EIGENVALUE OPTIMIZATION, ROBUST STABILITY, AND WELL-POSEDNESS

Adrian Lewis
Simon Fraser University
March 29, 2004

Collaborators: Jim Burke, Michael Overton
Paris, March 2004
1. THEME

The idea of robustness.
1. **THEME**

The idea of *robustness*.

Interplay between

- robustness in models;
1. THEME

The idea of robustness.

Interplay between

- robustness in models;
- conditioning of
  - inequality systems
1. THEME

The idea of robustness.

Interplay between

- robustness in models;
- conditioning of
  - inequality systems
  - generalized equations;
1. THEME

The idea of robustness.

Interplay between

• robustness in models;

• conditioning of

  – inequality systems
  – generalized equations;

• robust algebraic (spectral) properties.
1. THEME

The idea of robustness.

Interplay between

- robustness in models;
- conditioning of
  - inequality systems
  - generalized equations;
- robust algebraic (spectral) properties.

Key example: stability of dynamical system

\[
\frac{dx}{dt} = Ax.
\]
2. THREE IDEAS OF STABILITY

\[ A \in M^n = \{n \times n \text{ complex matrices}\} \]
\[ H^n_+ = \{\text{positive semidefinite Hermitians}\} \]

Equivalent conditions for stable \( A \):
2. THREE IDEAS OF STABILITY

\[ A \in \mathbf{M}^n = \{ n \times n \text{ complex matrices} \} \]
\[ \mathbf{H}_+^n = \{ \text{positive semidefinite Hermitians} \} \]

Equivalent conditions for stable \( A \):

- exponential decay for all trajectories of

\[ \frac{dx}{dt} = Ax; \]
2. THREE IDEAS OF STABILITY

\[ A \in \mathbf{M}^n = \{ n \times n \text{ complex matrices} \} \]
\[ \mathbf{H}^n_+ = \{ \text{positive semidefinite Hermitians} \} \]

Equivalent conditions for stable \( A \):

- exponential decay for all trajectories of
  \[
  \frac{dx}{dt} = Ax;
  \]
- all eigenvalues of \( A \) have real parts < 0;
2. THREE IDEAS OF STABILITY

\[ A \in \mathbf{M}^n = \{n \times n \text{ complex matrices}\} \]
\[ \mathbf{H}^n_+ = \{\text{positive semidefinite Hermitians}\} \]

Equivalent conditions for stable \( A \):

- exponential decay for all trajectories of 
  \[ \frac{dx}{dt} = Ax; \]
- all eigenvalues of \( A \) have real parts < 0;
- (Lyapunov, 1893) Surjectivity (onto \( \mathbf{H}^n \)) of 
  \[ X \in \mathbf{H}^n_+ \mapsto AX +XA^* + \mathbf{H}^n_+. \]
2. THREE IDEAS OF STABILITY

\[ A \in M^n = \{ n \times n \text{ complex matrices} \} \]
\[ H^n_+ = \{ \text{positive semidefinite Hermitians} \} \]

Equivalent conditions for stable \( A \): 

- exponential decay for all trajectories of 
  \[ \frac{dx}{dt} = Ax; \]
- all eigenvalues of \( A \) have real parts \(< 0\);
- (Lyapunov, 1893) Surjectivity (onto \( H^n \)) of 
  \[ X \in H^n_+ \mapsto AX + XA^* + H^n_+. \]
Choose parameters $u, v$ to stabilize $A = \begin{bmatrix} u - \epsilon & 1 & 0 \\ -u & -\epsilon & 1 \\ v & 0 & -\epsilon \end{bmatrix}$ $(\epsilon > 0 \text{ small})$. 

3. EXAMPLE
3. **EXAMPLE**

Choose parameters $u, v$ to stabilize

\[
A = \begin{bmatrix}
  u - \epsilon & 1 & 0 \\
  -u & -\epsilon & 1 \\
  v & 0 & -\epsilon
\end{bmatrix}
\]

($\epsilon > 0$ small).

Spectrum optimal (i.e. eigenvalues pushed maximally left) when $u = v = 0$, so

\[
\hat{A} = \begin{bmatrix}
  -\epsilon & 1 & 0 \\
  0 & -\epsilon & 1 \\
  0 & 0 & -\epsilon
\end{bmatrix}
\]

(stable).
3. EXAMPLE

Choose parameters $u, v$ to stabilize

$$A = \begin{bmatrix} u - \epsilon & 1 & 0 \\ -u & -\epsilon & 1 \\ v & 0 & -\epsilon \end{bmatrix} \quad (\epsilon > 0 \text{ small}).$$

Spectrum optimal (i.e. eigenvalues pushed maximally left) when $u = v = 0$, so

$$\hat{A} = \begin{bmatrix} -\epsilon & 1 & 0 \\ 0 & -\epsilon & 1 \\ 0 & 0 & -\epsilon \end{bmatrix} \quad (\text{stable}).$$

Typical to see multiple eigenvalues (B-L-O ’01): eigenvalues retain multiplicity for nearby problems.
3. **EXAMPLE**

Choose parameters $u, v$ to stabilize

$$A = \begin{bmatrix}
    u - \epsilon & 1 & 0 \\
    -u & -\epsilon & 1 \\
    v & 0 & -\epsilon
\end{bmatrix} \quad (\epsilon > 0 \text{ small}).$$

Spectrum optimal (i.e. eigenvalues pushed maximally left) when $u = v = 0$, so

$$\hat{A} = \begin{bmatrix}
    -\epsilon & 1 & 0 \\
    0 & -\epsilon & 1 \\
    0 & 0 & -\epsilon
\end{bmatrix} \quad \text{(stable)}.$$

Typical to see multiple eigenvalues (B-L-O ’01): eigenvalues retain multiplicity for nearby problems.

(Ulimate reason: “partial smoothness” (Lewis ’03))
4. NONROBUSTNESS

But, this solution has weaknesses:
4. NONROBUSTNESS

But, this solution has weaknesses:

- In model, trajectories of $\frac{dx}{dt} = \hat{A}x$ have big transient peaks.
4. NONROBUSTNESS

But, this solution has weaknesses:

- In model, trajectories of \( \frac{dx}{dt} = \hat{A}x \) have big transient peaks.

- Algebraically, the nearby matrix

\[
\begin{pmatrix}
-\epsilon & 1 & 0 \\
0 & -\epsilon & 1 \\
\epsilon^3 & 0 & -\epsilon
\end{pmatrix}
\]

is unstable.
4. NONROBUSTNESS

But, this solution has weaknesses:

- In model, trajectories of \( \frac{dx}{dt} = \hat{A}x \) have big transient peaks.

- Algebraically, the nearby matrix

\[
\begin{bmatrix}
-\epsilon & 1 & 0 \\
0 & -\epsilon & 1 \\
\epsilon^3 & 0 & -\epsilon
\end{bmatrix}
\]

is unstable.

- The inequality system

\[
\hat{A}X + X\hat{A}^* + Y = -I, \quad X, Y \in \mathbf{H}^n
\]

is ill-conditioned: \( \|X\| \) big.
5. DISTANCE TO INSTABILITY

How close to unstable is $A \in \mathcal{M}^n$?
5. DISTANCE TO INSTABILITY

How close to unstable is \( A \in \mathbb{M}^n \)?

Singular values

\[
\sigma_{\text{max}}(Y) = \|Y\| \quad \text{and} \quad \sigma_{\text{min}}(Y)
\]

are largest/smallest eigenvalues of \( \sqrt{Y^*Y} \).
5. DISTANCE TO INSTABILITY

How close to unstable is $A \in \mathbb{M}^n$?

Singular values

$$\sigma_{\text{max}}(Y) = \|Y\| \quad \text{and} \quad \sigma_{\text{min}}(Y)$$

are largest/smallest eigenvalues of $\sqrt{Y^*Y}$.

Goals for distance to instability

$$\beta(A) = \min\{\|X - A\| : X \text{ unstable}\} :$$
5. DISTANCE TO INSTABILITY

How close to unstable is $A \in \mathbb{M}^n$?

Singular values

$$\sigma_{\text{max}}(Y) = \|Y\| \quad \text{and} \quad \sigma_{\text{min}}(Y)$$

are largest/smallest eigenvalues of $\sqrt{Y^*Y}$.

Goals for distance to instability

$$\beta(A) = \min\{\|X - A\| : X \text{ unstable}\} :$$

- characterize it;
5. DISTANCE TO INSTABILITY

How close to unstable is $A \in \mathbb{M}^n$?

Singular values

$$\sigma_{\text{max}}(Y) = \|Y\| \quad \text{and} \quad \sigma_{\text{min}}(Y)$$

are largest/smallest eigenvalues of $\sqrt{Y^*Y}$.

Goals for distance to instability

$$\beta(A) = \min \{\|X - A\| : X \text{ unstable}\} :$$

- characterize it;
- compute it;
5. DISTANCE TO INSTABILITY

How close to unstable is $A \in \mathbb{M}^n$?

Singular values

$$\sigma_{\text{max}}(Y) = \|Y\| \quad \text{and} \quad \sigma_{\text{min}}(Y)$$

are largest/smallest eigenvalues of $\sqrt{Y^*Y}$.

Goals for distance to instability

$$\beta(A) = \min\{\|X - A\| : X \text{ unstable}\} :$$

- characterize it;
- compute it;
- optimize (i.e. maximize) it.
5. DISTANCE TO INSTABILITY

How close to unstable is $A \in \mathbb{M}^n$?

Singular values

$$\sigma_{\text{max}}(Y) = \|Y\| \quad \text{and} \quad \sigma_{\text{min}}(Y)$$

are largest/smallest eigenvalues of $\sqrt{Y^*Y}$.

Goals for distance to instability

$$\beta(A) = \min \{\|X - A\| : X \text{ unstable}\} :$$

- characterize it;
- compute it;
- optimize (i.e. maximize) it.

Equivalently, study (structured) distance to nonsurjectivity for the set-valued map

$$X \in H_+^n \quad \mapsto \quad AX + XA^* + H_+^n.$$
6. **EXAMPLE: LINEAR MAPS**

For $A \in \mathbb{M}^n$, consider distance to nonsurjectivity of

$$x \in \mathbb{C}^n \mapsto Ax \quad \text{(for } A \in \mathbb{M}^n).$$
For $A \in \mathbb{M}^n$, consider distance to nonsurjectivity of
\[ x \in \mathbb{C}^n \mapsto Ax \quad \text{(for } A \in \mathbb{M}^n). \]

- By (Eckart-Young '39), easy to characterize:
\[
\min \{ \|X - A\| : X \text{ singular} \} = \sigma_{\text{min}}(A).
\]
6. **EXAMPLE: LINEAR MAPS**

For \( A \in \mathbb{M}^n \), consider distance to nonsurjectivity of

\[
x \in \mathbb{C}^n \mapsto Ax \quad (\text{for } A \in \mathbb{M}^n).
\]

- By (Eckart-Young '39), easy to characterize:

\[
\min\{\|X - A\| : X \text{ singular}\} = \sigma_{\min}(A).
\]

- Hence easy to compute.
6. EXAMPLE: LINEAR MAPS

For $A \in M^n$, consider distance to nonsurjectivity of
\[ x \in \mathbb{C}^n \rightarrow Ax \quad \text{(for } A \in M^n). \]

- By (Eckart-Young '39), easy to characterize:
  \[ \min\{\|X - A\| : X \text{ singular}\} = \sigma_{\min}(A). \]

- Hence easy to compute.

- For optimization, $-\sigma_{\min}(\cdot)$ neither smooth nor convex, (but is Lipschitz and “Clarke regular”).
6. **EXAMPLE: LINEAR MAPS**

For \( A \in \mathbb{M}^n \), consider distance to nonsurjectivity of

\[
x \in \mathbb{C}^n \mapsto Ax \quad \text{(for } A \in \mathbb{M}^n \text{)}.
\]

- By (Eckart-Young ’39), easy to characterize:

\[
\min \{ \| X - A \| : X \text{ singular} \} = \sigma_{\min}(A).
\]

- Hence easy to compute.

- For optimization, \(-\sigma_{\min}(\cdot)\) neither smooth nor convex, 
  (but is Lipschitz and “Clarke regular”).

Analogously...
7. OPTIMIZING ROBUST STABILITY

- Distance to instability characterized by:

\[ \beta(A) = \min_{s \in \mathbb{R}} \sigma_{\text{min}}(A - isI) \]  

(Van Loan '85).
7. OPTIMIZING ROBUST STABILITY

- Distance to instability characterized by:
  \[ \beta(A) = \min_{s \in \mathbb{R}} \sigma_{\min}(A - isI) \]  (Van Loan '85).

- Computable via \(O(n^3)\) globally and quadratically convergent method of (Boyd et al. '90).
7. **OPTIMIZING ROBUST STABILITY**

- Distance to instability characterized by:
  \[
  \beta(A) = \min_{s \in \mathbb{R}} \sigma_{\min}(A - isI) \tag{Van Loan '85}.
  \]

- Computable via \(O(n^3)\) globally and quadratically convergent method of (Boyd et al. '90).

- \(\nabla \beta(A)\) follows easily (when it exists), so we can optimize (locally) by “gradient sampling”.
7. **OPTIMIZING ROBUST STABILITY**

- Distance to instability characterized by:
  \[ \beta(A) = \min_{s \in \mathbb{R}} \sigma_{\min}(A - isI) \]  
  (Van Loan '85).

- Computable via \( O(n^3) \) globally and quadratically convergent method of (Boyd et al. '90).

- \( \nabla \beta(A) \) follows easily (when it exists), so we can optimize (locally) by “gradient sampling”.

**Example:**

\[
\max_{u,v} \beta \begin{bmatrix}
  u - \epsilon & 1 & 0 \\
  -u & -\epsilon & 1 \\
  u & 0 & -\epsilon \\
\end{bmatrix} \sim \epsilon^2
\]
7. **OPTIMIZING ROBUST STABILITY**

- Distance to instability characterized by:

\[ \beta(A) = \min_{s \in \mathbb{R}} \sigma_{\min}(A - isI) \]  

(Van Loan '85).

- Computable via \( O(n^3) \) globally and quadratically convergent method of (Boyd et al. '90).

- \( \nabla \beta(A) \) follows easily (when it exists), so we can optimize (locally) by “gradient sampling”.

**Example:**

\[
\max_{u,v} \beta \begin{bmatrix}
    u - \epsilon & 1 & 0 \\
    -u & -\epsilon & 1 \\
    v & 0 & -\epsilon
\end{bmatrix} \sim \epsilon^2
\]

whereas optimal spectrum gives

\[
\beta \begin{bmatrix}
    -\epsilon & 1 & 0 \\
    0 & -\epsilon & 1 \\
    0 & 0 & -\epsilon
\end{bmatrix} \leq \epsilon^3.
\]
8. GRADIENT SAMPLING

We want to minimize almost-everywhere-smooth, regular $f$. 
8. GRADIENT SAMPLING

We want to minimize almost-everywhere-smooth, regular $f$.

Algorithm: Given current point $x$, repeat

1. Pick random points $x^j$ near $x$;
8. GRADIENT SAMPLING

We want to minimize almost-everywhere-smooth, regular $f$.

Algorithm: Given current point $x$, repeat

1. Pick random points $x^j$ near $x$;
2. Calculate shortest vector $d \in \text{conv}\{\nabla f(x^j)\}$;
8. GRADIENT SAMPLING

We want to minimize almost-everywhere-smooth, regular $f$.

Algorithm: Given current point $x$, repeat

1. Pick random points $x^j$ near $x$;
2. Calculate shortest vector $d \in \text{conv}\{\nabla f(x^j)\}$;
3. Update $x := x - td$, for $t \geq 0$ chosen by linesearch.
8. GRADIENT SAMPLING

We want to minimize almost-everywhere-smooth, regular $f$.

Algorithm: Given current point $x$, repeat

1. Pick random points $x^j$ near $x$;
2. Calculate shortest vector $d \in \text{conv}\{\nabla f(x^j)\}$;
3. Update $x := x - td$, for $t \geq 0$ chosen by linesearch.

Under reasonable conditions (B-L-O ’03),

$$\text{iterates } \rightarrow \hat{x} \text{ satisfying } f'(\hat{x}; d) \geq 0 \ \forall d.$$
8. GRADIENT SAMPLING

We want to minimize almost-everywhere-smooth, regular $f$.

Algorithm: Given current point $x$, repeat

1. Pick random points $x^j$ near $x$;
2. Calculate shortest vector $d \in \text{conv}\{\nabla f(x^j)\}$;
3. Update $x := x - td$, for $t \geq 0$ chosen by linesearch.

Under reasonable conditions (B-L-O '03),

iterates $\to \hat{x}$ satisfying $f'(\hat{x}; d) \geq 0 \ \forall d$.

Good if $f$ “subsmooth” (Rockafellar):

$$f(x) = \max_{v \in V} f_v(x) \quad (V \text{ compact}).$$
9. CONTROLLABILITY

Analogous equivalent conditions:

- In the model, trajectories for

\[
\frac{dx}{dt} = Ax + Bu
\]

can interpolate any endpoints using a control \( u(\cdot) \).
9. CONTROLLABILITY

Analogous equivalent conditions:

- In the model, trajectories for
  \[
  \frac{dx}{dt} = Ax + Bu
  \]
  can interpolate any endpoints using a control \( u(\cdot) \).

- Algebraically, \([A - zI, B]\) has full row rank \( \forall z \in \mathbb{C} \).
9. CONTROLLABILITY

Analogous equivalent conditions:

• In the model, trajectories for

\[ \frac{dx}{dt} = Ax + Bu \]

can interpolate any endpoints using a control \( u(\cdot) \).

• Algebraically, \([A - zI, B]\) has full row rank \( \forall z \in \mathbb{C} \).

• “Pole placement”: we can always find \( X \) solving generalized equation

\[ \text{spectrum}(A + BX) = \ldots. \]
9. **CONTROLLABILITY**

Analogous equivalent conditions:

- In the **model**, trajectories for

\[
\frac{dx}{dt} = Ax + Bu
\]

can interpolate any endpoints using a control \( u(\cdot) \).

- **Algebraically**, \([A - zI, B]\) has full row rank \( \forall z \in \mathbb{C} \).

- “Pole placement”: we can always find \( X \) solving generalized equation

\[
\text{spectum}(A + BX) \quad = \quad \ldots.
\]

Distance to uncontrollability (**Eising '84**)

\[
= \min_{z \in \mathbb{C}} \sigma_{\min}[A - zI, B].
\]
9. CONTROLLABILITY

Analogous equivalent conditions:

- In the model, trajectories for
  \[
  \frac{dx}{dt} = Ax + Bu
  \]
  can interpolate any endpoints using a control \( u(\cdot) \).

- Algebraically, \([A - zI, B]\) has full row rank \( \forall z \in \mathbb{C} \).

- “Pole placement”: we can always find \( X \) solving generalized equation
  \[
  \text{spectrum}(A + BX) = \ldots
  \]

Distance to uncontrollability (Eising ’84)

\[
= \min_{z \in \mathbb{C}} \sigma_{\min}[A - zI, B].
\]

Tractable via \( O(n^6) \) method of (Gu ’00).
10. THEORETICAL INTERLUDE

Recall: \( A \in M^n \) stable \( \iff \)

\[ X \in H^+_n \implies AX + XA^* + H^+_n \]

surjective.
10. THEORETICAL INTERLUDE

Recall: \( A \in M^n \) stable \( \iff \)

\[
X \in H^n_+ \quad \mapsto \quad AX + XA^* + H^n_+
\]

surjective.

So, how close is set-valued map

\[
F : R^n \Rightarrow R^m
\]

to nonsurjectivity?
10. THEORETICAL INTERLUDE

Recall: \( A \in \mathcal{M}^n \) stable \( \iff \)
\[
X \in \mathcal{H}^n_+ \mapsto AX + XA^* + \mathcal{H}^n_+
\]

surjective.

So, how close is set-valued map
\[
F : \mathbb{R}^n \Rightarrow \mathbb{R}^m
\]
to nonsurjectivity?

Distance to nonsurjectivity of linear \( A : \mathbb{R}^n \to \mathbb{R}^n \) is (by Eckart-Young)
\[
\sigma_{\min}(A) = \min_{\|y\| \leq 1} \max_x \left\{ \frac{1}{\|x\|} : Ax = y \right\}.
\]
10. THEORETICAL INTERLUDE

Recall: \( A \in \mathbb{M}^n \) stable \( \Leftrightarrow \)

\[
X \in \mathbb{H}^n_+ \quad \mapsto \quad AX + XA^* + \mathbb{H}^n_+
\]

surjective.

So, how close is set-valued map

\[
F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m
\]

to nonsurjectivity?

Distance to nonsurjectivity of linear \( A : \mathbb{R}^n \rightarrow \mathbb{R}^n \) is (by Eckart-Young)

\[
\sigma_{\min}(A) = \min_{\|y\| \leq 1} \max_x \left\{ \frac{1}{\|x\|} : Ax = y \right\}.
\]

Generalizing...
11. **SUBLINEAR MAPS**

**Theorem (Renegar '95)** Distance to nonsurjectivity for set-valued $F$ is

$$\min_{\text{linear } T} \left\{ \|T\| : F + T \text{ nonsurjective} \right\}$$

$$= \min_{\|y\| \leq 1} \max_x \left\{ \frac{1}{\|x\|} : y \in F(x) \right\}$$
11. **SUBLINEAR MAPS**

**Theorem (Renegar ’95)** Distance to nonsurjectivity for set-valued $F$ is

$$\min_{\text{linear } T} \left\{ \|T\| : F + T \text{ nonsurjective} \right\}$$

$$= \min_{\|y\| \leq 1} \max_{x} \left\{ \frac{1}{\|x\|} : y \in F(x) \right\}$$

whenever

$$F(x) = \begin{cases} Ax + P & (x \in Q) \\ \emptyset & (x \notin Q) \end{cases}$$

for linear $A$, closed convex cones $P, Q$. 
11. **SUBLINEAR MAPS**

**Theorem (Renegar '95)** Distance to nonsurjectivity for set-valued $F$ is

$$\min_{\text{linear } T} \{\|T\| : F + T \text{ nonsurjective}\}$$

$$= \min_{\|y\| \leq 1} \max_x \left\{ \frac{1}{\|x\|} : y \in F(x) \right\}$$

whenever

$$F(x) = \begin{cases} \text{Ax} + P & (x \in Q) \\ \emptyset & (x \notin Q) \end{cases}$$

for linear $A$, closed convex cones $P, Q$.

**Theorem (Lewis '99)** Same holds for closed sublinear $F$:

$$\text{graph} = \{(x, y) : y \in F(x)\}$$

a closed convex cone.
12. TWO FURTHER GENERALIZATIONS

Local version:

**Theorem (Dontchev-Lewis-Rockafellar ’03)**  Locally, around the graph of any closed \( F \),

\[
\text{distance to “nonregularity”} = \frac{1}{\text{regularity modulus}}.
\]
12. TWO FURTHER GENERALIZATIONS

Local version:

**Theorem (Dontchev-Lewis-Rockafellar ’03)** Locally, around the graph of any closed $F$,

$$\text{distance to “nonregularity”} = \frac{1}{\text{regularity modulus}}.$$  

(Regularity modulus at $(x, F(x))$ is $\|\nabla F(x)^{-1}\|$ if $F$ smooth.)
12. TWO FURTHER GENERALIZATIONS

Local version:

**Theorem (Dontchev-Lewis-Rockafellar ’03)** Locally, around the graph of any closed $F$,

$$\text{distance to “nonregularity” } = \frac{1}{\text{regularity modulus}}.$$ 

(Regularity modulus at $(x, F(x))$ is $\|\nabla F(x)^{-1}\|$ if $F$ smooth.)

Structured version, following (Peña ’03):
12. TWO FURTHER GENERALIZATIONS

Local version:

**Theorem (Dontchev-Lewis-Rockafellar ’03)**  Locally, around the graph of any closed $F$,

$$\text{distance to “nonregularity”} = \frac{1}{\text{regularity modulus}}.$$  

(Regularity modulus at $(x, F(x))$ is $\|\nabla F(x)^{-1}\|$ if $F$ smooth.)

Structured version, following (Peña ’03):

**Theorem (Lewis ’03)**  For closed sublinear $F$,

$$\min_{\text{linear } T_i} \left\{ \max_i \|T_i\| : F + \sum_i P_i T_i Q_i \text{ nonsurjective} \right\}$$

$$= \min_{\|v_i\| \leq 1} \sup_{x, w_i > 0} \left\{ \min_i \frac{w_i}{\|Q_i x\|} : \sum_i w_i P_i v_i \in F(x) \right\}.$$
13. PSEUDOSPECTRA

Good tool for visualizing robust spectral properties of $A \in \mathbb{M}^n$ (Trefethen\ldots). Available online as eigtool.
13. PSEUDOSPECTRA

Good tool for visualizing robust spectral properties of \( A \in \mathbb{M}_n \) (Trefethen...). Available online as eigtool.

Spectrum is \( \Lambda(A) = \{ \text{eigenvalues of } A \} \).

For \( \epsilon \geq 0 \), pseudospectrum is

\[
\Lambda_\epsilon(A) = \bigcup_{\| X - A \| \leq \epsilon} \Lambda(X)
\]
13. PSEUDOSPECTRA

Good tool for visualizing robust spectral properties of $A \in \mathbb{M}^n$ (Trefethen...). Available online as eigtool.

Spectrum is $\Lambda(A) = \{\text{eigenvalues of } A\}$.

For $\epsilon \geq 0$, pseudospectrum is

$$
\Lambda_\epsilon(A) = \bigcup_{\|X - A\| \leq \epsilon} \Lambda(X)
$$

$$
\quad = \{z \in \mathbb{C} : \sigma_{\min}(A - zI) \leq \epsilon\}.
$$
13. **PSEUDOSPECTRA**

Good tool for visualizing robust spectral properties of $A \in \mathbb{M}^n$ (Trefethen...). Available online as eigtool.

**Spectrum** is $\Lambda(A) = \{\text{eigenvalues of } A\}$.

For $\epsilon \geq 0$, **pseudospectrum** is

$$\Lambda_\epsilon(A) = \bigcup_{\|X - A\| \leq \epsilon} \Lambda(X)$$

$$= \{z \in \mathbb{C} : \sigma_{\text{min}}(A - zI) \leq \epsilon\}.$$ 

Note: distance to instability satisfies

$$\beta(A) \geq \epsilon \iff \Lambda_\epsilon(A) \subset \text{left halfplane}.$$
Pseudospectra for a random upper-triangular 10-by-10 real matrix (computed by eigtool):
15. EXAMPLE

Aeroelastic model \((A \in \mathbb{M}^{55})\) of Boeing 767 flutter condition, without feedback.
15. EXAMPLE

Aeroelastic model \((A \in \mathbb{M}^{55})\) of Boeing 767 flutter condition, without feedback.
16. B767 WITH FEEDBACK

“Static output feedback”:

\[ \frac{dx}{dt} = Ax + Bu, \quad y = Cx. \]

**Problem:** find stabilizing control \( u = Ky \).
“Static output feedback”:

\[
\frac{dx}{dt} = Ax + Bu, \quad y = Cx.
\]

**Problem:** find stabilizing control \( u = Ky \).

So, choose \( K \in \mathbb{M}^2 \) to optimize

- spectrum
- distance to instability

of \( A + BKC \).
16. B767 WITH FEEDBACK

“Static output feedback”:

\[
\frac{dx}{dt} = Ax + Bu, \quad y = Cx.
\]

**Problem:** find stabilizing control \( u = Ky \).

So, choose \( K \in \mathbb{M}^2 \) to optimize

- spectrum
- distance to instability

of \( A + BKC \).

Using gradient sampling...
17. B767: OPTIMIZED SPECTRUM
18. **B767: OPTIMIZED ROBUSTNESS**

![Diagram](image)
Pushing eigenvalues of $A$ left improves \textit{asymptotic} decay for

$$\frac{dx}{dt} = Ax$$

but not robustness.
19. DECAY VERSUS ROBUSTNESS

Pushing eigenvalues of $A$ left improves asymptotic decay for

$$\frac{dx}{dt} = Ax$$

but not robustness.

Maximizing distance to instability $\beta(A)$ improves robust stability and transient peaks, but not decay.
19. DECAY VERSUS ROBUSTNESS

Pushing eigenvalues of $A$ left improves asymptotic decay for

$$\frac{dx}{dt} = Ax$$

but not robustness.

Maximizing distance to instability $\beta(A)$ improves robust stability and transient peaks, but not decay.

Compromise — minimize pseudospectral abscissa

$$\alpha_\epsilon(A) = \max\{\Re z : z \in \Lambda_\epsilon(A)\}$$

$$= \max\{\Re z : \sigma_{\min}(A - zI) \leq \epsilon\}.$$
19. DECAY VERSUS ROBUSTNESS

Pushing eigenvalues of $A$ left improves asymptotic decay for

$$\frac{dx}{dt} = Ax$$

but not robustness.

Maximizing distance to instability $\beta(A)$ improves robust stability and transient peaks, but not decay.

Compromise — minimize pseudospectral abscissa

$$\alpha_\epsilon(A) = \max\{\text{Re } z : z \in \Lambda_\epsilon(A)\}$$

$$= \max\{\text{Re } z : \sigma_{\text{min}}(A - zI) \leq \epsilon\}.$$ 

Note: $\alpha_\epsilon(A) \leq 0 \iff \beta(A) \geq \epsilon.$
Pushing eigenvalues of $A$ left improves asymptotic decay for
\[
\frac{dx}{dt} = Ax
\]
but not robustness.

Maximizing distance to instability $\beta(A)$ improves robust stability and transient peaks, but not decay.

Compromise — minimize pseudospectral abscissa
\[
\alpha_\epsilon(A) = \max\{\Re z : z \in \Lambda_\epsilon(A)\} = \max\{\Re z : \sigma_{\min}(A - zI) \leq \epsilon\}.\]

Note: $\alpha_\epsilon(A) \leq 0 \Leftrightarrow \beta(A) \geq \epsilon$.

**Theorem (B-L-O '03)** Near any $A$ with geometrically simple eigenvalues, $\alpha_\epsilon$ is Lipschitz and regular.
19. DECAY VERSUS ROBUSTNESS

Pushing eigenvalues of $A$ left improves asymptotic decay for

$$\frac{dx}{dt} = Ax$$

but not robustness.

Maximizing distance to instability $\beta(A)$ improves robust stability and transient peaks, but not decay.

Compromise — minimize pseudospectral abscissa

$$\alpha_\epsilon(A) = \max \{ \Re z : z \in \Lambda_\epsilon(A) \}$$

$$= \max \{ \Re z : \sigma_{\min}(A - zI) \leq \epsilon \}.$$ 

Note: $\alpha_\epsilon(A) \leq 0 \iff \beta(A) \geq \epsilon.$

Theorem (B-L-O '03) Near any $A$ with geometrically simple eigenvalues, $\alpha_\epsilon$ is Lipschitz and regular.

(So gradient sampling should work well.)
20. PSEUDOSPECTRAL ABSCISSA
As $\epsilon$ varies, $\alpha_\epsilon(A)$ measures different aspects of model:

<table>
<thead>
<tr>
<th>$\epsilon$</th>
<th>aspect</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>asymptotic decay</td>
</tr>
<tr>
<td>large</td>
<td>initial growth</td>
</tr>
<tr>
<td>?</td>
<td>distance to instability</td>
</tr>
</tbody>
</table>
As $\epsilon$ varies, $\alpha_\epsilon(A)$ measures different aspects of model:

<table>
<thead>
<tr>
<th>$\epsilon$</th>
<th>aspect</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>asymptotic decay</td>
</tr>
<tr>
<td>large</td>
<td>initial growth</td>
</tr>
<tr>
<td>?</td>
<td>distance to instability</td>
</tr>
</tbody>
</table>

Minimizing $\alpha_\epsilon$ balances conflicting aspects for $\frac{dx}{dt} = Ax$: asymptotic versus transient.
As $\epsilon$ varies, $\alpha_\epsilon(A)$ measures different aspects of model:

<table>
<thead>
<tr>
<th>$\epsilon$</th>
<th>aspect</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>asymptotic decay</td>
</tr>
<tr>
<td>large</td>
<td>initial growth</td>
</tr>
<tr>
<td>?</td>
<td>distance to instability</td>
</tr>
</tbody>
</table>

Minimizing $\alpha_\epsilon$ balances conflicting aspects for $\frac{dx}{dt} = Ax$:

- asymptotic versus transient.

$\epsilon = \text{size of likely perturbations to } A$. 

20. PSEUDOSPECTRAL ABSCISSA
As \( \varepsilon \) varies, \( \alpha_{\varepsilon}(A) \) measures different aspects of model:

<table>
<thead>
<tr>
<th>( \varepsilon )</th>
<th>aspect</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>asymptotic decay</td>
</tr>
<tr>
<td>large</td>
<td>initial growth</td>
</tr>
<tr>
<td>?</td>
<td>distance to instability</td>
</tr>
</tbody>
</table>

Minimizing \( \alpha_{\varepsilon} \) balances conflicting aspects for \( \frac{dx}{dt} = Ax \):

asymptotic versus transient.

\( \varepsilon = \) size of likely perturbations to \( A \).

How do we compute \( \alpha_{\varepsilon}(A) \)?
21. **PSEUDOSPECTRAL GEOMETRY**

- Each point in pseudospectrum is accessible from an eigenvalue (using max modulus principle etc...).
21. PSEUDOSPECTRAL GEOMETRY

- Each point in pseudospectrum is accessible from an eigenvalue (using max modulus principle etc.).
- Finding intersections of lines with the pseudospectral boundary is easy via Hamiltonian eigenvalue computation (Van Loan ‘84, Benner ‘99).
21. PSEUDOSPECTRAL GEOMETRY

- Each point in pseudospectrum is accessible from an eigenvalue (using max modulus principle etc. . . ).

- Finding intersections of lines with the pseudospectral boundary is easy via Hamiltonian eigenvalue computation (Van Loan ’84, Benner ’99).
Motivated by $H^\infty$-norm algorithm of (Boyd et al. ’90)...

Algorithm (B-L-O ’03)

- Vertical sweeps, to find midpoints of each segment where a vertical line intersects pseudospectrum;
Motivated by $H^\infty$-norm algorithm of (Boyd et al. ’90)...

Algorithm (B-L-O ’03)

- **Vertical sweeps**, to find midpoints of each segment where a vertical line intersects pseudospectrum;

  ↓

- **Horizontal sweeps**, from each midpoint, to pseudospectral boundary.
23. CRISS-CROSS ALGORITHM

bisect

vertical

horizontal
23. CRISS-CROSS ALGORITHM

- Converges globally and (generically) quadratically.
23. **CRISS-CROSS ALGORITHM**

- Converges globally and (generically) quadratically.
- 100-by-100 matrix takes seconds.
23. **CRISS-CROSS ALGORITHM**

- Converges globally and (generically) quadratically.
- 100-by-100 matrix takes seconds.
- Available in eigtool.
23. CRISS-CROSS ALGORITHM

- Converges globally and (generically) quadratically.
- 100-by-100 matrix takes seconds.
- Available in eigtool.
- Easy to deduce $\nabla \alpha_\epsilon(A)$ (if it exists).
24. OPTIMAL PSEUDOSPECTRUM

Static output feedback stabilization of turbo-generator model: $A \in M^{10}$ with 4 parameters.
Static output feedback stabilization of turbo-generator model: $A \in \mathbb{M}^{10}$ with 4 parameters.
In discrete time, \( x_{r+1} = Ax_r \). Analogous equivalent properties:
25. THEME REVISITED

In discrete time, $x_{r+1} = Ax_r$. Analogous equivalent properties:

- $A^r \to 0$ exponentially;
In discrete time, $x_{r+1} = Ax_r$. Analogous equivalent properties:

- $A^r \to 0$ exponentially;
- spectral radius $< 1$;
In discrete time, $x_{r+1} = Ax_r$. Analogous equivalent properties:

- $A^r \rightarrow 0$ exponentially;
- spectral radius $< 1$;
- surjectivity of

$$X \in \mathbb{H}_+^n \iff A^*X^*A - X + \mathbb{H}_+^n.$$
In discrete time, \( x_{r+1} = A x_r \). Analogous equivalent properties:

- \( A^r \to 0 \) exponentially;
- spectral radius < 1;
- surjectivity of

\[
X \in \mathbb{H}_+^n \mapsto A^* X A - X + \mathbb{H}_+^n.
\]

**Theorem (Kreiss ’62)** Robust version:

\[
P = \sup_{r} \| A^r \|,
\]

\[
L = \inf_{X \in \mathbb{H}_+^n} \{ \text{cond}(X) : X \succ A^* X A \},
\]

\[
K = \sup_{\epsilon > 0, \ z \in \Lambda_\epsilon(A)} \frac{|z| - 1}{\epsilon}
\]

are all related.
25. THEME REVISITED

In discrete time, \( x_{r+1} = Ax_r \). Analogous equivalent properties:

- \( A^r \to 0 \) exponentially;
- spectral radius \( < 1 \);
- surjectivity of

\[
X \in \mathbb{H}^n_+ \iff A^*XA - X + \mathbb{H}^n_+.
\]

**Theorem (Kreiss ’62)** Robust version:

\[
P = \sup_r \|A^r\|,
\]
\[
L = \inf_{X\in \mathbb{H}^n_+} \{\text{cond}(X) : X \succ A^*XA\},
\]
\[
K = \sup_{\epsilon > 0, z \in \Lambda_\epsilon(A)} \frac{|z| - 1}{\epsilon}
\]

are all related. In fact (Spijker ’91),

\[
K \leq P \leq e n K.
\]
model behaviour
  eg: control

algebraic/spectral property

inequality system
generalized equation

robust

pseudospectral property

well-conditioned, far from nonsurjective