23. VL Einführung in die num. Math. 1.7.09 ①

Rundungsfehleranalyse von expliziten Einschrittverfahren

Zur numerischen Lösung eines AWP $y' = f(t,y),\ y(a) = y_0$ betrachten wir das Verfahren

$$y_{k+1} = y_k + h_k \Phi(t_k, y_k, h_k), \quad k = 0,1,2,\dots,N-1, \tag{6.36}$$

mit der Verfahrensfunktion Φ. Durch Rundungsfehler arbeitet man statt (6.36) mit einem Verfahren der Form

$$\hat{y}_{k+1} = \hat{y}_k + h_k \Phi(t_k, \hat{y}_k, h_k) + \rho_k, \quad k = 0,1,\dots,N-1, \tag{6.37}$$

$$\hat{y}_0 = y_0 + e_0, \quad |\rho_k| \le \delta,\ k = 0,1,\dots,N-1, \quad |e_0| \le \varepsilon,$$

mit gewissen Zahlen $e_0, \rho_k \in \mathbb{R}$.
Für die Rundungsfehler infolge des Verfahrens (6.37) gilt der folgende

Satz 6.16

Zur Lösung des AWP $y' = f(t,y),\ y(a) = y_0$, sei durch (6.36) ein Einschrittverfahren mit der Konsistenzordnung $p \ge 1$ gegeben, wobei die Verfahrensfunktion bezügl. der 2. Variablen Lipschitz-stetig mit der Konstanten $L > 0$ ist.
Dann gelten für die durch die fehlerbehaftete Verfahrensvorschrift (6.37) gewonnenen Approximationen die Abschätzungen

$$\max_{k=0,\dots,N} |\hat{y}_k - y(t_k)| \le K\left(h_{max}^p + \frac{\delta}{h_{min}}\right) + e^{L(b-a)}\varepsilon \tag{6.38}$$

mit der Konstanten $K = \frac{\max\{C,1\}}{L}\left[e^{L(b-a)} - 1\right]$.
C ist dabei die Konstante aus der Abschätzung
$|d_k| \leq C h_k^{p+1}$ für den lokalen Diskretisierungsfehler.

Beweis
Wir setzen

$$e_k = \hat{y}_k - y(t_k), \quad k = 0, 1, \dots, N,$$

und damit gilt für $k = 0, 1, \dots, N-1$

$$y(t_{k+1}) = y(t_k) + h_k \phi(t_k, y(t_k), h_k) + d_{k+1}$$

$$\hat{y}_{k+1} = \hat{y}_k + h_k \phi(t_k, \hat{y}_k, h_k) + \rho_k$$

bzw.

$$e_{k+1} = e_k + h_k\left[\phi(t_k, \hat{y}_k, h_k) - \phi(t_k, y(t_k), h_k)\right] + \rho_k - d_{k+1}$$

und folglich

$$|e_{k+1}| \leq |e_k| + h_k\left|\phi(t_k, \hat{y}_k, h_k) - \phi(t_k, y(t_k), h_k)\right| + |\rho_k| + |d_{k+1}|$$

$$\leq (1 + h_k L)|e_k| + h_k\left(C h_{max}^p + \frac{\delta}{h_{min}}\right).$$

Aus dem Hilfssatz 6.9
folgt mit $|e_0| \leq \varepsilon$ die Behauptung des Satzes. $\square$

Schrittweitensteuerung bei Einschrittverfahren

Bei der Konvergenzuntersuchung von Einschrittverfahren werden die lokalen Diskretisierungsfehler im gewissen Sinne summiert und deshalb erscheint eine Beschränkung des Absolutbetrags von d_k

durch die Wahl geeigneter Schrittweiten h_k sinnvoll. Man spricht hier von <u>Schrittweitensteuerung</u>.

Das Prinzip soll am Beispiel des Heun-Verfahrens

$$k_1 = f(t_k, y_k),$$
$$k_2 = f(t_k + h, y_k + h k_1),$$
$$y_{k+1} = y_k + \frac{h}{2}[k_1 + k_2]$$

erläutert werden. Als lokaler Diskretisierungsfehler ergibt sich

$$d_{k+1}^{(H)} = y(t_{k+1}) - y(t_k) - \frac{h}{2}[\bar{k}_1 + \bar{k}_2] \qquad (6.39)$$

mit $\bar{k}_1 = f(t_k, y(t_k))$, $\bar{k}_2 = f(t_k + h, y(t_k) + h\bar{k}_1)$.

Nun sucht man ein Verfahren höherer Ordnung, also mindestens dritter Ordnung (Heun-Verfahren ist Verf. zweiter Ordnung), dessen Steigungen k_1, k_2 mit den Steigungen des Heun-Verfahrens übereinstimmen.

Die Forderung der Gleichheit von k_1 und k_2 bedeutet $\alpha_2 = \beta_{21} = 1$. Die weiteren Parameter ergeben sich aus (6.32) bei der Wahl von $\alpha_3 = \frac{1}{2}$ zu

$$\gamma_3 = \frac{2}{3}, \quad \gamma_2 = \frac{1}{6}, \quad \gamma_1 = \frac{1}{6}, \quad \beta_{32} = \frac{1}{4}, \quad \beta_{31} = \alpha_3 - \beta_{32} = \frac{1}{4},$$

so dass sich das Runge-Kutta-Verfahren 3. Ordnung

$$k_1 = f(t_k, y_k),$$
$$k_2 = f(t_k + h, y_k + h k_1), \quad k_3 = f(t_k + \frac{h}{2}, y_k + \frac{h}{4}(k_1 + k_2)),$$
$$y_{k+1} = y_k + \frac{h}{6}[k_1 + k_2 + 4k_3] \qquad (6.40)$$

ergibt. Für den lokalen Diskretisierungsfehler des Verfahrens (6.39) ergibt sich

$$d_{k+1}^{(RK)} = y(t_{k+1}) - y(t_k) - \frac{h}{6}\left[\bar{k}_1 + \bar{k}_2 + 4\bar{k}_3\right] \qquad (6.41)$$

mit $\bar{k}_3 = f(t_k + \frac{h}{2}, y(t_k) + \frac{h}{4}(\bar{k}_1 + \bar{k}_2))$. Mit (6.39) und (6.41) ergibt sich die Darstellung des lokalen Diskretisierungsfehlers des Heun-Verfahrens

$$d_{k+1}^{(H)} = \frac{h}{6}\left[\bar{k}_1 + \bar{k}_2 + 4\bar{k}_3\right] - \frac{h}{2}\left[\bar{k}_1 + \bar{k}_2\right] + d_{k+1}^{(RK)}$$

Ersetzt man nun die unbekannten Werte von $\bar{k}_j$ durch die Näherungen k_j und berücksichtigt $d_{k+1}^{(RK)} = O(h^4)$, so erhält man

$$d_{k+1}^{(H)} = \frac{h}{6}\left[k_1 + k_2 + 4k_3\right] - \frac{h}{2}\left[k_1 + k_2\right] + O(h^4) =$$

$$\frac{h}{3}\left[2k_3 - k_1 - k_2\right] + O(h^4)$$

und damit kann der lokale Diskretisierungsfehler des Heun-Verfahren mit einer zusätzlichen Neigungsberechnung von k_3 durch den Ausdruck $\frac{h}{3}\left[2k_3 - k_1 - k_2\right]$ recht gut geschätzt werden. Aufgrund der Kontrolle des Betrags dieses Ausdruckes kann man eine vorgegebene Schranke $\varepsilon_{tol} > 0$ durch entsprechende Wahl von $h = h_k = t_{k+1} - t_k$

$$h_k < \frac{3\varepsilon_{tol}}{|2k_3 - k_1 - k_2|} \iff \frac{h_k}{3}\left[2k_3 - k_1 - k_2\right] < \varepsilon_{tol}$$

unterschreiten. D.h. man kann die aktuelle

Schrittweite voll. vergrößern oder muss sie verkleinern.

Die eben beschriebene Methode der Schrittweitensteuerung bezeichnet man auch als Einbettung des Heun-Verfahrens 2. Ordnung in das Runge-Kutta-Verfahren 3. Ordnung (6.40).

Mehrschrittverfahren

Die Klasse der Mehrschrittverfahren zur Lösung von AWP ist dadurch gekennzeichnet, dass man zur Berechnung des Näherungswertes y_{k+1} nicht nur den Wert y_k verwendet, sondern auch weiter zurückliegende Werte z.B. $y_{k-1}, y_{k-2}, y_{k-3}$.

Als Ausgangspunkt zur Konstruktion von Mehrschrittverfahren betrachten wir die zum AWP äquivalente Integralbeziehung

$$y(t_{k+1}) = y(t_k) + \int_{t_k}^{t_{k+1}} f(t, y(t))\,dt \quad . \qquad (6.42)$$

Kennt man die Werte $f_k = f(t_k, y_k)$, $f_{k-1} = f(t_{k-1}, y_{k-1})$, ..., $f_{k-3} = f(t_{k-3}, y_{k-3})$, dann kann man das Integral auf der rechten Seite von (6.42) i.d.R. besser approximieren als bei den Einschrittverfahren unter ausschließlicher Nutzung des Wertes $f_k = f(t_k, y_k)$. Für das Interpolationspolynom durch die Stützpunkte (t_j, f_j), $j = k-3, \ldots, k$

ergibt sich

$$p_3(t) = \sum_{j=0}^{3} f_{k-j}\, L_{k-j}(t)$$

mit den Lagrange'schen Basispolynomen

$$L_j(t) = \prod_{\substack{i=k-3 \\ i\neq j}}^{k} \frac{t-t_i}{t_j-t_i}, \quad j=k-3,\dots,k.$$

Die Idee der Mehrschrittverfahren besteht nun in der Nutzung von $p_3(t)$ als Approximation von $f(t,y(t))$ im Integral von (6.42), so dass man auf der Grundlage von (6.42) das Mehrschrittverfahren (4-Schritt-Vrf.)

$$y_{k+1} = y_k + \int_{t_k}^{t_{k+1}} \sum_{j=0}^{3} f_{k-j}\, L_{k-j}(t)\, dt$$

$$= y_k + \sum_{j=0}^{3} f_{k-j} \int_{t_k}^{t_{k+1}} L_{k-j}(t)\, dt$$

erhält. Im Fall äquidistanter Stützstellen $h=t_{k+1}-t_k$ erhält man für den 2. Integralsummanden $(j=1)$

$$I_1 = \int_{t_k}^{t_{k+1}} L_{k-1}(t)\,dt = \int_{t_k}^{t_{k+1}} \frac{(t-t_{k-3})(t-t_{k-2})(t-t_k)}{(t_{k-1}-t_{k-3})(t_{k-1}-t_{k-2})(t_{k-1}-t_k)}\,dt$$

und nach der Substitution $\xi = \frac{t-t_k}{h}$

$$I_1 = h\int_0^1 \frac{(\xi+3)(\xi+2)\xi}{2\cdot 1\cdot(-1)}\,d\xi = -\frac{h}{2}\int_0^1 (\xi^3+5\xi^2+6\xi)\,d\xi = -\frac{59}{24}h$$

Für die restlichen Integrale erhält man

$$I_0 = \frac{55}{24}h, \quad I_2 = \frac{37}{24}h, \quad I_3 = -\frac{9}{24}h,$$

so dass sich mit

$$y_{k+1} = y_k + \frac{h}{24}\left[55 f_k - 59 f_{k-1} + 37 f_{k-2} - 9 f_{k-3}\right] \qquad (6.43)$$

die Methode von Adams-Bashforth (kurz AB-Verfahren) ergibt.
Durch Taylor-Reihenentwicklung erhält man bei entsprechender Glattheit von f bzw. $y(t)$ den lokalen Diskretisierungsfehler

$$d_{k+1} = \frac{251}{720}h^5 y^{(5)} + O(h^6), \qquad (6.44)$$

d.h. das Verfahren (6.43) ist von 4. Ordnung.

Definition 6.17
Bei Verwendung von m Stützwerten $(t_k, f_k), \ldots, (t_{k-m+1}, f_{k-m+1})$ zur Berechnung eines Interpolationspolynoms p_{m-1} zur Approximation von f zwecks näherungsweiser Berechnung des Integrals in (6.42) spricht man von einem linearen m-Schrittverfahren.
Ein m-Schrittverfahren hat die Fehlerordnung p, falls für seinen lokalen Diskretisierungsfehler d_k die Abschätzung

$$\max_{m \le k \le N} |d_k| \le K = O(h^{p+1})$$

gilt.

Allgemein kann man für Adams-Bashforth-Verfahren (m-Schritt)

$$y_{k+1} = y_k + \sum_{j=0}^{m-1} f_{k-j} \int_{t_k}^{t_{k+1}} L_{k-j}(t)\,dt$$

bei ausreichender Glattheit der Lösung $y(t)$ zeigen, dass sie die Fehlerordnung m besitzen. Durch Auswertung der entsprechenden Integrale erhält man für $m = 2, 3, 4$ die folgenden 3-, 4- und 5-Schritt Adams-Bashforth-Verfahren

$$y_{k+1} = y_k + \frac{h}{12}\left[23 f_k - 16 f_{k-1} + 5 f_{k-2}\right], \quad (6.45)$$

$$y_{k+1} = y_k + \frac{h}{24}\left[55 f_k - 59 f_{k-1} + 37 f_{k-2} - 9 f_{k-3}\right],$$

$$y_{k+1} = y_k + \frac{h}{720}\left[1901 f_k - 2774 f_{k-1} + 2616 f_{k-2} - 1274 f_{k-3} + 251 f_{k-4}\right].$$

Die Formeln der Mehrschrittverfahren "funktionieren" erst ab dem Index $k = m$, d.h. bei einem 3-Schrittverfahren braucht man die Werte y_0, y_1, y_2 um y_3 mit der Formel (6.45) berechnen zu können.

Die Startwerte y_1, y_2 werden meist mit einem Runge-Kutta-Verfahren berechnet.

Es ist offensichtlich möglich, die Qualität der Lösungsverfahren für das AWP $y' = f(t,y)$, $y(a) = y_0$.

zu erhöhen, indem man das Integral

$$\int_{t_k}^{t_{k+1}} f(t,y(t))\,dt$$

aus der Beziehung (6.82) genauer berechnet. Das kann man durch Hinzunahme weiterer Stützpunkte zur Polynominterpolation tun. Nimmt man den (noch unbekannten) Wert $f_{k+1}=f(t_{k+1},y_{k+1})$ zu den Werten $f_k, f_{k-1}, f_{k-2}, f_{k-3}$ hinzu, dann erhält man mit

$$P_4(t)=\sum_{j=-1}^{3} f_{k-j}\,L_{k-j}(t)$$

in Analogie zur Herleitung der Adams-Bashforth-Verfahren mit

$$y_{k+1}=y_k+\int_{t_k}^{t_{k+1}}\sum_{j=-1}^{3} f_{k-j}\,L_{k-j}(t)\,dt=y_k+\sum_{j=-1}^{3} f_{k-j}\int_{t_k}^{t_{k+1}}L_{k-j}(t)\,dt$$

bzw. nach Auswertung der Integrale

$$y_{k+1}=y_k+\frac{h}{720}\left[251 f_{k+1}+646 f_k-264 f_{k-1}+106 f_{k-2}-19 f_{k-3}\right]. \qquad (6.86)$$

Das Verfahren (6.86) ist eine implizite 4-Schritt-Methode und heißt Methode von Adams-Moulton (kurz AM-Verfahren). Das 3-Schritt Adams-Moulton-Verfahren hat die Form

$$y_{k+1}=y_k+\frac{h}{24}\left[9 f(t_{k+1},y_{k+1})+19 f_k-5 f_{k-1}+f_{k-2}\right] \qquad (6.87)$$

Zur Bestimmung der Lösung der Gleichung (6.47) kann man z.B. eine Fixpunktiteration der Art

$$y_{k+1}^{(s+1)} = y_k + \frac{h}{24}\left[9 f(t_{k+1}, y_{k+1}^{(s)}) + 19 f_k - 5 f_{k-1} + f_{k-2}\right]$$

durchführen (Startwert z.B. $y_{k+1}^{(0)} = y_k$).

Bestimmt man den Startwert $y_{k+1}^{(0)}$ als Resultat eines 3-Schritt Adams-Bashforth-Verfahrens und führt nur eine Fixpunktiteration durch, dann erhält man das sogenannte Prädiktor-Korrektor-Verfahren

$$y_{k+1}^{(p)} = y_k + \frac{h}{12}\left[23 f_k - 16 f_{k-1} + 5 f_{k-2}\right] \qquad (6.48)$$

$$y_{k+1} = y_k + \frac{h}{24}\left[9 f(t_{k+1}, y_{k+1}^{(p)}) + 19 f_k - 5 f_{k-1} + f_{k-2}\right].$$

Diese Kombination von Adams-Bashforth- und Adams-Moulton-Verfahren bezeichnet man als Adams-Bashforth-Moulton-Verfahren (kurz ABM-Verfahren). Das ABM-Verfahren (6.48) hat ebenso wie das AM-Verfahren (6.47) den Diskretisierungsfehler $d_{k+1} = O(h^5)$ und damit die Fehlerordnung 4.

Generell kann man zeigen, dass m-Schritt-Verfahren vom AM- oder ABM-Typ jeweils die Fehlerordnung $p = m+1$ besitzen.