1. Projekt

Ableitungsfreie Methoden

im Fach

Numerische Optimierung

Mai 2020

Maximilian Gaul

Aufgabe 1

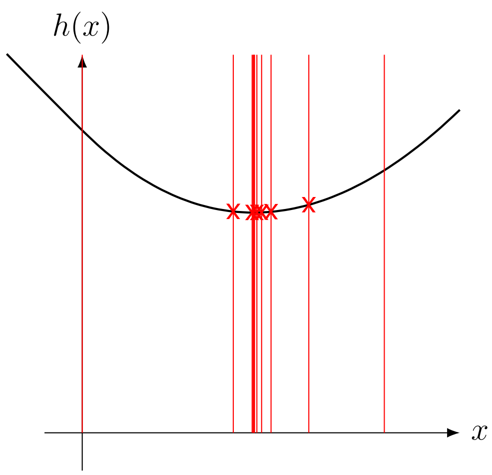
Teilintervalle und Funktionswerte des Bisektionsverfahrens aus `Bisektion.m` für das Minimum von

$$h(x) = e^{-x} + 0.5x^2$$

mit dem Startintervall [0, 1] nach 10 Schritten:

Schritt	**Intervall**	**Funktionswert**
1	[0.00, 1.00]	f(0.50)=0.7315
2	[0.50, 1.00]	f(0.75)=0.7536
3	[0.50, 0.75]	f(0.63)=0.7306
4	[0.50, 0.63]	f(0.56)=0.7280
5	[0.56, 0.63]	f(0.59)=0.7285
6	[0.56, 0.59]	f(0.58)=0.7281
7	[0.56, 0.58]	f(0.57)=0.7280
8	[0.56, 0.57]	f(0.57)=0.7280
9	[0.57, 0.57]	f(0.57)=0.7280
10	[0.57, 0.57]	f(0.57)=0.7280

Die Intervallgrenzen und Funktionsauswertungen sind wie folgt verteilt:



Aufgabe 2

Siehe `Mutation.m`.

Aufgabe 3

Eine Möglichkeit für ein Abbruchkriterium ist die Anzahl an Iterationen, d.h. der Algorithmus stoppt nach einer bestimmten Anzahl an Durchläufen und gibt den bis dahin berechneten Wert aus. Je nach Funktion und Zufallsvektoren kann die Genauigkeit des Ergebnisses stark schwanken. Während für eine festgelegte obere Grenze an Durchläufen für eine simple Funktion ein Minimum bereits ausreichend genau bestimmt werden kann, benötigt eine komplexere Funktion ggf. weitaus mehr Durchläufe. Die Anzahl an Iterationen ist also kein adaptives Kriterium.

Weiterhin könnte auch der Betrag aus der Differenz des aktuellen Funktionswertes und des letzten Funktionswertes als Abbruchkriterium funktionieren. Dabei aktualisiert man die zu überprüfenden Werte nur, sofern eine Verbesserung eingetreten ist (ansonsten wären letzter und aktueller Funktionswert eventuell gleich). Ist der Differenzbetrag kleiner als eine festgelegte Toleranz, endet der Algorithmus. Typische Werte sind hier z.B. 10^{-4} oder 10^{-8}.

Kriterium	**Funktion**	**Genauigkeit**	**Anzahl Schritte**
Differenzbetrag	f	10^{-6}	4.924.700
Differenzbetrag	f	10^{-6}	282.179
Differenzbetrag	f	10^{-6}	24.907
Differenzbetrag	g	10^{-6}	36.655
Differenzbetrag	g	10^{-6}	9.758
Differenzbetrag	g	10^{-6}	882

Durch die Wahl des Differenzbetrages kann man vom Zufall profitieren und mit verhältnismäßig wenigen Schritten zu einer guten Genauigkeit kommen. Wählt man dagegen die Anzahl der Iterationen, muss man ggf. neu rechnen wenn das Ergebnis noch zu ungenau ist:

Kriterium	**Funktion**	**Genauigkeit**	**Anzahl Schritte**
Anzahl Iterationen	f	$1 \cdot 10^{-5}$	225.000
Anzahl Iterationen	f	$7 \cdot 10^{-6}$	225.000
Anzahl Iterationen	f	$2 \cdot 10^{-8}$	225.000
Anzahl Iterationen	g	$2 \cdot 10^{-6}$	15.000
Anzahl Iterationen	g	$6 \cdot 10^{-7}$	15.000
Anzahl Iterationen	g	$1 \cdot 10^{-8}$	15.000

Der Parameter α bestimmt den Anteil des Zufallsvektors für den Wert von $\hat{x}$. Kleine Werte von α bedeuten eine höhere Wahrscheinlichkeit, dass sich der Funktionswert verbessert (die Verbesserung fällt aber möglicherweise kleiner aus) da sich besonders in der Nähe eines Minimums, bei quadratischen oder Funktionen 4. Grades der Funktionswert stark ändern und man leicht über das Ziel hinaus schießen kann wenn man zu große Schritte geht (ähnlich wie bei einem Abstiegsverfahren). Falls die Ableitung der Funktion zur Verfügung steht (tut sie hier wahrscheinlich nicht da das Projekt *Ableitungsfreie Methoden* heißt), könnte man ein Verfahren zur Schrittweitensteuerung implementieren (z.B. Armijo: $\varphi(\alpha) := f(x + \alpha r)$).

Aufgabe 4

Vergleich Rechenaufwand...

Aufgabe 5

Beispiel angeben bei dem Abbruchkriterium ungeeignet ist...

Aufgabe 6

Berechnet werden die ersten vier Iterationen des Nelder-Mead-Algorithmus von

$$g(x_1, x_2) = 100 \cdot (x_2 - 2)^4 + (x_1 - 2x_2)^2$$

mit den Parametern $n = 2$, $\alpha = \frac{1}{2}$, $\beta = 2$ und $\gamma = 1$.

$$x^{(0,0)} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}, e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Damit erhält man die Punkte $x^{(0,1)} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}, x^{(0,2)} = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$ und den Startsimplex

$$S_0 = (\begin{bmatrix} 4 \\ 2 \end{bmatrix}, \begin{bmatrix} 5 \\ 2 \end{bmatrix}, \begin{bmatrix} 4 \\ 3 \end{bmatrix})$$

k = 0

$$\max\{f(x^{(0,0)}) = 1600, f(x^{(0,1)}) = 8101, f(x^{(0,2)}) = 1604\} = f(x^{(0,1)})$$

$$s_0 = \frac{1}{2} \cdot (\begin{bmatrix} 4 \\ 2 \end{bmatrix} + \begin{bmatrix} 4 \\ 3 \end{bmatrix}) = \begin{bmatrix} 4 \\ \frac{5}{2} \end{bmatrix} \text{ und } x_0 = x^{(0,1)} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}$$

- Reflexion: $\hat{x}_0 = \begin{bmatrix} 4 \\ \frac{5}{2} \end{bmatrix} + 1 \cdot (\begin{bmatrix} 4 \\ \frac{5}{2} \end{bmatrix} - \begin{bmatrix} 5 \\ 2 \end{bmatrix}) = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$ mit $f(\hat{x}_0) = 25$
- Expansion: $\hat{x}_0^* = \begin{bmatrix} 4 \\ \frac{5}{2} \end{bmatrix} + 2 \cdot (\begin{bmatrix} 3 \\ 3 \end{bmatrix} - \begin{bmatrix} 4 \\ \frac{5}{2} \end{bmatrix}) = \begin{bmatrix} 2 \\ \frac{7}{2} \end{bmatrix}$ mit $f(\hat{x}_0^*) = 25$

Nach dem 1. Schritt erhält man den Simplex

$$S_1 = (\begin{bmatrix} 4 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ \frac{7}{2} \end{bmatrix}, \begin{bmatrix} 4 \\ 3 \end{bmatrix})$$

k = 1

$$\text{max}\{f(x^{(1,0)}) = 1600, f(x^{(1,1)}) = 25, f(x^{(1,2)}) = 1604\} = f(x^{(1,2)})$$

$$s_1 = \frac{1}{2} \cdot (\begin{bmatrix} 4 \\ 2 \end{bmatrix} + \begin{bmatrix} 2 \\ \frac{7}{2} \end{bmatrix}) = \begin{bmatrix} 3 \\ \frac{11}{4} \end{bmatrix} \text{ und } x_1 = x^{(1,2)} = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$$

- Reflexion: $\hat{x}_1 = \begin{bmatrix} 3 \\ \frac{11}{4} \end{bmatrix} + 1 \cdot (\begin{bmatrix} 3 \\ \frac{11}{4} \end{bmatrix} - \begin{bmatrix} 4 \\ 3 \end{bmatrix}) = \begin{bmatrix} 2 \\ \frac{5}{2} \end{bmatrix}$ mit $f(\hat{x}_1) = 9$
- Expansion: $\hat{x}_1^* = \begin{bmatrix} 3 \\ \frac{11}{4} \end{bmatrix} + 2 \cdot (\begin{bmatrix} 2 \\ \frac{5}{2} \end{bmatrix} - \begin{bmatrix} 3 \\ \frac{11}{4} \end{bmatrix}) = \begin{bmatrix} 1 \\ \frac{9}{4} \end{bmatrix}$ mit $f(\hat{x}_1^*) = 112.25$

Nach dem 2. Schritt erhält man den Simplex

$$S_2 = (\begin{bmatrix} 4 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ \frac{7}{2} \end{bmatrix}, \begin{bmatrix} 2 \\ \frac{5}{2} \end{bmatrix})$$

k = 2

$$\text{max}\{f(x^{(2,0)}) = 1600, f(x^{(2,1)}) = 25, f(x^{(2,2)}) = 9\} = f(x^{(2,0)})$$

$$s_2 = \frac{1}{2} \cdot (\begin{bmatrix} 2 \\ \frac{7}{2} \end{bmatrix} + \begin{bmatrix} 2 \\ \frac{5}{2} \end{bmatrix}) = \begin{bmatrix} 2 \\ 3 \end{bmatrix} \text{ und } x_2 = x^{(2,0)} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$$

- Reflexion: $\hat{x}_2 = \begin{bmatrix} 2 \\ 3 \end{bmatrix} + 1 \cdot (\begin{bmatrix} 2 \\ 3 \end{bmatrix} - \begin{bmatrix} 4 \\ 2 \end{bmatrix}) = \begin{bmatrix} 0 \\ 4 \end{bmatrix}$ mit $f(\hat{x}_2) = 1664$
- Innere Kontraktion: $\hat{x}_2^* = \begin{bmatrix} 2 \\ 3 \end{bmatrix} + \frac{1}{2} \cdot (\begin{bmatrix} 4 \\ 2 \end{bmatrix} - \begin{bmatrix} 2 \\ 3 \end{bmatrix}) = \begin{bmatrix} 3 \\ \frac{5}{2} \end{bmatrix}$ mit $f(\hat{x}_2^*) = 104$

Nach dem 3. Schritt erhält man den Simplex

$$S_3 = (\begin{bmatrix} 3 \\ \frac{5}{2} \end{bmatrix}, \begin{bmatrix} 2 \\ \frac{7}{2} \end{bmatrix}, \begin{bmatrix} 2 \\ \frac{5}{2} \end{bmatrix})$$

k = 3

$$\text{max}\{f(x^{(3,0)}) = 104, f(x^{(3,1)}) = 25, f(x^{(3,2)}) = 9\} = f(x^{(3,0)})$$

$$s_3 = \frac{1}{2}(\begin{bmatrix} 2 \\ \frac{7}{2} \end{bmatrix} + \begin{bmatrix} 2 \\ \frac{5}{2} \end{bmatrix}) = \begin{bmatrix} 2 \\ 3 \end{bmatrix} \text{ und } x_3 = x^{(3,0)} = \begin{bmatrix} 3 \\ \frac{5}{2} \end{bmatrix}$$

- Reflexion: $\hat{x}_3 = \begin{bmatrix} 2 \\ 3 \end{bmatrix} + 1 \cdot (\begin{bmatrix} 2 \\ 3 \end{bmatrix} - \begin{bmatrix} 3 \\ \frac{5}{2} \end{bmatrix}) = \begin{bmatrix} 1 \\ \frac{7}{2} \end{bmatrix}$ mit $f(\hat{x}_3) = 136$
- Innere Kontraktion: $\hat{x}_3^* = \begin{bmatrix} 2 \\ 3 \end{bmatrix} + \frac{1}{2} \cdot (\begin{bmatrix} 3 \\ \frac{5}{2} \end{bmatrix} - \begin{bmatrix} 2 \\ 3 \end{bmatrix}) = \begin{bmatrix} \frac{5}{2} \\ \frac{11}{4} \end{bmatrix}$ mit $f(\hat{x}_3^*) = 15.25$

Nach dem 4. Schritt erhält man den Simplex

$$S_4 = (\begin{bmatrix} \frac{5}{2} \\ \frac{11}{4} \end{bmatrix}, \begin{bmatrix} 2 \\ \frac{7}{2} \end{bmatrix}, \begin{bmatrix} 2 \\ \frac{5}{2} \end{bmatrix})$$

Aufgabe 7

Diskussion: Zuverlässigkeit und Rechenaufwand von *Mutation-Selektion* und *Nelder-Mead*