Pseudospectra, Nearest Matrices with Multiple Eigenvalues and Optimization of Symmetric Eigenvalues

Emre Mengi

Department of Mathematics
Koç University
Istanbul, Turkey

ICIAM VII
July 20, 2011

Outline

Pseudospectra

Definition (ϵ-pseudospectrum)

$$\Lambda_\epsilon(A) = \bigcup_{\|E\|_2 \le \epsilon} \Lambda(A+E)$$

$$= \left\{ \lambda \in \mathbb{C} : \|(A-\lambda I)^{-1}\|_2 \ge \frac{1}{\epsilon} \right\}$$

$$= \{ \lambda \in \mathbb{C} : \sigma_n(A-\lambda I) \le \epsilon \}$$

($\sigma_j(\cdot)$: jth largest singular value)

Properties

- $\Lambda_\epsilon(A)$ is compact.
- It has at most n disconnected components (one component around each eigenvalue).

Pseudospectra

Definition (ϵ-pseudospectrum)

$$\begin{aligned}\Lambda_\epsilon(A) &= \bigcup_{\|E\|_2 \leq \epsilon} \Lambda(A+E) \\ &= \left\{\lambda \in \mathbb{C} : \|(A-\lambda I)^{-1}\|_2 \geq \frac{1}{\epsilon}\right\} \\ &= \{\lambda \in \mathbb{C} : \sigma_n(A-\lambda I) \leq \epsilon\}\end{aligned}$$

($\sigma_j(\cdot)$: jth largest singular value)

Properties

- $\Lambda_\epsilon(A)$ is compact.
- It has at most *n* disconnected components (one component around each eigenvalue).

Pseudospectra

Definition (ϵ-pseudospectrum)

$$\begin{aligned}\Lambda_\epsilon(A) &= \bigcup_{\|E\|_2 \le \epsilon} \Lambda(A+E) \\ &= \left\{ \lambda \in \mathbb{C} : \|(A-\lambda I)^{-1}\|_2 \ge \frac{1}{\epsilon} \right\} \\ &= \{ \lambda \in \mathbb{C} : \sigma_n(A-\lambda I) \le \epsilon \}\end{aligned}$$

($\sigma_j(\cdot)$: jth largest singular value)

Properties

- $\Lambda_\epsilon(A)$ is compact.
- It has at most n disconnected components (one component around each eigenvalue).

Pseudospectra

Definition (ϵ-pseudospectrum)

$$\begin{aligned} \Lambda_\epsilon(A) &= \bigcup_{\|E\|_2 \le \epsilon} \Lambda(A+E) \\ &= \left\{ \lambda \in \mathbb{C} : \|(A-\lambda I)^{-1}\|_2 \ge \frac{1}{\epsilon} \right\} \\ &= \{ \lambda \in \mathbb{C} : \sigma_n(A - \lambda I) \le \epsilon \} \end{aligned}$$

($\sigma_j(\cdot)$: jth largest singular value)

Properties

- $\Lambda_\epsilon(A)$ is compact.
- It has at most n disconnected components (one component around each eigenvalue).

Pseudospectra

Definition (ϵ-pseudospectrum)

$$\begin{aligned}\Lambda_\epsilon(A) &= \bigcup_{\|E\|_2 \le \epsilon} \Lambda(A+E) \\ &= \left\{\lambda \in \mathbb{C} : \|(A-\lambda I)^{-1}\|_2 \ge \frac{1}{\epsilon}\right\} \\ &= \{\lambda \in \mathbb{C} : \sigma_n(A-\lambda I) \le \epsilon\}\end{aligned}$$

($\sigma_j(\cdot)$: jth largest singular value)

Properties

- $\Lambda_\epsilon(A)$ is compact.
- It has at most n disconnected components (one component around each eigenvalue).

Wilkinson Distance

Definition (Wilkinson Distance)

$$\mathcal{W}(A) = \inf\{\|\delta A\|_2 : \exists\lambda \ \ (A + \delta A) \text{ has } \lambda \text{ as a multiple eigenvalue}\}$$

Note: Above definition is equivalent to the distance to the nearest defective matrix.

Wilkinson Distance

Definition (Wilkinson Distance)

$\mathcal{W}(A) = \inf\{\|\delta A\|_2 : \exists\lambda \;\; (A + \delta A) \text{ has } \lambda \text{ as a multiple eigenvalue}\}$

Note: Above definition is equivalent to the distance to the nearest defective matrix.

Wilkinson Distance

Wilkinson distance measures the sensitivity of the worst-conditioned eigenvalue to perturbations.

- Any matrix close to being defective has an ill-conditioned eigenvalue.
- Conversely, Ruhe (1970) and Wilkinson (1971) showed that any matrix with an ill-conditioned eigenval. is close to being defective

Wilkinson's bound

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

Wilkinson Distance

Wilkinson distance measures the sensitivity of the worst-conditioned eigenvalue to perturbations.

- Any matrix close to being defective has an ill-conditioned eigenvalue.
- Conversely, Ruhe (1970) and Wilkinson (1971) showed that any matrix with an ill-conditioned eigenval. is close to being defective

  Wilkinson's bound

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

Wilkinson Distance

Wilkinson distance measures the sensitivity of the worst-conditioned eigenvalue to perturbations.

- Any matrix close to being defective has an ill-conditioned eigenvalue.
- Conversely, Ruhe (1970) and Wilkinson (1971) showed that any matrix with an ill-conditioned eigenval. is close to being defective

Wilkinson's bound

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

Outline

Geometric View of Wilkinson Distance

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

- It was conjectured by Demmel (1983), and later proven by Alam and Bora (2005) that
  - $\mathcal{W}(A) = \mathcal{C}(A)$.
  - Furthermore 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.

Geometric View of Wilkinson Distance

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

- It was conjectured by Demmel (1983), and later proven by Alam and Bora (2005) that
  - $\mathcal{W}(A) = \mathcal{C}(A)$.
  - Furthermore 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.

Geometric View of Wilkinson Distance

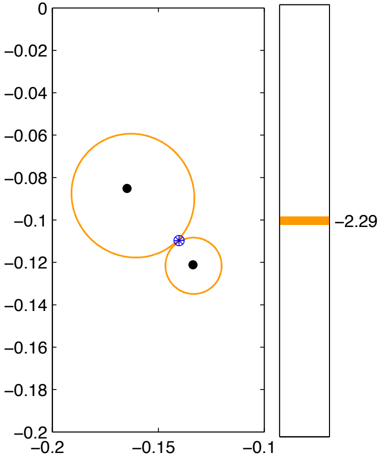
$$\mathcal{C}(A) := \inf\{\epsilon : \#comp\,(\Lambda_\epsilon(A)) \leq n-1\}$$

- It was conjectured by Demmel (1983), and later proven by Alam and Bora (2005) that
  - $\mathcal{W}(A) = \mathcal{C}(A)$.
  - Furthermore 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.

Geometric View of Wilkinson Distance

Orr-Sommerfeld matrix

- $\mathcal{W}(A) = \mathcal{C}(A) = 10^{-2.29}$
- $\lambda_* = -0.1402 - 0.1097i$ (point of coalescence marked with asterisk) is the multiple eigenvalue of a nearest matrix



Generalized Wilkinson distance

Definition (Generalized Wilkinson Distance)

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

The singular value characterization (M. 2011)

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

Problems

1. Geometric interpretation of $\mathcal{W}_r(A)$ in terms of pseudospectra
2. Optimizing symmetric eigenvalues numerically (e.g. computation of $\mathcal{W}_r(A)$)

Generalized Wilkinson distance

Definition (Generalized Wilkinson Distance)

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

The singular value characterization (M. 2011)

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

Problems

1. Geometric interpretation of $\mathcal{W}_r(A)$ in terms of pseudospectra
2. Optimizing symmetric eigenvalues numerically (e.g. computation of $\mathcal{W}_r(A)$)

Generalized Wilkinson distance

Definition (Generalized Wilkinson Distance)

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

The singular value characterization (M. 2011)

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

Problems

1. Geometric interpretation of $\mathcal{W}_r(A)$ in terms of pseudospectra
2. Optimizing symmetric eigenvalues numerically (e.g. computation of $\mathcal{W}_r(A)$)

Outline

Guess in terms of Pseudospectra

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

- $\mathcal{W}_2(A) = \mathcal{C}_2(A)$
- Is $\mathcal{W}_r(A) = \mathcal{C}_r(A)$ for $r > 2$? Turns out not true in general.

Guess in terms of Pseudospectra

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

- $\mathcal{W}_2(A) = \mathcal{C}_2(A)$
- Is $\mathcal{W}_r(A) = \mathcal{C}_r(A)$ for $r > 2$? Turns out not true in general.

Guess in terms of Pseudospectra

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

- $\mathcal{W}_2(A) = \mathcal{C}_2(A)$
- Is $\mathcal{W}_r(A) = \mathcal{C}_r(A)$ for $r > 2$? Turns out not true in general.

Guess in terms of Pseudospectra

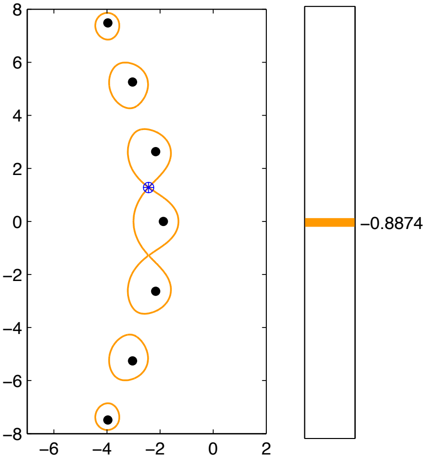
$$\mathcal{C}_r(A) := \inf\{\epsilon : \# comp\left(\Lambda_\epsilon(A)\right) \leq n - r + 1\}$$

- $\mathcal{W}_2(A) = \mathcal{C}_2(A)$
- Is $\mathcal{W}_r(A) = \mathcal{C}_r(A)$ for $r > 2$? Turns out not true in general.

Guess in terms of Pseudospectra

Matrix resulting from a discretization of the convection-diffusion operator

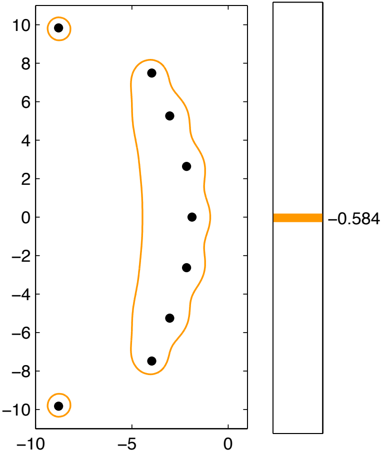
- $\mathcal{W}_2(A) = \mathcal{C}_2(A) = 10^{-0.887}$
- $\lambda_* = -2.4326 + 1.2803i$ (point of coalescence marked with asterisk) is the multiple eigenvalue of a nearest matrix



Guess in terms of Pseudospectra

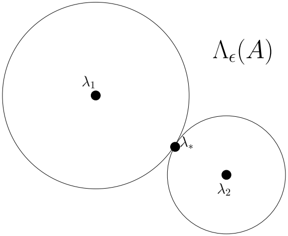
Matrix resulting from a discretization of the convection-diffusion operator

- $\mathcal{W}_3(A) = 10^{-0.584} > \mathcal{C}_3(A)$
- $\Lambda_\epsilon(A)$ is illustrated for $\epsilon = 10^{-0.584}$.



Guess in terms of Pseudospectra

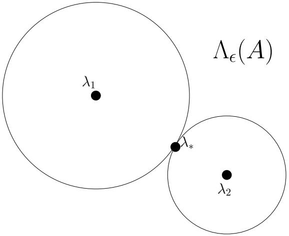
$\mathcal{W}_2(A) \leq \mathcal{C}_2(A)$



There exists a perturbation ΔA_* of norm ϵ s.t.
$\lambda_1(A + \Delta A_*) = \lambda_2(A + \Delta A_*) = \lambda_*$

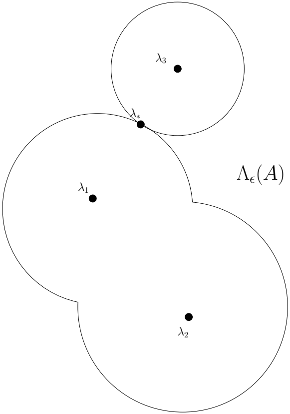
Guess in terms of Pseudospectra

$\mathcal{W}_2(A) \leq \mathcal{C}_2(A)$



There exists a perturbation ΔA_* of norm ϵ s.t. $\lambda_1(A + \Delta A_*) = \lambda_2(A + \Delta A_*) = \lambda_*$

$\mathcal{W}_r(A) \not\leq \mathcal{C}_r(A)$ for $r > 2$



There doesn't exist a perturbation ΔA_* of norm ϵ s.t. $\lambda_1(A + \Delta A_*) = \lambda_2(A + \Delta A_*) = \lambda_3(A + \Delta A_*) = \lambda_*$

Outline

A Special Singular Value Function

A special singular value function

For $r > 2$ the function

$$g_{(r)}(\lambda) := \sup_{\gamma} \sigma_{nr-r+1}\Bigg(\underbrace{\begin{bmatrix} A-\lambda I & \gamma_{1,2} I & \gamma_{1,3} I & \dots & \gamma_{1,r} I \\ 0 & A-\lambda I & \gamma_{2,3} I & & \vdots \\ 0 & 0 & \ddots & & \\ & & & A-\lambda I & \gamma_{r-1,r} I \\ & & & 0 & A-\lambda I \end{bmatrix}}_{\mathcal{A}(\lambda,\gamma)\in\mathbb{C}^{rn\times rn} \;:=}\Bigg)$$

takes the role of $g(\lambda) = \sigma_n(A - \lambda I)$.

Note : $g_{(r)}(\lambda)$ is the distance from A to the nearest matrix with λ as an eigenvalue with algebraic multiplicity $\geq r$.

A Special Singular Value Function

A special singular value function

For $r > 2$ the function

$$g_{(r)}(\lambda) := \sup_{\gamma} \sigma_{nr-r+1}\left(\underbrace{\begin{bmatrix} A-\lambda I & \gamma_{1,2} I & \gamma_{1,3} I & \dots & \gamma_{1,r} I \\ 0 & A-\lambda I & \gamma_{2,3} I & & \vdots \\ 0 & 0 & \ddots & & \\ & & & A-\lambda I & \gamma_{r-1,r} I \\ & & & 0 & A-\lambda I \end{bmatrix}}_{\mathcal{A}(\lambda,\gamma)\in\mathbb{C}^{rn\times rn} \quad :=} \right)$$

takes the role of $g(\lambda) = \sigma_n(A - \lambda I)$.

Note : $g_{(r)}(\lambda)$ is the distance from A to the nearest matrix with λ as an eigenvalue with algebraic multiplicity $\geq r$.

A Special Singular Value Function

Theorem (Alam and Bora)

Let $\lambda_* \in \mathbb{C}$ be a critical point of $g(\lambda) = \sigma_n(A - \lambda I)$ such that

(i) $g(\lambda_*) = \epsilon > 0$, *and*

(ii) *the multiplicity of $\sigma_n(A - \lambda_* I)$ is one.*

There exists a rank one perturbation δA with norm ϵ such that $A + \delta A$ has λ_* as a (defective) multiple eigenvalue.

A Special Singular Value Function

Theorem

Let $\lambda_* \in \mathbb{C}$ be a critical point of $g_{(r-1)}(\lambda)$ such that

(i) $g_{(r-1)}(\lambda_*) = \epsilon > 0$*, and*

(ii) *the multiplicity of the singular value $g_{(r-1)}(\lambda_*)$ is one.*

There exists a rank $(r-1)$ perturbation δA with norm ϵ such that $A + \delta A$ has λ_* as a (defective) eigenvalue with algebraic multiplicity $\geq r$.

Generalized Pseudospectra

Definition (Generalized (ϵ, r)-pseudospectrum)

$$\Lambda_{\epsilon,r}(A) = \bigcup_{\|E\|_2 \leq \epsilon} \{\lambda \in \mathbb{C} : \mathrm{rank}(A + E - \lambda I)^r \leq n - r\}$$

$$= \{\lambda \in \mathbb{C} : g_{(r)}(\lambda) \leq \epsilon\}$$

Examples

- $\Lambda_{\epsilon,1}(A) = \Lambda_\epsilon(A) = \{\lambda \in \mathbb{C} : \sigma_n(A - \lambda I) \leq \epsilon\}$
- $\Lambda_{\epsilon,2}(A) = \left\{\lambda \in \mathbb{C} : \sup_\gamma \sigma_{2n-1}\left(\begin{bmatrix} A - \lambda I & \gamma I \\ 0 & A - \lambda I \end{bmatrix}\right) \leq \epsilon\right\}$

Generalized Pseudospectra

Definition (Generalized (ϵ, r)-pseudospectrum)

$$\Lambda_{\epsilon,r}(A) = \bigcup_{\|E\|_2 \le \epsilon} \{\lambda \in \mathbb{C} : \mathrm{rank}(A + E - \lambda I)^r \le n - r\}$$
$$= \{\lambda \in \mathbb{C} : g_{(r)}(\lambda) \le \epsilon\}$$

Examples

- $\Lambda_{\epsilon,1}(A) = \Lambda_\epsilon(A) = \{\lambda \in \mathbb{C} : \sigma_n(A - \lambda I) \le \epsilon\}$
- $\Lambda_{\epsilon,2}(A) = \left\{\lambda \in \mathbb{C} : \sup_\gamma \sigma_{2n-1}\left(\begin{bmatrix} A - \lambda I & \gamma I \\ 0 & A - \lambda I \end{bmatrix}\right) \le \epsilon\right\}$

Generalized Pseudospectra

Definition (Generalized (ϵ, r)-pseudospectrum)

$$\Lambda_{\epsilon,r}(A) = \bigcup_{\|E\|_2 \leq \epsilon} \{\lambda \in \mathbb{C} : \mathrm{rank}(A + E - \lambda I)^r \leq n - r\}$$
$$= \{\lambda \in \mathbb{C} : g_{(r)}(\lambda) \leq \epsilon\}$$

Examples

- $\Lambda_{\epsilon,1}(A) = \Lambda_{\epsilon}(A) = \{\lambda \in \mathbb{C} : \sigma_n(A - \lambda I) \leq \epsilon\}$
- $\Lambda_{\epsilon,2}(A) = \left\{\lambda \in \mathbb{C} : \sup_{\gamma} \sigma_{2n-1}\left(\begin{bmatrix} A - \lambda I & \gamma I \\ 0 & A - \lambda I \end{bmatrix}\right) \leq \epsilon\right\}$

Generalized Pseudospectra Characterization

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

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

Conjecture

$\mathcal{W}_r(A) = \mathcal{G}_r(A)$

Generalized Pseudospectra Characterization

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

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

Conjecture

$\mathcal{W}_r(A) = \mathcal{G}_r(A)$

Generalized Pseudospectra Characterization

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

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

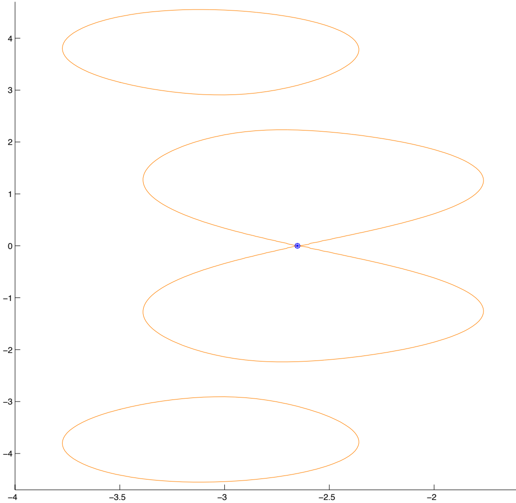
Conjecture

$$\mathcal{W}_r(A) = \mathcal{G}_r(A)$$

Generalized Pseudospectra Characterization

Matrix resulting from a discretization of the convection-diffusion operator

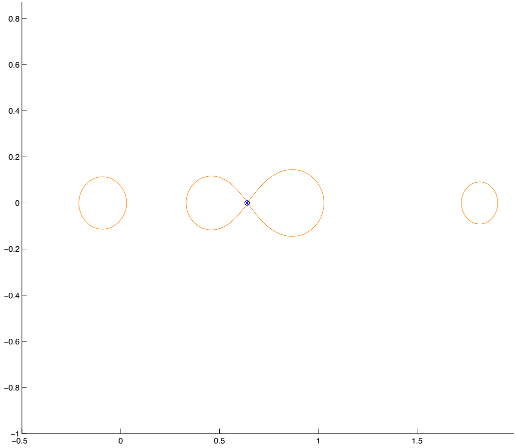
- $\mathcal{W}_3(A) = 10^{-0.584} = \mathcal{G}_3(A)$
- $\Lambda_{\epsilon,2}(A)$ is illustrated for $\epsilon = 10^{-0.584}$.
- $\lambda_* = -2.6520$ (point of coalescence marked with asterisk) is the triple eigenvalue of a nearest matrix



Generalized Pseudospectra Characterization

Hatano matrix

- $\mathcal{W}_3(A) = 10^{-1.1973} = \mathcal{G}_3(A)$
- $\Lambda_{\epsilon,2}(A)$ is illustrated for $\epsilon = 10^{-1.1973}$.
- $\lambda_* = 0.6421$ (point of coalescence marked with asterisk) is the triple eigenvalue of a nearest matrix



Proof of $\mathcal{W}_r(A) \leq \mathcal{G}_r(A)$

Theorem

Let $\lambda_* \in \mathbb{C}$ be a critical point of $g_{(r-1)}(\lambda)$ such that

(i) $g_{(r-1)}(\lambda_*) = \epsilon > 0$, *and*

(ii) *the multiplicity of the singular value $g_{(r-1)}(\lambda_*)$ is one.*

There exists a rank $(r-1)$ perturbation δA with norm ϵ such that $A + \delta A$ has λ_* as a (defective) eigenvalue with algebraic multiplicity $\geq r$.

Suppose two components of $\Lambda_{\epsilon,r-1}(A)$ coalesce at λ_* for $\epsilon = \mathcal{G}_r(A)$.

- Then $g_{(r-1)}(\lambda_*) = \mathcal{G}_r(A)$ and λ_* is a critical point of $g_{(r-1)}(\lambda)$.
- If the multiplicity of $g_{(r-1)}(\lambda_*) = \mathcal{G}_r(A)$ is one, the theorem above implies the existence of a perturbation ΔA_* of norm $\mathcal{G}_r(A)$ such that $(A + \Delta A_*)$ has λ_* as an eigenvalue with algebraic multiplicity $\geq r$.

Therefore $\mathcal{G}_r(A) \geq \mathcal{W}_r(A)$.

Proof of $\mathcal{W}_r(A) \leq \mathcal{G}_r(A)$

Theorem

Let $\lambda_* \in \mathbb{C}$ be a critical point of $g_{(r-1)}(\lambda)$ such that

(i) $g_{(r-1)}(\lambda_*) = \epsilon > 0$, *and*

(ii) *the multiplicity of the singular value $g_{(r-1)}(\lambda_*)$ is one.*

There exists a rank $(r-1)$ perturbation δA with norm ϵ such that $A + \delta A$ has λ_* as a (defective) eigenvalue with algebraic multiplicity $\geq r$.

Suppose two components of $\Lambda_{\epsilon,r-1}(A)$ coalesce at λ_* for $\epsilon = \mathcal{G}_r(A)$.

- Then $g_{(r-1)}(\lambda_*) = \mathcal{G}_r(A)$ and λ_* is a critical point of $g_{(r-1)}(\lambda)$.
- If the multiplicity of $g_{(r-1)}(\lambda_*) = \mathcal{G}_r(A)$ is one, the theorem above implies the existence of a perturbation ΔA_* of norm $\mathcal{G}_r(A)$ such that $(A + \Delta A_*)$ has λ_* as an eigenvalue with algebraic multiplicity $\geq r$.

Therefore $\mathcal{G}_r(A) \geq \mathcal{W}_r(A)$.

Proof of $\mathcal{W}_r(A) \leq \mathcal{G}_r(A)$

Theorem

Let $\lambda_* \in \mathbb{C}$ be a critical point of $g_{(r-1)}(\lambda)$ such that

(i) $g_{(r-1)}(\lambda_*) = \epsilon > 0$, *and*

(ii) *the multiplicity of the singular value $g_{(r-1)}(\lambda_*)$ is one.*

There exists a rank $(r-1)$ perturbation δA with norm ϵ such that $A + \delta A$ has λ_* as a (defective) eigenvalue with algebraic multiplicity $\geq r$.

Suppose two components of $\Lambda_{\epsilon, r-1}(A)$ coalesce at λ_* for $\epsilon = \mathcal{G}_r(A)$.

- Then $g_{(r-1)}(\lambda_*) = \mathcal{G}_r(A)$ and λ_* is a critical point of $g_{(r-1)}(\lambda)$.
- If the multiplicity of $g_{(r-1)}(\lambda_*) = \mathcal{G}_r(A)$ is one, the theorem above implies the existence of a perturbation ΔA_* of norm $\mathcal{G}_r(A)$ such that $(A + \Delta A_*)$ has λ_* as an eigenvalue with algebraic multiplicity $\geq r$.

Therefore $\mathcal{G}_r(A) \geq \mathcal{W}_r(A)$.

Proof of $\mathcal{W}_r(A) \leq \mathcal{G}_r(A)$

Theorem

Let $\lambda_* \in \mathbb{C}$ be a critical point of $g_{(r-1)}(\lambda)$ such that

(i) $g_{(r-1)}(\lambda_*) = \epsilon > 0$, *and*

(ii) *the multiplicity of the singular value $g_{(r-1)}(\lambda_*)$ is one.*

There exists a rank $(r-1)$ perturbation δA with norm ϵ such that $A + \delta A$ has λ_* as a (defective) eigenvalue with algebraic multiplicity $\geq r$.

Suppose two components of $\Lambda_{\epsilon,r-1}(A)$ coalesce at λ_* for $\epsilon = \mathcal{G}_r(A)$.

- Then $g_{(r-1)}(\lambda_*) = \mathcal{G}_r(A)$ and λ_* is a critical point of $g_{(r-1)}(\lambda)$.
- If the multiplicity of $g_{(r-1)}(\lambda_*) = \mathcal{G}_r(A)$ is one, the theorem above implies the existence of a perturbation ΔA_* of norm $\mathcal{G}_r(A)$ such that $(A + \Delta A_*)$ has λ_* as an eigenvalue with algebraic multiplicity $\geq r$.

Therefore $\mathcal{G}_r(A) \geq \mathcal{W}_r(A)$.

Proof of $\mathcal{W}_r(A) \leq \mathcal{G}_r(A)$

Theorem

Let $\lambda_* \in \mathbb{C}$ be a critical point of $g_{(r-1)}(\lambda)$ such that

(i) $g_{(r-1)}(\lambda_*) = \epsilon > 0$, *and*

(ii) *the multiplicity of the singular value $g_{(r-1)}(\lambda_*)$ is one.*

There exists a rank $(r-1)$ perturbation δA with norm ϵ such that $A + \delta A$ has λ_* as a (defective) eigenvalue with algebraic multiplicity $\geq r$.

Suppose two components of $\Lambda_{\epsilon,r-1}(A)$ coalesce at λ_* for $\epsilon = \mathcal{G}_r(A)$.

- Then $g_{(r-1)}(\lambda_*) = \mathcal{G}_r(A)$ and λ_* is a critical point of $g_{(r-1)}(\lambda)$.
- If the multiplicity of $g_{(r-1)}(\lambda_*) = \mathcal{G}_r(A)$ is one, the theorem above implies the existence of a perturbation ΔA_* of norm $\mathcal{G}_r(A)$ such that $(A + \Delta A_*)$ has λ_* as an eigenvalue with algebraic multiplicity $\geq r$.

Therefore $\mathcal{G}_r(A) \geq \mathcal{W}_r(A)$.

Outline

Analyticity Result

Theorem (Rellich)

Let $\mathcal{A}(\omega) : \mathbb{R} \rightarrow \mathbb{C}^{n \times n}$ be a Hermitian matrix function that depends on ω analytically. Each root of the characteristic polynomial of $\mathcal{A}(\omega)$ is an analytic function of ω.

- The eigenvalues $\lambda_1(\omega), \ldots, \lambda_n(\omega)$ ordered from largest to smallest of $\mathcal{A}(\omega)$ are piece-wise analytic.
- The result does not extend to non-Hermitian functions. e.g. the roots of the characteristic polynomial of

$$\mathcal{A}(\omega) = \begin{bmatrix} 0 & 1 \\ \omega & 0 \end{bmatrix}$$

(given by $\pm\sqrt{\omega}$) are not analytic.

Analyticity Result

Theorem (Rellich)

Let $\mathcal{A}(\omega) : \mathbb{R} \to \mathbb{C}^{n \times n}$ be a Hermitian matrix function that depends on ω analytically. Each root of the characteristic polynomial of $\mathcal{A}(\omega)$ is an analytic function of ω.

- The eigenvalues $\lambda_1(\omega), \dots, \lambda_n(\omega)$ ordered from largest to smallest of $\mathcal{A}(\omega)$ are piece-wise analytic.
- The result does not extend to non-Hermitian functions. e.g. the roots of the characteristic polynomial of

$$\mathcal{A}(\omega) = \begin{bmatrix} 0 & 1 \\ \omega & 0 \end{bmatrix}$$

(given by $\pm\sqrt{\omega}$) are not analytic.

Analyticity Result

Theorem (Rellich)

Let $\mathcal{A}(\omega) : \mathbb{R} \to \mathbb{C}^{n \times n}$ be a Hermitian matrix function that depends on ω analytically. Each root of the characteristic polynomial of $\mathcal{A}(\omega)$ is an analytic function of ω.

- The eigenvalues $\lambda_1(\omega), \dots, \lambda_n(\omega)$ ordered from largest to smallest of $\mathcal{A}(\omega)$ are piece-wise analytic.
- The result does not extend to non-Hermitian functions.

  e.g. the roots of the characteristic polynomial of

$$\mathcal{A}(\omega) = \begin{bmatrix} 0 & 1 \\ \omega & 0 \end{bmatrix}$$

  (given by $\pm\sqrt{\omega}$) are not analytic.

Analyticity Result

Theorem (Rellich)

Let $\mathcal{A}(\omega) : \mathbb{R} \to \mathbb{C}^{n \times n}$ be a Hermitian matrix function that depends on ω analytically. Each root of the characteristic polynomial of $\mathcal{A}(\omega)$ is an analytic function of ω.

- The eigenvalues $\lambda_1(\omega), \ldots, \lambda_n(\omega)$ ordered from largest to smallest of $\mathcal{A}(\omega)$ are piece-wise analytic.
- The result does not extend to non-Hermitian functions. e.g. the roots of the characteristic polynomial of

$$\mathcal{A}(\omega) = \begin{bmatrix} 0 & 1 \\ \omega & 0 \end{bmatrix}$$

(given by $\pm\sqrt{\omega}$) are not analytic.

Derivatives of Eigenvalues

Let $\tilde{\lambda}(\omega)$ be one of the unordered eigenvalues with the assoc. unit eigenvector $\tilde{v}(\omega)$ (which also varies analytically w.r.t ω).

First Derivative

$$\tilde{\lambda}'(\omega) = \tilde{v}^*(\omega)\frac{d\mathcal{A}(\omega)}{d\omega}\tilde{v}(\omega)$$

Second Derivative

$$\tilde{\lambda}''(\omega) = \tilde{v}^*(\omega)\frac{d^2\mathcal{A}(\omega)}{d\omega^2}\tilde{v}(\omega) + 2\tilde{v}^*(\omega)\frac{d\mathcal{A}(\omega)}{d\omega}\left[\tilde{\lambda}(\omega)I - \mathcal{A}(\omega)\right]^+\frac{d\mathcal{A}(\omega)}{d\omega}\tilde{v}(\omega)$$

Derivatives of Eigenvalues

Let $\tilde{\lambda}(\omega)$ be one of the unordered eigenvalues with the assoc. unit eigenvector $\tilde{v}(\omega)$ (which also varies analytically w.r.t ω).

First Derivative

$$\tilde{\lambda}'(\omega) = \tilde{v}^*(\omega)\frac{d\mathcal{A}(\omega)}{d\omega}\tilde{v}(\omega)$$

Second Derivative

$$\tilde{\lambda}''(\omega) = \tilde{v}^*(\omega)\frac{d^2\mathcal{A}(\omega)}{d\omega^2}\tilde{v}(\omega) + 2\tilde{v}^*(\omega)\frac{d\mathcal{A}(\omega)}{d\omega}\left[\tilde{\lambda}(\omega)I - \mathcal{A}(\omega)\right]^+\frac{d\mathcal{A}(\omega)}{d\omega}\tilde{v}(\omega)$$

Derivatives of Eigenvalues

Let $\tilde{\lambda}(\omega)$ be one of the unordered eigenvalues with the assoc. unit eigenvector $\tilde{v}(\omega)$ (which also varies analytically w.r.t ω).

First Derivative

$$\tilde{\lambda}'(\omega) = \tilde{v}^*(\omega)\frac{d\mathcal{A}(\omega)}{d\omega}\tilde{v}(\omega)$$

Second Derivative

$$\tilde{\lambda}''(\omega) = \tilde{v}^*(\omega)\frac{d^2\mathcal{A}(\omega)}{d\omega^2}\tilde{v}(\omega) + 2\tilde{v}^*(\omega)\frac{d\mathcal{A}(\omega)}{d\omega}\left[\tilde{\lambda}(\omega)I - \mathcal{A}(\omega)\right]^+\frac{d\mathcal{A}(\omega)}{d\omega}\tilde{v}(\omega)$$

Derivatives of Eigenvalues

Let $\tilde{\lambda}(\omega)$ be one of the unordered eigenvalues with the assoc. unit eigenvector $\tilde{v}(\omega)$ (which also varies analytically w.r.t ω).

First Derivative

$$\tilde{\lambda}'(\omega) = \tilde{v}^*(\omega)\frac{d\mathcal{A}(\omega)}{d\omega}\tilde{v}(\omega)$$

Second Derivative

$$\tilde{\lambda}''(\omega) = \tilde{v}^*(\omega)\frac{d^2\mathcal{A}(\omega)}{d\omega^2}\tilde{v}(\omega) + 2\sum_{j,\tilde{\lambda}_j(\omega)\neq\tilde{\lambda}(\omega)}\frac{1}{\tilde{\lambda}(\omega)-\tilde{\lambda}_j(\omega)}\left|\tilde{v}^*(\omega)\frac{d\mathcal{A}(\omega)}{d\omega}\tilde{v}_j(\omega)\right|^2$$

Derivatives of Eigenvalues

Some observations helpful algorithmically

- Analyticity implies the boundedness of derivatives. In particular we will exploit the existence of a γ such that

$$\left|\tilde{\lambda}''(\omega)\right| \leq \gamma \quad \forall \omega.$$

- Once $\tilde{\lambda}(\omega)$ is computed, $\tilde{\lambda}'(\omega)$ is available for free.

Derivatives of Eigenvalues

Some observations helpful algorithmically

- Analyticity implies the boundedness of derivatives. In particular we will exploit the existence of a γ such that

$$\left| \tilde{\lambda}''(\omega) \right| \leq \gamma \quad \forall \omega.$$

- Once $\tilde{\lambda}(\omega)$ is computed, $\tilde{\lambda}'(\omega)$ is available for free.

Derivatives of Eigenvalues

Some observations helpful algorithmically

- Analyticity implies the boundedness of derivatives. In particular we will exploit the existence of a γ such that

$$\left|\tilde{\lambda}''(\omega)\right| \leq \gamma \quad \forall \omega.$$

- Once $\tilde{\lambda}(\omega)$ is computed, $\tilde{\lambda}'(\omega)$ is available for free.

Outline

Quadratic Models

Some notation

- Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a piece-wise analytic function defined in terms of analytic functions $f_1, \dots, f_n : \mathbb{R} \rightarrow \mathbb{R}$,
- γ be an upper bound on the second derivatives (in absolute value) of f_j for $j = 1, \dots, n$,
- $x_k, x \in \mathbb{R}$ and $x_{k,1}, \dots, x_{k,m}$ be points in (x_k, x) where f is not analytic, and
- $\underline{f}'(x_k) := \min_{j=1,n} f_j'(x_k)$.

Quadratic Models

Some notation

- Let $f : \mathbb{R} \to \mathbb{R}$ be a piece-wise analytic function defined in terms of analytic functions $f_1, \dots, f_n : \mathbb{R} \to \mathbb{R}$,
- γ be an upper bound on the second derivatives (in absolute value) of f_j for $j = 1, \dots, n$,
- $x_k, x \in \mathbb{R}$ and $x_{k,1}, \dots, x_{k,m}$ be points in (x_k, x) where f is not analytic, and
- $\underline{f}'(x_k) := \min_{j=1,n} \; f_j'(x_k)$.

Quadratic Models

Some notation

- Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a piece-wise analytic function defined in terms of analytic functions $f_1, \ldots, f_n : \mathbb{R} \rightarrow \mathbb{R}$,
- γ be an upper bound on the second derivatives (in absolute value) of f_j for $j = 1, \ldots, n$,
- $x_k, x \in \mathbb{R}$ and $x_{k,1}, \ldots, x_{k,m}$ be points in (x_k, x) where f is not analytic, and
- $\underline{f}'(x_k) := \min_{j=1,n} f_j'(x_k)$.

Quadratic Models

Some notation

- Let $f : \mathbb{R} \to \mathbb{R}$ be a piece-wise analytic function defined in terms of analytic functions $f_1, \dots, f_n : \mathbb{R} \to \mathbb{R}$,
- γ be an upper bound on the second derivatives (in absolute value) of f_j for $j = 1, \dots, n$,
- $x_k, x \in \mathbb{R}$ and $x_{k,1}, \dots, x_{k,m}$ be points in (x_k, x) where f is not analytic, and
- $\underline{f}'(x_k) := \min_{j=1,n} f_j'(x_k)$.

Quadratic Models

Some notation

- Let $f : \mathbb{R} \to \mathbb{R}$ be a piece-wise analytic function defined in terms of analytic functions $f_1, \dots, f_n : \mathbb{R} \to \mathbb{R}$,
- γ be an upper bound on the second derivatives (in absolute value) of f_j for $j = 1, \dots, n$,
- $x_k, x \in \mathbb{R}$ and $x_{k,1}, \dots, x_{k,m}$ be points in (x_k, x) where f is not analytic, and
- $\underline{f}'(x_k) := \min_{j=1,n} \; f_j'(x_k)$.

Quadratic Models

$$f(x) = f(x_k) + \sum_{\ell=0}^{m} \int_{x_{k,\ell}}^{x_{k,\ell+1}} f'(t)dt$$

Note: $x_{k,0} = x_k$ and $x_{k,m+1} = x$

Quadratic Model Function about x_k

$$q_k(x) := f(x_k) + \underline{f}'(x_k)(x - x_k) - \frac{\gamma}{2}(x - x_k)^2$$

satisfies $f(x) \geq q_k(x)$ for all $x \in \mathbb{R}$.

Quadratic Models

$$\begin{aligned} f(x) &= f(x_k) + \sum_{\ell=0}^{m} \int_{x_{k,\ell}}^{x_{k,\ell+1}} f'(t)dt \\ &\geq f(x_k) + \sum_{\ell=0}^{m} \int_{x_{k,\ell}}^{x_{k,\ell+1}} \underline{f}'(x_k) - \gamma(t - x_k)dt \end{aligned}$$

Note: $f'(t) \geq \underline{f}'(x_k) - \gamma(t - x_k) \;\; \forall t \in (x_k, x) \setminus \{x_{k,1}, \ldots, x_{k,m}\}$

Quadratic Model Function about x_k

$$q_k(x) := f(x_k) + \underline{f}'(x_k)(x - x_k) - \frac{\gamma}{2}(x - x_k)^2$$

satisfies $f(x) \geq q_k(x)$ for all $x \in \mathbb{R}$.

Quadratic Models

$$\begin{aligned} f(x) &= f(x_k) + \sum_{\ell=0}^{m} \int_{x_{k,\ell}}^{x_{k,\ell+1}} f'(t)dt \\ &\geq f(x_k) + \int_{x_k}^{x} \underline{f}'(x_k) - \gamma(t - x_k)dt \end{aligned}$$

Quadratic Model Function about x_k

$$q_k(x) := f(x_k) + \underline{f}'(x_k)(x - x_k) - \frac{\gamma}{2}(x - x_k)^2$$

satisfies $f(x) \geq q_k(x)$ for all $x \in \mathbb{R}$.

Quadratic Models

$$
\begin{aligned}
f(x) &= f(x_k) + \sum_{\ell=0}^{m} \int_{x_{k,\ell}}^{x_{k,\ell+1}} f'(t)dt \\
&\geq f(x_k) + \int_{x_k}^{x} \underline{f}'(x_k) - \gamma(t - x_k)dt \\
&= f(x_k) + \underline{f}'(x_k)(x - x_k) - \frac{\gamma}{2}(x - x_k)^2
\end{aligned}
$$

Quadratic Model Function about x_k

$$q_k(x) := f(x_k) + \underline{f}'(x_k)(x - x_k) - \frac{\gamma}{2}(x - x_k)^2$$

satisfies $f(x) \geq q_k(x)$ for all $x \in \mathbb{R}$.

Quadratic Models

$$\begin{aligned} f(x) &= f(x_k) + \sum_{\ell=0}^{m} \int_{x_{k,\ell}}^{x_{k,\ell+1}} f'(t)dt \\ &\geq f(x_k) + \int_{x_k}^{x} \underline{f}'(x_k) - \gamma(t - x_k)dt \\ &= f(x_k) + \underline{f}'(x_k)(x - x_k) - \frac{\gamma}{2}(x - x_k)^2 \end{aligned}$$

Quadratic Model Function about x_k

$$q_k(x) := f(x_k) + \underline{f}'(x_k)(x - x_k) - \frac{\gamma}{2}(x - x_k)^2$$

satisfies $f(x) \geq q_k(x)$ for all $x \in \mathbb{R}$.

The Algorithm

Suppose that the global minimizer of f is in $[a, b]$.

1. Initially $x_0 = a$, $x_1 = b$ and $s = 1$. Evaluate $f(x_0)$, $f(x_1)$, $f'(x_0)$, and $f'(x_1)$.
2. Find the global minimizer of x_* of $q(x)$ where

$$q(x) = \max_{k=0,s} q_k(x).$$

3. Set $x_{s+1} = x_*$, evaluate $f(x_{s+1})$, $f'(x_{s+1})$.
4. Let $\ell = q(x_*)$ and $u = max_{k=0,s+1} f(x_k)$.
5. While $u - l > \epsilon$, increment s and repeat steps 2-4.

The Algorithm

Suppose that the global minimizer of f is in $[a, b]$.

1. Initially $x_0 = a$, $x_1 = b$ and $s = 1$. Evaluate $f(x_0)$, $f(x_1)$, $f'(x_0)$, and $f'(x_1)$.
2. Find the global minimizer of x_* of $q(x)$ where

$$q(x) = \max_{k=0,s} q_k(x).$$

3. Set $x_{s+1} = x_*$, evaluate $f(x_{s+1})$, $f'(x_{s+1})$.
4. Let $\ell = q(x_*)$ and $u = max_{k=0,s+1} f(x_k)$.
5. While $u - l > \epsilon$, increment s and repeat steps 2-4.

The Algorithm

Suppose that the global minimizer of f is in $[a, b]$.

1. Initially $x_0 = a$, $x_1 = b$ and $s = 1$. Evaluate $f(x_0)$, $f(x_1)$, $f'(x_0)$, and $f'(x_1)$.
2. Find the global minimizer of x_* of $q(x)$ where

$$q(x) = \max_{k=0,s} q_k(x).$$

3. Set $x_{s+1} = x_*$, evaluate $f(x_{s+1})$, $f'(x_{s+1})$.
4. Let $\ell = q(x_*)$ and $u = max_{k=0,s+1} f(x_k)$.
5. While $u - l > \epsilon$, increment s and repeat steps 2-4.

The Algorithm

Suppose that the global minimizer of f is in $[a, b]$.

1. Initially $x_0 = a$, $x_1 = b$ and $s = 1$. Evaluate $f(x_0)$, $f(x_1)$, $f'(x_0)$, and $f'(x_1)$.
2. Find the global minimizer of x_* of $q(x)$ where

$$q(x) = \max_{k=0,s} q_k(x).$$

3. Set $x_{s+1} = x_*$, evaluate $f(x_{s+1})$, $f'(x_{s+1})$.
4. Let $\ell = q(x_*)$ and $u = max_{k=0,s+1} f(x_k)$.
5. While $u - l > \epsilon$, increment s and repeat steps 2-4.

The Algorithm

Suppose that the global minimizer of f is in $[a, b]$.

1. Initially $x_0 = a$, $x_1 = b$ and $s = 1$. Evaluate $f(x_0)$, $f(x_1)$, $f'(x_0)$, and $f'(x_1)$.
2. Find the global minimizer of x_* of $q(x)$ where

$$q(x) = \max_{k=0,s} q_k(x).$$

3. Set $x_{s+1} = x_*$, evaluate $f(x_{s+1})$, $f'(x_{s+1})$.
4. Let $\ell = q(x_*)$ and $u = max_{k=0,s+1} f(x_k)$.
5. While $u - l > \epsilon$, increment s and repeat steps 2-4.

The Algorithm

Suppose that the global minimizer of f is in $[a, b]$.

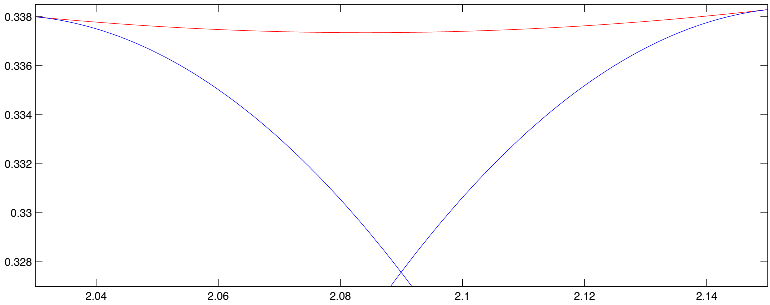
1. Initially $x_0 = a$, $x_1 = b$ and $s = 1$. Evaluate $f(x_0)$, $f(x_1)$, $f'(x_0)$, and $f'(x_1)$.
2. Find the global minimizer of x_* of $q(x)$ where

$$q(x) = \max_{k=0,s} q_k(x).$$

3. Set $x_{s+1} = x_*$, evaluate $f(x_{s+1})$, $f'(x_{s+1})$.
4. Let $\ell = q(x_*)$ and $u = max_{k=0,s+1} f(x_k)$.
5. While $u - l > \epsilon$, increment s and repeat steps 2-4.

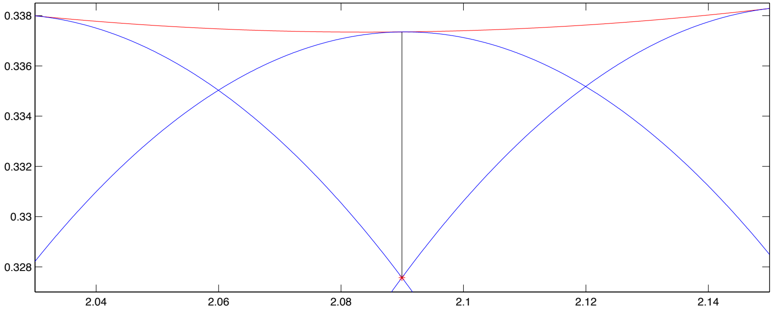
The Algorithm

Illustration of the algorithm on $\sigma_n(A - \omega i I)$ where σ_n denotes the smallest singular value.



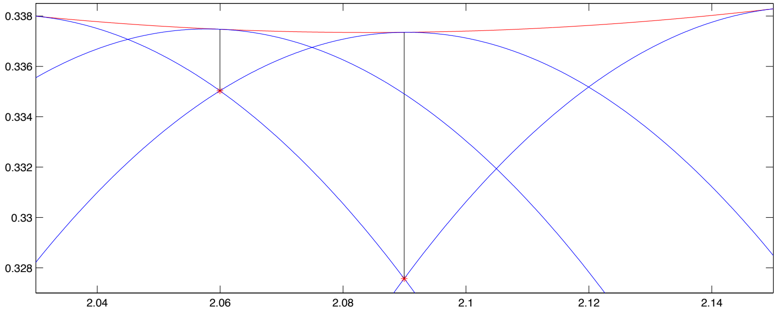
The Algorithm

Illustration of the algorithm on $\sigma_n(A - \omega iI)$ where σ_n denotes the smallest singular value.



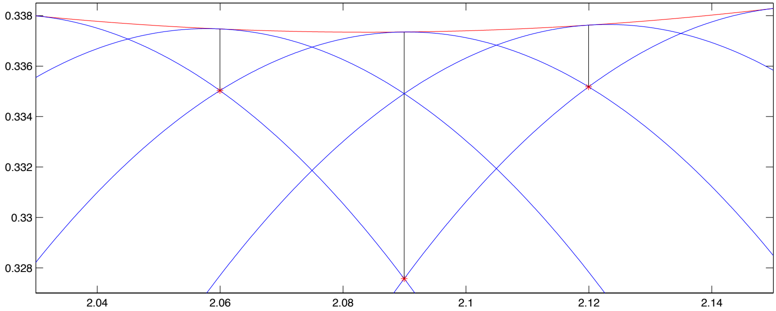
The Algorithm

Illustration of the algorithm on $\sigma_n(A - \omega iI)$ where σ_n denotes the smallest singular value.



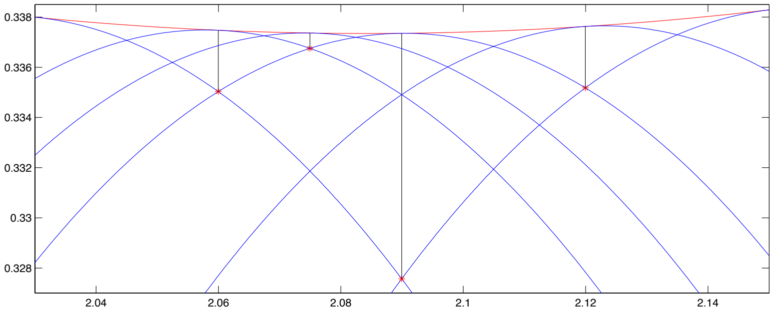
The Algorithm

Illustration of the algorithm on $\sigma_n(A - \omega iI)$ where σ_n denotes the smallest singular value.



The Algorithm

Illustration of the algorithm on $\sigma_n(A - \omega iI)$ where σ_n denotes the smallest singular value.



Case Study

Distance to Instability

$$\begin{aligned}\beta(A) &:= \inf\{\|\Delta A\|_2 : x'(t) = (A+\Delta A)x(t) \text{ is unstable}\} \\ &= \inf_{\omega\in\mathbb{R}} \sigma_n(A - \omega i I)\end{aligned}$$

Matrices result from a discretization of the Airy operator

# of function evaluations

n / ϵ	10^{-3}	10^{-5}	10^{-7}	10^{-9}	10^{-11}
25	40	48	56	62	68
100	46	54	65	74	82
400	46	54	65	74	81
1600	46	54	64	68	69

Case Study

Distance to Instability

$$\begin{aligned} \beta(A) &:= \inf\{\|\Delta A\|_2 : x'(t) = (A + \Delta A)x(t) \text{ is unstable}\} \\ &= \inf_{\omega \in \mathbb{R}} \sigma_n(A - \omega i I) \end{aligned}$$

Matrices result from a discretization of the Airy operator

# of function evaluations

n / ϵ	10^{-3}	10^{-5}	10^{-7}	10^{-9}	10^{-11}
25	40	48	56	62	68
100	46	54	65	74	82
400	46	54	65	74	81
1600	46	54	64	68	69

Case Study

Distance to Instability

$$\begin{aligned} \beta(A) \quad &:= \quad \inf\{\|\Delta A\|_2 : x'(t) = (A + \Delta A)x(t) \ \text{ is } \ \text{unstable}\} \\ &= \quad \inf_{\omega \in \mathbb{R}} \sigma_n(A - \omega i I) \end{aligned}$$

Matrices result from a discretization of the Airy operator

cpu-times

n / ϵ	10^{-3}	10^{-5}	10^{-7}	10^{-9}	10^{-11}
25	0.06	0.09	0.09	0.10	0.12
100	0.87	0.96	1.13	1.27	1.41
400	13.53	15.73	18.82	21.02	23.26
1600	362	409	474	505	506

Outline

Non-analyticity Result

- For a multivariate Hermitian function $\mathcal{A}(\omega) : \mathbb{R}^n \rightarrow \mathbb{C}^{n \times n}$ that depend on ω analytically an unordered eigenvalue $\tilde{\lambda}(\omega)$ is not analytic in general.

  e.g. The roots of the characteristic polynomial of

$$\mathcal{A}(\omega) = \begin{bmatrix} \omega_1 & \frac{\omega_1+\omega_2}{2} \\ \frac{\omega_1+\omega_2}{2} & \omega_2 \end{bmatrix}$$

  (given by $\omega_1 + \omega_2 \pm \sqrt{2}\sqrt{\omega_1^2 + \omega_2^2}$) are not analytic.

- But $\tilde{\lambda}(\omega)$ is analytic over any line in $\mathbb{R}^n$ (due to Rellich's result).

Non-analyticity Result

- For a multivariate Hermitian function $\mathcal{A}(\omega) : \mathbb{R}^n \to \mathbb{C}^{n \times n}$ that depend on ω analytically an unordered eigenvalue $\tilde{\lambda}(\omega)$ is not analytic in general.

  e.g. The roots of the characteristic polynomial of

$$\mathcal{A}(\omega) = \begin{bmatrix} \omega_1 & \frac{\omega_1+\omega_2}{2} \\ \frac{\omega_1+\omega_2}{2} & \omega_2 \end{bmatrix}$$

  (given by $\omega_1 + \omega_2 \pm \sqrt{2}\sqrt{\omega_1^2 + \omega_2^2}$) are not analytic.

- But $\tilde{\lambda}(\omega)$ is analytic over any line in $\mathbb{R}^n$ (due to Rellich's result).

Non-analyticity Result

- For a multivariate Hermitian function $\mathcal{A}(\omega) : \mathbb{R}^n \to \mathbb{C}^{n \times n}$ that depend on ω analytically an unordered eigenvalue $\tilde{\lambda}(\omega)$ is not analytic in general.

  e.g. The roots of the characteristic polynomial of

$$\mathcal{A}(\omega) = \begin{bmatrix} \omega_1 & \frac{\omega_1+\omega_2}{2} \\ \frac{\omega_1+\omega_2}{2} & \omega_2 \end{bmatrix}$$

  (given by $\omega_1 + \omega_2 \pm \sqrt{2}\sqrt{\omega_1^2 + \omega_2^2}$) are not analytic.

- But $\tilde{\lambda}(\omega)$ is analytic over any line in $\mathbb{R}^n$ (due to Rellich's result).

Model Functions

- Let $f : \mathbb{R}^n \to \mathbb{R}$ be analytic over any line in $\mathbb{R}^n$, and
- γ be an upper bound on the second derivative (on any line in $\mathbb{R}^n$) of f.

Quadratic Model Function about x_k

$$q_k(x) := f(x_k) + \nabla f(x_k)^T (x - x_k) - \frac{\gamma}{2}(x - x_k)^T (x - x_k)$$

satisfies $f(x) \geq q_k(x)$ for all $x \in \mathbb{R}^n$.

Model Functions

- Let $f : \mathbb{R}^n \to \mathbb{R}$ be analytic over any line in $\mathbb{R}^n$, and
- γ be an upper bound on the second derivative (on any line in $\mathbb{R}^n$) of f.

Quadratic Model Function about x_k

$$q_k(x) := f(x_k) + \nabla f(x_k)^T(x - x_k) - \frac{\gamma}{2}(x - x_k)^T(x - x_k)$$

satisfies $f(x) \geq q_k(x)$ for all $x \in \mathbb{R}^n$.

Model Functions

- Let $f : \mathbb{R}^n \to \mathbb{R}$ be analytic over any line in $\mathbb{R}^n$, and
- γ be an upper bound on the second derivative (on any line in $\mathbb{R}^n$) of f.

Quadratic Model Function about x_k

$$q_k(x) := f(x_k) + \nabla f(x_k)^T (x - x_k) - \frac{\gamma}{2}(x - x_k)^T (x - x_k)$$

satisfies $f(x) \geq q_k(x)$ for all $x \in \mathbb{R}^n$.

The Algorithm

The algorithm remains the same. But the calculation of a global minimizer of

$$q(x) = \max_{k=0,s} q_k(x)$$

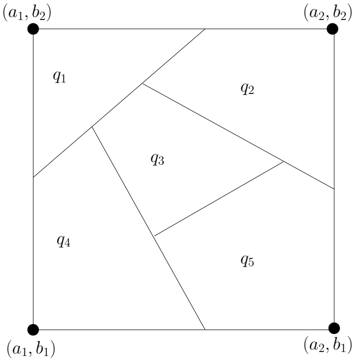
appears to be more difficult computationally.

The Algorithm

The algorithm remains the same. But the calculation of a global minimizer of

$$q(x) = \max_{k=0,s} q_k(x)$$

appears to be more difficult computationally.



- Split the region where a global minimizer is known to lie into subregions.
- In subregion q_k the quadratic function $q_k(x) \geq q_j(x) \;\; \forall j \neq k$.

The Algorithm

Finding a global minimizer of $q(x) = \max_{k=0,s} q_k(x)$

Solve the quadratic program (QP) for $k = 0, \ldots, s$.

$$\begin{array}{ll} \text{minimize}_{x \in \mathbb{R}^n} & q_k(x) \\ \text{subject to} & q_k(x) \geq q_j(x), \quad j \neq k \\ & x_\ell \in [a_\ell, b_\ell] \quad \ell = 1, \ldots, n \end{array}$$

The Algorithm

Notes on the quadratic program

- The constraints $q_k(x) \geq q_j(x)$ are linear.
- The fact that $q_k(x)$ is negative definite makes the QP NP-hard.
- The solution will be attained at a vertex. There are at most $\binom{s}{n}$ vertices.
- In practice number of vertices is much smaller; for $n = 2$ typically each QP has 5-6 vertices regardless of s.
- For small n each QP can be solved efficiently.

The Algorithm

Notes on the quadratic program

- The constraints $q_k(x) \geq q_j(x)$ are linear.
- The fact that $q_k(x)$ is negative definite makes the QP NP-hard.
- The solution will be attained at a vertex. There are at most $\binom{s}{n}$ vertices.
- In practice number of vertices is much smaller; for $n = 2$ typically each QP has 5-6 vertices regardless of s.
- For small n each QP can be solved efficiently.

The Algorithm

Notes on the quadratic program

- The constraints $q_k(x) \geq q_j(x)$ are linear.
- The fact that $q_k(x)$ is negative definite makes the QP NP-hard.
- The solution will be attained at a vertex. There are at most $\begin{pmatrix} s \\ n \end{pmatrix}$ vertices.
- In practice number of vertices is much smaller; for $n = 2$ typically each QP has 5-6 vertices regardless of s.
- For small n each QP can be solved efficiently.

The Algorithm

Notes on the quadratic program

- The constraints $q_k(x) \geq q_j(x)$ are linear.
- The fact that $q_k(x)$ is negative definite makes the QP NP-hard.
- The solution will be attained at a vertex. There are at most $\begin{pmatrix} s \\ n \end{pmatrix}$ vertices.
- In practice number of vertices is much smaller; for $n = 2$ typically each QP has 5-6 vertices regardless of s.
- For small n each QP can be solved efficiently.

The Algorithm

Notes on the quadratic program

- The constraints $q_k(x) \geq q_j(x)$ are linear.
- The fact that $q_k(x)$ is negative definite makes the QP NP-hard.
- The solution will be attained at a vertex. There are at most $\begin{pmatrix} s \\ n \end{pmatrix}$ vertices.
- In practice number of vertices is much smaller; for $n = 2$ typically each QP has 5-6 vertices regardless of s.
- For small n each QP can be solved efficiently.

The Algorithm

Notes on the quadratic program

- The constraints $q_k(x) \geq q_j(x)$ are linear.
- The fact that $q_k(x)$ is negative definite makes the QP NP-hard.
- The solution will be attained at a vertex. There are at most $\begin{pmatrix} s \\ n \end{pmatrix}$ vertices.
- In practice number of vertices is much smaller; for $n = 2$ typically each QP has 5-6 vertices regardless of s.
- For small n each QP can be solved efficiently.

Case Study

Wilkinson Distance

$$\mathcal{W}(A) := \inf\{\|\delta A\|_2 : \exists\lambda \;\; (A+\delta A) \text{ has } \lambda \text{ a multiple eigenvalue}\}$$
$$= \inf_{\lambda\in\mathbb{C}} \sup_{\gamma\in\mathbb{C}} \sigma_{2n-1}\left(\begin{bmatrix} A-\lambda I & \gamma I \\ 0 & A-\lambda I \end{bmatrix}\right)$$

Random matrices

# of function evaluations

n / ϵ	10^{-2}	10^{-3}	10^{-4}	10^{-5}
10	74	80	84	89
20	102	111	114	115
40	101	135	148	155

Case Study

Wilkinson Distance

$$\mathcal{W}(A) := \inf\{\|\delta A\|_2 : \exists\lambda \ (A+\delta A) \text{ has } \lambda \text{ a multiple eigenvalue}\}$$
$$= \inf_{\lambda\in\mathbb{C}} \sup_{\gamma\in\mathbb{C}} \sigma_{2n-1}\left(\begin{bmatrix} A-\lambda I & \gamma I \\ 0 & A-\lambda I \end{bmatrix}\right)$$

Random matrices

# of function evaluations

n / ϵ	10^{-2}	10^{-3}	10^{-4}	10^{-5}
10	74	80	84	89
20	102	111	114	115
40	101	135	148	155

Case Study

Wilkinson Distance

$$\mathcal{W}(A) := \inf\{\|\delta A\|_2 : \exists\lambda \ \ (A + \delta A) \text{ has } \lambda \text{ a multiple eigenvalue}\}$$
$$= \inf_{\lambda\in\mathbb{C}} \sup_{\gamma\in\mathbb{C}} \sigma_{2n-1}\left(\begin{bmatrix} A-\lambda I & \gamma I \\ 0 & A-\lambda I \end{bmatrix}\right)$$

Random matrices

cpu-times

n / ϵ	10^{-2}	10^{-3}	10^{-4}	10^{-5}
10	4.90	6.09	6.99	9.22
20	24.5	30.1	34.0	34.3
40	32.8	69.7	90.4	103.6

Summary

- A conjecture has been made relating the the psedospectra and generalized Wilkinson distance.
- In particular it is shown that $\mathcal{W}_r(A) \leq \mathcal{G}_r(A)$.
- A generic algorithm is introduced for the optimization of symmetric eigenvalues based on their analyticity.
- Future work
  - The direction $\mathcal{W}_r(A) \geq \mathcal{G}_r(A)$ remains open.
  - Improvements on the algorithm for the optimization of eigenvalues in the multivariate-case

Summary

- A conjecture has been made relating the the psedospectra and generalized Wilkinson distance.
- In particular it is shown that $\mathcal{W}_r(A) \leq \mathcal{G}_r(A)$.
- A generic algorithm is introduced for the optimization of symmetric eigenvalues based on their analyticity.
- Future work
  - The direction $\mathcal{W}_r(A) \geq \mathcal{G}_r(A)$ remains open.
  - Improvements on the algorithm for the optimization of eigenvalues in the multivariate-case

Summary

- A conjecture has been made relating the the psedospectra and generalized Wilkinson distance.
- In particular it is shown that $\mathcal{W}_r(A) \leq \mathcal{G}_r(A)$.
- A generic algorithm is introduced for the optimization of symmetric eigenvalues based on their analyticity.

- Future work
  - The direction $\mathcal{W}_r(A) \geq \mathcal{G}_r(A)$ remains open.
  - Improvements on the algorithm for the optimization of eigenvalues in the multivariate-case

Summary

- A conjecture has been made relating the the psedospectra and generalized Wilkinson distance.
- In particular it is shown that $\mathcal{W}_r(A) \leq \mathcal{G}_r(A)$.
- A generic algorithm is introduced for the optimization of symmetric eigenvalues based on their analyticity.

- Future work
  - The direction $\mathcal{W}_r(A) \geq \mathcal{G}_r(A)$ remains open.
  - Improvements on the algorithm for the optimization of eigenvalues in the multivariate-case

Summary

- A conjecture has been made relating the the psedospectra and generalized Wilkinson distance.
- In particular it is shown that $\mathcal{W}_r(A) \leq \mathcal{G}_r(A)$.
- A generic algorithm is introduced for the optimization of symmetric eigenvalues based on their analyticity.

- Future work
  - The direction $\mathcal{W}_r(A) \geq \mathcal{G}_r(A)$ remains open.
  - Improvements on the algorithm for the optimization of eigenvalues in the multivariate-case

Summary

- A conjecture has been made relating the the psedospectra and generalized Wilkinson distance.
- In particular it is shown that $\mathcal{W}_r(A) \leq \mathcal{G}_r(A)$.
- A generic algorithm is introduced for the optimization of symmetric eigenvalues based on their analyticity.

- Future work
  - The direction $\mathcal{W}_r(A) \geq \mathcal{G}_r(A)$ remains open.
  - Improvements on the algorithm for the optimization of eigenvalues in the multivariate-case

References

- 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*, 118:109-135, 2011
- F. Rellich, Perturbation Theory of Eigenvalue Problems, *Technical Report, NYU Institute of Mathematical Sciences*, 1954
- P. Lancaster, On eigenvalues of matrices dependent on a parameter, *Numer Math*, 6:377-387, 1964