A structured low rank approximation of the Sylvester resultant matrix

John D. Allan and Joab R. Winkler
Department of Computer Science
The University of Sheffield, United Kingdom

Overview

- The Sylvester resultant matrix
- Subresultants
- The method of structured total least norm (STLN)
- Results
- Conclusions

1. The Sylvester resultant matrix

Resultant matrices are used in computer aided geometric design for:

- Transforming between the implicit and parametric forms of a curve.
- The computation of the intersection points of curves and surfaces.

The Sylvester resultant matrix $S(f,g)$ of the polynomials

$$f(y) = \sum_{i=0}^{m} a_i y^{m-i} \qquad \text{and} \qquad g(y) = \sum_{j=0}^{n} b_j y^{n-j}$$

has two important properties:

- The determinant of $S(f,g)$ is equal to zero if and only if $f(y)$ and $g(y)$ have a non-constant common divisor.
- The rank of $S(f,g)$ is equal to $(m+n-d)$, where d is the degree of the greatest common divisor (GCD) of $f(y)$ and $g(y)$.

Are the forms of $f(y)$ and $g(y)$ unique?

- If $f(y)$ and $g(y)$ have a non-constant common divisor, then so do $f(y)$ and $\alpha g(y)$, where α is a non-zero constant.
- Instead of considering $S(f,g)$, the more general matrix $S(f,\alpha g)$ is considered.
- What is the effect of α on a structured low rank approximation of $S(f,\alpha g)$, and how is its value chosen?

The matrix $S(f, \alpha g)$ is

$$
\left[\begin{array}{cccc|cccc}
a_0 & & & & \alpha b_0 & & & \\
a_1 & a_0 & & & \alpha b_1 & \alpha b_0 & & \\
\vdots & a_1 & \ddots & & \vdots & \alpha b_1 & \ddots & \\
a_{m-1} & \vdots & \ddots & a_0 & \alpha b_{n-1} & \vdots & \ddots & \alpha b_0 \\
a_m & a_{m-1} & \ddots & a_1 & \alpha b_n & \alpha b_{n-1} & \ddots & \alpha b_1 \\
 & a_m & \ddots & \vdots & & \alpha b_n & \ddots & \vdots \\
 & & \ddots & a_{m-1} & & & \ddots & \alpha b_{n-1} \\
 & & & a_m & & & & \alpha b_n
\end{array}\right]
$$

Note that

$$S(f, \alpha g) \neq \alpha S(f, g)$$

- How can one compute a structured low rank approximation of $S(f, \alpha g)$?

Since the rank of $S(f, \alpha g)$ is equal to $(m + n - d)$, it follows that reducing the rank of $S(f, \alpha g)$ is equivalent to increasing d, the degree of the GCD of $f(y)$ and $\alpha g(y)$.

The computation of a structured low rank approximation of $S(f, \alpha g)$ is therefore obtained by:

- Computing perturbations $\delta f(y)$ and $\delta g(y)$ such that

$$\text{rank}\, S\Big(f + \delta f, \alpha(g + \delta g)\Big) = m + n - (d + 1)$$

- Use subresultants of a Sylvester resultant matrix to calculate these perturbations.

2. Subresultants

The k'th Sylvester matrix, or subresultant, $S_k \in \mathbb{R}^{(m+n-k+1)\times(m+n-2k+2)}$ is a submatrix of $S(f, \alpha g)$ that is formed by:

- Deleting the last $(k-1)$ rows of $S(f, \alpha g)$.
- Deleting the last $(k-1)$ columns of the coefficients of $f(y)$.
- Deleting the last $(k-1)$ columns of the coefficients of $\alpha g(y)$.

The integer k satisfies $1 \leq k \leq \min(m, n)$, and a subresultant matrix is defined for each value of k.

- Start with $k = k_0 = \min(m, n)$ and decrease by one until a solution is obtained.

Example 1

If $m = 4$ and $n = 3$, then

$$S_1 = S(f, \alpha g) = \begin{bmatrix} a_0 & & & \alpha b_0 & & & \\ a_1 & a_0 & & \alpha b_1 & \alpha b_0 & & \\ a_2 & a_1 & a_0 & \alpha b_2 & \alpha b_1 & \alpha b_0 & \\ a_3 & a_2 & a_1 & \alpha b_3 & \alpha b_2 & \alpha b_1 & \alpha b_0 \\ a_4 & a_3 & a_2 & & \alpha b_3 & \alpha b_2 & \alpha b_1 \\ & a_4 & a_3 & & & \alpha b_3 & \alpha b_2 \\ & & a_4 & & & & \alpha b_3 \end{bmatrix}$$

$$S_2 = \begin{bmatrix} a_0 & & \alpha b_0 & & \\ a_1 & a_0 & \alpha b_1 & \alpha b_0 & \\ a_2 & a_1 & \alpha b_2 & \alpha b_1 & \alpha b_0 \\ a_3 & a_2 & \alpha b_3 & \alpha b_2 & \alpha b_1 \\ a_4 & a_3 & & \alpha b_3 & \alpha b_2 \\ & a_4 & & & \alpha b_3 \end{bmatrix}, \qquad S_3 = \begin{bmatrix} a_0 & \alpha b_0 & \\ a_1 & \alpha b_1 & \alpha b_0 \\ a_2 & \alpha b_2 & \alpha b_1 \\ a_3 & \alpha b_3 & \alpha b_2 \\ a_4 & & \alpha b_3 \end{bmatrix}$$

$\square$

Each matrix S_k is partitioned into:

- A vector $c_k \in \mathbb{R}^{m+n-k+1}$, where c_k is the first column of S_k.
- A matrix $A_k = A_k(\alpha) \in \mathbb{R}^{(m+n-k+1)\times(m+n-2k+1)}$, where A_k is the matrix formed from the remaining columns of S_k.

$$
\begin{aligned}
S_k &= \left[\begin{array}{c|c} c_k & A_k \end{array}\right] \\
&= \left[\begin{array}{c|c|c} c_k & \text{coeffs. of } f(y) & \text{coeffs. of } \alpha g(y) \end{array}\right] \\
&\qquad \underbrace{}_{1} \quad \underbrace{\phantom{\text{coeffs. of } f(y)}}_{n-k} \quad \underbrace{\phantom{\text{coeffs. of } \alpha g(y)}}_{m-k+1}
\end{aligned}
$$

- The integer, $1 \leq k \leq \min(m, n)$, can be chosen arbitrarily.

The following theorem is required.

Theorem 1 Consider the polynomials $f(y)$ and $\alpha g(y)$, and let k be a positive integer, where $1 \leq k \leq \min(m, n)$. Then

1. The dimension of the null space of S_k is greater than or equal to one if and only if the over determined equation

$$A_k x = c_k$$

possesses a solution.

2. A necessary and sufficient condition for the polynomials $f(y)$ and $\alpha g(y)$ to have a common divisor of degree greater than or equal to k is that the dimension of the null space of S_k is greater than or equal to one.

Algorithm

1. Set $k = k_0 = \min(m, n)$.
2. Does the equation

$$A_k x = c_k$$

   have an exact solution?

   (a) Yes: GOTO 3.

   (b) No: Set $k := k - 1$ and GOTO 2.
3. The degree of the GCD is $k_0 := k$.

What happens if the polynomials are inexact?

- What are the smallest perturbations $\delta f(y)$ and $\alpha\delta g(y)$ such that $f(y)+\delta f(y)$ and $\alpha(g(y) + \delta g(y))$ have a non-constant GCD?

3. The method of structured total least norm (STLN)

The problem to be solved is:

For a given value of k, compute the smallest perturbations to $f(y)$ and $\alpha g(y)$ such that

$$(A_k + E_k)\, x = c_k + h_k$$

has a solution, where

- E_k has the same structure as A_k.
- h_k has the same structure as c_k.

- z_i be the perturbation of $a_i, i = 0, \ldots, m$, of $f(y)$.
- z_{m+1+j} be the perturbation of $\alpha b_j, j = 0, \ldots, n$, of $\alpha g(y)$.

The perturbed Sylvester resultant matrix is:

$$\left[\begin{array}{c|c} h_k & E_k \end{array}\right] = \left[\begin{array}{cccccc} z_0 & & & & z_{m+1} & & \\ z_1 & z_0 & & & z_{m+2} & & \\ \vdots & z_1 & \ddots & & \vdots & \ddots & \\ z_{m-1} & \vdots & \ddots & z_0 & z_{m+n} & \ddots & z_{m+1} \\ z_m & z_{m-1} & \ddots & z_1 & z_{m+n+1} & \ddots & z_{m+2} \\ & z_m & \ddots & \vdots & & \ddots & \vdots \\ & & \ddots & z_{m-1} & & \ddots & z_{m+n} \\ & & & z_m & & & z_{m+n+1} \end{array}\right]$$

The method of STLN is used to solve the following problem:

$$\min_z \|Dz\| \quad \text{where} \quad D = \begin{bmatrix} (n-k+1)I_{m+1} & 0 \\ 0 & (m-k+1)I_{n+1} \end{bmatrix}$$

such that

$$(A_k + E_k)\,x = c_k + h_k$$

and

- E_k has the same structure as A_k.
- h_k has the same structure as c_k.

Criteria for the acceptance of a structured low rank approximation

A solution is obtained for each value of α, and so how is the 'best' solution selected?

- Only the values of α for which the normalised residual $\|r_{norm}\|$ is less than 10^{-13} are retained.

$$\|r_{norm}\| = \frac{\|(A_k + E_k)\,x - (c_k + h_k)\|}{\|c_k + h_k\|}$$

- If the signal-to-noise ratio is μ, retain the values of α for which the perturbations satisfy

$$\|z_f\| \leq \frac{\|f\|}{\mu} \qquad \text{and} \qquad \frac{\|z_g\|}{|\alpha|} \leq \frac{\|g\|}{\mu}$$

- From the remaining values of α, select the one for which the ratio of the singular values

$$\frac{\sigma_{m+n-k}}{\sigma_{m+n-(k-1)}}$$

is a maximum.

4. Results

Example 1 Consider the polynomials

$$\hat{f}_1(y) = (y - 0.25)^8 (y - 0.5)^9 (y - 0.75)^{10} (y - 1)^{11} (y - 1.25)^{12}$$

and

$$\hat{g}_1(y) = (y + 0.25)^4 (y - 0.25)^5 (y - 0.5)^6$$

which have 11 common roots, from which it follows that rank $S(\hat{f}_1, \hat{g}_1) = 54$.

Noise was added to these polynomials, and the method of STLN was used to compute a structured low rank approximation of the Sylvester matrix formed from these noisy polynomials.

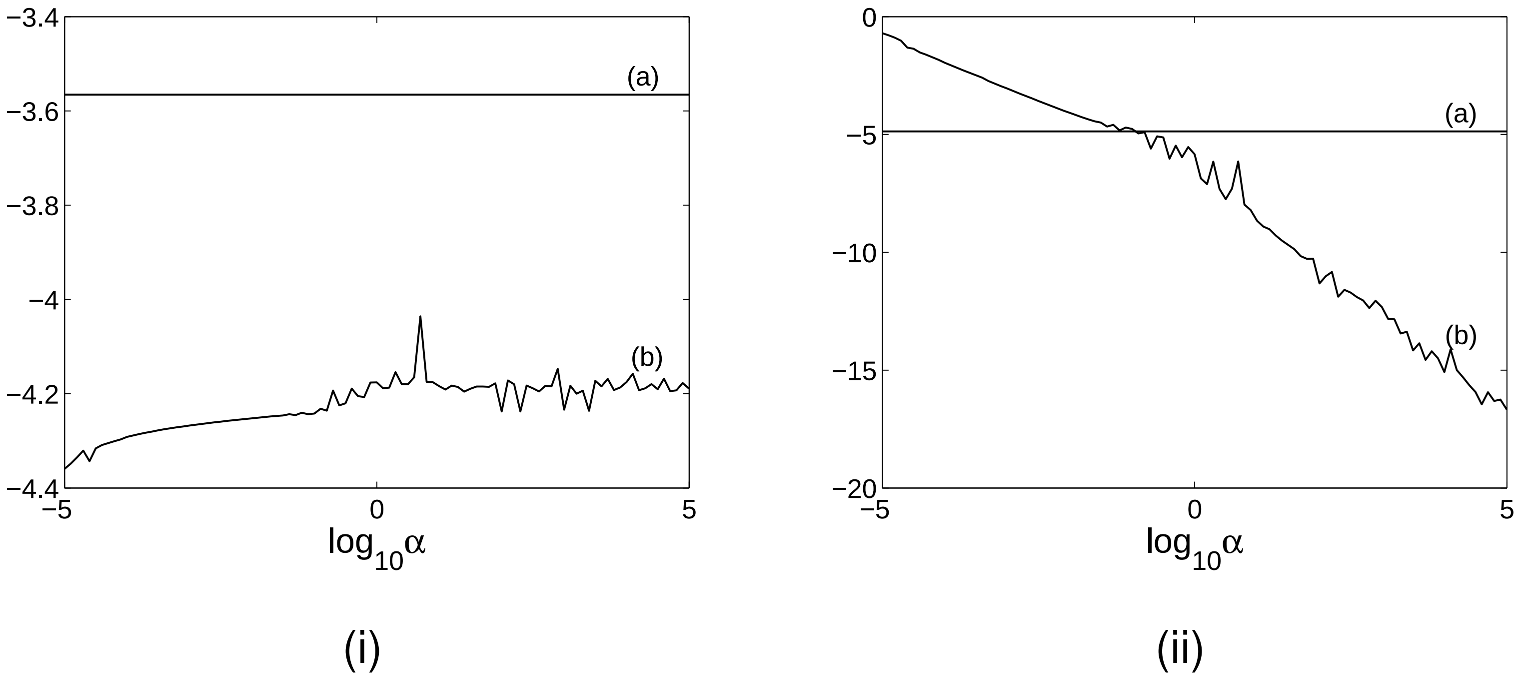
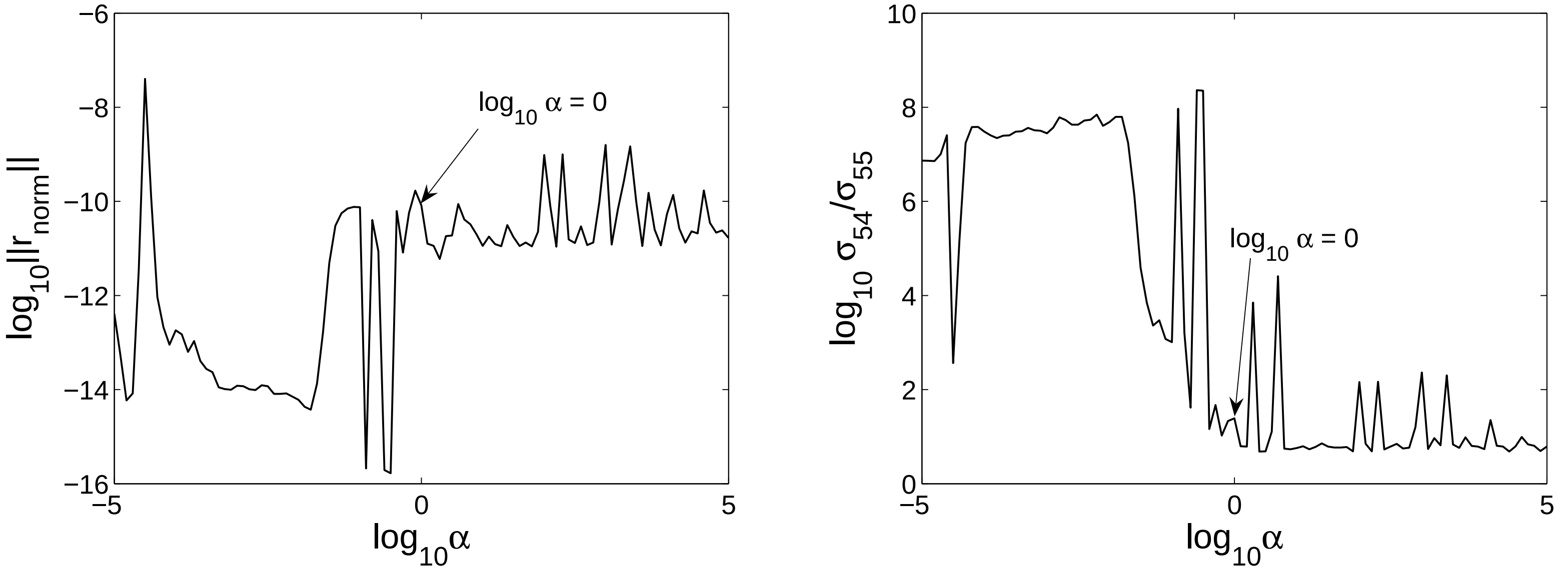


Figure 1: $\mu = 10^8$ and $k = 11$. (i)(a) The maximum allowable value of $\|z_{f_1}\|$, which is equal to $\|f_1\|/\mu$, (b) the computed value of $\|z_{f_1}\|$; (ii)(a) the maximum allowable value of $\|z_{g_1}\|/\alpha$, which is equal to $\|g_1\|/\mu$, (b) the computed value of $\|z_{g_1}\|/\alpha$.



(iii) (iv)

Figure 2: $\mu = 10^8$ and $k = 11$. (i) the normalised residual $\|r_{norm}\|$; (ii) the logarithm of the singular value ratio σ_{54}/σ_{55}.

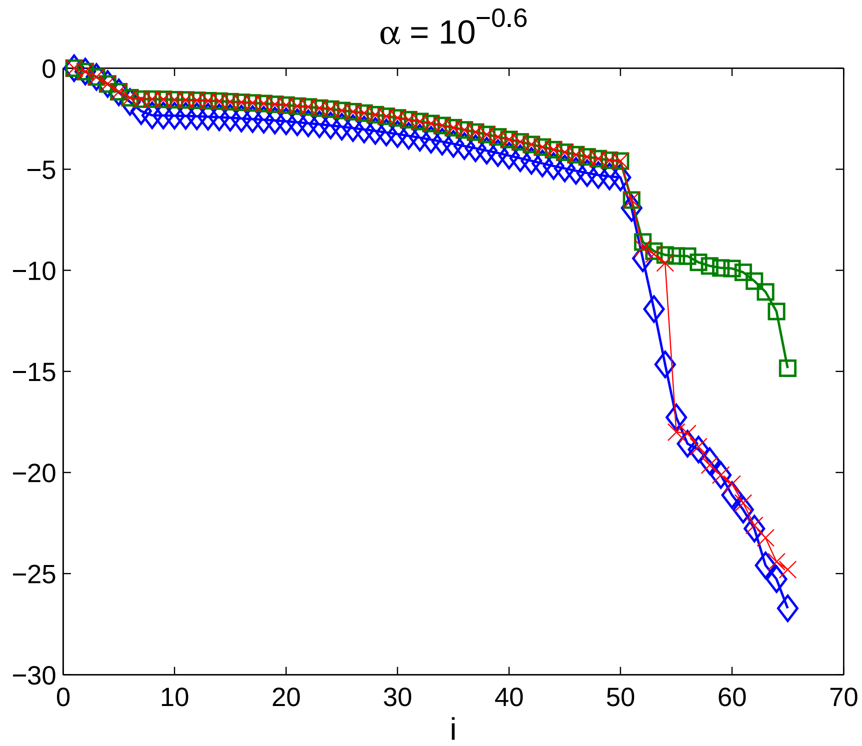


Figure 3: The normalised singular values, on a logarithmic scale, of the Sylvester matrix for (i) the theoretically exact data $\diamondsuit$; (ii) the given inexact data $\square$; (iii) the computed data $\times$, for $\alpha = 10^{-0.6}$.

Example 2 Consider the polynomials

$$\hat{f}_2(y) = (y-1)^8(y-2)^{16}(y-3)^{24}$$

and

$$\hat{g}_2(y) = (y-1)^{12}(y+2)^4(y-3)^8(y+4)^2$$

which have 16 common roots, and thus the rank of $S(\hat{f}_2, \hat{g}_2)$ is 58.

The polynomials were perturbed by noise, such that $\mu = 10^8$.

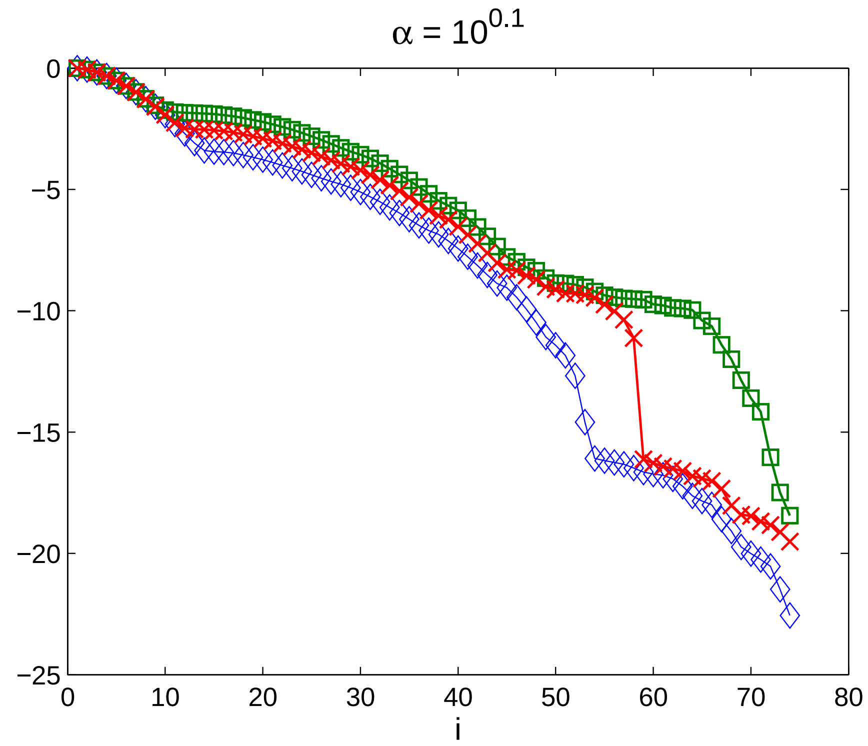


Figure 4: The normalised singular values, on a logarithmic scale, of the Sylvester matrix for (i) the theoretically exact data $\diamondsuit$; (ii) the given inexact data $\square$; (iii) the computed data $\times$, for $\alpha = 10^{0.1}$.

5. Conclusions

- A method for computing a structured low rank approximation of a Sylvester matrix has been described.
- The rank of the approximation can be selected.
- The introduction of an arbitrary scaling factor allows a family of low rank approximations to be constructed.
- Several criteria for accepting structured low rank approximations were used in order to eliminate unsatisfactory solutions.