Eigenvalue optimization: symmetric versus nonsymmetric matrices

Adrian Lewis

Cornell University

October 5, 2004

Collaborators: J.V. Burke, M.L. Overton

1. OUTLINE

- Stability for
  - symmetric matrices

1. OUTLINE

- Stability for
  - symmetric matrices
  - nonsymmetric matrices

1. OUTLINE

- Stability for
  - symmetric matrices
  - nonsymmetric matrices
  - polynomials.

1. OUTLINE

- Stability for
  - symmetric matrices
  - nonsymmetric matrices
  - polynomials.
- Structure and optimal stability

1. OUTLINE

- Stability for
  - symmetric matrices
  - nonsymmetric matrices
  - polynomials.
- Structure and optimal stability
- An example: the Belgian Chocolate problem

1. OUTLINE

- Stability for
  - symmetric matrices
  - nonsymmetric matrices
  - polynomials.
- Structure and optimal stability
- An example: the Belgian Chocolate problem
- Nonsmooth analysis: how to weaken convexity

1. OUTLINE

- Stability for
  - symmetric matrices
  - nonsymmetric matrices
  - polynomials.
- Structure and optimal stability
- An example: the Belgian Chocolate problem
- Nonsmooth analysis: how to weaken convexity
- Robust stability and pseudospectra

2. EIGENVALUES AND MODELLING

The **semidefinite cone**

$$\mathbf{S}^n_- \;=\; \{n \times n \text{ symmetric } X \text{ with negative eigenvalues}\}$$

has extraordinary modelling powers. (Recall Boyd's talk. . .)

2. EIGENVALUES AND MODELLING

The **semidefinite cone**

$$\mathbf{S}_-^n \ = \ \{n \times n \text{ symmetric } X \text{ with negative eigenvalues}\}$$

has extraordinary modelling powers. (Recall Boyd's talk. . .)

What about the cone of **stable** (nonsymmetric) matrices:

$$\mathbf{M}_-^n \ = \ \{X \text{ with negative eigenvalue real parts}\}?$$

2. EIGENVALUES AND MODELLING

The **semidefinite cone**

$$\mathbf{S}^n_- \;=\; \{n \times n \text{ symmetric } X \text{ with negative eigenvalues}\}$$

has extraordinary modelling powers. (Recall Boyd's talk. . .)

What about the cone of **stable** (nonsymmetric) matrices:

$$\mathbf{M}^n_- \;=\; \{X \text{ with negative eigenvalue real parts}\}?$$

Eigenvalues are (theoretically) very significant for models.

2. EIGENVALUES AND MODELLING

The **semidefinite cone**

$$\mathbf{S}^n_- \;=\; \{n \times n \text{ symmetric } X \text{ with negative eigenvalues}\}$$

has extraordinary modelling powers. (Recall Boyd's talk. . .)

What about the cone of **stable** (nonsymmetric) matrices:

$$\mathbf{M}^n_- \;=\; \{X \text{ with negative eigenvalue real parts}\}?$$

Eigenvalues are (theoretically) very significant for models. Eg:

trajectories of $\frac{dx}{dt} = Ax$ explode exponentially $\quad \Leftrightarrow \quad A \notin \mathbf{M}^n_-$.

2. EIGENVALUES AND MODELLING

The **semidefinite cone**

$$\mathbf{S}_{-}^{n} = \{n \times n \text{ symmetric } X \text{ with negative eigenvalues}\}$$

has extraordinary modelling powers. (Recall Boyd's talk. . .)

What about the cone of **stable** (nonsymmetric) matrices:

$$\mathbf{M}_{-}^{n} = \{X \text{ with negative eigenvalue real parts}\}?$$

Eigenvalues are (theoretically) very significant for models. Eg:

trajectories of $\frac{dx}{dt} = Ax$ explode exponentially $\quad \Leftrightarrow \quad A \notin \mathbf{M}_{-}^{n}$.

In optimization,

how do $\mathbf{M}_{-}^{n}$ and $\mathbf{S}_{-}^{n}$ compare?

3. POLYNOMIALS

An even simpler question: **stable polynomials** $\mathbf{P}_{-}^{n}$ =

$$\{\text{polynomials } p \text{ of degree} \le n \ : \ p(z)=0 \ \Rightarrow \ \mathrm{Re}\, z \le 0\ \}\,.$$

3. POLYNOMIALS

An even simpler question: **stable polynomials $\mathbf{P}_{-}^{n}$** $=$

$$\{\text{polynomials } p \text{ of degree} \le n \ : \ p(z) = 0 \ \Rightarrow \ \operatorname{Re} z \le 0 \ \}.$$

How do $\mathbf{P}_{-}^{n}$ and $\mathbf{S}_{-}^{n}$ compare?

3. POLYNOMIALS

An even simpler question: **stable polynomials** $\mathbf{P}^n_- =$

$$\{\text{polynomials } p \text{ of degree} \leq n \;:\; p(z) = 0 \;\Rightarrow\; \operatorname{Re} z \leq 0\,\}\,.$$

How do $\mathbf{P}^n_-$ and $\mathbf{S}^n_-$ compare?

Note stable polynomials and polynomials are closely related:

3. POLYNOMIALS

An even simpler question: **stable polynomials $\mathbf{P}_-^n$** $=$

$$\{\text{polynomials } p \text{ of degree} \le n \ : \ p(z) = 0 \ \Rightarrow \ \operatorname{Re} z \le 0 \ \}.$$

How do $\mathbf{P}_-^n$ and $\mathbf{S}_-^n$ compare?

Note stable polynomials and polynomials are closely related:

$$z^n - \sum_{j=0}^{n-1} a_j z^j \ \in \ \mathbf{P}_-^n$$

exactly when

$$\begin{bmatrix} a_{n-1} & 1 & 0 & \cdots & 0 \\ a_{n-2} & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_1 & 0 & 0 & \ddots & 1 \\ a_0 & 0 & 0 & \cdots & 0 \end{bmatrix} \ \in \ \mathbf{M}_-^n.$$

3. POLYNOMIALS

An even simpler question: **stable polynomials $\mathbf{P}^n_-$** $=$

$$\{\text{polynomials } p \text{ of degree} \leq n \ : \ p(z) = 0 \ \Rightarrow \ \mathrm{Re}\, z \leq 0 \ \}.$$

How do $\mathbf{P}^n_-$ and $\mathbf{S}^n_-$ compare?

Note stable polynomials and polynomials are closely related:

$$z^n - \sum_{j=0}^{n-1} a_j z^j \ \in \ \mathbf{P}^n_-$$

exactly when

$$\begin{bmatrix} a_{n-1} & 1 & 0 & \cdots & 0 \\ a_{n-2} & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_1 & 0 & 0 & \ddots & 1 \\ a_0 & 0 & 0 & \cdots & 0 \end{bmatrix} \ \in \ \mathbf{M}^n_-.$$

Conversely, $X \in \mathbf{M}^n_- \quad \Leftrightarrow \quad \det(zI - X) \in \mathbf{P}^n_-$.

4. CONVEXITY

Why does semidefinite programming work?

4. CONVEXITY

Why does semidefinite programming work?

Blend of global algebraic properties of $\mathbf{S}^n_-$, especially **convexity**:

$$\mathbf{S}^n_- \;=\; \bigcap_y \{X : y^T X y \leq 0\}.$$

4. CONVEXITY

Why does semidefinite programming work?

Blend of global algebraic properties of $\mathbf{S}^n_-$, especially **convexity**:

$$\mathbf{S}^n_- \;=\; \bigcap_y \{X : y^T X y \leq 0\}.$$

But $\mathbf{M}^n_-$ **isn't** convex:

4. CONVEXITY

Why does semidefinite programming work?

Blend of global algebraic properties of $\mathbf{S}^n_-$, especially **convexity**:

$$\mathbf{S}^n_- \;=\; \bigcap_y \{X : y^T X y \le 0\}.$$

But $\mathbf{M}^n_-$ **isn't** convex:

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

4. CONVEXITY

Why does semidefinite programming work?

Blend of global algebraic properties of $\mathbf{S}^n_-$, especially **convexity**:

$$\mathbf{S}^n_- \;=\; \bigcap_y \{X : y^T X y \le 0\}.$$

But $\mathbf{M}^n_-$ **isn't** convex:

$$\begin{array}{llll} & \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} & = & \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix} \\ \text{eigenvalues} & \{0,-2\} \qquad \{0,0\} & & \{-1 \pm \sqrt{2}\} \end{array}$$

4. CONVEXITY

Why does semidefinite programming work?

Blend of global algebraic properties of $\mathbf{S}^n_-$, especially **convexity**:

$$\mathbf{S}^n_- \;=\; \bigcap_y \{X : y^T X y \leq 0\}.$$

But $\mathbf{M}^n_-$ **isn't** convex:

$$\begin{array}{lccccc}
 & \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} & + & \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} & = & \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix} \\
\text{eigenvalues} & \{0,-2\} & & \{0,0\} & & \{-1 \pm \sqrt{2}\} \\
 & \text{stable} & + & \text{stable} & = & \text{unstable.}
\end{array}$$

4. CONVEXITY

Why does semidefinite programming work?

Blend of global algebraic properties of $\mathbf{S}^n_-$, especially **convexity**:

$$\mathbf{S}^n_- \;=\; \bigcap_y \{X : y^T X y \leq 0\}.$$

But $\mathbf{M}^n_-$ **isn't** convex:

$$\begin{array}{lccccc}
 & \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} & + & \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} & = & \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix} \\
\text{eigenvalues} & \{0, -2\} & & \{0, 0\} & & \{-1 \pm \sqrt{2}\} \\
 & \text{stable} & + & \text{stable} & = & \text{unstable.}
\end{array}$$

So, **global** optimization over $\mathbf{M}^n_-$ is hopeless.

4. CONVEXITY

Why does semidefinite programming work?

Blend of global algebraic properties of $\mathbf{S}^n_-$, especially **convexity**:

$$\mathbf{S}^n_- \;=\; \bigcap_y \{X : y^T X y \leq 0\}.$$

But $\mathbf{M}^n_-$ **isn't** convex:

$$\begin{array}{lccccc} & \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} & + & \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} & = & \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix} \\ \text{eigenvalues} & \{0,-2\} & & \{0,0\} & & \{-1 \pm \sqrt{2}\} \\ & \text{stable} & + & \text{stable} & = & \text{unstable.} \end{array}$$

So, **global** optimization over $\mathbf{M}^n_-$ is hopeless.

What about **local** optimization?

5. SENSITIVITY

For $A_1, \ldots, A_k, B \in \mathbf{S}^n$, suppose the semidefinite program

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \sum_j x_j A_j - B \;\in\; \mathbf{S}^n_- \end{array}$$

has optimal solution $x(c)$.

5. SENSITIVITY

For $A_1, \ldots, A_k, B \in \mathbf{S}^n$, suppose the semidefinite program

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \sum_j x_j A_j - B \;\in\; \mathbf{S}^n_- \end{array}$$

has optimal solution $x(c)$.

Usually

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

stays constant for small changes to c,

5. SENSITIVITY

For $A_1, \ldots, A_k, B \in \mathbf{S}^n$, suppose the semidefinite program

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \sum_j x_j A_j - B \ \in \ \mathbf{S}^n_- \end{array}$$

has optimal solution $x(c)$.

Usually

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

stays constant for small changes to c, and

$$c \mapsto x(c)$$

is smooth, and computable by solving **smooth equations**.

5. SENSITIVITY

For $A_1, \ldots, A_k, B \in \mathbf{S}^n$, suppose the semidefinite program

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \sum_j x_j A_j - B \ \in \ \mathbf{S}^n_- \end{array}$$

has optimal solution $x(c)$.

Usually

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

stays constant for small changes to c, and

$$c \mapsto x(c)$$

is smooth, and computable by solving **smooth equations**.

Optimization with stable matrices (or polynomials) is similar.

For example...

6. THE BELGIAN CHOCOLATE PROBLEM

Problem: “Simultaneous plant stabilization” (Blondel, 1994).
Find (strictly) stable polynomials $p(z), q(z)$, and

$$(*) \qquad r(z) = (z^2 - 2\delta z + 1)p(z) + (z^2 - 1)q(z)$$

with $\delta = 0.9$.

6. THE BELGIAN CHOCOLATE PROBLEM

Problem: "Simultaneous plant stabilization" (Blondel, 1994). Find (strictly) stable polynomials $p(z), q(z)$, and

$$(*) \qquad r(z) = (z^2 - 2\delta z + 1)p(z) + (z^2 - 1)q(z)$$

with $\delta = 0.9$. (Solved by Patel et al. 2002.)

6. THE BELGIAN CHOCOLATE PROBLEM

Problem: "Simultaneous plant stabilization" (Blondel, 1994). Find (strictly) stable polynomials $p(z), q(z)$, and

$$(*) \qquad r(z) = (z^2 - 2\delta z + 1)p(z) + (z^2 - 1)q(z)$$

with $\delta = 0.9$. (Solved by Patel et al. 2002.)

An optimization approach:

$$\begin{array}{lrcl} \text{minimize} & \alpha & & \\ \text{subject to} & p, q & \in & \mathbf{P}^3_- \\ & r(z+\alpha) & \in & \mathbf{P}^5_- \\ & (*) \text{ holds.} & & \end{array}$$

6. THE BELGIAN CHOCOLATE PROBLEM

Problem: "Simultaneous plant stabilization" (Blondel, 1994). Find (strictly) stable polynomials $p(z), q(z)$, and

$$(*) \qquad r(z) = (z^2 - 2\delta z + 1)p(z) + (z^2 - 1)q(z)$$

with $\delta = 0.9$. (Solved by Patel et al. 2002.)

An optimization approach:

$$\begin{array}{lrcl} \text{minimize} & \alpha & & \\ \text{subject to} & p, q & \in & \mathbf{P}^3_- \\ & r(z+\alpha) & \in & \mathbf{P}^5_- \\ & (*) \text{ holds.} & & \end{array}$$

Computationally, the optimal solution, for δ near 0.9, has

- q constant

6. THE BELGIAN CHOCOLATE PROBLEM

Problem: "Simultaneous plant stabilization" (Blondel, 1994). Find (strictly) stable polynomials $p(z), q(z)$, and

$$(*) \qquad r(z) = (z^2 - 2\delta z + 1)p(z) + (z^2 - 1)q(z)$$

with $\delta = 0.9$. (Solved by Patel et al. 2002.)

An optimization approach:

$$\begin{array}{lrcl} \text{minimize} & \alpha & & \\ \text{subject to} & p, q & \in & \mathbf{P}^3_- \\ & r(z+\alpha) & \in & \mathbf{P}^5_- \\ & (*) \text{ holds.} & & \end{array}$$

Computationally, the optimal solution, for δ near 0.9, has

- q constant
- $r(z) = (z - \alpha)^5$

6. THE BELGIAN CHOCOLATE PROBLEM

Problem: "Simultaneous plant stabilization" (Blondel, 1994). Find (strictly) stable polynomials $p(z), q(z)$, and

$$(*) \qquad r(z) = (z^2 - 2\delta z + 1)p(z) + (z^2 - 1)q(z)$$

with $\delta = 0.9$. (Solved by Patel et al. 2002.)

An optimization approach:

$$\begin{array}{lrcl} \text{minimize} & & \alpha & \\ \text{subject to} & p, q & \in & \mathbf{P}_-^3 \\ & r(z+\alpha) & \in & \mathbf{P}_-^5 \\ & (*) \text{ holds.} & & \end{array}$$

Computationally, the optimal solution, for δ near 0.9, has

- q constant
- $r(z) = (z - \alpha)^5$
- $\alpha < 0$.

6. THE BELGIAN CHOCOLATE PROBLEM

Problem: "Simultaneous plant stabilization" (Blondel, 1994). Find (strictly) stable polynomials $p(z), q(z)$, and

$$(*) \qquad r(z) = (z^2 - 2\delta z + 1)p(z) + (z^2 - 1)q(z)$$

with $\delta = 0.9$. (Solved by Patel et al. 2002.)

An optimization approach:

$$\begin{array}{lrcl} \text{minimize} & \alpha & & \\ \text{subject to} & p, q & \in & \mathbf{P}^3_- \\ & r(z+\alpha) & \in & \mathbf{P}^5_- \\ & (*) \text{ holds.} & & \end{array}$$

Computationally, the optimal solution, for δ near 0.9, has

- q constant
- $r(z) = (z - \alpha)^5$
- $\alpha < 0$. **CHOCOLATE!**

7. QUESTIONS

1. **Why does structure persist in optimal solutions?**

7. QUESTIONS

1. **Why does structure persist in optimal solutions?**

   Coming up. . .

7. QUESTIONS

1. **Why does structure persist in optimal solutions?**

   Coming up. . .

2. **Can we exploit it?**

7. QUESTIONS

1. **Why does structure persist in optimal solutions?**

   Coming up. . .

2. **Can we exploit it?**

   Yes: eg closed form chocolate solutions. Algorithms?

7. QUESTIONS

1. **Why does structure persist in optimal solutions?**

   Coming up. . .

2. **Can we exploit it?**

   Yes: eg closed form chocolate solutions. Algorithms?

3. **Practically, are multiple zeros worrisome?**

7. QUESTIONS

1. **Why does structure persist in optimal solutions?**

   Coming up...

2. **Can we exploit it?**

   Yes: eg closed form chocolate solutions. Algorithms?

3. **Practically, are multiple zeros worrisome?**

   Yes: multiple zeros may be very sensitive:

$$(z+\epsilon)^3 \;\text{ stable for small } \epsilon > 0$$
$$(z+\epsilon)^3 - \epsilon^3 \;\text{ isn't.}$$

7. QUESTIONS

1. **Why does structure persist in optimal solutions?**

   Coming up...

2. **Can we exploit it?**

   Yes: eg closed form chocolate solutions. Algorithms?

3. **Practically, are multiple zeros worrisome?**

   Yes: multiple zeros may be very sensitive:

   $$(z+\epsilon)^3 \ \text{ stable for small } \epsilon > 0$$
   $$(z+\epsilon)^3 - \epsilon^3 \ \text{ isn't.}$$

   Eigenvalues don't really predict the dynamics of

   $$\frac{dx}{dt} = \begin{bmatrix} -\epsilon & 1 & 0 \\ 0 & -\epsilon & 1 \\ 0 & 0 & -\epsilon \end{bmatrix} x, \qquad x(0) = \begin{bmatrix} 0 \\ 0 \\ 2e \end{bmatrix}.$$

7. QUESTIONS

1. **Why does structure persist in optimal solutions?**

   Coming up...

2. **Can we exploit it?**

   Yes: eg closed form chocolate solutions. Algorithms?

3. **Practically, are multiple zeros worrisome?**

   Yes: multiple zeros may be very sensitive:

   $$\begin{array}{l} (z+\epsilon)^3 \ \text{ stable for small } \epsilon > 0 \\ (z+\epsilon)^3 - \epsilon^3 \ \text{ isn't.} \end{array}$$

   Eigenvalues don't really predict the dynamics of

   $$\frac{dx}{dt} = \begin{bmatrix} -\epsilon & 1 & 0 \\ 0 & -\epsilon & 1 \\ 0 & 0 & -\epsilon \end{bmatrix} x, \qquad x(0) = \begin{bmatrix} 0 \\ 0 \\ 2e \end{bmatrix}.$$

   Here, $x(t) \to 0$ **eventually**, but $x(\epsilon^{-1}) = \epsilon^{-2}$.

8. TANGENTS AND OPTIMALITY

In nonlinear optimization, we recognize optimality by **linearization**.

8. TANGENTS AND OPTIMALITY

In nonlinear optimization, we recognize optimality by **linearization**.

For a point x in a closed set S, vectors

$$y = \lim_k \frac{y_k}{\|y_k\|} \quad \text{where} \quad y_k \to 0 \text{ in } S - x$$

are called **unit tangents**.

8. TANGENTS AND OPTIMALITY

In nonlinear optimization, we recognize optimality by **linearization**.

For a point x in a closed set S, vectors

$$y = \lim_k \frac{y_k}{\|y_k\|} \quad \text{where} \quad y_k \to 0 \text{ in } S - x$$

are called **unit tangents**.

If x maximizes $\langle c, \cdot \rangle$ over S, then $\langle c, y \rangle \leq 0$ for all such y.

8. TANGENTS AND OPTIMALITY

In nonlinear optimization, we recognize optimality by **linearization**.

For a point x in a closed set S, vectors

$$y = \lim_k \frac{y_k}{\|y_k\|} \quad \text{where} \quad y_k \to 0 \text{ in } S - x$$

are called **unit tangents**.

If x maximizes $\langle c, \cdot \rangle$ over S, then $\langle c, y \rangle \leq 0$ for all such y.

Unit tangents to **convex** sets can be realized by smooth paths:

$$y \;=\; \lim_{t \downarrow 0} \frac{y(t)}{t}.$$

8. TANGENTS AND OPTIMALITY

In nonlinear optimization, we recognize optimality by **linearization**.

For a point x in a closed set S, vectors

$$y = \lim_k \frac{y_k}{\|y_k\|} \quad \text{where} \quad y_k \to 0 \text{ in } S - x$$

are called **unit tangents**.

If x maximizes $\langle c, \cdot \rangle$ over S, then $\langle c, y \rangle \leq 0$ for all such y.

Unit tangents to **convex** sets can be realized by smooth paths:

$$y \ = \ \lim_{t \downarrow 0} \frac{y(t)}{t}.$$

Helpful to

- calculate tangents

8. TANGENTS AND OPTIMALITY

In nonlinear optimization, we recognize optimality by **linearization**.

For a point x in a closed set S, vectors

$$y = \lim_k \frac{y_k}{\|y_k\|} \quad \text{where} \quad y_k \to 0 \text{ in } S - x$$

are called **unit tangents**.

If x maximizes $\langle c, \cdot \rangle$ over S, then $\langle c, y \rangle \leq 0$ for all such y.

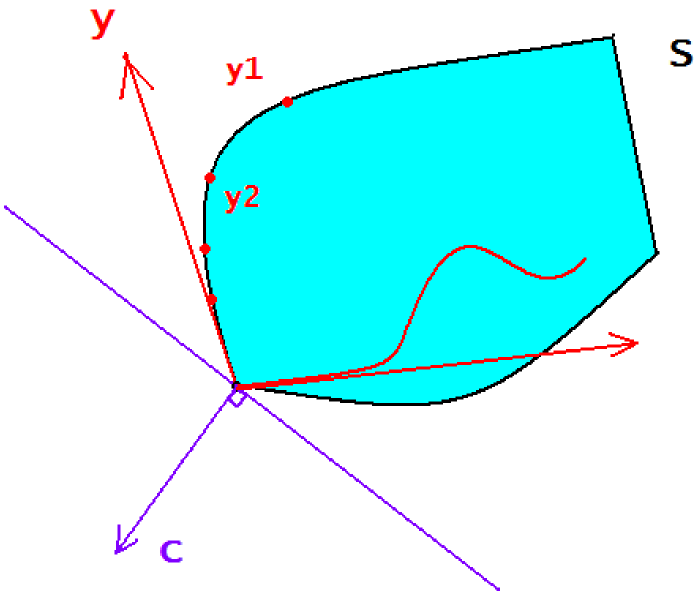
Unit tangents to **convex** sets can be realized by smooth paths:

$$y \;=\; \lim_{t \downarrow 0} \frac{y(t)}{t}.$$

Helpful to

- calculate tangents
- find improving paths.

9. GEOMETRY OF TANGENTS



10. TANGENTS TO NONCONVEX SETS

In general, we can't realize tangents with paths.

10. TANGENTS TO NONCONVEX SETS

In general, we can't realize tangents with paths. Eg:

$$1 \text{ is a tangent to } \left\{1, \frac{1}{2}, \frac{1}{3}, \dots, 0\right\} \text{ at } 0.$$

10. TANGENTS TO NONCONVEX SETS

In general, we can't realize tangents with paths. Eg:

$$1 \text{ is a tangent to } \left\{1, \frac{1}{2}, \frac{1}{3}, \ldots, 0\right\} \text{ at } 0.$$

But convexity is not crucial. Paths work for

$$\{x : g_i(x) \leq 0 \text{ for } i = 1, \ldots, m\}.$$

in two important cases.

10. TANGENTS TO NONCONVEX SETS

In general, we can't realize tangents with paths. Eg:

$$1 \text{ is a tangent to } \left\{1, \frac{1}{2}, \frac{1}{3}, \ldots, 0\right\} \text{ at } 0.$$

But convexity is not crucial. Paths work for

$$\{x : g_i(x) \leq 0 \text{ for } i = 1, \ldots, m\}.$$

in two important cases.

Smooth nonlinear constraints: The g_i are smooth with linearly independent gradients (by the implicit function theorem).

10. TANGENTS TO NONCONVEX SETS

In general, we can't realize tangents with paths. Eg:

$$1 \text{ is a tangent to } \left\{1, \frac{1}{2}, \frac{1}{3}, \ldots, 0\right\} \text{ at } 0.$$

But convexity is not crucial. Paths work for

$$\{x : g_i(x) \leq 0 \text{ for } i = 1, \ldots, m\}.$$

in two important cases.

Smooth nonlinear constraints: The g_i are smooth with linearly independent gradients (by the implicit function theorem). But we can't express stability this way...

10. TANGENTS TO NONCONVEX SETS

In general, we can't realize tangents with paths. Eg:

$$1 \text{ is a tangent to } \left\{1, \frac{1}{2}, \frac{1}{3}, \ldots, 0\right\} \text{ at } 0.$$

But convexity is not crucial. Paths work for

$$\{x : g_i(x) \leq 0 \text{ for } i = 1, \ldots, m\}.$$

in two important cases.

Smooth nonlinear constraints: The g_i are smooth with linearly independent gradients (by the implicit function theorem). But we can't express stability this way...

Semi-algebraic sets: The g_i are polynomials (by the "curve selection lemma").

10. TANGENTS TO NONCONVEX SETS

In general, we can't realize tangents with paths. Eg:

$$1 \text{ is a tangent to } \left\{1, \frac{1}{2}, \frac{1}{3}, \ldots, 0\right\} \text{ at } 0.$$

But convexity is not crucial. Paths work for

$$\{x : g_i(x) \leq 0 \text{ for } i = 1, \ldots, m\}.$$

in two important cases.

Smooth nonlinear constraints: The g_i are smooth with linearly independent gradients (by the implicit function theorem). But we can't express stability this way...

Semi-algebraic sets: The g_i are polynomials (by the "curve selection lemma").

The stable sets $\mathbf{P}^n_-$ and $\mathbf{M}^n_-$ **are** semi-algebraic (by "quantifier elimination"), so tangents come from (analytic) paths.

11. NORMALS TO CONVEX SETS

Rewriting the optimality condition for maximizing $\langle c, \cdot \rangle$ over a closed convex set S gives a **normal cone** condition:

$$c \in N_S(x) = \{d : \langle d, y \rangle \leq 0 \text{ for all } y \in S - x\}$$

11. NORMALS TO CONVEX SETS

Rewriting the optimality condition for maximizing $\langle c, \cdot \rangle$ over a closed convex set S gives a **normal cone** condition:

$$
\begin{aligned}
c \;\in\; N_S(x) \;&=\; \{d : \langle d, y \rangle \leq 0 \text{ for all } y \in S - x\} \\
&=\; \{d : \langle d, y \rangle \leq 0 \text{ for all unit tangents } y\}
\end{aligned}
$$

11. NORMALS TO CONVEX SETS

Rewriting the optimality condition for maximizing $\langle c, \cdot \rangle$ over a closed convex set S gives a **normal cone** condition:

$$\begin{aligned} c \in N_S(x) &= \{d : \langle d, y \rangle \leq 0 \text{ for all } y \in S - x\} \\ &= \{d : \langle d, y \rangle \leq 0 \text{ for all unit tangents } y\} \end{aligned}$$

Key property: the mapping $N_S(\cdot)$ is **closed** at every $x \in S$:

11. NORMALS TO CONVEX SETS

Rewriting the optimality condition for maximizing $\langle c, \cdot \rangle$ over a closed convex set S gives a **normal cone** condition:

$$\begin{aligned} c \;\in\; N_S(x) \;&=\; \{d : \langle d, y\rangle \le 0 \text{ for all } y \in S - x\} \\ &=\; \{d : \langle d, y\rangle \le 0 \text{ for all unit tangents } y\} \end{aligned}$$

Key property: the mapping $N_S(\cdot)$ is **closed** at every $x \in S$:

$$\begin{array}{rr} x_r \in S & x_r \to x \\ d_r \in N_S(x_r), & d_r \to d \end{array}$$

implies $d \in N_S(x)$.

11. NORMALS TO CONVEX SETS

Rewriting the optimality condition for maximizing $\langle c, \cdot \rangle$ over a closed convex set S gives a **normal cone** condition:

$$
\begin{aligned}
c \;\in\; N_S(x) \;&=\; \{d : \langle d, y \rangle \le 0 \text{ for all } y \in S - x\} \\
&=\; \{d : \langle d, y \rangle \le 0 \text{ for all unit tangents } y\}
\end{aligned}
$$

Key property: the mapping $N_S(\cdot)$ is **closed** at every $x \in S$:

$$
\begin{array}{rl}
x_r \in S & \quad x_r \to x \\
d_r \in N_S(x_r), & \quad d_r \to d
\end{array}
$$

implies $d \in N_S(x)$.

How can we relate

- normal cone behavior for $\mathbf{S}^n_-$
- solution rank persistence in semidefinite programming?

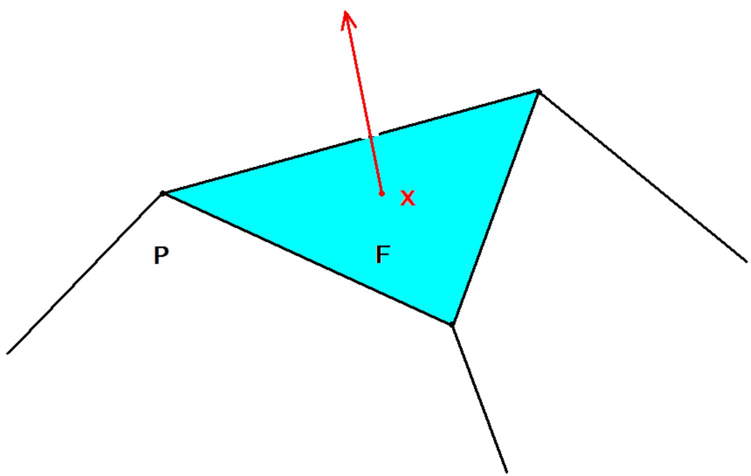
12. NORMALS AND FACIAL STRUCTURE

A polyhedron P decomposes into faces F of various dimensions.

12. NORMALS AND FACIAL STRUCTURE

A polyhedron P decomposes into faces F of various dimensions. As $x \in F$ varies,

- $N_P(x)$ stays constant
- $N_P(x) - N_P(x) = N_F(x)$ (so P is "**sharp**" around F).



13. DECOMPOSING THE SEMIDEFINITE CONE

Solution rank persists in semidefinite programming **not** because of faces.

13. DECOMPOSING THE SEMIDEFINITE CONE

Solution rank persists in semidefinite programming **not** because of faces.

Instead, $\mathbf{S}^n_-$ decomposes into **manifolds**

$$\mathcal{M}_r \;=\; \{X \in \mathbf{S}^n_- : \mathrm{rank}\, X = r\} \qquad (r = 0, 1, \ldots, n)$$

13. DECOMPOSING THE SEMIDEFINITE CONE

Solution rank persists in semidefinite programming **not** because of faces.

Instead, $\mathbf{S}^n_-$ decomposes into **manifolds**

$$\mathcal{M}_r \;=\; \{X \in \mathbf{S}^n_- : \text{rank}\, X = r\} \qquad (r = 0, 1, \ldots, n)$$

As $x \in \mathcal{M}_r$ varies, according to (Oustry, 2000),

- $N_{\mathbf{S}^n_-}(x)$ varies continuously
- $N_{\mathbf{S}^n_-}(x) - N_{\mathbf{S}^n_-}(x) = N_{\mathcal{M}_r}(x)$.

13. DECOMPOSING THE SEMIDEFINITE CONE

Solution rank persists in semidefinite programming **not** because of faces.

Instead, $\mathbf{S}^n_-$ decomposes into **manifolds**

$$\mathcal{M}_r \;=\; \{X \in \mathbf{S}^n_- : \text{rank}\, X = r\} \qquad (r = 0, 1, \ldots, n)$$

As $x \in \mathcal{M}_r$ varies, according to (Oustry, 2000),

- $N_{\mathbf{S}^n_-}(x)$ varies continuously
- $N_{\mathbf{S}^n_-}(x) - N_{\mathbf{S}^n_-}(x) = N_{\mathcal{M}_r}(x)$.

So, $\mathcal{M}_r$ is an **identifiable surface** (Wright, 1993) of $\mathbf{S}^n_-$.

14. NORMALS TO NONCONVEX SETS

Realizing tangents with paths is one way to weaken convexity. Another. . .

14. NORMALS TO NONCONVEX SETS

Realizing tangents with paths is one way to weaken convexity. Another...

For a point x in a closed set S, define

$$N_S(x) \;=\; \{d : \langle d, y\rangle \le 0 \text{ for all unit tangents } y\}.$$

14. NORMALS TO NONCONVEX SETS

Realizing tangents with paths is one way to weaken convexity. Another...

For a point x in a closed set S, define

$$N_S(x) \ = \ \{d : \langle d, y \rangle \leq 0 \text{ for all unit tangents } y\}.$$

F is **Clarke regular** at x if the mapping $N_F(\cdot)$ is closed at x.

14. NORMALS TO NONCONVEX SETS

Realizing tangents with paths is one way to weaken convexity. Another. . .

For a point x in a closed set S, define

$$N_S(x) \;=\; \{d : \langle d, y\rangle \le 0 \text{ for all unit tangents } y\}.$$

F is **Clarke regular** at x if the mapping $N_F(\cdot)$ is closed at x.

Then $N_F(x)$ is called the **normal cone**, and the **tangent cone**

$$\mathbf{R}_+\{\text{unit tangents}\}$$

is convex.

14. NORMALS TO NONCONVEX SETS

Realizing tangents with paths is one way to weaken convexity. Another...

For a point x in a closed set S, define

$$N_S(x) \;=\; \{d : \langle d, y\rangle \le 0 \text{ for all unit tangents } y\}.$$

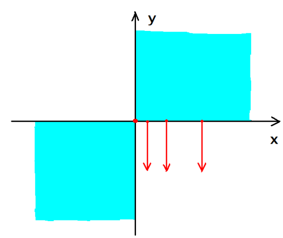
F is **Clarke regular** at x if the mapping $N_F(\cdot)$ is closed at x.

Then $N_F(x)$ is called the **normal cone**, and the **tangent cone**

$$\mathbf{R}_+\{\text{unit tangents}\}$$

is convex.

Convex sets are regular, but $\{(x,y) : xy \ge 0\}$ is **not** regular at the origin.



15. THE CHAIN RULE

Clarke regular sets have a nice calculus.

Theorem (Clarke, 1975) Suppose the map G is smooth and the set S is regular at $G(z)$, with

$$\mathrm{Null}(\nabla G(z)^*) \cap N_S(G(z)) \;=\; \{0\}.$$

Then the inverse image $G^{-1}(S)$ is regular at z, with

$$N_{G^{-1}(S)}(z) \;=\; \nabla G(z)^* N_S(G(z)).$$

15. THE CHAIN RULE

Clarke regular sets have a nice calculus.

Theorem (Clarke, 1975) Suppose the map G is smooth and the set S is regular at $G(z)$, with

$$\mathrm{Null}(\nabla G(z)^*) \cap N_S(G(z)) \;=\; \{0\}.$$

Then the inverse image $G^{-1}(S)$ is regular at z, with

$$N_{G^{-1}(S)}(z) \;=\; \nabla G(z)^* N_S(G(z)).$$

Example A smoothly constrained set

$$F \;=\; \{x : g_i(x) \leq 0 \text{ for } i = 1, \ldots, m\}$$

is regular at x, with

$$N_F(x) \;=\; \left\{ \sum_i \lambda_i \nabla g_i(x) : \lambda \geq 0 \right\},$$

providing $\{\nabla g_i(x)\}$ linearly independent.

16. STABILITY AND REGULARITY

The set of stable polynomials is **highly** nonconvex, and yet. . .

16. STABILITY AND REGULARITY

The set of stable polynomials is **highly** nonconvex, and yet...

Theorem (Burke/Overton, 2001) The set of stable monics

$$\{\text{degree } n-1 \text{ polynomials } p : z^n + p(z) \text{ stable}\}$$

is everywhere regular (with normal cone...).

16. STABILITY AND REGULARITY

The set of stable polynomials is **highly** nonconvex, and yet...

Theorem (Burke/Overton, 2001) The set of stable monics

$$\{\text{degree } n-1 \text{ polynomials } p : z^n + p(z) \text{ stable}\}$$

is everywhere regular (with normal cone...).

$X \in \mathbf{M}^n$ **nonderogatory** if every eigenspace is 1-dimensional. (The "typical" situation even for multiple eigenvalues, unlike $\mathbf{S}^n$.)

16. STABILITY AND REGULARITY

The set of stable polynomials is **highly** nonconvex, and yet...

Theorem (Burke/Overton, 2001) The set of stable monics

$$\{\text{degree } n-1 \text{ polynomials } p : z^n + p(z) \text{ stable}\}$$

is everywhere regular (with normal cone...).

$X \in \mathbf{M}^n$ **nonderogatory** if every eigenspace is 1-dimensional. (The "typical" situation even for multiple eigenvalues, unlike $\mathbf{S}^n$.)

Lemma The "characteristic polynomial" map G on $\mathbf{M}^n$,

$$G(X) \;=\; \det(zI - X) - z^n$$

has $\nabla G(X)$ onto if and only if X is nonderogatory.

16. STABILITY AND REGULARITY

The set of stable polynomials is **highly** nonconvex, and yet...

Theorem (Burke/Overton, 2001) The set of stable monics

$$\{\text{degree } n-1 \text{ polynomials } p : z^n + p(z) \text{ stable}\}$$

is everywhere regular (with normal cone...).

$X \in \mathbf{M}^n$ **nonderogatory** if every eigenspace is 1-dimensional. (The "typical" situation even for multiple eigenvalues, unlike $\mathbf{S}^n$.)

Lemma The "characteristic polynomial" map G on $\mathbf{M}^n$,

$$G(X) \;=\; \det(zI - X) - z^n$$

has $\nabla G(X)$ onto if and only if X is nonderogatory.

Corollary The stable cone $\mathbf{M}^n_-$ is regular at every nonderogatory matrix (with normal cone...).

17. PARTIAL SMOOTHNESS

Closed S is **partly smooth** relative to a manifold $\mathcal{M} \subset S$ if

17. PARTIAL SMOOTHNESS

Closed S is **partly smooth** relative to a manifold $\mathcal{M} \subset S$ if

- (regularity) S is regular throughout $\mathcal{M}$

17. PARTIAL SMOOTHNESS

Closed S is **partly smooth** relative to a manifold $\mathcal{M} \subset S$ if

- (regularity) S is regular throughout $\mathcal{M}$
- (sharpness) $N_S(x) - N_S(x) = N_{\mathcal{M}}(x)$ for all $x \in \mathcal{M}$

17. PARTIAL SMOOTHNESS

Closed S is **partly smooth** relative to a manifold $\mathcal{M} \subset S$ if

- (regularity) S is regular throughout $\mathcal{M}$
- (sharpness) $N_S(x) - N_S(x) = N_{\mathcal{M}}(x)$ for all $x \in \mathcal{M}$
- (continuity) $x \in \mathcal{M} \mapsto N_S(x)$ is continuous.

17. PARTIAL SMOOTHNESS

Closed S is **partly smooth** relative to a manifold $\mathcal{M} \subset S$ if

- (regularity) S is regular throughout $\mathcal{M}$
- (sharpness) $N_S(x) - N_S(x) = N_{\mathcal{M}}(x)$ for all $x \in \mathcal{M}$
- (continuity) $x \in \mathcal{M} \mapsto N_S(x)$ is continuous.

So, $\mathbf{S}^n_-$ is partly smooth relative to each $\mathcal{M}_r$

17. PARTIAL SMOOTHNESS

Closed S is **partly smooth** relative to a manifold $\mathcal{M} \subset S$ if

- (regularity) S is regular throughout $\mathcal{M}$
- (sharpness) $N_S(x) - N_S(x) = N_{\mathcal{M}}(x)$ for all $x \in \mathcal{M}$
- (continuity) $x \in \mathcal{M} \mapsto N_S(x)$ is continuous.

So, $\mathbf{S}^n_-$ is partly smooth relative to each $\mathcal{M}_r$

Can we decompose $\mathbf{P}^n_-$ and $\mathbf{M}^n_-$ analogously?

17. PARTIAL SMOOTHNESS

Closed S is **partly smooth** relative to a manifold $\mathcal{M} \subset S$ if

- (regularity) S is regular throughout $\mathcal{M}$
- (sharpness) $N_S(x) - N_S(x) = N_{\mathcal{M}}(x)$ for all $x \in \mathcal{M}$
- (continuity) $x \in \mathcal{M} \mapsto N_S(x)$ is continuous.

So, $\mathbf{S}^n_-$ is partly smooth relative to each $\mathcal{M}_r$

Can we decompose $\mathbf{P}^n_-$ and $\mathbf{M}^n_-$ analogously?

Theorem If $\mathcal{L}$ is any list of numbers, $\mathbf{P}^n_-$ is partly smooth relative to the manifold of those polynomials having imaginary zeros with multiplicities listed (downward) by $\mathcal{L}$.

17. PARTIAL SMOOTHNESS

Closed S is **partly smooth** relative to a manifold $\mathcal{M} \subset S$ if

- (regularity) S is regular throughout $\mathcal{M}$
- (sharpness) $N_S(x) - N_S(x) = N_{\mathcal{M}}(x)$ for all $x \in \mathcal{M}$
- (continuity) $x \in \mathcal{M} \mapsto N_S(x)$ is continuous.

So, $\mathbf{S}^n_-$ is partly smooth relative to each $\mathcal{M}_r$

Can we decompose $\mathbf{P}^n_-$ and $\mathbf{M}^n_-$ analogously?

Theorem If $\mathcal{L}$ is any list of numbers, $\mathbf{P}^n_-$ is partly smooth relative to the manifold of those polynomials having imaginary zeros with multiplicities listed (downward) by $\mathcal{L}$.

The same chain rule holds for partly smooth sets.
So, $\mathbf{M}^n_-$ is also partly smooth, near nonderogatory matrices.

18. SENSITIVITY

Partial smoothness unifies **active set** ideas in optimization. . .

18. SENSITIVITY

Partial smoothness unifies **active set** ideas in optimization. . .

Theorem Suppose S is partly smooth relative to $\mathcal{M}$, and

- "strict complementarity": $\bar{c} \in \operatorname{ri} N_S(\bar{x})$

18. SENSITIVITY

Partial smoothness unifies **active set** ideas in optimization...

Theorem Suppose S is partly smooth relative to $\mathcal{M}$, and

- "strict complementarity": $\bar{c} \in \mathrm{ri}\, N_S(\bar{x})$
- "second-order sufficiency" for maximizing $\langle \bar{c}, \cdot \rangle$ on $\mathcal{M}$.

18. SENSITIVITY

Partial smoothness unifies **active set** ideas in optimization. . .

Theorem Suppose S is partly smooth relative to $\mathcal{M}$, and

- "strict complementarity": $\bar{c} \in \mathrm{ri}\, N_S(\bar{x})$
- "second-order sufficiency" for maximizing $\langle \bar{c}, \cdot \rangle$ on $\mathcal{M}$.

Then for c near $\bar{c}$, the optimality conditions

$$c \in N_S(x)$$

has a unique solution $x(c) \in \mathcal{M}$ near $\bar{x}$,

18. SENSITIVITY

Partial smoothness unifies **active set** ideas in optimization...

Theorem Suppose S is partly smooth relative to $\mathcal{M}$, and

- "strict complementarity": $\bar{c} \in \operatorname{ri} N_S(\bar{x})$
- "second-order sufficiency" for maximizing $\langle \bar{c}, \cdot \rangle$ on $\mathcal{M}$.

Then for c near $\bar{c}$, the optimality conditions

$$c \in N_S(x)$$

has a unique solution $x(c) \in \mathcal{M}$ near $\bar{x}$, depending smoothly on c,

18. SENSITIVITY

Partial smoothness unifies **active set** ideas in optimization...

Theorem Suppose S is partly smooth relative to $\mathcal{M}$, and

- "strict complementarity": $\bar{c} \in \operatorname{ri} N_S(\bar{x})$
- "second-order sufficiency" for maximizing $\langle \bar{c}, \cdot \rangle$ on $\mathcal{M}$.

Then for c near $\bar{c}$, the optimality conditions

$$c \in N_S(x)$$

has a unique solution $x(c) \in \mathcal{M}$ near $\bar{x}$, depending smoothly on c, and computable from smooth equations.

18. **SENSITIVITY**

Partial smoothness unifies **active set** ideas in optimization...

Theorem Suppose S is partly smooth relative to $\mathcal{M}$, and

- "strict complementarity": $\bar{c} \in \mathrm{ri}\, N_S(\bar{x})$
- "second-order sufficiency" for maximizing $\langle \bar{c}, \cdot \rangle$ on $\mathcal{M}$.

Then for c near $\bar{c}$, the optimality conditions

$$c \in N_S(x)$$

has a unique solution $x(c) \in \mathcal{M}$ near $\bar{x}$, depending smoothly on c, and computable from smooth equations.

How can we guarantee $x(c)$ is a local maximizer?

19. PROX-REGULARITY

Another way to weaken convexity...

tangents realized by paths

19. PROX-REGULARITY

Another way to weaken convexity. . .

$$\text{tangents realized by paths} \;\Leftarrow\; \text{Clarke regularity}$$

19. PROX-REGULARITY

Another way to weaken convexity. . .

$$\begin{array}{rcl}
\text{tangents realized by paths} & \Leftarrow & \text{Clarke regularity} \\
 & & \Uparrow \\
\text{convexity} & \Rightarrow & \text{prox-regularity}
\end{array}$$

19. PROX-REGULARITY

Another way to weaken convexity...

$$\begin{array}{rcl} \text{tangents realized by paths} & \Leftarrow & \text{Clarke regularity} \\ & & \Uparrow \\ \text{convexity} & \Rightarrow & \text{prox-regularity} \end{array}$$

Theorem A set $S \subset \mathbf{R}^n$ is closed and convex $\Leftrightarrow$ every point $x \in \mathbf{R}^n$ has a unique nearest point in S.

19. PROX-REGULARITY

Another way to weaken convexity...

$$\begin{array}{ccc} \text{tangents realized by paths} & \Leftarrow & \text{Clarke regularity} \\ & & \Uparrow \\ \text{convexity} & \Rightarrow & \text{prox-regularity} \end{array}$$

Theorem A set $S \subset \mathbf{R}^n$ is closed and convex $\Leftrightarrow$ every point $x \in \mathbf{R}^n$ has a unique nearest point in S.

S is **prox-regular** at $\bar{x} \in S$ if every point x **near** $\bar{x}$ has a unique nearest point in S.

19. PROX-REGULARITY

Another way to weaken convexity...

$$\begin{array}{ccc} \text{tangents realized by paths} & \Leftarrow & \text{Clarke regularity} \\ & & \Uparrow \\ \text{convexity} & \Rightarrow & \text{prox-regularity} \end{array}$$

Theorem A set $S \subset \mathbf{R}^n$ is closed and convex $\Leftrightarrow$ every point $x \in \mathbf{R}^n$ has a unique nearest point in S.

S is **prox-regular** at $\bar{x} \in S$ if every point x **near** $\bar{x}$ has a unique nearest point in S.

Theorem Under the previous conditions, if S is also prox-regular, then $x(c)$ is a local maximizer of $\langle c, \cdot \rangle$ over S.

19. PROX-REGULARITY

Another way to weaken convexity...

$$\begin{array}{ccc} \text{tangents realized by paths} & \Leftarrow & \text{Clarke regularity} \\ & & \Uparrow \\ \text{convexity} & \Rightarrow & \text{prox-regularity} \end{array}$$

Theorem A set $S \subset \mathbf{R}^n$ is closed and convex $\Leftrightarrow$ every point $x \in \mathbf{R}^n$ has a unique nearest point in S.

S is **prox-regular** at $\bar{x} \in S$ if every point x **near** $\bar{x}$ has a unique nearest point in S.

Theorem Under the previous conditions, if S is also prox-regular, then $x(c)$ is a local maximizer of $\langle c, \cdot \rangle$ over S.

Question Are $\mathbf{P}^n_-$ and $\mathbf{M}^n_-$ prox-regular?

20. ROBUST STABILITY

Multiple eigenvalues persist, so should we worry about sensitivity?

20. ROBUST STABILITY

Multiple eigenvalues persist, so should we worry about sensitivity?

How can we check $X \in \mathbf{S}^n_-$ **robustly**?

20. ROBUST STABILITY

Multiple eigenvalues persist, so should we worry about sensitivity?

How can we check $X \in \mathbf{S}^n_-$ **robustly**? Given $\epsilon > 0$, check if

$$(*) \qquad \|Y - X\| \leq \epsilon \;\Rightarrow\; Y \in \mathbf{S}^n_-,$$

where $\|Z\|^2 = \text{trace}(Z^2)$.

20. ROBUST STABILITY

Multiple eigenvalues persist, so should we worry about sensitivity?

How can we check $X \in \mathbf{S}^n_-$ **robustly**? Given $\epsilon > 0$, check if

$$(*) \qquad \|Y - X\| \le \epsilon \;\Rightarrow\; Y \in \mathbf{S}^n_-,$$

where $\|Z\|^2 = \mathrm{trace}(Z^2)$.

Easy: $(*) \;\Leftrightarrow\; X + \epsilon I \in \mathbf{S}^n_-$.

20. ROBUST STABILITY

Multiple eigenvalues persist, so should we worry about sensitivity?

How can we check $X \in \mathbf{S}^n_-$ **robustly**? Given $\epsilon > 0$, check if

$$(*) \qquad \|Y - X\| \leq \epsilon \;\Rightarrow\; Y \in \mathbf{S}^n_-,$$

where $\|Z\|^2 = \text{trace}(Z^2)$.

Easy: $(*) \;\Leftrightarrow\; X + \epsilon I \in \mathbf{S}^n_-$.

Nonsymmetric version: Check if

$$(*) \qquad \|Y - X\| \leq \epsilon \;\Rightarrow\; Y \in \mathbf{M}^n_-,$$

where $\|Z\|^2 = \mathrm{Re}\,(\text{trace}(Z^* Z))$.

20. ROBUST STABILITY

Multiple eigenvalues persist, so should we worry about sensitivity?

How can we check $X \in \mathbf{S}_{-}^{n}$ **robustly**? Given $\epsilon > 0$, check if

$$(*) \qquad \|Y - X\| \leq \epsilon \;\Rightarrow\; Y \in \mathbf{S}_{-}^{n},$$

where $\|Z\|^2 = \text{trace}(Z^2)$.

Easy: $(*) \;\Leftrightarrow\; X + \epsilon I \in \mathbf{S}_{-}^{n}$.

Nonsymmetric version: Check if

$$(*) \qquad \|Y - X\| \leq \epsilon \;\Rightarrow\; Y \in \mathbf{M}_{-}^{n},$$

where $\|Z\|^2 = \mathrm{Re}\,(\text{trace}(Z^* Z))$.

Equivalently, calculate **distance to instability**

$$\inf\{\|Y - X\| : Y \notin \mathbf{M}_{-}^{n}\}.$$

20. ROBUST STABILITY

Multiple eigenvalues persist, so should we worry about sensitivity?

How can we check $X \in \mathbf{S}^n_-$ **robustly**? Given $\epsilon > 0$, check if

$$(*) \qquad \|Y - X\| \leq \epsilon \;\Rightarrow\; Y \in \mathbf{S}^n_-,$$

where $\|Z\|^2 = \text{trace}(Z^2)$.

Easy: $(*) \;\Leftrightarrow\; X + \epsilon I \in \mathbf{S}^n_-$.

Nonsymmetric version: Check if

$$(*) \qquad \|Y - X\| \leq \epsilon \;\Rightarrow\; Y \in \mathbf{M}^n_-,$$

where $\|Z\|^2 = \mathrm{Re}\,(\text{trace}(Z^* Z))$.

Equivalently, calculate **distance to instability**

$$\inf\{\|Y - X\| : Y \notin \mathbf{M}^n_-\}.$$

Aside: Checking stability of a "box" of polynomials is easy (Kharitonov, 1978).

21. PSEUDOSPECTRA

A powerful tool to visualize robust properties of eigenvalues of a nonsymmetric matrix X:

$$\Lambda_\epsilon(X) \;=\; \{\text{eigenvalues of } Y : \|Y - X\| \leq \epsilon\}$$

21. PSEUDOSPECTRA

A powerful tool to visualize robust properties of eigenvalues of a nonsymmetric matrix X:

$$\begin{aligned} \Lambda_\epsilon(X) &= \{\text{eigenvalues of } Y : \|Y - X\| \leq \epsilon\} \\ &= \{z \in \mathbf{C} : \sigma_{\mathsf{min}}(X - zI) \leq \epsilon\}, \end{aligned}$$

where the **smallest singular value**

$$\sigma_{\mathsf{min}}(X) = \min\{\|Xu\| : \|u\| = 1\}.$$

21. PSEUDOSPECTRA

A powerful tool to visualize robust properties of eigenvalues of a nonsymmetric matrix X:

$$\begin{aligned}\Lambda_\epsilon(X) &= \{\text{eigenvalues of } Y : \|Y - X\| \leq \epsilon\} \\ &= \{z \in \mathbf{C} : \sigma_{\min}(X - zI) \leq \epsilon\},\end{aligned}$$

where the **smallest singular value**

$$\sigma_{\min}(X) = \min\{\|Xu\| : \|u\| = 1\}.$$

In particular, $\Lambda_0(X)$ is the spectrum.

21. PSEUDOSPECTRA

A powerful tool to visualize robust properties of eigenvalues of a nonsymmetric matrix X:

$$\begin{aligned}\Lambda_\epsilon(X) &= \{\text{eigenvalues of } Y : \|Y - X\| \leq \epsilon\} \\ &= \{z \in \mathbf{C} : \sigma_{\min}(X - zI) \leq \epsilon\},\end{aligned}$$

where the **smallest singular value**

$$\sigma_{\min}(X) = \min\{\|Xu\| : \|u\| = 1\}.$$

In particular, $\Lambda_0(X)$ is the spectrum.

Denote the left halfplane by $\mathbf{C}_-$. Then

$$X \in \mathbf{M}^n_- \;\Leftrightarrow\; \Lambda_0(X) \subset \mathbf{C}_-$$

21. PSEUDOSPECTRA

A powerful tool to visualize robust properties of eigenvalues of a nonsymmetric matrix X:

$$\begin{aligned} \Lambda_\epsilon(X) &= \{\text{eigenvalues of } Y : \|Y - X\| \leq \epsilon\} \\ &= \{z \in \mathbf{C} : \sigma_{\text{min}}(X - zI) \leq \epsilon\}, \end{aligned}$$

where the **smallest singular value**

$$\sigma_{\text{min}}(X) = \min\{\|Xu\| : \|u\| = 1\}.$$

In particular, $\Lambda_0(X)$ is the spectrum.

Denote the left halfplane by $\mathbf{C}_-$. Then

$$X \in \mathbf{M}^n_- \;\Leftrightarrow\; \Lambda_0(X) \subset \mathbf{C}_-$$

whereas

$$X \in \mathbf{M}^n_- \text{ robustly } \;\Leftrightarrow\; \Lambda_\epsilon(X) \subset \mathbf{C}_-$$

21. PSEUDOSPECTRA

A powerful tool to visualize robust properties of eigenvalues of a nonsymmetric matrix X:

$$\begin{aligned} \Lambda_\epsilon(X) &= \{\text{eigenvalues of } Y : \|Y - X\| \leq \epsilon\} \\ &= \{z \in \mathbf{C} : \sigma_{\mathsf{min}}(X - zI) \leq \epsilon\}, \end{aligned}$$

where the **smallest singular value**

$$\sigma_{\mathsf{min}}(X) = \min\{\|Xu\| : \|u\| = 1\}.$$

In particular, $\Lambda_0(X)$ is the spectrum.

Denote the left halfplane by $\mathbf{C}_-$. Then

$$X \in \mathbf{M}^n_- \;\Leftrightarrow\; \Lambda_0(X) \subset \mathbf{C}_-$$

whereas

$$\begin{aligned} X \in \mathbf{M}^n_- \text{ robustly } &\Leftrightarrow \Lambda_\epsilon(X) \subset \mathbf{C}_- \\ &\Leftrightarrow \text{distance to instability } \geq \epsilon. \end{aligned}$$

21. PSEUDOSPECTRA

A powerful tool to visualize robust properties of eigenvalues of a nonsymmetric matrix X:

$$\begin{aligned}\Lambda_\epsilon(X) &= \{\text{eigenvalues of } Y : \|Y - X\| \leq \epsilon\} \\ &= \{z \in \mathbf{C} : \sigma_{\min}(X - zI) \leq \epsilon\},\end{aligned}$$

where the **smallest singular value**

$$\sigma_{\min}(X) = \min\{\|Xu\| : \|u\| = 1\}.$$

In particular, $\Lambda_0(X)$ is the spectrum.

Denote the left halfplane by $\mathbf{C}_-$. Then

$$X \in \mathbf{M}^n_- \;\Leftrightarrow\; \Lambda_0(X) \subset \mathbf{C}_-$$

whereas

$$\begin{aligned}X \in \mathbf{M}^n_- \text{ robustly} \;&\Leftrightarrow\; \Lambda_\epsilon(X) \subset \mathbf{C}_- \\ &\Leftrightarrow\; \text{distance to instability } \geq \epsilon.\end{aligned}$$

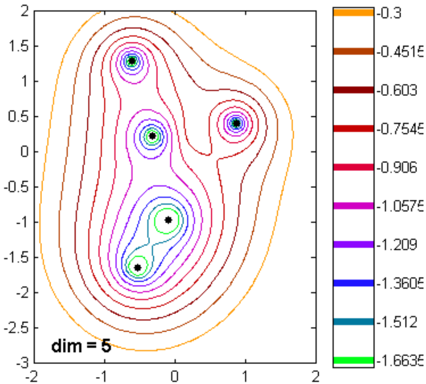
Quick to check via an algorithm of Boyd-Balakrishnan, 1990 .

22. EXAMPLES

Pseudospectra for a random 5×5 triangular complex matrix, plotted by T. Wright's **EigTool**:

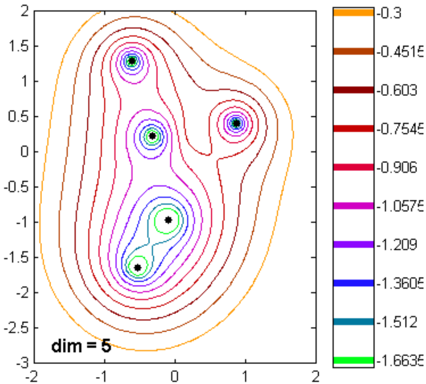
22. EXAMPLES

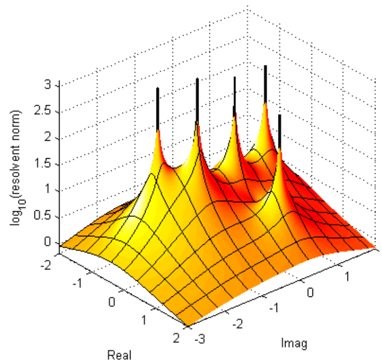
Pseudospectra for a random 5×5 triangular complex matrix, plotted by T. Wright's **EigTool**:



22. EXAMPLES

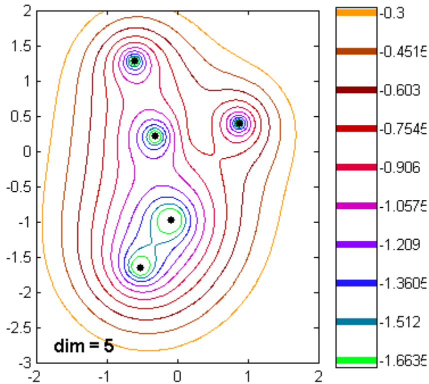
Pseudospectra for a random 5×5 triangular complex matrix, plotted by T. Wright's **EigTool**:

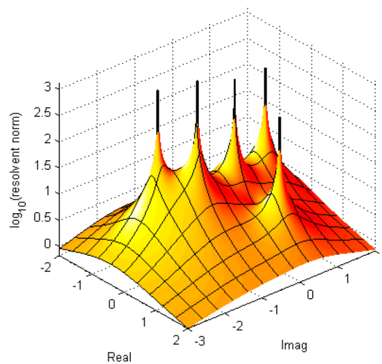




22. EXAMPLES

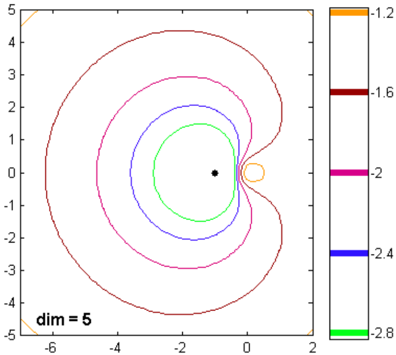
Pseudospectra for a random 5×5 triangular complex matrix, plotted by T. Wright's **EigTool**:



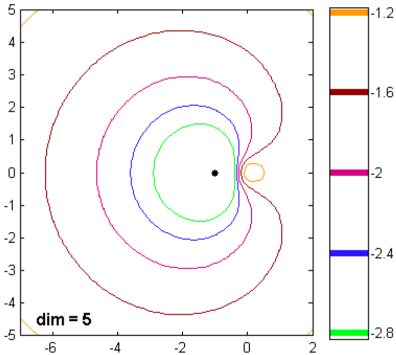


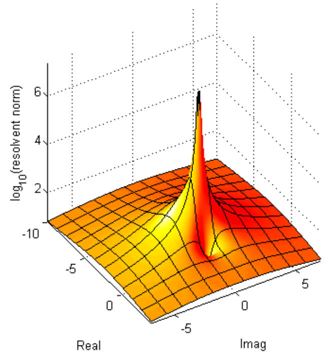
Demmel's example: $A = -\begin{bmatrix} 1 & 5 & 5^2 & 5^3 & 5^4 \\ 0 & 1 & 5 & 5^2 & 5^3 \\ 0 & 0 & 1 & 5 & 5^2 \\ 0 & 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \ldots$

23. A NEARLY UNSTABLE MATRIX

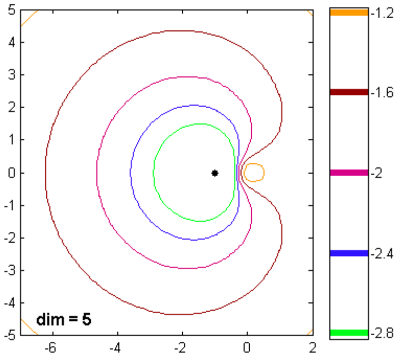


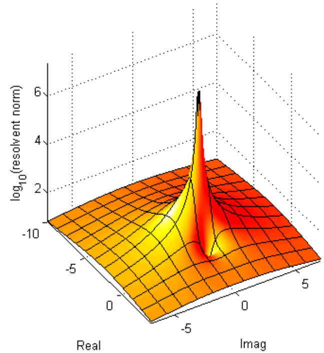
23. A NEARLY UNSTABLE MATRIX





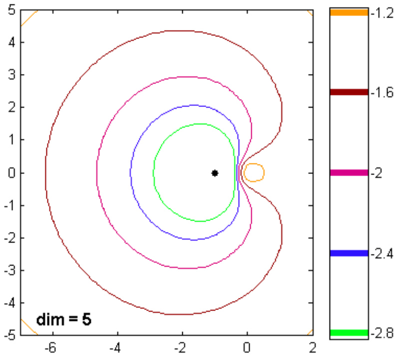
23. A NEARLY UNSTABLE MATRIX

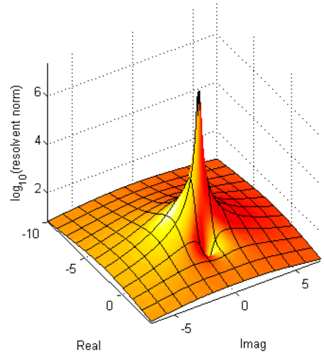




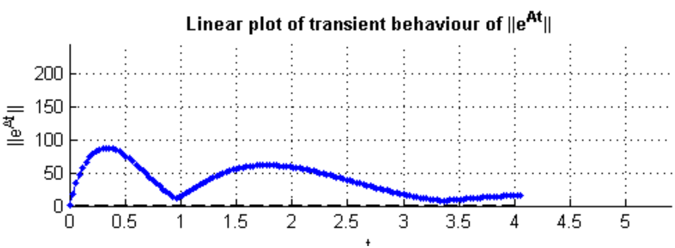
Notice $\Lambda_{.01}(A) \not\subset \mathbf{C}_-$: some unstable X satisfies $\|X - A\| \le .01$.

23. A NEARLY UNSTABLE MATRIX





Notice $\Lambda_{.01}(A) \not\subset \mathbf{C}_{-}$: some unstable X satisfies $\|X - A\| \leq .01$.



24. PSEUDOSPECTRA ARE NICER THAN SPECTRA

$\Lambda_\epsilon(X)$ encodes **robust** spectral properties of X.

24. PSEUDOSPECTRA ARE NICER THAN SPECTRA

$\Lambda_\epsilon(X)$ encodes **robust** spectral properties of X.

Symmetric $X \mapsto \Lambda_0(X)$ is Lipschitz (driven by convexity).

24. PSEUDOSPECTRA ARE NICER THAN SPECTRA

$\Lambda_\epsilon(X)$ encodes **robust** spectral properties of X.

Symmetric $X \mapsto \Lambda_0(X)$ is Lipschitz (driven by convexity).

Nonsymmetric $X \mapsto \Lambda_0(X)$ is highly nonlipschitz.

24. PSEUDOSPECTRA ARE NICER THAN SPECTRA

$\Lambda_\epsilon(X)$ encodes **robust** spectral properties of X.

Symmetric $X \mapsto \Lambda_0(X)$ is Lipschitz (driven by convexity).

Nonsymmetric $X \mapsto \Lambda_0(X)$ is highly nonlipschitz.

But the **pseudospectral** mapping $X \mapsto \Lambda_\epsilon(X)$ **does** have Lipschitz properties.

24. PSEUDOSPECTRA ARE NICER THAN SPECTRA

$\Lambda_\epsilon(X)$ encodes **robust** spectral properties of X.

Symmetric $X \mapsto \Lambda_0(X)$ is Lipschitz (driven by convexity).

Nonsymmetric $X \mapsto \Lambda_0(X)$ is highly nonlipschitz.

But the **pseudospectral** mapping $X \mapsto \Lambda_\epsilon(X)$ **does** have Lipschitz properties.

Trefethen's thesis: the **pseudospectral abscissa**

$$\alpha_\epsilon(X) \;=\; \max \operatorname{Re} \Lambda_\epsilon(X)$$

predicts practical dynamics of $\dot{y} = Xy$ better than α_0.

24. PSEUDOSPECTRA ARE NICER THAN SPECTRA

$\Lambda_\epsilon(X)$ encodes **robust** spectral properties of X.

Symmetric $X \mapsto \Lambda_0(X)$ is Lipschitz (driven by convexity).

Nonsymmetric $X \mapsto \Lambda_0(X)$ is highly nonlipschitz.

But the **pseudospectral** mapping $X \mapsto \Lambda_\epsilon(X)$ **does** have Lipschitz properties.

Trefethen's thesis: the **pseudospectral abscissa**

$$\alpha_\epsilon(X) \;=\; \max \operatorname{Re} \Lambda_\epsilon(X)$$

predicts practical dynamics of $\dot{y} = Xy$ better than α_0.
Eg: In Demmel's example, $\alpha_0 = -1$ but $\alpha_{.01} \approx .122$.

24. PSEUDOSPECTRA ARE NICER THAN SPECTRA

$\Lambda_\epsilon(X)$ encodes **robust** spectral properties of X.

Symmetric $X \mapsto \Lambda_0(X)$ is Lipschitz (driven by convexity).

Nonsymmetric $X \mapsto \Lambda_0(X)$ is highly nonlipschitz.

But the **pseudospectral** mapping $X \mapsto \Lambda_\epsilon(X)$ **does** have Lipschitz properties.

Trefethen's thesis: the **pseudospectral abscissa**

$$\alpha_\epsilon(X) \;=\; \max \operatorname{Re} \Lambda_\epsilon(X)$$

predicts practical dynamics of $\dot{y} = Xy$ better than α_0.
Eg: In Demmel's example, $\alpha_0 = -1$ but $\alpha_{.01} \approx .122$.

We can compute α_e with a globally and locally quadratically convergent algorithm (like Boyd-Balakrishnan's), available in EigTool. We can even optimize it: Overton's talk.

25. SUMMARY

- Unlike the semidefinite cone, the **stable** cones are nonconvex with nasty boundaries.

25. SUMMARY

- Unlike the semidefinite cone, the **stable** cones are nonconvex with nasty boundaries.
- However, $\mathbf{M}^n_-$ and $\mathbf{P}^n_-$ share key **local** properties with $\mathbf{S}^n$.

25. SUMMARY

- Unlike the semidefinite cone, the **stable** cones are nonconvex with nasty boundaries.
- However, $\mathbf{M}^n_-$ and $\mathbf{P}^n_-$ share key **local** properties with $\mathbf{S}^n$.
  - **Clarke regularity**: a local relaxation of convexity, crucial for calculus and optimality conditions.

25. SUMMARY

- Unlike the semidefinite cone, the **stable** cones are nonconvex with nasty boundaries.
- However, $\mathbf{M}^n_-$ and $\mathbf{P}^n_-$ share key **local** properties with $\mathbf{S}^n$.
  - **Clarke regularity**: a local relaxation of convexity, crucial for calculus and optimality conditions.
  - **Partial smoothness**: the key to understanding persistent structure in optimal solutions.

25. SUMMARY

- Unlike the semidefinite cone, the **stable** cones are nonconvex with nasty boundaries.
- However, $\mathbf{M}^n_-$ and $\mathbf{P}^n_-$ share key **local** properties with $\mathbf{S}^n$.
  - **Clarke regularity**: a local relaxation of convexity, crucial for calculus and optimality conditions.
  - **Partial smoothness**: the key to understanding persistent structure in optimal solutions.
  - **Robust interior**: identifiable computationally, having Lipschitz boundary, visualized using **pseudospectra**.

25. SUMMARY

- Unlike the semidefinite cone, the **stable** cones are nonconvex with nasty boundaries.
- However, $\mathbf{M}^n_-$ and $\mathbf{P}^n_-$ share key **local** properties with $\mathbf{S}^n$.
  - **Clarke regularity**: a local relaxation of convexity, crucial for calculus and optimality conditions.
  - **Partial smoothness**: the key to understanding persistent structure in optimal solutions.
  - **Robust interior**: identifiable computationally, having Lipschitz boundary, visualized using **pseudospectra**.

`orie.cornell.edu/∼aslewis`