Available online at www.sciencedirect.com
ScienceDirect
Journal of Computational and Applied Mathematics 213 (2008) 212–223
JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS
www.elsevier.com/locate/cam

A structured rank-revealing method for Sylvester matrix

Bingyu Li[a,*], Zhuojun Liu[b], Lihong Zhi[b]

[a] *School of Mathematics and Statistics, Northeast Normal University, Changchun 130024, China*
[b] *Key Laboratory of Mathematics Mechanization, AMSS, Beijing 100080, China*

Received 12 December 2004; received in revised form 18 December 2006

Abstract

We propose a fast algorithm for computing the numeric ranks of Sylvester matrices. Let S denote the Sylvester matrix and H denote the Hankel-like-Sylvester matrix. The algorithm is based on a fast Cholesky factorization of $S^{\mathrm{T}}S$ or $H^{\mathrm{T}}H$ and relies on a stabilized version of the generalized Schur algorithm for matrices with displacement structure. All computations can be done in $\mathrm{O}(r(n+m))$, where $n+m$ and r denote the size and the numerical rank of the Sylvester matrix, respectively.

MSC: 68W30; 65Y20

Keywords: Numeric rank; Singular value decomposition; Generalized Schur algorithm; Cholesky factorization

1. Introduction

Given univariate polynomials $f(x), g(x) \in \mathbb{R}[x]$, where $f(x) = f_n x^n + f_{n-1}x^{n-1} + \cdots + f_1 x + f_0$, $f_n \neq 0$ and $g(x) = g_m x^m + g_{m-1}x^{m-1} + \cdots + g_1 x + g_0$, $g_m \neq 0$. Let $S(f,g) \in \mathbb{R}^{(n+m)\times(n+m)}$ denote the Sylvester matrix of $f(x)$ and $g(x)$,

$$S(f,g) = \left[\begin{array}{cccc|cccc} f_n & & & & g_m & & & \\ f_{n-1} & f_n & & & g_{m-1} & g_m & & \\ \vdots & f_{n-1} & \ddots & & \vdots & g_{m-1} & \ddots & \\ f_1 & \vdots & \ddots & f_n & g_1 & \vdots & \ddots & g_m \\ f_0 & f_1 & & f_{n-1} & g_0 & g_1 & & g_{m-1} \\ & f_0 & \ddots & \vdots & & g_0 & \ddots & \vdots \\ & & \ddots & f_1 & & & \ddots & g_1 \\ & & & f_0 & & & & g_0 \end{array}\right].$$

$$\underbrace{\qquad\qquad\qquad}_{m}\qquad\underbrace{\qquad\qquad\qquad}_{n}$$

* Corresponding author. Tel.: +86 1062651369; fax: +86 1062630706.
E-mail addresses: liby@mmrc.iss.ac.cn (B. Li), zliu@mmrc.iss.ac.cn (Z. Liu), lzhi@mmrc.iss.ac.cn (L. Zhi).

0377-0427/$ - see front matter
doi:10.1016/j.cam.2007.01.032

Here and hereafter, we assume that the coefficients of input polynomials are only known inexactly. For a given input error tolerance ε, the computation of the ε-GCD (greatest common divisor) has been considered in [6–8,22,21]. For computing ε-GCD, a preliminary knowledge on the degree of ε-GCD is important. The following lemma in [6] provides us an upper bound on the degree of ε-GCD.

Lemma 1 (*Corless et al. [6]*). *If the singular values of $S(f,g)$ are σ_i, $i=1,\ldots,m+n$, and $\sigma_1 \geqslant \cdots \geqslant \sigma_r > \varepsilon\sqrt{m+n} > \varepsilon \geqslant \sigma_{r+1} \geqslant \cdots \geqslant \sigma_{m+n}$ (in other words, the numerical ε-rank of S is r), and if $\hat{d}$ is a common divisor of $f+\triangle f$ and $g+\triangle g$ with $\deg(\hat{d}) \geqslant m+n-r+1$, then one of $\|\triangle f\| > \varepsilon$ or $\|\triangle g\| > \varepsilon$ holds, where $\|\cdot\|$ denotes the Euclidean 2-norm of a polynomial coefficients vector.*

For the computation of numerical ε-rank of a Sylvester matrix, a full singular value decomposition (SVD) is undoubtedly most reliable but expensive. In this paper, a fast rank-revealing algorithm is obtained by exploiting the displacement structure of Sylvester matrices. The algorithm is based on the fast Cholesky factorization of $S^{\mathrm{T}}S$ or $H^{\mathrm{T}}H$ and costs $\mathrm{O}(r(n+m))$ flops, where r denotes the numerical ε-rank of the Sylvester matrix. Here H is the Hankel variation of the Sylvester matrix S. We introduce H in order to complete the fast Cholesky factorization without pivoting. Suppose $\hat{R}_r$ is the Cholesky factor of $S^{\mathrm{T}}S$ or $H^{\mathrm{T}}H$ produced by the first r steps of the generalized Schur algorithm [13,4,5],

$$\hat{R}_r = \begin{bmatrix} \hat{R}_{11}^{(r)} & \hat{R}_{12}^{(r)} \\ \mathbf{0} & \mathbf{0} \end{bmatrix},$$

where $\hat{R}_{11}^{(r)}$ is an $r \times r$ upper triangular matrix. We declare r is a candidate ε-rank of the Sylvester matrix when $\|S^{\mathrm{T}}S - \hat{R}_r^{\mathrm{T}}\hat{R}_r\|$ or $\|H^{\mathrm{T}}H - \hat{R}_r^{\mathrm{T}}\hat{R}_r\|$ is small enough.

Following the lead of [3–5,9], we derive a backward error bound for the fast Cholesky factorization. It allows us to set a termination tolerance in advance to compute ε-rank of a Sylvester matrix.

The organization of this paper is as follows. In Section 2, we give a brief introduction for the displacement structure and the generalized Schur algorithm. Then we discuss Cholesky factorizations of the semi-definite matrices $S^{\mathrm{T}}S$ and $H^{\mathrm{T}}H$. We also show how to avoid early breakdown of the generalized Schur algorithm. In Section 3, we present error analysis needed for our fast algorithm. In Section 4, we derive a fast rank-revealing algorithm. Numerical experiments are presented in Section 5.

1.1. Notation

In the discussion that follows we use $\|\cdot\|$ to denote the 2-norm of its argument. Let i be a positive integer, Z_i be an $i \times i$ lower shift matrix and I_i be an $i \times i$ identity matrix. For matrices, we denote their singular values by $\sigma_i(\cdot)$ and eigenvalues by $\lambda_i(\cdot)$. The $\hat{\cdot}$ notation denotes the computed quantities, and u denotes the machine precision.

2. Displacement structure and the generalized Schur algorithm

The displacement of an $n \times n$ Hermitian matrix T is defined as

$$\nabla T = T - ZTZ^{\mathrm{T}}, \tag{1}$$

where Z is an $n \times n$ lower-triangular matrix. The choice of Z depends on the matrix T, e.g., if T is a Toeplitz matrix, then Z is chosen equal to a lower shift matrix Z_n. If ∇T has a lower rank α ($\ll n$) independent of n, the size of T, then T is said to be structured with respect to the displacement defined by (1) and α is referred to as the displacement rank of T. It follows that ∇T can be factored as

$$\nabla T = GJG^{\mathrm{T}},$$

where G is an $n \times \alpha$ matrix and J is a signature matrix of the form

$$J = \begin{bmatrix} I_{p'} & \mathbf{0} \\ \mathbf{0} & -I_{q'} \end{bmatrix}, \quad p'+q'=\alpha.$$

The integers p', q' denote the numbers of positive eigenvalues and negative eigenvalues of ∇T, respectively. The pair (G, J) is said to be a generator pair for T.

When T is positive definite, the Cholesky factorization of T can be efficiently carried out by a generalized Schur algorithm that operates on the generator pair (G, J) directly and costs $O(\alpha n^2)$. The notion of displacement structure can be extended to non-Hermitian matrices. A more extensive description about the generalized Schur algorithm can be found in [13].

As discussed in [22,12], a Sylvester matrix $S(f, g)$ is a quasi-Toeplitz matrix with displacement rank at most 2. Furthermore, we have

Theorem 1 (*Zhi [22]*). *$S^{\mathrm{T}}S$ is a structured matrix with displacement rank at most* 4.

In fact, let $Z = \mathrm{diag}(Z_m, Z_n)$, $J = \mathrm{diag}(1, 1, -1, -1)$, then

$$S^{\mathrm{T}}S - ZS^{\mathrm{T}}SZ^{\mathrm{T}} = GJG^{\mathrm{T}}, \tag{2}$$

where $G = [\mathbf{x}_1\ \mathbf{x}_2\ \mathbf{x}_3\ \mathbf{x}_4]$ with

$$\begin{aligned}
&\mathbf{x}_1 = S^{\mathrm{T}}S(:, 1)/\|S(:, 1)\|,\\
&\mathbf{x}_2 = S^{\mathrm{T}}S(:, m+1)/\|S(:, m+1)\|, \quad \mathbf{x}_2[1] = 0,\\
&\mathbf{x}_3 = \mathbf{x}_1 \quad \text{except that } \mathbf{x}_3[1] = 0,\\
&\mathbf{x}_4 = \mathbf{x}_2 \quad \text{except that } \mathbf{x}_4[m+1] = 0.
\end{aligned}$$

2.1. Avoiding early breakdown of the generalized Schur algorithm

Assume that the numeric ε-rank of S is r. The generalized Schur algorithm applied to compute the fast Cholesky factorization of $S^{\mathrm{T}}S$ may break down during the first r steps due to the loss of positive definiteness of the $r \times r$ leading principal submatrix of $S^{\mathrm{T}}S$.

Example 1. Let

$$\begin{aligned}
&f(x) = 2x^2 + 3x - x^4 - 2x^3,\\
&g(x) = 3x + 2 + x^2,
\end{aligned}$$

then the Sylvester matrix $S(f, g)$ is of rank 5. Since the leading 5×5 principal submatrix of $S(f, g)^{\mathrm{T}}S(f, g)$ is of rank 4, the generalized Schur algorithm applied to $S^{\mathrm{T}}S$ completes the first four steps but breaks down at the fifth step.

However, for the above example, if we construct the Sylvester matrix as $S(g, f)$, we can check that the leading 5×5 principal submatrix of $S^{\mathrm{T}}(g, f)S(g, f)$ is of full rank and the generalized Schur algorithm completes the first five steps successfully. In general, we have the following theorem.

Theorem 2. *Let $S(f, g)$ be the Sylvester matrix for two univariate polynomials $f(x)$ and $g(x)$ with degrees n and m, respectively. Suppose $f(x) = x^p(f_n x^{n-p} + \cdots + f_p)$ and $g(x) = x^q(g_m x^{m-q} + \cdots + g_q)$, where $f_p \neq 0$ and $g_q \neq 0$. If $0 \leqslant p \leqslant q$ and $\mathrm{rank}(S) = m + n - d$ for $p \leqslant d \leqslant \min\{m, n\}$, then the first $m + n - d$ column vectors of $S(f, g)$ are of full column rank.*

Proof. Note that $\mathrm{rank}(S) = m + n - d$ implies $\gcd(f, g)$ is of degree d. Suppose the first $m + n - d$ column vectors of $S(f, g)$ are not of full column rank, according to the following equation

$$(x^{m-1}f, x^{m-2}f, \ldots, f,\ x^{n-1}g, \ldots, g) = (x^{m+n-1}, \ldots, x^n, \ldots, 1)S,$$

there exist not all zeros $s_i, t_j \in \mathbb{R}$ such that

$$(s_{m-1}x^{m-1} + \cdots + s_0)f + (t_{n-1}x^{n-1} + \cdots + t_d x^d)g = 0. \tag{3}$$

Set $f' = f/\gcd(f, g)$, $g' = g/\gcd(f, g)$, we have

$$(s_{m-1}x^{m-1} + \cdots + s_0)f' + x^d(t_{n-1}x^{n-d-1} + \cdots + t_d)g' = 0. \tag{4}$$

Since $p \leqslant q$, then $x \nmid f'$. Furthermore $\gcd(f', g') = 1$, then from (4), we have

$$f' | (t_{n-1}x^{n-d-1} + \cdots + t_d).$$

It implies that $t_j = 0$, $j = d, \ldots, n-1$ since the degree of f' is $n - d$. Then from (3), $s_i = 0$, $i = 0, \ldots, m-1$. It is a contradiction. □

Based on this theorem, to avoid early breakdown of the generalized Schur algorithm during the implementation of the Cholesky factorization of $S^{\mathrm{T}}S$, we construct the Sylvester matrix $S(f, g)$ with $f(x)$ having smaller number of near zero tailing terms than $g(x)$. However, for some cases such as $f(x)$ and $g(x)$ both have relatively small tailing terms, it will be difficult to determine which one has smaller number of near zero tailing terms. In order to deal with this numerical difficulty, we introduce a Hankel-like-Sylvester matrix.

Remark 1. As pointed out to us by Robert M. Corless, the rank deficiency of Barnett matrix introduced in [1,2] is also equal to the degree of GCD of univariate polynomials. Furthermore, if the Barnett matrix is of rank r then the first r columns of the Barnett matrix are linearly independent [14]. Our proof of Theorem 2 is very similar to the proof in [14].

2.2. Hankel-like-Sylvester matrix

Let P_i denote an $i \times i$ anti-diagonal permutation matrix. The Hankel-like-Sylvester matrix for $f(x)$ and $g(x)$ is defined as

$$H(f, g) = S(f, g)\,\mathrm{diag}(P_m, P_n)$$

$$= \left[\begin{array}{cccccccc}
 & & \iddots & f_n & & & \iddots & g_m \\
 & & & f_{n-1} & & & & g_{m-1} \\
 & f_n & \iddots & \vdots & & g_m & \iddots & \vdots \\
f_n & f_{n-1} & & f_1 & g_m & g_{m-1} & & g_1 \\
f_{n-1} & \vdots & & f_0 & g_{m-1} & \vdots & & g_0 \\
\vdots & f_1 & \iddots & & \vdots & g_1 & \iddots & \\
f_1 & f_0 & & & g_1 & g_0 & & \\
f_0 & & & & g_0 & & &
\end{array}\right].$$

$$\qquad\underbrace{\qquad\qquad\qquad\qquad}_{m}\quad\underbrace{\qquad\qquad\qquad\qquad}_{n}$$

The Hankel-like-Sylvester matrix has several properties similar to those of the Sylvester matrix.

Lemma 2. *Supposing $H = U\Sigma V^{\mathrm{T}}$ is the SVD of $H(f, g)$, where U and V are orthogonal matrices and $\Sigma = \mathrm{diag}(\sigma_1(H), \sigma_2(H), \ldots, \sigma_{m+n}(H))$ is a diagonal matrix, then $S = U\Sigma(PV)^{\mathrm{T}}$ is the SVD of $S(f, g)$, where $P = \mathrm{diag}(P_m, P_n)$.*

Proof. It follows directly from the definition of $H(f, g)$. □

Corollary 1. *The Hankel-like-Sylvester matrix $H(f, g)$ has the same rank as the Sylvester matrix $S(f, g)$.*

Theorem 3. *$H^{\mathrm{T}}H$ is a structured matrix with displacement rank at most* 4.

Proof. We can construct matrices Z, G, J such that

$$H^{\mathrm{T}}H - ZH^{\mathrm{T}}HZ^{\mathrm{T}} = GJG^{\mathrm{T}},$$

where $Z = \mathrm{diag}(Z_m, Z_n)$, $J = \mathrm{diag}(1, 1, -1, -1)$ and $G = [\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4]$ with

$$\begin{aligned}
&\mathbf{x}_1 = H^{\mathrm{T}} H(:, 1)/\|H(:, 1)\|,\\
&\mathbf{x}_2 = H^{\mathrm{T}} H(:, m+1)/\|H(:, m+1)\|, \quad \mathbf{x}_2[1] = 0,\\
&\mathbf{x}_3 = \mathbf{x}_1 \quad \text{except that } \mathbf{x}_3[1] = 0,\\
&\mathbf{x}_4 = \mathbf{x}_2 \quad \text{except that } \mathbf{x}_4[m+1] = 0. \qquad \square
\end{aligned}$$

Theorem 4. *Let $H(f, g)$ be the Hankel-like-Sylvester matrix for two univariate polynomials $f(x)$ and $g(x)$ with degrees n and m, respectively. If H is of rank $m+n-d$, with $0 < d \leqslant \min\{m, n\}$, then the first $m+n-d$ column vectors of H are of full column rank.*

Proof. Note that $\mathrm{rank}(H) = m+n-d$ implies that $\gcd(f, g)$ is of degree d. Similar to the proof of Theorem 2, we can derive from the following equation:

$$(f, xf, \ldots, x^{m-1}f,\ g, \ldots, x^{n-d-1}g, \ldots, x^{n-1}g) = (x^{m+n-1}, \ldots, x^n, \ldots, 1)H$$

that the first $m+n-d$ columns of H are linearly independent. $\square$

Theorem 4 tells us that for a rank r Hankel-like-Sylvester matrix H, the leading $r \times r$ principal submatrix of $H^{\mathrm{T}}H$ is always positive definite. We can apply the generalized Schur algorithm to obtain Cholesky factorization of $H^{\mathrm{T}}H$ without pivoting. In practice, we construct the Hankel-like-Sylvester matrix $H(f, g)$ with f_n having larger weight, i.e., $|f_n|/\|f\|$ being large.

3. Error analysis

3.1. Backward error analysis for the fast Cholesky factorization

Here and hereafter, we assume that $T = A^{\mathrm{T}}A$, where $A = H(f, g)$ or $S(f, g)$. According to Theorem 1 or 3, we compute (G, J) as a generator pair of T with respect to a shift-block matrix Z. After applying the first r steps of the generalized Schur algorithm to (G, J, Z), we obtain a truncated Cholesky factorization $\hat{T}_r = \hat{R}_r^{\mathrm{T}} \hat{R}_r$. In this section, we derive a backward error analysis for the fast Cholesky factorization along the line of [3–5,9,10].

Let $\hat{G}_i$ be the generator matrix at stage i of the generalized Schur algorithm, $\hat{\hat{G}}_i$ be the proper form of $\hat{G}_i$, i.e., the top nonzero row of $\hat{\hat{G}}_i$ has a single nonzero entry in its first column. It is formed via a sequence of two orthogonal transformations and a hyperbolic rotation in *OD* form [4].

Assume that

$$\hat{\hat{G}}_i J \hat{\hat{G}}_i^{\mathrm{T}} - \hat{G}_i J \hat{G}_i^{\mathrm{T}} = N_i, \quad i = 1, \ldots, r.$$

Summing up the equations, we get

$$\hat{T}_r - Z\hat{T}_r Z^{\mathrm{T}} + \hat{G}_{r+1} J \hat{G}_{r+1}^{\mathrm{T}} = \sum_{i=1}^{r} N_i + T - ZTZ^{\mathrm{T}}.$$

Noting that Z is nilpotent (without loss of generality, suppose $n \geqslant m$, then $Z^n = \mathbf{0}$), we derive

$$\hat{T}_r + \sum_{k=0}^{n-1} Z^k \hat{G}_{r+1} J \hat{G}_{r+1}^{\mathrm{T}} (Z^k)^{\mathrm{T}} = T + \sum_{k=0}^{n-1} Z^k \left(\sum_{i=1}^{r} N_i \right) (Z^k)^{\mathrm{T}}.$$

Setting

$$\hat{T}_s = \sum_{k=0}^{n-1} Z^k \hat{G}_{r+1} J \hat{G}_{r+1}^{\mathrm{T}} (Z^k)^{\mathrm{T}}, \quad E = \sum_{k=0}^{n-1} Z^k \left(\sum_{i=1}^{r} N_i \right) (Z^k)^{\mathrm{T}}, \tag{5}$$

we have

$$\hat{T}_r + \hat{T}_s = T + E.$$

Since the first r rows of entries in $\hat{G}_{r+1}$ are zero, by the definition of $\hat{T}_s$, we set

$$\hat{T}_s = \begin{bmatrix} \mathbf{0}_r & \\ & \hat{S}_{r+1} \end{bmatrix}, \tag{6}$$

where $\mathbf{0}_r$ is a zero matrix of size r, $\hat{S}_{r+1}$ is the true Schur complement of $T+E$. We aim to bound $\|E - \hat{T}_s\|$.

Firstly, along the line of [5], we derive that

$$\|E\| \leqslant c_1 u(6r^3\|\hat{R}_r^{\mathrm{T}}\|^2 + (24r+4)\|T\|) + \mathrm{O}(u^2),$$

where c_1 denotes a low order polynomial in n. Since

$$\|\hat{R}_r^{\mathrm{T}}\|^2 = \|\hat{T}_r\| \leqslant \|T\| + \|E\| + \|\hat{T}_s\|,$$

we obtain

$$\|E\| \leqslant \frac{c_1 u}{1 - 6r^3 c_1 u}((6r^3 + 24r + 4)\|T\| + 6r^3\|\hat{T}_s\|) + \mathrm{O}(u^2). \tag{7}$$

Thus, the backward error is bounded as

$$\|T - \hat{T}_r\| \leqslant \|E\| + \|\hat{T}_s\| \leqslant \frac{c_1 u(6r^3 + 24r + 4)\|T\| + \|\hat{T}_s\|}{1 - 6r^3 c_1 u} + \mathrm{O}(u^2). \tag{8}$$

In order to bound $\|\hat{T}_s\|$, we write T as

$$T = \tilde{T} + \Delta T, \tag{9}$$

where $\tilde{T}$ is the nearest matrix to T with exact rank r. Specifically, we assume that

$$T = V \operatorname{diag}(\sigma_1(T), \ldots, \sigma_{n+m}(T))V^{\mathrm{T}}$$

is the SVD of T, then $\tilde{T}$ is taken as

$$\tilde{T} = V \operatorname{diag}(\sigma_1(T), \ldots, \sigma_r(T), 0, \ldots, 0)V^{\mathrm{T}},$$

thus

$$\|\Delta T\| = \sigma_{r+1}(T).$$

Denote by $F = \Delta T + E$ and partition $\tilde{T}$ and F as

$$\tilde{T} = \begin{bmatrix} \tilde{T}_{11} & \tilde{T}_{12} \\ \tilde{T}_{12}^{\mathrm{T}} & \tilde{T}_{22} \end{bmatrix}, \quad F = \begin{bmatrix} F_{11} & F_{12} \\ F_{12}^{\mathrm{T}} & F_{22} \end{bmatrix},$$

where $\tilde{T}_{11}, F_{11} \in \mathbb{R}^{r\times r}$. Since $\hat{S}_{r+1}$ is the true Schur complement of $T + E = \tilde{T} + F$, following the lead of [9], we have

Lemma 3. *Assume that $\tilde{T}_{11}$ is invertible, then*

$$\hat{S}_{r+1} = F_{22} - \tilde{W}^{\mathrm{T}}F_{12} - F_{12}^{\mathrm{T}}\tilde{W} + \tilde{W}^{\mathrm{T}}F_{11}\tilde{W} + \mathrm{O}(\|F\|^2), \tag{10}$$

where

$$\tilde{W} = \tilde{T}_{11}^{-1}\tilde{T}_{12}. \tag{11}$$

Following the arguments in [9], we have

$$\|\hat{T}_s\| = \|\hat{S}_{r+1}\| \leqslant (1+\|\tilde{W}\|)^2(\|\Delta T\| + \|E\|) + \mathrm{O}(\|F\|^2).$$

Substituting from (7), and rearranging, we find

$$\|\hat{T}_s\| \leqslant \Omega\left(\frac{c_1u(6r^3+24r+4)\|T\|}{1-6r^3c_1u} + \|\Delta T\|\right)(1+\|\tilde{W}\|)^2 + \mathrm{O}(\|F\|^2), \tag{12}$$

where

$$\Omega = \left(1 - \frac{6r^3c_1u}{1-6r^3c_1u}(1+\|\tilde{W}\|)^2\right)^{-1}. \tag{13}$$

Finally, using (12) and (8), we have

Theorem 5. *Let $T = \tilde{T} + \Delta T$ be a symmetric matrix described as* (9). *Assume that*

$$\frac{\|\Delta T_{11}\|}{\|T_{11}\|} = \theta, \tag{14}$$

and T_{11} is positive definite with

$$2\theta\kappa_2(T_{11}) < 1. \tag{15}$$

After carrying out r steps of the generalized Schur algorithm, we get

$$\|T - \hat{R}_r^{\mathrm{T}}\hat{R}_r\| \leqslant \frac{c_1u\zeta}{1-6r^3c_1u} + \frac{\Omega}{1-6r^3c_1u}\left(\frac{c_1u\zeta}{1-6r^3c_1u} + \|\Delta T\|\right)(1+\|\tilde{W}\|)^2 + \mathrm{O}(\|F\|^2), \tag{16}$$

where $\zeta = (6r^3+24r+4)\|T\|$.

Proof. The proof is similar to that of Theorem 3.1 in [9]. □

Remark 2. According to Theorem 5, the size of the backward error of the fast Cholesky factorization of T depends on the size of $\|\tilde{W}\|$. The backward error is approximately bounded as

$$\|T - \hat{R}_r^{\mathrm{T}}\hat{R}_r\| \lessapprox (1+\|\tilde{W}\|)^2\|\Delta T\|. \tag{17}$$

For Cholesky factorization with complete pivoting, a nice upper bound of $\|\tilde{W}\|$ was given in [9]. However, pivoting is prohibited in our fast algorithm since it destroys the displacement structure of involved matrices, so we only have the following bound similar to Lemma 10.12 in [10].

Theorem 6. *For $\tilde{W}$ defined by* (11), *we have*

$$\|\tilde{W}\| \leqslant \sqrt{2\|T_{11}^{-1}\|\|T_{22}\|}. \tag{18}$$

Proof. Notice that the Schur complement $\tilde{T}_{22} - \tilde{T}_{12}^{\mathrm{T}}\tilde{T}_{11}^{-1}\tilde{T}_{12}$ is equal to $\mathbf{0}$, we derive that

$$\|\tilde{W}\| = \|\tilde{T}_{11}^{-1}\tilde{T}_{12}\| = \|\tilde{T}_{11}^{-1/2}\tilde{T}_{11}^{-1/2}\tilde{T}_{12}\| \leqslant \|\tilde{T}_{11}^{-1}\|^{1/2}\|\tilde{T}_{12}^{\mathrm{T}}\tilde{T}_{11}^{-1}\tilde{T}_{12}\|^{1/2} = \sqrt{\|\tilde{T}_{11}^{-1}\|\|\tilde{T}_{22}\|}.$$

Under the assumption of Theorem 5 we have

$$\|T_{11}^{-1}\Delta T_{11}\| \leqslant \|T_{11}^{-1}\|\,\|\Delta T_{11}\| = \kappa_2(T_{11})\frac{\|\Delta T_{11}\|}{\|T_{11}\|} < 1/2,$$

so that

$$\|\tilde{T}_{11}^{-1}\| \leqslant \frac{\|T_{11}^{-1}\|}{1 - \|T_{11}^{-1}\Delta T_{11}\|} < 2\|T_{11}^{-1}\|.$$

Meantime, we note that $\tilde{T}$ and ΔT are positive semi-definite, so that $\tilde{T}_{22}$ and ΔT_{22} are all positive semi-definite. Thus from $T_{22} = \tilde{T}_{22} + \Delta T_{22}$ we get $\|\tilde{T}_{22}\| \leqslant \|T_{22}\|$. $\square$

Remark 3. This theorem provides us with an upper bound on $\|\tilde{W}\|$, which depends on $\sqrt{\kappa_2(T_{11})}$ in some sense. In fact, it is easy to construct examples where $\|\tilde{W}\|$ is large and near to the derived upper bound. However, by applying some techniques for constructing T (see Algorithm HSylSRRA), the $\|\tilde{W}\|$ appearing during the fast Cholesky factorization of T, is typically less than 1000.

3.2. A forward error bound for the estimate of some singular values

Suppose a truncated Cholesky factorization $\hat{R}_r^{\mathrm{T}}\hat{R}_r$ for $T = A^{\mathrm{T}}A$ is obtained after carrying out r steps of the generalized Schur algorithm. The singular value $\sigma_r(\hat{R}_r)$ and a null space basis of $\hat{R}_r$ can be used to approximate $\sigma_r(A)$ and $\sigma_{r+1}(A)$, respectively.

Theorem 7. *Assume the truncated error after carrying out r steps of the generalized Schur algorithm is* $\bar{\varepsilon}$, *i.e.*,

$$\|T - \hat{R}_r^{\mathrm{T}}\hat{R}_r\| \leqslant \bar{\varepsilon}, \tag{19}$$

then

$$|\sigma_i(A) - \sigma_i(\hat{R}_r)| \leqslant \bar{\varepsilon}/\sigma_r(A), \quad i = 1, \ldots, r. \tag{20}$$

Furthermore, we assume that $V_r = [\mathbf{v}_{r+1}, \ldots, \mathbf{v}_{n+m}]$ *and* $W_r = [\mathbf{w}_{r+1}, \ldots, \mathbf{w}_{n+m}]$ *are two matrices with orthonormal column vectors, where* V_r *satisfies that*

$$A^{\mathrm{T}}AV_r = V_r\Lambda_r \quad \textit{with } \Lambda_r = \mathrm{diag}(\sigma_{r+1}^2(A), \ldots, \sigma_{n+m}^2(A)), \tag{21}$$

and W_r *consists of a null space basis of* $\hat{R}_r$, *then*

$$\|\sin\Theta(V_r, W_r)\| \leqslant \sqrt{n+m-r}\,\bar{\varepsilon}/\sigma_r^2(A), \tag{22}$$

and

$$|\|AW_r\| - \sigma_{r+1}(A)| \leqslant \frac{4\|A\|^2(n+m-r)\bar{\varepsilon}^2}{\sigma_r^4(A)\sigma_{r+1}(A)}. \tag{23}$$

Proof. Applying Corollary 7.3.8 in [11] to (19) we get

$$|\sigma_i(A^{\mathrm{T}}A) - \sigma_i(\hat{R}_r^{\mathrm{T}}\hat{R}_r)| = |\sigma_i^2(A) - \sigma_i^2(\hat{R}_r)| \leqslant \bar{\varepsilon}, \quad i = 1, 2, \ldots, r.$$

Consequently, we obtain

$$|\sigma_i(A) - \sigma_i(\hat{R}_r)| \leqslant \bar{\varepsilon}/(\sigma_i(A) + \sigma_i(\hat{R}_r)) \leqslant \bar{\varepsilon}/\sigma_i(A) \leqslant \bar{\varepsilon}/\sigma_r(A), \quad i = 1, 2, \ldots, r.$$

To prove the inequality (22) we note that

$$\sigma_j(\hat{R}_r) = 0, \quad j = r+1, \ldots, n+m.$$

We define

$$\bar{\delta} = \min_{i,j} |\sigma_i^2(A) - \sigma_j^2(\hat{R}_r)|, \quad i = 1, \ldots, r, \quad j = r+1, \ldots, n+m.$$

It is obvious that

$$\bar{\delta} = \sigma_r^2(A).$$

Meantime, denote by $\bar{R}$ the residual matrix:

$$\bar{R} = A^{\mathrm{T}} A W_r - W_r D_r \quad \text{where } D_r = \mathrm{diag}(\sigma_{r+1}^2(\hat{R}_r), \ldots, \sigma_{n+m}^2(\hat{R}_r)),$$

we have

$$\|\bar{R}\| \leqslant \|A^{\mathrm{T}} A W_r - \hat{R}_r^{\mathrm{T}} \hat{R}_r W_r\| + \|\hat{R}_r^{\mathrm{T}} \hat{R}_r W_r\| \leqslant \|A^{\mathrm{T}} A - \hat{R}_r^{\mathrm{T}} \hat{R}_r\| \|W_r\| \leqslant \bar{\varepsilon}.$$

According to the eigenproblem perturbation theories [19] we get

$$\|\sin \Theta(V_r, W_r)\| \leqslant \sqrt{n+m-r}\,\|\bar{R}\|/\bar{\delta} \leqslant \sqrt{n+m-r}\ \ \bar{\varepsilon}/\sigma_r^2(A).$$

Now we check the norm $\|AW_r\|$. Based on results on eigenvalues of Rayleigh quotient matrices in [15,17], we get

$$|\|AW_r\|^2 - \sigma_{r+1}^2(A)| = |\lambda_1(W_r^{\mathrm{T}} A^{\mathrm{T}} A W_r) - \lambda_1(V_r^{\mathrm{T}} A^{\mathrm{T}} A V_r)| \leqslant 4\|A^{\mathrm{T}} A\| \|\sin \Theta(V_r, W_r)\|^2,$$

where $\lambda_1(\cdot)$ denotes the largest eigenvalue of its argument. Combining with the inequality (22), we find

$$|\|AW_r\|^2 - \sigma_{r+1}^2(A)| \leqslant 4\|A^{\mathrm{T}} A\|(n+m-r)\bar{\varepsilon}^2/\sigma_r^4(A).$$

Finally, we have

$$|\|AW_r\| - \sigma_{r+1}(A)| \leqslant \frac{4\|A\|^2(n+m-r)\bar{\varepsilon}^2}{\sigma_r^4(A)\sigma_{r+1}(A)}. \qquad \square$$

In view of Theorem 7, when $\bar{\varepsilon}$ is small and $\sigma_r(A)$ is large, $\sigma_r(\hat{R}_r)$ and $\|AW_r\|$ are nice estimates to $\sigma_r(A)$ and $\sigma_{r+1}(A)$, respectively.

4. A fast rank-revealing scheme

We implement fast Cholesky factorization of T by running the generalized Schur algorithm on the generator pair (G, J) of T, where $T = A^{\mathrm{T}} A$ and A is $S(f, g)$ or $H(f, g)$. Assume that $\hat{R}_i$ is the Cholesky factor produced from the first i steps of the generalized Schur algorithm,

$$\hat{R}_i = \begin{bmatrix} \hat{R}_{11}^{(i)} & \hat{R}_{12}^{(i)} \\ \mathbf{0} & \mathbf{0} \end{bmatrix},$$

where $\hat{R}_{11}^{(i)}$ is an $i \times i$ upper triangular matrix. When $\|T - \hat{R}_{\hat{r}}^{\mathrm{T}} \hat{R}_{\hat{r}}\|$ is small, we terminate the Cholesky factorization and declare that the numeric rank of the Sylvester matrix is $\hat{r}$. This process costs $\mathrm{O}(\hat{r}(n+m))$ flops.

According to the backward error analysis in Section 3.1, we can set a Cholesky factorization termination tolerance γ in advance. Based on (17) and the fact that, when the Sylvester matrix is of ε-rank r, $\varepsilon^2 \geqslant \sigma_{r+1}(T) = \|\Delta T\|$, we take an appropriate multiple of ε^2 as a termination tolerance γ. As stated in [9,18], at each step of the generalized Schur algorithm, the norm of the Schur complement $\|\hat{S}_{i+1}\| (i \geqslant 1)$ is approximately equal to $\|T - \hat{R}_i^{\mathrm{T}} \hat{R}_i\|$, we can set γ directly on the norms of successive Schur complements. Furthermore, $\hat{S}_{i+1}$ can be efficiently computed from $\hat{G}_{i+1}$ by using a similar formula in (5) [13,20]:

$$\begin{bmatrix} \mathbf{0}_i & \\ & \hat{S}_{i+1} \end{bmatrix} = \sum_{k=0}^{n-1} Z^k \hat{G}_{i+1} J \hat{G}_{i+1}^{\mathrm{T}} (Z^k)^{\mathrm{T}},$$

where $\mathbf{0}_i$ is a zero matrix of size i.

Algorithm HSylSRRA.

Input: $f(x)$, $g(x)$ two univariate polynomials, ε a given error tolerance for the input polynomials, γ a termination tolerance for the Cholesky process.

Output: $\hat{r}$ a candidate numerical rank of the Sylvester matrix S, $\hat{R}_{\hat{r}}$ a truncated Cholesky factor, A a Sylvester matrix or a Hankel-like-Sylvester matrix.

1. Initialization
   (a) Select a suitable order of polynomials to form S and H.
      (i) Set $\bar{f} := f, \bar{g} := g$ or $\bar{f} := g, \bar{g} := f$ to guarantee that $|lc(\bar{f})|/\|\bar{f}\| \geqslant |lc(\bar{g})|/\|\bar{g}\|$, where $\mathrm{lc}(\cdot)$ denotes the leading coefficient of a polynomial;
      (ii) Assume that $f_p \neq 0$ and $g_q \neq 0$ are the first nonzero tailing coefficients of f and g, respectively. If $p \leqslant q$ then set $\tilde{f} := f, \tilde{g} := g$ or $\tilde{f} := g, \tilde{g} := f$ to satisfy $|\tilde{f}_p|/\|\tilde{f}\| \geqslant |\tilde{g}_p|/\|\tilde{g}\|$; otherwise set $\tilde{f} := g, \tilde{g} := f, p := q$.
   (b) If $|lc(\bar{f})|/\|\bar{f}\| \geqslant |\tilde{f}_p|/\|\tilde{f}\|$ then set $A := H(\bar{f}, \bar{g})$; otherwise, set $A := S(\tilde{f}, \tilde{g})$.
   (c) Set $T := A^{\mathrm{T}}A$, construct a shift-block matrix Z and a generator pair (G, J) according to Theorem 1 or 3.

2. Cholesky rank-revealing
   (a) Perform the generalized Schur algorithm on (G, J, Z), and compute the Schur complement $\hat{S}_{i+1}, i \geqslant 1$.
   (b) If $\|\hat{S}_{\hat{r}+1}\| \leqslant \gamma$, then terminate the Schur algorithm and return $\hat{r}$, $\hat{R}_{\hat{r}}$ and A.

Remark 4. When the Sylvester matrix S is of numerical rank r, and $\gamma \geqslant \sigma_r(T)$, it is possible to find a rank $r-1$ matrix in the γ-neighborhood of T. An appropriate γ must satisfy that $\gamma < \sigma_r(T) = \sigma_r^2(S)$. Combined with the numerical results as shown in the next section, it is necessary for our algorithm that the gap $\sigma_r(S)/\sigma_{r+1}(S)$ is sufficiently large.

Remark 5. Usually it is not necessary to compute all the Schur complements. We need to compute a few Schur complements in view of preliminary rank information.

As mentioned in Section 3.2, we can obtain approximate $\hat{r}$th and $(\hat{r}+1)$th largest singular values of the Sylvester matrix by computing $\sigma_{\hat{r}}(\hat{R}_{\hat{r}})$ and $\|AW_{\hat{r}}\|$, where $W_{\hat{r}}$ is a matrix consisting of an orthonormal null space basis of $\hat{R}_{\hat{r}}$. We can calculate $\sigma_{\hat{r}}(\hat{R}_{\hat{r}})$ and an orthonormal null space basis of $\hat{R}_{\hat{r}}$ by performing the algorithm given in [16] on $\hat{R}_{\hat{r}}$. Since $\hat{R}_{\hat{r}}$is an upper triangular matrix, the total cost of the procedure is $\mathrm{O}((n+m)^2)$ flops.

5. Numerical experiments

In the following table, we show the performance of Algorithm HSylSRRA and its application for computing some singular values of Sylvester matrices. Here we use Maple 9 and set Digits = 16.

We test 600 polynomial pairs at total. The sample polynomials are generated in the following way: for each polynomial pair, the prime parts and GCD are constructed by choosing polynomials with random integer coefficients in the range $-10 \leqslant c \leqslant 10$, and then adding a perturbation. For noise we choose a relative tolerance 10^{-e}, then randomly choose a polynomial that has the same degree as the product, and coefficients in $[-10^e, 10^e]$. Finally, we scale the perturbation so that the relative error is 10^{-e}. For each polynomial pair, we take the larger norm of the two perturbations as the input error tolerance ε. We perform Algorithm HSylSRRA to compute the ε-rank of Sylvester matrix for each polynomial pair.

In the tables, (n, m, d) denotes the degrees of the input polynomials and their GCDs, respectively. For each (n, m, d) we construct three sets of polynomials: each of them comprises 50 random cases with the same relative tolerance. We report the average of input error tolerances for each set of polynomial pairs, and denote it by $\varepsilon_{\mathrm{aver}}$.

As shown in Table 1,we take γ by multiplying ε^2 a multiply: 10^4, 10^5, 10^6. $\hat{r}$ is the computed numerical rank of the Sylvester matrix by applying our algorithm; r denotes the ε-rank obtained by the SVD. We make a comparison between $\hat{r}$ and r and report the results in Table 1. For cases where $\hat{r} = r$, the Schur complements $\hat{S}_{r+1}$ are all of rather small norms. Comparing with $\hat{S}_r$ the norms of $\hat{S}_{r+1}$ drop dramatically. The average of $\|\tilde{W}\|$ is smaller than 1000. For those cases where $\hat{r} < r$, the Cholesky factorization is broken down during the first r steps. For a number of such examples, especially for the cases with $d = 3$, $\|\tilde{W}\|$ is of small size; The early breakdown of the algorithm is due to an unsuitable chosen γ (see Remark 4). However, for cases $\hat{r} > r$, the norms of $\tilde{W}$ can be of large sizes, e.g. larger than 1000.

For cases $\hat{r} = r$, we also apply Algorithm HSylSRRA to calculate the rth and $(r+1)$th largest singular values of the Sylvester matrix. Let err-$\hat{\sigma}_r$, err-$\hat{\sigma}_{r+1}$ denote the average of the relative errors for the computed rth and $(r+1)$th

Table 1
Numerical rank tests

(n, m, d)	$\gamma = 10^4\varepsilon^2$			$\gamma = 10^5\varepsilon^2$			$\gamma = 10^6\varepsilon^2$		
	$\varepsilon_{\text{aver}}$	Cases	Num.	$\varepsilon_{\text{aver}}$	Cases	Num.	$\varepsilon_{\text{aver}}$	Cases	Num.
		$\hat{r} = r$	39		$\hat{r} = r$	47		$\hat{r} = r$	48
(71, 56, 11)	0.1006	$\hat{r} < r$	2	0.0099	$\hat{r} < r$	1	0.000101	$\hat{r} < r$	1
		$\hat{r} > r$	9		$\hat{r} > r$	2		$\hat{r} > r$	1
		$\hat{r} = r$	46		$\hat{r} = r$	48		$\hat{r} = r$	47
(68, 53, 8)	0.085	$\hat{r} < r$	1	0.0090	$\hat{r} < r$	2	0.000087	$\hat{r} < r$	0
		$\hat{r} > r$	3		$\hat{r} > r$	0		$\hat{r} > r$	3
		$\hat{r} = r$	46		$\hat{r} = r$	48		$\hat{r} = r$	50
(80, 78, 3)	0.064	$\hat{r} < r$	4	0.0068	$\hat{r} < r$	2	0.000065	$\hat{r} < r$	0
		$\hat{r} > r$	0		$\hat{r} > r$	0		$\hat{r} > r$	0
		$\hat{r} = r$	46		$\hat{r} = r$	45		$\hat{r} = r$	49
(43, 38, 8)	0.066	$\hat{r} < r$	1	0.0067	$\hat{r} < r$	0	0.000067	$\hat{r} < r$	0
		$\hat{r} > r$	3		$\hat{r} > r$	5		$\hat{r} > r$	1

Table 2
Relative errors of computed singular values

(n, m, d)	$\varepsilon_{\text{aver}}$	err $-\ \hat{\sigma}_r$ err $-\ \hat{\sigma}_{r+1}$	$\varepsilon_{\text{aver}}$	err $-\ \hat{\sigma}_r$ err $-\ \hat{\sigma}_{r+1}$	$\varepsilon_{\text{aver}}$	err $-\ \hat{\sigma}_r$ err $-\ \hat{\sigma}_{r+1}$
(71, 56, 11)	0.1006	0.33e − 1 0.147e − 2	0.0099	0.195e − 1 0.34e − 4	0.000101	0.22e − 1 0.33e − 6
(68, 53, 8)	0.085	0.43e − 1 0.24e − 3	0.0090	0.23e − 1 0.20e − 4	0.000087	0.39e − 1 0.17e − 6
(80, 78, 3)	0.064	0.32e − 1 0.24e − 5	0.0068	0.32e − 1 0.14e − 6	0.000065	0.24e − 1 0.63e − 6
(43, 38, 8)	0.066	0.36e − 1 0.20e − 2	0.0067	0.30e − 1 0.28e − 5	0.000067	0.25e − 1 0.91e − 6

largest singular values with respect to the singular values computed by SVD. The small relative errors in Table 2 show that the computed singular values are of satisfactory accuracy.

6. Conclusion

In this paper, a fast rank-revealing algorithm for the Sylvester matrix is derived. It costs $O(r(n + m))$ flops, where $n + m$ and r are the size and the numerical rank of the Sylvester matrix, respectively. This algorithm is based on the fast Cholesky factorization of $S^T S$ or $H^T H$, where S is the Sylvester matrix and H is the Hankel variation of S. Numerical tests show that the fast rank-revealing algorithm is valid for Sylvester matrices with low rank deficiency.

Acknowledgement

We would like to thank Robert M. Corless for his valuable suggestions. We also thank anonymous referees for useful comments.

References

[1] S. Barnett, Greatest common divisor of several polynomials, Proc. Cambridge Philos. Soc. 70 (1971) 263–268.
[2] S. Barnett, Polynomials and Linear Control Systems, Marcel Dekker, New York, 1983.

[3] A.W. Bojanczyk, R.P. Brent, F.R. de Hoog, D.R. Sweet, On the stability of the Bareiss and related Toeplitz factorization algorithm, SIAM J. Matrix Anal. Appl. 16 (1995) 40–57.
[4] S. Chandrasekaran, A.H. Sayed, Stabilizing the generalized Schur algorithm, SIAM J. Matrix Anal. Appl. 17 (1996) 950–983.
[5] S. Chandrasekaran, A.H. Sayed, A fast stable solver for nonsymmetric Toeplitz and quasi-Toeplitz systems of linear equations, SIAM J. Matrix Anal. Appl. 19 (1998) 107–139.
[6] R.M. Corless, P.M. Gianni, B.M. Trager, S.M. Watt, The singular value decomposition for polynomial systems, in: A.H.M. Levelt (Ed.), ISSAC '95 Proceedings of 1995 International Symposium on Symbolic Algebraic Computation, ACM Press, New York, NY, 1995, pp. 96–103.
[7] R.M. Corless, S.M. Watt, L. Zhi, QR factoring to compute the GCD of univariate approximate polynomials, IEEE Trans. Signal Process. 52 (2004) 3394–3402.
[8] I.Z. Emiris, A. Galligo, H. Lombardi, Certified approximate univariate GCDs, J. Pure Appl. Algebra 117/118 (1997) 229–251 (Special issue on Algorithms for Algebra).
[9] N.J. Higham, Analysis of the Cholesky decomposition of a semi-definite matrix, in: M.G. Cox, S.J. Hammarling (Eds.), Reliable Numerical Computation, Oxford University Press, Oxford, 1990, pp. 161–185.
[10] N.J. Higham, Accuracy and Stability of Numerical Algorithms, Society for Industrial and Applied Mathematics, Philadelphia, 1996.
[11] R.A. Horn, C.R. Johnson, Matrix Analysis, Cambridge University Press, Cambridge, 1985.
[12] C.P. Jeannerod, L. Zhi, Computing low rank generators of Sylvester matrix embeddings, Manuscript, 2002.
[13] T. Kailath, A.H. Sayed, Displacement structure: theory and applications, SIAM Rev. 37 (1995) 297–386.
[14] G.V. Laureano, An elementary proof of Barnett's theorem about the greatest common divisor of several univariate polynomials, Linear Algebra Appl. 247 (1996) 185–202.
[15] R. Li, On eigenvalues of a Rayleigh quotient matrix, Linear Algebra Appl. 169 (1992) 249–255.
[16] T.Y. Li, Z. Zeng, A rank-revealing method and its application, Manuscript available at ⟨http://www.neiu.edu/∼zzeng/papers.htm⟩, 2003.
[17] X. Liu, Y. Xu, On the Rayleigh quotient matrix, Math. Numer. Sinica 12 (1990) 208–213 (in Chinese).
[18] M. Stewart, Cholesky factorization of semi-definite Toeplitz matrices, Linear Algebra Appl. 254 (1997) 497–526.
[19] J. Sun, Matrix Perturbation Analysis, Science Press, 1987 (in Chinese).
[20] Z. Xu, K. Zhang, Q. Lu, The Fast Algorithm for Solving Toeplitz Matrices, Xibei Industry University Press, 1999 (in Chinese).
[21] Z. Zeng, The approximate GCD of inexact polynomials. part I: a univariate algorithm, Manuscript, 2004.
[22] L. Zhi, Displacement structure in computing approximate GCD of univariate polynomials, in: Proceedings of the Sixth Asian Symposium on Computer Mathematics (ASCM 2003), 2003, pp. 288–298.