Structured Low Rank Approximation of a Sylvester Matrix

Erich Kaltofen, Zhengfeng Yang and Lihong Zhi

Abstract. The task of determining the approximate greatest common divisor (GCD) of univariate polynomials with inexact coefficients can be formulated as computing for a given Sylvester matrix a new Sylvester matrix of lower rank whose entries are near the corresponding entries of that input matrix. We solve the approximate GCD problem by a new method based on structured total least norm (STLN) algorithms, in our case for matrices with Sylvester structure. We present iterative algorithms that compute an approximate GCD and that can certify an approximate ϵ-GCD when a tolerance ϵ is given on input. Each single iteration is carried out with a number of floating point operations that is of cubic order in the input degrees. We also demonstrate the practical performance of our algorithms on a diverse set of univariate pairs of polynomials.

Mathematics Subject Classification (2000). Primary 68W30; Secondary 65K10.

Keywords. Sylvester matrix, approximate greatest common divisor, structured total least norm, hybrid symbolic/numeric algorithm.

1. Introduction

The problem of perturbation errors in the scalars of the inputs to a symbolic computation task has been studied extensively in the recent past, giving rise of the subject of hybrid symbolic/numeric algorithms. Approximate GCDs and factors have been at the center of investigations. One can formulate the algorithm specifications as an optimization problem without appealing to floating point arithmetic [12, 5]. In the GCD case one has the following.

This research was supported in part by the National Science Foundation of the USA under Grants CCR-0305314 and CCF-0514585 (Kaltofen) and OISE-0456285 (Kaltofen, Yang and Zhi). This research was partially supported by NKBRPC (2004CB318000) and the Chinese National Natural Science Foundation under Grant 10401035 (Yang and Zhi).

PROBLEM 1.1. *Input are two univariate polynomials f, $g \in \mathbb{C}[x]$ with degree* $\deg(f) = m$ *and* $\deg(g) = n$. *For a positive integer k with $k \leq \min(m, n)$, we wish to compute $\triangle f$, $\triangle g \in \mathbb{C}[x]$ such that* $\deg(\triangle f) \leq m$, $\deg(\triangle g) \leq n$, $\deg(\mathrm{GCD}(f + \triangle f, g + \triangle g)) \geq k$ *and such that* $\|\triangle f\|_2^2 + \|\triangle g\|_2^2$ *is minimized.*

When $k = 1$, a polynomial time solution is presented in [16]; see also [5, Section 2.6]. One may restrict the above problem to polynomials with entirely real coefficients. Several authors assume that an error estimate ϵ is also input and either output a pair $\triangle f, \triangle g$ with $\|\triangle f\|_2^2 + \|\triangle g\|_2^2 \leq \epsilon$, yielding an ϵ-GCD equal $\mathrm{GCD}(f + \triangle f, g + \triangle g)$, or prove that no such pair exists or output "undecided," the latter when the used numerical techniques cannot settle the problem.

The computation of approximate GCDs of univariate polynomials has been extensively studied [26, 20, 5, 7, 15, 10, 2, 21, 25, 30, 6, 28, 29]. The singular value decomposition (SVD) of the Sylvester matrix derived from the input polynomials is used in [5, 7, 30, 6, 8, 29] to deduce approximate GCDs. By dropping insignificant singular values, the SVD yields a nearby matrix of lower rank, but that matrix has no longer the Sylvester structure. The approximate GCD can be found by additional manipulation, for example from the singular vectors.

Here we propose to approximate the given Sylvester matrix with a rank deficient matrix that also has Sylvester structure, which is an instance of the class of *structure preserving* total least squares problems. There are several methods at our disposal, and we have tested the STLN (structured total least norm) algorithm [22] and the iterated projection algorithm in [4]. STLN is an efficient method for obtaining an approximate solution $(A+E)X = B+H$ to an overdetermined linear system $AX \approx B$, preserving the given linear structure in the minimal perturbation $[E\, H]$. We show how to solve PROBLEM 1.1, at least for a local minimum, by applying STLN with L_2 norm to a submatrix of the Sylvester matrix. The algorithm in [4] projects the nearest rank deficient matrix by imposing Sylvester structure, thus destroying rank deficiency. Then it repeats the SVD/projection steps on the new Sylvester matrices.

We achieve excellent performance of STLN on our test cases, requiring only a handful of iterations and yielding a backward error that is comparable and better than earlier algorithms. For instance, our STLN-based approximations can have a relative backward error that is about 10 times smaller than the one achieved by the SVD-Gauss Newton approach [28]. In contrast, the algorithm in [4] does not perform well, exhibiting slow convergence similarly as experienced earlier on the factoring problem.

The organization of this paper is as follows. In Section 2, we introduce some notations and discuss the equivalence between the GCD problems and the low rank approximation of a Sylvester matrix. In Section 3, we consider solving an overdetermined system with Sylvester structure based on STLN. In Section 4, we describe our approximate GCD based on STLN and discuss the achieved practical performance on a number of benchmark pairs of univariate polynomials. Furthermore, we compare our results with preceding work. We conclude in Section 5 with remarks on the complexity and the rate of convergence of our algorithms.

2. Preliminaries

We first shall prove that the minimization Problem 1.1 in Section 1 always has a solution, in contrast to general and structured total least norm problems [17, Section 3.1]. Let $f, g \in \mathbb{C}[x] \setminus \{0\}$ with degree $\deg(f) = m$ and $\deg(g) = n$, namely

$$\begin{aligned} f &= a_m x^m + a_{m-1}x^{m-1} + \cdots + a_1 x + a_0, \quad a_m \neq 0, \\ g &= b_n x^n + b_{n-1}x^{n-1} + \cdots + b_1 x + b_0, \qquad b_n \neq 0. \end{aligned}$$

Theorem 2.1. *Let k be a positive integer with $k \le \min(m, n)$. There exist $\hat{f}$, $\hat{g} \in \mathbb{C}[x]$ with $\deg(\hat{f}) \le m$, $\deg(\hat{g}) \le n$ and $\deg \mathrm{GCD}(\hat{f}, \hat{g}) \ge k$ such that for all $\tilde{f}, \tilde{g} \in \mathbb{C}[x]$ with $\deg(\hat{f}) \le m$, $\deg(\hat{g}) \le n$ and $\deg \mathrm{GCD}(\tilde{f}, \tilde{g}) \ge k$ we have*

$$\|\hat{f} - f\|_2^2 + \|\hat{g} - g\|_2^2 \le \|\tilde{f} - f\|_2^2 + \|\tilde{g} - g\|_2^2.$$

Proof. Let $h \in \mathbb{C}[x]$ be monic with $\deg(h) = k$ and let $u, v \in \mathbb{C}[x]$ with $\deg(u) \le m - k$ and $\deg(v) \le n - k$. For the real and imaginary parts of the coefficients of h (excluding the leading coefficient, which is set to 1), and of u and v we consider the continuous objective function

$$F(h, u, v) = \|uh - f\|_2^2 + \|vh - g\|_2^2.$$

We prove that the function has a value on a closed and bounded set (with respect to the Euclidean metric) of its real argument vector that is smaller than elsewhere. Hence the function attains, by Weierstrass's theorem, a global minimum. Consider $\bar{f} = a_m x^m$ and $\bar{g} = b_n x^n$, which have a GCD of degree $\ge k$. Clearly, any triple h, u, v with $F(h, u, v) > \|\bar{f} - f\|_2^2 + \|\bar{g} - g\|_2^2$ can be discarded. So the coefficients of uh and vh can be bounded from above, and by any polynomial factor coefficient bound, so can the coefficients of h, u, v. Thus the domain of the function $F(h, u, v)$ can be restricted to a sufficiently large ball. It remains to exclude $u = v = 0$ as the minimal solution. We have $F(h, 0, 0) = \|f\|_2^2 + \|g\|_2^2 > \|\bar{f} - f\|_2^2 + \|\bar{g} - g\|_2^2$. □

Remark 2.2. The above theorem 2.1 remains valid when one restricts the input and perturbed polynomials to have real coefficients. We note that for real input polynomials the optimal complex solution may be nearer than the optimal real solution. In fact, for the pair $f = x^2 + 1$ and $g = x^2 + 2$, the optimal real solution for $k = 1$ is $\hat{f} = 0.723598\,x^2 + 1.170810$ and $\hat{g} = 1.170822\,x^2 + 1.894436$ with $\|\hat{f} - f\|_2^2 + \|\hat{g} - g\|_2^2 = 0.145898$, while there is a nearer pair of complex polynomials with a common root, namely

$$\begin{aligned}\hat{\hat{f}} = {} & (0.81227 - 0.0000016019\, i)\, x^2 \\ & + (0.00000038156 - 0.14812\, i)\, x \\ & + 1.1169 - 0.00000039519\, i\end{aligned}$$

and

$$\begin{aligned}\hat{\hat{g}} = {} & (1.1220 - 0.00000076026\, i)\, x^2 \\ & + (0.0000011810 + 0.096265\, i)\, x \\ & + 1.9240 + 0.0000013787\, i\end{aligned}$$

with $\|\hat{\hat{f}} - f\|_2^2 + \|\hat{\hat{g}} - g\|_2^2 = 0.1007615$. Moreover, in that case, the nearest pair is not unique. The second solution is

$$\begin{aligned}\hat{\hat{f}} = {} & (0.81227 - 0.0000010696\, i)\, x^2 \\ & + (0.0000018180 + 0.14812\, i)\, x \\ & + 1.1169 - 0.0000035348\, i\end{aligned}$$

and

$$\begin{aligned}\hat{\hat{g}} = {} & (1.1220 + 0.0000034875\, i)\, x^2 \\ & + (0.0000010267 - 0.096263\, i)\, x \\ & + 1.9240 + 0.00000055845\, i.\end{aligned}$$

A real example with an ambigous real nearest pair with a GCD is $f = x^2 - 2$ and $g = x^2 - 1$. We note that ambiguity of approximate solutions was already noted in [31, 13]. $\square$

Now suppose $S(f, g)$ is the Sylvester matrix of It is well-known that the degree of GCD of f and g is equal to the rank deficiency of S, we have

$$\min_{\deg(\mathrm{GCD}(\tilde{f},\tilde{g}))\geq k} \|\tilde{f} - f\|_2^2 + \|\tilde{g} - g\|_2^2 \iff \min_{\mathrm{rank}(\tilde{S})\leq \tilde{n}+\tilde{m}-k} \|\tilde{f} - f\|_2^2 + \|\tilde{g} - g\|_2^2 \quad (2.1)$$

where $\tilde{S}$ is the Sylvester matrix generated by $\tilde{f}$ and $\tilde{g}$, with $\deg \tilde{f} = \tilde{m} \leq m$ and $\deg \tilde{g} = \tilde{n} \leq n$. Note that for $\tilde{f} = 0$ or $\tilde{g} = 0$, the Sylvester matrix $\tilde{S}$ is not defined. If one polynomial is zero, we shall assume that $\mathrm{rank}(\tilde{S}) \leq \tilde{n} + \tilde{m} - k$ is satisfied since then GCD is the other non-zero polynomial. If both $\tilde{f} = \tilde{g} = 0$, we shall assume $\mathrm{rank}(\tilde{S}) = \infty$, as that solution is always sub-optimal (see proof of Theorem 2.1).

The k-th Sylvester matrix $S_k \in \mathbb{C}^{(m+n-k+1)\times(m+n-2k+2)}$ is a submatrix of S obtained by deleting the last $k-1$ rows of S and the last $k-1$ columns of coefficients of f and g separately in S.

$$S_k = \left[\begin{array}{cccccc} a_m & & & b_n & & \\ a_{m-1} & \ddots & & b_{n-1} & \ddots & \\ \vdots & \ddots & a_m & \vdots & \ddots & b_n \\ a_0 & & a_{m-1} & b_0 & & b_{n-1} \\ & \ddots & \vdots & & \ddots & \vdots \\ & & a_0 & & & b_0 \end{array}\right],$$
$$\underbrace{\qquad\qquad\qquad}_{n-k+1}\ \underbrace{\qquad\qquad\qquad}_{m-k+1}$$

For $k = 1$, $S_1 = S$ is the Sylvester matrix. In paper [8], we know the strong relationship between the Sylvester matrix S and its k-th submatrix S_k.

Theorem 2.3. [8] *Given univariate polynomials f, $g \in \mathbb{C}[x]$, $\deg(f) = m$, $\deg(g) = n$ and $1 \le k \le \min(m, n)$. $S(f, g)$ is the Sylvester matrix of f and g, S_k is the k-th Sylvester matrix of f and g. Then the following statements are equivalent:*

(a) $\operatorname{rank}(S) \le m + n - k$

(b) *Rank deficiency of S_k is greater than or equal to one.*

Having the above theorem, the formulation (2.1) can be transformed into:

$$\min_{\deg(\mathrm{GCD}(\tilde f,\tilde g))\ge k} \|\tilde f - f\|_2^2 + \|\tilde g - g\|_2^2$$
$$\iff \min_{\dim \mathrm{Nullspace}(\tilde S_k)\ge 1} \|\tilde f - f\|_2^2 + \|\tilde g - g\|_2^2, \quad (2.2)$$

where $\tilde S_k$ is the k-th Sylvester matrix generated by $\tilde f$ and $\tilde g$, with $\deg \tilde f \le m$ and $\deg \tilde g \le n$.

If we solve the following overdetermined system using STLN [24, 23, 22, 17]

$$A_k\,\mathbf{x} \approx \mathbf{b}_k, \quad (2.3)$$

for $S_k = [\mathbf{b}_k\, A_k]$, where $\mathbf{b}_k$ is the first column of S_k and A_k are the remainder columns of S_k, then we obtain a minimal perturbation $[\mathbf{h}_k\, E_k]$ of Sylvester structure such that

$$\mathbf{b}_k + \mathbf{h}_k \in \mathrm{Range}(A_k + E_k).$$

Therefore, $\tilde S_k = [\mathbf{b}_k + \mathbf{h}_k,\ A_k + E_k]$ is a solution with Sylvester structure (provided the highest order coefficients remain non-zero) and $\dim \mathrm{Nullspace}(\tilde S_k) \ge 1$.

The reason why we choose the first column to form the overdetermined system (2.3) can be seen from the following example and theorem.

Example 1. Suppose we are given two polynomials

$$f = x^2 + x = x\,(x+1),$$
$$g = x^2 + 4x + 3 = (x+3)\,(x+1),$$

S is the Sylvester matrix of f and g,

$$S = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 1 & 4 & 1 \\ 0 & 1 & 3 & 4 \\ 0 & 0 & 0 & 3 \end{bmatrix}.$$

The rank deficiency of S is 1. We partition S in two ways: $S = [\hat{A}_1\, \hat{\mathbf{b}}_1] = [\bar{\mathbf{b}}_1\, \bar{A}_1]$, where $\hat{\mathbf{b}}_1$ is the last column of S, whereas $\bar{\mathbf{b}}_1$ is the first column of S.

The overdetermined system

$$\hat{A}_1\mathbf{x} = \hat{\mathbf{b}}_1$$

has no solution, while the system

$$\bar{A}_1\mathbf{x} = \bar{\mathbf{b}}_1$$

has an exact solution as $\mathbf{x} = [-3, 1, 0]^T$.

Theorem 2.4. *Given univariate polynomials $f, g \in \mathbb{C}[x]$ with $\deg(f) = m$, $\deg(g) = n$ and a positive integer $k \le \min(m, n)$. Suppose S_k is the k-th Sylvester matrix of f and g. Partition $S_k = [\mathbf{b}_k\, A_k]$, where $\mathbf{b}_k$ is the first column of S_k and A_k consists of the last $n + m - 2k + 1$ columns of S_k. Then we have*

$$\textit{dim Nullspace}(S_k) \ge 1 \iff A_k\mathbf{x} = \mathbf{b}_k \textit{ has a solution.} \tag{2.4}$$

Proof. "$\Longleftarrow$": Let $A_k\mathbf{x} = \mathbf{b}_k$ have a solution, then $\mathbf{b}_k \in \mathit{Range}(A_k)$. Since $\mathbf{b}_k$ is the first column of S_k, the rank deficiency of $S_k = [\mathbf{b}_k\, A_k]$ is at least 1.

"$\Longrightarrow$:" Suppose the rank deficiency of $S_k = [\mathbf{b}_k\, A_k]$ is at least 1. Multiplying the vector $[x^{n+m-k}, \ldots, x, 1]$ to the two sides of the equation $A_k\mathbf{x} = \mathbf{b}_k$, it turns out to be

$$[x^{n-k-1}f, \ldots, f, x^{m-k}g, x^{m-k-1}g, \ldots, g]\mathbf{x} = x^{n-k}f. \tag{2.5}$$

The solution $\mathbf{x}$ of (2.5) corresponds to the coefficients of polynomials u, v, with $\deg(u) \le n - k - 1, \deg(v) \le m - k$ and satisfy

$$x^{n-k}f = u\, f + v\, g.$$

Let $d = \mathrm{GCD}(f, g)$, $f_1 = f/d, g_1 = g/d$. Since dim Nullspace$(S_k) \ge 1$, we have $\deg(d) \ge k$ and $\deg(f_1) \le m - k, \deg(g_1) \le n - k$. Dividing x^{n-k} by g_1, we have a quotient q and a remainder p such that

$$x^{n-k} = q\, g_1 + p,$$

where $\deg(q) \le \deg(d) - k, \deg(p) \le n - k - 1$. Now we can check that

$$u = p,\ v = q\, f_1,$$

are solutions of (2.5), since $\deg(u) \le n - k - 1$,

$$\deg(v) = \deg(q) + \deg(f_1) \le \deg(d) - k + \deg(f_1) \le m - k,$$

and

$$v\, g + u\, f = f_1\, q\, d\, g_1 + p\, f = f\, q\, g_1 + f\, p = f\, x^{n-k}. \qquad \square$$

Next, we show that for any given Sylvester matrix, when all the elements are allowed to be perturbed, it is always possible to find matrices $[\mathbf{h}_k \, E_k]$ with k-Sylvester structure (implying that the leading entries are non-zero) such that $\mathbf{b}_k + \mathbf{h}_k \in \mathrm{Range}(A_k + E_k)$, where $\mathbf{b}_k$ is the first column of S_k and A_k are the remainder columns of S_k.

Theorem 2.5. *Given the integers m, n and k, $k \le \min(m,n)$, then there exists a Sylvester matrix $S \in \mathbb{C}^{(m+n)\times(m+n)}$ with rank $m+n-k$.*

Proof. For all m and n, we always can construct polynomials $f, g \in \mathbb{C}[x]$ such that $\deg(f) = m$, $\deg(g) = n$, and the degree of $\mathrm{GCD}(f,g)$ is k. Hence S is the Sylvester matrix generating by f, g and its rank is $m+n-k$. □

Corollary 2.6. *Given the positive integers m, n, $k \le \min(m,n)$, and k-th Sylvester matrix $S_k = [\mathbf{b}_k \, A_k]$, where $A_k \in \mathbb{C}^{(m+n-k+1)\times(m+n-2k+1)}$ and $\mathbf{b}_k \in \mathbb{C}^{(m+n-k+1)\times 1}$, it is always possible to find a perturbation $[\mathbf{h}_k \, E_k]$ of k-th Sylvester structure such that $\mathbf{b}_k + \mathbf{h}_k \in \mathrm{Range}(A_k + E_k)$.*

3. STLN for Overdetermined Systems with Sylvester Structure

In this section, we illustrate how to solve the overdetermined system

$$A_k \, \mathbf{x} \approx \mathbf{b}_k, \tag{3.1}$$

where $A_k \in \mathbb{C}^{(m+n-k+1)\times(m+n-2k+1)}$ and $\mathbf{b}_k \in \mathbb{C}^{(m+n-k+1)\times 1}$, $S_k = [\mathbf{b}_k \, A_k]$ is the k-th Sylvester matrix. According to Theorem 2.5 and Corollary 2.6, there always exists k-th Sylvester structure perturbation $[\mathbf{h}_k \, E_k]$ such that $(\mathbf{b}_k + \mathbf{h}_k) \in$ *Range*$(A_k + E_k)$. In the following, we illustrate how to find the minimum solution using STLN.

First, the Sylvester-structure preserving perturbation $[\mathbf{h}_k \, E_k]$ of S_k

$$[\mathbf{h}_k \, E_k] = \underbrace{\left[\begin{array}{ccc} z_1 & & \\ z_2 & \ddots & \\ \vdots & \ddots & z_1 \\ z_{m+1} & & z_2 \\ & \ddots & \vdots \\ & & z_{m+1} \end{array}\right.}_{n-k+1} \underbrace{\left.\begin{array}{ccc} z_{m+2} & & \\ z_{m+3} & \ddots & \\ \vdots & \ddots & z_{m+2} \\ z_{m+n+2} & & z_{m+3} \\ & \ddots & \vdots \\ & & z_{m+n+2} \end{array}\right]}_{m-k+1}$$

can be represented by a vector $\mathbf{z} \in \mathbb{C}^{(m+n+2)\times 1}$:

$$\mathbf{z} = [z_1, z_2, \ldots, z_{m+n+1}, z_{m+n+2}]^T.$$

Since $\mathbf{h}_k$ is the first column of the above matrix, we can define a matrix P_k as

$$P_k = \begin{bmatrix} \mathbf{I}_{m+1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \in \mathbb{C}^{(m+n-k+1)\times(m+n+2)}, \tag{3.2}$$

where $\mathbf{I}_{m+1}$ is a $(m+1)\times(m+1)$ identity matrix, such that $\mathbf{h}_k = P_k\mathbf{z}$.

We solve the equality-constrained least squares problem:

$$\min_{\mathbf{z},\,\mathbf{x}} \|\mathbf{z}\|_2, \text{ subject to } \mathbf{r} = 0, \tag{3.3}$$

where the structured residual $\mathbf{r}$ is

$$\mathbf{r} = \mathbf{r}(\mathbf{z},\mathbf{x}) = \mathbf{b}_k + \mathbf{h}_k - (A_k + E_k)\mathbf{x}.$$

We do not know if the above STLN problem always has a solution. But even if we that were the case (cf. [17, Theorem 3.1.2]), the optimal solution may not correspond to a nearest GCD pair, as it may correspond to polynomials of smaller degrees that, for instance, remain relatively prime. The structured minimization problem (3.3) can be solved by using the penalty method in [1], transforming (3.3) into:

$$\min_{\mathbf{z}\,,\mathbf{x}} \left\| \begin{matrix} w\mathbf{r}(\mathbf{z},\mathbf{x}) \\ \mathbf{z} \end{matrix} \right\|_2, \qquad w \gg 1, \tag{3.4}$$

where w is a large penalty value between 10^8 and 10^{10}. It is shown in [1, 27] that an algorithm based on Givens rotations produces accurate results regardless of row sorting and even with extremely large penalty values.

Following [23, 22], we use a linear approximation to $\mathbf{r}(\mathbf{z},\mathbf{x})$ to solve the minimization problem. Let $\triangle\mathbf{z}$ and $\triangle\mathbf{x}$ represent a small change in $\mathbf{z}$ and $\mathbf{x}$ respectively, and $\triangle E_k$ represents the corresponding change in E_k. Then the first order approximation to $\mathbf{r}(\mathbf{z}+\triangle\mathbf{z}\,,\mathbf{x}+\triangle\mathbf{x})$ is

$$\begin{aligned}\mathbf{r}(\mathbf{z}+\triangle\mathbf{z}\,,\mathbf{x}+\triangle\mathbf{x}) &= \mathbf{b}_k + P_k(\mathbf{z}+\triangle\mathbf{z}) - (A_k + E_k + \triangle E_k)(\mathbf{x}+\triangle\mathbf{x}) \\ &\approx \mathbf{b}_k + P_k\mathbf{z} - (A_k+E_k)\mathbf{x} + P_k\triangle\mathbf{z} - (A_k+E_k)\triangle\mathbf{x} - \triangle E_k\mathbf{x} \\ &= \mathbf{r} + P_k\triangle\mathbf{z} - (A_k+E_k)\triangle\mathbf{x} - \triangle E_k\mathbf{x}.\end{aligned}$$

We introduce a matrix of Sylvester structure $Y_k \in \mathbb{C}^{\mu\times\nu}$, where $\mu = m+n-k+1$ and $\nu = m+n+2$, such that

$$Y_k\,\triangle\mathbf{z} = \triangle E_k\,\mathbf{x} \quad \text{with} \quad \mathbf{x} = [x_1\,, x_2,\,\ldots, x_{m+n-2k+1}]^T. \tag{3.5}$$

Now (3.4) can be approximated by

$$\min_{\triangle\mathbf{x},\,\triangle\mathbf{z}} \left\| \begin{bmatrix} w(Y_k - P_k) & w(A_k+E_k) \\ \mathbf{I}_{m+n+2} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \triangle\mathbf{z} \\ \triangle\mathbf{x} \end{bmatrix} + \begin{bmatrix} -w\mathbf{r} \\ \mathbf{z} \end{bmatrix} \right\|_2. \tag{3.6}$$

In the following, we propose a new method to construct the matrix Y_k. Suppose f, g, E, $\mathbf{z}$ and $\mathbf{x}$ are given above. Multiplying the vector

$$\mathbf{v} = [x^{m+n-k},\, x^{m+n-k-1}, \ldots,\, x^2,\, x^1,\, 1] \in \mathbb{C}[x]^{m+n-k+1}$$

to the two sides of the equation (3.5), we obtain the polynomial identity

$$\mathbf{v}\,Y_k\,\mathbf{z} = \mathbf{v}\,E_k\,\mathbf{x}.$$

For $\begin{bmatrix} 0 \\ \mathbf{x} \end{bmatrix}$ we obtain

$$\mathbf{v}\,E_k\,\mathbf{x} = \mathbf{v}\,[\mathbf{h}_k, E_k]\,\hat{\mathbf{x}} = \hat{g}_1\hat{u}_1 + \hat{g}_2\hat{u}_2, \tag{3.7}$$

where $\hat{g}_1$ is the polynomial of degree m, generated by the subvector of $\mathbf{z}$:

$$[z_1,\ z_2,\ \ldots,\ z_{m+1}],$$

$\hat{g}_2$ is the polynomial of degree n, generated by the subvector of $\mathbf{z}$:

$$[z_{m+2},\ z_{m+3},\ \ldots,\ z_{m+n+2}],$$

$\hat{u}_1$ is the polynomial of degree $n-k-1$, generated by the subvector of $\hat{\mathbf{x}}$:

$$[0,\ x_1,\ x_2,\ \ldots,\ x_{n-k}],$$

$\hat{u}_2$ is the polynomial of degree $m-k$, generated by the subvector of $\hat{\mathbf{x}}$:

$$[x_{n-k+1},\ x_{n-k+2},\ \ldots,\ x_{m+n-2k+1}].$$

Y_k is the coefficient matrix formed from the above linear system (3.7) with respect to powers of x and the variables z_i,

$$Y_k = \left[\begin{array}{ccccccc} 0 & & & x_{n+1-k} & & \\ x_1 & \ddots & & x_{n+2-k} & \ddots & \\ \vdots & \ddots & 0 & \vdots & \ddots & x_{n+1-k} \\ x_{n-k} & & x_1 & x_{m+n+1-2k} & & x_{n+2-k} \\ & \ddots & \vdots & & \ddots & \vdots \\ & & x_{n-k} & & & x_{m+n+1-2k} \end{array}\right].$$
$$\underbrace{\qquad\qquad\qquad}_{m+1}\ \underbrace{\qquad\qquad\qquad\qquad}_{n+1}$$

Example 2. Suppose $m = n = 3$, $k = 2$, then $S_2 = [\mathbf{b}_2\ A_2]$, where

$$A_2 = \begin{bmatrix} 0 & b_3 & 0 \\ a_3 & b_2 & b_3 \\ a_2 & b_1 & b_2 \\ a_1 & b_0 & b_1 \\ a_0 & 0 & b_0 \end{bmatrix},\ \mathbf{b}_2 = \begin{bmatrix} a_3 \\ a_2 \\ a_1 \\ a_0 \\ 0 \end{bmatrix},$$

$$P_2 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$Y_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & x_2 & 0 & 0 & 0 \\ x_1 & 0 & 0 & 0 & x_3 & x_2 & 0 & 0 \\ 0 & x_1 & 0 & 0 & 0 & x_3 & x_2 & 0 \\ 0 & 0 & x_1 & 0 & 0 & 0 & x_3 & x_2 \\ 0 & 0 & 0 & x_1 & 0 & 0 & 0 & x_3 \end{bmatrix}.$$

It is easy to see that the coefficient matrix in (3.6) is also of block Toeplitz structure. We could apply fast least squares method to solve it quickly. Preliminary results on that are reported in [19].

4. Approximate GCD Algorithm and Experiments

The following algorithm is designed for finding an approximate solution to PROBLEM 1.1.

Algorithm AppSylv-k

Input - A Sylvester matrix S generated by two polynomials $f,\, g \in \mathbb{C}[x]$ of total degrees $m \geq n$ respectively, an integer $1 \leq k \leq n$ and a tolerance tol.

Output - Polynomials $\tilde{f}$ and $\tilde{g}$ with dim Nullspace$(S(\tilde{f}, \tilde{g})) \geq k$ and the Euclidean distance $\|\tilde{f} - f\|_2^2 + \|\tilde{g} - g\|_2^2$ is reduced to a minimum.

1. Form the k-th Sylvester matrix S_k, choose the first column of S_k as $\mathbf{b}_k$, and A_k be the remainder columns of S_k. Let $E_k = \mathbf{0}$, $\mathbf{h}_k = \mathbf{0}$.
2. Compute $\mathbf{x}$ from $\min \|A_k\mathbf{x} - \mathbf{b}_k\|_2$ and $\mathbf{r} = \mathbf{b}_k - A_k\mathbf{x}$. Form P_k and Y_k as shown in the above sections.
3. Repeat
   (a) $\min_{\triangle\mathbf{x},\, \triangle\mathbf{z}} \left\| \begin{bmatrix} w(Y_k - P_k) & w(A_k + E_k) \\ \mathbf{I}_{m+n+2} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \triangle\mathbf{z} \\ \triangle\mathbf{x} \end{bmatrix} + \begin{bmatrix} -w\mathbf{r} \\ \mathbf{z} \end{bmatrix} \right\|_2.$
   (b) Set $\mathbf{x} = \mathbf{x} + \triangle\mathbf{x}$, $\mathbf{z} = \mathbf{z} + \triangle\mathbf{z}$.
   (c) Construct the matrix E_k and $\mathbf{h}_k$ from $\mathbf{z}$, and Y_k from $\mathbf{x}$. Set $A_k = A_k + E_k$, $\mathbf{b}_k = \mathbf{b}_k + \mathbf{h}_k$, $\mathbf{r} = \mathbf{b}_k - A_k\mathbf{x}$.
   until $(\|\triangle\mathbf{x}\|_2 \leq tol$ and $\|\triangle\mathbf{z}\|_2 \leq tol)$.
4. Output the polynomials $\tilde{f}$ and $\tilde{g}$ formed from $\mathbf{b}_k$ and A_k.

Given a tolerance ϵ, the algorithm AppSylv-k can be used to compute an ϵ-GCD of polynomials f and g with degrees $m \geq n$ respectively. The method starts with $k = n \leq m$, using AppSylv-k to compute the minimum $\mathcal{N} = ||\tilde{f} - f||_2 + ||\tilde{g} - g||_2$ with rank$(S(\tilde{f}, \tilde{g})) \leq m + n - k$. If $\mathcal{N} < \epsilon$, then we can compute the ϵ-GCD from the matrix $S_k(\tilde{f}, \tilde{g})$ [8, 28]; Otherwise, we reduce k by one and repeat the AppSylv-k algorithm. Another method tests the degree of ϵ-GCD by computing the singular value decomposition of Sylvester matrix $S(f, g)$, find the upper bound degree r of the ϵ-GCD as shown in [5, 7]. So we can start with $k = r$ rather than $k = n$ to compute the certified ϵ-GCD of the highest degree.

Example 3. The following example is given in Karmarkar and Lakshman's paper [16]. We wish to find the minimal polynomial perturbations $\triangle f$ and $\triangle g$

$$f = x^2 - 6x + 5 = (x - 1)(x - 5),$$
$$g = x^2 - 6.3x + 5.72 = (x - 1.1)(x - 5.2),$$

such that polynomials $f + \triangle f$ and $g + \triangle g$ have a common root. We consider this problem in two cases: the leading coefficients can be perturbed and the leading coefficients can not be perturbed.

Case 1: The leading coefficients can be perturbed. Applying the algorithm AppSylv-k to f, g with $k = 1$ and $tol = 10^{-3}$, after three iterations, we obtain the polynomials $\tilde{f}$ and $\tilde{g}$:

$$\tilde{f} = 0.9850x^2 - 6.0029x + 4.9994,$$
$$\tilde{g} = 1.0150x^2 - 6.2971x + 5.7206,$$

with a distance

$$\mathcal{N} = \|\tilde{f} - f\|_2^2 + \|\tilde{g} - g\|_2^2 = 0.0004663.$$

The common root of the $\tilde{f}$ and $\tilde{g}$ is 5.09890429.

Case 2: The leading coefficients can not be perturbed, i.e., the first and fourth terms of $\mathbf{q}$ are fixed as one. Running the algorithm AppSylv-k for $k = 1$ and $tol = 10^{-3}$, after three iterations, we get the polynomials $\tilde{f}$ and $\tilde{g}$:

$$\tilde{f} = x^2 - 6.0750x + 4.9853,$$
$$\tilde{g} = x^2 - 6.2222x + 5.7353,$$

with minimum distance as

$$\mathcal{N} = \|\tilde{f} - f\|_2^2 + \|\tilde{g} - g\|_2^2 = 0.01213604583.$$

The common root of $\tilde{f}$ and $\tilde{g}$ is 5.0969478.

In the paper [16] the perturbed polynomials are restricted to be monic. The minimum perturbation Karmarkar and Lakshman obtained is 0.01213605293, which corresponds to the perturbed common root 5.096939087.

Ex.	m, n	k	*it. (Chu)*	*it. (STLN)*	*error (Zeng)*	*error (STLN)*	σ_k	$\tilde{\sigma}_k$
1	2, 2	1	4.73	1.90	1.89*e*–4	2.87*e*–5	3.53*e*–3	10^{-9}
2	3, 3	2	8.49	1.98	1.36*e*–3	1.05*e*–4	8.21*e*–3	10^{-9}
3	5, 4	3	11.44	2.00	1.00*e*–3	1.25*e*–4	1.01*e*–2	10^{-9}
4	5, 5	3	13.64	2.00	7.43*e*–4	1.25*e*–4	9.57*e*–3	10^{-9}
5	6, 6	4	23.07	2.00	1.46*e*–3	1.41*e*–4	9.64*e*–3	10^{-9}
6	8, 7	4	32.64	2.00	6.53*e*–4	1.31*e*–4	8.04*e*–3	10^{-9}
7	10, 10	5	43.12	2.00	1.61*e*–3	2.01*e*–4	1.21*e*–2	10^{-9}
8	14, 13	7	58.16	2.00	1.23*e*–3	2.52*e*–4	1.51*e*–2	10^{-9}
9	28, 28	10	161.74	2.00	2.61*e*–3	3.41*e*–4	1.48*e*–2	10^{-10}
10	65, 65	15	633.64	2.00	6.19*e*–3	5.50*e*–4	1.90*e*–2	10^{-9}

TABLE 1. Algorithm performance on benchmarks (univariate case)

Remark 4.1. The above algorithm will for real inputs compute the optimal pair over the reals only. As stated in remark 2.2 even for real inputs the optimal complex pair may have complex coefficients. The following change in the initialization in Step 2 can accomplish that:

$2^{\mathbb{C}}$. Compute $\mathbf{x}$ from $\min \|(A_k+\delta A_k)\mathbf{x}-(\mathbf{b}_k+\delta\mathbf{b}_k\|_2$ and $\mathbf{r} = \mathbf{b}_k+\delta\mathbf{b}_k-A_k-\delta A_k\mathbf{x}$, where δA_k and $\delta\mathbf{b}_k$ are structured perturbations of random purely imaginary complex noise. Form P_k and Y_k as shown in the above sections.

For the first iteration in step 3 we use the original real input coefficients. The optimal complex solutions in remark 2.2 could be found by adding random noise of magnitute 10^{-2}.

We have also tested our algorithm on inputs where both polynomials have small leading coefficients and observed that the method can produce valid results provided the "tails," of the polynomials, i.e., the parts without the leading coefficients, are approximately relative prime. We present an example.

$$f = .1000000000 \cdot 10^{-9}x^2 + x,$$
$$g = .1000000000 \cdot 10^{-9}x^2 + x + 1.$$

The roots of f are -10^{10}, 0, the roots of g are -9999999999, -1. We compute two polynomials $\bar{f}$ and $\bar{g}$ by our algorithm:

$$\bar{f} = .100000000005000000 \cdot 10^{-9}x^2 + 1.x,$$
$$\bar{g} = .99999999994999994 \cdot 10^{-10}x^2 + 1.x + 1.$$

The roots of $\bar{f}$ are now -9999999999.5, 0 and the roots of $\bar{g}$ are now -9999999999.5, -1. The perturbation by our algorithm is

$$\|\bar{f} - f\|_2^2 + \|\bar{g} - g\|_2^2 = .500000000000000 \cdot 10^{-40}.$$

However, if the tails have a nearby GCD, the algorithm as stated does not find a good result due to the choice of $\mathbf{b}_k$ as the first column of the Sylvester matrix. That problem appears in a more general manner when applying our approach to multivariate approximate GCDs when some terms of maximal total degree can vanish in the nearest pair. The remedy is to determine the proper right side vector from the components of the first singular vector; see [14] for more detail. □

In Table 1, we show the performance of our algorithm for computing ϵ-GCDs of univariate polynomials randomly generated in Maple 9 with *Digits* = 10. For every example, we use 50 random cases for each (m, n), and report the average over all results. For each example, the prime parts and GCD of two polynomials are constructed by choosing polynomials with random integer coefficients in the range $-10 \le c \le 10$, and then adding a perturbation. For noise we choose a relative tolerance 10^{-e}, then randomly choose a polynomial that has the same degree as the product, and coefficients in $[-10^e, 10^e]$. Finally, we scale the perturbation so that the relative error is 10^{-e}. In our test cases we set $e = 2$. Here m and n denote the total degrees of polynomials f and g; k is the degree of approximate GCD of f and g; *it. (Chu)* is the number of the iterations needed by Chu's method[4]; whereas *it. (STLN)* denotes the number of iterations by AppSylv-k algorithm; *error (Zeng)* denotes the perturbation $\|\bar{f} - f\|_2^2 + \|\bar{g} - g\|_2^2$ computed by Zeng's algorithm [28]; whereas *error (STLN)* is the minimal perturbation $\|\tilde{f} - f\|_2^2 + \|\tilde{g} - g\|_2^2$ computed by AppSylv-k algorithm; σ_k and $\tilde{\sigma}_k$ are the last k-th singular values of $S(f, g)$ and

$S(\tilde{f}, \tilde{g})$, respectively. Riemannian SVD has been considered in [?] for computing approximate GCDs. We would like to compare with their implementation in the future.

5. Concluding Remarks

In this paper we present a practical and reliable way based on STLN to compute the approximate GCD of univariate polynomials. Note that the overall computational complexity of the algorithm AppSylv-k depends on the number of iterations needed for completing the Step 3 and the computational complexity of each iteration. If the starting values are good, then the iteration will converge quickly. This can be seen from the above table. For each iteration, Givens rotations are applied. Checking the size of the matrix involved in solving the minimization problem in the Step 3, we obtain that each iteration needs less than $(4\,m + 4\,n - k + 6)\,(2\,m + 2\,n - 2\,k + 3)^2$ operations. Since the matrices involved in the minimization problems are all structured matrix, they have low displacement rank [11]. It would be possible to apply the fast algorithm to solve these minimization problems [19]. This would reduce the complexity of our algorithm to be only quadratic with respect to the degrees of the given polynomials.

Our methods can be generalized to several polynomials and to several variables. Moreover, as observed in [9], arbitrary linear equational constraints can be imposed on the coefficients of the input and perturbed polynomials. Such constraints can be used to preserve monicity and sparsity, but also relations among the input polynomials' coefficients.

Acknowledgement

We thank the referees of an earlier version of this paper for their helpful remarks.

References

[1] A. A. Anda and H. Park, *Fast plane with dynamic scaling*, SIAM J. Matrix Anal. Applic., 15 (1994), pp. 162–174.

[2] B. Beckermann and G. Labahn, *When are two polynomials relatively prime?*, J. Symbolic Comput., (1998), pp. 677–689.

[3] B. Botting, M. Giesbrecht, and J. May, *Using Riemannian SVD for problems in approximate algebra*, in Proc. International Workshop on Symbolic-Numeric Computation, 2005, pp. 209–219.

[4] M. T. Chu, R. E. Funderlic, and R. J. Plemmons, *Structured low rank approximation*, Linear Algebra and Applications, 366 (2003), pp. 157–172.

[5] R. M. Corless, P. M. Gianni, B. M. Trager, and S. M. Watt, *The singular value decomposition for polynomial systems*, in Proc. 1995 Internat. Symp. Symbolic Algebraic Comput. ISSAC'95, A. H. M. Levelt, ed., New York, N. Y., 1995, ACM Press, pp. 96–103.

[6] R. M. Corless, S. M. Watt, and L. Zhi, *QR factoring to compute the GCD of univariate approximate polynomials*, IEEE Transactions on Signal Processing, 52 (2004), pp. 3394–3402.

[7] I. Z. Emiris, A. Galligo, and H. Lombardi, *Certified approximate univariate GCDs*, J. Pure Applied Algebra., 117 & 118 (1996), pp. 229–251. Special Issue on Algorithms for Algebra.

[8] S. Gao, E. Kaltofen, J. P. May, Z. Yang, and L. Zhi, *Approximate factorization of multivariate polynomials via differential equations*, in ISSAC 2004 Proc. 2004 Internat. Symp. Symbolic Algebraic Comput., J. Gutierrez, ed., New York, N. Y., 2004, ACM Press, pp. 167–174.

[9] M. A. Hitz and E. Kaltofen, *Efficient algorithms for computing the nearest polynomial with constrained roots*, in Proc. 1998 Internat. Symp. Symbolic Algebraic Comput. (ISSAC'98), O. Gloor, ed., New York, N. Y., 1998, ACM Press, pp. 236–243.

[10] V. Hribernig and H. J. Stetter, *Detection and validation of clusters of polynomials zeros*, J. Symbolic Comput., 24 (1997), pp. 667–681.

[11] T. Kailath and A. H. Sayed, *Displacement structure: theory and applications*, SIAM Review, 37 (1995), pp. 297–386.

[12] E. Kaltofen, *Polynomial factorization 1987-1991*, in Proc. LATIN '92, I. Simon, ed., vol. 583 of Lect. Notes Comput. Sci., Heidelberg, Germany, 1992, Springer Verlag, pp. 294–313.

[13] E. Kaltofen and J. May, *On approximate irreducibility of polynomials in several variables*, in ISSAC 2003 Proc. 2003 Internat. Symp. Symbolic Algebraic Comput., J. R. Sendra, ed., New York, N. Y., 2003, ACM Press, pp. 161–168.

[14] E. Kaltofen, Z. Yang, and L. Zhi, *Structured low rank approximation of a generalized Sylvester matrix*. Manuscript, 2005. 14 pages. Abstract accepted at the Asian Symposium on Computer Mathematics ASCM 2005.

[15] N. Karmarkar and Lakshman Y. N., *Approximate polynomial greatest common divisors and nearest singular polynomials*, in Proc. 1996 Internat. Symp. Symbolic Algebraic Comput. (ISSAC'96), Lakshman Y. N., ed., New York, N. Y., 1996, ACM Press, pp. 35–42.

[16] N. K. Karmarkar and Lakshman Y. N., *On approximate GCDs of univariate polynomials*, J. Symbolic Comput., 26 (1998), pp. 653–666. Special issue of the JSC on Symbolic Numeric Algebra for Polynomials S. M. Watt and H. J. Stetter, editors.

[17] P. Lemmerling, *Structured total least squares: analysis, algorithms and applications*, dissertation, Katholieke Universiteit Leuven, Belgium, 1999.

[18] B. Li, Z. Liu, and L. Zhi, *Fast implementation of low rank approximation of a Sylvester matrix*, in Proc. International Workshop on Symbolic-Numeric Computation, 2005, pp. 202–208.

[19] B. Li, Z. Yang, and L. Zhi, *Fast low rank approximation of a Sylvester matrix by structured total least norm*, J. JSSAC (Japan Society for Symbolic and Algebraic Computation), 11 (2005), pp. 165–174.

[20] M. T. Noda and T. Sasaki, *Approximate GCD and its application to ill-conditioned algebraic equations*, J. Comput. Appl. Math., 38 (1991), pp. 335–351.

[21] V. Y. Pan, *Numerical computation of a polynomial GCD and extensions*, Information and computation, 167 (2001), pp. 71–85.

[22] H. Park, L. Zhang, and J. B. Rosen, *Low rank approximation of a Hankel matrix by structured total least norm*, BIT, 39 (1999), pp. 757–779.

[23] J. B. Rosen, H. Park, and J. Glick, *Total least norm formulation and solution for structured problems*, SIAM J. Matrix Anal. Applic., 17 (1996), pp. 110–128.

[24] ———, *Structured total least norm for nonlinear problems*, SIAM J. Matrix Anal. Applic., 20 (1999), pp. 14–30.

[25] M. Sasaki and T. Sasaki, *Polynomial remaider sequence and approximate GCD*, ACM SIGSAM Bulletin, 31 (2001), pp. 4–10.

[26] A. Schönhage, *Quasi-gcd computations*, Journal of Complexity, 1 (1985), pp. 118–137.

[27] G. W. Stewart, *On the asymptotic behavior of scaled singular value and QR decompositions*, Mathematics of Compuatations, 43 (1983), pp. 488–489.

[28] Z. Zeng, *The approximate GCD of inexact polynomials. part I: a univariate algorithm.* Manuscript, 2004.

[29] Z. Zeng, *Computing multiple roots of inexact polynomials*, Math. Comput., 74 (2005), pp. 869–903.

[30] L. Zhi, *Displacement structure in computing approximate GCD of univariate polynomials*, in Proc. Sixth Asian Symposium on Computer Mathematics (ASCM 2003), Z. Li and W. Sit, eds., vol. 10 of Lecture Notes Series on Computing, Singapore, 2003, World Scientific, pp. 288–298.

[31] L. Zhi and W. Wu, *Nearest singular polynomial*, J. Symbolic Comput., 26 (1998), pp. 667–675. Special issue on Symbolic Numeric Algebra for Polynomials S. M. Watt and H. J. Stetter, editors.

Erich Kaltofen
Dept. of Mathematics
North Carolina State University
Raleigh, North Carolina 27695-8205, USA
e-mail: `kaltofen@math.ncsu.edu`

Zhengfeng Yang and Lihong Zhi
Key Lab of Mathematics Mechanization
AMSS, Beijing 10080 China
e-mail: {`zyang, lzhi`}`@mmrc.iss.ac.cn`