MATHEMATICAL ENGINEERING TECHNICAL REPORTS

Index Reduction for Differential-Algebraic Equations by Substitution Method

Mizuyo TAKAMATSU and Satoru IWATA

METR 2007–02

January 2007

DEPARTMENT OF MATHEMATICAL INFORMATICS GRADUATE SCHOOL OF INFORMATION SCIENCE AND TECHNOLOGY THE UNIVERSITY OF TOKYO BUNKYO-KU, TOKYO 113-8656, JAPAN

WWW page: http://www.i.u-tokyo.ac.jp/edu/course/mi/index_e.shtml

The METR technical reports are published as a means to ensure timely dissemination of scholarly and technical work on a non-commercial basis. Copyright and all rights therein are maintained by the authors or by other copyright holders, notwithstanding that they have offered their works here electronically. It is understood that all persons copying this information will adhere to the terms and constraints invoked by each author's copyright. These works may not be reposted without the explicit permission of the copyright holder.

Index Reduction for Differential-Algebraic Equations by Substitution Method

Mizuyo Takamatsu^{*} Satoru Iwata[†]

January, 2007

Abstract

Differential-algebraic equations (DAEs) naturally arise in many applications, but present numerical and analytical difficulties. The index of a DAE is a measure of the degree of numerical difficulty. In general, the higher the index is, the more difficult it is to solve the DAE. Therefore, it is desirable to transform the original DAE into an equivalent DAE with lower index.

In this paper, we propose an index reduction method for linear DAEs with constant coefficients. This is applicable to all DAEs having at most one derivative per equality, and always reduces the index exactly by one. Our method is different from other existing methods in that it does not introduce any additional variables.

1 Introduction

Dynamical systems such as electric circuits, mechanical systems, and chemical plants are often described by differential-algebraic equations (DAEs), which consist of algebraic equations and differential operations. DAEs present numerical and analytical difficulties which do not occur with ordinary differential equations (ODEs).

Several numerical methods have been developed for solving DAEs. For example, Gear [6] proposed the backward difference formulae (BDF), which were implemented in the DASSL code by Petzold (cf. [1]). Hairer and Wanner [9] implemented an implicit Runge-Kutta method in their RADAU5 code.

The *index* concept plays an important role in the analysis of DAEs. The index is a measure of the degree of difficulty in the numerical solution. In general, the higher the index is, the more difficult it is to solve the DAE. While many different concepts exist to assign an index to a DAE such as the *differentiation index* [1, 3, 9], the *perturbation index* [2], and the *tractability index* [14], we focus on the *nilpotency index* in this paper. In the case of linear DAEs with constant coefficients, all these indices are equal [2, 13].

^{*}Department of Mathematical Informatics, Graduate School of Information Science and Technology, University of Tokyo, Tokyo 113-8656, Japan (mizuyo_takamatsu@mist.i.u-tokyo.ac.jp).

[†]Research Institute for Mathematical Sciences, Kyoto University, Kyoto 606-8502, Japan.

In order to transform a DAE into an alternative form easier to solve, some index reduction methods have been developed [7, 10, 11]. These methods introduce additional variables, which leads to a drawback that the resulting DAE is a larger system than the original one.

This paper presents a new index reduction method, called the *substitution method*, for linear DAEs with constant coefficients

$$P\frac{\mathrm{d}\boldsymbol{x}(t)}{\mathrm{d}t} + Q\boldsymbol{x}(t) = \boldsymbol{f}(t), \qquad (1)$$

where P and Q are constant matrices. The method is shown to reduce the index of DAEs by one without introducing any additional variables, provided that P has at most one nonzero entry in each row. This class of DAEs includes the semi-explicit form and circuit equations of most linear time-invariant circuits free from mutual inductances.

The organization of this paper is as follows. In Section 2, we explain a matrix pencil and the definition of the nilpotency index. Section 3 introduces the substitution method. We describe the proposed method for index reduction in Section 4. Numerical examples are given in Section 5. Section 6 concludes this paper.

2 DAEs and Matrix Pencils

For a polynomial a(s), we denote the degree of a(s) by deg a, where deg $0 = -\infty$ by convention. A polynomial matrix $A(s) = (a_{ij}(s))$ with deg $a_{ij} \leq 1$ for all (i, j) is called a *matrix pencil*. Obviously, a matrix pencil A(s) can be represented as A(s) = sP + Q in terms of a pair of constant matrices P and Q. A matrix pencil A(s) is said to be *regular* if A(s) is square and det A(s) is a nonvanishing polynomial.

With the use of the Laplace transformation, the DAE in the form of (1) is expressed by the matrix pencil A(s) = sP + Q as $A(s)\tilde{\boldsymbol{x}}(s) = \tilde{\boldsymbol{f}}(s)$, where s is the variable for the Laplace transform that corresponds to d/dt, the differentiation with respect to time.

Theorem 2.1 ([1, Theorem 2.3.1]). The linear DAE with constant coefficients (1) is solvable if and only if A(s) is a regular matrix pencil.

The reader is referred to [1, Definition 2.2.1] for the precise definition of solvability. By Theorem 2.1, we assume that A(s) is a regular matrix pencil throughout this paper. It is known that a regular matrix pencil can be brought into the *Kronecker canonical form*, which determines the nilpotency index. Let N_{μ} denote a $\mu \times \mu$ matrix pencil defined by

$$N_{\mu} = \begin{pmatrix} 1 & s & 0 & \cdots & 0 \\ 0 & 1 & s & \ddots & \vdots \\ 0 & 0 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 1 & s \\ 0 & \cdots & 0 & 0 & 1 \end{pmatrix}$$

Theorem 2.2 ([5, Chapter XII, Theorem 3]). For an $n \times n$ regular matrix pencil A(s), there exist nonsingular constant matrices U and V which transform A(s) into the Kronecker canonical form:

$$UA(s)V = \begin{pmatrix} sI_{\mu_0} + H & O & O & \cdots & O \\ O & N_{\mu_1} & O & \cdots & O \\ O & O & N_{\mu_2} & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & O \\ O & O & \cdots & O & N_{\mu_b} \end{pmatrix},$$

where

 $\mu_1 \ge \mu_2 \ge \dots \ge \mu_b, \quad \mu_0 + \mu_1 + \mu_2 + \dots + \mu_b = n,$

and H is a $\mu_0 \times \mu_0$ constant matrix.

The matrices N_{μ_i} (i = 1, ..., b) are called the *nilpotent blocks*. The maximum size μ_1 of them is the *nilpotency index*, denoted by $\nu(A)$. It is obvious that ODEs have index zero, and algebraic equations have index one.

We denote by A[I, J] the submatrix of A(s) with row set $I \subseteq R$ and column set $J \subseteq C$, where R and C are the row set and the column set of A(s), respectively. Furthermore, we denote $w(I, J) = \deg \det A[I, J]$, where $w(\emptyset, \emptyset) = 0$ by convention. Then w enjoys the following property.

Lemma 2.3 ([12, pp. 287–289]). Let A(s) be a matrix pencil with row set R and column set C. For any $(I, J) \in \Lambda$ and $(I', J') \in \Lambda$, where $\Lambda = \{(I, J) \mid |I| = |J|, I \subseteq R, J \subseteq C\}$, both (VB-1) and (VB-2) below hold:

(VB-1) For any $i \in I \setminus I'$, at least one of the following two assertions holds:

(1a)
$$\exists j \in J \setminus J' : w(I, J) + w(I', J') \le w(I \setminus \{i\}, J \setminus \{j\}) + w(I' \cup \{i\}, J' \cup \{j\}),$$

(1b) $\exists h \in I' \setminus I : w(I, J) + w(I', J') \le w(I \setminus \{i\} \cup \{h\}, J) + w(I' \setminus \{h\} \cup \{i\}, J').$

(VB-2) For any $j \in J \setminus J'$, at least one of the following two assertions holds:

(2a)
$$\exists i \in I \setminus I' : w(I, J) + w(I', J') \le w(I \setminus \{i\}, J \setminus \{j\}) + w(I' \cup \{i\}, J' \cup \{j\}),$$

(2b) $\exists l \in J' \setminus J : w(I, J) + w(I', J') \le w(I, J \setminus \{j\} \cup \{l\}) + w(I', J' \setminus \{l\} \cup \{j\}).$

Let $\delta_k(A)$ denote the highest degree of a minor of order k in A(s):

$$\delta_k(A) = \max_{I,J} \{ w(I,J) \mid |I| = |J| = k, I \subseteq R, J \subseteq C \}.$$

The index $\nu(A)$ can be determined from $\delta_k(A)$ as follows.

Theorem 2.4 ([12, Theorem 5.1.8]). Let A(s) be an $n \times n$ regular matrix pencil. The nilpotency index $\nu(A)$ is given by

$$\nu(A) = \delta_{n-1}(A) - \delta_n(A) + 1.$$

3 Substitution Method

In this section, we introduce the substitution method for solving linear DAEs with constant coefficients. The substitution method eliminates some variables by replacement to obtain a smaller system than the original one. This is familiar as the solution method for simultaneous equations.

Let A(s) be an $n \times n$ regular matrix pencil with row set R and column set C, and B be a nonsingular constant submatrix of A with row set $X \subset R$ and column set $Y \subset C$. We transform A into \tilde{A} by row operations:

$$A = \begin{pmatrix} B & K \\ L & M \end{pmatrix} \to \tilde{A} = \begin{pmatrix} I & O \\ -LB^{-1} & I \end{pmatrix} \begin{pmatrix} B & K \\ L & M \end{pmatrix} = \begin{pmatrix} B & K \\ O & M - LB^{-1}K \end{pmatrix},$$
(2)

where $K = A[X, C \setminus Y]$, $L = A[R \setminus X, Y]$, and $M = A[R \setminus X, C \setminus Y]$. We denote $M - LB^{-1}K$ by D, which is not necessarily a matrix pencil.

Let \tilde{K} , \tilde{L} , \tilde{M} , and \tilde{D} denote the matrices obtained by replacing s with d/dt in K, L, M, and D, respectively. Consider the DAE

$$B\boldsymbol{x}_1(t) + \hat{K}\boldsymbol{x}_2(t) = \boldsymbol{f}_1(t), \qquad (3)$$

$$\hat{L}\boldsymbol{x}_1(t) + \hat{M}\boldsymbol{x}_2(t) = \boldsymbol{f}_2(t).$$
(4)

By applying the transformation shown in (2), we obtain

$$B\boldsymbol{x}_1(t) = \boldsymbol{f}_1(t) - \hat{K}\boldsymbol{x}_2(t), \tag{5}$$

$$\hat{D}\boldsymbol{x}_{2}(t) = \boldsymbol{f}_{2}(t) - \hat{L}B^{-1}\boldsymbol{f}_{1}(t).$$
(6)

The outline of the substitution method is as follows.

Phase 1: Solve the DAE (6) for $\boldsymbol{x}_2(t)$.

Phase 2: Solve the system of linear equations (5) for $x_1(t)$.

In the substitution method, the numerical difficulty is determined by the index $\nu(D)$ of the DAE (6). We show that $\nu(D)$ can be expressed in terms of the degrees of minors in A.

For each $i \in R$ and $j \in C$, let d_{ij} denote the degree of det $A[R \setminus \{i\}, C \setminus \{j\}]$. Then we have

$$d_{ij} = \deg \det \tilde{A}[R \setminus \{i\}, C \setminus \{j\}], \quad \forall i \in R \setminus X, \ \forall j \in C,$$

$$(7)$$

because we can transform $\tilde{A}[R \setminus \{i\}, C \setminus \{j\}]$ into $A[R \setminus \{i\}, C \setminus \{j\}]$ by row operations for each $i \in R \setminus X$ and $j \in C$. The index $\nu(D)$ can be rewritten as follows.

Theorem 3.1. For an $n \times n$ regular matrix pencil A(s), the index of D is given by

$$\nu(D) = \max_{i,j} \{ d_{ij} \mid i \in R \setminus X, j \in C \setminus Y \} - \delta_n(A) + 1.$$
(8)

Proof. We denote the size of D by m. By Theorem 2.4, we have $\nu(D) = \delta_{m-1}(D) - \delta_m(D) + 1$. Recall that $\tilde{A} = \begin{pmatrix} B & K \\ O & D \end{pmatrix}$ and that B is a constant matrix. It follows from det $A = \det \tilde{A}$ that

$$\delta_m(D) = \deg \det D = \deg \det A - \deg \det B = \deg \det A.$$

Moreover, we have

$$\delta_{m-1}(D) = \max_{I,J} \{ \deg \det D[I,J] \mid |I| = |J| = m-1 \}$$
$$= \max_{I,J} \{ \deg \det \tilde{A}[I,J] \mid |I| = |J| = n-1, I \supseteq X, J \supseteq Y \} - \deg \det B$$
$$= \max_{i,j} \{ d_{ij} \mid i \in R \setminus X, j \in C \setminus Y \},$$

where the last step is due to (7). Thus we obtain (8).

4 Index Reduction

Let A(s) = sP + Q be an $n \times n$ regular matrix pencil such that P has at most one nonzero entry in each row. We denote the row set of A(s) by R, and the column set by C. Moreover, we assume that $\nu(A)$ is positive. Let $Y \subseteq C$ be the set of indices such that their column vectors in P are zero vectors. Since A[R, Y] has full column rank by the regularity of A(s), we can find $X \subseteq R$ such that A[X, Y] is regular. Note that because B = A[X, Y] and $L = A[R \setminus X, Y]$ are constant matrices, $D = \tilde{A}[R \setminus X, C \setminus Y]$ is a matrix pencil. We prove that the index of D is one lower than that of A.

Lemma 4.1. For each $i \in R$ and each $j \in C \setminus Y$, we have $d_{ij} < \delta_{n-1}(A)$.

Proof. Suppose to the contrary that there exist $i \in R$ and $j \in C \setminus Y$ such that $d_{ij} = \delta_{n-1}(A)$. Let h be a row such that the (h, j) entry of P is nonzero. We put $(I, J) = (\{h\}, \{j\})$ and $(I', J') = (R \setminus \{i\}, C \setminus \{j\})$. By (VB-2) in Lemma 2.3, at least one of the following two assertions holds:

(2a) $h = i, w(\{h\}, \{j\}) + w(R \setminus \{i\}, C \setminus \{j\}) \le w(\emptyset, \emptyset) + w(R, C),$

(2b) $\exists l \in C \setminus \{j\} : w(\{h\}, \{j\}) + w(R \setminus \{i\}, C \setminus \{j\}) \le w(\{h\}, \{l\}) + w(R \setminus \{i\}, C \setminus \{l\}).$

Note that $w(\{h\}, \{j\}) = 1$ and $w(R \setminus \{i\}, C \setminus \{j\}) = d_{ij} = \delta_{n-1}(A)$.

If (2a) holds, then it follows from $w(\emptyset, \emptyset) = 0$ and $w(R, C) = \delta_n(A)$ that $1 + \delta_{n-1}(A) \le \delta_n(A)$, which implies $\nu(A) \le 0$ by Theorem 2.4. This contradicts $\nu(A) > 0$.

On the other hand, if (2b) holds, we have $1 + \delta_{n-1}(A) \leq w(\{h\}, \{l\}) + d_{il}$. Since *P* has at most one nonzero entry in each row, we have $w(\{h\}, \{l\}) = 0$. Thus we obtain $1 + \delta_{n-1}(A) \leq d_{il}$, which contradicts the definition of $\delta_{n-1}(A)$.

Theorem 4.2. The index of $D = \tilde{A}[R \setminus X, C \setminus Y]$ is exactly one lower than that of A.

Proof. By Theorems 2.4 and 3.1 and Lemma 4.1,

$$\nu(A) - \nu(D) = \delta_{n-1}(A) - \max_{i,j} \{ d_{ij} \mid i \in R \setminus X, j \in C \setminus Y \} > 0.$$

We now prove $\nu(D) \ge \nu(A) - 1$. It follows from Lemma 4.1 that there exist $i \in R$ and $j \in Y$ such that $d_{ij} = \delta_{n-1}(A)$.

Suppose that there exist $i \in R \setminus X$ and $j \in Y$ such that $d_{ij} = \delta_{n-1}(A)$. By applying (VB-2) in Lemma 2.3 to (X, Y) and $(R \setminus \{i\}, C \setminus \{j\})$, we have

$$\exists l \in C \setminus Y : w(X,Y) + w(R \setminus \{i\}, C \setminus \{j\}) \le w(X,Y \setminus \{j\} \cup \{l\}) + w(R \setminus \{i\}, C \setminus \{l\}).$$

Note that w(X,Y) = 0, because A[R,Y] is a constant matrix. Since A is a matrix pencil and A[X,Y] is a constant matrix, $w(X,Y \setminus \{j\} \cup \{l\}) \leq 1$. Therefore, we have $d_{ij} \leq d_{il} + 1$, which implies $\nu(D) \geq d_{il} - \delta_n(A) + 1 \geq d_{ij} - \delta_n(A) = \nu(A) - 1$ by Theorems 2.4 and 3.1.

We now consider the other case, which means that there exist $i \in X$ and $j \in Y$ such that $d_{ij} = \delta_{n-1}(A)$, and $d_{pq} < \delta_{n-1}(A)$ for any $p \in R \setminus X$ and $q \in Y$. By applying (VB-1) in Lemma 2.3 to (X, Y) and $(R \setminus \{i\}, C \setminus \{j\})$, at least one of the following assertions holds:

(1a) $w(X,Y) + w(R \setminus \{i\}, C \setminus \{j\}) \le w(X \setminus \{i\}, Y \setminus \{j\}) + w(R,C),$

(1b)
$$\exists h \in R \setminus X : w(X,Y) + w(R \setminus \{i\}, C \setminus \{j\}) \le w(X \setminus \{i\} \cup \{h\}, Y) + w(R \setminus \{h\}, C \setminus \{j\}).$$

Since A[R, Y] is a constant matrix, we have $w(X, Y) = w(X \setminus \{i\}, Y \setminus \{j\}) = w(X \setminus \{i\} \cup \{h\}, Y) = 0$. If (1a) holds, then we have $d_{ij} \leq \delta_n(A)$. Therefore, $\nu(A) = d_{ij} - \delta_n(A) + 1 \leq 1$ by Theorem 2.4.

It follows from the nonnegativity of $\nu(D)$ that $\nu(D) \ge \nu(A) - 1$. On the other hand, if (1b) holds, we have $d_{ij} \le d_{hj}$. This contradicts the assumption that

 $d_{pq} < \delta_{n-1}(A)$ for any $p \in R \setminus X$ and $q \in Y$.

Theorem 4.2 implies that the index of D is the same for any X with A[X, Y] being a nonsingular constant matrix.

5 Numerical Examples

In this section, we demonstrate the proposed method in numerical examples. We use RADAU5 [9] in Matlab as the DAE solver. RADAU5 is an implementation of a fifth order implicit Runge-Kutta method with three stages (RADAU IIA). This is applicable to ODEs and DAEs with index at most three.



Figure 1: Linear circuit described by circuit equations with index two.



Figure 2: Linear circuit described by circuit equations with index three.

Example 5.1 (Electric circuit with index two [4, 12]). Consider a circuit given in Figure 1, which is described by the circuit equations with index two:

The modified nodal analysis results in a DAE with index two. However, our method finds

$$X = \{r_1, r_4, r_5, r_6, r_7, r_8, r_9, r_{10}\} \text{ and } Y = \{c_1, c_2, c_3, c_5, c_6, c_7, c_8, c_9\},\$$

and the coefficient matrix pencil of the resulting DAE is

$$D = \begin{pmatrix} 1 + sL\left(\frac{1}{R_1} + \frac{1}{R_2}\right) & -\frac{1}{R_1} \\ 0 & -1 \end{pmatrix},$$

which has index one.

Example 5.2 (Electric circuit with index three [8]). Consider another circuit depicted in Figure 2,





Figure 3: The current through the inductance: numerical solutions of the original DAE (dashdotted line), the substitution method (solid line), and the exact solution (dotted line).

Figure 4: The error in the current through the inductance: the original DAE (dash-dotted line) and the substitution method (solid line).

which is described by the circuit equations with index three:

The modified nodal analysis results in a DAE with index three [8]. However, our method finds

$$X = \{r_1, r_4, r_5, r_6, r_7, r_8\}$$
 and $Y = \{c_1, c_2, c_3, c_5, c_7, c_8\},\$

and we obtain

$$D = \begin{pmatrix} 1 & saC \\ 0 & -1 \end{pmatrix},$$

which has index two.

Setting $C = 5[\mu F]$, L = 8[mH], a = 0.99, and $V(t) = 10 \sin(200t)[V]$, we numerically solve both the original and the resulting DAEs. Figure 3 presents these two numerical solutions and the exact

solution, which can be obtained analytically. In Figure 3, the exact solution coincides with the solution of the substitution method. Figure 4 shows the discrepancy between the two numerical solutions and the exact solution. It is observed that the index reduction effectively improves the accuracy of the numerical solution.

6 Conclusion

For linear DAEs with constant coefficients, we have proposed a new index reduction method. This method is applicable to all DAEs with at most one derivative per equality, and always reduces the index by one without introducing any additional variables. This class of DAEs includes the semi-explicit form and circuit equations of most linear time-invariant circuits.

Acknowledgements

The authors are grateful to Taketomo Mitsui, Kazuo Murota, and Caren Tischendorf for their helpful comments and suggestions.

References

- [1] K. E. Brenan, S. L. Campbell, and L. R. Petzold: Numerical Solution of Initial-Value Problems in Differential-Algebraic Equations, SIAM, Philadelphia, 2nd edition, 1996.
- [2] P. Bujakiewicz: Maximum Weighted Matching for High Index Differential Algebraic Equations, Doctor's dissertation, Delft University of Technology, 1994.
- [3] S. L. Campbell and C. W. Gear: The index of general nonlinear DAEs, Numerische Mathematik, vol. 72, pp. 173–196, 1995.
- [4] F. E. Cellier: Continuous System Modeling, Springer-Verlag, Berlin, 1991.
- [5] F. R. Gantmacher: The Theory of Matrices, Chelsea, New York, 1959.
- [6] C. W. Gear and L. R. Petzold: ODE methods for the solution of differential/algebraic systems, SIAM Journal on Numerical Analysis, vol. 21, pp. 716–728, 1984.
- [7] C. W. Gear: Differential-algebraic equation index transformations, SIAM Journal on Scientific and Statistical Computing, vol. 9, pp. 39–47, 1988.
- [8] M. Günther and P. Rentrop: The differential-algebraic index concept in electric circuit simulation, Zeitschrift für angewandte Mathematik und Mechanik, vol. 76, supplement 1, pp. 91–94, 1996.

- [9] E. Hairer and G. Wanner: Solving Ordinary Differential Equations II, Springer-Verlag, Berlin, 2nd edition, 1996.
- [10] P. Kunkel and V. Mehrmann: Index reduction for differential-algebraic equations by minimal extension, *Zeitschrift für angewandte Mathematik und Mechanik*, vol. 84, pp. 579–597, 2004.
- [11] S. E. Mattsson and G. Söderlind: Index reduction in differential-algebraic equations using dummy derivatives, SIAM Journal on Scientific Computing, vol. 14, pp. 677–692, 1993.
- [12] K. Murota: Matrices and Matroids for Systems Analysis, Springer-Verlag, Berlin, 2000.
- [13] S. Schulz: Four lectures on differential-algebraic equations, Technical Report 497, The University of Auckland, New Zealand, 2003.
- [14] D. E. Schwarz and C. Tischendorf: Structural analysis of electric circuits and consequences for MNA, International Journal of Circuit Theory and Applications, vol. 28, pp. 131–162, 2000.