Journal of Computational and Applied Mathematics 234 (2010) 3226–3242

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

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

Joab R. Winkler *, Madina Hasan

Department of Computer Science, The University of Sheffield, Regent Court, 211 Portobello Street, Sheffield S1 4DP, United Kingdom

ARTICLE INFO

Article history:
Received 10 August 2009
Received in revised form 31 March 2010

Keywords:
Sylvester matrix
Structured low rank approximation

ABSTRACT

A non-linear structure preserving matrix method for the computation of a structured low rank approximation $S(\tilde{f},\tilde{g})$ of the Sylvester resultant matrix $S(f,g)$ of two inexact polynomials $f = f(y)$ and $g = g(y)$ is considered in this paper. It is shown that considerably improved results are obtained when $f(y)$ and $g(y)$ are processed prior to the computation of $S(\tilde{f},\tilde{g})$, and that these preprocessing operations introduce two parameters. These parameters can either be held constant during the computation of $S(\tilde{f},\tilde{g})$, which leads to a linear structure preserving matrix method, or they can be incremented during the computation of $S(\tilde{f},\tilde{g})$, which leads to a non-linear structure preserving matrix method. It is shown that the non-linear method yields a better structured low rank approximation of $S(f,g)$ and that the assignment of $f(y)$ and $g(y)$ is important because $S(\tilde{f},\tilde{g})$ may be a good structured low rank approximation of $S(f,g)$, but $S(\tilde{g},\tilde{f})$ may be a poor structured low rank approximation of $S(g,f)$ because its numerical rank is not defined. Examples that illustrate the differences between the linear and non-linear structure preserving matrix methods, and the importance of the assignment of $f(y)$ and $g(y)$, are shown.

1. Introduction

Resultant matrices arise in several disciplines that require the processing of curves and surfaces, including computer graphics [1], computer vision [2] and computer aided geometric design. They are frequently used in geometric problems because they can be used to determine if two polynomial curves intersect, and thus the points of intersection are calculated only if the curves intersect. In particular, a resultant matrix, the entries of which are functions of the coefficients of the polynomials, is singular if and only if the curves intersect. Although design intent may require that the curves intersect, inexact data may imply they do not intersect, in which case the design intent is realised by perturbing the coefficients of the polynomials slightly such that their resultant matrix becomes singular, that is, a structured low rank approximation of the given resultant matrix is required. This paper compares the methods of structured total least norm (STLN) [3] and structured non-linear total least norm (SNTLN) [4] for the calculation of a structured low rank approximation of the Sylvester resultant matrix, which is one type of resultant matrix.

The Sylvester resultant matrix $S(f,g) \in \mathbb{R}^{(m+n)\times(m+n)}$ of the polynomials $f = f(y)$ and $g = g(y)$,

$$f(y) = \sum_{i=0}^{m} a_i y^{m-i} \quad \text{and} \quad g(y) = \sum_{i=0}^{n} b_i y^{n-i}, \quad a_0, b_0 \neq 0, \tag{1}$$

* Corresponding author. Tel.: +44 114 222 1834; fax: +44 114 222 1810.
E-mail addresses: j.winkler@dcs.shef.ac.uk (J.R. Winkler), acp07mah@sheffield.ac.uk (M. Hasan).

0377-0427/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.cam.2010.04.013

is

$$S(f,g)=\begin{bmatrix} a_0 & & & b_0 & & \\ a_1 & \ddots & & b_1 & \ddots & \\ \vdots & \ddots & a_0 & \vdots & \ddots & b_0 \\ a_{m-1} & \ddots & a_1 & b_{n-1} & \ddots & b_1 \\ a_m & \ddots & \vdots & b_n & \ddots & \vdots \\ & \ddots & a_{m-1} & & \ddots & b_{n-1} \\ & & a_m & & & b_n \end{bmatrix}, \tag{2}$$

where the coefficients a_i of $f(y)$ occupy the first n columns and the coefficients b_i of $g(y)$ occupy the last m columns.

The calculation of a structured low rank approximation of $S(f,g)$ is closely related to the calculation of an approximate greatest common divisor (AGCD) of $f(y)$ and $g(y)$. For example, Bini and Boito [5] discuss three methods, based on the structure of the Sylvester $S(f,g)$ and Bézout $B(f,g)$ resultant matrices, for AGCD computations. The QR decomposition of $S(f,g)$ is used by Corless et al. [6], and Zarowski et al. [7], and the singular value decomposition of $S(f,g)$ is used in [8]. The QR and singular value decompositions do not retain the structure of $S(f,g)$, and they must therefore be compared with methods that preserve the structure of $S(f,g)$, which are discussed in [9–12]. Other methods have also been used to calculate an AGCD of two polynomials. For example, optimisation techniques are used by Karmarkar and Lakshman [13], and Padé approximations are used by Pan [14].

Many methods for the calculation of an AGCD of two inexact polynomials involve two stages. In particular, the degree of an AGCD of the polynomials is determined initially, after which the coefficients of the AGCD are calculated. The computation of the degree of an AGCD of $f(y)$ and $g(y)$ is equivalent to the determination of the rank loss of a resultant matrix, and methods for this computation are considered in [15]. It is assumed in this paper, however, that the degree of an AGCD is known. This assumption is also made in [9–12], and a linear structure preserving method is used in these references to compute a structured low rank approximation of $S(f,g)$.

If the ratio of the maximum coefficient (in magnitude) to the minimum coefficient (in magnitude) of $\{f(y), g(y)\}$ is large, the polynomials must be processed before a structured low rank approximation $S(\tilde{f},\tilde{g})$ of $S(f,g)$ is computed. These preprocessing operations introduce two parameters, which can either be held constant, or incremented, during the computation of $S(\tilde{f},\tilde{g})$. A linear structure preserving matrix method is used if they are held constant, but a non-linear structure preserving matrix method is required if they are incremented. Considerably improved results are obtained when the preprocessing operations are included in the computation of $S(\tilde{f},\tilde{g})$, and the non-linear method yields better results than the linear method because the numerical rank of $S(\tilde{f},\tilde{g})$ is, in general, more clearly defined. Furthermore, it is shown that the assignment of the polynomials to $f(y)$ and $g(y)$ is important because the numerical rank of a structured low rank approximation of $S(f,g)$ may be defined, but the numerical rank of a structured low rank approximation of $S(g,f)$ may not be defined.

Subresultant matrices, which are derived from $S(f,g)$ and are important for the calculation of $S(\tilde{f},\tilde{g})$, are discussed in Section 2, and the preprocessing operations on $f(y)$ and $g(y)$ are considered in Section 3. Section 4 contains a brief comparison of STLN and SNTLN, and the application of SNTLN to the computation of $S(\tilde{f},\tilde{g})$ is discussed in Section 5. Section 6 contains examples that show the differences in the results using STLN and SNTLN, and the importance of the polynomial order, (f,g) or (g,f), for the computation of a structured low rank approximation of the Sylvester matrix of $f(y)$ and $g(y)$. A summary of the paper is contained in Section 7.

2. Subresultant matrices

This section discusses subresultant matrices, which are derived from $S(f,g)$ by deleting some of its rows and columns. These matrices are required for the calculation of $S(\tilde{f},\tilde{g})$, and they are most easily introduced by expressing the product of two polynomials as a matrix-vector product.

If $\hat{f}(y)$ and $\hat{g}(y)$ are the theoretically exact forms of $f(y)$ and $g(y)$ respectively, and the degree of their greatest common divisor (GCD) is $\hat{d}$, then there exist quotient polynomials $u_k(y)$ and $v_k(y)$, and a common divisor polynomial $d_k(y)$, such that for $k=1,\dots,\hat{d}$,

$$d_k(y)=\frac{\hat{f}(y)}{u_k(y)}=\frac{\hat{g}(y)}{v_k(y)}, \qquad \deg v_k < \deg \hat{g}=n, \qquad \deg u_k < \deg \hat{f}=m, \tag{3}$$

where

$$u_k(y)=\sum_{i=0}^{m-k} u_{k,i}y^{m-k-i} \quad \text{and} \quad v_k(y)=\sum_{i=0}^{n-k} v_{k,i}y^{n-k-i}.$$

It follows from (3) that there exists a non-zero polynomial $t_k(y)$ such that

$$t_k(y)=v_k(y)\hat{f}(y)=u_k(y)\hat{g}(y), \quad k=1,\dots,\hat{d},$$