Optimizing Over Hyperbolicity Cones
By Using Their Derivative Relaxations

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

$$\lambda \mapsto p(\lambda e - x)$$

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 \( e \):
\[ \frac{x_1}{e_1}, \ldots, \frac{x_n}{e_n} \]

SDP:
- \( p(x) = \det(x) \)
- \( e \succ 0 \)
- \[ \lambda \mapsto \det(\lambda e - x) = \det(e) \det(\lambda I - e^{-1/2}xe^{-1/2}) \]

Eigenvalues of \( x \) in direction \( e \)
\[ = \text{traditional eigenvalues of } e^{-1/2}xe^{-1/2} \]
\[ \lambda_1, e(x) \leq \lambda_2, e(x) \leq \cdots \leq \lambda_n, e(x) \quad \text{roots of } \lambda \mapsto p(x - \lambda e) \]

Hyperbolicity Cone:

\[ \Lambda_{++} := \{ x : 0 < \lambda_1, e(x) \} \]

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

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

Corollary: \( \Lambda_{++} \) is a convex cone

Corollary: \( x \mapsto \lambda_n, e(x) \) is a convex function
Bauschke, Güler, Lewis & Sendov:

If \( f : \mathbb{R}^n \rightarrow \mathbb{R} \) is a convex and permutation-invariant then \( x \mapsto f(\lambda_e(x)) \) 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.
a univariate polynomial
If $\phi$ has only real roots then:
  $\phi'$ has only real roots.
  Roots are interlaced: $\lambda_1 \leq \lambda'_1 \leq \lambda_2 \leq \cdots \leq \lambda'_{n-1} \leq \lambda_n$

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:

\[ p_{e}^{(i+1)}(x) = \langle \nabla p_{e}^{(i)}(x), e \rangle \]

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

\[
\begin{align*}
p_{e}^{(i)}(x) &= i! \ p(e) \ E_{n-i}(\vec{\lambda}_{e}(x)) \\
&\text{where } E_{k} = \text{elementary symmetric polynomial of degree } k
\end{align*}
\]

\[
\Lambda_{e, +}^{(i)} = \{ x : E_{k}(\vec{\lambda}_{e}(x)) \geq 0, \ k = 1, \ldots, n - i \}
\]
Hyperbolic Program (HP):

\[
\min \langle c, x \rangle \\
\text{s.t. } Ax = b \\
x \in \Lambda_+
\]

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

“Central Path” = \{x(\eta) : \eta > 0\}

where \(x(\eta)\) solves

\[
\min \eta \langle c, x \rangle - \ln \rho(x) \\
\text{s.t. } Ax = b
\]

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

to reduce \(\alpha := \langle c, x \rangle - \langle b, y \rangle\) to \(\epsilon \alpha\)
\[ \begin{align*}
\text{min} & \quad \langle c, x \rangle \\
\text{s.t.} & \quad Ax = b \\
& \quad x \in \Lambda_+ \\
\text{min} & \quad \langle c, x \rangle \\
\text{s.t.} & \quad Ax = b \\
& \quad x \in \Lambda_{+,e}^{(i)}
\end{align*} \]
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 $e$  □

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

Downside: Don’t gain facial structure along the way
\[
\begin{align*}
\min & \quad \langle c, x \rangle \\
\text{s.t.} & \quad Ax = b \\
& \quad x \in \Lambda_+ \\
\end{align*}
\]

\[z = \text{optimal solution}\]

If \( z \notin \partial \Lambda_{e,+} \) then \( z \) solves
\[
\begin{align*}
\min_x & \quad - \ln \langle c, e - x \rangle - \frac{p(x)}{p'(e(x))} \\
\text{s.t.} & \quad Ax = b \\
\end{align*}
\]

How good is Newton’s method at solving the latter problem?
**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 $x \mapsto -p(x)/p'_e(x)$
\[
\min \langle c, x \rangle \\
\text{s.t.} \quad Ax = b \\
x \in \Lambda_+
\]

\[z = \text{optimal solution}\]

If \( z \notin \partial \Lambda_{e,+} \) then \( z \) solves

\[
\min_x - \ln \langle c, e - x \rangle - \frac{p(x)}{p_e(x)} \\
\text{s.t.} \quad Ax = b
\]

How good is Newton’s method at solving the latter problem?
A general theorem on Newton’s method (Smale, Guler, ...)

\[
\begin{align*}
\min_{x} & \quad f(x) \\
\text{s.t.} & \quad Ax = b
\end{align*}
\]

Let \( z \) denote optimal solution

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

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

**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:

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

\( \gamma \) 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:

\[ \| \nabla^2 f(z) \| \text{ reflects curvature of } \partial \Lambda_+ \text{ at } z, \]

\[ \text{not shape of } \Lambda'_{e,+} \text{ around } z \]

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 \textit{truly} the largest within reason