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

Alexandro Jedaidi

Inhaltsverzeichnis

Zusammenfassung

Bachelorabeit 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*

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

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

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

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

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

Eine Lösung des Problems $\hat{x} : D \to \mathbb{R}$ muss also zusätzlich zu (I) *auch die Anfangswertbedingungen (vgl.* (II)*) erfüllen.*

Es ist möglich jede gewöhnliche Differentialgleichung $m - ter$ Ordnung zu einem System gewöhnlicher Differential gleichungen $erster$ 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. (I)). Diese ist mit Hilfe der Funktionen $x_j : D \to \mathbb{R}$ für $j \in \{1, \ldots, k\}$ äquivalent zu einem System erster Ordnung mit m Gleichungen:

$$\begin{aligned} x_1' &= x_2 \\ x_2' &= x_3 \\ &. \\ &. \\ &. \\ x_{m-1}' &= x_m \\ x_m' &= f(t, x_1, x_2, x_3, \ldots, x_m). \end{aligned} \tag{III}$$

Für ein AWP (II) gilt zusätzlich:

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

2.2 Existenz und Eindeutigkeit

In diesem Abschnitt betrachten wir ein Anfangswertproblem *erster* Ordnung

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

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

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

Qualitative Version: Seien $\mathcal{G} \in \mathbb{R} \times \mathbb{R}^n$ offen und $f : \mathcal{G} \to \mathbb{R}^n$ stetig. Dann besitzt das Anfangswertproblem (V) *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* (V) *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 (V). Da die rechte Seite f auf J stetig ist, gilt nach dem Hauptsatz der Differential- und Integralrechnung:

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

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 (V), wenn es die obige Integralgleichung erfüllt. für eine auf J stetige Funktion $\hat{x}(t)$ gilt also:

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

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

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

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

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

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

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

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

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

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

$$\begin{aligned}
\|v(t) - v_0\| &= \|\lambda x(t) + (1 - \lambda)y(t) - \lambda x(t_0) - (1 - \lambda)y(t_0)\| \\
&= \lambda \|x(t) - x_0\| + (1 - \lambda) \|y(t) - y_0\| \\
&\leq \beta(\lambda + 1 - \lambda) = \beta.
\end{aligned}$$

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

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

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

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

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

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

und

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

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 Maximimsnorm wodurch $A(K)$ relativ kompakt ist. □

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 $f : M \to \mathbb{R}^n$ 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 $(B, \|.\|)$ 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 $A : U \to U$ eine Selbstabbildung, für die*

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

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

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

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*

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

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

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

Qualitative Version: Seien $\mathcal{G} \in \mathbb{R} \times \mathbb{R}^n$ offen und $f : \mathcal{G} \to \mathbb{R}^n$ stetig und lokal Lipschitz-stetig bzgl. x auf $\mathcal{G}$. Dann besitzt das Anfangswertproblem (V) *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* (V) *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 $A : K \to K$ aus dem Beweis von Peano. Aus

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

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

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

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

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 diesem Sektion ist die rechte Seite f stetig und Lipschitz-stetig bezüglich der x-Variable, sodass wir die Eindeutigkeit der Lösung durch Picard-Lindelöff 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 (Gronwallsche Ungleichung) *Seien $D = [t_0, t_f]$ ein Intervall und die stetige, nicht-negative Funktion $u : D \to \mathbb{R}$ 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

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

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

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

Daraus folgt die Behauptung. □

Außerdem benötigen wir noch folgendes Lemma.

Lemma 2.3.2 *Sei $T \subset \mathbb{R}^{1+n}$ offen und $f : T \to \mathbb{R}$ 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 (V) *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*

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

(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. □

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 $f : T \to \mathbb{R}^n$ eine stetige Funktion, die Lipschitz-stetig bezüglich der x-Variable mit Konstante $L > 0$ gegeben. Dann hängt die Lösung x des Anfangswertproblems* (V) *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

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

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

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

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 $f : T \to \mathbb{R}^n$ 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 (V) *und dem maximalen Existenzintervall $I_{max}(t_0, x_0)$ heißt charakteristische Funktion der gewöhnlichen Differentialgleichung. Dabei nennt man*

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

und

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

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 $f : T \to \mathbb{R}^n$ eine stetige Funktion, die in der x-Variable stetig differenzierbar ist. Dann ist die charaktistische 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 $f : T \to \mathbb{R}^n$, 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 Vorraussetzung an die rechte Seite im Gegensatz zu Satz 2.3.3 verlangt wird.*

3 Numerischer Lösungsansatz

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 $i = 0 \ldots N$, wobei $t_N = t_f$, berechnet. Solche Verfahren werden mit Hilfe von Quadraturformeln der numerischen Integration [4, 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 Integralrechung 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 $D = [t_0, t_0 + \alpha]$ ein Zeitintervall und eine Zerlegung von D*

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

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

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

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

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

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

3.1.1 Fehlerdiskussion

Zur Analyse der Approximationsqualität der Verfahren werden neue Begriffe erläutert.

Definition 3.1.2 (lokaler Diskretisierungsfehler) *Sei das Anfangswertproblem* (V) *mit Lipschitz-stetiger rechten Seite und ein explizites Einschrittverfahren* (VI) *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 $u_1(\hat{t}) = x(\hat{t}) + h\phi\left(\hat{t}, x(\hat{t})\right., h, f)$ *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* m*) in der* x*-Variable gilt, dass für alle* $\hat{t} \in (t_0, t_0 + a]$

$$\lim_{h \to 0} \tau(\hat{t}, h) = 0,$$

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

Definition 3.1.3 (globaler Diskretisierungsfehler) *Mit* $\hat{t} = t_i = t_0 + ih, i = 1, \ldots, N$ *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* $\hat{t} \in (t_0, t_0 + a]$

$$\lim_{h \to 0} e(\hat{t}, h) = 0,$$

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

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* $i \in \{0, \ldots, N\}$ *und es gilt*

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

Dann gilt

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

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

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

Satz 3.1.5 *Sei das Einschrittverfahren* (VI) *und das Anfangswertproblem* (V) *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

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

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

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

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

Dann nutzen wir das Ergebnis für den globalen Fehler

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

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 $j = 0, \ldots, i-1$, also insbesondere

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

$\square$

3.2 Runge-Kutta-Verfahren

3.3 Mehrschrittverfahren

3.4 Fehlerdisskusion

3.4.1 Konsistenz und Konvergenz

3.4.2 Stabilität

4 neuronale Netze

4.1 Methodenbeschreibung

4.2 Gewichtsinitialisierung

4.3 Fehlerdisskussion (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] Guido Walz: *Lexikon der Mathematik*. 1. Bd. 1. Sprektum Akademischer Verlag. ISBN: 3-8274-0439-8.

Abbildungsverzeichnis