A non-linear structure preserving matrix method for the low rank approximation of the Sylvester resultant matrix

Joab R. Winkler
Department of Computer Science
The University of Sheffield
United Kingdom

Charles University, Prague. May 2009
CONTENTS:

• The Sylvester resultant matrix
  – Properties of the Sylvester resultant matrix
  – Subresultant matrices of the Sylvester resultant matrix

• Polynomial scaling
  – Scaling by the geometric mean
  – Relative scaling of the polynomials

• Non-linear structure preserving matrix methods
  – The order of the polynomials

• Results

• Summary
1. THE SYLVESTER RESULTANT MATRIX

The Sylvester resultant matrix is used in several applications:

- Computer graphics
- Computer vision
- Geometric modelling
- Control theory

Geometric applications: The solution of curve and surface intersection problems

Control theory: Approximate pole-zero cancellation
Consider the polynomials

\[ f(y) = \sum_{i=0}^{m} a_i y^i \quad \text{and} \quad g(y) = \sum_{i=0}^{n} b_i y^i \]

where \( a_m, b_n \neq 0 \).

- Each entry of the Sylvester resultant matrix \( S(f, g) \in \mathbb{R}^{(m+n) \times (m+n)} \) contains either a coefficient \( a_i \) or a coefficient \( b_i \).

- \( S(f, g) \) is used extensively in GCD computations.
\[
S(f, g) = \begin{bmatrix}
    a_m & & & & \\
    a_{m-1} & a_m & & & \\
    & \ddots & \ddots & & \\
    a_0 & & \ddots & a_m & \\
    a_0 & \ddots & a_{m-1} & & \\
    & \ddots & & & \\
    a_0 & & & & \\
\end{bmatrix}
\begin{bmatrix}
    b_n & & & & \\
    b_{n-1} & b_n & & & \\
    & \ddots & \ddots & & \\
    b_0 & & \ddots & b_n & \\
    b_0 & \ddots & b_{n-1} & & \\
    & \ddots & & & \\
    b_0 & & & & \\
\end{bmatrix}
\]

- The coefficients \( a_i \) of \( f(y) \) occupy the first \( n \) columns.
- The coefficients \( b_i \) of \( g(y) \) occupy the last \( m \) columns.
The Sylvester resultant matrix is linear, which makes it easy to use structure preserving matrix algorithms:

- If $p(y)$ and $q(y)$ are polynomials of degrees $m$ and $n$ respectively, then

$$S(f + p, g + q) = S(f, g) + S(p, q)$$

**Theorem 1.1** Let the degree of the GCD of $f(y)$ and $g(y)$ be $d$, then

1. The rank of $S'(f, g)$ is equal to $m + n - d$.

2. The coefficients of their GCD are given in the last non-zero row of $S'(f, g)^T$ after it has been reduced to upper triangular form by an LU or QR decomposition.
Example 1.1  Consider the polynomials $\hat{f}(y)$ and $\hat{g}(y)$

$$\hat{f}(y) = -3y^3 + \frac{25}{2}y^2 - \frac{23}{2}y + 3 \quad \text{and} \quad \hat{g}(y) = 6y^2 - 7y + 2$$

whose GCD is $\hat{g}(y)$

$$\hat{f}(y) = -\frac{1}{2}(y - 3)\hat{g}(y)$$

The transpose $S(\hat{f}, \hat{g})^T$ of $S(\hat{f}, \hat{g})$ is

$$S(\hat{f}, \hat{g})^T = \begin{bmatrix}
-3 & \frac{25}{2} & -\frac{23}{2} & 3 & 0 \\
0 & -3 & \frac{25}{2} & -\frac{23}{2} & 3 \\
6 & -7 & 2 & 0 & 0 \\
0 & 6 & -7 & 2 & 0 \\
0 & 0 & 6 & -7 & 2 \\
\end{bmatrix}$$
Its reduction to row echelon (upper triangular) form yields the matrix

\[
\begin{bmatrix}
-3 & \frac{25}{2} & -\frac{23}{2} & 3 & 0 \\
0 & -3 & \frac{25}{2} & -\frac{23}{2} & 3 \\
0 & 0 & 6 & -7 & 2 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
\end{bmatrix}
\]

- The rank loss of this matrix is two, which is equal to the degree of the GCD of \( \hat{f}(y) \) and \( \hat{g}(y) \).
- The non-zero coefficients in its last non-zero row are 6, -7, 2, and thus the GCD is \( 6y^2 - 7y + 2 \), which is equal to \( \hat{g}(y) \), as required.
1.1 Properties of the Sylvester resultant matrix

The Sylvester matrix satisfies

\[ S(f, \alpha g) \neq \alpha S(f, g), \quad \alpha \neq 1 \]

and the singular values \( \sigma_i (S(f, \alpha g)) \) of \( S(f, \alpha g) \) satisfy

\[ \sigma_i (S(f, \alpha g)) \neq \alpha \sigma_i (S(f, g)), \quad i = 1, \ldots, m + n, \quad \alpha \neq 1 \]

Compare

\[ \text{rank } S(f, \alpha g) \neq \text{rank } S(f, g), \quad \alpha \neq 1 \]

with

\[ \text{GCD}(f, \alpha g) \sim \text{GCD}(f, g), \quad \alpha \neq 0 \]
Example 1.2 Consider the exact polynomials

\[ \hat{f}(y) = (y - 1)^2(y - 2)^3(y - 4) \quad \text{and} \quad \hat{g}(y) = (y - 1)(y - 2)^2(y - 6)^9 \]

Figure 1: The normalised singular values \( \sigma_i / \sigma_1 \) of the Sylvester matrix \( S(\hat{f}, \alpha \hat{g}) \) for \( \alpha = \alpha_1 = 1, \alpha = \alpha_2 = 10 \) and \( \alpha = \alpha_3 = 50 \) in the absence of noise.
1.2 Subresultant matrices of the Sylvester resultant matrix

The product of two polynomials is equal to the convolution of their coefficients:

\[
\begin{bmatrix}
  r_{m+n} \\
r_{m+n-1} \\
  \vdots \\
r_1 \\
r_0
\end{bmatrix}
=\begin{bmatrix}
p_m \\
p_{m-1} & p_m \\
  \vdots & & \ddots & & \ddots \\
p_0 & & \ddots & & p_m \\
p_0 & \ddots & \ddots & & p_{m-1} \\
p_0 & & \ddots & \ddots & \ddots \\
p_0 & & & \ddots & \ddots
\end{bmatrix}
\begin{bmatrix}
  q_n \\
  q_{n-1} \\
  \vdots \\
  q_1 \\
  q_0
\end{bmatrix}
\]

\[
r = S_{n+1}(p) q = p \otimes q
\]

\[
r \in \mathbb{R}^{m+n+1}, \ p \in \mathbb{R}^{m+1}, \ q \in \mathbb{R}^{n+1} \text{ and } S_{n+1}(p) \in \mathbb{R}^{(m+n+1) \times (n+1)}
\]
Let:

- \( d_k(y) \) be a common divisor of degree \( k \) of the exact polynomials \( \hat{f}(y) \) and \( \hat{g}(y) \)
- The degree of the GCD of \( \hat{f}(y) \) and \( \hat{g}(y) \) be \( \hat{d} \)
- \( u_k(y) \) and \( v_k(y) \) be the quotient polynomials

\[
\hat{f}(y) = u_k(y)d_k(y) \quad \text{and} \quad \hat{g}(y) = v_k(y)d_k(y), \quad k = 1, \ldots, \hat{d}
\]

Thus

\[
\hat{f}(y)v_k(y) - \hat{g}(y)u_k(y) = 0 \quad \iff \quad S_{n-k+1}(\hat{f})v_k - S_{m-k+1}(\hat{g})u_k = 0
\]

where

\[
S_{n-k+1}(\hat{f}) \in \mathbb{R}^{(m+n-k+1) \times (n-k+1)} \quad \text{and} \quad v_k \in \mathbb{R}^{n-k+1}
\]
\[
S_{m-k+1}(\hat{g}) \in \mathbb{R}^{(m+n-k+1) \times (m-k+1)} \quad \text{and} \quad u_k \in \mathbb{R}^{m-k+1}
\]
\[
\begin{bmatrix}
S_{n-k+1}(\hat{f}) & S_{m-k+1}(\hat{g})
\end{bmatrix}
\begin{bmatrix}
v_k \\
u_k
\end{bmatrix}
= 0 \quad \text{or} \quad
S_k(\hat{f}, \hat{g})
\begin{bmatrix}
v_k \\
u_k
\end{bmatrix}
= 0
\]

- \(S_k(\hat{f}, \hat{g}) \in \mathbb{R}^{(m+n-k+1) \times (m+n-2k+2)}\) and it is rank deficient

- The nullspace vectors yield the coefficients of the quotient polynomials for \(k = 1, \ldots, \hat{d}\)

- Since the degree of the GCD of \(\hat{f}(y)\) and \(\hat{g}(y)\) is \(\hat{d}\), these polynomials possess common divisors of degrees 1, 2, \ldots, \(\hat{d}\), but not a divisor of degree \(\hat{d} + 1\):

  \[
  \begin{align*}
  \text{rank } S_k(\hat{f}, \hat{g}) &< m + n - 2k + 2, \quad k = 1, \ldots, \hat{d} \\
  \text{rank } S_k(\hat{f}, \hat{g}) &= m + n - 2k + 2, \quad k = \hat{d} + 1, \ldots, \min(m, n)
  \end{align*}
  \]

Calculating the degree of the GCD reduces to estimating the rank of a matrix.
Example 1.3 Consider $S_k(\hat{f}, \hat{f}^{(1)})$, for $k = 1, 2, 3$, where

\[
\hat{f}(y) = (y - 1)^2(y - 2)(y - 3) = y^4 - 7y^3 + 17y^2 - 17y + 6
\]
\[
\hat{f}^{(1)}(y) = 4y^3 - 21y^2 + 34y - 17
\]

Hence $S_1(\hat{f}, \hat{f}^{(1)}) = S(\hat{f}, \hat{f}^{(1)})$ is equal to

\[
\begin{bmatrix}
1 & 0 & 0 & 4 & 0 & 0 & 0 & 0 \\
-7 & 1 & 0 & -21 & 4 & 0 & 0 & 0 \\
17 & -7 & 1 & 34 & -21 & 4 & 0 & 0 \\
-17 & 17 & -7 & -17 & 34 & -21 & 4 & 0 \\
6 & -17 & 17 & 0 & -17 & 34 & -21 & 0 \\
0 & 6 & -17 & 0 & 0 & -17 & 34 & 0 \\
0 & 0 & 6 & 0 & 0 & 0 & -17 & 0
\end{bmatrix}
\]

and this matrix has unit loss of rank.
The subresultant matrix $S_2(\hat{f}, \hat{f}^{(1)})$ is

$$S_2(\hat{f}, \hat{f}^{(1)}) = \begin{bmatrix}
1 & 0 & 4 & 0 & 0 \\
-7 & 1 & -21 & 4 & 0 \\
17 & -7 & 34 & -21 & 4 \\
-17 & 17 & -17 & 34 & -21 \\
6 & -17 & 0 & -17 & 34 \\
0 & 6 & 0 & 0 & -17
\end{bmatrix}$$

and this matrix has full column rank.
The subresultant matrix $S_3(\hat{f}, \hat{f}^{(1)})$ is

$$S_3(\hat{f}, \hat{f}^{(1)}) = \begin{bmatrix} 4 & 0 & 1 \\ -21 & 4 & -7 \\ 34 & -21 & 17 \\ -17 & 34 & -17 \\ 0 & -17 & 6 \end{bmatrix}$$

and this matrix has full column rank.

It follows that the first rank deficient matrix in the sequence

$$S_3(\hat{f}, \hat{f}^{(1)}), S_2(\hat{f}, \hat{f}^{(1)}), S_1(\hat{f}, \hat{f}^{(1)})$$

is $S_1(\hat{f}, \hat{f}^{(1)})$, and thus the degree of the GCD of $\hat{f}(y)$ and $\hat{f}^{(1)}(y)$ is one. □
Recall

\[
\begin{bmatrix}
S_{n-k+1}(\hat{f}) & S_{m-k+1}(\hat{g})
\end{bmatrix}
\begin{bmatrix}
v_k \\
-u_k
\end{bmatrix}
= 0 \quad \text{or} \quad S_k(\hat{f}, \hat{g})
\begin{bmatrix}
v_k \\
-u_k
\end{bmatrix}
= 0
\]

Since \( \hat{f}(y) \) and \( \hat{g}(y) \) have common divisors of degrees \( k = 1, \ldots, \hat{d} \), this homogeneous equation can be transformed to a linear algebraic equation by setting \( v_{k,0} = -1 \), that is, the coefficient of \( y^{n-k} \) is set equal to \(-1\), in which case the homogeneous equation becomes

\[
A_k x = c_k, \quad k = 1, \ldots, \hat{d}
\]

where

- \( c_k \in \mathbb{R}^{m+n-k+1} \) is the first column of \( S_k \)
- \( A_k \in \mathbb{R}^{(m+n-k+1) \times (m+n-2k+1)} \) is formed from the remaining \( m + n - 2k + 1 \) columns of \( S_k \)
\[ S_k(\hat{f}, \hat{g}) = \begin{bmatrix} c_k & A_k \end{bmatrix} \]

and

\[ x = \begin{bmatrix} v_{k,1} & \cdots & v_{k,n-k} & -u_{k,0} & \cdots & -u_{k,m-k} \end{bmatrix}^T \in \mathbb{R}^{m+n-2k+1} \]

Properties of the equation \( A_k x = c_k \)

when exact polynomials are specified and all computations are performed exactly:

- It has an infinite number of solutions for \( k = 1, \ldots, \hat{d} - 1 \).
- It has exactly one solution for \( k = \hat{d} \).
- It has no solution for \( k = \hat{d} + 1, \ldots, \min(m,n) \).
When inexact polynomials are specified, the equation

\[ A_k x = c_k \]

does not have a solution for all values of \( k = 1, \ldots, \min(m, n) \).

If the polynomials are inexact and it is assumed that their theoretically exact forms have a non-constant GCD, it is necessary to construct a structured low rank approximation of their Sylvester matrix.
2. POLYNOMIAL SCALING

Computations on polynomials whose coefficients vary widely in magnitude are problematic.

- Perform two preprocessing operations on the polynomials before the AGCD computations. These operations minimise the ratio

\[
\frac{\max \{\max |\text{coeff. of } f(y)|, \max |\text{coeff. of } g(y)|\}}{\min \{\min |\text{coeff. of } f(y)|, \min |\text{coeff. of } g(y)|\}}
\]

The operations are:

- Scale the coefficients of each polynomial by the geometric mean of its coefficients
- Introduce a parameter substitution
2.1 Scaling by the geometric mean

- It is numerically advantageous for the coefficients of $S(f, g)$ to have the same magnitude.
- Many researchers scale the coefficients of $f(y)$ and $g(y)$ to have unit 2-norm because it yields a Sylvester matrix that is better conditioned.
- Normalisation by the geometric mean of the coefficients is better for polynomials whose coefficients vary greatly in magnitude because it yields a ‘better average’ than does normalisation by the 2-norm of the coefficients.
The polynomials are therefore redefined as

\[ f(y) = \sum_{i=0}^{m} a_i^* y^{m-i}, \quad a_i^* = \frac{a_i}{\left( \prod_{j=0}^{m} |a_j| \right)^{1/m+1}} \neq 0 \]

and

\[ g(y) = \sum_{i=0}^{n} b_i^* y^{n-i}, \quad b_i^* = \frac{b_i}{\left( \prod_{j=0}^{n} |b_j| \right)^{1/n+1}} \neq 0 \]

- Normalisation by the geometric mean is not required for the Bézout resultant matrix.
- The parameter \( \alpha \) in \( S(f, \alpha g) \) is interpreted as the weight of \( g(y) \) relative to the weight of \( f(y) \).
2.2 Relative scaling of the polynomials

Scale the independent variable \( y \) by

\[
y = \theta w
\]

and thus the transformed polynomials \( f^*(w) \) and \( g^*(w) \) are

\[
f^*(w) = \sum_{i=0}^{m} (a_i^* \theta^{m-i}) w^{m-i} \quad \text{and} \quad g^*(w) = \sum_{i=0}^{n} (b_i^* \theta^{n-i}) w^{n-i}
\]

The ratio of the maximum coefficient of \( \{f^*(w), \alpha g^*(w)\} \) to the minimum coefficient of \( \{f^*(w), \alpha g^*(w)\} \), that is, the arguments of \( S(f^*, \alpha g^*) \), is

\[
\frac{\max \{ \max_{i=0,...,m} |a_i^* \theta^{m-i}| , \max_{j=0,...,n} |\alpha b_j^* \theta^{n-j}| \} }{\min \{ \min_{i=0,...,m} |a_i^* \theta^{m-i}| , \min_{j=0,...,n} |\alpha b_j^* \theta^{n-j}| \} }
\]
The values $\alpha_0$ and $\theta_0$ that minimise this ratio are the optimal values of $\alpha$ and $\theta$, respectively.

Pose this minimisation as a linear programming (LP) problem:

Minimise $\frac{t}{s}$

Subject to

\[
\begin{align*}
  t &\geq |a_i^*| \theta^{m-i}, \quad i = 0, \ldots, m \\
  t &\geq \alpha |b_j^*| \theta^{n-j}, \quad j = 0, \ldots, n \\
  s &\leq |a_i^*| \theta^{m-i}, \quad i = 0, \ldots, m \\
  s &\leq \alpha |b_j^*| \theta^{n-j}, \quad j = 0, \ldots, n \\
  s &> 0 \\
  \theta &> 0 \\
  \alpha &> 0
\end{align*}
\]
The transformations

\[ T = \log t \quad S = \log s \quad \phi = \log \theta \]
\[ \mu = \log \alpha \quad \bar{\alpha}_i = \log |a_i^*| \quad \bar{\beta}_j = \log |b_j^*| \]

transform this constrained minimisation problem to:

Minimise \( T - S \)

Subject to

\[ T - (m - i)\phi \geq \bar{\alpha}_i, \quad i = 0, \ldots, m \]
\[ T - (n - j)\phi - \mu \geq \bar{\beta}_j, \quad j = 0, \ldots, n \]
\[ -S + (m - i)\phi \geq -\bar{\alpha}_i, \quad i = 0, \ldots, m \]
\[ -S + (n - j)\phi + \mu \geq -\bar{\beta}_j, \quad j = 0, \ldots, n \]
which is a standard linear programming (LP) problem, whose objective function is

\[
T - S = \begin{bmatrix} 1 & -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} T \\ S \\ \phi \\ \mu \end{bmatrix}
\]

The polynomials on which computations are performed are

\[
\tilde{f}(w) = \sum_{i=0}^{m} \tilde{a}_i w^{m-i} \quad \text{and} \quad \tilde{g}(w) = \sum_{i=0}^{n} \tilde{b}_i w^{n-i}
\]

where

\[
\tilde{a}_i = a_i^* \theta_0^{m-i} \quad \text{and} \quad \tilde{b}_i = \alpha_0 b_i^* \theta_0^{n-i}
\]
3. NON-LINEAR STRUCTURE PRESERVING MATRIX METHODS

Consider the inexact polynomials

\[ \bar{f}(w) = \sum_{i=0}^{m} (\bar{a}_i \theta^{m-i}) w^{m-i} \]

and

\[ \bar{g}(w) = \sum_{i=0}^{n} (\bar{b}_i \theta^{n-i}) w^{n-i} \]

where

\[ \bar{a}_i = \frac{\tilde{a}_i}{\left( \prod_{j=0}^{m} |\tilde{a}_j \theta^{m-j}| \right)^{\frac{1}{m+1}}} \]

and

\[ \bar{b}_i = \frac{\tilde{b}_i}{\left( \prod_{j=0}^{n} |\tilde{b}_j \theta^{n-j}| \right)^{\frac{1}{n+1}}} \]
Problem statement:

Given the inexact polynomials $\bar{f}(w)$ and $\bar{g}(w)$, which are assumed to be coprime, calculate the smallest perturbations to their coefficients such that the perturbed forms of $\bar{f}(w)$ and $\bar{g}(w)$ have a non-constant GCD.

Aim:

Compute the Sylvester matrix $S'(\delta \bar{f}, \alpha \delta \bar{g})$, such that

$$\| \delta \bar{f} \|^2 + \| \delta \bar{g} \|^2$$

is minimised, where

$$S(\bar{f} + \delta \bar{f}, \alpha (\bar{g} + \delta \bar{g})) = S(\bar{f}, \alpha \bar{g}) + S(\delta \bar{f}, \alpha \delta \bar{g})$$

is rank deficient.

Method:

Use the method of non-linear structured total least norm (SNTLN).
The Sylvester matrix $S(\bar{f}, \alpha \bar{g})$ is

$$
\begin{bmatrix}
\bar{a}_0 \theta^m & \alpha \bar{b}_0 \theta^n \\
\bar{a}_1 \theta^{m-1} & \alpha \bar{b}_1 \theta^{n-1} \\
\vdots & \vdots & \ddots & \vdots \\
\bar{a}_{m-1} \theta & \alpha \bar{b}_{n-1} \theta \\
\bar{a}_m & \alpha \bar{b}_n \\
\end{bmatrix}
$$
If the perturbations of the coefficients of $\bar{f}(w)$ and $\alpha \bar{g}(w)$ are

$$z_i \theta^{m-i}, i = 0, \ldots, m \quad \text{and} \quad \alpha z_{m+1+i} \theta^{n-i}, i = 0, \ldots, n$$

respectively, then $S(\delta \bar{f}, \alpha \delta \bar{g})$ is equal to

\[
\begin{bmatrix}
  z_0 \theta^m \\
  z_1 \theta^{m-1} & \ddots & \alpha z_{m+1} \theta^n \\
  \vdots & \ddots & \ddots & \ddots \\
  z_{m-1} \theta & \alpha z_{m+n} & \alpha z_{m+2} \theta^{n-1} \\
  z_m & \alpha z_{m+n+1} & \ddots & \ddots \\
  \vdots & \alpha z_{m+n} & \alpha z_{m+n+1} & \alpha z_{m+n+1}
\end{bmatrix}
\]
Recall

- \( S_k(\bar{f}, \bar{g}) = \begin{bmatrix} c_k & A_k \end{bmatrix} = \begin{bmatrix} c_k & \text{coeffs. of } \bar{f} & \text{coeffs. of } \alpha \bar{g} \end{bmatrix} \)

  where \( c_k = c_k(\theta) \) and \( A_k = A_k(\alpha, \theta) \).

- \( A_k x = c_k, \quad k = 1, \ldots, \min(m, n) \)

  does not have a solution because \( \bar{f}(w) \) and \( \bar{g}(w) \) are assumed to be coprime.

- In the presence of noise, it is necessary to perturb this equation to

  \((A_k(\alpha, \theta) + E_k(\alpha, \theta, z)) x = c_k(\theta) + h_k(\theta, z)\)

  where

  \( S_k(\delta \bar{f}, \delta \bar{g}) = \begin{bmatrix} h_k & E_k \end{bmatrix} = \begin{bmatrix} h_k & \text{coeffs. of } \delta \bar{f} & \text{coeffs. of } \alpha \delta \bar{g} \end{bmatrix} \)
The equation

\[
\left( A_k(\alpha, \theta) + E_k(\alpha, \theta, z) \right) x = c_k(\theta) + h_k(\theta, z)
\]

is non-linear in the unknowns \( \alpha, \theta \) and \( z \).

- \( A_k \) and \( E_k \) have the same structure, and \( c_k \) and \( h_k \) have the same structure:
  - Use the method of structured non-linear total least norm

- The initial values of \( \alpha \) and \( \theta \) are \( \alpha_0 \) and \( \theta_0 \)

- The initial vector of perturbations is \( z = 0 \)

- Solve this non-linear equation subject to the constraint that \( \|z\|^2 \) is minimised

- This leads to a least squares equality constrained problem
3.1 The order of the polynomials

The theory presented above considered the Sylvester matrix $S(f, \alpha g)$, but $S(g, f/\alpha)$ is also a Sylvester matrix of $f(y)$ and $g(y)$, where $\alpha$ retains its definition as the weight of $g(y)$ relative to the weight of $f(y)$.

- Are the results obtained with $S(f, \alpha g)$ the same as the results obtained with $S(g, f/\alpha)$?

  No: $S(\bar{f}, \alpha \bar{g})$ may be a structured low rank approximation of $S(f, \alpha g)$, but $S(\bar{g}, \bar{f}/\alpha)$ may not be a structured low rank approximation of $S(g, f/\alpha)$ because its numerical rank is not defined.

- Are the computations with $S(\bar{f}, \alpha \bar{g})$ and $S(\bar{g}, \bar{f}/\alpha)$ identical?

  Yes: The optimal values of $\alpha$ and $\theta$ in the matrix $S(\bar{g}, \bar{f}/\alpha)$ are $\alpha_0$ and $\theta_0$. 
4. RESULTS

Notation used in examples:

- $\hat{f}(y)$ and $\hat{g}(y)$ are the theoretically exact polynomials, and $S(\hat{f}, \hat{g})$ and $S(\hat{g}, \hat{f})$ are calculated by normalising each polynomial by the geometric mean of its coefficients.

- $f(y)$ and $g(y)$ are calculated from $\hat{f}(y)$ and $\hat{g}(y)$ by adding noise and normalising these inexact polynomials by the geometric mean of their coefficients.

- $\bar{f}(w)$ and $\bar{g}(w)$ are the polynomials, the coefficients of which form the entries of the Sylvester matrix whose structured low rank approximation is computed.
Example 4.1 The Sylvester matrix of the exact polynomials

\[
\hat{f}(y) = (y - 10^{-5})^3(y - 3.1 \times 10^{-3})^3(y - 3.2 \times 10^{-3})^3(y - 5)^{15}
\]
\[
\hat{g}(y) = (y - 3.1 \times 10^{-3})^4(y - 3.2 \times 10^{-3})^3(y + 3.3 \times 10^6)^{10}
\]

is of order $41 \times 41$, and since their GCD is of degree 6, it follows that $\text{rank } S(\hat{f}, \hat{g}) = 35$. Noise with a normwise signal-to-noise ratio of $10^8$ was added to the polynomials, which were then normalised, thereby yielding the polynomials $f(y)$ and $g(y)$.

Figure 2 shows the results obtained using a linear structure preserving matrix method, and $\alpha = 1$ and $\theta = 1$ are constant, but each polynomial is normalised by the geometric mean of its coefficients. It is seen that

\[
\text{rank } S(\hat{f}, \hat{g}) = \text{rank } S(f, g) = \text{rank } S(\tilde{f}, \tilde{g}) = 24
\]

which is the incorrect value.
Figure 2: (a) The normalised singular values of the Sylvester matrices $S(\hat{f}, \hat{g})$ $\Diamond$; $S(f, g)$ $\Box$; $S(\tilde{f}, \tilde{g})$ $\times$, and (b) the normalised residual. The preprocessing operations, apart from normalising each polynomial by the geometric mean of its coefficients, are omitted.
Figure 3: The magnitude of the coefficients of (a) $f(y)$ and (b) $g(y)$ before, $\Diamond$, and after, $\times$, scaling by $\alpha$ and $\theta$.
Figure 4: The normalised singular values of the Sylvester matrices (a) $S(\hat{f}, \hat{g}) \Diamond$; $S(f, g) \Box$; $S(\tilde{f}, \alpha^* \tilde{g}) \times$, and (b) $S(\hat{g}, \hat{f}) \Diamond$; $S(g, f) \Box$; $S(\tilde{g}, \tilde{f}/\alpha^*) \times$. 

37
Figure 5: The variation of the normalised residual with the number of iterations for (a) $S(f, g)$ and (b) $S(g, f)$. 
Example 4.2  The Sylvester matrix of the exact polynomials

\[
\hat{f}(y) = (y - 1.8722181 \times 10^7)^5(y - 0.3124444)^2 \\
\times (y - 4.4199430 \times 10^5)^7 \\
\hat{g}(y) = (y - 1.8722181 \times 10^7)^2(y - 0.3124444)^6(y - 8.8081342)^2
\]

is of order $40 \times 40$, and since their GCD is of degree 4, it follows that \( \text{rank } S(\hat{f}, \hat{g}) = 36 \). Noise with a normwise signal-to-noise ratio of $10^8$ was added to the polynomials, which were then normalised, thereby yielding the polynomials \( f(y) \) and \( g(y) \).

Figure 6 shows the results obtained using a linear structure preserving matrix method, and \( \alpha = 1 \) and \( \theta = 1 \) are constant, but each polynomial is normalised by the geometric mean of its coefficients. It is seen that

\[
\text{rank } S(\hat{f}, \hat{g}) = \text{rank } S(f, g) = \text{rank } S(\tilde{f}, \tilde{g}) = 26
\]

which is the incorrect value. A similar result is obtained for \( S(g, f) \).
Figure 6: (a) The normalised singular values of the Sylvester matrices $S(\hat{f}, \hat{g})$ ♦; $S(f, g)$ □; $S(\tilde{f}, \tilde{g})$ ×, and (b) the normalised residual. The preprocessing operations, apart from normalising each polynomial by the geometric mean of its coefficients, are omitted.
Figure 7: The magnitude of the coefficients of (a) $f(y)$ and (b) $g(y)$ before, ♦, and after, ×, scaling by $\alpha$ and $\theta$. 
Figure 8: (a) The normalised singular values of the Sylvester matrices $S(\hat{g}, \hat{f}) \Diamond$; $S(g, f) \Box$; $S(\tilde{g}, \tilde{f}/\alpha^*) \times$, and (b) the variation of the normalised residual with the number of iterations for $S(g, f)$. 
5. SUMMARY

- A non-linear structure preserving matrix method for the computation of a structured low rank approximation of the Sylvester matrix $S(f, g)$ has been discussed.

- The results show that it is important to preprocess the polynomials, thereby introducing the parameters $\alpha$ and $\theta$, in order to compute a structured low rank approximation of $S(f, g)$.

- The polynomial order, that is, $(f, g)$ or $(g, f)$, is important because the successful computation of a structured low rank approximation of $S(f, g)$ does not guarantee that a structured low rank approximation of $S(g, f)$ can be computed.

- It is necessary to determine *a priori* the best matrix, $S(f, g)$ or $S(g, f)$, to use to compute a structured low rank approximation of the Sylvester matrix of $f(y)$ and $g(y)$. 