

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
2.2	Existenz und Eindeutigkeit	3
2.2.1	Existenz von Lösungen	3
2.2.2	Eindeutigkeit von Lösungen	3
2.3	Abhängigkeit der Lösungen von den Daten	3
3	Numerischer Lösungsansatz	6
3.1	Einschrittverfahren	6
3.1.1	Fehlerdiskussion	7
3.2	Runge-Kutta-Verfahren	8
3.3	Mehrschrittverfahren	8
3.4	Fehlerdiskussion	8
3.4.1	Konsistenz und Konvergenz	8
3.4.2	Stabilität	8
4	neuronale Netze	8
4.1	Methodenbeschreibung	8
4.2	Gewichtsinitialisierung	8
4.3	Fehlerdiskussion (weighting function)	8
4.4	curriculum learning	8
5	Anwendungsbeispiele	8
	Abbildungsverzeichnis	9

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 1 Ein System gewöhnlicher Differentialgleichungen m -ter Ordnung hat die Form

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

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

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

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

Es ist möglich jede gewöhnliche Differentialgleichung m -ter Ordnung zu einem System gewöhnlicher 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. (I)). 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:

$$\left\{ \begin{array}{l} x'_1 = x_2 \\ x'_2 = x_3 \\ \cdot \\ \cdot \\ \cdot \\ x'_{m-1} = x_m \\ x'_m = f(t, x_1, x_2, x_3, \dots, x_m). \end{array} \right. \quad (\text{III})$$

Für ein AWP (II) gilt zusätzlich:

$$\left\{ \begin{array}{l} x_1(t_0) = x_{0,1} \\ x_2(t_0) = x_{0,2} \\ \cdot \\ \cdot \\ \cdot \\ x_{m-1}(t_0) = x_{0,m-1} \\ x_m(t_0) = x_{0,m}. \end{array} \right. \quad (\text{IV})$$

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 betrachten wir einen Satz, welcher zeigt, dass die Stetigkeit der rechten Seite f für die Existenz einer Lösung ausreicht.

Satz 3 (Existenzsatz von Peano, quantitative und qualitative Version) *Quantitative Version: Seien*

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

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

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

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 (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 [1, S. 52–55]

2.2.2 Eindeutigkeit von Lösungen

Ähnlich wie im vorherigen Kapitel existiert ein Satz, welcher zeigt, dass eine *Lipschitz*-stetige [2] rechte Seite f ausreicht, damit eine eindeutige Lösung für eine gewöhnliche Differentialgleichung erster Ordnung existiert.

Satz 4 (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 (V) genau eine Lösung \hat{x} auf dem Intervall $D = \{t_0 - a, t_0 + a\}$, wobei

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

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 (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.

Beweis [1, S. 56–58]

2.3 Abhängigkeit der Lösungen von den Daten

In dieser Arbeit betrachten in späteren Abschnitten gewöhnliche Differentialgleichungen, die zur Simulation/Vorhersage natürlicher Systeme genutzt wird, worin es üblich ist, dass Anfangsdaten durch Messfehler oder ähnlicher Fehler 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-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 5 (Gronwallsche Ungleichung) Seien $D = [t_0, t_f]$ ein Intervall und die stetige, nichtnegative 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

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

Daraus folgt die Behauptung. □

Außerdem benötigen wir noch folgendes Lemma.

Lemma 6 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 (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

$$\|\hat{x}'_a(t) - f(t, \hat{x}_a(t))\|_2 \leq d_e \quad t \in D,$$

$$|t_0 - \tilde{t}_0| \leq d_t,$$

$$\|x_0 - \hat{x}_a(\tilde{t}_0)\|_2 \leq d_a$$

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

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

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

$$\begin{aligned} \|\hat{x}(t_0) - \hat{x}_a(t_0)\|_2 &= \left\| \hat{x}_0 - \int_{\tilde{t}_0}^{t_0} \hat{x}'_a(s) ds - \hat{x}_a(\tilde{t}_0) \right\|_2 \\ &\leq \| \hat{x}_0 - \hat{x}_a(\tilde{t}_0) \|_2 \left\| \int_{\tilde{t}_0}^{t_0} [\hat{x}'_a(s) - f(s, \hat{x}_a(s))] ds \right\|_2 \left\| \int_{\tilde{t}_0}^{t_0} f(s, \hat{x}_a(s)) ds \right\|_2 \\ &\leq d_a + d_t(d_g + \sup_{s \in D} \|f(s, \hat{x}_a(s))\|_2). \end{aligned}$$

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

$$\begin{aligned} \|\hat{x}(t) - \hat{x}_a(t)\|_2 &= \left\| \hat{x}_0 + \int_{t_0}^t f(s, \hat{x}) ds - \hat{x}_a(t_0) - \int_{t_0}^t \hat{x}'_a(s) ds \right\|_2 \\ &\leq \|\hat{x}_0 - \hat{x}_a(t_0)\|_2 + \int_{t_0}^t [\|f(s, \hat{x}(s)) - f(s, \hat{x}_a(s))\|_2 + \|\hat{x}'_a(s) - f(s, \hat{x}_a(s))\|_2] ds \\ &\leq d_a + d_t(d_g + \sup_{s \in D} \|f(s, \hat{x}_a(s))\|_2) + \int_{t_0}^t [L \|\hat{x}(s) - \hat{x}_a(s)\|_2 + d_g] ds. \end{aligned}$$

Um das Gronwall'sche Lemma verwenden zu können, setzen wir $u(t) := \|\hat{x}(t) - \hat{x}_a(t)\|_2 + \frac{d_g}{L}$,

$$\beta := d_a + d_t(d_g + \sup_{s \in D} \|f(s, \hat{x}_a(s))\|_2) + \frac{d_g}{L}$$

und $\alpha := L$. Offensichtlich gilt

$$\begin{aligned} u(t) &\leq \alpha \int_{t_0}^t u(s) ds + \beta \\ \Leftrightarrow \|\hat{x}(t) - \hat{x}_a(t)\|_2 + \frac{d_g}{L} &\leq L \int_{t_0}^t \left[\|\hat{x}(s) - \hat{x}_a(s)\|_2 + \frac{d_g}{L} \right] ds + \beta \\ \Leftrightarrow \|\hat{x}(t) - \hat{x}_a(t)\|_2 &\leq d_a + d_t(d_g + \sup_{s \in D} \|f(s, \hat{x}_a(s))\|_2) + \int_{t_0}^t [L \|\hat{x}(s) - \hat{x}_a(s)\|_2 + d_g] ds - \frac{d_g}{L} \end{aligned}$$

Also können wir das Gronwall'sche Lemma anwenden und somit folgt

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

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 7 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 (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

$$\|x(t) - \hat{x}(t)\|_2 \leq \epsilon \quad t \in D$$

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

$$|t_0 - \hat{t}_0| \leq \delta, \quad \|x_0 - \hat{x}_0\|_2 \leq \delta, \quad \|f(t, x) - \hat{f}(t, x)\|_2 \leq \delta \quad \forall (t, x) \in U$$

gilt.

Beweis. [1, S. 67, 68]

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

Definition 8 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 (V) und dem maximalen Existenzintervall $I_{\max}(t_0, x_0)$ heißt charakteristische Funktion der gewöhnlichen Differentialgleichung. Dabei nennt man

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

und

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

die Lebensdauerfunktion.

Satz 7 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 9 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. [1, S. 69, 70]

Bemerkung 10 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. Daraus folgt, dass in Satz 9 eine verstärkte Voraussetzung an die rechte Seite im Gegensatz zu Satz 7 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 $i = 0 \dots N$, wobei $t_N = t_f$, berechnet. Solche Verfahren werden mit Hilfe von Quadraturformeln der numerischen Integration [3, 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 11 Sei $D = [t_0, t_0 + \alpha]$ ein Zeitintervall und eine Zerlegung von D

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

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

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

wobei die Lösung $x : D \rightarrow \mathbb{R}^n$ auf ganz D existiert. Dabei nennt man mit $K \supseteq \mathbb{R}^n$ $\phi : D_h \times \mathbb{R}^n \times \mathbb{R} \times C(D \times K, \mathbb{R}^n) \rightarrow \mathbb{R}^n$ die Inkrementfunktion. 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 Fehlerdiskussion

Zur Analyse der Approximationsqualität der Verfahren werden neue Begriffe erläutert.

Definition 12 (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 $u_1(\hat{t}) = x(\hat{t}) + h\phi(\hat{t}, x(\hat{t}), h, f)$ 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 m) in der x -Variable gilt, dass für alle $\hat{t} \in (t_0, t_0 + a]$

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

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

Definition 13 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 $\hat{t} \in (t_0, t_0 + a]$

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

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

Um eine obere Schranke für den globalen Fehler finden zu können, müssen wir zuerst die diskrete Version der Gronwall Ungleichung ?? beweisen.

Satz 14 (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:

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

3.2 Runge-Kutta-Verfahren

3.3 Mehrschrittverfahren

3.4 Fehlerdiskussion

3.4.1 Konsistenz und Konvergenz

3.4.2 Stabilität

4 neuronale Netze

4.1 Methodenbeschreibung

4.2 Gewichtsinitialisierung

4.3 Fehlerdiskussion (weighting function)

4.4 curriculum learning

5 Anwendungsbeispiele

Literatur

- [1] Lisa Beck: *Gewöhnliche Differentialgleichungen*. 12. Feb. 2018.
- [2] *Lipschitz-Stetigkeit* – Serlo „Mathe für Nicht-Freaks“ – Wikibooks, *Sammlung freier Lehr-, Sach- und Fachbücher*. URL: https://de.wikibooks.org/wiki/Mathe_f%C3%BCr_Nicht-Freaks:_Lipschitz-Stetigkeit [13. Jan. 2022].
- [3] Guido Walz: *Lexikon der Mathematik*. 1. Bd. 1. Spektrum Akademischer Verlag. ISBN: 3-8274-0439-8.

Abbildungsverzeichnis