Optimization Over Hyperbolicity Cones

Jim Renegar

- $p : \mathbb{R}^d \to \mathbb{R}$ homogeneous polynomial of degree n
- $p(e) > 0$

Defn: The polynomial p is

"hyperbolic in direction e"

if for all $x \in \mathbb{R}^d$, the univariate polynomial

$$\lambda \mapsto p(\lambda e - x) \text{ has only real roots.}$$

Roots: $\lambda_{1,e}(x) \leq \lambda_{2,e}(x) \leq \cdots \leq \lambda_{n,e}(x)$

"eigenvalues of x (in direction e)"

LP:

- $p(x) = x_1, \ldots, x_n$
- $e > 0$

$$\lambda \mapsto p(\lambda e - x) = (\lambda e_1 - x_1) \cdots (\lambda e_n - x_n)$$

Eigenvalues of x in direction e: $\quad \frac{x_1}{e_1}, \ldots, \frac{x_n}{e_n}$

SDP:

- $p(x) = \det(x)$
- $e \succ 0$

$$\lambda \mapsto \det(\lambda e - x) = \det(e) \det(\lambda I - e^{-1/2} x e^{-1/2})$$

Eigenvalues of x in direction e

= traditional eigenvalues of $e^{-1/2} x e^{-1/2}$

$\lambda_{1,e}(x) \leq \lambda_{2,e}(x) \leq \cdots \leq \lambda_{n,e}(x)$ roots of $\lambda \mapsto p(x - \lambda e)$

Hyperbolicity Cone:

$$\Lambda_{++} := \{x : 0 < \lambda_{1,e}(x)\}$$

$$= \text{connected component of } \{x : p(x) > 0\} \text{ containing } e$$

Gårding (1959): p is hyperbolic in direction e for all $e \in \Lambda_{++}$

Corollary: Λ_{++} is a convex cone

Corollary: $x \mapsto \lambda_{n,e}(x)$ is a convex function

Bauschke, Güler, Lewis & Sendov:

If $f : \mathbb{R}^n \to \mathbb{R}$ is a convex and permutation-invariant then $x \mapsto f(\vec{\lambda}_e(x))$ is convex

---

Lax, Vinnikov and Helton Theorem:

Every 3-dimensional hyperbolicity cone is a slice of a PSD cone.

Cor: Faces of hyperbolicity cones are exposed.

Chua: Every homogeneous cone is a slice of a PSD cone.

ϕ a univariate polynomial

If ϕ has only real roots then:

- ϕ' has only real roots.
- Roots are interlaced: $\lambda_1 \leq \lambda'_1 \leq \lambda_2 \leq \cdots \leq \lambda'_{n-1} \leq \lambda_n$

---

p a multivariate polynomial

$p'_e(x) := \langle \nabla p(x), e \rangle$ (directional derivative)

If p is hyperbolic in direction e then:

- p'_e is hyperbolic in direction e.
- $\Lambda_+ \subseteq \Lambda'_{e,+}$

Inductively:

$$p_e^{(i+1)}(x) = \langle \nabla p_e^{(i)}(x), e \rangle$$

$$\Lambda_+ = \Lambda_{e,+}^{(0)} \subseteq \Lambda_{e,+}^{(1)} \subseteq \cdots \subseteq \Lambda_{e,+}^{(n-1)} = \text{a halfspace}$$

$$p_e^{(i)}(x) = i!\, p(e)\, E_{n-i}(\vec{\lambda}_e(x))$$

where E_k = elementary symmetric polynomial of degree k

$$\Lambda_{e,+}^{(i)} = \{x : E_k(\vec{\lambda}_e(x)) \geq 0, \;\; k = 1, \ldots, n-i\}$$

Thm: If p is hyperbolic in direction e

then p/p'_e is a concave function on $\Lambda'_{e,++}$

Pf:

- $q(x,t) := tp(x)$ is hyperbolic in direction $(e,1)$
- Hence, $q'_{(e,1)}$ is hyperbolic in direction $(e,1)$
- Hyperbolicity cone of $q'_{(e,1)}$ is epigraph of $x \mapsto -p(x)/p'_e(x)$

Λ_+-Feasibility Problem: Find $x \in \Lambda_+$ satisfying $Ax = b$

Assume $\bar{x} \in \Lambda'_{e,++}$ satisfies $A\bar{x} = b$

Then Λ_+-feasibility attained by "solving"

$$\begin{array}{rl} \max & p(x)/p'_e(x) \\ \text{s.t.} & Ax = b \end{array}$$

More generally, assume $\bar{x}$ is $\Lambda^{(i)}_{e,++}$-feasible.

First find $\Lambda^{(i-1)}_{e,++}$-feasible point,

then find $\Lambda^{(i-2)}_{e,++}$-feasible point,

... and, finally, find Λ_+-feasible point.

Hyperbolic Program (HP):

$$\begin{array}{rl} \min & \langle c, x \rangle \\ \text{s.t.} & Ax = b \\ & x \in \Lambda_+ \end{array}$$

Introduced by Güler (mid-90's) in context of ipm's:

"Central Path" $= \{x(\eta) : \eta > 0\}$

where $x(\eta)$ solves

$$\begin{array}{rl} \min & \eta \langle c, x \rangle - \ln p(x) \\ \text{s.t.} & Ax = b \end{array}$$

$O(\sqrt{n}) \log(1/\epsilon)$ iterations suffice

to reduce $\alpha := \langle c, x \rangle - \langle b, y \rangle$ to $\epsilon\, \alpha$

Hyperbolic Program relaxation:

$$\underbrace{\begin{array}{rl} \min & \langle c, x\rangle \\ \text{s.t.} & Ax = b \\ & x \in \Lambda_+ \end{array}}_{\text{HP}} \quad \xrightarrow{\text{relax}} \quad \underbrace{\begin{array}{rl} \min & \langle c, x\rangle \\ \text{s.t.} & Ax = b \\ & x \in \Lambda_{e,+}^{(i)} \end{array}}_{\text{HP}_e^{(i)}}$$

Defn: The "i^{th} central swath" is the set of directions e satisfying

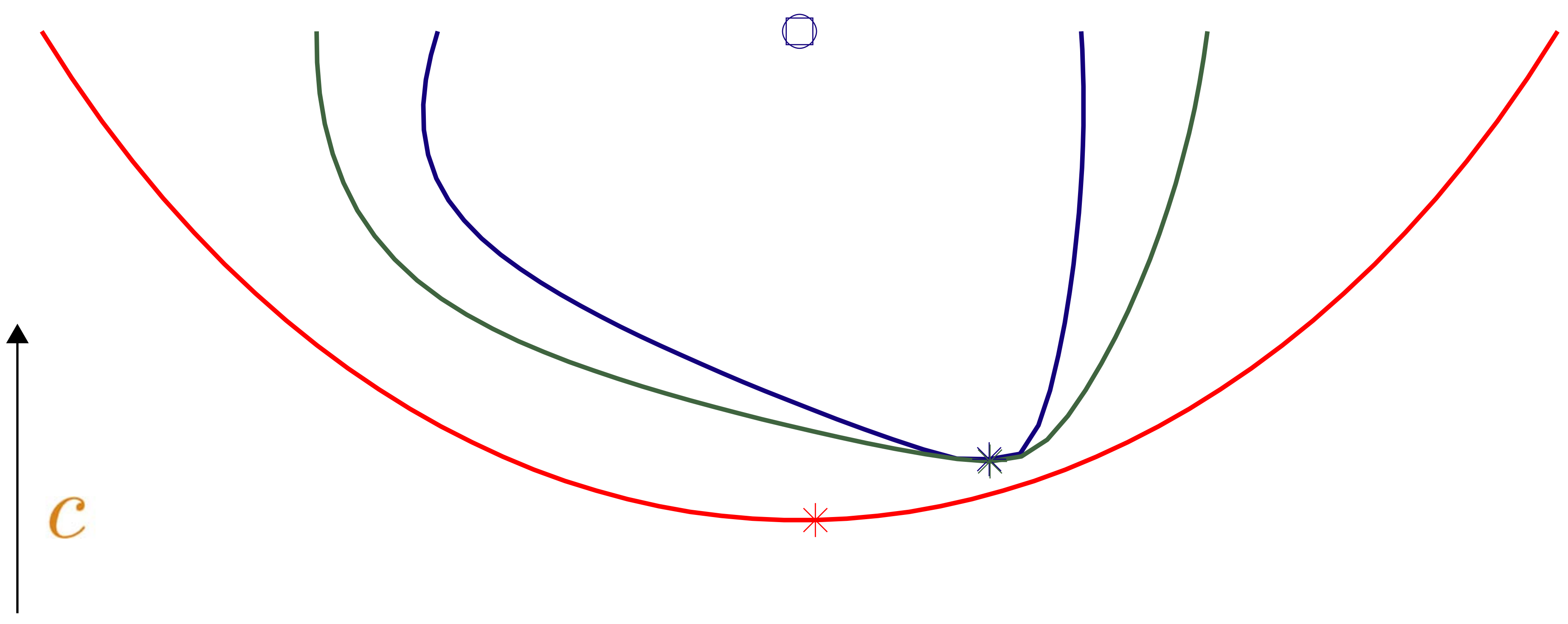
- $Ae = b, \quad e \in \Lambda_{e,++}^{(i)}$ (strict feasibility)
- $\text{HP}_e^{(i)}$ has an optimal solution

---

central path = $(n-1)^{th}$ central swath

$$\begin{aligned} \min \quad & \langle c, x \rangle \\ \text{s.t.} \quad & Ax = b \\ & x \in \Lambda_+ \end{aligned}$$

$$\begin{aligned} \min \quad & \langle c, x \rangle \\ \text{s.t.} \quad & Ax = b \\ & x \in \Lambda_{+,e}^{(i)} \end{aligned}$$



Hyperbolic Program relaxation:

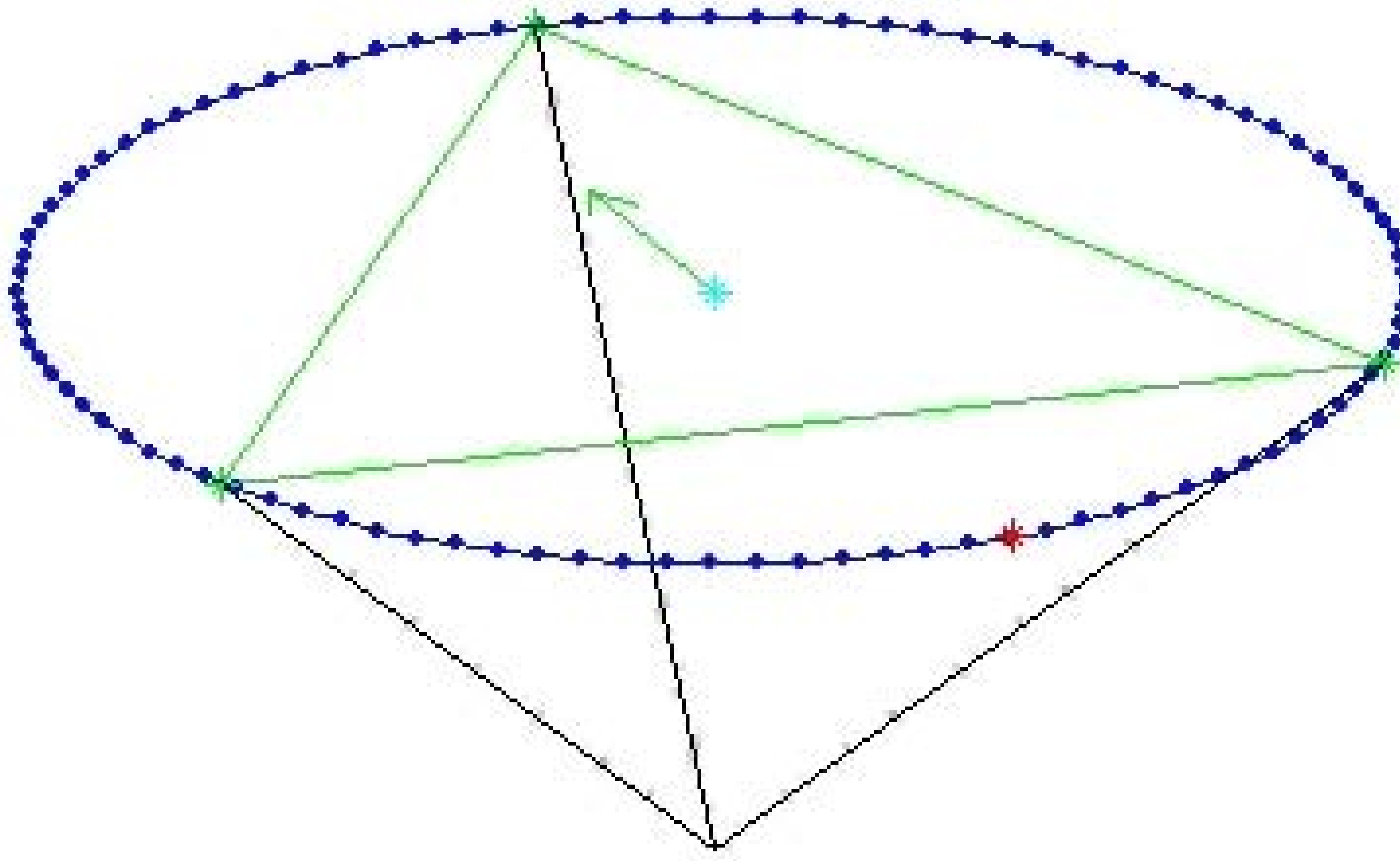
$$\underbrace{\begin{array}{rl} \min & \langle c, x\rangle \\ \text{s.t.} & Ax = b \\ & x \in \Lambda_+ \end{array}}_{\text{HP}} \quad \xrightarrow{\text{relax}} \quad \underbrace{\begin{array}{rl} \min & \langle c, x\rangle \\ \text{s.t.} & Ax = b \\ & x \in \Lambda_{e,+}^{(i)} \end{array}}_{\text{HP}_e^{(i)}}$$

Defn: The "i^{th} central swath" is the set of directions e satisfying

- $Ae = b, \quad e \in \Lambda_{e,++}^{(i)}$ (strict feasibility)
- $\text{HP}_e^{(i)}$ has an optimal solution

---

central path = $(n-1)^{th}$ central swath



see enclosed avi video by Y. Zinchenko

$e(t)$ time dependent

$z(t)$ optimal solution of $\mathrm{HP}^{(i)}_{e(t)}$

Dynamics: $\frac{d}{dt}e(t) = z(t) - e(t)$

If $i = n - 2$ and $e(0)$ is on the central path
then $e(t)$ traces the central path.

Thm: Assume dual of HP is strictly feasible and $i \leq n - 2$.

If $e(0)$ is in the i^{th} central swath, then:

- The dynamics are well-defined
  (in particular, $e(t)$ is in the swath for all $t \geq 0$)
- $e(t) \rightarrow$ optimality for HP (what about $z(t)$?)

$$\underbrace{\begin{array}{rl} \min & \langle c, x\rangle \\ \text{s.t.} & Ax = b \\ & x \in \Lambda_+ \end{array}}_{\text{HP}} \quad \xrightarrow{\text{relax}} \quad \underbrace{\begin{array}{rl} \min & \langle c, x\rangle \\ \text{s.t.} & Ax = b \\ & x \in \Lambda_{e,+}^{(i)} \end{array}}_{\text{HP}_e^{(i)}}$$

z optimal solution of $\text{HP}_e^{(i)}$

If $z \notin \Lambda_+$ then z is optimal also for

$$\begin{array}{rl} \min_x & -\ln\langle c, e - x\rangle - \frac{p_e^{(i)}(x)}{p_e^{(i+1)}(x)} \\ \text{s.t.} & Ax = b \end{array}$$

linearly-constrained optimization problem

with strictly convex objective function

$$\begin{aligned} \min \quad & \langle c, x \rangle \\ \text{s.t.} \quad & Ax = b \\ & x \in \Lambda_+ \end{aligned}$$

$$z = \text{optimal solution}$$

If $z \notin \partial\Lambda'_{e,+}$ then z solves

$$\begin{aligned} \min_x \quad & -\ln\langle c, e - x \rangle - \frac{p(x)}{p'_e(x)} \\ \text{s.t.} \quad & Ax = b \end{aligned}$$

How good is Newton's method at solving the latter problem?

A general theorem on Newton's method (Smale, Guler, ...)

$$\begin{array}{rl} \min & f(x) \\ \text{s.t.} & Ax = b \end{array} \qquad \text{Let } z \text{ denote optimal solution}$$

For u satisfying $Au = 0$, let $\phi_u(t) := f(z + tu)$, and define

$$\gamma := \sup_{u,\, k>2} \left| \frac{\phi_u^{(k)}(0)}{(k-2)!\, \phi_u^{(2)}(0)^{\frac{k}{2}}} \right|^{\frac{1}{k-2}}$$

Thm: If x satisfies $Ax = b$ and

$$\langle x - z, \nabla^2 f(z)(x - z) \rangle < \frac{1}{36\,\gamma^2}$$

then Newton's method initiated at x converges quadratically.

For interior-point methods:

$$f(x) = \eta \langle c, x \rangle - \ln p(x)$$

$$\gamma \leq 1$$

So $\|x - x(\eta)\|_{\nabla^2 f(x(\eta))} < \frac{1}{6} \;\Rightarrow\;$ quadratic convergence

---

For present context:

$$f(x) = -\ln\langle c, e - x\rangle - \frac{p(x)}{p'_e(x)}$$

γ can be arbitrarily large

("Inversely proportional to curvature of $\partial\Lambda_+$ at z")

$$f(x) = -\ln\langle c, e - x\rangle - \frac{p(x)}{p'_e(x)}$$

Nonetheless, something meaningful can be said ...

Thm:

$$\gamma \leq \frac{4}{\min\{\|x - z\|_{\nabla^2 f(z)} : Ax = b \text{ and } x \in \partial\Lambda'_{e,+}\}}$$

In other words, quadratic convergence occurs on
nearly the largest "ball" within reason.

Limitation of theorem:

$\| \ \|_{\nabla^2 f(z)}$ reflects curvature of $\partial \Lambda_+$ at z,

not shape of $\Lambda'_{e,+}$ around z

That shape is reflected by Hessian of $h(x) := -\ln p'_e(x)$

If $\| \ \|_{\nabla^2 f(z)}$ is (nearly) a scalar multiple of $\| \ \|_{\nabla^2 h(z)}$

then Newton's domain of convergence

is *truly* the largest within reason

$$\underbrace{\begin{array}{rl} \min & \langle c, x\rangle \\ \text{s.t.} & Ax = b \\ & x \in \Lambda_+ \end{array}}_{\text{HP}} \quad \xrightarrow{\text{relax}} \quad \underbrace{\begin{array}{rl} \min & \langle c, x\rangle \\ \text{s.t.} & Ax = b \\ & x \in \Lambda_{e,+}^{(i)} \end{array}}_{\text{HP}_e^{(i)}}$$

z optimal solution of $\text{HP}_e^{(i)}$

$$0 = \lambda_{1,e}^{(i)}(z) \le \ldots \le \lambda_{n-i,e}^{(i)}(z)$$

Cor (to Lax, Vinnikov and Helton Thm)**:**

If $0 < i < n-2$ then there exists a scalar κ such that

$$\kappa\, \lambda_{2,e}^{(i)}(z) \le \left(\frac{\|v\|_{\nabla^2 f(z)}}{\|v\|_{\nabla^2 h(z)}} \right)^2 \le \kappa\, (n-i)\, \lambda_{n-i,e}^{(i)}(z)$$

for all $v \neq 0$ $\underbrace{\text{satisfying } \frac{d}{dt}\lambda_{2,e}^{(i)}(z+tv)|_{t=0} = 0}_{\text{technicality}}$

$e(t)$ time dependent

$z(t)$ optimal solution of $\mathrm{HP}^{(i)}_{e(t)}$

Dynamics: $\frac{d}{dt} e(t) = z(t) - e(t)$

To implement, dynamics should be discretized:

$e_1, e_2, \ldots$ where $e_{j+1} = e_j + \delta\,(z_j - e_j)$ $(0 < \delta < 1)$

Open question: How large can we safely set the value δ?

Zinchenko:

Assume optimal solution z^* of HP is unique

and 0 is a root of multiplicity $i+1$ for $\lambda \mapsto p(\lambda e - z^*)$.

"Then"[1] in the limit, safe values for δ rapidly approach 1.

[1] Additional technical qualifications are used in the proof, but stating them here would take the present talk too far afield.