Computers Math. Applic. Vol. 19, No. 2, pp. 83–89, 1990
Printed in Great Britain. All rights reserved

0097-4943/90 $3.00 + 0.00
Copyright © 1990 Pergamon Press plc

AVOIDING THE EXACTNESS OF THE JACOBIAN MATRIX IN ROSENBROCK FORMULAE

H. ZEDAN
Computer Science Department, University of York, Heslington, York YO1 5DD, England

(Received 6 January 1989)

Abstract—A new class of methods, for solving stiff systems, which avoids the exactness of the Jacobian matrix is introduced. The order conditions for methods of order $p \leqslant 5$ are given. The linear stability properties for such methods are analysed; numerical testings are also included

1. INTRODUCTION

It is widely recognized that the linearly implicit Rosenbrock methods (with its variations) and the diagonally implicit methods are major competitors for the backward differentiation formulae for the numerical solution of the stiff systems:

$$\frac{\mathrm{d}\mathbf{y}}{\mathrm{d}x} = \mathbf{f}(\mathbf{y}(x)), \quad \mathbf{y}(x_0) = \mathbf{y}_0, \quad x \geqslant x_0. \tag{1}$$

However, the two classes of methods share a common difficulty and that is the greater need for the exact Jacobian matrix of the system considered. For example, any Rosenbrock method requires the evaluation of the "exact" Jacobian matrix at every integration step which makes it less attractive for integrating large systems with expensive function and Jacobian evaluations.

A new class of linearly implicit methods, in which the exact evaluation of the Jacobian matrix is avoided, in introduced. This class of methods is a modification to the well known ROW-methods [1]. An s-stage Modified ROW-method (MROW for short) takes the form

$$\mathbf{y}_{n+1} = \mathbf{y}_n + h_{n+1}\sum_{i=1}^{s} b_i\mathbf{k}_i, \tag{2a}$$

where

$$\mathbf{M}\mathbf{k}_i = \mathbf{f}\left(\mathbf{y}_n + h_{n+1}\sum_{j=1}^{i-1} a_{ij}\mathbf{k}_j\right) + h_{n+1}\mathbf{A}_n\sum_{j=1}^{i-1} d_{ij}\mathbf{k}_j, \quad i = 1, 2, \ldots, s, \tag{2b}$$

$$\mathbf{M} = \mathbf{I} - h_{n+1}d\mathbf{A}_n$$

and

$$\mathbf{A}_n = \mathbf{J}_n + h_{n+1}\mathbf{B}. \tag{2c}$$

The coefficients b_i, a_{ij}, d_{ij} and d are assumed to be real and h_{n+1} is the step length. $\mathbf{J}_n$ is the Jacobian matrix evaluated at $\mathbf{y}_n$ and $\mathbf{B}$ is any real square matrix that is to represent a perturbation in $\mathbf{J}_n$ where it has been computed numerically. This, however, represents an assumption that the error in the approximate Jacobian is $O(h)$ as $h \to 0$; and that in practice the matrix $\mathbf{B}$ is comparable in size to $\mathbf{J}_n$ for range of step sizes that occur. In addition, we may view the matrix $\mathbf{B}$ as a representation of the accumulative error that is to be expected when the same Jacobian matrix is used over several steps.

Rosenbrock–Wanner (ROW) methods can be obtained by setting $\mathbf{B} = 0$ in equation (2c) above. This class of methods has been thoroughly examined by many authors: Kaps and Rentrop [2], have constructed a fourth order method with local error estimate, and Kaps and Wanner [3], have considered the construction of higher orders ROW-type methods.

Codes which are based on any ROW-method employ strategies such as keeping the Jacobian value fixed for some integration steps. Further, an approximation to the Jacobian matrix is often used. These strategies, however, are incompatible with the way the methods were originally derived. In other words, keeping the Jacobian fixed has an effect on both the order and the stability properties of the method concerned.

The class of methods proposed in this paper is an attempt to deal with this problem. In Section 2 the order conditions of MROW-methods (of order $p \leqslant 5$) are given. The linear stability properties of MROW-methods are discussed in Section 3. A second order two-parameter family and a third order MROW-methods, each with local error estimate are given in Section 4. Some numerical results are also included.

2. ORDER CONDITIONS

The order of the method described by formula (2) can be defined in the usual way as follows. Let $h_{n+1} = h$ and $\Delta y = y(x + h) - y(x)$, where $y(x)$ is the true solution of equation (1). Assume that $y_n = y(x)$. Consider

$$\phi_i(h) = \Delta y - h \sum_{i=1}^{s} b_i \mathbf{k}_i .$$

Using Taylor expansion of $\phi_i(h)$ we get $(0 < \theta < 1)$

$$\phi_i(h) = \sum_{j=0}^{p} \left(\frac{h^j}{j!}\right)\phi_i^{(j)}(0) + \left(\frac{h^{p+1}}{(p+1)!}\right)\phi_i^{(p+1)}(\theta h).$$

Thus, an MROW-method is of order p when

$$\phi_i^{(j)}(0) = 0, \quad \text{for } j = 1(1)p,$$

$$\phi_i^{(p+1)}(0) \neq 0.$$

As in the case of ROW-method, it can be easily shown that the maximum order of an s-stage MROW-method is $s + 1$. Table 1 contains the order conditions for MROW-methods for $p \leqslant 5$ each

Table 1

Order	Elementary differential	Order conditions
1	f	$\Sigma b_i = 1$
2	$f'f$	$\Sigma b_i N_i = \frac{1}{2} - d$
3	$f'f'f$	$\Sigma b_i K_{ij} N_j = \frac{1}{6} - d + d^2$
	$f''ff$	$\Sigma b_i M_i^2 = \frac{1}{3}$
	Bf	$\Sigma b_i D_i = -d$
4	$f'f'f'f$	$\Sigma b_i K_{ij} K_{jk} N_k = \frac{1}{24} - \frac{1}{2}d + \frac{3}{2}d^2 - d^3$
	$f'f''ff$	$\Sigma b_i K_{ij} M_j^2 = \frac{1}{12} - \frac{1}{3}d$
	$f''ff'f$	$\Sigma b_i M_i a_{ij} N_j = \frac{1}{8} - \frac{1}{3}d$
	$f'''fff$	$\Sigma b_i M_i^3 = \frac{1}{4}$
	$f'Bf$	$\Sigma b_i K_{ij} D_j = d(d - \frac{1}{2})$
	$Bf'f$	$\Sigma b_i d_{ij} N_j = d(d - \frac{1}{2})$
5	$f'f'f'f'f$	$\Sigma b_i K_{ij} K_{jk} K_{kl} N_l = \frac{1}{120} - \frac{1}{6}d + d^2 - 2d^3 + d^4$
	$f'f'f''ff$	$\Sigma b_i K_{ij} K_{jk} M_k^2 = \frac{1}{60} - \frac{1}{6}d + \frac{1}{3}d^2$
	$f'f''ff'f$	$\Sigma b_i K_{ij} M_j a_{jk} N_k = \frac{1}{40} - \frac{5}{24}d + \frac{1}{3}d^2$
	$f'f'''fff$	$\Sigma b_i K_{ij} M_j^3 = \frac{1}{20} - \frac{1}{4}d$
	$f''ff'f'f$	$\Sigma b_i M_i a_{ij} K_{jk} N_k = \frac{1}{30} - \frac{1}{4}d + \frac{1}{3}d^2$
	$f''ff''ff$	$\Sigma b_i M_i a_{ij} M_j^2 = \frac{1}{15}$
	$f''f'ff'f$	$\Sigma b_i (a_{ij} N_j)^2 = \frac{1}{20} - \frac{1}{4}d + \frac{1}{3}d^2$
	$f'''fff'f$	$\Sigma b_i M_i^2 a_{ij} N_j = \frac{1}{10} - \frac{1}{4}d$
		$f^{(4)}ffff$
	$f'f'Bf$	$\Sigma b_i K_{ij} K_{jk} D_k = d[-d^2 + d - \frac{1}{6}]$
	$f'Bf'f$	$\Sigma b_i K_{ij} d_{jk} D_k = d[-\frac{1}{2} + \frac{3}{2}d - d^2]$
	$f''fBf$	$\Sigma b_i M_i a_{ij} d_j = -\frac{1}{3}d$
	$Bf'f'f$	$\Sigma b_i D_i K_{ij} N_j = d[-\frac{1}{6} + d - d^2]$
	$Bf''ff$	$\Sigma b_i d_{ij} M_j^2 = -\frac{1}{3}d$
	BBf	$\Sigma b_i d_{ij} D_j = d$

$M_i = \sum_{j=1}^{i-1} a_{ij}, \quad N_i = \sum_{j=1}^{i-1} K_{ij}, \quad D_i = \sum_{j=1}^{i-1} d_{ij}$

with its associated elementary differential. The derivation of the table is lengthy and is therefore omitted.

3. STABILITY

The application of an s-stage MROW-method of order $p \geqslant s$ to the scalar test equation

$$y' = \lambda, \quad \lambda \in C^-$$

yields

$$y_{n+1} = \hat{R}(z;\alpha)y_n,$$

with the stability function

$$\hat{R}(z;\alpha) = (1-\alpha dz)^{-s}\sum_{i=0}^{s}(-\alpha dz)^i L_i^{(s-i)}\left(\frac{1}{\alpha d}\right), \tag{3}$$

where $z = h\lambda$, $\alpha z = ha_n$ and $L_s(z)$ is the Laguerre polynomial of order s such that

$$L_j^{(k)}(y) = \sum_{i=0}^{j}(-1)^i\left(\frac{j-i}{j+k}\right)\frac{y^i}{i!}.$$

In general α is a complex number, where

$$\alpha = 1 + \frac{\delta}{\lambda}$$

$$= 1 + \frac{h\delta}{z}$$

and

$$\delta = a_n - \lambda.$$

Therefore, we view $\alpha - 1$ as the *relative error* in the exact Jacobian which is introduced either by its numerical approximation or by keeping its value fixed for some integration steps.

Similar to the ROW-method, the stability properties of the MROW-methods can be related to the *acceptability* properties of the rational function $\hat{R}(z;\alpha)$. We define the $\hat{A}$-acceptability of $\hat{R}(z;\alpha)$ as follows.

Definition

For a given $\alpha \in C$, $\hat{R}(z;\alpha)$ is said to be $\hat{A}$-acceptable iff

$$|\hat{R}(z;\alpha)| \leqslant 1, \quad \forall z \in C^-. \qquad \blacksquare$$

Consequently, an MROW-method is $\hat{A}$-stable iff $\hat{R}(z;\alpha)$ is $\hat{A}$-acceptable. The above definition requires the boundedness of the numerical solutions $\{y_n\}$ for a given perturbation to the exact Jacobian.

Similarly, an MROW-method is said to be $\hat{L}$-stable iff it is $\hat{A}$-stable and $|\hat{R}(z;\alpha)| \to 0$ as $z \to \infty$.

The natural question is therefore: how much is the error in the Jacobian matrix allowed to accumulate so that $\hat{R}(z;\alpha)$ is $\hat{A}$-acceptable? Generally, the answer depends on the number of stages of the particular method. For example, the following results establish an upper bound for the error in the Jacobian matrix for the cases $s = 1, 2$.

Theorem

(1) A general 1-stage MROW-method is $\hat{A}$-stable iff $\frac{1}{2} \leqslant d < \infty$ and $\kappa(=\alpha - 1)$ is such that $\mathrm{Re}(\kappa) \geqslant 1/2d - 1$.

(2) A general 2-stage MROW-method is $\hat{A}$-stable iff $\frac{1}{4} \leqslant d < \infty$ and κ is such that $\mathrm{Re}(\kappa) \geqslant 1/4d - 1$.

Proof. (1) From equation (3) with $s = 1$ we get

$$\hat{R}(z;\alpha) = \frac{1 + z(1 - d\alpha)}{1 - zd\alpha},$$

$$= \frac{\hat{P}(z;\alpha)}{\hat{Q}(z;\alpha)}.$$

Set $\alpha = u + iv$, where, $u, v \in \mathbf{R}$. To establish the $\hat{A}$-acceptability of $\hat{R}(z;\alpha)$, and hence the $\hat{A}$-stability for the method, we construct the corresponding E-polynomial [4], where for all $y \in \mathbf{R}$

$$E(iy;\alpha) = |\hat{P}(iy;\alpha)|^2 - |\hat{Q}(iy;\alpha)|^2$$

$$= y^2(2du - 1).$$

For $\hat{A}$-acceptability, we require that

$$E(iy;\alpha) \geqslant 0, \quad \forall y \in \mathbf{R},$$

which gives the stated bound.

(2) For $s = 2$, we follow similar argument as in (1) to get

$$E(iy;\alpha) = y^4[[4(du)^3 - 5(du)^2 + 2(du) - \tfrac{1}{4}] + (dv)^2[4(du) - 3]] + y^3[2(dv)[2(du) - 1]].$$

We require

$$E(iy;\alpha) \geqslant 0, \forall \in \mathbf{R},$$

i.e.

$$2(dv)[2(du) - 1] = 0,$$

for which $(dv) = 0$. Thus, $E(iy;\alpha) \geqslant 0$, $\forall y \in \mathbf{R}$ if and only if

$$[4(du)^3 - 5(du)^2 + 2(du) - \tfrac{1}{4}] \geqslant 0,$$

which implies

$$u \geqslant \frac{1}{4d}$$

and hence the stated results. ■

Notes

Similar results may be obtained for higher values of s. The lack of space does not permit us to do so.

4. FORMULAE

Using the order condition of Table 1 we can construct MROW-methods with several orders of accuracy (for $p \leqslant 5$). For each method and estimator of the local truncation error is provided. The well known embedded technique will be adopted. In order to reduce the computational cost involved, an optimal number of stages for both the basic integration method and the local error estimator must be used.

4.1. Second order formula

A second order MROW-method with a third order error estimate may be designed by solving seven order conditions [cf. Table (1)]. However, in order to keep the number of function-evaluations in the formulae optimal (i.e. two function-evaluations only) the following constraint is applied

$$M_2 = M_3,$$

i.e.

$$a_{31} = a_{21}$$

and

$$a_{32} = 0.$$

The following solutions give us a two-parameter family of second order MROW-method with built-in error estimate

$$
\begin{aligned}
&a_{21} = \tfrac{2}{3}, \quad b_2^{[2]} = \gamma, \quad b_1^{[2]} = 1 - \gamma,\\
&d_{21} = \frac{3 - 6d - 4\gamma}{6\gamma}, \quad d_{32} = \beta,\\
&b_1^{[3]} = 0.25, \quad b_3^{[3]} = \frac{6d^2 - 6d + 1}{\beta(4 + 6d_{21})}, \quad b_2^{[3]} = 0.75 - b_3^{[3]},\\
&d_{31} = -\frac{d + b_2^{[3]} d_{21} + b_3^{[3]}\beta}{b_3^{[3]}},
\end{aligned}
$$

where γ, β and d are free parameters such that $\gamma \neq 0$, $\beta \neq 0$, $b_3^{[3]} \neq 0$ and $N_2 \neq 0$.

The choice of γ, β and d depends on several factors: Stability, minimization of error constants and computational efforts of the formulae. The results of the previous section showed that the value $d = 1 - 1/\sqrt{2}$ made the second and third order formulae, $\hat{L}$-stable and $\hat{A}$-stable, respectively. The parameter γ on the other hand was chosen to minimize the error constant in the basic integration formula. The local truncation error is given by

$$\mathbf{l}_{n+1} = h^3 G_{3,n} + h^4 G_{4,n} + O(h^5),$$

where

$$G_{3n} = A_1 f'f'f + A_2 f''ff + A_3 Bf$$

and

$$G_{4,n} = A_4 f'Bf + A_5 Bf'f + A_6 f'''fff + A_7 f''f'ff + A_8 f'f''ff + A_9 f'^3 f,$$

where the coefficients A_i, $\forall i = 1, 2, \ldots, 9$ are functions of the formula's parameters.

Since the formula's error estimate will calculate $h^3 G_{3,n}$, then there is no need to choose γ or β such that the coefficients Bf in G_{3n} is cancelled. Also the error estimate works well under the assumption that $G_{4,n}$ is small enough. However, this assumption is no longer valid in our case. As we are planning to use an approximate Jacobian and to keep its value fixed for as long as we can, the coefficients A_4 and A_5 will be quite large if the system is highly stiff and heavily nonlinear. Thus, the free parameter γ is chosen to cancel those coefficients from $G_{4,n}$. Hence, the value $\gamma = 3(1 - d)/2$. As for β it was chosen to reduce the computational effort involved. Thus the value of $\beta = 1$ was chosen.

4.2. *Third order formula*

For a third order method with a fourth order error estimate, 16 order conditions must be satisfied [see Table (1)]. The following was assumed:

$$a_{4i} = b_i, \quad \forall i = 1, 2, 3.$$

In this case the formulae use only three function-evaluations per every accepted step otherwise four function-evaluations are used. The following is the analytic solution for the formulae's coefficients:

$$
\begin{aligned}
&a_{21} = \frac{4d(1 - 3d)}{1 - 12d^2}, \quad d_{21} = \frac{-24d^2(1 - 3d)(1 - 4d)}{(1 - 12d^2)^2},\\
&a_{31} = 0.5 - \frac{(1 - 4d)(1 - 12d^2)^2}{32d(1 - 6d + 12d^2)(1 - 3d)}, \quad a_{32} = 0.5 - a_{31},\\
&d_{31} = -1.5d + \frac{(1 - 12d^2)^2}{8(1 - 6d + 12d^2)}, \quad d_{32} = -1.5d - d_{31},\\
&a_{42} = \frac{-(1 - 12d^2)^2}{24d(1 - 2d)(1 - 3d)(1 - 6d)}, \quad a_{43} = \frac{8(d^2 - d + \frac{1}{6})}{(1 - 2d)(1 - 6d)},\\
&a_{41} = 1 - a_{42} - a_{43},\\
&d_{42} = \frac{2a_{42}(1 - 24d + 144d^2 - 324d^3 + 216d^4)}{1 - 6d + 12d^2}.
\end{aligned}
$$

$$d_{43} = \frac{2(1 - 24d + 96d^2 - 72d^3)}{3(1 - 2d)(1 - 6d)}, \quad d_{41} = -(d_{42} + d_{43}),$$

$$\hat{b}_1 = \tfrac{1}{6}, \quad \hat{b}_2 = 0, \quad \hat{b}_3 = \tfrac{2}{3}, \quad \hat{b}_4 = \tfrac{1}{6}$$

Results similar to that given in Section 3 showed that the value $d = 0.435\,866\,5215$ gives an $\hat{A}$-stable fourth order method and an $\hat{L}$-stable third order method.

5. NUMERICAL TESTS

The methods derived in the previous section were implemented in the codes MROW23 and MROW34. The codes were tested over a wide range of stiff systems. For the lack of space, we shall only present results obtained by MROW34 on two sample test problems. Our stepsize strategy (for MROW34) can be summarized as follows. For the user specified tolerance, TOL

(a) Calculate RATIO = $(\text{TOL EST})^{\frac{1}{4}}$
(b) $H_{\text{new}} = H_{\text{old}} * \text{RATIO}$
(c.1) EST $\leqslant$ 2. TOL then accept step and advance with H_{new}. Otherwise reject step and repeat step with H_{new}.
(c.2) If 0.2 * TOL $\leqslant$ EST $\leqslant$ 2. *TOL then accept step and advance with $H_{\text{new}} = H_{\text{old}}$.
(d) LU-decomposition was updated whenever a new h was used. The Jacobian matrix was approximated by numerical differences and updated whenever there was a change in stepsize.

(Note that EST $= |y_{n+1}^{[4]} - y_{n+1}^{[3]}|$.) The two sample examples are

Example 1

$$y_1' = 0.01 - [1 + (y_1 + 1000)(y_1 + 1)][0.01 + y_1 + y_2],$$
$$y_2' = 0.01 - [1 + y_2^2][0.01 + y_1 + y_2],$$
$$y_1(0) = y_2(0) = 0,$$
$$\bar{y}_1(100) = -0.99164207, \quad \bar{y}_2(100) = 0.98333636.$$

Example 2

$$y_1' = 0.04 - 0.04(y_1 + y_2) + 10^4 y_1 y_2 - 3 * 10^7 y_1^2,$$
$$y_2' = 3 * 10^7 y_1^2,$$
$$y_1(0) = y_2(0) = 0,$$
$$\bar{y}_1(10) = 0.1623391063 * 10^{-4}, \quad \bar{y}_2(10) = 0.1586138424.$$

The values of $\bar{y}$ were obtained using the NAG library routine C02QBF [5], which is based on the GEAR method with TOL = 10^{-9}. The testing results are given in Tables 2 and 3 below.

Table 2 Results of Problem 1

TOL	NSTEP	NFCN	NJAC	NLU	Y_{appr}
10^{-1}	7	22	7	7	−0.976 120 548 0.966 645 051
10^{-3}	19	76	15	15	−0.995 752 267 0.987 006 487
10	224	703	22	22	−0.991 163 748 0.983 567 180

Table 3 Results of Problem 2

TOL	NSTEP	NFCN	NJAC	NLU	Y_{appr}
10^{-1}	5	5	16	5	0.000 016 090 0.160 615 315
10^{-3}	10	8	31	8	0.000 016 187 0.158 894 169
10^{-6}	24	11	82	11	0.000 016 222 0.158 601 851

6. RELATED WORK AND REMARKS

The Steihaug and Wolfbrandt paper [1], was the first to consider the problem of using inexact Jacobian matrix in ROW-methods. In Ref. [1] they have considered formula (2) without condition (2c) and constructed a second order method with built-in error estimate for any square matrix **A**; higher order methods of this type are not possible.

Day and Murthy [6], have introduced two classes of Rosenbrock-type methods (called generalized Runge–Kutta) and derived a second and a third order processes which are internally S-stable only when an accurate Jacobian matrix is used. However, it was claimed that the processes remain stable when an approximate Jacobian is used.

The class of methods presented in this paper is an attempt towards overcoming the limitations in any linearly implicit Rosenbrock-type formula for solving stiff systems. A class of MROW-methods was introduced which avoids the exactness of the Jacobian matrix. A second and a third order MROW-methods were derived. These methods remain consistent and highly stable when an inexact Jacobian matrix is used and or its value kept unchanged for some integration steps. Initial test results showed that the new methods have performed quite well. However, more tests are needed to investigate the performance of MROW-methods on large stiff systems. Strategy for the automatic update of the Jacobian matrix, which is based on an upper bound of the error κ (see Section 3), needs further investigation.

REFERENCES

1. T. Steihung and A. Wolfbrandt, An attempt to avoid exact Jacobian and non-linear equations in the numerical solution of stiff differential equations. *Math. Comput.* **33,** 521–534 (1979)
2. P. Kaps and P Rentrop, Generalized Runge–Kutta methods of order 4 with stepsize control for stiff ordinary differential equations *Numer. Math.* **33,** 55–68 (1979)
3. P. Kaps and G. Wanner, A study of Rosenbrock-type methods of high order *Num. Math.* **38,** 279–298 (1981).
4. S. P. Norsett, C-Polynomials for rational approximation to the exponential function. *Num. Math.* **25,** 39–56 (1975)
5. Numerical Algorithm Group Library Manual, Mark 7 NAG Central Office, 7 Banbury Rd, Oxford, England (1979)