

Ausgleichsprobleme

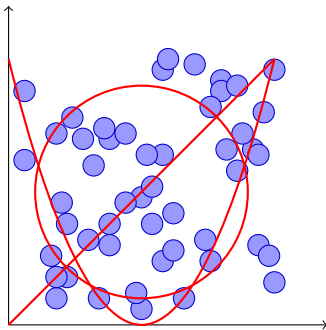
Lineare Ausgleichsprobleme

Motivation

Ausgleichsrechn

Lineare
Ausgleichsprobleme

Nichtlineare
Ausgleichsprobleme



Ausgleichsproblem

Definition 10

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 , die die Wertepaare bestmöglich annähert, d.h. dass möglichst genau gilt: $f(x_i) \approx y_i$ für $i = 1, \dots, n$.

Bemerkungen:

- Ist die Menge der stetigen Funktionen nicht eingeschränkt, so ist die optimale Funktion, die sogar $f(x_i) = y_i$ erfüllt, das Interpolationspolynom.
- Aber: Beim Ausgleichsproblem ist nicht verlangt, dass die Funktion genau durch die gegebenen Wertepaare verläuft. In vielen Anwendungen sind die Wertepaare Messdaten, die mit einem Fehler behaftet sind. Dann ist die Forderung, dass die Funktion durch die Wertepaare verläuft, wenig sinnvoll.

Ausgleichsproblem

Definition 11

Gegeben sei eine Menge \mathcal{F} von stetigen Funktionen auf einem Intervall $[a, b]$ ($\mathcal{F} \subset \mathcal{C}([a, b])$) und n Wertepaare $(x_i, y_i), i = 1, \dots, n$. Eine Funktion $f \in \mathcal{F}$ heißt Ausgleichsfunktion zu den Wertepaaren, falls das Fehlerfunktional

$$E(f) := \sum_{i=1}^n \left(y_i - f(x_i) \right)^2$$

für f minimal wird, d.h. $E(f) = \min\{E(g) | g \in \mathcal{F}\}$. Die Menge \mathcal{F} nennt man auch Ansatzfunktionen.

Ausgleichsproblem

Ausgleichsrechn

Lineare
Ausgleichsprobleme
Nichtlineare
Ausgleichsprobleme

Bemerkungen:

- Das gefundene f ist optimal im Sinne der *kleinsten Fehlerquadrate*. Dies entspricht der Forderung, dass die 2-Norm des Fehlervektors $(f(x_1) - y_1, f(x_2) - y_2, \dots, f(x_n) - y_n)^T$ minimal werden soll. Alternativ können andere Normen des Fehlervektors zur Minimierung berechnet werden, jedoch wird die 2-Norm am häufigsten verwendet.
- Die Fehler können mit Gewichten ω_i belegt werden, d.h. $\sum_{i=1}^n \omega_i (f(x_i) - y_i)^2$, wobei $\omega_i > 0$ Konstante sind. Hierdurch können z.B. Messwerte aus qualitativ „schlechterer“ Messung geringer gewichtet werden, indem sie mit einem Faktor $\omega_i < 1$ multipliziert werden.
- In Anwendungen wählt man oft \mathcal{F} als die Menge aller Geraden. Das so gefundene f nennt man *Ausgleichsgerade* oder *Regressionsgerade*.

Lineare Ausgleichsprobleme

Beispiel

Zu den Wertepaaren $\begin{array}{c|c|c|c|c} x_i & 1 & 2 & 3 & 4 \\ y_i & 6 & 6.8 & 10 & 10.5 \end{array}$ soll die Ausgleichsgerade bestimmt werden.

Lösung: Gesucht ist die Ausgleichsgerade in der Form $y = ax + b$, also $\mathcal{F} := \{a_1 f_1 + a_2 f_2 | a_1, a_2 \in \mathbb{R}\}$ mit den Ansatzfunktionen $f_1(x) = x$ und $f_2(x) = 1$.

Das Fehlerfunktional hat dann die Form

$$E(f(a, b)) := \sum_{i=1}^4 (y_i - f(x_i))^2 = \sum_{i=1}^4 (y_i - (ax_i + b))^2$$

Lineare Ausgleichsprobleme

Das Fehlerfunktional $E(f(a, b))$ soll minimal werden, d.h. die partiellen Ableitungen nach den Parametern a und b müssen verschwinden:

$$0 = \frac{\partial E(f(a, b))}{\partial a} = -2 \sum_{i=1}^n (y_i - (ax_i + b))x_i$$

$$0 = \frac{\partial E(f(a, b))}{\partial b} = -2 \sum_{i=1}^n (y_i - (ax_i + b))$$

Dies liefert 2 Gleichungen für die beiden Unbekannten a und b . Nach wenigen Umformungen erhält man:

$$a \sum_{i=1}^n x_i^2 + b \sum_{i=1}^n x_i = \sum_{i=1}^n y_i x_i$$

$$a \sum_{i=1}^n x_i + b \sum_{i=1}^n 1 = \sum_{i=1}^n y_i$$

Lineare Ausgleichsprobleme

Dies ist ein lineares Gleichungssystem für die beiden Unbekannten a und b , das sich in Matrix-Vektor-Form schreiben lässt als:

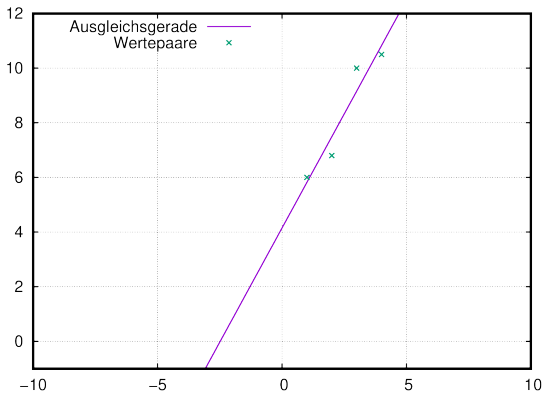
$$\begin{pmatrix} \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & n \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n y_i x_i \\ \sum_{i=1}^n y_i \end{pmatrix}$$

Nach Einsetzen der Werte für x_i und y_i erhält man:

$$\begin{pmatrix} 30 & 10 \\ 10 & 4 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 91.6 \\ 33.3 \end{pmatrix} \Rightarrow a = 1.67, b = 4.15$$

Die Ausgleichsgerade ist: $y = 1.67x + 4.15$.

Lineare Ausgleichsprobleme



Lineare Ausgleichsprobleme

Definition 12

Gegeben sind Basisfunktionen f_1, \dots, f_m ,

$\mathcal{F} := \{\lambda_1 f_1 + \lambda_2 f_2 + \dots \lambda_m f_m \mid \lambda_i \in \mathbb{R} \text{ für alle } i = 1, \dots, m\}$
sowie n Wertepaare $(x_i, y_i), i = 1, \dots, n$. Man sagt, dass ein
lineares Ausgleichsproblem vorliegt.

Weiter sei für $f = \sum_{j=1}^m \lambda_j f_j \in \mathcal{F}$:

$$\begin{aligned} E(\lambda_1, \lambda_2, \dots, \lambda_m) &:= \sum_{i=1}^n (y_i - f(x_i))^2 \\ &= \sum_{i=1}^n \left(y_i - \sum_{j=1}^m \lambda_j f_j(x_i) \right)^2 \\ &= \|\mathbf{y} - \mathbf{A}\boldsymbol{\lambda}\|_2^2. \end{aligned}$$

Lineare Ausgleichsprobleme

Hierbei ist

$$\mathbf{A} = \begin{pmatrix} f_1(x_1) & f_2(x_1) & \dots & f_m(x_1) \\ f_1(x_2) & f_2(x_2) & \dots & f_m(x_2) \\ \vdots & \vdots & \vdots & \vdots \\ f_1(x_n) & f_2(x_n) & \dots & f_m(x_n) \end{pmatrix} \text{ und}$$

$$\mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \quad \boldsymbol{\lambda} = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_m \end{pmatrix}.$$

Das System $\mathbf{A} \boldsymbol{\lambda} = \mathbf{y}$ heißt **Fehlergleichungssystem**.

Normalgleichungen

Definition 13

Die Gleichungen

$$0 = \frac{\partial E(f(\lambda_1, \lambda_2, \dots, \lambda_m))}{\partial \lambda_i}, \quad i = 1, \dots, m$$

heißen **Normalgleichungen** zum linearen Ausgleichsproblem.

Das System der Normalgleichungen heißt
Normalgleichungssystem; es lässt sich als lineares
Gleichungssystem in der Form

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

schreiben.

Normalgleichungen

Bemerkungen:

- Das Fehlergleichungssystem besitzt n Gleichungen und m Unbekannte. In der Regel ist $n > m$, d.h. es gibt mehr Wertepaare als zu bestimmende Parameter. Das Gleichungssystem ist überbestimmt. Es gibt nicht immer eine Lösung.
- Zur Lösung des Ausgleichsproblems muss das Fehlerfunktional in den Parametern minimal werden, d.h. die partiellen Ableitungen nach den Parametern müssen verschwinden. Die Lösungen der Normalgleichungen sind die gesuchten Parameter des Ausgleichsproblems.
- Das Lösen der Normalgleichungen eines linearen Ausgleichsproblems entspricht dem Lösen eines linearen Gleichungssystems mit symmetrischer $m \times m$ Matrix $\mathbf{A}^T \mathbf{A}$, rechter Seite $\mathbf{A}^T \mathbf{y}$ und gesuchtem Lösungsvektor $\boldsymbol{\lambda} \in \mathbb{R}^m$.

Lineare Ausgleichsprobleme

Beispiel

Gesucht ist das lineare Gleichungssystem bestehend aus den Normalgleichungen zu den Wertepaaren:

x_i	1	2	3	4
y_i	6	6.8	10	10.5

Lösung: Gegeben sind $f_1(x) = x$, $f_2(x) = 1$ und $n = 4$ Wertepaare. Die Matrix \mathbf{A} ist also eine 4×2 Matrix

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

Lineare Ausgleichsprobleme

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

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

Damit ist $\mathbf{A}^T \mathbf{A} \boldsymbol{\lambda} = \mathbf{A}^T \mathbf{y}$ dasselbe Gleichungssystem wie aus dem vorherigen Beispiel.

Lineare Ausgleichsprobleme

Beispiel

Gegeben sind die Messdaten

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

Gesucht ist eine Funktion $f(x) = ae^x + b$, die die Daten bestmöglich bzgl. der kleinsten Fehlerquadrate approximiert.

Lösung: Die Ansatzfunktionen lauten $f_1(x) = e^x$ und $f_2(x) = 1$. Das Fehlergleichungssystem hat dann die Form:

$$\mathbf{A}\boldsymbol{\lambda} = \mathbf{y} \iff \begin{pmatrix} 1.0 & 1.0 \\ 2.7183 & 1.0 \\ 7.3891 & 1.0 \\ 20.0865 & 1.0 \\ 54.5982 & 1.0 \end{pmatrix} \boldsymbol{\lambda} = \begin{pmatrix} 6 \\ 12 \\ 30 \\ 80 \\ 140 \end{pmatrix}$$

Lineare Ausgleichsprobleme

Hieraus ergibt sich das zugehörige Normalgleichungssystem

$$\mathbf{A}^T \mathbf{A} \boldsymbol{\lambda} = \mathbf{A}^T \mathbf{y} \iff \begin{pmatrix} 3447.37 & 85.79 \\ 85.79 & 5.0 \end{pmatrix} \boldsymbol{\lambda} = \begin{pmatrix} 9510.88 \\ 268.0 \end{pmatrix}$$

Dieses System hat die Lösung

$$\begin{pmatrix} 2.49 \\ 10.93 \end{pmatrix}$$

und damit die Lösung für die Ausgleichsfunktion

$$f(x) = 2.49e^x + 10.93.$$

Nichtlineare Ausgleichsprobleme

Nichtlineares Ausgleichsproblem

Definition 14

Gegeben sei eine Menge \mathcal{F} von Ansatzfunktionen f_p , die von m Parametern $\lambda_1, \lambda_2, \dots, \lambda_m$ abhängen:

$\mathcal{F} := \{f_p(\lambda_1, \lambda_2, \dots, \lambda_m, x) \mid \lambda_1, \lambda_2, \dots, \lambda_m, x \in \mathbb{R}\}$, sowie n Wertepaare (x_i, y_i) , $i = 1, \dots, n$.

Das Fehlerfunktional lautet dann

$$E(\lambda_1, \lambda_2, \dots, \lambda_m) = \sum_{i=1}^n (y_i - f_p(\lambda_1, \lambda_2, \dots, \lambda_m, x_i))^2$$

Nichtlineare Ausgleichsprobleme

Mit den Bezeichnungen

$$\begin{aligned}\mathbf{f}(\lambda_1, \dots, \lambda_m) &:= \begin{pmatrix} f_1(\lambda_1, \dots, \lambda_m) \\ \vdots \\ f_n(\lambda_1, \dots, \lambda_m) \end{pmatrix} \\ &:= \begin{pmatrix} y_1 - f_p(\lambda_1, \dots, \lambda_m, x_1) \\ \vdots \\ y_n - f_p(\lambda_1, \dots, \lambda_m, x_n) \end{pmatrix}\end{aligned}$$

lässt sich das Fehlerfunktional $E(\lambda_1, \dots, \lambda_m)$ schreiben als:

$$\begin{aligned}E(\lambda_1, \dots, \lambda_m) &= \sum_{i=1}^n f_i(\lambda_1, \dots, \lambda_m)^2 \\ &= \|\mathbf{f}(\lambda_1, \dots, \lambda_m)\|_2^2 = \|\mathbf{f}(\boldsymbol{\lambda})\|_2^2\end{aligned}$$

Nichtlineare Ausgleichsprobleme

Die Lösung des **Ausgleichsproblems** entspricht der Aufgabe:
Bestimme $\lambda_1, \lambda_2, \dots, \lambda_m$ so, dass $E(\lambda_1, \lambda_2, \dots, \lambda_m)$ minimal wird unter allen zulässigen Parameterbedingungen.

Bemerkung:

- Für Ansatzfunktionen f_p , die linear in den Parametern sind, entspricht die Definition dem linearen Ausgleichsproblem.
- Das allgemeine Ausgleichsproblem erfordert die Bestimmung des Minimums der Funktion $E : \mathbb{R}^m \rightarrow \mathbb{R}$. Hierzu können die Normalengleichungen aufgestellt werden, indem die partiellen Ableitungen von \mathbf{f} nach den Parametern λ_i gleich Null gesetzt und das resultierende nichtlineare Gleichungssystem gelöst wird.

Nichtlineare Ausgleichsprobleme

Beispiel

An folgende Daten

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

 soll eine Ansatzfunktion $f(x) = ae^{bx}$ bestmöglich im Sinne der kleinsten Fehlerquadrate angepasst werden.

Lösung: Die Parameter a und b der Ansatzfunktion $f_p(a, b, x) = ae^{bx}$ müssen so bestimmt werden, dass

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

minimal wird.

Nichtlineare Ausgleichsprobleme

Die partiellen Ableitungen von $E(a, b)$ lauten

$$0 = \frac{\partial E(a, b)}{\partial a} = -2 \sum_{i=1}^5 \left(y_i - a e^{b x_i} \right) e^{b x_i}$$
$$0 = \frac{\partial E(a, b)}{\partial b} = -2 \sum_{i=1}^5 \left(y_i - a e^{b x_i} \right) a e^{b x_i} x_i.$$

Dies ist ein nichtlineares Gleichungssystem mit zwei Unbekannten. Die Lösung kann mit einem mehrdimensionalen Newton-Verfahren berechnet werden.

Hier hängt die 1. Gleichung linear von a ab und kann daher direkt nach a aufgelöst und in die 2. Gleichung eingesetzt werden. Dann hängt die 2. Gleichung nur noch von b ab und kann über ein eindimensionales Newton-Verfahren gelöst werden. Als Ergebnis ergibt sich:

$$a = 2.981658972 \quad \text{und} \quad b = -1.003281352$$

Nichtlineare Ausgleichsprobleme

Bemerkungen:

Das Vorgehen hat in der Praxis verschiedene Nachteile:

- Es werden die partiellen Ableitungen benötigt. Diese sind nicht unbedingt bekannt.
- Das Newton-Verfahren konvergiert nur, wenn eine genügend gute Anfangsnäherung gegeben ist. Hierzu muss man bereits eine Näherung für die optimalen Parameter kennen. Das ist ebenfalls in der Praxis meistens nicht gegeben.
- Eine Alternative zur Behandlung von Minimierungsverfahren bietet daher das sog. Gauß-Newton-Verfahren an, das diese Nachteile nicht besitzt.

Gauß-Newton-Verfahren

Ausgangspunkt ist ein nichtlineares Ausgleichsproblem, das zu minimieren ist.

Definition 15

Gegeben ist eine Funktion $\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}$ und das zugehörige Fehlerfunktional $E : \mathbb{R}^n \rightarrow \mathbb{R}$ gegeben durch $E(\boldsymbol{\lambda}) := \|\mathbf{f}(\boldsymbol{\lambda})\|_2^2$. Das Problem, einen Lösungsvektor $\boldsymbol{\lambda}$ zu finden, für den $E(\boldsymbol{\lambda})$ minimal wird, nennt man *Quadratmittelproblem*.

Das Gauß-Newton-Verfahren besteht aus einer Kombination von linearer Ausgleichsrechnung und Newton-Verfahren.

Hierzu wird in $\|\mathbf{f}(\boldsymbol{\lambda})\|_2^2$ der Vektor $\mathbf{f}(\boldsymbol{\lambda})$ durch einen linearen Ausdruck ersetzt und dieser wird minimiert. Das nichtlineare Ausgleichsproblem wird in ein lineares überführt. Der Prozess wird anschließend ähnlich dem Newton-Verfahren iteriert.

Gauß-Newton-Verfahren

Rückblick: Newton-Verfahren

Eine differenzierbare Funktion kann sich in der Nähe einer Stelle x_0 durch ihre Tangentengleichung (bzw. durch die beiden ersten Summanden der Taylorentwicklung) recht gut nähern:

$$f(\lambda) \approx f(\lambda_0) + f'(\lambda_0)(\lambda - \lambda_0).$$

Eine entsprechende Linearisierung ist auch für eine mehrdimensionale Funktionen $\mathbf{f} : \mathbb{R}^m \rightarrow \mathbb{R}^n$ möglich in der Form

$$\mathbf{f}(\boldsymbol{\lambda}) \approx \mathbf{f}(\boldsymbol{\lambda}^{(0)}) + \mathbf{D}\mathbf{f}(\boldsymbol{\lambda}^{(0)})(\boldsymbol{\lambda} - \boldsymbol{\lambda}^{(0)}),$$

wobei $\mathbf{D}\mathbf{f}(\boldsymbol{\lambda}^{(0)})$ die sogenannte *Jacobi-Matrix* in $\boldsymbol{\lambda}^{(0)}$ bezeichnet mit

$$\mathbf{D}\mathbf{f}(\boldsymbol{\lambda}^{(0)}) = \begin{pmatrix} \frac{\partial f_1}{\partial \lambda_1}(\boldsymbol{\lambda}^{(0)}) & \frac{\partial f_1}{\partial \lambda_2}(\boldsymbol{\lambda}^{(0)}) & \dots & \frac{\partial f_1}{\partial \lambda_m}(\boldsymbol{\lambda}^{(0)}) \\ \frac{\partial f_2}{\partial \lambda_1}(\boldsymbol{\lambda}^{(0)}) & \frac{\partial f_2}{\partial \lambda_2}(\boldsymbol{\lambda}^{(0)}) & \dots & \frac{\partial f_2}{\partial \lambda_m}(\boldsymbol{\lambda}^{(0)}) \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_n}{\partial \lambda_1}(\boldsymbol{\lambda}^{(0)}) & \frac{\partial f_n}{\partial \lambda_2}(\boldsymbol{\lambda}^{(0)}) & \dots & \frac{\partial f_n}{\partial \lambda_m}(\boldsymbol{\lambda}^{(0)}) \end{pmatrix}$$

Gauß-Newton-Verfahren

Zurück zum Ausgleichsproblem und Gauß-Newton-Verfahren:

Wir nehmen an, dass eine Näherung $\lambda^{(0)}$ vorliegt und definieren das in $\lambda^{(0)}$ linearisierte Ausgleichsproblem über

$$\tilde{E}(\lambda) := \|f(\lambda^{(0)}) + Df(\lambda^{(0)})(\lambda - \lambda^{(0)})\|_2^2.$$

Das Minimieren von \tilde{E} ist ein lineares Ausgleichsproblem, dessen Lösung eine bessere Näherung als $\lambda^{(0)}$ ist und mit $\lambda^{(1)}$ bezeichnet wird.

Gauß-Newton-Verfahren

Definition 16

Sei $\lambda^{(0)}$ ein Startvektor in der Nähe des Minimums von E . Das *Gauß-Newton-Verfahren* zur näherungsweisen Bestimmung des Minimums lautet:

Für $k = 0, 1, \dots$

- Berechne $\delta^{(k)} := (\lambda - \lambda^{(k)})$ als Lösung des linearen Ausgleichsproblems: Minimiere
$$\|f(\lambda^{(k)}) + Df(\lambda^{(k)})\delta^{(k)}\|_2^2$$
- Setze $\lambda^{(k+1)} = \lambda^{(k)} + \delta^{(k)}$.

Aufgrund der schnelleren und für einen größeren Bereich an Startwerten gültigen Konvergenz wird in Anwendungen eine bessere Variante, das sogenannte *gedämpfte* Gauß-Newton-Verfahren verwendet.

Gedämpftes Gauß-Newton-Verfahren

Definition 17

Sei $\lambda^{(0)}$ ein Startvektor in der Nähe des Minimums von E . Das *gedämpfte* Gauß-Newton-Verfahren zur näherungsweisen Bestimmung des Minimums lautet:

Für $k = 0, 1, \dots$

- Berechne $\delta^{(k)}$ als Lösung des linearen Ausgleichsproblems:
Minimiere $\|\mathbf{f}(\lambda^{(k)}) + \mathbf{D}\mathbf{f}(\lambda^{(k)})\delta^{(k)}\|_2^2$ durch Lösung von

$$\mathbf{D}\mathbf{f}(\lambda^{(k)})^T \mathbf{D}\mathbf{f}(\lambda^{(k)}) \delta^{(k)} = -\mathbf{D}\mathbf{f}(\lambda^{(k)})^T \mathbf{f}(\lambda^{(k)})$$

- Bestimme die Schrittweite für die Dämpfung: Wähle die größte Zahl $t \in \{1, \frac{1}{2}, \frac{1}{4}, \dots\}$, für die mit
 $\varphi(t) := \|\mathbf{f}(\lambda^{(k)} + t\delta^{(k)})\|_2^2$ gilt $\varphi(t) < \varphi(0)$.
- Setze $\lambda^{(k+1)} := \lambda^{(k)} + t\delta^{(k)}$.

Gauß-Newton-Verfahren

Bemerkungen:

- Bei dem Dämpfungsschritt wird die Korrekturrichtung $\delta^{(k)}$ mit einem Faktor t multipliziert. Dieser Faktor entsteht aus fortlaufender Halbierung. Hierdurch konvergiert das Verfahren für eine größere Menge an Startvektoren als beim ungedämpften Verfahren. Außerdem wird ein echter Abstieg der Funktionswerte sichergestellt, so dass

$$\|f(\lambda^{(k+1)})\|_2^2 = \varphi(t) < \varphi(0) = \|f(\lambda^{(k)})\|_2^2$$

- Die Konvergenz muss durch weitere Bedingungen garantiert werden.
- Als Abbruchkriterium der Iteration kann z.B. $\|t\delta^{(k)}\| < \epsilon$ wie bei dem Newton-Verfahren gewählt werden.

Gauß-Newton-Verfahren

Beispiel

Zu den Messdaten

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

und der Ansatzfunktion $f(x) = ae^{bx}$ soll das ungedämpfte Gauß-Newton-Verfahren angewandt werden.

Lösung: Die Funktion f besitzt die Komponenten $f_i(a, b) := y_i - ae^{bx_i}$. Daher lautet die i -te Zeile der Jacobi-Matrix $(Df)_i(a, b) = (-e^{bx_i}, -ax_ie^{bx_i})$.

Mit dem Startwert $\lambda^{(0)} = (1, -1.5)$ ergibt das ungedämpfte Gauß-Newton-Verfahren die Iterationswerte:

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

Gauß-Newton-Verfahren

Bei der Newton-Iteration ändert sich ab $\lambda^{(13)} = \begin{pmatrix} 2.981658972 \\ -1.003281352 \end{pmatrix}$ der Lösungsvektor im Bereich der angegebenen Ziffern nicht mehr.

Bemerkung:

Für den Startvektor $\lambda^{(0)} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}^T$ tritt bei dem ungedämpften Gauß-Newton-Verfahren keine Konvergenz ein, wohl aber bei dem gedämpften Gauß-Newton-Verfahren.