MAST90053 EXPERIMENTAL MATHEMATICS

LECTURER: DR ANDREA BEDINI

2014, SEMESTER 1, WEEK 4

Gröbner basis

This section is a very brief introduction to Gröbner bases, largerly adapted from Adams and Loustaunau (1994).

In many instances we are interested in the set of roots of a system of nonlinear equations. For example, let $f_i \in \mathbb{R}[x_1, \ldots, x_n]$ be polynomials in n variables with real coefficients. In geometry, a variety $V(f_1, \ldots, f_s)$ can be defined as the set of all solutions of the equations

$$f_1 = 0, \quad f_2 = 0, \quad \ldots, \quad f_s = 0,$$

or more formally, if $S \subseteq \mathbb{R}[x_1, \ldots, x_n]$,

$$V(S) = \{(a_1, \ldots, a_n)^n \in \mathbb{R} \mid f(a_1, \ldots, a_n) = 0 \text{ for all } f \in S\}.$$

There are many (numerical) algorithms for solving systems of equations, but these generally do not take the geometry of the variety into account. The computation of solutions for may drastically improve if the given system of equations is transformed into a different system with the same solutions. This will be done by considering the ideal generated by the polynomials $f_1, \ldots, f_s$, denoted by $\langle f_1, \ldots, f_s \rangle$.

$$\langle f_1, \ldots, f_s \rangle = \left\{ \sum_{i=1}^{s} u_i f_i \mid u_i \in \mathbb{R}[x_1, \ldots, x_n], i = 1, \ldots, s \right\}$$

The linear case

In this section we consider the system

$$f_1 = 0, \quad f_2 = 0, \quad \ldots, \quad f_s = 0, \text{ where each } f_i \text{ is linear.}$$

In this case, the algorithmic method to solve such system is the row reduction which changes the above system to row echelon form. Let's consider an example.

EXAMPLE 1. Let $f_1 = x + y - z$ and $f_2 = 2x + 3y + 2z$ be linear polynomials in $\mathbb{R}[x, y, z]$. We consider the ideal $I = \langle f_1, f_2 \rangle$ and the variety $V(f_1, f_2)$, that is, the solution to the system

$$\begin{cases} x + y & - z & = 0 \\ 2x + 3y & + 2z & = 0. \end{cases}$$

We now perform row reduction on the matrix associated with this sytem

$$\begin{pmatrix} 1 & 1 & -1 \\ 2 & 3 & 2 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 1 & -1 \\ 0 & 1 & 4 \end{pmatrix}.$$

The last matrix is in row echelon form. The solution of the original system are the same as those of the following system

$$\begin{cases} x + y & - z & = 0 \\ \quad + y & + 4z & = 0. \end{cases}$$

and are now easily obtained parametrically as $x = 5z$ and $y = -4z$.

The reduction process is, in fact, a method to change a generating set for the ideal $I = \langle f_1, f_2 \rangle$ into another generating set. In this case we created a new polynomial $f_3 = f_2 - 2f_1 = y + 4z$, and replaced f_2 by f_3. Since f_2 can be written as $2f_1 + f_3$, the ideal $I = \langle f_1, f_2 \rangle$ can also be generated by $\langle f_1, f_3 \rangle$.

The process by which the polynomial f_2 was replaced by f_3 using f_1 is called reduction of f_2 by f_1, and we write

$$f_2 \xrightarrow{f_1} f_3.$$

The polynomial f_3 can be viewed as a reminder of a certain division: we used the first term of f_1, namely x, to eliminate a term from f_2, namely $2x$. Since the first term of f_1 cannot eliminate any other terms, the division stops and the reminder is exactly f_3.

One variable case

In this section we consider polynomials in $\mathbb{R}[x]$. If $f = a_x x^n + a_{n-1}x^{n-1} + \ldots + a_0$, with $a_n \neq 0$, then: the *degree* of f is $\deg(f) = n$; the *leading term* is $\mathrm{lt}(f) = a_n x^n$; the *leading power* is $\mathrm{lp}(f) = x^n$; and the *leading coefficient*, denoted by $\mathrm{lc}(f) = a_n$.

Example 2. Let $f = x^3 - 2x^2 + 2x + 8$ and $g = 2x^2 + 3x + 1$. We can divide f by g to get the quotient $\frac{1}{2}x - \frac{7}{4}$ and remainder $\frac{27}{4}x + \frac{39}{4}$ as follows:

$$\begin{array}{rl} & \frac{1}{2}x - \frac{7}{4} \\ \cline{2-2} 2x^2 + 3x + 1 \big) & x^3 - 2x^2 + 2x + 8 \\ & -x^3 - \frac{3}{2}x^2 - \frac{1}{2}x \\ \cline{2-2} & -\frac{7}{2}x^2 + \frac{3}{2}x + 8 \\ & \frac{7}{2}x^2 + \frac{21}{4}x + \frac{7}{4} \\ \cline{2-2} & \frac{27}{4}x + \frac{39}{4} \end{array}$$

so we have $f = (\frac{1}{2}x - \frac{7}{4})g + \frac{27}{4}x + \frac{39}{4}$. The idea was to multiply g by an appropriate term, namely $\frac{1}{2}x$, so taht the leading term of g times

this term cancelled the leading term of f. After this cancellation we obtained the first reminder $h = f - \frac{1}{2}xg = -\frac{7}{2}x^2 + \frac{3}{2}x + 8$. Using the notation introduced above we note that the factor of g in the product is $\frac{lt(f)}{lt(g)}$ and so we get $h = f - \frac{lt(f)}{lt(g)}g$ as the first reminder. Again, we call h a *reduction* of f by g and we write

$$f \xrightarrow{g} h.$$

At variance with the example in previous section, we can now repeat this process on h obtaining the second, and final, reminder $r = \frac{27}{4}x + \frac{39}{4}$. This can be written using the notation

$$f \xrightarrow{g} h \xrightarrow{g} r$$

or, simply,

$$f \xrightarrow{g}_{+} r$$

If we now consider the ideal $I = \langle f, g\rangle$, it is clear that a new generating set would be $\langle g, r\rangle$. In fact, reducing g by r and repeating the process one more it follows that the ideal I is generated by one element, which in this case is the constant function (hence there are no solutions to the equations $f(x) = g(x) = 0$).

THEOREM 1. Every ideal of $I = \langle f_1, \ldots, f_s\rangle \subseteq \mathbb{R}[x]$ is generated by one element $g = \gcd(f_1, \ldots, f_s)$.

This theorem shows that a variety defined by s equations

$$f_1 = f_2 = \ldots = f_s = 0,$$

where $f_i \in \mathbb{R}[x]$ and $i = 1, \ldots, s$ has precisely the same set of solutions as the single equation $g = \gcd(f_1, \ldots, f_s) = 0$, where $\langle f_1, \ldots, f_s\rangle = \langle g\rangle$. The function $\gcd(f_1, \ldots, f_s)$ is therefore the "best" generating set for $V(I)$.

Algorithm 1 One variable division algorithm

Input: $f, g \in \mathbb{R}[x]$, with $g \neq 0$
Output: q, r such that $f = qg + r$ and $r = 0$ or $\deg(r) \leq \deg(g)$

```
q ← 0
r ← f
while r ≠ 0 and deg(g) ≤ deg(r) do
    q ← q + lt(r)/lt(g)
    r ← r − lt(r)/lt(g) g
end while
```

Term ordering

In the previous examples we implicitly used a choice of ordering of the monomials: we used the highest degree term first in the reduction pro-

cess. When dealing with more than one variable, also a choice of order among the various variables has to be made. Two common choices of order among variables are the *lexicographical* and *degree lexicographical* order.

We will denote $x_1^{\beta_1} \cdots x_n^{\beta_n}$ by $\boldsymbol{x}^{\boldsymbol{\beta}}$ where $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_n) \in \mathbb{N}^n$. Also, by convention, $x_1 > x_2 > \ldots > x_n$.

DEFINITION 1. Lexicographic (lex) ordering.

$$\boldsymbol{x}^{\boldsymbol{\alpha}} < \boldsymbol{x}^{\boldsymbol{\beta}} \Leftrightarrow \begin{cases} \text{the first coordinates } \alpha_i \text{ and } \beta_i \text{ in } \boldsymbol{\alpha} \\ \text{and } \boldsymbol{\beta}, \text{ which are different, satisfy} \\ \alpha_i < \beta_i \end{cases}$$

In the case of two variables x_1 and x_2, the lex order gives

$$1 < x_2 < x_2^2 < \ldots < x_1 < x_1x_2 < x_1x_2^2 < \ldots < x_1^2 < \ldots .$$

DEFINITION 2. Lexicographic (lex) ordering.

$$\boldsymbol{x}^{\boldsymbol{\alpha}} < \boldsymbol{x}^{\boldsymbol{\beta}} \Leftrightarrow \begin{cases} \sum_{i=1}^{n} \alpha_i < \sum_{i=1}^{n} \beta_i \\ \text{or} \\ \sum_{i=1}^{n} \alpha_i = \sum_{i=1}^{n} \beta_i \text{ and } \boldsymbol{x}^{\boldsymbol{\alpha}} < \boldsymbol{x}^{\boldsymbol{\beta}} \text{ with} \\ \text{respect to the lexicographic order.} \end{cases}$$

In the case of two variables x_1 and x_2, the deglex order gives

$$1 < x_2 < x_1 < x_2^2 < x_1x_2 < x_1^2 < x_2^3 < x_1x_2^2 < x_1^2x_2 < x_1^3 < \ldots .$$

Multivariate reduction

In the previous sections we had a division algorithm, also referred to as a reduction process. We now need to extend this algorithm to the multivariate case.

The basic idea behind the algorithm is the same as for linear and one variable polynomials: when dividing f by $f_1, \ldots, f_s$, we want to cancel terms of f using the leading terms of the f_i's (so the new terms which are introduced are smaller than the cancelled terms). We want to continue this process until it cannot be done anymore.

DEFINITION 3. Given $f, g, h \in \mathbb{R}[x_1, \ldots, x_s]$, with $g \neq 0$, we say that f reduces to h modulo g in one step, written

$$f \xrightarrow{g} h,$$

if and only if $\mathrm{lp}(f)$ divides a non-zero term X of f and

$$h = f - \frac{X}{\mathrm{lt}(g)} g.$$

EXAMPLE 3. Let $f = y^2x + 4yx - 3x^2$, $g = 2y + x + 1$, and let the term order be deglex with $y > x$. Then we can write the following chain of reductions:

$$f \xrightarrow{g} -\frac{1}{2}yx^2 + \frac{7}{2}yx - 3x^2 \xrightarrow{g} \frac{1}{4}x^3 + \frac{7}{2}yx - \frac{11}{4}x^2 \xrightarrow{g} \frac{1}{4}x^3 - \frac{9}{2}x^2 - \frac{7}{4}x$$

Note that in the last polynomial no term is divisible by $\mathrm{lp}(g) = y$ and so this procedure cannot continue.

DEFINITION 4. A polynomial r is called *reduced* with respect to a set of non-zero polynomials $F = \{f_1, \ldots, f_s\}$ if $r = 0$ or no term in r is divisible by any one of the $\mathrm{lp}(f_i)$, $i = 1, \ldots, s$.

Algorithm 2 Multivariate division algorithm

Input: $f, f_1, \ldots, f_s \in \mathbb{R}[x_1, \ldots, x_n]$, with $f_i \neq 0$, $i = 1, \ldots, s$

Output: $u_1, \ldots, u_s, r$ such that $f = u_1 f_1 + \ldots\ u_s f_s + r$ and r is reduced w.r.t. $\{f_1, \ldots, f_s\}$ and $\max(\mathrm{lp}(u_1)\,\mathrm{lp}(f_1), \ldots, \mathrm{lp}(u_s)\,\mathrm{lp}(f_s), \mathrm{lp}(r)) = \mathrm{lp}(f)$.

```
for i = 1 to s do
    u_i ← 0
end for
r ← 0
h ← f
while h ≠ 0 do
    if there exist i such that lp(f_i) divides lp(h) then
        u_i ← u_i + lt(h)/lt(f_i)
        h ← h − lt(h)/lt(f_i) f_i
    else
        r ← r + lt(h)
        h ← h − lt(h)
    end if
end while
```

Given a term order, we use tha above algorithm to reduce a system of equations by a set of predefined functions, by systematically removing the largest monomials. This is by no means a unique process, as can be seen in the next example.

EXAMPLE 4. Let $f(x, y) = y^2x - x$, and let $I = \langle f_1, f_2 \rangle$ where $f_1(x, y) = yx - y$, $f_2(x, y) = y^2 - x$. Using the deglex ordering with $y > x$, we find

$$f(x, y) = yf_1(x, y) + f_2(x, y),$$

and hence has zero remainder. However, if we first reduce with respect to f_2, we find

$$f(x, y) = xf_2(x, y) + x^2 - x.$$

In this case, the remainder $x^2 - x$ is non-zero but reduced with respect to $\{f_1, f_2\}$.

The ambiguity in this example can be resolved by finding another generating set for I. This finally leads to the definition of a Gröbner basis:

DEFINITION 5. A set of nonzero polynomials $G = \{g_1, \ldots, g_s\}$ contained in an ideal I, is called a Gröbner basis for I if for all $f \in \mathbb{R}[x_1, \ldots, x_n]$ the remainder of the division of f by G is unique.

The following observaton provides a pathway to the construction of such bases. In the division of f by $f_1, \ldots, f_s$, it may happen, as it does in the above example, that some term X in f is divisible by both $\mathrm{lp}(f_i)$ and $\mathrm{lp}(f_j)$ for $i \neq j$. Hence X is divisible by $L = \mathrm{lcm}(\mathrm{lp}(f_i), \mathrm{lp}(f_j))$. If we reduce f using f_i, we get the polynomial $h_1 = f - \frac{X}{\mathrm{lt}(f_i)} f_i$, and if we reduce f using f_j, we get $h_2 = f - \frac{X}{\mathrm{lt}(f_j)} f_j$. The ambiguity introduced is

$$h_2 - h_1 = \frac{X}{\mathrm{lt}(f_i)} f_i - \frac{X}{\mathrm{lt}(f_j)} f_j = \frac{X}{L} S(f_i, f_j),$$

where $S(f, g)$ is given by

$$S(f, g) = \frac{L}{\mathrm{lt}(f)} f - \frac{L}{\mathrm{lt}(g)} g$$

and is called the S-polynomial of f and g.

The classical algorithm for computing Gröbner bases is the Buchberger's algorithm. It constructs a Gröbner bases by making sure all S-polynomials can be reduced to zero.

Algorithm 3 Buchberger's algorithm for computing Gröbner bases

Input: $F = \{f_1, \ldots, f_s\} \subseteq \mathbb{R}[x_1, \ldots, x_n]$, with $f_i \neq 0$, $i = 1, \ldots, s$
Output: $G = \{g_1, \ldots, g_t\}$, a Gröbner bases for $\langle f_1, \ldots, f_s \rangle$

```
G ← F
𝒢 ← {{f_i, f_j} | 1 ≤ i < j ≤ s}
while 𝒢 ≠ ∅ do
    Choose any {f, g} ∈ 𝒢
    𝒢 ← 𝒢 − {{f, g}}
    S(f, g) −G→+ h, where h is reduced w.r.t. G
    if h ≠ 0 then
        𝒢 ← 𝒢 ∪ {{u, h} | for all u ∈ G}
        G ← G ∪ {h}
    end if
end while
```

EXERCISE 1. Implement the multivariate division and Buchberger's algorithms in `Mathematica`. You can use the following definitions for lp, lt, and lc.

```
lt[poly_] := MonomialList[f, variables, order] // First
lp[poly_] := FactorTermsList[lt[poly]][[2]]
lc[poly_] := FactorTermsList[lt[poly]][[1]]
```

These definitions are based on global definitions of `variables` and `order`.

```
variables={x,y};
order=Lexicographic; (* or DegreeLexicographic *)

f = 3 x^3 y^3 + 2 y^2;

lp[f] (* returns x^3 y^3 *)
```

Test your code against `Mathematica` functions `PolynomialReduce` and `GroebnerBasis`. Make sure `Mathematica` is using the same term ordering as your code (using your definition of `variables` and `order`).

```
PolynomialReduce[f,F,variables,
  MonomialOrder->order]
GroebnerBasis[F,variables,MonomialOrder->order]
```

The 3-colour problem

This section is an adaption of Section 2.7 from Adams and Loustaunau (1994), in which we illustrate how the technique of Gröbner bases can be used to determine whether a graph can be 3-coloured.

Consider the following problem. Given a graph G with n vertices with at most one edge between any two vertices. Can we colour the vertices, using only three colours, in such a way that no two vertices connected by an edge have the same colour? In order to use Gröbner bases we have to capture this problem in a set of polynomial equations, which is done as follows. Let the three colours be represented by the three cube roots of unity $1, \xi, \xi^2$ where $\xi^3 = 1$. Each vertex is to be assigned one of these three colours, and this can be represented by the equations

$$x_i^3 - 1 = 0, \quad i = 1, \ldots, n$$

However, if vertex i and j are adjacent their colours need to be different, i.e. $x_i \neq x_j$. This can be transformed into an equation by observing that $0 = x_i^3 - x_j^3 = (x_i - x_j)(x_i^2 + x_i x_j + x_j^2)$. Hence, if i and j are adjacent vertices with different colours, we must have that

$$x_i^2 + x_i x_j + x_j^2 = 0.$$

Let I be the ideal of $\mathbb{C}[x_1, \ldots, x_n]$ generated by the above two sets of polynomials. Then, the graph G is 3-colourable if $V(I) \neq \emptyset$. To establish whether or not $V(I) = \emptyset$, we compute a Gröbner basis B for I. If $1 \in B$, then $V(I) = \emptyset$ and otherwise $V(I) \neq \emptyset$.

EXERCISE 2. Consider the graph on the right. Using `GroebnerBasis`, determine whether or not G is 3-colourable. If G is 3-colourable, find a 3-colouring.

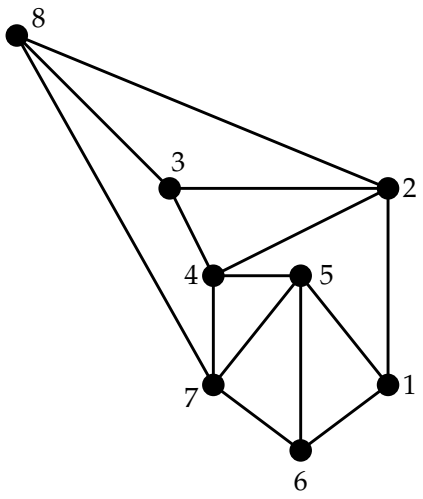

Figure 1: The graph G.

References

Adams, William Wells, and Philippe Loustaunau. 1994. *An Introduction to Gröbner Bases*. 3. American Mathematical Soc.