

**Numerische Behandlung
von gewöhnlichen Differentialgleichungen
Eine Einführung**

K. Taubert
SoSe07

Universität Hamburg

Numerische Behandlung von gewöhnlichen Differentialgleichungen

Eine Einführung

Universität Hamburg
SoSe07

K. Taubert

Numerik gewöhnlicher Anfangswertaufgaben

1 EINFÜHRENDE BEMERKUNGEN

Durch einfach nachvollziehbare Bedingungen kann die Frage nach der Existenz und Eindeutigkeit von Lösungen von Anfangswertaufgaben

$$y' = f(x, y), \quad y(x_0) = y_0,$$

geklärt werden. Auch Aussagen über die Veränderung der Lösungen bei Veränderungen von f und den Anfangsbedingungen können so erbracht werden.

Die explizite Lösung einer konkreten Anfangswertaufgabe kann unmöglich oder mit einem nichtvertretbaren Aufwand verbunden sein. Die Alternative und manchmal auch einzige Möglichkeit ist dann die Bestimmung einer numerischen Approximation.

Verfahren zur numerischen Approximation von Lösungen von Anfangswertaufgaben, insbesondere Differenzenverfahren, finden deshalb eine breite Anwendung.

Vernünftige Differenzenverfahren sind zwar stets konvergent, aber nicht jedes konvergente Differenzenverfahren ist gleich gut für jede Anfangswertaufgabe. Die numerische Integration einer speziellen Anfangswertaufgabe kann besondere (Stabilitäts-) Eigenschaften der Verfahren erfordern.

Beispiele geben einen Einblick in das Konstruktionsprinzip von Differenzenverfahren und in grundsätzliche Aspekte für die Auswahl von Verfahren.

§ 1.1 Gewöhnliche Anfangswertaufgaben

Gegeben sei die Anfangswertaufgabe

$$y' = f(x, y), \quad y(x_0) = y_0, \quad (*)$$

mit einer stetigen Funktion $f: \mathbf{D} \subset \mathbf{R}^* \times \mathbf{R}^n \rightarrow \mathbf{R}^n$ auf einer offenen Teilmenge \mathbf{D} mit $(x_0, y_0) \in \mathbf{D}$.

Eine stetig differenzierbare Funktion $y(\cdot)$ auf dem Intervall I (offen, abgeschlossen oder halboffen) heißt Lösung von (*) auf I , wenn

- 1) $y(x_0) = y_0$ mit $x_0 \in I$,
- 2) $(x, y(x)) \in D$ für alle $x \in I$ und
- 3) $y'(x) = f(x, y(x))$ für alle $x \in I$.

Der Stetigkeit von f wegen ist die Funktion $y(\cdot)$ auf dem Intervall I genau dann eine Lösung der Anfangswertaufgabe, wenn

$$y(x) = y_0 + \int_{x_0}^x f(s, y(s)) ds \quad \text{für alle } x \in I.$$

Für einige Anfangswertaufgaben kann die Lösung explizit angegeben werden:
Die allgemeine Lösung eines Systems von linearen Differentialgleichungen

$$y' = Ay$$

mit einer reellen $n \times n$ Matrix kann dargestellt werden als Linearkombination

$$y(x) = \sum_{i=1}^n a_i Y_i(x)$$

mit einem Hauptsystem von Lösungen $[Y_1(x), Y_2(x), \dots, Y_n(x)]$. Die Funktionen $Y_i(x)$ ergeben sich aus den Eigenwerten und zugehörigen Hauptvektoren der Matrix A . Um die allgemeine Lösung einer Anfangsbedingung anzupassen, muss dann noch ein zugehöriges Gleichungssystem gelöst werden.

Kleinere Aufgaben wird man so oder besser mit Hilfe der Laplace-Transformation lösen können.

In einem Buch von E. Kamke¹ werden viele Klassen von nichtlinearen Differentialgleichungen angegeben die mit passenden Tricks explizit gelöst werden können. Ist eine konkrete Aufgabe dort nicht aufgeführt spricht vieles dafür, dass eine explizite Lösung nicht angegeben werden kann.

Die allgemeine Frage nach der Existenz einer Lösung von (*) ist schnell beantwortet: Es gibt stets eine Umgebung von (x_0, y_0) , in der die Aufgabe eine Lösung besitzt. Der Beweis kann z.B. mit dem Schauderschen Fixpunktsatz oder mit dem Eulerschen Polygonzugverfahren erbracht werden.

Die Frage nach dem größtmöglichen Definitionsbereich einer Lösung einer Anfangswertaufgabe und der Eindeutigkeit kann im konkreten Fall schwierig sein. Einfache Beispiele zeigen:

Selbst wenn f für alle (x, y) aus $\mathbf{R} \times \mathbf{R}^n$ definiert ist, muss es nicht für jeden Anfangswert eine Lösung geben, die auf ganz \mathbf{R} definiert ist. Auch muss es nicht notwendig nur eine Lösung für jeden Anfangswert geben. Beispiele dafür sind:

$$\begin{aligned} y' &= y^2, & y(0) &= 1, \\ y' &= |y|^{1/2}, & y(0) &= 0. \end{aligned}$$

Die folgenden Sätze sind Beispiele für Aussagen zur globalen Existenz, zur Eindeutigkeit und zum Verhalten von Lösungen bei Veränderungen von f und (x_0, y_0) .

¹ Lösungsmethoden und Lösungen. 10 ed. Teubner Stuttgart 1983

Als robustes Mittel zur Bestimmung eines Existenzintervalls von Lösungen von Anfangswertaufgaben haben sich Vergleichssätze bewährt. Ein Beispiel dafür ist:

Satz 1.1

Es sei $\omega(x,u)$ eine stetige und reellwertige Funktion auf

$$D = \{ (x, u) / x \in [x_0, x_0 + a), u \geq 0, a > 0 \}.$$

Die Anfangswertaufgabe $u' = \omega(x, u), u(x_0) = u_0$ habe eine eindeutige nichtnegative Lösung für alle $x \in [x_0, x_0 + a)$.

Es sei $f : [x_0, x_0 + a) \times \mathbf{R}^n \rightarrow \mathbf{R}^n$ eine stetige Funktion und

$$\| f(x, y) \| \leq \omega(x, \| y \|)$$

für alle $x \in [x_0, x_0 + a)$ und alle $y \in \mathbf{R}^n$ mit irgendeiner Vektornorm $\| \cdot \|$.

Dann existieren die Lösungen von

$$y' = f(x, y), \quad \| y(x_0) \| \leq u_0,$$

auf dem gesamten Intervall $[x_0, x_0 + a)$.

Zum Nachweis der Eindeutigkeit der Lösungen von Anfangswertaufgaben haben sich „Wachstumsbedingungen“ für f bewährt. Ein Beispiel dafür ist:

Satz 1.2

Es sei $Q = \{ (x, y) / |x - x_0| \leq a \text{ und } \|y - y_0\| \leq b \}$, $a, b > 0$, $\| \cdot \|$ eine beliebige Norm im \mathbf{R}^n und (\cdot, \cdot) ein beliebiges inneres Produkt auf $\mathbf{R}^n \times \mathbf{R}^n$.

Gegeben sei die Anfangswertaufgabe

$$y' = f(x, y), \quad y(x_0) = y_0.$$

Ist f eine stetige Funktion $f : Q \subset \mathbf{R} \times \mathbf{R}^n \rightarrow \mathbf{R}^n$ mit der Eigenschaft

$$(f(x, u) - f(x, v), u - v) \leq K (u - v, u - v).$$

mit einem $K \in \mathbf{R}$ und hat die Anfangswertaufgabe eine Lösung auf dem Intervall $I = [x_0, x_0 + a]$, dann ist diese auch die einzige.

Aussagen über die Veränderung der Lösungen bei Veränderung von f und (x_0, y_0) können oft mit Hilfe von Differentialungleichungen erzielt werden:

Satz 1.3

Es sei (\cdot, \cdot) das euklidische innere Produkt auf $\mathbf{R}^n \times \mathbf{R}^n$ und $\| \cdot \|_2$ die zugehörige euklidische Norm im \mathbf{R}^n .

Gegeben seien die Anfangswertaufgaben

$$\begin{aligned} y_1' &= f(x, y_1), & y_1(x_0) &= u_0, \\ y_2' &= f(x, y_2) + \delta(x), & y_2(x_0) &= u_0 + \delta, \end{aligned}$$

mit einer stetigen Funktion $\delta(\cdot) : [x_0, \infty) \rightarrow \mathbf{R}^n$, $x_0 \in \mathbf{R}$ und $u_0, \delta \in \mathbf{R}^n$.

Es sei f eine stetige Funktion $f: [x_0, \infty) \times \mathbf{R}^n \rightarrow \mathbf{R}^n$ mit der Eigenschaft

$$(f(x,u) - f(x,v), u-v) \leq K(u-v, u-v)$$

für alle zulässigen x, u, v und einem $K < 0$.

Sind $y_1(\cdot)$ und $y_2(\cdot)$ Lösungen der jeweiligen Anfangswertaufgaben auf $[x_0, \infty)$, dann gilt für die Differenz $z(\cdot) = y_1(\cdot) - y_2(\cdot)$ die Differentialungleichung

$$\frac{1}{2} \frac{d}{dx} \|z(x)\|_2^2 \leq K \|z(x)\|_2^2 + \|\delta(x)\|_2^2 \|z\|_2, \quad \|z(x_0)\|_2^2 = \|\delta\|_2^2$$

und

$$\|z(x)\|_2^2 \leq u(x) \quad \text{für alle } x \in [x_0, \infty).$$

Dabei ist $u(\cdot)$ die Lösung der Riccati-Differentialgleichung

$$u'(x) = 2Ku(x) + 2\delta(x)u(x)^{1/2}, \quad u(x_0) = \|\delta\|_2^2.$$

§ 1.2 Elementare Differenzenverfahren für gewöhnliche Anfangswertaufgaben Ein Konvergenzsatz

Gegeben sei eine für alle „Zeiten“ $x \geq 0$ und alle Anfangsbedingungen y_0 eindeutig lösbare Anfangswertaufgabe

$$y' = f(x, y), \quad y(0) = y_0,$$

mit einer stetigen Funktion $f: \mathbf{R} \times \mathbf{R} \rightarrow \mathbf{R}$.

Ist $y(\cdot)$ die Lösung der Aufgabe, dann gelten für alle $x > 0$ und für alle $h > 0$ die folgenden Gleichungen

$$\lim_{h \rightarrow 0} (y(x+h) - y(x))/h = f(x, y(x)),$$

$$\lim_{h \rightarrow 0} (y(x+h) - y(x))/h = \lim_{h \rightarrow 0} f(x+h, y(x+h)),$$

$$\lim_{h \rightarrow 0} (y(x+h) - y(x))/h = \lim_{h \rightarrow 0} [f(x, y(x)) + f(x+h, y(x+h))]/2,$$

$$\lim_{h \rightarrow 0} (y(x+h) - y(x))/h = \lim_{h \rightarrow 0} f(x+h/2, (y(x) + y(x+h))/2).$$

Wird nun auf den Grenzübergang verzichtet, dann können die obigen Zusammenhänge auch so gedeutet werden:

Jedes vorgegebene $h > 0$ und $y(x)$ liefern eine Näherung für $y(x+h)$ durch

$$\begin{aligned} y(x+h) &\cong y(x) + hf(x, y(x)), \\ y(x+h) &\cong y(x) + hf(x+h, y(x+h)), \\ y(x+h) &\cong y(x) + (h/2)(f(x, y(x)) + f(x+h, y(x+h))), \\ y(x+h) &\cong y(x) + hf(x+h/2, (y(x) + y(x+h))/2). \end{aligned}$$

Diese Näherungsformeln führen zu wohlbekannten Differenzenverfahren zur Integration von Anfangswertaufgaben, nämlich:

$y_{n+1} = y_n + hf(x_n, y_n)$	Explizites Euler-Verfahren
$y_{n+1} = y_n + hf(x_{n+1}, y_{n+1})$	Implizites Euler-Verfahren
$y_{n+1} = y_n + (h/2)(f(x_n, y_n) + f(x_{n+1}, y_{n+1}))$	Trapezregel
$y_{n+1} = y_n + hf((x_n+x_{n+1})/2, (y_n+y_{n+1})/2)$	„one-leg“ Trapezregel

Dabei ist y_n ein Näherungswert für $y(nh)$ an der Stelle $x_n = nh$.

Für alle vier Verfahren gilt der folgende Konvergenzsatz:

Satz 1.4

Gegeben sei eine für alle $x \geq 0$ und alle Anfangsbedingungen y_0 eindeutig lösbare Anfangswertaufgabe

$$y' = f(x, y), \quad y(0) = y_0,$$

mit einer stetigen Funktion $f: \mathbf{R} \times \mathbf{R}^n \rightarrow \mathbf{R}^n$.

Dann gilt für alle $n \rightarrow \infty$ und alle $h \rightarrow 0$ mit $nh \rightarrow x$

$$\lim_{\substack{n \rightarrow \infty \\ h \rightarrow 0 \\ n \cdot h \rightarrow x}} y_n = y(x).$$

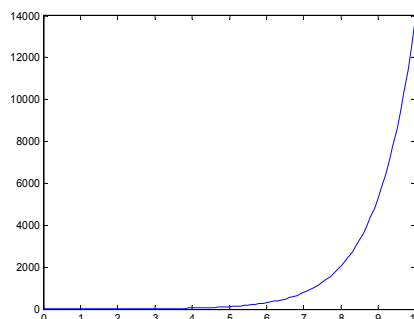
Dieser schöne Satz liefert leider keinen Hinweis darauf, ob die angegebenen Verfahren für eine konkrete Aufgabe auch wirklich brauchbar sind. Dieses zeigen die folgenden Beispiele:

Numerisch integriert wurde die Aufgabe

$$y' = \lambda y, \quad y(0) = 1, \quad \lambda = -1000,$$

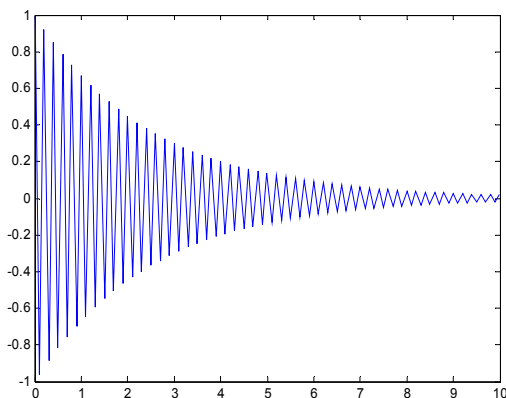
auf dem Intervall von $[0, 10]$ mit dem expliziten Euler-Verfahren.

Das Bild zeigt die völlig falsche Näherungslösung für die Schrittweite $h = 0.1$. Erst Schrittweiten, die kleiner als $1/500$ sind, führen zu akzeptablen Ergebnissen. Wird beachtet, dass in der Praxis Aufgaben auftreten können mit Lösungsanteilen, die ein $\lambda = -10^{12}$ besitzen, dann zeigt sich schnell, dass dieses Verfahren zu Schrittweiten führen kann, die nicht mehr akzeptabel sind.

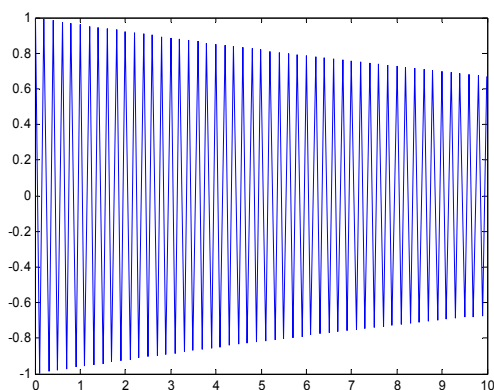


Integriert man die Aufgabe mit der Trapezregel oder deren „one-leg“ Version, dann können die Näherungslösungen zwar nicht mehr – wie bei dem expliziten Euler-Verfahren - explodieren,

jedoch zeigen sich stets unangemessene Oszillationen. Das nachfolgende Bild zeigt die Näherungslösung für $h = 0.1$.



Lässt man $\lambda \rightarrow -\infty$ gehen, dann fällt – im Gegensatz zur Lösung der Anfangswertaufgabe – die Näherungslösung immer noch oszillierend und obendrein auch noch langsamer gegen Null. Das folgende Bild gibt die Näherungslösung für $\lambda = -10000$ wieder.



Nur das implizite Euler-Verfahren würde mit jeder Schrittweite noch qualitativ richtige Lösungen liefern (Probieren Sie es!).

Eine Integration der Aufgabe

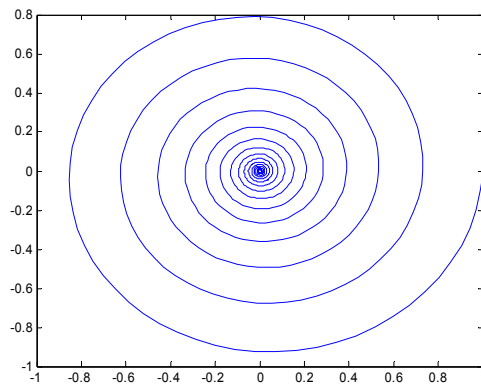
$$\begin{aligned} y' &= x, & y(0) &= 1, \\ x' &= -y, & x(0) &= 0, \end{aligned}$$

mit dem impliziten Euler

```
h= 0.1;
z(:,1) = [1;0];
A = [1 -h;h 1]
B = inv(A)
for i = 1:1000
    z(:,i+1) = inv(A)*z(:,i)
end
plot(z(1,:),z(2,:))
```

führt zu einem qualitativ völlig falschen Ergebnis. Man beachte: Für die Lösung gilt

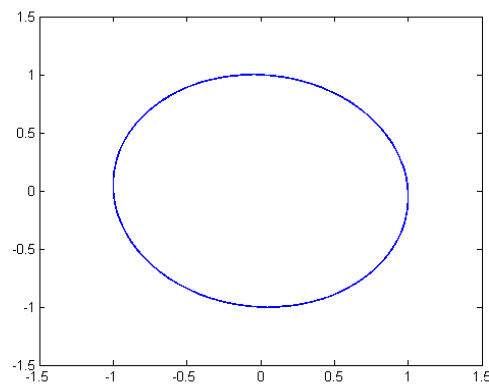
$x^2(t)+y^2(t) = 2$, d.h. im Phasenraum muss die Lösung auf einem Kreis liegen.



Dieses Fehlverhalten tritt übrigens auch mit dem expliziten Euler-Verfahren auf, aber nicht mit der Trapezregel oder deren one-leg Version (Probieren Sie es!). Auch eine Mischung des expliziten und des impliziten Euler-Verfahren

$$\begin{aligned}y_{n+1} &= y_n + hx_n \\ x_{n+1} &= x_n - hy_{n+1}\end{aligned}$$

führt zu einem qualitativ richtigen Ergebnis (Beweisen Sie es!).



Das zugehörige Programm ist:

```
h= 0.1;
z(:,1) = [1;0];
A = [1 h;-h 1-h*h]
for i = 1:1000
    z(:,i+1) = A*z(:,i)
end
plot(z(1,:),z(2,:))
```

§ 1.3 Ein elementarer Konvergenzbeweis

Alle bisher behandelten Beispiele können als Anfangswertaufgaben vom Typ

$$y'(x) = f(x,y(x)), \quad y(x_0) = y_0,$$

mit der „Wachstumsbedingung“ $\|f(x,u)-f(x,v)\| \leq K \|u-v\|$ angesehen werden. Man sagt: die Aufgaben sind Lipschitzstetig.

Für diese Klasse von Aufgaben soll, am einfachen Beispiel des expliziten Euler-Verfahrens, der Einfluss unterschiedlicher Parameter auf die Differenz zwischen Näherungslösung und Lösung der Anfangswertaufgabe eingehender untersucht werden.

Es gilt der

Satz 1.5

Es sei $f: \mathbf{R} \times \mathbf{R}^n \rightarrow \mathbf{R}^n$ eine stetige Funktion mit

$$\|f(x,u) - f(x,v)\| \leq K \|u-v\| \quad \text{für alle } x, u, v \in \mathbf{R} \times \mathbf{R}^n \times \mathbf{R}^n \text{ und } K \geq 0.$$

Auf dem Intervall $[0, T]$ sei die Lösung $y(\cdot)$ der Anfangswertaufgabe

$$y' = f(x, y), \quad y(0) = y_0,$$

zweimal stetig differenzierbar und $M = \max_{t \in [0, T]} \|y''(x)\|$.

Gegeben sei das explizite Euler-Verfahren

$$y_{n+1} = y_n + hf_n$$

mit dem Startwert y_0^h und es sei $F_0 = y(0) - y_0^h$ und $F_n = y(x_n) - y_n$.

Für jedes $h > 0$ und alle n mit $nh \leq T$ gilt dann

$$\|F_n\| \leq e^{KT} \|F_0\| + (1/2)(M/K)(e^{KT} - 1)h.$$

Ein ähnliches Resultat ergibt sich auch für das implizite Euler-Verfahren. Auch die Trapezregeln liefern ein solches Resultat, allerdings mit der (besseren) Potenz h^2 , sofern die Lösung der Anfangswertaufgabe dreimal stetig differenzierbar ist.

Bemerkung

Alle Verallgemeinerungen der oben genannten Verfahren (Adams-Verfahren, Runge-Kutta-Verfahren und Simplektische-Verfahren) führen zu einem ähnlichen Resultat. Lediglich kann (dieses ist allerdings wichtig für die Praxis) die Potenz von h höher ausfallen.

Der letzte Satz deutet - auf den ersten Blick - auch wieder an, dass alle fünf Verfahren für die numerische Integration aller Anfangswertaufgaben geeignet sind. Was offensichtlich nicht der Fall ist!

Man bekommt jedoch schon einige Hinweise. Zu beachten ist offenbar, dass vielleicht die Konstante K sehr groß ist, über lange Zeiträume T integriert werden muss oder die Lösungen der Differentialgleichung ein spezielles Verhalten aufweisen. Bei diesen Problemen können sich die Verfahren dann sehr unterschiedlich verhalten, zu einem prohibitiv großen Aufwand führen oder sogar völlig versagen. Diese Erfahrung hat dazu geführt, die Differenzenverfahren auf unterschiedlichste (Stabilitäts-) Eigenschaften zu untersuchen.

§ 1.4 Stabilitätsbedingungen für Differenzenverfahren

Die in den Beispielen aufgetretenen Probleme konnten (§1.2) in der Klasse der Lipschitzstetigen Anfangswertaufgaben nicht geklärt werden. Für weitere Untersuchungen betrachten wir jetzt lediglich die Klasse der linearen und asymptotisch stabilen Anfangswertaufgaben

$$y' = Ay, \quad y(0) = y_0,$$

mit einer reellen $n \times n$ -Matrix A .

Unter der Annahme dass A diagonalisierbar ist, lässt sich das System mit einer Ähnlichkeitstransformation H überführen in

$$Hy' = HAH^{-1}Hy$$

oder in eine Differentialgleichung der Form

$$z' = Jz \quad \text{oder} \quad z' = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix} z$$

mit möglicherweise komplexen (Eigenwerten) λ_i .

Es genügt deshalb, die eindimensionale Testgleichung (G. Dahlquist)

$$y' = \lambda y \quad \text{mit} \quad \operatorname{Re} \lambda < 0,$$

deren Lösungen $y(x) = y_0 \cdot \exp(\lambda \cdot x)$ alle für $x \rightarrow \infty$ gegen Null gehen, zu betrachten.

Die Anwendung der vier Verfahren auf diese Differentialgleichung liefert:

$$\begin{array}{ll} y_{n+1} = y_n + h\lambda y_n & \text{oder} \quad y_{n+1} = (1 + h\lambda)y_n \\ y_{n+1} = y_n + h\lambda y_{n+1} & \text{oder} \quad y_{n+1} = (1/(1 - h\lambda))y_n \\ y_{n+1} = y_n + (h/2)\lambda(y_n + y_{n+1}) & \text{oder} \quad y_{n+1} = [(1 + (h\lambda/2))/(1 - (h\lambda/2))]y_n \\ y_{n+1} = y_n + (h/2)\lambda(y_n + y_{n+1}) & \text{oder} \quad y_{n+1} = [(1 + (h\lambda/2))/(1 - (h\lambda/2))]y_n \end{array}$$

Die Formeln auf der rechten Seite zeigen, dass neben h und λ das Produkt dieser beiden Zahlen eine besondere Rolle spielt.

- Für festes h und hinreichend kleine $|h\lambda|$ liefern alle Verfahren (Stabilität)

$$\lim_{n \rightarrow \infty} y_n = 0.$$

- Für festes h und alle $h\lambda$ liefert nur das erste Verfahren nicht (A-Stabilität)

$$\lim_{n \rightarrow \infty} y_n = 0.$$

- Bei festem h und $\operatorname{Re}(\lambda) \rightarrow -\infty$ liefert nur das implizite Euler-Verfahren das gewünschte Verhalten (L-Stabilität)

$$\lim_{n \rightarrow \infty} y_n = 0.$$

Auch numerisch konnten wir diese verschiedenen Verhaltensweisen eindrucksvoll und schnell verifizieren.

Dieses Verhalten hat zu einer Klassifizierung der Verfahren geführt:

Expliziter Euler	Impliziter Euler	“one-leg” und Trapezregel
Stabil	Stabil	Stabil
Nicht A-Stabil	A-Stabil	A-Stabil
Nicht L-Stabil	L-Stabil	Nicht L-Stabil

Bemerkung:

Diese Konzepte oder diese unterschiedlichen (Stabilitäts-) Verhalten der Differenzenverfahren können für eine praktische und effiziente numerische Integration einer Anfangswertaufgabe offenbar nicht unberücksichtigt bleiben. Sie führen umgangssprachlich zu dem folgenden Vorschlag:

„Für moderate Probleme sind stabile Verfahren geeignet, für durchschnittlich steife Probleme sind A-stabile oder Verfahren mit hinreichend großem Bereich der absoluten Stabilität geeignet und bei steifen Problemen muss mit L-stabilen Verfahren integriert werden“.

In **MATLAB** werden einige numerische Verfahren zur Lösung von Anfangswertaufgaben oder **ODE-Solver** angeboten. Der Katalog der angebotenen Verfahren berücksichtigt den obigen Vorschlag.

Was unter moderat, durchschnittlich steif, steif und Bereich der absoluten Stabilität zu verstehen ist, präzisieren die nachfolgenden Betrachtungen:

Steifheit

Die Lösung der (2*2) Anfangswertaufgabe

$$y' = Ay, \quad y(0) = y_0,$$

lasse eine Darstellung der Form

$$y(x) = \alpha_1 e^{\lambda_1 x} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} + \alpha_2 e^{\lambda_2 x} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

(dabei seien λ_1 und λ_2 Eigenwerte von A mit den Eigenvektoren $\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$ und $\begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$) zu.

Die „Steifheit“ des Systems wird dann durch das Verhältnis $|\lambda_1 / \lambda_2|$ festgelegt. Diese hat Einfluss auf das einzusetzende Verfahren.

Sind z.B. die $\lambda_i < 0$, $i = 1, 2$, von unterschiedlicher Größenordnung (langsame und schnelle Dynamik), dann spricht man von einem steifen System.

Zum betragsgrößten Eigenwert gehört dann langfristig ein Lösungsanteil, der in Relation zum zweiten Anteil „unbedeutend“ ist. Trotzdem müsste man beim expliziten Euler-Verfahren die Schrittweite nach dem „unbedeutenden“ Anteil richten, was aufwendig und der Sache unangemessen ist.

In einem zweiten Versuch könnte man dieses Problem mit der Trapezregel in Angriff nehmen. Es zeigte sich aber, dass auch die Trapezregel nicht immer zu adäquaten Näherungslösungen führen muss. Wären die beiden Eigenwerte von sehr unterschiedlicher Größenordnung, dann würde auch die Trapezregel (zumindest in einer Komponente) völlig unakzeptable Ergebnisse liefern. Außerdem neigt die Trapezregel zu unangemessenen Oszillationen.

Als Ausweg bleibt das implizite Euler-Verfahren oder die Forderung nach der „L-Stabilität“. Die L-Stabilität verlangt im Wesentlichen die „sachgerechte“ Integration aller Aufgaben

$$\varepsilon y' = -y, \quad y(0) = y_0, \quad \text{mit } \varepsilon > 0 \quad \text{und} \quad \varepsilon \rightarrow 0.$$

Die bisher angegebenen Differenzenverfahren sind nur Repräsentanten von großen Klassen von Differenzenverfahren. Die Forderung nach A-Stabilität bei einem allgemeinen Differenzenverfahren hat sich als schwierig erwiesen. Den Verfahren wurden deshalb Bereiche der absoluten Stabilität zugewiesen. Die Gestalt dieser Bereiche sagt dann etwas darüber aus, für welchen Typ von Anfangswertaufgaben die zugehörigen Differenzenverfahren geeignet sind.

Bereiche der absoluten Stabilität

Einen Hinweis auf einen weiteren theoretischen und praktischen Zugang für die Beurteilung von Differenzenverfahren gibt der Bereich der absoluten Stabilität der Verfahren.

Für die von uns betrachteten vier Differenzenverfahren kann der jeweilige Bereich der absoluten Stabilität schnell angegeben werden:

Wird das Verfahren auf die Testgleichung

$$y' = \lambda y \quad \text{mit} \quad \text{Re } \lambda < 0,$$

angewandt, dann ist der Bereich der absoluten Stabilität die Menge der $\tilde{h} = \lambda h$, für die gilt

$$\lim_{n \rightarrow \infty} y_n = 0.$$

Beispiele

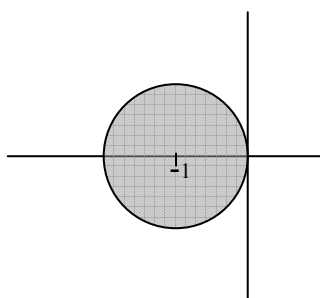
Das explizite Euler Verfahren

$$y_{n+1} = y_n + hf_n$$

führt bei der Integration von $y' = \lambda y$ zu

$$y_{n+1} - y_n = \lambda h y_n \quad \Leftrightarrow \quad y_{n+1} - (1 + \tilde{h})y_n = 0.$$

Damit ist dieses Verfahren absolut stabil für alle \tilde{h} mit $|1 + \tilde{h}| < 1$ und „geeignet“ bei akzeptablen Schrittweiten h und λ derart, dass $|1 + \tilde{h}| < 1$ ist.



Bereich der absoluten Stabilität
(Explizites Euler-Verfahren)

Damit die Asymptotik stimmt, muss die Schrittweite h so gewählt werden, dass $h\lambda$ im Bereich der absoluten Stabilität liegt (Grauer Bereich)

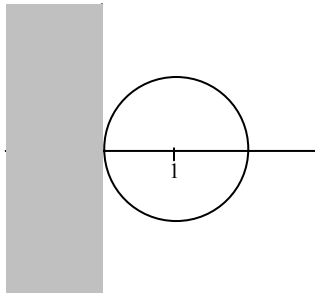
Das implizite Euler-Verfahren

$$y_{n+1} = y_n + hf_{n+1}$$

führt auf die charakteristische Gleichung

$$\xi - 1/(1+h) = 0.$$

Damit ist die gesamte linke Halbebene von \mathbb{C} der Bereich der absoluten Stabilität.



Bereich der absoluten Stabilität
(Implizites Euler-Verfahren)

Die Asymptotik stimmt für alle Schrittweiten h ,
da $h\lambda$ stets im Bereich der absoluten Stabilität
liegt (Grauer Bereich)

Ähnliches gilt auch für die Trapezregeln.

Definition (A-stabil)

Eines der angegebenen Verfahren heißt A-stabil, wenn dessen Bereich der absoluten Stabilität die gesamte linke komplexe Ebene umfasst.

Bemerkung

Die Unterscheidung zwischen A-stabilen und L-stabilen Verfahren kann aus den Bereichen der absoluten Stabilität nicht erkannt werden. Hierfür sind offenbar noch weitere Qualitäten erforderlich, nämlich das Verhalten bei $\operatorname{Re}(\lambda) \rightarrow -\infty$. Ein Stichwort in diesem Zusammenhang sind die BDF- oder Gear-Verfahren.

G-Stabile Verfahren

Offen bleibt die Frage, ob die vorgestellten Konzepte auch für die Beurteilung von Differenzenverfahren bei der Integration nichtlinearer Anfangswertaufgaben geeignet sind.

Die Praxis hat gezeigt, dass dieses weitestgehend der Fall ist. Für eine Integration von nicht-linearen Problemen werden also geeignete Verfahren über den Umweg der „linearisierten“ Aufgaben festgelegt.

Ein zugehöriges theoretisches Ergebnis aus diesem Bereich soll noch angegeben werden.

Gegeben sei eine (im Sinne von Minty) monotone Funktion f , d.h.

$$(f(x,z) - f(x,w), (z-w)) \leq K \|z-w\|_2^2 \quad \text{für alle } x, z, w \text{ und ein } K \in \mathbf{R}.$$

Für die Differenz von zwei Lösungen $u(\cdot)$, $v(\cdot)$ der Differentialgleichung $y'=f(x,y)$ gilt dann (Stabilitätsungleichung):

$$\| u(x)-v(x) \|_2 \leq \| u_0-v_0 \|_2 e^{K(x-x_0)}.$$

Ist $K \leq 0$, dann bleibt die Differenz der Lösungen beschränkt oder strebt ($K < 0$) mit zunehmendem x sogar gegen Null.

Es stellt sich also die Frage: Welche der oben angegebenen Verfahren erhalten diese Qualität?

Definition (G-stabil)

Eine „one-leg“-Methode heißt G-stabil (G. Dahlquist), wenn für zwei numerische Lösungen $\{u_n\}_{n=0,\dots,N+1}$ und $\{v_n\}_{n=0,\dots,N+1}$ gilt

$$\| u_{n+1}-v_{n+1} \|_2 \leq \| u_n-v_n \|_2$$

für alle Schrittweiten h und $K \leq 0$.

Der folgende schöne Satz zeigt, dass für die nichtlinearen Probleme die A-stabilen Verfahren durch ihre „one-leg“-Versionen ersetzt werden müssen!

Satz 1.6

Die Verfahren sind genau dann A-stabil, wenn deren „one-leg“-Versionen G-stabil sind.

Der aufmerksame Leser wird noch einige Bemerkungen zu der Differentialgleichung

$$\begin{aligned} y' &= x, \\ x' &= -y \end{aligned}$$

und deren numerische Behandlung vermissen.

Die Differentialgleichung ist wohl das einfachste Beispiel für ein Hamiltonsystem. Neben der oben angegebenen Erhaltungsgröße haben diese Systeme die Eigenschaft, dass der sogenannte „Fluss“ zeitlich volumeninvariant ist. Es kann nachgewiesen werden, dass die one-leg Version der Trapezregel diese Eigenschaft beibehält. Allgemein spricht man von einem symplektischen Verfahren.

Numerische Behandlung von gewöhnlichen Differentialgleichungen

Eine Einführung

Universität Hamburg
SoSe07

K. Taubert

Numerik gewöhnlicher Anfangswertaufgaben

2 EINSCHRITTVERFAHREN, insbesondere RUNGE-KUTTA-VERFAHREN

Es werden allgemeine Einschrittverfahren und insbesondere (R)unge-(K)utta-(V)erfahren zur numerischen Approximation von Lösungen von Anfangswertaufgaben eingeführt. Mit den BUTCHER-Arrays wird ein übersichtliches Mittel zur Darstellung von allgemeinen expliziten und impliziten RKV zur Verfügung gestellt.

Die Begriffe Konsistenz und lokaler (Abbruch-) Fehler von Einschrittverfahren werden eingeführt. Diese Begriffe sagen etwas über die „Güte“ der Verfahren aus und liefern für geeignete Klassen von Anfangswertaufgaben auch die „Konvergenz“ dieser Verfahren. Einschrittverfahren eignen sich besonders für Schrittweitensteuerungen die zu erheblichen Recheneinsparungen führen können.

Später zeigt sich, dass viele RKV sich auch für die Approximation von Lösungen von Algebra-Differentialgleichungen eignen.

§ 2.1 Einschrittverfahren

Gegeben sei die skalare Anfangswertaufgabe

$$y' = f(x, y), \quad y(x_0) = y_0, \quad (*)$$

mit einer stetigen Funktion $f: \mathbf{R} \times \mathbf{R} \rightarrow \mathbf{R}$, die außerdem die Lipschitzbedingung

$$|f(x, u) - f(x, v)| \leq K |u - v| \quad \text{für alle } x, u, v \in \mathbf{R}$$

mit einem $K \geq 0$ erfüllt.

Bemerkung

Die Lipschitzbedingung wird häufig auch durch die Forderung der stetigen Differenzierbarkeit von f in einer Umgebung (x_0, y_0) ersetzt.

Die Voraussetzungen führen dazu, dass es eine eindeutige Lösung auf jedem Intervall I mit $x_0 \in I$ gibt.

Ein allgemeines explizites Einschrittverfahren für (*) hat die Form (mit $h > 0$)

$$y_{n+1} = y_n + h \Phi(x_n, y_n, h, f) \quad (**)$$

mit einer reellwertigen Funktion $\Phi(x, u, h, f)$. Außerdem verlangen wir noch, dass Φ eine Lipschitzbedingung

$$|\Phi(x, u, h, f) - \Phi(x, v, h, f)| \leq L_f |u - v|$$

für alle zulässigen x, u, v und h mit einer Lipschitzkonstanten $L_f > 0$ erfüllt.

Beispiel 1 (Das Euler-Verfahren)

Ein explizites Einschrittverfahren kennen wir bereits, nämlich das explizite Euler-Verfahren:

$$y_{n+1} = y_n + hf(x_n, y_n) \quad \text{mit} \quad \Phi(x, y, h, f) = f(x, y).$$

Beispiel 2 (Das Euler-Cauchy-Verfahren)

Das Euler-Cauchy-Verfahren baut auf das Euler-Verfahren auf und hat die Gestalt

$$y_{n+1} = y_n + h \left[\frac{1}{2} f(x_n, y_n) + \frac{1}{2} f(x_n + h, y_n + hf(x_n, y_n)) \right]$$

oder

$$\Phi(x, y, h, f) = \frac{1}{2} f(x, y) + \frac{1}{2} f(x+h, y+hf(x, y)).$$

Beispiel 3 (Verbessertes Euler-Verfahren)

Hierbei handelt es sich um eine Art Mittelpunktsformel

$$y_{n+1} = y_n + h f\left(x_n + \frac{h}{2}, y_n + \frac{1}{2} hf(x_n, y_n)\right)$$

oder

$$\Phi(x, y, h, f) = f\left(x + \frac{h}{2}, y + \frac{1}{2} hf(x, y)\right).$$

Beispiel 4 (Klassisches Runge-Kutta-Verfahren)

Das klassische Runge-Kutta-Verfahren (der Ordnung 4) hat die Form

$$y_{n+1} = y_n + \frac{1}{6} h(k_1 + 2k_2 + 2k_3 + k_4)$$

mit

$$k_1 = f(x_n, y_n)$$

$$k_2 = f\left(x_n + \frac{1}{2}h, y_n + \frac{1}{2}hk_1\right)$$

$$k_3 = f\left(x_n + \frac{1}{2}h, y_n + \frac{1}{2}hk_2\right)$$

$$k_4 = f(x_n + h, y_n + hk_3).$$

und entsprechendem Φ .

Bemerkung

Man beachte, dass die angegebenen Funktionen Φ stets ein gewichtetes Mittel von Steigungen sind. Da f eine Lipschitzbedingung erfüllt, kann leicht verifiziert werden, dass die bisher angegebenen Funktionen Φ auch die geforderte Lipschitzbedingung erfüllen.

Gegeben sei nun die Anfangswertaufgabe (*), ein explizites Einschrittverfahren (**) und ein Intervall (x_0, x_0+T) mit $0 < T < \infty$.

Es sei $y(\cdot)$ die Lösung von (*) auf $[x_0, x_0+T)$, $x_n = x_0 + nh$ für $h > 0$ mit $n = 0, 1, 2, \dots, N_h$ und N_h die maximale ganze Zahl, so dass $x_{N_h} \leq x_0+T$.

Für jedes $0 \leq n \leq N_h - 1$ gilt dann

$$y(x_{n+1}) = y(x_n) + h \Phi(x_n, y(x_n), h, f) + \varepsilon_{n+1}.$$

mit einem geeigneten ε_{n+1} .

Definition 2.1 (Lokaler Fehler)

Der Defekt $E_{n+1}(h) = y(x_{n+1}) - y(x_n) - h \Phi(x_n, y(x_n), h, f)$ heißt lokaler Fehler im Knoten x_{n+1} .

Bemerkung

Der Begriff des lokalen Fehlers ist in der Literatur nicht ganz einheitlich. In der Darstellung

$$E_{n+1}(h) = h \tau_{n+1}(h).$$

wird auch $\tau_{n+1}(h)$ oft als lokaler (Abbruch-) Fehler im Knoten x_{n+1} bezeichnet

Definition 2.2 (Ordnung)

Ein explizites Einschrittverfahren (**) hat die Ordnung $p \geq 1$ ($p \in \mathbf{N}$), wenn für alle Anfangswertaufgaben (x_0, y_0) mit hinreichend oft differenzierbaren Lösungen gilt

$$|y(x_0+h) - y_1| \leq Mh^{p+1}.$$

Dabei sei $y(x_0+h)$ die Lösung zum Anfangswert $y(x_0) = y_0$ an der Stelle x_0+h und

$$y_1 = y_0 + h \Phi(x_0, y_0, h, f).$$

Bemerkung

Häufig wird für die Einschrittverfahren auch noch ein Konsistenzbegriff eingeführt. Der Begriff Konsistenz wird in der Literatur nicht ganz einheitlich formuliert. Gelegentlich wird die Konsistenz auch mit der Forderung nach einer Mindestordnung 1 gleichgesetzt.

Man beachte, dass bei der Definition der Ordnung für explizite RKV die Konstante M zunächst von den Anfangsbedingungen und f abhängt.

Die Ordnung p kann auch so interpretiert werden: Die Taylorentwicklung von $y(x_0+h)$ an der Stelle x_0 muss einschließlich dem Term h^p mit y_1 übereinstimmen.

Beispiel 5

Gegeben sei eine Anfangswertaufgabe (*) mit einer sogar stetig differenzierbaren Funktion f . Für alle Anfangsbedingungen sind dann die Lösungen der Anfangswertaufgaben zweimal stetig differenzierbar.

Der lokale Fehler im Knoten x_{n+1} des expliziten Euler-Verfahrens ergibt sich aus

$$\begin{aligned} E_{n+1}(h) &= y(x_{n+1}) - y(x_n) - h \Phi(x_n, y(x_n); h, f) \\ &= y(x_{n+1}) - y(x_n) - hf(x_n, y(x_n)) \\ &= y(x_n) + y'(x_n)h + y''(\zeta) \frac{h^2}{2!} - y(x_n) - hy'(x_n) = y''(\zeta) \frac{h^2}{2} \end{aligned}$$

mit einem $\zeta \in (x_n, x_{n+1})$.

Damit ist der lokale Fehler im Knoten x_{n+1}

$$E_{n+1}(h) = \frac{h^2}{2} y''(\zeta)$$

Das Euler-Verfahren ist nach Konstruktion von der Ordnung 1.

Bemerkungen

Im Kapitel 3 wird sich zeigen, dass die Ermittlung der Ordnung eines Einschrittverfahrens durchaus mühsam und langwierig sein kann. Besondere Techniken für die Ermittlung der Ordnung von RKV sind entwickelt worden.

Den einzelnen Verfahren wird also eine Ordnung p zugeordnet. Diese Ordnung plus Eins liefert auch die Größenordnung des lokalen Fehler, wenn f hinreichend oft differenzierbar ist.

Im schraffierten Bereich der nachfolgenden Tabelle ist die Ordnung p für das jeweilige Verfahren angegeben. Wird zu der Zahl in der Tabelle eine 1 addiert, dann liefert diese Zahl die Größenordnung des lokalen Fehlers für die jeweilige Klasse von Funktionen f .

Euler Verfahren	Verbessertes Euler Verfahren	Euler-Cauchy Verfahren	Klassisches Runge-Kutta Verfahren	Klassen f
1	1	1	1	f 1*stet. diff
1	2	2	2	f 2*stet. diff
1	2	2	4	f 4*stet. diff

§ 2.2 Ein erster Konvergenzsatz. Bemerkungen

Es soll hier ein erster Konvergenzsatz für Einschrittverfahren angegeben werden.

Satz 2.1

Gegeben sei das Intervall $[x_0, x_0+T]$ und die skalare Anfangswertaufgabe

$$y' = f(x, y), \quad y(x_0) = y_0^a,$$

mit einer p -mal stetig differenzierbaren Funktion $f: \mathbf{R} \times \mathbf{R} \rightarrow \mathbf{R}$, die außerdem die Lipschitzbedingung

$$|f(x, u) - f(x, v)| \leq K |u - v| \quad \text{für alle } x, u, v \in \mathbf{R}$$

mit einem $K \geq 0$ erfüllt. Es sei $y(\cdot)$ die Lösung der Anfangswertaufgabe auf $[x_0, x_0+T]$.

Gegeben sei ein Einschrittverfahren mit der Ordnung p

$$y_{n+1} = y_n + h \Phi(t_n, y_n, h, f) + \delta_n, \quad y_0 = y_0^h,$$

das außerdem die Lipschitzbedingung.

$$|\Phi(x, u, h, f) - \Phi(x, v, h, f)| \leq L_f |u - v|$$

erfüllt.

Es sei $\delta \geq |\delta_n|$ eine obere Schranke für die während der Durchführung des Einschrittverfahrens auftretenden (Rundungs-)fehler δ_n

Für die in Frage kommenden lokalen Fehler $E_{n+1}(h)$ gelte $|E_{n+1}(h)| \leq Mh^{p+1}$ mit einem einheitlichen M .

Es sei $F_n = y(x_n) - y_n^h$ und insbesondere $F_0 = y(x_0) - y_0^h$.

Für jedes $h > 0$ und alle n mit $x_0 + nh \leq x_0 + T$ gilt dann

$$|F_n| \leq e^{L_f T} |F_0| + \frac{M}{L_f} (e^{L_f T} - 1) h^p + \frac{1}{L_f} (e^{L_f T} - 1) \frac{\delta}{h}.$$

Beweis

Wir führen den Beweis nur für das explizite Euler-Verfahren und $\delta_n = 0$ durch. Einsetzen der Lösung der Anfangswertaufgabe in das Verfahren liefert:

$$\begin{aligned} E_{n+1}(h) &= y(t_{n+1}) - y(t_n) - hf(t_n, y(t_n)) \\ &= y(t_n) + y'(t_n)h + y''(\zeta) \frac{h^2}{2!} - y(t_n) - hy'(t_n) \\ &= y''(\zeta) \frac{h^2}{2} \end{aligned}$$

mit einem geeigneten $\zeta \in (t_n, t_{n+1})$.

Damit ist

$$y(t_{n+1}) - y_{n+1} = y(t_n) - y_n + h(f(t_n, y(t_n)) - hf(t_n, y_n)) + y''(\zeta) \frac{h^2}{2}$$

oder

$$|F_{n+1}| \leq |F_n| + hK|F_n| + \varepsilon(h)$$

mit

$$\varepsilon(h) = \frac{h^2}{2!} \max_{\eta \in [t_0, t_0+T]} |y''(\eta)| =: Mh^2.$$

Die Rekursionsformel für $|F_n|$ liefert

$$\begin{aligned} |F_{n+1}| &\leq (1+hK)^{n+1} |F_0| + \sum_{v=0}^n (1+hK)^v \varepsilon(h) \\ &\leq (1+hK)^{n+1} |F_0| + \frac{(1+hK)^{n+1} - 1}{(1+hK) - 1} \varepsilon(h) \end{aligned}$$

$$\leq e^{K(x_{n+1}-x_0)} |F_0| + h \frac{M}{K} (e^{K(x_{n+1}-x_0)} - 1).$$

Werden noch die Rundungsfehler δ_n berücksichtigt, dann ergibt sich in ähnlicher Form

$$|F_{n+1}| \leq e^{K\tau} |F_0| + h \frac{M}{K} (e^{K\tau} - 1) + \frac{\delta}{h} \frac{1}{K} (e^{K\tau} - 1).$$

Bemerkungen

Es ist zweckmäßig, die angegebene Abschätzung für den globalen Fehler genauer zu untersuchen:

Ein Fehlerverhalten der Größenordnung $O(h^p)$ ist dann zu erwarten, wenn

1. $|F_0| \leq C_1 h^p$,
2. $|\delta_n| \leq C_2 h^{p+1}$,
3. das Einschrittverfahren von der Ordnung p ist und
4. f p -mal stetig differenzierbar ist.

Ist auch nur eine der obigen Bedingungen nicht erfüllt, dann muss von einem schlechteren Fehlerverhalten ausgegangen werden.

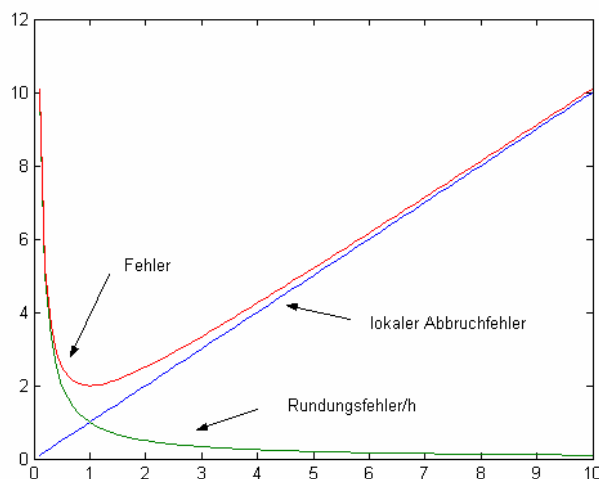
Von praktischer Bedeutung ist auch noch die folgende Beobachtung:

Angenommen, es ist $|F_0| = 0$, dann setzt sich der globale Fehler den zwei Komponenten

$$h^p * L \quad \text{und} \quad (1/h) * \bar{L} * \delta$$

zusammen.

Das qualitative Fehlerverhalten, welches sich daraus ergibt, kann aus der folgenden Zeichnung entnommen werden:



Sind die Rundungsfehler nicht der Schrittweite angepasst, dann ist unterhalb einer kritischen Schrittweite mit einer Zunahme des globalen Fehlers zu rechnen. D.h., eine ständige Verringerung der Schrittweite ist nicht nur aufwendig, sondern letztlich auch völlig kontraproduktiv.

§ 2.3 Explizite Runge-Kutta-Verfahren

Ein M-stufiges explizites Runge-Kutta-Verfahren hat die allgemeine Form

$$y_{n+1} = y_n + h \sum_{i=1}^M b_i Y_i'$$

mit

$$\begin{array}{ll}
 Y_1 = y_n & Y_1' = f(x_n, y_n) \\
 Y_2 = y_n + h[a_{21} Y_1'] & Y_2' = f(x_n + c_2 h, Y_2) \\
 Y_3 = y_n + h[a_{31} Y_1' + a_{32} Y_2'] & Y_3' = f(x_n + c_3 h, Y_3) \\
 \vdots & \vdots \\
 \vdots & \vdots \\
 Y_M = y_n + h[a_{M1} Y_1' + a_{M2} Y_2' + \dots + a_{M,M-1} Y_{M-1}'] & Y_M' = f(x_n + c_M h, Y_M)
 \end{array}$$

Als Kurzschreibweise haben sich (seit 1964) die zugehörigen BUTCHER-Arrays bewährt:

$$\begin{array}{c|cccc}
 0=c_1 & & & & \\
 c_2 & a_{21} & & & \\
 c_3 & a_{31} & a_{32} & & \\
 \vdots & \vdots & \vdots & \ddots & \\
 \vdots & \vdots & \vdots & & \ddots \\
 c_M & a_{M1} & a_{M2} & \dots & \dots & a_{M-1,M} \\
 \hline
 & b_1 & b_2 & \dots & \dots & b_{M-1} & b_M
 \end{array}
 \quad \text{oder} \quad
 \begin{array}{c|c}
 c & A \\
 \hline
 & b^T
 \end{array}$$

Einige Beispiele für Butcher-Arrays mit den zugehörigen Ordnungen M der RKV sind:

Euler-Verfahren

$$\begin{array}{c|c}
 0 & \\
 \hline
 1 &
 \end{array}$$

M=p=1

Euler-Cauchy

$$\begin{array}{c|cc}
 0 & & \\
 1 & 1 & \\
 \hline
 & 1/2 & 1/2
 \end{array}$$

M=p=2

Verbessertes Euler

$$\begin{array}{c|cc}
 0 & & \\
 1/2 & 1/2 & \\
 \hline
 & 0 & 1
 \end{array}$$

M=p=2

Heun-Verfahren

Klassisches RKV

3/8-Formel

0			
1/3	1/3		
2/3	0	2/3	
	1/4	0	3/4

M=p=3

0				
1/2	1/2			
1/2	0	1/2		
1	0	0	1	
	1/6	2/6	2/6	1/6

M=p=4

0				
1/3	1/3			
2/3	-1/3	1		
1	1	-1	1	
	1/8	3/8	3/8	1/8

M=p=4

Bemerkung:

Die Erwartung, dass mit einem p-stufigen expliziten RKV auch stets die Ordnung p erreicht werden kann, ist leider falsch. Die in besten Fall erreichbare Ordnung wird in der folgenden Tabelle genannt:

Ordnung	1	2	3	4	5	6	7	8
Mindeststufe	1	2	3	4	6	7	9	11

Beachten Sie, dass mit zunehmender Stufe auch die Anzahl der Funktionsauswertungen pro Schritt steigt.

§ 2.4 Ein zweiter Konvergenzsatz. Bemerkungen

Neben den bisher behandelten skalaren Anfangswertaufgaben

$$y' = f(x,y), \quad y(x_0) = y_0, \quad (***)$$

mit einer stetigen Funktion $f : \mathbf{R}^* \mathbf{R} \rightarrow \mathbf{R}$, sind auch Aufgaben mit der Eigenschaft

$$(f(x,u) - f(x,v))(u-v) \leq K(u-v)(u-v) \quad \text{für alle } x, u, v \in \mathbf{R}$$

und ein $K \in \mathbf{R}$ von praktischer und theoretischer Bedeutung.

Für $K < 0$ bedeutet diese Bedingung, dass f eine in der zweiten Komponente streng monoton fallende Funktion ist. Für $K \leq 0$ haben die Anfangswertaufgaben stets eine eindeutige Lösung für alle $x \geq x_0$ und für die Differenz von zwei Lösungen $y_1(\cdot)$ und $y_2(\cdot)$ gilt offenbar (!)

$$|y_1(x) - y_2(x)| \leq e^{K(x-x_0)} |y_1(x_0) - y_2(x_0)| \quad \text{für alle } x \geq x_0.$$

Einer numerischen Integration dieser Aufgaben mit Einschrittverfahren steht zunächst nichts im Wege. Der folgende Satz liefert sogar eine auf den ersten Blick adäquate Abschätzung für den globalen Fehler.

Satz 2.2

Gegeben sei die Anfangswertaufgabe

$$y' = f(x,y), \quad y(x_0) = y_0$$

mit einer stetigen Funktion $f : \mathbf{R}^* \mathbf{R} \rightarrow \mathbf{R}$ und der Eigenschaft

$$(f(x,u) - f(x,v))(u-v) \leq K(u-v)(u-v) \quad \text{für alle } x, u, v \in \mathbf{R}$$

mit einem $K < 0$.

Es sei $y_0(\cdot)$ die Lösung der Anfangswertaufgabe mit $y_0(x_0) = y_0$ auf dem Intervall $I = [x_0, x_0+T]$ und U eine Umgebung von $\{(x, y_0(x)) / x \in I\}$.
Gegeben sei ein RKV

$$y_{n+1} = y_n + h \Phi(x_n, y_n, h, f)$$

der Ordnung p und für alle zulässigen (x, v) aus U gelte

$$|y(x+h) - y_1| \leq Mh^{p+1}$$

mit einem einheitlichen M .

Dabei sei $y(\cdot)$ die Lösung der Anfangswertaufgabe mit $y(x) = v$ und $y_1 = v + h \Phi(x, v, h, f)$.
Dann gilt

$$|F_n| \leq \frac{M}{K} e^{(-Kh)} (e^{K(x_n - x_0)} - 1) h^p$$

für alle hinreichend kleinen h , so dass die numerische Approximation in U bleibt.

Beweis

Es sei $x_i = x_0 + ih$ für $i = 0, 1, 2, \dots, N$ mit $x_0 + Nh \in I$.

Es sei $y_1(x)$ die Lösung der Anfangswertaufgabe mit $y_1(x_1) = y_1$. Dann gilt

$$E_1 = |y_0(x_N) - y_1(x_N)| \leq Mh^{p+1} e^{K(x_N - x_1)}.$$

Es sei $y_2(x)$ die Lösung der Anfangswertaufgabe mit $y_2(x_2) = y_2$. Dann gilt

$$E_2 = |y_1(x_N) - y_2(x_N)| \leq Mh^{p+1} e^{K(x_N - x_2)}.$$

Ist allgemein $y_i(x)$ die Lösung der Anfangswertaufgabe mit $y_i(x_i) = y_i$ dann gilt

$$E_i = |y_{i-1}(x_N) - y_i(x_N)| \leq Mh^{p+1} e^{K(x_N - x_i)}.$$

Damit kann $|y_0(x_N) - y_N|$ abgeschätzt werden durch

$$|y_0(x_N) - y_N| \leq Mh^p \sum_{i=1}^N E_i$$

und die Summe durch

$$\int_{x_0}^{x_N} e^{K(x_N - h - x)} dx.$$

Bemerkung

Die im ersten Kapitel durchgeführte numerische Integration der Anfangswertaufgabe $y' = -1000y$, $y(0) = 1$, mit dem Euler-Verfahren zeigt, dass der obige Satz seine Tücken hat. Da die Funktion f der Bedingung

$$(f(x, u) - f(x, v))(u - v) \leq 0$$

genügt, ist es nahe liegend, zu fordern bzw. zu erwarten, dass auch Φ die Bedingung

$$(\Phi(x, u, h, f) - \Phi(x, v, h, f))(u - v) \leq 0$$

erfüllt. Dieses erweist sich als problematisch und wird später in Kapitel 5 gesondert behandelt.

§ 2.5 Implizite Runge-Kutta Verfahren

Mit den bisherigen Butcher-Arrays können nun leicht auch implizite M-stufige RKV eingeführt werden:

c_1	a_{11}	a_{21}	\dots	\dots	$a_{1,M-1}$	a_{1M}
c_2	a_{21}	a_{22}	\dots	\dots	$a_{2,M-1}$	a_{2M}
c_3	a_{31}	a_{32}	\dots	\dots	$a_{3,M-1}$	a_{3M}
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
c_M	a_{M1}	a_{M2}	\dots	\dots	$a_{M-1,M}$	$a_{M,M}$
	b_1	b_2	\dots	\dots	b_{M-1}	b_M

Diese Arrays führen zu

$$y_{n+1} = y_n + h \sum_{i=1}^M b_i f(x_n + c_i h, g_i)$$

$$g_i = y_n + h \sum_{j=1}^M a_{ij} f(x_n + c_j h, g_j).$$

Bei einem echten impliziten RKV ist also in jedem Schritt eine Fixpunktaufgabe $G = F(G)$ mit $F : \mathbf{R}^M \rightarrow \mathbf{R}^M$ (!) zu lösen.

Beispiel (M=2)

Ein zweistufiges implizites RKV hat die Gestalt

$$y_{n+1} = y_n + hb_1 f(x_n + c_1 h, g_1) + hb_2 f(x_n + c_2 h, g_2)$$

mit

$$g_1 = y_n + ha_{11} f(x_n + c_1 h, g_1) + ha_{12} f(x_n + c_2 h, g_2)$$

oder $G = F(G)$

$$g_2 = y_n + ha_{21} f(x_n + c_1 h, g_1) + ha_{22} f(x_n + c_2 h, g_2)$$

Mit $G = (g_1, g_2)^T$ ist also in jedem Schritt die Fixpunktaufgabe $G = F(G)$ zu lösen.

Damit kann der Aufwand für die Berechnung einer neuen Näherung erheblich anwachsen. Für hinreichend kleine Schrittweiten h und Lipschitzstetige f können diese Aufgaben mit dem Fixpunktsatz für kontrahierende Abbildungen gelöst werden. Bei Aufgaben, bei denen f eine große Lipschitzkonstante hat, wird man die Fixpunktaufgabe mit dem Newton-Verfahren lösen müssen.

Einige Beispiele für implizite RKV sind

Impliziter Euler Mittelpunktfornel

Gauß-Legendre

$$\begin{array}{c|c} 1 & 1 \\ \hline & 1 \end{array}$$

$$\begin{array}{c|c} 1/2 & 1/2 \\ \hline & 1 \end{array}$$

$$\begin{array}{c|cc} \frac{3-\sqrt{3}}{6} & \frac{1}{4} & \frac{3-2\sqrt{3}}{12} \\ \frac{3+\sqrt{3}}{6} & \frac{3+2\sqrt{3}}{12} & \frac{1}{4} \\ \hline & \frac{1}{2} & \frac{1}{2} \end{array}$$

Ordnung 1

Ordnung 2

Ordnung 4

Das Butcher-Array für das implizite Euler-Verfahren führt zu

$$y_{n+1} = y_n + hf(x_n+h, g_1)$$

oder

$$y_{n+1} = y_n + hf(x_n+h, y_{n+1}).$$

$$g_1 = y_n + hf(x_n+h, g_1)$$

Das Butcher-Array für die Mittelpunktfornel führt zu

$$y_{n+1} = y_n + hf(x_n+h/2, g_1)$$

oder

$$y_{n+1} = y_n + hf(x_n+h/2, (y_n + y_{n+1})/2).$$

$$g_1 = y_n + (h/2)f(t_n+h/2, g_1)$$

Numerische Behandlung von gewöhnlichen Differentialgleichungen

Eine Einführung

Universität Hamburg
SoSe07

K. Taubert

Numerik gewöhnlicher Anfangswertaufgaben

3 SCHRITTWEITENSTEUERUNGEN FÜR RUNGE-KUTTA-VERFAHREN

Die numerische Integration eines Anfangswertproblems

$$y' = f(x,y), \quad y(x_0) = y_0,$$

mit einem Einschrittverfahren

$$y_{n+1} = y_n + h \Phi(x_n, y_n; h, f)$$

liefert eine Folge von Punkten

$$(x_0, y_0), (x_1, y_1), \dots, (x_n, y_n),$$

wobei die y_i mehr oder weniger gute Näherungen von $y(x_i)$ darstellen.

Dass der Fehler $y_i - y(x_i)$ sehr deutlich von der verwendeten Schrittweite abhängt, zeigt schon die einfache Anfangswertaufgabe

$$y' = -y, \quad y(0) = 1,$$

und das explizite Euler-Verfahren:

- Die Schrittweite $h = 3$ liefert eine Folge (y_i) mit wachsenden $|y_i|$ und alternierenden Vorzeichen,
- die Schrittweite $h = 2$ eine Folge (y_i) mit $|y_i| = 1$ und alternierenden Vorzeichen,
- die Schrittweite $h = 1$ die Folge $(1, 0, 0, \dots)$ und
- eine Schrittweite $h < 1$ eine monoton fallende Folge (y_i) .
- Wird noch berücksichtigt, dass bei kleiner werdenden Schrittweiten der durch Rundungsfehler verursachte Fehler steigen kann, wird die Situation noch komplizierter.

Dieses zeigt, dass Verfahren zur Kontrolle des Fehlers $y_n - y(x_n)$ und Verfahren zur Bestimmung von geeigneten Schrittweiten nötig sind.

Der Wunsch, die Lösung einer Anfangswertaufgabe mit einer vorgegebenen Genauigkeit zu bestimmen und dafür nur so viele Schritte durchzuführen, wie unbedingt nötig sind, führt

zu einer Schrittweitensteuerung, bei der die Schrittweite h des RKV dann dem (aktuellen) Verhalten der Lösungen angepasst wird. Zu beachten ist allerdings auch, dass eine Schrittweitensteuerung stets mit einem gewissen „Overhead“ verbunden ist.

Da der globale Fehler

$$y(x_n) - y_n$$

i.A. schwer zugänglich ist, wird bei den Schrittweitensteuerungen für Einschrittverfahren mit dem lokalen Fehler

$$E(x+h) = y(x+h) - y(x) - h \Phi(x, y(x); h, f)$$

an den Stellen $x+h$ gearbeitet. Hierbei ist $y(x+h)$ die exakte Lösung des Anfangswertproblems, die an der Stelle x den Wert $y(x)$ hat und die wir hier als eindeutig bestimmt voraussetzen.

Unter der Annahme, dass die Lösung des Anfangswertproblems hinreichend oft differenzierbar ist, kann der lokale Fehler an der Stelle x_{n+1} in der Form

$$E(x_{n+1}) = \Psi(x_n, y(x_n))h^{p+1} + O(h^{p+2})$$

dargestellt werden, falls das Einschrittverfahren von der Ordnung p ist.

In der englischen Literatur wird $\Psi(x, y)$ als „principal error function“ und $\Psi(x_n, y(x_n))h^{p+1}$ als „principal local truncation error“ bezeichnet.

Wir geben drei Arten von Schrittweitensteuerungen an:

- Schrittweitensteuerungen, die auf einer direkten Abschätzung von $|\Psi(x, y)|$ beruhen
- Schrittweitensteuerungen durch Verwendung eines RKV, aber mit verschiedenen Schrittweiten (üblicherweise $2h$ und h)
- Schrittweitensteuerungen mit Paaren von RKV von unterschiedlicher Ordnung, aber gleicher Stufenzahl s . Um die Anzahl der Funktionsauswertungen klein zu halten, werden zusätzlich solche Paare benutzt, welche die gleichen Steigungen Y_i' verwenden.

§ 3.1 Das Verfahren von Lotkin

Wird einmal von Fragen der Definitionsbereiche abgesehen und eine hinreichende Regularität der Funktion f vorausgesetzt, dann kann mittels Taylorentwicklung und Koeffizientenvergleich eine Darstellung

$$\Psi(x, y) = D(f, f_x, f_y, f_{xx}, f_{xy}, f_{yy}, \dots)(x, y)$$

für die „principal error function“ $\Psi(x, y)$ gefunden werden.

Abschätzungen für $\Psi(x, y)$ können für die Bestimmung geeigneter Schrittweiten benutzt werden. Da sie bei der rechnergestützten Integration keine große Rolle mehr spielen, gehen wir hier nur kurz und exemplarisch auf das Verfahren von Lotkin ein und zwar im Zusammenhang mit dem Euler-Cauchy-Verfahren.

Das Euler-Cauchy-Verfahren hat die Gestalt

$$y_{n+1} = y_n + h \Phi(x_n, y_n; h, f)$$

mit
$$\Phi(x, y; h, f) = (1/2)f(x, y) + (1/2)f(x+h, y+hf(x, y)).$$

Die Taylorentwicklung von $y(x+h) - y(x)$ liefert

$$y(x+h) - y(x) = y(x) + y'(x)h + (1/2)y''(x)h^2 + (1/6)y'''(x)h^3 + O(h^4) - y(x).$$

Als Taylorentwicklung von $\Phi(x, y; h, f)$ ergibt sich

$$\begin{aligned} \Phi(x, y; h, f) &= f(x, y) + (1/2)(f_x(x, y)h + f_y(x, y)hf(x, y)) \\ &\quad + (1/4)(f_{xx}(x, y)h^2 + 2f_{xy}(x, y)h^2f(x, y) + f_{yy}(x, y)h^2f^2(x, y)) + O(h^3). \\ &= f + (1/2)h(f_x + f_yf) + (1/4)h^2(f_{xx} + 2f_{xy}f + f_{yy}f^2) + O(h^3) \end{aligned}$$

Die Differentialgleichung $y'(x) = f(x, y(x))$ liefert außerdem

$$y''(x) = f_x(x, y(x)) + f_y(x, y(x))y'(x) = f_x(x, y(x)) + f_y(x, y(x))f(x, y(x)) = f_x + f_yf$$

$$\begin{aligned} y'''(x) &= f_{xx}(x, y(x)) + f_{xy}(x, y(x))f(x, y(x)) + (f_{yx}(x, y(x)) + f_{yy}(x, y(x))y'(x))y'(x) \\ &\quad + f_y(x, y(x))y''(x) \end{aligned}$$

$$= f_{xx} + f_{xy}f + (f_{yx} + f_{yy}f)f + f_y(f_x + f_yf) = f_{xx} + 2f_{xy}f + f_{yy}f^2 + f_xf_y + (f_y)^2f.$$

Ein Koeffizientenvergleich in den Potenzen von h ergibt

$$E(x+h) = y(x+h) - y(x) - h \Phi(x, y; h, f) = \Psi(x, y)h^3 + O(h^4)$$

mit
$$\Psi(x, y) = (1/4 - 1/6)(f_{xx} + 2f_{xy}f + f_{yy}f^2) - (1/6)(f_xf_y + (f_y)^2f).$$

Mit den Abkürzungen

$$F = f_x + f_yf \quad \text{und} \quad G = f_{xx} + 2f_{xy}f + f_{yy}f^2$$

ist
$$\Psi(x, y) = (1/12)G - (1/6)Ff_y.$$

Das Verfahren von Lotkin geht nun davon aus, dass man zwei Schranken P, Q findet mit

$$|f(x, y)| < Q$$

und
$$\left| \frac{\partial^{i+j}f(x, y)}{\partial x^i \partial x^j} \right| < \frac{P^{i+j}}{Q^{j-1}} \quad \text{mit } i+j \leq 2 \text{ (Die Ordnung des Verfahrens).}$$

Man hat dann

$$\begin{aligned} |f_y| &< P/1 = P \\ |F| &= |f_x + ff_y| < P/Q^{-1} + Q(P/1) = 2PQ \\ |G| &= |f_{xx} + 2ff_{xy} + f^2f_{yy}| < P^2/Q^{-1} + 2QP^2/1 + Q^2(P^2/Q) = 4P^2Q \end{aligned}$$

und damit

$$|\Psi(x_n, y(x_n))h^3| < (2/3)h^3P^2Q$$

oder eine Abschätzung für den lokalen Fehler $E(x_{n+1})$.

Für das klassische RKV der Ordnung 4 ergibt sich analog

$$|\Psi(x_n, y(x_n))h^5| < (73/720)h^5P^4Q.$$

Beispiel

Die Anfangswertaufgabe $y' = e^{10(x-y)}$, $y(0) = 0.1$, soll auf dem Intervall $[0,1]$ mit dem klassischen Runge-Kutta-Verfahren integriert werden. Der lokale Fehler soll kleiner als 10^{-8} ausfallen.

Es ist:

$$f(x,y) = e^{10(x-y)} \quad \text{und} \quad \frac{\partial^{i+j}f}{\partial x^i \partial y^j} = (-1)^j 10^{i+j} e^{10(x-y)} \quad i+j \leq 4.$$

Gelingt es ein Q zu finden mit $|f(x,y)| < Q < 1$, dann gilt

- $|f_x| < 10Q$, $|f_y| < 10Q$,
- $|f_{xx}| < 10^2Q$, $|f_{xy}| < 10^2Q$, $|f_{yy}| < 10^2Q$,
-
- $|f_{xxxx}| < 10^4Q$,, $|f_{yyyy}| < 10^4Q$

und es kann $P = 10$ gewählt werden.

Wird nun $Q = f(0,0.1) = e^{-1} \approx 0.368$ gewählt, und davon ausgegangen, dass sich $f(x,y)$ während der Integration wenig ändert (nur dadurch ist die Wahl von Q gerechtfertigt), so erhalten wir für den lokalen Fehler die Abschätzung

$$|E(x+h)| < (73/720) \cdot 10^4 \cdot 0.368 \cdot h^5$$

Die Forderung, dass der lokale Fehler kleiner als 10^{-8} sein soll, führt damit auf

$$h < 0.8 \cdot 10^{-2}.$$

Bei dem Verfahren von Lotkin wird die Schrittweite einmalig festgelegt. Die anschließenden Verfahren sind so genannte adaptive Verfahren. Die Schrittweite wird jetzt dem aktuellen Verhalten der Lösung der Anfangswertaufgabe angepasst.

§ 3.2 Schrittweitensteuerung mit einfacher und doppelter Schrittweite (Richardson-Extrapolation)

Die folgende Schrittweitensteuerung basiert auf der Verwendung von nur einem RKV

$$y_{n+1} = y_n + h \Phi(x_n, y_n; h, f),$$

mit verschiedenen Schrittweiten (typischerweise h und $2h$).

Ist das RKV ein Verfahren der Ordnung p und wird von x_n nach x_{n+1} mit der Schrittweite h numerisch integriert, dann gilt für den lokalen Fehler

$$E(x_{n+1}) = y(x_{n+1}) - y_{n+1} = \Psi(x_n, y_n)h^{p+1} + O(h^{p+2}).$$

Wird von x_{n-1} nach x_{n+1} mit der Schrittweite $2h$ mit dem Ergebnis y_{n+1}^* integriert, dann gilt für den lokalen Fehler

$$\begin{aligned} E^*(x_{n+1}) &= y(x_{n+1}) - y_{n+1}^* = \Psi(x_{n-1}, y_{n-1})(2h)^{p+1} + O(h^{p+2}) \\ &= \Psi(x_n, y_n)(2h)^{p+1} + O(h^{p+2}), \end{aligned}$$

wobei auch $\Psi(x_{n-1}, y_{n-1})$ bezüglich (x_n, y_n) nach Taylor entwickelt wurde. Werden die beiden Gleichungen voneinander abgezogen, dann ergibt sich

$$(2^{p+1}-1)h^{p+1} \Psi(x_n, y_n) = y_{n+1} - y_{n+1}^* + O(h^{p+2})$$

oder

$$y(x_{n+1}) - y_{n+1} \approx \frac{y_{n+1} - y_{n+1}^*}{(2^{p+1} - 1)} = E.$$

Ist $|E|$ kleiner als eine vorgegebene Toleranz ε , geht das Schema zum nächsten Schritt über, andernfalls wird die Abschätzung mit der halben Schrittweite wiederholt. Im Allgemeinen wird die Schrittweite verdoppelt, falls $|E|$ kleiner als $\varepsilon/2^{p+1}$ ist.

Dieses Verfahren zur Schrittweitensteuerung ist mit einem erheblichen Mehraufwand verbunden. Allerdings ist die Methode nicht auf Einschrittverfahren beschränkt.

§ 3.3 Schrittweitensteuerung mit Paaren von RKV

Die nun vorgestellte Schrittweitensteuerung arbeitet mit zwei RKV

$$\begin{aligned} y_{n+1} &= y_n + h \Phi(x_n, y_n; h, f), \\ y_{n+1}^* &= y_n + h \Phi^*(x_n, y_n; h, f). \end{aligned}$$

Dabei sei durch Φ ein Verfahren der Ordnung p und durch Φ^* ein Verfahren mindestens der Ordnung $p+1$ gegeben

Der Ausdruck $|y_{n+1} - y_{n+1}^*|$ wird als Näherung von $|y_{n+1} - y(x_{n+1})|$ angesehen und für die Akzeptanz oder Ablehnung der aktuellen Schrittweite h verwendet.

Bei einer vorgegebenen Toleranz ε wird die Schrittweite h akzeptiert, falls

$$|y_{n+1} - y_{n+1}^*| \leq h\varepsilon.$$

Ist f eine hinreichend oft differenzierbare Funktion, dann ist für eine beliebige Schrittweite h :

1. $y(x_{n+1}) - y_n - h \Phi(x_n, y_n; h, f) = \alpha_{n+1} h^{p+1} + O(h^{p+2})$
2. $y(x_{n+1}) - y_n - h \Phi^*(x_n, y_n; h, f) = O(h^{p+2})$

3. $y_{n+1} = y_n + h \Phi(x_n, y_n; h, f)$
4. $y_{n+1}^* = y_n + h \Phi^*(x_n, y_n; h, f)$

und damit

$$|y_{n+1} - y_{n+1}^*| = |\alpha_{n+1}| h^{p+1} + O(h^{p+2}).$$

Wird $O(h^{p+2})$ vernachlässigt, dann wäre die optimale Schrittweite $h_n = sh$ jene mit der Eigenschaft

$$|\alpha_{n+1}| h_n^{p+1} = h_n \varepsilon$$

gewesen.

Den Streckungsfaktor s liefern die folgenden Gleichungen

$$h_n \varepsilon = sh \varepsilon = |\alpha_{n+1}| s^{p+1} h^{p+1} = s^{p+1} |y_{n+1} - y_{n+1}^*| \quad \Rightarrow \quad h \varepsilon = s^p |y_{n+1} - y_{n+1}^*|$$

oder

$$s = \sqrt[p]{\frac{h \varepsilon}{|y_{n+1} - y_{n+1}^*|}}.$$

Damit ergibt sich:

Die gewählte Schrittweite war genau dann richtig, wenn $s = 1$ ist.

Für $s > 1$ wäre eine größere Schrittweite möglich gewesen. Im nächsten Schritt wird dann

$$h_{n+1} = \min(2, s)h$$

gewählt.

Für $s < 1$ muss der Schritt mit einer kleineren Schrittweite wiederholt werden. Man wählt z.B.

$$h_n = \max(1/2, s)h.$$

Bemerkung

In der Praxis kann es zweckmäßig sein, einen von $s = 1$ etwas abweichenden Wert zu wählen.

Einbettungsformeln

Es gibt Paare von RKV, bei denen die Berechnung von Φ^* zurückgeführt werden kann auf die Daten, die bereits bei der Berechnung Φ benötigt wurden. Beide Verfahren benötigen also dieselben Steigungen, jedoch nur mit unterschiedlichen Gewichten b .

Ein bekanntes Beispiel stammt von Fehlberg. Es ist ein Verfahren mit RKV der Ordnungen 4 und 5 und wird häufig als RKV45 bezeichnet.

0		
$\frac{1}{4}$	$\frac{1}{4}$	
$\frac{3}{8}$	$\frac{3}{32}$	$\frac{9}{32}$

12	$\frac{1932}{2197}$	$-\frac{7200}{2197}$	$\frac{7296}{2197}$			
13	$\frac{439}{216}$	-8	$\frac{3680}{513}$	$-\frac{845}{4104}$		
1	$\frac{1}{27}$	2	$-\frac{3544}{2565}$	$-\frac{1859}{4104}$	$-\frac{11}{40}$	
2	$\frac{25}{216}$	0	$\frac{1408}{2565}$	$\frac{2197}{4104}$	$-\frac{1}{5}$	0
y	$\frac{16}{135}$	0	$\frac{6656}{12825}$	$\frac{28561}{56430}$	$-\frac{9}{50}$	$\frac{2}{55}$
y*						

Bekannt und häufig benutzt werden auch die eingebetteten Paare von Dormand&Price 5(4) (DOPRI5) und Dormand&Price 8(6) (DOPRI8).

Numerische Behandlung von gewöhnlichen Differentialgleichungen Eine Einführung

Universität Hamburg
SoSe07

K. Taubert

Numerik gewöhnlicher Anfangswertaufgaben

4 LINEARE MEHRSCHRITTVERFAHREN

Es werden allgemeine Mehrschrittverfahren (MSV) definiert und insbesondere die Adams- und BDF-Verfahren bis zur Ordnung 6 angegeben. Das Konstruktionsprinzip der Adams- und BDF-Verfahren wird erläutert. Hinreichende Bedingungen für die Auflösung impliziter MSV werden angegeben und die damit zusammenhängenden Prädiktor-Korrektor- und Rosenbrock-Wanner-Verfahren eingeführt.

Es werden die Begriffe Konsistenz, Ordnung und (starke) Stabilität für lineare Mehrschrittverfahren eingeführt. Es werden zwei Konvergenzsätze angegeben und Grenzen der numerischen Methoden durch Beispiele belegt.

§ 4.1 Allgemeine lineare Mehrschrittverfahren

Für die numerische Integration von gewöhnlichen Anfangswertaufgaben

$$y' = f(t, y), \quad y(t_0) = y_0,$$

gibt es neben den RKV und den Extrapolationsverfahren auch noch die linearen Mehrschrittverfahren oder lineare k -Schrittverfahren

$$a_k y_{n+k} + \dots + a_0 y_n = h(b_k f_{n+k} + \dots + b_0 f_n),$$

$$|a_0| + |b_0| \neq 0, \quad a_k \neq 0$$

$$a_i, b_i \in \mathbf{R}, \quad i = 0, 1, 2, \dots, k, \quad f_n = f(t_n, y_n), \quad t_n = t_0 + nh.$$

Im Falle $b_k \neq 0$ spricht man von einem impliziten und anderenfalls von einem expliziten Mehrschrittverfahren.

Bereits bekannte Beispiele sind:

- | | | |
|---|---------|----------|
| 1. $y_{n+1} - y_n = hf_n$ | $k = 1$ | explizit |
| 2. $y_{n+1} - y_n = hf_{n+1}$ | $k = 1$ | implizit |
| 3. $y_{n+1} - y_n = h((1/2)f_n + (1/2)f_{n+1})$ | $k = 1$ | implizit |
| 4. $y_{n+2} - y_n = hf_{n+1}$ | $k = 2$ | explizit |

Bisher nicht aufgetretene Beispiele für MSV sind

$$5. \quad y_{n+2} - \frac{4}{3}y_{n+1} + \frac{1}{3}y_n = h \frac{2}{3} f_{n+2} \quad k=2 \quad \text{implizit}$$

$$6. \quad y_{n+4} - y_{n+3} = h \left(\frac{55}{24} f_{n+3} - \frac{59}{24} f_{n+2} + \frac{37}{24} f_{n+1} - \frac{9}{24} f_n \right) f_n \quad k=4 \quad \text{explizit}$$

In der Menge der linearen k-Schrittverfahren gibt es zwei besondere Klassen von Verfahren: die Adams- und die BDF- oder Gear-Verfahren.

Die Adams-Verfahren

Die **Adams-Bashforth-Verfahren** sind explizite lineare k-Schrittverfahren ($k = 1, 2, 3, \dots$) der speziellen Form

$$y_n - y_{n-1} = h \sum_{i=1}^k b_{ki} f_{n-i}.$$

(Man beachte: Es wurde eine zweckmäßige Verschiebung der Indizes $n+k \rightarrow n$ vorgenommen.)

Die Koeffizienten b_{ki} für $k = 1, 2, 3, \dots, 6$ können aus der folgenden Tabelle entnommen werden:

i	1	2	3	4	5	6	Ordnung
b_{1i}	1						1
$2b_{2i}$	3	-1					2
$12b_{3i}$	23	-16	5				3
$24b_{4i}$	55	-59	37	-9			4
$720b_{5i}$	1901	-2774	2616	-1274	251		5
$1440 b_{6i}$	4277	-7923	9982	-7298	2877	-475	6

Beispiele

Für $k=1$ liegt das explizite Euler-Verfahren vor. Für $k=2$ liefert die Tabelle

$$y_n - y_{n-1} = h(b_{21}f_{n-1} + b_{22} f_{n-2}) = h \left(\frac{3}{2} f_{n-1} - \frac{1}{2} f_{n-2} \right)$$

oder in der ursprünglichen Indizierung

$$y_{n+2} - y_{n+1} = h \left(\frac{3}{2} f_{n+1} - \frac{1}{2} f_n \right).$$

Das Konstruktionsprinzip für die Adams-Bashforth-Verfahren kann aus den folgenden Ausführungen entnommen werden.

Eine Lösung der Differentialgleichung

$$y' = f(t, y)$$

genügt der Gleichung

$$y(t_n) - y(t_{n-1}) = \int_{t_{n-1}}^{t_n} y'(s) ds = \int_{t_{n-1}}^{t_n} f(s, y(s)) ds.$$

Die Funktion $F(s) = f(s, y(s))$ wird hier durch ein (Interpolations-) Polynom $P_k^B(s)$ $(k-1)$ -ten Grades ersetzt, das an den Stellen

$$t_{n-1}, \quad t_{n-2}, \quad t_{n-3}, \quad \dots, \quad t_{n-k}$$

die Werte $F(t_{n-1}), F(t_{n-2}), F(t_{n-3}), \dots, F(t_{n-k})$ hat.

Damit gilt

$$y(t_n) - y(t_{n-1}) \cong \int_{t_{n-1}}^{t_n} P_k^B(s) ds.$$

Die Ausführung der Interpolation und Integration führt auf die angegebenen Differenzenverfahren bzw. auf die Koeffizienten der Tabelle.

Beispiele

Für $k=1$ ist $F(s) = f(s, y(s))$ zu ersetzen durch $P_1^B(s) = f(t_{n-1}, y(t_{n-1}))$. Damit ist

$$y(t_n) - y(t_{n-1}) \cong \int_{t_{n-1}}^{t_n} f(t_{n-1}, y(t_{n-1})) ds = h f(t_{n-1}, y(t_{n-1}))$$

und es ergibt sich erneut das bekannte Euler-Verfahren

$$y_n - y_{n-1} = h f_{n-1}.$$

Das Polynom $(2-1)$ -ten Grades (gerade), das an der Stelle t_{n-1} den Wert $F(t_{n-1})$ und an der Stelle t_{n-2} den Wert $F(t_{n-2})$ hat, ist:

$$P_2^B(s) = \frac{s - t_{n-2}}{h} F(t_{n-1}) - \frac{s - t_{n-1}}{h} F(t_{n-2}).$$

Damit ist

$$\begin{aligned} y(t_n) - y(t_{n-1}) &\cong \int_{t_{n-1}}^{t_n} P_2^B(s) ds = \frac{F(t_{n-1})}{h} \int_{t_{n-1}}^{t_n} (s - t_{n-2}) ds - \frac{F(t_{n-2})}{h} \int_{t_{n-1}}^{t_n} (s - t_{n-1}) ds \\ &= \frac{F(t_{n-1})}{h} \left(\frac{1}{2} (t_n - t_{n-2})^2 - \frac{1}{2} (t_{n-1} - t_{n-2})^2 \right) - \frac{F(t_{n-2})}{h} \frac{1}{2} (t_n - t_{n-1})^2 \\ &= \frac{F(t_{n-1})}{h} \left(\frac{1}{2} (2h)^2 - \frac{1}{2} h^2 \right) - \frac{F(t_{n-2})}{h} \frac{1}{2} h^2 \\ &= h \left(\frac{3}{2} F(t_{n-1}) - \frac{1}{2} F(t_{n-2}) \right) = h \left(\frac{3}{2} f(t_{n-1}, y(t_{n-1})) - \frac{1}{2} f(t_{n-2}, y(t_{n-2})) \right). \end{aligned}$$

Dieses führt also für $k = 2$ auf das Verfahren

$$y_n - y_{n-1} = h \left(\frac{3}{2} f_{n-1} - \frac{1}{2} f_{n-2} \right).$$

Die **Adams-Moulton-Verfahren** sind implizite lineare k-Schrittverfahren ($k = 1, 2, 3, \dots$) der speziellen Form

$$y_n - y_{n-1} = h \sum_{i=0}^{k-1} b_{ki}^* f_{n-i}.$$

Die Koeffizienten b_{ki}^* für $k = 1, 2, 3, \dots, 6$ können aus der folgenden Tabelle entnommen werden:

i	0	1	2	3	4	5	Ordnung
b_{1i}^*	1						1
$2b_{2i}^*$	1	1					2
$12b_{3i}^*$	5	8	-1				3
$24b_{4i}^*$	9	19	-5	1			4
$720b_{5i}^*$	251	646	-264	106	-19		5
$1440b_{6i}^*$	475	1427	-798	482	-173	27	6

Beispiele

$k = 1$

$$y_n - y_{n-1} = hf_n$$

$k = 2$

$$y_n - y_{n-1} = h \left(\frac{1}{2} f_n + \frac{1}{2} f_{n-1} \right)$$

$k = 3$

$$y_n - y_{n-1} = h \left(\frac{5}{12} f_n + \frac{8}{12} f_{n-1} - \frac{1}{12} f_{n-2} \right)$$

usw.

Das Konstruktionsprinzip für die Adams-Moulton-Verfahren ist ähnlich dem der Adams-Bashforth-Verfahren:

In der Gleichung

$$y(t_n) - y(t_{n-1}) = \int_{t_{n-1}}^{t_n} f(s, y(s)) ds$$

wird $F(s) = f(s, y(s))$ erneut durch ein (Interpolations-) Polynom $P_k^M(s)$ ($k-1$)-ten Grades ersetzt. Im Gegensatz zu den Adams-Bashforth Verfahren werden als Stützstellen jetzt die um einen Index erhöhten Stellen gewählt, d.h.

$$(t_n, F(t_n)), (t_{n-1}, F(t_{n-1})), (t_{n-2}, F(t_{n-2})), \dots, (t_{n-k+1}, F(t_{n-k+1})).$$

Beispiel

Für $k = 1$ ist $F(s) = f(s, y(s))$ zu ersetzen durch $P_1^M(s) = f(t_n, y(t_n))$. Damit ergibt sich

$$y(t_n) - y(t_{n-1}) \cong \int_{t_{n-1}}^{t_n} f(t_n, y(t_n)) ds = h f(t_n, y(t_n))$$

und somit das implizite Euler-Verfahren

$$y_n - y_{n-1} = hf_n.$$

Aufgabe

Verifizieren Sie den Fall $k = 2$.

Die BDF- oder Gear-Verfahren

Die BDF-Verfahren sind implizite lineare k -Schrittverfahren ($k = 1, 2, 3, \dots, 6$) der speziellen Form

$$y_n = a_1 y_{n-1} + a_2 y_{n-2} + \dots + a_k y_{n-k} + h b_0 f_n.$$

Die Koeffizienten a_i, b_0 können aus der folgenden Tabelle entnommen werden:

k	1	2	3	4	5	6
b_0	1	2/3	6/11	12/25	60/137	60/147
a_1	1	4/3	18/11	48/25	300/137	360/147
a_2		-1/3	-9/11	-36/25	-300/137	-450/147
a_3			2/11	16/25	200/137	400/147
a_4				-3/25	-75/137	-225/147
a_5					12/137	72/147
a_6						-10/147
Ordnung	1	2	3	4	5	6

Zur Konstruktion der BDF-Verfahren wird die linke Seite der Differentialgleichung

$$y' = f(t, y)$$

durch die Ableitung geeigneter Interpolationspolynome ersetzt.

Für jedes k wird $y(t)$ zunächst durch ein Interpolationspolynom $P_k^G(t)$ k -ten Grades mit den Stützstellen

$$(t_n, y(t_n)), (t_{n-1}, y(t_{n-1})), (t_{n-2}, y(t_{n-2})), \dots, (t_{n-k}, y(t_{n-k}))$$

approximiert. Anschließend wird die Ableitung dieses Polynoms gebildet und $y'(t_n)$ durch $P_k^G'(t_n)$ ersetzt.

Beispiel

Das Interpolationspolynom ersten Grades ($k=1$) an den Stellen t_n, t_{n-1} ist

$$P_1^G(t) = \frac{t - t_{n-1}}{t_n - t_{n-1}} y(t_n) - \frac{t - t_n}{t_n - t_{n-1}} y(t_{n-1})$$

Die Ableitung ist (unabhängig von t)

$$P_1^G'(t) = \frac{1}{t_n - t_{n-1}} y(t_n) - \frac{1}{t_n - t_{n-1}} y(t_{n-1}) = \frac{1}{h} y(t_n) - \frac{1}{h} y(t_{n-1}).$$

Damit ergibt sich

$$\frac{1}{h} y(t_n) - \frac{1}{h} y(t_{n-1}) \cong f(t_n, y(t_n))$$

oder erneut das implizite Euler-Verfahren

$$y_n - y_{n-1} = hf_n.$$

Aufgabe

Verifizieren Sie die Formel für $k = 2$.

§ 4.2 Lösbarkeit impliziter linearer k-Schrittverfahren Prädiktor-Korrektor-Verfahren und Rosenbrock-Wanner- Methoden

Gegeben sei ein allgemeines lineares implizites k-Schrittverfahren

$$a_k y_{n+k} + \dots + a_0 y_n = h(b_k f_{n+k} + \dots + b_0 f_n),$$

$$|a_0| + |b_0| \neq 0, \quad a_k \neq 0 \quad \text{und} \quad b_k \neq 0.$$

Sind die Werte $y_n, y_{n+1}, \dots, y_{n+k-1}$ und $f(t_n, y_n), f(t_{n+1}, y_{n+1}), \dots, f(t_{n+k-1}, y_{n+k-1})$ bekannt, dann muss für die Ermittlung von y_{n+k} die Fixpunktgleichung

$$y^* = h \frac{b_k}{a_k} f(t_{n+k}, y^*) + \text{Rest}_{\text{bekannt}} \quad (*),$$

$$y_{n+k} = y^*,$$

gelöst werden.

Für den Fall, dass f nichtlinear in der zweiten Komponente ist, können zwei hinreichende Bedingungen für die eindeutige Lösbarkeit der Fixpunktaufgabe angegeben werden:

Satz 4.1

Es sei f eine stetige Funktion $f: \mathbf{R}^* \mathbf{R} \rightarrow \mathbf{R}$, die außerdem eine Lipschitzbedingung

$$|f(t, u) - f(t, v)| \leq K |u - v|$$

für alle $t, u, v \in \mathbf{R}$ erfüllt.

Dann gibt es ein $H > 0$, so dass für alle $0 < h \leq H$ und für beliebige $y_n, y_{n+1}, \dots, y_{n+k-1}$ die Fixpunktaufgabe (*) für y_{n+k} eindeutig lösbar ist.

Beweis

Betrachte die Iteration

$$z^{v+1} = h \frac{b_k}{a_k} f(t_{n+k}, z^v) + \text{Rest}_{\text{bekannt}}$$

mit einem beliebigen $z^0 \in \mathbf{R}$ und $v = 0, 1, 2, 3, \dots$

Dann gilt

$$|z^{v+1} - z^v| \leq h \left| \frac{b_k}{a_k} \right| K |z^v - z^{v-1}|.$$

Wird H so gewählt, dass $0 < L := H \left| \frac{b_k}{a_k} \right| K < 1$ ist, dann gilt für alle $0 < h \leq H$ und $p \geq 1$

$$\begin{aligned} |z^{v+p} - z^v| &\leq |z^{v+p} - z^{v+p-1}| + \dots + |z^{v+2} - z^{v+1}| + |z^{v+1} - z^v| \\ &\leq L^{v+p} |z^1 - z^0| + \dots + L^{v+1} |z^1 - z^0| \\ &\leq L^v (L^p + L^{p-1} + \dots + L + 1) |z^1 - z^0| \end{aligned}$$

$$\leq \frac{L^v}{1-L} |z^1 - z^0|.$$

Wegen $0 < L < 1$, ist (z^v) eine Cauchy-Folge und es gibt ein z^* mit $z^v \rightarrow z^*$. Wegen der Stetigkeit von f gilt

$$z^* = h \frac{b_k}{a_k} f(t_{n+k}, z^*) + \text{Rest}_{\text{bekannt}}.$$

Einen weiteren Fixpunkt z^{**} kann es wegen

$$|z^{**} - z^*| \leq L |z^{**} - z^*|$$

und $0 < L < 1$ nicht geben.

Bemerkung

Der Beweis zeigt:

Es gibt nicht nur eine eindeutige Lösung der Fixpunktaufgabe, sondern – sofern h so gewählt

wird, dass $0 < h \leq H \left| \frac{b_k}{a_k} \right| K < 1$ ist – sogar eine iterative Vorschrift

$$z^{v+1} = h \frac{b_k}{a_k} f(t_{n+k}, z^v) + \text{Rest}_{\text{bekannt}},$$

um diesen zu bestimmen.

Satz 4.2

Es sei f eine stetige Funktion $f: \mathbf{R}^* \mathbf{R} \rightarrow \mathbf{R}$, die außerdem monoton fallend ist, d.h.

$$(f(t,u) - f(t,v))(u-v) \leq 0$$

für alle $t, u, v \in \mathbf{R}$.

Ist $b_k/a_k > 0$, dann ist für beliebige $y_n, y_{n+1}, \dots, y_{n+k-1}$ und alle $h > 0$ die Fixpunktaufgabe (*) eindeutig lösbar.

Beweis

Trivial.

Bei impliziten Differenzenverfahren wird es nur in Ausnahmefällen möglich sein, die Werte y_{n+k} exakt zu ermitteln. Man wird sich mit Näherungswerten begnügen müssen und es liegt mithin ein anderes Verfahren vor. Unterschiedliche Methoden zur Ermittlung von Näherungswerten von y_{n+k} führen also auch zu anderen Verfahren.

Prädiktor-Korrektor-Verfahren

Das implizite Adams-Moulton-Verfahren k -ter Ordnung wird im Allgemeinen mit einem Adams-Bashforth-Verfahren $(k-1)$ -ter Ordnung in folgender Weise kombiniert:

$$\left(\begin{array}{l} \text{Adams} \\ \text{Bashforth} \end{array} \right) \quad y_n^{(0)} = y_{n-1} + h \sum_{i=1}^{k-1} b_{k-1,i} f_{n-i} \quad (\text{P}) \text{ rädiktor}$$

$$f_n^{(0)} = f(t_n, y_n^{(0)}) \quad (\text{E}) \text{ valuator}$$

$$\begin{pmatrix} \text{Adams} \\ \text{Moulton} \end{pmatrix} \quad y_n = y_{n-1} + h \left(b_{k0}^* f_n^{(0)} + \sum_{i=1}^{k-1} b_{ki}^* f_{n-i} \right) \quad \text{(C) orrector}$$

$$f_n = f(t_n, y_n) \quad \text{(E) vualation}$$

Diese Art von Verfahren (Prädiktor-Korrektor) wird als PECE-Verfahren bezeichnet. Selbstverständlich sind auch weitere Kombinationen von Adams-Bashforth und Adams-Moulton-Verfahren denkbar, z.B. PECEC oder PECECE usw.

Rosenbrock-Wanner-Methode

Im Zusammenhang mit den BDF-Verfahren

$$y_n = a_1 y_{n-1} + a_2 y_{n-2} + \dots + a_k y_{n-k} + h b_0 f_n$$

werden Näherungen für y_{n+k} häufig wie folgt ermittelt:

$$y_n^{(0)} = a_1 y_{n-1} + a_2 y_{n-2} + \dots + a_k y_{n-k} + h b_0 f(t_n, y_{n-1})$$

und anschließend einige Newton-Schritte

$$y_n^v = y_n^{v-1} - (F')^{-1}(y_n^{v-1}) F(y_n^{v-1})$$

mit

$$F(y) = y - (a_1 y_{n-1} + a_2 y_{n-2} + \dots + a_k y_{n-k} + h b_0 f(t_n, y))$$

durchgeführt.

Schließlich wird $y_n = y_n^p$ gesetzt.

Bemerkung

Manchmal wird auch das vereinfachte Newton-Verfahren gewählt.

§ 4.3 Konsistenz, Ordnung, Stabilität und Konvergenz

Für die gewöhnlichen Anfangswertaufgaben

$$y' = f(t, y), \quad y(t_0) = y_0,$$

betrachten wir wieder die linearen Mehrschrittverfahren oder linearen k-Schrittverfahren

$$a_k y_{n+k} + \dots + a_0 y_n = h(b_k f_{n+k} + \dots + b_0 f_n),$$

$$|a_0| + |b_0| \neq 0, \quad a_k \neq 0$$

$$a_i, b_i \in \mathbf{R}, \quad i = 0, 1, 2, \dots, k, \quad f_n = f(t_n, y_n), \quad t_n = t_0 + nh.$$

Für eine stetig differenzierbare Funktion $y(\cdot)$ sei nun

$$L[y;h](t) := a_k y(t+kh) + \dots + a_0 y(t) - h(b_k y'(t+kh) + \dots + b_0 y'(t)).$$

Ist $y(\cdot)$ hinreichend oft differenzierbar, dann kann für $L[y;h](t)$ die folgende (Taylor-Entwicklung) Darstellung angegeben werden:

$$\begin{aligned}
L[y;h](t) = & a_k [y(t) + kh y'(t) + \dots + \frac{(kh)^p}{p!} y^{(p)}(t)] + \\
& + a_{k-1} [y(t) + (k-1)h y'(t) + \dots + \frac{((k-1)h)^p}{p!} y^{(p)}(t)] + \\
& + \dots \dots \dots + \\
& + a_0 [y(t)] \\
& - b_k h [kh y'(t) + \dots + \frac{(kh)^p}{(p)!} y^{(p-1)}(t)] - \\
& - \dots \dots \dots - \\
& - b_0 h [y'(t)] + O(h^{p+1})
\end{aligned}$$

oder kurz

$$L[y;h](t) = C_0 y(t) + C_1 h y'(t) + \dots + C_p h^p y^{(p)}(t) + O(h^{p+1}).$$

Ein einfacher Koeffizientenvergleich liefert, in Abhängigkeit von a_i und b_i , für die Konstanten C_q

$$C_0 = a_0 + a_1 + a_2 + \dots + a_k$$

$$C_1 = a_1 + 2a_2 + \dots + ka_k - (b_0 + b_1 + b_2 + \dots + b_k)$$

.....

$$C_p = \frac{1}{p!} (a_1 + 2^p a_2 + \dots + k^p a_k) - \frac{1}{(p-1)!} (b_1 + 2^{p-1} b_2 + \dots + k^{p-1} b_k).$$

Definition (Konsistenz, Ordnung)

Ein lineares Mehrschrittverfahren heißt von der Ordnung $q \geq 1$, wenn

$$C_0 = C_1 = C_2 = \dots = C_q = 0$$

und $C_{q+1} \neq 0$.

Das Verfahren heißt konsistent, wenn $C_0 = C_1 = 0$ ist.

Beispiel

Für die Mittelpunkstformel

$$y_{n+2} - y_n = 2hf_{n+1}$$

gilt $C_0 = C_1 = C_2 = 0$ und $C_3 = 1/3$.

Damit ist die Mittelpunkstformel konsistent und von der Ordnung 2.

Dass die Konsistenz oder sogar eine höhere Ordnung eines MSV keine Garantie dafür sind, dass auch ein numerisch brauchbares Verfahren vorliegt, zeigt das folgende einfache

Beispiel

Das Verfahren

$$y_{n+2} + 4y_{n+1} - 5y_n = 4hf_{n+1} + 2hf_n$$

hat die Ordnung 3.

Die Anwendung dieses Verfahrens auf die triviale Anfangswertaufgabe

$$y' = 0, \quad y(0) = y^0,$$

liefert für das Startfeld $y_0 = y_1 = y^0$ die exakte Lösung.

Für das Startfeld

$$y_0 = y^0 + h \quad \text{und} \quad y_1 = y^0$$

ergibt sich jedoch

$$y_n = (y^0 + \frac{5}{6}h) + h \frac{(-5)^n}{6}.$$

Mit $h = \frac{a}{N}$, $N \in \mathbf{N}$, $N \geq 2$ und $n = N$ ist

$$y_N = (y^0 + \frac{5}{6} \frac{a}{N}) + \frac{a}{N} \frac{(-5)^N}{6}.$$

D.h., der erste „Lösungsast“ $y^0 + \frac{5}{6} \frac{a}{N}$ liefert für $N \rightarrow \infty$ den gewünschten Wert der Lösung.

Der zweite Lösungsast $\frac{a}{N} \frac{(-5)^N}{6}$ geht aber für $N \rightarrow \infty$ gegen $\pm \infty$.

Da sich Störungen in den Startfeldern oder Anfangsbedingungen während der Rechnung nicht vermeiden lassen, ist dieses Verfahren der Ordnung 3 numerisch völlig unbrauchbar.

Definition

Das Differenzenverfahren

$$a_k y_{n+k} + \dots + a_0 y_n = h(b_k f_{n+k} + \dots + b_0 f_n)$$

$|a_0| + |b_0| \neq 0$, $a_k \neq 0$, heißt stabil, wenn die Wurzeln des charakteristischen Polynoms

$$p(\xi) = a_k \xi^k + a_{k-1} \xi^{k-1} + \dots + a_0$$

dem Betrage nach kleiner oder gleich 1 sind und jene Wurzeln mit dem Betrag 1 einfach sind.

Beispiele

MSV	$p(\xi)$	Wurzeln	Stabil
$y_{n+1} - y_n = hf_n$	$\xi - 1$	$\xi \equiv 1$	ja
$y_{n+2} - y_n = 2hf_{n+1}$	$\xi^2 - 1$	$\xi_{1,2} = \pm 1$	ja
$y_{n+2} + 4y_{n+1} - 5y_n = 4hf_{n+1} + 2hf_n$	$\xi^2 + 4\xi - 5$	$\xi_1 = 1, \xi_2 = -5$	nein

Satz 4.3

Gegeben sei eine beliebige Anfangswertaufgabe

$$y' = f(t,y), \quad y(t_0) = y_0,$$

mit einer stetigen Funktion $f: \mathbf{R}^* \times \mathbf{R} \rightarrow \mathbf{R}$ und der Eigenschaft

$$|f(t,u) - f(t,v)| \leq K |u - v|$$

für alle $x, u, v \in \mathbf{R}$ mit einem $K \geq 0$. Es sei $y(\cdot)$ die Lösung der Anfangswertaufgabe auf $[t_0, t_0 + T_0]$.

Gegeben sei ein stabiles und konsistentes lineares k -Schrittverfahren

$$a_k y_{n+k} + \dots + a_0 y_n = h(b_k f_{n+k} + \dots + b_0 f_n),$$

$$|a_0| + |b_0| \neq 0, \quad a_k \neq 0,$$

$$a_i, b_i \in \mathbf{R}, \quad i = 0, 1, 2, \dots, k, \quad f_n = f(t_n, y_n), \quad t_n = t_0 + nh.$$

Dann gilt für alle Anfangsfelder $y_{k-1}^h, y_{k-2}^h, \dots, y_0^h$ mit der Eigenschaft

$$\lim_{h \rightarrow 0} y_i^h = y(t_0) \quad i = 0, 1, 2, \dots, k-1$$

und für alle $t \in [t_0, t_0 + T_0]$

$$\lim_{\substack{h \rightarrow 0 \\ n \rightarrow \infty \\ t_0 + nh \rightarrow t}} y_n^h = y(t).$$

Beweis:

Zur Vereinfachung nehmen wir an, dass $b_k = 0$.

Es ist zweckmäßig, das Differenzenverfahren in ein System von Gleichungen umzuschreiben.

$$\begin{pmatrix} y_{n+k} \\ y_{n+k-1} \\ \vdots \\ y_{n+1} \end{pmatrix} = \underbrace{\begin{pmatrix} -\frac{a_{k-1}}{a_k} & -\frac{a_{k-2}}{a_k} & \dots & -\frac{a_0}{a_k} \\ 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 1 & 0 \end{pmatrix}}_{=: A} \begin{pmatrix} y_{n+k-1} \\ y_{n+k-2} \\ \vdots \\ y_n \end{pmatrix} + h \underbrace{\begin{pmatrix} \frac{b_{k-1}}{a_k} & \dots & \frac{b_0}{a_k} \\ 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{pmatrix}}_{=: B} \begin{pmatrix} f_{n+k-1} \\ f_{n+k-2} \\ \vdots \\ f_n \end{pmatrix}.$$

Setzt man die Lösung der Anfangswertaufgabe ein, dann ergibt sich wegen der Konsistenz des Verfahrens

$$\begin{pmatrix} y(t_{n+k}) \\ y(t_{n+k-1}) \\ \vdots \\ y(t_{n+1}) \end{pmatrix} = A \begin{pmatrix} y(t_{n+k-1}) \\ y(t_{n+k-2}) \\ \vdots \\ y(t_n) \end{pmatrix} + hB \begin{pmatrix} f(t_{n+k-1}, y(t_n + (k-1)h)) \\ f(t_{n+k-2}, y(t_n + k-2)h) \\ \vdots \\ f(t_n, y(t_n)) \end{pmatrix} + \begin{pmatrix} o(h) \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Es sei nun

$$F_q = \begin{pmatrix} y(t_q + (k-1)h) \\ \vdots \\ y(t_q) \end{pmatrix} - \begin{pmatrix} y_{q+k-1} \\ \vdots \\ y_q \end{pmatrix}.$$

Für eine beliebige Vektornorm und eine zugeordnete Matrixnorm gilt dann

$$\|F_{n+1}\| \leq \|A\| \|F_n\| + h \|B\| K \|F_n\| + \varepsilon(h)$$

für alle n mit $t_0 + (n+1)h \leq t_0 + T_0$ und ein $\varepsilon(\cdot)$ mit der Eigenschaft $\lim_{h \rightarrow 0} \frac{\varepsilon(h)}{h} = 0$.

Es gibt nun stets eine Vektornorm $\| \cdot \|_H$ und eine zugeordnete Matrixnorm mit $\| A \|_H = 1$.
Hierzu:

Die charakteristischen Zahlen der Matrix A

$$A = \begin{pmatrix} -\frac{a_{k-1}}{a_k} & -\frac{a_{k-2}}{a_k} & \dots & -\frac{a_0}{a_k} \\ 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 1 & 0 \end{pmatrix}$$

sind die Nullstellen von

$$a_k \xi^k + \dots + a_0 = 0.$$

Diese Nullstellen sind aber wegen der Stabilität des Verfahrens dem Betrage nach kleiner oder gleich 1 und jene Wurzeln mit dem Betrag 1 sind einfach.

D.h. es gibt eine Ähnlichkeitstransformation

$$A \rightarrow HAH^{-1}$$

mit $HAH^{-1} =$

$$\begin{pmatrix} \xi_1 & & & & & & & & & \\ & \ddots & & & & & & & & \\ & & \xi_\ell & & & & & & & \\ & & & \ddots & & & & & & \\ & & & & \xi_\mu & & & & & \\ & & 0 & & \varepsilon_\mu & & & & & \\ & & & & & \ddots & & & & \\ & & & & & & \varepsilon_\mu & & & \\ & & & & & & & \xi_\mu & & \\ & & & & & & & & \ddots & \end{pmatrix}$$

←

←

Einfache Nullstellen ξ
mit Betrag ≤ 1

Jordan Kästchen
mit $|\xi| < 1$, $\varepsilon \geq 0$
und $|\xi_\mu| + \varepsilon < 1$

wobei die ξ_i die Nullstellen von $a_k \xi^k + \dots + a_0 = 0$ sind.

Mit der Vektornorm $\| x \|_H := \| Hx \|_\infty$ und der ihr zugeordneten Matrixnorm

$\| A \|_H = \| HAH^{-1} \|_\infty$ ergibt sich mithin

$$\| F_{n+1} \|_H \leq \| F_n \|_H + h \tilde{K} \| F_n \|_H + \tilde{\varepsilon}(h)$$

mit einer geeigneten Konstante \tilde{K} und einen $\tilde{\varepsilon}(\cdot)$ mit der Eigenschaft $\lim_{h \rightarrow 0} \frac{\tilde{\varepsilon}(h)}{h} = 0$.

Mit unserer bereits bekannten Methode liefert die Rekursionsformel dann auch

$$\| F_{n+1} \|_H \leq \| F_0 \|_H e^{h \tilde{K} T_0} + \frac{\tilde{\varepsilon}(h)}{h} \frac{1}{\tilde{K}} (e^{h \tilde{K} T_0} - 1).$$

Wegen der Voraussetzung an die Anfangsfelder ist $\lim_{h \rightarrow 0} \| F_0 \| = 0$ und wegen der Konsistenz

ist $\lim_{h \rightarrow 0} \frac{\tilde{\varepsilon}(h)}{h} = 0$. D.h. das Verfahren ist konvergent.

Bemerkung

Man beachte, dass sogar eine gleichmäßige Konvergenz vorliegt. Es wird dem Leser überlassen auch den Fall $b_k \neq 0$ zu behandeln.

Dass der angegebene Satz nicht ohne Weiteres verallgemeinert werden kann, zeigt das folgende einfache

Beispiel:

Gegeben sei die monotone Anfangswertaufgabe

$$y' = f(t, y), \quad y(0) = 0, \quad t \geq 0,$$

mit

$$f(t, y) = \begin{cases} 0 & t = 0 \\ 2t & y < 0 \\ 2t - 4\frac{y}{t} & 0 \leq y \leq t^2 \\ -2t & y \geq t^2. \end{cases} \quad \text{für}$$

Die Lösung der Anfangswertaufgabe ist $y(t) = (1/3)t^2$.

Die Mittelpunkstformel

$$y_{n+2} - y_n = 2f_{n+1} \quad \text{mit} \quad y_0 = 0 \quad \text{und} \quad y_1 = -h^2$$

liefert ($v = 0, 1, 2, \dots$)

$$y_{2v} = (2vh)^2 \\ y_{2v+1} = ((2v+1)h)^2$$

und mithin keine Konvergenz!

Definition

Das Differenzenverfahren

$$a_k y_{n+k} + \dots + a_0 y_n = h(b_k f_{n+k} + \dots + b_0 f_n)$$

$|a_0| + |b_0| \neq 0$, $a_k \neq 0$, heißt stark stabil, wenn die Wurzeln des charakteristischen Polynoms

$$p(\xi) = a_k \xi^k + a_{k-1} \xi^{k-1} + \dots + a_0 = 0$$

dem Betrage nach kleiner oder gleich 1 sind. Außer der einfachen Wurzel 1 darf es keine weiteren Wurzeln mit dem Betrag 1 geben.

Satz 4.4

Gegeben sei eine beliebige Anfangswertaufgabe

$$y' = f(t, y), \quad y(t_0) = y_0,$$

mit einer stetigen Funktion $f: \mathbf{R}^* \times \mathbf{R} \rightarrow \mathbf{R}$ und der Eigenschaft

$$(f(t,u) - f(t,v))(u - v) \leq K(u - v)(u - v)$$

für alle t, u, v mit einem $K \in \mathbf{R}$. Es sei $y(\cdot)$ die Lösung der Anfangswertaufgabe auf $[t_0, t_0 + T_0]$. Gegeben sei ein stark stabiles und konsistentes lineares k -Schrittverfahren

$$a_k y_{n+k} + \dots + a_0 y_n = h(b_k f_{n+k} + \dots + b_0 f_n),$$

$$|a_0| + |b_0| \neq 0, \quad a_k \neq 0,$$

$$a_i, b_i \in \mathbf{R}, \quad i = 0, 1, 2, \dots, k, \quad f_n = f(t_n, y_n), \quad t_n = t_0 + nh.$$

Dann gilt für alle Anfangsfelder $y_{k-1}^h, y_{k-2}^h, \dots, y_0^h$ mit der Eigenschaft

$$y_i^h - y(t_0) = O(h) \quad i = 0, 1, 2, \dots, k-1$$

und für alle $t \in [t_0, t_0 + T_0]$

$$\lim_{\substack{h \rightarrow 0 \\ n \rightarrow \infty \\ t_0 + nh \rightarrow t}} y_n^h = y(t)$$

Beweis

Ein nicht so einfach zu beweisender Satz!

Numerische Behandlung von gewöhnlichen Differentialgleichungen

Eine Einführung

Universität Hamburg
SoSe07

K. Taubert

Numerik gewöhnlicher Anfangswertaufgaben

5 DIE STABILITÄT VON DIFFERENZENVERFAHREN

Ein grundsätzlicher Ansatz zur Beurteilung von Differenzenverfahren ist neben dem lokalen Fehler (Ordnung) die Stabilität.

Bei der „Stabilität“ wird i.A. verlangt, dass die von einem Differenzenverfahren gelieferten Näherungslösungen für alle hinreichend kleinen oder sogar alle h ein einheitliches und ähnliches „Stabilitätsverhalten“ wie die Lösungen der Anfangswertaufgabe aufweisen. Vier Klassen von Anfangswertaufgaben

$$y' = f(x,y), \quad y(x_0) = y_0$$

stehen hier im Vordergrund und zwar meistens solche mit einem stetigen $f : \mathbf{R}^* \mathbf{R} \rightarrow \mathbf{R}$ der Art:

1. $|f(x,u) - f(x,v)| \leq K|u - v|$ für alle zulässigen x,u,v .
2. Die Anfangswertaufgaben sind eindeutig lösbar für alle Anfangsbedingungen.
3. $(f(x,u) - f(x,v))(u - v) \leq 0$ für alle zulässigen x,u,v .
4. $f(x,y) = \lambda y$ mit $\operatorname{Re}(\lambda) < 0$.

Schon diese Klassen von Anfangswertaufgaben führen zu einer größeren Menge von Stabilitätsbegriffen. Dazu gehören:

1. (Starke) Stabilität,
2. A-Stabilität,
3. B-Stabilität,
4. G-Stabilität,
5. (Bereiche der) absoluten Stabilität und
6. L-Stabilität.

§ 5.1 Stabilität von Differenzenverfahren für lipschitzstetige Aufgaben

Es sei $f : \mathbf{R}^* \mathbf{R} \rightarrow \mathbf{R}$ eine stetige Funktion mit

$$|f(x,u) - f(x,v)| \leq K|u - v| \quad \text{für alle } x,u,v \text{ und einem } K \geq 0.$$

Gegeben seien die zwei Anfangswertaufgaben

$$\begin{aligned} y' &= f(x,y), & y(x_0) &= y_1, \\ y' &= f(x,y) + \delta(x), & y(x_0) &= y_2, \end{aligned}$$

mit einer stetigen Funktion $\delta : \mathbf{R} \rightarrow \mathbf{R}$.

Die Differenz der Lösungen $y_1(\cdot)$ und $y_2(\cdot)$ der beiden Anfangswertaufgaben auf dem Intervall $I = [x_0-a, x_0+a]$, $a > 0$, kann mit der Gronwallschen Ungleichung in der folgenden Form abgeschätzt werden :

$$\text{Max}_{x \in I} | y_1(x) - y_2(x) | \leq | y_1 - y_2 | e^{Ka} + \frac{M}{K} (e^{Ka} - 1).$$

Dabei sei $M = \text{Max}_{x \in I} | \delta(x) |$.

D.h., die Differenz der Lösungen hängt (lipschitz-) stetig von Störungen in den Anfangsbedingungen und der rechten Seite der Anfangswertaufgabe ab.

Die einfachste Stabilitätsbedingung für Differenzenverfahren bei Anfangswertaufgaben besteht nun darin, von den Lösungen der Differenzenverfahren eine ähnliche Eigenschaft zu fordern und zwar gleichmäßig in h für alle $0 < h < h_0$.

Satz 5.1 (Stabilität bei RKV)

Gegeben sei das RKV

$$\frac{y_{n+1} - y_n}{h} = \Phi(x_n, y_n; h, f), \quad y_0 = y^1,$$

und das gestörte RKV

$$\frac{y_{n+1} - y_n}{h} = \Phi(x_n, y_n; h, f) + \delta_n, \quad y_0 = y^2.$$

Dabei sei f eine lipschitzstetige Funktion und Φ genüge der Bedingung

$$| \Phi(x_n, u; h, f) - \Phi(x_n, v; h, f) | \leq L_f | u - v |$$

für alle zulässigen x_n, u, v und alle $0 < h < h_0$ auf $I = [x_0-a, x_0+a]$, $a > 0$. Es sei außerdem $M = \text{Max}_n | \delta_n |$ und $L_f > 0$.

Dann gilt für die Differenz der zwei Lösungen (y_n^1) und (y_n^2) der RKV

$$\text{Max}_n | y_n^1 - y_n^2 | \leq | y^1 - y^2 | e^{L_f a} + \frac{M}{L_f} (e^{L_f a} - 1).$$

Beweis

Siehe § 2.

Bemerkung

Man beachte, dass mit einer Zusatzvoraussetzung (Konsistenz) im Satz 5.1 sofort die Konvergenz eines RKV gefolgert werden kann :

Erfüllt die Lösung einer Lipschitzstetigen Anfangswertaufgabe die Bedingung

$$\frac{y(x_{n+1}) - y(x_n)}{h} = \Phi(x_n, y(x_n); h, f) + \tau_{n+1}(h)$$

mit $\tau_{n+1}(h) = O(h)$ für alle $x \in I$, dann liegt bei gleicher Anfangsbedingung für das RKV offenbar Konvergenz auf I vor!

Satz 5.2 (Stabilität bei MSV)

Gegeben seien ein lineares k -Schrittverfahren

$$a_k y_{n+k} + \dots + a_0 y_n = h(b_k f_{n+k} + \dots + b_0 f_n)$$

mit dem Anfangsfeld $y_{k-1}^1, y_{k-2}^1, \dots, y_0^1$ und das gestörte Verfahren

$$a_k y_{n+k} + \dots + a_0 y_n = h(b_k f_{n+k} + \dots + b_0 f_n) + h \delta_n$$

mit dem Anfangsfeld $y_{k-1}^2, y_{k-2}^2, \dots, y_0^2$.

Dabei sei $|a_0| + |b_0| \neq 0$, $a_k \neq 0$, $a_i, b_i \in \mathbf{R}$, $i = 0, 1, 2, \dots, k$, $f_n = f(t_n, y_n)$.

Die Wurzeln des charakteristischen Polynoms seien alle dem Betrage nach kleiner oder gleich 1 und jene Wurzeln mit Betrag Eins seien einfach.

Es sei f eine Lipschitzstetige Funktion mit der Lipschitzkonstanten K und

$$Y_n^i = \begin{pmatrix} y_{n+k-1}^i \\ \vdots \\ y_{n+1}^i \\ y_n^i \end{pmatrix}, \quad i=1,2, \quad n=0,1,2, \dots$$

Dabei sei $(y_n^1)_n$ die Lösung des ungestörten linearen k -Schrittverfahrens und $(y_n^2)_n$ die Lösung des gestörten linearen k -Schrittverfahrens.

Dann gilt für alle zulässigen n und alle h mit $0 < h < h_0$ auf dem Intervall $I = [x_0 - a, x_0 + a]$, $a > 0$,

$$\max_n \|Y_n^1 - Y_n^2\|_\infty \leq \|Y_0^1 - Y_0^2\|_\infty e^{L_f a} + \frac{\delta}{K_f} (e^{L_f a} - 1).$$

Dabei seien $\delta = \max_n |\delta_n|$ und $L_f > 0$ eine geeignete von der Lipschitzstetigen Funktion f abhängige Konstante.

Beweis

Siehe § 4.

Bemerkung

Es bleibt dem Leser überlassen, mittels des Satzes 5.2 auch die Konvergenz eines linearen und konsistenten Mehrschrittverfahrens zu beweisen.

Bemerkung

In der modernen Literatur werden die Ungleichungen aus Satz 5.1 und 5.2 häufig als Stabilitätsdefinitionen für die Differenzenverfahren verwendet. Hinreichend für diese Stabilitätsforderung ist dann, in der Klasse der Lipschitzstetigen Funktionen f , die gegebene Lipschitzbedingung bei den RKV oder die Wurzelbedingung bei den linearen Mehrschrittverfahren.

§ 5.2 (Starke) Stabilität bei Differenzenverfahren

Die Anfangswertaufgabe

$$y' = f(x,y), \quad y(x_0) = y_0,$$

mit einer stetigen Funktion $f: \mathbf{R}^* \times \mathbf{R} \rightarrow \mathbf{R}$ habe eine eindeutige Lösung $y(\cdot)$ im Intervall $I = [x_0, x_0+a]$, $a > 0$.

Gegeben sei außerdem ein RKV

$$\frac{y_{n+1} - y_n}{h} = \Phi(x_n, y_n; h, f), \quad y_0 = y^1.$$

Ist f keine Lipschitzstetige oder monotone Funktion, dann stellt sich die Frage, welche (Stabilitäts-) Eigenschaften das RKV erfüllen muss, damit es ein brauchbares Verfahren für die numerische Approximation der Lösung der Anfangswertaufgabe darstellt.

Wir erinnern uns daran, dass das explizite Euler-Verfahren eine Folge von gleichmäßig konvergenten Polygonzügen lieferte, deren Grenzwert die Lösung der Anfangswertaufgabe war. Die gleichmäßige Konvergenz der Polygonzüge wurde dadurch erzwungen, dass die Polygonzüge eine Familie von gleichgradig stetigen und gleichmäßig beschränkten Funktionen auf I waren. Genau dies ist die Stabilitätsforderung, die an das Differenzenverfahren zu stellen ist!

Satz 5.3

Es sei $\Phi(u,v;h,f)$ eine in einer Umgebung U von $\{(x,y(x)) \mid x \in I\}$ für alle $0 < h \leq h_0$ beschränkte Funktion. Liegen die durch das RKV gelieferten Näherungen y_n^h , $nh \leq a$, für eine beschränkte Umgebung von Anfangsbedingungen y_0^h von $y(x_0)$ alle in U , dann erzeugt die Familie $y^h(\cdot, y_0^h)$, $0 < h \leq h_0$, ein System von gleichgradig stetigen Polygonzügen auf I .

Beweis

Trivial.

Bemerkung

Man beachte:

Wegen der Stetigkeit von f und der Konstruktionsvorschrift für explizite RKV ist die Beschränktheit von Φ in einer Umgebung von U stets gewährleistet.

Soll für $h \rightarrow 0$ und $y_0^h \rightarrow y(x_0)$ zusätzlich noch die gleichmäßige Konvergenz der Familie $y^h(\cdot, y_0^h)$ gegen die Lösung $y(\cdot)$ der Anfangswertaufgabe auf I gewährleistet werden, dann ist die gleichmäßige Konvergenz $\Phi(u,v;h,f) \xrightarrow{h \rightarrow 0} f(u,v)$ eine dafür ausreichende Bedingung. Diese letzte Bedingung wird häufig als Konsistenz eines allgemeinen Einschrittverfahrens bezeichnet.

Etwas komplizierter gestaltet sich die obige Problematik, wenn lineare k -Schrittverfahren zur numerischen Integration der Anfangswertaufgabe verwendet werden sollen:

Gegeben sei ein lineares k -Schrittverfahren

$$a_k y_{n+k} + \dots + a_0 y_n = h(b_k f_{n+k} + \dots + b_0 f_n)$$

mit dem Startfeld $Y_0^h = (y_{k-1}^h, y_{k-2}^h, \dots, y_0^h)$ und es sei $|y_i^h - y(x_0)| \leq Lh$ für $i = 0, 1, \dots, k-1$ mit einem $L > 0$.

Definition

Ein lineares konsistentes Mehrschrittverfahren heißt stark stabil, wenn die Wurzeln des charakteristischen Polynoms $p(\xi) = a_k \xi^k + a_{k-1} \xi^{k-1} + \dots + a_0$ alle dem Betrage nach kleiner oder gleich 1 sind. Es gibt genau eine Wurzel mit dem Betrag Eins, nämlich die Eins.

Satz 5.4

Gegeben seien ein stark stabiles k - Schrittverfahren und eine Familie von Startfeldern

$$Y_0^h = (y_{k-1}^h, y_{k-2}^h, \dots, y_0^h) \text{ mit } |y_i^h - y(x_0)| \leq Lh \text{ für } i = 0, 1, 2, \dots, k-1 \text{ und } L > 0.$$

Es sei $(y_n^h)_n$ die Lösung des Differenzenverfahrens für das Startfeld Y_0^h und $y^h(\cdot, Y_0^h)$ der zugehörige Polygonzug auf $I = [x_0, x_0+a]$, $a > 0$.

Die Familie der Funktionen $y^h(\cdot, Y_0^h)$ ist dann auf I gleichmäßig beschränkt und gleichgradig stetig.

Zum Beweis

Beim Beweis spielt natürlich die starke Stabilität des Differenzenverfahrens eine entscheidende Rolle. Versuchen Sie es! Gegenbeispiele zeigen, dass ein Verzicht auf diese Bedingung nicht möglich ist.

Bemerkung

Offen bleibt bis jetzt noch die Frage, ob die Funktionen $y^h(\cdot, Y_0^h)$ für $h \rightarrow 0$ auch die Lösung der Anfangswertaufgabe liefern. Dieses ist dann der Fall, wenn neben

$$1. \sum_{n=0}^k a_n = 0$$

auch noch die Bedingung

$$2. \sum_{n=0}^k n a_n = \sum_{n=0}^k b_n \quad b_n \geq 0$$

erfüllt ist.

§ 5.3 Stabilität von Differenzenverfahren für monotone Aufgaben B und G-Stabilität

Es sei $f: \mathbf{R}^* \times \mathbf{R} \rightarrow \mathbf{R}$ eine stetige Funktion mit

$$(f(x,u) - f(x,v))(u-v) \leq 0 \text{ für alle zulässigen } x, u, v.$$

Gegeben seien die zwei Anfangswertaufgaben

$$\begin{aligned} y' &= f(x,y) & y(x_0) &= y_0 \\ y' &= f(x,y) & y(x_0) &= \tilde{y}_0. \end{aligned}$$

Für die Differenz der Lösungen $y(\cdot)$ und $\tilde{y}(\cdot)$ der beiden Anfangswertaufgaben gilt:

$$|y(x) - \tilde{y}(x)| \leq |y_0 - \tilde{y}_0|$$

und zwar für alle x aus dem Intervall $I = [x_0, \infty)$.

Werden die beiden Anfangswertaufgaben mit dem impliziten Euler-Verfahren approximativ gelöst, dann gilt für alle Schrittweiten $h > 0$ und alle n

$$|y_n - \tilde{y}_n| \leq |y_0 - \tilde{y}_0|.$$

Definition :

Ein implizites RKV $y_{n+1} = y_n + h \Phi(x_n, y_n, h, f)$ heißt B-stabil, wenn für alle $h > 0$ und alle Anfangsbedingungen y_0, \tilde{y}_0 gilt

$$|y_n - \tilde{y}_n| \leq |y_0 - \tilde{y}_0|.$$

Beispiele

Schnell kann man sich davon überzeugen, dass die one-leg Version der Trapezregel

$$y_{n+1} = y_n = hf((1/2)x_n + (1/2)x_{n+1}, (1/2)y_{n+1} + (1/2)y_n)$$

ein B-stabiles Verfahren ist.

Die Butcher-Arrays

γ	γ	0
$1 - \gamma$	$1 - 2\gamma$	γ
	$1/2$	$1/2$

$(\gamma = \frac{3 \pm \sqrt{3}}{6})$ liefern implizite RKV der Ordnung 3. Mit $\gamma = \frac{3 + \sqrt{3}}{6}$ liegt ein B-stabiles Verfahren vor.

Bei den linearen k-Schrittverfahren

$$a_k y_{n+k} + \dots + a_0 y_n = h(b_k f_{n+k} + \dots + b_0 f_n)$$

ist es zweckmäßig deren one-leg Versionen

$$a_k y_{n+k} + \dots + a_0 y_n = hf(b_k t_{n+k} + \dots + b_0 t_n, a_k y_{n+k} + \dots + a_0 y_n)$$

zu nehmen. Dabei sei $\sum_{i=0}^k b_i = 1$ (Normalisierung).

Bei den BDF-Verfahren fallen beide Klassen zusammen!

Bemerkung

Werden die monotonen Aufgaben mit linearen k-Schrittverfahren integriert, dann sollten – auf den ersten Blick – Näherungslösungen $(y_n^1)_n$ und $(y_n^2)_n$ die Bedingung

$$|y_{n+1}^1 - y_{n+1}^2| \leq |y_n^1 - y_n^2|$$

für alle n erfüllen.

Da die Werte y_n jedoch vom Startfeld $y_0^h, y_1^h, \dots, y_{k-1}^h$ abhängen, wird stattdessen eine ähnliche Bedingung gefordert:

Definition

Die one-leg Verfahren heißen G-stabil, wenn es eine reelle, symmetrische und positiv definite Matrix G gibt, so dass für je zwei numerische Lösungen $(y_n^1)_n$ und $(y_n^2)_n$ gilt

$$\| Y_{n+1}^1 - Y_{n+1}^2 \|_G \leq \| Y_n^1 - Y_n^2 \|_G \quad \text{für alle } n.$$

Dies gilt im speziellen für alle $h > 0$ und alle monotonen Anfangswertaufgaben.

Dabei sei

$$Y_n = (y_{n+k-1}, \dots, y_n)^T,$$

$\| z \|_G = (Gz, z)^{1/2}$ und (\cdot, \cdot) das gewöhnliche innere Produkt auf dem \mathbf{R}^k .

Beispiele

1. Das implizite Euler-Verfahren ist G-stabil.
2. Die one leg-Version der Trapezregel ist G-stabil.

3. Die BDF-Formel der Ordnung zwei ist G-stabil mit $G = \frac{1}{4} \begin{pmatrix} 5 & -2 \\ -2 & 1 \end{pmatrix}$.

Bemerkung

Es ist ein Ergebnis von G. Dahlquist (1978), dass es in der Klasse der linearen k-Schrittverfahren (one-leg Versionen) außer den genannten Beispielen keine weiteren wichtigen Formeln gibt, die G-stabil sind.

§ 5.4 A-Stabilität und Bereiche der absoluten Stabilität von RKV und MSV

Die Lösungen der Anfangswertaufgabe

$$y' = f(x, y) = \lambda y \quad \text{mit } \operatorname{Re}(\lambda) < 0, \\ y(x_0) = y_0,$$

hat die Eigenschaft

$$\lim_{x \rightarrow \infty} y(x) = 0.$$

Gesucht sind deshalb Differenzenverfahren, die für alle festen Schrittweiten h und alle λ mit $\operatorname{Re}(\lambda) < 0$ dieses asymptotische Verhalten wiedergeben.

Gegeben sei ein RKV (explizit oder implizit)

$$y_{n+1} = y_n + h \Phi(x_n, y_n; h, f).$$

Wird dieses Verfahren zur numerischen Integration der Anfangswertaufgabe $y' = \lambda y$ mit $\operatorname{Re}(\lambda) < 0$ verwendet, dann ergibt sich eine Rekursionsformel der Form

$$y_{n+1} = R(\tilde{h}) y_n \quad \text{mit } \tilde{h} = \lambda h.$$

Dabei ist $R(z)$ ein Quotient von zwei Polynomen vom Grade $\leq s$, wenn das RKV von der Stufe s ist.

Beispiele

Das explizite Euler-Verfahren führt zu $R(z) = 1+z$ und das implizite Euler-Verfahren zu $R(z) = 1/(1-z)$.

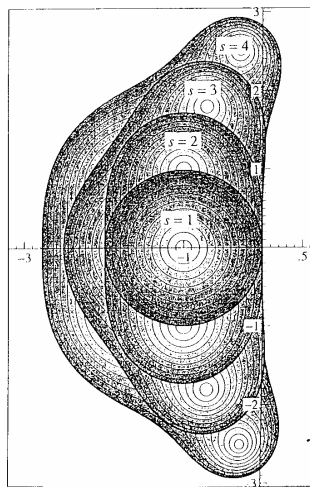
Für ein allgemeines explizites RKV der Ordnung p hat $R(\cdot)$ die Gestalt

$$R(z) = 1+z + \frac{z^2}{2!} + \frac{z^3}{3!} + \dots + \frac{z^p}{p!}.$$

Definition

Für ein RKV heißt die Menge der $z \in \mathbf{C}$ mit der Eigenschaft $|R(z)| \leq 1$ der Bereich der absoluten Stabilität des Verfahrens. Umfasst dieser Bereich die Menge aller z mit $\text{Re}(z) \leq 0$, dann ist das Verfahren A-stabil.

Anschließend finden Sie die Bereiche der absoluten Stabilität der expliziten RKV der Ordnungen 1,2,3 und 4. Die Zeichnung wurde aus dem umfassenden und sehr zu empfehlenden Buch von E. Hairer, S.P. Nörsett und G. Wanner¹ entnommen:



Aus der Darstellung der Übertragungsfunktion $R(z)$ für explizite RKV der Ordnung p kann sofort erkannt werden, dass es in der Klasse der expliziten RKV keine A-stabilen Verfahren geben kann.

Implizite A-stabile RKV können jedoch angegeben werden. Beispiele dafür sind:

γ	γ	0
$1-\gamma$	$1-2\gamma$	γ
	$1/2$	$1/2$

$\gamma = (3 + \sqrt{3}) / 6$ SDIRK Ordnung 3

0	0	0	0	Lobatto 3A Ordnung 4
$1/2$	$5/24$	$1/3$	$-1/24$	
1	$1/6$	$1/3$	$1/6$	
	$1/6$	$2/3$	$1/6$	

Die zugehörigen Übertragungsfunktionen sind:

¹ Solving Ordinary Differential Equations II. Springer Series in Computational Mathematics 14. 1996

$$R(z) = \frac{1 + (1 - 2\gamma)z + (1/2 - 2\gamma + \gamma^2)}{(1 - \gamma z)^2}$$

bzw.

$$R(z) = \frac{1 + z/2 + z^2/12}{1 - z/2 + z^2/12}$$

Bevor der Frage nach der A-Stabilität bzw. den Bereichen der absoluten Stabilität bei linearen Mehrschrittverfahren nachgegangen wird, ist es zweckmäßig, einige Bemerkungen über lineare Differenzgleichungen zu machen:

Eine Gleichung der Form

$$y_{n+k} + a_1 y_{n+k-1} + \dots + a_k y_n = 0, \quad n \in \mathbf{N}_0 \quad (*)$$

($a_i, i = 1, 2, \dots, k, a_k \neq 0$, aus \mathbf{R} oder \mathbf{C}) heißt homogene lineare Differenzgleichung mit konstanten Koeffizienten der Ordnung k .

Eine homogene lineare Differenzgleichung mit konstanten Koeffizienten wird durch eine Zahlenfolge $(y_j)_j$ mit $y_j = cu^j$ ($c \neq 0$) gelöst, wenn

$$cu^{n+k} + a_1 cu^{n+k-1} + \dots + a_k cu^n = 0$$

für alle $n \in \mathbf{N}_0$ ist oder u die Gleichung

$$\varphi(u) = u^k + a_1 u^{k-1} + \dots + a_k = 0$$

erfüllt.

Die Gleichung $\varphi(u) = 0$ heißt charakteristische Gleichung der Differenzgleichung und es gilt der

Satz

Ist s eine Wurzel der charakteristischen Gleichung

$$\varphi(u) = u^k + a_1 u^{k-1} + \dots + a_k = 0$$

mit der Vielfachheit r so sind die Folgen

$$(s^j)_i, (is^j)_i, \dots, (i^{r-1}s^j)_i,$$

Lösungen der linearen Differenzgleichung (*) und jede Lösung von (*) lässt sich als Linearkombination von Lösungen der obigen Form darstellen.

Die allgemeine Lösung der Differenzgleichung der Ordnung k hat noch k freie Parameter c_i . Diese können durch die Vorgabe von

$$y_0, y_1, \dots, y_{k-1}$$

festgelegt werden.

Beispiel

Die numerische Integration der linearen Differentialgleichung

$$y' = \lambda y \quad \text{mit} \quad \lambda \in \mathbf{C}$$

durch ein stabiles und konsistentes k -Schrittverfahren

$$a_k y_{n+k} + \dots + a_0 y_n = h(b_k f_{n+k} + \dots + b_0 f_n)$$

führt zu einer homogenen Differenzgleichung der Ordnung k , nämlich:

$$\sum_{j=0}^k (a_j - h\lambda b_j) y_{n+j} = 0, \quad n \in \mathbf{N}_0.$$

Die zugehörige (von $h\lambda$ abhängige) charakteristische Gleichung lautet dann

$$\varphi(u) = \sum_{j=0}^k (a_j - h\lambda b_j) u^j = 0.$$

Definition

Gegeben sei ein stabiles und konsistentes lineares k -Schrittverfahren

$$a_k y_{n+k} + \dots + a_0 y_n = h(b_k f_{n+k} + \dots + b_0 f_n).$$

Die Menge aller $\tilde{h} = h\lambda$, $h > 0$, $\operatorname{Re}(\lambda) < 0$, für die die charakteristische Gleichung

$$(a_k - \tilde{h}b_k)u^k + \dots + (a_0 - \tilde{h}b_0) = 0$$

nur Wurzeln mit dem Betrag kleiner als 1 hat, heißt Bereich der absoluten Stabilität des Mehrschrittverfahrens.

Beispiele

Das explizite Euler-Verfahren führt auf die charakteristische Gleichung

$$u + (-1 - \tilde{h}) = 0$$

mit der Wurzel $1 + \tilde{h}$. Das explizite Euler-Verfahren ist deshalb absolut stabil für alle \tilde{h} aus der offenen Scheibe mit dem Mittelpunkt -1 und dem Radius 1 in \mathbf{C} (Siehe auch Kapitel 1).

Die Mittelpunktsformel $y_{n+2} = y_n + 2hf_{n+1}$ führt auf die charakteristische Gleichung

$$u^2 - 2\tilde{h}u - 1 = 0$$

mit den Wurzeln $s_{1,2} = \tilde{h} \pm \sqrt{\tilde{h}^2 + 1}$. Damit besitzt die Mittelpunktsformel kein Bereich der absoluten Stabilität.

Die Trapezregel $y_{n+1} = y_n + (h/2)f_{n+1} + (h/2)f_n$ führt auf die charakteristische Gleichung

$$\left(1 - \frac{\tilde{h}}{2}\right)u - \left(1 + \frac{\tilde{h}}{2}\right) = 0$$

mit der Wurzel

$$s = (1 + \tilde{h}/2)/(1 - \tilde{h}/2).$$

Diese Wurzel ist für alle \tilde{h} dem Betrage nach kleiner als 1.

Die Bereiche der absoluten Stabilität einiger Adams-Verfahren und Ihrer zugehöriger PECE-Version können aus den folgenden Bildern (wieder aus dem Buch von Hairer, Nørsett und Wanner) entnommen werden:

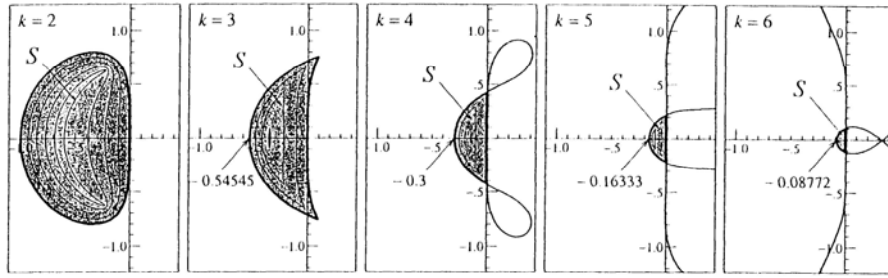


Fig. 1.2. Stability domains for explicit Adams methods

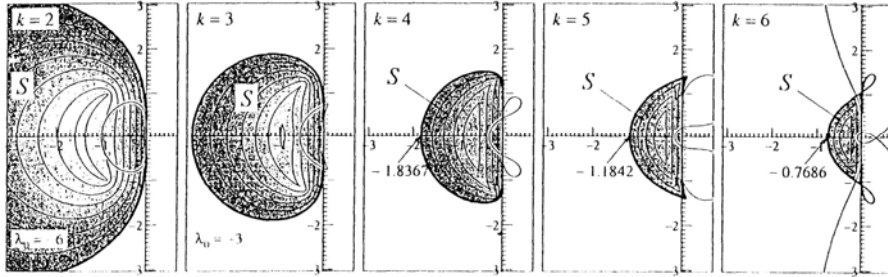


Fig. 1.3. Stability domains of implicit Adams methods, compared to those of the explicit ones

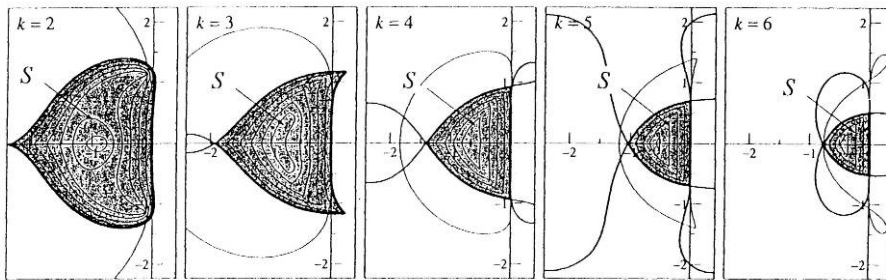


Fig. 1.4. Stability domains for PECE compared to original implicit methods

Anschließend noch die Bereiche der absoluten Stabilität der BDF-Formeln

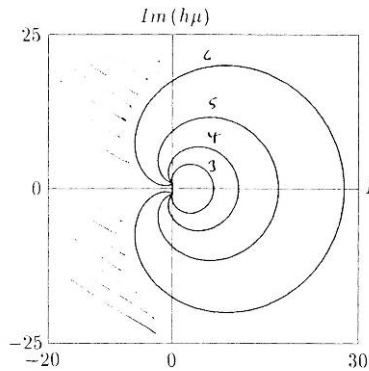


Abbildung 1.5: Stabilitätsgebiete für Ordnung $k = 3, 4, 5, 6$

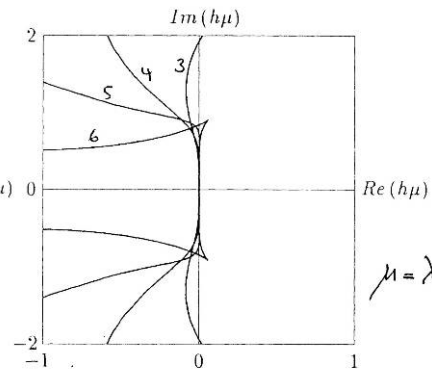


Abbildung 1.6: Ausschnitt für Ordnung $k = 3, 4, 5, 6$

Bemerkung:

Es wird i.A. schwierig sein, für alle h , alle Wurzeln einer charakteristischen Gleichung

$$(a_k - hb_k)u^k + \dots + (a_0 - hb_0) = 0$$

zu ermitteln. Da hilft der folgende Trick:

Bestimmen Sie die h (Kurve in \mathbf{C}) für die das charakteristische Polynom Wurzeln mit dem Betrag 1 hat. Die Wurzeln mit dem Betrag 1 sind gegeben durch

$$\sum_{j=0}^k (a_j - hb_j)e^{jj\varphi} = 0 \quad \Rightarrow \quad h = \frac{\sum_{j=0}^k a_j e^{jj\varphi}}{\sum_{j=0}^k b_j e^{jj\varphi}} .$$

Ein einfacher Test liefert dann den Bereich der absoluten Stabilität.

Definition

Ein stabiles und konsistentes Mehrschrittverfahren heißt A-stabil, wenn die Wurzeln der charakteristischen Gleichung für alle h , $\text{Re}(h) < 0$, dem Betrage nach kleiner als 1 sind. Kurz: Wenn der Bereich der absoluten Stabilität die linke komplexe Halbebene ist.

Beispiele

Das implizite Euler-Verfahren, die Trapezregel und die BDF-Formel der Ordnung 2 sind stabile Differenzenverfahren.

§ 5.5 L-Stabile Verfahren

Bereits im Kapitel 1 hatten wir gesehen, dass die Trapezregel (obwohl A-Stabil) nicht in der Lage war, die Anfangswertaufgabe $y' = \lambda y$, $y(0) = y_0$, mit negativen und betragsgroßen λ adäquat zu integrieren. Dieses hat zu einem weiteren Stabilitätsbegriff geführt.

Definition

Ein implizites Runge-Kutta-Verfahren heißt L-stabil, wenn es A-Stabil ist und für seine Übertragungsfunktion $R(z)$ gilt

$$\lim_{z \rightarrow \infty} R(z) = 0.$$

Beispiele für L-stabile Verfahren sind wieder das implizite Euler-Verfahren, aber auch die sogenannte Radau IIA-Formel.

Auch für lineare konsistente und stabile Mehrschrittverfahren wird L-Stabilität definiert. Die L-Stabilität der Verfahren wird allerdings nicht verlangt.

Definition

Ein lineares, konsistentes und stabiles Mehrschrittverfahren heißt L-stabil, wenn für die Wurzeln $s_i(h)$ der charakteristischen Gleichung

$$\varphi(u) = \sum_{j=0}^k (a_j - hb_j)u^j = 0$$

gilt

$$\limsup_{\operatorname{Re}(\hbar) \rightarrow -\infty} |s_i(\hbar)| < 1 \quad \text{für } i = 1, 2, \dots, k.$$

Als Beispiele für L-Stabile Verfahren können die BDF-Formeln genannt werden. Die Trapezregel ist nicht L-stabil!

Numerische Behandlung von gewöhnlichen Differentialgleichungen

Eine Einführung

Universität Hamburg
SoSe07

K. Taubert

Besondere Aufgaben

6 UNSTETIGE-, SYMPLEKTISCHE- und ALGEBRO-DIFFERENTIALGLEICHUNGEN

§ 6.1 Einführung

Es werden zunächst einige Beispiele für Differentialgleichungen mit Unstetigkeiten angegeben. Ein allgemeiner Existenzsatz garantiert die Existenz von Lösungen von Differentialgleichungen mit Unstetigkeiten bzw. von entsprechenden mehrwertigen Differentialgleichungen. Eine sehr große Klasse von Differenzenverfahren liefert in völlig unproblematischer Weise, ausreichend gute Näherungen für die Lösungen.

Aufgaben der klassischen Mechanik führen oft auf Differentialgleichungen die mit Erhaltungsgrößen einhergehen. Werden diese Differentialgleichungen mit klassischen Differenzenverfahren und über längere Zeiten integriert, wird meistens die Erhaltungsgröße im Laufe der numerischen Integration verletzt. Symplektische Verfahren überwinden diese Problematik z.B. im Bereich der Volumeninvarianz.

Die formale Aufstellung der Gleichungen für mechanische oder elektrische Systeme führen meistens zu hybriden Systemen von Differentialgleichungen und Gleichungen. Erst weitere Überlegungen führen dann zu einem vollständigen System von Differentialgleichungen. Bei modernen Simulationsprogrammen sind solche weitergehenden Überlegungen unerwünscht oder wegen des Umfangs der Probleme sogar oft nicht durchführbar. Es müssen Verfahren zur Integration von Algebro-Differentialgleichungen eingesetzt werden. Algebro-Differentialgleichungen mit dem Index 1 sollen hier kurz behandelt werden.

§ 6.2 Differentialgleichungen mit Unstetigkeiten. Beispiele

1. Stabilisierung einer beschleunigten Masse

Ein Körper mit der Masse m genüge dem Newton'schen Bewegungsgesetz

$$mx'' = u$$

mit der äußeren Kraft u .

Für den speziellen Fall einer Masse $m = 1$, der Kraft $u = \varepsilon b$ (mit $\varepsilon = +1$ oder $\varepsilon = -1$) und $b > 0$ ergeben sich die beiden Differentialgleichungen

$$x'' = \varepsilon b.$$

Für das Weitere ist es zweckmäßig die Geschwindigkeiten $x' = v$ in Abhängigkeit vom Ort x zu bestimmen:

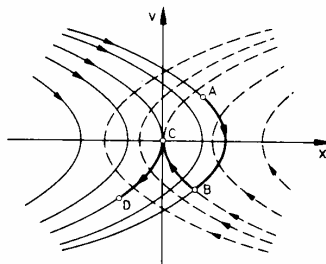
Für jede Lösung der Differentialgleichung gilt

$$\frac{d}{dt} \left(\left(\frac{1}{2} \right) x'^2 - \varepsilon b x \right) = 0$$

und deshalb, mit geeigneten Konstanten C und C' , notwendigerweise auch

$$x'^2 - 2 \varepsilon b x = C \quad \text{oder} \quad v^2 = 2 \varepsilon b (x - C')$$

Eine Auswahl dieser Parabeln zeigen die folgenden Bilder (Phasenportraits). Die Richtung der Bewegungsabläufe ($x(t)$, $x'(t)$) wird durch die Pfeile angezeigt:

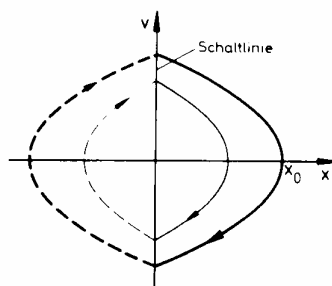


Die Frage nach einer Stabilisierung dieses Systems führt sofort zu unstetigen Differentialgleichungen:

Beschränkte Bewegungsabläufe ergeben sich z.B. mit $u = -\text{sgn}(x) \cdot b$. Die Differentialgleichung

$$x' = -\text{sgn}(x) \cdot b$$

führt zum Phasenportrait



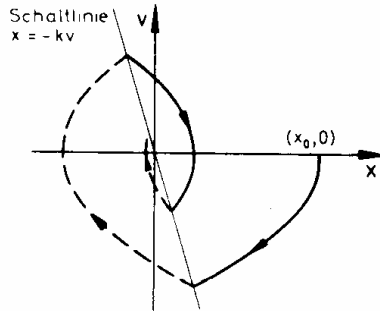
Die oben gewählte Ortsabhängige äußere Kraft $u = -\text{sgn}(x) \cdot b$ führt zwar zu beschränkten Bewegungsabläufen aber noch nicht zu solchen Bewegungen die das System stets in die Ruhelage bringen. Letzteres kann z.B. mit der Orts- und Geschwindigkeitsabhängigen Kraft

$$u = -\text{sgn}(x + k \cdot x') \cdot b, \quad k > 0$$

erreicht werden. Diese Wahl ergibt dann die Differentialgleichung

$$x'' = -\text{sgn}(x + k \cdot x') \cdot b$$

mit dem Phasenportrait



Außerhalb der Geraden $x + k \cdot x'$ bewegt sich das System auf den oben angegebenen Parabeln und (klar!) mit zunehmender Zeit gegen den Nullpunkt.

Ist die Gerade $x + k \cdot x'$ sehr „flach“, dann muss der zu erwartender Bewegungsablauf etwas genauer untersucht werden.

Es kann dann passieren, dass sich das System ab einem gewissen Zeitpunkt t_0 entlang der Geraden $x + kx' = 0$ bewegen muss. Dann muss

$$x(t) = x(t_0)e^{-(t-t_0)/k}$$

sein.

Soll schließlich das System nicht nur gegen Null streben, sondern den Nullpunkt auch noch möglichst schnell (zeitoptimal) erreichen, dann bietet sich der Übergang von einer Parabel zu einer durch Null gehenden Parabel an. Auch dann entsteht ein System

$$\begin{aligned} x' &= y \\ y' &= \Psi(x, y) \end{aligned}$$

mit einer unstetigen Funktion $\Psi(x, y)$ oder eine unstetige Differentialgleichung.

2. Erzwungene Schwingungen mit trockener Reibung.

Erzwungene Schwingungen mit trockener Reibung führen auf die Differentialgleichung ($D, \omega_0, \mu > 0$)

$$x'' + 2D \omega_0 x' + \mu \text{sgn}(x') + \omega_0^2 x = \alpha \cos(\Omega t).$$

Wenn die äußeren Kräfte und die Rückstellkraft die trockene Reibung nicht überwinden können, kann das System zeitweilig ruhen. Orte möglicher Ruhezustände x sind dann

$$-\mu + \alpha \cos(\Omega t) \leq \omega_0^2 x \leq +\mu + \alpha \cos(\Omega t).$$

Auch hier liegt eine unstetige Differentialgleichung $z' = F(t,z)$ vor:

$$\begin{aligned} x' &= y \\ y' &= -2D\omega_0 y - \mu \operatorname{sgn}(y) - \omega_0^2 x + \alpha \cos(\Omega t). \end{aligned}$$

Die unstetige Funktion $F : \mathbf{R} \times \mathbf{R}^2 \rightarrow \mathbf{R}$ erfüllt übrigens die Monotoniebedingung

$$(F(t,u) - F(t,v)) \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1/\omega_0^2 \end{pmatrix} (u - v) \leq 0$$

für alle $t, u, v \in \mathbf{R}$.

Es war eine Idee von A.F. Filippov (1960) unstetige Differentialgleichungen als mehrwertige Differentialgleichungen zu interpretieren. Gegeben sei also jetzt eine (mehrwertige) Differentialgleichung der Form

$$y' \in F(t,y), \quad y(t_0) = y_0.$$

Dabei sei F eine Abbildung

$$F : \mathbf{R} \times \mathbf{R}^n \rightarrow P(\mathbf{R}^n)$$

und $P(\mathbf{R}^n)$ die Menge aller nichtleeren, konvexen und abgeschlossenen Teilmengen aus \mathbf{R}^n .

Definition

Eine Funktion $y : [t_0, \infty) \rightarrow \mathbf{R}^n$ heißt Lösung der mehrwertigen Anfangswertaufgabe

$$y' \in F(t,y), \quad y(t_0) = y_0,$$

sofern

1. $y(t_0) = y_0$
2. $y(\cdot)$ absolut stetig ist und
3. für fast alle t gilt $y'(t) \in F(t, y(t))$.

Definition

Eine Funktion $y : [t_0, \infty) \rightarrow \mathbf{R}^n$ heißt absolut stetig, wenn die folgenden Eigenschaften gegeben sind:

1. $y(\cdot)$ ist fast überall differenzierbar,
2. $y'(\cdot)$ ist summierbar und
3. $y(t) = y(t_0) + \int_0^t y'(s) ds$.

Beispiel

Jede stetig differenzierbare Funktion $y(t)$ ist absolut stetig. Aber auch jede Funktion $y : [t_0, \infty) \rightarrow \mathbf{R}^n$ mit

$$\|y(t) - y(t')\| \leq K \|t - t'\|$$

hat diese Eigenschaft.

Beispiel

Gegeben sei die (unstetige) Differentialgleichung

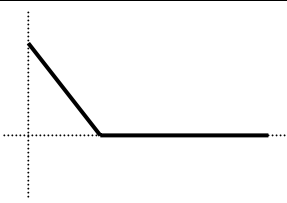
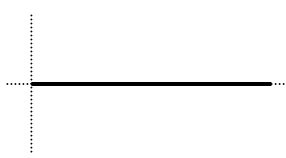
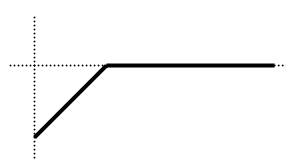
$$y' = -\operatorname{sgn}(y) + \alpha, \quad -1 < \alpha < 1$$

Viele Gründe sprechen dafür, anstatt dieser Aufgabe die folgende mehrwertige Differentialgleichung

$$y' \in F(y) = \begin{cases} \{1 + \alpha\} & \text{für } y < 0 \\ [-1 + \alpha, 1 + \alpha] & \text{für } y = 0 \\ \{-1 + \alpha\} & \text{für } y > 0 \end{cases}$$

zu betrachten.

Die folgende Tabelle zeigt die Lösungen der mehrwertigen Differentialgleichung zu verschiedenen Anfangsbedingungen

$y' \in F(t,y), \quad y(0) = y_0.$		
$y_0 = 1$	$y_0 = 0$	$y_0 = -1$
$1 + (-1 + \alpha)t$ für $0 \leq t \leq 1/(\alpha - 1)$ 0 für $t \geq 1/(\alpha - 1)$	$y(t) = 0$ für alle $t \geq 0$	$-1 + (1 + \alpha)t$ für $0 \leq t \leq 1/(1 + \alpha)$ 0 für $t \geq 1/(1 + \alpha)$
		

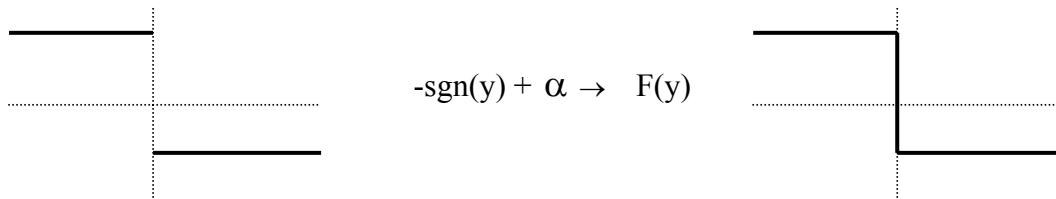
Dass die angegebenen Funktionen auch wirklich Lösungen der Anfangswertaufgaben sind, kann schnell eingesehen werden:

Lediglich die Abschnitte mit $y(t) = 0$ sind etwas ungewöhnlich. Für diese gilt aber ($-1 < \alpha < 1$) die Relation

$$y'(t) = y'(0) = 0 \in F(0) = [-1 + \alpha, 1 + \alpha].$$

Bemerkung

Im Beispiel ist die mehrwertige Differentialgleichung offenbar dadurch entstanden, dass der „Graph“ von $-\operatorname{sgn}(y) + \alpha$ „abgeschlossen“ wurde.



Das dieses sehr vernünftig ist zeigt sich daran, das $y(t) = 0$ die Beziehung

$$y' = -\text{sgn}(y) + \alpha$$

nur dann erfüllt, wenn $\text{sgn}(0) = \alpha$ ist. D.h. $\text{sgn}(0)$ muss dem Wert von α angepasst werden. Die Mehrwertigkeit von $\text{sgn}(0)$ erledigt diese Aufgabe von selber.

Satz 6.1 (Existenz- und Eindeutigkeit)

Gegeben sei eine (mehrwertige) Differentialgleichung der Form

$$y' \in F(t,y), \quad y(t_0) = y_0.$$

Dabei sei

$$F : [t_0, \infty) * \mathbf{R}^n \rightarrow P(\mathbf{R}^n)$$

eine nach oben halbstetige mengenwertige Abbildung in die Menge der nichtleeren, konvexen und abgeschlossenen Teilmengen aus \mathbf{R}^n .

Genügt die Abbildung F außerdem noch der Monotonieeigenschaft ($K \in \mathbf{R}$)

$$(\xi - \eta, u - v) \leq K (u - v, u - v)$$

für alle $u, v \in \mathbf{R}^n$ und alle $\xi \in F(t,u)$ und $\eta \in F(t,v)$, dann besitzt die Aufgabe eine eindeutig bestimmte Lösung für alle $t \geq t_0$.

Gegeben sei nun die unstetige Differentialgleichung

$$y' = f(t,y)$$

mit einer unstetigen Funktion

$$f : [t_0, \infty) * \mathbf{R}^n \rightarrow \mathbf{R}^n.$$

Es war ein Vorschlag von A.F. Filippov, durch die Konstruktion

$$F(t,y) = \bigcap_{\delta > 0} \bigcap_{\sigma(N)=0} \overline{\text{Konf}(t, U_\delta(y) - N)}$$

diese Aufgabe in eine mehrwertige zu überführen.

Von großer praktischer Bedeutung ist, dass i.a. die Ermittlung von Lösungen von unstetigen Differentialgleichungen ohne eine Überführung in die entsprechende mehrwertige Differentialgleichung möglich sein wird. Differenzenverfahren liefern regelhaft eine Lösung des entsprechenden mehrwertigen Problems.

Satz 6.2 (Taubert/1975)

Die mehrwertige (unstetige) Anfangswertaufgabe genüge den Bedingungen des Existenzsatzes. F sei außerdem lokal beschränkt, d.h. es existiert eine Konstante K , so dass für alle $(t,u) \in [t_0, t_0+T] \times \mathbf{R}$ und alle $s \in F(t,u)$ gilt $|s| \leq K$.
Gegeben sei das explizite Euler Verfahren

$$y_{n+1} = y_n + hg_n, \quad g_{n+i} \in F(t_{n+i}, y_{n+i}).$$

Dann gilt auf jedem kompakten Intervall $[t_0, t_0+T]$

$$\lim_{\substack{h \rightarrow 0 \\ n \rightarrow \infty \\ t_0 + nh \rightarrow t}} y^h = y(t),$$

d.h., es liegt Konvergenz vor.

Abschließende Bemerkungen

Im Allgemeinen können für die Integration von unstetigen Aufgaben die folgenden Verfahren gewählt werden:

1. Explizite RKV
2. Explizite und stark stabile MSV.

Bei den folgenden Methoden ist Vorsicht geboten:

1. Implizite RKV und stark stabile implizite MSV können verwendet werden, allerdings können die Verfahren bei der Auflösung der impliziten Gleichungen versagen
2. Extrapolationsverfahren auf der Basis der Mittelpunkformel (nicht stark stabil) sind ungeeignet.
3. Schrittweitensteuerungen können die Integration behindern oder sogar ausschließen.

§ 6.3 Symplektische Integration

Die Bewegung eines Körpers in einem Kraftfeld mit dem Potential V führt auf die Hamiltonfunktion

$$H(q,p) = \frac{p^2}{2m} + V(q)$$

mit der verallgemeinerten Koordinaten q und den verallgemeinerten Impuls $p = m\dot{q}$, $m > 0$,
Die totale Ableitung von H nach der Zeit

$$\frac{d}{dt} H = H(q, p)\dot{q} + H(q, p)\dot{p}.$$

und Nullsetzen führt zu den Hamiltonsystem

$$\begin{aligned}\dot{q} &= H_p(q, p) \\ \dot{p} &= -H_q(q, p).\end{aligned}$$

Entlang einer Lösung des Hamiltonsystems gilt dann (Erhaltungsgröße)

$$H(\xi) = \text{const}$$

Die Praxis zeigt, dass es mit Differenzenverfahren i.a. nicht möglich sein wird, diese Erhaltungsgröße über längere Zeiten zu gewährleisten. Erstaunlicherweise kann aber eine andere Erhaltungsgröße eingehalten werden.

Gegeben sei jetzt eine zweimal stetig differenzierbare Funktion $H : \mathbf{R}^2 \rightarrow \mathbf{R}^2$

$$H(q, p)$$

und das zugehörige Hamilton System.

Es bezeichne $q(t; t_0, (q^0, p^0))$ und $p(t; t_0, (q^0, p^0))$ die Lösung des Hamiltonsystems zu den Anfangsbedingungen $q(t_0) = q^0$ und $p(t_0) = p^0$.

Der Fluss φ_t des Hamiltonsystems wird dann definiert als

$$\varphi_t : \mathbf{R}^2 \rightarrow \mathbf{R}^2$$

mit

$$\varphi_t(q^0, p^0) = (q(t; 0, (q^0, p^0)), p(t; 0, (q^0, p^0)))$$

Für eine Menge von Anfangswerten $A \subset \mathbf{R}^2$, heißt

$$\varphi_t(A) = \{ \varphi_t(a) \mid a \in A \}$$

dann der Fluss von A zur Zeit t.

Ein Theorem (Liouville) besagt, dass für solche Systeme der Fluss „Volumeninvariant“ ist, d.h.:

$$\text{Vol}(\varphi_t(A)) = \text{Vol}(A)$$

Es soll lediglich angedeutet werden, wie dieses bewiesen werden kann:

Ist $y = (q, p)$ und $y' = f(y)$ das Hamiltonsystem, dann gilt

$$\text{Vol}(\varphi_t(A)) = \int_{\varphi_t(A)} dy = \int_A \left| \det \left(\frac{\partial y}{\partial y^0}(t, 0, y^0) \right) \right| dy^0.$$

Über „bekannte“ Sätze über Differentialgleichungen kann dann gezeigt werden, dass

$$\det \left| \frac{\partial y}{\partial y^0} (t, 0, y^0) \right| = \exp \left(\int_0^t \text{Spur}(f_y(y(s, 0, y^0))) ds \right)$$

Bei Hamiltonsysteme ist die Spur ($f_y(\cdot)$) stets gleich Null und damit die „Volumeninvarianz“ gezeigt.

Diese Erhaltungsgröße bleibt auch bei einer numerischen Integration mit der „one-leg“ Version der Trapezregel erhalten. In diesem Zusammenhang wird dieses Differenzenverfahren häufig auch als implizite Mittelpunktsregel bezeichnet

Für das Hamiltonsystem

$$y' = f(y), \quad y(0) = y_0$$

hat die „one leg“ Trapezregel die Form

$$y_{n+1} = y_n + hf((y_{n+1}+y_n)/2)$$

oder

$$y_{n+1} = y_n + hk \quad \text{mit} \quad k = f(y_n + (1/2)hk)$$

Ähnlich wie im kontinuierlichen Fall, benötigt der Nachweis der „Volumeninvarianz“ einige „bekannte“ Sätze aus der Theorie der Differentialformen. Unermüdliches Rechnen führt zum Ziel. Wir wollen hier darauf verzichten und stattdessen ein Beispiel angeben.

Beispiel

An einen masselosen Band der Länge l , sei eine Punktmasse der Masse m befestigt. Wir wollen das Pendel durch den Winkel α zwischen dem Band und dem Lot zum Boden beschreiben. J bezeichne das Trägheitsmoment des Massenpunktes. Mit der Erdbeschleunigung g wirkt dann ein Drehmoment $M = g m \sin(\alpha)$. Damit lautet die Bewegungsgleichung

$$J\ddot{\alpha} = g m \sin(\alpha) l,$$

Die Gesamtenergie des Systems ist gegeben durch

$$E(\alpha, \dot{\alpha}) = (1/2)J\dot{\alpha}^2 + mg(l - \cos(\alpha)l)$$

Mit den Bezeichnungen $p = J\dot{\alpha}$ (Drehimpuls) und $q = \alpha$ ergibt sich die Hamiltonfunktion

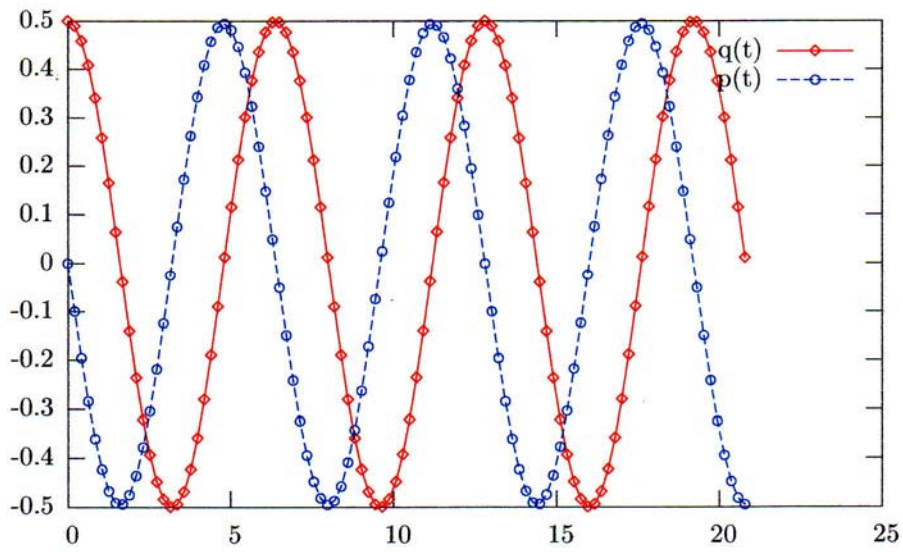
$$H(q,p) = \frac{p^2}{2J} + mg(l - \cos(q)l)$$

Mit g, l und $m = 1$ ergibt sich das Hamiltonsystem

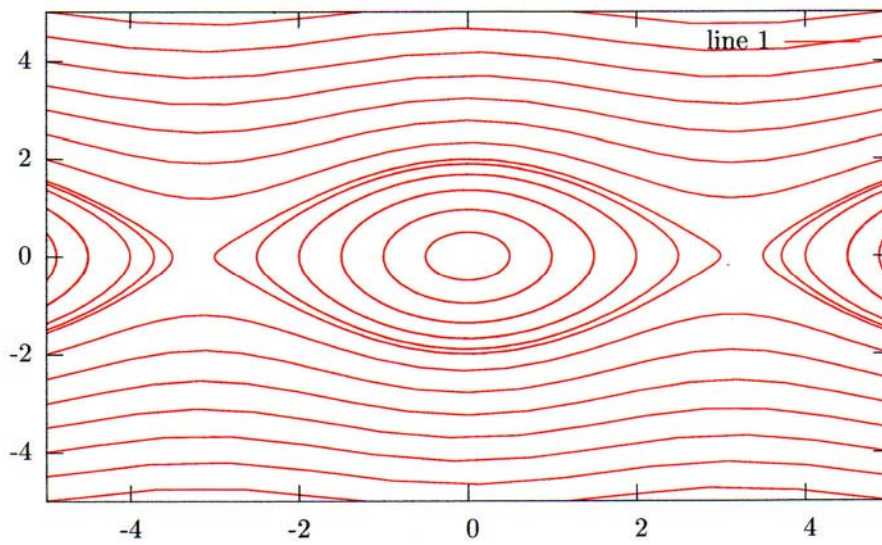
$$\begin{aligned} \dot{q} &= p \\ \dot{p} &= -\sin(q) \end{aligned}$$

Wird diese Aufgabe mit der impliziten Mittelpunktsformel integriert, dann ergibt sich*:

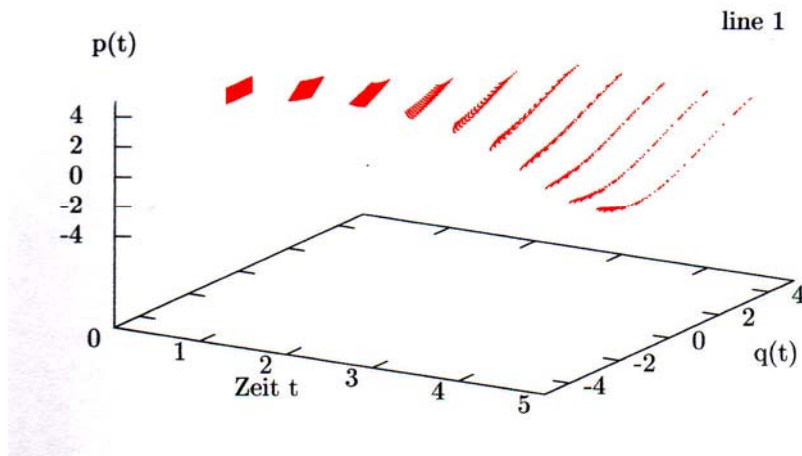
* Die Bilder wurden von Herrn Eike Scholz im Rahmen eines Proseminars im SS2003 angefertigt.



Der zugehörige Phasenplot hat die Form



Die „Volumeninvarianz“ zeigt das folgende Bild



§ 6.4 Algebra-Differentialgleichungen

Wir betrachten ein System von Gleichungen ($F : \mathbf{R}^n * \mathbf{R}^n * \mathbf{R} \rightarrow \mathbf{R}^n$)

$$F(u', u, t) = 0 \quad (*)$$

und nehmen an, dass f hinreichend oft differenzierbar ist.

Die Gleichung (*) hat entlang einer „Lösung“ den Differentiationsindex m , wenn m die kleinste Anzahl von Differentiationen

$$F(u', u, t) = 0, \quad \frac{d}{dt} F(u', u, t) = 0, \quad \dots, \quad \frac{d^m}{dt^m} F(u', u, t) = 0 \quad (**)$$

ist, so dass aus den $n * m$ Gleichungen (**) ein explizites System von n Differentialgleichungen

$$u' = \varphi(u, t)$$

entsteht.

Beispiel

Mit $f, g : \mathbf{R} * \mathbf{R} \rightarrow \mathbf{R}$ sei

$$F(y', z', y, z) = 0 \quad \Leftrightarrow \quad \begin{aligned} y' &= f(y, z) \\ 0 &= g(y, z) \end{aligned}$$

Ist $(y(t), z(t))$ eine „Lösung“ dieser Algebra-Differentialgleichung, dann gilt

$$\begin{aligned} y'(t) - f(y(t), z(t)) &= 0 \\ g(y(t), z(t)) &= 0. \end{aligned}$$

Andererseits ist

$$\frac{d}{dt} F(y', z', y, z) = 0 \quad \Leftrightarrow \quad \begin{aligned} y'' - f_y y' - f_z z' &= 0 \\ g_y y' + g_z z' &= 0. \end{aligned}$$

Ist $g_z \neq 0$, dann kann ein zugehöriges System von Differentialgleichungen erster Ordnung erzeugt werden

$$\begin{aligned} y' &= f(y, z) \\ z' &= -g_z^{-1} g_y y' = -g_z^{-1} g_y f(y, z). \end{aligned}$$

Die folgende Tabelle gibt Stellvertreter für Algebra-Differentialgleichungen vom Index 1-3 an:

Index	Nichtlinear	Bedingungen		Linear
1	$y' = f(y, z)$ $0 = g(y, z)$	$g_z \neq 0$	$a_{22} \neq 0$	$y' = a_{11}y + a_{12}z$ $0 = a_{21}y + a_{22}z$
2	$y' = f(y, z)$ $0 = g(y)$	$g_y f_z \neq 0$	$a_{12}a_{21} \neq 0$	$y' = a_{11}y + a_{12}z$ $0 = a_{21}y$
3	$y' = f(y, z)$ $z' = k(y, z, u)$ $0 = g(y)$	$g_y f_z k_u \neq 0$	$a_{12}a_{31}a_{23} \neq 0$	$y' = a_{11}y + a_{12}z$ $z' = a_{21}y + a_{22}z + a_{23}u$ $0 = a_{21}y$

Neben dem Differentiationsindex gibt es bei Algebra-Differentialgleichungen noch den Perturbationsindex:

Die Gleichung

$$F(u', u, t) = 0$$

hat den Perturbationsindex m entlang einer Lösung $u(\cdot)$ auf $[0, T]$, wenn m die kleinste Zahl ist mit

$$\| \hat{u}(t) - u(t) \| \leq C (\| \hat{u}(0) - u(0) \| + \text{Max}_{0 \leq \xi \leq T} \| \delta(\xi) \| + \dots + \text{Max}_{0 \leq \xi \leq T} \| \delta^{m-1}(\xi) \|)$$

für alle \hat{u} mit $F(\hat{u}'(t), \hat{u}(t), t) = \delta(t)$.

Bemerkung

Differentiationsindex und Perturbationsindex können sich im Fall $m > 1$ unterscheiden. Für lineare Systeme mit konstanten Koeffizienten stimmen Sie aber überein.

Algebra-Differentialgleichungen können sich – im Unterschied zu gewöhnlichen Differentialgleichungen – merkwürdig Verhalten. Dieses zeigen die folgenden Beispiele:

$Ax' + Bx = f(t)$		
$\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}x' + \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}x = \begin{pmatrix} f_1(t) \\ f_2(t) \end{pmatrix}$	$\begin{pmatrix} 5 & 0 \\ 0 & 0 \end{pmatrix}x' + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}x = \begin{pmatrix} 0 \\ 5 + \sin(t) \end{pmatrix}$	$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}x' + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}x = \begin{pmatrix} f_1(t) \\ f_2(t) \end{pmatrix}$
I.a. nicht lösbar	Lösbar nur für bestimmte Anfangswerte	Lösungen hängen von Ableitungen ab (f_2')

Die einfachste Methode zur numerischen Integration einer Algebra-Differentialgleichung

$$F(u', u, t) = 0$$

besteht darin, dass implizite Euler Verfahren zu benutzen

$$F\left(\frac{y_n - y_{n-1}}{h_n}, y_n, t_n\right) = 0.$$

Bei einem linearen System

$$Au' + Bu = f(t)$$

entsteht das lineare Gleichungssystem

$$(A + h_n B)y_n = Ay_{n-1} + h_n f(t_n).$$

Damit ergibt sich für y_n i.a. ein System von n nichtlinearen Gleichungen. Leider funktioniert diese Methode nicht immer (!), jedoch bei Index 1 Problemen gut.

Satz 6.3

Die kanonische Index 1 Aufgabe

$$\begin{aligned} y' &= f(y, z, t) \\ 0 &= g(y, z, t) \end{aligned}$$

werde mit dem impliziten Euler Verfahren gelöst. Der Anfangsfehler sei von der Größenordnung $O(h)$, dann gilt für den globalen Fehler

$$y_n - y(t_n) = O(h) \quad \text{und} \quad z_n - z(t_n) = O(h)$$

für $t_n - t_0 = nh < K$.

Dieses Ergebnis ist leicht einsehbar. Die numerische Integration mit dem impliziten Euler Verfahren entspricht einer Integration von gewöhnlichen Differentialgleichungen mit diesen Verfahren (!):

Betrachte

$$\begin{aligned} y' &= f(y, z, t) \\ 0 &= g(y, z, t) \end{aligned}$$

mit invertierbarem g_z . Mit dem Satz über implizite Funktionen gilt dann

$$y' = f(y, h(y,t), t).$$

Das implizite Euler Verfahren hat die Form

$$\begin{aligned} \frac{y_n - y_{n-1}}{h_n} &= f(y_n, z_n, t_n) \\ 0 &= g(y_n, z_n, t_n). \end{aligned}$$

Auflösung der zweiten Gleichung nach z_n liefert

$$\frac{y_n - y_{n-1}}{h_n} = f(y_n, h(y_n, t_n), t_n).$$

Damit liegt aber eine bereits bekannte Situation vor.

Bemerkungen:

Auf zwei Probleme soll noch hingewiesen werden:

Bei der praktischen Integration einer Algebro-Differentialgleichung sind Anfangsbedingungen i.a. nicht bekannt (!)

Bei der Auflösung der Gleichungen mit dem Newton-Verfahren (nichtlinearer Fall) hat die Funktionalmatrix die Neigung singular zu werden (!)

In der Regel sind (implizite) Verfahren für die Integration von steifen Systemen auch für die Integration von Algebro-Differentialgleichungen „gut“.

Numerische Behandlung von gewöhnlichen Differentialgleichungen

Eine Einführung

Universität Hamburg
SoSe07

K. Taubert

Numerik gewöhnlicher Randwertaufgaben

7 DIFFERENZENVERFAHREN ZUR AUFLÖSUNG VON GEWÖHNLICHEN LINEAREN RANDWERTAUFGABEN

Eine in vielen Anwendungen auftretende gewöhnliche Differentialgleichung hat die Form

$$L(y) = - (p(x)y')' + q(x)y = r(x)$$

Dabei sei (zunächst) $p(\cdot) > 0$ eine stetig differenzierbare Funktion von $[a,b]$ in \mathbf{R} . Die Funktionen $r(\cdot)$ und $q(\cdot)$ seien stetige Funktionen von $[a,b]$ in \mathbf{R} und $q(\cdot) \geq 0$.

Um die gesuchte Lösungsfunktion dieser Differentialgleichungen eindeutig festzulegen, müssen noch Randbedingungen angegeben werden. Typische Randbedingungen sind:

1. $y(a) = d_1, \quad y(b) = d_2$ (Dirichlet-Bedingung)
2. $-y'(a) = d_1, \quad y'(b) = d_2$ (Neumann-Bedingung)
3. $-y'(a) + cy(a) = d_1, \quad y'(b) + dy(b) = d_2$ (Cauchy-Bedingung).

Dabei seien die d_i reelle Zahlen und $c, d \geq 0$. Der Fall $q = 0$ mit $c = d = 0$ (Neumann-Bedingung mit $q=0$) sei ausgeschlossen.

Bei der Diskretisierung von Randwertaufgaben mit Differenzenverfahren spielen „Stabilitätsungleichungen“ eine zentrale Rolle. Dieses soll hier nur an einem einfachen Beispiel vorgeführt werden.

Exemplarisch behandeln wir also eine Diskretisierung der Aufgabe

$$-y'' = f(x)$$

mit den Randbedingungen

$$y(0) = \gamma_1, \quad y(1) = \gamma_2.$$

Es sei $h = 1/N$ eine vorgegebene Schrittweite mit $N \in \mathbf{N}$ und $x_i = ih$ die Menge der zugehörigen Gitterpunkte aus $[0,1]$ mit $i = 0, 1, 2, \dots, N$.

Wird die zweite Ableitung an den Stellen x_1, x_2, \dots, x_{N-1} durch den zweiten Differenzenquotient

$$y''(x_j) \cong \frac{y(x_{j+1}) - 2y(x_j) + y(x_{j-1}))}{h^2}$$

ersetzt, dann entsteht ein System von N-1 Gleichungen

$$-\frac{y_{j+1} - 2y_j + y_{j-1}}{h^2} = f(x_j), \quad j=1,2,\dots, N-1.$$

Durch Berücksichtigung der Randbedingungen entsteht insgesamt ein Gleichungssystem der Form

$$\frac{1}{h^2} \begin{pmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & 0 \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & \ddots \\ 0 & & & -1 & 2 & -1 \\ & & & & -1 & 2 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ \vdots \\ y_{N-2} \\ y_{N-1} \end{pmatrix} = \begin{pmatrix} f(x_1) \\ f(x_2) \\ \vdots \\ \vdots \\ f(x_{N-2}) \\ f(x_{N-1}) \end{pmatrix} + \frac{1}{h^2} \begin{pmatrix} \gamma_1 \\ 0 \\ \vdots \\ \vdots \\ 0 \\ \gamma_2 \end{pmatrix}$$

oder kurz

$$\frac{1}{h^2} A_h y_h = f_h + \frac{1}{h^2} b_h$$

Einen Zugang zu diesem Gleichungssystem liefert u.a. der

Satz 7.1 (Stabilitätsungleichung)

Für alle $y_h = (y_0^h, y_1^h, \dots, y_N^h) \in \mathbf{R}^{N+1}$ und alle $h = 1/N, N \in \mathbf{N}$, sei

$$\|y_h\|_h := \text{Max}_{i=0,1,\dots,N} |y_i^h|,$$

$$\|y_h\|^h := \text{Max}_{i=1,\dots,N-1} |y_i^h|,$$

$$L_h y_h := \frac{1}{h^2} \begin{pmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & 0 \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & \ddots \\ 0 & & & -1 & 2 & -1 \\ & & & & -1 & 2 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ \vdots \\ y_{N-2} \\ y_{N-1} \end{pmatrix} \quad \text{und} \quad \begin{matrix} R_1 y_h := y_0^h \\ R_2 y_h := y_N^h \end{matrix}.$$

Dann ist

$$\|y_h\|_h \leq K (\|L_h y_h\|^h + |R_1 y_h| + |R_2 y_h|)$$

mit einer von h unabhängigen Konstanten K.

Beweis

Für jedes $y_h \in \mathbf{R}^{N+1}$ gilt

$$-y_{j+1}^h + 2y_j^h - y_{j-1}^h = h^2 \varepsilon_j, \quad j = 1, 2, \dots, N-1,$$

mit geeigneten ε_j .

Der Vektor y_h lässt also eine Darstellung der Form

$$y_j^h = c_1(1)^j + c_2j + z_j, \quad j = 0, 1, 2, \dots, N,$$

zu. Dabei ist $(z_j)_{j=0,1,\dots,N}$ eine spezielle Lösung der Differenzgleichung und $c_1(1)^j + c_2j$ die allgemeine Lösung der homogenen Differenzgleichung.

Eine spezielle Lösung der Differenzgleichung ergibt sich aus

$$\begin{pmatrix} z_{n+1} \\ z_n \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} z_n \\ z_{n-1} \end{pmatrix} + h^2 \begin{pmatrix} \varepsilon_n \\ 0 \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix}^n \begin{pmatrix} z_1 \\ z_0 \end{pmatrix} + h^2 \sum_{v=0}^{n-1} \begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix}^v \begin{pmatrix} \varepsilon_{n-v} \\ 0 \end{pmatrix}.$$

Die spezielle Wahl $(z_1, z_0)^T = (0, 0)^T$ liefert

$$\begin{pmatrix} z_{n+1} \\ z_n \end{pmatrix} = h^2 \sum_{v=0}^{n-1} \begin{pmatrix} v+1 & -v \\ v & -(v-1) \end{pmatrix} \begin{pmatrix} \varepsilon_{n-v} \\ 0 \end{pmatrix}.$$

Mit $\|u\| = \max(|u_1|, |u_2|)$ für $u = (u_1, u_2)^T$ ergibt sich

$$\left\| \begin{pmatrix} z_{n+1} \\ z_n \end{pmatrix} \right\| \leq h^2 N(2N+1) \max_v \left\| \begin{pmatrix} \varepsilon_{n-v} \\ 0 \end{pmatrix} \right\| = \frac{1}{N^2} N(2N+1) \|L_h y_h\|^h \leq \tilde{K} \|L_h y_h\|^h,$$

mit einer von h unabhängigen Konstanten \tilde{K} .

Die noch fehlenden Koeffizienten c_1 und c_2 werden festgelegt durch

$$\begin{aligned} y_0^h &= c_1 1^0 + c_2 0, \\ y_N^h &= c_1 1^N + c_2 N + z_N. \end{aligned}$$

Es gilt mithin

$$y_j^h = y_0^h + \left(\frac{y_N^h - y_0^h}{N} - \frac{z_N}{N} \right) j + z_j$$

und somit die gewünschte Abschätzung

$$\|y_h\|_h \leq K (\|L_h y_h\|^h + |R_1 y_h| + |R_2 y_h|)$$

mit einer von h unabhängigen Konstanten K .

Wichtig ist nun, was alles aus der Stabilitätsungleichung gefolgert werden kann:

Korollar 7.1 (Eindeutige Lösbarkeit)

Die Aufgabe

$$\frac{1}{h^2} A_h y_h = f_h + \frac{1}{h^2} b_h$$

ist eindeutig lösbar.

Beweis:

Die vollhomogene Aufgabe

$$\begin{aligned} L_h y_h &= 0, \\ R_1 y_h &= 0, \\ R_2 y_h &= 0, \end{aligned}$$

besitzt wegen der Stabilitätsungleichung nur die triviale Lösung.

Korollar 7.2 (Gleichmäßige Beschränktheit der Inversen)

Es gilt

$$\| h^2 A_h^{-1} \| \leq K$$

für alle h . Dabei sei die Matrixnorm $\| \cdot \|$ die zur Vektornorm $\| \cdot \|$ zugeordnete Matrixnorm.

Beweis:

Die Wahl $\gamma_1 = \gamma_2 = 0$ und die eindeutige Lösbarkeit des Gleichungssystems

$$\frac{1}{h^2} A_h y_h = f_h,$$

liefert $\| (\frac{1}{h^2} A_h)^{-1} f_h \| = \| y_h \| \leq K \| L_h y_h \| = K \| f_h \|$ und damit die Behauptung.

Korollar 7.3 (Konvergenzsatz)

Es sei $y(\cdot)$ die Lösung von $-y'' = f(x)$ mit den Randbedingungen $y(0) = \gamma_1$, $y(1) = \gamma_2$.

Dabei sei f eine stetige Funktion $f: [0,1] \rightarrow \mathbf{R}$ und $\gamma_1, \gamma_2 \in \mathbf{R}$.

Für die Lösungen y_h des Differenzenverfahrens

$$\frac{1}{h^2} A_h y_h = f_h + \frac{1}{h^2} b_h$$

gilt dann

$$\lim_{h \rightarrow 0} \max_{i=0, \dots, N} | y(x_i) - y_i^h | = 0.$$

Ist $y(\cdot)$ sogar viermal stetig differenzierbar, dann gilt

$$\max_{i=0, \dots, N} | y(x_i) - y_i^h | = O(h^2).$$

Beweis

Einsetzen der Lösung der Randwertaufgabe in das Differenzenverfahren und Taylorentwicklung liefern

$$-y(x_j+h) + 2y(x_j) - y(x_j-h) - h^2 f(x_j) = o(h^2), \quad j = 1, 2, 3, \dots, N-1$$

(bzw. $O(h^4)$).

Es sei \bar{y} die Restriktion der Lösung der Randwertaufgabe auf die Gitterpunkte $0, x_1, x_2, \dots, x_{N-1}, 1$.

Dann folgt

$$A_h \bar{y} = h^2 f_h + b_h + o(h^2) \quad (\text{bzw. } O(h^4)).$$

Außerdem gilt

$$A_h y_h = h^2 f_h + b_h.$$

Für die Differenz gilt also

$$A_h (\bar{y} - y_h) = o(h^2) \quad (\text{bzw. } O(h^4)).$$

Wegen $\|h^2 A_h^{-1}\|^h \leq K$, folgt die Behauptung des Korollars.

Bemerkungen

Die hier am einfachen Beispiel vorgeführte Technik ist sehr verallgemeinerungsfähig¹.

In unserem Beispiel hat das Gleichungssystem auch noch die Eigenschaft von Monotoner Art zu sein, d.h. es gilt

$$A_h u_h \geq 0 \Rightarrow u_h \geq 0,$$

wobei die Ungleichungen für jede Komponente zu verstehen sind. Diese Eigenschaft liefert - unabhängig von den Stabilitätsungleichungen - ebenfalls einen mächtigen Hebel zur Behandlung von (auch nichtlinearen) Randwertaufgaben.

¹ H.-O. Kreiss: Difference Approximations for Boundary and Eigenvalue Problems for Ordinary Differential Equations. Mathematics of Computations. 1972

Numerische Behandlung von gewöhnlichen Differentialgleichungen

Eine Einführung

Universität Hamburg
SoSe07

K. Taubert

Numerik gewöhnlicher Randwertaufgaben

8 VARIATIONSGLEICHUNGEN und VARIATIONS-AUFGABEN. VARIATIONSUNGLEICHUNGEN

Wir betrachten erneut die gewöhnliche Differentialgleichung

$$-(p(x)y')' + q(x)y = r(x) \quad x \in (a,b)$$

mit den homogenen Randbedingungen (Dirichlet-Aufgabe)

$$y(a) = 0, \quad y(b) = 0.$$

Dabei sei (zunächst) $p(\cdot) > 0$ eine stetig differenzierbare Funktion von $[a,b]$ in \mathbf{R} . Die Funktionen $r(\cdot)$ und $q(\cdot)$ seien stetige Funktionen von $[a,b]$ in \mathbf{R} und $q(\cdot) \geq 0$.

Lösungen solcher Aufgaben sind Funktionen $y(\cdot) \in C^2[a,b]$, welche punktweise sowohl der Differentialgleichung als auch den Randbedingungen genügen.

Entscheidend in diesem Kapitel ist nun, dass diese Randwertaufgaben auch anders (schwach) formuliert werden können. Diese neue Formulierung beinhaltet nicht nur die alte Aufgabenstellung sondern führt sogar dazu, dass die Bedingungen an $p(\cdot)$, $q(\cdot)$ und $r(\cdot)$ abgeschwächt werden können. Vom physikalischen Standpunkt her, muss sogar gesagt werden, dass die ursprüngliche Formulierung häufig zu einschränkend ist.

Es sei $V = H_0^1[a,b] \subset H^1[a,b]$ der Raum der absolut stetigen Funktionen mit $v(a)=v(b) = 0$, deren Ableitungen quadratisch integrierbar sind.

Mit der Bilinearform $a : V \times V \rightarrow \mathbf{R}$

$$a(u,v) = \int_a^b (pu'v' + quv) dx$$

und der Linearform $F : V \rightarrow \mathbf{R}$

$$F(v) = \int_a^b rv dx$$

kann die Dirichlet-Aufgabe in die folgenden zwei äquivalenten Aufgaben umgeschrieben werden:

$$\left\{ \begin{array}{l} \text{Gesucht ist ein } u \in V \text{ mit} \\ a(u, v) = F(v) \quad \forall v \in V. \end{array} \right. \quad (\text{Variationelle Aufgabe})$$

und

$$\left\{ \begin{array}{l} \text{Gesucht ist ein } u \in V \text{ mit} \\ E(u) \leq E(v) \quad \forall v \in V. \end{array} \right. \quad (\text{Variationsaufgabe})$$

Dabei ist $E : V \rightarrow \mathbf{R}$ das „Energiefunktional“

$$E(v) = (1/2)a(v, v) - F(v).$$

Bemerkung:

In der klassischen Literatur findet man häufig die folgende ausgeschriebene Form der Variationsaufgabe:

„Gesucht ist ein $\varphi \in V$ mit

$$J[\varphi] = \text{Min!}$$

und

$$J[\varphi] = \int_a^b (p\varphi'^2 + q\varphi^2 - 2r\varphi) dx \text{ “.}$$

Betrachtet man die „schwachen“ Aufgaben jetzt einmal unabhängig von der ursprünglichen Randwertaufgabe, dann fällt folgendes auf:

- Bei den „schwachen“ Formulierungen kann auf die stetige Differenzierbarkeit von $p(\cdot) > 0$ verzichtet werden. Die Funktionen $q(\cdot) \geq 0$ und $r(\cdot)$ brauchen außerdem nur quadratisch integrierbar zu sein. Die ursprüngliche Aufgabenstellung ist damit verallgemeinert worden.
- Die ursprüngliche Aufgabe wird durch die variationelle Formulierung und den Minimierungsaspekt völlig anders interpretiert.

Die „schwachen“ Formulierungen führen aber auch zu völlig anderen Ansätzen zur Lösung der ursprünglichen Aufgabe:

Für einen endlich dimensionalen Teilraum V_h von V können z.B. die folgenden Aufgaben betrachtet werden

$$\left\{ \begin{array}{l} \text{Gesucht ist ein } u \in V_h \text{ mit} \\ a(u, v) = F(v) \quad \forall v \in V_h. \end{array} \right. \quad (\text{Galerkin-Verfahren})$$

$$\left\{ \begin{array}{l} \text{Gesucht ist ein } u \in V_h \text{ mit} \\ E(u) \leq E(v) \quad \forall v \in V_h. \end{array} \right. \quad (\text{Ritz-Verfahren})$$

oder entsprechend ausgeschrieben

$$\text{„Gesucht ist ein } \varphi \in V_h \text{ mit } J[\varphi] = \text{Min! und } J[\varphi] = \int_a^b (p\varphi'^2 + q\varphi^2 - 2r\varphi) dx \text{ “.}$$

Bemerkung

Diese Interpretation und die damit verbundene Möglichkeit zur approximativen Lösung der Dirichlet-Aufgabe ist nicht auf diese Klasse von Aufgaben beschränkt. Besonders wirksam wird diese Methode bei den partiellen Differentialgleichungen (Finite Elemente).

§ 8.1 Abstrakte lineare Variationsprobleme

Es sei V ein Vektorraum mit einem inneren Produkt (\cdot, \cdot) und $\|\cdot\|$ die durch $(\cdot, \cdot)^{1/2}$ gegebene Norm.

Es seien außerdem $a : V \times V \rightarrow \mathbb{R}$ und $F : V \rightarrow \mathbb{R}$ zwei Abbildungen mit den folgenden Eigenschaften:

- a ist bilinear, d.h. eine lineare Abbildung in jedem Argument;
 - a ist beschränkt, d.h. es existiert ein M mit $a(u, v) \leq M \|u\| \|v\| \quad \forall u, v \in V$;
 - a ist symmetrisch, d.h. $a(u, v) = a(v, u) \quad \forall u, v \in V$;
 - a ist koerziv, d.h. es existiert ein $\mu > 0$ mit $a(u, u) \geq \mu \|u\|^2 \quad \forall u \in V$.
-
- F ist linear;
 - F ist beschränkt.

Wir betrachten nun die beiden Aufgaben

$$\left\{ \begin{array}{l} \text{Gesucht ist ein } u \in V \text{ mit} \\ a(u, v) = F(v) \quad \forall v \in V. \end{array} \right. \quad (\text{Variationelle Aufgabe})$$

$$\left\{ \begin{array}{l} \text{Gesucht ist ein } u \in V \text{ mit} \\ E(u) \leq E(v) \quad \forall v \in V. \end{array} \right. \quad (\text{Variationsaufgabe})$$

Dabei sei $E : V \rightarrow \mathbb{R}$ das „Energiefunktional“

$$E(v) = (1/2)a(v, v) - F(v).$$

Der folgende Satz zeigt, dass die beiden Aufgaben äquivalent sind

Satz 8.1

Ist V ein Hilbertraum, dann sind die beiden Aufgaben äquivalent und haben eine eindeutige Lösung

Beweis

Durch $a(\cdot, \cdot)$ wird ein inneres Produkt auf V definiert. Die zugehörige Norm $\|\cdot\|_a = a(\cdot, \cdot)^{1/2}$ ist wegen

$$\mu \|v\|^2 \leq \|v\|_a^2 \leq M \|v\|^2 \quad \forall v \in V;$$

äquivalent zu $\|\cdot\|$. Damit ist V auch ein Hilbertraum bezüglich $\|\cdot\|_a$.

Der Darstellungssatz von Riesz liefert deshalb ein eindeutiges $v_F^a \in V$ mit

$$F(v) = a(v_F^a, v) \quad \forall v \in V.$$

Damit löst v_F^a in eindeutiger Weise die Aufgabe

$$a(u, v) = F(v) = a(v_F^a, v) \quad \forall v \in V.$$

Ist u eine Lösung von

$$a(u, v) = F(v) \quad \forall v \in V,$$

dann gilt auch

$$a(u, u-v) = F(u-v) \quad \forall v \in V.$$

Die Ungleichung

$$a(u, u-v) \leq F(u-v) + (1/2)a(u-v, u-v) \quad \forall v \in V,$$

liefert

$$a(u, u) - a(u, v) \leq F(u) - F(v) + (1/2)a(u, u) - a(u, v) + (1/2)a(v, v) \quad \forall v \in V$$

und damit

$$E(u) \leq E(v) \quad \text{für alle } v \in V.$$

Es bleibt noch zu zeigen, dass u die einzige Lösung der Minimierungsaufgabe ist:

Es sei also $E(u_1) = E(u_2) \leq E(v) \quad \forall v \in V$. Dann ist auch

$$E(u_1) \leq E((1/2)(u_1+u_2)) \quad \text{und} \quad E(u_2) \leq E((1/2)(u_1+u_2)).$$

Die Addition der beiden Ungleichungen ergibt

$$(1/2)a(u_1, u_1) + (1/2)a(u_2, u_2) \leq (1/4)(a(u_1, u_1)+a(u_2, u_2)+a(u_1, u_2)+a(u_2, u_1))$$

und damit den Widerspruch

$$(1/4)a(u_1 - u_2, u_1 - u_2) \leq 0.$$

Ist u eine Lösung der Minimierungsaufgabe, dann erfüllt (Differentiation) notwendig die Bedingung

$$a(u, v) = F(v) \quad \text{für alle } v \in V.$$

Bemerkung

Das ursprüngliche Dirichlet Problem und $V = H_0^1[a, b]$ führen zu den im Satz 8.1 geforderten Bedingungen. Unabhängig von der ursprünglichen Aufgabe sind die Bedingungen des Satzes aber auch dann erfüllt, wenn $p(\cdot) > 0$ lediglich stetig ist und die Funktionen $q(\cdot) \geq 0$ und $r(\cdot)$ quadratisch integrierbar sind.

§ 8.2 „Finite Elementmethode“ mit stückweise linearen Funktionen

Wir betrachten wiederum die Dirichlet Aufgabe

$$\begin{aligned} -(p(x)y')' + q(x)y &= r(x) & x \in (a, b) \\ y(a) &= 0 \quad \text{und} \quad y(b) = 0. \end{aligned}$$

in seiner schwachen Formulierung:

„Gesucht ist ein $u \in V = H_0^1[a, b]$ mit

$$a(u, v) = \int_a^b (pu'v' + quv)dx = \int_a^b r v dx = F(v) \quad \text{für alle } v \in V = H_0^1[a, b]$$

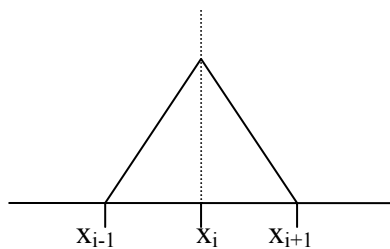
Dabei sei $p \in C[a, b]$ mit $p(x) > 0$ und $q, r \in L^2[a, b]$ mit $q \geq 0$.

Die „Finite Elementmethode“ approximiert die Lösung dieser Aufgaben über besondere endlichdimensionale Teilräume $V_h \subset V$. Exemplarisch soll als endlichdimensionaler Raum V_h der Raum der stückweise linearen Funktionen gewählt werden:

Das Intervall $[a, b]$ wird durch $a = x_0 < x_1 < \dots < x_{n-1} < x_n = b$ in Teilintervalle $I_j = (x_{j-1}, x_j)$ mit $j = 1, \dots, n$ zerlegt. Die x_j werden dabei als Knoten bezeichnet. Die Feinheit der Unterteilung wird durch $h = \max_{j=1, \dots, n} h_j$ gemessen, wobei $h_j = x_j - x_{j-1}$ die Länge des Teilintervalls I_j angibt. Der Teilraum V_h ist dann:

$$V_h = \{ v \in C[a, b] \mid v \text{ ist linear auf } I_j, j = 1, \dots, n \} \subset V.$$

Als Basis für V_h können die Funktionen $\varphi_j \in V_h, j = 1, 2, \dots, n-1$ mit



$$\varphi_j(x_i) = \begin{cases} 1 & \text{falls } i = j \\ 0 & \text{falls } i \neq j. \end{cases}$$

verwendet werden. Eine beliebige Funktion $v \in V_h$ mit $v(x_i) = x_i$ besitzt dann die Darstellung

$$v(x) = \sum_{i=1}^{n-1} v_i \varphi_i(x) \quad \text{mit } x \in [a, b].$$

Bemerkung

Typisch für die „Finite Elementmethode“ ist, dass die Basiselemente φ_j von V_h nur auf wenigen (zusammenhängenden) Teilintervallen I_j , den so genannten Elementen, von Null verschieden sind. Dies hat zur Folge, dass die meisten Produkte $\varphi_j \cdot \varphi_i$ identisch Null sind und damit eine später entstehende Systemmatrix zur Berechnung der Näherungslösung dünn besetzt ist.

Galerkin-Verfahren

Die Lösung der Aufgabe

„Gesucht ist ein $u_h \in V_h \subset H_0^1[a, b]$ mit

$$a(u_h, v) = \int_a^b (p u_h' v' + q u_h v) dx = \int_a^b r v dx = F(v) \quad \text{für alle } v \in V_h \subset H_0^1[a, b]$$

ist gegeben durch

$$a(u_h, \varphi_j) = (r, \varphi_j), \quad j=1, \dots, n-1.$$

Besitzt die gesuchte Lösung u_h die Darstellung

$$u_h(x) = \sum_{i=1}^{n-1} \xi_i \varphi_i(x),$$

dann ergeben sich die gesuchten ξ_i aus dem folgendem $(n-1) \times (n-1)$ linearen Gleichungssystem in $\xi = (\xi_1, \xi_2, \dots, \xi_{n-1})^T$

$$A \xi = b$$

mit den Lastvektor

$$b = ((r, \varphi_1), \dots, (r, \varphi_{n-1}))^T$$

und der Steifigkeitsmatrix

$$A = (a(\varphi_i, \varphi_j))_{i,j=1,2,\dots,n-1}$$

Bemerkungen

Die Matrix A ist wegen der Koerzitivität von $a(\cdot, \cdot)$ und nach Konstruktion natürlich eine symmetrische positiv definite Matrix.

Aufgrund der Definition der Basisfunktionen φ_i erhält man

$$a(\varphi_i, \varphi_j) = \begin{cases} \neq 0 & \text{für } |i-j| \leq 1 \\ = 0 & \text{sonst,} \end{cases}$$

d.h. die Steifigkeitsmatrix A besitzt Tridiagonalform.

Die Bezeichnungen von A und b gehen auf Anwendungen in der Strukturmechanik zurück, in denen die finite Elementmethode ihren Ursprung besitzt.

Ritz Verfahren

Gegeben sei

$$E(v) = (1/2) \int_a^b (p v' v' + q v v) dx - \int_a^b r v dx$$

und die Variationsaufgabe

„Gesucht ist ein $u_h \in V_h \subset H_0^1[a, b]$ mit

$$E(u_h) \leq E(v) \quad \text{für alle } v \in V_h \subset H_0^1[a, b]$$

Besitzt die gesuchte Lösung u_h die Darstellung

$$u_h(x) = \sum_{i=1}^{n-1} \xi_i \varphi_i(x),$$

dann ergeben sich die gesuchten ξ_i aus der Minimierung folgendes quadratischen Funktionals in $\xi = (\xi_1, \xi_2, \dots, \xi_{n-1})^T$

$$\frac{1}{2} \xi^T A \xi - b^T \xi = \text{Min!}$$

mit der Steifigkeitsmatrix und dem Lastvektor aus dem Galerkin Verfahren.

§ 8.3 Eine Fehlerabschätzung

Wir betrachten wiederum die Dirichletaufgabe

$$-(p(x)y')' + q(x)y = r(x)$$

$$y(a) = y(b) = 0.$$

in seiner schwachen Formulierung. Es sei also $p \in C[a,b]$ mit $p(x) > 0$ und $q, r \in L^2[a,b]$ mit $q(x) \geq 0$.

Wir bemerken noch, dass durch die Bilinearform

$$a(u,v) = \int_a^b (p(x)u'v' + q(x)uv)dx \quad \text{für alle } u,v \in V = H_0^1[a,b]$$

ein inneres Produkt auf $H_0^1[a,b]$ gegeben und durch $\|v\|_a = \sqrt{a(v,v)}$ eine Norm in $H_0^1[a,b]$.

Satz 8.2

Es sei u die Lösung des Dirichlet Problems in seiner variationellen Formulierung und u_h eine mit dem Ritz- oder dem Galerkin Verfahren ermittelte Näherungslösung, dann gilt für alle $v \in V_h \subset H_0^1[a,b]$

1. $a(u-u_h, v) = 0,$
2. $\|u-u_h\|_a^2 + \|u_h\|_a^2 = \|u\|_a^2,$
3. $\|u-u_h\|_a \leq \|u-v\|_a.$

Beweis:

Es gilt $a(u,v) = F(v)$ und $a(u_h, v) = F(v)$ für alle $v \in V_h \subset V \subset H_0^1[a,b]$ und durch Differenzenbildung die erste Behauptung

$$a(u-u_h, v) = 0 \quad \text{für alle } v \in V_h.$$

$$\begin{aligned} \text{Es ist } \|u-u_h\|_a^2 + \|u_h\|_a^2 &= a(u-u_h, u-u_h) + a(u_h, u_h) = a(u-u_h, u) - a(u-u_h, u_h) + a(u_h, u_h) \\ &= a(u, u) - a(u_h, u) + a(u_h, u_h) = a(u, u) = \|u\|_a^2. \end{aligned}$$

Es sei $w = u_h - v \in V_h$, dann gilt

$$\|u-u_h\|_a^2 = a(u-u_h, u-u_h) + a(u-u_h, w) = a(u-u_h, u-u_h+w) = a(u-u_h, u-v) \leq \|u-v\|_a \|u-u_h\|_a.$$

Bemerkungen

Die erste Behauptung des Satzes besagt, dass $u-u_h$ senkrecht auf V_h bezüglich des Skalarproduktes $a(\cdot, \cdot)$ steht.

Die letzte Behauptung besagt, dass u_h diejenige Funktion aus V_h ist, die der Lösung u bezüglich des durch $\|\cdot\|_a$ gemessenen Abstandes am nächsten liegt.

Die Abschätzung $\|u-u_h\|_a \leq \|u-v\|_a$ für alle $v \in V_h$ kann zur quantitativen Abschätzung des Fehlers der Näherungslösung u_h genutzt werden. Die Idee besteht darin, ein geeignetes $v \in V_h$ zu finden, bzgl. dessen man den Fehler berechnen kann. Wählt man den linearen Spline π_u , der die Funktionswerte von u in den Knoten x_i interpoliert, so kann der Fehler über den Interpolationsfehler $f_\pi(h)$ abgeschätzt werden.

Hilfssatz 8.1

Ist u zweimal stetig differenzierbar, so gilt für den, bzgl. der äquidistanten Knoten $a = x_0 < x_1 < \dots < x_n = b$, interpolierenden linearen Spline π_u mit $h = (b-a)/n$

$$\max_{x \in [a, b]} |u(x) - \pi_u(x)| \leq \frac{h^2}{8} \max_{x \in [a, b]} |u''(x)| \quad \text{und}$$
$$\max_{x \in [a, b]} |u'(x) - \pi'_u(x)| \leq h \max_{x \in [a, b]} |u''(x)|.$$

Satz 8.3

Für den Fehler der „Finite Elementlösung“ u_h mit linearen Elementen gilt für zweimal stetig differenzierbares u

$$\|u-u_h\|_a^2 = a(u-u_h, u-u_h) \leq Ch^2 \max_{x \in [a, b]} |u''(x)|^2.$$

Beweis:

Es ist

$$\begin{aligned} a(u-u_h, u-u_h) &\leq a(u-\pi_u, u-\pi_u) = \int_a^b (p(x)(u'(x) - \pi'_u(x))^2 + q(x)(u(x) - \pi_u(x))^2) dx \\ &\leq \max_{x \in [a, b]} |p(x)(u'(x) - \pi'_u(x))^2 + q(x)(u(x) - \pi_u(x))^2| |b-a| \\ &\leq (b-a) \left(\max_{x \in [a, b]} |p(x)| + \frac{h^2}{64} \max_{x \in [a, b]} |q(x)| \right) h^2 \max_{x \in [a, b]} |u''(x)|^2 \\ &\leq C h^2 \max_{x \in [a, b]} |u''(x)|^2. \end{aligned}$$

§ 8.4 Variationsungleichungen

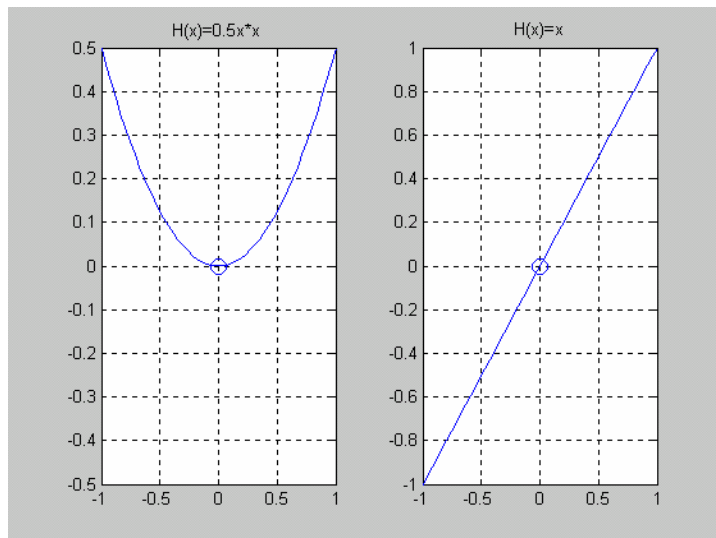
Bei den Variationsungleichungen handelt es sich um Minimierungsaufgaben mit Nebenbedingungen.

Es ist vielleicht zweckmäßig, den elementaren theoretischen Hintergrund von Variationsungleichungen zunächst an einer ganz einfachen Aufgabe zu erläutern:

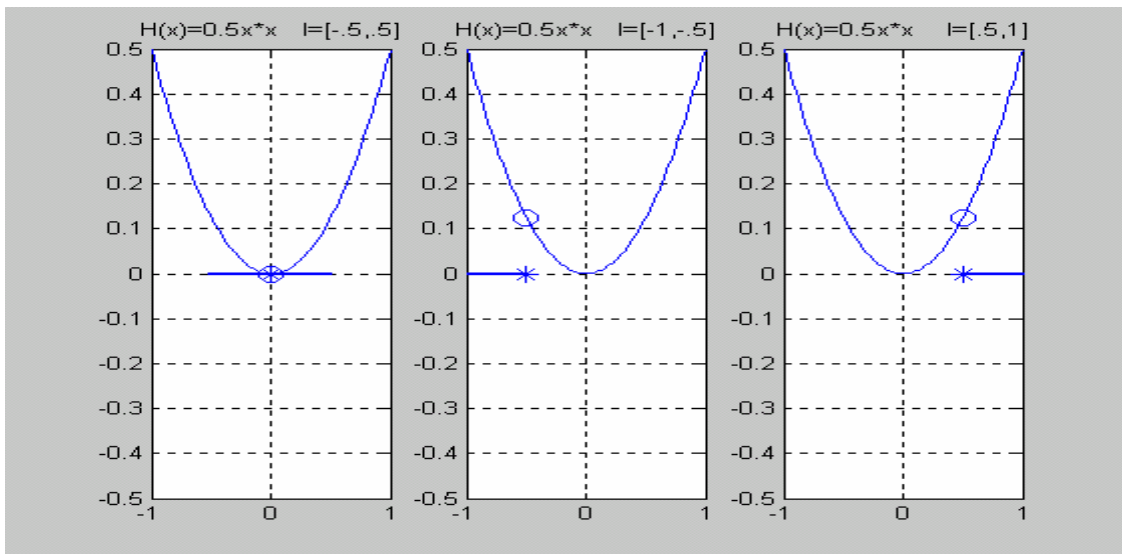
Das Minimum der Funktion $H: \mathbf{R} \rightarrow \mathbf{R}$ mit $H(x) = 0.5x^2$ wird an der Stelle x_0 mit

$$H'(x_0) = h(x_0) = 0$$

angenommen und mithin über die Ableitung h der Funktion H charakterisiert.



Diese Charakterisierung bleibt nicht mehr richtig, wenn das Minimum der Funktion H lediglich auf einer (konvexen und abgeschlossenen) Teilmenge I aus \mathbf{R} gesucht wird. Dennoch kann auch in diesem Fall der Punkt x_0 , in dem das Minimum angenommen wird, leicht bestimmt werden. Dieses zeigt das folgende Beispiel:
 Für verschiedene Intervalle $I = [a, b] \subset \mathbf{R}$ ($I = [-.5, .5]$, $[-1, -.5]$ oder $[.5, 1]$) können die zugehörigen Minimalwerte (0) von H und die zugehörigen Argumente $x_0 = *$ aus folgenden Bild entnommen werden.



Schnell überzeugt man sich davon, dass die zum jeweiligen Minimum (0) gehörenden Argumente $x_0 = *$ durch die Bedingung (Variationsungleichung)

$$H'(x_0)(v-x_0) \geq 0 \quad \text{für alle } v \in I$$

bestimmt werden.

Obwohl diese Variationsungleichung durch einfache Inspektion bestimmt wurde, kann diese auch formal hergeleitet werden:

Es sei x_0 das Minimum der Funktion H auf $I = [a, b]$. Die Funktion $f_v : [0, 1] \rightarrow \mathbf{R}$ mit

$$f_v(t) = H(x_0 + t(v-x_0)) \quad v \in I$$

ist auf $[0,1]$ stetig differenzierbar und es gilt

$$f'_v(t) = H'(x_0 + t(v-x_0))(v-x_0) \quad \text{für alle } v \in I$$

Da $f'_v(t)$ als Funktion von t sein Minimum an der Stelle $t = 0$ annimmt, muss $f'_v(0) \geq 0$ für alle $v \in I$ sein, d.h.

$$H'(x_0)(v-x_0) \geq 0 \quad \text{für alle } v \in I$$

Damit wird das Minimum nicht mehr durch die Bedingung

$$H'(x_0) = 0$$

sondern durch die Variationsungleichung

$$H'(x_0)(v-x_0) \geq 0 \quad \text{für alle } v \in I$$

charakterisiert. Dieses hat weitgehende Konsequenzen und Anwendungen. Was bereits am folgenden Beispiel erkannt werden kann.

Wir betrachten die Differentialgleichung

$$L(y) := -(p(x)y')' + q(x)y = f(x)$$

mit geeigneten Randbedingungen.

Dabei seien $p(\cdot) > 0$ und $q(\cdot) \geq 0$ stetige Funktionen von $[a,b]$ in \mathbf{R} . Die Funktion $f(\cdot)$ sei eine quadratisch integrierbare Funktion von $[a,b]$ in \mathbf{R} .

Multiplikation der Differentialgleichung mit Funktionen $v \in C_0^\infty(a, b)$ und Integration über das Intervall $[a,b]$ führt auf die Bilinearform

$$a(u, v) = \int_a^b (pu'v' + quv) dx$$

und die Linearform

$$F(v) = \int_a^b f v dx .$$

Mit geeigneten Hilberträumen V

$$C_0^\infty(a, b) \subset V \subset H^1(a, b)$$

führen diese Formen auf die Aufgabe

$$\begin{cases} \text{Gesucht ist ein } u \in V \text{ mit} \\ a(u, v) = F(v) \quad \forall v \in V. \end{cases}$$

und die dazu äquivalente Minimierungsaufgabe

$$\begin{cases} \text{Gesucht ist ein } u \in V \text{ mit} \\ E(u) \leq E(v) \quad \forall v \in V. \end{cases}$$

mit dem „Energiefunktional“ $E : V \rightarrow \mathbf{R}$

$$E(v) = (1/2)a(v, v) - F(v).$$

Die entstandene Minimierungsaufgabe kann nun eingeschränkt werden. Es kann vernünftig sein, ein Minimum nicht auf ganz V sondern „nur“ auf einer abgeschlossenen und konvexen Teilmenge M von V zu suchen.

Die folgenden Zeilen zeigen, dass solche Minimalpunkte auch zu einer Variationsungleichung führen.

Es sei also M eine abgeschlossene und konvexe Teilmenge $M \subset V$ und $u \in M$ mit

$$E(u) \leq E(v) \quad \text{für alle } v \in M \subset V$$

Für ein beliebiges $v \in M$ sei $f_v : [0,1] \rightarrow \mathbf{R}$

$$\begin{aligned} f_v(t) &= (1/2)a(u+t(v-u), u+t(v-u)) - F(u+t(v-u)) = \\ &= (1/2)a(u,u) + (1/2)ta(u,v-u) + (1/2)ta(v-u,u) + (1/2)t^2a(v-u,v-u) - F(u) - tF(v-u). \end{aligned}$$

Die Funktion $f_v : [0,1] \rightarrow \mathbf{R}$ ist wieder auf $[0,1]$ stetig differenzierbar und es gilt

$$f'_v(t) = a(u,v-u) - F(v-u) + t a(v-u,v-u) \quad \text{für alle } v \in I$$

Als notwendige Bedingung für ein Minimalpunkt u ergibt sich also die Variationsungleichung

$$a(u,v-u) - F(v-u) \geq 0 \quad \text{für alle } v \in M.$$

§ 8.5 Beispiele

Gegeben sei die Differentialgleichung

$$-u'' = f \quad \text{auf } \Omega$$

Ist M eine abgeschlossene und konvexe Teilmenge aus $H^1(0,1)$, dann führt diese Aufgabe zur Variationsungleichung:

„Gesucht ist ein $u \in M$ mit

$$\int_0^1 u'(x)(v'(x) - u'(x)) dx \geq \int_0^1 f(x)(v(x) - u(x)) dx \quad \forall v \in M$$

Es sei jetzt $\Omega = (0,1)$ und

$$M = \{ v \in H^1(0,1) / v(0) \geq 0, v(1) \geq 0 \}.$$

Interessant ist nun, dass dieser Variationsungleichung wieder eine „klassische Randwertaufgabe“ zugeordnet werden kann. Aber: Welche ist diese „Randwertaufgabe“?

Es zeigt sich, dass die Variationsungleichung zu folgender „Randwertaufgabe“ führt:

$$-u'' = f \quad \text{auf } \Omega = (0,1)$$

$$u(0) \geq 0, u(1) \geq 0, u'(0) \leq 0, u'(1) \geq 0 \\ u(0)u'(0) = u(1)u'(1) = 0.$$

Die spezielle Wahl von M führt dazu, dass $0 \in M$ und mit u auch $2u \in M$ ist. Damit liefert die Variationsungleichung die beiden Relationen

$$\int_0^1 u' v' dx \geq \int_0^1 f v dx \quad \text{für alle } v \in M$$

$$\int_0^1 u'^2 dx = \int_0^1 f u dx.$$

Wir nehmen nun an, dass eine Funktion $u : [0,1] \rightarrow \mathbf{R}$ existiert welche die Variationsungleichung erfüllt, deren erste Ableitung absolut stetig und deren zweite Ableitung quadratisch integrierbar ist.

Es sei $v \in C_0^\infty(0,1)$, dann ist $v(0) = v(1) = 0$ und mit v liegt auch $-v$ in M . Daraus folgt

$$\int_0^1 u' v' dx = \int_0^1 f v dx.$$

Partielle Integration liefert

$$\int_0^1 (-u'' - f)v dx = 0 \quad \text{für alle } v \in C_0^\infty(0,1)$$

und damit

$$-u'' = f \quad \text{f.ü. auf } \Omega = (0,1).$$

Partielle Integration liefert für alle $v \in M$

$$\int_0^1 -u'' v dx + [u'(1)v(1) - u'(0)v(0)] \geq \int_0^1 f v dx.$$

Daraus ergibt sich die Ungleichung

$$u'(1)v(1) - u'(0)v(0) \geq 0 \quad \text{für alle } v \in M.$$

Durch ein v mit den Eigenschaften $v(1) = 1$ und $v(0) = 0$ (oder $v(1) = 0$ und $v(0) = 1$) ergibt sich

$$u'(1) \geq 0 \quad (\text{oder } -u'(0) \geq 0).$$

Aus

$$\int_0^1 u'^2 dx = \int_0^1 f u dx$$

folgt

$$u'(1)u(1) - u'(0)u(0) = 0.$$

Wegen $u'(1) \geq 0$ folgt

$$0 \leq u'(1)u(1) = u'(0)u(0) \leq 0$$

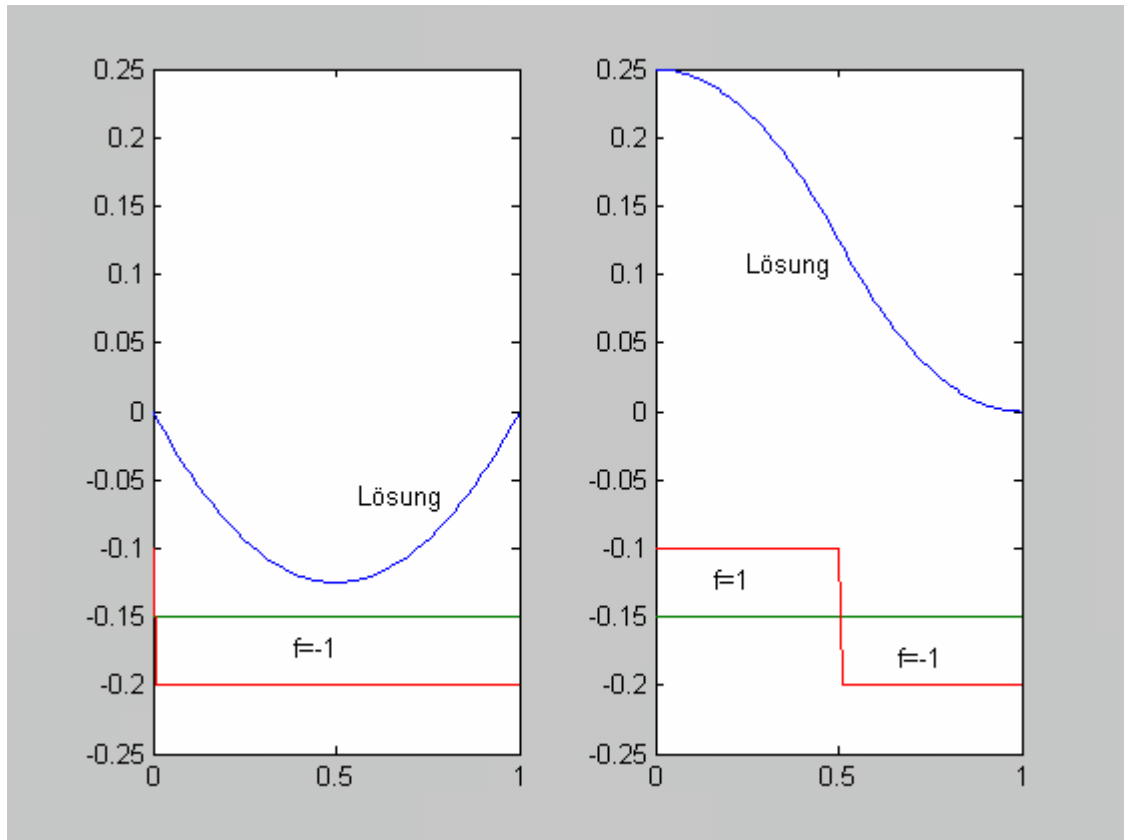
und damit die noch fehlenden Gleichungen

$$u'(1)u(1) = u'(0)u(0) = 0.$$

Die konkrete Wahl

$$f_0 = -1 \quad \text{für} \quad 0 \leq x \leq 1 \quad f_{1/2} = \begin{cases} -1 & \text{für} \quad 0 \leq x \leq 0.5 \\ 1 & \text{für} \quad 0.5 \leq x \leq 1 \end{cases}$$

und eifriges Rechnen führt zu folgenden Lösungen:



Es sei jetzt $\Omega = (0,2)$, $f = 6(x-1)$ und $M = M_a$

$$M_a = \{ v \in H^1(0,2) / v(0) = a, \quad v(2) = a, \quad v(x) \geq 0 \}$$

Die zugehörige Variationsungleichung hat wieder die Form

„Gesucht ist ein $u \in M_a$ mit

$$\int_0^1 u'(x)(v'(x) - u'(x)) dx \geq \int_0^1 f(x)(v(x) - u(x)) dx \quad \forall v \in M_a$$

und wir stellen uns erneut die Frage: Welche „klassischen Randwertaufgabe“ wird durch die zugehörige Variationsungleichung gelöst?

Wir nehmen nun wieder an, dass eine Funktion $u : [0,2] \rightarrow \mathbf{R}$ existiert welche die Variations-ungleichung erfüllt, deren erste Ableitung absolut stetig und deren zweite Ableitung quadratisch integrierbar ist. Dann genügt diese Funktion der „Randwertaufgabe“

$$\begin{aligned} -u'' &= f & \text{auf} & \quad \Omega = (0,2) \\ u(0) &= u(2) = a \end{aligned}$$

mit den zusätzlichen Bedingungen:

Ist Ω_0 die Menge der $x \in \Omega$ mit $u(x) = 0$ und Ω_1 die Menge der $x \in \Omega$ mit $u(x) > 0$, dann gilt für (die unbekannt) Punkte x_i aus $\partial\Omega_0 \cap \partial\Omega_1$

$$u(x_i) = u'(x_i) = 0.$$

Es sei $\varphi \in C_0^\infty(0,2)$ eine Funktion deren Träger M in Ω_1 liegt. M ist kompakt und deshalb $u(x) \geq \eta > 0$.

Es sei jetzt $v = u \pm \varepsilon\varphi$. Für hinreichend kleine ε ist dann $v(x) \geq 0$. Aus der Variationsungleichung folgt dann

$$\pm \varepsilon \int_{\Omega_1} u'(x)\varphi'(x)dx \geq \pm \varepsilon \int_{\Omega_1} u(x)\varphi(x)dx \quad \forall \varphi.$$

Partielle Integration liefert

$$\pm \varepsilon \int_{\Omega_1} [-u'' - f]\varphi dx \geq 0 \quad \forall \varphi.$$

D.h. es muss gelten

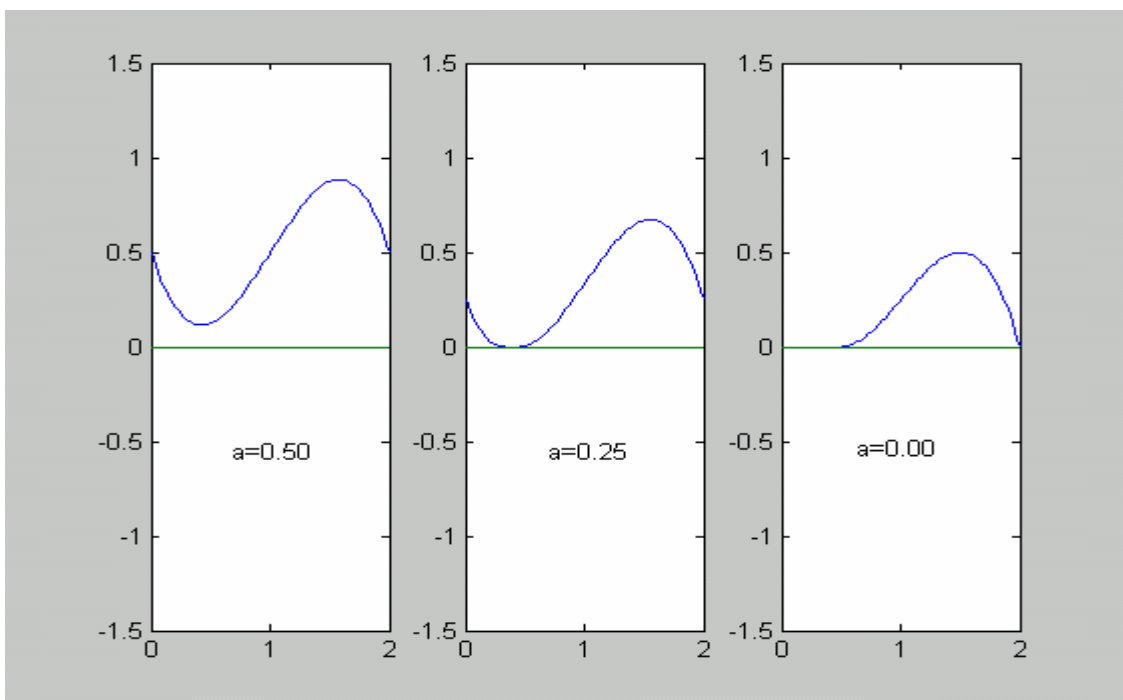
$$-u'' - f = 0 \quad \text{für} \quad x \in \Omega_1.$$

Da wir angenommen haben, dass die Funktion aus u mindestens einmal stetig differenzierbar ist, muss an den Übergängen x_i die Bedingung

$$u(x_i) = u'(x_i) = 0.$$

erfüllt sein.

Erneut führt eifriges Rechnen zu den folgenden Lösungen (in Abhängigkeit von a)



§ 8.6 Etwas Theorie

Wir benötigen einige theoretischen Hilfsmittel:

Es sei M eine abgeschlossene konvexe Teilmenge eines Hilbertraumes über \mathbf{R} , dann gibt es zu jedem $x \in H$ ein eindeutig bestimmtes $y \in M$ mit

$$\|x - y\| = \inf_{\eta \in M} \|x - \eta\|.$$

Dieser Punkt heißt Projektion von x auf M

$$y := P_M x.$$

Die Projektion von y auf M kann wie folgt charakterisiert werden:

Satz 8.4

Es sei M eine abgeschlossene Teilmenge eines Hilbertraumes über \mathbf{R} . Dann ist $y = P_M x$ genau dann, wenn

$$(y, v-y) \geq (x, v-y) \quad \text{für alle } v \in M.$$

Beweis :

Es sei $x \in H$ und $y = P_M x$. Da M konvex ist, folgt

$$(1-t)y + tv = y + t(v-y) \in M \quad \text{für alle } v \in M \text{ und alle } t \text{ mit } 0 \leq t \leq 1.$$

Die Funktion

$$f(t) = \|x - y - t(v-y)\|^2.$$

ist stetig differenzierbar auf $[0,1]$ und die Ableitung an der Stelle $t = 0$ ist

$$-2(x-y, v-y).$$

Wegen der Minimalität von y muss die Ableitung an der Stelle $t = 0$ aber nichtnegativ sein, d.h.

$$(y, v-y) \geq (x, v-y) \quad \text{für alle } v \in M.$$

Hat umgekehrt $y \in M$ die Eigenschaft $(y, v-y) \geq (x, v-y)$ für alle $v \in M$. Dann gilt

$$0 \leq (y-x, v-y) = (y-x, (v-x)+(x-y)) = -\|x-y\|^2 + (y-x, v-x).$$

Damit ist aber für alle $v \in M$ auch

$$\|x-y\|^2 \leq \|y-x\| \|v-y\| \quad \text{oder} \quad \|x-y\| \leq \|v-x\|.$$

Korollar

Sei M eine konvexe abgeschlossene Teilmenge eines Hilbertraumes H , dann ist der Operator P_M nichtexpansiv, d.h.

$$\|P_M x_1 - P_M x_2\| \leq \|x_1 - x_2\| \quad \text{für alle } x_1, x_2 \in H.$$

Beweis:

Es sei $y_i = P_M x_i$, $i=1,2$. Dann gilt

$$\begin{aligned}(y_1, v-y_1) &\geq (x_1, v-y_1) && \text{für alle } v \in M \\ (y_2, v-y_2) &\geq (x_2, v-y_2) && \text{für alle } v \in M.\end{aligned}$$

Wähle $v = y_2$ in der ersten und $v = y_1$ in der zweiten Ungleichung. Addition liefert dann

$$- \|y_1 - y_2\|^2 \geq (x_2 - x_1, y_1 - y_2)$$

oder

$$\|y_1 - y_2\|^2 \leq \|x_1 - x_2\| \|y_1 - y_2\|.$$

§ 8.7 Numerische Methoden

Es sei A eine symmetrische und positiv definite $n \times n$ Matrix und f ein Vektor aus dem \mathbf{R}^n . Bekanntlich sind dann die beiden folgenden Aufgaben äquivalent:

$$\left\{ \begin{array}{l} \text{Gesucht ist ein } u \in \mathbf{R}^n \text{ mit} \\ Au = f \end{array} \right.$$

$$\left\{ \begin{array}{l} \text{Gesucht ist } u \in \mathbf{R}^n \text{ mit} \\ J(u) \leq J(x) \text{ für alle } x \in \mathbf{R}^n \end{array} \right.$$

Dabei sei $J(u) = (1/2)(Au, u) - (f, u)$ und (\cdot, \cdot) das gewöhnliche Skalarprodukt im \mathbf{R}^n .

Es sei jetzt M eine konvexe und abgeschlossene Teilmenge aus dem \mathbf{R}^n und die folgende Minimierungsaufgabe

$$\left\{ \begin{array}{l} \text{Gesucht ist } u \in M \text{ mit} \\ J(u) \leq J(x) \text{ für alle } x \in M \end{array} \right.$$

gegeben. Dieser gesuchte Minimalpunkt existiert und kann wie folgt charakterisiert werden:

Satz 8.5

Ist $u \in M$ eine Lösung von

$$J(u) \leq J(v) \quad \text{für alle } v \in M$$

Dann erfüllt u die „Variationsungleichung“

$$(Au, v-u) \geq (f, v-u) \quad \text{für alle } v \in M$$

und umgekehrt.

Beweis

Auf den Beweis der ersten Richtung verzichten wir. u sei also eine Lösung der Aufgabe

$$(Au, v-u) \geq (f, v-u) \quad \text{für alle } v \in M.$$

Aus

$$J(v) = J(u+(v-u)) = (1/2)(Au, u) - (f, u) + (Au, v-u) - (f, v-u) + (1/2)(A(v-u), v-u)$$

folgt dann die Behauptung.

Bemerkenswerter Weise kann die Aufgabe

$$\text{„Gesucht ist ein } u \in M \text{ mit } (Au-f, v-u) \geq 0 \quad \text{für alle } v \in M \text{“}$$

in eine Fixpunktaufgabe überführt werden.

Die Fixpunktaufgabe ergibt sich aus der folgenden Folge von äquivalenten Relationen:

$$\begin{aligned} & (Au-f, v-u) \geq 0 \quad \text{für alle } v \in M \\ & \quad \quad \quad \updownarrow \\ & (-p(Au-f), v-u) \geq 0 \quad \text{für alle } v \in M \text{ und } p > 0 \\ & \quad \quad \quad \updownarrow \\ & (u-p(Au-f)-u, v-u) \geq 0 \quad \text{für alle } v \in M \\ & \quad \quad \quad \updownarrow \\ & \text{„Gesucht ist ein } u \in K \text{ mit } \| (u-p(Au-f)-u) \| \leq \| (u-p(Au-f)-v) \| \quad \text{für alle } v \in M \text{“} \\ & \quad \quad \quad \updownarrow \\ & F(u) := P_M(u-p(Au-f)) = u \end{aligned}$$

P_M bildet nach Definition die konvexe Menge M in sich ab. Außerdem ist die Abbildung

$$F : M \rightarrow M$$

für hinreichend kleine p eine kontrahierende Abbildung:

Es sei

$$\begin{aligned} F(v_1) &= P_M(v_1 - p(Av_1 - f)) \\ F(v_2) &= P_M(v_2 - p(Av_2 - f)). \end{aligned}$$

Da P_M nichtexpansiv ist, folgt

$$\| F(v_1) - F(v_2) \|^2 \leq \| v_1 - v_2 - p(A(v_1 - v_2)) \|^2.$$

Ist $(Ax, x) \geq a(x, x)$ ($a > 0$) dann ist F für alle $p < 2a/\|A\|^2$ kontrahierend und der Fixpunktsatz für kontrahierende Abbildungen liefert auch die Existenz und Eindeutigkeit von y . Außerdem liefert der Fixpunktsatz auch eine konstruktive Möglichkeit zur Ermittlung des Fixpunktes. Die Erfahrung zeigt allerdings, dass hier ein außerordentlich träges Verfahren vorliegt.

Es gehört zu den neueren Entwicklungen (2003) im Bereich der Variationsungleichungen diese Aufgaben mit einem verallgemeinerten Newton Verfahren zu lösen.

Man betrachtet dazu die Funktion

$$H : M \rightarrow M$$

mit

$$H(v) = P_M(v - (Av - f)) - v$$

und sucht eine Nullstelle $H(v) = 0$ dieses Problems.

Formal hat das Newton Verfahren dann die Form

$$((DP_M)_w^*(E-A) - E)z = -H(v_n) \quad w = v_n - (Av_n - f)$$

$$v_{n+1} = v_n + z.$$

Leider ist die Projektion P_M i.a. keine differenzierbare Abbildung. D.h. der Ausdruck DP_M ist an der Stelle w i.a. nicht definiert. Man wird die sogenannte Clark'sche Ableitung wählen und für $(DP_M)_w$ eine Auswahl aus dieser mengenwertigen „Ableitung“ treffen. In der Praxis kann dieses sehr einfach sein:

Wir betrachten die einfache Aufgabe

$$-u'' = -1, \quad u(0) = u(2) = 0, \quad u \geq -0.25$$

Es sei h eine vorgegebene Schrittweite $h = 1/N$, $N \in \mathbb{N}$, und $x_i = ih$ die zugehörigen Gitterpunkte aus $[0,2]$ mit $i = 0,1,2, \dots, N$.

Eine kanonische Diskretisierung der Differentialgleichung führt auf das Gleichungssystem $Av = f$ mit

$$A = \frac{1}{h^2} \begin{pmatrix} 2 & -1 & 0 & \dots & 0 & 0 \\ -1 & 2 & -1 & \dots & 0 & 0 \\ \cdot & \cdot & \cdot & \dots & \cdot & \cdot \\ \cdot & \cdot & \cdot & \dots & \cdot & \cdot \\ \cdot & \cdot & \cdot & \dots & \cdot & \cdot \\ 0 & 0 & 0 & \dots & 2 & -1 \\ 0 & 0 & 0 & \dots & -1 & 2 \end{pmatrix} \quad \text{und} \quad f = \begin{pmatrix} -1 \\ -1 \\ \cdot \\ \cdot \\ \cdot \\ -1 \\ -1 \end{pmatrix}$$

Zu lösen ist dann das Gleichungssystem $Av = f$ unter der Nebenbedingung $v_i \geq -0.25$ d.h.

$$P_M(v - (Av - f)) - v = 0.$$

Die Komponenten des Projektionsoperators $(P_M v)_i$ haben die Gestalt

$$(P_M v)_i = \begin{cases} v_i & \text{falls } v_i > -0.25 \\ -0.25 & \text{falls } v_i \leq -0.25 \end{cases}$$

Damit kann $(DP_M)_w$ als Diagonalmatrix mit den Einträgen $a_{ii} = (\text{sign}(w_i + 0.25) + 1)/2$ gewählt werden.

Das folgende MATLAB-Programm löst die Aufgabe mit dem Newton Verfahren und liefert ein Ergebnis welches von der exakten Lösung nicht zu unterscheiden ist:

```
function z = newtondata(u,psi,f,N)
%Liefert die Daten Dfu und fu für die Newtoniteration
%Dfu(u1-u)=-fu bei vorgegebenem u
%Die Aufgabe ist -u''= f, u(0) =u(2) = 0 u>=psi
%N gibt die Anzahl der Knoten in [0,2] an
```

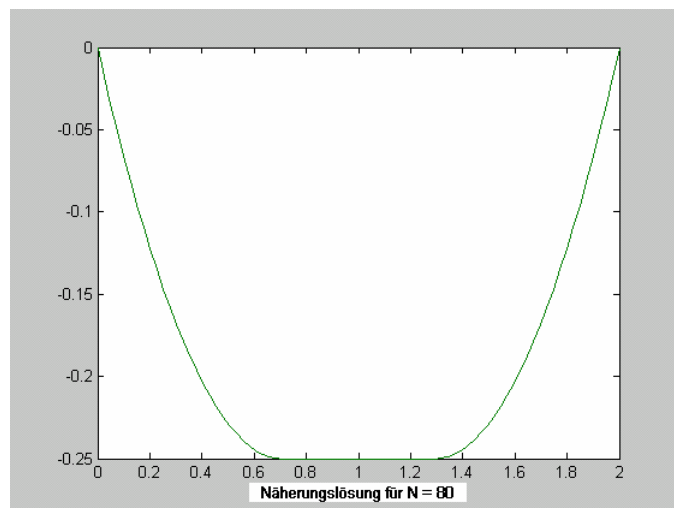


```

h = 2/(N-1);
A = 2*eye(N-2);
if N>3
    B =-diag(linspace(1,1,N-3),1)-diag(linspace(1,1,N-3),-1);
    A = (1/h^2)*(A+B);
end
H1 = u'-A*u'+f';
a = (sign(H1-psi')+1)/2;
DPK = diag(a');
Dfu = DPK*(eye(N-2)-A)-eye(N-2);
H2 = u'-A*u'+f';
PKH2 = (H2'>=psi).*(H2')+(H2'<psi).*psi;
fu = PKH2'-u';
z = [Dfu,fu];

N=80
psi= linspace(-.25,-.25,N-2);
u0 = linspace(-1,-1,N-2);
f = linspace(-1,-1,N-2);
v = u0;
for i=1:160
    z = newtondata(v,psi,f,N);
    DFu = z(:,1:N-2)
    fu = z(:,N-1);
    y=-DFu\fu;
    vv = [0 v 0];
    v= y'+v;
end
x = (0:2/(N-1):2)
w = [0 v 0]
plot(x,w,x,vv)

```



Übungen
zur Theorie und Numerik gewöhnlicher Differentialgleichungen
SoSe 2007

Taubert, Heitmann

Abgabe 8.6.2007 vor den Übungen

Aufgabe 1

Gegeben sei die Anfangswertaufgabe

$$y' = 0.5y^2, \quad y(0) = 1.$$

Es sei y_N^h ein, mit der Schrittweite h und den expliziten Euler-Verfahren ermittelter, Näherungswert für die Lösung der Anfangswertaufgabe $y(\cdot)$ an der Stelle $x_N = Nh$.

Tabelliere die Differenz

$$E(1,h) = y(1) - y_N^h, \quad Nh = 1$$

für $h = 2^{-p}$ ($p = 1, 2, 3, \dots, 12$).

Tabellieren Sie anschließend die Differenzen $h^{-1}E(1,h)$ und $h^{-2}E(1,h)$. Stehen die Ergebnisse im Einklang mit der in der Vorlesung angegebenen Fehlerabschätzung?

Aufgabe 2

Es sollen Näherungslösungen für

$$y' = y, \quad y(0) = 1,$$

mit Hilfe des expliziten Euler-Verfahrens und den Schrittweiten $h = 2^{-p}$ ($p = 1, 2, 3, \dots$) bestimmt werden.

Zeichnen Sie für das Intervall $[0,4]$, in einem Bild, sowohl die Lösung der Anfangswertaufgabe als auch die Näherungslösungen für $p = 1, 2, 3$ und 4 .

Wiederholen Sie obiges mit der Trapezregel.

Bestimme mit der Trapezregel und den Schrittweiten $h = 2^{-p}$ ($p = 1, 2, 3, \dots$) Näherungen y^h für $y(1)$ und die Werte

$$P(h) = \log_2 \left(\left| \frac{y^h - y^{h/2}}{y^{h/2} - y^{h/4}} \right| \right).$$

Liefere Sie eine Erklärung für das Verhalten von $P(h)$ bei kleiner werdenden h .

Aufgabe 3

Die Differentialgleichung

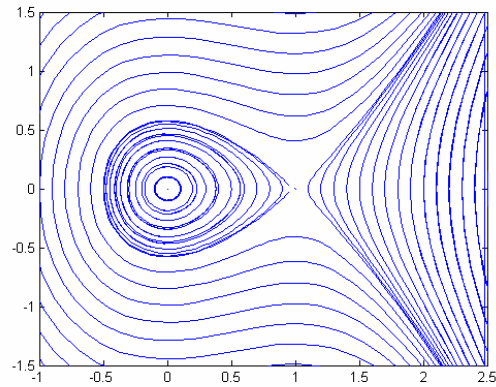
$$y'' = -y + y^2$$

ist äquivalent zum System von Differentialgleichungen erster Ordnung

$$y' = z$$

$$z' = -y + y^2$$

Mit einem numerischen Verfahren wurde ein zugehöriges Phasenporträt (Lösungen $(y(t), y'(t))$) für den Bereich $-1.5 < y < 2.5$ und $-1.5 < y' < 1.5$ mit folgender Gestalt erstellt:



Versuchen Sie dieses Phasenporträt mit der Trapezregel zu reproduzieren.

Die numerische Integration - mit einer moderaten Schrittweite - von 50 Lösungen über das Intervall $[0, 5]$ mit äquidistant verteilten Startwerten $(y(0), y'(0))$ aus $([-1.8, 2.5], 0)$ sollte das gewünschte Ergebnis liefern. Bei jedem Schritt der Trapezregel muss eine nichtlineare Gleichung gelöst werden. Es bietet sich das Newton-Verfahren an und eine Iteration sollte bereits ausreichend gute Näherungen liefern. Allerdings kann es auch Ärger geben (!) Wie sieht das entsprechende, mit dem Euler-Verfahren erzeugte, Phasenporträt aus?

Übungen
zur Theorie und Numerik gewöhnlicher Differentialgleichungen
SoSe 2007

Taubert, Heitmann

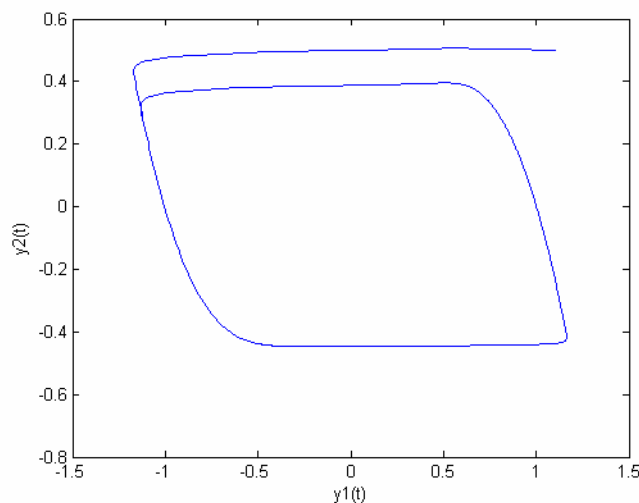
Abgabe 15.6.2007 vor den Übungen

Aufgabe 4

Gegeben sei die Anfangswertaufgabe

$$y' = f(y) = \begin{cases} -100(y_1^3 - y_1 + y_2) \\ y_1 - 0.5 \end{cases}, \quad y(0) = \begin{pmatrix} 1.1 \\ 0.5 \end{pmatrix}.$$

Integrieren Sie die Aufgabe mit dem impliziten Euler-Verfahren im Intervall $[0,5]$. Die Auflösung der nichtlinearen Gleichungen soll mit dem Newton Verfahren erfolgen. Der Phasenplot $(y_1(t), y_2(t))$ der numerischen Lösung sollte die folgende Gestalt



annehmen.

Rechnen Sie mit verschiedenen Schrittweiten ($h \in [0.01, 0.1]$) und mit unterschiedlicher Anzahl von Newton Schritten (2-10).

Wiederholen Sie die Berechnungen mit der Trapezregel oder der one-leg Version der Trapezregel.

Versuchen Sie die gewünschte Näherungslösung mit dem expliziten Euler-Verfahren zu berechnen.

Aufgabe 5

Gegeben sei das System von N linearen Anfangswertaufgaben

$$y' = Ay, \quad y(0) = y_0 = (y^1(0), y^2(0), y^3(0), \dots, y^N(0)).$$

Die Matrix A habe die Tridiagonalgestalt

$$A = (N+1)^2 \begin{pmatrix} -2 & 1 & & & \\ 1 & -2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & & 1 & -2 & 1 \\ & & & & 1 & -2 \end{pmatrix}$$

und es sei $y^i(0) = \sin(2\pi \frac{i}{N+1})$ für $i=1, \dots, N$.

Bestimmen Sie den betragsgrößten Eigenwert von A für $N = 1$ bis 50. Plotten Sie das Ergebnis.

Integrieren Sie mit dem expliziten Euler-Verfahren die Anfangswertaufgabe mit $N=10$ für $t \in [0,10]$ und den verschiedenen Schrittweiten $h = 0.01, 0.009, 0.008, \dots, 0.001$.

Beachten Sie besonders den Bereich um 0.004.

Plotten Sie die berechneten Näherungswerte für $y^i(10)$ in Abhängigkeit von $i = 1, 2, 3, \dots, 10$.

Aufgabe 6

Bestimmen Sie die Ordnung vom verbesserten Euler-Verfahren

Übungen
zur Theorie und Numerik gewöhnlicher Differentialgleichungen
SoSe 2007

Taubert, Heitmann

Abgabe 22.6.2007 vor den Übungen

Aufgabe 7

Die Ordnung eines impliziten RKV wird so definiert wie bei dem expliziten RKV.
Bestimmen Sie die Ordnung des Verfahrens

0	0	0
1	1/2	1/2
	1/2	1/2

Aufgabe 8

Gegeben sei die Anfangswertaufgabe

$$y' = f(y) = 1 + |y|, \quad y(0) = -0.5.$$

Geben Sie die Lösung für die Anfangswertaufgabe im Intervall $[0,1]$ an und plotten Sie die Lösung und deren Ableitung.

Integrieren Sie die Aufgabe auf dem Intervall $[0,1]$ mit dem expliziten Euler-Verfahren und dem Euler-Cauchy-Verfahren mit den Schrittweiten $h = 2^{-M}$, $M = 1,2,3, \dots, 14$.

Geben Sie den Fehler an der Stelle $x=1$ an. Zu erwarten ist beim Euler-Verfahren eine Halbierung des Fehlers bei Halbierung der Schrittweite und beim Euler-Cauchy-Verfahren eine Viertelung des Fehlers bei Halbierung der Schrittweite, oder?

Versuchen Sie, entsprechend der Aufgabe 2, die Ordnung 1 bzw. 2 der Verfahren mit dieser Anfangswertaufgabe zu verifizieren

Aufgabe 9

Gegeben sei die Anfangswertaufgabe

$$y' = y \cos(x), \quad y(-8) = \exp(\sin(-8)).$$

Mit dem klassischen RKV soll die Lösung des Anfangswertproblems im Intervall $[-8,10]$ berechnet werden. Wähle als Schrittweite $h = 0.1$.

Bestimme die Näherungslösung mit dem Euler-Verfahren bei gleicher Anzahl von Funktionsauswertungen. Wiederhole dieses mit dem Euler-Cauchy-Verfahren.

Plote die Fehlerkurven oder die Differenz zwischen der exakten und der Näherungslösung in den drei Fällen und vergleiche diese.

Übungen
zur Theorie und Numerik gewöhnlicher Differentialgleichungen
SoSe 2007

Taubert, Heitmann

Abgabe 29.6.2007 vor den Übungen

Aufgabe 10

Gesucht sind alle Runge-Kutta-Verfahren der Ordnung 2 mit dem Butcher-Array

0			
c ₂	c ₂		
c ₃	0	c ₃	
	0	0	1

Zeigen Sie: Das Runge-Kutta Verfahren mit dem Butcher-Array

0			
1/3	1/3		
2/3	0	2/3	
	1/4	0	3/4

hat die Ordnung 3.

Aufgabe 11

Integrieren Sie die Anfangswertaufgabe

$$y' = 1 + |y|, \quad y(0) = -0.5$$

im Intervall von 0 bis 1.

Zur Integration soll ein Verfahren mit variabler Schrittweite, auf der Basis des RKV der Ordnung 2

$$y_{n+1} = y_n + h * f(x_n + h/2, y_n + (h/2) * (f(x_n, y_n)))$$

und des Runge-Kutta-Verfahrens der Ordnung 3 aus Aufgabe 10, benutzt werden.

Die vorgegebene Toleranz sei $\epsilon = 0.0001$ und die Anfangsschrittweite $h = 0.1$.

Beobachten Sie das Verhalten der Schrittweiten in Laufe der Integration.

Aus theoretischen Gründen wurde in der Vorlesung vorgeschlagen die neuen Schrittweiten in Abhängigkeit von $s > 1$ oder $s < 1$ zu wählen. Gleichzeitig wurde darauf hingewiesen, dass in dieser Bedingung die 1 häufig durch eine andere Zahl zu ersetzen sei. Die neue Schrittweite wird dann z.B. in Abhängigkeit von $s > 0.98$ oder $s < 0.98$ gewählt. Probieren Sie es!

Aufgabe 12

Bestimmen Sie das explizite lineare 2-Schrittverfahren

$$y_{n+2} + a_1 y_{n+1} + a_2 y_n = h(b_2 f_{n+2} + b_1 f_{n+1} + b_0 f_n)$$

mit maximaler Ordnung.

Zeigen Sie: Die Verfahren

$$y_{n+2} - y_n = 2h(f_{n+1}) \quad \text{und} \quad y_{n+2} - y_{n+1} = h\left(\frac{3}{2}f_{n+1} - \frac{1}{2}f_n\right)$$

haben die Ordnung 2.

Integrieren Sie mit den drei Verfahren die Anfangswertaufgabe $y' = -y$, $y(0) = 1$ mit der Schrittweite $h = 0.1$ und den exakten Startwerten $y_0 = 1$ und $y_1 = \exp(-h)$ auf den Intervallen $[0,1]$, $[0,5]$ und $[0,10]$.

Geben Sie eine Erklärung für das unterschiedliche Verhalten.

Aufgabe 13

Zeichnen Sie den Bereich der absoluten Stabilität von

$$y_{n+1} = y_n + h * f(x_n + h/2, y_n + (h/2) * f(x_n, y_n)).$$

Zeichnen Sie den Bereich der absoluten Stabilität für das Verfahren

$$y_{n+2} - y_{n+1} = h\left(\frac{5}{12}f_{n+2} + \frac{8}{12}f_{n+1} - \frac{1}{12}f_n\right).$$

Zum Vergleich zeichne gleichzeitig den Bereich der absoluten Stabilität vom expliziten Euler-Verfahren in die einzelnen Zeichnungen.

Übungen zur Theorie und Numerik gewöhnlicher Differentialgleichungen SoSe 2007

Taubert, Heitmann

Abgabe 6.7.2007 vor den Übungen

Aufgabe 14

Gegeben sei die Anfangswertaufgabe

$$y' = -3x^2 y, \quad x \in [0,8], \quad y(0) = 5.$$

Berechnen Sie Näherungslösungen mit den Verfahren von Gear der Ordnung 2,4 und 6. Als Startfeld wähle man stets die exakte Lösung und als Schrittweite $h = 0.1, 0.01$ und 0.001 . Beobachten Sie den Fehler der Näherungslösungen bei Verkleinerung der Schrittweite. Entspricht dieser den Erwartungen?

Aufgabe 15

Die Bewegung einer verschiebbaren Hülse auf einer Schräggestellten rotierenden Stange führt auf eine unstetige Differentialgleichung der Form (Siehe Bild, Sommerfeld-Mechanik)

$$r'' = (-g \sin \alpha \cos \alpha + r \omega^2 \cos^2 \alpha) - \tan \varphi (g \cos^2 \alpha + r \omega^2(t) \sin \alpha \cos \alpha) \text{sign}(r')$$

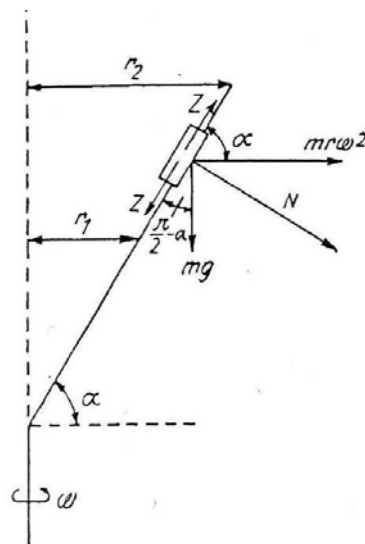


Fig. 22. Verschiebbare Hülse auf einer schräggestellten rotierenden Stange. Gleichgewicht gegen Reibung

Durch Einführung der Variablen $y_1 = r$ und $y_2 = r'$ kann diese Differentialgleichung in ein System erster Ordnung umgeschrieben werden

$$\dot{y}_1 = y_2$$

$$\dot{y}_2 = (-g \sin \alpha \cos \alpha + \omega^2 y_1 \cos^2 \alpha) - \tan \varphi (g \cos^2 \alpha + y_1 \omega^2(t) \sin \alpha \cos \alpha) \text{sign}(y_2).$$

Integrieren Sie diese Aufgabe im Intervall $[0,7]$ mit dem expliziten Euler-Verfahren und der Schrittweite $h = 0.001$. Was beobachten Sie? Versuchen Sie eine Integration mit der Mittelpunktsformel.

Es sei $g = 9.81$, $\alpha = \pi/4$, $\tan \varphi = 0.267$ und $\omega^2 = g/(1 - 0.5 \sin(2\pi t))$. Die Anfangsbedingungen seien $r(0) = 1$ und $r'(0) = 0$.

Aufgabe 16

Gegeben seien die Anfangswertaufgaben (Hamilton System)

$$\begin{aligned}x' &= y & x(0) &= \cos(\varphi) \\y' &= -\sin(x) & y(0) &= \sin(\varphi)\end{aligned}$$

mit $\varphi \in [0, 2\pi]$.

Die Aufgaben sollen für die Anfangsbedingungen $\varphi = (0 : 2\pi/N : 2\pi)$ mit $N = 50$ numerisch über das Intervall $[0, 10]$ integriert werden.

Integrieren Sie die Aufgaben mit dem expliziten Euler-Verfahren, dem impliziten Euler-Verfahren, der Trapezregel und die one-leg Version der Trapezregel mit der Schrittweite $h = 0.1$.

Durch jedes Verfahren werden die $N+1$ (A)nfangsbedingungen aus dem Einheitskreis in eine neue Punktmenge zum Zeitpunkt $t = 10$ überführt. Zeichnen Sie für jedes (V)erfahren diese Punkte $(x_A^V(10), y_A^V(10))_A$ auf und vergleichen Sie das Ergebnis mit dem Einheitskreis.

Übungen
zur Theorie und Numerik gewöhnlicher Differentialgleichungen
SoSe 2007

Taubert, Heitmann

Behandlung: 13.7.2007

Aufgabe 17

Ein einfaches Modell für eine isotherme chemische Reaktion führt auf die Randwertaufgabe

$$\begin{aligned} -y'' + \lambda y &= \lambda, & \lambda > 0, \\ y(0) &= y(1) = 0. \end{aligned}$$

Für die eindeutigen Lösungen der Randwertaufgaben gilt

1. $0 \leq y_\lambda \leq 1$
2. $y_\lambda(t) = y_\lambda(1-t)$
3. $0 \leq y_\lambda(t) \leq y_\mu(t) \leq 1, \quad \lambda \leq \mu.$

Zeige numerisch:

Die kanonische Diskretisierung

$$-\frac{(y_{i+1} - 2y_i + y_{i-1}))}{h^2} + \lambda y_i = \lambda$$

führt auf Näherungslösungen mit den genannten Eigenschaften der kontinuierlichen Lösungen.

Wähle die Schrittweite $h = 0.1$ und $\lambda = 10, 20, 30, 300, 10000$. Vergleiche die Näherungslösungen mit der exakten Lösung.

Aufgabe 18

Gegeben sei die Randwertaufgabe

$$\begin{aligned} -u'' &= f \quad \text{auf } (0,1) \\ u(0) &= u'(1) = 0 \end{aligned}$$

mit f gleich

$$\begin{aligned} f &= (x \leq 0.5) .* (2*x) + (x > 0.5) .* (2-2*x) \\ f &= 100*(x-1/3).^2 .* (x-2/3).^2 \\ f &= -100*(x.*(x-1)).*(x-1/3).*(x-2/3))+1 \\ f &= \sin(2*pi*x) \end{aligned}$$

Die Aufgabe besitzt eine Lösung im Raum

$$V = \{ v / v \in H^1[0,1] \text{ mit } v(0) = 0 \}.$$

Zur Approximation der Lösung wählen wir die Räume

$$V_h = \text{span} \{ (1/i)x^i \}_{i=1}^N$$

Obwohl die zugehörigen Gleichungssysteme eine schlechte Kondition (Hilbert-Matrizen) haben, liefern diese für die angegebenen f dennoch gute Approximationen.

Untersuchen Sie:

- Die Besetzung der Matrizen
- Die Kondition der Matrizen
- Die Verteilung der gesuchten Koeffizienten
- Ab welchem N wird die zweite Randbedingung gut erfüllt?

und geben Sie eine Näherungslösung an.

Aufgabe 19

Gegeben sei die Randwertaufgabe

$$\begin{aligned} -u'' &= f \quad \text{auf } (0,1) \\ u(0) &= u(1) = 0 \end{aligned}$$

mit f gleich

$$\begin{aligned} f &= (x \leq 0.5) \cdot (2 \cdot x) + (x > 0.5) \cdot (2 - 2 \cdot x) \\ f &= 100 \cdot (x - 1/3)^2 \cdot (x - 2/3)^2 \\ f &= -100 \cdot (x \cdot (x - 1) \cdot (x - 1/3) \cdot (x - 2/3)) + 1 \\ f &= \sin(2 \cdot \pi \cdot x) \end{aligned}$$

Die Aufgabe besitzt eine Lösung im Raum

$$V = \{ v \mid v \in H^1[0,1] \text{ mit } v(0) = v(1) = 0 \}.$$

Zur Approximation der Lösung wählen wir die Räume

$$V_h = \text{span} \{ \sin(i \pi x) \}_{i=1}^N$$

Untersuchen Sie:

- Die Besetzung der Matrizen
- Die Kondition der Matrizen
- Der Abfall der gesuchten Koeffizienten

und geben Sie eine Näherungslösung an.

Behandeln Sie 8 der 10 Aufgaben

- 1) Es sei $y(\cdot)$ eine dreimal stetig differenzierbare Lösung von $y' = f(x,y)$ auf dem Intervall $I = [x_0-a, x_0+a]$, $a > 0$.
 Zeige: Für den lokalen Fehler der Trapezregel gilt

$$|y(x_{n+1}) - y(x_n) - (h/2)(f(x_{n+1}, y(x_{n+1})) + f(x_n, y(x_n)))| \leq (5/12)h^3 \max_{\xi \in I} |y'''(\xi)|$$

Tipp: Beachten Sie $f(x, y(x)) = y'(x)$.

- 2) Gegeben sei die Anfangswertaufgabe

$$y' = \begin{pmatrix} -6 & -4 & 0 \\ -4 & -6 & 0 \\ 100 & 2 & -5 \end{pmatrix} y, \quad y(0) = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

Diese soll mit dem PECE-Verfahren der Ordnung 3 über lange Zeiten numerisch integriert werden. Welche Schrittweite h darf nicht überschritten werden? Begründen Sie die Wahl der Schrittweite.

Es dürfen die Abbildungen von Seite 56 verwendet werden!

- 3) Erläutern Sie Anhand einer Skizze die Konstruktionsvorschrift für die Adams-Moulton Verfahren.
 Warum sind diese Verfahren immer stabil und konsistent?

- 4) Gegeben sei die Anfangswertaufgabe

$$y' = -x^2 y, \quad y(0) = y_0, \quad x \geq 0.$$

Warum kann es vernünftig sein, diese Aufgabe mit einem BDF-Verfahren zu lösen?
 Zeigen Sie - mit einer Skizze - dass für die gegebene Anfangswertaufgabe das Verfahren

$$y_{n+3} - (18/11)y_{n+2} + (9/11)y_{n+1} - (2/11)y_n = h(6/11)f_{n+3}$$

für beliebige Startfelder, für alle h und jedes n eine Lösung besitzt.

- 5) Beweisen Sie, dass das RKV

$$\begin{array}{c|c} 1 & 1/2 \\ \hline & 1 \end{array}$$

ein B-stabiles Verfahren ist

6) Beschreiben Sie eine Schrittweitensteuerung auf der Basis des expliziten Euler-Verfahrens und der Trapezregel.

7) Eine Lipschitzstetige Anfangswertaufgabe

$$y' = f(x,y), \quad y(x_0) = y_0,$$

soll mit einem MSV numerisch integriert werden.

Welche Anforderungen sind an die Startfelder für das MSV, an die Differentialgleichung und an die Rundungsfehler beim Rechnen zu stellen, damit auf jedem kompakten Intervall die maximale Differenz zwischen der exakten und der numerischen Lösung mindestens von der Größenordnung $O(h^p)$ ist?

8) Erläutern Sie anhand einer Zeichnung, dass das verallgemeinerte explizite Euler-Verfahren bei der unstetigen Anfangswertaufgabe $y' = -\text{sign}(y) + 0.5$, $y(0) = 1$ für $x > 0$ zu einer Lösung im Sinne von Filippov führen muss.

9) Zeigen Sie: Die Integration von $y' = -10y$, $y(0) = 1$, $x > 0$, mit dem impliziten Euler-Verfahren führt zu der Fehlerabschätzung

$$|F_n| \leq -15 \exp(10h)(\exp(-10x_n) - 1)h.$$

Tipp: Beachten Sie den Satz 2.2 und deren Beweis.

10) a) Geben Sie ein explizites 2-Schrittverfahren der Ordnung 3 an.

b) Zeigen Sie, dass das Verfahren aus a) nicht stabil ist.

c) Zeigen Sie: Es gibt keine expliziten stabilen 2-Schrittverfahren mit der Ordnung größer als 2.