1. Lecture 1

A polynomial $p \in \mathbb{R}[x_1, \dots, x_n]$ is nonnegative if $p(x) \geq 0$ for all $x \in \mathbb{R}^n$. If p is a sum of squares of polynomials $p = \sum q_i^2$ then p is obviously nonnegative. Moreover this representation gives us a *certificate* for nonnegativity of p.

The main question that we will study is the relationship between nonnegative (psd) polynomials and sums od squares (sos) polynomials. We will look at both global nonnegativity and polynomials nonnegative on some set.

1.1. **Optimizational Motivation.**

Testing whether a polynomial is nonnegative is a very hard computational problem. If we can test for nonnegativity then we can quickly minimize polynomials: the minimum of p is the largest γ such that $p - \gamma$ is nonnegative.

The question of relationship between psd and sos polynomials has been receiving renewed attention because it was realized that the question of testing whether a polynomial is sos is actually computationally tractable and it can be solved via semi-definite programming.(Parrilo-Sturmfels,Powers-W'ormann)

This is not the whole story, as semi-definite programs (SDPs) coming from sos testing grow rather quickly in size as the number of variables or degree increase. However these semi-definite programs have special structure. Also we can exploit special structure in the polynomials that we test for being sos, such as sparseness and symmetry to reduce the size of the SDPs. There are currently several packages (SosTools, Yalmip) that allow efficient computation of sum of squares representation of polynomials.

We can relax testing for nonnegativity to test whether a polynomial is a sum of squares. (This is usually called *sos relaxation*). For optimization, instead of computing the true minimum γ we can compute the maximal γ^* such that $p - \gamma^*$ is sos.

One advantage of this approach is that if we can show that p is sos with an explicit representation then we have a certificate for nonnegativity of p.

1.2. **Lyapunov Functions.**

(Parrilo) We can extend the idea of substituting sums of squares conditions for nonnegativity conditions to other settings. Suppose that we are given a system of ordinary differential equations with $x_1, \dots, x_n$ being functions of time t:

$$\frac{dx_1}{dt} = f_1(x_1, \dots, x_n)$$
$$\vdots$$
$$\frac{dx_n}{dt} = f_n(x_1, \dots, x_n).$$

This is usually written shorthand as $\dot{x}(t) = f(x)$. Suppose that the system has a steady state at the origin (we can always translate a steady state, if it exists, to the origin). Often we want to know whether the steady state is globally stable or we want to compute its basin of attraction or at least some approximation of it. For now we will focus on global stability.

By a theorem of Lyapunov the origin is globally stable if there exists a function $V(x) : \mathbb{R}^n \to \mathbb{R}$ such that

$$V(0) = 0 \text{ and } V(x) > 0 \text{ for all } x \neq 0, \tag{1.1}$$

$$\frac{dV}{dt} < 0 \text{ for all } x \in \mathbb{R}^n, \tag{1.2}$$

and additionally V is radially unbounded. The first condition (1.1) states that V has a global minimum at the origin. We can rewrite the second condition as

$$\frac{dV}{dt} = \sum_{i=1}^{n} \frac{dV}{dx_i}\frac{dx_i}{dt} = \sum_{i=1}^{n} \frac{dV}{dx_i} f_i = \langle \nabla V, \dot{x} \rangle < 0,$$

where ∇V is the gradient of V and $\langle , \rangle$ is the standard inner product. In other words condition (1.2) states that if $x(t)$ is a solution of our system then by following the trajectory of $x(t)$ the value of V strictly decreases as $\langle \nabla V, \dot{x} \rangle < 0$.
Since V has a unique global minimum at the origin it follows that we must settle there by following any solution and therefore the origin is globally stable. Moreover if the functions f_i are polynomials then we can find a polynomial Lyapunov function V.

Exercise 1.1. *Given the system:*

$$\dot{x}_1 = -x_1$$
$$\dot{x}_2 = -2x_2 + x_1,$$

check that the origin is a steady state and the function $V(x_1, x_2) = x_1^2 + x_2^2$ *is a Lyapunov function for this system, certifying global stability of the origin.*

There are no systematic ways of computing Lyapunov functions. However, we can relax both conditions (1.1) and (1.2) to sums of squares conditions by requiring that V and $-\frac{dV}{dt}$ are both sos. This allows to systematically search for Lyapunov functions of given degree by formulating the search as a semi-definite program.

1.3. **Hilbert's Theorem.** A nonnegative polynomial can be homogenized and it will remain nonnegative. The same holds for sums of squares. Therefore we will restrict ourselves to the case of homogeneous polynomials (forms). We need some notation: Let $H_{n,d}$ be the vector space of forms in n variable of degree d. It is not hard to show that the dimension of $H_{n,d} = \binom{n+d-1}{d}$.

In order to be nonnegative a polynomial must have even degree and therefore our forms will have even degree $2d$. Inside $H_{n,2d}$ sit two closed convex cones, the cone of nonnegative polynomials:

$$P_{n,2d} = \{p \in H_{n,2d} \mid p(x) \geq 0 \text{ for all } x \in \mathbb{R}^n\},$$

and the cone of sums of squares:

$$\Sigma_{n,2d} = \left\{p \in H_{n,2d} \mid p(x) = \sum q_i^2 \text{ for some } q_i \in H_{n,d}\right\}.$$

The first fundamental result about the relationship between $P_{n,2d}$ and $\Sigma_{n,2d}$ was shown by Hilbert in 1888:

Theorem 1.2. $P_{n,2d} = \Sigma_{n,2d}$ *in the following three cases:* $n = 2$ *(univariate non-homogeneous case),* $2d = 2$ *(quadratic forms), and* $n = 3$, $2d = 4$ *(ternary quartics). In all other cases there exists nonnegative forms that are not sums of squares.*

There are several parts to proof. First we must show the equality in all of the above cases and then we need to demonstrate the there exist psd forms that are not sos in all other cases.

Proof of the case $n = 2$. This is the only time when we will prefer to work with polynomials rather than forms. We need to show that a nonnegative polynomial in 1 variable is always a sum of squares. In fact we will show that it is always a sum of just 2 squares.
Let $p(x) = a_{2d}x^{2d} + \ldots + a_0$ be a nonnegative polynomial; we must have $a_{2d} > 0$. By Fundamental theorem of algebra p factors completely over the complex numbers:

$$p = a_{2d}(x - r_1) \ldots (x - r_{2d}),$$

where r_i are the roots of p. Since p is psd the real roots must have even degree. Complex roots come in conjugate pairs r_i and $\bar{r}_i$. Multiplying out $(x-r_i)(x-\bar{r}_i)$ we get a monic (leading coefficient is equal to 1) quadratic with no real roots. Therefore we can write p as

$$p = a_{2d}(x-r_1)^{2\alpha_1}\dots(x-r_m)^{2\alpha_m}\cdot q_1\dots q_k,$$

where $r_1,\dots,r_k$ are real roots and q_i are monic polynomials with no real roots. To show that p is a sum of 2 squares we only need to analyze the product of q_i.

Let $q_i = x^2+bx+c$. Since q_i has no real roots we know that the discriminant is negative $b^2-4c<0$. We can complete the square and write q_i as a sum of two squares: $q_i = (x-b/2)^2+(c-b^2/4)$.

To show that entire product of q_i is a sum of two squares we repeatedly use the following identity: $(a^2+b^2)(c^2+d^2) = (ac+bd)^2+(ad-bc)^2$. □

Proof of the case $2d=2$. In this case we are dealing with quadratic forms. Given a quadratic form $p(x)$ we can write it as $p = x^TAx$ where $x=(x_1,\dots,x_n)$ and A is a symmetric matrix. Symmetric matrices have real eigenvalues and p being psd is equivalent to A having all nonnegative eigenvalues.

Symmetric matrices are diagonalizable using orthogonal matrices: we can always write $A = M^TDM$ where M is an orthogonal matrix ($M^T=M^{-1}$) and D is a diagonal matrix with eigenvalues $\lambda_1,\dots,\lambda_n$ on the diagonal.

Let $m_1,\dots,m_n$ be the rows of M. Then multiplying $p=(x^TM^T)D(Mx)$ out we see that

$$p = \lambda_1\langle m_1,x\rangle^2+\ldots+\lambda_n\langle m_n,x\rangle^2,$$

where $\langle,\rangle$ is the standard inner product. Therefore p is a sum of at most n squares and it not hard too see that the least number of squares needed to represent p is equal to the rank of A.

□

The proof for the case $n=3, 2d=4$ (so called ternary quartics) is more complicated. We will use the following lemma whose proof we postpone for now:

Lemma 1.3. *Suppose that $\Sigma_{n,2d}\subsetneq P_{n,2d}$. Then there exists a form p such that the zero set of p cannot be realized by an sos form in $P_{n,2d}$.*

This is also a good time to stop for discussion of what zeroes do to psd forms. Suppose that a form p has a (projective) zero at a point $v\in\mathbb{R}^n$. In this case v is a global minimum for p. Therefore the gradient of p at v must be zero. This imposes n linear conditions on p:

$$\frac{\partial p}{\partial x_1}(v)=0,\dots,\frac{\partial p}{\partial x_n}(v)=0. \tag{1.3}$$

You might wonder about the condition that $p(v)=0$ but by *Euler's Relation* since p is homogeneous we have

$$\langle\nabla p(v),v\rangle = 2d\cdot p(v).$$

Therefore vanishing of the gradient at v implies that $p(v)=0$.

Exercise 1.4. *Prove Euler's Relation: for $p\in H_{n,d}$ and all $x\in\mathbb{R}^n$:*

$$\langle\nabla p,x\rangle = d\cdot p(x).$$

Now suppose that p has a zero at a standard basis vector e_i. In this case it is easy to check that the gradient vanishing conditions (1.3) transform into the following conditions:

$$p \text{ has no terms of the form } x_i^{d-1}x_j \text{ for } j=1,\dots,n. \tag{1.4}$$

Proof of the case of ternary quartics. Suppose that $\Sigma_{3,4} \subsetneq P_{3,4}$ and let p be a psd form whose zero set $Z(p)$ can be realized any sum of squares. It is easy to see that p must have at least 3 projective zeroes v_1, v_2 and v_3.

If v_i all lie in the same plane in $\mathbb{R}^3$ then we can rotate the plane to (x_1, x_2) plane. If we restrict p to the first 2 variables then the restriction $\bar{p}$ is a fourth degree form in2 variables with 3 distinct projective zeroes. If we dehomogenize appropriately then we obtain a fourth degree univariate polynomial with 3 distinct zeroes. Since zeroes of univariate polynomials must have even degree, it follows $\bar{p}$ identically zero. It follows that x_3 divides p and since p is nonnegative x_3^2 must divide p. Thus $p = x_3^2 q$ where q is a nonnegative quadratic form and therefore p is a sum of squares.

Therefore we may assume that the three zeroes v_i are linearly independent. Applying a non-singular linear transformation that maps v_i to the standard basis vectors e_i we may assume that p has zeroes at $[1:0:0]$, $[0:1:0]$ and $[0:0:1]$. In this case p by (1.4) p has no terms involving x_i^3 or x_i^4. We can write

$$p = a_1 x_1^2 x_2^2 + a_2 x_1^2 x_3^2 + a_3 x2^2 x_3^2 + 2b_1 x_1^2 x_2 x_3 + 2b_2 x_1 x_2^2 x_3 + 2b_3 x_1 x_2 x_3^2.$$

Let A be the following symmetric matrix:

$$A = \begin{pmatrix} a_1 & b_1 & b_2 \\ b_1 & a_2 & b_3 \\ b_2 & b_3 & a_3 \end{pmatrix},$$

and let $s = (x_1x_2, x_1x_3, x_2x_3)$. It is easy to check that $p = s^T A s$. To show that p is sos it will suffice to show that A must be positive semi-definite. Suppose that A is not psd. Then there exists a point $w = (w_1, w_2, w_3)$ such that $w^T A w < 0$. By continuity we can find w with all w_i non-zero. Taking $-w$ if necessary we can also assume that at least two w_i are positive. Then the system

$$x_1x_2 = w_1, \;\; x_1x_3 = w_2 \;\; x_2x_3 = w_3$$

has a solution:

$$x_1 = \sqrt{\frac{w_1w_2}{w_3}}, x_2 = \sqrt{\frac{w_1w_3}{w_2}}, x_3 = \sqrt{\frac{w_2w_3}{w_1}}.$$

It follows that p is not psd and therefore we obtain a contradiction.

□

Coming Up Next: The proof of existence nonnegative polynomials that are not sums of squares in all other cases. Hilbert's original non-constructive proof and explicit examples. Also proof of Lemma 1.3.

2. Lecture 2

We begin with the proof of Lemma 1.3. Although it seems like an algebraic statement we prove using techniques of convexity. First we need the notion of the dual cone: Let K be a convex cone in a real vector space V. The dual cone of K is formally defined as the set of all linear functionals in the dual space that are nonnegative on K:

$$K^* = \{l \in V^* \;\; | \;\; l(x) \geq 0 \text{ for all } x \in K\}.$$

Informally, K^* is the set of all linear inequalities that cut out K.

If K is a closed cone then we the double-dual theorem

Theorem 2.1. *Let K be a closed cone in a finite-dimensional real vector space V. Then we have*

$$(K^*)^* = K$$

.

For a point $v \in \mathbb{S}^{n-1}$ let $l_v \in H^*_{n,2d}$ be the the linear function given by evaluation at v:

$$l_v(p) = p(v) \text{ for } p \in H_{n,2d}.$$

It is not hard to show that the dual cone of $P_{n,2d}$ is spanned by the functionals l_v, it essentially states that a form is nonnegative if and only if is is nonnegative on the unit sphere:

Proposition 2.2. *$P^*_{n,2d}$ is the conical hull of functionals l_v for all $v \in \mathbb{S}^{n-1}$.*

An extreme ray r of a convex cone is called *exposed* if $r = K \cap W$ for some hyperplane $W \subset V$. The following is a consequence of Straszewicz's Theorem:

Proposition 2.3. *Let K be a proper convex cone. Then every extreme ray of K can be obtained as a limit of exposed extreme rays.*

Proof of Lemma 1.3. Suppose that $P_{n,2d}$ is strictly larger than $\Sigma_{n,2d}$. Then there exists an extreme ray of $P_{n,2d}$ that is not in $\Sigma_{n,2d}$. By Lemma 2.3 and the fact that the cones $P_{n,2d}$ and $\Sigma_{n,2d}$ are proper closed it follows that there exists an exposed extreme ray r of $P_{n,2d}$ that is not in $\Sigma_{n,2d}$. By the description of the dual cone from Lemma 2.2 it follows that r is cut from $P_{n,2d}$ by requiring the form spanning r to vanish on a set of points in the unit sphere $\mathbb{S}^{n-1}$. Therefore r spans the only nonnegative forms with this zero and since r is a sums of squares, it follows that its zero set cannot be realized by an sos form. □

2.1. **Explicit Examples.** Only seventy years later after Hilbert's original proof the first explicit example was constructed by Motzkin. The following is known as the Motzkin form:

$$M(x, y, z) = x^4y^2 + x^2y^4 + z^6 - 3x^2y^2z^2.$$

M is psd by arithmetic mean/geometric mean inequality applied to the first 3 terms. The fact that M is not a sum of squares is best shown by the Newton Polytope argument, which is contained in exercise 4 of the handout.

Here are two more explicit examples due to Choi and Lam that employ the same ideas:

$$x^4y^2 + y^4z^2 + z^4x^2 - 3x^2y^2z^2$$

$$x^2y^2 + y^2z^2 + x^2z^2 + w^4 - 4xyzw.$$

Exercise 2.4. *Show that both of these forms are nonnegative but not sums of squares.*

2.2. **Uniform Denominators.** For $p \in H_{n,2d}$ define

$$\epsilon(p) = \frac{\min_{x \in \mathbb{S}^{n-1}} p(x)}{\min_{x \in \mathbb{S}^{n-1}} p(x)}.$$

Define $r^2 = x_1^2 + \ldots + x_n^2$.

Theorem 2.5. *(Reznick) For a strictly positive $p \in H_{n,2d}$ the polynomial $r^{2m}p$ is a sum of $2m+2d$-th powers for all*

$$m \geq \frac{nd(2d-1)}{(2\log 2)\epsilon(p)} - \frac{n+2d}{2}.$$

2.3. **Pólya's Theorem.** Define $\lambda(p) = \min_{\Delta_n} p(x)$. Let $p = \sum a_\alpha x^\alpha$. Define $c_\alpha = \binom{d}{\alpha}$, $L(p) = \max_\alpha \frac{a_\alpha}{c_\alpha}$.

Theorem 2.6. *Let $p \in H_{n,d}$ be a polynomial strictly positive on $\mathbb{R}^n_+$. If $N > \frac{d(d-1)}{2}\frac{L}{\lambda} - d$ then $(x_1 + \ldots + x_n)^N p$ is a polynomial with strictly positive coefficients.*