Available online at www.sciencedirect.com
SCIENCE DIRECT®
Applied Mathematics and Computation 171 (2005) 596–603
APPLIED MATHEMATICS AND COMPUTATION
www.elsevier.com/locate/amc

An index reduction method for linear Hessenberg systems

M.M. Hosseini

Department of Mathematics, Yazd University, P.O. Box 89195-741, Yazd, Iran

Abstract

In [E. Babolian, M.M. Hosseini, Reducing index, and pseudospectral methods for differential–algebraic equations, Appl. Math. Comput. 140 (2003) 77–90] a reducing index method has been proposed for some cases of semi-explicit DAEs (differential algebraic equations). In this paper, this method is generalized to more cases. Also, it is focused on Hessenberg index 2 systems and proposed reduction index method will be illustrated for this problem. The Hessenberg system and its obtained reduced index system are numerically solved through pseudospectral method. In addition, aforementioned methods will be considered by one example.

Keywords: Differential–algebraic equations; Hessenberg systems; Index reduction techniques; Pseudospectral method

1. Introduction

Consider a linear (or linearized) semi-explicit DAEs

$$X^{(m)} = \sum_{j=1}^{m} A_j X^{(j-1)} + By + q, \tag{1a}$$

E-mail address: hosse_m@yazduni.ac.ir

0096-3003/$ - see front matter
doi:10.1016/j.amc.2005.01.074

$$0 = CX + r, \tag{1b}$$

where A_j and C are smooth functions t, $t_0 \leqslant t \leqslant t_f$, $A_j(t) \in R^{n\times n}$, $j = 1,\ldots,m$, $B(t) \in R^{n\times k}$, $C(t) \in R^{k\times n}, n > k$, and CB is nonsingular (the DAEs has index $m+1$). The homogeneities are $q(t) \in R^n$ and $r(t) \in R$. It is well-known that the DAEs (1) can be difficult to solve when it has a higher index (index greater than one, [1]). In this case an alternative treatment is the use of index reduction methods (see, e.g., [4,5,7,9]), until a well-posed problem (index-1 DAEs or ordinary differential equations) is obtained. Early in [3] and [8], for $k = 1$, the index of problem (1) has been reduced by introducing a simple formulation. In this paper, we will reduce the index of (1) when $k > 1$. For this reason, we put

$$y = (CB)^{-1}C\left[X^{(m)} - \sum_{j=1}^{m} A_j X^{(j-1)} - q\right], \tag{2}$$

and by substituting (2) in (1a), we obtain an implicit DAE which has index m, as follows:

$$\sum_{j=0}^{m} E_j X^{(j)} = \hat{q}, \tag{3}$$

where $E_j(t) \in R^{n\times n}$, $j = 0,1,\ldots,m$, and except $E_0(t)$, others are singular matrices. Note that system (3) has k equations less than system (1).

It is known that the eigenfunctions of certain singular Sturm–Liouville problems allow the approximation of functions $C^\infty[a,b]$ where truncation error approaches zero faster than any negative power of the number of basic functions used in the approximation, as that number (order of truncation N) tends to infinity [6]. This phenomenon is usually referred to as "spectral accuracy" [6]. The accuracy of derivatives obtained by direct, term-by-term differentiation of such truncated expansion naturally deteriorates [2], but for low-order derivatives and sufficiently high order truncations this deterioration is negligible, compared to the restrictions in accuracy introduced by typical difference approximations (for more details, refer to [3,8]). Throughout, we are using first kind orthogonal Chebyshev polynomials $\{T_k\}_{k=0}^{+\infty}$ which are eigenfunctions of singular Sturm–Liouville problem

$$\left(\sqrt{1-x^2}\,T'(x)\right)' + \frac{k^2}{\sqrt{1-x^2}}T_k(x) = 0.$$

2. A simple formulation for index reduction

In this section, DAEs (1) is considered when $m = 1$ and $k = 2$. To extend it to general case (1) is easy. Now consider the Hessenberg index-2 system,

$$X' = AX + By + q, \tag{4a}$$

$$0 = CX + r, \tag{4b}$$

where $A = (a_{ij})_{n\times n}$, $B = (b_{ij})_{n\times 2}$, $C = (c_{ij})_{2\times n}$, $n \geqslant 3$ and

$$\det(CB(t)) \neq 0, \quad t \in [t_0, t_f]. \tag{5}$$

From (4a) and (5), we can write

$$y = (CB)^{-1}C[X' - AX - q], \quad t \in [t_0, t_f], \tag{6}$$

and substituting (6) into (4a), implies,

$$X' = AX + B(CB)^{-1}C[X' - AX - q] + q.$$

So, problem (4) transforms to the system

$$\det(CB(t))[I - B(CB)^{-1}C][X' - AX - q] = 0, \tag{7a}$$

$$CX + r = 0. \tag{7b}$$

Here, the overdetermined system (7) will be transformed to a full rank DAE system with n equation and n unknowns which has index one.

Theorem 1. *The index*-2 *DAE system* (4), *with* $n = 3$, *is equivalent to index*-1 *DAE system* (8),

$$\begin{bmatrix} \overline{M} \\ 0 \end{bmatrix} X' + \begin{bmatrix} -\overline{M}A \\ C \end{bmatrix} X = \begin{bmatrix} \overline{M}q \\ -r \end{bmatrix}, \tag{8}$$

such that,

$$\overline{M} = [\, b_{21}b_{32} - b_{22}b_{31} \quad b_{12}b_{31} - b_{11}b_{32} \quad b_{11}b_{22} - b_{12}b_{21} \,]. \tag{9}$$

Proof. As it is seen, the DAE system (4) is transformed to overdetermined system (7) by using (6). Since $n = 3$, we have,

$$\begin{aligned}\det(CB(t)) = {} & (c_{11}c_{22} - c_{12}c_{21})(b_{11}b_{22} - b_{12}b_{21}) \\ & + (c_{11}c_{23} - c_{13}c_{21})(b_{11}b_{32} - b_{12}b_{31}) \\ & + (c_{12}c_{23} - c_{13}c_{22})(b_{21}b_{32} - b_{22}b_{31}).\end{aligned}$$

Now, if we define

$$\begin{aligned}&c_1 = c_{11}c_{22} - c_{12}c_{21}, \quad c_2 = c_{11}c_{23} - c_{13}c_{21}, \quad c_3 = c_{12}c_{23} - c_{13}c_{22}, \\ &b_1 = b_{11}b_{22} - b_{12}b_{21}, \quad b_2 = b_{11}b_{32} - b_{12}b_{31}, \quad b_3 = b_{21}b_{32} - b_{22}b_{31},\end{aligned}$$

then we can rewrite $\det(CB(t))$ as below,

$$\det(CB(t)) = (c_1b_1 + c_2b_2 + c_3b_3) \neq 0. \quad t \in [t_0, t_f]. \tag{10}$$

Also, we have

$$M_{n\times n} = \det(CB(t))\left[I - B(CB)^{-1}C\right] = \begin{bmatrix} c_3b_3 & -c_3b_2 & c_3b_1 \\ -c_2b_3 & c_2b_2 & -c_2b_1 \\ c_1b_3 & -c_1b_2 & c_1b_1 \end{bmatrix}, \tag{11}$$

(10) and (11) imply that $\operatorname{rank}(M) = 1$. In addition, if we define $\overline{M} = [\,b_3 \quad -b_2 \quad b_1\,]$ by considering (7), we have

$$\begin{aligned} \overline{M}[X' - AX - q] &= 0, \\ CX + r &= 0 \end{aligned}$$

and it implies that

$$\begin{bmatrix} \overline{M} \\ 0 \end{bmatrix} X' + \begin{bmatrix} -\overline{M}A \\ C \end{bmatrix} X = \begin{bmatrix} \overline{M}q \\ -r \end{bmatrix}.$$

So, the overdetermined system (7) is transformed to system (8), with 3 equations and unknowns. In continuation, we must show that the system (8) is full rank and has index 1. For this reason, it is sufficient to show that $\begin{bmatrix} \overline{M} \\ C \end{bmatrix}_{n\times n}$ is nonsingular (according to Algorithm (4.1) mentioned in [7]). But by computing the determinant of $\begin{bmatrix} \overline{M} \\ C \end{bmatrix}_{n\times n}$, we have

$$\det\left(\begin{bmatrix} \overline{M} \\ C \end{bmatrix}\right) = \det(CB(t)) \neq 0, \quad t \in [t_0, t_f].$$

Hence, by Theorem 1 a simple formulation is presented to reduce the index of DAE system (4) when $n = 3$. For $n > 3$, to present a simple formulation, as well as $n = 3$, is not possible. But this reduction index can be done by using a simple Maple program. For example, when $n = 4$, through using a Maple program, we conclude that

$$\det(CB(t)) = \sum_{i=1}^{4-1} \sum_{j=i+1}^{4} \left(c_{1i}c_{2j} - c_{1j}c_{2i}\right)\left(b_{i1}b_{j2} - b_{j1}b_{i2}\right). \tag{12}$$

Also, by introducing $M = \det(CB(t))\lfloor I - B(CB)^{-1}C\rfloor$, we have $\operatorname{rank}(M) = 2$ and if $\overline{M}$ is obtained by eliminating the first and second rows of M, then

$$\det\left(\begin{bmatrix} \overline{M} \\ C \end{bmatrix}\right) = (c_{11}c_{22} - c_{12}c_{21}) \times \det{}^2(CB(t)).$$

$$\text{Hence rank}\left(\begin{bmatrix}\overline{M}\\ C\end{bmatrix}_{4\times 4}\right) = 4 \text{ when } (c_{11}c_{22} - c_{12}c_{21}) \neq 0, \quad t \in [t_0, t_f].$$

In general case if $\overline{M}$ is obtained by eliminating the ith and jth rows of M, then

$$\text{Rank}\left(\begin{bmatrix}\overline{M}\\ C\end{bmatrix}_{4\times 4}\right) = 4 \text{ when } \left(c_{1i}c_{2j} - c_{1j}c_{2i}\right)(t) \neq 0, \quad t \in [t_0, t_f]. \tag{13}$$

Since $\det(CB(t)) \neq 0$, hence according to (12), there exist i and j such that condition (13) is hold.

3. Implementation of numerical method

Here, the implementation of pseudospectral method is presented for DAEs systems (4) and (8). This discussion can simply be extended to general forms and (3). Now consider the DAEs systems,

$$\sum_{j=1}^{3} f_j(t)x'_j + \sum_{j=4}^{6} f_j(t)x_{j-3} = \hat{q}(t), \tag{14a}$$

$$\sum_{j=1}^{3} c_{ij}(t)x_j = -r_i(t), \quad i = 1, 2, \tag{14b}$$

with initial condition,

$$x_1(t_0) = \alpha. \tag{15}$$

For an arbitrary natural number m, we suppose that the approximate solution of DAEs systems (14) is as below,

$$x_j(t) = \sum_{i=0}^{v} a_{i+(j-1)\times(v+1)}T_i(s), \quad j = 1, 2, 3, \quad s \in [-1, 1], \tag{16}$$

where

$$t = h(s) = \frac{t_f - t_0}{2}s + \frac{t_f + t_0}{2}, \tag{17}$$

where $a = (a_0, a_1, \ldots, a_{3m+2})^t \in R^{3m+3}$ and $\{T_k\}_{k=0}^{+\infty}$ is sequence of Chebyshev polynomials of the first kind. Here, the main purpose is to find vector a. Now, by using (17), we rewrite system (14) and (15), as below,

$$\sum_{j=1}^{3} \frac{\mathrm{d}s}{\mathrm{d}t} f_j(h(s)) x_j' + \sum_{j=4}^{6} f_j(h(s)) x_{j-3} = \hat{q}(h(s)), \quad s \in [-1, 1], \tag{18a}$$

$$\sum_{j=1}^{3} c_{ij}(h(s)) x_j = -r_i(h(s)), \quad i = 1, 2, \tag{18b}$$

$$x_1(-1) = \alpha. \tag{18c}$$

By substituting (16) into (18) we have (for more details refer to [2,3])

$$\sum_{i=0}^{3v+2} a_i \Phi_i(s) \approx \hat{q}(s), \tag{19a}$$

$$\sum_{i=0}^{3v+2} a_i \Psi_{ij}(s) \approx -r_j(s), \quad j = 1, 2, \tag{19b}$$

$$\sum_{i=0}^{v} a_i T_i(-1) = \alpha \Rightarrow \sum_{i=0}^{v} a_i (-1)^i = \alpha. \tag{19c}$$

Relation (19c) forms a system with one equation and $3m + 3$ unknowns, to construct the remaining $3m + 2$ equations we substitute Chebyshev–Gauss points,

$$s_j = \cos\left(\frac{2\pi j}{2v}\right), \quad j = 0, 1, \ldots, v - 1,$$

in (19a) and

$$s_j = \cos\left(\frac{2\pi j}{2v}\right), \quad j = 0, 1, \ldots, v,$$

in (19b) to obtain $3m + 2$ equations.

4. Numerical example

Here, we use "e_x" and "e_y" to denote the maximum absolute error in vectors $X = (x_1, x_2, x_3)$ and $y = (y_1, y_2)$. These values are approximately obtained through their graphs. Results show the advantages of techniques, mentioned in Sections 2 and 3. Also, the presented algorithm in Section 3, is performed by using Maple 8 with 25 digits precision.

Example. Consider for $-1 \leqslant t \leqslant 1$,

$$X' = AX + By + q, \tag{20a}$$

Table 1
Maximum norm error for problem (20)

m	Without index reduction		With index reduction	
	e_x	e_y	e_x	e_y
6	2.7 (−3)	4.2 (−3)	4.5 (−4)	2.3 (−4)
10	1.6 (−4)	2.5 (−4)	1.4 (−6)	1.0 (−6)
12	2.7 (−6)	4.4 (−6)	8.0 (−12)	3.0 (−12)
16	1.4 (−7)	2.3 (−7)	5.0 (−17)	3.4 (−17)

$$0 = CX + r, \tag{20b}$$

where

$$A = \begin{bmatrix} \frac{3-2t}{2-t} & 0 & 0 \\ \frac{1}{2-t} & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}, \quad B = \begin{bmatrix} 4-2t & 0 \\ 0 & 1 \\ \sin(2t) & \cos(2t) \end{bmatrix}, \quad C = B^T,$$

with initial condition, $x_1(0) = 1$, and exact solutions, $x_1(t) = x_2(t) = x_3(t) = e^t$ and $y_1(t) = y_2(t) = \frac{e^t}{t-2}$. $q(t)$ and $r(t)$ are compatible with above exact solutions. This problem has index 2.

$$\det(CB) = 1 + \sin^2(2t) + 4(2-t)^2 + 4(t-2)^2\cos^2(2t) \neq 0, \quad t \in [0, 1].$$

Also according to (9),

$$\overline{M} = [-\sin(2t) \quad (2t-4)\cos(2t) \quad 4-2t].$$

Hence, by Theorem 1 the index-2 DAE (20) converts to index-1 DAE (21) as below,

$$\begin{bmatrix} \overline{M} \\ 0 \end{bmatrix} X' + \begin{bmatrix} -\overline{M}A \\ C \end{bmatrix} X = \begin{bmatrix} \overline{M}q \\ -r \end{bmatrix}. \tag{21}$$

In Table 1, we record the results of running pseudospectral method for this example with and without index reduction.

The advantage of using index reduction method (proposed in Section 2) is clearly demonstrated for above example.

References

[1] U.M. Ascher, P. Lin, Sequential regularization methods for higher index differential–algebraic equations with constraint singularities: the linear index-2 case, SIAM J. Anal. 33 (1996) 1921–1940.

[2] E. Babolian, M.M. Hosseini, A modified spectral method for numerical solution of ordinary differential equations with non-analytic solution, Appl. Math. Comput. 132 (2002) 341–351.

[3] E. Babolian, M.M. Hosseini, Reducing index, and pseudospectral methods for differential–algebraic equations, Appl. Math. Comput. 140 (2003) 77–90.
[4] K.E. Brenan, S.L. Campbell, L.R. Petzold, Numerical Solution of Initial-value Problems in Differential–algebraic Equations, Elsevier, New York, 1989.
[5] S.L. Campbell, A computational method for general higher index singular systems of differential equations, IMACS, Trans. Sci. Comput. 89 (1989) 555–560.
[6] C. Canuto, M.Y. Hussaini, A. Quarteroni, A. Zang, Spectral Methods in Fluid Dynamics, Springer-Verlag, 1998.
[7] C.W. Gear, L.R. Petzold, ODE methods for the solution of differential/algebraic systems, SIAM J. Numer. Anal. 21 (1984) 716–728.
[8] M.M. Hosseini, Reducing index method for differential–algebraic equations with constraint singularities, Appl. Math. Comput. 153 (2004) 205–214.
[9] H. Wang, Y. Song, Regularization methods for solving differential–algebraic equations, Appl. Math. Comput. 119 (2001) 283–296.