Computational Algebra

5. S-polynomials and Buchberger's algorithm

So far, we proved that Gröbner bases solve the problem of confluence of the reduction relation, which means that they provide algorithmic ways to test ideal membership and to have a unique division remainder.
However, so far we have not seen any method to produce a Gröbner bases, for example starting from any set of generators of an ideal.
The answer to this problem is Buchberger's Algorithm.

Remark: we notice that one issue in being able to satisfy the defining condition of Gröbner bases is that we can be in the situation:

$I = (x^2-1,\ xy-1) \subset \mathbb{Q}[x,y]$ and $\leq$ = DegLex

then the polynomial

$$y(x^2-1) - x(xy-1) = x-y$$

belongs to I, but its leading term cannot be generated by the leading terms of the two given generators of I

The situation portrayed by the remark is prototypical:

Def: let $f, g \in k[x_1, \dots, x_n]$ be non-zero polynomials; set

$$S(f,g) := \frac{lcm(LT_{\leq}(f), LT_{\leq}(g))}{LM_{\leq}(f)} \cdot f - \frac{lcm(LT_{\leq}(f), LT_{\leq}(g))}{LM_{\leq}(g)} \cdot g$$

we say that $S(f,g)$ is the S-polynomials of f and g

The intuition by Buchberger was that S-polynomials are the "only source" of issues that cause an obstruction to being a Gröbner basis.

In other words:

Theorem: Let $G := \{g_1, \dots, g_s\} \subset k[x_1, \dots, x_n]$ be a set of non-zero polynomials; then G is a Gröbner basis for $I := (g_1, \dots, g_s)$ if and only if for all $i \neq j$, we have $S(g_i, g_j) \xrightarrow{G} 0$.

To prove this result, we first establish a technical lemma.

Lemma: Let $f_1, \dots, f_s \in k[x_1, \dots, x_n]$ be polynomials with the same leading term t; let $c_1, \dots, c_s \in k$ and set $f := \sum_{i=1}^{s} c_i f_i$; if $LT_<(f) < t$, then f is a linear combination, with coefficients in k, of

$$\{ S(f_i, f_j) : i, j \in \{1, \dots, s\} \}$$

Proof: We write $f_i = a_i t + \tilde{f}_i$ with $a_i \in k$ for all $i \in \{1, \dots, s\}$; the hypothesis then says that $\sum_{i=1}^{s} a_i c_i = 0$; by assumption, we have

$$S(f_i, f_j) = \frac{1}{a_i} f_i - \frac{1}{a_j} f_j$$

then

$$\begin{aligned} f &= c_1 f_1 + \dots + c_s f_s \\ &= c_1 a_1 \left(\frac{1}{a_1} f_1\right) + \dots + c_s a_s \left(\frac{1}{a_s}\right) f_s \\ &= c_1 a_1 \left(\frac{1}{a_1} f_1 - \frac{1}{a_2} f_2\right) + \left(c_1 a_1 + c_2 a_2\right)\left(\frac{1}{a_2} f_2 - \frac{1}{a_3} f_3\right) + \dots \\ &\dots + \left(c_1 a_1 + \dots + c_{s-1} a_{s-1}\right)\left(\frac{1}{a_{s-1}} f_{s-1} - \frac{1}{a_s} f_s\right) + \underbrace{\left(c_1 a_1 + \dots + c_s a_s\right)}_{=0}\left(\frac{1}{a_s} f_s\right) \end{aligned}$$

$$= c_1 a_1 S(f_1, f_2) + (c_1 a_1 + c_2 a_2) S(f_2, f_3) + \dots + (c_1 a_1 + \dots + c_{s-1} a_{s-1}) S(f_{s-1}, f_s)$$

which proves the statement

We can now proceed with the proof of the theorem by Buchberger.

Proof: "$\Rightarrow$" suppose that G is a Gröbner basis for $I = (G)$; then $S(f_i, f_j) \in I$, then $S(f_i, f_j) \xrightarrow{G} 0$.

"$\leftarrow$" we assume that $S(f_i, f_j) \xrightarrow{G} 0$ for all $i \neq j$ and we are going to prove that condition A_2 holds, namely, that $f \in I$ if and only if $f = \sum_{i=1}^{s} f_i g_i$ with $LT_<(f) = \max\{LT_<(f_i g_i) : f_i g_i \neq 0\}$

pick $f \in I$, then by assumption f can be written in the form $\sum_{i=1}^{s} f_i g_i$; out of all possible ways to write f as $\sum_{i=1}^{s} f_i g_i$ we choose the one that minimizes the term

$$t := \max\{LT_<(f_i g_i) : f_i g_i \neq 0\}$$

if $t = LT_<(f)$, then we are done; suppose that this is not the case, then $LT_<(f) < t$; we are going to find another representation of f that yields a smaller t, thus giving us a contradiction; let

$$S := \{i \in \{1, \dots, s\} : LT_<(f_i g_i) = t\}$$

for $i \in S$, we write $f_i = c_i t_i + \tilde{f}_i$ with $c_i \in k$, $t_i \in \mathbb{T}^n$; and $LT_<(\tilde{f}_i) < LT_<(f_i)$

we set

$$g := \sum_{i \in S} c_i t_i g_i$$

by construction, $LT_{\leq}(t_i g_i) = t$ for all $i \in S$; however, since $f = \sum_{i=1}^{s} f_i g_i$ and $LT_{\leq}(f) < t$, then the term t must dis- appear in the sum, and so it must disappear in the sum that defines g_i then $LT_{\leq}(g) < t$, so we can apply the previous lemma, that says that there exist $a_{ij} \in k$ such that

$$g = \sum_{i,j \in S,\, i \neq j} a_{ij}\, S(t_i g_i, t_j g_j)$$

now $LT_{\leq}(t_i g_i) = LT_{\leq}(t_j g_j) = t$, so

$$S(t_i g_i, t_j g_j) = \frac{t}{LT_{\leq}(t_i g_i)} t_i g_i - \frac{t}{LT_{\leq}(t_j g_j)} t_j g_j =$$

$$= \frac{t}{LT_{\leq}(g_i)} \cdot g_i - \frac{t}{LT_{\leq}(g_j)} g_j =$$

$$= \frac{t}{lcm(LT_{\leq}(g_i), LT_{\leq}(g_j))} \cdot S(g_i, g_j)$$

since $S(g_i, g_j) \xrightarrow{G} 0$, then $\frac{t}{lcm(\ldots)} S(g_i, g_j) \xrightarrow{G} 0$, which means that $S(t_i g_i, t_j g_j) \xrightarrow{G} 0$; therefore

$$S(t_i g_i, t_j g_j) = \sum_{k=1}^{s} h_{ijk}\, g_k$$

where, since the way how $S(t_i g_i, t_j g_j)$ is written as a sum comes from a reduction process, we have

$$\max\{LT_\leq(h_{ij}^k g_k)\} = LT_\leq(S(t_i g_i, t_j g_j))$$

$$\leq \max\{LT_\leq(t_i g_i), LT_\leq(t_j g_j)\} = t$$

now, we substitute the expressions of $S(t_i g_i, t_j g_j)$ into g, and then we substitute this expression for g into the one of f_i by the way g was obtined, we obtain an expression of f in the form $\sum_{i=1}^{s} f_i' g_i$ with $\max\{LT_\leq(f_i' g_i)\} < t$, which is a contradiction.

This theorem provides the theoretical grand for Buchberger's Algorithm for the computation of Gröbner bases.

Algorithm Buchberger GB

Input: $F = \{f_1, \dots, f_r\} \subseteq k[x_1, \dots, x_n]$, $f_i \neq 0 \ \forall i$; a term order $\leq$

Output: $G = \{g_1, \dots, g_s\}$ for $(f_1, \dots f_r)$ with respect to $\leq$

1. Set $G := F$, $S_{pairs} := \{\{f_i, f_j\} : f_i, f_j \in G, f_i \neq f_j\}$
2. While $S_{pairs} \neq \emptyset$
3. Pick $\{h_1, h_2\} \in S_{pairs}$
4. $S_{pairs} := S_{pairs} \setminus \{h_1, h_2\}$
5. $h :=$ remainder of the reduction of $S(h_1, h_2)$ by G

6. If $h \neq 0$

7. $S_{pairs} := S_{pairs} \cup \{\{u, h\};\ u \in G\}$

8. $G := G \cup \{h\}$

9. Return G

Theorem: Algorithm BuchbergerGB is correct (i.e., given $f_1, \ldots, f_s$ with $f_i \neq 0$ $\forall i$, produces a Gröbner basis for $(f_1, \ldots, f_r)$) and ends in a finite amount of time

Proof: suppose that the algorithm does not terminate; then what happens is that we construct an increasing sequence

$$G_1 \subsetneq G_2 \subsetneq \ldots$$

where each G_i is obtained from G_{i-1} by adding an element $h \in (f_1, \ldots, f_r)$ which is the non-zero reduction with respect to G_{i-1} of an S-polynomial of two elements of G_i; since h is reduced with respect to G_{i-1}, then

$$LT_{\leq}(h) \notin \left(LT_{\leq}(g) : g \in G_{i-1}\right)$$

so we get a strictly ascending chain

$$LT_{\leq}(G_1) \subsetneq LT_{\leq}(G_2) \subsetneq \ldots$$

which contradicts Dickson's Lemma; so the algorithm terminates in finite time; when the algorithm is over, we construct G that by definition contains F, so it is a generating set for $(f_1, \dots, f_r)$; moreover, for any two $g_i, g_j \in G$, we have $S(g_i, g_j) \xrightarrow{G} 0$, so by Buchberger's theorem the set G is a Gröbner basis.

Example: take $f_1 := xy - y$, $f_2 := y^2 - x$ with the DegLex term order; then Buchberger GB outputs $\{f_1, f_2, f_3\}$, where $f_3 := x^2 - x$

Notice that adding a polynomial to a Gröbner basis produces another Gröbner basis, so maybe one may wonder whether there is a notion of "minimality" for Gröbner bases. This is what we achieve by introducing the concept of reduced Gröbner bases.