Ján Eliaš; Jan Zítko
Approximate polynomial GCD

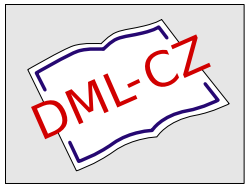
In: Jan Chleboun and Karel Segeth and Jakub Šístek and Tomáš Vejchodský (eds.): Programs and Algorithms of Numerical Mathematics, Proceedings of Seminar. Dolní Maxov, June 3-8, 2012. Institute of Mathematics AS CR, Prague, 2013. pp. 63--68.

Persistent URL: http://dml.cz/dmlcz/702708

Terms of use:

© Institute of Mathematics AS CR, 2013

Institute of Mathematics of the Academy of Sciences of the Czech Republic provides access to digitized documents strictly for personal use. Each copy of any part of this document must contain these *Terms of use*.



This paper has been digitized, optimized for electronic delivery and stamped with digital signature within the project *DML-CZ: The Czech Digital Mathematics Library*
http://project.dml.cz

Programs and Algorithms of Numerical Matematics 16
J. Chleboun, K. Segeth, J. Šístek, T. Vejchodský (Eds.)
Institute of Mathematics AS CR, Prague 2013

APPROXIMATE POLYNOMIAL GCD

Ján Eliaš, Jan Zítko

Department of Numerical Mathematics,
Faculty of Mathematics and Physics, Charles University in Prague
Sokolovská, Prague, Czech Republic
janelias@ymail.com, zitko@karlin.mff.cuni.cz

Abstract

The computation of polynomial greatest common divisor (GCD) ranks among basic algebraic problems with many applications, for example, in image processing and control theory. The problem of the GCD computing of two exact polynomials is well defined and can be solved symbolically, for example, by the oldest and commonly used Euclid's algorithm. However, this is an ill-posed problem, particularly when some unknown noise is applied to the polynomial coefficients. Hence, new methods for the GCD computation have been extensively studied in recent years.
The aim is to overcome the ill-posed sensitivity of the GCD computation in the presence of noise. We show that this can be successively done through a TLS formulation of the solved problem, [1, 5, 7].

1. Approximate greatest common divisor

Suppose a pair of two polynomials f and g of degrees m and n,

$$f(x) = \sum_{i=0}^{m} a_i x^{m-i} \quad (a_0 a_m \neq 0) \quad \text{and} \quad g(x) = \sum_{j=0}^{n} b_j x^{n-j} \quad (b_0 b_n \neq 0) \tag{1}$$

with a non-trivial GCD h of degree d is given, $1 \leq d \leq n \leq m$. Vectors of polynomial coefficients are denoted by bold lower-case Latin letters, e.g. $\mathbf{f} = [a_0, a_1, \ldots, a_m]^T$ represents the vector of coefficients of f. Similarly, $\mathbf{g}$, $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{h}$ will denote the vectors of coefficients of involved polynomials g, u, v and h.

Then there exist polynomials u and v of degrees $m - d$ and $n - d$, respectively, so that

$$uh = f \quad \text{and} \quad vh = g. \tag{2}$$

Equations in (2) can be rewritten to the matrix-vector notation as

$$S_d(f, g) \begin{bmatrix} \mathbf{v} \\ -\mathbf{u} \end{bmatrix} = \mathbf{0}, \tag{3}$$

where

$$S_d(f,g) = \underbrace{\begin{bmatrix} a_0 & & & & b_0 & & & \\ a_1 & a_0 & & & b_1 & b_0 & & \\ \vdots & a_1 & \ddots & & \vdots & b_1 & \ddots & \\ a_m & \vdots & \ddots & a_0 & b_n & \vdots & \ddots & b_0 \\ & a_m & & a_1 & & b_n & & b_1 \\ & & \ddots & \vdots & & & \ddots & \vdots \\ & & & a_m & & & & b_n \end{bmatrix}}_{n-d+1 \text{ col.}} \underbrace{\qquad\qquad}_{m-d+1 \text{ col.}} \in \mathbb{R}^{(m+n-d+1)\times(m+n-2d+2)}$$

is, except the case $d = 1$, rectangular $m+n-d+1$ by $m+n-2d+2$ matrix called the dth Sylvester subresultant matrix. The coefficients $\{a_i\}$ of f occupy the first $n-d+1$ columns and the coefficients $\{b_j\}$ of g occupy the last $m-d+1$ columns. Hence, $S_d(f,g)$ is the block matrix consisting of the two Cauchy matrices, $S_d(f,g) = [C_{n-d+1}(f),\, C_{m-d+1}(g)]$.[1] The Sylvester matrix is then the matrix $S(f,g) = S_1(f,g) = [C_n(f),\, C_m(g)] \in \mathbb{R}^{(m+n)\times(m+n)}$.

The most important relations between the GCD and the Sylvester matrices are summarised in the following theorem.[2]

Theorem 1. *Suppose that f and g are polynomials of degrees m and n, $m \geq n$, and $h = GCD(f,g)$. Then*

i) $rank\,(S(f,g)) = m+n-d \iff \deg h = d$,

ii) $rank\,(S_d(f,g)) = m+n-2d+1 \iff \deg h = d$,

iii) the coefficient vector $\mathbf{h}$ *is a solution of the linear system*

$$\begin{bmatrix} C_{d+1}(u) \\ C_{d+1}(v) \end{bmatrix} \mathbf{h} = \begin{bmatrix} \mathbf{f} \\ \mathbf{g} \end{bmatrix}. \tag{4}$$

Moreover, if $\deg h = d$, *then*

iv) $rank\,(S_j(f,g)) < m+n-2j+2, \quad j = 1,\ldots,d$,

v) $rank\,(S_j(f,g)) = m+n-2j+2, \quad j = d+1,\ldots,n$. □

Hence, if $\deg h = d$, then $S_d = S_d(f,g)$ is rank deficient by 1 since S_d has $m+n-2d+2$ columns and rank $m+n-2d+1$ by recalling the property *ii)* from the theorem. Therefore,

$$S_d \begin{bmatrix} \mathbf{v} \\ -\mathbf{u} \end{bmatrix} = \mathbf{0} \Longrightarrow \exists\, s \in \mathbb{R} : \begin{bmatrix} \mathbf{v} \\ -\mathbf{u} \end{bmatrix} = s\, \mathbf{v}_{\min}(S_d),$$

where $\mathbf{v}_{\min}(S_d)$ is the right singular vector associated with $\sigma_{\min}(S_d) = 0$.

[1] The subscripts $n-d+1$ and $m-d+1$ in $C_{n-d+1}(f)$ and $C_{m-d+1}(g)$ represent the number of columns filled with the coefficients of f and g, respectively.

[2] A proof is outlined in the second authors' paper of these proceedings.

The coefficients of h can be now easily computed. For this purpose we have to calculate the smallest singular pair $\{\sigma_{\min}, \mathbf{v}_{\min}\}$ of every matrix in the sequence $S_n, S_{n-1}, \ldots, S_1$ until the first rank deficient matrix is found.[3] Once, the rank deficient matrix S_d is revealed, $\mathbf{v}$ and $\mathbf{u}$ can be extracted from the singular vector $\mathbf{v}_{\min}(S_d)$. The coefficients of h are then computed from (4).

The smallest singular value and its corresponding right singular vector of S_d can be computed by the Gauss-Newton method, [4], i.e. by the iteration process

$$\mathbf{x}_{i+1} = \mathbf{x}_i - \begin{bmatrix} 2\tau\mathbf{x}_i^T \\ S_d \end{bmatrix}^{\dagger} \begin{bmatrix} \tau\mathbf{x}_i^T\mathbf{x}_i - \tau \\ S_d\mathbf{x}_i \end{bmatrix} \quad \text{and} \quad \zeta_{i+1} = \frac{\|S_d\mathbf{x}_{j+1}\|_2}{\|\mathbf{x}_{j+1}\|_2}$$

for τ sufficiently large.[4] Then

$$\mathbf{x}_i \xrightarrow[i\to\infty]{} \mathbf{v}_{\min}(S_d) \quad \text{and} \quad \zeta_i \xrightarrow[i\to\infty]{} \sigma_{\min}(S_d).$$

The GCD solver in [4, 6] is based on this iteration process. However, note that in real computations some threshold θ must be applied to ζ_i to reveal the rank deficiency. Assuming that the level of noise is not known, the solver in [4, 6] cannot be used, since the numerical rank cannot be computed reliably, [1, 5].

Whether the level of imposed noise is known or not, $\mathbf{v}_{\min}(S_d)$, $\mathbf{u}$ and $\mathbf{v}$ are computed approximately and so the coefficients of h are not calculated exactly. Hence, an approximate greatest common divisor (AGCD) is only computed.

2. Impact of noise

Numerically, S_d is considered to be rank deficient whenever $\sigma_{\min}(S_d) \leq \theta$ for a prescribed threshold θ. If rounding errors are only assumed, then $\theta = \varepsilon\|S_d\|_2$ with a machine precision ε is usually used, [2] p. 261. However, if some additional noise of unknown level is considered, then computations with all similar choices of θ usually fail. In this case a different approach has to be developed.

Dependence of the GCD computation on noise can be seen from the following example. Consider two polynomials f and g of degree 32,

$$\begin{aligned} f(x) &= \prod_{i=1}^{8}\left[(x - r_1\alpha_i)^2 + r_1^2\beta_i^2\right]\prod_{i=9}^{16}\left[(x - r_2\alpha_i)^2 + r_2^2\beta_i^2\right], \\ g(x) &= \prod_{i=1}^{16}\left[(x - r_1\alpha_i)^2 + r_1^2\beta_i^2\right], \end{aligned} \tag{5}$$

where $\alpha_i = \cos\left(\frac{\pi i}{m}\right), \beta_i = \sin\left(\frac{\pi i}{m}\right), i = 1, \ldots, n, r_1 = 0.5$ and $r_2 = 1.5$. These polynomials have the exact GCD of degree 16. So the rank of the Sylvester matrix $S(f, g)$ is 48 by recalling Theorem 1 *i)*.

[3] Note that if S_d is the first rank deficient matrix and $d < n \leq m$, then every S_j in $S_n, \ldots, S_{d+1}$ has full column rank using Theorem 1 *v)*.

[4] The symbol $(\cdot)^{\dagger}$ denotes the Moore-Penrose inverse of $(\cdot)$.

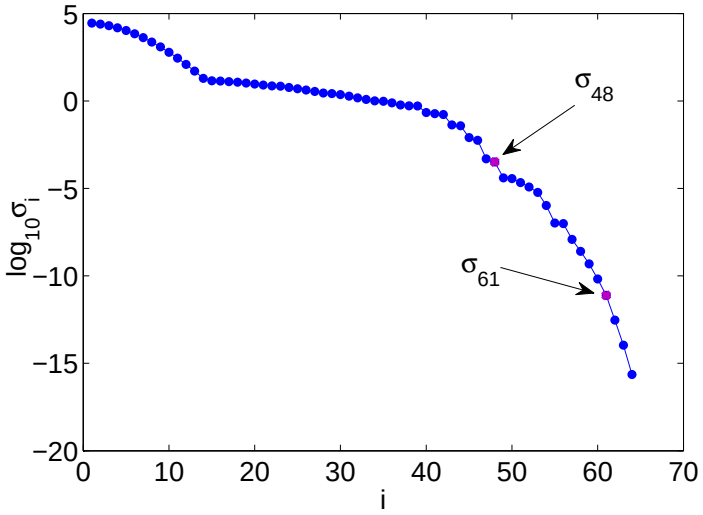


Figure 1: Singular values of the Sylvester matrix $S(f,g)$ for f and g perturbed componentwisely by the noise of the SNR $= 10^8$.

The numerical rank of $S(f,g)$ is well defined and can be revealed by using Gauss-Newton iteration for the choice $\theta = \varepsilon\|S(f,g)\|_2 \approx 10^{-12}$ in case when only rounding errors are considered.

Suppose now, that a noise of the signal-to-noise ratio SNR $= 10^8$ is componentwisely imposed to the coefficients of f and g. Figure 1 shows the singular values of the Sylvester matrix of perturbed polynomials. For the choice $\theta = 10^{-12}$ the numerical rank is 61 that is incorrect. The correct numerical rank 48 can be revealed with $\theta = 10^{-4}$. The question, however, is how to estimate this θ only from the involved data.

3. TLS formulation, methods for AGCD

For the exact polynomials the system of equations (3) can be transformed to the system

$$A_d\mathbf{x} = \mathbf{c_d}, \tag{6}$$

where $\mathbf{c_d}$ is the first column of S_d and A_d is formed from the remaining $m+n-2d+1$ columns of S_d, $S_d = [\mathbf{c_d}, A_d]$.

While the system (6) possesses exactly one solution $\mathbf{x}$ for the exact polynomials, it does not possess any solution for the inexact polynomials, since the perturbed polynomials are coprime with probability almost one, i.e. $\mathbf{c_d} \notin \mathrm{Range}(A_d)$ for the inexact polynomials. However, if the polynomials f and g are coprime, we can demand to compute the minimal corrections of their coefficients, i.e. polynomials δf and δg so that $f+\delta f$ and $g+\delta g$ have a non-trivial GCD with the highest possible degree. Then, $\mathrm{AGCD}(f,g) = \mathrm{GCD}(f+\delta f, g+\delta g)$.

Let us denote the Sylvester matrix of δf and δg by $\delta S_d = \delta S_d(\delta f, \delta g)$, $\delta S_d =$

$[\mathbf{h_d}, E_d]$, and let $\mathbf{z} = [\delta\mathbf{f}^T, \delta\mathbf{g}^T]^T$ be the vector of the coefficients of δf and δg. Then δf and δg can be computed so that

$$(A_d + E_d)\mathbf{x} = \mathbf{c_d} + \mathbf{h_d}$$

has exactly one solution $\mathbf{x}$ and $\|\mathbf{z}\|_2$ is minimal. Hence, the problem, that is finally solved, is the structured TLS problem:

$$\begin{aligned} &\min_{\mathbf{z},\mathbf{x}} \|\mathbf{z}\|_2 \\ &\text{subject to } (A_d + E_d)\mathbf{x} = \mathbf{c_d} + \mathbf{h_d} \\ &\text{and } [\mathbf{h_d}, E_d] \text{ is of the same structure as } [\mathbf{c_d}, A_d]. \end{aligned} \tag{7}$$

Two methods for solving (7) are presented in [3]. These methods are modified and customised for the AGCD computation in [1, 5].

Methods for the AGCD computation are not discussed in this paper, however note, that Sylvester matrices are badly conditioned, for example, if considered polynomials have coefficients that differ by several orders in magnitude. It is therefore necessary to apply some preprocessing operations on polynomials before a method is used. Particularly, these operations include

- normalisation of the coefficients by the geometric mean that preserves the propagation of noise,
- variable substitution $x = \gamma w$ for minimising the ratio of the maximum to the minimum coefficient of both polynomials f and g,
- considering a parameter α in $S(f, \alpha g)$ for weighting the coefficients of one polynomial with respect to the coefficients of the second polynomial,
- column pivoting, i.e. a column of S_d for which the residual $\|A_d\mathbf{y} - \mathbf{c_d}\|_2$ is minimal replaces $\mathbf{c_d}$ in (6).

There are several possible ways how to compute α and γ, for example, they can be computed as values that minimise the ratio

$$\frac{\max\left\{\max_{i=0,\ldots,m} |a_i\gamma^{m-i}|, \max_{j=0,\ldots,n} |\alpha b_j\gamma^{n-j}|\right\}}{\min\left\{\min_{i=0,\ldots,m} |a_i\gamma^{m-i}|, \min_{j=0,\ldots,n} |\alpha b_j\gamma^{n-j}|\right\}}.$$

More information on the preprocessing operations is provided in [5].

Finally, Figure 2 shows the singular values of $S(f, g) + \delta S(\delta f, \delta g)$ for the polynomials f and g in (5) perturbed componentwisely by the noise of the SNR $= 10^8$. The polynomials δf and δg are obtained by solving (7). We can see that the numerical rank is now perfectly defined and so further computation of the GCD by the procedure discussed in Section 1 can be processed.

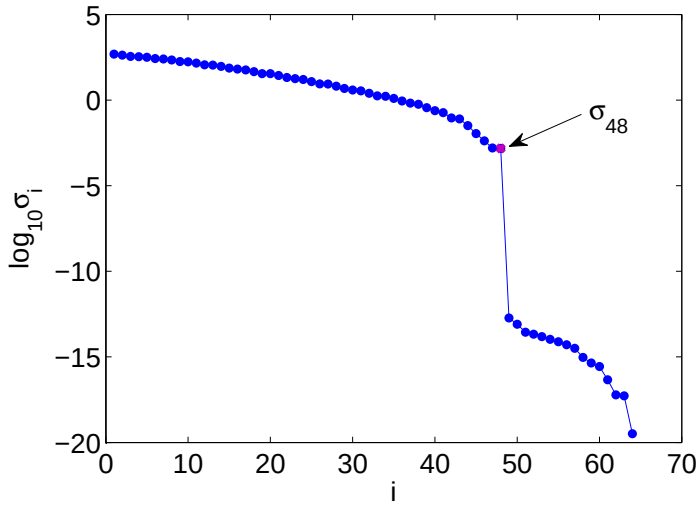


Figure 2: Singular values of $S(f,g)+\delta S(\delta f,\delta g)$ where f and g in (5) are perturbed componentwisely by the noise of the SNR $= 10^8$, and δf and δg are obtained by solving (7).

Acknowledgements

This work was supported by the Department of Numerical Mathematics, Charles University in Prague. The authors thank for this support.

References

[1] Eliaš, J.: *Approximate Polynomial Greatest Common Divisor*. Master Thesis, Charles University in Prague, 2012.

[2] Golub, G. H. and Van Loan, C. F.: *Matrix Computations*. 3rd Ed. The John Hopkins University Press, Baltimore, USA, 1996.

[3] Lemmerling, P., Mastronardi, N., and Van Huffel, S.: Fast algorithm for solving the Hankel/Toeplitz Structured Total Least Squares Problem. Numer. Algorithms **23** (2000), 371–392.

[4] Li, T. Y. and Zeng, Z.: A rank-revealing method with updating, downdating and applications. SIAM J. Matrix Anal. Appl. **26** (2005), 918–946.

[5] Winkler, J. R. and Hasan, M.: A non-linear structure preserving matrix method for the low rank approximation of the Sylvester resultant matrix. J. Comput. Appl. Math. **234** (2010), 3226-3242.

[6] Zeng, Z.: The approximate GCD of inexact polynomials, Part I: univariate algorithm. Preprint (2004).

[7] Zítko, J. and Eliaš, J.: Application of the rank revealing algorithm for the calculation of the GCD. In: *Winter School and SNA'12*, pp. 175–180. Technická Univerzita v Liberci, Liberec, 2012.