APPRIOPXIMATE GREATEST COMMON DIVISOR OF MANY POLYNOMIALS AND GENERALISED RESULTSANTs

N. Karcanias*, S. Fatouros*, M. Mitrouli† and G. Halikias*

* Control Engineering Centre, School of Engineering, City University, Northampton Square, London, ECI0HB, UK., email: n.karcanias@city.ac.uk
† Department of Mathematics, University of Athens, Panepistemiopolis 15784, Athens, Greece, email:mmitroul@cc.uoa.gr

Keywords: approximate, computations, greatest common divisor

Abstract

In this paper, a new characterisation of the approximate GCD of many polynomials is given that also allows the evaluation of accuracy of the corresponding ‘approximate GCD computation’. This new approach is based on recent results on the factorisation of the generalised resultant of a set of polynomials into reduced resultants and appropriate Toeplitz matrices representing the exact GCD [1]. This allows the reduction of ‘approximate GCD’ computation to an equivalent ‘approximate factorisation’ of generalised resultants. This new approach may be formulated as a structured optimization problem (distance between structured matrices). We use this new framework to evaluate the ‘accuracy’ of the ‘approximate GCD’ of a certain degree. This evaluation is equivalent to finding the minimal perturbation on the original set of polynomials, which makes the selected given degree ‘approximate GCD’ exact for the perturbed set. The later makes precise the meaning of approximate GCD, since it relates it to the exact notion on a perturbed set.

1 Generalised Resultant, GCD and the GCD representation problem

We consider the set of polynomials $\mathcal{P}_{h+1,n} = \{p_i(s) \in \mathbb{R}[s], i = 0, 1, \ldots, h\}$ where we assume that $\partial[p_i(s)] = n$ and $p_0(s)$ monic and the elements are represented as:

$$p_0(s) = s^n + a_{n-1}s^{n-1} + \cdots + a_1s + a_0$$

$$p_i(s) = b_{i,n}s^n + \cdots + b_{i,1}s + b_{i,0}, \quad i = 1, 2, \ldots, h$$

With this notation we can introduce the generalised resultant of the set $\mathcal{P}_{h+1,n}$ as:

$$p_0(s)$$

$$S_0 = \begin{bmatrix}
1 & a_{n-1} & a_{n-2} & \cdots & a_1 & a_0 & 0 & \cdots & 0 \\
0 & 1 & a_{n-1} & \cdots & a_1 & a_0 & \cdots & \cdots & \cdots \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \cdots & \cdots & \cdots \\
0 & \cdots & 0 & 1 & a_{n-1} & \cdots & a_1 & a_0
\end{bmatrix} \quad (2a)$$

and a $n \times (n + m)$ matrix associated with $p_i(s), i = 1, 2, \ldots, h$ as:

$$S_i = \begin{bmatrix}
b_{i,m} & b_{i,m-1} & \cdots & b_{i,1} & b_{i,0} & 0 \cdots & 0 \\
0 & b_{i,m} & b_{i,m-1} & \cdots & b_{i,1} & b_{i,0} & \cdots & \cdots & \cdots \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \cdots & \cdots & \cdots \\
0 & \cdots & 0 & b_{i,m} & b_{i,m-1} & \cdots & b_{i,1} & b_{i,0}
\end{bmatrix} \quad (2b)$$

for every $i = 1, 2, \ldots, h$. For the set $\mathcal{P}_{h+1,n}$ we define the extended Sylvester matrix, or generalised resultant the matrix

$$S_{\mu} = \begin{bmatrix}
S_0 \\
S_1 \\
\vdots \\
S_h
\end{bmatrix} \in \mathbb{R}^{n \times (n + m)}, \quad \mu = m + nh \quad (2c)$$

The properties of the greatest common divisor (GCD) of $\mathcal{P}_{h+1,n}$, which shall be denoted as $\text{GCD}\{\mathcal{P}_{h+1,n}\}$ are linked to the properties of the generalised resultant and are summarised by the following results:

**Theorem 1** Consider the set of polynomials $\mathcal{P}_{h+1,n} = \{p_i(s), i = 0, 1, \ldots, h\}$ as described before, let $S_{\mu}$ be the generalised resultant and $\phi(s) = \text{GCD}\{\mathcal{P}_{h+1,n}\}$. The following properties hold true:

(i) $\text{rank}(S_{\mu}) = n + m - \deg \phi(s)$ \quad (3)

(ii) The set $\mathcal{P}_{h+1,n}$ is coprime if and only if $\text{rank}(S_{\mu}) = n + m$.

(iii) The GCD of $\mathcal{P}_{h+1,n}$ is invariant under elementary row operations on $S_{\mu}$. Furthermore, if we reduce $S_{\mu}$ to its row echelon form, then the last nonvanishing row defines the coefficients of the gcd.
The above result suggest clearly a procedure for computing the “approximate $\text{GCD}$” based on the use of Gaussian transformations with partial pivoting (part of the ERES method \cite{6}, \cite{7}, without the use of shifting), which reduce $S_P$ to a triangular form. We may use the notion of numerical rank for the computation of degree of the $\text{GCD}$ and then evaluate if from the last row.

\textbf{Note 1} ERES method reduces computation of $\text{GCD}$ to a best approximation of a given “almost rank $1$” matrix. A method that starts on the resultant and computes the approximate $\text{GCD}$ on the basis of triangularisation has to answer the question of determining the “best” triangularisation, given that there is no unique way of reducing a matrix to a triangular form. This problem requires some investigation, since different triangularisation will produce different “approximate $\text{GCD}$”.

\textbf{Note 2} An alternative method for computing the $\text{GCD}$ (and its approximate evaluation) may be based on the Matrix Pencil Algorithm \cite{4} applied on the generalized Sylvester \cite{5}. This may be improved by using a framework based on the numerical range.

Clearly, those two procedures require farther investigation and we will come back to this. We consider next same important property that can be used for the evaluation of some error that may be associated with the computation of “almost”, “approximate $\text{GCD}$”. This is based on result characterising factorisation of resultants \cite{1}.

\textbf{Theorem 2} Let $\mathcal{P}_{h+1,n}$ be a set of polynomials with the two maximal degree values $(n,m)$, $S_P$ be the corresponding Sylvester matrix and let $\text{GCD}(\mathcal{P}_{h+1,n}) = \phi(s) = \lambda_k s^k + \cdots + \lambda_1 s + \lambda_0$, $\lambda_k \neq 0$. Then, there exists transformation $\phi \in \mathcal{R}^{(n+m) \times (n+m)}$ such that

$$S_P^{(k)} = S_P \cdot \Phi$$

\begin{align}
\begin{bmatrix}
O & S_0^{(k)} \\
oalign{\vspace{1em}}
O & S_1^{(k)} \\
oalign{\vspace{1em}}
\vdots & \\
oalign{\vspace{1em}}
O & S_h^{(k)} \\
\end{bmatrix} = \begin{bmatrix} y_0 & 0 & \cdots & \cdots & 0 \\
y_1 & y_0 & \cdots & \cdots & \vdots \\
y_2 & y_1 & \cdots & \cdots & \vdots \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
y_{n+m-2} & y_{n+m-3} & \cdots & y_0 & 0 \\
y_{n+m-1} & y_{n+m-2} & \cdots & y_1 & y_0 \\
\end{bmatrix} = \Phi
\end{align}

(5a)

and for $i = \ldots, h$

$$S_i^{(k)} = \begin{bmatrix} \phi_{b_{i,m-k}}^{(k)} & \cdots & b_i^{(k)} & 0 & \cdots & 0 \\
O & \phi_{b_{i,m-k}}^{(k)} & \cdots & b_i^{(k)} & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
O & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\
\end{bmatrix}$$

(5b)

where

$$y_0 = 1/\lambda_0, \ y_i = -\lambda_1/\lambda_0 y_0, \ldots, \ y_{n+m-1} = \lambda_{h}/\lambda_0 y_0$$

(5c)

$$y_j = -1/k_0 \sum_{i=1}^{\min(j,k)} k_j y_{j-i}, \ j = 2, \ldots, n+m-1$$

(5d)

and

$$a_i^{(k)} = \{a_{m-k}^{(k)}, \ldots, a_0^{(k)}\}, \ b_j^{(k)} = \{b_{j,m-k}^{(k)}, \ldots, b_{j,0}^{(k)}\}$$

are the coprime polynomials obtained from the original set after division of $P_i(s)$ by the $\text{gcd}$. The matrix $\Phi$ has all information associated with the $\text{GCD}$ and this is established by the following results \cite{1}.

\textbf{Lemma 1} Let $\mathcal{T}_n = \{A_i : A_i \in \mathcal{R}^{n \times n}\}$, where $A_i$ are Toeplitz matrices of the type

$$A_i = \begin{bmatrix} a_0^i & 0 & 0 & \cdots & 0 \\
a_1^i & a_0^i & 0 & \cdots & \vdots \\
a_2^i & a_1^i & a_0^i & \cdots & \vdots \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & \cdots & \cdots & a_0^i \\
\end{bmatrix}$$

(6)

The set $\mathcal{T}_n$ is a multiplicative group and for every $A \in \mathcal{T}_n$ with $a_0 \neq 0$ there is an inverse in $\mathcal{T}_n$ of the form

$$A^{-1} = \begin{bmatrix} \ell_0 & 0 & 0 & \cdots & 0 \\
\ell_1 & \ell_0 & 0 & \cdots & \vdots \\
\ell_2 & \ell_1 & \ell_0 & \cdots & \vdots \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
\ell_{n-2} & \ell_{n-3} & \cdots & \ell_0 & 0 \\
\ell_{n-1} & \ell_{n-2} & \cdots & \ell_1 & \ell_0 \\
\end{bmatrix}$$

(7a)
where
\[ \ell_0 = 1/a_0, \quad \ell_1 = -1/a_0 \sum_{j=0}^{i-1} \ell_j a_{i-j}. \] (7b)

Using the above result we have:

**Corollary 1**  Let \( \phi(s) = \lambda_k s^k + \cdots + \lambda_1 s + \lambda_0, \lambda_k \neq 0 \) be the \( \text{GCD} \) of \( P_{h+1,n} \) and \( \Phi \) the matrix in the factorisation (4). Then
\[
\hat{\Phi} = \Phi^{-1} = \begin{bmatrix}
\lambda_0 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\
\lambda_1 & \lambda_0 & \cdots & \cdots & \cdots & \cdots & \vdots \\
\lambda_2 & \lambda_1 & \lambda_0 & \cdots & \cdots & \cdots & \vdots \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
\lambda_k & \lambda_{k-1} & \lambda_{k-2} & \cdots & \lambda_1 & \lambda_0 & 0 \\
0 & \lambda_k & \lambda_{k-1} & \cdots & \lambda_2 & \lambda_1 & 0 \\
0 & 0 & \lambda_v & \lambda_{v-1} & \lambda_2 & \lambda_1 & 0 \\
\end{bmatrix}
\tag{8}
\]

and thus
\[
S_P = S^{(k)}_{\hat{\Phi}} \tag{9}
\]

Theorem 2 and Corollary 1 establish a very important property of factorisation of resultants. In fact, the Sylvester matrix \( S_P \) can be factored as a product of a reduced resultant matrix that consists of the coefficients of the coprime polynomials, which are obtained by division of the original set by the \( \text{GCD} \). The extraction of the Toeplitz matrix \( \hat{\Phi} \) is equivalent to factorisation of \( \text{GCD} \) from the original set of polynomials and this equation (9) is a matrix representation of the classical factorisation of the original set of polynomials. We may summarise this discussion as a restatement of the last result as:

**Corollary 2** Let \( P_{h+1,n} = \{p_i(s), i = 0, 1, \ldots, h\} \) be a set of polynomials with the two maximal degrees \( (n, m; n \geq m) \) and let \( \phi(s) = \text{GCD}\{P_{h+1,n}\} \) with \( \deg \phi(s) = k \). If
\[ p_i(s) = p_i^{(k)}(s) \cdot \phi(s), \forall i \in h+1 \text{ and } P_{h+1,n}^{(k)} = \{p_i^{(k)}(s), i = 0, 1, \ldots, h\} \]
and denote this factorisation as
\[
P_{h+1,n} = P_{h+1,n}^{(k)} \cdot \phi(s) \tag{10}
\]
then a matrix representation of the above algebraic factorisation is defined by
\[
S_P = S^{(k)}_{\hat{\Phi}} \cdot \hat{\Phi} \tag{11}
\]
where \( S_P, S^{(k)}_{\hat{\Phi}} \) are the resultant representations of \( P_{h+1,n} \), \( P_{h+1,n}^{(k)} \) and \( \hat{\Phi} \) is the Toeplitz representation of \( \phi(s) \).

**Note 3** In the description of \( P_{h+1,n} \) set we assume that the first \( k \) columns are zero that is \( S^{(k)}_{\hat{\Phi}} = [O_k, \tilde{S}^{(k)}_{\hat{\Phi}}] \) and \( \tilde{S}^{(k)}_{\hat{\Phi}} \) corresponds to the coprime \( P_{h+1,n}^{(k)} \) set; clearly, factorisation like (11) way be also established for factors, which are not necessarily the \( \text{GCD} \). In fact, algebraic factorisation may be represented as Sylvester matrices factorisations with Toeplitz factors.

The representation of the factorisation (10) of a set of polynomials as a factorisation of associated resultants in the form (4), or (11) allows the development of means for properly defining the notion of “approximate \( \text{GCD} \)” as well as quantifying how “good” such an approximation is. The following analysis also allows development of alternative means for computing “almost zeros”.

## 2 Algorithm for evaluating the strength of approximation of a given approximate \( \text{GCD} \)

Given the set \( P_{h+1,n} \) of polynomials with ordered degree set
\[ D = \{d_i, i = 0, 1, \ldots, h : d_0 = n \geq d_1 = m > \cdots \geq d_h\} \]
Let \( S_P \in \Psi_D \) be the generalised resultant representation of \( P_{h+1,n} \) and \( \Psi_D \) the set of all resultants parametrical by \( D \). Let us assume that \( \omega(s) = \lambda_k s^k + \cdots + \lambda_1 s + \lambda_0 \) is an approximate \( \text{GCD} \), where \( k \leq d_h \) and let \( \hat{\Phi} \) be the Toeplitz representation
\[
\hat{\Phi}_{\omega} = \Phi^{-1}_{\omega} = \begin{bmatrix}
\lambda_0 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\
\lambda_1 & \lambda_0 & \cdots & \cdots & \cdots & \cdots & \vdots \\
\lambda_2 & \lambda_1 & \lambda_0 & \cdots & \cdots & \cdots & \vdots \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
\lambda_k & \lambda_{k-1} & \lambda_{k-2} & \cdots & \lambda_1 & \lambda_0 & 0 \\
0 & \lambda_k & \lambda_{k-1} & \cdots & \lambda_2 & \lambda_1 & 0 \\
0 & 0 & \lambda_v & \lambda_{v-1} & \lambda_2 & \lambda_1 & 0 \\
\end{bmatrix}
\tag{12}
\]

**Step 1** Define the family of all structured dynamic perturbations
\[ Q_{h+1,n} = \{q_i(s), i = 0, 1, \ldots, h, \partial\{q_i(s)\} = \sigma \leq d_i, \forall i = 0, 1, \ldots, h\} \]
with the same degree set \( D \) as \( P_{h+1,n} \), which has \( \phi(s) \) as an exact \( \text{GCD} \). The set of all such dynamic perturbations \( Q_{h+1,n} \) has a Sylvester representation:
\[
S_Q = S_P - [O_n, \hat{S}^{(k)}_{\hat{\Phi}_{\omega}}] \hat{\Phi}_{\omega} = S_P - S^{(k)}_{\hat{\Phi}_{\omega}} \tag{13}
\]
where \( S^{(k)} \) is a resultant from \( \Psi_D \) set with the first \( k \)-columns zero and the rest of the elements arbitrary (see (5a), with \( a_i^{(k)}, i = 0, 1, \ldots, n-k \) and \( u_j^{(k)}, j = 1, \ldots, h, i = 0, 1, \ldots, m-h \) arbitrary parameters.
Step 2 Using as parameters those free variables in $S_p^{(k)} = S_w$
define the function

$$f(P, W) = \| S_p - [0_k, S_W] \Phi_w \|_F$$

and solve the minimisation problem

$$\min_W f(P, W)$$

where $w$ denotes the composite vector of all parameters in $S_W$.

Example

We consider the set of polynomials

$$\mathcal{P}_{2,2} = \{ p_0(s) = (s-1)(s-2) = s^2 + 3s + 2, \quad p_1(s) = s - 0.99999 \}$$

Then we have: $n = 2, \quad m = 1$ and

$$S_p = \begin{bmatrix} 1 & -3 & 2 \\ 1 & -0.99999 & 0 \\ 0 & 1 & -0.99999 \end{bmatrix}$$

An approximate GCD of the set using ERES method is: $\phi(s) = s - 1$. Then

$$\hat{\Phi} = \begin{bmatrix} -1 & 0 & 0 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix}$$

and

$$S_W = \begin{bmatrix} a & b \\ c & 0 \\ 0 & c \end{bmatrix}$$

where $w = (a, b, c)^t$. $f(P, W) = \| S_p - [0_k, S_W] \Phi_w \|_F$

Solving the minimization problem with MATLAB we get:

$$a = 1.00008737833788, \quad b = -1.99996490822102, \quad c = 0.999967283843.$$ 

This set of values attains a minimum of order $0(10^{-8})$. Thus, the approximate GCD that we computed, becomes an exact GCD of the set derived from the original data if we add perturbations of order $0(10^{-8})$.

References


