

8 Iterationsverfahren zur Lösung von Gleichungen

Nichtlineare Gleichungssysteme (sogar eine nichtlineare Gleichung in einer Unbekannten) müssen fast immer iterativ gelöst werden (vgl. Kapitel 2.2).

Große dünn besetzte **lineare** Gleichungssysteme müssen iterativ gelöst werden, weil direkte Verfahren wie die Gauß-Elimination sowohl bez. Rechenaufwand (Komplexität $\sim n^3$) als auch bez. Speicherbedarf („fill-in“) in der Regel zu kostspielig sind.

Wir betrachten nichtlineare Gleichungen stets in einer der beiden (nicht-eindeutigen) Standardformen:

$$f(x) = 0,$$

$$x = g(x),$$

Nullstellenform,

Fixpunktform.

8.1 Fixpunktiteration

8.2 Konvergenzordnung

8.3 Nullstellen reellwertiger Funktionen

8.4 Das Newton-Verfahren im \mathbb{R}^n

8.5 Modifikationen des Newton-Verfahrens

8.6 Nichtlineare Ausgleichsprobleme

8.7 Klassische Iterationsverfahren für lineare Systeme

8.8 Gradientenverfahren für lineare Systeme

8.1 Fixpunktiteration

Satz 8.1. *Es sei $\|\cdot\|$ eine beliebige Norm im \mathbb{R}^n und $D \subset \mathbb{R}^n$ abgeschlossen. Die Selbstabbildung $f : D \rightarrow D$ sei **stark kontrahierend**, d.h. es existiere eine Konstante $L < 1$ mit*

$$\|f(x) - f(y)\| \leq L\|x - y\| \quad \forall x, y \in D. \quad (8.1)$$

Dann folgt:

- (a) *f besitzt genau einen Fixpunkt x^* in D ($f(x^*) = x^*$),*
- (b) *für jeden Startvektor $x_0 \in D$ konvergiert die durch $x_{m+1} := f(x_m)$ (**Fixpunktiteration**) definierte Folge $\{x_m\}_{m \geq 0}$ gegen x^* ,*
- (c) $\|x_m - x^*\| \leq \frac{L^m}{1-L} \|x_0 - x_1\|$ *(a-priori-Abschätzung),*
- (d) $\|x_m - x^*\| \leq \frac{L}{1-L} \|x_m - x_{m-1}\|$ *(a-posteriori-Abschätzung).*

Der Spezialfall $n = 1$:

Korollar 8.2. $I \subseteq \mathbb{R}$ sei ein abgeschlossenes Intervall, $f : I \rightarrow I$ sei stark kontrahierend, d.h. $\exists L < 1$ mit

$$|f(x) - f(y)| \leq L |x - y| \quad \forall x, y \in I.$$

Dann besitzt f genau einen Fixpunkt x^* in I . Die Iterationsfolge $\{x_m\}_{m \geq 0}$,

$$x_{m+1} := f(x_m) \quad (m = 0, 1, 2, \dots),$$

konvergiert für beliebiges $x_0 \in I$ gegen x^* . Es gelten die Abschätzungen

$$|x^* - x_m| \leq \frac{L^m}{1 - L} |x_0 - x_1| \quad \text{sowie} \quad |x^* - x_m| \leq \frac{L}{1 - L} |x_m - x_{m-1}|.$$

Ein Kriterium für starke Kontraktion:

Ist $f : D \rightarrow D$ aus $C^1(D)$, d.h. für die Komponentenabbildungen $f_i : D \rightarrow \mathbb{R}$ sind alle partiellen Ableitungen $\frac{\partial f_i}{\partial x_j} : D \rightarrow \mathbb{R}$ stetig ($1 \leq i, j \leq n$), und gibt es eine Matrixnorm $\|\cdot\|_M$, so dass für die Funktionalmatrix von f ,

$$f'(\mathbf{x}) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}) & \cdots & \frac{\partial f_1}{\partial x_n}(\mathbf{x}) \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1}(\mathbf{x}) & \cdots & \frac{\partial f_n}{\partial x_n}(\mathbf{x}) \end{bmatrix} \in \mathbb{R}^{n \times n} \quad (\text{eigentlich } f' : D \rightarrow \mathbb{R}^{n \times n}),$$

$L := \max_{\mathbf{x} \in D} \|f'(\mathbf{x})\|_M < 1$ gilt, dann folgt

$$\|f(\mathbf{x}) - f(\mathbf{y})\| \leq L \|\mathbf{x} - \mathbf{y}\| \quad \forall \mathbf{x}, \mathbf{y} \in D,$$

d.h. f ist stark kontrahierend bez. jeder Vektornorm $\|\cdot\|$, die mit $\|\cdot\|_M$ verträglich ist.

Für den Spezialfall $n = 1$ bedeutet das: Ist $f : I \rightarrow I$ aus $C^1(I)$ mit $L := \max_{x \in I} |f'(x)| < 1$, dann folgt

$$|f(x) - f(y)| = |f'(\xi)| \cdot |x - y| \leq L|x - y| \quad \forall x, y \in I,$$

d.h. f ist stark kontrahierend.

Beispiel 1. Gesucht ist eine Lösung $x^* = [\xi, \eta]^\top$ des Gleichungssystems

$$x_1 = 0.7 \sin x_1 + 0.2 \cos x_2,$$

$$x_2 = 0.7 \cos x_1 - 0.2 \sin x_2.$$

Äquivalent: Gesucht ist ein Fixpunkt x^* der Abbildung

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \mapsto \begin{bmatrix} 0.7 \sin x_1 + 0.2 \cos x_2 \\ 0.7 \cos x_1 - 0.2 \sin x_2 \end{bmatrix}.$$

Die Frobenius-Norm der Funktionalmatrix von f ,

$$f'(\mathbf{x}) = f' \left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right) = \begin{bmatrix} 0.7 \cos x_1 & -0.2 \sin x_2 \\ -0.7 \sin x_1 & -0.2 \cos x_2 \end{bmatrix}$$

ist (für alle $\mathbf{x} \in \mathbb{R}^2$) durch

$$\begin{aligned} \|f'(\mathbf{x})\|_F &= (0.49 \cos^2 x_1 + 0.04 \sin^2 x_2 + 0.49 \sin^2 x_1 + 0.04 \cos^2 x_2)^{1/2} \\ &= (0.49 + 0.04)^{1/2} =: L \approx 0.729 \end{aligned}$$

gegeben. Also ist f auf ganz \mathbb{R}^2 stark kontrahierend (mit der Kontraktionskonstanten L). f besitzt deshalb in \mathbb{R}^2 genau einen Fixpunkt \mathbf{x}^* und die Fixpunktiteration

$$\begin{bmatrix} x_1^{(m+1)} \\ x_2^{(m+1)} \end{bmatrix} := f \left(\begin{bmatrix} x_1^{(m)} \\ x_2^{(m)} \end{bmatrix} \right) = \begin{bmatrix} 0.7 \sin x_1^{(m)} + 0.2 \cos x_2^{(m)} \\ 0.7 \cos x_1^{(m)} - 0.2 \sin x_2^{(m)} \end{bmatrix}$$

konvergiert für jeden Startvektor $(x_1^{(0)}, x_2^{(0)})^\top \in \mathbb{R}^2$ gegen \mathbf{x}^* .

Wir wählen $\mathbf{x}_0 = [x_1^{(0)}, x_2^{(0)}]^\top = [0, 0]^\top$.

Frage: Wieviele Iterationsschritte m sind erforderlich, um garantieren zu können, dass $\|\mathbf{x}^* - \mathbf{x}_m\|_2 \leq 10^{-4}$?

Die a-priori-Abschätzung liefert

$$\|\mathbf{x}^* - \mathbf{x}_m\|_2 \leq \frac{L^m}{1-L} \|\mathbf{x}_0 - \mathbf{x}_1\|_2 = \frac{0.53^{m/2}}{1-0.53^{1/2}} 0.53^{1/2} \stackrel{!}{\leq} 10^{-4},$$

was auf

$$m + 1 \geq (-4 + \log_{10}(1 - 0.53^{1/2})) / \log_{10}(0.53^{1/2}) = 33.11 \dots$$

führt. Nach 33 Iterationsschritten ist also $\|\mathbf{x}^* - \mathbf{x}_m\|_2 \leq 10^{-4}$ garantiert.

Die numerische Rechnung zeigt, dass dieser Fehler schon nach 21 Schritten erreicht wird. Man wird die Fixpunktiteration in der Praxis dann abbrechen, wenn die a-posteriori-Schranke unter 10^{-4} liegt, was hier nach 22 Schritten der Fall ist.

m	$x_1^{(m)}$	$x_2^{(m)}$	a-priori	a-posteriori	tats. Fehler
0	0	0	—	—	7.3158e-01
1	2.0000e-01	7.0000e-01	1.9486e+00	1.9486e+00	3.7883e-01
2	2.9204e-01	5.5720e-01	1.4186e+00	4.5472e-01	2.3961e-01
3	3.7128e-01	5.6460e-01	1.0328e+00	2.1303e-01	1.6527e-01
⋮					
20	5.2642e-01	5.0796e-01	4.6811e-03	1.6374e-04	1.1206e-04
21	5.2646e-01	5.0795e-01	3.4079e-03	1.0592e-04	7.2487e-05
22	5.2648e-01	5.0794e-01	2.4810e-03	6.8518e-05	4.6889e-05
⋮					
32	5.2652e-01	5.0792e-01	1.0375e-04	8.7864e-07	6.0124e-07
33	5.2652e-01	5.0792e-01	7.5533e-05	5.6834e-07	3.8891e-07

Definition 8.3. $f : D \rightarrow \mathbb{R}^n$, $D \subseteq \mathbb{R}^n$, sei eine Abbildung mit dem Fixpunkt $x^* \in D$ ($f(x^*) = x^*$). Der Fixpunkt x^* heißt **anziehend** (attraktiv), wenn es eine Kugel $K_\varepsilon(x^*) := \{x \in \mathbb{R}^n : \|x - x^*\| < \varepsilon\}$ mit $\varepsilon > 0$ gibt, so dass die Fixpunktiteration für jeden Startvektor aus $K_\varepsilon(x^*) \cap D$ gegen x^* strebt (die Fixpunktiteration ist **lokal konvergent**).

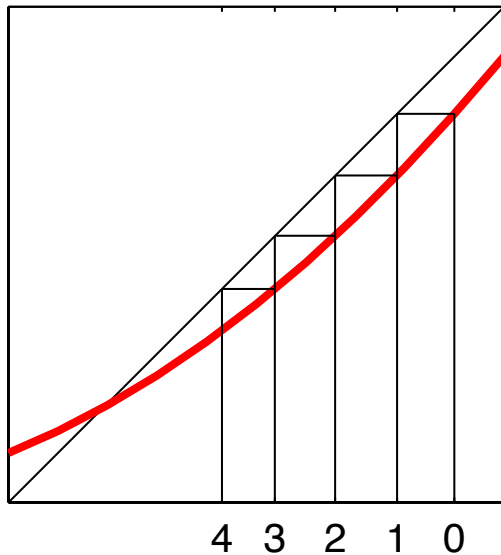
Kriterium: (Satz von Ostrowski) $x^* \in D$ ist ein anziehender Fixpunkt der stetig differenzierbaren Abbildung $f \in C^1(D)$, wenn

- (a) $f(x^*) = x^*$ und
- (b) $\rho(f'(x^*)) := \max\{|\lambda| : \lambda \text{ ist Eigenwert von } f'(x^*)\} < 1$

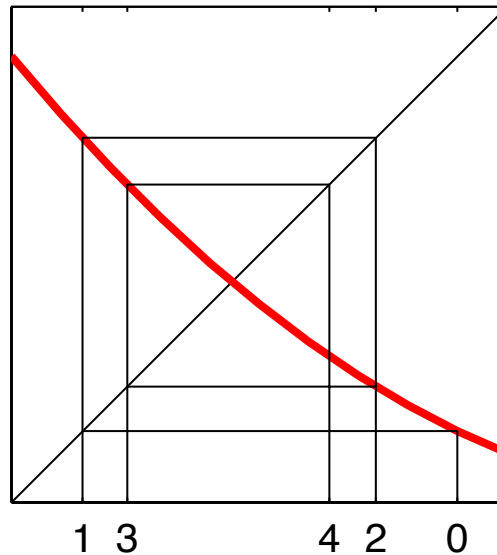
gelten.

Hinreichend für (b): Existenz einer Matrixnorm $\|\cdot\|_M$ mit $\|f'(x^*)\|_M < 1$.

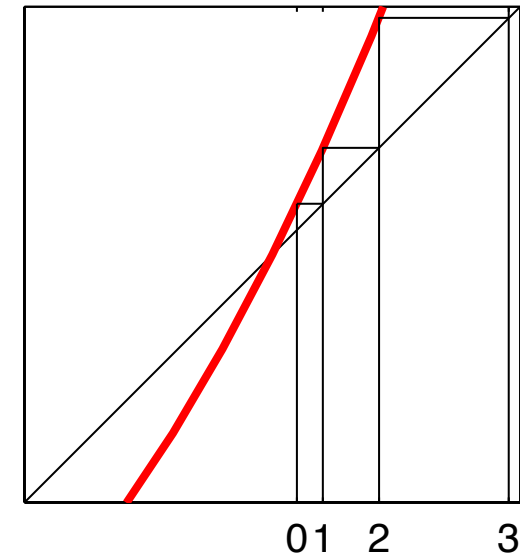
Im Spezialfall $n = 1$ ist x^* ein anziehender Fixpunkt der stetig differenzierbaren Funktion $f : I \rightarrow \mathbb{R}$, wenn $f(x^*) = x^*$ und $|f'(x^*)| < 1$ erfüllt sind.



$$0 < f'(x^*) < 1$$



$$-1 < f'(x^*) < 0$$



$$|f'(x^*)| > 1$$

8.2 Konvergenzordnung

Es sei $\{a_m\}_{m \geq 0}$ eine Nullfolge positiver reeller Zahlen. Die Folge $\{a_m\}_{m \geq 0}$ besitzt (mindestens) die **Konvergenzordnung** $p \geq 1$, wenn es ein $C > 0$ gibt mit

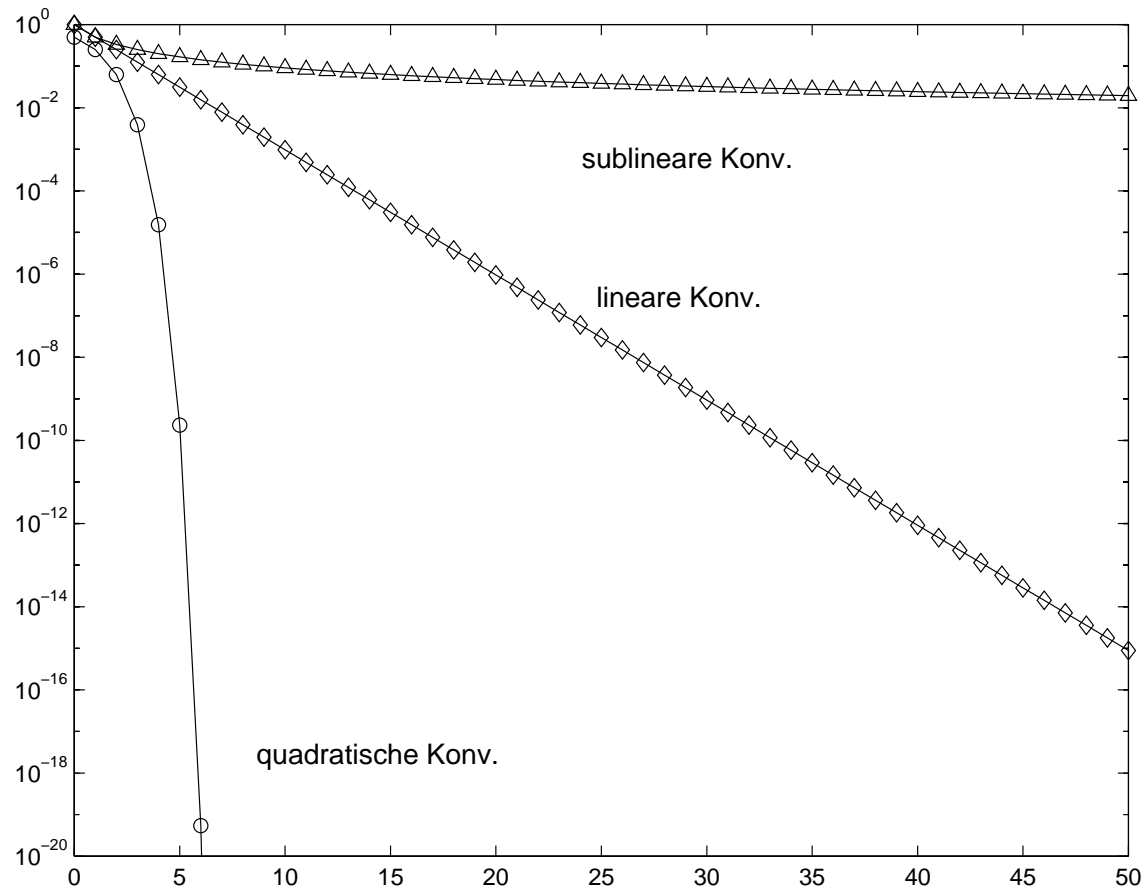
$$a_{m+1} \leq C a_m^p \quad \text{für alle } m \geq m_0.$$

Für $p = 1$ (**lineare Konvergenz**) wird zusätzlich $C < 1$ gefordert.

Die konvergente Vektorfolge $\{\mathbf{x}_m\}_{m \geq 0} \subset \mathbb{R}^n$ ($\mathbf{x}^* = \lim_{m \rightarrow \infty} \mathbf{x}_m$) besitzt (mindestens) die Konvergenzordnung p , wenn dies für die Fehlerfolge $\{\|\mathbf{x}^* - \mathbf{x}_m\|\}_{m \geq 0}$ gilt. (Dabei ist es irrelevant, in welcher Norm die Fehler gemessen werden.)

Eine Folge wie $a_m = 1/(m+1)$ konvergiert **sublinear** (langsamer als linear). Gilt $\lim_{m \rightarrow \infty} a_{m+1}/a_m = 0$, so spricht man von **superlinearer Konvergenz** (schneller als linear). Jede Folge mit Konvergenzordnung $p > 1$ konvergiert superlinear.

Faustregel: Eine Folge konvergiert umso schneller, je größer ihre Konvergenzordnung ist.



Kriterium (für $n = 1$):

$f : [a, b] \rightarrow \mathbb{R}$ sei aus $C^p[a, b]$ (p -mal stetig differenzierbar).

Es sei $f(x^*) = x^*$ für ein $x^* \in [a, b]$.

Gilt

$$0 = f'(x^*) = f''(x^*) = \dots = f^{(p-1)}(x^*),$$

so konvergiert die Fixpunktiteration

$$x_{m+1} = f(x_m) \quad (m = 0, 1, 2, \dots)$$

lokal mit (mindestens) der Ordnung p .

Beispiel.

Die Iterationsvorschrift des **Newton-Verfahrens** zur Bestimmung einer Nullstelle einer Funktion $g \in C^2[a, b]$ lautet:

$$x_{m+1} = x_m - \frac{g(x_m)}{g'(x_m)} =: f(x_m)$$

mit der **Iterationsfunktion**

$$f(x) := x - \frac{g(x)}{g'(x)}.$$

Sei $g'(x^*) \neq 0$: Dann ist x^* Nullstelle von $g \Leftrightarrow x^*$ ist Fixpunkt von f .

$$f'(x) = 1 - \frac{g'(x)g'(x) - g(x)g''(x)}{g'(x)^2} = \frac{g(x)g''(x)}{g'(x)^2} \Rightarrow f'(x^*) = 0.$$

Für **einfache** Nullstellen konvergiert das Newton-Verfahren **lokal** (mindestens) **quadratisch**, d.h. mit der Ordnung $p = 2$.

8.3 Nullstellen reellwertiger Funktionen

Gesucht sind Nullstellen einer (zumindest) stetigen Funktion $g : [a, b] \rightarrow \mathbb{R}$. Gilt $g(a)g(b) < 0$, dann besitzt g in (a, b) (mindestens) eine Nullstelle (Zwischenwertsatz).

Wir beschreiben und vergleichen drei Verfahren zur näherungsweisen Berechnung einer Nullstelle von g :

- das Intervallhalbierungsverfahren,
- das Newton-Verfahren,
- das Sekantenverfahren.

Intervallhalbierungsverfahren:

```
Setze  $a_0 = a$  und  $b_0 = b$   
 $m = 0$   
while  $b_m - a_m > \text{tol}$   
     $x = .5 * (a_m + b_m)$   
     $m = m + 1$   
    if  $g(x) * g(a_{m-1}) < 0$   
         $a_m = a_{m-1}$   
         $b_m = x$   
    else  
         $a_m = x$   
         $b_m = b_{m-1}$   
    end if  
end while
```

Das Intervallhalbierungsverfahren konvergiert linear mit dem Konvergenzfaktor $1/2$. Pro Iterationsschritt ist eine Funktionsauswertung erforderlich.

Ist g differenzierbar, so können wir das [Newton-Verfahren](#) einsetzen (vgl. Kapitel 1.2):

$$x_{m+1} = x_m - \frac{g(x_m)}{g'(x_m)} \quad (m = 0, 1, 2, \dots).$$

Wir wissen bereits, dass es bei einfachen Nullstellen lokal quadratisch konvergiert. Der Rechenaufwand beträgt i.W. zwei Funktionsauswertungen pro Schritt (insbesondere die Auswertung der Ableitung ist in der Praxis problematisch, da g' nur in Ausnahmefällen explizit bekannt ist).

Ersetzt man im Newton-Verfahren die Tangentensteigung $g'(x_m)$ durch die Sekantensteigung $\frac{g(x_m) - g(x_{m-1})}{x_m - x_{m-1}}$, so erhält man das **Sekantenverfahren**

$$\begin{aligned}x_{m+1} &= x_m - \frac{x_m - x_{m-1}}{g(x_m) - g(x_{m-1})} g(x_m) \\ &= \frac{x_{m-1}g(x_m) - x_mg(x_{m-1})}{g(x_m) - g(x_{m-1})} \quad (m = 0, 1, 2, \dots).\end{aligned}$$

Hier sind zwei Startwerte erforderlich. Ist $g \in C^2[a, b]$ und besitzt g eine einfache Nullstelle $\zeta \in [a, b]$, so konvergiert das Sekantenverfahren lokal gegen ζ mit der Konvergenzordnung

$$p = (1 + \sqrt{5})/2 = 1.61 \dots$$

Jeder Iterationsschritt erfordert eine Funktionsauswertung.

Beispiel. $g(x) = x - \exp(-x/2)$ besitzt wegen $g(0.6) = -0.14 \dots$ und $g(0.8) = 0.12 \dots$ eine Nullstelle in $[0.6, 0.8]$.

Intervallhalbierungsverfahren:

m	a_m	b_m
0	5.999999999999999998e-001	8.000000000000000004e-001
1	7.000000000000000000e-001	8.000000000000000004e-001
2	7.000000000000000000e-001	7.500000000000000000e-001
⋮		
10	7.033203125000000004e-001	7.035156250000000000e-001
⋮		
20	7.0346736907958984e-001	7.0346755981445308e-001
⋮		
30	7.0346742235124116e-001	7.0346742253750572e-001

Erwartungsgemäß konvergieren Sekantenverfahren (Startwerte $x_0 = 0.6$, $x_1 = 0.8$) und Newton-Verfahren (Startwert $x_0 = 0.6$) wesentlich schneller. Wir tabellieren die resultierenden Fehler:

m	Sekantenverfahren	Newton-Verfahren
0	1.0346742249839169e-001	1.0346742249839169e-001
1	9.6532577501608374e-002	7.1108527096874763e-004
2	6.5026396360778893e-004	3.2897961022015920e-008
3	4.0432478531915450e-006	< eps = 2.220446048e-016
4	1.7102297356075269e-010	
5	< eps = 2.220446048e-016	

8.4 Das Newton-Verfahren im \mathbb{R}^n

Gesucht: Nullstelle x^* einer stetig differenzierbaren Funktion $f : \mathbb{R}^n \supset D \rightarrow \mathbb{R}^n$, d.h. $f(x^*) = 0$. Iterationsvorschrift:

$$\mathbf{x}_{m+1} = \mathbf{x}_m - [\mathbf{f}'(\mathbf{x}_m)]^{-1} \mathbf{f}(\mathbf{x}_m) \quad (m = 0, 1, 2, \dots)$$

mit der Funktionalmatrix $\mathbf{f}'(\mathbf{x}) = [\partial f_i / \partial(x_j)(\mathbf{x})]_{1 \leq i, j \leq n}$.

Praxis: Die Inverse von $\mathbf{f}'(\mathbf{x}_m)$ wird **nicht** berechnet, sondern:

(a) Löse das $n \times n$ lineare Gleichungssystem (nach \mathbf{h})

$$\mathbf{f}'(\mathbf{x}_m) \mathbf{h} = -\mathbf{f}(\mathbf{x}_m),$$

(b) Setze

$$\mathbf{x}_{m+1} = \mathbf{x}_m + \mathbf{h}.$$

Beispiel 2 (vgl. Beispiel 1). Um die Lösung $\mathbf{x}^* \in \mathbb{R}^n$ von

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{f} \left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right) := \begin{bmatrix} 0.7 \sin x_1 + 0.2 \cos x_2 \\ 0.7 \cos x_1 - 0.2 \sin x_2 \end{bmatrix}$$

mit dem Newton-Verfahren berechnen zu können, müssen wir das Problem zuerst in **Nullstellenform** formulieren. \mathbf{x}^* ist Nullstelle von etwa $\mathbf{g}(\mathbf{x}) = \mathbf{x} - \mathbf{f}(\mathbf{x})$,

$$\mathbf{g} : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \mapsto \begin{bmatrix} x_1 - 0.7 \sin x_1 - 0.2 \cos x_2 \\ x_2 - 0.7 \cos x_1 + 0.2 \sin x_2 \end{bmatrix}.$$

Es gilt:

$$\mathbf{g}'(\mathbf{x}) = \mathbf{g}' \left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right) = \begin{bmatrix} 1 - 0.7 \cos x_1 & 0.2 \sin x_2 \\ 0.7 \sin x_1 & 1 + 0.2 \cos x_2 \end{bmatrix}.$$

Folglich

$$[\mathbf{g}'(\mathbf{x})]^{-1} = \left[\mathbf{g}' \left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right) \right]^{-1} = \frac{1}{D} \begin{bmatrix} 1 + 0.2 \cos x_2 & -0.2 \sin x_2 \\ -0.7 \sin x_1 & 1 - 0.7 \cos x_1 \end{bmatrix}$$

mit der Funktionaldeterminanten

$$D = \det \mathbf{g}'(\mathbf{x}) = (1 - 0.7 \cos x_1)(1 + 0.2 \cos x_2) - 0.14 \sin x_1 \sin x_2$$

(beachte

$$D \geq 0.3 \cdot 0.8 - 0.14 = 0.1 > 0 \quad \forall \mathbf{x} \in \mathbb{R}^2,$$

so dass $\mathbf{g}'(\mathbf{x})$ für jedes $\mathbf{x} \in \mathbb{R}^2$ invertierbar ist).

Iterationsvorschrift in unserem Beispiel:

$$\begin{bmatrix} x_1^{(m+1)} \\ x_2^{(m+1)} \end{bmatrix} = \begin{bmatrix} x_1^{(m)} \\ x_2^{(m)} \end{bmatrix} -$$

$$\frac{1}{D} \begin{bmatrix} 1 + 0.2 \cos x_2^{(m)} & -0.2 \sin x_2^{(m)} \\ -0.7 \sin x_1^{(m)} & 1 - 0.7 \cos x_1^{(m)} \end{bmatrix} \begin{bmatrix} x_1^{(m)} - 0.7 \sin x_1^{(m)} - 0.2 \cos x_2^{(m)} \\ x_2^{(m)} - 0.7 \cos x_1^{(m)} + 0.2 \sin x_2^{(m)} \end{bmatrix}$$

(wir weisen noch einmal darauf hin, dass die Invertierung von $g'(\mathbf{x}_m)$ nur im Ein- oder Zweidimensionalen durchgeführt werden kann — i.A. wird man, wie oben beschrieben, ein lineares Gleichungssystem mit der Matrix $g'(\mathbf{x}_m)$ lösen).

Die numerische Rechnung liefert:

m	$x_1^{(m)}$	$x_2^{(m)}$	$\ \mathbf{x}^* - \mathbf{x}_m\ _2$
0	0	0	7.3158e-01
1	0.6666666666666667	0.5833333333333333	1.5915e-01
2	0.53623996436649	0.50884900677977	9.7617e-03
3	0.52656199042839	0.50793192959763	4.1219e-05
4	0.52652262259473	0.50791971922729	7.0284e-10
5	0.52652262191818	0.50791971903685	< eps = 2.2204e-16

(vgl. die Konvergenzgeschwindigkeit der gewöhnlichen Fixpunktiteration).

8.5 Konvergenz des Newton-Verfahrens

Lemma 8.4. *Die Funktion $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ sei stetig differenzierbar und genüge in der konvexen Menge D der Bedingung*

$$\|f'(u) - f'(v)\| \leq \gamma \|u - v\|$$

für beliebige $u, v \in D$, d.h. f' sei Lipschitz-stetig in D . Dann gilt für $x, y \in D$

$$\|f(y) - f(x) - f'(x)(y - x)\| \leq \frac{\gamma}{2} \|x - y\|^2.$$

Lemma 8.5. *Die Funktion $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ sei differenzierbar an der Stelle $x^* \in D$ mit $f(x^*) = 0$. Sei ferner die matrixwertige Funktion $A : D \rightarrow \mathbb{R}^{n \times n}$ stetig in x^* und die Matrix $A(x^*)$ invertierbar. Dann ist die Funktion*

$$g : D \rightarrow \mathbb{R}^n, \quad x \mapsto x - A(x)^{-1}f(x)$$

in einer Umgebung von x^ wohldefiniert und differenzierbar in x^* mit $g'(x^*) = I - A(x^*)^{-1}f'(x^*)$.*

Satz 8.6 (Lokale Konvergenz des Newton-Verfahrens). *Die Funktion $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ besitze die Nullstelle x^* , sei stetig differenzierbar an der Stelle x^* und $f'(x^*)$ sei invertierbar. Dann ist x^* anziehender Fixpunkt der Newton-Iteration*

$$\mathbf{x}_{m+1} = \mathbf{x}_m - \mathbf{f}'(\mathbf{x}_m)^{-1} \mathbf{f}(\mathbf{x}_m) \quad m = 0, 1, \dots,$$

und es gilt

$$\lim_{m \rightarrow \infty} \frac{\|\mathbf{x}_{m+1} - \mathbf{x}^*\|}{\|\mathbf{x}_m - \mathbf{x}^*\|} = 0. \quad (8.2)$$

Falls in einer Umgebung $U(x^)$ von x^* gilt*

$$\|\mathbf{f}'(\mathbf{x}) - \mathbf{f}'(\mathbf{x}^*)\| \leq \alpha \|\mathbf{x} - \mathbf{x}^*\| \quad \forall \mathbf{x} \in U(\mathbf{x}^*), \quad (8.3)$$

dann existieren $C > 0$ und $m_0 \in \mathbb{N}$ mit

$$\|\mathbf{x}_{m+1} - \mathbf{x}^*\| \leq C \|\mathbf{x}_m - \mathbf{x}^*\|^2 \quad \text{für } m \geq m_0.$$

Satz 8.7 (Newton-Kantorovich). *Es sei $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ differenzierbar in $D \subset \mathbb{R}^n$, D konvex; für $x, y \in D$ gelte*

$$\|f'(x) - f'(y)\| \leq \gamma \|x - y\|.$$

Es gebe ein $x_0 \in D$ mit

$$\beta := \|f'(x_0)^{-1}\|, \quad \eta := \|f'(x_0)^{-1}f(x_0)\|$$

sodass $\alpha := \beta\gamma\eta < \frac{1}{2}$ ist. Ferner sei $S := \{x \in \mathbb{R}^n : \|x - x_0\| \leq t^\} \subset D$ mit $t^* = (1 - \sqrt{1 - 2\alpha})/(\beta\gamma)$. Dann gilt: mit dem in den Voraussetzungen genannten x_0 sind die Iterierten des Newton-Verfahrens*

$$x_{m+1} = x_m - f'(x_m)^{-1}f(x_m), \quad m = 1, 2, \dots,$$

wohldefiniert und konvergieren gegen eine Lösung $x^ \in D$ von $f(x) = 0$. Es gilt die Fehlerabschätzung*

$$\|x^* - x_1\| \leq 2\beta\gamma \|x_1 - x_0\|^2.$$

Bemerkungen:

1. Der Satz von Newton-Kantorovich liefert eine quantitative Charakterisierung der lokal quadratischen Konvergenz des Newton-Verfahrens im \mathbb{R}^n .
2. Neben der lokal quadratischen Konvergenz sichert er auch die Existenz einer Nullstelle in einer t^* -Kugel um x^* .
3. Die (recht starken) Voraussetzungen sind in der Praxis selten zu verifizieren. Es kommt letztlich immer darauf an, einen hinreichend guten Startwert x_0 zu wählen.

Einfache (eindimensionale) Beispiele zeigen, dass das Newton-Verfahren nur in Ausnahmefällen global konvergiert. Unter speziellen Voraussetzungen kann globale Konvergenz gesichert werden.

Definition 8.8. Eine Menge $D \subset \mathbb{R}^n$ heißt *konvex*, wenn sie mit zwei Punkten x, y auch deren Verbindungsstrecke enthält:

$$x, y \in D \Rightarrow \alpha x + (1 - \alpha)y \in D \quad \forall \alpha \in [0, 1].$$

Die Abbildung $f : \mathbb{R}^n \supset D \rightarrow \mathbb{R}^n$, D konvex, heißt *konvex in D* , wenn

$$f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha)f(y)$$

für alle $x, y \in D$ und alle $\alpha \in [0, 1]$ gilt. Die Relation \leq ist hier *komponentenweise zu verstehen*.

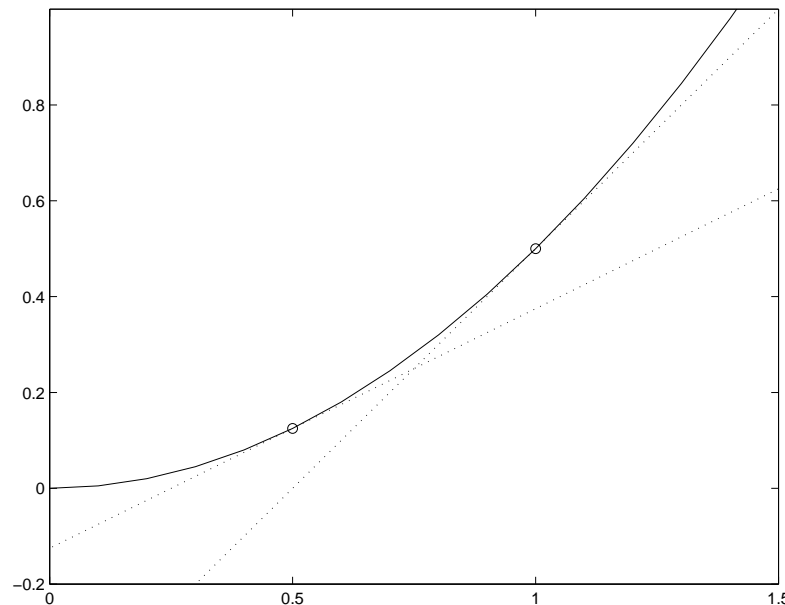
Interpretation im Eindimensionalen: Der Graph von f liegt unterhalb aller Sekanten.

Lemma 8.9. *Ist f stetig differenzierbar in der konvexen Menge D , so ist f genau dann konvex, wenn für alle $x, y \in D$*

$$f(y) - f(x) \geq f'(x)(y - x) \quad (8.4)$$

gilt.

Interpretation im Eindimensionalen: Der Graph von f liegt oberhalb aller Tangenten.



Satz 8.10 (Newton-Baluev). *Es sei $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ stetig differenzierbar und konvex. $f'(x)$ sei invertierbar mit $[f'(x)]^{-1} \geq O$ (komponentenweise) für alle $x \in \mathbb{R}^n$. Außerdem besitze $f(x) = \mathbf{0}$ eine Lösung x^* . Dann gilt:*

(a) x^* ist die einzige Lösung von $f(x) = \mathbf{0}$.

(b) Die Newton-Folge

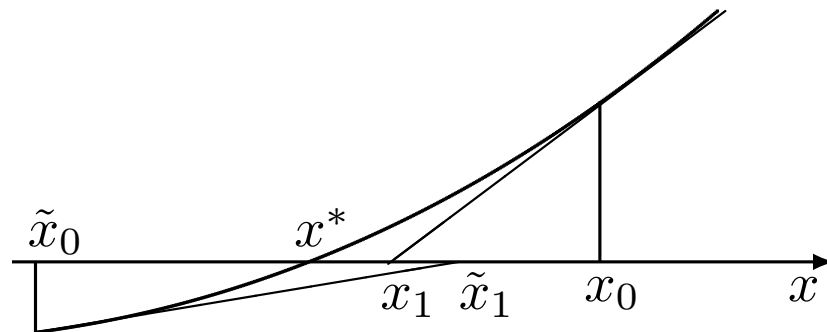
$$\mathbf{x}_{m+1} = \mathbf{x}_m - f'(\mathbf{x}_m)^{-1} f(\mathbf{x}_m), \quad m = 0, 1, \dots$$

konvergiert für beliebiges $x_0 \in \mathbb{R}^n$ gegen x^* (globale Konvergenz!),

(c) Es gilt (komponentenweise monotone Konvergenz)

$$\mathbf{x}^* \leq \mathbf{x}_{m+1} \leq \mathbf{x}_m \quad \text{für } m = 1, 2, \dots$$

1-D:



Bemerkung. Genügt f zusätzlich der Monotoniebedingung

$$f'(x) \leq f'(y) \quad \text{falls } x \leq y,$$

so lässt sich eine Einschliessung von x^* konstruieren:

Neben der Newton-Folge wird noch die Folge

$$\hat{x}_{m+1} := \hat{x}_m - f'(x_m)^{-1} f(\hat{x}_m) \quad (m = 1, 2, \dots)$$

(Funktionalmatrix an der Stelle x_m , nicht an der Stelle \hat{x}_m !) berechnet.

Ist $\hat{x}_1 \in D$ so gewählt, dass $f(\hat{x}_1) \leq 0$ gilt, so folgt

$$\hat{x}_m \leq \hat{x}_{m+1} \leq x^* \leq x_{m+1} \leq x_m \quad (m = 1, 2, \dots)$$

und

$$\lim_{m \rightarrow \infty} \hat{x}_m = \lim_{m \rightarrow \infty} x_m = x^*.$$

8.6 Modifikationen des Newton-Verfahrens

Problem beim Newton-Verfahren: Nur lokale Konvergenz.

Idee des **gedämpften Newton-Verfahrens**:

Iteriere gemäß

$$\mathbf{x}_{m+1} = \mathbf{x}_m - \alpha_m \mathbf{f}'(\mathbf{x}_m)^{-1} \mathbf{f}(\mathbf{x}_m) \quad (m = 0, 1, 2, \dots),$$

dabei ist $\alpha_m \in (0, 1]$ ein Dämpfungsfaktor, der so gewählt wird, dass

$$\|\mathbf{f}(\mathbf{x}_0)\| \geq \|\mathbf{f}(\mathbf{x}_1)\| \geq \|\mathbf{f}(\mathbf{x}_2)\| \geq \dots$$

erfüllt ist.

Programmwurf:

```
m = 0
while norm(f(x_m)) > tol
    Loese f'(x_m)* h = -f(x_m)
    alpha = 1
    x = x_m + alpha*h
    while norm(f(x)) > (1-0.25*alpha)*norm(f(x_m))
        alpha = 0.5*alpha
        if alpha < 2^(-10), break, end if
        x = x + alpha*h
    end while
    m = m+1
    x_m = x
end while
```

Problem beim Newton-Verfahren: Berechnung von $f'(\mathbf{x}_m)$ teuer.

Idee der **Quasi-Newton-Verfahren:**

Iteriere gemäß

$$\mathbf{x}_{m+1} = \mathbf{x}_m - A_m^{-1} \mathbf{f}(\mathbf{x}_m) \quad (m = 0, 1, 2, \dots), \quad (8.5)$$

dabei ist A_m eine Näherung an $f'(\mathbf{x}_m)$, die der **Quasi-Newton-Bedingung**

$$A_m(\mathbf{x}_m - \mathbf{x}_{m-1}) = \mathbf{f}(\mathbf{x}_m) - \mathbf{f}(\mathbf{x}_{m-1}) \quad (8.6)$$

genügt. Populär ist der Ansatz

$$A_m = A_{m-1} + \mathbf{u}_m \mathbf{v}_m^\top$$

(**Rang-Eins-Verfahren**), weil sich A_m^{-1} dann leicht aus A_{m-1}^{-1} berechnen lässt:

Beim **Broyden-Verfahren** wird neben

$$(A_m - A_{m-1})(\mathbf{x}_m - \mathbf{x}_{m-1}) = \mathbf{f}(\mathbf{x}_m) - \mathbf{f}(\mathbf{x}_{m-1}) - A_{m-1}(\mathbf{x}_m - \mathbf{x}_{m-1})$$

(folgt aus der Quasi-Newton-Bedingung) noch

$$(A_m - A_{m-1})\mathbf{h} = \mathbf{0} \text{ für alle } \mathbf{h} \perp \mathbf{x}_m - \mathbf{x}_{m-1}$$

gefordert. Damit ist A_m eindeutig bestimmt. Es ergibt sich

$$A_m = A_{m-1} + \mathbf{u}_m \mathbf{v}_m^\top$$

mit

$$\mathbf{v}_m := \mathbf{x}_m - \mathbf{x}_{m-1}, \quad \mathbf{u}_m := \frac{(\mathbf{f}(\mathbf{x}_m) - \mathbf{f}(\mathbf{x}_{m-1}) - A_{m-1}(\mathbf{x}_m - \mathbf{x}_{m-1})) \mathbf{v}_m^\top}{\mathbf{v}_m^\top \mathbf{v}_m}.$$

Die Berechnung der nächsten Iterierten \mathbf{x}_{m+1} nach (8.5) erfordert zunächst $O(n^3)$ Operationen. Durch Aufdatieren der Inversen kann man den Aufwand auf $O(n^2)$ reduzieren:

Satz 8.11 (Sherman-Morrison-Formel). *Sind $u, v \in \mathbb{R}^n$ und $A \in \mathbb{R}^{n \times n}$ invertierbar. Dann ist $A + uv^\top$ genau dann invertierbar, wenn $1 + v^\top A^{-1}u \neq 0$ gilt. In diesem Fall ist*

$$(A + uv^\top)^{-1} = A^{-1} - \frac{A^{-1}uv^\top A^{-1}}{1 + v^\top A^{-1}u}.$$

Satz 8.12. *Die Abbildung $f : \mathbb{R}^n \supseteq D \rightarrow \mathbb{R}^n$ besitze eine Nullstelle $x^* \in D$. f sei differenzierbar in einer Umgebung U von x^* und f' sei in U Lipschitzstetig. Außerdem sei $f'(x^*)$ invertierbar.*

Dann gibt es positive Konstanten δ und ρ mit den folgenden Eigenschaften: Für alle Startwerte x_0 mit $\|x_0 - x^\| \leq \delta$ und alle Näherungen A_0 (für $f'(x_0)$) mit $\|A_0 - f'(x_0)\| \leq \rho$ konvergiert das Broyden-Verfahren superlinear gegen x^* .*

8.7 Nichtlineare Ausgleichsprobleme

Gegeben sind $f : \mathbb{R}^n \supseteq D \rightarrow \mathbb{R}^m$ ($m \geq n$) mit den Komponenten $f_1, \dots, f_m : D \rightarrow \mathbb{R}$ und $\mathbf{y} \in \mathbb{R}^m$.

Gesucht ist ein Vektor $\mathbf{x}^* \in \mathbb{R}^n$ mit

$$\|\mathbf{y} - \mathbf{f}(\mathbf{x}^*)\|_2 = \min_{\mathbf{x} \in D} \|\mathbf{y} - \mathbf{f}(\mathbf{x})\|_2 \quad (8.7)$$

In Kapitel 5 wurde der lineare Spezialfall dieses Problems behandelt, in dem

$$\mathbf{f}(\mathbf{x}) = A\mathbf{x}, \quad A \in \mathbb{R}^{m \times n}.$$

Beispiel. n Parameter einer Modellfunktion φ sollen an m Messwerte angepasst werden. Etwa:

$$\varphi(t) = \varphi(t; x_1, x_2, x_3) = x_1 \exp(-x_2 t) + x_3.$$

Als Daten seien folgende fünf Messungen gegeben:

j	1	2	3	4	5
t_j	0	1	2	3	5
y_j	3	2	1.5	1	0.5

Hier ist also:

$$\mathbf{f} = \begin{bmatrix} f_1 \\ \vdots \\ f_5 \end{bmatrix} : \mathbb{R}^3 \rightarrow \mathbb{R}^5, \quad f_j(x_1, x_2, x_3) = x_1 \exp(-x_2 t_j) + x_3, \quad j = 1, \dots, 5.$$

Setzt man $F : \mathbb{R}^n \supseteq D \rightarrow \mathbb{R}$, $\mathbf{x} \mapsto \|\mathbf{f}(\mathbf{x}) - \mathbf{y}\|_2^2$, so ist

$$\frac{1}{2} \nabla F(\mathbf{x}^*) = \mathbf{f}'(\mathbf{x}^*)^\top [\mathbf{f}(\mathbf{x}^*) - \mathbf{y}] = \mathbf{0}$$

eine notwendige Bedingung für die Lösung \mathbf{x}^* von (8.7) (falls $\mathbf{f} \in C^1(D)$).

Man könnte dieses Gleichungssystem mit dem Newton-Verfahren attackieren. Dies ist i.A. zu kompliziert (um F' zu bestimmen, benötigt man \mathbf{f}'') und man linearisiert (8.7) direkt. Ausgangspunkt ist die Taylorreihe von \mathbf{f} ($\mathbf{f} \in C^2(D)$)

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{x}_0) + \mathbf{f}'(\mathbf{x}_0)[\mathbf{x} - \mathbf{x}_0] + o(\|\mathbf{x} - \mathbf{x}_0\|).$$

Ist x_0 eine gute Näherung von x^* , so erwartet man, dass die Lösung x_1 von

$$\min_{x \in \mathbb{R}^n} \|y - f(x_0) - f'(x_0)[x - x_0]\|_2 \quad \left(\approx \min_{x \in \mathbb{R}^n} \|y - f(x)\|_2 \right) \quad (8.8)$$

eine bessere Approximation an x^* ist. (8.8) ist ein **lineares** Ausgleichsproblem, kann also mit den Methoden aus Kapitel 5 gelöst werden.

Gauß-Newton-Verfahren:

Wähle x_0 .

for $k = 0, 1, 2, \dots$:

Berechne $A = f'(x_k)$ und $b = y - f(x_k)$,

bestimme die Lösung h_k des linearen Ausgleichsproblems

$\|b - A h_k\|_2 = \min$,

setze $x_{k+1} = x_k + h_k$.

end for

In der Praxis muss dieses Verfahren gedämpft werden ($\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha \mathbf{h}_k$, vgl. Abschnitt 8.6). Die folgende Tabelle zeigt – für unser Beispiel – das Residuum $\|\mathbf{y} - \mathbf{f}(\mathbf{x}_k)\|_2$ der Iterierten \mathbf{x}_k des Gauß-Newton-Verfahrens mit und ohne Dämpfung ($\mathbf{x}_0 = \mathbf{0}$, in Klammern: Dämpfungsfaktoren):

k	ohne Dämpfung	mit Dämpfung	
0	4.0620	4.0620	
1	1.9235	1.9235	[1.0]
2	1.1281	1.1281	[1.0]
3	1.7878	0.8317	[0.5]
4	3.8386	0.0937	[1.0]
5	2.3371	0.0931	[1.0]
6	$\approx 10^{10}$!!	0.0931	[1.0]

Satz 8.13. Sei $h = x - x_0$ die Lösung des linearen Ausgleichsproblems (8.8), $f'(x_0)$ besitze vollen Rang und $y - f(x_0) \neq 0$. Dann existiert $\hat{\alpha} > 0$ so dass die Funktion $\varphi(\alpha) = \|y - f(x_0 + \alpha h)\|_2^2$ für alle $\alpha \in [0, \hat{\alpha}]$ streng monoton fällt. Insbesondere gilt

$$\varphi(\alpha) = \|y - f(x_0 + \alpha h)\|_2^2 < \varphi(0) = \|y - f(x_0)\|_2^2.$$

Die Gauß-Newton-Korrektur h ist also eine **Abstiegsrichtung**. Ein geeignet gedämpftes Gauß-Newton-Verfahren führt somit immer zu einer monoton abnehmenden Residualnorm der Iterierten.

Das **Levenberg-Marquardt-Verfahren** wählt bei gegebenem $\lambda_k > 0$ die Korrekturrichtung \mathbf{h}_k für \mathbf{x}_k als die Lösung von

$$\|[\mathbf{y} - \mathbf{f}(\mathbf{x}_k)] - \mathbf{f}'(\mathbf{x}_k)\mathbf{h}_k\|_2^2 + \lambda_k^2 \|\mathbf{h}_k\|_2^2 \rightarrow \min_{\mathbf{h}_k \in \mathbb{R}^n} . \quad (8.9)$$

Die Idee ist, λ_k klein zu wählen, wenn \mathbf{x}_k weit von einer Lösung entfernt ist und λ_k groß zu wählen, wenn \mathbf{x}_k eine gute Näherung für die Lösung ist.

Eine äquivalente Formulierung von (8.9) ist

$$\left\| \begin{bmatrix} \mathbf{y} - \mathbf{f}(\mathbf{x}_k) \\ \mathbf{0} \end{bmatrix} - \begin{bmatrix} \mathbf{f}'(\mathbf{x}_k) \\ \lambda_k I_n \end{bmatrix} \mathbf{h}_k \right\|_2^2 \rightarrow \min_{\mathbf{h}_k \in \mathbb{R}^n} .$$

Es handelt sich also um ein lineares Ausgleichsproblem der Dimension $(m + n) \times n$, dessen Koeffizientenmatrix stets vollen Rang n besitzt.

Die sehr spezielle Struktur dieses Kleinsten-Quadrate-Problems muss bei der Rechnung ausgenutzt werden.

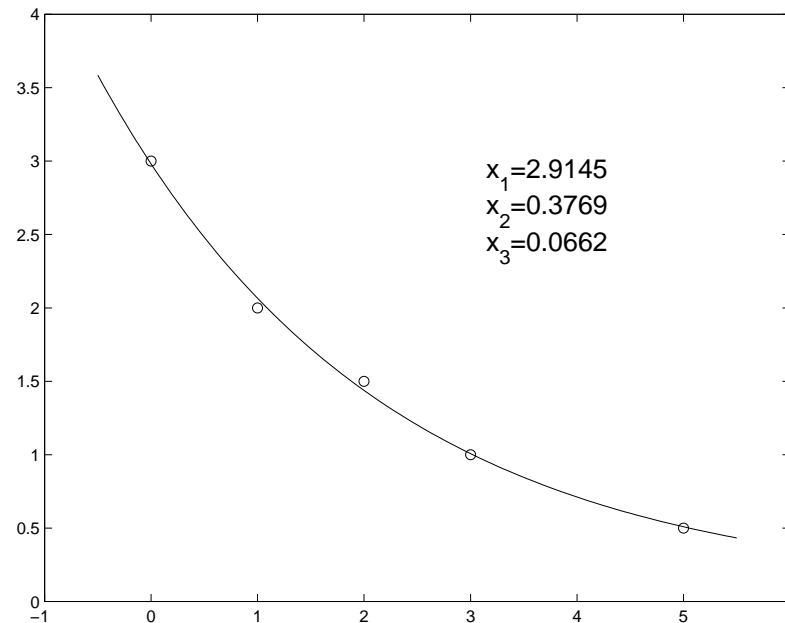
Auch die Korrektur \mathbf{h}_k des Marquardt-Verfahrens ist eine Abstiegsrichtung.

Marquardt-Verfahren:

```
Waehle  $x_0$  und  $\lambda_0 > 0$ .
for  $k = 0, 1, 2, \dots$ :
    Berechne  $A = f'(x_k)$  und  $b = y - f(x_k)$ ,
    bestimme die Loesung  $h_k$  des linearen Ausgleichsproblems
 $\|b - A h_k\|^2 + \lambda_k^2 \|h_k\|^2 = \min$ ,
    setze  $x_{\text{test}} = x_k + h_k$ .
    while  $\|y - f(x_{\text{test}})\| \geq \|y - f(x_k)\|$ 
        Setze  $\lambda_k = 2 \lambda_k$  und loese
 $\|b - A h_k\|^2 + \lambda_k^2 \|h_k\|^2 = \min$ ,
        setze  $x_{\text{test}} = x_k + h_k$ .
    end while
    Setze  $x_{k+1} = x_{\text{test}}$ 
    und  $\lambda_{k+1} = 0.5 \lambda_k$ .
end for
```

In unserem Beispiel ist das Marquardt-Verfahren etwa so schnell wie das gedämpfte Gauß-Newton-Verfahren ($r_k := y - f(x_k)$, $x_0 = \mathbf{0}$, $\lambda_0 = 1$):

k	0	1	2	3	4	5	6
$\ r_k\ _2$	4.0620	1.9508	1.0737	0.5059	0.0983	0.0933	0.0931
λ_k	1	1/2	1/4	1/8	1/16	1/32	1/64



8.8 Klassische Iterationsverfahren für lineare Systeme

Zu lösen: $Ax = b$, $A \in \mathbb{R}^{n \times n}$ invertierbar, aber n so groß, dass Gauß-Elimination nicht durchführbar ist.

Mit einer **Zerlegung** von A ,

$$A = M - N \quad (M, N \in \mathbb{R}^{n \times n}, M \text{ invertierbar}),$$

lässt sich das LGS äquivalent als Fixpunktaufgabe schreiben:

$$Mx = Nx + b \quad \text{oder} \quad x = M^{-1}Nx + M^{-1}b.$$

Zugehörige Fixpunktiteration

$$Mx_{m+1} = Nx_m + b \quad \text{oder} \quad x_{m+1} = M^{-1}Nx_m + M^{-1}b \quad (m = 0, 1, 2, \dots).$$

Bezeichnungen:

$$\mathbf{e}_m = A^{-1} \mathbf{b} - \mathbf{x}_m \text{ (**Fehler**)},$$

$$\mathbf{r}_m = \mathbf{b} - A\mathbf{x}_m \text{ (**Residuum**)},$$

$$\tilde{\mathbf{r}}_m = M^{-1}(\mathbf{b} - A\mathbf{x}_m) \text{ (**Pseudo-Residuum**)}.$$

Mit $T = M^{-1}N$ (Iterationsmatrix) folgt

$$\mathbf{e}_m = T\mathbf{e}_{m-1} = \cdots = T^m \mathbf{e}_0 \text{ und } \tilde{\mathbf{r}}_m = T\tilde{\mathbf{r}}_{m-1} = \cdots = T^m \tilde{\mathbf{r}}_0.$$

Satz 8.14. *Das Iterationsverfahren*

$$\mathbf{x}_{m+1} = M^{-1}N\mathbf{x}_m + M^{-1}\mathbf{b} \quad (m = 0, 1, 2, \dots) \quad (8.10)$$

konvergiert genau dann – für jeden Startvektor $\mathbf{x}_0 \in \mathbb{R}^n$ – gegen $A^{-1}\mathbf{b}$, wenn es eine Matrixnorm $\|\cdot\|_M$ gibt mit

$$\|T\|_M = \|M^{-1}N\|_M < 1.$$

Korollar 8.15. *Das Iterationsverfahren (8.10) konvergiert genau dann für jeden Startvektor $\mathbf{x}_0 \in \mathbb{R}^n$, wenn $\rho(T) < 1$.*

Eine Matrix T mit $\rho(T) < 1$ heißt **konvergente Matrix**.

Es gilt:

$$\mathbf{x}_{m+1} = M^{-1}N\mathbf{x}_m + M^{-1}\mathbf{b} = \mathbf{x}_m + M^{-1}\mathbf{r}_m$$

M wird i.a. nicht invertiert:

```
m = 0
r_0 = b - A*x_0
while norm(r_m) > tol*norm(b)
    m = m+1
    Loese M*h = r_{m-1}
    x_m = x_{m-1}+h
    r_m = b-A*x_m
end while
```

Lineare Gleichungssysteme mit Koeffizientenmatrix M müssen sehr viel kostengünstiger zu lösen sein als solche mit A !

Die Wahl $M = \text{diag}(A)$ führt zum **Jacobi-** oder **Gesamtschrittverfahren**:

$$x_j^{(m)} = \frac{1}{a_{j,j}} \left[b_j - \sum_{k \neq j} a_{j,k} x_k^{(m-1)} \right] \quad (1 \leq j \leq n, m > 0).$$

Die Wahl $M = \text{tril}(A)$ führt zum **Gauß-Seidel-** oder **Einzelschrittverfahren**:

$$x_j^{(m)} = \frac{1}{a_{j,j}} \left[b_j - \sum_{k < j} a_{j,k} x_k^{(m)} - \sum_{k > j} a_{j,k} x_k^{(m-1)} \right] \quad (1 \leq j \leq n, m > 0).$$

Unterschied: Beim Jacobi-Verfahren wird x_m ausschließlich aus Komponenten von x_{m-1} berechnet. Beim Gauß-Seidel-Verfahren werden zur Berechnung von $x_j^{(m)}$ alle verfügbaren Komponenten von x_m , d.h. $x_k^{(m)}$ für $k < j$, verwendet.

Definition 8.16. *Eine Matrix $A \in \mathbb{R}^{n \times n}$ heißt **streng diagonaldominant**, falls*

$$\sum_{k \neq j} |a_{j,k}| < |a_{j,j}| \quad \forall j = 1, 2, \dots, n.$$

Satz 8.17. *Ist A streng diagonaldominant, so konvergieren sowohl das Jacobi-Verfahren als auch das Gauß-Seidel Verfahren für jeden Startvektor.*

Dieses Kriterium nennt man auch **starkes Zeilensummenkriterium**. Analog kann man ein starkes Spaltensummenkriterium herleiten.

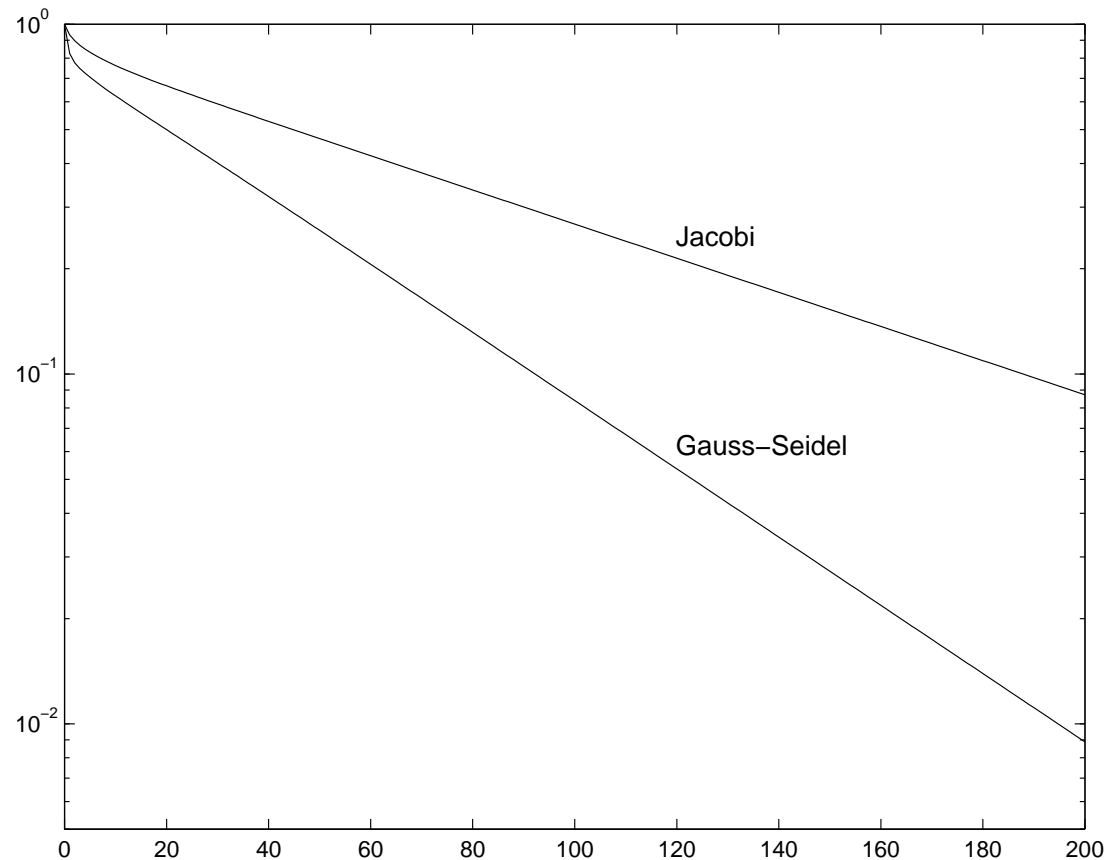
Definition 8.18. $A = M - N$, $A, M, N \in \mathbb{R}^{n \times n}$ heißt **reguläre Zerlegung** von A falls $M^{-1} \geq O$ und $N \geq O$. (Die Ungleichungen sind wieder komponentenweise zu verstehen.)

Insbesondere gilt für reguläre Zerlegungen $T = M^{-1}N \geq O$.

Satz 8.19. Ist $A \in \mathbb{R}^{n \times n}$ nichtsingulär, gilt $A^{-1} \geq O$ und ist $A = M - N$ eine reguläre Zerlegung, so gilt $\rho(M^{-1}N) < 1$.

Definition 8.20. Eine Matrix $A = [a_{i,j}] \in \mathbb{R}^{n \times n}$ heißt **M-Matrix**, falls $A^{-1} \geq O$ und $a_{i,j} \leq 0$ falls $i \neq j$.

Korollar 8.21. Ist A eine M-Matrix, so konvergieren Jacobi- und Gauß-Seidel Verfahren für jeden Startvektor.



Hier: $n = 20$. Die Konvergenz ist in beiden Fällen linear mit den Konvergenzfaktoren $.989 \dots$ (Jacobi) bzw. $.978 \dots$ (Gauß-Seidel).
Vergrößert man n , so konvergieren beide Verfahren (noch) langsamer.

8.9 Gradientenverfahren

In diesem Abschnitt sei $A \in \mathbb{R}^{n \times n}$ stets symmetrisch und positiv definit, $\mathbf{b} \in \mathbb{R}^n$ beliebig und (\cdot, \cdot) das Euklidische Innenprodukt in \mathbb{R}^n .

Grundlegendes Resultat: Unter dieser Voraussetzung löst $\mathbf{x}^* \in \mathbb{R}^n$ das lineare Gleichungssystem $A\mathbf{x} = \mathbf{b}$ genau dann, wenn \mathbf{x}^* die Funktion

$$\phi : \mathbb{R}^n \rightarrow \mathbb{R}, \quad \mathbf{x} \mapsto \phi(\mathbf{x}) := \frac{1}{2} (A\mathbf{x}, \mathbf{x}) - (\mathbf{b}, \mathbf{x})$$

minimiert, d.h. wenn

$$\phi(\mathbf{x}^*) = \min_{\mathbf{x} \in \mathbb{R}^n} \phi(\mathbf{x})$$

gilt.

Dies folgt unmittelbar aus den Identitäten

$$\phi(\mathbf{x}) = \frac{1}{2} \|\mathbf{x}^* - \mathbf{x}\|_A^2 + \phi(\mathbf{x}^*) = \frac{1}{2} (\|\mathbf{b} - A\mathbf{x}\|_{A^{-1}}^2 - \|\mathbf{b}\|_{A^{-1}}^2).$$

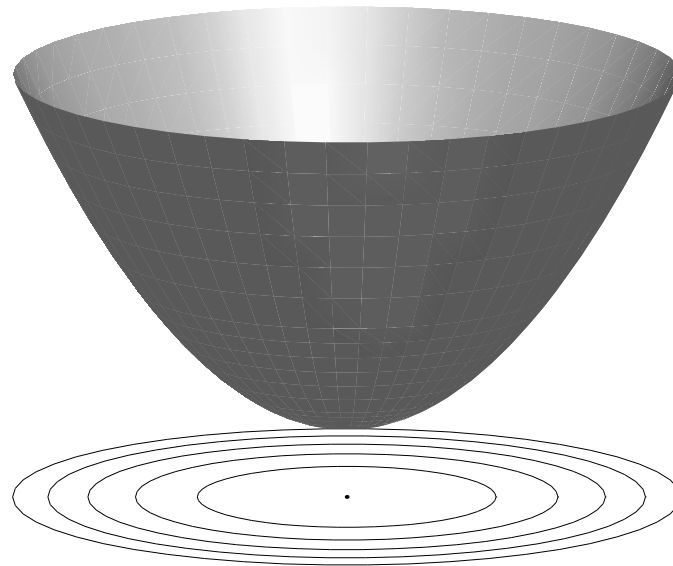
Hierbei sind $\|\cdot\|_A$ bzw. $\|\cdot\|_{A^{-1}}$ die zu den Innenprodukten

$$(\mathbf{x}, \mathbf{y})_A := \mathbf{y}^\top A\mathbf{x} \quad \text{bzw.} \quad (\mathbf{x}, \mathbf{y})_{A^{-1}} := \mathbf{y}^\top A^{-1}\mathbf{x}, \quad \mathbf{x}, \mathbf{y} \in \mathbb{R}^n,$$

gehörenden Normen auf \mathbb{R}^n . („ A -Norm“ und „ A^{-1} -Norm“).

Fazit: Das Lösen des linearen Gleichungssystems $A\mathbf{x} = \mathbf{b}$ mit symmetrisch positiv definiten Koeffizientenmatrix A ist äquivalent mit der Minimierung des quadratischen Funktionals $\phi(\mathbf{x}) = \frac{1}{2}(A\mathbf{x}, \mathbf{x}) - (\mathbf{b}, \mathbf{x})$. Daher können auch Minimierungsverfahren zur Lösung eingesetzt werden.

Graph und Niveaulinien der Funktion ϕ (für $n = 2$):



Eindimensionale Minimierung

Gegeben: \mathbf{x} : Näherung an \mathbf{x}^* , zugehöriges Residuum $\mathbf{r} = \mathbf{b} - A\mathbf{x}$,
 \mathbf{p} : „Suchrichtung“ ,

Gesucht: $\min\{\phi(\mathbf{x} + \alpha\mathbf{p}) : \alpha \in \mathbb{R}\}$

$$\begin{aligned}\text{Es ist} \quad \phi(\mathbf{x} + \alpha\mathbf{p}) &= \frac{1}{2}(A(\mathbf{x} + \alpha\mathbf{p}), \mathbf{x} + \alpha\mathbf{p}) - (\mathbf{b}, \mathbf{x} + \alpha\mathbf{p}) \\ &= \phi(\mathbf{x}) - \alpha(\mathbf{b} - A\mathbf{x}, \mathbf{p}) + \frac{\alpha^2}{2}(A\mathbf{p}, \mathbf{p})\end{aligned}$$

Da A spd wird das Minimum erzielt für

$$\alpha = \frac{(\mathbf{b} - A\mathbf{x}, \mathbf{p})}{(A\mathbf{p}, \mathbf{p})} = \frac{(\mathbf{r}, \mathbf{p})}{(A\mathbf{p}, \mathbf{p})}, \quad \text{was auf } \phi(\mathbf{x} + \alpha\mathbf{p}) = \phi(\mathbf{x}) - \frac{1}{2} \frac{(\mathbf{r}, \mathbf{p})^2}{(A\mathbf{p}, \mathbf{p})}$$

führt. Notwendig ist somit $\mathbf{p} \not\perp \mathbf{r}$.

Strategie: Ist \mathbf{x}_{m-1} eine Näherung für \mathbf{x}^* , so bestimmt man eine Suchrichtung $\mathbf{p}_m \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ und minimiert ϕ entlang der Geraden $\{\mathbf{x}_{m-1} + \alpha \mathbf{p}_m : \alpha \in \mathbb{R}\}$. Dies führt auf die nächste Approximation

$$\mathbf{x}_m = \mathbf{x}_{m-1} + \frac{(\mathbf{r}_{m-1}, \mathbf{p}_m)}{(A\mathbf{p}_m, \mathbf{p}_m)} \mathbf{p}_m$$

(ersetze die *n-dimensionale* Minimierungsaufgabe $\phi(\mathbf{x}) = \min_{\mathbf{x} \in \mathbb{R}^n}$ durch eine Folge *eindimensionaler* Probleme $\phi(\mathbf{x}_{m-1} + \alpha \mathbf{p}_m) = \min_{\alpha \in \mathbb{R}}$).

Wahl der Suchrichtung: Naheliegender ist die (lokal optimale) Richtung des steilsten Abstiegs von ϕ , also die des (negativen) Gradienten

$$-\nabla \phi(\mathbf{x}) = -\nabla \left(\frac{1}{2} (A\mathbf{x}, \mathbf{x}) - (\mathbf{b}, \mathbf{x}) \right) = \mathbf{b} - A\mathbf{x} = \mathbf{r}.$$

Dies führt auf $\alpha_m = (\mathbf{r}_{m-1}, \mathbf{r}_{m-1}) / (A\mathbf{r}_{m-1}, \mathbf{r}_{m-1})$ und zum **Verfahren des steilsten Abstiegs**.

Verfahren des steilsten Abstiegs

$m := 0, x_0$ gegeben, $r_0 := b - Ax_0$

while $r_k \neq 0$

$m := m + 1$

$\alpha_m := (r_{m-1}, r_{m-1}) / (Ar_{m-1}, r_{m-1})$

$x_m := x_{m-1} + \alpha_m r_{m-1}$

$r_m := b - Ax_m$

Beachte: Wegen $b - Ax_m = b - A(x_{m-1} + \alpha_m r_{m-1}) = r_{m-1} - \alpha_m Ar_{m-1}$ ist nur ein Matrix-Vektorprodukt zu berechnen.

Satz 8.22. *Das Verfahren des steilsten Abstiegs zur Lösung von $Ax = b$ ist global linear konvergent.*

Bemerkung: Eine schärfere Abschätzung als im Beweis von Satz 8.22 liefert die Kantorovich-Ungleichung

$$\frac{(\mathbf{x}, \mathbf{x})^2}{(A\mathbf{x}, \mathbf{x})(A^{-1}\mathbf{x}, \mathbf{x})} \geq \frac{4\lambda_{\max}(A)\lambda_{\min}(A)}{(\lambda_{\max}(A) + \lambda_{\min}(A))^2}, \quad \forall \mathbf{x} \in \mathbb{R}^n,$$

was zur linearen Konvergenzrate

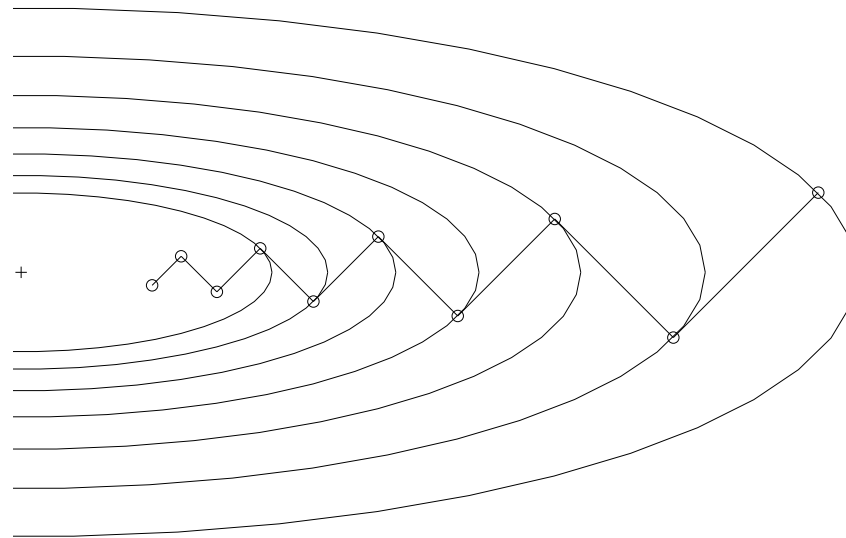
$$\frac{\kappa - 1}{\kappa + 1} < 1 - \frac{1}{\kappa}, \quad \kappa = \text{cond}_2(A)$$

des Verfahrens des steilsten Abstiegs führt.

Es zeigt sich allerdings, dass diese (lokal optimale) Wahl der Suchrichtung i.A. zu einem sehr langsam konvergenten Verfahren führt.

Konvergenz beim Verfahren des steilsten Abstiegs:

Beispiel: $A = \begin{bmatrix} 1 & 0 \\ 0 & \varepsilon \end{bmatrix}$, $b = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\kappa = \frac{1}{\varepsilon}$, $1 - \frac{1}{\kappa} = 1 - \varepsilon$.



Allgemeinere Suchrichtungen

Idee: Wähle m linear unabhängige Suchrichtungen $\mathbf{p}_1, \dots, \mathbf{p}_m$ so, dass die m -te Iterierte die Funktion ϕ minimiert über

$$\{\mathbf{x} = \mathbf{x}_0 + \mathbf{p} : \mathbf{p} \in \text{span}\{\mathbf{p}_1, \dots, \mathbf{p}_m\}\}.$$

Folge: Minimum nach spätestens n Schritten gefunden, da dann $\text{span}\{\mathbf{p}_1, \dots, \mathbf{p}_n\} = \mathbb{R}^n$.

Frage: Ist es möglich, eine globale Minimierung durch eine Folge lokaler Minimierungen zu realisieren?

A-konjugierte Suchrichtungen

Ist $P_m = [\mathbf{p}_1, \dots, \mathbf{p}_m] \in \mathbb{R}^{n \times m}$ die Matrix der ersten m Suchrichtungen, so ist $\mathbf{x} \in \mathbf{x}_0 + \text{span}\{\mathbf{p}_1, \dots, \mathbf{p}_m\}$ äquivalent mit

$$\mathbf{x} = \mathbf{x}_0 + P_{m-1} \mathbf{y} + \alpha \mathbf{p}_m, \quad \mathbf{y} \in \mathbb{R}^{m-1}, \alpha \in \mathbb{R}.$$

Dann gilt

$$\phi(\mathbf{x}) = \phi(\mathbf{x}_0 + P_{m-1} \mathbf{y}) + \underbrace{\alpha (A \mathbf{p}_m, P_{m-1} \mathbf{y})}_{(*)} + \frac{\alpha^2}{2} (A \mathbf{p}_m, \mathbf{p}_m) - \alpha (\mathbf{r}_0, \mathbf{p}_m)$$

Ohne den gemischten Term (*) würde die globale Minimierung zerfallen in

- (i) Minimierung über $\mathbf{x}_0 + \text{span}\{\mathbf{p}_1, \dots, \mathbf{p}_{m-1}\}$ und
- (ii) Minimierung längs der Geraden $\{\mathbf{x}_{m-1} + \alpha \mathbf{p}_m : \alpha \in \mathbb{R}\}$.

Wir fordern also

$$\begin{aligned} (A\mathbf{p}_m, P_{m-1}\mathbf{y}) &= 0 & \forall \mathbf{y} \in \mathbb{R}^{m-1}, \\ \Leftrightarrow (A\mathbf{p}_m, \mathbf{p}) &= 0 & \forall \mathbf{p} \in \text{span}\{\mathbf{p}_1, \dots, \mathbf{p}_{m-1}\} \end{aligned}$$

d.h. \mathbf{p}_m ist A -orthogonal („ A -konjugiert“) zu $\text{span}\{\mathbf{p}_1, \dots, \mathbf{p}_{m-1}\}$.

Gilt dann

- (i) $\phi(\mathbf{x}_{m-1}) = \min\{\phi(\mathbf{x}_0 + P_{m-1}\mathbf{y}) : \mathbf{y} \in \mathbb{R}^{m-1}\}$ und
- (ii) $\alpha_m = (\mathbf{r}_0, \mathbf{p}_m) / (A\mathbf{p}_m, \mathbf{p}_m)$,

so ist die Lösung von

$$\begin{aligned} &\min_{\substack{\mathbf{y} \in \mathbb{R}^{m-1} \\ \alpha \in \mathbb{R}}} \phi(\mathbf{x}_0 + P_{m-1}\mathbf{y} + \alpha\mathbf{p}_m) \\ &= \min_{\mathbf{y} \in \mathbb{R}^{m-1}} \phi(\mathbf{x}_0 + P_{m-1}\mathbf{y}) + \min_{\alpha \in \mathbb{R}} \left(\frac{\alpha^2}{2} (A\mathbf{p}_m, \mathbf{p}_m) - \alpha (\mathbf{r}_0, \mathbf{p}_m) \right) \end{aligned}$$

gegeben durch $P_{m-1} \mathbf{y} = \mathbf{x}_{m-1}, \alpha = \alpha_m$.

Beachte: Wegen $\mathbf{x}_{m-1} \in \mathbf{x}_0 + \text{span}\{\mathbf{p}_1, \dots, \mathbf{p}_{m-1}\}$ ist $\mathbf{r}_{m-1} = \mathbf{b} - A\mathbf{x}_{m-1} \in \mathbf{r}_0 + A \text{span}\{\mathbf{p}_1, \dots, \mathbf{p}_{m-1}\}$ und damit $(\mathbf{r}_0, \mathbf{p}_m) = (\mathbf{r}_{m-1}, \mathbf{p}_m)$. Demzufolge gilt auch $\alpha_m = (\mathbf{r}_{m-1}, \mathbf{p}_m) / (A\mathbf{p}_m, \mathbf{p}_m)$.

Bisherige Forderungen an \mathbf{p}_m :

- (i) \mathbf{p}_m A -konjugiert zu $\text{span}\{\mathbf{p}_1, \dots, \mathbf{p}_{m-1}\}$,
- (ii) $\mathbf{p}_m \not\perp \mathbf{r}_{m-1}$.

Noch offen: Existenz eines solchen \mathbf{p}_m in jedem Schritt.

Man kann zeigen: ein solches \mathbf{p}_m liegt in $\text{span}\{\mathbf{r}_{m-1}, \mathbf{p}_{m-1}\}$.

Das Verfahren der konjugierten Gradienten [Hestenes & Stiefel, 1952]:

$$m := 0$$

$$\mathbf{r}_0 := \mathbf{b} - A\mathbf{x}_0$$

while $\|\mathbf{r}_m\| > tol \cdot \|\mathbf{b}\|$

$$m := m + 1$$

if $m = 1$ **then**

$$\mathbf{p}_1 := \mathbf{r}_0$$

else

$$\beta_m := (\mathbf{r}_{m-1}, \mathbf{r}_{m-1}) / (\mathbf{r}_{m-2}, \mathbf{r}_{m-2})$$

$$\mathbf{p}_m := \mathbf{r}_{m-1} + \beta_m \mathbf{p}_{m-1}$$

$$\alpha_m := (\mathbf{r}_{m-1}, \mathbf{r}_{m-1}) / (A\mathbf{p}_m, \mathbf{p}_m)$$

$$\mathbf{x}_m := \mathbf{x}_{m-1} + \alpha_m \mathbf{p}_m$$

$$\mathbf{r}_m := \mathbf{r}_{m-1} - \alpha_m A\mathbf{p}_m$$

Aufwand pro Schritt: 1 Matrix-Vektor-Multiplikation, 2 Innenprodukte, 3 Vektor-Aufdatierungen (saxpys).

Satz 8.23. Für die Iterierten des cg-Verfahrens gilt die Fehlerabschätzung

$$\frac{\|\mathbf{x}^* - \mathbf{x}_m\|_A}{\|\mathbf{x}^* - \mathbf{x}_0\|_A} \leq 2 \left(\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^m.$$

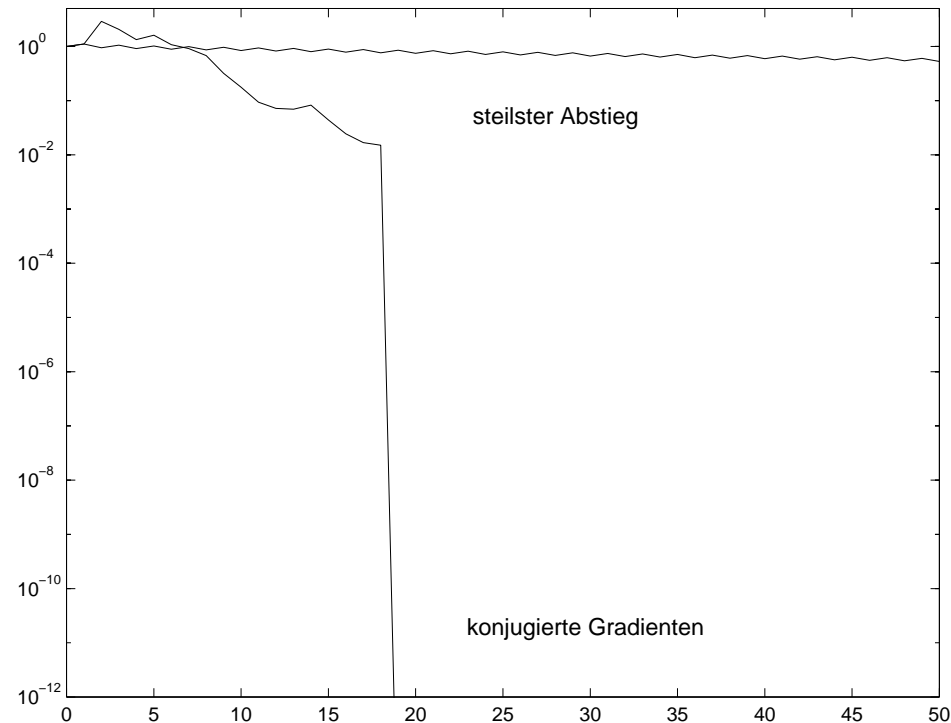
Dabei ist $\kappa = \text{cond}_2(A) = \lambda_{\max}(A)/\lambda_{\min}(A)$ die Konditionszahl von A bez. der Euklid-Norm.

In der Praxis wird das cg-Verfahren häufig auf das **vorkonditionierte** Gleichungssystem

$$L^{-1}AL^{-\top} \mathbf{y} = L^{-1} \mathbf{b} \text{ mit } \mathbf{y} = L^{\top} \mathbf{x}$$

angewandt. Dabei ist L eine invertierbare untere Dreiecksmatrix mit (im Idealfall) $\text{cond}_2(L^{-1}AL^{-\top}) = \text{cond}_2((LL^{\top})^{-1}A) \ll \text{cond}_2(A)$.

Abschließend illustrieren wir beide Gradientenverfahren am Testbeispiel $Ax = b$ mit $A = \text{tridiag}(-1, 2, -1) \in \mathbb{R}^{20 \times 20}$:



(Bei exakter Rechnung liefert das cg-Verfahren nach $n = \dim(A)$ Schritten immer die exakte Lösung.)