Julia Eaton
jreaton@uw.edu
University of Washington Tacoma

Workshop on Numerical Linear Algebra and Optimization
University of British Columbia
Vancouver, BC Canada

Variational properties of polynomial root functions and spectral max functions

Abstract 1

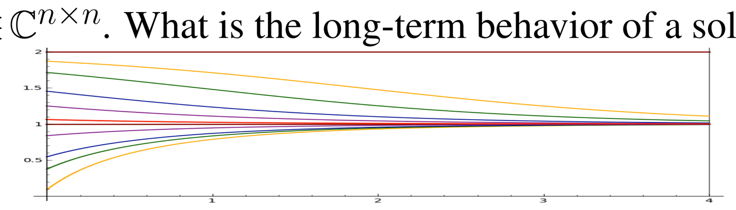
Eigenvalue optimization problems arise in the control of continuous and discrete time dynamical systems. The spectral abscissa (largest real part of an eigenvalue) and spectral radius (largest eigenvalue in modulus) are examples of functions of eigenvalues, or spectral functions, connected to these problems. A related class of functions are polynomial root functions. In 2001, Burke and Overton showed that the abscissa mapping on polynomials is subdifferentially regular on the monic polynomials of degree n. In 2012 we extended these results to a class of max polynomial root functions which includes both the polynomial abscissa and the polynomial radius. We are currently working to extend these results to the matrix setting.

Motivation 2

Consider the linear continuous-time dynamical system

$$\text{(DE)} \qquad y' = Ay.$$

where $A\in\mathbb{C}^{n\times n}$. What is the long-term behavior of a solution?



Let y_e be an equilibrium solution to $\quad dy/dt = f(t,y)$.

- y_e is *Lyapunov stable* if for any $\varepsilon > 0$, $\exists\ \delta_\varepsilon > 0$ such that $|y(t) - y_e| \le \varepsilon$ whenever $|y(0) - y_e| \le \delta_\varepsilon$
- y_e is *asymptotically stable* if it is LS and $\exists \gamma > 0$ such that $|y(t) - y_e| \to 0$ whenever $|y(t) - y_e| < \gamma$

A stable solution y_e to (DE) is asymptotically stable if the eigenvalues of A lie in the left-half plane.

- Define the *spectral abscissa* $\alpha : \mathbb{C}^{n\times n} \to \mathbb{R}$ by
$$\alpha(A) = \max\{\mathrm{Re}(\lambda) \mid \det(\lambda I - A) = 0\}$$
- Then y_e is asymptotically stable if $\alpha(A) < 0$.
- Also, the rate of decay of a soln. depends on $\alpha(A)$.

Next consider the discrete-time analog: $y^{k+1} = Ay^k,\ \ A\in\mathbb{C}^{n\times n}$.

- Asymptotic stability tied to max modulus of eigenvalues of A.
- The *spectral radius* $\rho : \mathbb{C}^{n\times n} \to \mathbb{R}$ is given by
$$\rho(A) = \max\{|\lambda| \mid \det(\lambda I - A) = 0\}.$$
- A stable soln is asymptotically stable if $\rho(A) < 1$.

Example 3

(Damped oscillator) Consider the system $v'' + xv' + v = 0$.

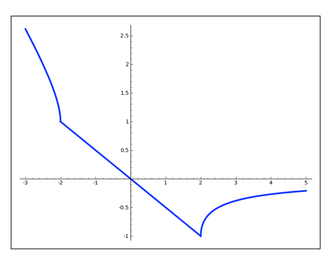
Linearizing gives $\begin{bmatrix} v \\ v' \end{bmatrix}' = \begin{bmatrix} 0 & 1 \\ -1 & -x \end{bmatrix}\begin{bmatrix} v \\ v' \end{bmatrix}$,

The matrix (call it $A(x)$) has eigenvalues $(-x \pm \sqrt{x^2-4})/2$

Computing, $\alpha(A(x)) = \begin{cases} (-x + \sqrt{x^2-4})/2 & |x| \ge 2 \\ -x/2 & |x| < 2 \end{cases}$

What value of x minimizes $\alpha(A(x))$?

It's easy: $\underset{x\in\mathbb{R}}{\arg\min}\ \ \alpha(A(x)) = 2$



Spectral Functions 4

We say $\psi : \mathbb{C}^{n\times n} \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ is a *spectral function* if it

- depends only on the eigenvalues of its argument
- is invariant under permutations of those eigenvalues

A *spectral max function* $\varphi : \mathbb{C}^{n\times n} \to \overline{\mathbb{R}}$ is a spectral fcn. of the form
$$\varphi(A) = \max\{f(\lambda) \mid \det(\lambda I - A) = 0\}$$
where $f : \mathbb{C} \to \overline{\mathbb{R}}$. The spectral abscissa and radius are spectral max functions. (Note: $f\ \mathbb{C} \to \overline{\mathbb{R}}$ is treated as a function from $\mathbb{R}^2$ to $\overline{\mathbb{R}}$.)

Optimizing Spectral Functions 5

We saw an example of a simple abscissa minimization problem where the matrix depended on a single parameter. More generally we can consider
$$\min_{x\in U} \psi(A(x))$$
where $U \subset \mathbb{C}^k$ is some parameter set. How do we verify a candidate soln is optimal? We need to know about the structure of ψ.

- Gradients can give necessary conditions, provided they exist.
- Parameterization means a composite fcn—use a chain rule?

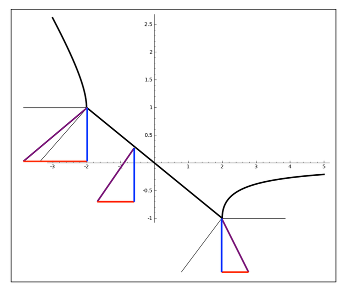
Some variational analysis 6

Let $h : E \to \overline{\mathbb{R}}$.

The *regular subdifferential of h at $\widetilde{x}$* is the set
$$\hat{\partial} h(\widetilde{x}) = \{z \mid h(x) - h(\widetilde{x}) \ge \langle z, x - \widetilde{x}\rangle + o(\|x - \widetilde{x}\|),\ \forall x\}$$
We say $\widetilde{x}$ is *sharp* if $\exists\ \delta, \varepsilon > 0$ such that $h(x) - h(\widetilde{x}) \ge \delta\,\|x - \widetilde{x}\|$ for all x for which $\|x - \widetilde{x}\| < \varepsilon$.

Fact: (Burke, Overton, 2000): $\widetilde{x}$ is a sharp local minimizer of h iff $0 \in \mathrm{int}(\hat{\partial} h(\widetilde{x}))$.

Example reg. subgradients of the damped oscillator abscissa:
Set $h(\cdot) = \alpha(A(\cdot))$:
$z \in \{-1/2\} = \hat{\partial} h(-3/5)$
$z \in \hat{\partial} h(-2), \quad z \in \hat{\partial} h(2)$



The *subderivative* (generalizes the directional derivative)
$$dh(\widetilde{x})(\widetilde{v}) = \liminf_{t\searrow 0, v\to\widetilde{v}} (\,h(\widetilde{x} + tv) - h(\widetilde{x})\,)/t.$$

A key subderivative & regular subdifferential relationship:
$$\hat{\partial} h(\widetilde{x}) = \{z \mid \langle z, v\rangle \le dh(\widetilde{x})(v)\ \forall\, v\}$$

7

The *general subdifferential* (limits of regular subgradients)
$$\partial h(\widetilde{x}) = \left\{ z \,\middle|\, \begin{array}{c} \exists\{x_\nu\} \subseteq \mathrm{dom}\,(h) \text{ and } \{z_\nu\} \subseteq E \text{ such that} \\ z_\nu \in \hat{\partial} h(x_\nu)\ \forall\,\nu,\ x_\nu \to \widetilde{x}, h(x_\nu) \to h(\widetilde{x}),\ \text{and } z_\nu \to z \end{array} \right\}$$

The *horizon subdifferential* (captures non-Lipschitz-ness)
$$\partial^\infty h(\widetilde{x}) = \left\{ z \,\middle|\, \begin{array}{l} \exists\{x_\nu\} \subseteq \mathrm{dom}\,(h), \{t_\nu\} \subseteq [0,\infty), \{z_\nu\} \subseteq E,\ \text{s.t.}\ \ \forall\,\nu \\ z_\nu \in \hat{\partial} h(x_\nu)\ x_\nu \to \widetilde{x}, h(x_\nu) \to h(\widetilde{x}),\ t_\nu \downarrow 0 \text{ and } t_\nu z_\nu \to z \end{array} \right\}$$

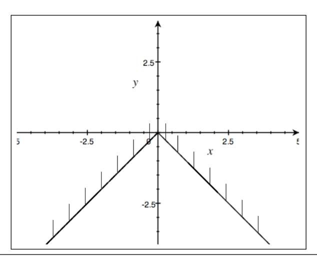
h is *regular* at $\widetilde{x}$ if $\partial h(\widetilde{x}) \ne \emptyset$, $\ \hat{\partial} h(\widetilde{x}) = \partial h(\widetilde{x})$ and $\hat{\partial} h(\widetilde{x})^\infty = \partial^\infty h(\widetilde{x})$

(All convex functions are regular on their domains.)

8

Negative absolute value: $y = -|x|$.

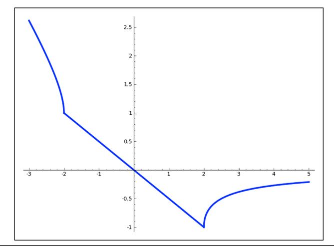
- piecewise linear, non-diff at 0
- $\hat{\partial}(-|\cdot|)(0) = \emptyset$
- $\partial(-|\cdot|)(0) = \{-1, 1\}$
- not subdifferentially regular



9

Damped oscillator abscissa (see panel 3)

- Smooth everywhere but -2 and 2.
- Non-Lipschitz at -2 and 2
- Non-convex
- $\hat{\partial} h(2) = [-1/2, \infty) = \partial h(2)$
- subdifferentially regular at 2



Variational properties of spectral functions 10

Familiar spectral functions: abscissa α and radius ρ.

- α and ρ continuous on $\mathbb{C}^{n\times n}$, not Lipschitz continuous
- α and ρ are not differentiable in general

Compute regular subgradients? If subdifferentially regular...

(Chain Rule, Rockafellar & Wets '98)

Let $g : \mathbb{R}^m \to \overline{\mathbb{R}}$ be regular at $H(\widetilde{x})$, H smooth, $h = g \circ H$. CQ: *the only vector $z \in \partial^\infty g(H(\widetilde{x}))$ with $\nabla H(\widetilde{x})^* z = 0$ is $z = 0$.* If (CQ) is satisfied, then
$$\underline{\partial h(\widetilde{x}) = \nabla H(\widetilde{x})^* \partial g(H(\widetilde{x}))}, \qquad \partial^\infty h(\widetilde{x}) = \nabla H(\widetilde{x})^* \partial^\infty g(H(\widetilde{x})).$$

α regular at X iff X's active evals are nonderogatory (Burke, Overton 01)

Decomposition 11

- $\mathcal{P}^n$: linear space of polynomials of degree $0, 1, \ldots, n$
- $\mathcal{M}^n \subset \mathcal{P}^n$: polynomials of degree n
- Define $\Phi : \mathbb{C}^{n\times n} \to \mathcal{P}^n$ by $\quad \Phi(A) = \det(\lambda I - A)$.

Define the *max root function* $\mathbf{f} : \mathcal{P}^n \to \overline{\mathbb{R}}$ by
$$\mathbf{f}(p) = \max\{f(\lambda) \mid p(\lambda) = 0\}.$$
Then $\quad \varphi = \mathbf{f} \circ \Phi$.

$$\begin{tikzcd} \mathbb{C}^{n\times n} \arrow[d, "\Phi"'] \arrow[dr, "\varphi"] & \\ \mathcal{P}^n \arrow[r, "\mathbf{f}"'] & \overline{\mathbb{R}} \end{tikzcd}$$

Properties of polynomial root functions 12

If f is continuous on its domain,

- $\mathbf{f}$ continuous on $\mathcal{M}^n \cap \mathrm{dom}\,(\mathbf{f})$, not generally Lipschitz
  - $\mathbf{a}(\lambda^n - \varepsilon) = \sqrt[n]{\varepsilon}$
- $\mathbf{f}$ not differentiable in general
- $\mathbf{f}$ is unbounded in a neighborhood of q for all $q(\lambda) \in \mathcal{P}^n \setminus \mathcal{M}^n$
  - let $q(\lambda) \in \mathcal{P}^{n-1}$, then $q(\lambda)(\varepsilon\lambda - 1) \to q(\lambda)$, but root $1/\varepsilon$ is unbounded

Restrict our attention to $\mathcal{M}^n$.

Polynomial max root history

Let $f : \mathbb{C} \to \overline{\mathbb{R}}$ be proper, convex, lsc.
$$\mathbf{f}(q) = \max\{f(q) \mid q(\lambda) = 0\}$$

- If $f = \mathrm{Re}(\cdot)$, the polynomial abscissa $\mathbf{a}$. Well studied for general $p = \prod(\lambda - \lambda_j)^{n_j} \in \mathcal{M}_1^n$ (Burke & Overton '01)
- For general f, $\mathbf{f}$ is studied for polynomials $(\lambda - \lambda_0)^n$, regularity not shown. (Burke, Lewis, Overton '05)
- **Done**: fill the gap, and clarify process, go beyond max root fcn.
- **To do**: push through to the matrix setting.

Polynomial radius subdifferential 13

$\mathbf{r}(p) = \max\{|\lambda| \mid p(\lambda) = 0\}$

2001: for $\lambda_0 \ne 0$,
$$\partial\mathbf{r}((\lambda - \lambda_0)^n) = \left\{ z \,\middle|\, \begin{array}{c} z = \sum_{s=1}^n \mu_s(\lambda - \lambda_0)^{n-s},\ \ \mu_1 = -\lambda_0/(n\,|\lambda_0|) \\ \mathrm{Re}(\overline{\lambda_0^2}\,\mu_2) \le |\lambda_0|\,/n \end{array} \right\}$$

2012: For $p = \prod_{j=1}^m(\lambda - \lambda_j)^{n_j}$, $\lambda_1, \ldots, \lambda_m$ distinct,
$$\partial\mathbf{r}(p) = \left\{ z \,\middle|\, \begin{array}{c} z = \sum_{j=1}^m \prod_{k\ne j}(\lambda - \lambda_k)^{n_k} \sum_{s=1}^{n_j} \mu_{js}(\lambda - \lambda_j)^{n_j - s} \\ \exists\{\gamma_j\}_{j\in\mathcal{I}(p)} \subset [0,1] \ \text{ with } \ \sum\gamma_j = 1 \text{ such that} \\ \mu_{j1} = -\gamma_j\lambda_j/(n_j\,|\lambda_j|) \text{ and } \mathrm{Re}(\overline{\lambda_j^2}\,\mu_{j2}) \le \gamma_j\,|\lambda_j|\,/n_j\ \ \forall\, j \in \mathcal{I}(p) \end{array} \right\}$$

Applications 14

Recall damped oscillator problem in panel 3. The characteristic polynomial of $A(x)$ is: $\ p(\lambda, x) = \lambda(\lambda + x) + 1 = \lambda^2 + x\lambda + 1$. and abscissa graph is shown in panel 3.

Consider the parameterization $H : \mathbb{R} \to \mathcal{P}^2$, where
$H(x) = (\lambda - \lambda_0)^2 + h_1(x)(\lambda - \lambda_0) + h_2(x), \quad \lambda_0 \in \mathbb{C},\ h_1, h_2 : \mathbb{R} \to \mathbb{R}$.

Require $\ p(\lambda, x) = \lambda^2 + x\lambda + 1 = (\lambda - \lambda_0)^2 + h_1(x)(\lambda - \lambda_0) + h_2(x)$ so
$$h_1(x) = 2\lambda_0 + x, \quad h_2(x) = \lambda_0^2 + x\lambda_0 + 1, \quad \nabla H(x) = [0, 1, \lambda_0]^T$$

Set $h = \mathbf{a} \circ H$, so $\ h(x) = \mathbf{a}(p(\lambda, x))$. By chain rule (panel 10),
$$\hat{\partial} h(\widetilde{x}) = \nabla H(\widetilde{x})^T \hat{\partial}\mathbf{a}(p(\lambda, \widetilde{x})).$$

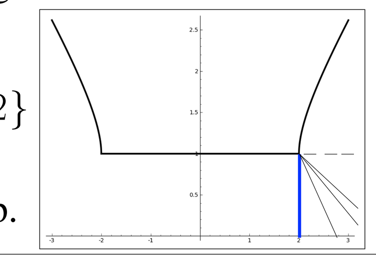
Global min of h is 2, so $\ p(\lambda, 2) = \lambda^2 + 2\lambda + 1 = (\lambda + 1)^2$ and
$$\hat{\partial} h(2) = \nabla H(2)^T \partial\mathbf{a}(\,(\lambda + 1)^2\,).$$

From 2001, $\partial\mathbf{a}(\,(\lambda + 1)^2\,) = \{z \mid z = -(\lambda + 1)/2 + \mu_2,\ \mathrm{Re}\,(\mu_2) \le 0\}$.

Set $\lambda_0 = -1$: $\quad \hat{\partial} h(2) = \left\langle [0,\ 1,\ -1]^T, \partial\mathbf{a}(\,(\lambda + 1)^2\,) \right\rangle$
$$= \{\mathrm{Re}(-1/2 - \mu_2) \mid \mathrm{Re}(\mu_2) \le 0\} = [-1/2, \infty)$$
Therefore $\ 0 \in \mathrm{int}\,(\hat{\partial} h(2))$. Minimizer 2 is sharp.

15

Applying the polynomial radius to the family
$$\{p(\lambda, x)\}_{x\in\mathbb{R}} = \{\lambda^2 + x\lambda + 1\}_{x\in\mathbb{R}},$$
we have
$$\mathbf{r}(p(\lambda, x)) = \begin{cases} (|x| + \sqrt{x^2 - 4})/2 & |x| \ge 2 \\ 1 & |x| < 2 \end{cases}$$
Let $h = \mathbf{r} \circ H$, so $h(x) = \mathbf{r}(p(\lambda, x))$. A global minimum is at 2.
$$\begin{aligned} \hat{\partial} h(2) &= \left\langle [0, 1, -1]^T, \partial\mathbf{r}(\,(\lambda + 1)^2\,) \right\rangle \\ &= \{\mathrm{Re}(1/2 - \mu_2) \mid \mathrm{Re}(\mu_2) \le 1/2\} \\ &= [0, \infty) \end{aligned}$$
$0 \notin \mathrm{int}(\hat{\partial} h(2))$, so minimizer is not sharp.



Matrix Setting 16

In 2001 Burke & Overton derived many properties of a broad class of spectral functions as well as the subdifferential regularity of the spectral abscissa at a matrix with nonderogatory active eigenvalues. In 2013 Grundel & Overton derived the formula for the subdifferential of the simplest example of a derogatory, defective matrix: a 3×3 matrix w/ one eigenvalue and two Jordan blocks. We have been working to extend the abscissa results to a broader class of spectral functions. We have shown the subdifferential regularity in the case of nonderogatory active eigenvalues for a very narrow class of functions which includes the abscissa but not the radius. Extending the methods in the 2001 paper has proven difficult. As demonstrated in G & O's 2013 paper, even the 3×3 case is extremely challenging and does not easily lend itself to generalization. We are currently exploring other avenues to derive variational properties of spectral functions.

References

Grundel, Overton. Var. Analy. of Spect. Absc. at a Matrix w/ Nongeneric Mult. Eval. *Set-Val. & Var. Analy.* 2013.
Burke, Eaton. On the subdifferential regularity of polynomial max root functions. *J. Nonlinear Analysis* 2012.
Burke, Lewis, Overton. Variational Analysis of functions of the roots of polynomials. *Mathematical Prog.*, 2005.
Burke, Lewis, Overton. Variational Analysis of Abscissa Mapping...via the Gauss-Lucas Theorem. *J. G Opt.*, 2004.
Burke, Overton. Variational Analysis of the Abscissa Mapping for Polynomials. *SIAM J. on Control & Opt.*, 2001.
Burke, Overton. Variational Analysis of non-Lipschitz spectral functions. *Mathematical Programming*, 2001.
Burke, Lewis, Overton. Optimizing Matrix Stability. *AMS Proceedings*, 2000.