Computer algebra for hyperbolic programming

Simone Naldi

(joint with **D. Plaumann**)

Technische Universität Dortmund

JNCF 2017

Luminy, 17/01/2017

technische universität dortmund

fakultät für mathematik

Hyperbolic polynomials

Definition of hyperbolic polynomial

$f \in \mathbb{R}[x]_d$ is *hyperbolic w.r.t.* $e = (e_1, \ldots, e_n) \in \mathbb{R}^n$ if

- $f(e) \neq 0$ (we suppose w.l.o.g. $f(e) = 1$)
- $\forall\, a \in \mathbb{R}^n \quad t \mapsto ch_a(t) := f(t\, e - a)$ has **only real roots**

If such e exists, f is called a *hyperbolic polynomial*.

Hyperbolic polynomials

Definition of hyperbolic polynomial

$f \in \mathbb{R}[x]_d$ is *hyperbolic w.r.t.* $e = (e_1, \dots, e_n) \in \mathbb{R}^n$ if

- $f(e) \neq 0$ (we suppose w.l.o.g. $f(e) = 1$)
- $\forall\, a \in \mathbb{R}^n \quad t \mapsto ch_a(t) := f(t\, e - a)$ has **only real roots**

If such e exists, f is called a *hyperbolic polynomial*.

Fundamental examples:

(1) **Products of real linear forms**: $f = x_1 \cdots x_d$
For $e = \mathbf{1} = (1, \dots, 1) \qquad ch_a(t) = (t - a_1) \cdots (t - a_d)$

Hyperbolic polynomials

Definition of hyperbolic polynomial

$f \in \mathbb{R}[x]_d$ is *hyperbolic w.r.t.* $e = (e_1, \dots, e_n) \in \mathbb{R}^n$ if

- $f(e) \neq 0$ (we suppose w.l.o.g. $f(e) = 1$)
- $\forall\, a \in \mathbb{R}^n \quad t \mapsto ch_a(t) := f(t\,e - a)$ has **only real roots**

If such e exists, f is called a *hyperbolic polynomial*.

Fundamental examples:

(1) **Products of real linear forms**: $f = x_1 \cdots x_d$
For $e = \mathbf{1} = (1, \dots, 1) \qquad ch_a(t) = (t - a_1) \cdots (t - a_d)$

(2) **Symmetric determinant**: $f = \det(X)$, X symmetric matrix
For $e = \mathbb{I}_d \qquad ch_a(t)$ is the characteristic polynomial of $a \in \mathbb{S}_d(\mathbb{R})$

Convex optimization: (1) Linear Programming and (2) Semidefinite Programming

Hyperbolicity cones

Definition of hyperbolicity cone

The *hyperbolicity cone* of $f \in \mathbb{R}[x]_d$ (*w.r.t.* e) is

$$\Lambda_+(f, e) = \{a \in \mathbb{R}^n : ch_a(t) = 0 \Rightarrow t \geq 0\}$$

Hyperbolicity cones

Definition of hyperbolicity cone

The *hyperbolicity cone* of $f \in \mathbb{R}[x]_d$ (*w.r.t.* e) is

$$\Lambda_+(f, e) = \{a \in \mathbb{R}^n : ch_a(t) = 0 \Rightarrow t \geq 0\}$$

Base-cases:

(1) For $f = x_1 \cdots x_d$, $e = \mathbf{1}$: $\Lambda_+(f, \mathbf{1}) = \mathbb{R}^n_+$ (LP)

(2) $f = \det(X)$, $e = \mathbb{I}_d$: $\Lambda_+(f, \mathbb{I}_d) =$ PSD cone (SDP)

Hyperbolicity cones

Definition of hyperbolicity cone

The *hyperbolicity cone* of $f \in \mathbb{R}[x]_d$ (*w.r.t.* e) is

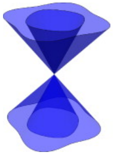
$$\Lambda_+(f, e) = \{a \in \mathbb{R}^n : ch_a(t) = 0 \Rightarrow t \geq 0\}$$

Base-cases:

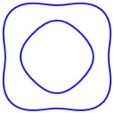
(1) For $f = x_1 \cdots x_d$, $e = \mathbf{1}$: $\Lambda_+(f, \mathbf{1}) = \mathbb{R}^n_+$ (LP)

(2) $f = \det(X)$, $e = \mathbb{I}_d$: $\Lambda_+(f, \mathbb{I}_d) = $ PSD cone (SDP)

(3) General case (here $d = 4$) :



$\mathbb{R}^3$



$\mathbb{P}^2$

Hyperbolicity cones

Definition of hyperbolicity cone

The *hyperbolicity cone* of $f \in \mathbb{R}[x]_d$ (*w.r.t.* e) is

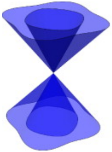
$$\Lambda_+(f,e) = \{a \in \mathbb{R}^n : ch_a(t) = 0 \Rightarrow t \geq 0\}$$

Base-cases:

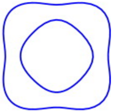
(1) For $f = x_1 \cdots x_d$, $e = \mathbf{1}$: $\quad \Lambda_+(f,\mathbf{1}) = \mathbb{R}^n_+$ (LP)

(2) $f = \det(X)$, $e = \mathbb{I}_d$: $\quad \Lambda_+(f,\mathbb{I}_d) =$ PSD cone (SDP)

(3) General case (here $d = 4$) :



$\mathbb{R}^3$



$\mathbb{P}^2$

A hierarchy of convex opt. problems:

Linear Programming (LP)

$\Downarrow$

Semidefinite Programming (SDP)

$\Downarrow$

Hyperbolic Programming (HP)

Can we design Algebraic/Exact methods?

The Lax conjecture

Examples of hyperbolic polynomials

- Elementary Symmetric polynomials $f = \sum x_i, \sum x_i x_j, \ldots$
- Derivatives along hyperbolic directions: f hyperb. $\Rightarrow \sum e_i \frac{\partial f}{\partial x_i}$ hyperb.
- $f = \det(A_1 x_1 + \cdots + A_n x_n)$, where $\exists e$ with $e_1 A_1 + \cdots + e_n A_n \succ 0$

Example (Brändén)

There exists $f \in \mathbb{R}[x_1, \ldots, x_8]$ hyperbolic but no symmetric matrices $A_1, \ldots, A_8$ with $f = \det(A_1 x_1 + \cdots + A_8 x_8)$ and $e_1 A_1 + \cdots + e_8 A_8 \succ 0$

The Lax conjecture

Examples of hyperbolic polynomials

- Elementary Symmetric polynomials $f = \sum x_i, \sum x_i x_j, \dots$
- Derivatives along hyperbolic directions: f hyperb. $\Rightarrow \sum e_i \frac{\partial f}{\partial x_i}$ hyperb.
- $f = \det(A_1 x_1 + \cdots + A_n x_n)$, where $\exists e$ with $e_1 A_1 + \cdots + e_n A_n \succ 0$

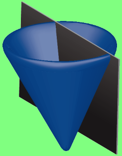
Example (Brändén)

There exists $f \in \mathbb{R}[x_1, \dots, x_8]$ hyperbolic but no symmetric matrices $A_1, \dots, A_8$ with $f = \det(A_1 x_1 + \cdots + A_8 x_8)$ and $e_1 A_1 + \cdots + e_8 A_8 \succ 0$

Generalized Lax conjecture

Every hyperbolicity cone is a **linear section of the cone of PSD symmetric matrices**, that is $\exists A_1, \dots, A_n$ such that

$$\Lambda_+(f, e) = \{x \in \mathbb{R}^n : A_1 x_1 + \ldots + A_n x_n \succeq 0\}$$



If the conjecture holds, then HP coincides with SDP.

Origin

From *hyperbolic PDE theory*

The *Cauchy problem* (given $f \in \mathbb{R}[x]_{\leq d}$ and $\Omega \subset \mathbb{R}^n$ open) :

Given $p \in C^\infty(\Omega)$ compute $u \in C^\infty(\Omega)$ such that $f(\partial_1, \ldots, \partial_n)u = p$.

Theorem (Lax, Mizohata)

Decompose $f = \sum_{i \leq d} f_i$ *with* $f_i \in \mathbb{R}[x]_i$.
If *the Cauchy problem is well-posed* **then** f_d *is hyperbolic.*

Origin

From *hyperbolic PDE theory*

The *Cauchy problem* (given $f \in \mathbb{R}[x]_{\leq d}$ and $\Omega \subset \mathbb{R}^n$ open) :

Given $p \in C^\infty(\Omega)$ compute $u \in C^\infty(\Omega)$ such that $f(\partial_1, \ldots, \partial_n)u = p$.

Theorem (Lax, Mizohata)

Decompose $f = \sum_{i \leq d} f_i$ *with* $f_i \in \mathbb{R}[x]_i$.
If *the Cauchy problem is well-posed* **then** f_d *is hyperbolic.*

EX: The **Wave operator** $(\partial_t^2 - \sum_i \partial_i^2)u = p$ corresponds to the polynomial

$$f = x_{n+1}^2 - \sum_{i=1}^{n} x_i^2$$

hyperbolic in direction $e = (1, 0, \ldots, 0)$. Its hyp. cone is the **second-order (or Lorentz) cone**

$$\lambda_+(f, (1, 0, \ldots, 0)) = \{x \in \mathbb{R}^{n+1} : x_{n+1} \geq \sqrt{x_1^2 + \cdots + x_n^2}\}$$

Optimization over Multiplicity sets

Problem 1 (Hyperbolic Programming).
Given $f \in \mathbb{R}[x]_d$ hyperbolic in dir. e, and ℓ linear, solve

$$\inf\{\ell(a) : a \in \Lambda_+(f, e)\} \tag{1}$$

Optimization over Multiplicity sets

Problem 1 (Hyperbolic Programming).
Given $f \in \mathbb{R}[x]_d$ hyperbolic in dir. e, and ℓ linear, solve

$$\inf\{\ell(a) : a \in \Lambda_+(f, e)\} \tag{1}$$

Multiplicity: For $a \in \mathbb{R}^n$, we define

$$\text{mult}(a) := \text{multiplicity of } 0 \text{ as root of } \text{ch}_a(t) = f(te - a)$$

Multiplicity set: For $m \leq d$, $\Gamma_m = \{a \in \mathbb{R}^n : \text{mult}(a) \geq m\}$

Optimization over Multiplicity sets

Problem 1 (Hyperbolic Programming).
Given $f \in \mathbb{R}[x]_d$ hyperbolic in dir. e, and ℓ linear, solve

$$\inf\{\ell(a) : a \in \Lambda_+(f, e)\} \tag{1}$$

Multiplicity: For $a \in \mathbb{R}^n$, we define

$$\text{mult}(a) := \text{multiplicity of } 0 \text{ as root of } \text{ch}_a(t) = f(te - a)$$

Multiplicity set: For $m \leq d$, $\Gamma_m = \{a \in \mathbb{R}^n : \text{mult}(a) \geq m\}$

Remark: The set Γ_m is **real algebraic**.
Indeed, if $\text{ch}_a(t) = t^d + g_1(a)t^{d-1} + \cdots + g_{d-1}(a)t + g_d(a)$ then

$$\Gamma_m = \{a : g_i(a) = 0, i \geq d - m + 1\}$$

Optimization over Multiplicity sets

Problem 1 (Hyperbolic Programming).
Given $f \in \mathbb{R}[x]_d$ hyperbolic in dir. e, and ℓ linear, solve

$$\inf\{\ell(a) : a \in \Lambda_+(f, e)\} \tag{1}$$

Multiplicity: For $a \in \mathbb{R}^n$, we define

$$\text{mult}(a) := \text{multiplicity of } 0 \text{ as root of } \text{ch}_a(t) = f(te - a)$$

Multiplicity set: For $m \leq d$, $\Gamma_m = \{a \in \mathbb{R}^n : \text{mult}(a) \geq m\}$

Remark: The set Γ_m is **real algebraic**.
Indeed, if $\text{ch}_a(t) = t^d + g_1(a)t^{d-1} + \cdots + g_{d-1}(a)t + g_d(a)$ then

$$\Gamma_m = \{a : g_i(a) = 0, i \geq d - m + 1\}$$

Theorem 1. If $\overline{x}$ is a minimizer of (2), and $\overline{m} = \text{mult}(\overline{x})$, then $\overline{x}$ is a local minimum of ℓ on $\Gamma_{\overline{m}}$

Optimization over Multiplicity sets

Sketch of Algorithm for Problem 1:

INPUT

$f \in \mathbb{R}[x]_d, e \in \mathbb{R}^n, \ell \in \mathbb{R}[x]_1$

OUTPUT

A finite set (parametrized by Rational Univ. Repres.) containing the minimizer

PROCEDURE

For $m = 0, \ldots, d$ do

- Compute the ideal $I_m = \mathrm{crit}(\ell, \Gamma_m)$ of critical points of ℓ on Γ_m
- Compute a Rational Univ. Repres. of I_m

Optimization over Multiplicity sets

Sketch of Algorithm for Problem 1:
INPUT

$f \in \mathbb{R}[x]_d, e \in \mathbb{R}^n, \ell \in \mathbb{R}[x]_1$

OUTPUT

A finite set (parametrized by Rational Univ. Repres.) containing the minimizer

PROCEDURE

For $m = 0, \dots, d$ do

- Compute the ideal $I_m = \mathrm{crit}(\ell, \Gamma_m)$ of critical points of ℓ on Γ_m
- Compute a Rational Univ. Repres. of I_m

Problem 2 (Maximum multiplicity).
Given $f \in \mathbb{R}[x]_d$ hyperbolic in dir. e, solve the *non-convex* opt. prob.:

$$\max\{\mathrm{mult}(a) : a \in \Lambda_+(f, e)\} \tag{2}$$

Optimization over Multiplicity sets

Sketch of Algorithm for Problem 1:
INPUT
$f \in \mathbb{R}[x]_d, e \in \mathbb{R}^n, \ell \in \mathbb{R}[x]_1$
OUTPUT
A finite set (parametrized by Rational Univ. Repres.) containing the minimizer
PROCEDURE
For $m = 0, \ldots, d$ do

- Compute the ideal $I_m = \text{crit}(\ell, \Gamma_m)$ of critical points of ℓ on Γ_m
- Compute a Rational Univ. Repres. of I_m

Problem 2 (Maximum multiplicity).
Given $f \in \mathbb{R}[x]_d$ hyperbolic in dir. e, solve the *non-convex* opt. prob.:

$$\max\{\text{mult}(a) : a \in \Lambda_+(f, e)\} \tag{2}$$

Theorem 2. Suppose $m^* = \max(\text{mult}(a))$ in $\Lambda_+(f, e)$. Then one of the real connected components of Γ_{m^*} is a subset of $\Lambda_+(f, e)$.

The special case of LMI/SDP

$$f = \det(\mathbf{A}(x)) \text{ with } A(x) = x_1 A_1 + \cdots + x_n A_n$$

- $\Lambda_+(f, e) = \{x \in \mathbb{R}^n : \mathbf{A}(x) \succeq 0\}$ (HP reduces to an SDP)
- $\text{mult}(a) \equiv \text{corank}(A(a))$.
- Multiplicity set $\leftrightarrow$ Determinantal variety $\Gamma_m = \{x \in \mathbb{R}^n : \text{rank}\mathbf{A}(x) \leq d - m\}$

Optimality conditions:

$$x \in \Gamma_m \iff \text{rank}\mathbf{A}(x) \leq d - m \iff \mathbf{A}(\mathbf{x})\mathbf{Y}(\mathbf{y}) = \mathbf{0}$$

The special case of LMI/SDP

$$f = \det(\mathbf{A}(x)) \text{ with } A(x) = x_1 A_1 + \cdots + x_n A_n$$

- $\Lambda_+(f, e) = \{x \in \mathbb{R}^n : \mathbf{A}(x) \succeq 0\}$ (HP reduces to an SDP)
- $\text{mult}(a) \equiv \text{corank}(A(a))$.
- Multiplicity set $\leftrightarrow$ Determinantal variety $\Gamma_m = \{x \in \mathbb{R}^n : \text{rank}\mathbf{A}(x) \leq d - m\}$

Optimality conditions:

$$x \in \Gamma_m \iff \text{rank}\mathbf{A}(x) \leq d - m \iff \mathbf{A}(\mathbf{x})\mathbf{Y}(\mathbf{y}) = \mathbf{0}$$

Henrion, N., Safey El Din (2015-2016)
Exact algorithms for linear matrix inequalities: $\mathbf{A}(x) \succeq 0$
N. (ISSAC 2016)
Rank-constrained SDP (poly-time if n **or** $d = \text{size}(\mathbf{A})$ is fixed)

SPECTRA: Maple library for linear matrix inequalities

Work in progress!
Can we get the same complexity bounds for general hyperbolic polynomials?

Renegar's derivative relaxations

Fundamental remark:

$f \in \mathbb{R}[x]_d$ hyperbolic in direction $e \quad \Rightarrow \quad D_e f = \sum_i e_i \frac{\partial f}{\partial x_i}$ still hyperbolic

This gives a nested sequence of convex hyperbolicity cones:

$$\Lambda_+(f, e) \subset \Lambda_+(D_e f, e) \subset \cdots \subset \Lambda_+(D_e^{(n-1)} f, e)$$

(the last one being a half-space), giving a sequence of *lower bounds* for the linear function to optimize:

$$\inf_{\Lambda_+(f,e)} \ell(a) \geq \inf_{\Lambda_+(D_e f,e)} \ell(a) \geq \cdots \geq \inf_{\Lambda_+(D_e^{(d-1)} f,e)} \ell(a)$$

Renegar's derivative relaxations

Fundamental remark:

$f \in \mathbb{R}[x]_d$ hyperbolic in direction $e \quad \Rightarrow \quad D_e f = \sum_i e_i \frac{\partial f}{\partial x_i}$ still hyperbolic

This gives a nested sequence of convex hyperbolicity cones:

$$\Lambda_+(f,e) \subset \Lambda_+(D_e f, e) \subset \cdots \subset \Lambda_+(D_e^{(n-1)} f, e)$$

(the last one being a half-space), giving a sequence of *lower bounds* for the linear function to optimize:

$$\inf_{\Lambda_+(f,e)} \ell(a) \geq \inf_{\Lambda_+(D_e f,e)} \ell(a) \geq \cdots \geq \inf_{\Lambda_+(D_e^{(d-1)} f,e)} \ell(a)$$

Why Renegar's method is useful from a computational viewpoint:

- At each step of the relaxation, the **degree** of the polynomial **decreses** by 1
- One of the $\Lambda_+(D_e^{(j)} f, e)$ could be a section of the PSD cone (solution set of a **LMI**), in which case a lower bound can be computed by solving a *single SDP*.

N-ellipse

Given N points $P_1, \ldots, P_n$ in $\mathbb{R}^2$, and $D \in \mathbb{R}_+$.
The *N-ellipse* is the set $\mathcal{E}_N$ of $Q \in \mathbb{R}^2$ satisfying

$$\sum_{i=1}^{N} \mathrm{dist}(Q, P_i) = D.$$

Fact: The polynomial f vanishing on the boundary of $\mathcal{E}_N$ is *hyperbolic* (for all N, for general P_i).

Remark! The degree of f is 2^N.

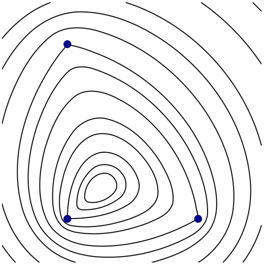


Figure: $3-$ellipse for many D

N-ellipse

Given N points $P_1, \ldots, P_n$ in $\mathbb{R}^2$, and $D \in \mathbb{R}_+$.
The *N-ellipse* is the set $\mathcal{E}_N$ of $Q \in \mathbb{R}^2$ satisfying

$$\sum_{i=1}^{N} \mathrm{dist}(Q, P_i) = D.$$

Fact: The polynomial f vanishing on the boundary of $\mathcal{E}_N$ is *hyperbolic* (for all N, for general P_i).

Remark! The degree of f is 2^N.

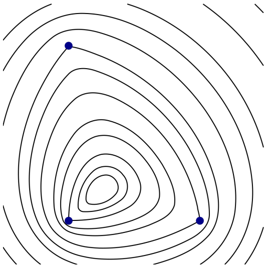


Figure: $3-$ellipse for many D

For $N = 3$, $P_1 = (0, 4)$, $P_2 = (0, 0)$, $P_3 = (3, 0)$ (using Renegar's derivatives) :

k	$\approx x^*$	m^*	$\ell(x^*)$	Degree of ex. repr.
0	$(0.750, 0.000, 0.250)$	2	5.500000000	56
1	$(0.759, -0.018, 0.258)$	1	5.499158216	42
2	$(0.797, -0.051, 0.250)$	1	5.456196445	30
3	$(0.862, -0.116, 0.254)$	1	5.392044926	20
4	$(0.981, -0.254, 0.273)$	1	5.292250029	12
5	$(1.336, -0.762, 0.426)$	1	5.090555573	6

Pre-print `arXiv:1612.07340` (2016)
[N., Plaumann] *Symbolic computation in hyperbolic programming*

Conclusions

- An exact algorithm for hyperbolic programming
- We can compute the maximum multiplicity on a hyp. cone $\Lambda_+(f, e)$
- Combined with Renegar derivatives, one can certify lower bounds for HP

Questions/Perspectives

- Extend complexity bounds from SDP to HP ($poly(\cdot)$ when n or d is fixed?)
- Hyperbolicity test? Complexity of determinantal representations?

Merci

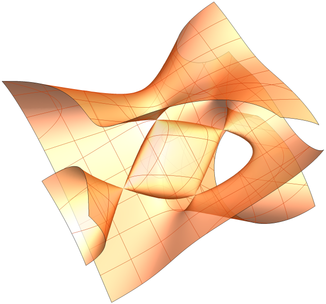


Figure: A non-determinantal quartic hyperbolic surface