REAL ALGEBRAIC GEOMETRY

KENNETH R. DRIESSEL

Sums of Squares of Real Polynomials (Continued)

References: In this section my ideas are based on the following article: Powers and Woermann (1998) "An algorithm for sums of squares of real polynomials". Powers and Woermann attribute the connection between sums of squares of real polynomials and positive semi-definite matrices to Choi, Lam and Reznick (1995).

Notation: Throughout this section, we use the following notation: The symbol $\mathbb{R}$ denotes the field of real numbers and $R := \mathbb{R}[X_1, \ldots, X_k]$ denotes the algebra of polynomials with coefficients in $\mathbb{R}$ and variables $X_1, \ldots, X_k$. We also use the standard multi-index notation: $X^\alpha := X_1^{\alpha_1} \ldots X_k^{\alpha_k}$.

Sums of squares and positive semi-definite matrices. Recall the following result (which we discussed earlier).

Proposition 1. Sums of squares and positive semi-definite matrices. *Let $f \in R$ be a real polynomial and let $b_1, \ldots, b_m \in R$ be independent real polynomials. Then f is a sum of squares of elements in the span of $b_1, \ldots, b_m$ iff there is a positive semi-definte matrix M such that $f = \mathbf{b}^T M \mathbf{b} = \Sigma m_{ij} b_i b_j$ where $\mathbf{b}$ is the column vector with components b_i.*

Here we consider more applications of this result.

We shall regard the entries of M as unknown. We shall want M to be positive semi-definite and we shall want the entries of M to satisfy certain linear equations. In particular, we shall consider the following problem:

Problem: Given a multivariate polynomial can it be written as a nontrivial sum of squares?

We can use the following result to reduce the number of monomials that we need to consider. (This result appears in Powers and Woermann(1998).)

Proposition 2. The monomial reduction lemma. *Let Λ be a set of multi-indices with length k and let b_i, for $i = 1, \ldots, m$, be a list of the monomials in the set $\{X^\lambda : \lambda \in \Lambda\}$. Suppose that the polynomial f can be written as a nontrivial sum of squares using polynomials h_j in the span of the b_i. Let α and β be multi-indices which satisfy the following conditions: β is an element of Λ, $\alpha = 2\beta$ and α cannot be written in any other way as a sum of elements in Λ. If the coefficient of α in f is 0 then the monomial X^β is not needed in the nontrivial sum of squares representation of f.*

Date: June 4, 2007.

Proof. Let M denote the positive semi-definite matrix corresponding to the nontrivial sum of squares representation of f. Note that $b_i = X^\beta$ for some i. Then $m_{ii} = 0$. Since $0 \leq M$ the ith row and column of M must be zero. $\square$

Recall the following two closely related results. We shall use the triangle inquality in the proof of the degree limitation lemma below.

Proposition 3. The Cauchy-Schwarz inequality. *Let V be an inner product space. Then, for all vectors u and v in V,*

$$|\langle u, v\rangle| \leq \|u\|\|v\|$$

with equality iff u and v are linearly dependent.

Proposition 4. The triangle inequality. *Let V be an inner product space. Then, for all vectors u and v in V,*

$$\|u + v\| \leq \|u\| + \|v\|$$

with equality iff u and v are linearly dependent.

Proof. This result follows easily from the Cauchy-Schwarz inequality. $\square$

Proposition 5. The degree limitation lemma. *Let f be a polynomial with total degree $2d$. If f can be written as a nontrivial sum of squares then it can be written as a nontrivial sum of squares of polynomials with degrees at most d.*

Proof. Suppose that we have f written as a sum of squares of real polynomials. Let X^α where $\alpha := (\alpha_1, \ldots, \alpha_k)$ be a monomial with maximal degree $|\alpha| := \alpha_1 + \cdots + \alpha_k$. Note that there exists β such that $\alpha = 2\beta$ (since f is weakly positive). Suppose $2d < |\alpha|$. Let Λ be the set of multi-indices with length k and degree at most $|\alpha|$. By the triangle inequality in $\mathbb{R}^k$, $\alpha = 2\beta$ cannot be written in any other way as a sum of elements of Λ. Now apply the monomial reduction lemma. In particular, since the coefficient of X^α is zero in f the monomial X^α is not needed in the nontrivial sum of squares representation of f. $\square$

There is an analogous result for homogeneous polynomials which appears below. Recall the definition:

Definition: Let f be a polynomial in R. Then f is **homogeneous with degree** e if, for all $\lambda \in \mathbb{R}$, and all $x \in \mathbb{R}^k$, $\lambda \in \mathbb{R}$, $f(\lambda x) = \lambda^e f(x)$.

For any polynomial g and positive integer e, we shall use Homog(g, e) to denote the part of g which is homogeneous with degree e. For example, if $g(X, Y) := aX^2 + bXY + CY^2 + dX + eY$ where a, b, c and d are real numbers, then

- Homog$(g, 2)(X, Y) = aX^2 + bXY + cY^2$ and
- Homog$(g, 1)(X, Y) = dX + eY$.

Proposition 6. *Let g and h be polynomials in R.Then $g = h$ iff for every positive integer e, Homog(g, e) = Homog(h, e).*

Proof. $\Leftarrow$: This implication is trivial since $g = \sum_e \text{Homog}(g, e)$.

$\Rightarrow$: Suppose $0 = g = \sum_e \text{Homog}(g, e) = 0$. Then, for all $\lambda \in \mathbb{R}$, $0 = \sum_e \text{Homog}(g, e)(\lambda x) = \sum_e \lambda^e \text{Homog}(g, e)(x)$. We have a polynomial in λ. It follows (since the field $\mathbb{R}$ has characteristic 0) that its coefficients are zero: $\text{Homog}(g, e) = 0$ for all e. □

Proposition 7. Homogeneity of products. *Let f_1 and f_2 be polynomials in R. If f_1 is homogeneous with degree e_1 and f_2 is homogeneous with degree e_2 then $f_1 f_2$ is homogeneous with degree $e_1 + e_2$.*

Proof. Use the definition of homogeneous. □

Proposition 8. Homogeneity of squares. *Let f be a homogeneous polynomial with degree d. Then f^2 is homogeneous with degree $2d$.*

Proposition 9. *Let f be a homogeneous element of R with degree $2d$. If f can be written as a sum of squares of polynomials then f can be written as a sum of squares of polynomials which are homogeneous with degree d.*

Proof. Suppose $f = h_1^2 + \cdots + h_m^2$. Then

$$f = \text{Homog}(f, 2d) = \text{Homog}(h_1^2, 2d) + \cdots + \text{Homog}(h_m^2, 2d) \tag{1}$$

$$= \text{Homog}(h_1, d)^2 + \cdots + \text{Homog}(h_m, d)^2. \tag{2}$$

□

We shall use the "geometric-arithmetic" inequality below. We recall it now.

Definition: Let $a_1, \ldots, a_m$ be a list of real numbers. Then the quantity $(1/m)(a_1 + a_2 + \cdots + a_m)$ is called the **arithmetic mean** of the a_i. If the a_i are positive then the quantity $\sqrt[m]{a_1 a_2 \ldots a_m}$ is called the **geometric mean** of the a_i.

Proposition 10. The inequality of the arithmetic and geometric means. *For every list $a_1, \ldots, a_m$ of positive real numbers,*

$$\sqrt[m]{a_1 a_2 \ldots a_m} \leq (1/m)(a_1 + a_2 + \cdots + a_m).$$

Equality holds iff $a_1 = a_2 = \cdots = a_m$.

Proof. See, for example, Beckenbach and Bellman (1961) or (1971). □

Example: The Motzkin polynomial. It is clear that every sum of squares of real polynomials is always a weakly positive function. But there are polynomials which are always weakly positive and are not sums of squares. Motzkin presented the first explicit example of such a polynomial in 1967. He proved the following result. The polynomial defined in this result is called the **Motzkin polynomial**.

Proposition 11. *Let $s(X, Y, Z)$ denote the following polynomial.*

$$X^4 Y^2 + X^2 Y^4 + Z^6 - 3X^2 Y^2 Z^2.$$

Then for every triple x, y, z of real numbers, $0 \le s(x, y, z)$. But s cannot be written as a sum of squares.

Proof. Claim: The polynomial s is always weakly positive.

Note that $0 \le s(x, y, z)$ is equivalent to

$$x^2y^2z^2 \le (1/3)(x^4y^2 + x^2y^4 + z^6)$$

which follows from the inequality of the arithmetic and geometric means.

Claim: The polynomial is not the sum of squares.

Note that the polynomial s is homogeneous with degree 6. If it is the sum of squares then it must be the sum of squares of homogeneous polynomials with degree 3.

Here is the list of multi-indices for monomials with degree 3:

$$300, 210, 201, 120, 111, 102, 030, 021, 012, 003.$$

For $i := 1, \ldots, 10$ let α_i denote the ith multi-index in this list and let b_i denote the corresponsing monomial. We have the following table of sums of these multi-indices. The entries with asterisks are filled by symmetry.

	+	300	210	201	120	111	102	030	021	012	003
b_1	300	600	510	501	420	411	402	330	321	312	303
b_2	210	*	420	411	330	321	312	240	231	222	213
b_3	201	*	*	402	321	312	303	231	222	213	204
b_4	120	*	*	*	240	231	222	150	141	132	123
b_5	111	*	*	*	*	222	213	141	132	123	114
b_6	102	*	*	*	*	*	204	132	123	114	105
b_7	030	*	*	*	*	*	*	060	051	042	033
b_8	021	*	*	*	*	*	*	*	042	033	024
b_9	012	*	*	*	*	*	*	*	*	024	015
b_{10}	003	*	*	*	*	*	*	*	*	*	006

Suppose that there is a sum of squares representation of the polynomial s in terms of polynomials in the span of the monomials with multi-index exponents given in the above list. Let M be the corresponding 10-by-10 positive semi-definite matrix.

(We now play a variant of the Sudoku puzzle.) Note that $m_{11} = 0$ since the exponent 006 appears uniquely; hence the first row (and column) of M must be zero. Note that $m_{77} = 0$ since the exponent 060 appears uniquely; hence the seventh row of M must be zero. Note that $m_{33} = 0$ since the exponent 402 appears in only places 33 and 16 in the upper triangle and we already have $m_{16} = 0$. Hence the third row of M must be zero. The same sort of argument shows that $m_{66} = 0, m_{88} = 0$ and $m_{99} = 0$. It follows that the third, sixth, eighth and ninth rows must be zero. Note that we must then have $m_{55} = -3$ since the 222 exponent appears only in places 55, 45, 38, and 29 and we already have $m_{45} = m_{38} = m_{29} = 0$. This is a contradiction since $0 \le M$ implies that all the diagonal entries of M must be weakly positive. $\square$