

Vergleich der Integrationsmethoden und der Methoden des maschinellen Lernens für gewöhnliche Differentialgleichungen

Alexandro Jedaidi

Inhaltsverzeichnis

1	Einleitung	2
2	Problemstellung	2
2.1	Gewöhnliche Differentialgleichungen und Anfangswertprobleme	2
2.1.1	lineare gewöhnliche Differentialgleichungen	3
2.2	Existenz und Eindeutigkeit	3
2.2.1	Existenz von Lösungen	4
2.2.2	Eindeutigkeit von Lösungen	5
2.3	Abhängigkeit der Lösungen von den Daten	6
3	Numerischer Lösungsansatz	9
3.1	Einschrittverfahren	9
3.1.1	Konsistenz und Konvergenz	9
3.2	Runge-Kutta-Verfahren	11
3.2.1	Konsistenz	13
3.2.2	Schrittweitensteuerung	15
3.3	Extrapolation	16
3.4	Mehrschrittverfahren	16
3.4.1	Konsistenz und Konvergenz	16
3.4.2	Stabilität	16
3.5	Verfahren für steife Differentialgleichungen	16
4	neuronale Netze	16
4.1	Methodenbeschreibung	16
4.2	Gewichtsinitialisierung	16
4.3	Fehlerdiskussion (weighting function)	16
4.4	curriculum learning	16
5	Anwendungsbeispiele	16
	Abbildungsverzeichnis	17

Zusammenfassung

Bachelorarbeit WiSe 2021/2022

1 Einleitung

2 Problemstellung

In diesem Abschnitt wird die Theorie zu den gewöhnlichen Differentialgleichungen erläutert. Diese ist Grundlage für das Verständnis der Schlussfolgerungen und Ergebnisse dieser Arbeit.

2.1 Gewöhnliche Differentialgleichungen und Anfangswertprobleme

Definition 2.1.1 Ein System gewöhnlicher Differentialgleichungen m -ter Ordnung hat die Form

$$x^{(m)} = f(t, x, x', x'', \dots, x^{(m-1)}) \quad (2.1)$$

mit der gegebenen Funktion $f : D \times \mathbb{R}^n \times \mathbb{R}^n \times \dots \times \mathbb{R}^n \rightarrow \mathbb{R}^n$, wobei $D \subseteq \mathbb{R}$ ein Zeitintervall ist. Eine dazugehörige Lösung (falls existent) $\hat{x} : D \rightarrow \mathbb{R}$ ist eine m -mal differenzierbare Funktion und erfüllt die Bedingung

$$\hat{x}^{(m)} = f(t, \hat{x}, \hat{x}', \hat{x}'', \dots, \hat{x}^{(m-1)}).$$

Definition 2.1.2 Ein Anfangswertproblem für eine Differentialgleichung (2.1) mit gegebenen Anfangswerten $x_{0,1}, \dots, x_{0,m} \in \mathbb{R}^n$ hat die Form

$$x^{(m)} = f(t, x, x', x'', \dots, x^{(m-1)}), \quad x(t_0) = x_{0,1}, \quad x'(t_0) = x_{0,2}, \dots, x^{(m-1)}(t_0) = x_{0,m}. \quad (2.2)$$

Eine Lösung des Problems $\hat{x} : D \rightarrow \mathbb{R}$ muss also zusätzlich zu (2.1) auch die Anfangswertbedingungen (vgl. (2.2)) erfüllen.

Es ist möglich jede gewöhnliche Differentialgleichung m -ter Ordnung zu einem System gewöhnlicher Differentialgleichungen erster Ordnung umzuwandeln. Dies erleichtert uns in späteren Abschnitten Aussagen über die Existenz und Eindeutigkeit der Lösung(en) \hat{x} zu treffen. Betrachte hierzu eine gewöhnliche Differentialgleichung m -ter Ordnung (vgl. (2.1)). Diese ist mit Hilfe der Funktionen $x_j : D \rightarrow \mathbb{R}$ für $j \in \{1, \dots, m\}$ äquivalent zu einem System erster Ordnung mit m Gleichungen:

$$\begin{aligned} x_1' &= x_2 \\ x_2' &= x_3 \\ &\vdots \\ x_{m-1}' &= x_m \\ x_m' &= f(t, x_1, x_2, x_3, \dots, x_m). \end{aligned} \quad (2.3)$$

Für ein AWP (2.2) gilt zusätzlich:

$$\begin{aligned} x_1(t_0) &= x_{0,1} \\ x_2(t_0) &= x_{0,2} \\ &\vdots \\ x_{m-1}(t_0) &= x_{0,m-1} \\ x_m(t_0) &= x_{0,m}. \end{aligned} \quad (2.4)$$

Definition 2.1.3 Ein System gewöhnlicher Differentialgleichungen (2.1) heißt autonom, wenn die rechte Seite f nicht explizit von der unabhängigen Variable t abhängt. Das bedeutet, es gilt

$$x^m = f(x, x', x'', \dots, x^{(m-1)})$$

auf ganz D .

In ähnlichem Stil zu der Ordnungsreduktion können wir auch eine nichtautonome Differentialgleichung zu einem autonomen System gewöhnlicher Differentialgleichungen umwandeln. Dazu führen wir eine neue Variable z ein, für die gilt

$$z' = 1, \quad z(t_0) = t_0.$$

Dann ist das Anfangswertproblem (2.2) äquivalent zu dem autonomen Anfangswertproblem $m+1$ -ter Ordnung

$$\begin{aligned} z' &= 1, & z(t_0) &= t_0 \\ x' &= f(z, x), & x(t_0) &= x_0. \end{aligned}$$

Die Lösung hat also die Form $\begin{bmatrix} t \\ x(t) \end{bmatrix}$.

2.1.1 lineare gewöhnliche Differentialgleichungen

Wir definieren in diesem Abschnitt noch eine Sonderform einer gewöhnlichen Differentialgleichung.

Definition 2.1.4 Ein System gewöhnlicher Differentialgleichungen m -ter Ordnung mit der Form

$$x^m(t) = \sum_{i=0}^{m-1} a_i(t)x^i(t) + g(t). \quad (2.5)$$

heißt lineares System gewöhnlicher Differentialgleichungen. Dabei sind $a_k : D \rightarrow \mathbb{R}^{n \times n}$ und $g : D \rightarrow \mathbb{R}^n$ stetige Funktionen.

Falls $g \equiv 0$, so nennt man (2.5) homogen, sonst inhomogen.

Die Lösung einer linearen Differentialgleichung erster Ordnung erfüllt also die Gleichung

$$x'(t) = A(t)x(t) + g(t),$$

worin $A : D \rightarrow \mathbb{R}^{n \times n}$ und $g : D \rightarrow \mathbb{R}^n$ stetig sind.

Definition 2.1.5 Eine lineare Differentialgleichung (2.5) mit $A \in \mathbb{R}^{n \times n}$ und $b(t) \in \mathbb{R}^n$ heißt steif, wenn

$$1. \quad \operatorname{Re}(\lambda_j(A)) < 0, \quad \text{für alle Eigenwerte } \lambda_j(A) \text{ von } A, \quad (2.6)$$

$$2. \quad \min_{1 \leq j \leq n} \operatorname{Re}(\lambda_j(A)) \ll \max_{1 \leq j \leq n} \operatorname{Re}(\lambda_j(A)). \quad (2.7)$$

Wir werden später ein Verfahren betrachten, welches sich speziell dafür eignet, eine steife gewöhnliche Differentialgleichung zu lösen.

2.2 Existenz und Eindeutigkeit

In diesem Abschnitt betrachten wir ein Anfangswertproblem erster Ordnung

$$\begin{aligned} x' &= f(t, x) \\ x(t_0) &= x_0 \end{aligned} \quad (2.8)$$

und zeigen, unter welchen Bedingungen der rechten Seite $f(t, x(t))$ eine Lösung existiert und ggf. eindeutig ist.

2.2.1 Existenz von Lösungen

Hier beweisen wir einen Satz, welcher zeigt, dass die Stetigkeit der rechten Seite f für die Existenz einer Lösung ausreicht. Dazu benötigen wir noch einen Satz aus der Funktionalanalysis.

Satz 2.2.1 (Fixpunktsatz von Schauder, 2. Version) *Sei M eine nichtleere, abgeschlossene und konvexe Teilmenge eines Banachraums X und $A : M \rightarrow M$ stetig. Dann besitzt A wenigstens einen Fixpunkt x , sofern die Bildmenge $A(M)$ relativ kompakt ist.*

Beweis. [1, S. 13, 14]

Satz 2.2.2 (Existenzsatz von Peano, quantitative und qualitative Version) *Quantitative Version: Seien*

$$\mathcal{G} = \{(t, x) \in \mathbb{R} \times \mathbb{R}^n : |t - t_0| \leq \alpha, \quad \|x - x_0\|_2 \leq \beta, \quad \alpha, \beta > 0\}$$

und $f : \mathcal{G} \rightarrow \mathbb{R}^n$ stetig. Dann besitzt das Anfangswertproblem (2.8) mindestens eine Lösung \hat{x} auf dem Intervall $D = \{t_0 - a, t_0 + a\}$, wobei

$$a = \min\left\{\alpha, \frac{\beta}{M}\right\}, \quad M = \max_{(t,y) \in \mathcal{G}} \|f(t, x)\|_2.$$

Qualitative Version: Seien $\mathcal{G} \subset \mathbb{R} \times \mathbb{R}^n$ offen und $f : \mathcal{G} \rightarrow \mathbb{R}^n$ stetig. Dann besitzt das Anfangswertproblem (2.8) für jedes Paar $(t_0, x_0) \in \mathcal{G}$ mindestens eine lokale Lösung, d.h., es existiert ein $a = a(t_0, x_0) \geq 0$, sodass das Anfangswertproblem (2.8) auf dem Intervall $[t_0 - a, t_0 + a]$ mindestens eine Lösung \hat{x} hat.

Beweis. (Quantitative Lösung) Sei $J \subset D$ abgeschlossen mit $t_0 \in J$ und $\hat{x}(t)$ eine Lösung von (2.8). Da die rechte Seite f auf J stetig ist, gilt nach dem Hauptsatz der Differential- und Integralrechnung:

$$\hat{x}(t) = \hat{x}_0 + \int_{t_0}^t f(s, \hat{x}(s)) ds \quad \text{für alle } t \in J.$$

Außerdem gilt für jede auf J stetige Funktion $\hat{x}(t)$, die obige Gleichung erfüllt, dass sie auf J differenzierbar ist. Dies bedeutet, dass eine auf J stetige Funktion $\hat{x}(t)$ löst genau dann das Anfangswertproblem (2.8), wenn es die obige Integralgleichung erfüllt. für eine auf J stetige Funktion $\hat{x}(t)$ gilt also:

$$\hat{x}(t) \text{ löst das AWP (2.8)} \Leftrightarrow \hat{x}(t) \text{ erfüllt obige Gleichung.}$$

Betrachte nun die Menge $K := \{x \in C(D) : \|x(t) - x_0\| \leq \beta \text{ für alle } t \in D\}$, wobei $C(D) = \{f : f(t) \text{ ist stetig für alle } t \in D\}$. K ist offensichtlich nicht leer (Betrachte bspw. die konstante Funktion $x \equiv x_0$). Jedem $x \in K$ wird nun eine Funktion $Ax \in C(D)$ zugewiesen mit der Definition

$$(Ax)(t) := x_0 + \int_{t_0}^t f(s, x(s)) ds \quad \text{für alle } t \in D.$$

Also löst Ax nach obiger Äquivalenz das Anfangswertproblem auf D . Außerdem gilt:

$$\begin{aligned} \|(Ax)(t) - x_0\| &= \left\| \int_{t_0}^t f(s, x(s)) ds \right\| \\ &\leq \int_{t_0}^t \|f(s, x(s))\| ds \\ &\leq |t - t_0| \max_{(t,y) \in \mathcal{G}} \|f(t, x)\|_2 \\ &\leq |t_0 + a - t_0| M \\ &\leq \frac{\beta}{M} M = \beta \quad \text{für alle } t \in D \end{aligned}$$

Das bedeutet, dass $Ax \in K$, bzw. $AK = K$. Nun können wir die hergeleitete Äquivalenz umformulieren: Für eine Funktion $\hat{x} \in K$ gilt

$$\hat{x} \text{ löst auf } D \text{ das Anfangswertproblem (2.8)} \Leftrightarrow \hat{x} \text{ ist ein Fixpunkt der Abbildung } A : K \rightarrow K.$$

Dazu nutzen wir den Fixpunktsatz von Schauder. Betrachte den Banachraum $C(D)$ mit der Maximumsnorm

$$\|x\|_\infty := \max_{t \in D} |x(t)|.$$

Sei $(x_n)_{n \in \mathbb{N}} \subset K$ eine konvergente Folge mit Grenzwert $x \in C(D)$. Dann ist

$$\|x - x_0\| = \left\| \lim_{n \rightarrow \infty} (x_n) - x_0 \right\| = \lim_{n \rightarrow \infty} \|(x_n) - x_0\| \leq \lim_{n \rightarrow \infty} \beta = \beta.$$

Also ist $x \in K$ und somit ist K abgeschlossen. Sei $x, y \in K$ bel. und betrachte $v := \lambda x + (1 - \lambda)y$ mit $0 \leq \lambda \leq 1$. Dann ist $v \in K$, da

$$\begin{aligned} \|v(t) - v_0\| &= \|\lambda x(t) + (1 - \lambda)y(t) - \lambda x(t_0) - (1 - \lambda)y(t_0)\| \\ &= \lambda \|x(t) - x(t_0)\| + (1 - \lambda) \|y(t) - y(t_0)\| \\ &\leq \beta(\lambda + 1 - \lambda) = \beta. \end{aligned}$$

Das heißt, K ist außerdem konvex. Die Stetigkeit der Abbildung $A : K \rightarrow K$ lässt sich wie folgt zeigen. Wegen der gleichmäßigen Stetigkeit von f auf \mathcal{G} können wir ein $\delta > 0$ so bestimmen, sodass gilt:

$$\|f(t, x) - f(t, y)\| < \frac{\epsilon}{a} \quad \text{für } \|x - y\| < \delta.$$

Betrachte nun $x, y \in K$ mit $\|x - y\|_\infty$ also auch $\|x(t) - y(t)\|$ für alle $t \in D$, so ist für $t \in D$ immer $\|f(t, x(t)) - f(t, y(t))\| < \frac{\epsilon}{a}$ und es folgt

$$\|(Ax)(t) - (Ay)(t)\| \leq |t - t_0| \frac{\epsilon}{a} = \epsilon \quad \text{für jedes } t \in D,$$

also auch $\|Ax - Ay\|_\infty \leq \epsilon$. Es bleibt jetzt nur noch zu zeigen, dass die Bildmenge $A(K)$ relativ kompakt ist. Sei hierfür $x \in K$ beliebig, dann ist

$$\|(Ax)(t)\| \leq \|x_0\| + aM \quad \text{für jedes } t \in D$$

und

$$\|(Ax)(t_1) - (Ax)(t_2)\| \leq |t_1 - t_2| M \quad \text{für alle } t_1, t_2 \in D.$$

Somit ist $A(K)$ punktweise beschränkt und gleichgradig stetig auf dem kompakten Intervall D . Nach dem Satz von Arzela-Ascoli [2, S. 49] hat also jede Folge $(Ax)_n \subset A(K)$ eine gleichmäßig konvergente Teilfolge. Dies gilt offensichtlich auch für die Maximumsnorm wodurch $A(K)$ relativ kompakt ist. \square

2.2.2 Eindeutigkeit von Lösungen

Ähnlich wie im vorherigen Kapitel existiert ein Satz, welcher zeigt, dass eine Lipschitz-stetige rechte Seite f ausreicht, damit eine eindeutige Lösung für eine gewöhnliche Differentialgleichung erster Ordnung existiert. Dazu definieren wir zuerst eine Lipschitz-stetige Funktion.

Definition 2.2.3 Sei $M \subset \mathbb{R} \times \mathbb{R}^n$ und $f : M \rightarrow \mathbb{R}^n$ eine Funktion. Dann heißt genau dann f Lipschitz-stetig auf M bezüglich der x -Variable, wenn ein $L > 0$ existiert, sodass gilt:

$$\|f(t, x) - f(t, y)\| \leq L \|x - y\|$$

für alle $(t, x), (t, y) \in M$.

Außerdem benötigen wir für den Beweis noch den Fixpunktsatz von Weissinger.

Satz 2.2.4 (Fixpunktsatz von Weissinger) Sei $(B, \|\cdot\|)$ ein Banachraum und $U \subset B$ abgeschlossen und nichtleer. Ferner ist $\sum_{j=1}^{\infty} \alpha_j$ eine konvergente Reihe positiver Zahlen, also $\sum_{j=1}^{\infty} \alpha_j < \infty$, $\alpha_j \geq 0$, sowie $A : U \rightarrow U$ eine Selbstabbildung, für die

$$\|A^j u - A^j v\| \leq \alpha_j \|u - v\| \quad \text{für alle } u, v \in U \quad \text{und } j \in \mathbb{N}$$

gilt. Dann besitzt A genau einen Fixpunkt in U , nämlich

$$u = \lim_{n \rightarrow \infty} A^n u_0$$

mit beliebigem $u_0 \in U$. u ist also der Grenzwert der rekursiven Folge $u_j := Au_{j-1}$ für $j \geq 1$ mit $u_0 \in U$ beliebig. Außerdem gilt die Fehlerabschätzung

$$\|u - u_n\| \leq \sum_{j=n}^{\infty} \alpha_j \|u_1 - u_0\|.$$

Beweis. [3, S. 139]

Nun können wir einen grundlegenden Satz in der Theorie für gewöhnliche Differentialgleichungen beweisen.

Satz 2.2.5 (Existenzsatz- und Eindeutigkeitssatz von Picard-Lindelöf) *Quantitative Version:* Seien

$$\mathcal{G} = \{(t, x) \in \mathbb{R} \times \mathbb{R}^n : |t - t_0| \leq \alpha, \quad \|x - x_0\|_2 \leq \beta, \quad \alpha, \beta \geq 0\},$$

$(t_0, x_0) \in \mathcal{G}$ und $f : \mathcal{G} \rightarrow \mathbb{R}^n$ stetig und Lipschitz-stetig bzgl. x . Dann besitzt das Anfangswertproblem (2.8) genau eine Lösung \hat{x} auf dem Intervall $D = \{t_0 - a, t_0 + a\}$, wobei

$$a = \min\left\{\alpha, \frac{\beta}{M}\right\}, \quad M = \max_{(t, y) \in \mathcal{G}} \|f(t, y)\|_2.$$

Qualitative Version: Seien $\mathcal{G} \subset \mathbb{R} \times \mathbb{R}^n$ offen und $f : \mathcal{G} \rightarrow \mathbb{R}^n$ stetig und lokal Lipschitz-stetig bzgl. x auf \mathcal{G} . Dann besitzt das Anfangswertproblem (2.8) für jedes Paar $(t_0, x_0) \in \mathcal{G}$ genau eine lokale Lösung, d.h., es existiert ein $a = a(t_0, x_0) \geq 0$, sodass das Anfangswertproblem (2.8) auf dem Intervall $[t_0 - a, t_0 + a]$ genau eine Lösung \hat{x} hat.

! TODO: Ungleichung beweisen ! *Beweis* Nun ist f Lipschitz-stetig, also setzen wir $B = C(D)$ mit der Maximumsnorm, $U = K$ und $A : K \rightarrow K$ aus dem Beweis von Peano. Aus

$$\|(A^j u)(t) - (A^j v)(t)\| \leq \frac{|t - t_0|^j}{j!} L^j \|u - v\|_{\infty}$$

für alle $j \in \mathbb{N}$ und $t \in D$, folgt direkt

$$\|A^j u - A^j v\|_{\infty} \leq \frac{(aL)^j}{j!} \|u - v\|_{\infty}.$$

Die Reihe $\sum_{j=1}^{\infty} \frac{(aL)^j}{j!}$ ist nach dem Quotientenkriterium offensichtlich konvergent. Also lässt sich der Fixpunktsatz von Weisinger anwenden. Der daraus gewonnene Fixpunkt ist eindeutig und mit ähnlicher Äquivalenz zu dem Beweis von Peano folgt, dass der Fixpunkt die gesuchte eindeutige Lösung ist. \square

2.3 Abhängigkeit der Lösungen von den Daten

In späteren Abschnitten werden wir gewöhnliche Differentialgleichungen betrachten, die zur Simulation/Vorhersage natürlicher Systeme genutzt werden. Darin ist es üblich, dass Anfangsdaten durch Messfehler oder unüblicher Verhalten von tatsächlichen Daten abweichen. Deshalb ist es sinnvoll Aussagen zu betrachten, die zeigen, welche Auswirkungen kleine Störungen auf die Lösung der Differentialgleichungen haben. In dieser Sektion ist die rechte Seite f stetig und Lipschitz-stetig bezüglich der x -Variable, sodass wir die Eindeutigkeit der Lösung durch Picard-Lindelöf garantieren. Der große Vorteil hierfür ist, dass wir keine Maximal- und Minimallösungen betrachten müssen. Um eine Aussage über die stetige Abhängigkeit der Anfangsdaten treffen zu können, beweisen wir zuerst einen wichtigen Hilfssatz.

Satz 2.3.1 (Gronwall'sche Ungleichung) Seien $D = [t_0, t_f]$ ein Intervall und die stetige, nicht-negative Funktion $u : D \rightarrow \mathbb{R}$ sowie $a \geq 0, b > 0$ gegeben. Des Weiteren gilt folgende Ungleichung:

$$u(t) \leq \alpha \int_{t_0}^t u(s) ds + \beta$$

für alle $t \in D$. Dann gilt:

$$u(t) \leq e^{\alpha(t-t_0)} \beta$$

für alle $t \in D$.

Beweis. Definiere zuerst eine Hilfsfunktion

$$v(t) := \alpha \int_{t_0}^t u(s) ds + \beta.$$

Für diese gilt

$$v'(t) = \alpha u(t) \leq \alpha v(t)$$

für alle $t \in D$. Daraus folgt

$$(e^{-\alpha t} v(t))' = e^{-\alpha t} (v'(t) - \alpha v(t)) \leq 0, \quad t \in D.$$

Die Funktion $e^{-\alpha t} v(t)$ ist also monoton fallend, das bedeutet

$$e^{-\alpha t} u(t) \leq e^{-\alpha t} v(t) \stackrel{t \geq t_0}{\leq} e^{-\alpha t_0} v(t_0) = e^{-\alpha t_0} \beta.$$

Daraus folgt die Behauptung. \square

Außerdem benötigen wir noch folgendes Lemma.

Lemma 2.3.2 Sei $T \subset \mathbb{R}^{1+n}$ offen und $f : T \rightarrow \mathbb{R}$ eine stetige Funktion, die zusätzlich Lipschitz-stetig bezüglich der x -Variable ist mit

$$\|f(t, x) - f(t, y)\|_2 \leq L \|x - y\|_2$$

für alle $(t, x), (t, y) \in T$ mit $L > 0$. Ist \hat{x} eine stetig-differenzierbare Funktion auf dem Intervall $D \subset \mathbb{R}$ und eine Lösung des Anfangswertproblems (2.8) und ist \hat{x}_a eine stetig-differenzierbare Funktion und eine Näherungslösung mit $(t, \hat{x}_a(t)) \in T$ für alle $t \in D$ und es gilt

$$\begin{aligned} \|\hat{x}'_a(t) - f(t, \hat{x}_a(t))\|_2 &\leq d_e \quad t \in D, \\ |t_0 - \tilde{t}_0| &\leq d_t, \\ \|x_0 - \hat{x}_a(\tilde{t}_0)\|_2 &\leq d_a \end{aligned}$$

(d_g representiert die Störung der rechten Seite, d_t die Störung der Anfangszeit und d_a die Störung des Anfangswerts). Dann gilt die Abschätzung

$$\|\hat{x}(t) - \hat{x}_a(t)\|_2 \leq e^{L|t-t_0|} (d_a + d_t(d_g + \sup_{s \in D} \|f(s, \hat{x}_a(s))\|) + \frac{d_g}{L}) + \frac{d_g}{L}.$$

Beweis. Betrachte zuerst die Differenz der Lösung \hat{x} und \hat{x}_a bei $t = t_0$.

$$\begin{aligned} \|\hat{x}(t_0) - \hat{x}_a(t_0)\| &= \left\| \hat{x}_0 - \int_{\tilde{t}_0}^{t_0} \hat{x}'_a(s) ds - \hat{x}_a(\tilde{t}_0) \right\| \\ &\leq \|\hat{x}_0 - \hat{x}_a(\tilde{t}_0)\| \left\| \int_{\tilde{t}_0}^{t_0} [\hat{x}'_a(s) - f(s, \hat{x}_a(s))] ds \right\| \left\| \int_{\tilde{t}_0}^{t_0} f(s, \hat{x}_a(s)) ds \right\| \\ &\leq d_a + d_t(d_g + \sup_{s \in D} \|f(s, \hat{x}_a(s))\|). \end{aligned}$$

Nun können wir mit Hilfe der Lipschitz-Eigenschaft der rechten Seite f die Differenz für allgemeines $t \in D, t > t_0$ abschätzen:

$$\begin{aligned} \|\hat{x}(t) - \hat{x}_a(t)\| &= \left\| \hat{x}_0 + \int_{t_0}^t f(s, \hat{x}) ds - \hat{x}_a(t_0) - \int_{t_0}^t \hat{x}'_a(s) ds \right\| \\ &\leq \|\hat{x}_0 - \hat{x}_a(t_0)\| + \int_{t_0}^t [\|f(s, \hat{x}(s)) - f(s, \hat{x}_a(s))\| + \|\hat{x}'_a(s) - f(s, \hat{x}_a(s))\|] ds \\ &\leq d_a + d_t(d_g + \sup_{s \in D} \|f(s, \hat{x}_a(s))\|) + \int_{t_0}^t [L \|\hat{x}(s) - \hat{x}_a(s)\| + d_g] ds. \end{aligned}$$

Um das gronwallsche Lemma verwenden zu können, setzen wir $u(t) := \|\hat{x}(t) - \hat{x}_a(t)\| + \frac{d_g}{L}$,

$$\beta := d_a + d_t(d_g + \sup_{s \in D} \|f(s, \hat{x}_a(s))\| + \frac{d_g}{L})$$

und $\alpha := L$. Offensichtlich gilt

$$\begin{aligned} u(t) &\leq \alpha \int_{t_0}^t u(s) ds + \beta \\ \Leftrightarrow \|\hat{x}(t) - \hat{x}_a(t)\| + \frac{d_g}{L} &\leq L \int_{t_0}^t \left[\|\hat{x}(s) - \hat{x}_a(s)\| + \frac{d_g}{L} \right] ds + \beta \\ \Leftrightarrow \|\hat{x}(t) - \hat{x}_a(t)\| &\leq d_a + d_t(d_g + \sup_{s \in D} \|f(s, \hat{x}_a(s))\|) + \int_{t_0}^t [L \|\hat{x}(s) - \hat{x}_a(s)\| + d_g] ds - \frac{d_g}{L} \end{aligned}$$

Also können wir das Gronwallsche Lemma anwenden und somit folgt

$$\|\hat{x}(t) - \hat{x}_a(t)\| + \frac{d_g}{L} \leq e^{L(t-t_0)} \left[d_a + d_t(d_g + \sup_{s \in D} \|f(s, \hat{x}_a(s))\| + \frac{d_g}{L}) \right].$$

Der Beweis für $t \in D$ mit $t < t_0$ funktioniert analog. \square

Mit diesem Lemma können wir etwas über die stetige Abhängigkeit der Lösung von der Zeitvariable aussagen.

Satz 2.3.3 Sei $T \subset \mathbb{R}^{1+n}$ offen und $f : T \rightarrow \mathbb{R}^n$ eine stetige Funktion, die Lipschitz-stetig bezüglich der x -Variable mit Konstante $L > 0$ gegeben. Dann hängt die Lösung x des Anfangswertproblems (2.8) stetig von den Anfangsdaten $(t_0, x_0) \in T$ und der rechten Seite f ab. Darunter versteht man: ist eine Lösung x auf einem kompakten Intervall $D \subset \mathbb{R}$, eine Umgebung U des Graphen $\{(t, x(t)) : t \in D\}$ in T und ein $\epsilon > 0$ gegeben, dann existiert ein $\delta(\epsilon, U, f, D) > 0$ in, sodass die Lösung \hat{x} des gestörten Anfangswertproblems

$$\begin{aligned} \hat{x}' &= \hat{f}(t, \hat{x}) \\ \hat{x}(\hat{t}_0) &= \hat{x}_0 \end{aligned}$$

auf ganz D existiert und die Abschätzung

$$\|x(t) - \hat{x}(t)\| \leq \epsilon \quad t \in D$$

erfüllt. Voraussetzungen hierfür sind, dass $\hat{t}_0 \in D$, $(\hat{t}_0, \hat{x}_0) \in T$, \hat{f} stetig auf U , Lipschitz-stetig bezüglich der x -Variable und

$$|t_0 - \hat{t}_0| \leq \delta, \quad \|x_0 - \hat{x}_0\| \leq \delta, \quad \|f(t, x) - \hat{f}(t, x)\| \leq \delta \quad \forall (t, x) \in U$$

gilt.

Beweis. [2, S. 67, 68]

Um die Abhängigkeit der Anfangsdaten (t_0, x_0) formulieren zu können, wird im Folgenden eine Notation eingeführt.

Definition 2.3.4 Seien $T \subset \mathbb{R}^{1+n}$ offen und $f : T \rightarrow \mathbb{R}^n$ eine stetige Funktion, die Lipschitz-stetig bezüglich der x -Variable ist. Die Abbildung

$$(t, t_0, x_0) \mapsto x(t; t_0, x_0)$$

mit $(t_0, x_0) \in T$ und $t \in I_{\max}(t_0, x_0) = (I^-(t_0, x_0), I^+(t_0, x_0))$ bezüglich der maximalen Lösung des Anfangswertproblems (2.8) und dem maximalen Existenzintervall $I_{\max}(t_0, x_0)$ heißt charakteristische Funktion der gewöhnlichen Differentialgleichung. Dabei nennt man

$$I^-(t_0, x_0) = \sup\{t \in \mathbb{R} : \text{das AWP (2.8) ist auf } [t_0, t] \text{ lösbar}\}$$

und

$$I^+(t_0, x_0) = \inf\{t \in \mathbb{R} : \text{das AWP (2.8) ist auf } [t, t_0] \text{ lösbar}\}$$

die Lebensdauerfunktion.

Satz 2.3.3 besagt, dass die charakteristische Funktion in allen Argumenten (t, t_0, x_0) stetig ist und dass $I^-(t_0, x_0)$ bzw. $I^+(t_0, x_0)$ ober- bzw. unterhalbstetig sind. Dies bedeutet, dass durch kleine Störungen der Anfangswerte (t_0, x_0) sich das maximale Existenzintervall nur stetig verkleinern kann.

Satz 2.3.5 Seien $T \subset \mathbb{R}^{1+n}$ und $f : T \rightarrow \mathbb{R}^n$ eine stetige Funktion, die in der x -Variable stetig differenzierbar ist. Dann ist die charakteristische Funktion $x(t; t_0, x_0)$ stetig differenzierbar in $(t_0, x_0) \in T$ und $t \in I_{\max}(t_0, x_0)$.

Beweis. [2, S. 69, 70]

Bemerkung 2.3.6 Man kann zeigen, dass für $T \subset \mathbb{R}^{1+n}$ und eine stetige Funktion $f : T \rightarrow \mathbb{R}^n$, die stetig differenzierbar in der x -Variable ist, gilt: f ist lokal Lipschitz-stetig in der x -Variable. Das bedeutet, dass in Satz 2.3.5 eine verstärkte Voraussetzung an die rechte Seite im Gegensatz zu Satz 2.3.3 verlangt wird.

3 Numerischer Lösungsansatz

Der Großteil dieses Abschnitts beruht im Wesentlichen auf [4, S. 63–79]. Unser Ziel ist es, eine effektive Methode zu verwenden, um eine Lösung eines Anfangswertproblem finden zu können. Da es nicht immer möglich ist, eine ODE analytisch zu lösen, gibt es sogenannte Ein- und Mehrschrittverfahren, welche eine Lösung der ODE approximiert. Der Grundgedanke dieser Verfahren ist es das Zeitintervall $D = [t_0, t_f]$ zu diskretisieren und beginnend mit dem Anfangswert t_0 eine Näherung $u_i \approx x(t_i)$ für $i = 0 \dots N$, wobei $t_N = t_f$, berechnet. Solche Verfahren werden mit Hilfe von Quadraturformeln der numerischen Integration [5, Numerische Integration] hergeleitet. Dies basiert darauf, dass die Differentialgleichung auf einem Teilintervall $[t_i, t_{i+1}] \subset D$ integriert wird und nach dem Hauptsatz der Differential- und Integralrechnung auf folgende Gleichung gebracht werden kann:

$$x(t_{i+1}) - x(t_i) = \int_{t_i}^{t_{i+1}} x'(t) dt = \int_{t_i}^{t_{i+1}} f(t, x(t)) dt.$$

Das Integral kann nun mit einer bereits genannten Quadraturformel approximiert werden, welches je nach Wahl der Quadraturformel ein Einschrittverfahren bildet.

3.1 Einschrittverfahren

Definition 3.1.1 Sei $D = [t_0, t_0 + a]$ ein Zeitintervall und eine Zerlegung von D

$$t_0 < t_1 < \dots < t_N = t_0 + a$$

mit der Schrittweite $h = \frac{a}{N}$, also $D_h = \{t_i = t_0 + ih \text{ für } i = 0, \dots, N\}$ eine Intervallzerlegung. Ein explizites Einschrittverfahren für das Anfangswertproblem (2.8) hat die Form

$$\begin{aligned} u_0 &= x_0 \\ u_{i+1} &= u_i + h\phi(t_i, u_i, h, f), \quad i = 0, \dots, N-1, \end{aligned} \quad (3.1)$$

wobei die Lösung $x : D \rightarrow \mathbb{R}^n$ auf ganz D existiert. Dabei nennt man $\phi : D_h \times \mathbb{R}^n \times \mathbb{R} \times C(D \times K, \mathbb{R}^n) \rightarrow \mathbb{R}^n$ die Inkrementfunktion, wobei $K \subseteq \mathbb{R}^n$. Implizite Einschrittverfahren haben die Form

$$\begin{aligned} u_0 &= x_0 \\ u_{i+1} &= u_i + h\phi(t_i, u_i, u_{i+1}, h, f), \quad i = 0, \dots, N-1, \end{aligned}$$

mit $\phi : D_h \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \times C(D \times K, \mathbb{R}^n) \rightarrow \mathbb{R}^n$.

3.1.1 Konsistenz und Konvergenz

Zur Analyse der Approximationsqualität der Verfahren werden neue Begriffe erläutert.

Definition 3.1.2 (lokaler Diskretisierungsfehler) Sei das Anfangswertproblem (2.8) mit Lipschitz-stetiger rechten Seite und ein explizites Einschrittverfahren (3.1) gegeben. Für $\hat{t} \in [t_0, t_0 + a]$ und $0 < h < t_0 + a - \hat{t}$ ist der lokale Diskretisierungsfehler definiert als

$$\tau(\hat{t}, h) := \frac{x(\hat{t} + h) - u_1(\hat{t})}{h},$$

wobei $u_1(\hat{t}) = x(\hat{t}) + h\phi(\hat{t}, x(\hat{t}), h, f)$ ist die Approximation der exakten Lösung nach einem Schritt und mit $x(\hat{t})$ als Startpunkt.

Falls zusätzlich für alle f mit stetiger und beschränkter Ableitung (bis zur Ordnung p) in der x -Variable gilt, dass für alle $\hat{t} \in (t_0, t_0 + a]$

$$\lim_{h \rightarrow 0} \tau(\hat{t}, h) = 0 \quad (\tau(\hat{t}, h) = \mathcal{O}(h^p) \text{ für } h \rightarrow 0),$$

dann nennt man das Einschrittverfahren konsistent (von der Ordnung m).

Definition 3.1.3 (globaler Diskretisierungsfehler) Mit $\hat{t} = t_i = t_0 + ih, i = 1, \dots, N$ ist der globale Diskretisierungsfehler definiert als

$$e(\hat{t}, h) := x(\hat{t}) - u_i.$$

Falls zusätzlich für alle f mit stetiger und beschränkter Ableitung (bis zur Ordnung m) in der x -Variable gilt, dass für alle $\hat{t} \in (t_0, t_0 + a]$

$$\lim_{h \rightarrow 0} e(\hat{t}, h) = 0 \quad (e(\hat{t}, h) = \mathcal{O}(h^p) \text{ für } h \rightarrow 0),$$

dann nennt man das Einschrittverfahren konvergent (von der Ordnung p).

Es ist uns möglich, eine obere Schranke für den globalen Fehler zu finden. Dazu müssen wir zuerst die diskrete Version der Gronwall Ungleichung 2.3.1 beweisen.

Satz 3.1.4 (diskrete Gronwallsche Ungleichung) Gegeben seien die Folgen $(\alpha)_i, (\beta)_i, (\gamma)_i \geq 0$ mit $i \in \{0, \dots, N\}$ und es gilt

$$\gamma_{i+1} \leq (1 + \alpha_i)\gamma_i + \beta_i \quad \text{für } i = 0, \dots, N-1.$$

Dann gilt

$$\gamma_{i+1} \leq \left(\gamma_0 + \sum_{j=0}^i \beta_j \right) e^{\alpha_0 + \dots + \alpha_i} \quad \text{für } i = 0, \dots, N-1.$$

Beweis. Da $i = 1, \dots, N$ bietet sich ein Induktionsbeweis nach i an:

$$\begin{aligned} (i=0) : \quad \gamma_1 &\leq \underbrace{(1 + \alpha_0)}_{\leq e^{\alpha_0}} \gamma_0 + \underbrace{\beta_0}_{=\beta_0 e^0 \leq \beta_0 e^{\alpha_0}} \leq (\gamma_0 + \beta_0) e^{\alpha_0} \\ (IV) : \quad \gamma_i &\leq \left(\gamma_0 + \sum_{j=0}^{i-1} \beta_j \right) e^{\alpha_0 + \dots + \alpha_{i-1}} \\ (i \rightarrow i+1) : \quad \gamma_{i+1} &\leq (1 + \alpha_i)\gamma_i + \beta_i \\ &\stackrel{IV}{\leq} (1 + \alpha_i) \left(\gamma_0 + \sum_{j=0}^{i-1} \beta_j \right) e^{\alpha_0 + \dots + \alpha_{i-1}} + \beta_i \\ &\leq e^{\alpha_i} \left(\gamma_0 + \sum_{j=0}^{i-1} \beta_j \right) e^{\alpha_0 + \dots + \alpha_{i-1}} + \beta_i e^{\alpha_0 + \dots + \alpha_i} \\ &\leq \left(\gamma_0 + \sum_{j=0}^i \beta_j \right) e^{\alpha_0 + \dots + \alpha_i} \end{aligned}$$

Satz 3.1.5 Sei das Einschrittverfahren (3.1) und das Anfangswertproblem (2.8) gegeben. Ist die Inkrementfunktion ϕ in der x -Variable Lipschitz-stetig, also

$$\|\phi(t, x_1, h, f) - \phi(t, x_2, h, f)\| \leq L \|x_1 - x_2\|$$

für alle $(t, x_1, h), (t, x_2, h) \in D \times \mathbb{R}^n \times [0, h_0]$ mit $h_0, L > 0$, dann gilt für den globalen Fehler in $t_i = t_0 + ih$ die obere Schranke

$$\|e(t_i, h)\| \leq (\|e_0\| + (t_i - t_0)\tau_h) e^{L(t_i - t_0)}, \quad i = 1, \dots, N.$$

Dabei ist $e_0 = x(t_0) - u_0$ und $\tau_h = \max_{k=1, \dots, N} \|\tau(t_k, h)\|$.

Beweis. Löse zuerst die Gleichung des lokalen Diskretisierungsfehlers auf $x(t_{j+1})$ auf:

$$x(t_{j+1}) = x(t_j) + h\phi(t_j, x(t_j), h, f) + h\tau(t_j), \quad j = 0, \dots, i-1$$

Dann nutzen wir das Ergebnis für den globalen Fehler

$$e(t_{j+1}, h) = x(t_{j+1}) - u_{j+1} = x(t_j) - u_j + h\tau(t_j) + h(\phi(t_j, x(t_j), h, f) - \phi(t_j, u_j, h, f))$$

und schätzen ab

$$\|e(t_{j+1}, h)\| \leq \|e(t_j, h)\| + h\tau_h + hL\|e(t_j, h)\| = (1 + hL)\|e(t_j, h)\| + h\tau_h.$$

Nun können wir mit $\alpha_j = hL$, $\gamma_j = \|e(t_j, h)\|$ und $\beta_j = h\tau_h$ die diskrete Gronwall Ungleichung anwenden und erhalten

$$\begin{aligned} \|e(t_{j+1}, h)\| &\leq \left(\|e_0\| \sum_{i=0}^j h\tau_h e^{\sum_{i=0}^j hL} \right) \\ &= (\|e_0\| + (j+1)h\tau_h) e^{(j+1)hL} \\ &= (\|e_0\| + (t_{j+1} - t_0)\tau_h) e^{L(t_{j+1} - t_0)} \end{aligned}$$

für alle $j = 0, \dots, i-1$, also insbesondere

$$\|e(t_i, h)\| = (\|e_0\| + (t_i - t_0)\tau_h) e^{L(t_i - t_0)}$$

für $j = i-1$. □

Der Satz besagt also, dass die Konsistenz des Einschrittverfahrens bereits die Konvergenz impliziert. Im allgemeinen ist ein Verfahren genau dann konvergent, wenn es konsistent und stabil ist, wie wir später beweisen werden.

3.2 Runge-Kutta-Verfahren

Runge-Kutta-Verfahren sind Einschrittverfahren höherer Ordnung. Um ein RK-Verfahren herleiten zu können, müssen wir zuerst definieren, was eine Quadraturformel ist.

Definition 3.2.1 Eine Quadratur ist eine Abbildung $Q_n : C([a, b]) \rightarrow \mathbb{R}$ für die gilt

$$Q_n(f) = (b-a) \sum_{i=0}^n \alpha_i f(x_i) \quad \text{für alle } f \in C([a, b]),$$

wobei $\Delta = \{a = x_0, \dots, x_n = b\}$ die Einteilung Stützstellen und die Zahlen $\alpha_0, \dots, \alpha_n$ (Quadratur Gewichte genannt werden.

Für eine interpolische Quadratur gilt zusätzlich

$$\alpha_i = \int_0^1 \prod_{\substack{j=0 \\ j \neq i}}^n \frac{t - t_j}{t_i - t_j} dt, \quad t_j = \frac{x_j - a}{b - a}.$$

In unserem Fall ist $(b-a) = h$, da wir das Integral $\int_{t_i}^{t_{i+1}} f(s, x(s)) ds$ mit einer Quadratur approximieren:

$$\int_{t_i}^{t_{i+1}} f(s, x(s)) ds \approx h \sum_{l=1}^m \alpha_l f(s_l, x(s_l)).$$

Die Stützstellen sind also $s_1 = t_i$, $s_l = t_i + c_l h$ für $l = 2, \dots, m$. Dazu approximieren wir $f(s_l, x(s_l)) \approx k_{l,i}$ mit

$$\begin{aligned} k_{1,i} &= f(t_i, u_i) \\ k_{l,i} &= f(t_i + c_l h, u_i + h \sum_{o=1}^{l-1} a_{lo} k_{o,i}), \quad l = 2, \dots, m. \end{aligned} \tag{3.2}$$

Dies ergibt das m -stufige explizite Runge-Kutta-Verfahren

$$u_{i+1} = u_i + h \sum_{l=1}^m b_l k_{l,i}.$$

Wir werden später ein klassisches 4-stufiges RK-Verfahren verwenden, welches wir jetzt herleiten. Dazu benötigen wir die Simpsonregel

$$Q(f) = \frac{b-a}{6} (f(a) + 4f(m) + f(b))$$

mit $m = \frac{a+b}{2}$ und $a = (t_i, x(t_i))$, $b = (t_i + h, x(t_i + h))$. Dabei ist !!!TODO: m korrigieren!!!

$$m = \frac{a+b}{2} = \frac{(t_i, x(t_i)) + (t_i + h, x(t_i + h))}{2} = \frac{(2t_i + h, x(t_i) + x(t_i + h))}{2} = (t_i + \frac{h}{2}, x(t_i + \frac{h}{2}))$$

Es gilt also die Approximation

$$\begin{aligned} \int_{t_i}^{t_{i+1}} f(s, x(s)) ds &\approx \frac{h}{6} (f(t_i, x(t_i)) + 4f(m) + f(t_i + h, x(t_i + h))) \\ &= h \left(\frac{1}{6} f(t_i, x(t_i)) + \frac{2}{3} f\left(t_i + \frac{h}{2}, x\left(t_i + \frac{h}{2}\right)\right) + \frac{1}{6} f(t_{i+1}, x(t_{i+1})) \right). \end{aligned}$$

Damit ergibt sich das 4-stufige klassische Runge-Kutta-Verfahren

$$u_{i+1} = u_i + h \left(\frac{1}{6} k_{1,i} + \frac{1}{3} k_{2,i} + \frac{1}{3} k_{3,i} + \frac{1}{6} k_{4,i} \right)$$

mit den Stufen

$$\begin{aligned} k_{1,i} &= f(t_i, u_i) \\ k_{2,i} &= f\left(t_i + \frac{h}{2}, u_i + \frac{h}{2} k_{1,i}\right) \\ k_{3,i} &= f\left(t_i + \frac{h}{2}, u_i + \frac{h}{2} k_{2,i}\right) \\ k_{4,i} &= f\left(t_i + h, u_i + h k_{3,i}\right). \end{aligned}$$

In diesem Fall haben wir die Koeffizienten

$$b := (b_1, \dots, b_m)^\top = \begin{pmatrix} \frac{1}{6} \\ \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{6} \end{pmatrix}, \quad c := (c_1, \dots, c_m)^\top = \begin{pmatrix} 0 \\ \frac{1}{2} \\ \frac{1}{2} \\ 1 \end{pmatrix}, \quad A := (a)_{lo} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Im Allgemeinen gilt für explizite RK-Verfahren, dass $a_{lo} = 0$ für $l \leq o$.

Bemerkung 3.2.2 Die Koeffizienten charakterisieren ein Runge-Kutta-Verfahren, weshalb sie für üblich in sogenannten Butcher-Tabellen zusammengefasst werden:

$$\begin{array}{c|c} c & A \\ \hline & b^\top \end{array}$$

Da RK-Verfahren Einschrittverfahren sind, gibt es auch implizite Runge-Kutta-Verfahren mit der Form

$$\begin{aligned} k_{1,i} &= f(t_i + c_1 h, u_i + h(a_{11} k_{1,i} + \dots + a_{1m} k_{m,i})) \\ &\vdots \\ k_{m,i} &= f(t_i + c_m h, u_i + h(a_{m1} k_{1,i} + \dots + a_{mm} k_{m,i})) \\ u_{i+1} &= u_i + h(b_1 k_{1,i} + \dots + b_m k_{m,i}) \end{aligned}$$

Hier kann A im Gegensatz zu expliziten RK-Verfahren eine volle Matrix sein.

3.2.1 Konsistenz

Es wurde bereits gezeigt, dass die Konvergenz eines Einschrittverfahrens aus der Konsistenz folgt. Deshalb genügt es sich die Konsistenz(ordnung) der Runge-Kutta-Verfahren zu betrachten. Dabei spielt die Wahl der Koeffizienten eine große Rolle. Ziel ist es, die Koeffizienten so zu wählen, dass die Konsistenzordnung unseres Runge-Kutta-Verfahrens möglichst hoch ist. Dazu zeigen wir zuerst, dass unter bestimmten Voraussetzungen der Koeffizienten ein explizites RK-Verfahren invariant gegenüber Autonomisierung ist, das bedeutet, dass ein RK-Verfahren genau dann ein Anfangswertproblem (2.8) mit der Näherungslösung u_i approximiert, wenn es das autonomisierte Anfangswertproblem

$$\begin{aligned} z' &= 1, & z(t_0) &= t_0 \\ x' &= f(z, x), & x(t_0) &= x_0. \end{aligned} \quad (3.3)$$

mit der Näherungslösung $v_i := \begin{bmatrix} t_i \\ u_i \end{bmatrix}$ approximiert.

Satz 3.2.3 *Gegeben sei ein explizites Runge-Kutta-Verfahren. Dies ist genau dann invariant gegenüber Autonomisierung, wenn für die Koeffizienten gilt:*

$$\sum_{j=1}^m b_j = 1 \quad \text{und} \quad c_l = \sum_{j=1}^{l-1} a_{lj}, \quad l = 1, \dots, m.$$

Beweis. Für ein autonomes Anfangswertproblem (??) hat das Runge-Kutta-Verfahren die Form

$$\begin{aligned} \hat{k}_{l,i} &= \begin{bmatrix} \hat{k}_{l,i}^z \\ \hat{k}_{l,i}^x \end{bmatrix} = \begin{bmatrix} 1 \\ f\left(t_i + h \sum_{o=1}^{l-1} a_{lo} \cdot \hat{k}_{o,i}^z, u_i + h \sum_{o=1}^{l-1} a_{lo} \hat{k}_{o,i}^x\right) \end{bmatrix} \quad l = 1, \dots, m, \\ v_{i+1} &= v_i + h \sum_{l=1}^m b_l \hat{k}_{l,i} = \begin{bmatrix} t_i + h \sum_{l=1}^m b_l \cdot 1 \\ u_i + h \sum_{l=1}^m b_l \hat{k}_{l,i}^x \end{bmatrix} \stackrel{!}{=} \begin{bmatrix} t_{i+1} \\ u_{i+1} \end{bmatrix}. \end{aligned}$$

Es gilt also genau dann $t_i + h \sum_{l=1}^m b_l \cdot 1 = t_{i+1}$, wenn $\sum_{l=1}^m b_l = 1$. Für die zweite Komponente gilt genau dann $u_i + h \sum_{l=1}^m b_l \hat{k}_{l,i}^x = u_{i+1}$, wenn $\hat{k}_{l,i}^x = k_{l,i}$, $l = 1, \dots, m$. Äquivalent dazu muss gelten

$$f\left(t_i + h \sum_{o=1}^{l-1} a_{lo} \cdot 1, u_i + h \sum_{o=1}^{l-1} a_{lo} \hat{k}_{o,i}^x\right) = f\left(t_i + h c_l, u_i + h \sum_{o=1}^{l-1} a_{lo} k_{o,i}\right), \quad l = 1, \dots, m,$$

was genau dann erfüllt ist, wenn $c_l = \sum_{o=1}^{l-1} a_{lo}$, $l = 1, \dots, m$. □

Im Folgenden werden wir uns auf autonome Anfangswertprobleme der Form

$$x'^0 = f(x), \quad x(t_0) = x_0 \quad (3.4)$$

beschränken. Unser Ziel ist es Voraussetzungen an die Koeffizienten des RK-Verfahrens zu setzen, so dass wir eine größtmögliche Konsistenzordnung erhalten. Dazu betrachten wir den lokalen Diskretisierungsfehler

$$\begin{aligned} \tau(t_i, h) &= \frac{x(t_i + h) - x(t_i)}{h} - \phi(t_i, x(t_i), h, f) \\ &= \frac{x(t_i + h) - x(t_i)}{h} - \sum_{l=1}^m b_l k_{l,i}. \end{aligned} \quad (3.5)$$

Wenden wir nun die Taylorentwicklung auf die exakte Lösung x für das AWP (3.4) an, erhalten wir

$$\begin{aligned} x(t_{i+1}) &= x(t_i) + hx'(t_i) + \frac{1}{2}h^2x''(t_i) + \mathcal{O}(h^3) \\ &= x(t_i) + hf(x(t_i)) + \frac{1}{2}h^2Df(x(t_i))f(x(t_i)) + \mathcal{O}(h^3), \end{aligned}$$

wobei für die Ableitungen der Lösung x die Kettenregel benutzt wird

$$\begin{aligned} x'(t_i) &= f(x(t_i)) \\ x''(t_i) &= Df(x(t_i))x'(t_i) = Df(x(t_i))f(x(t_i)) \end{aligned}$$

und $Df(t_i)$ die Jacobi-Matrix von f in $x(t_i)$ ist. Da wir uns nur auf autonome AWP's (3.4) beschränken, hat das explizite Runge-Kutta-Verfahren die Stufen

$$\begin{aligned} k_{i,1} &= f(x(t_i)), \\ k_{i,2} &= f(x(t_i) + ha_{21}k_{1,i}), \\ &\vdots \\ k_{i,m} &= f(x(t_i) + ha_{m1}k_{1,i} + \cdots + ha_{m,m-1}k_{m-1,i}). \end{aligned}$$

Nun wenden ähnlich wie oben die Taylorentwicklung im Punkt $x(t_i)$ an und erhalten

$$\begin{aligned} k_{i,1} &= f(x(t_i)), \\ k_{i,2} &= f(x(t_i)) + ha_{21}Df(x(t_i))k_{1,i} + \mathcal{O}(h^2), \\ &\vdots \\ k_{i,m} &= f(x(t_i)) + ha_{m1}Df(x(t_i))k_{1,i} + \cdots + ha_{m,m-1}Df(x(t_i))k_{m-1,i} + \mathcal{O}(h^2) \end{aligned}$$

Im nächsten Schritt setzen wir die Stufen sukzessive ein und verwenden !!!TODO: checken ob das stimmt:!!!

$$hk_{j,i} = hf(x(t_i)) + h^2a_{j-1,l}Df(x(t_i))f(x(t_i)) + \mathcal{O}(h^2) = hf(x(t_i)) + \mathcal{O}(h^2)$$

$$\begin{aligned} k_{i,1} &= f(x(t_i)), \\ k_{i,2} &= f(x(t_i)) + ha_{21}Df(x(t_i))f(x(t_i)) + \mathcal{O}(h^2), \\ &\vdots \\ k_{i,m} &= f(x(t_i)) + ha_{m1}Df(x(t_i))f(x(t_i)) + \cdots + ha_{m,m-1}Df(x(t_i))f(x(t_i)) + \mathcal{O}(h^2). \end{aligned}$$

Einsetzen in (3.5) ergibt

$$\begin{aligned} \tau(t_i, h) &= f(x(t_i)) + \frac{1}{2}hD(f(x(t_i)))f(x(t_i)) + \mathcal{O}(h^2) \\ &\quad - \sum_{l=1}^m b_l \left(f(x(t_i)) + h \sum_{j=1}^{m-1} a_{lj}D(f(x(t_i)))f(x(t_i)) + \mathcal{O}(h^2) \right) \\ &= \left(1 - \sum_{l=1}^m b_l \right) f(x(t_i)) + h \left(\frac{1}{2} - \sum_{l=1}^m \sum_{j=1}^{l-1} b_l a_{lj} \right) Df(x(t_i))f(x(t_i)) + \mathcal{O}(h^2) \end{aligned}$$

Hieraus kann schnell abgelesen werden, dass für die gegenüber Autonomisierung invariante Runge-Kutta-Verfahren die Konsistenzordnung $p = 1$ gilt. Außerdem bilden sich weitere Voraussetzungen für höhere Konsistenzordnungen, wie beispielsweise für $p = 2$:

$$\sum_{l=1}^m \sum_{j=1}^{l-1} b_l a_{lj} = \sum_{l=1}^m b_l c_l = \frac{1}{2}.$$

Allgemein lassen sich für höhere Konsistenzordnungen weitere Voraussetzung durch Taylorentwicklungen mit höheren h -Potenzen finden.

Satz 3.2.4 (Butcherschränken) Für die maximale erreichbare Ordnung p eines expliziten Runge-Kutta-Verfahrens mit m Stufen gilt

$$p \leq s.$$

Außerdem kann man folgende Abschätzungen finden

Stufe m	1	2	3	4	5	6	7	8	9	≥ 9
Ordnung p	1	2	3	4	4	5	6	6	7	$\leq m - 2$

TODO: Beweis

3.2.2 Schrittweitensteuerung

Wir haben im letzten Abschnitt gesehen, unter welchen Bedingungen die Konvergenz der Runge-Kutta-Verfahren die exakte Lösung der gegebenen Differentialgleichung gewährleistet ist. Unser Ziel ist es nun die bisher konstante Schrittweite h so zu wählen, dass der Rechenaufwand zur Approximation der Lösung so gering wie möglich gehalten wird. Dazu gibt es im Grunde zwei Fälle: Ist der lokale Diskretisierungsfehler größer als eine gegebene Toleranz, so sollten wir die Schrittweite verkleinern, um den Fehler zu verringern. Ein Beispiel hierfür wäre eine oszillierende Lösung, bei der eine zu große Schrittweite zu hohem lokalen Fehler führen würde, da das Einschrittverfahren Wendungspunkte überspringen würde.

Ist der lokale Diskretisierungsfehler aber klein in Relation zur Toleranz, dann können wir die Schrittweite vergrößern, um den Rechenaufwand zu minimieren. Dies tritt beispielsweise bei nahezu linearen Lösungsfunktionen auf.

Also müssen wir uns zuerst überlegen, wie wir den lokalen Diskretisierungsfehler sinnvoll mit minimalen Rechenaufwand abschätzen, ohne die tatsächliche Lösung der DGL zu kennen. Dazu verwenden wir sogenannte *eingebettete Runge-Kutta-Verfahren*, welche aus zwei Einschrittverfahren

$$u_{i+1}^{(1)} = u_i + h_i \Phi_p(t_i, u_i, h_i, f)$$

und

$$u_{i+1}^{(2)} = u_i + h_i \Phi_{p+1}(t_i, u_i, h_i, f)$$

mit den Konsistenzordnungen p und $p + 1$ bestehen. Die Butcherschränken implizieren, dass die Anzahl der Stufen schneller steigt, als die Konsistenzordnung, weshalb es nur bis zu einem gewissen Grad sinnvoll ist, ein Runge-Kutta-Verfahren höherer Ordnung für $u^{(2)}$ zu verwenden. Deshalb beschränken wir uns deshalb auf Verfahren mit Konsistenzordnung $p + 1$. Das Verfahren nennt man *eingebettet*, wenn die für $u^{(1)}$ berechneten Stufen bei der Berechnung von $u^{(2)}$ wieder verwendet werden. Dabei werden Funktionsauswertungen gespart, was den Rechenaufwand und die Laufzeit reduziert. Die dazugehörige Butcher-Tabelle wird dementsprechend zu einer *kombinierten* Butcher-Tabelle erweitert

b	A
	$(b^{(1)})^\top$
	$(b^{(2)})^\top$

Dabei sind $b^{(1)}$ Koeffizienten von $u^{(1)}$ und $b^{(2)}$ Koeffizienten von $u^{(2)}$.

Wir definieren

$$\Delta_p := x(t_i + h_i) - u_{i+1}^{(1)} = x(t_i + h_i) - u_i - h_i \Phi_p(t_i, u_i, h_i, f) = h_i \tau^{(1)}(t_i, h_i) = \mathcal{O}(h_i^{p+1})$$

$$\Delta_{p+1} := x(t_i + h_i) - u_{i+1}^{(2)} = x(t_i + h_i) - u_i - h_i \Phi_{p+1}(t_i, u_i, h_i, f) = h_i \tau^{(2)}(t_i, h_i) = \mathcal{O}(h_i^{p+2})$$

Daraus erhalten wir eine Abschätzung des lokalen Diskretisierungsfehlers mit

$$\begin{aligned} h_i \tau_p &= \Delta_p = \Delta_{p+1} + h_i(\Phi_{p+1}(t_i, u_i, h_i, f) - \Phi_p(t_i, u_i, h_i, f)) \\ &\approx h_i(\Phi_{p+1}(t_i, u_i, h_i, f) - \Phi_p(t_i, u_i, h_i, f)), \end{aligned}$$

da $\Delta_{p+1} \approx 0$ für kleine h_i gilt. Also erhalten wir mit der alten Schrittweite h_{alt}

$$\Delta_{p,\text{alt}} \approx h_i(\Phi_{p+1}(t_i, u_i, h_i, f) - \Phi_p(t_i, u_i, h_i, f)) = \|u^{(2)} - u^{(1)}\|_2.$$

Da wir nun den lokalen Fehler abgeschätzt haben, können wir uns darum kümmern, die Schrittweite anzupassen. Sei die Toleranz mit ϵ gegeben. Bei zu großem Fehler, also $\Delta_{p,\text{alt}} > \epsilon$, wählen wir eine neue Schrittweite h_{neu} und wiederholen den aktuellen Schritt. Unser Ziel dabei ist es, dass der neue Fehler $\Delta_{p,\text{neu}}$ etwas kleiner ist als ϵ , was mit der Approximation $\Delta_{p,\text{neu}} \approx \rho\epsilon$ mit $0 < \rho < 1$ gesichert ist. Da $\Delta_p = \mathcal{O}(h_i^{p+1}) \approx ch_i^{p+1}$ für jede Schrittweite gilt, kann man folgern

$$\begin{aligned} \frac{\rho\epsilon}{h_{\text{neu}}^{p+1}} &\approx \frac{\Delta_{p,\text{neu}}}{h_{\text{neu}}^{p+1}} \approx c \approx \frac{\Delta_{p,\text{alt}}}{h_{\text{alt}}^{p+1}} \\ \Leftrightarrow h_{\text{neu}} &\approx h_{\text{alt}} \sqrt[p+1]{\frac{\rho\epsilon}{\Delta_{p,\text{alt}}}} \end{aligned}$$

Im anderen Fall ist der Fehler zu klein $\Delta_{p,\text{alt}} \approx 0$. Dann können wir die Schrittweite mit Hilfe einer Hochschaltungsbeschränkung anpassen. Dabei setzen wir $h_{\text{neu}} = \min(qh_{\text{alt}}, h_{\text{max}})$, wobei h_{max} die maximale Schrittweite und $q > 1$ der Hochschaltfaktor ist. Wird der Schritt akzeptiert, so setzen wir $u_{i+1} = u^{(2)}$ und führen die Approximation mit der neuen Schrittweite fort.

3.3 Extrapolation

TODO

3.4 Mehrschrittverfahren

3.4.1 Konsistenz und Konvergenz

3.4.2 Stabilität

3.5 Verfahren für steife Differentialgleichungen

4 neuronale Netze

4.1 Methodenbeschreibung

4.2 Gewichtsinitialisierung

4.3 Fehlerdiskussion (weighting function)

4.4 curriculum learning

5 Anwendungsbeispiele

Literatur

- [1] Jan Frederik Sundermeier: „Der Fixpunktsatz von Schauder“. In: (), S. 31.
- [2] Lisa Beck: *Gewöhnliche Differentialgleichungen*. 12. Feb. 2018.
- [3] Harro Heuser: *Gewöhnliche Differentialgleichungen*. 6. Auflage. ISBN: 978-3-8348-0705-2.
- [4] Tatjana Stykel: *Skript zur Vorlesung Numerik Gewöhnlicher Differentialgleichungen*. Sommersemester 2020.
- [5] Guido Walz: *Lexikon der Mathematik*. 1. Bd. 1. Sprektum Akademischer Verlag. ISBN: 3-8274-0439-8.

Abbildungsverzeichnis