Karlsruher Institut für Technologie

Institut für Analysis
Prof. Dr. Michael Plum
M.Sc. Jonathan Wunderlich

Numerische Methoden - Sommersemester 2016

Lösungsvorschlag zu Übungsblatt 6

Aufgabe 12 (Newton-Verfahren)**:**
Gegeben sei die Funktion $f\colon (0,\infty) \to \mathbb{R}$ durch

$$f(x) = \log(x) + 0{,}5x - 2 \quad (x \in (0,\infty)),$$

wobei $\log(\cdot)$ den natürlichen Logarithmus bezeichnet.

(i) Zeigen Sie, dass f eine Nullstelle x^* in $(2,3)$ besitzt. Gibt es noch weitere Nullstellen?

(ii) Formulieren Sie das Newton-Verfahren zur Bestimmung einer Nullstelle von f.

(iii) Führen Sie die ersten drei Iterationen zum Startwert $x^{(0)} = 2{,}9$ durch und geben Sie eine approximative Nullstelle von f in $(2,3)$ an.

(iv) Bestimmen Sie anhand von Satz 5.1 im Skript eine Konstante η, sodass die Ungleichung

$$\left|x^* - x^{(k+1)}\right| \leq \eta \left|x^* - x^{(k)}\right|^2$$

für alle $k \in \mathbb{N}_0$ und $x^{(k)} \in (x^* - 1, x^* + 1)$ erfüllt ist und zeigen Sie:

$$x^{(k)} \in (x^* - 1, x^* + 1) \;\Rightarrow\; x^{(k+1)} \in (x^* - 1, x^* + 1).$$

(v) Für welche Startwerte $x^{(0)} \in (2,3)$ konvergiert das Newton-Verfahren?

Lösung:

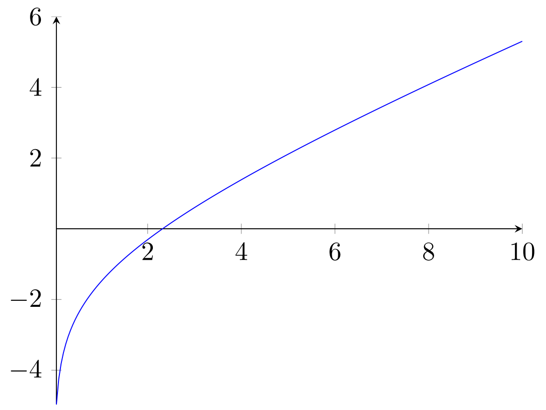

Abbildung 1: Graph von f

(i) Durch Einsetzen der Intervallgrenzen erhält man $f(2) \approx -0{,}306 < 0$ und $f(3) \approx 0{,}598 > 0$. Wegen der Stetigkeit von f liegt also nach dem Zwischenwertsatz eine Nullstelle in $(2,3)$.

Da f streng monoton wachsend ist, kann es nur eine Nullstelle geben.

(ii) Es sei $x^{(0)}$ der Startwert für das Newton-Verfahren. Dann kann die k-te Iterierte wie folgt berechnet werden:

$$x^{(k)} = x^{(k-1)} - \frac{f(x^{(k-1)})}{f'(x^{(k-1)})} = x^{(k-1)} - \frac{\log(x^{(k-1)}) + 0{,}5x^{(k-1)} - 2}{\frac{1}{x^{(k-1)}} + 0{,}5} \quad (k \geq 1).$$

(iii) Die ersten drei Iterierten des Verfahrens lauten:

$$x^{(1)} \approx 2{,}290751, \quad x^{(2)} \approx 2{,}318240, \quad x^{(3)} \approx 2{,}318317.$$

(iv) Wir wissen nach Teil (i), dass die gesuchte Nullstelle im Intervall $(2,3)$ liegt. Wir verwenden daher für die weiteren Berechnungen das Intervall $I := (1,4) \supset (x^* - 1, x^* + 1)$. In I gelten die folgenden Abschätzungen:

$$\begin{aligned} \left|f'(x)\right| &\geq \alpha := \frac{3}{4} = \left|f'(4)\right| > 0 \quad (x \in I), \\ \left|f''(x)\right| &\leq \beta := 1 = \left|f''(1)\right| \quad (x \in I). \end{aligned}$$

Nach Satz 5.1 aus dem Skript gilt die geforderte Abschätzung also mit $\eta := \frac{\beta}{2\alpha} = \frac{2}{3}$.

Um die angegebene Implikation zu zeigen verwenden wir, dass die Voraussetzung $x^{(k)} \in (x^* - 1, x^* + 1)$ besagt, dass $\left|x^* - x^{(k)}\right| < 1$ gilt. Also liefert obige Abschätzung

$$\left|x^* - x^{(k+1)}\right| \leq \frac{2}{3}\left|x^* - x^{(k)}\right|^2 < \frac{2}{3} < 1$$

und somit folgt die Behauptung.

(v) Nach Satz 5.1 konvergiert das Newton-Verfahren für alle $x^{(0)} \in I$ mit $\left|x^* - x^{(0)}\right| < \frac{1}{\eta} = \frac{3}{2}$, also für alle $x^{(0)} \in (2,3)$.

Aufgabe 13 (Newton-Verfahren, Matlab)**:**

(i) Formulieren Sie das Newton-Verfahren zur Bestimmung einer Nullstelle einer auf $D \subset \mathbb{R}^m$ differenzierbaren Funktion $F\colon D \subset \mathbb{R}^m \to \mathbb{R}^m$.

(ii) Gegeben sei die Funktion $F\colon \mathbb{R}^2 \to \mathbb{R}^2$ durch

$$F(x_1, x_2) := \begin{pmatrix} -{x_1}^2 + x_2 \\ {x_1}^2 + {x_2}^2 - 1 \end{pmatrix}, \quad \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2.$$

Bestimmen Sie die Ableitung von F. Warum ist der Vektor $(1, -\frac{1}{2})^T$ nicht als Startvektor geeignet?

(iii) Schreiben Sie ein Matlab-Programm das für eine Funktion $F\colon \mathbb{R}^m \to \mathbb{R}^m$ das Newton-Verfahren numerisch realisiert. Das Verfahren soll abbrechen, sobald $\left\|F(x^{(k)})\right\|_2$ unterhalb einer gewissen Toleranz ε liegt, d.h. wenn $\left\|F(x^{(k)})\right\|_2 < \varepsilon$ erfüllt ist.

(iv) Berechnen Sie mit Ihrem Programm aus (iii) approximative Nullstellen für die Funktion aus Aufgabe 12. Verwenden Sie die Toleranz $\varepsilon = 10^{-10}$.

(v) Bestimmen Sie mit Ihrem Programm aus (iii) eine approximative Nullstelle von F. Verwenden Sie dafür den Startvektor $(1,1)^T$ und die Toleranz $\varepsilon = 10^{-10}$.

Lösung:

(i) Das allgemeine Newton-Verfahren im $\mathbb{R}^m$ mit Startvektor $x^{(0)} \in D$ lautet:

$$x^{(k)} = x^{(k-1)} - \left(F'(x^{(k-1)})\right)^{-1} F(x^{(k-1)}) \quad (k \geq 1).$$

(ii) Die Ableitung von F ist gegeben durch die Jacobi-Matrix

$$J(x) = \begin{pmatrix} -2x_1 & 1 \\ 2x_1 & 2x_2 \end{pmatrix}.$$

An der Stelle $(1, -\frac{1}{2})^T$ ist die Jacobi-Matrix J nicht invertierbar und somit ist dieser Vektor als Startvektor ungeeignet.

(iii) Eine mögliche Implementierung findet man auf der Vorlesungshomepage.

(iv) Die approximative Nullstelle von f aus Aufgabe 12 lautet $\tilde{x} \approx 2{,}318317009243$.

(v) Die approximative Nullstelle von F lautet $\tilde{x} \approx (0{,}786151377762; 0{,}618033988750)$.

Aufgabe 14 (Quadratur)**:**
Gegeben seien die Integrale

$$\int_0^1 \sin(\pi x)\; dx \quad \text{und} \quad \int_{-1}^1 (4x^3 - 2x + \frac{1}{2}e^x)\; dx.$$

(i) Bestimmen Sie für die obigen Integrale den exakten Wert.

(ii) Geben Sie eine obere Schranke des Quadraturfehlers der Trapezregel für die obigen Integrale an.

(iii) Berechnen Sie für die Integrale einen Näherungswert unter Verwendung der Trapezregel und geben Sie den tatsächlichen absoluten Fehler an.

(iv) Zeigen Sie, dass für den Quadraturfehler der zusammengesetzten Trapezregel zur Intervalllänge $h = \frac{b-a}{N}$ die Fehlerdarstellung

$$R_Z(f) = -\frac{b-a}{12} \cdot h^2 \cdot f''(\eta) \quad \text{mit einem } \eta \in [a, b]$$

gilt, falls f zweimal stetig differenzierbar auf $[a, b]$ ist.

(v) Wiederholen Sie die Schritte (ii) und (iii) mit der zusammengesetzten Trapezregel zur Intervalllänge $h = \frac{1}{2}$.

Lösung:

(i) Es gilt:

$$\int_0^1 \sin(\pi x)\; dx = -\frac{1}{\pi}\cos(\pi x)\Big|_0^1 = -\frac{1}{\pi}(-1-1) = \frac{2}{\pi} \approx 0{,}64,$$
$$\int_{-1}^1 (4x^3 - 2x + \frac{1}{2}e^x)\; dx = x^4 - x^2 + \frac{1}{2}e^x\Big|_{-1}^1 = \frac{1}{2}(e - e^{-1}) = \sinh(1) \approx 1{,}18.$$

(ii) Die Trapezregel lautet

$$\int_a^b f(x)\; dx = \underbrace{\frac{b-a}{2}(f(a) + f(b))}_{:=Q_T(f)} + R_T(f).$$

Falls f zweimal stetig differenzierbar ist erhalten wir eine obere Schranke für den Quadraturfehler der Trapezregel aus der Fehlerdarstellung (siehe Skript S. 77):

$$|R_T(f)| \leq \frac{|b-a|^3}{12} \max_{\xi\in[a,b]} \left|f''(\xi)\right|.$$

Wir verwenden die Abkürzungen $g_1(x) := \sin(\pi x)$ und $g_2(x) := 4x^3 - 2x + \frac{1}{2}e^x$. Damit erhalten wir die Fehlerabschätzungen:

$$\begin{aligned}
|R_T(g_1)| &\leq \frac{1}{12} \max_{\xi\in[0,1]} \left|\pi^2 \sin(\pi\xi)\right| = \frac{\pi^2}{12} \approx 0{,}82,\\
|R_T(g_2)| &\leq \frac{2^3}{12} \max_{\xi\in[-1,1]} \left|24\xi + \frac{1}{2}e^\xi\right| = 16 + \frac{1}{3}e \approx 16{,}91.
\end{aligned}$$

(iii) Mit der Trapezregel berechnet man:

$$\begin{aligned}
Q_T(g_1) &= \frac{1}{2}(\sin(0) + \sin(\pi)) = 0,\\
Q_T(g_2) &= \frac{2}{2}\left((-4 + 2 + \frac{1}{2}e^{-1}) + (4 - 2 + \frac{1}{2}e^1)\right) = \frac{1}{2}(e + e^{-1}) = \cosh(1) \approx 1{,}54.
\end{aligned}$$

Damit ergeben sich die absoluten Fehler:

$$\begin{aligned}
\left|Q_T(g_1) - \frac{2}{\pi}\right| &= \frac{2}{\pi} \approx 0{,}64,\\
|Q_T(g_2) - \sinh(1)| &= \frac{1}{e} \approx 0{,}37.
\end{aligned}$$

(iv) Wir zerlegen das Intervall $[a,b]$ in äquidistante Teilintervalle $[x_{i-1}, x_i]$ $(1 \leq i \leq N)$ mit $x_i := a + ih$ $(0 \leq i \leq N)$. Damit können wir das gesuchte Integral zerlegen und auf jedem Teilintervall die Trapezregel anwenden. Außerdem verwenden wir auf jedem Teilintervall die Fehlerdarstellung der Trapezregel aus der Vorlesung (Skript S. 77):

$$\begin{aligned}
\int_a^b f(x)\ dx &= \sum_{i=1}^N \int_{x_{i-1}}^{x_i} f(x)\ dx\\
&= \sum_{i=1}^N \left[\frac{x_i - x_{i-1}}{2}(f(x_{i-1}) + f(x_i)) - \frac{(x_i - x_{i-1})^3}{12} f''(\eta_i)\right]
\end{aligned}$$

mit $\eta_i \in [x_{i-1}, x_i]$. Wir verwenden jetzt, dass die Teilintervalle äquidistant sind und formen weiter um

$$\begin{aligned}
\int_a^b f(x)\ dx &= \frac{h}{2}\left[f(a) + f(b) + 2\sum_{i=1}^{N-1} f(x_i)\right] - \frac{h^3}{12}\sum_{i=1}^N f''(\eta_i)\\
&= \frac{h}{2}\left[f(a) + f(b) + 2\sum_{i=1}^{N-1} f(x_i)\right] - \frac{h^2}{12} \cdot \frac{b-a}{N}\sum_{i=1}^N f''(\eta_i)\\
&= \frac{h}{2}\left[f(a) + f(b) + 2\sum_{i=1}^{N-1} f(x_i)\right] - \frac{b-a}{12} \cdot h^2 f''(\eta)
\end{aligned}$$

mit einem $\eta \in [a,b]$. Da f'' nach Voraussetzung stetig ist, konnte im letzten Schritt der Zwischenwertsatz angewendet werden.

(v) Eine obere Schranke für den Quadraturfehler der zusammengesetzten Trapezregel erhalten wir aus (iv):

$$|R_Z(f)| \leq \frac{|b-a|}{12} h^2 \max_{\xi \in [a,b]} \left| f''(\xi) \right|.$$

Damit erhalten wir zunächst Fehlerabschätzungen:

$$\begin{aligned} |R_Z(g_1)| &\leq \frac{1}{12} \cdot \frac{1}{4} \max_{\xi \in [0,1]} \left| \pi^2 \sin(\pi\xi) \right| = \frac{\pi^2}{48} \approx 0{,}21, \\ |R_Z(g_2)| &\leq \frac{2}{12} \cdot \frac{1}{4} \max_{\xi \in [-1,1]} \left| 24\xi + \frac{1}{2} e^{\xi} \right| = 1 + \frac{1}{48} e \approx 1{,}06. \end{aligned}$$

Die Approximationen an die Integrale lauten mit der zusammengesetzten Trapezregel:

$$\begin{aligned} Q_Z(g_1) &= \frac{1}{2} \left(\frac{1}{2} \sin(0) + \sin(\tfrac{\pi}{2}) + \frac{1}{2} \sin(\pi) \right) = \frac{1}{2}, \\ Q_Z(g_2) &= \frac{1}{2} \left(\frac{1}{2} g_2(-1) + g_2(-\tfrac{1}{2}) + g_2(0) + g_2(\tfrac{1}{2}) + g_2(1) \right) \\ &= \frac{1}{2} \left(\frac{1}{2} + \frac{1}{2} \cosh(1) + \cosh(\tfrac{1}{2}) \right) \approx 1{,}20. \end{aligned}$$

Die absoluten Quadraturfehler ergeben jetzt:

$$\begin{aligned} \left| Q_Z(g_1) - \frac{2}{\pi} \right| &\approx 0{,}14, \\ |Q_Z(g_2) - \sinh(1)| &\approx 0{,}02. \end{aligned}$$