MIT 6.7230 - Algebraic techniques and semidefinite optimization April 10, 2024

Lecture 13

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

Today we will see a few more examples and applications of Groebner bases, and we will develop in detail the zero-dimensional case.

1 Zero-dimensional ideals

In practice, we are often interested in polynomial systems that have only a finite number of solutions (the "zero-dimensional" case), and as we will see, many interesting things happen in this case.

Definition 1. *An ideal I is* zero-dimensional *if the associated variety $V(I)$ is a finite set.*

Given a system of polynomial equations, how to decide if it has a finite number of solutions (i.e., if the corresponding ideal is zero-dimensional)? If the system has finitely many solutions, then the quotient ring $\mathbb{C}[x]/I$ is finite-dimensional, and the number of roots (counted with multiplicity) is equal to the dimension of this vector space (why?). Furthermore, this is also equal to the number of *standard monomials*, i.e., the monomials "under the staircase."

When do we have a finite number of standard monomials? We can state a simple criterion for this in terms of a Groebner basis.

Lemma 2. *Let G be a Groebner basis of the ideal $I \subset \mathbb{C}[x_1, \ldots, x_n]$. The ideal I is zero-dimensional if and only if for each i ($1 \leq i \leq n$), there exists an element in the Groebner basis whose initial term is a pure power of x_i.*

Among other important consequences, when I is a zero-dimensional ideal the quotient ring $\mathbb{C}[x]/I$ is a *finite dimensional* vector space, with its dimension being equal to the number of *standard monomials*. These are the monomials that are not in the initial ideal $in(I)$ (i.e., the monomials "under the staircase"). In fact, the condition in the lemma can be easily interpreted as ensuring that the number of standard monomials remains finite.

In the zero-dimensional case we can produce *two* natural bases of the coordinate ring $\mathbb{C}[x]/I$ (for simplicity, we assume no multiplicities). The first one, as already explained, is given by the standard monomials (and thus, functions on the variety can be represented in terms of linear combinations of standard monomials). The other basis is perhaps more natural, and simply corresponds to the values that a function takes on the points of the variety. The existence of these two bases (and the coordinate change between them) will allow us to use linear algebra techniques to solve polynomial equations.

Polynomial systems via eigenvalues In the zero-dimensional case, we can use Groebner bases to reduce a zero-dimensional polynomial system to a standard eigenvalue problem, generalizing the "companion matrix" notion from the univariate case. We sketch this below.

Recall that in this case, the quotient $\mathbb{C}[x]/I$ is a finite dimensional vector space. The main idea is to consider the homomorphisms given by the n linear maps $M_{x_i} : \mathbb{C}[x]/I \to \mathbb{C}[x]/I$, $f \mapsto \widehat{(x_i f)}$ (that is, multiplication by the coordinate variables, followed by normal form). Choosing as a basis the set of standard monomials, we can effectively compute a matrix representation of these linear maps. This defines n matrices M_{x_i}, that commute with each other (why?).

Assume for simplicity that all the roots have single multiplicity. Since all the M_{x_i} commute, they can be simultaneously diagonalized by a single matrix V, and the kth diagonal entry of $VM_{x_i}V^{-1}$ contains the ith coordinate of the kth solution, for $1 \leq k \leq \#\{V(I)\}$.

(In general, we can block-diagonalize this commutative algebra, splitting into its semisimple and nilpotent components. The nilpotent part is trivial if and only if the ideal is radical.)

To understand these ideas a bit better, let's recall the univariate case.

Example 3. *Consider the ring $\mathbb{C}[x]$ of polynomials in a single variable x, and an ideal $I \subset \mathbb{C}[x]$. Since every ideal in this ring is principal, I can be generated by a single polynomial $p(x) = p_n x^n + \cdots + p_1 x + p_0$. Then, we can write $I = \langle p(x) \rangle$, and $\{p(x)\}$ is a Groebner basis for the ideal (why?). The quotient $\mathbb{C}[x]/I$ is an n-dimensional vector space, with a suitable basis given by the standard monomials $\{1, x, \ldots, x^{n-1}\}$.*

Consider as before the linear map $M_x : \mathbb{C}[x]/I \to \mathbb{C}[x]/I$. The matrix representation of this linear map in the given basis is given by

$$\begin{bmatrix} 0 & 0 & 0 & \cdots & -p_0/p_n \\ 1 & 0 & 0 & \cdots & -p_1/p_n \\ 0 & 1 & 0 & \cdots & -p_2/p_n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & -p_{n-1}/p_n \end{bmatrix},$$

which is the standard companion matrix $\mathcal{C}_p$ associated with $p(x)$. Its eigenvalues are exactly the roots of $p(x)$.

We present next a multivariate example.

Example 4. *Consider the ideal $I \subset \mathbb{C}[x, y, z]$ given by*

$$I = \langle xy - z, yz - x, zx - y \rangle.$$

Choosing a term ordering (e.g., lexicographic, where $x \prec y \prec z$), we obtain the Groebner basis

$$G = \{x^3 - x, yx^2 - y, y^2 - x^2, z - yx\}.$$

We can directly see from this that I is zero-dimensional (why?). A basis for the quotient space is given by $\{1, x, x^2, y, yx\}$. Considering the maps M_x, M_y, and M_z, we have that their corresponding matrix representations are given by

$$M_x = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}, \qquad M_y = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix}, \qquad M_z = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \end{bmatrix}.$$

It can be verified that these three matrices commute. A simultaneous diagonalizing transformation is given by the matrix:

$$V = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & 1 & -1 \\ 1 & -1 & 1 & -1 & 1 \end{bmatrix}, \qquad V^{-1} = \frac{1}{4} \begin{bmatrix} 4 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & -1 & -1 \\ -4 & 1 & 1 & 1 & 1 \\ 0 & 1 & -1 & 1 & -1 \\ 0 & 1 & -1 & -1 & 1 \end{bmatrix}.$$

The corresponding transformed matrices are:

$$VM_xV^{-1} = \mathrm{diag}(0, 1, 1, -1, -1)$$
$$VM_yV^{-1} = \mathrm{diag}(0, 1, -1, 1, -1),$$
$$VM_zV^{-1} = \mathrm{diag}(0, 1, -1, -1, 1)$$

from where the coordinates of the five roots can be read.

In the general (radical) case, the matrix V is a generalized Vandermonde matrix, with rows indexed by roots (points in the variety) and columns indexed by the standard monomials. The V_{ij} entry contains the j-th monomial evaluated at the ith root. Since $VV^{-1} = I$, we can also interpret the jth column of V^{-1} as giving the coefficients of a Lagrange interpolating polynomial $p_j(x)$, that vanishes at all the points in the variety, except at r_j, where it takes the value 1 (i.e., $p_j(r_k) = \delta_{jk}$).

```
Generalize Hermite form, etc
```
ToDo

Remark 5. *In practice, a better alternative to a full diagonalization (which is in general numerically unstable) is a Schur-like approach, where we find a unitary matrix U that simultaneously triangularizes the matrices in the commuting family; see [CGT97] for details.*

2 Hilbert series

Consider an ideal $I \subset \mathbb{C}[x]$ and the corresponding quotient ring $\mathbb{C}[x]/I$. We have seen that, once a particular Groebner basis is chosen, we could associate to every element of $\mathbb{C}[x]/I$ a unique representative, namely a $\mathbb{C}$-linear combination of *standard monomials*, obtained as the remainder after division with the corresponding Groebner basis. We are interested in studying, for every integer k, the dimension of the vector space of remainders of degree less than or equal to k. Expressed in a simpler way, we want to know how many standard monomials of degree k there are, for any given k.

Rather than studying this for different values of k separately, it is convenient to collect (or bundle) all these numbers together in a single object (this general technique is usually called "generating function"). The *Hilbert series* of I, denoted $H_I(t)$, is then defined as the generating function of the dimension of the space of residues of degree k, i.e.,

$$H_I(t) = \sum_{k=0}^{\infty} \dim(\mathbb{C}[x]/I \cap P_{n,k}) \cdot t^k, \tag{1}$$

where $P_{n,k}$ denotes the set of homogeneous polynomials is n variables of degree k.

Notice that, if the ideal is zero-dimensional, the corresponding Hilbert series is actually a finite sum, and thus a polynomial. The number of solutions is then equal to $H_I(1)$.

Example 6. *For the ideal I in Example 4, the corresponding Hilbert function is $H_I(t) = 1+2t+2t^2$.*

To compute the Hilbert series, we use the fact that for *graded* orderings (i.e., those for which $|\alpha| < |\beta|$ implies $\alpha \prec \beta$), the Hilbert series of I and of the initial ideal $in(I)$ are the same. For the monomial ideal $in(I)$, the Hilbert function can be easily obtained via inclusion-exclusion. Thus, all the relevant algebraic and geometric properties of the ideal (e.g., the number of roots, dimension, etc) can be obtained from a Groebner basis via the Hilbert function.

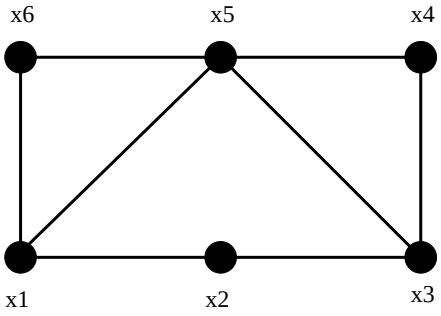

Figure 1: A six-node graph.

3 Examples

3.1 Graph ideals

Consider a graph $G = (V, E)$, and define the associated *edge ideal* $I_G = \langle x_i x_j : (i, j) \in E \rangle$. Notice that I_G is a monomial ideal. For instance, for the graph in Figure 1, the corresponding ideal is given by:

$$I_G := \langle x_1x_2, x_2x_3, x_3x_4, x_4x_5, x_5x_6, x_1x_6, x_1x_5, x_3x_5 \rangle.$$

One of the motivations for studying this kind of ideals is that many graph-theoretic properties (e.g., bipartiteness, acyclicity, connectedness, etc) can be understood in terms of purely algebraic properties of the corresponding ideal. This enables the extension and generalization of these notions to much more abstract settings (e.g., simplicial complexes, resolutions, etc).

For our purposes here, rather than studying I_G directly, we will instead study the ideal obtained when restricting to zero-one solutions[1]. For this, consider the ideal I_b defined as

$$I_b := \langle x_1^2 - x_1, \ldots, x_n^2 - x_n \rangle. \tag{2}$$

Clearly, this is a zero-dimensional radical ideal, with the corresponding variety having 2^n distinct points, namely $\{0, 1\}^n$. Its corresponding Hilbert series is $H_{I_b}(t) = (1 + t)^n = \sum_{k=0}^{n} \binom{n}{k} t^k$.

Since we want to study the intersection of the corresponding varieties, we must consider the *sum* of the ideals, i.e., the ideal $I := I_G + I_b$. It can be shown that the given set of generators (i.e., the ones corresponding to the edges, and the quadratic relations in (2)) are always a Groebner basis of the corresponding ideal. What are the standard monomials? How can they be interpreted in terms of the graph?

The *Hilbert function* of the ideal I can be obtained from the Groebner basis. In this case, the corresponding Hilbert function is given by

$$H_I(t) = 1 + 6t + 7t^2 + t^3,$$

and we can read from the coefficient of t^k the number of stable sets of size k. In particular, the degree of the Hilbert function (which is actually a polynomial, since the ideal is zero-dimensional) indicates the size of the maximum stable set, which is equal to three in this example (for the subset $\{x_2, x_4, x_6\}$).

Remark 7. *An important generalization of the edge ideals of a graph is the* Stanley-Reisner *ideal associated to a simplicial complex; see e.g. [Sta04, MS04] for details and applications.*

[1]There are more efficient ways of doing this, that would not require adding generators. We adopt this approach to keep the discussion relatively straightforward.

3.2 Integer programming

Another interesting application of Groebner bases deals with integer programming. For more details, see the papers [CT91, ST97, TW97].

Consider the integer programming problem

$$\min c^T \mathbf{x} \qquad \text{s.t.} \quad \begin{cases} A\mathbf{x} = b \\ \mathbf{x} \geq 0 \\ \mathbf{x} \in \mathbb{Z}^n \end{cases} \tag{3}$$

where $A \in \mathbb{Z}^{m \times n}$, $b \in \mathbb{Z}^m$, and $c \in \mathbb{Z}^n$. For simplicity, we assume that $A, c \geq 0$, and that we know a feasible solution $\mathbf{x}_0$. These assumptions can be removed.

The main idea to solve (3) will be to interpret the nonnegative integer decision variables $\mathbf{x}$ as the *exponents* of a monomial. The integer program will be modeled in terms of a *binomial ideal*, i.e., where the generating polynomials have only two nonzero terms. The associated Groebner basis can then be interpreted in terms of a *integer test set*, i.e., a set of lattice reduction operations under which the objective function improves monotonically, and that produce the optimal solution.

`Complete`	ToDo

Example 8. *Consider the problem data given by*

$$A = \begin{bmatrix} 4 & 5 & 6 & 1 \\ 1 & 2 & 7 & 3 \end{bmatrix}, \qquad b = \begin{bmatrix} 750 \\ 980 \end{bmatrix}, \qquad c^T = \begin{bmatrix} 1 & 2 & 3 & 4 \end{bmatrix}.$$

An initial feasible solution is given by $\mathbf{x}_0 = [0, 30, 80, 120]^T$. *Notice that given a feasible* $\mathbf{x}_0$ *we can compute the right-hand side b (so we don't really need b). We will work on the ring* $\mathbb{C}[z_1, z_2, w_1, w_2, w_3, w_4]$. *Thus, we need to compute a Groebner basis G of the binomial ideal*

$$\langle z_1^4 z_2 - w_1, z_1^5 z_2^2 - w_2, z_1^6 z_2^7 - w_3, z_1 z_2^3 - w_4 \rangle,$$

for a term ordering that combines elimination of the z_i *with the weight vector c. To obtain the solution, we compute the normal form of the monomial given by the initial feasible point, i.e.,* $w_2^{30} w_3^{80} w_4^{120}$. *This reduction process yields the result* $w_2^8 w_3^{106} w_4^{74}$, *and thus the optimal solution is* $[0, 8, 106, 74]$. *The corresponding costs of the initial and optimal solutions are* $c^T \mathbf{x}_0 = 780$ *and* $c^T \mathbf{x}_{opt} = 630$.

We should remark that there are more efficients ways of implementing this than the one described. Also, although this basic method cannot currently compete with specialized techniques used in integer programming for most problems, there are some particular cases where it is very efficient, mostly related with the solution of parametric problems. Additionally, these techniques have been used to prove the existence of polynomial-time algorithms for certain subclasses of integer programs; see e.g. [DLHK13] and the references therein.

References

[CGT97] R. M. Corless, P. M. Gianni, and B. M. Trager. A reordered Schur factorization method for zero-dimensional polynomial systems with multiple roots. In *ISSAC '97: Proceedings of the 1997 international symposium on Symbolic and algebraic computation*, pages 133–140, New York, NY, USA, 1997.

[CT91] P. Conti and C. Traverso. Buchberger algorithm and integer programming. In *Applied algebra, algebraic algorithms and error-correcting codes (New Orleans, LA, 1991)*, volume 539 of *Lecture Notes in Comput. Sci.*, pages 130–139. Springer, Berlin, 1991.

[DLHK13] J. A. De Loera, R. Hemmecke, and M. Köppe. *Algebraic and geometric ideas in the theory of discrete optimization*, volume 14. SIAM, 2013.

[MS04] E. Miller and B. Sturmfels. *Combinatorial commutative algebra*, volume 227. Springer Science & Business Media, 2004.

[ST97] B. Sturmfels and R. Thomas. Variation of cost functions in integer programming. *Math. Programming*, 77(3, Ser. A):357–387, 1997.

[Sta04] R. P. Stanley. *Combinatorics and commutative algebra*, volume 41. Springer Science & Business Media, 2 edition, 2004.

[TW97] R. Thomas and R. Weismantel. Truncated Gröbner bases for integer programming. *Appl. Algebra Engrg. Comm. Comput.*, 8(4):241–256, 1997.