange Curve – $\Lambda_{\varepsilon}(H)$ for $\varepsilon = W_3(H)$

Red Square (λ_*) – Eigenvalue of multiplicity three of the nearest matrix

References

- J.H. Wilkinson, Note on matrices with a very ill-conditioned eigenproblem, *Numer Math*, 19:176-178, 1972
- J.W. Demmel, Computing Stable Eigendecompositions of matrices, *Linear Algebra Appl*, 79:163-193, 1986.
- A.N. Malyshev, A formula for the 2-norm distance from a matrix to the set of matrices with multiple eigenvalues, *Numer Math*, 83:443-454, 1999.
- R. Alam and S. Bora, On sensitivity of eigenvalues and eigendecompositions of matrices, *Linear Algebra Appl*, 396:273-301, 2005
- E. Mengi, Locating a nearest matrix with an eigenvalue of prespecified algebraic multiplicity, *Numer Math*, 2009 (accepted subject to minor revision).