

REAL ALGEBRAIC GEOMETRY

KENNETH R. DRIESSEL

SUMS OF SQUARES OF REAL POLYNOMIALS

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, $R := \mathbb{R}[X_1, \dots, X_k]$ denotes an algebra of polynomials with coefficients in \mathbb{R} , I denotes an ideal in R , $A := R/I$ denotes the quotient algebra determined by R and I , $n := \dim A$ denotes the dimension of A , which we assume to be finite, and

$$\text{Variety}(I, \mathbb{R}^k) := \{x \in \mathbb{R}^k : (\forall p \in I) p(x) = 0\}$$

denotes the set (variety) of real solutions of the polynomials in I .

Introduction. One of the main algorithmic tasks in real algebraic geometry is the counting of the number of real solutions in the variety $\text{Variety}(I, \mathbb{R}^k)$. We can use Hermite’s quadratic form to accomplish this task. However, the corresponding algorithm has complexity $\mathcal{O}(n^4)$. (See Basu, Pollack and Roy(2003) .) In many applications the dimension n is large. So it we want to have methods for reducing the size of n . The set of complex solutions influences this size. But we normally do not care about the complex solutions. So we want to eliminate some of the complex solutions without changing the set of real solutions. We shall see that we can do this in a systematic way. In particular, we shall see that, by considering sums of squares of real polynomials, we can enlarge the ideal I without changing the set of real solutions. Enlarging the ideal will reduce the size of n . (In fact we would really like to enlarge I to the real radical of I . But I do not know any nice algorithms to compute the real radical.)

This section consists of several subsections. In the subsection with title “Using sums of square to enlarge ideals” we shall see how to use sums of squares to enlarge an ideal without changing the set of real solutions. In the subsection with title “Positive semi-definite matrices” we review some of the properties of such matrices. Later we shall use some of these properties. In the subsection with title “Sums of squares and positive semi-definite

Date: May 15, 2007.

matrices” we shall see that there is a bijective correspondence between sums of squares of real polynomials and positive semi-definite matrices. In the subsection with title “Examples” we shall consider a number of applications of these methods.

Using sums of squares to enlarge ideals.

Proposition 1. Enlarging ideals. *Let a_1, \dots, a_m be elements of the polynomial algebra R . If the polynomial $a_1^2 + \dots + a_m^2$ is in the ideal I , then*

$$\text{Variety}(I, \mathbb{R}^k) = \text{Variety}(J, \mathbb{R}^k).$$

where $J := \text{Ideal}(I \cup \{a_1, \dots, a_m\})$ is the ideal in the algebra R generated by $I \cup \{a_1, \dots, a_m\}$.

This proposition says that if a sum of squares of polynomials a_i is in the ideal I , then adding the a_i to the ideal does not change the set of real solutions.

Proof. \supseteq : Note that, for any subsets P and Q of R we have $P \subseteq Q$ implies $\text{Variety}(P, \mathbb{R}^k) \supseteq \text{Variety}(Q, \mathbb{R}^k)$. (This is an example of the standard inclusion reversing which happens in Galois connections.)

\subseteq : Consider any x in $\text{Variety}(I, \mathbb{R}^k)$. Note that $a_1^2(x) + \dots + a_m^2(x) = 0$. It follows (since the evaluation only involves element of \mathbb{R}) that $a_1(x) = \dots = a_m(x) = 0$. Hence $x \in \text{Variety}(I \cup \{a_1, \dots, a_m\}, \mathbb{R}^k)$. \square

Here is an example. Let $R := \mathbb{R}[X]$ be the algebra of real polynomials in a single variable X . Let I be the ideal generated by $X^2 + 1$. Then the algebra $A := R/I$ has dimension 2; in particular, the polynomials 1 and X provide a basis for it. Note that $X^2 + 1^2$ is in the ideal I . It follows that we can add X and 1 to the ideal without changing the set of real solutions. But the ideal $J := \text{Ideal}(1, X, X^2)$ generated by 1, X, X^2 is the whole polynomial algebra R and hence $R/J = R/R$ has dimension 0.

Here is another example. Again let $R := \mathbb{R}[X]$. Let I be the ideal generated by $X(X^2 + 1)$. Note that the algebra $A := R/I$ has dimension 3; in particular, the polynomials 1, X, X^2 form a basis. Note that $X^4 + X^2 = X^2(X^2 + 1)$ is in I . We can add X^2 and X to the ideal. The enlarged ideal is $J := \text{Ideal}(X, X^2, X(X^2 + 1)) = \text{Ideal}(X)$. Note R/J has dimension 1.

Positive semi-definite matrices. We begin this subsection by recalling the definition of “positive semi-definite”.

Definition: Let V be a real vector space with inner product $\langle \cdot, \cdot \rangle$. Let M be a symmetric matrix (self-adjoint linear map $M : V \rightarrow V$). Then M is **positive semi-definite (weakly positive)** if, for all v in V , $0 \leq \langle v, Mv \rangle$. We write $0 \leq M$ to briefly say that M is positive semi-definite.

There are many alternative characterizations of the set of positive semi-definite matrices. The next proposition gives some of them. (This result appears, for example, in the undergraduate book *Linear Algebra and Its Applications* by G. Strang(1980,2005).)

Proposition 2. Characterizations of positive semi-definite. *Let \mathbb{R}^n have its standard inner product. Let M be an n -by- n real symmetric matrix. Then the following conditions are equivalent:*

- *The matrix M is positive semi-definite (that is, $0 \leq M$).*
- *The eigenvalues of M are weakly positive (that is, $0 \leq \lambda_i$).*
- *The principal sub-matrices of M have weakly positive determinants (that is, $0 \leq$ principal minors).*
- *The leading principal sub-matrices of M have weakly positive determinants (that is, $0 \leq$ leading principal minors).*
- *In Gaussian elimination (without row exchanges) all the diagonal pivots are weakly positive (that is, $0 \leq p_i$).*
- *There is an n -by- n real matrix W such that $M = W^T W$ (that is, the signature of M equals its rank).*
- *There exist n -by- n real matrices L and D such that $M = LDL^T$ where L is lower triangular with ones on the diagonal and $D := \text{diag}(d_1, \dots, d_n)$ is a diagonal matrix with weakly positive diagonal entries (that is, $0 \leq d_i$).*

Note that congruence is the associated group action. In particular, Sylvester’s law of inertia is relevant. The decomposition in the last equivalent characterization is often called the “Cholesky factorization”

Proof. I omit the proof. See, for example, Strang(1980,2005). □

The following result says that if a positive semi-definite matrix has a diagonal entry equal to zero then every entry in that row and column must be zero.

Corollary 1. *Let M be a positive semi-definite matrix. If $M_{ii} = 0$ then, for all j , $m_{ij} = 0$ and $m_{ji} = 0$.*

Proof. Consider any j different than i . Consider the 2-by-2 principal sub-matrix determined by the pair i, j :

$$\begin{pmatrix} 0 & m_{ij} \\ m_{ij} & m_{jj} \end{pmatrix}.$$

(I assumed $j > i$. The treatment of the case $j < i$ is the same.) The determinant of this sub-matrix must be weakly positive: $0 \leq -m_{ij}^2$. It follows that $m_{ij} = 0$. □

So far, my treatment of the set of positive semi-definite matrices has been very algebraic. But I also need to mention some of the geometric properties of this set.

Proposition 3. *The set of positive semi-definite matrices forms a convex cone; that is,*

- *If M_1 and M_2 are positive semi-definite matrices then so is $M_1 + M_2$.*
- *If ρ is a weakly positive real number, $0 \leq \rho$, and M is positive semi-definite then so is ρM .*

Proof. I leave the (easy) proof as an exercise for the reader. \square

Example: Consider a 2-by-2 symmetric matrix:

$$M := \begin{pmatrix} x & z \\ z & y \end{pmatrix}.$$

Note that $0 \leq M$ iff $0 \leq x$ and $0 \leq xy - z^2$. The set $\{(x, y, z) \in \mathbb{R}^3 : 0 \leq x \wedge 0 \leq xy - z^2\}$ is a (“solid”) right circular cone. The ray $\{(x, x, 0) \in \mathbb{R}^3 : 0 \leq x\}$ is the axis of symmetry for this cone.

Sums of squares and positive semi-definite matrices. Powers and Woermann(1998) attribute the following result connecting sums of squares of real polynomials and positive semi-definite matrices to Choi, Lam and Reznick(1995). (I have not looked at this 1995 paper.)

Proposition 4. Sums of squares and positive semi-definite matrices. *Let $f \in R$ be a real polynomial and let $b_1, \dots, b_m \in R$ be independent real polynomials. Then f is a sum of squares of elements in the span of b_1, \dots, b_m iff there is a positive semi-definite matrix M such that $f = \mathbf{b}^T M \mathbf{b} = \sum m_{ij} b_i b_j$ where \mathbf{b} is the column vector with components b_i .*

Proof. \Rightarrow : Let p be any element in the span of \mathbf{b} . Then

$$p = \mathbf{p}^T \mathbf{b} = \mathbf{b}^T \mathbf{p}$$

where \mathbf{p} is the column vector of coefficients of p with respect to \mathbf{b} . Note that

$$p^2 = \mathbf{b}^T \mathbf{p} \mathbf{p}^T \mathbf{b}.$$

Also note that $0 \leq \mathbf{p} \mathbf{p}^T$. (A rank one matrix $\mathbf{p} \mathbf{q}^T$ is often called the “outer product” of the vectors \mathbf{p} and \mathbf{q} .)

Now consider a list of polynomials p_1, \dots, p_m in the span of \mathbf{b} . We have

$$p_1^2 + \dots + p_m^2 = \mathbf{b}^T (\mathbf{p}_1 \mathbf{p}_1^T + \dots + \mathbf{p}_m \mathbf{p}_m^T) \mathbf{b}.$$

\Leftarrow : We use the Cholesky decomposition of M . Here are some of the details. We get a lower triangular matrix L with ones on the diagonal and a diagonal matrix $D := \text{diag}(d_1^2, \dots, d_m^2)$ such that $M = LDL^T$. Let $\mathbf{l}_1, \dots, \mathbf{l}_m$ be the columns of L . Then we have an “outer product” expansion of M :

$$M = (d_1 \mathbf{l}_1)(d_1 \mathbf{l}_1)^T + \dots + (d_m \mathbf{l}_m)(d_m \mathbf{l}_m)^T.$$

\square

Examples. Example: Let $R := \mathbb{R}[X]$ and let I be the ideal generated by $X(X^2 + 1)$. (We considered this example previously. Here we treat it in a more systematic way.) Here is the multiplication table for the algebra $A := R/I$ for the basis $\mathbf{b} := (1, X, X^2)^T$:

·	1	X	X ²
1	1	X	X ²
X	X	X ²	-X
X ²	X ²	-X	-X ²

(Note we have the production (replacement) rule $X^3 \mapsto -X$.) Let M be a symmetric 3-by-3 matrix:

$$M := \begin{pmatrix} m_{00} & m_{01} & m_{02} \\ m_{01} & m_{11} & m_{12} \\ m_{02} & m_{12} & m_{22} \end{pmatrix}$$

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.

We have

$$\mathbf{b}^T M \mathbf{b} = m_{00} + 2m_{01}X + (m_{11} + 2m_{02})X^2 + 2m_{12}X^3 + m_{22}X^4.$$

Note that this polynomial is in the ideal I iff it equals 0 in the algebra A . Now in the algebra A this polynomial reduces to the following one:

$$m_{00} + 2(m_{01} - m_{12})X + (m_{11} - m_{22} + 2m_{02})X^2.$$

This element of A is 0 iff the following linear equations are satisfied:

$$m_{00} = 0, m_{01} = m_{12}, m_{11} - m_{22} + 2m_{02} = 0.$$

In other words, the matrix M must have the following form:

$$M = \begin{pmatrix} 0 & s & u - t \\ s & 2t & s \\ u - t & s & 2u \end{pmatrix}.$$

We now impose the constraint $0 \leq M$. Since $m_{00} = 0$ the first row and column of M must be zero - that is, we must have $s = 0$ and $u = t$. Hence M must have the following form:

$$M = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 2t & 0 \\ 0 & 0 & 2t \end{pmatrix}.$$

Note that $0 \leq M$ iff $0 \leq t$. With this M we get the following sum of squares in the ideal I :

$$\mathbf{b}^T M \mathbf{b} = 2t(X^2 + X^4).$$

It follows that we can add the polynomials X and X^2 to the ideal I to get the ideal J generated by X without changing the set of real solutions.