BIT 16 (1976), 237–240

ON THE IMPLEMENTATION OF IMPLICIT RUNGE-KUTTA METHODS

J. C. BUTCHER

Abstract.

The modified Newton iterations in the implementation of an s stage implicit Runge-Kutta method for an n dimensional differential equation system require $2s^3n^3/3+O(n^2)$ operations for the LU factorisations and $2s^2n^2+O(n)$ operations for the back substitutions. This paper describes a method for transforming the linear system so as to reduce these operation counts.

In the numerical solution of an n dimensional stiff differential equation system

(1) $$y'(x) = f(y(x)) ,$$

using an s stage implicit Runge-Kutta method, the solution at $x_m = x_{m-1}+h$ is computed as

(2) $$y_m = y_{m-1}+h \sum_{j=1}^{s} b_j f(Y_j)$$

where

(3) $$Y_i = y_{m-1}+h \sum_{j=1}^{s} a_{ij} f(Y_j), \quad i = 1,2,\ldots,s .$$

To evaluate $Y_1, Y_2, \ldots, Y_s$ satisfying the system (3), it is usual to use a modification of the Newton-Raphson method so that at the end of a current iteration, Y_i is to be replaced by Y_i+w_i where $w_1,w_2,\ldots,w_s$ are given by

(4) $$w_i-h \sum_{j=1}^{s} a_{ij} J w_j - Z_i = 0, \quad i = 1,2,\ldots,s ,$$

with J, the $n \times n$ Jacobian matrix of f, evaluated at a recent point on the solution trajectory and

(5) $$Z_i = -Y_i+y_{m-1}+h \sum_{j=1}^{s} a_{ij} f(Y_j), \quad i = 1,2,\ldots,s .$$

Since a major part of the computation time is expended in the evaluation of J and the treatment of the linear system (4), it is standard practice to evaluate J as seldom as possible and to carry out preliminary work on the linear system (4) so that the actual iterations can be per-

Received March 19, 1976.

formed efficiently. It is with this preliminary work that this paper is mainly concerned.

Let $w, Z \in \mathsf{R}^{ns}$ and the $s \times s$ matrix A be defined by

$$w = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_s \end{bmatrix}, \quad Z = \begin{bmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_s \end{bmatrix}, \quad A = \begin{bmatrix} a_{11} a_{12} \dots a_{1s} \\ a_{21} a_{22} \dots a_{2s} \\ \vdots \quad \vdots \quad\quad \vdots \\ a_{s1} a_{s2} \dots a_{ss} \end{bmatrix}$$

and let $M = \bar{I} \otimes I - hA \otimes J$ be the matrix of coefficients in (4) where $\bar{I}$ is the $s \times s$ unit matrix and I the $n \times n$ unit matrix. Thus (4) can be written as

$$Mw - Z = 0 . \tag{6}$$

Throughout this paper we will assume that A is non-singular. This assumption holds for most implicit Runge-Kutta methods that have been proposed as suitable for stiff problems, and leads to some simplifications in this paper.

We will regard it as the standard practice to compute the LU factorisation of M as the preliminary treatment of (4). In this case, the number of multiplicative and additive calculations to perform are each $C(n^3/3) + O(n^2)$ (for large n), where $C = s^3$, the number of operations in the back substitution for each iteration is $Dn^2 + O(n)$ where $D = s^2$. We will consider how the factors C, D can be lowered, either through the choice of parameters or else through a suitable organisation of the work.

Let P, Q be non-singular $s \times s$ matrices and let

$$\tilde{w} = (Q^{-1} \otimes I)w, \quad \tilde{Z} = (P \otimes I)Z ,$$

$\tilde{M} = (P \otimes I)M(Q \otimes I) = (PQ) \otimes I - h\tilde{A} \otimes J$ where $\tilde{A} = PAQ$ so that (6) is equivalent to

$$\tilde{M}\tilde{w} - \tilde{Z} = 0 . \tag{7}$$

Since the computation of $\tilde{Z}$ from Z and of w from $\tilde{w}$ each require $O(n)$ multiplicative and additive calculations, we might just as well use this transformed version of these equations if this leads to some advantage.

We now consider how to make a judicious choice of P and Q. Let the Jordan canonical form of A^{-1} be

$$T^{-1}A^{-1}T = \begin{bmatrix} \lambda_1^{-1} & 0 & 0 & \dots & 0 \\ \mu_1 & \lambda_2^{-1} & 0 & \dots & 0 \\ 0 & \mu_2 & \lambda_3^{-1} & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & \lambda_s^{-1} \end{bmatrix}$$

where each subdiagonal element $\mu_i\,(i=1,2,\ldots,s-1)$ is zero if $\lambda_i \neq \lambda_{i+1}$ and is either zero or an arbitrary non-zero number if $\lambda_i=\lambda_{i+1}$. Where it is non-zero we suppose that $\mu_i=\lambda_i^{-1}$. Let $D=\operatorname{diag}(\lambda_1,\lambda_2,\ldots,\lambda_s)$. We select $P=DT^{-1}A^{-1}$, $Q=T$ so that

$$PQ = \begin{bmatrix} 1 & 0 & 0 & \ldots & 0 \\ \varepsilon_1 & 1 & 0 & \ldots & 0 \\ 0 & \varepsilon_2 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \ldots & 1 \end{bmatrix}$$

where each of the subdiagonals $\varepsilon_1,\varepsilon_2,\ldots$ is either 0 or 1, and $PAQ=D$.

The matrix $\tilde{M}$ now consists of diagonal blocks of the form $I-h\lambda J$ together with subdiagonal blocks of 0 (the zero matrix) or I. The preliminary treatment of (4) now consists of the LU factorisation of each of the *distinct* diagonal blocks and the back substitutions break into s separate blocks with the subdiagonal elements of PQ contributing only a further $O(n)$ operations.

To assess the factors C and D, we must take into account the possible presence of non-real eigenvalues of A. Let α denote the number of distinct real eigenvalues and β the total number of real zeros of the characteristic polynomial of A. Also let γ denote the number of distinct conjugate complex eigenvalue pairs and δ the total number of conjugate pairs of zeros of the characteristic polynomial of A. Thus $\alpha \leqq \beta, \gamma \leqq \delta$, $\beta+2\delta=s$. Since complex multiplications and additions require the time of 4 real multiplications and additions, we have $C=\alpha+4\gamma, D=\beta+4\delta$.

Consider, for example, the implicit Runge-Kutta methods of order $2s$, [1]. For these methods A has at most one real eigenvalue and all zeros of the characteristic polynomial are distinct. Hence, $\alpha=\beta=0$ (s even) and $\alpha=\beta=1$ (s odd); $\gamma=\delta=[s/2]$. In this case $C=D=2s$ (s even) and $C=D=2s-1$ (s odd), a marked improvement for s greater than 2 over the standard values of $C=s^3, D=s^2$.

For the semi-explicit methods of Nørsett [2], where all diagonals of A are equal, we find $C=1, D=s$. However, if the transformation described here is not applied we would have $C=1, D=s(s+1)/2$. It is interesting to note that values of C and D identical to those for Nørsett's methods could also be obtained for any method which is not necessarily semi-explicit but for which the characteristic polynomial of A has only a single s-fold zero.

REFERENCES

1. J. C. Butcher, *Implicit Runge-Kutta Processes*, Math. Comp. 18 (1964), 50–64.
2. S. P. Nørsett, *Semi Explicit Runge-Kutta Methods*, Mathematics Department, University of Trondheim, Reprint No. 6/74.

DEPARTMENT OF MATHEMATICS
THE UNIVERSITY OF AUCKLAND
AUCKLAND, NEW ZEALAND