Hyperbolic Polynomials, Riemannian Geometry and Optimization

Jim Renegar
School of Operations Research and Industrial Engineering
Cornell University
renegar@orie.cornell.edu
\( \mathcal{E} \) Euclidean space

\( p : \mathcal{E} \rightarrow \mathbb{R} \) homogeneous polynomial of degree \( n \)

**Defn:** \( p \) is **hyperbolic**

if there exists \( e \) such that \( p(e) \neq 0 \) and

for every \( x \in \mathcal{E} \), all roots of

\( t \mapsto p(x + te) \) are real.

“hyperbolic in direction \( e \)”

**Examples:**

LP \( p(x) := x_1 x_2 \cdots x_n \quad e := (1, \ldots, 1) \)

SDP \( p(X) := \det(X) \quad e := I \)

Introduced into optimization by:

Güler

Bauschke, Güler, Lewis & Sendov
\[ \lambda \mapsto p(\lambda e - x) \]

"characteristic polynomial of \( x \)"

roots: \[ \lambda(x)_{\text{min}} := \lambda_1(x) \leq \ldots \leq \lambda_n(x) \]

"eigenvalues of \( x \)"

\[ \Lambda_+ := \{ x : \lambda(x) \geq 0 \} \]

"hyperbolicity cone"

**Thm** (Gårding, 1959): \( \Lambda_+ \) is convex

Proof shows \( p \) is hyperbolic for **all** \( e \in \Lambda_{++} \)

**Cor:** For all \( e \in \Lambda_{++} \), \( x \in \mathcal{E} \),

\( \gamma \mapsto p(e - \gamma x) \) has only real roots.

**Pf:** By homogeneity,

\[ p(e - \gamma x) = (-\gamma)^n p(x - (1/\gamma)e). \]
Cor: \[ x \mapsto \lambda(x)_{\text{min}} \text{ is concave} \]

Pf: Use \[ \lambda_i(\alpha x + \beta e) = \alpha \lambda_i(x) + \beta. \]

Of course \( \partial \Lambda_+ = \{ x : \lambda(x)_{\text{min}} = 0 \} \)

Defns:

\[ \text{mult}(x) := \text{multiplicity of 0 as eigenvalue for } x \]

(independent of \( e \in \Lambda_{++} \))

\( \partial_m \Lambda_+ := \{ x : \text{mult}(x) = m \} \quad m = 1, \ldots, n \)

\( \partial_0 \Lambda_+ := \Lambda_{++} \)

Facts:

\[ \partial(\partial_m \Lambda_+) \subseteq \bigcup_{m' > m} \partial_{m'} \Lambda_+ \]

\( \partial_m \Lambda_+ \) is a submanifold of \( \mathcal{E} \)

\( x \mapsto \lambda(x)_{\text{min}} \) is analytic on \( \partial_m \Lambda_+ \)
For \( m > 0, \ x \in \partial_m \Lambda_+ \),
let \( T_x := \) tangent space at \( x \)

Note: if \( v \in T_x \) then
\[
\frac{d}{dt} \lambda(x + tv)_{\text{min}} \bigg|_{t=0} = 0, \\
\frac{d^2}{dt^2} \lambda(x + tv)_{\text{min}} \bigg|_{t=0} \leq 0.
\]

**Thm:** \( \text{span}(T_x \cap \partial_m \Lambda_+) = \{ v \in T_x : \frac{d^2}{dt^2} = 0 \} \)

**Cor:** All faces of \( \Lambda_+ \) are exposed

(generalizes(?) Truong & Tuncel: homogeneous cones)

**Lax Conjecture:** \( \exists \ n \ & \ \text{subsp. } L \text{ s.t. } \Lambda_+ = S^{n \times n} \cap L \)

**Thm** (Chua; also, Faybusovich): True if \( \Lambda_+ \) homogeneous
Derivative polynomial:

\[ p'(x) := \frac{d}{dt} p(x + te)|_{t=0} = \langle \nabla p(x), e \rangle \]

\[ p' (\lambda e - x) = \frac{d}{d\lambda} p(\lambda e - x) \]

\(p\) hyperbolic \(\Rightarrow\) \(p'\) hyperbolic

Eigenvalues of \(x\) w.r.t. \(p'\): \(\lambda'_1(x) \leq \ldots \leq \lambda'_{n-1}(x)\)

Interlacing:
\(\lambda_1(x) \leq \lambda'_1(x) \leq \lambda_2(x) \leq \ldots \leq \lambda'_{n-1}(x) \leq \lambda_n(x)\)

Consequence: \(\Lambda_+ \subseteq \Lambda'_+\)

Moreover,

- for \(m \geq 2\), \(\partial_m \Lambda'_+ = \partial_{m+1} \Lambda_+\)
- \((\partial_1 \Lambda'_+) \cap \Lambda_+ = \partial_2 \Lambda_+\)
- \(x \in \partial \Lambda'_+ \& x \notin \Lambda_+ \Rightarrow x\) extreme for \(\Lambda'_+\)
Higher derivatives:

\[ p^{(k+1)}(x) := \frac{d}{dt} p^{(k)}(x + te) \big|_{t=0} \]

Then

\[ \Lambda_+ = \Lambda_+^{(0)} \subseteq \Lambda_+^{(1)} \subseteq \cdots \subseteq \Lambda_+^{(n-1)} \]

and for \( m \geq 2 \),

\[ \partial_m \Lambda_+^{(k)} = \partial_{m+1} \Lambda_+^{(k-1)} = \cdots = \partial_{m+k} \Lambda_+ \]

Thus,

\[ p(\lambda e - x) = \sum_{k=0}^{n} \frac{1}{k!} p^{(k)}(-x) \lambda^k \]

\[ = \sum_{k=0}^{n} \frac{(-1)^{n-k}}{k!} p^{(k)}(x) \lambda^k \]

Thus, \( \frac{1}{k!} p^{(k)}(x) = E_{n-k}(\lambda(x)) \)

where

\[ E_j(\lambda_1, \ldots, \lambda_n) := \sum_{i_1 < i_2 < \ldots < i_j} \lambda_{i_1} \lambda_{i_2} \cdots \lambda_{i_j} \]
Consequence:

\[ \Lambda_+^{(k)} = \{ x : p^{(j)}(x) \geq 0, \ j = k, \ldots, n \} \]
\[ = \{ x : E_i(x) \geq 0, \ i = 1, \ldots, n - k \} \]

Defn: \( \text{trace}_e(x) := E_1(x) = \lambda_1(x) + \cdots + \lambda_n(x) \)

Fact: \( (\Lambda_+^*)^o = \{ \text{trace}_e : e \in \Lambda_{++} \} \)

Hyperbolic program (HP):

\[ \begin{align*}
\min & \quad \text{trace}(x) \\
\text{s.t.} & \quad Ax = b \\
& \quad x \in \Lambda_+ 
\end{align*} \]

Relaxations (HP\(^{(k)}\)):
\[ x \in \Lambda_+^{(k)} \]

HP feasible \(\Leftrightarrow\) HP has optimal solution
\[ \Rightarrow \text{HP}^{(k)} \text{ has optimal solution} \]
Morphing

Observation:

Eigenvalues of $x \mapsto \text{trace}(x) p'(x)$ are

$$\lambda'_1(x), \ldots, \lambda'_{n-1}(x), \frac{1}{n-1} \sum_j \lambda'_j(x)$$

Thus,

$$\Lambda'_+ = \text{hyperbolicity cone for } x \mapsto \text{trace}(x)p'(x)$$

Thm: $x \mapsto (1 - \epsilon) p(x) + \epsilon \text{trace}(x)p'(x)$ is hyperbolic if $0 < \epsilon < 1$

Hyperbolicity cone: $\Lambda_+^{(\epsilon)}$

Facts:

$$\Lambda_+ \subseteq \Lambda_+^{(\epsilon)} \subseteq \Lambda'_+$$

$$\Lambda_+ \cap \Lambda_+^{(\epsilon)} = \Lambda'_+ \cap \Lambda_+^{(\epsilon)} = \Lambda_+ \cap \Lambda'_+$$

$$x \in \partial \Lambda_+^{(\epsilon)} \& x \notin \Lambda_+ \Rightarrow x \text{ extreme for } \Lambda_+^{(\epsilon)}$$
\( \text{Opt}^{(\epsilon)} \) := optimal solution set for \( \text{HP}^{(\epsilon)} \)

**Thm:**

All \( 0 < \epsilon \leq 1 \) satisfy

\[
\text{Opt}^{(\epsilon)} = \text{Opt}^{(0)}
\]

or all \( 0 < \epsilon \leq 1 \) satisfy

\[
\text{Opt}^{(\epsilon)} = \{ z^{(\epsilon)} \} \quad \text{where} \quad \frac{d}{d\epsilon} z^{(\epsilon)} \neq 0
\]

Higher derivatives:

\[
\epsilon p^{(k)}(x) + (1 - \epsilon) \text{trace}(x) p^{(k+1)}(x)
\]

\[
\text{Opt}^{(k+\epsilon)} = \{ z^{(k+\epsilon)} \}
\]

**Follow the path to optimality**

starting point: \( z^{(n-2)} \) (easily computed)
Thm: Fix $\alpha, \beta > 0$.

If $q_1, q_2$ are hyperbolic in direction $e$

and $k < \deg(q_1) + \deg(q_2)$

then

$$\sum_{j=0}^{k} \binom{k}{j} \alpha^j \beta^{k-j} q_1^{(j)} q_2^{(k-j)}$$

is hyperbolic in direction $e$.

Pf:

• $Q(x, t) := q_1(x + t\alpha e) q_2(x + t\beta e)$

• Hyperbolic in direction $(0, 1)$

• $(e, 0)$ in hyperbolicity cone of $Q$, hence of $Q^{(k)}$

• Thus, $x \mapsto Q^{(k)}(x, 0)$ is hyperbolic in direction $e$ \hfill \Box

Consequence: Can morph directly from $\Lambda_+^{(k)}$ to $\Lambda_+$

Downside: Don’t gain facial structure along the way
Path Following

\[ p_\epsilon := (1 - \epsilon) \, p + \epsilon \, \text{trace} \, p' \]

HP\(^{(\epsilon)}\): \min \text{trace}(x), \text{ s.t. } Ax = b, \ x \in \Lambda^{(\epsilon)}_+ \]

\[ z^{(\epsilon)} = \text{optimal solution} \]

Predictor step:

1. Compute \( d \), the path’s tangent direction at \( z^{(\epsilon)} \)

2. Compute \( t := \min \{ t : p(z^{(\epsilon)} + td) = 0 \} \)
   (then \( z^{(\epsilon)} + td \in \partial \Lambda_+ \))

3. Let \( x := z^{(\epsilon)} + (.99)td \)

Corrector steps:

4. Let \( \epsilon' := p(x)/(p(x) - p'(x)) \)
   (then \( 0 < \epsilon' < \epsilon \) and \( x \in \partial \Lambda^{(\epsilon')}_+ \))

5. Move from \( x \) along \( \partial \Lambda^{(\epsilon')}_+ \) to \( z^{(\epsilon')} \)
Compute gradients and Hessians with FFT:

Relies on the identity

$$\nabla p^{(k)}(x) = \frac{k!}{n} \sum_{i=1}^{n} \omega_i^{n-k} \nabla p(x + \omega_i e)$$

where $\omega_1, \ldots, \omega_n$ are the $n^{th}$ roots of unity.

Cost per coordinate beyond evaluating $\nabla p(x + \omega_i e)$:

$O(n \log^2 n)$ arithmetic operations

(computes the coordinate for all $k = 1, \ldots, n$)
Riemannian Metrics

The ubiquitous metric on \((\Lambda_+')^\circ\):

\[
\langle u, v \rangle_x := \langle u, \nabla^2 f(x)v \rangle
\]

where \(f(x) := -\ln p'(x)\)

**Shortcoming:** Does not encode \(\partial \Lambda_+\)

Potentially useful for corrector steps:

“Move from \(x\) along \(\partial \Lambda_+^{(\epsilon)}\) to \(z^{(\epsilon)}\)”

\(f|_{\partial_1 \Lambda_+^{(\epsilon)}}\): a “log barrier fn” for the boundary

e.g., if \(\epsilon = 0\) & \(\Lambda_+ = \mathbb{R}_+^n\) then \(-\sum_{j=1}^{n-1} \ln x_j\)
A Corrector Strategy:

Assume $A e = 0$ (unrestrictive)

1. Compute $\tilde{e}$ satisfying $\text{trace}(y) = \langle \tilde{e}, y \rangle_x \ \forall y$  
   (i.e., $\tilde{e} = \nabla_R \text{trace}$, the Riemannian gradient)

2. Orthogonally project: $d := -P_{T,R} \tilde{e}$  
   ($T$ = tang. sp. of $\partial \Lambda_+^{(e)} \cap \{ x : Ax = b \}$ at $x$)

3. Choose $\tau > 0$. Let $\bar{x} := x + \tau d$.

4. Compute $t := \max \{ t : p_e(\bar{x} + te) = 0 \}$.  
   Let $x_+ := \bar{x} + te$.

Prop: If $\text{dist}(x, z^{(e)}) < .1$ then

$$\text{dist}(x_+, z^{(e)}) \leq 10 \text{dist}(x, z^{(e)})^2$$
Towards a second metric . . .

\[ h(x) := -\frac{p(x)}{p'(x)} = \frac{-1}{\sum_j (1/\lambda_j(x))} \]

**Thm:** \( h : (\Lambda'_+)^\circ \to \mathbb{R} \) is convex

**Pf:**

- Let \( Q(x, t) := t \, p(x) \).
- Hyperbolic in direction \((e, 1)\), hence so is \( Q' \).
- But hyperbolicity cone for \( Q' = \text{epigraph of } h \).

For \( x \in (\Lambda'_+)^\circ \setminus \Lambda_+ \):

0 is an eigenvalue for \( \nabla^2 h(x) \) of multiplicity 1;

But \( x \) is the eigenvector;

Hence irrelevant because \( Ax = b \neq 0 \).

**The second metric:** For \( u, v \in \text{null}(A) \),

\[ \langle u, v \rangle_x := \langle u, \nabla^2 h(x)v \rangle \]
Observe:

\[ \partial \Lambda_+^{(\varepsilon)} = \{ x : (1 - \varepsilon)h(x) = \varepsilon \text{trace}(x) \} \]

Consequently, \( z^{(\varepsilon)} \) is also the optimal solution for

\[ \min h(x), \text{ s.t. } Ax = b, \quad x \in \Lambda_+^{(\varepsilon)} \]

**Corrector Strategy:**

1. Compute \( \tilde{e} \) satisfying \( \text{trace}(y) = \langle \tilde{e}, y \rangle_x \quad \forall y \)

2. Orthogonally project: \( d := -P_{T,R}\tilde{e} \)

   (then \( d = -\frac{\varepsilon}{1-\varepsilon}P_{T,R}\nabla_R h(x) \))

3. Line search: \( \bar{\tau} := \arg \min h(x + \tau d) \)

   Let \( \bar{x} := x + \bar{\tau}d \)

4. Compute \( t := \max \{ t : p_{\varepsilon}(\bar{x} + te) = 0 \} \).

   Let \( x_+ := \bar{x} + te \).

**Prop:** Converges to \( z^{(\varepsilon)} \) if initial point is in

connected component of \( (\partial_1 \Lambda_+^{(\varepsilon)}) \cap \{ x : Ax = b \} \)

containing \( z^{(\varepsilon)} \).
Something with the flavor of duality . . .

Assume \( e \in \text{null}(A) \) (unrestrictive)

Let \( C \) be a connected component of

\[
(\partial_1 \Lambda_+) \cap \{ x : Ax = b \}
\]

Assume some \( x \in C \) is an extreme point of

\[
\Lambda_+ \cap \{ x : Ax = b \}
\]

Then all \( x \in C \)

- are extreme points
- satisfy \( \text{null}(\nabla^2 h(x)) \cap \text{null}(A) = \{0\} \)

Moreover,

- \( e \perp_x T_x \) for all \( x \in C \),
- \( x \mapsto P_{\text{null}(A)} \nabla h(x) \) maps \( C \) diffeomorphically onto a convex set.
A final metric:

Replace $\nabla^2 h$ with $\nabla^2 h + \frac{1}{h}(\nabla h)(\nabla h)^T$

 Metric on $(\Lambda'_+)^{\circ} \setminus \Lambda_+$

Some Properties:

- $\|x\|_x = h(x)$

- For $x \in \partial_1 \Lambda_+^{(\varepsilon)}$, let $T$ be tangent space. Let $u(x) = -P_{T,\mathcal{R}}x$
  (so $x + u(x)$ solves $\min \|y\|_x$, s.t. $y \in x + T$)
  Then the induced flow on $\Lambda_+^{(\varepsilon)}$ converges to $z^{(\varepsilon)}$
  (i.e., to where we want to be)

- On the other hand, let $v(z^{(\varepsilon)}) := -P_{T,\mathcal{R}}z^{(\varepsilon)}$.
  Then $v(z^{(\varepsilon)})$ is tangent to path of optimal solutions, but in the direction which increases $h$
  (i.e., away from where we want to be)