Aplikace matematiky

Karl-Heinz Bachmann

# Fehlerabschätzung bei numerischer Lösung von Anfangswertproblemen gewöhnlicher Differentialgleichungssysteme

*Aplikace matematiky*, Vol. 10 (1965), No. 3, 200–205

Persistent URL: http://dml.cz/dmlcz/102949

# FEHLERABSCHÄTZUNG BEI NUMERISCHER LÖSUNG VON ANFANGSWERTPROBLEMEN GEWÖHNLICHER DIFFERENTIALGLEICHUNGSSYSTEME

K.-H. BACHMANN

(zum Thema a)

Ausgangspunkt der Fehlerabschätzung ist das Picard-Lindelöffsche Iterationsverfahren und eine darauf aufbauende von TOLLMIEN 1938 [1] veröffentlichte Fehlerabschätzung, die hier in einer etwas abgeänderten Form benutzt wird. Gegeben sei das Anfangswertproblem

$$\mathbf{y}' = f(x, \mathbf{y}) \quad \text{mit} \quad \mathbf{y}(a) = \mathbf{y}_0 \,. \tag{1}$$

Hier und im folgenden seien $n$-dimensionale Vektoren durch halbfette Kursiva gekennzeichnet. Für Abschätzungszwecke werde irgendeine der gebräuchlichen Normen (z.B. Summe der Absolutbeträge der Komponenten) benutzt und durch $\|\mathbf{y}\|$ bezeichnet. Durch ein numerisches Verfahren $V$ werden Näherungswerte $\mathbf{y}_k$ für die exakte Lösung $\mathbf{u}(x)$ an den Stellen $x_k \in [a, b]$ (mit $x_0 = a$) erhalten, wobei $x_{k+1} > x_k$. Eine stetige und differenzierbare Funktion $\mathbf{v}_1(x)$ ist eine Näherungsfunktion, wenn $\mathbf{v}_1(x_k) = \mathbf{y}_k$; $\mathbf{v}_1(x)$ kann z.B. ein Interpolationspolynom sein. Durch Einsetzen von $\mathbf{v}_1(x)$ in die Ausgangsgleichung (1) erhält man den Defekt der Näherungslösung

$$\mathbf{d}_1(x) = \mathbf{f}(x, \mathbf{v}_1(x)) - \mathbf{v}_1'(x) \,. \tag{2}$$

Das Picard-Lindelöffsche Iterationsverfahren ergibt nun eine zweite Näherung durch

$$\mathbf{v}_2(x) = \mathbf{y}_0 + \int_a^x \mathbf{f}(x, \mathbf{v}_1(t))\, \mathrm{d}t = \mathbf{v}_1(x) + \int_a^x \mathbf{d}_1(t)\, \mathrm{d}t \,. \tag{3}$$

Man vergleiche hierzu auch [2].

Es werde wie üblich die Gültigkeit einer Lipschitzbedingung in einer genügend großen Umgebung von $\mathbf{v}_1(x)$ vorausgesetzt, die alle durch das Iterationsverfahren erhaltenen Näherungen und $\mathbf{u}(x)$ enthält:

$$\|\mathbf{f}(x, \mathbf{y}(x)) - \mathbf{f}(x, \mathbf{z}(x))\| \leqq L\|\mathbf{y}(x) - \mathbf{z}(x)\| \quad \text{für} \quad x \in [a, b] \,. \tag{4}$$

Dabei liegen $\mathbf{y}(x)$ und $\mathbf{z}(x)$ in der erwähnten Umgebung von $\mathbf{v}_1(x)$. Unter der Voraussetzung

$$(5) \qquad \|\mathbf{v}_1(x) - \mathbf{v}_2(x)\| \leqq \varepsilon \quad \text{für} \quad x \in [a, b]$$

gilt dann die Abschätzung

$$(6) \qquad \|\mathbf{u}(x) - \mathbf{v}_2(x)\| \leqq \varepsilon(e^{L|x-a|} - 1) \quad \text{für} \quad x \in [a, b] .$$

Sei weiter $\mathbf{u}_k(x)$ die exakte Lösung des Anfangswertproblems

$$(7) \qquad \mathbf{y}' = \mathbf{f}(x, \mathbf{y}) \quad \text{mit} \quad \mathbf{y}(x_k) = \mathbf{y}_k ,$$

so gilt die Abschätzung

$$(8) \qquad \|\mathbf{u}(x) - \mathbf{u}_k(x)\| \leqq \|\mathbf{u}(x_k) - \mathbf{u}_k(x_k)\| \cdot e^{L|x-x_k|} .$$

Hiermit läßt sich eine rekursive Fehlerabschätzung aufbauen: Sei

$$(9) \qquad \|\mathbf{u}(x_k) - \mathbf{v}_2(x_k)\| \leqq \delta_k$$

und

$$(10) \qquad \|\mathbf{v}_1(x) - \mathbf{v}_2(x)\| \leqq \varepsilon_k \quad \text{für} \quad x \in [x_k, x_{k+1}] .$$

Dann ist

$$(11) \qquad \delta_{k+1} = \delta_k e^{L(x_{k+1}-x_k)} + \varepsilon_k(e^{L(x_{k+1}-x_k)} - 1)$$

eine Schranke für $\|\mathbf{u}(x_{k+1}) - \mathbf{v}_2(x_{k+1})\|$.

Ein Algorithmus zur numerischen Berechnung der $\mathbf{v}_2(x_k)$ und der $\delta_k$ wird im allgemeinen nur Näherungswerte für die $\mathbf{v}_2(x_k)$ liefern. Hierzu kann folgendes Verfahren angewandt werden: Für $\mathbf{v}_1(x)$ wird ein Interpolationspolynom eines Grades verwendet, der der Ordnung des zur Gewinnung der $\mathbf{y}_k$ verwendeten Verfahrens $V$ entspricht. Ist $V$ ein Verfahren vierter Ordnung, z.B. das übliche Runge-Kutta-Verfahren, so verwendet man für $\mathbf{v}_1(x)$ ein Polynom vierten Grades. Damit läßt sich der Defekt $\mathbf{d}_1(x)$ berechnen, für die numerische Rechnung wird man ihn auch nur an diskreten Stellen berechnen. Damit kann durch eine Quadraturformel das benötigte Integral über den Defekt und dadurch $\mathbf{v}_2(x)$ genähert bestimmt werden. Hierzu wird der Defekt ebenfalls durch ein Interpolationspolynom angenähert und integriert, die erhaltene Näherung für $\mathbf{v}_2(x)$ sei $\mathbf{V}_2(x)$. $\|\mathbf{V}_2(x) - \mathbf{v}_2(x)\|$ läßt sich mittels einer Restgliedabschätzung nach oben abschätzen. Wird noch $\|\mathbf{V}_2(x) - \mathbf{v}_1(x)\|$ abgeschätzt, kann (11) angewandt werden, um den Fehler der tatsächlich verwendeten Näherung $\mathbf{V}_2(x)$ zu bestimmen.

Diese Art des Vorgehens kann also kurz dadurch charakterisiert werden, daß eine genäherte Berechnung des Fehlers einer Ausgangsnäherung $\mathbf{v}_1$ erfolgt, und der hierbei auftretende Fehler abgeschätzt wird. Insofern weicht das Verfahren von den reinen

Fehlerabschätzungen ab, die den Fehler der Ausgangsnäherung $\mathbf{v}_1$ selbst abschätzen. Hierdurch verspricht das angewandte Verfahren genauere Resultate ohne wesentlich erhöhten Rechenaufwand. Defektabschätzungen sind in der Literatur bereits mehrfach gegeben worden, eine auch für den Einsatz automatischer Rechenanlagen geeignete Form hat J. SCHRÖDER [3] für die Lösung einer gewöhnlichen Differentialgleichung beschrieben. Hier werden sogleich Systeme von Differentialgleichungen behandelt, statt des Betrages des Defektes ist das Integral über den Defekt die maßgebliche Größe. Im folgenden sei der eingeschlagene Weg kurz erläutert, wenn $V$ ein Verfahren höchstens vierter Ordnung ist.

Die Näherungsfunktion $\mathbf{v}_1(x)$ ist stückweise aus Polynomen zusammengesetzt, die durch Interpolation über die Intervalle $[x_0, x_2]$, $[x_2, x_4]$ usw. bestimmt werden. $\mathbf{v}_1(x)$ wird im Intervall $[x_{k-1}, x_{k+1}]$ ($k$ ungerade) als Polynom vierten Grades so bestimmt, daß

$$(12) \qquad \begin{aligned} \mathbf{v}_1(x_{k-1}) &= \mathbf{y}_{k-1}\,, & \mathbf{v}_1'(x_{k-1}) &= \mathbf{f}(x, \mathbf{y}_{k-1})\,, \\ \mathbf{v}_1(x_k) &= \mathbf{y}_k\,, & \mathbf{v}_1'(x_k) &= \mathbf{f}(x, \mathbf{y}_k)\,, \\ \mathbf{v}_1(x_{k+1}) &= \mathbf{y}_{k+1}\,. \end{aligned}$$

Der Defekt verschwindet somit an den Stellen $x_{k-1}$ und $x_k$, für $x_{k+2}$ gilt, wenn noch $x_{k+1} - x_k = x_k - x_{k-1} = h$ vorausgesetzt wird,

$$(13) \quad h\,\mathbf{d}_1(x_{k+1}) = h\,\mathbf{f}(x, \mathbf{y}_{k+1}) + 4h\,\mathbf{f}(x, \mathbf{y}_k) + h\,\mathbf{f}(x, \mathbf{y}_{k-1}) - 3(\mathbf{y}_{k+1} - \mathbf{y}_{k-1})\,.$$

Taylorsche Entwicklung von $\mathbf{d}_1(x)$ ergibt bei der Voraussetzung genügender Differenzierbarkeitseigenschaften, daß $\mathbf{d}_1(x)$ zweckmäßig durch ein spezielles Polynom vierten Grades approximiert wird:

$$(14) \qquad h\,\mathbf{d}_1(x) = \mathbf{P}(z) + \mathbf{R}(z) = \boldsymbol{\alpha}_0 P_0(z) + \boldsymbol{\alpha}_1 P_1(z) + \boldsymbol{\alpha}_2 P_2(z) + \mathbf{R}(z)$$

mit

$$z = \frac{x - x_k}{h}$$

und

$$\begin{aligned} P_0(z) &= \quad 5z^4 + 4z^3 - 3z^2 - 2z\,, \\ P_1(z) &= -\,4z^3 + 4z\,, \\ P_2(z) &= -\,z^3 - \tfrac{3}{2}z^2 - \tfrac{1}{2}z\,. \end{aligned}$$

Das Restglied $R(z)$ wird in der Langrangeschen Form verwendet:

$$(15) \qquad \mathbf{R}(z) = \frac{h^6}{120}(z^2 - 1)(z^2 - \tfrac{1}{4})\,z\,\frac{d^5}{dx^5}\mathbf{d}_1(\xi) \quad \text{mit} \quad \xi \in [(z_{k-1}, x_{k+1}]\,.$$

Die hierin vorkommende fünfte Ableitung des Defektes ist identisch mit der fünften Ableitung von $\boldsymbol{f}(x, \boldsymbol{v}_1(x))$, da $\boldsymbol{v}_1(x)$ ein Polynom vierten Grades ist. Es ergibt sich

$$(16) \qquad \begin{aligned} \boldsymbol{v}_2(x) &= \boldsymbol{v}_1(x) + \int_{-1}^{z} \boldsymbol{P}(t)\,\mathrm{d}t + \int_{-1}^{z} \boldsymbol{R}(t)\,\mathrm{d}t = \\ &= \boldsymbol{v}_1(x) + \boldsymbol{\alpha}_0 Q_0(z) + \boldsymbol{\alpha}_1 Q_1(z) + \boldsymbol{\alpha}_2 Q_2(z) + \int_{-1}^{z} \boldsymbol{R}(t)\,\mathrm{d}t = \\ &= \boldsymbol{V}_2(x) + \int_{-1}^{z} \boldsymbol{R}(t)\,\mathrm{d}t \end{aligned}$$

mit

$$\begin{aligned} Q_0(z) &= z^2(z+1)^2\,(z-1)\,, \\ Q_1(z) &= -\,(z+1)^2\,(z-1)^2\,, \\ Q_2(z) &= -\tfrac{1}{4}z^2(z+1)^2\,. \end{aligned}$$

Hieraus folgt

$$(17) \qquad \begin{aligned} \|\boldsymbol{v}_2(x) - \boldsymbol{v}_1(x)\| &\leqq \|\boldsymbol{\alpha}_0\| \operatorname*{Max}_{z\in[-1,1]} |Q_0(z)| + \|\boldsymbol{\alpha}_1\| \operatorname*{Max}_{z\in[-1,1]} |Q_1(z)| + \\ &+ \|\boldsymbol{\alpha}_2\| \operatorname*{Max}_{z\in[-1,1]} |Q_2(z)| + q_k = 0{,}435\,\|\boldsymbol{\alpha}_0\| + \|\boldsymbol{\alpha}_1\| + \|\boldsymbol{\alpha}_2\| + q_k = \varepsilon_k\,, \end{aligned}$$

wobei

$$(18) \qquad \left\| \int_{-1}^{z} \boldsymbol{R}(t)\,\mathrm{d}t \right\| \leqq q_k \quad \text{für} \quad z \in [-1, 1]$$

verwendet wird. Aus (16) und (18) folgt noch

$$(19) \qquad \|\boldsymbol{V}_2(x) - \boldsymbol{v}_2(x)\| \leqq q_k\,.$$

Damit läßt sich die durch (11) gegebene Fehlerabschätzung auf die hier vorliegenden Verhältnisse anwenden:

Sei $\|\boldsymbol{V}_2(x_{k-1}) - \boldsymbol{u}(x_{k-1})\| \leqq D_{k-1}$, so wird

$$(20) \qquad \|\boldsymbol{V}_2(x_k) - \boldsymbol{u}(x_k)\| \leqq D_k = D_{k-1}\,.\,e^{Lh} + \varepsilon_k(e^{Lh} - 1) + q_k\,,$$

$$(21) \qquad \|\boldsymbol{V}_2(x_{k+1}) - \boldsymbol{u}(x_{k+1})\| \leqq D_{k+1} = D_{k-1}\,.\,e^{2Lh} + \varepsilon_k(e^{2Lh} - 1) + q_k\,.$$

Zu bestimmen sind also die Parametervektoren $\boldsymbol{\alpha}_0$, $\boldsymbol{\alpha}_1$ und $\boldsymbol{\alpha}_2$ sowie die Norm $q_k$ des Quadraturfehlers und die Lipschitzkonstante $L$, um die neue Näherung $\boldsymbol{V}_2$ zu errechnen und deren Fehler abzuschätzen. $\alpha_0$, $\alpha_1$ und $\alpha_2$ lassen sich z.B. so bestimmen, daß der Defekt an den Stellen $z = -1$, $z = -\frac{1}{2}$, $z = 0$, $z = \frac{1}{2}$ und $z = 1$ berechnet wird (bei $z = -1$ und $z = 0$ verschwindet er), und das Interpolationspolynom vierten

Grades für $\boldsymbol{d}_1(x)$ durch die erhaltenen Punkte gelegt wird. Es folgt, wenn die Funktionswerte von $\boldsymbol{f}(x, \boldsymbol{v}_1(x))$ direkt statt des Defektes benutzt werden:

$$(22) \quad \boldsymbol{\alpha}_0 = \frac{2h}{15}\left(\boldsymbol{f}_{k-1} - 4\boldsymbol{f}_{k-\frac{1}{2}} + 6\boldsymbol{f}_k - 4\boldsymbol{f}_{k+\frac{1}{2}} + \boldsymbol{f}_{k+1}\right),$$

$$\boldsymbol{\alpha}_1 = -\boldsymbol{y}_{k-1} + \boldsymbol{y}_k - \tfrac{29}{180}h\boldsymbol{f}_{k-1} - \tfrac{31}{45}h\boldsymbol{f}_{k-\frac{1}{2}} - \tfrac{2}{15}h\boldsymbol{f}_k - \tfrac{1}{45}h\boldsymbol{f}_{k+\frac{1}{2}} + \frac{h}{180}\boldsymbol{f}_{k+1},$$

$$\boldsymbol{\alpha}_2 = -\boldsymbol{y}_{k-1} + \boldsymbol{y}_{k+1} - \tfrac{7}{45}h\boldsymbol{f}_{k-1} - \tfrac{32}{45}h\boldsymbol{f}_{k-\frac{1}{2}} - \tfrac{12}{45}h\boldsymbol{f}_k - \tfrac{32}{45}h\boldsymbol{f}_{k+\frac{1}{2}} - \tfrac{7}{45}h\boldsymbol{f}_{k+1}.$$

Dabei bedeuten

$$(23) \quad \begin{aligned}
\boldsymbol{f}_{k-1} &= f(\boldsymbol{x}_{k-1}, \boldsymbol{y}_{k-1}), \\
\boldsymbol{f}_{k-\frac{1}{2}} &= f(\boldsymbol{x}_k - h/2, \boldsymbol{y}_{k-\frac{1}{2}}) \quad \text{mit} \\
\boldsymbol{y}_{k-\frac{1}{2}} &= \tfrac{27}{64}\boldsymbol{y}_{k-1} + \tfrac{9}{16}\boldsymbol{y}_k + \tfrac{1}{64}\boldsymbol{y}_{k+1} + \tfrac{3}{32}h\boldsymbol{f}_{k-1} - \tfrac{3}{16}h\boldsymbol{f}_k, \\
\boldsymbol{f}_k &= f(x_k, \boldsymbol{y}_k), \\
\boldsymbol{f}_{k+\frac{1}{2}} &= f(x_k + h/2, \boldsymbol{y}_{k+\frac{1}{2}}) \quad \text{mit} \\
\boldsymbol{y}_{k+\frac{1}{2}} &= \tfrac{19}{64}\boldsymbol{y}_{k-1} + \tfrac{9}{16}\boldsymbol{y}_k + \tfrac{9}{64}\boldsymbol{y}_{k+1} + \tfrac{3}{32}h\boldsymbol{f}_{k-1} + \tfrac{9}{16}h\boldsymbol{f}_k, \\
\boldsymbol{f}_{k+1} &= f(x_k, \boldsymbol{y}_{k+1}).
\end{aligned}$$

Zur Abschätzung der Norm $q_k$ des Quadraturfehlers werden Schranken für die fünften totalen Ableitungen der einzelnen Komponenten von $\boldsymbol{f}(x, \boldsymbol{v}_1(x))$ benötigt. Numerisch kann man sich Näherungen dafür durch Differenzenquotienten verschaffen, wenn das nicht auf analytischem Wege einfach möglich ist. Entsprechendes gilt für die Lipschitzkonstante, für die man durch Normenbildung für einen Differenzenquotienten bzw. aus mehreren Differenzenquotienten eine Näherung gewinnen kann, wenn sie nicht analytisch einfach zu bestimmen ist. Aus (16) ergeben sich die verbesserten Näherungswerte
$\boldsymbol{Y}_k = \boldsymbol{V}_2(x_k)$ und $\boldsymbol{Y}_{k+1} = \boldsymbol{V}_2(x_{k+1})$ zu

$$(24) \quad \begin{aligned}
\boldsymbol{Y}_k &= \boldsymbol{y}_k - \boldsymbol{\alpha}_1, \\
\boldsymbol{Y}_{k+1} &= \boldsymbol{y}_{k+1} - \boldsymbol{\alpha}_2.
\end{aligned}$$

Es ist noch folgendes zu bemerken:

1. Der Fehler kann bei jedem Schritt parallel zur Rechnung abgeschätzt werden, Rundungsfehler lassen sich dabei berücksichtigen. Damit ist eine automatische Steuerung der Schrittweite möglich. Z.B. kann man verlangen, daß die errechnete Fehlerschranke zwischen zwei linearen Funktionen verläuft. Ist sie zu klein, wird die Schrittweite vergrößert; ist sie zu groß, wird die Schrittweite verkleinert. In diesem Sinne kann das Verfahren zur optimalen Wahl der Schrittweite ausgenutzt werden.

2. Das Verfahren beschreibt eine Korrektor-Formel, die zugehörige, Prediktor-Formel soll von vierter Ordnung sein. Die Stabilität des Verfahrens hängt von der Wahl des Prediktors ab. Benutzt man dazu z.B. das übliche Runge-Kutta-Verfahren, so bleibt das Verfahren auch mit dem Korrektor ein Einschrittverfahren, Stabilitätsbetrachtungen sind dann nicht weiter nötig. Durch ungünstige Wahl eines Prediktors auftretende Instabilitäten sollten mittels der Fehlerabschätzung erkannt werden.

3. Der Rechenaufwand ist ziemlich hoch, er hängt davon ab, wieweit durch analytische Vorbetrachtungen Informationen über die auftretende fünfte Ableitung und die Lipschitzkonstante erhalten werden können. Ist das nicht der Fall, ist man dabei auf Näherungen angewiesen, die jedoch ziemlich grob sein dürfen.

*Literatur*

[1] *W. Tollmien:* Über die Fehlerabschätzung beim Adamsschen Verfahren zur Integration gewöhnlicher Differentialgleichungen, Zeitschr. angew. Math. Mech. 18 (1938), 83—90.

[2] *L. Collatz:* Numerische Behandlung von Differentialgleichungen, Berlin, Göttingen, Heidelberg 1955, § 4.9.

[3] *J. Schröder:* Fehlerabschätzung mit Rechenanlagen bei gewöhnlichen Differentialgleichungen erster Ordnung, Numerische Mathematik 3 (1961), 39—61.

*K.-H. Bachmann*, Institut für angewandte Mathematik und Mechanik der Deutschen Akademie der Wissenschaften, Mohrenstr. 39, 108 Berlin, DDR.