Parameterschätzung bei Systemen gewöhnlicher Differentialgleichungen mit Anwendung in der Naturwissenschaft

Wissenschaftliche Arbeit
im Fachbereich Angewandte Mathematik
an der
Universität des Saarlandes
(Erstes Staatsexamen)

von

Katharina Elisabeth Pirrung

Saarbrücken
31. Juli 2010

Betreuer: Universitätsprofessor Dr. A. K. Louis

Selbstständigkeitserklärung

Hiermit erkläre ich an Eides statt, dass ich die vorliegende Arbeit eigenständig und ausschließlich mit den angegebenen Hilfsmitteln und Quellen angefertigt habe.

Saarbrücken, den 31. Juli 2010

Inhaltsverzeichnis

Kapitel 1

Einleitung

In den Naturwissenschaften werden viele zeitabhängige Phänomene mit gewöhnlichen Differentialgleichungen modelliert. Häufig sind dabei die Modellparameter nicht aus theoretischen Analysen zu bestimmen; vielmehr muss anhand von Messdaten versucht werden, die unbekannte Größe abzuleiten. In diesem Fall ist also nicht das *direkte Problem*, d.h. die numerische Simulation zu lösen, sondern das *inverse Problem*, bei welchem von der zu beobachtenden Wirkung eines Systems auf ihre Ursache geschlossen wird [5].
Der Kern der vorliegenden Arbeit bildet die Bestimmung der Modellparameter in nichtlinearen gewöhnlichen Differentialgleichungssystemen. Aus experimentell erhobenen bzw. synthetisch generierten Daten wird mit verschiedenen Algorithmen versucht, eine optimale Schätzung des Parameters zu liefern und gleichzeitig eine Aussage über die Effektivität bzw. Genauigkeit der Algorithmen zu treffen.
Nach einem Überblick über die verwendeten mathematischen Grundlagen im anschließenden Kapitel wird eine kurze Einführung in die Theorie und Numerik der Differentialgleichungen gegeben. Daraufhin werden die Ideen zur Parameterschätzung vorgestellt und basierend auf den Ergebnissen von Wikström [14] ein erster Algorithmus zum Auffinden des gesuchten Parametervektors vorgestellt.
Diese Methode bildet die Grundlage zur Entwicklung eigener Algorithmen, die in den Kapiteln 4 und 5 erläutert werden. Zunächst wird hierbei ein iteratives Verfahren vorgestellt: Die Idee ist, mit einem bereits berechneten Parametersatz abschnittsweise das Vorwärtsmodell zu lösen, um weitere Funktionswerte generieren zu können. Diese wiederum dienen einer verbesserten Integration der rechten Seite, was im Algorithmus von Wikström lediglich mit Hilfe der Messdaten erfolgt. Basierend auf der verbesserten Integration wird eine exaktere Näherung des Parametervektors erzielt.
In Kapitel 5 wird ein weiterer Ansatz zur Parameterschätzung präsentiert, der speziell auf verrauschte Daten abzielt. Hierbei löst man sich von der Integral-

auswertung mittels Quadraturformeln und integriert stattdessen exakt über eine glättende Splinefunktion.
Das resultierende Gleichungssystem zur Bestimmung des Momentevektors geht dabei über in eine *Tikhonov-Philips* regularisierte Version. Der Regularisierungsparameter ist in diesem Zusammenhang gekoppelt an eine *a priori*-Information bezüglich des Rauschmodells und kann mit Hilfe des *Verfahrens von Hebden* explizit bestimmt werden.
Die numerischen Experimente in dieser Arbeit geben Hinweise darauf, dass das in der Literatur beschriebene Verfahren überregularisiert. Durch Einführung eines Skalierungsfaktors konnten die Resultate jedoch verbessert werden.
Der Bezug zur Praxis wird durch Beispiele aus der Chemie und Biologie geliefert: Bei der analytisch-numerischen Umsetzung wird exemplarisch der Zerfall radioaktiver Substanzen sowie ein modifiziertes Räuber-Beute-Modell aus der Biologie betrachtet. Die vorgestellten Methoden lassen sich - sofern die Differentialgleichungssysteme die geforderte Struktur aufweisen - ohne Weiteres auf analoge Fragestellungen in Medizin, Physik und andere naturwissenschaftliche Sparten übertragen.
In Kapitel 7 werden die entwickelten Algorithmen der Methode von Wikström gegenübergestellt und entsprechende Vergleiche durchgeführt. Dabei stellt sich heraus, dass die neuen Varianten dem Grundalgorithmus nur zum Teil überlegen sind. Dies wird schließlich in Kapitel 8 diskutiert und im Anschluss daran ein Ausblick auf weitere Forschungsaktivitäten gegeben.

Kapitel 2

Mathematische Grundlagen

Um in den sich anschließenden Kapiteln ohne Weiteres von bestimmten Sätzen, Definitionen und Umformungen Gebrauch zu machen, wird zunächst die dazugehörige Theorie geklärt.

2.1 Die Normalgleichung

Das Ziel ist die Minimierung des Funktionals $I(x) := \|Ax - b\|_2^2$ mit $A \in \mathbb{R}^{n \times m}$, welche vollen Rang besitzt. Dabei bezeichnet $\| \, . \, \|_2$ die euklidische Norm in $\mathbb{R}^n$. Zunächst werden dazu die wichtigsten Eigenschaften des Standardskalarprodukts erläutert.

Definition 2.1 (siehe [9]). *Sei V ein reeller Vektorraum. Unter einem **Skalarprodukt auf V** versteht man die Abbildung*

$$V \times V \longrightarrow \mathbb{R}, \quad (v, w) \mapsto \langle v, w \rangle$$

mit folgenden Eigenschaften:

1. *Für jedes $w \in V$ ist $\langle \, \cdot \, , w \rangle ; V \longrightarrow \mathbb{R}$ linear, das heißt es ist stets $\langle v_1 + v_2, w \rangle = \langle v_1, w \rangle + \langle v_2, w \rangle$ und $\langle \lambda v, w \rangle = \lambda \langle v, w \rangle$, analog für festes v und $\langle v, \, \cdot \, \rangle ; V \longrightarrow \mathbb{R}$ („Bilinearität“).*
2. *$\langle v, w \rangle = \langle w, v \rangle$ für alle v, w („Symmetrie“).*
3. *$\langle v, v \rangle \geq 0$ für alle v, und $\langle v, v \rangle = 0$ nur für $v = 0$ („Positive Definitheit“).*

Die Ableitung einer vektorwertigen Funktion ist definiert als:

Definition 2.2 (Differenzierbarkeit (siehe [8])). *Sei G eine offene Teilmenge des $\mathbb{R}^p$ und ξ ein Punkt aus G. Dann heißt die Funktion $f : G \longrightarrow \mathbb{R}^q$ differenzierbar in ξ, wenn es eine $(q \times p)$-Matrix D gibt, so dass für alle $\xi + h$*

aus einer δ-Umgebung $U \subset G$ das Inkrement $f(\xi + h) - f(\xi)$ die Darstellung gestattet:

$$f(\xi + h) - f(\xi) = Dh + r(h) \quad mit \quad \lim_{\|h\| \to 0} \frac{r(h)}{\|h\|_2} = 0. \tag{2.1}$$

Dabei ist Dh definiert als $Dh := \nabla f \cdot h$.

Um $I(x)$ minimieren zu können, schreibt man das Funktional mit Hilfe des Skalarprodukts:

$$I(x) = \|Ax - b\|_2^2 = \langle Ax - b, Ax - b \rangle. \tag{2.2}$$

Für beliebige Vektoren $h \in \mathbb{R}^n$ berechnet man

$$\begin{aligned} I(x + h) &= \|A(x + h) - b\|_2^2 \\ &= \langle A(x + h) - b, A(x + h) - b_i \rangle \\ &= \langle Ax - b + Ah, Ax - b + Ah \rangle. \end{aligned}$$

Der Zähler des Differenzenquotienten aus Definiton 2.2 ergibt sich wie folgt:

$$\begin{aligned} I(x + h) - I(x) &= 2\langle Ax - b, Ah \rangle + \langle Ah, Ah \rangle \\ &= 2\langle Ax - b, Ah \rangle + \|Ah\|_2^2 \\ &= 2\langle A^T Ax - A^T b, h \rangle + \|Ah\|_2^2 \\ &= \underbrace{2(A^T Ax - A^T b)^T}_{=:D} h + \underbrace{\|Ah\|_2^2}_{r(h)}. \end{aligned} \tag{2.3}$$

Es gilt die Abschätzung

$$\begin{aligned} r(h) &= \|Ah\|_2^2 \\ &\leq \|A\|_2^2 \cdot \|h\|_2^2, \end{aligned}$$

wobei $\|A\|_2^2$ die mit der euklidischen Vektornorm verträgliche Spektralnorm bezeichnet. Damit ergibt sich für das Restglied:

$$\lim_{h \to 0} \frac{r(h)}{\|h\|_2} = \lim_{h \to 0} \|A\|_2^2 \cdot \|h\|_2 = 0.$$

Das Funktional $I(x)$ ist also differenzierbar in x mit der Ableitung $\nabla I(x) = 2(A^T(Ax - b))^T = D$. Damit folgt:

$$\begin{aligned} \nabla I(x) = 0 \quad &\Leftrightarrow \quad A^T Ax \quad = \quad A^T b \\ &\Leftrightarrow \quad x \quad = (A^T A)^{-1} A^T b \quad = A^+ b. \end{aligned} \tag{2.4}$$

Die Gleichung 2.4 wird als *Normalgleichung* bezeichnet und mit

$$A^+ = (A^T A)^{-1} A^T \tag{2.5}$$

als *Pseudoinverse* gelöst.

2.2 Das Hebden Verfahren

Im weiteren Verlauf dieser Arbeit spielt das sogenannte Verfahren von Hebden eine wichtige Rolle. Es repräsentiert ein Iterationsverfahren, das dazu genutzt wird, Nullstellen spezieller, nichtlinearer Funktionen zu bestimmen. Neben den bekanntesten Verfahren, dem Newton- und dem Sekantenverfahren, nimmt das Hebden-Verfahren eine Sonderstellung ein, da die Funktionen zunächst gewisse Voraussetzungen erfüllen müssen, damit dieses Verfahren angewendet werden darf. Sind diese erfüllt, konvergiert es schneller gegen die gesuchte Nullstelle als das Newton-Verfahren.

Zunächst zu den Voraussetzungen, die die Funktion erfüllen muss:

Satz 2.3. *Sei*

$$r(x) = \sum_{i=1}^{n} \frac{z_i^2}{(d_i + x)^2} \stackrel{!}{=} \rho \tag{2.6}$$

mit $z_i, d_i, \rho > 0$ *und* $d_1 > d_2 > ... > d_n$.
Dann gilt: $r(x)$ *hat genau dann eine Lösung* x_0*, wenn* $r(0) > \rho$ *erfüllt ist.*

Beweis:
Wegen $d_1 > d_2 > ... > d_n$ besitzt $r(x)$ genau n Polstellen $x_i = -d_i, i = 1, ..., M$. Aufgrund des quadratischen Nenners in r ist jeder Pol von der Ordnung Zwei, d.h. an den Polstellen findet kein Vorzeichenwechsel statt. Weiter gilt $-d_1 < -d_2 < ... < -d_n$ woraus ersichtlich ist, dass $-d_n < 0$ die größte Polstelle darstellt. Auf dem offenen Intervall $]-d_n, \infty[$ gelten offensichtlich folgende Grenzwerte:

$$\lim_{x \to -\infty} r(x) = 0 \quad \text{und}$$
$$\lim_{x \to -d_n^+} r(x) = \infty$$

Für $x > 0$ nimmt die Funktion r Werte zwischen 0 und $r(0)$ an. Das bedeutet: Liegt ρ im Intervall $]0, r(0)[$, so existiert ein $x \in]0, \infty[$ mit $r(x) = \rho$. Gilt dagegen $\rho > r(0)$, so gibt es keine Lösung der Gleichung (2.6). Da nach Voraussetzung ρ positiv ist, besitzt die Gleichung $r(x) = \rho$ genau dann eine Lösung, wenn $r(0) > \rho$ erfüllt ist. □

Um zum Hebden-Verfahren zu gelangen, betrachtet man zunächst das Newton-Verfahren zur Lösung eines Nullstellenproblems $f(x) = 0$. Für gegebenen Startwert x_0 schreibt sich das Iterationsverfahren in der Form

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}, \qquad k = 0, 1, ... \tag{2.7}$$

Da die Tangenten beim Newtonverfahren für x nahe 0 sehr steil sind, konvergiert das Iterationsverfahren zunächst relativ langsam. Allerdings lässt sich für große x folgendes asymptotisches Verhalten für r beobachten:

$$\begin{aligned} r(x) &= \sum_{i=1}^{n} \frac{z_i^2}{(d_i + x)^2} \\ &\approx \sum_{i=1}^{n} \frac{z_i^2}{x^2} \\ &= \frac{\|z\|_2^2}{x^2} \end{aligned}$$

Damit folgt:

$$\begin{aligned} & r(x) && \approx \frac{\|z\|_2^2}{x^2} \\ \Leftrightarrow \qquad & \frac{1}{\sqrt{r(x)}} && = \frac{x}{\|z\|_2}. \end{aligned}$$

Das bedeutet, dass sich $\frac{1}{\sqrt{r(x)}}$ für große x wie eine Ursprungsgerade verhält. Anschaulich ist klar, dass das Newton-Verfahren für eine fast lineare Funktion sehr schnell konvergiert. Wegen

$$\begin{aligned} & r(x) && = \rho && = 0 \\ \Leftrightarrow \qquad & \frac{1}{\sqrt{r(x)}} && = \frac{1}{\sqrt{\rho}} && = 0 \end{aligned}$$

und aufgrund des asymptotischen Verhaltens von $\frac{1}{\sqrt{r(x)}}$ wendet man das Newton-Verfahren auf

$$k(x) = \frac{1}{\sqrt{r(x)}} - \frac{1}{\sqrt{\rho}} \stackrel{!}{=} 0 \tag{2.8}$$

an und erwartet eine rasche Konvergenz. Mit

$$k'(x) = -\frac{r'(x)}{2r^{3/2}(x)}$$

berechnet man schließlich die Hebden-Iterierten:

$$\begin{aligned} x_{k+1} &= x_k - \frac{k(x_k)}{k'(x_k)} \\ &= x_k + \frac{2r(x_k)(1 - \sqrt{\frac{r(x_k}{\rho}})}{r'(x_k))}. \end{aligned} \tag{2.9}$$

2.3 Vorüberlegungen zu glättenden Splines

In Kapitel 5 tritt unter anderem folgendes Problem auf:

$$r(\lambda) := \left\|(I + \lambda T^{-1} G T^{-1})^{-1} y\right\|_2^2 \stackrel{!}{=} (l-1)\delta^2 \tag{2.10}$$

mit $l \in \mathbb{N}$. Dabei ist vorauszusetzen, dass $G \in \mathbb{R}^{n\times n}$ eine symmetrische, positiv definite (=: spd) Matrix sei.
In dieser Gleichung soll λ mit Hilfe des Hebdenverfahrens bestimmt werden. Dazu muss zunächst gezeigt werden, dass $r(\lambda)$ der Form aus Abschnitt 2.2 entspricht mit

$$r(\lambda) = \sum_{i=1}^{n} \frac{z^2}{(d_i + \lambda)^2} \stackrel{!}{=} \delta.$$

Zunächst jedoch folgendes Lemma vorweg:

Lemma 2.4. *Sei $T^{-1} \in \mathbb{R}^{n\times n}$ eine reguläre Matrix und $G \in \mathbb{R}^{n\times n}$ spd. Dann ist die Matrix $A := T^{-1}GT^{-1}$ symmetrisch und positiv definit.*

Beweis:
Es ist zu zeigen, dass

$$x^T A x > 0 \qquad \text{für alle} \quad x \in \mathbb{R}^n \backslash \{0\}.$$

Sei also $0 \neq x \in \mathbb{R}^n$ beliebig. Es gilt:

$$\begin{aligned} x^T A x &= x^T T^{-1} G T^{-1} x \\ &= x^T (T^{-1})^T G T^{-1} x \\ &= (T^{-1}x)^T G T^{-1} x \\ &= y^T G y > 0 \qquad\qquad \text{mit} \quad y = T^{-1}x. \end{aligned}$$

Da T^{-1} regulär ist und $x \neq 0$ gilt: $y = T^{-1}x \neq 0$. Aufgrund der positiven Definitheit von G ist $y^T G y = x^T A x > 0$ für alle $x \neq 0$. □

Satz 2.5. *Sei* $r(\lambda) = \|(I + \lambda T^{-1} G T^{-1})^{-1} y\|_2^2$ *mit einer symmetrischen, regulären Matrix* $T^{-1} \in \mathbb{R}^{n \times n}$*, einer spd-Matrix* $G \in \mathbb{R}^{n \times n}$ *und* $y \in \mathbb{R}^n$*. Dann lässt sich* $r(\lambda)$ *schreiben als:*

$$r(\lambda) = \sum_{i=1}^{n} \frac{z^2}{(d_i + \lambda)^2}$$

Beweis:
Zur Notation: Mit $A := T^{-1} G T^{-1}$ schreibt sich $r(\lambda)$ in der Form

$$r(\lambda) = \left\|(I + \lambda A)^{-1} y\right\|_2^2.$$

Man betrachtet nun die Matrix A. Da mit Lemma 2.4 gezeigt wurde, dass A spd ist, gilt dies auch für A^{-1}.
Sei $\{u_1, ..., u_n\}$ die Orthonormalbasis des $\mathbb{R}^n$ mit Eigenvektoren u_i der Matrix $A^{-1} = TG^{-1}T$. Es gilt also für beliebige Eigenvektoren u_i zum Eigenwert $d_i > 0$, dass

$$\begin{aligned}
& A^{-1}u_i && = d_i u_i \\
\Leftrightarrow \quad & Au_i && = \frac{1}{d_i} u_i \\
\overset{\lambda > 0}{\Leftrightarrow} \quad & \lambda A u_i && = \frac{\lambda}{d_i} u_i \\
\overset{+u_i}{\Leftrightarrow} \quad & (I + \lambda A) u_i && = (1 + \frac{\lambda}{d_i}) u_i \\
\Leftrightarrow \quad & (\frac{1}{1 + \frac{\lambda}{d_i}}) u_i && = (I + \lambda A)^{-1} u_i.
\end{aligned}$$

Schreibe $y \in \mathbb{R}^n$ als Linearkombination der u_i:

$$y = \sum_{i=1}^{n} w_i u_i. \tag{2.11}$$

Es folgt:

$$\begin{aligned}
(I + \lambda A)^{-1} y &= \sum_{i=1}^{n} w_i (I + \lambda A)^{-1} u_i \\
&= \sum_{i=1}^{n} \frac{w_i}{1 + \frac{\lambda}{d_i}} u_i.
\end{aligned}$$

Dann gilt:

$$\begin{aligned}
\left\|(I+\lambda A)^{-1}y\right\|_2^2 &= \langle (I+\lambda A)^{-1}y, (I+\lambda A)^{-1}y\rangle \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{w_i w_j}{(1+\frac{\lambda}{d_i})(1+\frac{\lambda}{d_i})} \underbrace{\langle u_i, u_j\rangle}_{\delta_{ij}=1 \text{ für } i=j, \text{ sonst } 0} \\
&= \sum_{i=1}^{n} \frac{w_i^2}{(1+\frac{\lambda}{d_i})^2} \\
&= \sum_{i=1}^{n} \frac{(d_i w_i)^2}{(d_i+\lambda)^2} \\
&= \sum_{i=1}^{n} \frac{z_i^2}{(d_i+\lambda)^2} \qquad \text{mit} \qquad z_i = d_i \cdot w_i.
\end{aligned}$$

□

Um nun das Hebdenverfahren anwenden zu können, müssen noch einige Überlegungen zur Darstellung von $r'(\lambda)$ getätigt werden.

Satz 2.6. *Sei*

$$r(\lambda) = \left\|(I+\lambda T^{-1}GT^{-1})^{-1}y\right\|_2^2$$

mit $T^{-1}, G \in \mathbb{R}^{n\times n}$ regulär, symmetrisch und G zusätzlich positiv definit. Dann hat die Ableitung die Darstellung

$$r'(\lambda) = -2\langle (I+\lambda A)^{-1}y, (I+\lambda A)^{-1}A(I+\lambda A)^{-1}y\rangle \tag{2.12}$$

mit $A = T^{-1}GT^{-1}$.

Vor dem Beweis des Satzes 2.6 noch ein technisches Lemma:

Lemma 2.7. *Seien B, C, B+C reguläre, quadratische Matrizen aus $\mathbb{R}^{n\times n}$, dann gilt:*

$$(B+C)^{-1} - B^{-1} = -(B+C)^{-1}\cdot CB^{-1}.$$

Beweis:

$$\begin{aligned}
(B+C)^{-1} - B^{-1} &= (B+C)^{-1} - (B+C)^{-1}(B+C)B^{-1} \\
&= (B+C)^{-1}\cdot(I-(B+C)B^{-1}) \\
&= (B+C)^{-1}\cdot(I-I-CB^{-1}) \\
&= -(B+C)^{-1}\cdot CB^{-1}.
\end{aligned}$$

□

Nun zum Beweis des Satzes 2.6.

Beweis:
Die Kettenregel liefert:

$$\begin{aligned} r'(\lambda) &= \sum_{i=1}^{n} \frac{d}{d\lambda} \left(((I+\lambda A)^{-1}y)_i\right)^2 \\ &= \sum_{i=1}^{n} 2((I+\lambda A)^{-1}y)_i \cdot \frac{d}{d\lambda}((I+\lambda A)^{-1}y)_i \\ &= \sum_{i=1}^{n} 2((I+\lambda A)^{-1}y)_i \cdot (\frac{d}{d\lambda}(I+\lambda A)^{-1}y)_i. \end{aligned}$$

Die innere Ableitung bezüglich der skalaren Größe λ berechnet sich wie folgt:

$$\begin{aligned} \frac{d}{d\lambda}(I+\lambda A)^{-1} &= \lim_{h\to 0} \frac{(I+(\lambda+h)A)^{-1} - (I+\lambda A)^{-1}}{h} \\ &= \lim_{h\to 0} \frac{(I+\lambda+h\cdot A)^{-1} - (I+\lambda A)^{-1}}{h}. \end{aligned}$$

Aus Lemma 2.7 folgt mit $B = I + \lambda A$ und $C = h \cdot A$

$$\begin{aligned} \frac{d}{d\lambda}(I+\lambda A)^{-1} &= \lim_{h\to 0} -\frac{(I+\lambda A)^{-1} \cdot h \cdot A \cdot (I+\lambda A)^{-1}}{h} \\ &= -(I+\lambda A)^{-1} \cdot A \cdot (I+\lambda A)^{-1}. \end{aligned}$$

$$\begin{aligned} \Rightarrow r'(\lambda) &= -2\sum_{i=1}^{n}((I+\lambda A)^{-1}y)_i \cdot ((I+\lambda A)^{-1} \cdot A \cdot (I+\lambda A)^{-1}y)_i \\ &= -2\,\langle\,(I+\lambda A)^{-1}y, (I+\lambda A)^{-1} \cdot A \cdot (I+\lambda A)^{-1}y\,\rangle. \end{aligned}$$

□

2.4 Glättende Splines

Splines bestehen im Gegensatz zu interpolierenden Polynomen aus stückweise zusammengesetzten Polynomen niedrigen Grades. Dadurch haben sie den Vorteil, die gesuchte Funktion qualitativ besser zu approximieren, da man jeweils

nur auf den Subintervallen I_i operiert, wobei für jedes I_i ein eigenes Polynom berechnet wird. Der Grad des Polynoms ist abhängig von der gewünschten Genauigkeit bzw. Größenordnung des Fehlers und den Eigenschaften der zugrundeliegenden Funktion.
Auch wenn lineare und quadratische Splines zur Verfügung stehen, werden in den meisten Fällen kubische Splines verwendet, da der Mensch diese visuell als glatt empfindet.
Im Zusammenhang mit der Parameterschätzung ist besonders der Fall der sogenannten *geglätteten* oder *glättenden Splines* interessant. Im Folgenden werden die allgemeinen Kenntnisse zu interpolierenden kubischen Splines vorausgesetzt.
Um näher auf das Thema der glättenden Splines eingehen zu können, werden zunächst die wichtigsten Eigenschaften eines natürlichen kubischen Splines dargelegt.

Satz 2.8. *Ein natürlicher kubischer Spline $s \in S_{3,\triangle}$ mit $x \in I_i = [x_{i-1}, x_i] \subset [a, b], i = 1, ..., l$ hat die Darstellung:*

$$s(x) = y_i + s_i'(x - x_i) + \gamma_i \frac{(x - x_i)^2}{2} + \frac{\gamma_i - \gamma_{i-1}}{h_i} \frac{(x - x_i)^3}{6},$$
$$s''(a) = s''(b) = 0.$$

Der reelle Vektorraum $s \in S_{3,\triangle}$ stellt einen Unterraum der $\mathcal{C}^2$- Kurven dar, wobei $\triangle = x_0 < x_1 < ... < x_l = b$ ein Gitter über $[a, b]$ und $s'(x)$ die erste Ableitung des Splines s bezeichnet.

Die Koeffizienten $\gamma_i = s''(x_i)$ repräsentieren die sogenannten Momente des kubischen Splines und $h_i = x_i - x_{i-1}$ bezeichnet die Länge des Intervalls I_i.

Der Beweis kann in [6] nachgelesen werden.
Um nun zu einem glättenden Spline zu gelangen, überlegt man sich folgendes: Wurden zuvor die Messwerte y_i als Interpolationspunkte genommen, so geht man jetzt davon aus, dass diese in einer gewissen Fehlerbandbreite des Intervalls $[-\delta, \delta]$ zu der exakten Funktion y liegen. Hierbei wird zum einen die Funktion y als hinreichend glatt vorausgesetzt, zum anderen homogene Werte $y(a) = y(b) = 0$ an den Rändern gefordert. Der Spline wird jetzt nicht mehr über die Messwerte y_i bestimmt, sondern fordert eine „glatte“ Funktion s, welche die Interplationsaufgabe nach der „*Kleinste-Fehlerquadrat-Methode*“ löst. Das Problem lautet also folgendermaßen:

Problem 2.9. *Minimiere* $\|s''\|_{\mathcal{L}^2(a,b)}$ *unter allen hinreichend glatten Funktionen* s *mit* $s(a) = s(b) = 0$ *unter der Nebenbedingung*

$$\frac{1}{l-1}\sum_{i=1}^{l-1}|y_i - s(x_i)|^2 \leq \delta^2. \tag{2.13}$$

Dies führt zu folgendem Satz:

Satz 2.10. *Sei* s *ein natürlicher kubischer Spline über dem Gitter* $\triangle$ *mit* $s(a) = s(b) = 0$*, der die Nebenbedingung* (2.13) *mit Gleichheit erfüllt,*

$$\sum_{i=1}^{l-1}|y_i - s(x_i)|^2 = (l-1)\delta^2 \tag{2.14}$$

und dessen dritte Ableitung an den Gitterknoten für ein $\lambda > 0$ *die folgenden Sprünge aufweist:*

$$[s'''\,]_{x_i} := s'''(x_i+) - s'''(x_i-) = \lambda(y_i - s(x_i)) \tag{2.15}$$

$i = 1, ..., l-1$*. Dann ist* s *die eindeutig bestimmte Lösung von* (2.13)*.*

Der Beweis sowie auch der Satz selbst ist [6] zu entnehmen.

Nun ist zwar gezeigt, dass die Lösung eindeutig ist, jedoch ist noch keine Aussage getroffen worden, ob und unter welcher Voraussetzung sie überhaupt existiert! Die Frage wird im folgenden Satz 2.11 beantwortet, dessen konstruktiver Beweis zugleich einen Lösungsansatz liefert. Der Satz sowie der Beweis sind zum größten Teil ebenfalls aus dem Kapitel *„Geglättete kubische Splines“* von [6] übernommen.

Satz 2.11. *Unter der Voraussetzung*

$$\frac{1}{l-1}\sum_{i=1}^{l-1}|y_i|^2 > \delta^2$$

existiert ein Spline s, der alle Bedingungen aus Satz 2.10 erfüllt.

Beweis:
Zunächst nimmt man sich einen Vektor $c \in \mathbb{R}^{l-1}$ als Vektor der Momente $\gamma_i = s''(x_i)$ mit $i = 1, ..., l-1$ eines natürlichen kubischen Splines s. Aus der Theorie der kubischen Splines ist bekannt, dass der Zusammenhang

$$Gc = Ts \tag{2.16}$$

zwischen c und den Funktionswerten $s = (s_1, ..., s_{l-1})^T \in \mathbb{R}$, $s_i = s(x_i)$ besteht. Die $(l-1) \times (l-1)$-Matrizen G und T sind dabei wie folgt definiert:

$$G = \frac{1}{6} \begin{bmatrix} 2(h_1+h_2) & h_2 & & 0 \\ h_2 & 2(h_2+h_3) & \ddots & \\ & \ddots & \ddots & h_{l-1} \\ 0 & & h_{l-1} & 2(h_{l-1}+h_l) \end{bmatrix} \tag{2.17}$$

und

$$T = - \begin{bmatrix} h_1^{-1}+h_2^{-1} & -h_2^{-1} & & 0 \\ -h_2^{-1} & h_2^{-1}+h_3^{-1} & \ddots & \\ & \ddots & \ddots & -h_{l-1}^{-1} \\ 0 & & -h_{l-1}^{-1} & h_{l-1}^{-1}+h_l^{-1} \end{bmatrix}. \tag{2.18}$$

Ebenso gehen die Einträge von T bei der Berechnung der Sprünge $[s''']_{x_i}$ an den Gitterknoten mit ein. Man berechnet in diesem Fall für $i = 1, ..., l$

$$[s''']_{x_i} := s'''(x_i+) - s'''(x_i-) = \frac{\gamma_{i+1} - \gamma_i}{h_{i+1}} - \frac{\gamma_i - \gamma_{i-1}}{h_i}. \tag{2.19}$$

Mit Hilfe von (2.19) schreibt sich (2.15) in vektorieller Schreibweise als

$$Tc = \lambda(y - s). \tag{2.20}$$

Da T regulär ist, kann man nach Multiplikation mit T von links und mit Hilfe von (2.16) die Gleichung umformen zu

$$\begin{aligned} & T^2 c = \lambda(Ty - Ts) = \lambda Ty - \lambda Gc \\ \Leftrightarrow \quad & (G + \alpha T^2)c = Ty \qquad \text{mit} \quad \alpha = \frac{1}{\lambda}. \end{aligned} \tag{2.21}$$

Die Bedingung (2.15) wird erfüllt, denn G und T^2 sind beide symmetrisch und positiv definit, weswegen $G + \alpha T^2$ für jedes $\alpha > 0$ invertierbar ist und ein natürlicher kubischer Spline mit homogenen Randwerten existiert. Abschließend muss nur noch nachgewiesen werden, dass der Parameter $\lambda > 0$ so gewählt werden kann, dass s die Nebenbedingung

$$\sum_{i=1}^{l-1} |y_i - s(x_i)|^2 = (l-1)\delta^2$$

erfüllt. Einsetzen von (2.20) in die Nebenbedingung (2.14) liefert die äquivalente vektorielle Notation

$$(l-1)\delta^2 \overset{!}{=} \frac{1}{\lambda^2} \|Tc\|_2^2 .$$

Weiterhin lässt sich mit obigen Überlegungen der Momentenvektor explizit darstellen:

$$c = (G + \frac{1}{\lambda}T^2)^{-1}Ty.$$

Damit gilt:

$$\begin{aligned}(l-1)\delta^2 &\stackrel{!}{=} \frac{1}{\lambda^2}\|Tc\|_2^2 \\ &= \left\|\frac{1}{\lambda}T(G+\frac{1}{\lambda}T^2)^{-1}Ty\right\|_2^2 \\ &= \left\|\lambda^{-1}(T^{-1})^{-1}(G+\frac{1}{\lambda}T^2)^{-1}(T^{-1})^{-1}y\right\|_2^2 \qquad (2.22) \\ &= \left\|(I+\lambda T^{-1}GT^{-1})^{-1}y\right\|_2^2 := \quad r(\lambda) \qquad (2.23)\end{aligned}$$

Da schon in 2.5 gezeigt wurde, dass $r(\lambda)$ eine Darstellung der Form $\sum_{i=1}^{n} \frac{z_i^2}{(d_i+\lambda)^2}$ besitzt und der Parameter λ genau dann die eindeutige positive Lösung von $r(\lambda) = (l-1)\delta^2$ ist, wenn

$$r(0) = \|y\|_2^2 > (l-1)\delta^2 \qquad (2.24)$$

erfüllt ist, hat man somit den Beweis erbracht, denn mit (2.24) wurde gerade die Voraussetzung aus Satz 2.11 erfüllt. □

Bemerkung 2.12. *Es ist offensichtlich, dass* (2.21) *genau einer* ***Tikhonov-Philipps-Regularisierung*** *[10] der Gleichung $Gc = Ty$ mit Regularisierungsterm αT^2 entspricht. Im vorliegenden Fall liefert das Hebdenverfahren eine explizite Wahl des Regularisierungsparameters $\alpha = \frac{1}{\lambda}$.*

2.5 Rechteckregel vs. Trapezregel

Mit Hilfe der Rechteckregel kann man das Integral einer Funktion $f \in \mathcal{C}([a,b],\mathbb{R})$ approximieren.

Definition 2.13 (Rechteckregel)**.** *Mit Hilfe der Rechteckregel wird eine Funktion $f \in \mathcal{C}([a,b],\mathbb{R})$ folgendermaßen integriert:*

$$\int_a^b f(x)dx = (b-a)\cdot f(a) + R_R \qquad (2.25)$$

mit R_R als Fehler.

Eine ebenfalls häufig verwendete Approximation des Integrals lässt sich mit der sogenannten Trapezregel realisieren:

Definition 2.14 (Trapezregel)**.** *Mit Hilfe der Form eines Trapezes wird eine Funktion $f \in \mathcal{C}([a,b],\mathbb{R})$ folgendermaßen integriert:*

$$\int_a^b f(x)dx = \frac{b-a}{2}(f(a)+f(b)) + R_T \tag{2.26}$$

mit R_T als Fehler.

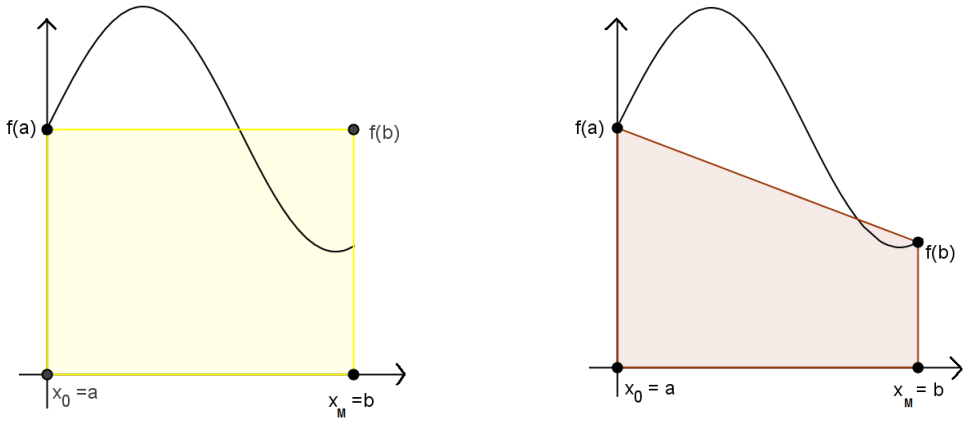


Abbildung 2.1: Eine Funktion $f(x)$ links mit der Rechteckregel integriert, rechts mit der Trapezregel.

Um eine genauere Abschätzung des Integrals über die Funktion f im Intervall $[a,b]$ zu erhalten, betrachtet man nun eine Zerlegung des Integrationsgebiets in M-Subintervalle $I_i = [x_i, x_{i+1}], a = x_0 < x_1 < ... < x_M = b$. Das gesamte Integral wird über die Summe jedes einzelnen Integrals über I_i wie folgt berechnet:

Zusammengesetzte Rechteckregel:

$$\int_a^b f(x)dx = h[f(x_0)+f(x_1)+...+f(x_{M-1})] + R_R, \tag{2.27}$$
$$h = \frac{b-a}{M}, \qquad x_i = a+ih, \quad i = 0,1,...,M-1.$$

Zusammengesetzte Trapezregel:

$$\int_a^b f(x)dx = \frac{h}{2}[f(x_0) + 2f(x_1) + ... + 2f(x_{M-1}) + f(x_M)] + R_T, \qquad (2.28)$$

$$h = \frac{b-a}{M}, \qquad x_i = a + ih, \quad i = 0, 1, ..., M.$$

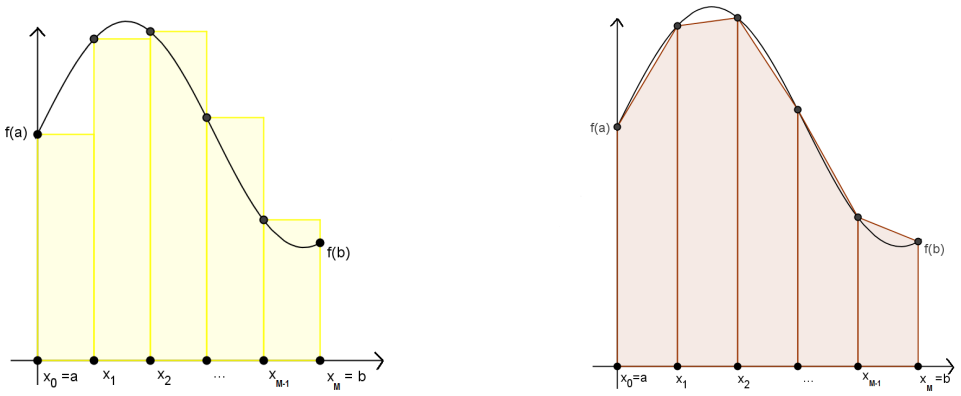


Abbildung 2.2: Die gleiche Funktion $f(x)$ aus Abbildung 2.1 auf den Intervallen I_i links mit der zusammengesetzten Rechteckregel, rechts mit der zusammengesetzten Trapezregel integriert.

In beiden Fällen wird in dieser Arbeit von einer äquidistanten Zerlegung in $M-1$ Subintervalle I_i ausgegangen, wobei die Schrittweite jeweils h beträgt. Die Fehler R_R und R_T sollen im weiteren Verlauf nicht näher beachtet werden; es ist jedoch aus Abbildung 2.1 ersichtlich, dass die Rechteckregel unter bestimmten Voraussetzungen für spezielle Schrittweite und Funktionen nicht automatisch schlechter approximiert als die Trapezregel.

Bemerkung 2.15. *Im weiteren Verlauf dieser Arbeit werden ausschließlich die Begriffe Trapez- und Rechteckregel verwendet; gemeint ist jedoch immer die zusammengesetzte Version der Quadraturformeln.*

Kapitel 3

Parameterbestimmung bei Systemen autonomer Differentialgleichungen

Zur Parameterschätzung wird in diesem Fall der Weg über die numerische Integration von Systemen gewöhnlicher Differentialgleichungen gewählt. Dies ist möglich, da der gesuchte Parametervektor linear eingeht. Der Ausgangspunkt der sich anschließenden Kapitel 4 und 5 basiert auf den Arbeiten von G. Wikström, die sich mit dem Thema der Parameterschätzung in ihrer Dissertation [14] befasst hat. Im Folgenden werden diese Ergebnisse kurz erläutert.

3.1 Grundlagen zu gewöhnlichen Differentialgleichungen

Bevor nun im Speziellen auf das Thema der Parameterschätzung eingegangen wird, noch ein paar allgemeine Erklärungen zu Systemen gewöhnlicher Differentialgleichungen vorweg.

Definition 3.1. *Die allgemeine Form eines* ***Anfangswertproblems*** *(AWP) 1. Ordnung ist definiert als*

$$\dot{y} = f(t, y(t)), \qquad y(t_0) = y_0 \tag{3.1}$$

mit $f \in \mathcal{C}(\mathbb{R} \times \mathbb{R}^m, \mathbb{R}^m)$ *und* $y_o \in \mathbb{R}^m$.

Wie unter anderem in [2] nachgelesen werden kann, existiert nach dem Satz von Picard-Lindelöf genau eine eindeutige Lösung $y \in \mathcal{C}^1(I, \mathbb{R}^n)$ welche das AWP löst. Dabei bezeichne $I \subset \mathbb{R}$ ein Intervall, welches t_0 enthält. Die Voraussetzungen für die Anwendung dieses Satzes sind erfüllt, denn die rechte

Seite (3.1) ist stetig als Komposition stetiger Funktionen auf jedem Intervall I.

Im weiteren Verlauf werden ausschließlich *autonome* Differentialgleichungssysteme betrachtet, welche durch folgende Definition charakterisiert sind:

Definition 3.2. *Ein System*

$$\dot{y} = f(t, y)$$

heißt ***autonom****, wenn die rechte Seite nicht von der unabhängigen Veränderlichen t abhängt, d.h.*

$$f(t, y) = f(y).$$

Im Fokus dieser Arbeit stehen nichtlineare Systeme, deren Parametervektor $k \in \mathbb{R}^n$ linear eingeht. Bei dieser Art von Problemen lässt sich f faktorisieren in die Form

$$f(t, y) = G(t, y) \cdot k \tag{3.2}$$

mit $G(t, y) \in \mathbb{R}^{m \times n}$. Wenn die DGL

$$\dot{y} = G(t, y) \cdot k \tag{3.3}$$

integriert wird, lässt sich der Parameter offensichtlich aus dem Integranten herausziehen. Dies bildet die Grundlage der vorgestellten Methoden zur Parameterschätzung.
Trotz dieser vereinfachten Darstellung sei an dieser Stelle darauf hingewiesen, dass das System nichtlinear in y ist. Für Probleme dieser Art existieren in der Regel keine geschlossenen Lösungsdarstellungen. Lediglich in wenigen Spezialfällen kann man eine solche Lösung explizit darstellen [1].

3.2 Parameterschätzung bei Systemen gewöhnlicher Differentialgleichungen

Häufig werden dynamische Prozesse, wie sie in der Natur auftreten, durch gewöhnliche Differentialgleichungen (DGL) dargestellt [4], [7], [11]. Dabei ist es zunächst unerheblich, ob die beeinflussenden Parameter bekannt sind oder nicht: In Experimenten kann man Werte der abhängigen Variablen für gegebene Werte der unabhängigen Variable messen und durch Nutzung dieser Messwerte im Anschluss den Parameter schätzen.

Der übliche Weg eine Parameterschätzung bei gewöhnlichen DGL ist die numerische Integration der Gleichung, um die Differenzen zwischen den numerisch gewonnenen und den gemessenen Daten zu minimieren. Ein gängiger Ansatz hierzu stellt die *Methode der kleinsten Fehlerquadrate* dar, welche auch die Grundlage dieser Arbeit bildet.

Das allgemeine mathematische Modell, welches den Ausgangspunkt für alle weiteren Berechnungen bildet, lautet wie folgt:

$$\dot{y}(t, v) = f(t, y(t, v), k) \qquad y(t_0) = y_0 \tag{3.4}$$

mit $v^T = (y_0^T, k^T)$ als erweiterter Parametervektor und $y = y(t, v)$ als Lösung zum Zeitpunkt t.

Bemerkung 3.3. *In dieser Arbeit wird der Startwert y_0 als bekannt vorausgesetzt, auch wenn sich diese Bedingung in der Realität nicht immer erfüllen lässt. Sollte dieser Fall eintreten, kann man den Wert anhand der übrigen Messdaten „abschätzen“ oder ihn durch eine Variation der unten beschriebenen Verfahren bestimmen. Für eine ausführliche Beschreibung und algorithmische Umsetzung wird auf [14] verwiesen.*

Das zu lösende Minimierungsproblem lässt sich formulieren als

$$\min_{k, y(t,k)} = \frac{1}{2} \sum_{i=1}^{M} \|\tilde{y}_i - y(t_i, k)\|_2^2 \tag{3.5}$$

mit Messwerten $\tilde{y}_i$ zum Zeitpunkt t_i, $i = 1, ..., M$ und der von k abhängigen Funktion y. Dieser Ausdruck lautet in diskreter Notation

$$\min_{k,y} = \frac{1}{2} \sum_{i=1}^{M} \|\tilde{y}_i - y_i\|_2^2 \tag{3.6}$$

wobei $y_i \approx y(t_i)$ die numerisch bestimmte Lösung des AWP 's (3.4) bezeichnet. Wenn nun die numerische Lösung des gewöhnlichen Differentialgleichungsystems in den Termen für k berechnet ist, gelangt man zu

$$\min_{k} = \frac{1}{2} \sum_{i=1}^{M} \|\tilde{y}_i - y_i(k)\|_2^2 \tag{3.7}$$

mit $y_i(k)$ als numerischer Lösung von 3.4 bezüglich des Parameters k. Das Minimierungsproblem lässt sich somit auf die Bestimmung des Parameters k reduzieren.

3.3 Algorithmus von Wikström

Wie in (3.3) hergeleitet, bildet folgender spezieller Typ von Anfangswertproblemen den Ausgangspunkt der Arbeit von Wikström:

$$\dot{y}(t,k) = G(t,y(t))k, \tag{3.8}$$

wobei der Parametervektor k linear auf der rechten Seite eingeht. Eine *globale* Integration von Gleichung (3.8) über dem Intervall $I_i = [t_0, t_i]$ liefert

$$y(t_i) = y_0 + k \int_{t_0}^{t_i} G(\tau, y(\tau))d\tau, \tag{3.9}$$

Alternativ kann auch *lokal* über das Intervall $I_i = [t_{i-1}, t_i]$ integriert werden:

$$y(t_i) - y(t_{i-1}) = k \int_{t_{i-1}}^{t_i} G(\tau, y(\tau))d\tau.$$

Ist statt dem Anfangswert y_0 nur die Größe y_j zum Zeitpunkt t_i bekannt, so kann etwa mit folgender Integraldarstellung gearbeitet werden:

$$y(t_i) - y(t_j) = k \int_{t_j}^{t_i} G(\tau, y(\tau))d\tau.$$

Da diese Integrale im Regelfall nicht exakt bestimmt werden können, bedient man sich einer numerischen Integration wie beispielsweise der zusammengesetzten Rechteck- bzw. Trapezregel aus Kapitel 2.5. Als einzusetzende Funktionswerte dienen die zuvor erhobenen und als exakt angenommenen Messwerte $\tilde{y}_i$.

Bemerkung 3.4. *Auf die Analyse von Rundungsfehlern, die bei einer numerischen Approximation der Integrale entstehen, wird hier verzichtet und auf [14] sowie auf die Lehrwerke [2] bzw. [6] verwiesen. Ebenso wird im weiteren Verlauf ausschließlich die globale Integration betrachtet. Die Methoden und Algorithmen übertragen sich analog auf den Fall der lokalen Integration.*

3.4 Lösung des Minimierungsproblems

Zunächst konstruiert man sich Matrizen $J_i \in \mathbb{R}^{m\times n}$ mit $i = 1, ..., M$, welche die jeweiligen Integrationswerte speichern, die man durch die Verwendung der

entsprechenden Quadraturformeln (Rechteck-/ Trapezregel) erhält:

$$J_i = \begin{cases} h(G_0 + ... + G_{i-1}) & \text{(Rechteckregel)}, \\ \frac{h}{2}(G_0 + 2G_1 + ... + 2G_{i-1} + G_i) & \text{(Trapezregel)}, \end{cases} \tag{3.10}$$

mit $G_i = G(t_i, \tilde{y_i}) \in \mathbb{R}^{m \times n}$ und Schrittweite $h = \frac{b-a}{M}$. Gleichung (3.9) kann explizit nach y_i aufgelöst werden, wobei je nach Verwendung der Rechteck- oder Trapezregel die entsprechende Matrix J_i auftritt:

$$y_i = y_0 + J_i k, \qquad i = 1, ..., M. \tag{3.11}$$

Die Messwerte $\tilde{y_i} = y(t_i, k)$ dienen also zur numerischen Approximation der Integralgleichung in (3.9). Mit der exakten Lösung stimmen diese in der Regel nicht exakt überein, da die Fehler (sowohl Mess- als auch Rundungsfehler) in der Praxis unbekannt sind. Wie bereits oben erwähnt werden diese im Algorithmus vernachlässigt, sodass die numerisch bestimmten y_i als exakte Werte behandelt werden.

Um nun das gesamte Minimierungsproblem formulieren zu können, werden die einzelnen Matrizen J_i in eine „erweiterte“ Matrix $H \in \mathbb{R}^{Mm \times n}$ geschrieben:

$$H_p = \begin{pmatrix} J_1 \\ J_2 \\ \vdots \\ J_M \end{pmatrix} \qquad p = \begin{cases} 1, & \text{für die Rechteckregel}, \\ 2, & \text{für die Trapezregel} \end{cases} \tag{3.12}$$

Das nun zu lösende *„Kleinste-Fehlerquadrat-Problem“* lautet daher:

$$\min_k \frac{1}{2} \left\| H_p k - By + y_0 \right\|_2^2 = \min_k \frac{1}{2} \left\| H_p k - b \right\|_2^2 \tag{3.13}$$

mit $b = By - y_0$. Die Matrix $B \in \mathbb{R}^{Mm \times Mm}$ enthält auf der Hauptdiagonalen die Einheitsmatrizen $I_m \in \mathbb{R}^{m \times m}$; y_j, $y_0 \in \mathbb{R}^{Mm \times 1}$ repräsentieren die Vektoren mit den Messdaten bzw. Startwerten:

$$B = \begin{pmatrix} I_m & & \\ & \ddots & \\ & & I_m \end{pmatrix}, \qquad y_j = \begin{pmatrix} y_1 \\ \vdots \\ y_M \end{pmatrix}, \qquad y_0 = \begin{pmatrix} y_0 \\ \vdots \\ y_0 \end{pmatrix}.$$

Offensichtlich ist das Minimierungsproblem auf ein Problem der Form

$$\min_k \frac{1}{2} \left\| Ax - b \right\|_2^2,$$

zurückzuführen, welches in Kapitel 2.1 thematisiert wurde. Mit Bezug auf (2.5) ist die Lösung gegeben durch:

$$k = H^{+}b, \tag{3.14}$$

wobei $H^{+} = (H^{T}H)^{-1}H^{T}$ die Pseudoinverse von H bezeichnet.

Bemerkung 3.5. *Um reale Daten zu simulieren werden künstliche Messfehler eingeführt. Dazu addiert man zu den exakten Daten sogenannte Pseudozufallszahlen, die in den meisten Programmiersprachen bereitgestellt werden. Auf diese Weise entstehen künstlich verrauschte Daten, deren Fehlerniveau und Art (zum Beispiel: Normalverteilung, Gleichverteilung) vom Anwender selbst bestimmt werden kann.*

3.5 Pseudocode

Zur Parameterschätzung:
Gegeben:

1. Messdaten $y_i \in \mathbb{R}^m$, $i = 1, ..., M$,
2. Zeitpunkte t_i, $i = 1, ..., M$,
3. Abbildung $G : \mathbb{R} \times \mathbb{R}^m \longrightarrow \mathbb{R}^{m \times n}$.

Gesucht: $k \in \mathbb{R}^n$

→ Initialisiere J_i nach (3.10), $i = 1, ..., M$.
→ Übergebe J_i in H_p
→ Initialisiere b nach 3.4,$i = 1, ..., M$.
→ Löse nun die Normalgleichung

$$H^{T}Hk = H^{T}b \tag{3.15}$$

Kapitel 4

Ein iteratives Verfahren zur Parameterschätzung

Der Algorithmus von Wikström basiert im Wesentlichen auf der Approximation der Integrale

$$\int_{t_{i-1}}^{t_i} G(\tau, y(\tau))d\tau \tag{4.1}$$

durch die Rechteck- bzw. Trapezregel, d.h. für $h = \frac{b-a}{M}$ berechnet sich (4.1) als

$$\begin{cases} \int_{t_{i-1}}^{t_i} G(\tau, y(\tau))d\tau = h \cdot G(t_i, \tilde{y}_i) & \text{Rechteckregel,} \\ \int_{t_{i-1}}^{t_i} G(\tau, y(\tau))d\tau = \frac{h}{2} \cdot G(t_i, \tilde{y}_i) & \text{Trapezregel.} \end{cases}$$

Hier wird vorausgesetzt, dass der Messwert $\tilde{y}_i$ rauschfrei ist und zu dem exakten Parameter k^* gehört. Formal liefert dies die Notation:

$$\tilde{y}_i = y(t_i, k^*).$$

In diesem Fall bezeichnet y die exakte Lösung des AWP's mit korrektem Parametervektor k^*.
Allerdings können in der Praxis die Knotenpunkte t_{i-1} und t_i weit auseinander liegen, was lediglich eine grobe Näherung des Integrals (4.1) möglich macht. Um die Qualität der Approximation zu erhöhen, löst man jetzt mit Hilfe des bereits geschätzten Parameters $k = k^{(0)}$ auf jedem Intervall $I_i = [t_{i-1}, t_i]$ das AWP

$$\dot{y} = G(t, y, k^{(0)}) \quad \text{mit} \quad y(t_i) = y_i \tag{4.2}$$

für $t \in [t_{i-1}, t_i]$. Somit lässt sich mit den neu gewonnenen Funktionswerten zum Parametervektor $k = k^{(0)}$ das Integral (4.1) genauer bestimmen.

4.1 Der Algorithmus im Speziellen

Es wird davon ausgegangen, dass der Parameter $k = k^{(0)}$ bereits mit der in Abschnitt 3.4 thematisierten Methode bestimmt ist. Zu diesem Parameter werden nun N weitere „Messdaten“ $u_j^{(i)}$, $j = 1, ..., N+1$ auf jedem Intervall I_i gewonnen. Dazu wird ein explizites Einschrittverfahren gewählt, in diesem Fall das Runge-Kutta-Verfahren vierter Ordnung [7] bzw. [**?**]. Als Startwerte dienen jeweils die bekannten Messdaten $\tilde{y}_i \in I_i$ und als Parameter wird $k^{(0)}$ eingesetzt (vgl. (4.2)). Die hierfür benötigte Schrittweite beträgt $h_{klein} = \frac{h}{N}$. Insgesamt ergeben sich $(M-1) \cdot (N+1)$ zusätzliche Werte zu den bereits bekannten Messdaten hinzu. Hat es also zuvor genügt

$$\begin{cases} \int_{t_{i-1}}^{t_i} G(\tau, y(\tau)) d\tau k = h \cdot \tilde{y}_i \cdot k & \text{Rechteckregel} \\ \int_{t_{i-1}}^{t_i} G(\tau, y(\tau)) d\tau k = \frac{h}{2} \cdot (\tilde{y}_{i-1} + \tilde{y}_i) \cdot k & \text{Trapezregel} \end{cases} \tag{4.3}$$

zu berechnen, geht man jetzt dazu über die Werte $u_j^{(i)}$ als Argumente der Funktion G einzusetzen und in eine summierte Rechteck- bzw. Trapezregel einzubauen. Mittels der verwendeten Quadraturformeln werden nun erneut die Matrizen J_i generiert (vgl. (3.10)). Diese werden wiederum in der Matrix J gemäß (3.12) konkateniert. Wird also

$$\int_{t_{i-1}}^{t_i} G(\tau, y(\tau)) d\tau k = h_{klein} \sum_{j=1}^{N+1} G(t_i + h_{klein}(j-1))k \tag{4.4}$$

auf die summierte Rechteckregel übertragen, ergibt dies:

$$y_i - y_0 = J_i = \int_{t_0}^{t_i} G(\tau, y(\tau)) d\tau k = h_{klein} \cdot \sum_{l=1}^{i} \sum_{j=1}^{N+1} G(u_j^{(l)}) \cdot k. \tag{4.5}$$

Dieser Ausdruck überträgt sich analog auf die Trapezregel (vgl. (3.10)). Die Lösung ist gegeben durch:

$$k^{(1)} = H^+ b.$$

In der Grafik 4.1 sind im oberen Bild die beiden Komponenten der Barnesfunktion (Erläuterungen und Eigenschaften sind in Kapitel 6.2 gegeben) mit exaktem Parameter $k^* = [2, 5, 7]^t$ abgebildet. Aus ihnen werden jeweils 10 äquidistante *Messwerte* genommen, welche als (**x**) markiert sind. Mit diesen Vorgaben und dem Parameter $k^{(0)}$ wird nun im unteren Bild versucht, mit

jeweils $N = 40$ zusätzlichen Messwerten pro Intervall I_i die Barnesfunktion zu rekonstruieren. Die hier sichtbaren Abstände der Funktionswerte an den Intervallgrenzen $y(t_{N+1}) \in I_i$ und $y(t_1) \in I_{i+1}$ sollen durch wiederholte Iteration reduziert werden.

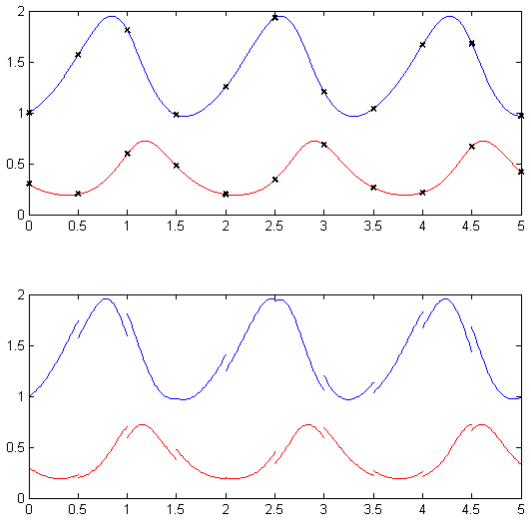


Abbildung 4.1: Ein Beispiel zur iterativen Methode anhand der Barnesfunktion.

Insgesamt besteht das Ziel also darin, die Fehler durch die numerische Integration in den Matrizen J_i zu reduzieren und sich so eine genauere Abschätzung des Parameters $k^{(0)}$ bzw. k^* zu erhoffen. Nach Bestimmung des neuen Parameters $k^{(1)}$ ist es möglich, die gesamte Vorgehensweise zu wiederholen und mit einer Folge von Parametern $k = k^{(0)}$, $k^{(1)}$, $k^{(2)}$,... ein *iteratives Verfahren* zu konstruieren.

Die Fragen, die an dieser Stelle auftauchen, lauten wie folgt:

1. Liefert die Folge $k = k^{(0)}$, $k^{(1)}$, $k^{(2)}, \ldots$ eine verbesserte Schätzung des Parametervektors?

2. Wie wirkt es sich aus, wenn die Anzahl M der Messdaten relativ klein ist, die Anzahl N der „künstlichen“ Zwischenstellen hingegen sehr groß?

3. Wie wirkt es sich aus, wenn eine Streuung statt mit einer Gleichverteilung mit einer Normalverteilung verursacht wird?

4.2 Pseudocode

→ Pseudocode wie in Abschnitt 3.5, um $k^{(0)}$ zu bestimmen

Gegeben

1. Anzahl N der zusätzlichen „künstlichen“ Messwerte pro Intervall I_i.
2. Parameter $k^{(0)}$ aus 3.5

Gesucht: $k^{(1)} \in \mathbb{R}^n$

→ Berechne den Vektor als Lösung des AWP's (4.2) $u_j^{(i)}$, $j = 1, ..., N+1$.
→ Initialisiere J_i nach (4.5), $i = 1, ..., M$ basierend auf den Werten $u_j^{(i)}$ zur Schrittweite h_{klein}.
→ Übergebe J_i in H_p
→ Initialisiere b nach 3.5
→ Löse die Normalgleichung

$$H^T H k^{(1)} = H^T b. \tag{4.6}$$

Kapitel 5

Parameterschätzung mit Hilfe glättender Splines

In diesem Kapitel wird ein weiterer Algorithmus zur Parameterbestimmung vorgestellt, welcher speziell auf stark verrauschte Daten abzielt. Die Idee ist, mit Hilfe von glättenden, kubischen Splines Regressionskurven zu berechnen, die auf den einzelnen Intervallen exakt integriert werden können. Die grundlegende Theorie und Eigenschaften glättender Splines wurde bereits in Kapitel 2.4 näher erläutert und wird somit als bekannt vorausgesetzt.

Zur Erinnerung:
Ein glättender Spline s ist wie der interpolierende über einem Gitter $\triangle$ definiert, jedoch fordert man statt Gleichheit in den Knoten eine Fehlergleichung der Art:

$$\sum_{i=1}^{M-1} |y_i - s(x_i)|^2 = (l-1)\delta^2.$$

Dabei wird in den Daten ein gleichverteiltes Rauschen über den Intervallen $[-\delta, \delta]$ angenommen.

Zur Existenz- und Eindeutigkeit der glättenden Splines ist folgende Ungleichung der Messwerte notwendig:

$$\frac{1}{M-1} \sum_{i=1}^{M-1} |y_i|^2 > \delta^2.$$

Interessant für die hier durchgeführte Anwendung ist die Tatsache, dass sich Splines als stückweise kubische Polynome stets exakt integieren lassen: Der

Fehler der Integration liegt demnach bei Null und die Zuhilfenahme der Trapez- und Rechteckregel ist nicht mehr von Nöten.

5.1 Der Algorithmus im Speziellen

Zunächst werden die Messwerte $y_0, ..., y_M$ mittels einer affin-linearen Verschiebung auf homogene Randdaten transformiert. Dazu wird eine lineare Funktion

$$l(t) = y_0 + \frac{y_m - y_0}{t_M - t_0}(t - t_0)$$

von den Messdaten subtrahiert. Mit diesen Werten wird anschließend bezüglich des Gitters $\triangle = \{t_0, ..., t_M\}$ gemäß dem Algorithmus aus Kapitel 2.4 der glättende Spline berechnet.
Nach Rücktransformation auf die usprünglichen Randdaten können die Regressionspolynome zur Approximation der Funktion G verwendet werden. Die rechte Seite setzt sich nun aus stückweise kubischen Polynomen zusammen und die exakt approximierte rechte Seite G kann exakt integriert werden. Die entsprechenden Integrale dienen damit der Initialisierung der Matrizen J_i

$$J_i = h \cdot \sum_{j=1}^{i-1} G_j = h \cdot \sum_{j=1}^{i-1} S_j \tag{5.1}$$

$$\text{mit} \qquad S_j = \int_{t_0}^{t_i} s(t)dt. \tag{5.2}$$

Veranschaulicht ergibt sich die Grafik:

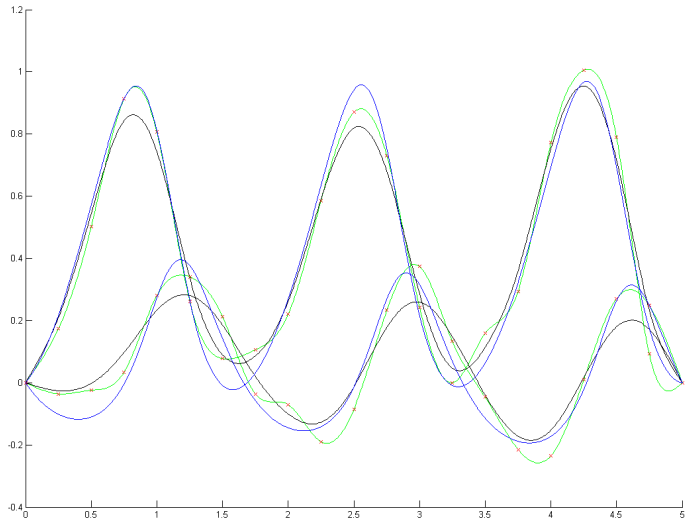


Abbildung 5.1: Splines im Vergleich: Zu sehen ist ein natürlicher, kubischer Spline (grün), die Originalfunktion (blau), die 40 Messpunkte (braun) und ein glättender Spline (schwarz) nach Transformation auf homogene Randdaten.

Die einzelnen Matrizen J_i werden wie üblich in die Matrix $H \in \mathbb{R}^{Mm \times n}$ überführt. Das bereits bekannte Minimierungsproblem $\|Hk - (By - y_0)\|_2^2$ löst sich wieder mit

$$k = H^+ \cdot (By - y_0). \tag{5.3}$$

Bemerkung 5.1. *Da in der Realität die Messdaten niemals mit maximalem Fehler eingehen, wird die Voraussetzung der Gleichheit der Nebenbedingung* (2.14) *als zu einschränkend angesehen. Der maximale Fehler sollte mit einem Skalierungsfaktor $\beta \in (0, 1]$ gewichtet werden, um eine Überglättung zu vermeiden. Die "neue" Nebenbedingung lautet von daher:*

$$\sum_{i=1}^{l-1} |y_i - s(x_i)|^2 = \beta \cdot (l-1)\delta^2. \tag{5.4}$$

Numerische Experimente haben gezeigt, dass zumindest in diesem verwendeten Beispiel die Wahl für ein $\beta \in [0.2, 0.4]$ optimale Ergebnisse liefert. Dies verifiziert anhand konkreter Beispiele die Einführung des Skalierungsfaktors zur Definition der glättenden Splines, falls von einer Gleichverteilung des Rauschens ausgegangen wird.

Die numerisch am besten approximierten Ergebnisse eines Zerfallsprozesses (Erläuterung folgt in Kapitel 6.1) mit $k = 0.25$ und Transformation auf homogene Randdaten im Vergleich:

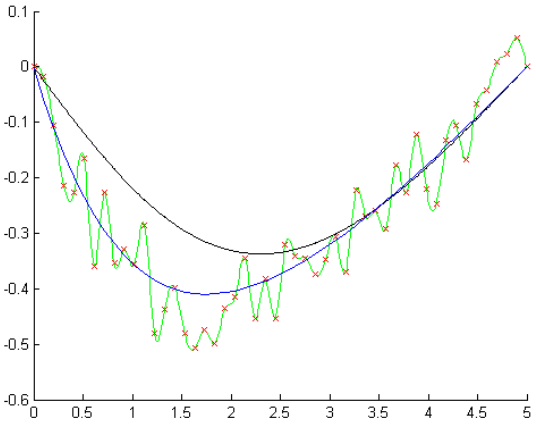


Abbildung 5.2: Skalierungsfaktor $\beta = 1$

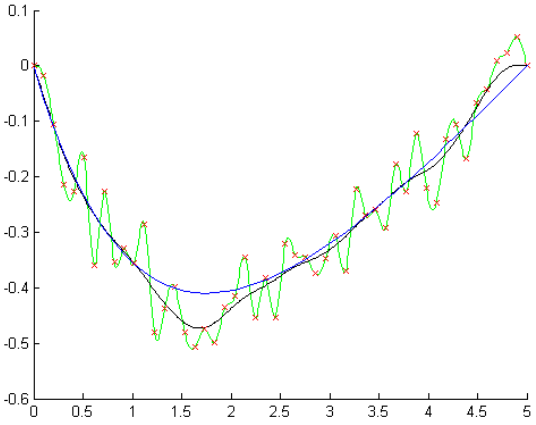


Abbildung 5.3: Skalierungsfaktor $\beta = 1/5$

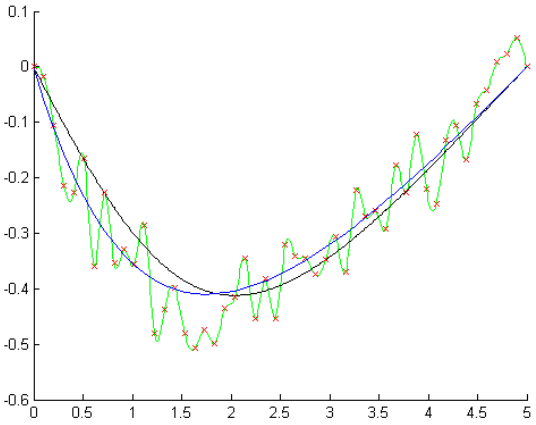


Abbildung 5.4: Skalierungsfaktor $\beta = 2/5$

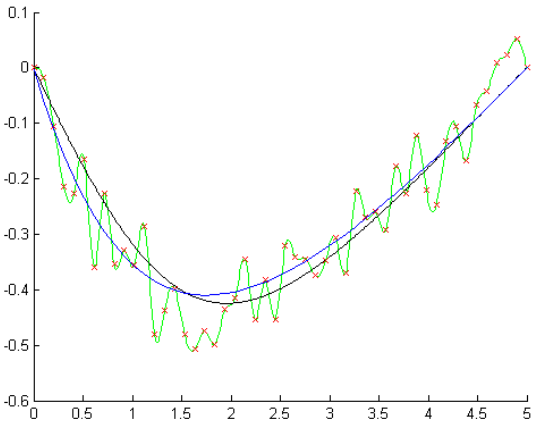


Abbildung 5.5: Skalierungsfaktor $\beta = 1/3$

Die weiteren Berechnungen in den folgenden Kapiteln werden sowohl mit dem Faktor $\beta = \frac{1}{3}$ als auch mit $\beta = \frac{1}{5}$ durchgeführt.

5.2 Pseudocode

Gegeben:

1. Anzahl M der Messewerte $\tilde{y_i}$.
2. exakter Parameter k
3. Funktion y

Gesucht: $k^{(0)} \in \mathbb{R}^n$

$\rightarrow$ Transformation von $\tilde{y_i}$, y auf homogene Randdaten.
$\rightarrow$ Initialisiere mit $h = \frac{b-a}{M}$ als Schrittweite die Matrizen G, T aus (2.17), (2.18).
$\rightarrow$ Durchführung des Hebdenverfahrens zur Bestimmung von λ aus der Gleichung $r(\lambda)$; Skalierungsfaktor beachten!
$\rightarrow$ Initialisiere den den glättenden Spline $s(x)$.
$\rightarrow$ Rücktransformation auf ursprüngliche Randdaten.
$\rightarrow$ Exakte Integration des Splines auf jeden Intervall I_i und Abspeicherung nach 5.1 in die Matrizen J_i.
$\rightarrow$ Initialisiere $b = Bs_i - y_0$ mit den rücktransformierten Werten.
$\rightarrow$ Löse nun die Normalgleichung.

$$H^T H k = H^T b \tag{5.5}$$

Kapitel 6

Klassische Beispiele aus den Naturwissenschaften

In diesem Abschnitt wird die zuvor erarbeitete Theorie auf die Praxis angewendet. Dazu werden zum einen die Modellgleichungen (radioaktiver Zerfall, Barnes-Problem, Robertson-Problem) vorgestellt, zum anderen werden die Parameter und ihre (Aus-)Wirkungen auf die Modelle näher erläutert.

6.1 Zerfallsprozesse

Zerfallsprozesse treten sowohl in der Biologie als auch in der Chemie und Physik auf. Während es sich im erstgenannten Fachgebiet häufig darum dreht, die Größe bestimmter Zellkulturen zum Zeitpunkt t zu bestimmen [11], [7] (je nach Eigenschaften des Vorzeichens von α in (6.1) werden in diesem Fall vorwiegend Wachstumsprozesse beschrieben), betrachtet man in der Chemie/Physik hauptsächlich den Vorgang eines sogenannten *radioaktiven Zerfalls* [7], mit dem sich auch dieser Abschnitt beschäftigt.

Definition 6.1 (Zerfalls- bzw. Wachstumsgesetz)**.** *Die Anzahl der noch nicht umgewandelten Atome radioaktiver Substanzen zum Zeitpunkt t und einer gegebenen Anfangsmenge N_0 zum Zeitpunkt $t = 0$ lässt sich schreiben als*

$$N(t) = N_0 \cdot e^{-\alpha t}, \quad \alpha > 0. \tag{6.1}$$

Die physikalische Halbwertszeit (Zeitpunkt zu dem die Anfangsmenge sich halbiert hat) beträgt

$$t_{\frac{1}{2}} = \frac{\ln 2}{\alpha}. \tag{6.2}$$

Radioaktiver Zerfall bedeutet entgegen der eigentlichen Bedeutung des Begriffs „zerfallen“ nicht, dass sich eine Materie teilt oder Atome in sich zerfallen. Vielmehr steht der Terminus für die über das Zerfallsgesetz (6.1) definierte Umwandlung von Nukliden. Während dieses Prozesses wird eine ionisierende Strahlung freigesetzt, wobei hier je nach ihrer steigenden Fähigkeit eine feste Masse zu durchdringen in α-, β- und γ- Strahlung unterschieden wird. Diese Strahlung (lateinisch: radius = Strahl) gibt dem dynamischen Prozess ihren Namen.

In diesem Abschnitt wird insbesondere der Zerfall der radioaktiven chemischen Substanz *Cäsium* (^{137}Cs) in den (metastabilen) Stoff *Barium* (^{137}Ba) untersucht.
Das Alkalimetall ^{137}Cs ist strahlenbiologisch gesehen das bedeutendste unter den 35 bekannten Cäsium-Isotopen. Mit einer physikalischen Halbwertszeit von 30,2 Jahren zerfällt es sehr langsam. Dabei werden 6,5% direkt in ein stabiles ^{137}Ba umgewandelt, 93,5% indirekt über das metastabile Zwischenprodukt $^{137}Ba - m$, welches wiederum mit einer Halbwertszeit von 2,55 min in den stabilen Zustand übergeht.
Exemplarisch wird im Folgenden die Umwandlung eines metastabilen zu einem stabilen ^{137}Ba beschrieben. Entsprechende Daten dieses radioaktiven Zerfallsprozesses sind in [13]wiedergegeben.

In diesem Fall handelt es sich um eine gewöhnliche Differentialgleichung erster Ordnung der Form

$$\dot{y}(t,k) = k \cdot y \quad \text{mit} \quad y(0) = y_0, \quad k > 0. \tag{6.3}$$

Lemma 6.2. *Die Lösung dieses Anfangswertproblems lautet:*

$$y(t) = y_0 \cdot e^{-kt}.$$

Beweis:
Integration der Variablen liefert

$$\begin{aligned} -k \cdot y &= \frac{\mathrm{d}y}{\mathrm{d}t} \\ -k \cdot \mathrm{d}t &= \frac{1}{y} \cdot \mathrm{d}y \\ \int_0^t -k \cdot \mathrm{d}t &= \int_{y_0}^{y} \frac{1}{y} \cdot \mathrm{d}y \\ -kt &= \ln\left(\frac{y}{y_0}\right) \\ e^{-kt} &= \frac{y}{y_0}. \end{aligned}$$

$$\Rightarrow y(t) = y_0 \cdot e^{-kt} \tag{6.4}$$

□

Bemerkung 6.3. *Ersetzt man $y(t)$ durch $N(t)$, k durch α und $y_0 = N_0$ als Anfangswert, so erhält man genau die Zerfallsgleichung* (6.1)

6.1.1 Das Gauß-Newton Verfahren

Im Falle eines Zerfallsprozesses kann die explizite Lösung des Anfangswertproblems

$$\dot{y} = -ky, \qquad y(0)) = y_0$$

mit $t > t_0$ explizit angegeben werden. Es gilt:

$$y(t) = y(t,k) = y_0 \cdot e^{-kt}.$$

Die Bestimmung des unbekannten Parameters k führt auf ein *nichtlineares Regressionsploblem* zurück welches mit Hilfe des **Gauß-Newton Verfahrens** gelösenst werden kann [3].
Das Ziel des Verfahrens besteht darin, den Parameter k im Sinne der *kleinsten Fehlerquadratmethode* an die zur Verfügung stehenden Messdaten $\tilde{y}$ anzupassen:

Lemma 6.4. *Es gilt das zum Zerfallsprozess 6.1 gehörende Funktional*

$$y(t,k) := \|F(k)\|_2^2 \xrightarrow{k} min \tag{6.5}$$

über k zu minimieren, wobei

$$F(k) = \begin{bmatrix} y_0 \cdot e^{-kt_1} - \tilde{y_1} \\ y_0 \cdot e^{-kt_2} - \tilde{y_2} \\ \vdots \\ y_0 \cdot e^{-kt_M} - \tilde{y_M} \end{bmatrix} \in \mathbb{R}^M.$$

Die Jacobimatrix F' von F bezüglich k ist gegeben durch

$$F'(k) = \begin{bmatrix} -t_1 \cdot y_0 \cdot e^{-kt_1} \\ -t_2 \cdot y_0 \cdot e^{-kt_2} \\ \vdots \\ -t_M \cdot y_0 \cdot e^{-kt_M} \end{bmatrix} \in \mathbb{R}^M.$$

Die Lösung der Normalgleichung

$$(F')^T \cdot F' \cdot \Delta k \stackrel{!}{=} -(F')^T \cdot F$$

bildet das Gauß-Newton-Inkrement Δk. Die Iterationsvorschrift lautet

$$k^{i+1} = k^i + \Delta k^i \quad mit \quad \Delta k^i = -F'(k^i)^+ F(k^i),$$

wobei der obere Index $^+$ die Pseudoinverse kennzeichnet.

Der Beweis von Lemma 6.4 kann im allgemeinen Fall in [3] nachgelesen werden. In der aktuellen Anwendung interessiert man sich für die Bestimmung des Parameters k im Cäsium-Barium-Zerfall. Anhand dieses Beispiels gelingt die Gegenüberstellung der Resultate des Gauß-Newton-Verfahrens, der Methode von Wikström und der in dieser Arbeit vorgestellten Algorithmen. Der zum Gauß-Newton-Verfahren zugehörige Pseudocode lautet wie folgt:

Pseudocode

Gegeben:

1. Initialer Parameter $k^{(0)}$.
2. Maximale Gauß-Newton-Korrektur abs(Δk) als Abbruchkriterium.
3. Messwerte $\tilde{y}_i$ für $i = 1, ..., M$.
4. Knotenpunkte t_i, $i = 1, ..., M$.

Gesucht: $k \in \mathbb{R}$ als von $\|F(k)\|_2^2 \xrightarrow{k} \min$

while abs(Δk) $> eps$ und $i < i_{max}$

$\rightarrow i = i + 1$ (Schleifenzähler)
$\rightarrow$ Berechne $F'(k) \in \mathbb{R}^n$ (Jacobimatrix)
$\rightarrow$ Löse die Normalgleichung $(F')^T \cdot F' \cdot \Delta k \stackrel{!}{=} -(F')^T \cdot F$
$\rightarrow$ Setze $k = k + \Delta k$

6.2 Das Barnes Problem

Das sogenannte Barnes-Problem tritt ebenfalls in der Chemie auf. Viel häufiger jedoch ist diese modifizierte *Lotka-Volterra-Gleichung* (LV-Gleichung) als **Räuber-Beute-Modell** in der Biologie bekannt (siehe z.B. [12], [7]). Dabei geht man von zwei Populationsgrößen aus, die in einem interspezifischen Abhängigkeitsverhältnis stehen.
Die Population R ($y_1(t)$ = Anzahl der Räuber zum Zeitpunkt t) nutzen die Mitglieder der anderen Population B ($y_2(t)$ = Anzahl der Beutetiere zum Zeitpunkt t) als Nahrung, wodurch sich auf einen längeren Zeitraum gesehen

gewisse quantitative Populationsschwankungen entwickeln. Das Basismodell der zugrunde liegenden LV-Gleichung hat die Gestalt

$$\begin{aligned}\dot{y}_1(t) &= k_1 y_1(t) - k_2 y_1(t) y_2(t),\\ \dot{y}_2(t) &= k_4 y_1(t) y_2(t) - k_3 y_2(t).\end{aligned}$$

Während hier vier unbekannte Parameter auftreten, werden in dem Barnes-Problem k_2 und k_4 gleichgesetzt. Es bleiben also drei Parameter zu bestimmen:

$$\begin{aligned}\dot{y}_1(t) &= k_1 y_1(t) - k_2 y_1(t) y_2(t),\\ \dot{y}_2(t) &= k_2 y_1(t) y_2(t) - k_3 y_2(t).\end{aligned} \tag{6.6}$$

Da keine expliziten Lösungsverfahren exisiteren, ist man zur Berechnung der Lösung auf numerische Methoden angewiesen. In dieser Arbeit wurde dazu auf ein klassisches Runge-Kutta-Verfahren 4. Ordnung zurückgegriffen.

Kapitel 7

Numerische Experimente

Die in Kapitel 6 vorgestellten Probleme werden auf den nächsten Seiten mit den in den Kapitel 3.3 bis 5 beschriebenen Algorithmen numerisch ausgewertet. Die Werte in den Tabellen repräsentieren den dazugehörigen relativen Fehler

$$F_{rel} = \frac{\left\|k - k^{(i)}\right\|}{\|k\|}, \quad i = 0, 1, ..., D \tag{7.1}$$

mit D = Anzahl der Iterationen. Dabei wird unterschieden, ob es sich um eine gleichmäßige Erzeugung der auf die Funktion aufaddierten Zufallszahlen handelt oder ob sie normalverteilt sind.
Darüber hinaus werden auf jedes Verfahren zwei verschiedene Parameter k angewendet, um zu untersuchen, ob die Größe im Zusammenhang mit der Funktion eine Rolle spielt oder eher unerheblich ist. Im Zuge der glättenden Splines fließt zusätzlich noch der Streuparameter δ mit in die Analyse ein.

7.1 Einmalige Integration mittels der Rechteckregel/ Trapezregel

7.1.1 Zerfallsprozesse

Zunächst betrachtet man einen Zerfallsprozess mit synthetisch generierten Daten.

Synthetisch generierte Daten mit $k = 0.8$, $\delta = 0.1$

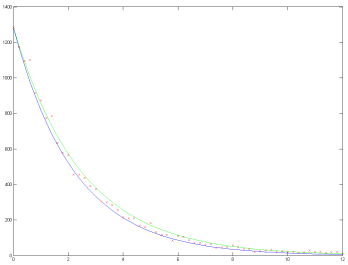
einmaliges Iterieren				
	gleichmäßige Verteilung		**Normalverteilung**	
M	Rechteckregel	Trapezregel	Rechteckregel	Trapezregel
10	0.0645	0.1122	0.9215	0.9654
20	0.0601	0.0364	0.8568	0.9433
50	0.0271	0.0089	0.3978	0.3897
100	0.0014	0.0063	0.5079	0.5098
250	0.0031	0.0018	0.4110	0.4129

Synthetisch generierte Daten mit $k = 0.25$, $\delta = 0.1$

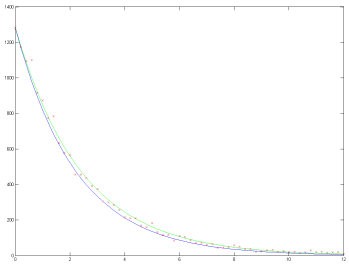
einmaliges Iterieren				
	gleichmäßige Verteilung		**Normalverteilung**	
M	Rechteckregel	Trapezregel	Rechteckregel	Trapezregel
10	0.0929	0.1193	2.0948	2.1822
20	0.1783	0.1607	1.6929	1.9111
50	0.0538	0.0223	0.7048	0.6876
100	$9.7528 \cdot 10^{-4}$	0.0049	1.0166	1.0229
250	0.0090	$5.5410 \cdot 10^{-4}$	0.7983	0.8046

Cäsium-Barium-Zerfall

Die genaue Zerfallskonstante des Cäsium-Barium-Zerfalls ist $k = 0.45$. Die Ergebnisse, die mittels des in [13] vorgestellten Messverfahrens gewonnen wurden, liefern unter Anwendung der Rechteckregel eine Approximation von $k = 0.40$. Im Gegenzug ergibt das Trapezverfahren einen Wert von $= 0.41$. Der relative Fehler liefert also für die 60 realen Messpunkte einmal für die Rechteckformel 10.49% und für die Trapezformel 7.87%.



(a) Rechteckregel



(b) Trapezregel

Abbildung 7.1: Cäsium-Barium Zerfall nach Wikström ausgewertet; zu sehen sind die Originalfunktion (blau), die Messdaten (rot) und die Funktion mit den approximierten Parametern k (grün).

Grafik 7.2 zeigt deutlich, dass das Gauß-Newton Verfahren den Parameter k schlechter approximiert, als die Methode Wikströms nach der Trapezregel. Mit einer Abschätzung für $k = 0.3869$ nach 12 Operationsschritten beträgt der relative Fehler $14,07\%$.

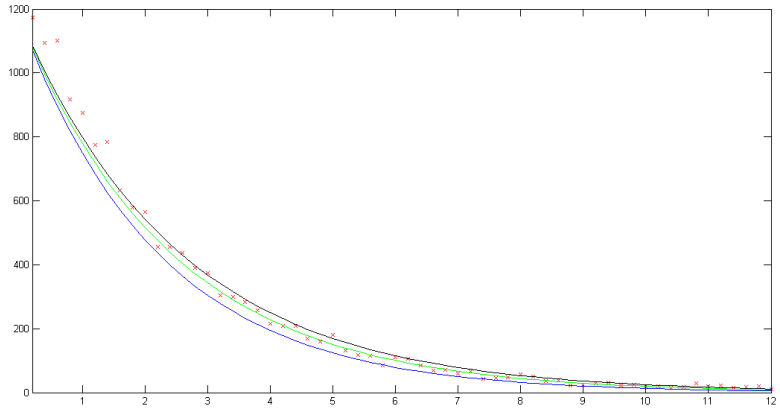


Abbildung 7.2: Cäsium-Barium Zerfall nach Gauß-Newton ausgewertet; zu sehen sind die Originalfunktion (blau), die Messdaten (rot), die Funktion mit dem approximierten Parameter k aus der Wikström-Methode (Trapezregel-grün) und die Auswertung mit dem Parameter der Gauß-Newton Funktion (schwarz).

7.1.2 Das Barnes-Problem

Für das zweidimensionale Barnesproblem werden zum Abschätzen einmal der Parametervektor $k^* = [2, 5, 7]^T$ eingesetzt, ein anderes Mal $k^* = [1, 1, 1]^T$.

Synthetisch generierte Daten für $k^* = [2, 5, 7]^T$

einmaliges Iterieren				
	Gleichverteilung		**Normalverteilung**	
M	Rechteckregel	Trapezregel	Rechteckregel	Trapezregel
10	0.5109	0.1013	0.8691	0.7931
20	0.2532	0.1966	0.8101	0.8177
50	0.3008	0.2806	0.4215	0.3990
100	0.1206	0.1148	0.2151	0.2354
250	0.0569	0.0060	0.1703	0.1773

Synthetisch generierte Daten für $k^* = [1, 1, 1]^T$

einmaliges Iterieren				
	Gleichverteilung		**Normalverteilung**	
M	Rechteckregel	Trapezregel	Rechteckregel	Trapezregel
10	0.0724	0.0307	0.2789	0.2316
20	0.0606	0.0463	0.1673	0.1566
50	0.0336	0.0319	0.1003	0.1005
100	0.0171	0.0168	0.0994	0.0996
250	0.0140	0.0139	0.1290	0.1299

7.2 Integration durch mehrfaches Iterieren

Zur mehrfachen Iteration für $D = 1$ und $D = 10$ wird zusätzlich die Anzahl Z der zusätzlichen Messwerte auf jedem Intervall I_i variiert. Auf das eindimensionale Zerfallsproblem wird hier nicht näher eingegangen, da die zuvor erreichten Werte schon gute Näherungen an den Parameter k^* gezeigt haben.

7.2.1 Barnes-Problem

Vektor $[2, 5, 7]^T$, Gleichverteilung

mehrfaches Iterieren (Gleichverteilung)								
	$Z = 40$				$Z = 100$			
	Rechteckregel		**Trapezregel**		**Rechteckregel**		**Trapezregel**	
M	$D = 1$	$D = 10$	$D = 1$	$D = 10$	$D = 1$	$D = 10$	$D = 1$	$D = 10$
10	0.3349	0.1867	0.1972	0.1761	0.3330	0.1829	0.2000	0.1013
20	0.1904	0.1839	0.2003	0.2000	0.1899	0.1834	0.2012	0.2009
50	0.3009	0.3007	0.2903	0.2900	0.3010	0.3008	0.2806	0.2902
100	0.1218	0.1219	0.1199	0.1200	0.1218	0.1220	0.1200	0.1202
250	0.0058	0.0058	0.0054	0.0054	0.0058	0.0058	0.0053	0.0054

Vektor $[2, 5, 7]^T$, Normalverteilung

mehrfaches Iterieren (Normalverteilung)								
	$Z = 40$				$Z = 100$			
	Rechteckregel		**Trapezregel**		**Rechteckregel**		**Trapezregel**	
M	$D = 1$	$D = 10$	$D = 1$	$D = 10$	$D = 1$	$D = 10$	$D = 1$	$D = 10$
10	0.8334	0.8178	0.7202	0.7235	0.8328	0.8165	0.7177	0.7208
20	0.8069	0.8072	0.8279	0.8275	0.8069	0.8072	0.8280	0.8276
50	0.4023	0.4050	0.3910	0.3946	0.4021	0.4049	0.3909	0.3946
100	0.2248	0.2256	0.2457	0.2465	0.2250	0.2258	0.2459	0.2466
250	0.1744	0.1744	0.1817	0.1819	0.1745	0.1747	0.1818	0.1820

Vektor $[1, 1, 1]^T$, Gleichverteilung

mehrfaches Iterieren (Gleichverteilung)								
	$Z = 40$				$Z = 100$			
	Rechteckregel		**Trapezregel**		**Rechteckregel**		**Trapezregel**	
M	$D = 1$	$D = 10$	$D = 1$	$D = 10$	$D = 1$	$D = 10$	$D = 1$	$D = 10$
10	0.0298	0.0316	0.0094	0.0117	0.0294	0.0313	0.0099	0.0126
20	0.0484	0.0505	0.0491	0.0514	0.0483	0.0505	0.0493	0.0517
50	0.0338	0.0343	0.0356	0.0361	0.0339	0.0344	0.0357	0.0362
100	0.0171	0.0172	0.0182	0.0184	0.0171	0.0173	0.0183	0.0184
250	0.0139	0.0139	0.0141	0.0141	0.0139	0.0139	0.0141	0.0141

Vektor $[1,1,1]^T$, Normalverteilung

mehrfaches Iterieren (Normalverteilung)								
	$Z = 40$				$Z = 100$			
	Rechteckregel		**Trapezregel**		**Rechteckregel**		**Trapezregel**	
M	$D = 1$	$D = 10$	$D = 1$	$D = 10$	$D = 1$	$D = 10$	$D = 1$	$D = 10$
10	0.2221	0.2216	0.2048	0.2091	0.2216	0.2211	0.2047	0.2091
20	0.1570	0.1577	0.1598	0.1610	0.1570	0.1577	0.1600	0.1612
50	0.0994	0.0997	0.1013	0.1016	0.0994	0.0998	0.1013	0.1017
100	0.0982	0.0982	0.0988	0.0988	0.0982	0.0982	0.0988	0.0988
250	0.1290	0.1290	0.1299	0.1300	0.1290	0.1290	0.1299	0.1299

7.3 Exakte Integration durch glättende Splines

In diesem Fall werden die Fehler in den Messdaten ausschließlich durch eine Gleichverteilung erzeugt. Darüber hinaus geht der Streuparameter für $\delta = 0.1$ und $\delta = 0.01$ in die Untersuchungen mit ein sowie die in Bemerkung 5.1 angesprochenen Skalierungsfaktoren $\beta = \frac{1}{3}$ und $\beta = \frac{1}{5}$

7.3.1 Zerfallsprozesse

	$k = 0.8, \delta = 0.1$	
M	$\beta = \frac{1}{3}$	$\beta = \frac{1}{5}$
10	0.2085	0.1446
20	0.1434	0.0627
50	0.0728	0.0188
100	0.0897	0.0092
250	0.0258	$8.1348 \cdot 10^{-4}$

	$k = 0.8, \delta = 0.01$	
M	$\beta = \frac{1}{3}$	$\beta = \frac{1}{5}$
10	0.1308	0.1264
20	0.0704	0.0643
50	0.0309	0.0260
100	0.0202	0.0144
250	0.0105	0.0058

	$k = 0.25, \delta = 0.1$	
M	$\beta = \frac{1}{3}$	$\beta = \frac{1}{5}$
10	0.1554	0.1780
20	0.1407	0.0592
50	0.0772	0.0057
100	0.1535	0.0047
250	0.0026	0.0078

	$k = 0.25, \delta = 0.01$	
M	$\beta = \frac{1}{3}$	$\beta = \frac{1}{5}$
10	0.1548	0.1420
20	0.0864	0.0705
50	0.0392	0.0288
100	0.0304	0.0150
250	0.0116	0.0054

7.3.2 Das Barnes-Problem

$k^* = [2,5,7]^T, \delta = 0.1$		
M	$\beta = \frac{1}{3}$	$\beta = \frac{1}{5}$
10	0.2647	0.2532
20	0.1192	0.1225
50	0.2701	0.1710
100	0.0845	0.0933
250	0.0327	0.0245

$k^* = [2,5,7]^T, \delta = 0.01$		
M	$\beta = \frac{1}{3}$	$\beta = \frac{1}{5}$
10	0.0095	0.0130
20	0.0151	0.0148
50	0.0135	0.0136
100	0.0166	0.0166
250	0.0164	0.0161

Synthetisch generierte Daten für $k^* = [1,1,1]^T$

$k^* = [1,1,1]^T, \delta = 0.1$		
M	$\beta = \frac{1}{3}$	$\beta = \frac{1}{5}$
10	0.0828	0.0807
20	0.0983	0.0959
50	0.1054	0.1011
100	0.1156	0.1098
250	0.1222	0.1177

$k^* = [1,1,1]^T, \delta = 0.01$		
M	$\beta = \frac{1}{3}$	$\beta = \frac{1}{5}$
10	0.0821	0.0818
20	0.1037	0.1033
50	0.1164	0.1159
100	0.1213	0.1207
250	0.1243	0.1238

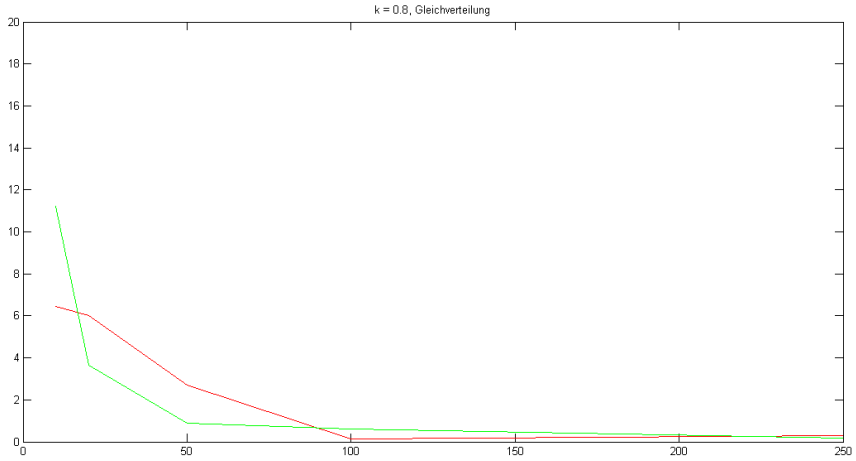
Kapitel 8

Auswertung der Ergebnisse

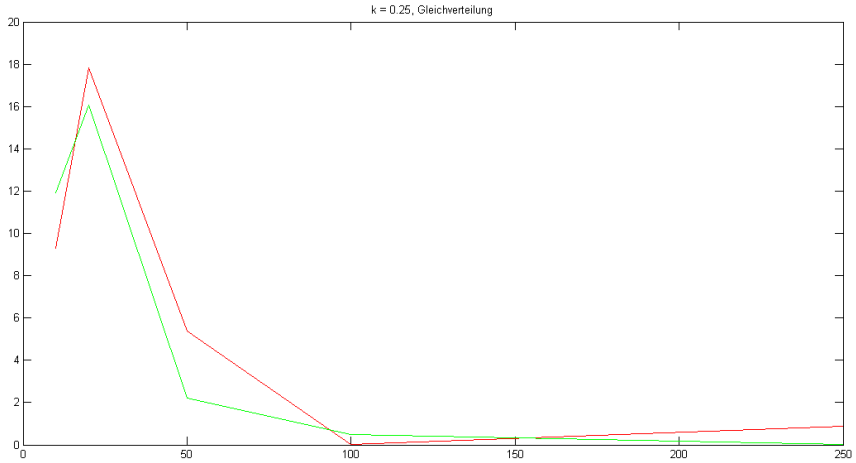
8.1 Methode Wikström

8.1.1 Zerfallsprozesse

Betrachtet man das eindimensionale Zerfallsproblem mit $k^* = -0.8$, so stellt man fest, dass bei einer gleichmäßigen Verteilung der Messwerte um die eigentliche Funktion y und einer Streubreite von $\delta = 0.1$ der relative Fehler der Parameterschätzung schon bei wenigen Messwerten unter 10% liegt. Bei einer großen Anzahl der Messwerte werden mit einer Fehlerungenauigkeit von weniger als 0.5% hervorragende Schätzungen erzielt, wobei die Trapezregel für Messwerte $M > 10$ in den Regel genauere Approximationen liefert, was die Erwartungen bezüglich der Fehleranalyse in [14] bestätigt.
Die gleichen Voraussetzungen für $k^* = -0.25$ lassen keine eindeutigen Aussagen über die Effizienz bezüglich der verwendeten Quadraturformeln treffen. Für $M = 10$ Messwerte liegt der relative Fehler bei ca. 10%, steigt anschließend fast auf das Doppelte an und fällt dann rapide auf weniger als 5% ab. Allerdings liegt der Ausreißer bei $M = 250$ für die Rechteckregel unter 1%, jedoch ist anzumerken, dass dies wieder das Neunfache des relativen Fehlers für $M = 100$ ist.



(a) Gleichverteilung, $k = 0.8$



(b) Gleichverteilung, $k = 0.25$

Abbildung 8.1: Zu sehen ist der relative Fehler (Angaben in Prozent). Dabei steht der rote Graph für die Rechteckregel, der grüne für die Trapezregel.

Anders sieht es bei einer normalverteilten Streuung der Funktionswerte aus. Mit Fehlern für wenige Messwerte von fast 100% bzw. 200% für den Parameter $k^* = -0.25$ ist diese Methode gänzlich ungeeignet. Zwar dezimiert sich der relative Fehler für $k^* = -0.8$ und $M = 250$ der F_{rel} auf ca. 40%, doch eine gute Approximation lässt sich in diesem Fall nicht feststellen. Insgesamt sinkt der relative Fehler für $k^* = -0.25$ nicht unter 60%.
Die Methode Wikströms auf die realen Messdaten (siehe Anhang) angewendet

ergibt einmal einen relativen Fehler von 10.49% für die Rechteckregel und für die Trapezformel 7.87%. Die Trapezregel approximiert also für M=60 Messwerte den gesuchten Parameter $k^* = 0.45$ besser als die Rechteckregel. Man erkennt zudem, dass das Gauß-Newton Verfahren für nichtlineare Regressionsprobleme in diesem Fall deutlich schlechter den gesuchten Parameter approximiert, als der zuvor verwendete Algorithmus.

8.1.2 Barnes-Problem

Vergrößert man nun die Dimension sowohl des Differentialgleichungssystems als auch die des Parameters und betrachtet exemplarisch das zweidimensionale Barnes-Problem, so erkennt man, dass für einen „komplizierten“ Parametervektor $k^* = [2, 5, 7]^T$ die Rechteckregel allgemein schlechtere Resultate liefert. Bei einer Gleichverteilung des Fehlers ist diese Methode für wenige Daten mit einem relativen Fehler von 50% und sogar über 80% unter einer Normalverteilung außerhalb des verwertbaren Bereichs anzusiedeln. Für $M = 250$ Messwerte wird bei der Gleichverteilung mit der Rechteckregel ein Fehler von ca. 5% erzielt. Dies ist mit dem Ergebnis von 0.6% unter Verwendung der Trapezregel qualitativ nicht zu vergleichen. Probleme ergaben sich bei dieser Quadraturformel jedoch für $M \in \{20, ..., 100\}$; hier steigt der relative Fehler auf bis zu 30% an.
Auf normalverteilte Datenfehler angewendet liefert die Trapezregel vergleichbare Ergebnisse wie die Rechteckregel: Beide Resultate weisen inakzeptable Fehler auf.
Für $k^* = [1, 1, 1]^T$ wird der Parameter bei einer Gleichverteilung des Datenfehlers und steigender Anzahl der Messwerte M hingegen sehr gut ausgewertet. Angefangen bei einem maximalem relativen Fehler von $7,5\%$ ($M = 10$) sinkt er auf $1,4\%$ ($M = 250$) herab - sowohl unter Verwendung der Rechteck- als auch der Trapezregel.
Zwar liefert der relative Fehler bei einer normalverteilten Streuung des Datenfehlers für diesen Parametervektor bessere Resultate als für $k^* = [2, 5, 7]^T$, weniger als 10% sind aber auch hier nicht zu erwarten.

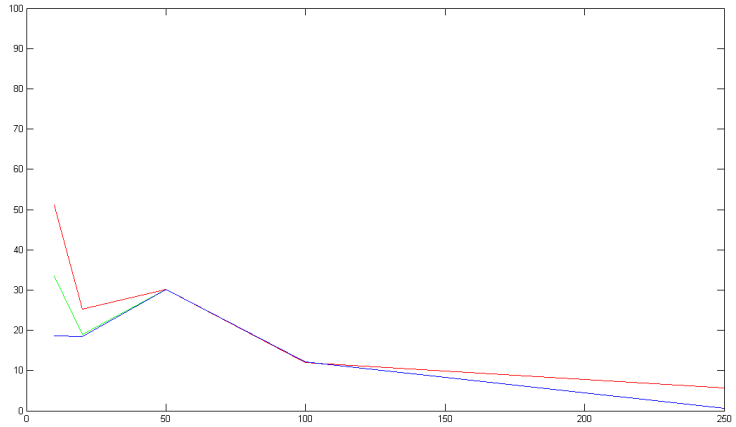
8.2 Ergebnisse des Iterationsverfahrens

Da die Auswertung mittels der Methode von Wikström bereits sehr gute Ergebnisse gezeigt hat, wird der einfache Zerfallsprozess in diesem Fall nicht näher betrachtet und direkt auf das mehrdimensionale Problem von Barnes eingegangen.
Bei der Auswertung der mehrfachen Iteration sind nun zwei weitere Faktoren zu beachten: Zum einen die Anzahl Z der zusätzlichen Zwischenstellen, zum

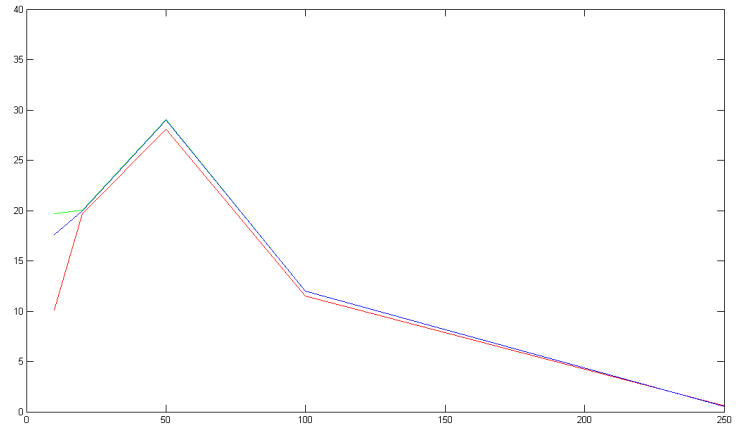
anderen die Anzahl D der durchgeführten Iterationen.

8.2.1 Barnes-Problem

Zunächst wird die Schätzung des Parametervektors $k^* = [2, 5, 7]^T$ mit gleichverteilten Messwerten betrachtet. Hier ist unter Verwendung der Rechteckregel bereits für die erste Iterierte $k^{(1)}$ im Vergleich zu $k^{(0)}$ eine sehr gute Verringerung des relativen Fehlers von ca. 51% auf 33% zu erkennen. Je höher jedoch die Anzahl der Messwerte wird, desto weniger zieht eine wiederholte Iteration eine signifikante Verbesserung nach sich. Ausgenommen ist der untersuchte Maximalwert $M = 250$: Bereits eine einfache Iteration bringt eine Verbesserung von 5.69% auf 0.58%. Bemerkenswert ist die Tatsache, dass für $M = 100$ der relative Fehler 12% beträgt, für das zweieinhalbfache der Werte der Fehler kleiner ist als 1%, der Parameter also sehr gut approximiert wird.
Bei der Trapezregel wird die Schätzung von $k^{(1)}$ schlechter im Vergleich zu $k^{(0)}$, d.h. die zusätzlichen Iterationen bringen hier keine Verbesserung. Die Ausnahme liegt wieder bei $M = 250$: hier ist eine minimale Verbesserung von 0.06% von $k^{(0)}$ zu $k^{(1)}$ bzw. $k^{(10)}$ festzustellen. Interessant ist die Tatsache, dass die Erhöhung der Zwischenstellen von $Z = 40$ auf $Z = 100$ keinen nennenswerten Unterschied erzeugt.
Ähnlich sieht die relative Fehlerauswertung des Vergleichs von $k^{(0)}$ und $k^{(1)}$ bzw. $k^{(0)}$ und $k^{(10)}$ im Hinblick auf die normalverteilten Datenfehler aus. Teilweise werden sogar schlechtere Resultate erzielt, sodass man sagen kann, dass das Verfahren in diesem Fall keine Nutzen in sich birgt.



(a) Gleichverteilung, Rechteckregel, $Z = 40$

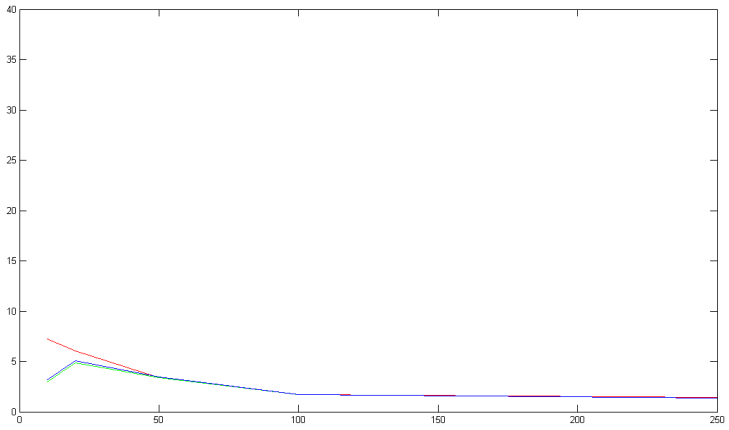


(b) Gleichverteilung, Trapezregel, $Z = 40$

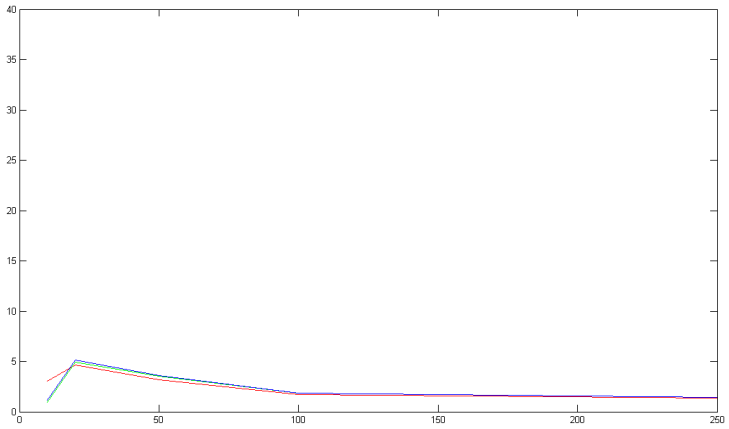
Abbildung 8.2: Zu sehen ist der relative Fehler (Angaben in Prozent) des Parameters $k = [2, 5, 7]^T$. Dabei steht der rote Graph für $k^{(0)}$, der grüne für $k^{(1)}$ und der blaue für $k^{(10)}$.

Bei Betrachtung von $k^* = [1, 1, 1]^T$ zeigt sich bei gleichverteiltem Zufallsrauschen, einer wiederholten Iteration und wenigen Messdaten wieder ein positives Ergebnis, besonders im Hinblick auf die Rechteckregel. Auch wenn bei zunehmender Anzahl an Messwerten die Approximation zunächst leicht schlechter wird, so ist sie insgesamt besser als die Schätzung bei $k^{(0)}$ und auch als bei $k^{(10)}$.
Unter Verwendung der Trapezregel nützt diese Methode lediglich etwas für

$D = 1$. Ansonsten wird für mehrere Messwerte der relative Fehler des geschätzten Parameters größer für $D >> 1$. Insgesamt werden für $k^{(0)}$ durch die Methode Wikströms immer noch bessere Ergebnisse erzielt, sodass die Mehrfachiteration in diesem Fall „versagt“.



(a) Gleichverteilung, Rechteckregel, $Z = 40$

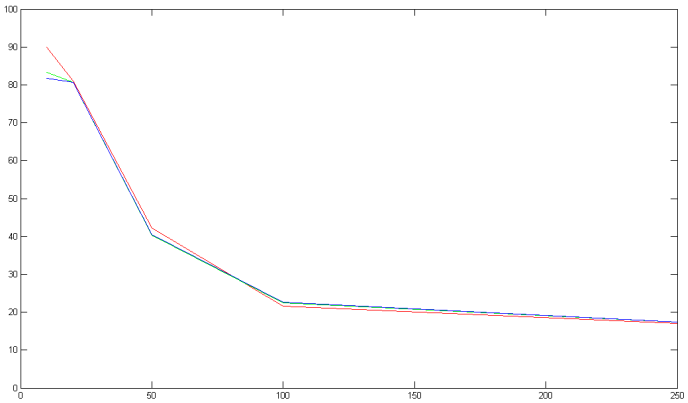


(b) Gleichverteilung, Trapezregel, $Z = 40$

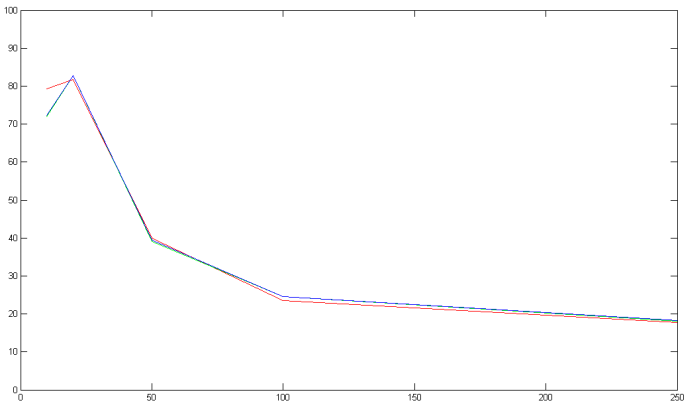
Abbildung 8.3: Zu sehen ist der relative Fehler (Angaben in Prozent) des Parameters $k = [1, 1, 1]^T$. Dabei steht der rote Graph für $k^{(0)}$, der grüne für $k^{(1)}$ und der blaue für $k^{(10)}$.

Unerheblich ist auch hier wieder die Anzahl der zusätzlichen Stützstellen Z. Die Fehlerbeträge sind für $Z = 40$ und $Z = 100$ annähernd austauschbar und

mit einem relativen Fehler zwischen 5% und 1.4% im Allgemeinen gut ausgewertet. Die Iterationsmethode ausgewertet auf normalverteilte Datenfehler zeigt wie im Fall zuvor nur für wenige Messdaten einen Nutzen und auch nur bei einer Wiederholung der Iteration zu $k^{(1)}$. Bei $M = 250$ ist der relative Fehler identisch mit $k^{(0)}$ bezüglich der initialen Parameterschätzung; die hier erzielten 13% des relativen Fehlers sind im Vergleich zu dem Parameter $k = [2, 5, 7]^T$ jedoch ca. 5% kleiner.

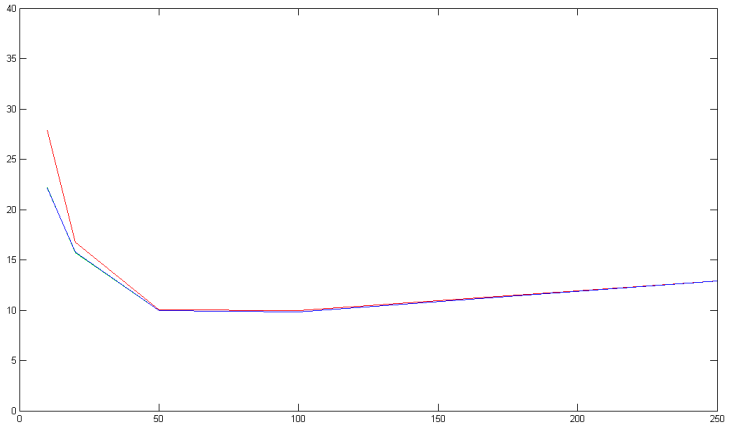


(a) Normalverteilung, Rechteckregel, $Z = 40$

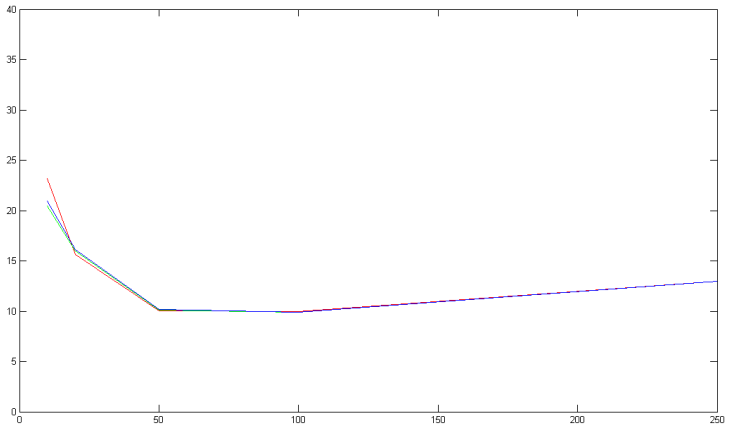


(b) Normalverteilung, Trapezregel, $Z = 40$

Abbildung 8.4: Zu sehen ist der relative Fehler (Angaben in Prozent) des Parameters $k = [2, 5, 7]^T$. Dabei steht der rote Graph für $k^{(0)}$, der grüne für $k^{(1)}$ und der blaue für $k^{(10)}$.



(a) Normalverteilung, Rechteckregel, $Z = 40$



(b) Normalverteilung, Trapezregel, $Z = 40$

Abbildung 8.5: Zu sehen ist der relative Fehler (Angaben in Prozent) des Parameters $k = [1, 1, 1]^T$. Dabei steht der rote Graph für $k^{(0)}$, der grüne für $k^{(1)}$ und der blaue für $k^{(10)}$.

8.3 Glättende Splines

8.3.1 Zerfallsprozesse

Für einen einfachen Zerfallsprozess mit Parameter $k^* = -0.8$ und gleichverteiltem Maximalfehler $\delta = 0.1$ ist der relative Parameterfehler bezüglich des Skalierungsfaktors $\beta = \frac{1}{3}$ mit ca. 20% signifikant größer als für $\beta = \frac{1}{5}$ mit 15%.

Während im letztgenannten Fall schon für $M = 50$ der relative Fehler bei 2% liegt, sind für die gleiche Anzahl M und $\beta = \frac{1}{3}$ noch 9% Fehlerbandbreite zu erkennen. Die numerischen Experimente zeigen, dass bei einem $\beta = \frac{1}{5}$ eine bestmögliche Approximation der Funktionswerte erzielt wird.
Interessant ist, dass für $\delta = 0.1$ gerade bei vielen Messdaten eine Verschlechterung des zuvor abgeschätzten relativen Fehlers zu erkennen ist und zwar sowohl für $k^* = -0.8$ als auch für $k^* = -0.25$. Dieses Phänomen ist für $\delta = 0.01$ augenscheinlich nicht zu erkennen.
Während die Auswertung beider Parameter $k^* = -0.8$ und $k^* = -0.25$ für wenige Messdaten und $\beta = \frac{1}{3}$ mit einem relativen Fehler zwischen 14% und 21% relativ hoch ist, wird im Vergleich dazu bei $\beta = \frac{1}{5}$ die Bandbreite mit ca. $6\% - 18\%$ niedriger begonnen. Für große M wird der Fehler kleiner als 1%, bleibt dennoch insgesamt schlechter als der nach der Wikström-Methode gemessene Wert. Auch eine Verkleinerung der Streubreite zu $\delta = 0.01$ zeigt zwar bedingt Verbesserungen, diese sind jedoch nicht weiter beachtenswert.
Insgesamt kann man für das eindimensionale Problem behaupten, dass das Verfahren sehr sensibel auf Variation der Skalierungsparameter reagiert. Dagegen zeigt eine Streuung der Daten innerhalb gewisser Grenzen eher einen geringen Effekt, was vermutlich auf die glättende Wirkung der numerischen Integration zurückzuführen ist.
Die numerischen Experimente zeigen, dass ein kompletter Verzicht auf den Skalierungsfaktor (d.h. $\beta = 1$) schlechte Resultate liefert. Die Annahme, dass der maximale Fehler in jedem Datenpunkt angenommen wird, scheint zu pessimistisch und entsprechend erfolgt im Verfahren eine Überregularisierung, d.h. der „Regularisierungsparameter“ wird zu hoch angesetzt.

8.3.2 Barnes-Problem

Im zweidimensionalen Fall für $k^* = [2, 5, 7]^T$ kann man bis auf $M = 10$ beobachten, dass die Methode für beide Skalierungsfaktoren besser funktioniert als die Methode Wikströms.
Für $k^* = [1, 1, 1]^T$ hingegen ist der Fehler für zunehmende Anzahl an Messwerten nach oben hin ansteigend, was für beide verwendete Streu- und Skalierungsfaktoren in diesem Fall eine relative Fehlerbandbreite von $8\% - 12\%$ ausmacht, also insgesamt etwa 10% schlechter auswertet, als der Algorithmus Wikströms.

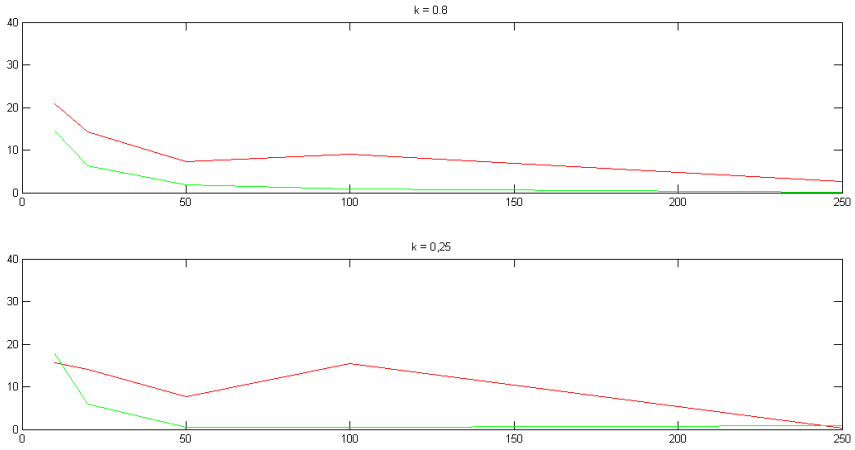


Abbildung 8.6: Zu sehen ist der relative Fehler (Angaben in Prozent) des Parameters eines Zerfallsprozesses für $\delta = 0.1$. Im oberen Bild erfolgte die Auswertung für $k = 0.8$, im unteren für $k = 0.25$. Der rote Graph steht für den Skalierungsfaktor $\beta = \frac{1}{3}$, der grüne für $\beta = \frac{1}{5}$.

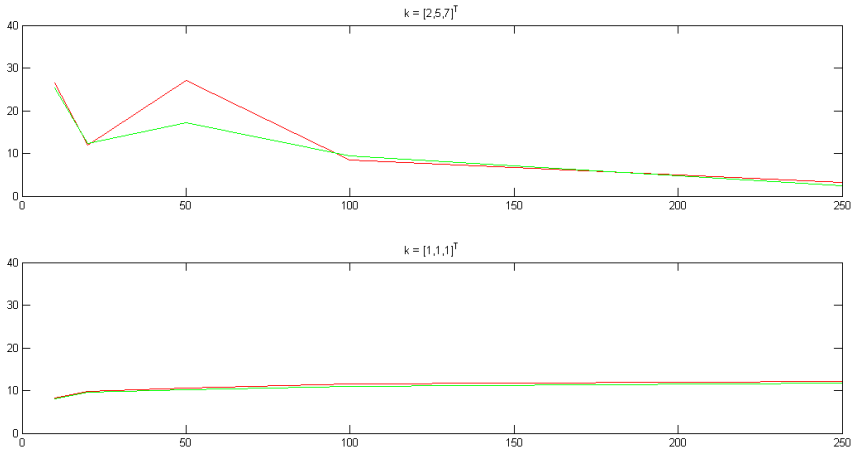


Abbildung 8.7: Zu sehen ist der relative Fehler (Angaben in Prozent) des Parameters des Barnesproblems für $\delta = 0.1$. Im oberen Bild erfolgte die Auswertung für $k = [2, 5, 7]^T$, im unteren für $k = [1, 1, 1]^T$. Der rote Graph steht für den Skalierungsfaktor $\beta = \frac{1}{3}$, der grüne für $\beta = \frac{1}{5}$.

Kapitel 9

Fazit und Ausblick

Abschließend lässt sich nach der experimentellen Analyse der Algorithmen festhalten, dass für das eindimensionale Problem eines Zerfallsprozesses die Methode Wikströms am effektivsten den gesuchten Parameter abschätzt, sofern die auf die Messwerte aufaddierten Fehler gleichverteilt sind. Der gleiche Algorithmus liefert bei einer Normalverteilung des Rauschens schlechtere Ergebnisse. Ein Grund könnte darin liegen, dass der Algorithmus sehr sensibel auf Ausreißer reagiert.
Obwohl bei der Durchführung des Gauß-Newton Verfahrens die exakte Lösung der Differentialgleichung bekannt ist, liefert es eine schlechtere Approximation des Parameters k als die Wikström-Methode. Dieses Resultat zeigt die gute Verwertbarkeit der letztgenannten Methode bezüglich dieser Art von Anfangswertproblemen.

Bei der Barnes-Funktion spielt der zu approximierende Parameter eine weitaus größere Rolle: Für gleichverteilte Fehler in den Eingabedaten ist für den Parameter $k^* = [2, 5, 7]^T$ die Wikström-Methode mit 10 bis 50 Messdaten in der Praxis nicht verwertbar, da mit einer Fehlerbandbreite von 30% bis 50% gerechnet werden muss. Dahingegen zeigt sich für $k^* = [1, 1, 1]^T$ eine sehr gute Auswertung mit maximalem relativen Fehler $F_{rel} = 7.24\%$, der auf ca. 1% sinkt, wenn die Anzahl M der Messwerte erhöht wird. Der kleinste relative Fehler für $k^* = [2, 5, 7]^T$ beträgt unter der Annahme einer normalverteilten Streuung der Daten ca. 17%, für $k^* = [1, 1, 1]^T$ etwa 10%. Für diesen Rauschtyp lässt sich die Tendenz festhalten, dass mit zunehmender Größe M der Parametervektor besser approximiert wird. Die Auswertung des relativen Fehlers bleibt jedoch im Vergleich zu den bereits gewonnenen Ergebnissen einer Gleichverteilung eher schlecht.
Insgesamt zeichnet sich ab, dass die Anwendung der Trapezregel ab 20 Messwerten die kleineren Abweichungen an den Originalparameter liefert. Diese Beobachtung deckt sich mit der theoretischen Fehleranalyse in den Arbeiten

von [14].
Zumindest in den numerischen Experimenten dieser Arbeit setzt sich die Überlegenheit der Trapezregel in der mehrmaligen Iteration nicht zwangläufig fort: Nur für $M = 10$ wertet die Trapezregel noch besser aus als die Rechteckregel, für größere M werden die relativen Fehlerbeträge der Trapezregel schlechter. Wie schon erwähnt birgt eine höhere Anzahl Z an zusätzlichen Funktionswerten keine Vorteile in sich: Mit $Z = 40$ Zwischenstellen auf jedem Intervall werden Resultate erzielt, die denen für $Z = 100$ in etwa entsprechen.
Offensichtlich bringt eine zusätzliche Iteration nur dann etwas, wenn wirklich sehr wenige Messwerte vorliegen, denn schon ab $M = 20$ ist auch die Veränderung von $k^{(1)}$ zu $k^{(10)}$ nur noch auf den Nachkommastellen sichtbar. Da in der Realität mitunter nur wenige oder unvollständige Daten erhoben werden können, ermöglicht die vorgestellte Iterationsmethode eine Verbesserung der Resultate im Vergleich zum ursprünglichen Wikström-Verfahren.
Die relative Fehlerauswertung bei der Iterationsmethode für eine normalverteilte Fehlerstreuung auf die exakten Werte einer Funktion y zeigt im Vergleich zu $k^{(0)}$ keine signifikante Verbesserung.
Der dritte Algorithmus der Integration mittels glättenden Splines, in dem gleichverteilte Messfehler betrachtet werden, offenbart im eindimensionalen Problem im Vergleich zur Methode Wikströms für lediglich 10 bis 50 Messwerte zwar eine schlechtere Approximation des Parameters, allerdings sind die relativen Fehler für viele Messwerte vergleichbar und teilweise sogar besser.
In der Analyse des zweidimensionalen Barnes-Problems wird der „komplizierte“ Parameter für $k = [2, 5, 7]^T$ mit dem Splineansatz besser bestimmt, als mit der entsprechenden Wikström-Methode: Die relativen Fehler unter Verwendung der Rechteckregel sind in diesem Fall grundsätzlich besser ausgewertet; für die Trapezregel gilt dies nur bedingt.
Dahingegen ist bei der „einfachen“ Barnes-Funktion mit $k^* = [1, 1, 1]^T$ ein ausschließlich negativer Effekt zu beobachten: Der relative Fehler wächst mit steigender Zahl M an Messwerten. Zwar sind 12% je nach Fragestellung noch vertretbar, denn in den vorangehenden Algorithmen treten weitaus größere Fehler auf. Dennoch muss man sagen, dass die Methode den numerischen Aufwand in diesem Fall nicht rechtfertigt.
Interessant ist in diesem Algorithmus, dass eine Verkleinerung der Streubreite δ nicht unbedingt impliziert, dass die Approximation des Parameters genauer wird. Die numerischen Experimente haben gezeigt, dass die Qualität der Resultate stark von dem in Kapitel 5.1 eingeführten Skalierungsfaktor β abhängt. In der Regel gewinnt man für $\beta = \frac{1}{5}$ eine genauere Approximationen als für $\beta = \frac{1}{3}$.
Lediglich beim Barnes-Problem für $k^* = [2, 5, 7]^T$ ist bei einer Reduktion von δ eine deutliche Verbesserung der Approximationsgüte zu erkennen.

Zusammenfassend ist hervorzuheben, dass die Methode Wikströms in der Mehrheit der untersuchten Funktionen gute Resultate bezüglich der Parameterschätzung liefert. Für sehr wenige Messwerte sollten einige Iterationen durchgeführt werden; zu viele Iterationen ziehen jedoch wieder eine Verschlechterung des abgeschätzten Parameters nach sich. Eine Normalverteilung der fehlerhaften Messdaten bewirkt eine allgemein relativ schlechte Schätzung des Parameters. Für den Ansatz der glättenden Splines ist die Fehlerannahme nicht ohne Weiteres umsetzbar. Die exakte Integration über die glättenden Splines löst erst bei komplizierteren Differentialgleichungssystemen und einer hohen Zahl an zur Verfügung stehenden Messwerten einen positiven Effekt aus. Ist man an einem schnellen Algorithmus interessiert und steht eine Vielzahl an Messwerten zur Verfügung, so sollte im Allgemeinen jedoch eher auf die Methode Wikströms zurückgegriffen werden, da keine nennenswerte Verbesserung mehr erzielt werden kann.

Für zukünftige Forschungsansätze bezüglich dieser Thematik könnte man untersuchen, was passiert, wenn die Messdaten nicht äquidistant verteilt sind beziehungsweise wenn Messdaten fehlen. Darüber hinaus könnte man eine weitere Quadraturformel wie etwa die Simpsonregel implementieren und untersuchen, wie die Methode Wikströms anschließend funktioniert bzw. ob eine wiederholte Iteration weitere Verbesserungen bewirkt. Zudem wäre es interessant zu erfahren, ob bei steifen Differentialgleichungssystemen die hier vorgestellten Algorithmen möglicherweise Vorteile gegenüber der Standardmethode mit sich bringen.
Im Zuge der glättenden Splines sollte die Wirkung des Skalierungsfaktors noch intensiver analysiert werden, um Überregularisierungen zu vermeiden.
Weder in den Publikationen von Wikström noch in dieser Arbeit wurde untersucht, wie sich eine Gewichtung bei der Summe der Fehlerquadrate auf die Parameterbestimmung auswirkt. Denkbar ist es zum Beispiel statt des absoluten Fehlers $\|y_i - \tilde{y}_i\|_2^2$ den relativen Fehler

$$\frac{\|y_i - \tilde{y}_i\|_2^2}{\|\tilde{y}_i\|_2^2}$$

zu minimieren.

Experimentell erhobene Daten eines Caesium-137/Barium-137m-Zerfalls.

Zeit (s)	*Messdaten*	*Zeit (s)*	*Messdaten*	*Zeit (s)*	*Messdaten*
20	1174	420	209	820	49
40	1095	440	209	840	36
60	1101	460	169	860	39
80	917	480	159	880	22
100	874	500	181	900	24
120	775	520	132	920	28
140	783	540	117	940	32
160	633	560	116	960	22
180	578	580	84	980	25
200	565	600	111	1000	20
220	456	620	106	1020	19
240	456	640	86	1040	14
260	437	660	69	1060	18
280	390	680	71	1080	28
300	374	700	59	1100	20
320	305	720	66	1120	22
340	299	740	44	1140	16
360	284	760	45	1160	18
380	257	780	47	1180	19
400	215	800	57	1200	11

Quelle: http://www.buetzer.info/fileadmin/pb/pdf-Dateien/Ba-137m.pdf,
(zuletzt eingesehen: 29.07.2010)

Für diese Messungen wurde der „Vernier Radiation Monitor“ verwendet. Die Messdauer pro Probe beträgt 20 Sekunden, die gesamte Messdauer 1200 Sekunden.

DANKSAGUNG

An dieser Stelle bleibt mir nur noch Danke zu sagen an all diejenigen, die mir geholfen haben, diese Arbeit zu dem entstehen zu lassen, was sie letzten Endes geworden ist.
Zunächst gilt dies Herrn Prof. Dr. A.K. Louis für seinen interessanten Themenvorschlag und die umfassende Betreuung, die er mir hat zukommen lassen.
Aus dem allgemein sehr hilfsbereiten und kollegialen Team des Lehrstuhls bei dem ich mich vor allem bei Dipl.-Math. Martin Riplinger für sein Korrekturlesen und seine Ausdauer in der Lösung meiner technischen Fragen bedanken möchte, gilt mein besonderer Dank Dr. Andreas Groh, der mir mit seinen Anregungen, Korrekturvorschlägen und seiner immerwährenden Hilfe eine großartige Stütze in den letzten Monaten war!

Doch auch die „Uniauswärtigen“ möchte ich nicht vergessen:
Florian Schneider, der mir des öfteren bei Programmierschwierigkeiten aus der Ferne zur Verfügung stand.
Meine Freunde, die in den letzten Wochen ein wenig zurückstehen mussten und dennoch gerade in der Phase der Fertigstellung und beim Korrekturlesen der Arbeit für mich da waren.
Mein Bruder, der es nicht aufgibt mich immer wieder zu lehren, dass man im Leben ein wenig weiter vorausschauen sollte (und damit *manchmal* Recht hat...).
Am Wichtigsten und deswegen auch am Schluss erwähnt: Meine Eltern auf die ich mich immer verlassen kann und die mich in meiner gesamten Laufbahn unterstützt haben. Ohne eure Hilfe hätte ich das Studium nicht in dieser Art und Weise durchziehen können!

Danke Euch allen!

Literaturverzeichnis

[1] Mostafa A. Abdelkader. Exact solutions of lotka-volterra equations. *Mathematical Biosciences*, 20:293–297, 1974.

[2] Peter Deuflhard and Folkmar Bornemann. *Numerische Mathematik II.* Walter de Gruyter, 2008.

[3] Peter Deuflhard and Andreas Hohmann. *Numerische Mathematik I: eine algorithmisch orientierte Einführung.* Walter de Gruyter, 2002.

[4] Heinz W. Engl, Christoph Flamm, Philipp Kügler, James Lu, Stefan Müller, and Peter Schuster. Inverse problems in systems biology. *Inverse Problems*, 25(12), 2009.

[5] Martin Hanke and Otmar Scherzer. Inverse problems light: Numerical differentiation. *The American Mathematical Monthly*, 108(6):512–521, 2001.

[6] Martin Hanke-Bourgeois. *Grundlagen der numerischen Mathematik und des wissenschaftlichen Rechnens.* Mathematische Leitfäden - Lehrbuch Mathematik. Vieweg, Wiesbaden, 3 edition, 2009.

[7] Harro Heuser. *Gewöhnliche Differentialgleichungen: Einführung in Lehre und Gebrauch.* Teubner Verlag, Stuttgart, 6 edition, 2009.

[8] Harro Heuser. *Lehrbuch der Analysis Teil 2.* B.G. Teubner, Mathematische Leitfäden. Teubner, Stuttgart, 17 edition, 2009.

[9] Klaus Jänich. *Lineare Algebra.* Springer, Berlin, Heidelberg, 11 edition, 2008.

[10] Alfred K. Louis. *Inverse und schlecht gestellte Probleme.* Teubner Studienfächer Mathematik. Teubner, Stuttgart, 1989.

[11] James D. Murray. *Mathematical biology: an introduction.* Springer Verlag, Berlin, Heidelberg, 2002.

[12] Reinhard Schuster. *Grundkurs Biomathematik.* Teubner, Stuttgart, 1995.

[13] Peter Bützer und Markus Roth. Radioaktiver Zerfall: Ba-137m. Website, August 2009. http://www.buetzer.info/fileadmin/pb/pdf-Dateien/Ba-137m.pdf.

[14] Gunilla Wikström. *Computation of Parameters in some Mathematical Models*. PhD thesis, Department of Computing Science, Umeå University, 2002.