MIT 6.7230 - Algebraic techniques and semidefinite optimization April 12, 2024

Lecture 14

Lecturer: Pablo A. Parrilo *Scribe: ???*

1 Generalizing the Hermite matrix

Recall the basic construction of the Hermite matrix $H_q(p)$ in the univariate case, whose signature gave important information on the signs of the polynomial $q(x)$ on the real roots of $p(x)$.

In a very similar way to the extension of the companion matrix to the multivariate case, we can construct an analogue of the Hermite form for general zero-dimensional ideals. The basic idea is again to consider the zero-dimensional ideal $I \subset \mathbb{R}[x_1, \ldots, x_n]$, and an associated basis of the quotient ring $B = \{x^{\alpha_1}, \ldots, x^{\alpha_m}\}$, where the elements of B are standard monomials.

For simplicity, we assume first that I is radical. In this case, the corresponding finite variety is given by m distinct points, i.e., $V(I) = \{r_1, \ldots, r_m\} \subset \mathbb{C}^n$. Notice that by the definition of the multiplication matrices M_{x_i}, we have $\sum_{i=1}^m r_i^\beta = \mathrm{Tr}[M_{x_1}^{\beta_1} \cdots M_{x_n}^{\beta_n}]$. Thus, in a similar way as we did in the univariate case, for any polynomial $q = \sum_\beta q_\beta x^\beta$ we have

$$\sum_{i=1}^m q(r_i) = \mathrm{Tr}[q(M_{x_1}, \ldots, M_{x_n})]. \tag{1}$$

Once again, this implies that if we have access to matrix representations $M_{x_1}, \ldots, M_{x_n}$, then we can explicitly compute the sum of q at all roots, by evaluating the trace of the polynomial q at the matrices M_{x_i}. Notice also that, if both q and the generators of the ideal have rational coefficients, then the expression above is also a rational number (even if the roots are not).

Example 1. *Consider the system in Example 4 of the previous lecture, and the polynomial* $p(x, y, z) = (x+y+z)^2$. *To evaluate the sum of the values that this polynomial takes on the variety, we compute:*

$$p(M_x, M_y, M_z) = \mathrm{Tr}\,(M_x + M_y + M_z)^2 = \mathrm{Tr} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 2 & 3 & 2 & 2 & 2 \\ 3 & 2 & 3 & 2 & 2 \\ 2 & 2 & 2 & 3 & 2 \\ 2 & 2 & 2 & 2 & 3 \end{bmatrix} = 12.$$

As expected, the squares of the sum of the coordinates of each of the five roots are $\{0, 9, 1, 1, 1\}$, *with the total sum being equal to 12.*

Given any $q \in \mathbb{R}[x_1, \ldots, x_n]$, we can then define a Hermite-like matrix $H_q(I)$ as

$$[H_q(I)]_{jk} := \sum_{i=1}^m q(r_i) r_i^{\alpha_j + \alpha_k}. \tag{2}$$

Notice that the rows and columns of $H_q(I)$ are indexed by standard monomials.

Consider now a vector $f = [f_1, \ldots, f_m]^T$, and the quadratic form

$$
\begin{aligned}
f^T H_q(I) f &:= \sum_{j,k=1}^{m} \sum_{i=1}^{m} q(r_i)(f_j r_i^{\alpha_j})(f_k r_i^{\alpha_k}) \\
&= \sum_{i=1}^{m} q(r_i)(f_1 r_i^{\alpha_1} + \cdots + f_m r_i^{\alpha_m})^2 \\
&= \mathrm{Tr}[(qf^2)(M_{x_1}, \ldots, M_{x_n})].
\end{aligned} \tag{3}
$$

As we see, the matrix $H_q(I)$ is a specific representation, in a basis given by standard monomials, of a quadratic form $H_q : \mathbb{C}[x]/I \to \mathbb{C}$, with $H_q : f \mapsto \sum_{i=1}^{m} (qf^2)(r_i)$. The expressions in (3) allow us to explicitly compute a matrix representation of this quadratic map. (What is the other "natural" representation of this map?)

The following theorem then generalizes the results of the univariate case, and enables, among other things, to do root counting.

Theorem 2. *The signature of the matrix $H_q(I)$ is equal to the number of real points r_i in $V(I)$ for which $q(r_i) > 0$, minus the number of real points for which $q(r_i) < 0$.*

Corollary 3. *Consider a zero dimensional ideal I. The signature of the matrix $H_1(I)$ is equal to the number of real roots, i.e., $|V(I) \cap \mathbb{R}^n|$.*

In the general (non-radical) case, we would take the property (3) as the definition of $H_q(I)$, instead of (2). Also, in Theorem 2, multiple real zeros are counted only once.

2 Parametric versions

One of the most appealing properties of Groebner-based eigenvalue methods is that they allow us to extend many of the results to the *parametric* case, i.e., when we are interested in obtaining all solutions of a polynomial system as a function of some additional parameters η_i.

Consider for simplicity the case of a single parameter η, and a polynomial system defined by $p_i(x, \eta) = 0$. In order to solve this for any fixed η, we need to compute a Groebner basis of the corresponding ideal. However, when η changes, it is possible that the resulting set of polynomials is no longer a GB. A way of fixing this inconvenience is to compute instead a *comprehensive Groebner basis*, which is a set of polynomials with the the property that it remains a Groebner basis of I for all possible specializations of the parameters. Using the corresponding monomials as a basis for the quotient space, we can give an eigenvalue characterization of the solutions for all values of η.

3 Sums of squares on quotient rings

We describe next a natural modification of the standard sos methods, that will allow us to compute sum of squares decompositions on quotient rings. This can be done by using essentially the same SDP techniques as in the standard case. Since we will need to do effective computations on the quotient, we assume that a Gröbner basis $\mathcal{G} = \{b_1, \ldots, b_k\}$ of the polynomial ideal I is available.

The method will be basically the same as in the standard case (expressing the polynomial as a quadratic form on a vector of monomials and writing linear equations to obtain a semidefinite program), but with two main differences:

- Instead of indexing the rows and columns of the matrix Q in the semidefinite program by the usual monomials, we use *standard* monomials corresponding to the Gröbner basis $\mathcal{G}$ of the

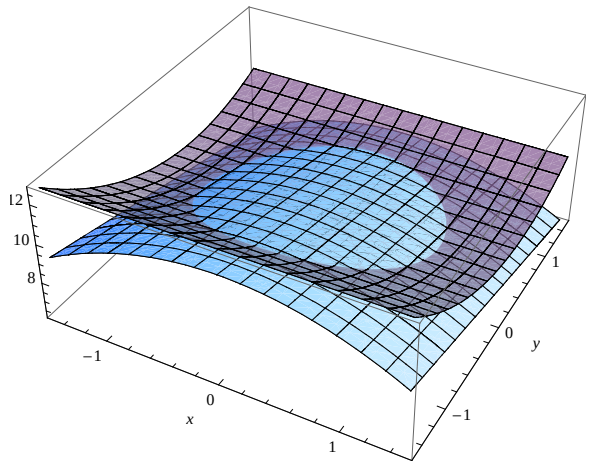

Figure 1: The polynomials $p = 10 - x^2 - y$ and $(3 - \frac{y}{6})^2 + \frac{35}{36}\, y^2$ take exactly the same values on the unit circle $x^2 + y^2 = 1$. Thus, p is nonnegative on the circle.

ideal I. These are the monomials that are not divisible by any leading term of the polynomials b_i in the Gröbner basis.

- When equating the left- and right-hand sides to form linear equations defining the subspace of valid Gram matrices, all operations are performed in the quotient ring, i.e., we rewrite the terms in *normal form* after multiplication.

Rather than giving a formal description, it is more transparent to explain the methodology via a simple example:

Example 4. *Consider the problem of deciding if the polynomial $p := 10 - x^2 - y$ is nonnegative on the variety defined by $f := x^2 + y^2 - 1 = 0$ (the unit circle). We will check whether p is a sum of squares in $\mathbb{R}[x, y]/I$, where I is the ideal $I = \langle f \rangle$. Since the ideal I is principal (generated by a single polynomial), we already have a Gröbner basis, which is simply $\mathcal{G} = \{f\}$. We use a graded lexicographic monomial ordering, where $x \prec y$. The corresponding set of standard monomials is then $\mathcal{B} = \{1, x, y, x^2, xy, x^3, x^2y, \ldots\}$.*

To formulate the corresponding semidefinite program, we pick a partial basis of the quotient ring (i.e., a subset of monomials in $\mathcal{B}$). In this example, we take only $\{1, x, y\}$, and as before, we write p as a quadratic form in these monomials:

$$\begin{aligned} 10 - x^2 - y &= \begin{bmatrix} 1 \\ x \\ y \end{bmatrix}^T \begin{bmatrix} q_{11} & q_{12} & q_{13} \\ q_{12} & q_{22} & q_{23} \\ q_{13} & q_{23} & q_{33} \end{bmatrix} \begin{bmatrix} 1 \\ x \\ y \end{bmatrix} \\ &= q_{11} + q_{22}x^2 + q_{33}y^2 + 2q_{12}x + 2q_{13}y + 2q_{23}xy \\ &\equiv (q_{11} + q_{33}) + (q_{22} - q_{33})x^2 + 2q_{12}x + 2q_{13}y + 2q_{23}xy \qquad \text{mod } I, \end{aligned}$$

where in the last line, we used reduction modulo the ideal to rewrite some terms as linear combinations of standard monomials only (e.g., the term $q_{33}y^2$ is replaced by $q_{33} - q_{33}x^2$). Matching coefficients between left and right, we obtain the linear equations that define the semidefinite program. Solving it, we have for this example

$$Q = \begin{bmatrix} 9 & 0 & -\frac{1}{2} \\ 0 & 0 & 0 \\ -\frac{1}{2} & 0 & 1 \end{bmatrix} = L^T L, \qquad L = \frac{1}{\sqrt{2}} \begin{bmatrix} 3 & 0 & -\frac{1}{6} \\ 0 & 0 & \frac{\sqrt{35}}{6} \end{bmatrix},$$

and therefore

$$10 - x^2 - y \equiv \left(3 - \frac{y}{6}\right)^2 + \frac{35}{36}\, y^2 \qquad mod\ I,$$

which shows that p is indeed a sum of squares on $\mathbb{R}[x, y]/I$. A simple geometric interpretation is shown in Figure 1. As expected, by the condition above, p coincides with an sos polynomial on the variety, and thus it is obviously nonnegative on that set.

Remark 5. *Despite the similarities between the "standard" case of sum of squares on the polynomial ring $\mathbb{R}[x]$ vs. the quotient ring $\mathbb{R}[x]/I$, there are a few important differences. A key distinction is related to computational complexity issues. Consider an sos decomposition $p(x) = \sum_i q_i(x)^2$. When working on $\mathbb{R}[x]$, we can always bound a priori the degree of the polynomials q_i in terms of the degree of p (namely,* $\deg(q_i) \leq \frac{1}{2} deg(p)$*). This is* not *true when working on a quotient ring, since monomials can "wrap around" when computing normal forms. This is the reason why when working on $\mathbb{R}[x]/I$ we typically have some freedom in choosing a finite set of standard monomials to index the matrix Q (unless it is feasible to include all of them).*

In fact, since for the ideal $I = \langle x_1^2 - 1, \ldots, x_n^2 - 1\rangle$ every polynomial nonnegative on $V(I)$ is a sum of squares on $\mathbb{R}[x]/I$ (see below), it directly follows that, in the general case, deciding whether a polynomial is sum of squares modulo I is NP-hard.

Nonnegative polynomials on a finite variety are SOS mod I For simplicity, we assume throughout that the ideal I is radical. Then, if a polynomial is nonnegative on a finite variety, it is a sum of squares on the quotient ring; see [Par02].

Theorem 6. *Let $f(x)$ be nonnegative on $\{x \in \mathbb{R}^n | h_i(x) = 0\}$. If the ideal $I = \langle h_1, \ldots, h_m\rangle$ is radical, then $f(x)$ is a sum of squares in the quotient ring $\mathbb{R}[x]/I$, i.e.. there exist polynomials q_i, λ_i, such that*

$$f(x) = \sum_i q_i^2(x) + \sum_{i=1}^{m} \lambda_i(x) h_i(x).$$

Remark 7. *The assumption that I is radical (or a suitable local modification) is necessary when $f(x)$ is nonnegative but not strictly positive. For instance, the polynomial $f = x$ is nonnegative on the variety defined by the (non-radical) ideal $\langle x^2\rangle$, although no decomposition of the form $x = s_0(x) + \lambda(x)x^2$ (where s_0 is SOS), can possibly exist.*

References

[Par02] P. A. Parrilo. An explicit construction of distinguished representations of polynomials nonnegative over finite sets. Technical Report IfA Technical Report AUT02-02. Available from `https://www.mit.edu/~parrilo`, ETH Zürich, 2002.