Vorlesung Numerische Mathematik 2

Kapitel 9: Ausgleichrechnung

17. Februar 2019

Zürcher Hochschule
für Angewandte Wissenschaften

Gliederung des Kapitels

Einführung

- In der Auswertung von Daten ist man häufig mit dem Problem konfrontiert, Datenpunkte mit einer gewissen Streuung durch eine relativ einfache Funktion anzunähern.
- Im Gegensatzu zur Interpolation sucht man dann eine Funktion, die in einem gewissen Sinne möglichst nahe bei den Datenpunkten durchläuft, ohne diese exakt zu interpolieren.
- In diesem Kapitel werden wir einige dieser Verfahren kennen lernen.

Lernziele:

- Sie können die Begriffe Ausgleichproblem, Ansatzfunktion, Ausgleichsfunktion und Fehlerfunktioinal definieren.
- Sie kennen die Optimierung des Fehlerfunktionals im Sinne der kleinsten Fehlerquadrate (least squares).
- Sie können das lineare sowie das allgemeine Ausgleichproblem definieren und für spzeifische Beispiele lösen.
- Sie können das gedämpfte Gauss-Newton Verfahren anwenden und in Matlab implementieren.

9.1 Zur historischen Entwicklung[1]

[1] Übernommen aus Wikipedia: http://de.wikipedia.org/wiki/Methode_der_kleinsten_Quadrate

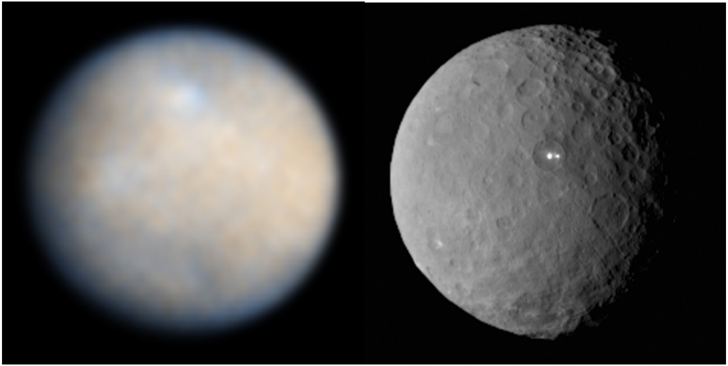
Zur historischen Entwicklung

- Die Ausgleichrechnung geht zurück auf die Methode der kleinsten Quadrate von Gauss (1777-1855).
- Am Neujahrstag 1801 entdeckte der italienische Astronom Giuseppe Piazzi (1746-1826) den Zwergplaneten Ceres. 40 Tage lang konnte er die Bahn verfolgen, dann verschwand Ceres hinter der Sonne.
- Im Laufe des Jahres versuchten viele Wissenschaftler erfolglos, anhand von Piazzis Beobachtungen die Bahn zu berechnen – unter der Annahme einer Kreisbahn, denn nur für solche konnten damals die Bahnelemente aus beobachteten Himmelspositionen mathematisch ermittelt werden.

Zur historischen Entwicklung

Der Zwergplanet Ceres: von der Erde aus (links) und von der Raumsonde Dawn

(NASA, ESA, J. Parker (Southwest Research Institute), P. Thomas (Cornell University), and L. McFadden (University of Maryland, College Park, siehe http://de.wikipedia.org/wiki/(1)_Ceres)



Grössenvergleich von Ceres:

Zur historischen Entwicklung

- Der 24-jährige Gauß hingegen konnte auch elliptische Bahnen aus drei Einzelbeobachtungen berechnen.
- Da aber deutlich mehr Bahnpunkte vorlagen, wandte er seine Methode der kleinsten Quadrate an, um so die Genauigkeit zu erhöhen.
- Als Astronomen im Dezember 1801 den Kleinplaneten genau an dem von Gauß vorhergesagten Ort wiederfanden, war das nicht nur ein großer Erfolg für Gauß' Methode: Piazzis Ruf, der aufgrund seiner nicht zu einer Kreisbahn passen wollenden Bahnpunkte stark gelitten hatte, war ebenfalls wiederhergestellt.

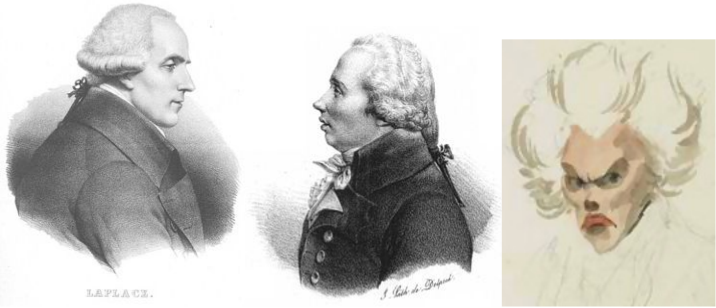
Zur historischen Entwicklung

Carl Friedrich Gauss (1777-1855). Ausschnitt aus einem Gemälde von Gottlieb Biermann (1887).

Zur historischen Entwicklung

- Die Grundlagen seines Verfahrens hatte Gauß schon 1795 im Alter von 18 Jahren entwickelt.
- Basis war eine Idee des französischen Mathematikers Pierre-Simon Laplace (1749-1827), die Beträge von Fehlern aufzusummieren, so dass sich die Fehler zu Null addieren.
- Gauß nahm stattdessen die Fehlerquadrate und konnte die Nullsummen-Anforderung an die Fehler weglassen.
- Unabhängig davon entwickelte der französische Mathematiker Adrien-Marie Legendre (1752-1833) dieselbe Methode erstmals im Jahre 1805 am Schluss eines kleinen Werkes über die Berechnung der Kometenbahnen und veröffentlichte eine zweite Abhandlung darüber im Jahr 1810.
- Von ihm stammt der Name Méthode des moindres carrés (Methode der kleinsten Quadrate).

Zur historischen Entwicklung



Links: Pierre-Simon Laplace (Kupferstich aus dem 19. Jhr.). Mitte: der Politiker Louis Legendre (1752-1797), dessen Bild offenbar während fast 200 Jahren fälschlicherweise für dasjenige des Mathematikers Adrien-Marie Legendre gehalten worden war, bis 2005 der Irrtum entdeckt wurde.Rechts: Karrikatur des Mathematikers Adrien-Marie Legendre (1752-1833).

Zur historischen Entwicklung

- 1809 publizierte Gauß dann im zweiten Band seines himmelsmechanischen Werkes Theoria motus corporum coelestium in sectionibus conicis solem ambientium (Theorie der Bewegung der Himmelskörper, welche in Kegelschnitten die Sonne umlaufen) sein Verfahren, inklusive der Normalgleichungen und des Gaußschen Eliminationsverfahrens.
- Dabei erwähnte er, dass er es schon vor Legendre entdeckt und benutzt habe, was zu einem Prioritätsstreit zwischen den beiden führte.
- Die Methode der kleinsten Quadrate wurde nun schnell das Standardverfahren zur Behandlung von astronomischen oder geodätischen Datensätzen.

Zur historischen Entwicklung

- Gauß benutzte dann das Verfahren intensiv bei seiner Vermessung des Königreichs Hannover durch Triangulation.
- 1821 und 1823 erschien die zweiteilige Arbeit sowie 1826 eine Ergänzung zur Theoria combinationis observationum erroribus minimis obnoxiae (Theorie der den kleinsten Fehlern unterworfenen Kombination der Beobachtungen), in denen Gauß eine Begründung liefern konnte, weshalb sein Verfahren im Vergleich zu den anderen so erfolgreich war: Die Methode der kleinsten Quadrate ist in einer breiten Hinsicht optimal, also besser als andere Methoden.

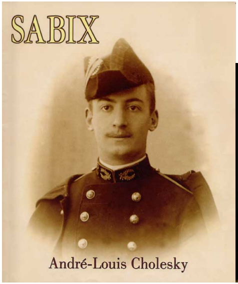
Zur historischen Entwicklung

- Die genaue Aussage ist als der Satz von Gauß-Markow bekannt, da die Arbeit von Gauß wenig Beachtung fand und schließlich im 20. Jahrhundert vom russischen Mathematiker Andrei Andrejewitsch Markow (1856-1922) wiederentdeckt und bekannt gemacht wurde.
- Theoria Combinationis enthält ferner wesentliche Fortschritte beim effizienten Lösen der auftretenden linearen Gleichungssysteme, wie das Gauß-Seidel-Verfahren und die LR-Zerlegung.

Zur historischen Entwicklung

- Der französische Vermessungsoffizier und Mathematiker André-Louis Cholesky (1875-1918) entwickelte während des Ersten Weltkrieges die Cholesky-Zerlegung, die gegenüber den Lösungsverfahren von Gauß nochmal einen erheblichen Effizienzgewinn darstellte.
- In den 1960er Jahren entwickelte der amerikanische Mathematiker Gene Golub (1932-2007) die Idee, die auftretenden linearen Gleichungssysteme mittels QR-Zerlegung zu lösen.

Zur historischen Entwicklung



Links: Andrei Andrejewitsch Markow (1856-1922). Mitte: André-Louis Cholesky (1875-1918). Rechts: Gene Golub (1932-2007).

9.2 Problemstellung

Problemstellung

- Im Unterschied zur Interpolation versuchen wir bei der Ausgleichsrechnung nicht, eine Funktion f zu finden, die exakt durch sämtliche Wertepaare geht, sondern diese nur möglichst gut approximiert.
- Dies ist insbesondere dann sinnvoll, wenn es eine grosse Anzahl von Datenpunkten gibt (die meist zusätzlich noch fehlerbehaftet sind) und die durch eine Funktion mit nur wenigen Parametern beschrieben werden sollen.

Problemstellung

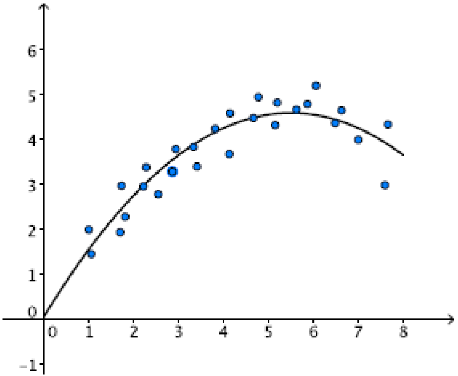


Abbildung: Beispiel einer "Wolke" von Datenpunkten und einer quadratischen Näherungsfunktion [10].

Problemstellung: Definition Ausgleichproblem

Definition 9.1: Ausgleichproblem [1]

- Gegeben sind n Wertepaare (x_i, y_i), $i = 1, \dots, n$ mit $x_i \neq x_j$ für $i \neq j$.
- Gesucht ist eine stetige Funktion $f : \mathbb{R} \longrightarrow \mathbb{R}$, die die Wertepaare in einem gewissen Sinn bestmöglich annähert, d.h. dass möglichst genau gilt

$$f(x_i) \approx y_i$$

für alle $i = 1, \dots, n$.

Problemstellung: Definition Fehlerfunktional

Definition 9.2: Ansatzfunktionen / Ausgleichsfunktion / Fehlerfunktional / kleinste Fehlerquadrate [1]

- Gegeben sei eine Menge F von stetigen **Ansatzfuntkionen** f auf dem Intervall $[a,b]$ sowie n Wertepaare (x_i, y_i), $i = 1, \ldots, n$.
- Ein Element $f \in F$ heisst **Ausgleichsfunktion** von F zu den gegebenen Wertepaaren, falls das **Fehlerfunktional**

$$E(f) := \| \boldsymbol{y} - f(\boldsymbol{x}) \|_2^2 = \sum_{i=1}^{n} (y_i - f(x_i))^2$$

für f minimal wird, d.h. $E(f) = \min\{E(g) \mid g \in F\}$.
- Man nennt das so gefundene f dann optimal im Sinne der **kleinsten Fehlerquadrate** (least squares fit).

Problemstellung: Bemerkungen

- Diese Forderung der kleinsten Fehlerquardrate bedeutet nichts anderes, als dass das Quadrat der 2-Norm (vgl. Kap. 4.6.1) des Fehlervektors

$$\begin{pmatrix} y_1 - f(x_1) \\ \vdots \\ y_n - f(x_n) \end{pmatrix}$$

minimal sein soll. Andere Normen sind auch möglich, was zu anderen Ergebnissen führen kann. Die in der Praxis am häufigsten verwendete Norm ist allerdings die 2-Norm.

Problemstellung: Bemerkungen

- Es ist möglich, die Fehler noch mit Gewichten $w_i > 0$ zu versehen, d.h. man minimiert

$$\sum_{i=1}^{n} w_i \cdot (y_i - f(x_i))^2 .$$

- Wenn z.B. gewisse Datenpunkte (x_i, y_i) grosse Messungenauigkeiten aufweisen, können diese mit einem kleineren Gewicht versehen werden, als Datenpunkte mit höherer Messgenauigkeit. Die Summe der Gewichte sollte dabei normiert sein, also $\sum_{i=1}^{n} w_i = 1$.

Problemstellung: Bemerkungen

- Wählt man F als Menge aller Geraden, dann nennt man das so gefundene f *Ausgleichsgerade* oder in statistischen Zusammenhängen auch *Regressionsgerade*.

9.3 Lineare Ausgleichprobleme

Lineare Ausgleichprobleme

- Gegeben sei die Wertetabelle

x_i	x_1	x_2	$\dots$	x_n
y_i	y_1	y_2	$\dots$	y_n

- Wir suchen die Ausgleichsgerade der Form $f(x) = ax + b$, also $F := \{a_1 f_1 + a_2 f_2 \mid a_1, a_2 \in \mathbb{R}\}$ mit den Ansatzfunktionen $f_1(x) = x$ und $f_2(x) = 1$.
- Das Fehlerfunktional hat dann die Form

$$E(f)(a,b) := E(f) = \sum_{i=1}^{n} (y_i - f(x_i))^2 = \sum_{i=1}^{n} (y_i - (ax_i + b))^2 .$$

Lineare Ausgleichprobleme

- Dieses soll minimal werden, d.h. die partiellen Ableitung nach den Parametern a und b müssen verschwinden (in Analogie zum eindimensionalen Fall):

$$
\begin{aligned}
0 &\stackrel{!}{=} \frac{\partial E}{\partial a} = \sum_{i=1}^{n} \frac{\partial}{\partial a} (y_i - (ax_i + b))^2 \\
&= \sum_{i=1}^{n} 2 \cdot (y_i - (ax_i + b)) \cdot (-x_i) \\
&= -2 \sum_{i=1}^{n} (y_i - (ax_i + b)) \cdot (x_i) \\
0 &\stackrel{!}{=} \frac{\partial E}{\partial b} = \sum_{i=1}^{n} \frac{\partial}{\partial b} (y_i - (ax_i + b))^2 \\
&= \sum_{i=1}^{n} 2 \cdot (y_i - (ax_i + b)) \cdot (-1) \\
&= -2 \sum_{i=1}^{n} (y_i - (ax_i + b))
\end{aligned}
$$

Lineare Ausgleichprobleme

- Wir haben also zwei Gleichungen für die Unbekannten a, b. Durch Umformung erhalten wir

$$\begin{aligned}
0 &\stackrel{!}{=} \sum_{i=1}^{n} (y_i - (ax_i + b)) \cdot (x_i) = \sum_{i=1}^{n} \left(x_i y_i - a x_i^2 - b x_i\right) \\
&= \sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} a x_i^2 - \sum_{i=1}^{n} b x_i = \sum_{i=1}^{n} x_i y_i - a \sum_{i=1}^{n} x_i^2 - b \sum_{i=1}^{n} x_i \\
0 &\stackrel{!}{=} \sum_{i=1}^{n} (y_i - (ax_i + b)) = \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} a x_i - \sum_{i=1}^{n} b \\
&= \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} a x_i - \sum_{i=1}^{n} b = \sum_{i=1}^{n} y_i - a \sum_{i=1}^{n} x_i - b \sum_{i=1}^{n} 1
\end{aligned}$$

Lineare Ausgleichprobleme

- Daraus folgt

$$\begin{aligned} a\sum_{i=1}^n x_i^2 + b\sum_{i=1}^n x_i &= \sum_{i=1}^n x_i y_i \\ a\sum_{i=1}^n x_i + b\underbrace{\sum_{i=1}^n 1}_{=n} &= \sum_{i=1}^n y_i \end{aligned}$$

oder in Matrixschreibweise erhalten wir das lineare Gleichungssystem für die Unbekannten (a, b)

$$\begin{pmatrix} \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & n \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n x_i y_i \\ \sum_{i=1}^n y_i \end{pmatrix}$$

und können dieses nach a, b auflösen.

Beispiel 9.1

- Wir bestimmen die Ausgleichsgerade $f(x) = ax + b$ für die folgende Wertetabelle:

x_i	1	2	3	4
y_i	6	6.8	10	10.5

Beispiel 9.1: Lösung

- Mit $n = 4$ und

$$\sum_{i=1}^{4} x_i^2 = 30, \ \sum_{i=1}^{4} x_i = 10, \ \sum_{i=1}^{4} x_i y_i = 91.6, \ \sum_{i=1}^{4} y_i = 33.3$$

erhalten wir das lineare Gleichungssystem

$$\begin{pmatrix} 30 & 10 \\ 10 & 4 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 91.6 \\ 33.3 \end{pmatrix}$$

mit der Lösung

$$\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 1.67 \\ 4.15 \end{pmatrix}.$$

Also ist die gesuchte Ausgleichsgerade $f(x) = 1.67x + 4.15$ (siehe folgende Abb.).

Beispiel 9.1: Lösung

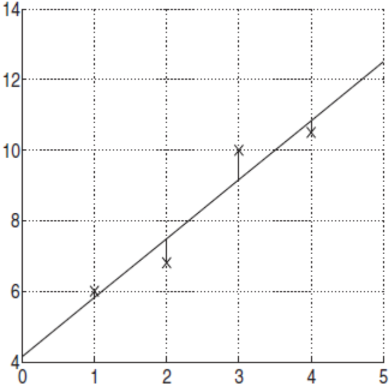


Abbildung: Ausgleichsgerade $f(x) = 1.67x + 4.15$ zu Bsp. 9.1 (aus [1]).

Lineare Ausgleichprobleme: Definition 9.3

Definition 9.3: Lineares Ausgleichproblem [1]

- Gegeben seien n Wertepaare $(x_i, y_i), i = 1, \dots, n$, und m Basisfunktionen $f_1, \dots, f_m$ auf einem Intervall $[a, b]$. Wir wählen F als die Menge der Ansatzfunktionen $f := \lambda_1 f_1 + \dots + \lambda_m f_m$ mit $\lambda_j \in \mathbb{R}$, also $F = \{ f = \lambda_1 f_1 + \dots + \lambda_m f_m \mid \lambda_j \in \mathbb{R},\ j = 1, \dots, m \}$.
- Es liegt dann ein **lineares Ausgleichproblem** vor mit dem Fehlerfunktional

$$\begin{aligned} E(f) &= \| \boldsymbol{y} - f(\boldsymbol{x}) \|_2^2 = \sum_{i=1}^{n} (y_i - f(x_i))^2 \\ &= \sum_{i=1}^{n} \left(y_i - \sum_{j=1}^{m} \lambda_j f_j(x_i) \right)^2 = \| \boldsymbol{y} - \boldsymbol{A}\lambda \|_2^2 \end{aligned}$$

$$\boldsymbol{A} = \begin{pmatrix} f_1(x_1) & f_2(x_1) & \cdots & f_m(x_1) \\ f_1(x_2) & f_2(x_2) & \cdots & f_m(x_2) \\ \vdots & \vdots & \vdots & \vdots \\ f_1(x_n) & f_2(x_n) & \cdots & f_m(x_n) \end{pmatrix}, \ \boldsymbol{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \ \lambda = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_m \end{pmatrix}$$

- Das System $\boldsymbol{A}\lambda = \boldsymbol{y}$ heisst **Fehlergleichungssystem**.

Lineare Ausgleichprobleme: Bemerkungen

- Das Fehlergleichungssystem besitzt n Gleichungen und m Unbekannte.
- In der Regel ist $n > m$, wir haben also mehr Gleichungen als Unbekannte, d.h. das Gleichungssystem ist überbestimmt.
- Man kann daher nicht damit rechnen, dass es eine Lösung gibt (bzw. dass das Fehlerfunktional $E(f) = 0$ erreicht werden kann).

Lineare Ausgleichprobleme: Bemerkungen

- Für $m = n$ haben wir eine eindeutige Lösung und $E(f) = 0$.
- Dies ist gleichbedeutend damit, dass die Ausgleichfunktion f exakt durch sämtliche Wertepaare geht.
- Dies ist der Spezialfall der Interpolation.

Lineare Ausgleichprobleme: Normalgleichungen

- Wir können nun versuchen, das Fehlerfunktional zu minimieren. Dafür müssen sämtliche partiellen Ableitungen von $E(f)$ nach den unbekannten Koeffizienten λ_j verschwinden, also
$$\frac{\partial E(f)}{\partial \lambda_j} = 0.$$
- Dies führt uns zu den Begriffen der Normalgleichungen und des Normalgleichungssystem.

Definition 9.4: Normalgleichungen / Normalgleichungssystem [1]

- Die Gleichungen

$$\frac{\partial E(f)(\lambda_1, \ldots, \lambda_m)}{\partial \lambda_j} = 0, \quad j = 1, \ldots, m$$

  heissen **Normalgleichungen** des linearen Ausgleichproblems.

- Das System sämtlicher Normalgleichungen heisst **Normalgleichungssystem** und lässt sich als lineares Gleichungssystem schreiben

$$\boldsymbol{A}^T \boldsymbol{A} \lambda = \boldsymbol{A}^T \boldsymbol{y}$$

Lineare Ausgleichprobleme: Normalgleichungen

Bemerkungen [1]:

- Die Lösungen des Normalgleichungssystem sind die gesuchten Parameter des linearen Ausgleichproblems.
- Die $m \times m$ Matrix $\boldsymbol{A}^T\boldsymbol{A}$ ist symmetrisch und positiv definit, das Normalgleichungssystem kann deshalb mit der Cholesky-Zerlegung (Kap. 4.5.2) gelöst werden.
- Numerisch stabiler ist die sogn. QR- Zerlegung, auf die wir hier aber nicht eingehen werden.

Beispiel 9.2

- Stellen Sie für das lineare Ausgleichproblem in Bsp. 9.1 das Normalgleichungssystem auf.

Beispiel 9.2: Lösung

- Mit $f_1(x) = x$ und $f_2(x) = 1$ erhalten wir

$$\boldsymbol{A} = \begin{pmatrix} f_1(x_1) & f_2(x_1) \\ f_1(x_2) & f_2(x_2) \\ f_1(x_3) & f_2(x_3) \\ f_1(x_4) & f_2(x_4) \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 4 & 1 \end{pmatrix}$$

$$\Rightarrow \boldsymbol{A}^T\boldsymbol{A} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 4 & 1 \end{pmatrix} = \begin{pmatrix} 30 & 10 \\ 10 & 4 \end{pmatrix}$$

$$\boldsymbol{A}^T\boldsymbol{y} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 6 \\ 6.8 \\ 10 \\ 10.5 \end{pmatrix} = \begin{pmatrix} 91.6 \\ 33.3 \end{pmatrix}$$

Damit entspricht $\boldsymbol{A}^T\boldsymbol{A}\lambda = \boldsymbol{A}^T\boldsymbol{y}$ aus Bsp. 9.1.

Beispiel 9.3

- Bestimmen Sie mit den bisher behandelten Techniken eine Funktion der Form $f(x) = ae^{bx}$, die die folgenden Daten möglichst gut approximiert:

x_i	0	1	2	3	4
y_i	3	1	0.5	0.2	0.05

Beispiel 9.3: Lösung

- Auf den ersten Blick liegt hier kein lineares Ausgleichproblem vor, da wir $f(x)$ nicht als Summe $af_1(x)+bf_2(x)$ schreiben können.
- Wir können aber durch beidseitiges Logarithmieren einen linearen Ansatz erreichen:

$$\ln f(x) = \ln\left(ae^{bx}\right) = \ln(a) + b \cdot \ln\left(e^x\right) = \ln(a) \cdot \underbrace{1}_{f_1(x)} + b \cdot \underbrace{x}_{f_2(x)}$$

- Wir werden durch die Lösung des Normalgleichungssystems also $\ln(a)$ und b berechnen können, wenn wir die y_i logarithmieren.

Beispiel 9.3: Lösung

$$
\begin{aligned}
\boldsymbol{A}^T\boldsymbol{A} &= \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 & 4 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{pmatrix} = \begin{pmatrix} 5 & 10 \\ 10 & 30 \end{pmatrix} \\
\boldsymbol{A}^T \cdot \tilde{\boldsymbol{y}} &= \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 & 4 \end{pmatrix} \begin{pmatrix} \ln 3 \\ \ln 1 \\ \ln 0.5 \\ \ln 0.2 \\ \ln 0.05 \end{pmatrix} = \begin{pmatrix} -4.1997 \\ -18.1975 \end{pmatrix} \\
\boldsymbol{A}^T\boldsymbol{A}\lambda &= \boldsymbol{A}^T\tilde{\boldsymbol{y}} \Rightarrow \lambda = \begin{pmatrix} \ln a \\ b \end{pmatrix} = \begin{pmatrix} 1.11968\ldots \\ -0.97981\ldots \end{pmatrix} \\
\Rightarrow f(x) &= \mathrm{e}^{\ln a} \cdot \mathbf{e}^{bx} = 3.06388\ldots \cdot \mathbf{e}^{-0.97981\ldots x}
\end{aligned}
$$

Wir werden dieses Beispiel im nächsten Abschnitt bei nichtlinearen Ausgelichsproblemen gleich nochmals aufgreifen.

Beispiel 9.3: Lösung

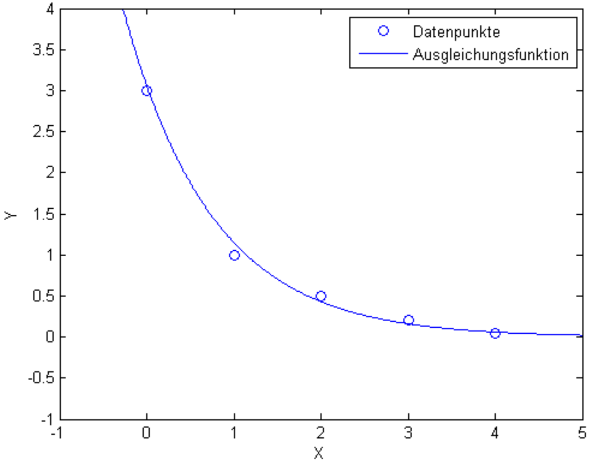


Abbildung: Ausgleichsfunktion $f(x) = 3.0638 \cdot e^{-0.9798x}$ zu Bsp. 9.3.

Aufgabe 9.1

- Bestimmen Sie für die folgenden Daten die Funktion der Form $f(x) = ae^x + b$, die diese Daten bestmöglich im Sinner der kleinsten Fehlerquadrate approximiert:

x_i	0	1	2	3	4
y_i	6	12	30	80	140

Aufgabe 9.1: Lösung

Aufgabe 9.1: Lösung

Aufgabe 9.1: Lösung

9.4 Nichtlineare Ausgleichprobleme

Nichtlineare Ausgleichprobleme

- Wir haben in Beispiel 9.3 gesehen, dass man allenfalls in Spezialfällen ein nichtlineares Ausgleichsproblem in ein lineares Ausgleichproblem umformen kann.
- Trotzdem wollen wir natürlich in der Lage sein, allgemeine nichtlineare Ausgleichprobleme zu lösen.

Nichtlineare Ausgleichprobleme

Definition 9.5: Allgemeines Ausgleichproblem [1]

- Gegeben seien n Wertepaare $(x_i, y_i), i = 1, \dots, n$, und die Menge F der Ansatzfunktionen $f_p = f_p(\lambda_1, \lambda_2, \dots, \lambda_m, x)$ mit m Parametern $\lambda_j \in \mathbb{R}$, $j = 1, \dots, m$, also $F = \{f_p(\lambda_1, \lambda_2, \dots, \lambda_m, x) \mid \lambda_j \in \mathbb{R},\ j = 1, \dots, m\}$.
- Das **allgemeine Ausgleichproblem** besteht darin, die m Parameter $\lambda_1, \dots, \lambda_m$ zu bestimmen, so dass das Fehlerfunktional E

$$\begin{aligned} E(f) &= \sum_{i=1}^{n} \left(y_i - f_p(\lambda_1, \lambda_2, \dots, \lambda_m, x_i)\right)^2 = \left\| \begin{pmatrix} y_1 - f_p(\lambda_1, \lambda_2, \dots, \lambda_m, x_1) \\ y_2 - f_p(\lambda_1, \lambda_2, \dots, \lambda_m, x_2) \\ \vdots \\ y_n - f_p(\lambda_1, \lambda_2, \dots, \lambda_m, x_n) \end{pmatrix} \right\|_2^2 \\ &= \| \mathbf{y} - \mathbf{f}(\lambda) \|_2^2 \end{aligned}$$

minimal wird unter allen zulässigen Parameterbelegungen.

Definition 9.5: Fortsetzung

- Dabei ist

$$
\begin{aligned}
\mathbf{f}(\lambda) &= \mathbf{f}(\lambda_1, \lambda_2, \ldots, \lambda_m) := \begin{pmatrix} f_1(\lambda_1, \lambda_2, \ldots, \lambda_m) \\ f_2(\lambda_1, \lambda_2, \ldots, \lambda_m) \\ \vdots \\ f_n(\lambda_1, \lambda_2, \ldots, \lambda_m) \end{pmatrix} \\
&= \begin{pmatrix} f_p(\lambda_1, \lambda_2, \ldots, \lambda_m, x_1) \\ f_p(\lambda_1, \lambda_2, \ldots, \lambda_m, x_2) \\ \vdots \\ f_p(\lambda_1, \lambda_2, \ldots, \lambda_m, x_n) \end{pmatrix} \\
\boldsymbol{y} &= \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_{\boldsymbol{n}} \end{pmatrix}, \ \lambda = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_{\boldsymbol{m}} \end{pmatrix}
\end{aligned}
$$

Nichtlineare Ausgleichprobleme: Bemerkungen

- Falls die Ansatzfunktionen f_p linear in den Parametern sind, haben wir den Spezialfall des linearen Ausgleichsproblems mit $\boldsymbol{f}(\lambda) = \boldsymbol{A}\lambda$ (vgl. Def. 9.4).

Nichtlineare Ausgleichprobleme: Bemerkungen

- Das allg. Ausgleichproblem ist also äquivalent zur Bestimmung des Minimums einer Funktion

$$E : \mathbb{R}^m \longrightarrow \mathbb{R}$$

  und wir könnten wieder die Normalgleichungen aufstellen , indem wir die partiellen Ableitungen von f nach den Parametern λ_i gleich Null setzen und das entstehende, diesmal nicht lineare Gleichungssystem lösen.
- Wir werden im folgenden Beispiel dies einmal ansatzweise machen und werden dann aber mit Gauss-Newton-Verfahren im nächsten Abschnitt 9.5 ein anderes, effektiveres Verfahren kennen lernen.

Beispiel 9.4

- Wir nehmen wieder das Problem aus Beispiel 9.3, d.h. wir wollen eine Ansatzfunktion $f(x) = ae^{bx}$ bestmöglich im Sinn der kleinsten Fehlerquadrate an die gegebenen Daten anpassen:

x_i	0	1	2	3	4
y_i	3	1	0.5	0.2	0.05

Beispiel 9.4: Lösung

- Wir haben also zwei Parameter a und b in der Ansatzfunktion $f_p(a, b, x) = ae^{bx}$ so zu bestimmen, dass

$$E(f) = \sum_{i=1}^{5} (y_i - f_p(a, b, x_i))^2 = \sum_{i=1}^{5} (y_i - ae^{bx_i})^2$$

minimal wird.

Beispiel 9.4: Lösung

- Nullsetzen der partiellen Ableitungen liefert

$$
\begin{aligned}
0 &= \frac{\partial E(f)}{\partial a} = \sum_{i=1}^{5} \frac{\partial}{\partial a}(y_i - ae^{bx_i})^2 \\
&= \sum_{i=1}^{5} 2 \cdot (y_i - ae^{bx_i}) \cdot (-e^{bx_i}) \\
&= -2 \sum_{i=1}^{5} (y_i - ae^{bx_i}) \cdot e^{bx_i} \\
0 &= \frac{\partial E(f)}{\partial b} = \sum_{i=1}^{5} \frac{\partial}{\partial b}(y_i - ae^{bx_i})^2 \\
&= \sum_{i=1}^{5} 2 \cdot (y_i - ae^{bx_i}) \cdot (-ae^{bx_i}) \cdot x_i \\
&= -2 \sum_{i=1}^{5} (y_i - ae^{bx_i}) \cdot ae^{bx_i} \cdot x_i
\end{aligned}
$$

Beispiel 9.4: Lösung

- Dies sind zwei nichtlineare Gleichungen, die mit dem Newton-Verfahren für Systeme (vlg. Kap. 5) gelöst werden können.
- Wegen der Länge der Ausdrücke verzichten wir hier an dieser Stelle auf die Details , ein funktionierendes Matlab Skript ist unten angegeben.
- Für die Startwerte $a_0 = 3$ und $b_0 = -1$ und einer Fehlerschranke von 10^{-5} liefert es nach zweimaliger Iteration

$$a = 2.98165..., \; b = -1.00328...$$

- Zum Vergleich: in Bsp. 9.3 hatten wir durch Logarithmieren das nichtlineare Ausgleichsproblem auf ein lineares zurückgeführt und erhielten die Parameter

$$a = 3.06388..., \; b = -0.97981...$$

Beispiel 9.4: Lösung

- Der Unterschied wird nicht durch Rechenungenauigkeiten verursacht, tatsächlich handelt es sich um zwei verschiedene Modellanpasungen.

Beispiel 9.4: Lösung

```
% Matlab Skript zu Bsp. 9.4
x = [0:1:4]; y = [3 1 0.5 0.2 0.05];
syms a b
f = @(x) a*exp(b*x);
sum1 = 0;
for i=1:5
sum1 = sum1+(y(i)-a*exp(b*x(i)))*exp(b*x(i));
end
sum1 = -2*sum1;
sum2 = 0;
for i=1:5
sum2 = sum2+(y(i)-a*exp(b*x(i)))*a*exp(b*x(i))*x(i);
end
sum2 = -2*sum2;

Df = jacobian([f1,f2],[a,b]);
Df = matlabFunction(Df);
f1 = matlabFunction(f1);
f2 = matlabFunction(f2);
fvec = @(x) [f1(x(1),x(2));f2(x(1),x(2))];
Dfvec = @(x) Df(x(1),x(2));

tol = 1e-5; x0 = [3,-1]'; xn = x0; err = tol + 1; n = 0;
while err > tol
Dx = -Dfvec(xn)\fvec(xn);
xn = xn + Dx;
err = norm(fvec(xn),2);
n = n+1;
end
```

Beispiel 9.4: Lösung

- Das obige Verfahren führt zwar zu einem Ergebnis, hat jedoch mehrere Nachteile.
- Die partiellen Ableitungen des Fehlerfunktionals müsen bekannt sein, und es konvergiert nur, wenn man genügend gute Anfangsnäherungen für die Parameter zur Verfügung hat, was häufig nicht der Fall ist.
- So würden wir mit einem Startwert $x_0 = [1, -1]$ im obigen Skript das Resultat $a = 3.00000...$, $b = -13.60235...$ erhalten.
- Das folgende iterative Verfahren hat diesbezüglich einige Vorteile.

9.5 Gauss- Newton-Verfahren

Gauss-Newton-Verfahren

- Wir defnieren:

Definition 9.6: Quadratmittelproblem [1]

- Gegeben ist eine Funktion $\boldsymbol{g} : \mathbb{R}^m \longrightarrow \mathbb{R}^n$ und das zugehörige Fehlerfunktional $E : \mathbb{R}^m \longrightarrow \mathbb{R}$, definiert durch $E(\boldsymbol{x}) := \| \boldsymbol{g}(\boldsymbol{x}) \|_2^2$. Das Problem, einen Vektor $\boldsymbol{x} \in \mathbb{R}^m$ zu finden, für den $E(\boldsymbol{x})$ minimal wird, nennt man Quadratmittelproblem.

Gauss-Newton-Verfahren

- Nichtlineare Ausgleichsprobleme sind also Quadratmittelprobleme, wobei wir $\boldsymbol{g}(\mathbf{x})$ ersetzen mit $\boldsymbol{g}(\lambda)$:

$$\boldsymbol{g}(\lambda) := \boldsymbol{y} - \boldsymbol{f}(\lambda), \ \mathbf{y} \in \mathbb{R}^n, \lambda \in \mathbb{R}^m$$

  wie in Def. 9.5.

- Das Fehlerfunktional lautet entsprechend

$$E(\lambda) := \| \boldsymbol{g}(\lambda) \|_2^2 = \| \boldsymbol{y} - \boldsymbol{f}(\lambda) \|_2^2$$

  und der Vektor $\lambda = (\lambda_1, ..., \lambda_m)^T$ beinhaltet die gesuchten Parameter.

Gauss-Newton-Verfahren

- Das folgende Gauss-Newton-Verfahren besteht aus einer Kombination von linearer Ausgleichrechnung und dem Newton-Verfahren.
- Wir ersetzen dabei in $\| \boldsymbol{g}(\lambda) \|_2^2$ den Vektor $\boldsymbol{g}(\lambda)$ durch die Linearkombination

$$\boldsymbol{g}(\lambda) \approx \boldsymbol{g}(\lambda^{(0)}) + \boldsymbol{Dg}(\lambda^{(0)}) \cdot (\lambda - \lambda^{(0)})$$

Dies entspricht der 'verallgemeinerten Tangentengleichung' aus Def. 5.3 mit der $n \times m$ Jacobi-Matrix $\boldsymbol{Dg}(\lambda^{(0)})$ in der Umgebung eines Startwertes $\lambda^{(0)}$:

$$\boldsymbol{Dg}(\lambda^{(0)}) := \begin{pmatrix} \frac{\partial g_1}{\partial \lambda_1}(\lambda^{(0)}) & \frac{\partial g_1}{\partial \lambda_2}(\lambda^{(0)}) & \cdots & \frac{\partial g_1}{\partial \lambda_m}(\lambda^{(0)}) \\ \frac{\partial g_2}{\partial \lambda_1}(\lambda^{(0)}) & \frac{\partial g_2}{\partial \lambda_2}(\lambda^{(0)}) & \cdots & \frac{\partial g_2}{\partial \lambda_m}(\lambda^{(0)}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_n}{\partial \lambda_1}(\lambda^{(0)}) & \frac{\partial g_n}{\partial \lambda_2}(\lambda^{(0)}) & \cdots & \frac{\partial g_n}{\partial \lambda_m}(\lambda^{(0)}) \end{pmatrix}$$

Gauss-Newton-Verfahren

- Das heisst, wir suchen statt des Minimums des ursprünglichen Fehlerfunktionals

$$E(\lambda) \quad = \quad \| \boldsymbol{g}(\lambda) \|_2^2$$

nun das Minimum des 'linearisierten Fehlerfunktionals'

$$\tilde{E}(\lambda) \quad = \quad \| \boldsymbol{g}(\lambda^{(0)}) + \boldsymbol{Dg}(\lambda^{(0)}) \cdot (\lambda - \lambda^{(0)}) \|_2^2$$

und haben so das nichtlineare Ausgleichproblem zurückgeführt auf ein lineares Ausgleichproblem.

Gauss-Newton-Verfahren

- Mit den Subsitutionen

$$\begin{aligned}\widetilde{\boldsymbol{y}} &:= \boldsymbol{g}(\lambda^{(0)})\\ \widetilde{\boldsymbol{A}} &:= -\boldsymbol{Dg}(\lambda^{(0)})\\ \delta &:= \lambda - \lambda^{(0)}\end{aligned}$$

haben wir die gleiche Form wie in Def. 9.3

$$\tilde{E}(\lambda) = \| \widetilde{\boldsymbol{y}} - \widetilde{A}\delta \|_2^2$$

und die Lösung berechnet sich gemäss Def. 9.4

$$\widetilde{\boldsymbol{A}}^T \widetilde{\boldsymbol{A}} \delta = \widetilde{\boldsymbol{A}}^T \widetilde{\boldsymbol{y}}$$

bzw.

$$\boldsymbol{Dg}(\lambda^{(0)})^T \boldsymbol{Dg}(\lambda^{(0)})\delta = -\boldsymbol{Dg}(\lambda^{(0)})^T \cdot \boldsymbol{g}(\lambda^{(0)}).$$

Die Lösung dieses linearen Gleichungssystems liefert uns $\delta = \lambda - \lambda^{(0)}$ und damit können wir nach der eigentlich gesuchten Grösse

$$\lambda = \lambda^{(0)} + \delta$$

auflösen.

Gauss-Newton-Verfahren

- Unter der Annahme, dass λ eine bessere Näherung ist als der Startwert $\lambda^{(0)}$, können wir $\lambda^{(1)} := \lambda$ setzen und diesen Schritt durch rekursives Einsetzen wiederholen. Damit erhalten wir das Gauss-Newton Verfahren:

Definition 9.7: Gauss-Newton-Verfahren [1]

- Sei $\lambda^{(0)}$ ein Startvektor in der Nähe des Minimums von E. Das Gauss-Newton-Verfahren zur näherungsweisen Bestimmung des Minimums lautet:
- Für $k = 0, 1, \ldots$:
  1. Berechne $\delta^{(k)}$ als Lösung des linearen Ausgleichproblems
     $$\min \| \boldsymbol{g}(\lambda^{(k)}) + \boldsymbol{Dg}(\lambda^{(k)}) \cdot \delta^{(k)} \|_2^2,$$
     d.h. löse konkret:
     $$\boldsymbol{Dg}(\lambda^{(k)})^T \boldsymbol{Dg}(\lambda^{(k)}) \delta^{(k)} = -\boldsymbol{Dg}(\lambda^{(k)})^T \cdot \boldsymbol{g}(\lambda^{(k)})$$
     nach $\delta^{(k)}$ auf.
  2. Setze
     $$\lambda^{(k+1)} = \lambda^{(k)} + \delta^{(k)}.$$

Gauss-Newton-Verfahren

- Die Grösse $\delta^{(k)}$ kann als 'Korrekturrichtung' der Näherung $\lambda^{(k)}$ aufgefasst werden.
- In der Praxis verwendet man besser das folgende gedämpfte Gauss-Newton-Verfahren.
- Analog zum gedämpften Newton-Verfahren aus Kapitel 5.4.3 akzeptieren wir die Korrekutr $\delta^{(k)}$ nur, wenn sie wirklich zu einer Abnahem des Fehlerfunktionals führt, also

$$E(\lambda^{(k+1)}) = \| \boldsymbol{g}(\lambda^{(k+1)}) \|_2^2 < \| \boldsymbol{g}(\lambda^{(k)}) \|_2^2 = E(\lambda^{(k)}).$$

erfüllt ist. Dies können wir durch die Dämpfung wie folgt anstreben:

Definition 9.8: Gedämpftes Gauss-Newton-Verfahren [1]

- Sei $\lambda^{(0)}$ ein Startvektor in der Nähe des Minimums von E. Das gedämpfte Gauss-Newton-Verfahren zur näherungsweisen Bestimmung des Minimums lautet:
- Für $k = 0, 1, \ldots$:
  1. Berechne $\delta^{(k)}$ als Lösung des linearen Ausgleichproblems
  $$\min \| \boldsymbol{g}(\lambda^{(k)}) + \boldsymbol{Dg}(\lambda^{(k)}) \cdot \delta^{(k)} \|_2^2,$$
  d.h. als Lösung des linearen Gleichungssystems
  $$\boldsymbol{Dg}(\lambda^{(k)})^T \boldsymbol{Dg}(\lambda^{(k)}) \delta^{(k)} = -\boldsymbol{Dg}(\lambda^{(k)})^T \cdot \boldsymbol{g}(\lambda^{(k)})$$
  2. Finde das minimale $p \in \{0, 1, \ldots, p_{max}\}$ mit
  $$\| \boldsymbol{g}\underbrace{\left(\lambda^{(k)} + \frac{\delta^{(k)}}{2^p}\right)}_{\lambda^{(k+1)}} \|_2^2 < \| \boldsymbol{g}\left(\lambda^{(k)}\right) \|_2^2$$
  3. Falls kein minimales p gefunden werden kann, rechne mit $p = 0$ weiter
  4. Setze
  $$\lambda^{(k+1)} = \lambda^{(k)} + \frac{\delta^{(k)}}{2^p}.$$

Bemerkungen:

- Ein Vorteil des gedämpften Gauss-Newton-Verfahrens ist, dass es i.d.R. für einen grösseren Bereich von Startvektoren konvergiert als das ungedämpfte Gauss-Newton-Verfahren.
- Durch das fortlaufende Anpassen der 'Korrekturrichtung' kann es sozusagen noch mit Startvektoren umgehen, die weiter entfernt sind vom Optimum.
- Die Dämpfung ist aber keine Garantie für Konvergenz. Man benötigt noch weitere Bedingungen, auf die wir hier nicht eingehen werden.

Gauss-Newton-Verfahren

Bemerkungen:

- Als Abbruchkriterium des Algorithmus kann z.B.

$$\| \frac{\delta^{(k)}}{2^p} \|_2 < TOL$$

  verwendet werden. Wiederum ist das aber keine Garantie, dass die berechnete Näherung einen maximalen Abstand von TOL zum gesuchten Minimum besitzt.

Beispiel 9.5

- Wir nehmen nochmals das Problem aus Beispiel 9.4.
- Wir wollen also eine Ansatzfunktion $f(x) = ae^{bx}$ bestmöglich im Sinn der kleinsten Fehlerquadrate an die gegebenen Daten anpassen und vergleichen die Resultate des ungedämpften und des gedämpften Gauss-Newton-Verfahrens, einmal für den Startvektor $\lambda^{(0)} = (a^{(0)}, b^{(0)})^T = (1, -1.5)^T$ und dann nochmals für den Startvektor $\lambda^{(0)} = (a^{(0)}, b^{(0)})^T = (2, 2)^T$

x_i	0	1	2	3	4
y_i	3	1	0.5	0.2	0.05

Beispiel 9.5: Lösung

$$\boldsymbol{g}(a,b) := \boldsymbol{y} - \boldsymbol{f}(a,b) = \begin{pmatrix} g_1(a,b) \\ \vdots \\ g_5(a,b) \end{pmatrix} = \begin{pmatrix} y_1 - a\mathbf{e}^{bx_1} \\ \vdots \\ y_5 - a\mathbf{e}^{bx_5} \end{pmatrix} = \begin{pmatrix} 3 - a\mathbf{e}^{b\cdot 0} \\ 1 - a\mathbf{e}^{b\cdot 1} \\ 0.5 - a\mathbf{e}^{b\cdot 2} \\ 0.2 - a\mathbf{e}^{b\cdot 3} \\ 0.05 - a\mathbf{e}^{b\cdot 4} \end{pmatrix}$$

$$\boldsymbol{Dg}(a,b) := \begin{pmatrix} \frac{\partial g_1}{\partial a}(a,b) & \frac{\partial g_1}{\partial b}(a,b) \\ \frac{\partial g_2}{\partial a}(a,b) & \frac{\partial g_2}{\partial b}(a,b) \\ \vdots & \vdots \\ \frac{\partial g_5}{\partial a}(a,b) & \frac{\partial g_5}{\partial b}(a,b) \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ -\mathbf{e}^{b\cdot 1} & -a\mathbf{e}^{b\cdot 1} \\ -\mathbf{e}^{b\cdot 2} & -2a\mathbf{e}^{b\cdot 2} \\ -\mathbf{e}^{b\cdot 3} & -3a\mathbf{e}^{b\cdot 3} \\ -\mathbf{e}^{b\cdot 4} & -4a\mathbf{e}^{b\cdot 4} \end{pmatrix}$$

Beispiel 9.5: Lösung

- Für den ersten Schritt des ungedämpften Verfahrens erhalten wir also

$$\boldsymbol{Dg}(1,-1.5) = \begin{pmatrix} -1 & 0 \\ -\mathbf{e}^{-1.5} & -\mathbf{e}^{-1.5} \\ -\mathbf{e}^{-3} & -2\mathbf{e}^{-3} \\ -\mathbf{e}^{-4.5} & -3\mathbf{e}^{-4.5} \\ -\mathbf{e}^{-6} & -4\mathbf{e}^{-6} \end{pmatrix}, \quad \boldsymbol{g}(1,-1.5) = \begin{pmatrix} 3-1 \\ 1-\mathbf{e}^{-1.5} \\ 0.5-\mathbf{e}^{-3} \\ 0.2-\mathbf{e}^{-4.5} \\ 0.05-\mathbf{e}^{-6} \end{pmatrix}$$

Beispiel 9.5: Lösung

- Für das lineare Gleichungssystem:

$$\begin{pmatrix} -1 & -\mathbf{e}^{-1.5} & -\mathbf{e}^{-3} & -\mathbf{e}^{-4.5} & -\mathbf{e}^{-6} \\ 0 & -\mathbf{e}^{-1.5} & -2\mathbf{e}^{-3} & -3\mathbf{e}^{-4.5} & -4\mathbf{e}^{-6} \end{pmatrix} \begin{pmatrix} -1 & 0 \\ -\mathbf{e}^{-1.5} & -\mathbf{e}^{-1.5} \\ -\mathbf{e}^{-3} & -2\mathbf{e}^{-3} \\ -\mathbf{e}^{-4.5} & -3\mathbf{e}^{-4.5} \\ -\mathbf{e}^{-6} & -4\mathbf{e}^{-6} \end{pmatrix} \begin{pmatrix} \delta_1^{(0)} \\ \delta_2^{(0)} \end{pmatrix}$$

$$= -\begin{pmatrix} -1 & -\mathbf{e}^{-1.5} & -\mathbf{e}^{-3} & -\mathbf{e}^{-4.5} & -\mathbf{e}^{-6} \\ 0 & -\mathbf{e}^{-1.5} & -2\mathbf{e}^{-3} & -3\mathbf{e}^{-4.5} & -4\mathbf{e}^{-6} \end{pmatrix} \cdot \begin{pmatrix} 3-1 \\ 1-\mathbf{e}^{-1.5} \\ 0.5-\mathbf{e}^{-3} \\ 0.2-\mathbf{e}^{-4.5} \\ 0.05-\mathbf{e}^{-6} \end{pmatrix}$$

Beispiel 9.5: Lösung

Vereinfacht:

$$\begin{pmatrix} 1.0524 & 0.0551 \\ 0.0551 & 0.0609 \end{pmatrix} \begin{pmatrix} \delta_1^{(0)} \\ \delta_2^{(0)} \end{pmatrix} = \begin{pmatrix} 2.1980 \\ 0.2249 \end{pmatrix}$$

mit der Lösung

$$\begin{pmatrix} \delta_1^{(0)} \\ \delta_2^{(0)} \end{pmatrix} = \begin{pmatrix} 1.9894 \\ 1.8920 \end{pmatrix}$$

und damit

$$\begin{aligned} \lambda^{(1)} &= \lambda^{(0)} + \delta^{(0)}. \\ &= \begin{pmatrix} 1 \\ -1.5 \end{pmatrix} + \begin{pmatrix} 1.9894 \\ 1.8920 \end{pmatrix} = \begin{pmatrix} 2.9894 \\ 0.3920 \end{pmatrix} \end{aligned}$$

Beispiel 9.5: Lösung

- Die weiteren Lösungen des ungedämpften Verfahrens sind

i	0	1	2	5	10
$\lambda^{(i)}$	$\begin{pmatrix} 1 \\ -1.5 \end{pmatrix}$	$\begin{pmatrix} 2.99 \\ 0.392 \end{pmatrix}$	$\begin{pmatrix} 1.26 \\ 0.279 \end{pmatrix}$	$\begin{pmatrix} 2.91 \\ -0.856 \end{pmatrix}$	$\begin{pmatrix} 2.981658705 \\ -1.003280776 \end{pmatrix}$

- Ab $\lambda^{(13)} = (2.981658972, -1.003281352)^T$ tritt keine Veränderung der Ziffern mehr ein.
- Dies stimmt mit dem Ergebnis aus Bsp. 9.4 überein.
- Für den Startvektor $\lambda^{(0)} = (2,2)^T$ tritt hingegen keine Konvergenz ein, so erhalten wir z.B. $\lambda^{(13)} = (0.3698, 33.6868)^T$.

Beispiel 9.5: Lösung

- Dagegen sind die Ergebnisse des gedämpften Gauss-Newton-Verfahrens:

i	0	1	2	3	4
$\lambda^{(i)}$	$\begin{pmatrix} 1 \\ -1.5 \end{pmatrix}$	$\begin{pmatrix} 1.99 \\ -0.554 \end{pmatrix}$	$\begin{pmatrix} 2.919 \\ -0.951 \end{pmatrix}$	$\begin{pmatrix} 2.980 \\ -0.999 \end{pmatrix}$	$\begin{pmatrix} 2.981516868 \\ -1.002965939 \end{pmatrix}$

- Ab $\lambda^{(7)} = (2.981658324, -1.003279952)^T$ tritt keine Veränderung der Ziffern mehr ein.
- Man sieht dass das gedämpfte Gauss-Newton Verfahren bei gleichem Startvektor wesentlich schneller konvergiert.

Beispiel 9.5: Lösung

- Für den Startvektor $\lambda^{(0)} = (2,2)^T$ tritt beim gedämpften Verfahren ebenfalls Konvergenz ein:

i	0	1	2	5	10
$\lambda^{(i)}$	$\begin{pmatrix} 2 \\ 2 \end{pmatrix}$	$\begin{pmatrix} 0.00384 \\ 2.00 \end{pmatrix}$	$\begin{pmatrix} 0.00384 \\ 1.75 \end{pmatrix}$	$\begin{pmatrix} 0.207 \\ -0.752 \end{pmatrix}$	$\begin{pmatrix} 2.981652024 \\ -1.003266310 \end{pmatrix}$

- Ab $\lambda^{(14)} = (2.981658971, -1.003281352)^T$ tritt keine Veränderung der Ziffern mehr auf.
- Man sieht, dass das gedämpfte Gauss-Newton-Verfahren auch für Startvektoren konvergiert, für die das ungedämpfte Gauss-Newton-Verfahren nicht konvergiert.

Aufgabe 9.2

- Bestimmen Sie für die folgenden Daten die Funktion der Form $f(x) = a \ln(x+b)$, die diese Daten bestmöglich im Sinne der kleinsten Fehlerquadrate approximiert:

x_i	1	2	3	4
y_i	7.1	7.9	8.3	8.8

Aufgabe 9.2: Lösung