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 4

3 Numerischer Lösungsansatz **5**
3.1 Methodenbeschreibung 5
3.2 Fehlerdisskusion 5
3.2.1 Konsistenz und Konvergenz 5
3.2.2 Stabilität 5

4 neuronale Netze **5**
4.1 Methodenbeschreibung 5
4.2 Gewichtsinitialisierung 5
4.3 Fehlerdisskussion (weighting function) 5
4.4 curriculum learning 5

5 Anwendungsbeispiele **5**

Abbildungsverzeichnis **6**

Zusammenfassung

Bachelorabeit WiSe 2021/2022

1 Einleitung

Citing template[**Google**]

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*

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

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 *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:

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

Für ein AWP (II) gilt zusätzlich:

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

2.2 Existenz und Eindeutigkeit

In diesem Abschnitt betrachten wir ein Anfangswertproblem *erster* Ordnung

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

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

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

(Quantitative 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. (Qualitative Version)*

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 reicht, damit eine eindeutige Lösung für eine gewöhnliche Differentialgleichung erster Ordnung existiert.

Satz 4 *(Existenzsatz- und Eindeutigkeitssatz von Picard-Lindelöf, lokale quantitative und qualitative 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*

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

(Quantitative 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. (Qualitative Version)*

Beweis [1, S. 56–58] [3]

2.3 Abhängigkeit der Lösungen von den Daten

TODO: Motivation für den Abschnitt Um eine Aussage über die stetige Abhängigkeit der Anfangsdaten teffen 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 $u : D \to \mathbb{R}$ 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 6 *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

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

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

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

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 gronwallsche 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 gronwallsche 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. □

Nun können wir eine Abschätzung für eine Lösung eines Anfangswertproblems u und eine Lösung mit gestörten Anfangswerten

3 Numerischer Lösungsansatz

3.1 Methodenbeschreibung

3.2 Fehlerdisskusion

3.2.1 Konsistenz und Konvergenz

3.2.2 Stabilität

4 neuronale Netze

4.1 Methodenbeschreibung

4.2 Gewichtsinitialisierung

4.3 Fehlerdisskussion (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] Tatjana Stykel: *Skript zur Vorlesung Numerik Gewöhnlicher Differentialgleichungen*. Sommersemester 2020.

Abbildungsverzeichnis