

2. Projekt

***Quasi-Newton-Verfahren & Gauß-Newton-Verfahren***

im Fach

Numerische Optimierung

Juni 2020

Maximilian Gaul

## Aufgabe 1

Siehe Programmcode in Project2.m und InverseBFGS.m.

## Aufgabe 2

Für  $A^0$  wählt man im BFGS-Verfahren üblicherweise die Einheitsmatrix, d.h. für den Induktionsanfang gilt:

$$A^0 = I$$

Wenn man das nun in die Formel für die zwei Rang-1-Modifikationen und in die BFGS-Update-Formel einsetzt erhält man den Ansatz (für  $B = A^{-1} = I$ ):

$$\left( I + \frac{ss^T - ys^T + ss^T - sy^T}{y^T s} - \frac{s^T y s s^T - y^T y s s^T}{(y^T)^2 s^2} \right) \cdot \left( I - \frac{ss^T}{s^T s} + \frac{yy^T}{y^T s} \right)$$

Der Übersicht halber teile ich die einzelnen Produkte auf und füge sie nach dem Kürzen am Ende wieder zusammen. Es ist zu beachten, dass  $y^T s = s^T y$ . Andere skalare Werte werden gekürzt.

$$I \cdot I = I$$

$$I \cdot \left( -\frac{ss^T}{s^T s} \right) = -\frac{ss^T}{s^T s}$$

$$I \cdot \left( \frac{yy^T}{y^T s} \right) = \frac{yy^T}{y^T s}$$

$$\left( \frac{ss^T - ys^T + ss^T - sy^T}{y^T s} \right) \cdot I = \frac{ss^T}{y^T s} - \frac{ys^T}{y^T s} + \frac{ss^T}{y^T s} - \frac{sy^T}{y^T s}$$

$$\left( \frac{ss^T - ys^T + ss^T - sy^T}{y^T s} \right) \cdot \left( -\frac{ss^T}{s^T s} \right) = -\frac{ss^T ss^T - ys^T ss^T + ss^T ss^T - sy^T ss^T}{y^T ss^T s}$$

$$= -\frac{ss^T ss^T}{y^T ss^T s} + \frac{ys^T ss^T}{y^T ss^T s} - \frac{ss^T ss^T}{y^T ss^T s} + \frac{sy^T ss^T}{y^T ss^T s} = -\frac{ss^T}{y^T s} + \frac{ys^T}{y^T s} - \frac{ss^T}{y^T s} + \frac{ss^T}{s^T s}$$

$$\left( \frac{ss^T - ys^T + ss^T - sy^T}{y^T s} \right) \cdot \left( \frac{yy^T}{y^T s} \right) = \frac{ss^T y - ys^T yy^T + ss^T yy^T - sy^T yy^T}{(y^T s)^2}$$

$$= \frac{ss^T yy^T}{(y^T s)^2} - \frac{ys^T yy^T}{(y^T s)^2} + \frac{ss^T yy^T}{(y^T s)^2} - \frac{sy^T yy^T}{(y^T s)^2} = \frac{sy^T}{y^T s} - \frac{yy^T}{y^T s} + \frac{sy^T}{y^T s} - \frac{sy^T yy^T}{(y^T s)^2}$$

$$\begin{aligned}
\left( -\frac{s^T y s s^T - y^T y s s^T}{(y^T s)^2} \right) \cdot I &= -\frac{\textcolor{red}{s}^T y s s^T}{(\textcolor{red}{y}^T \textcolor{red}{s})^2} + \frac{y^T y s s^T}{(y^T s)^2} = -\frac{s s^T}{y^T s} + \frac{y^T y s s^T}{(y^T s)^2} \\
\left( -\frac{s^T y s s^T - y^T y s s^T}{(y^T s)^2} \right) \cdot \left( -\frac{s s^T}{s^T s} \right) &= \frac{s^T y s s^T s s^T - y^T y s s^T s s^T}{(y^T s)^2 s^T s} \\
&= \frac{\textcolor{red}{s}^T y s \textcolor{blue}{s}^T s s^T}{(\textcolor{red}{y}^T \textcolor{red}{s})^2 \textcolor{blue}{s}^T \textcolor{blue}{s}} - \frac{y^T y s \textcolor{blue}{s}^T s s^T}{(y^T s)^2 \textcolor{blue}{s}^T \textcolor{blue}{s}} = \frac{s s^T}{y^T s} - \frac{y^T y s s^T}{(y^T s)^2} \\
\left( -\frac{s^T y s s^T - y^T y s s^T}{(y^T s)^2} \right) \cdot \left( \frac{y y^T}{y^T s} \right) &= -\frac{s^T y s s^T y y^T + y^T y s s^T y y^T}{(y^T s)^3} \\
&= -\frac{\textcolor{red}{s}^T y s \textcolor{red}{s}^T y y^T}{(\textcolor{red}{y}^T \textcolor{red}{s})^3} + \frac{y^T y s \textcolor{red}{s}^T y y^T}{(\textcolor{red}{y}^T \textcolor{red}{s})^3} = -\frac{s y^T}{y^T s} + \frac{y^T y s y^T}{(y^T s)^2}
\end{aligned}$$

Zusammenfassen:

$$\begin{aligned}
I - \frac{\textcolor{red}{s} \textcolor{red}{s}^T}{\textcolor{red}{s}^T \textcolor{red}{s}} + \frac{\textcolor{blue}{y} \textcolor{blue}{y}^T}{\textcolor{blue}{y}^T \textcolor{blue}{s}} + \frac{\textcolor{green}{s} \textcolor{green}{s}^T}{\textcolor{green}{y}^T \textcolor{green}{s}} - \frac{\textcolor{orange}{y} \textcolor{orange}{s}^T}{\textcolor{orange}{y}^T \textcolor{orange}{s}} + \frac{s s^T}{y^T s} - \frac{\textcolor{red}{s} \textcolor{red}{y}^T}{\textcolor{red}{y}^T \textcolor{red}{s}} - \frac{\textcolor{green}{s} \textcolor{green}{s}^T}{\textcolor{green}{y}^T \textcolor{green}{s}} + \frac{\textcolor{orange}{y} \textcolor{orange}{s}^T}{\textcolor{orange}{y}^T \textcolor{orange}{s}} - \frac{s s^T}{y^T s} + \frac{\textcolor{red}{s} \textcolor{red}{s}^T}{\textcolor{red}{s}^T \textcolor{red}{s}} + \frac{\textcolor{red}{s} \textcolor{red}{y}^T}{\textcolor{red}{y}^T \textcolor{red}{s}} - \frac{\textcolor{blue}{y} \textcolor{blue}{y}^T}{\textcolor{blue}{y}^T \textcolor{blue}{s}} + \frac{\textcolor{blue}{s} \textcolor{blue}{y}^T}{\textcolor{blue}{y}^T \textcolor{blue}{s}} \\
- \frac{\textcolor{red}{s} \textcolor{red}{y}^T y y^T}{(y^T s)^2} - \frac{s s^T}{y^T s} + \frac{y^T y s s^T}{(y^T s)^2} + \frac{s s^T}{y^T s} - \frac{y^T y s s^T}{(y^T s)^2} - \frac{\textcolor{blue}{s} \textcolor{blue}{y}^T}{\textcolor{blue}{y}^T \textcolor{blue}{s}} + \frac{y^T y s y^T}{(y^T s)^2} = I
\end{aligned}$$

Im Induktionsschritt geht man von  $k = 0$  auf allgemeines  $k$ , dann erhält man folgenden Ansatz (für  $B = A^{-1}$ ,  $k$  wurde zur besseren Übersichtlichkeit weggelassen):

$$\left( A^{-1} + \frac{(s - A^{-1}y)s^T + s(s - A^{-1}y)^T}{y^T s} - \frac{(s - A^{-1}y)^T y s s^T}{(y^T s)^2} \right) \cdot \left( A - \frac{As(As)^T}{s^T As} + \frac{y y^T}{y^T s} \right)$$

Die einzelnen Produkte werden wieder aufgeteilt:

$$(A^{-1}) \cdot (A) = I$$

$$(A^{-1}) \cdot \left( -\frac{As(As)^T}{s^T As} \right) = -\frac{s(As)^T}{s^T As} = -\frac{s s^T A}{s^T As}$$

$$(A^{-1}) \cdot \left( \frac{y y^T}{y^T s} \right) = \frac{A^{-1} y y^T}{y^T s}$$

$$\left( \frac{(s - A^{-1}y)s^T + s(s - A^{-1}y)^T}{y^T s} \right) \cdot (A) = \frac{s s^T A - A^{-1} y s^T A + s s^T A - s y^T A^{-1} A}{y^T s}$$

$$= \frac{s s^T A}{y^T s} - \frac{A^{-1} y s^T A}{y^T s} + \frac{s s^T A}{y^T s} - \frac{s y^T}{y^T s}$$

$$\begin{aligned}
& \left( \frac{(s - A^{-1}y)s^T + s(s - A^{-1}y)^T}{y^T s} \right) \cdot \left( -\frac{As(As)^T}{s^T As} \right) = -\frac{ss^T As s^T A - A^{-1}ys^T As s^T A + s(s^T - y^T A^{-1})As s^T A}{y^T s s^T As} \\
& = -\frac{ss^T As s^T A}{y^T s s^T As} + \frac{A^{-1}ys^T As s^T A}{y^T s s^T As} - \frac{ss^T As s^T A}{y^T s s^T As} + \frac{sy^T s s^T A}{y^T s s^T As} \\
& = -\frac{ss^T A}{y^T s} + \frac{A^{-1}ys^T A}{y^T s} - \frac{ss^T A}{y^T s} + \frac{ss^T A}{s^T As}
\end{aligned}$$

$$\begin{aligned}
& \left( \frac{(s - A^{-1}y)s^T + s(s - A^{-1}y)^T}{y^T s} \right) \cdot \left( \frac{yy^T}{y^T s} \right) = \frac{ss^T yy^T - A^{-1}ys^T yy^T + ss^T yy^T - sy^T A^{-1}yy^T}{(y^T s)^2} \\
& = \frac{ss^T yy^T}{(y^T s)^2} - \frac{A^{-1}ys^T yy^T}{(y^T s)^2} + \frac{ss^T yy^T}{(y^T s)^2} - \frac{sy^T A^{-1}yy^T}{(y^T s)^2} \\
& = \frac{sy^T}{y^T s} - \frac{A^{-1}yy^T}{y^T s} + \frac{sy^T}{y^T s} - \frac{sy^T A^{-1}yy^T}{(y^T s)^2} \\
& \left( -\frac{(s - A^{-1}y)^T y s s^T}{(y^T s)^2} \right) \cdot (A) = -\frac{s^T y s s^T A - y^T A^{-1}y s s^T A}{(y^T s)^2} \\
& = -\frac{s^T y s s^T A}{(y^T s)^2} + \frac{y^T A^{-1}y s s^T A}{(y^T s)^2} = -\frac{ss^T A}{y^T s} + \frac{y^T A^{-1}y s s^T A}{(y^T s)^2}
\end{aligned}$$

$$\begin{aligned}
& \left( -\frac{(s - A^{-1}y)^T y s s^T}{(y^T s)^2} \right) \cdot \left( -\frac{As(As)^T}{s^T As} \right) = \frac{s^T y s s^T As s^T A - y^T A^{-1}y s s^T As s^T A}{(y^T s)^2 s^T As} \\
& = \frac{s^T y s s^T As s^T A}{(y^T s)^2 s^T As} - \frac{y^T A^{-1}y s s^T As s^T A}{(y^T s)^2 s^T As} = \frac{ss^T A}{y^T s} - \frac{y^T A^{-1}y s s^T A}{(y^T s)^2}
\end{aligned}$$

$$\begin{aligned}
& \left( -\frac{(s - A^{-1}y)^T y s s^T}{(y^T s)^2} \right) \cdot \left( \frac{yy^T}{y^T s} \right) = -\frac{s^T y s s^T yy^T - y^T A^{-1}y s s^T yy^T}{(y^T s)^3} \\
& = -\frac{s^T y s s^T yy^T}{(y^T s)^3} + \frac{y^T A^{-1}y s s^T yy^T}{(y^T s)^3} = -\frac{sy^T}{y^T s} + \frac{y^T A^{-1}y sy^T}{(y^T s)^2}
\end{aligned}$$

Zusammenfassen

$$\begin{aligned}
I - \frac{ss^T A}{s^T As} + \frac{A^{-1}yy^T}{y^T s} + \frac{ss^T A}{y^T s} - \frac{A^{-1}ys^T A}{y^T s} + \frac{ss^T A}{y^T s} - \frac{sy^T}{y^T s} - \frac{ss^T A}{y^T s} + \frac{A^{-1}ys^T A}{y^T s} - \frac{ss^T A}{y^T s} \\
+ \frac{ss^T A}{s^T As} + \frac{sy^T}{y^T s} - \frac{A^{-1}yy^T}{y^T s} + \frac{sy^T}{y^T s} - \frac{sy^T A^{-1}yy^T}{(y^T s)^2} - \frac{ss^T A}{y^T s} + \frac{y^T A^{-1}y s s^T A}{(y^T s)^2}
\end{aligned}$$

$$+\frac{ss^T A}{y^T s} - \frac{y^T A^{-1} y s s^T A}{(y^T s)^2} - \frac{sy^T}{y^T s} + \frac{y^T A^{-1} y s y^T}{(y^T s)^2} = I$$

Da  $y^T A^{-1} y$  ein Skalar ist, kann es beliebig verschoben werden, d.h.  $sy^T A^{-1} y y^T = y^T A^{-1} y s y^T$ .

Für  $k + 1$  ändert sich an der oben gezeigten Rechnung nichts. Jeder Vektor bzw. jede Matrix erhält dann einfach nur den Index  $k + 1$ .

### Aufgabe 3

Wenn die Suchrichtung des BFGS Verfahrens:

$$d = -B \cdot \nabla f(x)$$

keine Abstiegsrichtung ist, d.h. die Bedingung:

$$\nabla f(x^k)^T \cdot d^k \leq -\rho \cdot \|d^k\|^p$$

nicht erfüllt ist, muss das Verfahren 'resettet' werden. Das Newton-Verfahren sucht nach einem stationären Punkt (d.h. Gradient = 0). Da der Gradient aber nicht nur in einem Minimum sondern auch in einem Maximum oder Sattelpunkt gleich Null sein kann, gibt das BFGS-Verfahren unter Umständen keine Abstiegsrichtungen aus. In diesem Fall bietet es sich an, die Suchrichtung auf den negativen Gradienten zu setzen:

$$d = -\nabla f(x)$$

Da nun die Abstiegsrichtung nicht mehr zur approximierten Inversen der Hesse-Matrix  $B$  passt, muss diese ebenfalls für den nächsten Schritt neu bestimmt werden (bzw. das nächste Update erfolgt dann mit dieser Matrix).

Hierzu bieten sich verschiedene Möglichkeiten an:

1. Wie beim Start des BFGS-Verfahrens  $B = I$  setzen
  - Hierbei geht jeglicher berechnete Fortschritt verloren, es handelt sich um einen recht naiven Ansatz
2. Die Hesse-Matrix einmalig aus Differenzenquotienten des Gradienten bestimmen und anschließend invertieren
  - Hoher Rechenaufwand von  $\mathcal{O}(n^2)$  für die Hesse-Matrix und nochmal  $\mathcal{O}(n^3)$  für das Invertieren
  - Problem wenn die Hesse-Matrix nicht invertierbar ist bzw. aufgrund von Auslöschung oder anderen numerischen Fehlern die Inverse schlecht konditioniert ist
  - Bisher berechneter Fortschritt geht ebenfalls verloren aber die Approximation der Hesse-Matrix ist sehr genau

3. Man könnte, wie in den Vorlesungsfolien beschrieben,  $\frac{y^T s}{y^T y} \cdot I_n$  als positiv-definitive Matrix verwenden

- Der Fortschritt geht zwar ähnlich wie bei  $B = I$  weitestgehend verloren, allerdings überträgt man noch Informationen von  $y$  und  $s$  auf den nächsten Schritt, erscheint also sinnvoller als die erste Variante

Da (2) gerade durch das BFGS-Verfahren vermieden werden soll da zu aufwändig und (1) keine Informationen über den vorherigen Verlauf des Algorithmus enthält, ist (3) die beste Wahl für die ich mich auch in meiner Implementierung entschieden habe.

#### Aufgabe 4

Auswertungen der Himmelblau-Funktion

$$f(x_1, x_2) = (x_1^2 + x_2 - 11)^2 + (x_1 + x_2^2 - 7)^2$$

Mit Gradient:

$$\nabla f(x_1, x_2) = \begin{bmatrix} 2 \cdot (x_1^2 + x_2 - 11) \cdot 2x_1 + 2 \cdot (x_1 + x_2^2 - 7) \\ 2 \cdot (x_1^2 + x_2 - 11) + 2 \cdot (x_1 + x_2^2 - 7) \cdot 2x_2 \end{bmatrix}$$

Schritt	$\mathbf{x}$	$f(\mathbf{x})$
1	$[0.00, -1.00]^T$	180
2	$[4.50, -1.00]^T$	70.31
3	$[2.54, -1.86]^T$	41.79
4	$[2.68, -2.26]^T$	37.84
5	$[4.16, -1.94]^T$	20.24
$\vdots$	$\vdots$	$\vdots$
13	$[3.58, -1.85]^T$	$6.37 \cdot 10^{-20}$

Abbildung 1: Verlauf von InverseBFGS für  $f$  bei einer Genauigkeit von  $10^{-8}$ . Man erkennt gut die superlineare Konvergenz.

Schritt	$\mathbf{x}$	$f(\mathbf{x})$
1	$[0.00, -1.00]^T$	180
2	keine Info	06.38
3	keine Info	05.91
4	keine Info	00.26
5	keine Info	0.00
$\vdots$	$\vdots$	$\vdots$
9	$[3.58, -1.85]^T$	$7.05 \cdot 10^{-13}$

Abbildung 2: Verlauf von `fminunc` für  $f$  bei einer Genauigkeit von  $10^{-8}$  mit Einstellungen `Algorithm = 'quasi-newton'`, `HessUpdate = 'bfgs'` und `Display = 'iter-detailed'`.

Die Rosenbrock-Funktion ist in `Projekt2.m` als `f_rosen_mult` definiert. Beispielhaft für  $N = 3$

$$g(x_1, x_2, x_3) = \sum_{i=1}^2 (1 - x_i)^2 + 100 \cdot (x_{i+1} - x_i^2)^2$$

$$=$$

$$(1 - x_1)^2 + 100 \cdot (x_2 - x_1^2)^2 + (1 - x_2)^2 + 100 \cdot (x_3 - x_2^2)^2$$

habe ich den Gradienten bestimmt:

$$\nabla g_3 = \begin{bmatrix} -2 \cdot (1 - x_1) + 200 \cdot (x_2 - x_1^2) \cdot (-2x_1) \\ 200 \cdot (x_2 - x_1^2) - 2 \cdot (1 - x_2) + 200 \cdot (x_3 - x_2^2) \cdot (-2x_2) \\ 200 \cdot (x_3 - x_2^2) \end{bmatrix}$$

Man erkennt eine Regel: Der Gradient besteht aus drei Teilen. Der erste Eintrag im Gradienten ist:

$$-2 \cdot (1 - x_1) + 200 \cdot (x_2 - x_1^2) \cdot (-2x_1)$$

Alle weiteren Einträge (bis auf den letzten an Position  $N - 1$ ) sind:

$$200 \cdot (x_i - x_{i-1}^2) - 2 \cdot (1 - x_i) + 200 \cdot (x_{i+1} - x_i^2) \cdot (-2x_i)$$

Der letzte Eintrag ist:

$$200 \cdot (x_N - x_{N-1}^2)$$

Die Ableitung ist in `Projekt2.m` in der Funktion `f_rosen_mult_deriv_func` definiert.

Schritt	$\mathbf{x}$	$f(x)$
1	$[-1.0, -1.0, \dots, -1.0]^T$	3636.00
2	$[0.57, 1.35, \dots, -0.22]^T$	682.94
3	$[0.94, 1.25, \dots, 0.20]^T$	508.69
4	$[1.33, 1.29, \dots, 0.52]^T$	419.10
5	$[1.45, 1.40, \dots, 1.43]^T$	379.44
$\vdots$	$\vdots$	$\vdots$
49	$[1.00, 1.00, \dots, 1.00]^T$	$2.2 \cdot 10^{-21}$

Abbildung 3: Verlauf von `InverseBFGS` für die 10-dim. Rosenbrock-Funktion bei einer Genauigkeit von  $10^{-8}$ .

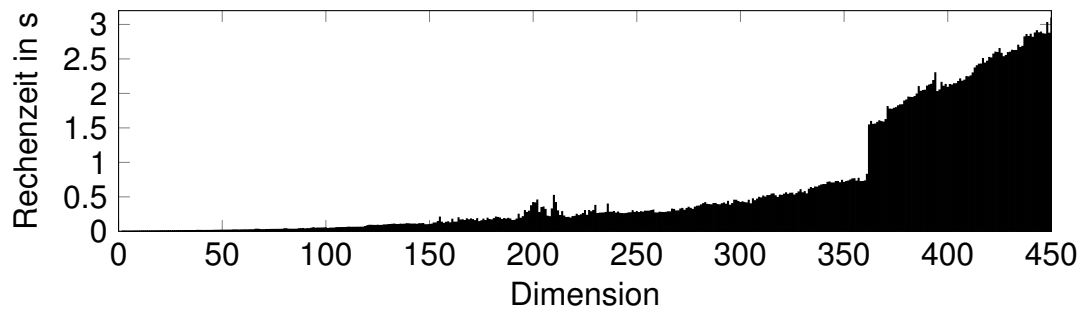


Abbildung 4: Rechenzeit der N-dimensionalen Rosenbrock-Funktion mit Startwert  $[-1, \dots, -1]^T$  für eine Genauigkeit von  $10^{-8}$  gemittelt über 100 Durchläufe von `InverseBFGS` (Intel Core i3-7100U, 8GB RAM, Windows 10 64-Bit)

Schritt	$\mathbf{x}$	$f(x)$
1	$[-1.0, -1.0, \dots, -1.0]^T$	3636.00
2	keine Info	55.59
3	keine Info	41.02
4	keine Info	9.56
5	keine Info	9.49
$\vdots$	$\vdots$	$\vdots$
79	$[1.00, 1.00, \dots, 1.00]^T$	$6.07 \cdot 10^{-11}$

Abbildung 5: Verlauf von `fminunc` für die 10-dim. Rosenbrock-Funktion bei einer Genauigkeit von  $10^{-8}$  mit Einstellungen `Algorithm = 'quasi-newton'`, `HessUpdate = 'bfgs'` und `Display = 'iter-detailed'`.



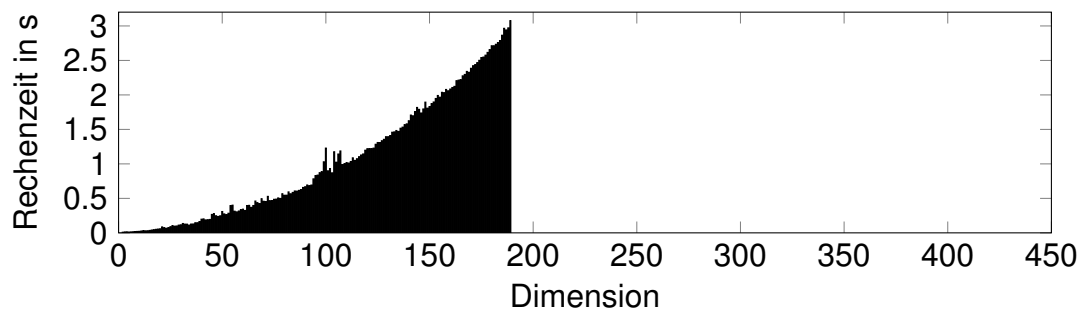


Abbildung 6: Rechenzeit der N-dimensionalen Rosenbrock-Funktion mit Startwert  $[-1, \dots, -1]^T$  für eine Genauigkeit von  $10^{-8}$  gemittelt über 100 Durchläufe von `fminunc` mit `Algorithm = 'quasi-newton'`, `HessUpdate = 'bfgs'` (Intel Core i3-7100U, 8GB RAM, Windows 10 64-Bit)

Man erkennt, dass mit höheren Dimensionen die Rechenzeit für beide Verfahren drastisch zunimmt und `fminunc` deutlich früher 3s pro Durchlauf benötigt.

Verfahren	Dimensionen	Iterationen	Genauigkeit	Anmerkung
InverseBFGS	750	941	$10^{-8}$	Zielfunktionswert: 3.98, die ersten 8 Werte bei 0.9967
InverseBFGS	850	1095	$10^{-8}$	Zielfunktionswert: $1.4 \cdot 10^{-20}$ , keine Werte $\neq 1$
InverseBFGS	925	1079	$10^{-8}$	Zielfunktionswert: $5.8 \cdot 10^{-21}$ , keine Werte $\neq 1$
fminunc	750	3936	$10^{-8}$	Zielfunktionswert: $1.8 \cdot 10^{-10}$ , keine Werte $\neq 1$
fminunc	850	4455	$10^{-8}$	Zielfunktionswert: $1.7 \cdot 10^{-10}$ , keine Werte $\neq 1$
fminunc	925	4852	$10^{-8}$	Zielfunktionswert: $1.9 \cdot 10^{-10}$ , keine Werte $\neq 1$

Abbildung 7: Unterschiede zwischen beiden Verfahren bei der Berechnung der N-dimensionalen Rosenbrock-Funktion bei hohen Dimensionen

Möglichkeiten um ein höheres  $N$  zu erreichen:

- Nach den Regeln der Analysis lässt sich eine Summe aufteilen in zwei Summen:

$$\begin{aligned}
 & \sum_{i=1}^{N-1} (1 - x_i)^2 + 100 \cdot (x_{i+1} - x_i^2)^2 \\
 &= \\
 & \sum_{i=1}^a (1 - x_i)^2 + 100 \cdot (x_{i+1} - x_i^2)^2 + \sum_{i=a+1}^{N-1} (1 - x_i)^2 + 100 \cdot (x_{i+1} - x_i^2)^2
 \end{aligned}$$

Die Berechnung des Minimums dieser zwei Summen kann auf z.B. zwei Threads aufgeteilt und am Ende wieder zusammengefügt werden. Bei ungeradem  $N$  muss man sich entscheiden wie die Summe aufgeteilt wird, ein Thread bearbeitet dann eine Dimension mehr als der andere. Bei besonders großem  $N$  können auch diese beiden Summen wiederum aufgeteilt und somit auf noch mehr Threads verteilt werden.

- Verwenden von sparse-Matrizen und Vektoren durch die sich die Rechenzeit unter Umständen reduzieren kann. Sparse-Datenstrukturen verwenden eine spezielle Repräsentation der Werte in denen Einträge mit 0 effizienter gespeichert werden. Aufgrundessen beschleunigt sich die Berechnung von Matrix-Vektor-Produkten (die beim BFGS-Verfahren sehr oft verwendet werden)
- Gegebenfalls genauere Untersuchungen über Kondition und Stabilität der Operationen bzw. wie diese verbessert werden können. Durch Auslöschung in der Update-Formel könnte es unter Umständen zu fehlerhaften Richtungsvektoren kommen die sich bei besonders großen Problemen potenzieren und somit die Konvergenz verlangsamen.

## Aufgabe 5

Die gegebene symmetrische Rang-1 Formel erhält man durch eine Rang-1 Modifikation

$$A^{k+1} = A^k + c \cdot u \cdot u^T$$

mit  $c = \frac{1}{u^T s^k}$  und  $u = y^k - A^k s^k$

$$A^{k+1} = A^k + \frac{1}{u^T s^k} \cdot u \cdot u^T = A^k + \frac{(y^k - A^k s^k)(y^k - A^k s^k)^T}{(y^k - A^k s^k)^T s^k}$$

Die Sherman-Morrison-Woodbury-Formel hat auf der linken Seite allerdings stehen  $(A + uv^T)^{-1}$ . Mit einem zusätzlichen  $v = c \cdot u$  bekommt den selben Ausdruck (Achtung:  $u^T s = s^T u$ )

$$A^k + uv^T = A^k + u \cdot (c \cdot u)^T = A^k + u \cdot \left( \frac{u}{u^T s^k} \right)^T = A^k + \frac{u \cdot u^T}{(s^k)^T u} = A^k + \frac{(y^k - A^k s^k)(y^k - A^k s^k)^T}{(y^k - A^k s^k)^T s^k}$$

Nach dem die Werte für  $u, v, c$  gefunden wurden, kann man diese nun in die Sherman-Morrison-Woodbury-Formel einsetzen und erhält somit den Ansatz:

$$A^{-1} - \frac{A^{-1}(y - As) \left( \frac{y - As}{(y - As)^T s} \right)^T A^{-1}}{1 + \left( \frac{y - As}{(y - As)^T s} \right)^T A^{-1}(y - As)}$$

Wenn man das T in den Bruch im Nenner und Zähler einfließen lässt, erhält man diesen Doppelbruch:

$$A^{-1} - \frac{\frac{A^{-1}(y-As)(y-As)^T A^{-1}}{((y-As)^T s)^T}}{1 + \frac{(y-As)^T A^{-1}(y-As)}{((y-As)^T s)^T}}$$

Nun den Zähler vom Zähler ausmultipliziert:

$$A^{-1} - \frac{\frac{(A^{-1}y - A^{-1}As)(y^T - (As)^T)A^{-1}}{((y-As)^T s)^T}}{1 + \frac{(y-As)^T A^{-1}(y-As)}{((y-As)^T s)^T}}$$

$$A^{-1} - \frac{\frac{(A^{-1}y - s)(y^T - s^T A)A^{-1}}{((y-As)^T s)^T}}{1 + \frac{(y-As)^T A^{-1}(y-As)}{((y-As)^T s)^T}}$$

$$A^{-1} - \frac{\frac{(A^{-1}y - s)(y^T A^{-1} - s^T A A^{-1})}{((y-As)^T s)^T}}{1 + \frac{(y-As)^T A^{-1}(y-As)}{((y-As)^T s)^T}}$$

$$A^{-1} - \frac{\frac{(A^{-1}y - s)(y^T A^{-1} - s^T)}{((y-As)^T s)^T}}{1 + \frac{(y-As)^T A^{-1}(y-As)}{((y-As)^T s)^T}}$$

Ebenfalls den Zähler im Nenner ausmultiplizieren:

$$A^{-1} - \frac{\frac{(A^{-1}y - s)(y^T A^{-1} - s^T)}{((y-As)^T s)^T}}{1 + \frac{(y^T - (As)^T)A^{-1}(y-As)}{((y-As)^T s)^T}}$$

$$A^{-1} - \frac{\frac{(A^{-1}y - s)(y^T A^{-1} - s^T)}{((y-As)^T s)^T}}{1 + \frac{(y^T - s^T A)A^{-1}(y-As)}{((y-As)^T s)^T}}$$

$$A^{-1} - \frac{\frac{(A^{-1}y - s)(y^T A^{-1} - s^T)}{((y-As)^T s)^T}}{1 + \frac{(y^T A^{-1} - s^T)(y-As)}{((y-As)^T s)^T}}$$

$$A^{-1} - \frac{\frac{(A^{-1}y - s)(y^T A^{-1} - s^T)}{((y-As)^T s)^T}}{1 + \frac{y^T A^{-1}y - y^T A^{-1}As - s^T y + s^T As}{((y-As)^T s)^T}}$$

$$A^{-1} - \frac{\frac{(A^{-1}y - s)(y^T A^{-1} - s^T)}{((y-As)^T s)^T}}{1 + \frac{y^T A^{-1}y - y^T s - s^T y + s^T As}{((y-As)^T s)^T}}$$

$$A^{-1} - \frac{\frac{(A^{-1}y - s)(y^T A^{-1} - s^T)}{((y-As)^T s)^T}}{1 + \frac{y^T(A^{-1}y - s) - s^T(y - As)}{((y-As)^T s)^T}}$$

Die 1 im Nenner erweitern:

$$A^{-1} = \frac{\frac{(A^{-1}y-s)(y^T A^{-1}-s^T)}{((y-As)^T s)^T}}{\frac{((y-As)^T s)^T}{((y-As)^T s)^T} + \frac{y^T(A^{-1}y-s)-s^T(y-As)}{((y-As)^T s)^T}}$$

mit dem zweiten Term verbinden:

$$A^{-1} = \frac{\frac{(A^{-1}y-s)(y^T A^{-1}-s^T)}{((y-As)^T s)^T}}{\frac{((y-As)^T s)^T + y^T(A^{-1}y-s)-s^T(y-As)}{((y-As)^T s)^T}}$$

ausmultiplizieren:

$$A^{-1} = \frac{\frac{(A^{-1}y-s)(y^T A^{-1}-s^T)}{((y-As)^T s)^T}}{\frac{s^T(y-As)+y^T A^{-1}y-y^T s-s^T y+s^T As}{((y-As)^T s)^T}}$$

$$A^{-1} = \frac{\frac{(A^{-1}y-s)(y^T A^{-1}-s^T)}{((y-As)^T s)^T}}{\frac{s^T y-s^T As+y^T A^{-1}y-y^T s-s^T y+s^T As}{((y-As)^T s)^T}}$$

und kürzen:

$$A^{-1} = \frac{\frac{(A^{-1}y-s)(y^T A^{-1}-s^T)}{((y-As)^T s)^T}}{\frac{y^T A^{-1}y-y^T s}{((y-As)^T s)^T}}$$

$$A^{-1} = \frac{\frac{(A^{-1}y-s)(y^T A^{-1}-s^T)}{((y-As)^T s)^T}}{\frac{y^T(A^{-1}y-s)}{((y-As)^T s)^T}}$$

Den Doppelbruch auflösen:

$$A^{-1} = \frac{(A^{-1}y-s)(y^T A^{-1}-s^T)}{((y-As)^T s)^T} \cdot \frac{((y-As)^T s)^T}{y^T(A^{-1}y-s)}$$

Wenn man sich die nun gegenüberstehenden Terme ansieht:

$$((y-As)^T s)^T$$

erkennt man, dass es sich hierbei um eine Zahl handelt (beispielhaft für ein  $2 \times 2$ -System):

$$\left( \left( \begin{bmatrix} 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 2 & 3 \\ 3 & -1 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ 2 \end{bmatrix} \right)^T \cdot \begin{bmatrix} 3 \\ 2 \end{bmatrix} \right)^T$$

$$\left( \left( \begin{bmatrix} 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 12 \\ 7 \end{bmatrix} \right)^T \cdot \begin{bmatrix} 3 \\ 2 \end{bmatrix} \right)^T$$

$$\begin{aligned} & \left( \left( \begin{bmatrix} -11 \\ -6 \end{bmatrix} \right)^T \cdot \begin{bmatrix} 3 \\ 2 \end{bmatrix} \right)^T \\ & \left( \begin{bmatrix} -11 & -6 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ 2 \end{bmatrix} \right)^T \\ & (-45)^T = -45 \end{aligned}$$

D.h. man darf den Doppelbruch kürzen, dann erhält man:

$$A^{-1} - \frac{(A^{-1}y - s)(y^T A^{-1} - s^T)}{y^T (A^{-1}y - s)}$$

Um den Ausdruck noch schöner zu machen, kann man jeweils ein Minus aus den zwei Termen im Zähler herausziehen:

$$A^{-1} - \frac{(-A^{-1}y + s)(-y^T A^{-1} + s^T)}{y^T (A^{-1}y - s)}$$

und Summanden vertauschen:

$$A^{-1} - \frac{(s - A^{-1}y)(s^T - y^T A^{-1})}{y^T (A^{-1}y - s)}$$

Aus dem rechten Term kann man aufgrund der Symmetrie das Transponieren herausziehen:

$$A^{-1} - \frac{(s - A^{-1}y)(s - A^{-1}y)^T}{y^T (A^{-1}y - s)}$$

Weiterhin kann man aus dem Nenner ebenfalls ein Minus entwerfen und die Summanden vertauschen:

$$A^{-1} + \frac{(s - A^{-1}y)(s - A^{-1}y)^T}{y^T (s - A^{-1}y)}$$

Da es sich beim Nenner ebenfalls um eine Zahl handelt (beispielhaft für ein  $2 \times 2$ -System):

$$\begin{bmatrix} 1 & 2 \end{bmatrix} \left( \begin{bmatrix} 3 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 & -2 \\ -2 & 3 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right)$$

kann man ihn auch so schreiben:

$$A^{-1} + \frac{(s - A^{-1}y)(s - A^{-1}y)^T}{(s - A^{-1}y)^T y}$$

Damit erhält man nun die Update-Formel für das Inverse-Verfahren:

$$B^{(k+1)} = B^k + \frac{(s^k - B^k y^k)(s^k - B^k y^k)^T}{(s^k - B^k y^k)^T y^k}$$

## Aufgabe 6

Siehe GaussNewton.m.

## Aufgabe 7

Die erste Funktion

$$f(t, x_1, x_2) = x_1 \cdot e^{x_2 \cdot t}$$

mit partiellen Ableitungen:

$$\frac{df}{dx_1} = e^{x_2 \cdot t}, \quad \frac{df}{dx_2} = t \cdot x_1 \cdot e^{x_2 \cdot t}$$

hat mit dem gegebenen Datensatz folgenden Residuen-Vektor:

$$r = \begin{bmatrix} f(0, x_1, x_2) - 2 \\ f(1, x_1, x_2) - 0.7 \\ f(2, x_1, x_2) - 0.3 \\ f(3, x_1, x_2) - 0.1 \end{bmatrix}$$

und diese Jacobi-Matrix ( $r'$ ):

$$J = \begin{bmatrix} \frac{df}{dx_1}(0, x_1, x_2) & \frac{df}{dx_2}(0, x_1, x_2) \\ \frac{df}{dx_1}(1, x_1, x_2) & \frac{df}{dx_2}(1, x_1, x_2) \\ \frac{df}{dx_1}(2, x_1, x_2) & \frac{df}{dx_2}(2, x_1, x_2) \\ \frac{df}{dx_1}(3, x_1, x_2) & \frac{df}{dx_2}(3, x_1, x_2) \end{bmatrix}$$

Schritt	$\mathbf{x}$	$\sum_{i=1}^M (r_i)^2$
1	$[0.50, 1.00]^T$	113.07
2	$[0.39, 0.74]^T$	16.64
3	$[0.69, 0.14]^T$	2.99
4	$[1.54, -0.66]^T$	0.24
5	$[1.98, -0.99]^T$	$2.5 \cdot 10^{-3}$
$\vdots$	$\vdots$	$\vdots$
10	$[2.00, -1.0]^T$	$2.0 \cdot 10^{-3}$

Abbildung 8: Verlauf der aufsummierten quadratischen Residuen mit dem gegebenen Datensatz und GaussNewton.m bei Abbruchbedingung  $\|\nabla f\| \leq 10^{-8}$

Für  $f$  erhält man somit die Werte  $x_1 = 1.9950$  und  $x_2 = -1.0095$  nach 10 Iterationen.

Weiterhin gilt für die zweite Funktion

$$g(t, x_1, x_2, x_3) = x_1 \cdot e^{-(x_2^2 + x_3^2) \cdot t} \cdot \frac{\sinh(x_3^2 \cdot t)}{x_3^2}$$

mit den partiellen Ableitungen

$$\frac{dg}{dx_1} = e^{-(x_2^2 + x_3^2) \cdot t} \cdot \frac{\sinh(x_3^2 \cdot t)}{x_3^2}$$

$$\frac{dg}{dx_2} = -2t \cdot x_2 x_1 \cdot e^{-(x_2^2 + x_3^2) \cdot t} \cdot \frac{\sinh(x_3^2 \cdot t)}{x_3^2}$$

$$\frac{dg}{dx_3} = \frac{2 \cdot x_1 \cdot e^{-(x_2^2 + x_3^2) \cdot t} \cdot (t \cdot x_3^2 \cdot \cosh(x_3^2 \cdot t) - (x_3^2 \cdot t + 1) \cdot \sinh(x_3^2 \cdot t))}{x_3^2}$$

und dem gegebenen Datensatz folgender Residuen-Vektor:

$$r = \begin{bmatrix} g(6, x_1, x_2, x_3) - 24.19 \\ g(12, x_1, x_2, x_3) - 35.34 \\ \vdots \\ g(180, x_1, x_2, x_3) - 55.21 \end{bmatrix}$$

und diese Jakobi-Matrix ( $r'$ ):

$$J = \begin{bmatrix} \frac{dg}{dx_1}(6, x_1, x_2, x_3) & \frac{dg}{dx_2}(6, x_1, x_2, x_3) & \frac{dg}{dx_3}(6, x_1, x_2, x_3) \\ \frac{dg}{dx_1}(12, x_1, x_2, x_3) & \frac{dg}{dx_2}(12, x_1, x_2, x_3) & \frac{dg}{dx_3}(12, x_1, x_2, x_3) \\ \vdots & \vdots & \vdots \\ \frac{dg}{dx_1}(180, x_1, x_2, x_3) & \frac{dg}{dx_2}(180, x_1, x_2, x_3) & \frac{dg}{dx_3}(180, x_1, x_2, x_3) \end{bmatrix}$$

Schritt	<b>x</b>	$\sum_{i=1}^M (r_i)^2$
1	$[10.0, 0.05, 0.10]^T$	$1.81 \cdot 10^6$
2	$[2.99, 0.05, 0.23]^T$	$2.45 \cdot 10^4$
3	$[3.02, 0.05, 0.15]^T$	$1.39 \cdot 10^3$
4	$[3.53, 0.06, 0.15]^T$	916.73
5	$[3.52, 0.06, 0.16]^T$	866.10
$\vdots$	$\vdots$	$\vdots$
25	$[3.54, 0.05, 0.15]^T$	776.75
$\vdots$	$\vdots$	$\vdots$
157	$[3.54, 0.05, 0.15]^T$	776.75

Abbildung 9: Verlauf der aufsummierten quadratischen Residuen mit dem gegebenen Datensatz und GaussNewton bei Abbruchbedingung  $\|\nabla f\| \leq 10^{-8}$

Für  $g$  erhält man mit dem gegebenen Datensatz und den Startwerten:

- $x_0^{(1)} = [10, 0.05, 0.1]^T$
- $x_0^{(2)} = [5, 0.145, 0.125]^T$
- $x_0^{(3)} = [3, 0.1, 0.05]^T$

somit den Wert

- $x_1 = 3.5355, x_2 = 0.0546, x_3 = 0.1539$

nach  $157^{(1)}$ ,  $154^{(2)}$  bzw.  $155^{(3)}$  Iterationen. Die Wolfe-Powell Schrittweitensteuerung musste aufgrund der Funktion  $g$  nach unten hin begrenzt werden da die Kombination aus sehr kleiner Schrittweite und sehr kleiner Richtung  $d$  sonst zu keiner Konvergenz geführt hätte.

Allgemein reagiert die Funktion  $g$  sehr empfindlich auf kleine Änderungen in den Startwerten. Das liegt vermutlich an den stark unterschiedlichen Werten der einzelnen Faktoren ( $x_1 \cdot e^{-(x_2^2 + x_3^2) \cdot t}$  ist sehr klein während  $\frac{\sinh(x_3^2 \cdot t)}{x_3^2}$  sehr groß wird) aufgrundessen es zu numerischer Instabilität kommt.

## Aufgabe 8

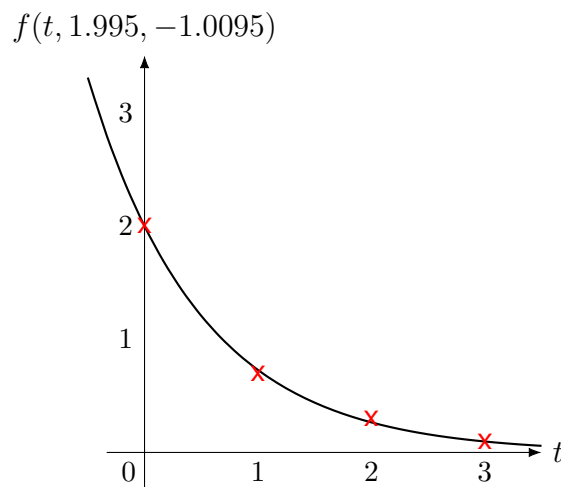


Abbildung 10: Modellfunktion und Datensatz für  $f(t, x_1, x_2) = x_1 \cdot e^{x_2 \cdot t}$



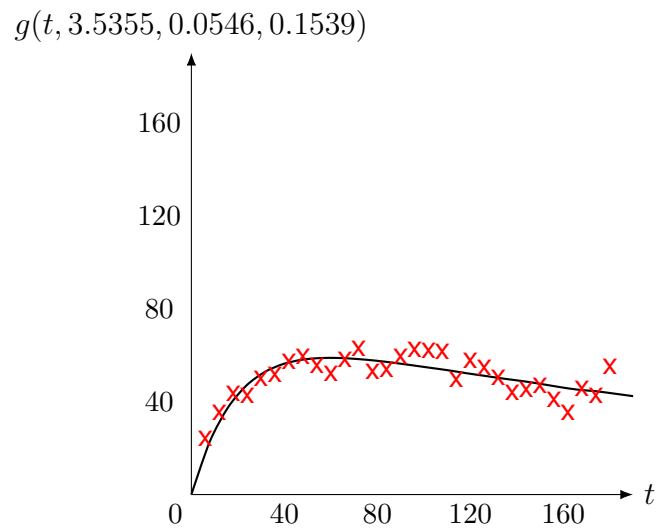


Abbildung 11: Modellfunktion und Datensatz für  $g(t, x_1, x_2, x_3) = x_1 \cdot e^{-(x_2^2 + x_3^2) \cdot t} \cdot \frac{\sinh(x_3^2 \cdot t)}{x_3^2}$

### Aufgabe 9

Bei einem Least-Squares-Problem wird der Abstand von einer Funktion  $f$  zu den gegebenen Datenpunkten  $(t, y)$  minimiert, die Zielfunktion (nach Folien-satz) ist daher:

$$f_{LeastSquares}(x) = \sum_{i=1}^m (f(t_i, x_1, x_2, \dots, x_n) - y_i)^2$$

Es werden also alle Residuen:

$$r(x) = \begin{bmatrix} f(t_1, x_1, \dots, x_n) - y_1 \\ f(t_2, x_1, \dots, x_n) - y_2 \\ \vdots \end{bmatrix}$$

quadriert und aufsummiert.  $f_{LeastSquares}(x)$  ist somit der erste Parameter, den InverseBFGS erhält.

Die BFGS-Funktion erhält ebenfalls den Gradienten als Parameter, der als:  $\nabla f(x) = 2 \cdot J(x)^T \cdot r(x)$  definiert ist (mit  $J$  aus Aufgabe 7).

Schritt	$\mathbf{x}$	$f(x)$
1	$[0.50, 1.00]^T$	113.07
2	$[-6.53, -9.17]^T$	73.42
3	$[4.82, -17.14]^T$	8.55
4	$[2.00, -15.16]^T$	0.59
5	$[2.00, -15.16]^T$	0.59
$\vdots$	$\vdots$	$\vdots$
19	$[1.9950, -1.0095]^T$	$2.0 \cdot 2.0^{-3}$

Abbildung 12: Verlauf von InverseBFGS für die Funktion  $f$  aus Aufgabe 7 bei einer Genauigkeit von  $10^{-8}$

Für die erste Funktion  $f(t, x_1, x_2) = x_1 \cdot e^{x_2 \cdot t}$  liefert das BFGS-Verfahren den selben Wert wie GaussNewton, allerdings benötigt es etwas mehr Iterationen.

Bei der zweiten Funktion bin ich mit der Definition  $\nabla f(x) = 2 \cdot J(x)^T \cdot r(x)$  nicht zum Ziel gekommen. Der Grund waren zu große Werte des Gradienten vor allem in der ersten Iteration von InverseBFGS die entweder zu NaN oder  $x$ -Werten geführt haben bei denen die Funktion über den Bereich  $[6 - 180]$  gleich Null war.

Ich habe den Gradienten daher durch numerische Differentiation ersetzt:

$$\nabla g = \begin{bmatrix} g_{LeastSquares}(x_1 + 10^{-10}, x_2, x_3) - g_{LeastSquares}(x_1, x_2, x_3) \\ g_{LeastSquares}(x_1, x_2 + 10^{-10}, x_3) - g_{LeastSquares}(x_1, x_2, x_3) \\ g_{LeastSquares}(x_1, x_2, x_3 + 10^{-10}) - g_{LeastSquares}(x_1, x_2, x_3) \end{bmatrix} \cdot \begin{bmatrix} 10^{10} \\ 10^{10} \\ 10^{10} \end{bmatrix}$$

Schritt	$\mathbf{x}$	$g_{LeastSquares}(x)$
1	$[5.00, 0.15, 0.13]^T$	$2.98 \cdot 10^4$
2	$[5.00, -0.18, -0.02]^T$	$2.83 \cdot 10^4$
3	$[5.00, -0.16, -0.08]^T$	$2.73 \cdot 10^4$
4	$[7.06, -0.16, -0.08]^T$	$2.45 \cdot 10^4$
5	$[3.21, -0.11, -0.06]^T$	$8.28 \cdot 10^3$
$\vdots$	$\vdots$	$\vdots$
15	$[2.93, -0.06, -0.13]^T$	923.52
$\vdots$	$\vdots$	$\vdots$
189	$[3.54, -0.05, -0.15]^T$	776.75

Abbildung 13: Verlauf von InverseBFGS für die Funktion  $g$  aus Aufgabe 7 bei einer Genauigkeit von  $10^{-8}$  und Startwert  $x_0^{(2)}$

Das Verfahren liefert mit  $x = [3.5356, -0.0546, -0.1539]^T$  also nicht den selben Wert wie GaussNewton, da  $x_2$  und  $x_3$  in  $g$  aber nur im Quadrat vorkommen ergibt sich die selbe Funktion (das Residuum am Ende ist daher gleich groß). InverseBFGS benötigt ca. 30 Schritte mehr.