Hairer, E. and Wanner, G.

On a Generalization of a Theorem of von Neumann

Studying the error growth function of implicit Runge-Kutta methods, we are lead in a natural way to a problem that can be viewed as a generalization of a well-known theorem of von Neumann.

1. Introduction

We consider nonlinear systems of differential equations $y' = f(t,y)$, where the function $f : \mathbb{R} \times \mathbb{C}^n \to \mathbb{C}^n$ is supposed to satisfy the one-sided Lipschitz condition

$$\Re\langle f(t,y) - f(t,\widehat{y}), y - \widehat{y}\rangle \le \nu \|y - \widehat{y}\|. \tag{1}$$

Such systems can be arbitrarily stiff. They are characterized by the fact that any two solutions satisfy for $h > 0$

$$\|y(t+h) - \widehat{y}(t+h)\| \le e^{h\nu}\|y(t) - \widehat{y}(t)\|.$$

We are interested in an analoguous estimate for the difference of any two numerical solutions.

Applying a Runge-Kutta method to $y' = f(t,y)$ (with two different initial values y_0 and $\widehat{y}_0$), we get the relation

$$\Delta y_1 = \Delta y_0 + h\sum_{i=1}^{s} b_i \Delta f_i, \qquad \Delta g_i = \Delta y_0 + h\sum_{j=1}^{s} a_{ij}\Delta f_j, \tag{2}$$

where $\Delta y_0 = y_0 - \widehat{y}_0$, $\Delta y_1 = y_1 - \widehat{y}_1$, $\Delta g_i = g_i - \widehat{g}_i$ denote the differences of initial values, numerical solution after one step and internal stage values, respectively, and $\Delta f_i = f(t_0 + c_i h, g_i) - f(t_0 + c_i h, \widehat{g}_i)$. For the following analysis it is useful to write this difference (multiplied by h) as

$$h\Delta f_i = Z_i \Delta g_i \qquad \text{with} \qquad Z_i = h\int_0^1 \frac{\partial f}{\partial y}(t_0 + c_i h, \widehat{g}_i + \tau\Delta g_i)\, d\tau. \tag{3}$$

As a consequence of (1) the matrices Z_i satisfy

$$\mu(Z_i) \le x \qquad \text{with} \qquad x = h\nu, \tag{4}$$

where $\mu(Z) = \sup_{\|v\|=1}\langle v, Zv\rangle$ denotes the logarithmic norm of Z. Inserting (3) into the second relation of (2), we can expess Δg_i and then also Δy_1 in terms of $Z_1, \ldots, Z_s$ as follows:

$$\Delta y_1 = K(Z_1, \ldots, Z_s)\Delta y_0, \tag{5}$$

where

$$K(Z_1, \ldots, Z_s) = I + (b^T \otimes I)\big(I \otimes I - (A \otimes I)Z\big)^{-1}(\mathbb{1} \otimes I), \tag{6}$$

Z is the block diagonal matrix with $Z_1, \ldots, Z_s$ as entries in the diagonal, $b^T = (b_1, \ldots, b_s)$, $A = (a_{ij})$, $\mathbb{1}^T = (1, \ldots, 1)$, and I is the identity matrix. From (5) we immediately get the estimate

$$\|\Delta y_1\| \le \varphi(h\nu)\|\Delta y_0\| \qquad \text{with} \qquad \varphi(x) = \sup_{\mu(Z_1)\le x, \ldots, \mu(Z_s)\le x} \|K(Z_1, \ldots, Z_s)\|. \tag{7}$$

This estimate is uniform in all problems $y' = f(t,y)$ satisfying the condition (1). The function $\varphi(x)$ is called *error growth function* of the method.

As an example consider the two-stage Radau IIA method, for which we have

$$K(Z_1, Z_2) = \Big(I - \frac{5}{12}Z_1 - \frac{1}{4}Z_2 + \frac{1}{6}Z_1 Z_2\Big)^{-1}\Big(I + \frac{1}{3}Z_1\Big).$$

Observe that, in general, the matrices Z_1 and Z_2 do not commute and that the nominator and denominator of the "rational function" depend only linearly on each of Z_1 and Z_2.

For a computation of $\varphi(x)$ the formula in Eq. (7) is not very practical, because one has to search the supremum over the s matrices $Z_1, \ldots, Z_s$, whose dimension is not limited.

2. Formulation of the Problem

In the classical papers [2], [3], Burrage and Butcher have given upper bounds of the error growth function $\varphi(x)$ for various low-stage Runge-Kutta methods. We observed that, always when their estimate is optimal, it is equal to

$$\varphi_K(x) = \sup_{\Re z_1 \le x, \ldots, \Re z_s \le x} |K(z_1, \ldots, z_s)|. \tag{8}$$

This motivates the study of the following question.

Problem. Let $K(Z_1, \ldots, Z_s)$ be given by (6). In which situations is it true that the functions $\varphi(x)$ and $\varphi_K(x)$ of (7) and (8), respectively, are identical for all x?

In the case $s = 1$ the answer of this question is affirmative. Indeed, it is a consequence of a theorem of von Neumann (see e.g., [4], Sect. IV.11). The same will be proved below for $s = 2$, if the nominator and denominator of the rational function $K(z_1, \ldots, z_s)$ do not have a common factor. For the moment, it is not clear to us whether $\varphi(x)$ and $\varphi_K(x)$ are equal for every irreducible Runge-Kutta method.

3. A Related Optimization Problem

The computation of $\varphi(x)$, defined in (7), is of course equivalent to searching the maximum of $\|\Delta y_1\|$ under the restriction (1). It is therefore natural to consider the following inequality constrained optimization problem:

$$\frac{1}{2}\|\Delta y_1\|^2 \to \max, \qquad \Re\langle \Delta f_i, \Delta g_i\rangle \le x\|\Delta g_i\|^2, \quad i = 1, \ldots, s. \tag{9}$$

Here $\Delta f_1, \ldots, \Delta f_s$ are regarded as variables in $\mathbb{C}^n$, Δy_1 and Δg_i are defined by (2), and Δy_0 is considered as a parameter (without loss of generality we have put $h = 1$).

A classical approach for solving the optimization problem (9) is to introduce Lagrange multipliers $d_1, \ldots, d_s$ and to consider the Lagrangian

$$\begin{aligned}\mathcal{L}(\Delta f_1, \ldots, \Delta f_s) &= \frac{1}{2}\|\Delta y_1\|^2 - \sum_{i=1}^{s} d_i\Big(\Re\langle \Delta f_i, \Delta g_i\rangle - x\|\Delta g_i\|^2\Big) \\ &= -\frac{1}{2}\left(\overline{\Delta y_0}, \overline{\Delta f}\right)^T \left(\begin{pmatrix} \alpha & u^T \\ u & W\end{pmatrix} \otimes I\right)\begin{pmatrix}\Delta y_0 \\ \Delta f\end{pmatrix}.\end{aligned} \tag{10}$$

Here we have used the notation $\Delta f = (\Delta f_1, \ldots, \Delta f_s)^T$, $D = \mathrm{diag}(d_1, \ldots, d_s)$, and

$$\alpha = -1 - 2x\mathbb{1}^T D\mathbb{1}, \qquad u = D\mathbb{1} - b - 2xA^T D\mathbb{1}, \qquad W = DA + A^T D - bb^T - 2xA^T DA.$$

Necessary Condition. Assuming that the Lagrange multipliers exist (see [1], Chap. 3), the derivatives of the Lagrangian $\mathcal{L}(\Delta f_1, \ldots, \Delta f_s)$ with respect to Δf_i have to vanish at the solution point, i.e.,

$$(u \otimes I)\Delta y_0 + (W \otimes I)\Delta f = 0. \tag{11}$$

Moreover, the Lagrange multipliers have to be non-negative and it holds $d_i(\Re\langle \Delta f_i, \Delta g_i\rangle - x\|\Delta g_i\|^2) = 0$ for all i.

Sufficient Condition. Subtracting $k^2\|\Delta y_0\|^2/2$ from both sides of Eq. (10), we get the following result of [3]: if the matrix

$$\begin{pmatrix} \alpha + k^2 & u^T \\ u & W\end{pmatrix} \quad \text{is positive semi-definite for some } \ d_1 \ge 0, \ldots, d_s \ge 0, \tag{12}$$

then the inequalities of (9) imply $\|\Delta y_1\| \le k\|\Delta y_0\|$. In this way we are able to get upper bounds of $\varphi(x)$.

4. Algorithmic Verification

The function $\varphi_K(x)$ of (8) is, for irreducible Runge-Kutta methods, a lower bound of the error growth function (7). For nonconfluent methods (i.e., all c_i are distinct), this is seen by considering problems of the form $y' = \lambda(t)y$, and for confluent methods it follows by using the techniques of Hundsdorfer & Spijker [5] (see also [4], Sect. IV.12).

In order to verify whether $\varphi_K(x)$ is also an upper bound of $\varphi(x)$, we can proceed as follows:

1. compute $z_1^0, \ldots, z_s^0$ with $\Re z_j^0 = x$ such that $|K(z_1^0, \ldots, z_s^0)| \geq |K(z_1, \ldots, z_s)|$ for all z_j with $\Re z_j \leq x$ (observe that some of the z_j^0 may be infinite);
2. with $Z_0 := \mathrm{diag}(z_1^0, \ldots, z_s^0)$ we put $F = Z_0(I - AZ_0)^{-1}1\!\!1$ and compute $d_1, \ldots, d_s$ from the relation $u + WF = 0$ (it turns out that $d_1, \ldots, d_s$ exist uniquely, are real and positive; see Lemma ?? below);
3. with $k := \varphi_K(x)$ we check the sufficient condition (12). Since, for this choice of K, the matrix in (12) is singular, it is sufficient to check that W is positive semi-definite.

5. Technical Details

In this part we shall prove some results that are useful for the application of the above algorithm. Let $z_1^0, \ldots, z_s^0$ with $\Re z_j^0 = x$ be such that $|K(z_1^0, \ldots, z_s^0)| \geq |K(z_1, \ldots, z_s)|$ for all z_j with $\Re z_j \leq x$, and assume for the moment that all z_j^0 are finite. From the formula

$$K(z_1, \ldots, z_s) = \frac{\det(I - (A - 1\!\!1 b^T)Z)}{\det(I - AZ)} \qquad \text{with} \quad Z = \mathrm{diag}(z_1, \ldots, z_s) \tag{13}$$

(see [4], page 197) it then follows that $(I - AZ)$ is invertible in a neighbourhood of $Z_0 = \mathrm{diag}(z_1^0, \ldots, z_s^0)$, provided that the fraction in (13) cannot be reduced.

Lemma 1. *The derivative of $K(Z) = K(z_1, \ldots, z_s)$ with respect to the jth argument satisfies (with $e_j^T = (0, \ldots, 0, 1, 0, \ldots, 0)$)*

$$\partial_j K(Z) = b^T(I - ZA)^{-1} e_j e_j^T (I - AZ)^{-1} 1\!\!1. \tag{14}$$

Furthermore, $\partial_j K(Z_0) \neq 0$ (hence also $b^T(I - ZA)^{-1}e_j \neq 0$ and $e_j^T(I - AZ)^{-1}1\!\!1 \neq 0$) and

$$0 < \partial_j K(Z_0)/K(Z_0) < \infty.$$

Proof. Differentiating $K(Z) = 1 + b^T Z(I - AZ)^{-1}1\!\!1$ with respect to z_j yields (14). Observe that the inveribility of $(I - ZA)$ follows from that of $(I - AZ)$, because $(I - ZA)(I + Z(I - AZ)^{-1}A) = \ldots = I$.

6. References

1 Avriel, M.: Nonlinear Programming: Analysis and Methods, Prentice-Hall, 1976.
2 Burrage, K., Butcher, J.C.: Stability critera for implicit Runge-Kutta methods, SIAM J. Numer. Anal. **16** (1979), 46–57.
3 Burrage, K., Butcher, J.C.: Non-linear stability of a general class of differential equation methods, BIT **20** (1980), 185–203.
4 Hairer, E., Wanner, G.: Solving Ordinary Differential Equations II, Stiff and Differential-Algebraic Problems, Springer-Verlag, 1991.
5 Hundsdorfer, W.H., Spijker, M.N.: A note on B-stability of Runge-Kutta methods, Numer. Math. **36** (1981), 319–331.

Addresses: Hairer Ernst, Wanner Gerhard, Université de Genève, Section de Mathématiques, 2-4 rue du Lièvre, Case postale 240, CH-1211 Genève 24, Switzerland.