MIT 6.972 Algebraic techniques and semidefinite optimization	March 2, 2006
Lecture 7	
Lecturer: Pablo A. Parrilo	*Scribe: ???*

In this lecture we introduce a special class of multivariate polynomials, called *hyperbolic*. These polynomials were originally studied in the context of partial differential equations. As we will see, they have many surprising properties, and are intimately linked with convex optimization problems that have an algebraic structure. A few good references about the use of hyperbolic polynomials in optimization are [Gül97, BGLS01, Ren].

1 Hyperbolic polynomials

Consider a homogeneous multivariate polynomial $p \in \mathbb{R}[x_1, \ldots, x_n]$ of degree d. Here *homogeneous of degree d* means that the sum of degrees of each monomial is constant and equal to d, i.e.,

$$p(x) = \sum_{|\alpha|=d} c_\alpha x^\alpha,$$

where $\alpha = (\alpha_1, \ldots, \alpha_n), \alpha_i \in \mathbb{N} \cup \{0\}$, and $|\alpha| = \alpha_1 + \cdots + \alpha_n$. A homogeneous polynomial satisfies $p(tw) = t^d p(w)$ for all real t and vectors $w \in \mathbb{R}^n$. We denote the set of such polynomials by $\mathcal{H}_n(d)$. By identifying a polynomial with its vector of coefficients, we can consider $\mathcal{H}_n(d)$ as a normed vector space of dimension $\binom{n+d-1}{d}$.

Definition 1. *Let e be a fixed vector in $\mathbb{R}^n$. A polynomial $p \in \mathcal{H}_n(d)$ is* hyperbolic with respect to e *if $p(e) > 0$ and, for all vectors $x \in \mathbb{R}^n$, the univariate polynomial $t \mapsto p(x - te)$ has only real roots.*

A natural geometric interpretation is the following: consider the hypersurface in $\mathbb{R}^n$ given by $p(x) = 0$. Then, hyperbolicity is equivalent to the condition that every line in $\mathbb{R}^n$ parallel to e intersects this hypersurface at exactly d points (counting multiplicities), where d is the degree of the polynomial.

Example 2. *The polynomial $x_1 x_2 \cdots x_n$ is hyperbolic with respect to the vector $(1, 1, \ldots, 1)$, since the univariate polynomial $t \mapsto (x_1 - t)(x_2 - t)\cdots(x_n - t)$ has roots $x_1, x_2, \ldots, x_n$.*

Hyperbolic polynomials enjoy a very surprising property, that connects in an unexpected way algebra with convex analysis. Given a hyperbolic polynomial $p(x)$, consider the set defined as:

$$\Lambda_{++} := \{x \in \mathbb{R}^n : p(x - te) = 0 \;\Rightarrow\; t > 0\}.$$

Geometrically, this condition says that if we start at the point $x \in \mathbb{R}^n$, and slide along a line in the direction parallel to e, then we will never encounter the hypersurface $p(x) = 0$, while if we move in the opposite direction, we will cross it exactly n times. Figure 1 illustrates a particular hyperbolicity cone.

It is immediate from homogeneity and the definition above that $\lambda > 0, x \in \Lambda_{++} \Rightarrow \lambda x \in \Lambda_{++}$. Thus, we call Λ_{++} the *hyperbolicity cone* associated to p, and denote its closure by Λ_+. As we will see shortly, it turns out that these cones are actually *convex cones*. We prove this following the arguments in Renegar [Ren]; the original results are due to Gårding [Går59].

Lemma 3. *The hyperbolicity cone Λ_{++} is the connected component of $p(x) > 0$ that includes e.*

Example 4. *The hyperbolicity cone Λ_{++} associated with the polynomial $x_1 x_2 \cdots x_n$ discussed in Example 2 is the open positive orthant $\{x \in \mathbb{R}^n \mid x_i > 0\}$.*

The first step is to show that we can replace e with any vector in the hyperbolicity cone.

Lemma 5. *If $p(x)$ is hyperbolic with respect to e, then it is also hyperbolic with respect to every direction $v \in \Lambda_{++}$. Furthermore, the hyperbolicity cones are the same.*

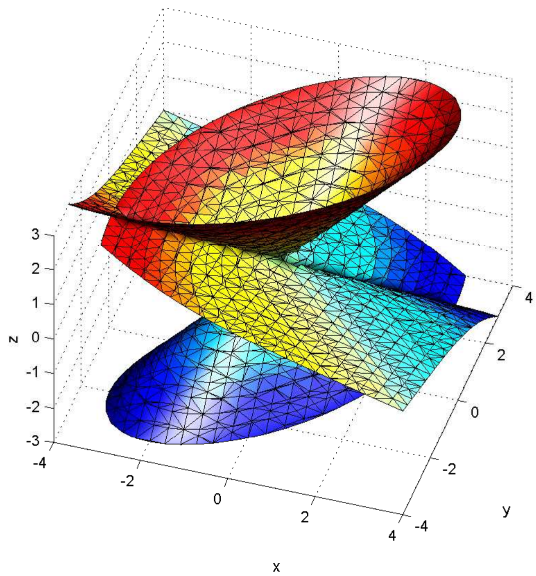

Figure 1: Hyperbolicity cone corresponding to the polynomial $p(x, y, z) = 4xyz + xz^2 + yz^2 + 2z^3 - x^3 - 3zx^2 - y^3 - 3zy^2$. This polynomial is hyperbolic with respect to $(0, 0, 1)$.

Proof. By Lemma 3 we have $p(v) > 0$. We need to show that for every $x \in \mathbb{R}^n$, the polynomial $\beta \mapsto p(\beta v + x)$ has only real roots if $v \in \Lambda_{++}$.

Let $\alpha > 0$ be fixed, and consider the polynomial $\beta \mapsto p(\alpha ie + \beta v + \gamma x)$, where i is the imaginary unit. We claim that if $\gamma \geq 0$, this polynomial has only roots in the lower half-plane. Let's look at the $\gamma = 0$ case first. It is clear that $\beta \mapsto p(\alpha ie + \beta v)$ cannot have a root at $\beta = 0$, since $p(\alpha ie) = (\alpha i)^d p(e) \neq 0$. If $\beta \neq 0$, we can write

$$p(\alpha ie + \beta v) = 0 \quad \Leftrightarrow \quad p(\alpha\beta^{-1}ie + v) = 0 \quad \Rightarrow \quad \alpha\beta^{-1}i < 0 \quad \Rightarrow \quad \beta \in i\mathbb{R}_-,$$

and thus the roots of this polynomial are on the strict negative imaginary axis (we have used $v \in \Lambda_{++}$ in the second implication). If by increasing γ there is ever a root in the upper half-plane, then there must exist a $\gamma_\star$ for which $\beta \mapsto p(\alpha ie + \beta v + \gamma_\star x)$ has a real root $\beta_\star$, and thus $p(\alpha ie + \beta_\star v + \gamma_\star x) = 0$. However, this contradicts hyperbolicity, since $\beta_\star v + \gamma_\star x \in \mathbb{R}^n$. Thus, for all $\gamma \geq 0$, the roots of $\beta \mapsto p(\alpha ie + \beta v + \gamma x)$ are in the lower half-plane.

The conclusion above was true for any $\alpha > 0$. Letting $\alpha \to 0$, by continuity of the roots we have that the polynomial $\beta \mapsto p(\beta v + \gamma x)$ must also have its roots in the lower closed half-plane. However, since it is a polynomial with real coefficients (and therefore its roots always appear in complex-conjugate pairs), then all the roots must actually be real. Taking now $\gamma = 1$, we have that $\beta \mapsto p(\beta v + x)$ has real roots for all x, or equivalently, p is hyperbolic in the direction v. □

The following result shows that this set is actually convex:

Theorem 6 ([Går59]). *The hyperbolicity cone Λ_{++} is convex.*

Proof. We want to show that $u, v \in \Lambda_{++}$, $\beta, \gamma > 0$ implies that $\beta u + \gamma v \in \Lambda_{++}$. The previous result implies that we can always assume $v = e$. But then the roots of $t \mapsto p(\beta u + \gamma e - te)$ are just a nonnegative affine scaling of the roots of $t \mapsto p(u - te)$, since

$$p(u - t_\star e) = 0 \qquad \Leftrightarrow \qquad p(\beta u + \gamma e - (\beta t_\star + \gamma)e) = 0,$$

and $u \in \Lambda_{++}$ implies that $t_\star > 0$, hence $\beta t_\star + \gamma > 0$, and as a consequence, $\beta u + \gamma e \in \Lambda_{++}$. □

Hyperbolic polynomials are of interest in convex optimization, because they unify in a quite appealing way many facts about the most important tractable classes: linear, second order, and semidefinite programming.

Example 7 (SOCP)**.** *Let $p(x) = x_{n+1}^2 - \sum_{k=1}^n x_k^2$. This is a homogeneous quadratic polynomial, hyperbolic in the direction $e = (0, \dots, 0, 1)$, since*

$$p(x - te) = (x_{n+1} - t)^2 - \sum_{k=1}^n x_k^2 = t^2 - 2tx_{n+1} + \left(x_{n+1}^2 - \sum_{k=1}^n x_k^2\right),$$

and the discriminant of this quadratic equation is equal to

$$4x_{n+1}^2 - 4\left(x_{n+1}^2 - \sum_{k=1}^n x_k^2\right) = 4\sum_{k=1}^n x_k^2,$$

which is always nonnegative, so the polynomial $t \mapsto p(x - te)$ has only real roots. The corresponding hyperbolicity cone is the Lorentz or second order cone given by

$$\Lambda_+ = \left\{x \in \mathbb{R}^{n+1} \mid x_{n+1} \geq 0, \quad \sum_{k=1}^n x_k^2 \leq x_{n+1}^2\right\}.$$

Example 8 (SDP)**.** *Consider the homogeneous polynomial*

$$p(x) = \det(x_1 A_1 + \cdots + x_n A_n),$$

where $A_i \in \mathcal{S}^d$ are given symmetric matrices, with $A_1 \succ 0$. The polynomial $p(x)$ is homogeneous of degree d. Letting $e = (1, 0, \dots, 0)$, we have

$$p(x - te) = \det\left(\sum_{k=1}^n x_k A_k - tA_1\right) = \det A_1 \cdot \det\left(\sum_{k=1}^n x_k A_1^{-\frac{1}{2}} A_k A_1^{-\frac{1}{2}} - tI\right),$$

and as a consequence the roots of $p(x - te)$ are always real since they are the eigenvalues of a symmetric matrix. Thus, $p(x)$ is hyperbolic with respect to e. The corresponding hyperbolicity cone is

$$\Lambda_{++} = \{x \in \mathbb{R}_n \mid x_1 A_1 + \cdots + x_n A_n \succ 0\}.$$

Thus, by Lemma 5, $p(x)$ is hyperbolic with respect to every $x \in \Lambda_{++}$.

Based on the results discussed earlier regarding the number of real roots of a univariate polynomial, we have the following lemma.

Lemma 9. *The polynomial $p(x)$ is hyperbolic with respect to e if and only if the Hermite matrix $H_1(p) \in \mathcal{S}^n[x]$ is positive semidefinite for all $x \in \mathbb{R}^n$.*

Lemma 10. *The hyperbolicity cone Λ_+ is basic closed semialgebraic, i.e., it can be described by unquantified polynomial inequalities.*

The two following results are of importance in optimization and the formulation of interior-point methods.

Theorem 11. *A hyperbolic cone Λ_+ is facially exposed.*

Theorem 12 ([Gül97])**.** *The function $-\log p(x)$ is a logarithmically homogeneous self-concordant barrier[1] for the hyperbolicity cone Λ_{++}, with barrier parameter equal to d.*

[1] A function $f : \mathbb{R} \to \mathbb{R}$ is *self-concordant* if it satisfies $f''(x) \geq |\frac{1}{2} f'''(x)|^{\frac{2}{3}}$. A function $f : \mathbb{R}^n \to \mathbb{R}$ is self-concordant if the univariate function obtained when restricting to any line is self-concordant. Self-concordance implies strict convexity, and is a crucial property in the analysis of the polynomial-time global convergence of Newton's method; see [NN94] or [BV04, Section 9.6] for more details.

One of the main open issues regarding hyperbolic cones is about their generality. As Example 8 shows, the cone associated with a semidefinite program is a hyperbolic cone. An open question (known as the generalized Lax conjecture) is whether the converse holds, more specifically, whether every hyperbolic cone is a "slice" of the semidefinite cone, i.e., it can be represented as the intersection of an affine subspace and $\mathcal{S}^n_+$. As we will see in the next lecture, a special case of the conjecture has been settled recently.

2 SDP representability

Recall that in the previous lecture, we encountered a class of convex sets in $\mathbb{R}^2$ that lacked certain desirable properties (namely, being basic semialgebraic, and facially exposed). As we will see, hyperbolic polynomials will play a fundamental role in the characterization of the properties a set in $\mathbb{R}^2$ must satisfy for it to be the feasible set of a semidefinite program.

References

[BGLS01] H. H. Bauschke, O. Güler, A. S. Lewis, and H. S. Sendov. Hyperbolic polynomials and convex analysis. *Canad. J. Math.*, 53(3):470–488, 2001.

[BV04] S. Boyd and L. Vandenberghe. *Convex Optimization.* Cambridge University Press, 2004.

[Går59] L. Gårding. An inequality for hyperbolic polynomials. *J. Math. Mech.*, 8:957–965, 1959.

[Gül97] O. Güler. Hyperbolic polynomials and interior point methods for convex programming. *Math. Oper. Res.*, 22(2):350–377, 1997.

[NN94] Y. E. Nesterov and A. Nemirovski. *Interior point polynomial methods in convex programming*, volume 13 of *Studies in Applied Mathematics.* SIAM, Philadelphia, PA, 1994.

[Ren] J. Renegar. Hyperbolic programs, and their derivative relaxations. Preprint, 2005.