STABILITY FOR SYMMETRIC VERSUS NONSYMMETRIC MATRICES: A COMPARISON

Adrian Lewis

Simon Fraser University
Canada

October 2, 2002

(Including joint work with
Jim Burke and Michael Overton)

OUTLINE

- Two matrix cones:
  - negative-semidefinite symmetrics
  - stable nonsymmetrics

  What do they have in common?

- "Convex-like" properties: regularity
- "Active-set-like" properties: partial smoothness
- "Robustness" properties:
  - Lyapunov characterizations
  - Pseudospectra
  - Kreiss matrix theorem
  - stability radii and the $\mathbf{H}_\infty$-norm
  - Lipschitz effects

THE SEMIDEFINITE CONE

$\boxed{\mathbf{S}^n}$ — the space of n-by-n real symmetric matrices.

$\boxed{\mathbf{S}^n_-}$ — the set of negative-semidefinites.

In this space

$$X \succeq Y \quad \text{means} \quad Y - X \in \mathbf{S}^n_-.$$

Properties of $\mathbf{S}^n_-$:

- closed convex cone
- rich algebraic structure (homogeneous, self-dual ...)
- powerful modeling tool in modern optimization.

THE STABLE CONE

$\boxed{\mathbf{M}^n}$ — the space of n-by-n complex matrices.

$\boxed{\mathbf{M}^n_-}$ — the set of stable matrices: all eigenvalues have real part ≤ 0.

Properties of $\mathbf{M}^n_-$:

- closed cone (eigenvalues continuous)
- basic modeling tool: all trajectories of $\dot{x} = Ax$ go to zero exponentially $\Leftrightarrow \quad A \in \text{int}\, \mathbf{M}^n_-$.

But **not** convex:

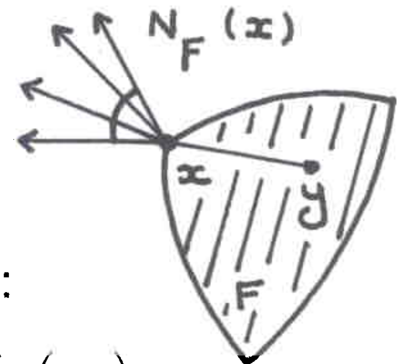
$$\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix}.$$

stable + stable = unstable

So no nice **global** structure.

LOCAL PROPERTIES

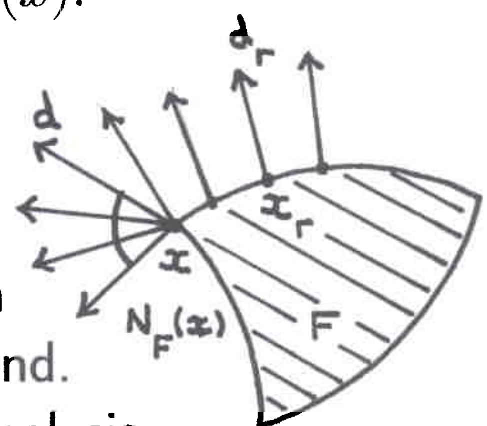
Key property of a closed convex set F:
normal cone map



$$x \in F \mapsto N_F(x) = \bigcap_{y \in F} \{d : d^T(y - x) \leq 0\}$$

is a **closed** multifunction:

$$x_r \in F, \quad d_r \in N_F(x_r)$$
$$x_r \to x, \quad d_r \to d$$
$$\Rightarrow \quad d \in N_F(x).$$



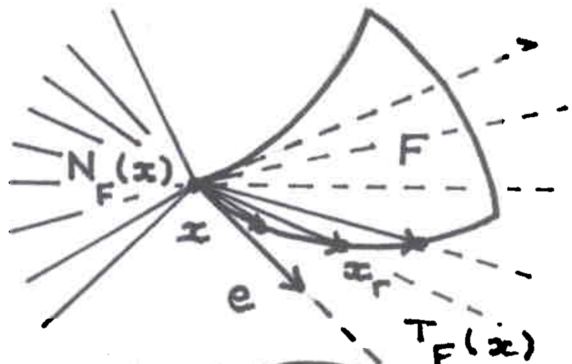
But many nonconvex sets have this property (for suitable N_F). Eg: $\{x : g_i(x) \leq 0 \ \forall i\}$ when $\{\nabla g_i(x) : i \text{ active}\}$ lin. ind.
Crucial for nonsmooth analysis.

NORMALS AND REGULARITY

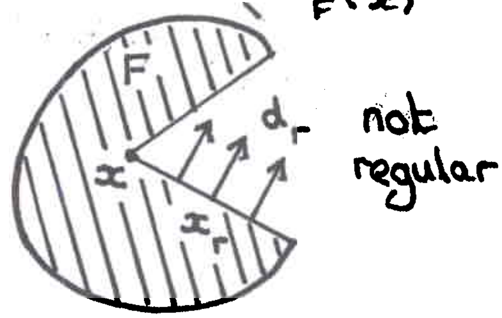
$d \in N_F(x)$ (a regular normal) means $d^T e \leq 0$ for all tangent e, ie. whenever

$x^r \rightarrow x$ in F and

$$\frac{x_r - x}{\|x_r - x\|} \rightarrow e.$$



F Clarke regular at x means regular normal map N_F closed at x.



(Clarke. . . or Rockafellar/Wets, 1998)

REGULARITY OF STABLE CONE

Thm (Burke/Overton '99)
If $A \in \mathbf{M}^n_-$ satisfies

$$(*) \qquad \begin{cases} \text{every imaginary eigenvalue} \\ \text{has geometric multiplicity } 1 \end{cases}$$

then $\mathbf{M}^n_-$ is regular at A.

Eg: Jordan block

$$J_n = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ddots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}$$

Note: the "nonderogatory" assumption $(*)$ is generic—within the matrices having imaginary eigenvalues of any given algebraic multiplicities, those failing $(*)$ have lower dimension.

ACTIVE SET IDEAS

Eg Given $A_1, \ldots, A_k, B \in \mathbf{S}^n$, denote optimal solution of SDP

$$\min\left\{c^T x : \sum_j x_j A_j - B \in \mathbf{S}^n_-\right\}$$

by $x(c)$.

Usually,

$$\text{rank}\left(\sum_j x_j A_j - B\right)$$

stays constant for small changes in c, and

$$c \mapsto x(c)$$

smooth.

PERSISTENCE OF JORDAN FORM

Eg Given $A_1, \ldots, A_k, B \in \mathbf{M}^n$, denote optimal solution of

$$\min\left\{c^T x : \sum_j x_j A_j - B \in \mathbf{M}^n_-\right\}$$

by $x(c)$.

Usually, each imaginary eigenvalue of

$$\sum_j x_j A_j - B$$

has

- geometric multiplicity 1
- algebraic multiplicity constant

for small changes in c, and

$$c \mapsto x(c)$$

smooth.

EXAMPLE

Minimize $c^T x$ subject to

$$\mathcal{A}(x) = \begin{bmatrix} -x_2 & 0 & 0 & \cdots & 0 \\ x_2 & 0 & 0 & \cdots & 0 \\ x_3 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_n & 0 & 0 & \cdots & 0 \end{bmatrix} - x_1 I_n + J_n$$

stable.

For $c = [1, 0, 0, \ldots, 0]^T$, optimal $x = 0$, so $\mathcal{A}(x) = J_n$, (single Jordan block).

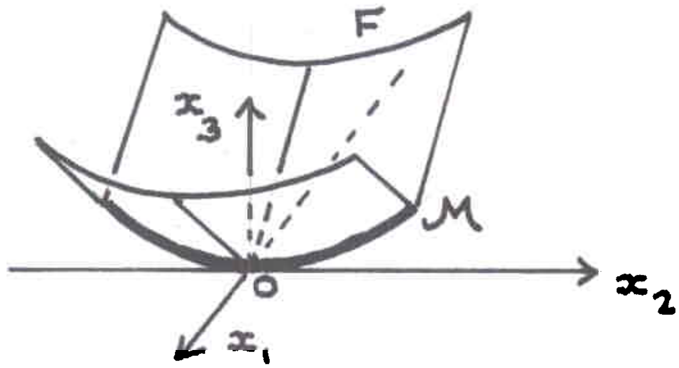
For c close to $[1, 0, 0, \ldots, 0]^T$, optimal $\mathcal{A}(x)$ remains a single Jordan block.

WHY?

PARTLY SMOOTH SETS

A framework for "active-set" ideas.

Eg Consider $\qquad \min\{c^T x : x \in F\}$,
where $F = \{x \in \mathbf{R}^3 : x_3 \geq |x_1| + x_2^2\}$.



As c varies near $[0, 0, 1]^T$, optimal $x(c)$ varies smoothly on active manifold

$$\mathcal{M} = \{x : x_1 = 0,\ x_3 = x_2^2\}.$$

Reason:

- N_F continuous on $\mathcal{M}$, and
- F "sharp" around $\mathcal{M}$.

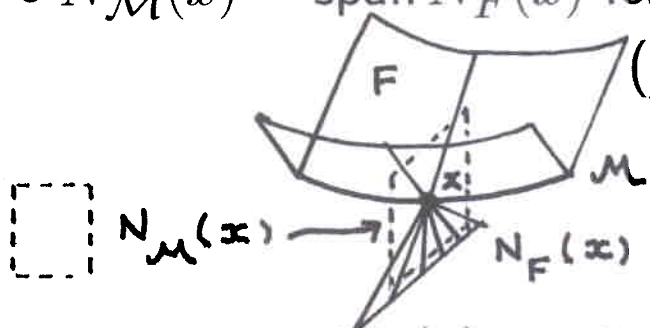
PARTLY SMOOTH: THE DEFINITION

Suppose $\mathcal{M} \subset F$ a manifold:

- defined locally by smooth equations
- linearly independent gradients.

Closed F partly smooth relative to $\mathcal{M}$ if regular throughout $\mathcal{M}$, and normal cone satisfies

- $N_{\mathcal{M}}(x) = \text{span}\, N_F(x)$ for $x \in \mathcal{M}$ (sharpness)



- $x \in \mathcal{M} \mapsto N_F(x)$ continuous:

$$d \in N_F(x),\ x^r \to x \text{ in } \mathcal{M} \quad \Rightarrow \quad \exists d^r \in N_F(x^r) \text{ with } d^r \to d.$$

Partly smooth sets have good calculus.

THE CONES

Thm (Oustry, '00) For any r,
$\mathbf{S}^n_-$ is partly smooth relative to

$$\{X : \text{rank}\, X = r\}.$$

For any list of numbers $\mathcal{L}$ and $X \in \mathbf{M}^n$

$$X \in \mathbf{M}^n_{\mathcal{L}}$$

means X has imaginary eig/vals with

- geometric multiplicities 1
- algebraic multiplicities constant, listed (downward) by $\mathcal{L}$.

Thm (Arnold, '71) $\mathbf{M}^n_{\mathcal{L}}$ a manifold.

Thm $\mathbf{M}^n_-$ partly smooth rel. to $\mathbf{M}^n_{\mathcal{L}}$.

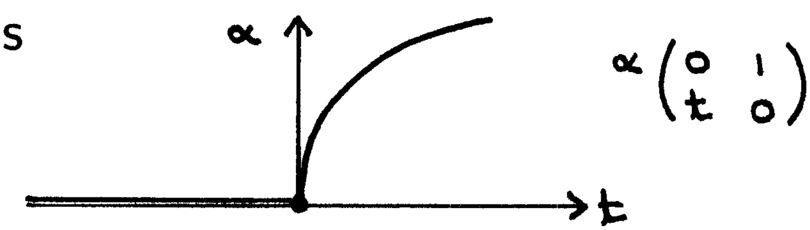
SPECTRAL ABSCISSA

Asymptotic decay for trajectories of $\dot{x} = Ax$ measured by spectral abscissa:

$$\alpha(A) = \max\{\text{Re}\, z : z \text{ eig/val of } A\}.$$

(So, $A \in \mathbf{M}^n_- \Leftrightarrow \alpha(A) \leq 0$.)

Unlike maximum eig/val $\lambda_{\max}$ on $\mathbf{S}^n$, α is



- not convex
- not Lipschitz:

$$\alpha(X) - \alpha(Y) \not\leq k\|X - Y\|.$$

α **is** differentiable a.e., since semi-algebraic: its epigraph defined by polynomial equations and inequalities.

PEAKING AND APPROXIMATION

Classical characterization $\alpha(A) = \inf\{\gamma : 2\gamma P \succeq PA + A^*P,\ P \in \mathbf{H}^n_{++}\}$ (P Hermitian) suggests a "Lyapunov" approximation: for small $\delta > 0$,

$$\alpha^\delta(A) = \inf\{\gamma : 2\gamma P \succeq PA + A^*P,\ I \succeq P \succeq \delta I\}.$$

Advantages

- Stops transient peaks for $\dot{x} = Ax$.
- Characterization attained.
- α^δ Lipschitz and "more" convex:

$$\alpha^1(A) = \lambda_{\max}\left(\frac{A + A^*}{2}\right).$$

Disadvantage (computationally): large auxiliary variable P.

ROBUSTNESS

How do we check

$$A \in \mathbf{S}^n_- \quad \underline{\textbf{robustly}}?$$

Suppose we mean, for small $\epsilon > 0$,

$$\|X - A\| \leq \epsilon \quad \Rightarrow \quad X \in \mathbf{S}^n_-.$$

In terms of functions, we need

$$\lambda_{\max}(A) \leq 0 \quad \text{robustly.}$$

In other words:

$$0 \geq \max_{\|X-A\| \leq \epsilon} \lambda_{\max}(X)$$

(the robust regularization)

$$= \lambda_{\max}(A) + \epsilon.$$

What is robust regularization of spectral abscissa α?

PSEUDOSPECTRA

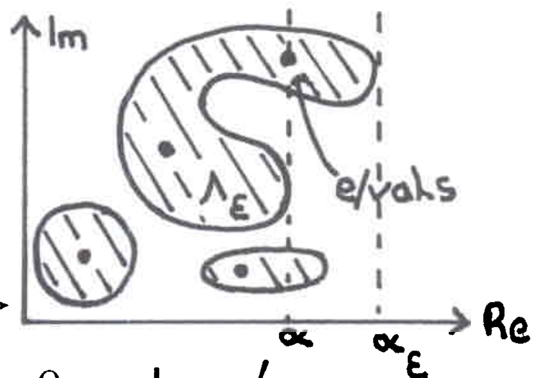
$$\Lambda_\epsilon(A) = \{\text{e/vals of } X : \|X - A\| \le \epsilon\}$$
$$= \{z : \sigma_{\min}(A - zI) \le \epsilon\}$$

(where $\sigma_{\min}$ is smallest singular value).

Robust regularization of α is pseudospectral abscissa:

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

Computing α_ϵ is relatively easy (like finding $\mathbf{H}_\infty$-norms).



Kreiss matrix thm $\Rightarrow$

$$\forall \delta > 0 \;\; \exists \epsilon > \epsilon' > 0 \text{ and } c, c' \text{ so}$$
$$c'\alpha_{\epsilon'} < \alpha^\delta < c\alpha_\epsilon.$$

Hence reducing α_ϵ also discourages transient peaks.

ASIDE: KREISS MATRIX THEOREM

$\mathcal{F} \subset \mathbf{M}^n$ is:

(a) L_2-stable if $p =$

$$\sup_{k \in \mathbf{N},\ A \in \mathcal{F}} \|A^k\| < \infty.$$

(b) strictly Hermitian-stable if $q =$

$$\inf\{\gamma : H_\gamma(A) \neq \emptyset \ \forall A \in \mathcal{F}\} < \infty$$

where $H \in H_\gamma(A)$ means

$$H \succeq A^* H A, \ \gamma I \succeq H \succeq \gamma^{-1} I.$$

(c) pseudospectral stable if $r =$

$$\sup_\epsilon \frac{1}{\epsilon}(|\Lambda_\epsilon(\mathcal{F})| - 1) < \infty.$$

Thm (Kreiss '62) $(a) \Leftrightarrow (b) \Leftrightarrow (c)$, and p, q, r related.

Thm (Spijker '91) $r \leq p \leq enr$.

PSEUDOSPECTRAL ABSCISSA

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

Like the Lyapunov approximation α^δ,

- reducing α_ϵ discourages transient peaks for $\dot{x} = Ax$ (by Kreiss)
- optimum in characterization is attained (but easier to compute)
- α_ϵ approximates α, because $\alpha_\epsilon \longrightarrow \alpha$ "variationally" as $\epsilon \downarrow 0$, so

$$\text{argmin}_F\, \alpha_\epsilon \longrightarrow \text{argmin}_F\, \alpha.$$

- enhances convexity, because, as $\epsilon \longrightarrow \infty$,

$$\alpha_\epsilon(X) - \epsilon \longrightarrow \lambda_{\max}\left(\frac{X + X^*}{2}\right)$$

variationally.

$\mathbf{H}_\infty$-NORM

For stable A, complex stability radius

$$\beta(A) = \inf\{\|X - A\| : X \text{ not stable}\}$$

$$= \frac{1}{\|s \mapsto (sI - A)^{-1}\|_{\mathbf{H}_\infty}}$$

Prop If A maximizes β over a set F, then it also minimizes $\alpha_{\beta(A)}$ over F.

So as ϵ decreases through $\mathbf{R}_+$, set of minimizers $\text{argmin}_F\, \alpha_\epsilon$ evolves through

- minimizers of $\lambda_{\max}$(symmetric part)
- minimizers of $\mathbf{H}_\infty$-norm
- minimizers of spectral abscissa.

LIPSCHITZ PROPERTIES

An advantage of Lyapunov approximation: it's Lipschitz. What about pseudospectral abscissa α_ϵ?

Thm Consider set E and function f:

- E closed, semi-algebraic (or ...)
- f locally Lipschitz off E and "grows sharply" from E:

$$T_E(x) \subset \text{cl}\{y : d(-f)(x)(y) < 0\}.$$

Then robust regularization (for $\epsilon > 0$)

$$g_\epsilon(x) = \sup\{g(y) : \|y - x\| \leq \epsilon\}$$

loc. Lip. around any point, if ϵ small.

Hence α_ϵ loc. Lip. around any fixed nonderogatory matrix, once ϵ small.

SUMMARY

Unlike largest eigenvalue $\lambda_{\max}$ on $\mathbf{S}^n$, the spectral abscissa α on $\mathbf{M}^n$ is

- not convex
- not Lipschitz.

But it <u>**is**</u>:

- regular at nonderogatory matrices
- partly smooth at nonderog. matrices ($\longrightarrow$ easy sensitivity analysis)
- a.e. differentiable ($\longrightarrow$ sampling algorithms).
- We can approximate α, robustly and computably, by the pseudospectral abscissa α_ϵ, which has enhanced convexity and Lipschitz properties.