APPROXIMATE GREATEST COMMON DIVISOR OF MANY POLYNOMIALS AND GENERALISED RESULTANTS

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

* Control Engineering Centre, School of Engineering, City University, Northampton Square, London, EC1V0HB, 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 some 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 make 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 $\mathcal{P}_{h+1,h} = \{p_i(s) \in \mathcal{R}[s], i = 0,1,\ldots,h\}$ where we assume that $\partial[p_0(s)] = n$ and $p_0(s)$ monic and the elements are represented as:

$$\begin{aligned} p_0(s) &= s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0 \\ p_i(s) &= b_{i,n}s^n + \cdots\cdots + b_{i,1}s + b_{i,0}, \qquad i = 1,2,\ldots,h \end{aligned} \tag{1}$$

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

Definition 1 *If we assume that* $\max\{\partial[p_i(s)], i = 1,2,\ldots,h\} = n$, $m \leq n$, *i.e.* $b_{k,m} \neq 0$ *for at least one value of* k, *but* $b_{k,j} = 0$ *for all* $j = m+1,\ldots,n$ *and all* k, *then we can define an* $m \times (n+m)$ *matrix associated with* $p_0(s)$ *as:*

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

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

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

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

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

The properties of the greatest common divisor (GCD) *of* $\mathcal{P}_{h+1,n}$, *which shall be denoted as* $GCD\{\mathcal{P}_{h+1,n}\}$ *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* $S_{\mathcal{P}}$ *be the generalized resultant and* $\phi(s) = GCD\{\mathcal{P}_{h+1,n}\}$. *The following properties hold true:*

(i) $$rank(S_{\mathcal{P}}) = n + m - \deg \phi(s) \tag{3}$$

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

(iii) The GCD *of* $\mathcal{P}_{h+1,n}$ *is invariant under elementary row operations on* $S_{\mathcal{P}}$. *Furthermore, if we reduce* $S_{\mathcal{P}}$ *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 GCD" based on the use of Gaussian transformations with partial pivoting (part of the ERES method [6], [7], without the use of shifting), which reduce $S_{\mathcal{P}}$ to a triangular form. We may use the notion of numerical rank for the computation of degree of the GCD and then evaluate if from the last row.

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

Note 2 *An alternative method for computing the GCD (and its approximate evaluation) may be based on the Matrix Pencil Algorithm [4] applied on the generalized Sylvester [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 GCD". This is based on result characterising factorisation of resultants [1].

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

$$S_{\mathcal{P}}^{(k)} = S_{\mathcal{P}} \cdot \Phi \tag{4}$$

where $S_{\mathcal{P}}^{(k)}$ and Φ are given by:

$$S_{\mathcal{P}}^{(k)} = \begin{bmatrix} \mathcal{O} & S_0^{(k)} \\ \mathcal{O} & S_1^{(k)} \\ & \vdots \\ \mathcal{O} & S_h^{(k)} \end{bmatrix} = [O \;\; \tilde{S}_p^{(k)}] \tag{5a}$$

and

$$S_0^{(k)} = \begin{bmatrix} \mathcal{O} a_{n-k}^{(k)} a_{n-k-1}^{(k)} \cdots a_1^{(k)} \; a_0^{(k)} & 0 & \cdots & 0 \\ \mathcal{O} \;\; 1 \;\; a_{n-k}^{(k)} \;\; \cdots \;\; \cdots \;\; a_1^{(k)} a_0^{(k)} & \ddots & & \vdots \\ \vdots \qquad \ddots \qquad \ddots & & \ddots & 0 \\ \mathcal{O} \;\; \cdots \;\; 0 \;\; 1 \; a_{n-k}^{(k)} \cdots \cdots a_1^{(k)} a_0^{(k)} & & & \end{bmatrix}$$

and for $i = \ldots, h$

$$S_i^{(k)} = \begin{bmatrix} \mathcal{O} b_{i,m-k}^{(k)} b_{i,m-k-1}^{(k)} & \cdots & b_{i,1}^{(k)} & b_{i,0}^{(k)} & 0 & \cdots\cdots & 0 \\ \mathcal{O} b_{i,m-k}^{(k)} b_{i,m-k-1}^{(k)} & \cdots & \cdots & b_{i,1}^{(k)} b_{i,0}^{(k)} & 0 & \cdots & 0 \\ \vdots \quad \ddots \quad \ddots & & & & \ddots & \ddots & \vdots \\ \vdots \quad \ddots \quad \ddots & \ddots & & & & \ddots \; \ddots & 0 \\ \mathcal{O} \;\; \cdots \;\; 0 & b_{i,m-k}^{(k)} b_{i,m-k-1}^{(k)} & \cdots\cdots & & & b_{i,1}^{(k)} b_{i,0}^{(k)} & \end{bmatrix}$$

$$\Phi = \begin{bmatrix} y_0 & 0 & \cdots & \cdots & \cdots & 0 \\ y_1 & y_0 & \ddots & & & \vdots \\ y_2 & y_1 & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ y_{n+m-2} & y_{n+m-3} & \cdots & \ddots & y_0 & 0 \\ y_{n+m-1} & y_{n+m-2} & \cdots & \cdots & y_1 & y_0 \end{bmatrix} \tag{5b}$$

where

$$y_0 = 1/\lambda_0, \quad , y_1 = -\lambda_1/\lambda_0 y_0, \ldots,$$

$$y_j = -1/k_0 \sum_{i=1}^{\min(j,k)} k_i y_{j-i}, \; j = 2, \ldots n+m-1 \tag{5c}$$

and

$$\{a_{m-k}^{(k)}, a_{m-k-1}^{(k)}, \ldots, a_0^{(k)}\},$$

$$\{b_{j,m-k}^{(k)}, b_{j,m-k-1}^{(k)}, \ldots, b_{j,0}^{(k)}\} \; j = 1, \ldots, h$$

are the coprime polynomials obtained from the original set after division of $P_i(s)$ by the gcd.

The matrix Φ has all information associated with the GCD and this is established by the following results [1].

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*

$$A_i = \begin{bmatrix} a_0^i & 0 & 0 & \cdots & \cdots & 0 \\ a_1^i & a_0^i & 0 & & & \vdots \\ a_2^i & a_1^i & a_0^i & \ddots & & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ a_{n-2}^i & a_{n-3}^i & a_{n-4}^i & \cdots & a_0^i & 0 \\ a_{n-1}^i & a_{n-2}^i & a_{n-3}^i & \cdots & a_1^i & a_0^i \end{bmatrix} \tag{6}$$

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

$$A^{-1} = \begin{bmatrix} \ell_0 & 0 & 0 & \cdots & \cdots & 0 \\ \ell_1 & \ell_0 & 0 & & & \vdots \\ \ell_2 & \ell_1 & \ell_0 & \ddots & & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ \ell_{n-2} & \ell_{n-3} & \cdots & \cdots & \ell_0 & 0 \\ \ell_{n-1} & \ell_{n-2} & \cdots & \cdots & \ell_1 & \ell_0 \end{bmatrix} \tag{7a}$$

where

$$\ell_0 = 1/a_0,\ \ell_1 = -1/a_0 \sum_{j=0}^{i-1} \ell_j a_{i-j}. \tag{7b}$$

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 GCD of $\mathcal{P}_{h+1,n}$ and Φ the matrix in the factorisation (4) Then*

$$\hat{\Phi} = \Phi^{-1} = \begin{bmatrix} \lambda_0 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ \lambda_1 & \lambda_0 & \ddots & & & & & \vdots \\ \lambda_2 & \lambda_1 & \ddots & \ddots & & & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & & & \vdots \\ \lambda_k & \vdots & \ddots & \ddots & \ddots & \ddots & & \vdots \\ 0 & \lambda_k & & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & & \ddots & \lambda_1 & \lambda_0 & 0 \\ 0 & \cdots & 0 & \lambda_v & \cdots & \lambda_2 & \lambda_1 & \lambda_0 \end{bmatrix} \tag{8}$$

and thus

$$S_{\mathcal{P}} = S_{\mathcal{P}}^{(k)} \hat{\Phi} \tag{9}$$

Theorem 2 and Corollary 1 establish a very important property of factorisation of resultants. In fact, the Sylvester matrix $S_{\mathcal{P}}$ can be factorised 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 GCD.The extraction of the Toeplitz matrix $\hat{\Phi}$ is equivalent to factorisation of 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 $\mathcal{P}_{h+1,n} = \{p_i(s), i = 0, 1, \ldots, h\}$ be a set of polynomials with the two maximal degrees $(n, m; n \geq m)$ and let $\phi(s) = GCD\{\mathcal{P}_{h+1,n}\}$ with $\deg \phi(s) = k$. If $p_i(s) = p_i^{(k)}(s) \cdot \phi(s)$, $\forall\, i \in h+1$ and $\mathcal{P}_{h+1,n}^{(k)} = \{p_i^{(k)}(s), i = 0, 1, \ldots, h\}$ and denote this factorisation as*

$$\mathcal{P}_{h+1,n} = \mathcal{P}_{h+1,n}^{(k)} \cdot \phi(s) \tag{10}$$

then a matrix representation of the above algebraic factorisation is defined by

$$S_{\mathcal{P}} = S_{\mathcal{P}}^{(k)} \cdot \hat{\Phi} \tag{11}$$

where $S_{\mathcal{P}}$, $S_{\mathcal{P}}^{(k)}$ are the resultant representations of $\mathcal{P}_{h+1,n}$, $\mathcal{P}_{h+1,n}^{(k)}$ and $\hat{\Phi}$ is the Toeplitz representation of $\phi(s)$.

Note 3 *In the description of $\mathcal{P}_{h+1,n}^{(k)}$ set we assume that the first k columns are zero that is $S_{\mathcal{P}}^{(k)} = [O_k, \tilde{S}_{\mathcal{P}}^{(k)}]$ and $\tilde{S}_{\mathcal{P}}^{(k)}$ corresponds to the coprime $\mathcal{P}_{h+1,n}^{(k)}$ set; clearly, factorisation like (11) way be also established for factors, which are not necessarily the 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 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 *GCD*

Given the set $\mathcal{P}_{h+1,n}$ of polynomials with ordered degree set $\mathcal{D} = \{d_i, i = 0, 1, \ldots, h : d_0 = n \geq d_1 = m \geq d_2 \geq \cdots \geq d_h\}$. Let $S_{\mathcal{P}} \in \Psi_{\mathcal{D}}$ be the generalised resultant representation of $\mathcal{P}_{h+1,n}$ and Ψ_D the set of all resultants parametrical by $\mathcal{D}$. Let us assume that $\omega(s) = \lambda_k s^k + \cdots + \lambda_1 s + \lambda_0$ is an approximate GCD, where $k \leq d_h$ and let $\hat{\Phi}$ be the Toeplitz representation

$$\hat{\Phi}_\omega = \Phi_\omega^{-1} = \begin{bmatrix} \lambda_0 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ \lambda_1 & \lambda_0 & \ddots & & & & & \vdots \\ \lambda_2 & \lambda_1 & \ddots & \ddots & & & & \vdots \\ \vdots & \ddots & \ddots & \lambda_0 & \ddots & & & \vdots \\ \lambda_k & \vdots & \ddots & \lambda_1 & \ddots & \ddots & & \vdots \\ 0 & \lambda_k & & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & & \ddots & \ddots & \lambda_0 & 0 \\ 0 & \cdots & 0 & \lambda_k & \cdots & \cdots & \lambda_1 & \lambda_0 \end{bmatrix} \tag{12}$$

Step 1 *Define the family of all structured dynamic perturbations*

$\mathcal{Q}_{h+1,n} = \{q_i(s), i = 0, 1, \ldots, h, \partial\{q_i(s)\} = \sigma \leq d_i,\ \forall\, i = 0, 1, \ldots, h\}$ *such that*

$$\mathcal{P}'_{h+1,n} = \mathcal{P}_{h+1,n} + \mathcal{Q}_{h+1,n} = \{p'_i(s) = p_i(s) + q_i(s), \\ \forall\, i = 0, \ldots, h\}$$

with the same degree set $\mathcal{D}$ as $\mathcal{P}_{h+1,n}$ which has $\phi(s)$ as an exact GCD. The set of all such dynamic perturbations $\mathcal{Q}_{h+1,n}$ has a Sylvester representation:

$$S_{\mathcal{Q}} = S_{\mathcal{P}} - [O_n, \tilde{S}_{\mathcal{P}}^{(k)}]\hat{\Phi}_\omega = S_{\mathcal{P}} - S^{(k)}\hat{\Phi}_\omega \tag{13}$$

where $S^{(k)}$ is a resultant from $\Psi_{\mathcal{D}}$ 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 $b_{j,i}^{(k)}$, $j = 1, \ldots, h$, $i = 0, 1, \ldots, m-h$ arbitrary parameters.

Step 2 *Using as parameters those free variables in* $\tilde{S}_{\mathcal{P}}^{(k)} = S_w$ *define the function*

$$f(\mathcal{P}, W) = \|S_{\mathcal{P}} - [O_k, S_W]\hat{\Phi}_\omega\|_F$$

and solve the minimisation problem

$$\min_W f(\mathcal{P}, W)$$

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

Example We consider the set of polynomials

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

Then we have: $n = 2, \;\; m = 1$ 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 $w = (a, b, c)^t$. $f(\mathcal{P}, W) = \|S_{\mathcal{P}} - [O_k, S_W]\hat{\Phi}_\omega\|_F$

Solving the minimization problem with MATLAB we get:

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

[1] S. Fatouros and N. Karcanias, The GCD of many polynomials and the factorisation of generalised resultants, *Research Report*, Control Engineering Centre, City University, London, 2002.

[2] N. Karcanias, G. Giannakopoulos and M. Hubbard, Almost zeros of a set of polynomials, *Int. Journ. Control*, 40 (1984), 673-698.

[3] N. Karcanias and M. Mitrouli, Numerical computation of the least common multiple of a set of polynomials, *Reliable Computing*, Issue 4, 6 (2000b), 439-457.

[4] N. Karcanias N. and M. Mitrouli, A Matrix Pencil Based Numerical Method for the Computation of the GCD of Polynomials, *IEEE Trans. Autom. Cont.*, 39 (1994) , 977-981.

[5] N. Karcanias, M. Mitrouli and S. Fatouros, Computation of normal normal factorisation of polynomials using resultant sets , *IFAC Symposium on System Structure and Control.*, Prague, September 2001.

[6] M. Mitrouli and N. Karcanias, Computation of the GCD of polynomials using Gaussian transformation and shifting, *Int. Journ. Control*, 58 (1993) , 211-228.

[7] M. Mitrouli, N. Karcanias and C. Koukouvinos, Further numerical aspects of the ERES algorithm for the computation of the greatest common divisor of polynomials and comparison with other existing methodologies, *Utilitas Mathematica*, 50 (1996), 65-84.

[8] M. Noda M. and T. Sasaki, Approximate GCD and its applications to ill-conditioned algebraic equations, *Jour. of Comp. and Appl. Math.*, 38 (1991), 335-351.

[9] I. S. Pace and S. Barnett, Comparison of algorithms for calculation of GCD of polynomials, *Int. Journ. System Scien.*, 4 (1973) , 211-226.