Sylvester Matrix and GCD for Several Univariate Polynomials

Manuela Wiesinger-Widi*
Doctoral Program Computational Mathematics
Johannes Kepler University Linz
4040 Linz, Austria
manuela.wiesinger@dk-compmath.jku.at

Abstract

We extend Habicht's result on computing a gcd of two univariate polynomials by triangularization of the Sylvester matrix to the case of several univariate polynomials. Similar extensions have been done in the past, but the size of our matrix is smaller.

1 Introduction and Notation

In his PhD thesis [1], Buchberger introduced the notion of Gröbner bases and gave the first algorithm for computing them. Since then, extensive research has been done in order to reduce the complexity of the computation. But nevertheless, even for small examples the computation sometimes does not terminate in reasonable time.

There are basically two approaches for computing a Gröbner basis. The first is the one pursued by the Buchberger algorithm: We start from the initial set F, execute certain reduction steps (consisting of multiplication of polynomials by terms — called shifts — and subtraction of polynomials) and due to Buchberger's theorem, which says that the computation is finished if all the s-polynomials reduce to zero, we know that after finitely many iterations of this procedure we obtain a Gröbner basis of the ideal generated by F. The second approach is to start from F, execute certain shifts of the initial polynomials in F, arrange them as rows in a matrix, triangularize this matrix and from the resulting matrix extract a Gröbner basis.

In project DK1 of the Doctoral Program, which was proposed by Buchberger, we pursue the second approach and seek to improve the theory in order to speed up the Gröbner bases computation. This approach has been studied a couple of times in the past, but never thoroughly. The immediate question is: Does there exist a finite set of shifts such that a triangularization of the matrix built by these shifts yields a Gröbner basis and, if so, how can we construct these shifts? We give results for the univariate case.

Let $\mathbb{K}$ be a field. In the following, when we talk about a shift of a polynomial $f \in \mathbb{K}[x]$, what we mean is f multiplied by a power product x^i for some $i \in \mathbb{N}$.

*This project is funded by the Austrian Science Fund (FWF) under grant W1214/DK1.

Let $f = \sum_{i=0}^{d} c_i x^i \in \mathbb{K}[x]$ with $c_d \neq 0$. For any natural number $m \geq d$ we associate to f the vector $v_m(f) = (0, \ldots, 0, c_d, c_{d-1}, \ldots, c_0) \in \mathbb{K}^{m+1}$. Conversely to any vector $v = (c_m, c_{m-1}, \ldots, c_0) \in \mathbb{K}^{m+1}$ we associate the polynomial $pol(v) = \sum_{i=0}^{m} c_i x^i$ and we define $\deg(v) := deg(pol(v))$. For any sequence of polynomials $f_1, \ldots, f_r \in \mathbb{K}[x]$ we define the matrix

$$mat(f_1, \ldots, f_r) := \begin{pmatrix} v_m(f_1) \\ \vdots \\ v_m(f_r) \end{pmatrix},$$

where $m = \max_{i \in \{1,\ldots,r\}} (\deg(f_i))$.

For any set of polynomials $F = \{f_1, \ldots, f_r\} \subset \mathbb{K}[x]$ we define

$$\langle f_1, \ldots, f_r \rangle_{\mathbb{K}} := \Big\{ \sum_{i=1}^{r} c_i f_i : c_1, \ldots, c_r \in \mathbb{K} \Big\}.$$

2 GCD Computation of Two Univariate Polynomials

We consider the following problem.

Problem 1 (Gcd computation of two univariate polynomials)**.**
Given Two polynomials $f_1, f_2 \in \mathbb{K}[x]$ of degrees $d_1, d_2 \geq 1$, respectively.
Find A gcd of f_1 and f_2.

In [3] (see also [4] for a good overview on this topic), Habicht establishes a connection between the computation of polynomial remainder sequences and linear algebra. In particular, Problem 1 can be solved by triangularizing the Sylvester matrix of f_1 and f_2.

Theorem 2. *Let* $M := \mathrm{mat}(x^{d_2-1} f_1, x^{d_2-2} f_1, \ldots, f_1, x^{d_1-1} f_2, x^{d_1-2} f_2, \ldots, f_2)$ *and let* M' *be a matrix obtained by triangularizing* M*.*
Then the polynomial corresponding to the non-zero row of lowest degree in M' *is a* gcd *of the polynomials* f_1 *and* f_2*.*

We give our own proof for this.

Proof. Assume without loss of generality $d_1 \geq d_2$ and let $f_3, \ldots, f_s, f_{s+1}$ $(s \geq 2)$ be the intermediate polynomials generated by the Euclidean algorithm applied to f_1 and f_2, where $f_{s+1} = 0$. Hence, f_s is the gcd of f_1 and f_2 computed by the algorithm. Let $d_i := \deg(f_i)$ for $i \in \{3, \ldots, s\}$. From d_2 on, the d_i are strictly decreasing.

In the first part of the proof we show:

1. $f_s \in \langle x^{d_2-1} f_1, x^{d_2-2} f_1, \ldots, f_1, x^{d_1-1} f_2, x^{d_1-2} f_2, \ldots, f_2 \rangle_{\mathbb{K}} =: A$
2. $\forall f \in A \setminus \{0\} : (\exists c \in \mathbb{K} \setminus \{0\} : f = c\, f_s \vee \deg(f) > d_s)$.

For the case $s = 2$ the statement in 1. trivially holds. In order to show 1. for the case $s \geq 3$, we prove

$$\forall n \in \{3, \ldots, s\} : f_n \in \langle x^{d_2-d_{n-1}} f_1, x^{d_2-d_{n-1}-1} f_1, \ldots, f_1, x^{d_1-d_{n-1}} f_2, x^{d_1-d_{n-1}-1} f_2, \ldots, f_2 \rangle_{\mathbb{K}}$$

by induction on n.

In the Euclidean algorithm a multiple of f_1 of degree d_1 and multiples of f_2 of degrees up to d_1 are used to compute f_3, so

$$f_3 \in \langle f_1, x^{d_1-d_2} f_2, \ldots, f_2 \rangle_{\mathbb{K}}.$$

We assume that $n \in \{3, \ldots, s-1\}$ and that the statement holds for all $n' \in \{3, \ldots, n\}$. We show that it holds for $n+1$. In the Euclidean algorithm a multiple of f_{n-1} of degree d_{n-1} and multiples of f_n of degrees up to d_{n-1} are used to compute f_{n+1}, so

$$f_{n+1} \in \langle f_{n-1}, x^{d_{n-1}-d_n} f_n, \ldots, f_n \rangle_{\mathbb{K}}.$$

But since

$$f_{n-1} \in \langle f_2 \rangle_{\mathbb{K}}$$

in the case that $n = 3$, and

$$f_{n-1} \in \langle x^{d_2-d_{n-2}} f_1, \ldots, f_1, x^{d_1-d_{n-2}} f_2, \ldots, f_2 \rangle_{\mathbb{K}},$$

otherwise by the induction hypothesis, and

$$f_n \in \langle x^{d_2-d_{n-1}} f_1, \ldots, f_1, x^{d_1-d_{n-1}} f_2, \ldots, f_2 \rangle_{\mathbb{K}},$$

also by the induction hypothesis, this means that

$$f_{n+1} \in \langle x^{d_2-d_n} f_1, \ldots, f_1, x^{d_1-d_n} f_2, \ldots, f_2 \rangle_{\mathbb{K}}.$$

This completes the induction step.

It follows that

$$f_s \in \langle x^{d_2-d_{s-1}} f_1, \ldots, f_1, x^{d_1-d_{s-1}} f_2, \ldots, f_2 \rangle_{\mathbb{K}},$$

and since $d_{s-1} \geq 1$,

$$f_s \in \langle x^{d_2-1} f_1, \ldots, f_1, x^{d_1-1} f_2, \ldots, f_2 \rangle_{\mathbb{K}} = A.$$

We now prove the statement in 2. Let $f \in A \setminus \{0\}$. Since $f_s \mid f_1$ and $f_s \mid f_2$, we obtain that f_s divides every multiple of f_1 and f_2, respectively. From $f \in A \setminus \{0\}$ it follows now that $f_s \mid f$. Hence, $\deg(f) \geq d_s$. Now assume that $\deg(f) = d_s$. Since f is a multiple of f_s, it follows that there exists a $c \in \mathbb{K} \setminus \{0\}$ such that $f = c\, f_s$, which conludes this part of the proof.

So f_s has lowest degree among all the non-zero elements in A. Since f_s is an element in A we can triangularize the matrix M in such a way that f_s occurs in the resulting matrix M^*. So the statement in the theorem holds for M^*. Note that the following holds: Let $B \neq 0$ be a matrix over $\mathbb{K}$ and let B_1 and B_2 be matrices obtained from B by triangularization. Let b_1 and b_2 be the non-zero rows of lowest degree of B_1 and B_2, respectively. Then there exists a $c \in \mathbb{K} \setminus \{0\}$ such that $b_1 = c\, b_2$. Since the gcd is unique up to multiplication by a non-zero element in $\mathbb{K}$, it follows that the statement in the theorem holds for all M' that can be obtained via triangularization of M. $\square$

3 GCD Computation of Several Univariate Polynomials

We consider the following problem.

Problem 3 (Gcd computation of several univariate polynomials)**.**
Given r polynomials $f_1, \ldots, f_r \in \mathbb{K}[x]$, $r \geq 2$, with minimum degree 1.
Find A gcd of $f_1, \ldots, f_r$.

Let d and p be the highest and second highest degrees among the polynomials $f_1, \ldots, f_r$ and let m be the smallest degree.

In [5] and [2] we found direct generalizations of the Sylvester matrix to the case of r univariate polynomials with $r \geq 2$. In [5] the block size for every polynomial is d, resulting in a $d\,r \times 2d$ matrix. In [2] the block size of one of the polynomials with degree d is p and the block sizes of all of the other polynomials are d, resulting in a $(p + (r-1)d) \times (d + p)$ matrix. As can been seen in the next theorem, our block sizes are d for a polynomial with smallest degree m and m for the other polynomials, resulting in a $(d + (r-1)m) \times (d + m)$ matrix. This matrix is much smaller in general and it is equal to the other two only if all of the polynomials have the same degree.

Theorem 4. *Assume f_r has minimal degree m.*
Let $M := \mathrm{mat}(x^{m-1}f_1, x^{m-2}f_1, \ldots, f_1, \ \ldots\ , x^{m-1}f_{r-1}, x^{m-2}f_{r-1}, \ldots, f_{r-1}, x^{d-1}f_r, x^{d-2}f_r, \ldots, f_r)$ and let M' be a matrix obtained by triangularizing M.
Then the polynomial corresponding to the non-zero row of lowest degree in M' is a gcd *of the polynomials in F.*

Proof. Let $d_i := \deg(f_i)$ for $i = 1, \ldots, r$. Assume w.l.o.g. that $d_1 \geq \cdots \geq d_r$, i.e. $d_1 = d$ and $d_r = m$. We prove the theorem by induction on r.

We have proved case $r = 2$ in Theorem 2.

Now assume that the claim holds for a fixed number r of polynomials, $r \geq 2$. We show, that it also holds for $r + 1$. We want to compute a gcd of $f_1, \ldots, f_{r+1}$. First we compute a gcd g of $f_1, \ldots, f_{r-1}, f_{r+1}$ and then a gcd of g and f_r.

By the induction hypothesis we can obtain a gcd g by triangularizing the matrix

$$A := \mathrm{mat}(x^{d_{r+1}-1}f_1, \ldots, f_1, \ \ldots\ , x^{d_{r+1}-1}f_{r-1}, \ldots, f_{r-1}, x^{d_1-1}f_{r+1}, \ldots, f_{r+1}).$$

Let A' be the resulting matrix of this triangularization process. The non-zero rows of A' correspond to polynomials with degrees $k := \deg(g)$ up to $d_1 + d_{r+1} - 1$, each being a multiple of g and hence being a polynomial that can be used for reducing f_r and its multiples during the gcd computation of f_r and g.

Let us now have a look at the problem of adding rows to matrix A' such that we can compute a gcd of g and f_r by triangularizing the then obtained matrix. For this we use the arguments from case $r = 2$.

In order to compute a gcd of f_r and g, multiples of f_r and g up to degree $d_r + d_{r+1} - 1$ suffice since $k \leq d_{r+1}$. Since we have multiples of g up to degree $d_1 + d_{r+1} - 1$ in A' and since $d_1 \geq d_r$, we also have multiples up to degree $d_r + d_{r+1} - 1$ of g, so no additional shifts of g are necessary. But we can already add the shifts of f_r up to $x^{d_{r+1}-1}f_r$ to the matrix A itself instead of first triangularizing to A' and then adding them to A'.

So by triangularizing the matrix

$$M := \mathrm{mat}(x^{d_{r+1}-1}f_1, \ldots, f_1, \ \ldots \ , x^{d_{r+1}-1}f_r, \ldots, f_r, x^{d_1-1}f_{r+1}, \ldots, f_{r+1})$$

to M' we get a gcd of $f_1, \ldots, f_r$ and f_{r+1} as the polynomial corresponding to the non-zero row of M' with smallest degree by the same arguments as in the case for $r = 2$. This concludes the induction step. □

From this result we can infer the following two corollaries. Let M be as in Theorem 4.

Corollary 1. *$f_1, \ldots, f_r$ are co-prime if and only if M has rank $d + m$.*

Corollary 2. *Let g be a gcd of $f_1, \ldots, f_r$. Then* $\deg(g) = d + m - rank(M)$.

4 Conclusion

We gave a proof for the correctness of Habichts method for computing a gcd of two univariate polynomials by triangularization of the Sylvester matrix. We then directly extended the Sylvester matrix in such a way that a gcd of several univariate polynomials can be computed by triangularization. Other generalizations in this style have been done before, however those resulted in a bigger matrix.

References

[1] Bruno Buchberger. An Algorithm for Finding the Basis Elements in the Residue Class Ring Modulo a Zero Dimensional Polynomial Ideal (german). Mathematical Institute, University of Innsbruck, Austria. PhD Thesis. 1965. English translation in *J. of Symbolic Computation, Special Issue on Logic, Mathematics, and Computer Science: Interactions*. Vol. 41, Number 3-4, pages 475–511, 2006.

[2] S. Fatouros and N. Karcanias. Resultant properties of gcd of many polynomials and a factorization representation of gcd. *International Journal of Control*, Vol. 76, Issue 16, pp. 1666–1683, 2003.

[3] Walter Habicht. Eine Verallgemeinerung des Sturmschen Wurzelzählverfahrens. *Comm. Math. Helvetici* 21, pp. 99–116, 1948.

[4] Rüdiger Loos. Generalized Polynomial Remainder Sequences. In *Computer Algebra: Symbolic and Algebraic Computation*, pp. 115–137, Springer-Verlag, 1982.

[5] A.I.G. Vardulakis and P.N.R. Stoyle. Generalized Resultant Theorem. *IMA Journal of Applied Mathematics*, Vol. 22, Issue 3, pp. 331–335, 1978.