Approximate Solutions to Root Clustering Problem

S. Fatouros*, N. Karcanias* and M. Mitrouli**
s.fatouros@city.ac.uk, N.Karcanias@city.ac.uk, mmitroul@math.uoa.gr

*Control Engineering Centre, School of Engineering and Mathematical Science, CITY UNIVERSITY LONDON, U.K.
** Faculty of Mathematics, University of Athens, Greece

Abstract

The algebra of polynomials provides the basis for the development of algebraic control approaches [5],[6],[7] etc and issues such as computation of Smith forms, solvability of Diophantine equations, solution of general matrix are essential parts of algorithms, and procedures linked to algebraic design.

In this paper we investigate the problem of defining approximate solutions to the root clustering problem, by deploying some resent results on the representation of the greatest common divisor (gcd) of many polynomials [3] and by using a new definition for the notion of the approximate gcd [4]. The results provide different order approximations to the root clustering problem. The basis of our analysis is an algebraic framework which is based on the normal factorisation of polynomials [9] and the algebraic framework for gcd representation and its approximate version are briefly summarised below.

Keywords: Approximate GCD of polynomials, Normal Factorisation, Sylvester Resultant Matrix

1. Definitions and preliminary results

Consider a set of polynomials:

$$P=\{a(s),b_i(s)\in\mathbb{R}[s],\ i\in\underset{\sim}{h}\} \qquad (1.1)$$

which has $h+1$ elements and with the two largest values of degrees (n,p). Without loss of generality we may assume $a(s)$ monic and represent the polynomials with respect to the n degree as

$a(s)=s^n+a_{n-1}s^{n-1}+\ldots+a_1s+a_0$ and

$b_i(s)=b_{i,p}s^p+\ldots+b_{i,1}s+b_{i,0},\ i=1,2,,\ldots,h$.

Whenever we want to denote the number of elements and the maximal degree we shall use the notation $P_{h+1,n}$. The greatest common divisor (gcd) of P will be denoted by $\varphi(s)\triangleq\gcd\{P\}$. With the set P we may associate the Generalised Sylvester Resultant [2]:

Definition (1) [2]: We can define a $p\times(n+p)$ matrix associated with $a(s)$:

$$S_0=\begin{bmatrix}1 & a_{n-1} & a_{n-2} & \cdots & a_1 & a_0 & 0 & \cdots & 0\\ 0 & 1 & a_{n-1} & \cdots & a_2 & a_1 & a_0 & \cdots & 0\\ \vdots & & \ddots & & & & & & \vdots\\ 0 & 0 & \cdots & 1 & a_{n-1} & \cdots & & a_1 & a_0\end{bmatrix} \qquad (1.2a)$$

and an $n\times(n+p)$ matrix associated with $b_i(s)$:

$$S_i=\begin{bmatrix}b_{i,p} & b_{i,p-1} & b_{i,p-2} & \cdots & b_{i,1} & b_{i,0} & 0 & \cdots & 0\\ 0 & b_{i,p} & b_{i,p-1} & \cdots & b_{i,2} & b_{i,1} & b_{i,0} & \cdots & 0\\ \vdots & & \ddots & & & & & & \vdots\\ 0 & \cdots & 0 & b_{i,p} & b_{i,p-1} & \cdots & & b_{i,1} & b_{i,0}\end{bmatrix} \qquad (1.2b)$$

for each $i=1,2,\cdots,h$. An ***extended Sylvester matrix*** for the set P is then defined by:

$$S_P=\begin{bmatrix}S_0\\ S_1\\ \vdots\\ S_h\end{bmatrix}\in\mathbb{R}^{(p+hn)\times(n+p)} \qquad (1.2c)$$

■

The Sylvester matrix is used for the evaluation of the GCD of two or more polynomials. This property is expressed by the Generalised Resultant Theorem: [2],[8]:

Theorem (1): Given a set of polynomials $P = \left\{a(s)=s^n+a_{n-1}s^{n-1}+\ldots+a_0,\ b_i(s)=b_{i,p}s^p+\ldots+b_{i,1}s+b_{i,0}\right.$, $\left. i=1,2,\ldots,h,\ \max\left\{\deg b_i(s)\right\}=p\right\}$ with a generalised resultant S_P the following properties hold true:

i) Necessary and sufficient condition for a set of polynomials to be coprime is that:

$$\operatorname{rank}\left(S_P\right)=n+p \tag{1.3}$$

ii) Let $\varphi(s)$ be the g.c.d. of P . Then:

$$\operatorname{rank}\left(S_P\right)=n+p-\deg\varphi(s) \tag{1.4}$$

iii) If we reduce S_P, by using elementary row operations, to its row echelon form, the last non vanishing row defines the coefficients of the g.c.d..

■

The Toeplitz block based Generalised Resultant is crucial in the representation of the GCD, which is defined by the following factorization of resultants result [3]:

Theorem (2): Let $P=\left\{a(s)b_1(s),\ldots,b_h(s)\right\}$ be a 0-order set, $\deg a(s)=n$, $\deg b_i(s)\le p\le n$, $i=1,\ldots,h$ be a polynomial set, S_P the respective Sylvester matrix, $\varphi(s)=\lambda_k s^k+\cdots+\lambda_1 s+\lambda_0$ be the greatest common divisor of the set and let k be its degree. Then there exists transformation matrix $\Phi_\varphi\in\mathbb{R}^{(n+p)\times(n+p)}$ such that:

$$\overline{S}_{P^*}^{(k)}=S_P\Phi_\varphi=\left[\mathbf{0}_k\ \middle|\ \overline{S}_{P^*}\right], \tag{1.5}$$

where $\overline{S}_{P^*}^{(k)}$, Φ_φ are given by:

$$\Phi_\varphi=\begin{bmatrix}
y_0 & 0 & \cdots & & & & \cdots & 0\\
y_1 & y_0 & & & & & & \vdots\\
y_2 & y_1 & \ddots & \ddots & & & & \\
\vdots & \vdots & \ddots & y_0 & 0 & & & \\
 & & & y_1 & y_0 & \ddots & & \\
 & & & \vdots & \vdots & \ddots & & \\
y_{n+p-2} & y_{n+p-3} & \cdots & y_{n+p-j-2} & y_{n+p-j-3} & \cdots & y_0 & 0\\
y_{n+p-1} & y_{n+p-2} & \cdots & y_{n+p-j-1} & y_{n+p-j-2} & \cdots & y_1 & y_0
\end{bmatrix} \tag{3.24a}$$

(1.6a)

where

$$y_0=\frac{1}{\lambda_0},\quad y_1=-\frac{\lambda_1}{\lambda_0}y_0,\ \ldots,\ y_j=-\frac{1}{\lambda_0}\sum_{i=1}^{\min\{j,k\}}\lambda_i y_{j-i},\ j=2,\ldots,n+p-1 \tag{3.24b}$$

(1.6b)

$$\overline{S}_{P^*}^{(k)}=\left[\begin{array}{ccc|cccccccc}
0 & \cdots & 0 & a_{n-k}^{(k)} & a_{n-k-1}^{(k)} & \cdots & a_1^{(k)} & a_0^{(k)} & 0 & \cdots & 0\\
0 & \cdots & 0 & 0 & a_{n-k}^{(k)} & & \cdots & a_1^{(k)} & a_0^{(k)} & \ddots & \vdots\\
\vdots & & \vdots & \vdots & \ddots & \ddots & & & \ddots & \ddots & 0\\
0 & \cdots & 0 & 0 & \cdots & 0 & & a_{n-k}^{(k)} & \cdots & a_1^{(k)} & a_0^{(k)}\\
0 & \cdots & 0 & b_{1,p-k}^{(k)} & b_{1,p-k-1}^{(k)} & \cdots & b_{1,0}^{(k)} & 0 & \cdots & & 0\\
0 & \cdots & 0 & 0 & b_{1,p-k}^{(k)} & b_{1,p-k-1}^{(k)} & \cdots & b_{1,0}^{(k)} & 0 & \cdots & 0\\
\vdots & & \vdots & & & & & & & \vdots & \\
 & & & & & \ddots & \ddots & & & \ddots & \\
0 & \cdots & 0 & 0 & \cdots & & 0 & b_{1,p-k}^{(k)} & \cdots & b_{1,1}^{(k)} & b_{1,0}^{(k)}\\
0 & \cdots & 0 & b_{2,p-k}^{(k)} & b_{2,p-k-1}^{(k)} & \cdots & b_{2,0}^{(k)} & 0 & \cdots & & 0\\
0 & \cdots & 0 & 0 & b_{2,p-k}^{(k)} & b_{2,p-k-1}^{(k)} & \cdots & b_{2,0}^{(k)} & 0 & \cdots & 0\\
 & & & & & & & & & & \\
 & & & & & & & & & & \\
 & & & & & & & & & & \\
0 & \cdots & 0 & b_{h,p-k}^{(k)} & b_{h,p-k-1}^{(k)} & \cdots & b_{h,0}^{(k)} & 0 & \cdots & & 0\\
0 & \cdots & 0 & 0 & b_{h,p-k}^{(k)} & b_{h,p-k-1}^{(k)} & & b_{h,0}^{(k)} & 0 & \cdots & 0\\
\vdots & & \vdots & \vdots & \ddots & \ddots & \ddots & & \ddots & \ddots & \\
 & & & & & & & & & b_{h,0}^{(k)} & 0\\
0 & \cdots & 0 & 0 & \cdots & 0 & b_{h,p-k}^{(k)} & b_{h,p-k-1}^{(k)} & \cdots & b_{h,1}^{(k)} & b_{h,0}^{(k)}
\end{array}\right] \tag{1.7}$$

where $\left[a_{p-k}^{(k)},a_{p-k-1}^{(k)},\ldots,a_0^{(k)}\right]$ $\left[b_{j,p-k}^{(k)},b_{j,p-k-1}^{(k)},\ldots,b_{j,0}^{(k)}\right]$ $j=1,\ldots,h$ are the coefficients of the coprime polynomials obtained from the original set after the division by the gcd, which define the set $P_{h+1,n-k}^*$ and $\overline{S}_{P^*}$ is called the corresponding *Expanded Resultant*.

■

Theorem (2) is important for the characterisation of approximate GCD of fixed degree for many polynomials. This approach is based on relaxation of conditions on Theorem (2). The Optimal GCD is the one that corresponds to the *minimum perturbation* that has to be applied on the coefficients of the polynomials so that it becomes exact GCD of the

perturbed set [4]. This is a new method introduced in [10] and the results are summarized in Theorem (4).

The following analysis assumes that the degree of the GCD is fixed. In the case of the exact GCD the degree is equal to the nullity of the Resultant. The generalisation the notion of nullity in the approximate case is the *Numerical Nullity* of a matrix [11]:

Definition (2): The numerical ε-rank of a matrix $A \in \mathbb{R}^{m\times n}$ is defined by

$$\rho_{\varepsilon}(A) = \min_{B}\left\{\rho(B):\ \|A-B\| \le \varepsilon,\ \varepsilon > 0\right\}$$

and the numerical ε-right nullity

$$N_{\varepsilon}(A) = \max_{B}\left\{N(B):\ \|A-B\| \le \varepsilon,\ \varepsilon > 0\right\} = n - \rho_{\varepsilon}(A)$$

■

The evaluation of the numerical nullity is based on the singular values of the matrix:

Lemma (1): [11]: For a matrix $A \in \mathbb{R}^{m\times n}$

$\rho_{\varepsilon}(A) =$ number of singular values of A that are $\le \varepsilon$

■

Theorem (3): The Optimal approximate GCD of degree k $\varphi(s) = \lambda_k s^k + ... + \lambda_1 s + \lambda_0$, $\lambda_0 \neq 0$ of a set of many polynomials corresponds to the minimisation of $\left\|\hat{S}_Q\right\|$ where

$$\hat{S}_Q = \left[\begin{array}{cccccc}
z_{0,n} & \cdots & z_{0,n-k+1} & 0 & \cdots & 0 \\
0 & z_{0,n} & \cdots & z_{0,n-k+1} & \ddots & \vdots \\
\vdots & \ddots & \ddots & & \ddots & 0 \\
0 & \cdots & 0 & z_{0,n} & \cdots & z_{0,n-k+1} \\
\hdashline
z_{i,p} & \cdots & z_{i,p-k+1} & 0 & & 0 \\
0 & z_{i,p} & \cdots & z_{i,p-k+1} & \ddots & \\
\vdots & \ddots & & & \ddots & 0 \\
0 & \cdots & 0 & z_{i,p} & \cdots & z_{i,p-k+1} \\
\hdashline
z_{h,p} & \cdots & z_{h,p-k+1} & 0 & \cdots & 0 \\
0 & z_{h,p} & & z_{h,p-k+1} & \ddots & \vdots \\
\vdots & \ddots & \ddots & & \ddots & 0 \\
0 & \cdots & 0 & z_{h,p} & \cdots & z_{h,p-k+1}
\end{array}\right] \qquad (1.8a)$$

where the $z_{0,n-\theta}$, $z_{0,p-\theta}$ elements are defined by

$$z_{0,n-\theta} = \sum_{j=\theta}^{k-1} \bar{c}_{0,n-j}\lambda_j = \sum_{j=\theta}^{k-1}\left[\left(\sum_{\mu=j}^{n} b_{0,n-\mu}y_{\mu-j}\right)\lambda_j\right] = \sum_{j=\theta}^{k-1}\sum_{\mu=1}^{n}\left(b_{0,n-\mu}y_{\mu-j}\lambda_j\right)$$

$$z_{i,n-\theta} = \sum_{j=\theta}^{k-1}\sum_{\mu=1}^{p}\left(b_{0,p-\mu}y_{\mu-j}\lambda_j\right) \qquad (1.8b)$$

■

The above analysis assumes that the degree of the GCD is fixed. In the case of the exact GCD the degree is equal to the nullity of the Resultant. The generalisation the notion of nullity in the approximate case is the *Numerical Nullity* of a matrix [11]:

2. Root Clustering of Polynomials

For every polynomial $b(s) \in \mathbb{R}[s]$ there exist positive integers $d_1,...,d_\sigma$ where $d_1 > d_2 > ... > d_\sigma \ge 1$ such that $b(s)$ may be expressed as [9]:

$$b(s) = e_1(s)^{d_1} e_2(s)^{d_2} \cdots e_\sigma(s)^{d_\sigma} \qquad (2.1)$$

where the polynomials $e_1(s), e_2(s), ..., e_\sigma(s)$ are pairwise coprime and the polynomial

$$\hat{b}(s) = e_1(s)e_2(s)\cdots e_\sigma(s) \qquad (2.2)$$

has distinct roots. This factorisation is known as normal factorisation [9].

Proposition (1): Consider a polynomial $b(s) \in \mathbb{R}[s]$, $b(s) = s^n + b_{n-1}s^{n-1} + ... + b_1 s + b_0$. We assume that $(s+\lambda)^\tau$ is an elementary divisor of $b(s)$ over C. The following properties hold true:

i) The first derivative $b^{(1)}(s)$ has $(s+\lambda)^{\tau-1}$ as elementary devisor.

ii) The k-th derivative $b^{(k)}(s)$, $k < \tau$, has $(s+\lambda)^{\tau-k}$ as elementary divisor.

iii) The $b^{(\tau)}(s)$ derivative is the smallest order derivative that has no roots at $s = -\lambda$.

■

The significance of the above result is that introduces a framework for factorisation of polynomials without root finding. An exact elementary divisor $\left(s+\lambda_j\right)$ of multiplicity $h \le n$ is also an exact divisor of the polynomial set consisted of the initial polynomial and its h first derivatives. We denote this set as

$$D_n^{h+1}=\left\{b_0(s)=s^n+\ldots+b_{0,1}s+b_{0,0},\ b_i(s)=b_0^{(i)}(s),\ i=1,\ldots,h,\ h\le n\right\}$$

The Sylvester matrix S_D of the D_n^{h+1} set will be a $\left[(h+1)n-1\right]\times\left(2n-1\right)$ matrix of the form:

$$S_D=\begin{bmatrix} S_D^{(0)} \\ S_D^{(i)} \\ S_D^{(h)} \end{bmatrix} \tag{2.3a}$$

where

$$S_D^{(0)}=\begin{bmatrix} 1 & b_{0,p-1} & b_{0,p-2} & \cdots & b_{0,1} & b_{0,0} & 0 & \cdots & 0 \\ 0 & 1 & b_{0,p-1} & \cdots & b_{0,2} & b_{0,1} & b_{0,0} & \cdots & 0 \\ \vdots & & \ddots & & & & & & \vdots \\ 0 & \cdots & 0 & 1 & b_{0,p-1} & \cdots & & b_{0,1} & b_{0,0} \end{bmatrix} \tag{2.3b}$$

$$S_D^{(i)}=\left[\begin{array}{ccccccc|ccc} b_{i,p-i} & \cdots & b_{i,1} & b_{i,0} & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \ddots & & & & \ddots & \vdots & \vdots & & \vdots \\ \vdots & \ddots & b_{i,p-i} & b_{i,p-i-1} & \cdots & b_{i,0} & 0 & 0 & & 0 \\ 0 & \cdots & 0 & b_{i,p-i} & b_{i,p-i-1} & \cdots & b_{i,0} & 0 & \cdots & 0 \end{array}\right] \tag{2.3c}$$

Note that $b_i(s)$ is the i-th derivative of $b(s)$ and thus

$$b_i(s)=b_0^{(i)}(s)=\sum_{j=i}^{n}\frac{j!}{(j-i)!}b_{0,j}s^{j-i}=\sum_{k=0}^{n-i}b_{i,k}s^k \tag{2.4a}$$

which in terms of the coefficients is:

$$b_{i,k}=\frac{(k+i)!}{(k)!}b_{0,k+i},\ k=0,1,\ldots,n-i,\ i\le h \tag{2.4b}$$

or

$$b_{i,k}=F_{k+i,k}b_{0,k+i},\ \ k=0,1,\ldots,n-i\ ,\ \ i\le h \tag{2.4c}$$

where by $F_{a,b}$ we denote the factorial

$$F_{a,b}=\frac{a!}{(a-b)!},\ \ a\ge b\ge 0 \tag{2.4d}$$

An approximate factorisation of the polynomial $b(s)$ can be evaluated by combining the above results with the algorithm for approximate GCD on the set of the derivatives, for all approximate divisors of degree 2 or higher. This procedure includes three basic phases:

Step1: Estimation of the degree $h_1 \le n$ for which there exists an approximate elementary divisor of multiplicity h_1, i.e. $\left(s+\lambda_1\right)^{h_1}$ is an approximate factor of $b(s)$. To this aim we investigate whether $N_\varepsilon\left(D_n^{h+1}\right)\ge 1$.

Step 2: Evaluation of best approximate common factor $\left(s+\lambda_1\right)$ for the set D_n^{h+1}. The combination of the nature of the elements of D_n^{h+1} with Theorem 3 implies the form of the optimisation:

Proposition (2): The best approximate factor $e_j=\left(s+\lambda_j\right)$ of multiplicity h is obtained from the optimisation of $\min\left\|\tilde{S}_D'\hat{\Phi}_{e_j}\right\|$ where (b)

$$\tilde{S}_D'\hat{\Phi}_{e_j}=\left[\begin{array}{ccccccc} z_{0,n} & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & z_{0,n} & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & & \ddots & \vdots \\ 0 & \cdots & 0 & z_{0,n} & 0 & \cdots & 0 \\ \hline & & \vdots & & \vdots & & \\ \hline z_{h,p} & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & z_{h,p} & 0 & \cdots & 0 & & 0 \\ \vdots & \ddots & & \ddots & & \ddots & \vdots \\ 0 & \cdots & 0 & z_{h,p} & 0 & \cdots & 0 \end{array}\right] \tag{2.5a}$$

and

$$z_{0,n}=\sum_{\theta=0}^{n}(-1)^\theta\,\lambda_j^{\,1-\theta}b_{n-\theta}$$

$$z_{i,p}=\sum_{\theta=0}^{p}(-1)^\theta\,\lambda_j^{\,1-\theta}F_{(n-\theta),i}b_{n-\theta} \tag{2.5b}$$

■

Step 3: Division of the initial polynomial by the factor $\left(e_1(s)\right)^h = \left(s+\lambda_1\right)^{h_1}$ such that

$$b(s) = \left(s+\lambda_1\right)^{h_1} \cdot f_1(s) + \varepsilon_1(s) \qquad (2.6)$$

$$n_1 = \deg\left\{ f_1(s)\right\} = n - h_1$$

where $-\varepsilon_1(s)$ is the minimum perturbation that is required for $\left(s+\lambda_1\right)^{h_1}$ to become exact factor of $b(s)$.

The polynomial pair $\left(f_1(s), -\varepsilon_1(s)\right)$ can be found with matrix operations described in [4].

Step 4: We repeat the procedure from Step 1, for the quotient polynomial of the division (2.6). We terminate it when $N_{\varepsilon}\left(D^2_{n_k}\right) = 0$.

The result of the above algorithm will be a sequence of polynomials $f_1(s),..., f_k(s), f_{k+1}(s)$ where $f_j(s) = \left(s+\lambda_j\right)$, $j = 1,...,k$ and $f_{k+1}(s)$ will be the irreducible part. The approximate factorisation of $b(s)$ is then given by:

$$f_1(s)^{h_1}\ f_2(s)^{h_2} \cdots f_k(s)^{h_k}\ f_{k+1}(s) \qquad (2.7a)$$

where

$$h_1 \geq h_2 \geq \cdots \geq h_k \geq 2 \qquad (2.7b)$$

and

$$h_1 + h_2 + ... + h_k + \deg\left\{ f_{k+1}(s)\right\} = n \qquad (2.7c)$$

3. Discussion

This paper has introduced a theoretic framework for approximate factorisation of polynomials based on recent matrix algorithms for the evaluation of the exact [3] and the approximate GCD [4], [10] of many polynomials. The disadvantage of the algorithm is its operational complexity, which may be reduced with a more appropriate sequence of operations.

4. References

[1] **KARCANIAS, N., and MITROULI, M.** (2000): "Numerical computation of the least common multiple of a set of polynomials", *Reliable Computing*, **4**(6), 439-457.

[2] **BARNETT, S.** (1990): "Matrices Methods and Applications", *Clarendon Press*, Oxford

[3] **FATOUROS, S. and KARCANIAS, N.** (2002): "Resultant properties of GCD of many polynomials, GCD representation and its computation" *City University Control Engineering Center*, *Research report*, London

[4] **FATOUROS, S., KARCANIAS, N., MITROYLI, M. and HALIKIAS G.** (2002): "Approximate Greatest Common Divisor of many polynomials and generalised resultants" *ACA' 2002 Conference*, June 25-28, 2002, Volos, Greece

[5] **ROSENBROCK, H.H.** (1970): "State Space and Multivariable Theory", *Nelson*, London

[6] **KAILATH, T.** (1980): "Linear Systems", *Pentice Hall*, Englewood Cliffs N.J.

[7] **KUCERA, V. (**1984): "Discrete Linear Control: The Polynomial Equation Approach", *Willey*, Chichester. UK

[8] **VARDOULAKIS, A.I.G. and STOYLE, P.N.R** (1978): "Generalised Resultant Theorem", *J. Inst. of Maths. and its Appl.*, **22**, 331-335.

[9] **KARCANIAS, N. and MITROULI, M.** (2002): "Normal factorisation of polynomials and computational issues", *An Inter. Journal of Computers and Mathematics with Applications, Submitted 2002*

[10] **FATOUROS, S. and KARCANIAS, N.** "Optimal Approximate GCD of many polynomials", (2003): *City University, Control Engineering Center*, *Research report*, London

[11] **FOSTER, L**. (1986): "Rank and Null Space calculations using matrix using matrix decomposition without column interchanges", *Lin. Alg. And its Appl.*, **74**, 47-71.