

Numerische Mathematik 1

Wintersemester 2020, Übungsblatt 8

Ausarbeitung über Moodle bis 27. November 2020. Nach diesem Datum erscheinen die nachträglichen Kommentare und die [Lösungen](#) der Teilnehmer.

1. Seien $g \in C^1(\mathbb{R})$ und $f(x) = x - g(x)$ gegeben. Das Bisektionsverfahren auf Seite 164 im Skriptum mit Iterierten $\{a_k, b_k, c_k\}_{k \in \mathbb{N}_0}$ für f und die Fixpunktiteration auf Seite 168 im Skriptum mit Iterierten $\{x_k\}_{k \in \mathbb{N}_0}$ für g werden untersucht. Kreuzen Sie bei den wahren Behauptungen an.
 - (a) Wenn $\{x_k\}_{k \in \mathbb{N}_0}$ der Fixpunktiteration zu einem eindeutigen Fixpunkt $x^* = g(x^*)$ konvergieren, dann gibt es $i, j \in \mathbb{N}$ mit denen $\{c_k\}$ des Bisektionsverfahrens mit Startwerten $a_0 = x_i$ und $b_0 = x_j$ zu einer eindeutigen Nullstelle $f(x^*) = 0$ konvergieren.
 - (b) Wenn $\{c_k\}_{k \in \mathbb{N}_0}$ des Bisektionsverfahrens zu einer eindeutigen Nullstelle $f(x^*) = 0$ konvergieren, an der $f'(x^*) = 0$ gilt, dann konvergieren $\{x_k\}_{k \in \mathbb{N}_0}$ der Fixpunktiteration zu einem eindeutigen Fixpunkt $x^* = g(x^*)$, wenn $|x^* - x_0|$ ausreichend klein ist.
 - (c) Angenommen hat g genau einen Fixpunkt x^* , und $h(x) = g(g(x))$ hat noch zwei x^\dagger, x^\ddagger mit $x^\dagger < x^* < x^\ddagger$ und $|h'(x^\dagger)|, |h'(x^\ddagger)| < 1 < |g'(x^*)|$. Dann konvergieren die Iterierten $\{x_k\}_{k \in \mathbb{N}_0}$ zu x^* nicht, sondern nähern sich der Menge $\{x^\dagger, x^\ddagger\}$ mit $g(x^\dagger) = x^\ddagger$ und $g(x^\ddagger) = x^\dagger$. Weiters existieren Startwerte $\{a_0, b_0\}$ des Bisektionsverfahrens, mit denen die Iterierten $\{c_k\}_{k \in \mathbb{N}_0}$ zu x^* konvergieren.
 - (d) Angenommen hat g genau einen Fixpunkt $x^* = g(x^*)$, in dem $g'(x^*) = 1$ ein lokales Extremum von g' ist. Dann existieren Startwerte $\{a_0, b_0\}$ des Bisektionsverfahrens, mit denen die Iterierten $\{c_k\}_{k \in \mathbb{N}_0}$ zu x^* konvergieren.

Kommentare: Beispiel (2a) zeigt, dass Teil (a) falsch ist. Beispiel (2b) zeigt, dass Teil (b) falsch ist.

Teil (c) ist die allgemeine Situation vom Beispiel (2c), die folgendermaßen argumentiert wird. Aus $1 < |g'(x^*)|$ folgt mit dem 3. Satz auf Seite 170 im Skriptum, dass eine Fixpunkt Iteration mit g nie zum einzigen Fixpunkt x^* konvergieren kann. Da x^* für g eindeutig ist, sind x^\dagger und x^\ddagger keine Fixpunkte für g . Da x^\dagger und x^\ddagger die einzigen Fixpunkte für $h(x) = g(g(x))$ sind, folgen aus $g(x^\dagger) = g(h(x^\dagger)) = g(g(g(x^\dagger))) = h(g(x^\dagger))$ und $g(x^\ddagger) = g(h(x^\ddagger)) = g(g(g(x^\ddagger))) = h(g(x^\ddagger))$, dass $g(x^\dagger) = x^\ddagger$ und $g(x^\ddagger) = x^\dagger$ gelten. Laut dem 1. Satz auf Seite 170 im Skriptum folgt aus $|h'(x^\dagger)|, |h'(x^\ddagger)| < 1$, dass x^\dagger und x^\ddagger für eine Fixpunkt-Iteration mit h anziehend sind. Mit $|g'(x^*)| > 1$ folgt $f'(x^*) = 1 - g'(x^*) \neq 0$, und daher hat f unterschiedliche Vorzeichen in einer Umgebung von x^* .

Analog zum letzten Satz wird für Teil (d) folgendermaßen argumentiert. Wenn $g'(x^*) = 1$ ein lokales Extremum für $g'(x)$ ist, ist $f'(x) = 1 - g'(x)$ immer positiv oder immer negativ in einer Umgebung $B(x^*, \epsilon) \setminus \{x^*\}$, $\epsilon > 0$. In beiden Fällen hat $f(x)$ unterschiedliche Vorzeichen in $B(x^*, \epsilon) \setminus \{x^*\}$. Daher existieren $a_0, b_0 \in B(x^*, \epsilon) \setminus \{x^*\}$ mit denen $\{c_k\}_{k \in \mathbb{N}_0}$ des Bisektionsverfahrens zum x^* konvergieren. Teil (d) hängt mit Beispielen (2b) und (2d) zusammen.

2. Seien $g \in \mathcal{C}^1(\mathbb{R})$ und $f(x) = x - g(x)$ gegeben. Das Bisektionsverfahren auf Seite 164 im Skriptum mit Iterierten $\{a_k, b_k, c_k\}_{k \in \mathbb{N}_0}$ für f und die Fixpunktiteration auf Seite 168 im Skriptum mit Iterierten $\{x_k\}_{k \in \mathbb{N}_0}$ für g werden untersucht. Kreuzen Sie bei den wahren Behauptungen an.

- (a) Seien $f(x) = x^2 - 2$, $x_0 = 3$, $x_1 = g(x_0)$, $a_0 = x_0$ und $b_0 = x_1$. Der Fixpunkt für g und die Nullstelle für f ist $x^* = \sqrt{2}$. Die Iterierten $\{x_k\}$ und $\{c_k\}$ der Fixpunktiteration bzw. des Bisektionsverfahrens konvergieren zu x^* .
- (b) Seien $g(x) = (x^2 - x - 1)/(x^2 - 2x)$, $a_0 = 0$, $b_0 = 2$ und $x_0 \in (0, 2)$. Der Fixpunkt für g und die Nullstelle für f ist $x^* = 1$. Die Iterierten $\{x_k\}$ und $\{c_k\}$ der Fixpunktiteration bzw. des Bisektionsverfahrens konvergieren zu x^* .
- (c) Seien $g(x) = 16x(1 - x)/5$, $a_0 = (21 - \sqrt{21})/32$, $b_0 = (21 + \sqrt{21})/32$ und $x_0 = (a_0 + b_0)/2$. Der Fixpunkt für g und die Nullstelle für f ist $x^* = 11/16$. Die Iterierten $\{x_k\}$ der Fixpunktiteration konvergieren zu x^* nicht, aber die Iterierten $\{c_k\}$ des Bisektionsverfahrens konvergieren zu x^* .
- (d) Seien $g(x) = (1 - (x - 1)^2/2) \ln(e/x)$, $a_0 = 1/2$ und $b_0 = 3/2$. Der Fixpunkt für g und die Nullstelle für f ist $x^* = 1$. Die Iterierten $\{c_k\}$ des Bisektionsverfahrens konvergieren zu x^* . Für die Iterierten $\{x_k\}$ der Fixpunktiteration gibt es ein $\delta : (0, 1) \rightarrow \mathbb{R}_+$, wobei $|x_k - x^*| < \epsilon$ aus $|x_0 - x^*| < \delta(\epsilon)$ folgt.

Kommentare: Sehen Sie diesen [Code](#). Solang $f(g(x_0)) > 0$ im Teil (a) gilt, haben $f(x_0)$ und $f(x_1)$ das gleiche Vorzeichen. Im Teil (b) ist $g'(x^*) = 1$ ein lokales Maximum für $g'(x)$, und ein Argument mit dem Mittelwertsatz zeigt, $g(\bar{B}(x^*, \epsilon)) \subset \bar{B}(x^*, \epsilon)$ wird für jedes beliebig kleine $\epsilon > 0$ verletzt. Auf der anderen Seite im Teil (d) ist $g'(x^*) = 1$ ein lokales Minimum für $g'(x)$, und ein Argument mit dem Mittelwertsatz zeigt, es existiert ein $\epsilon > 0$ mit $g(\bar{B}(x^*, \epsilon)) \subset \bar{B}(x^*, \epsilon)$. Teil (c) folgt aus Beispiel (1c).

3. Fixpunktiterationen für die Funktion $f(x) = 2e^{-x} - 2 + 2x - x^2$ mit Nullstelle $x^* = 0$ werden untersucht. Kreuzen Sie bei den wahren Behauptungen an.

- (a) Gegeben sei $g(x) = x - f(x)/f'(x)$ für das Newton Verfahren auf Seite 173 im Skriptum. Laut dem Satz auf Seite 170 im Skriptum erfüllt die Fixpunktiteration $x_{k+1} = g(x_k)$, $k \in \mathbb{N}_0$, $x_0 \in [-1, +1]$, die Abschätzung $|x_k - x^*| \leq (e/2 - 1)\gamma^k/(1 - \gamma)$, wobei $\gamma = g'(-1)$.
- (b) Die Newton Iterierten $x_{k+1} = g(x_k)$, $k \in \mathbb{N}_0$, konvergieren zum Fixpunkt x^* für jedes $x_0 \in \mathbb{R}$, und es gilt $|x_{k+1} - x^*|/|x_k - x^*| \xrightarrow{k \rightarrow \infty} \frac{2}{3}$.
- (c) Gegeben sei $h(x) = x - 3f(x)/f'(x)$ für das modifizierte Newton Verfahren auf Seite 181 im Skriptum. Die Fixpunktiteration $x_{k+1} = h(x_k)$, $k \in \mathbb{N}_0$, konvergiert für jedes $x_0 \in [-1, +1]$, und es gilt $|x_{k+1} - x^*|/|x_k - x^*|^3 \xrightarrow{k \rightarrow \infty} \frac{1}{6}$.
- (d) Gegeben sei die Sonderfunktion $\phi(x) = x + \ln(1 - x + x^2/2)$ mit dem Fixpunkt x^* . Die Fixpunktiteration $x_{k+1} = \phi(x_k)$, $k \in \mathbb{N}_0$, konvergiert für jedes $x_0 \in (-1, +1)$, und es gilt $|x_{k+1} - x^*|/|x_k - x^*|^4 \xrightarrow{k \rightarrow \infty} 1$.

Kommentare: Sehen Sie diesen [Code](#). Für Teil (a) gilt $\{f(x), f^{(1)}(x), f^{(2)}(x), f^{(3)}(x)\} \rightarrow \{0, 0, 0, -2\}$ für $x \rightarrow x^*$, und daher hat die Nullstelle x^* Multiplizität 3. Daher gilt $g'(x^*) = 0$ für das Newton Verfahren nicht. Laut dem 2. Satz auf Seite 170 im Skriptum folgt $|x_k - x^*| \leq \gamma^k |x_1 - x_0|/(1 - \gamma)$, wobei $\gamma = \max_{x \in [-1, +1]} |g'(x)| = g'(1)$. Weiters gilt $|x_1 - x_0| \leq (e/2 - 1)$, $\forall x_0, x_1 \in [-1, +1]$.

Für Teil (b) gilt $\{g(x), g^{(1)}(x), g^{(2)}(x)/2!, g^{(3)}(x)/3!\} \rightarrow \{0, 2/3, -1/36, 1/540\}$ für $x \rightarrow x^*$. Laut dem Beweis auf Seite 179 im Skriptum gilt $|x_{k+1} - x^*|/|x_k - x^*| \rightarrow g'(x^*) = 2/3$.

Für Teil (c) gilt $\{h(x), h^{(1)}(x), h^{(2)}(x)/2!, h^{(3)}(x)/3!\} \rightarrow \{0, 0, -1/12, 1/180\}$ für $x \rightarrow x^*$. Laut dem Beweis auf Seite 180 im Skriptum gilt $|x_{k+1} - x^*|/|x_k - x^*|^2 \rightarrow h''(x^*) = 1/12$.

Für Teil (d) gilt $\{\phi(x), \phi^{(1)}(x), \phi^{(2)}(x)/2!, \phi^{(3)}(x)/3!\} \rightarrow \{0, 0, 0, 1/6\}$ für $x \rightarrow x^*$. Laut dem Argument im Beweis auf Seite 180 im Skriptum gilt $|x_{k+1} - x^*|/|x_k - x^*|^3 \rightarrow \phi^{(3)}(x^*) = 1/6$.

Wenn diese Konvergenzraten mit Rechnungen überprüft werden, muss darauf geachtet werden, dass die verwendeten Differenzen $|x_k - x^*|$ ausreichend größer als $\text{eps} \approx 2.2\text{e-}16$ sind, da sonst Quotienten $|x_{k+1} - x^*|/|x_k - x^*|^\alpha$ nicht zuverlässig sind. Sehen Sie die zutreffenden Details für Beispiel (4d).

4. Sei x^* eine Nullstelle einer Funktion f mit $f \in \mathcal{C}^2(B(x^*, \epsilon))$ für ein $\epsilon > 0$ und $f'(x^*) \neq 0$. Das Sekant-Verfahren

$$x_{k+1} = x_k - f(x_k)(x_k - x_{k-1})/(f(x_k) - f(x_{k-1})), \quad k \in \mathbb{N}$$

wird untersucht. Angenommen konvergieren die Iterierten $\{x_k\}_{k \in \mathbb{N}_0}$ zur Nullstelle x^* . Kreuzen Sie bei den wahren Behauptungen an.

- (a) Die Differenzen $\epsilon_k = x_k - x^*$ erfüllen

$$\epsilon_{k+1}/(\epsilon_k \epsilon_{k-1}) \xrightarrow{k \rightarrow \infty} \frac{1}{2} f''(x^*)/f'(x^*).$$

- (b) Mit $p = \sqrt{5}/2$ gilt

$$|\epsilon_{k+1}|/|\epsilon_k|^p \xrightarrow{k \rightarrow \infty} |\frac{1}{2} f''(x^*)/f'(x^*)|^{1/p}.$$

- (c) Wenn $\mathbf{F} \in \mathcal{C}^1(\mathbb{R}^2, \mathbb{R}^2)$ durch die stetige Ergänzung von

$$\mathbf{F}(x, y) = \begin{bmatrix} y \\ y - f(y)(y - x)/(f(y) - f(x)) \end{bmatrix}$$

gegeben ist, wird das Sekant Verfahren mit $\mathbf{x}_k = (x_{k-1}, x_k)$ durch $\mathbf{x}_{k+1} = \mathbf{F}(\mathbf{x}_k)$, $k \in \mathbb{N}$, gegeben, und mit $\mathbf{x}^* = (x^*, x^*)$ gilt $\rho(\mathbf{F}'(\mathbf{x}^*)) = 0$.

- (d) Für $f(x) = ex + \ln(|x|)$ mit der Nullstelle $x^* = 1/e$ ergibt sich $|x_6 - x^*|/|x_5 - x^*| < 10^{-5}$ mit $x_0 = \frac{1}{2}x^*$ und $x_1 = \frac{3}{2}x^*$.

Kommentare: Sehen Sie diesen [Code](#). Für Teile (a) und (b) seien

$$\begin{aligned} f(x_k) &= f(x_{k-1}) + f'(\eta_k)(x_k - x_{k-1}), & \eta_k \text{ zwischen } x_k \text{ und } x_{k-1} \\ f(x_k) &= f(x^*)_0 + f'(x^*)(x_k - x^*) + \frac{1}{2} f''(\zeta_k)(x_k - x^*)^2, & \zeta_k \text{ zwischen } x_k \text{ und } x^*. \end{aligned}$$

Zuerst muss festgestellt werden, dass

$$\frac{\epsilon_{k+1}}{\epsilon_k} = 1 - \frac{f(x_k) - f(x^*)}{x_k - x^*} \frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})} \xrightarrow{k \rightarrow \infty} 1 - \frac{f'(x^*)}{f'(x^*)} = 0.$$

Dann mit

$$\begin{aligned} \epsilon_{k+1} &= \epsilon_k - f(x_k) \frac{\epsilon_k - \epsilon_{k-1}}{f(x_k) - f(x_{k-1})} = \frac{\epsilon_k f(x_{k-1}) - \epsilon_{k-1} f(x_k)}{f(x_k) - f(x_{k-1})} \\ &= \left[\frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})} \right] \left[\frac{f(x_{k-1})/\epsilon_{k-1} - f(x_k)/\epsilon_k}{x_k - x_{k-1}} \right] \epsilon_k \epsilon_{k-1} \end{aligned}$$

und mit η_k und ζ_k ergibt sich

$$\begin{aligned}\epsilon_{k+1} &= \frac{1}{f'(\eta_k)} \left[\frac{\frac{1}{2}f''(\zeta_{k-1})\epsilon_{k-1} - \frac{1}{2}f''(\zeta_k)\epsilon_k}{\epsilon_k - \epsilon_{k-1}} \right] \epsilon_k \epsilon_{k-1} \\ &= \frac{1}{f'(\eta_k)} \left[\frac{f''(\zeta_{k-1}) - f''(\zeta_k)\epsilon_k/\epsilon_{k-1}}{2(\epsilon_k/\epsilon_{k-1} - 1)} \right] \epsilon_k \epsilon_{k-1}.\end{aligned}$$

Denn $\epsilon_k/\epsilon_{k-1} \xrightarrow{k \rightarrow \infty} 0$ gilt, folgt Teil (a),

$$\frac{\epsilon_{k+1}}{\epsilon_k \epsilon_{k-1}} = \frac{1}{f'(\eta_k)} \left[\frac{f''(\zeta_{k-1}) - f''(\zeta_k)\epsilon_k/\epsilon_{k-1}}{2(\epsilon_k/\epsilon_{k-1} - 1)} \right] \xrightarrow{k \rightarrow \infty} \frac{f''(x^*)}{2f'(x^*)}.$$

Wie schon festgestellt, gilt $|\epsilon_{k+1}|/|\epsilon_k|^p \xrightarrow{k \rightarrow \infty} C$ mindestens für $p = 1$ und $C = 0$. Wenn es für dieseen Grenzwert Konstanten $p > 1$ und $C > 0$ geben sollte, dann aus Teil (a) gelten notwendigerweise

$$\frac{C|\epsilon_k|^{p-1}}{|\epsilon_{k-1}|} = \left(\frac{|\epsilon_{k+1}|}{|\epsilon_k||\epsilon_{k-1}|} \right) \left(\frac{C|\epsilon_k|^p}{|\epsilon_{k+1}|} \right) \xrightarrow{k \rightarrow \infty} \frac{1}{2}|f''(x^*)/f'(x^*)|$$

und mit $|\epsilon_k|^{p(p-1)-1} \xrightarrow{k \rightarrow \infty} E \in \{0, 1, \infty\}$

$$\frac{C|\epsilon_k|^{p-1}}{|\epsilon_{k-1}|} = C \left(\frac{|\epsilon_k|}{|\epsilon_{k-1}|^{\frac{1}{p-1}}} \right)^{p-1} = C \left(\frac{|\epsilon_k|}{|\epsilon_{k-1}|^p} \right)^{p-1} |\epsilon_{k-1}|^{p(p-1)-1} \xrightarrow{k \rightarrow \infty} C^p E.$$

Im nicht trivialen Fall gelten $E = 1$, $p(p-1) = 1$ und daher $p = (1 + \sqrt{5})/2$. Da die Grenzwerte übereinstimmen müssen, folgt $C = [\frac{1}{2}|f''(x^*)/f'(x^*)|]^{1/p}$.

Für Teil (d) gelten mit Matlab

$$|x_5 - x^*| = 7.4e-8, \quad |x_6 - x^*| = 2.2e-12, \quad |x_7 - x^*| = 5.6e-17$$

zu 2 signifikanten Ziffern. Also mit nur doppelter Genauigkeit kann $|x_7 - x^*|$ mit keinem Rechnungssystem (auch nicht mit Octave) zuverlässig bearbeitet werden. Mit Mathematica können die Iterierten aber *exakt* berechnet werden, und es gelten

$$\frac{|x_5 - x^*|}{|x_4 - x^*|} = 1.7e-3, \quad \frac{|x_6 - x^*|}{|x_5 - x^*|} = 2.9e-5, \quad \frac{|x_7 - x^*|}{|x_6 - x^*|} = 5.1e-8$$

zu 2 signifikanten Ziffern. Also gilt $|\epsilon_k|/|\epsilon_{k-1}| \xrightarrow{k \rightarrow \infty} 0$ wie erwartet. Um die Konvergenzrate mit Matlab abzuschätzen, können nur die Iterierten $\{x_k\}_{k=0}^6$ verwendet werden. Mit dem Goldenen Schnitt `gs = (1+sqrt(5))/2` werden mit Matlab berechnet,

```

...
k = 3, |x(k+1)-x*|/|x(k)-x*|^gs = 6.7e-01
k = 4, |x(k+1)-x*|/|x(k)-x*|^gs = 8.7e-01
k = 5, |x(k+1)-x*|/|x(k)-x*|^gs = 7.4e-01

```

wobei aus Teilen (a) und (b) der Grenzwert $7.9e-01$ erwartet wird.

Für Teil (c) ist die Ableitung von \mathbf{F} gegeben durch die stetige Ergänzung von

$$\mathbf{F}'(\mathbf{x}) = \begin{bmatrix} 0 & 1 \\ f(y) \frac{f(y) - f(x) + (x - y)f'(x)}{[f(y) - f(x)]^2} & f(x) \frac{f(x) - f(y) + (y - x)f'(y)}{[f(y) - f(x)]^2} \end{bmatrix}$$

also mit ξ, η, ζ zwischen x und y gilt

$$\mathbf{F}'(\mathbf{x}) = \begin{bmatrix} 0 & 1 \\ f(y) \frac{f''(\xi)}{2f'(\eta)^2} & f(x) \frac{f''(\zeta)}{2f'(\eta)^2} \end{bmatrix}$$

und an der Stelle $\mathbf{x}^* = (x^*, x^*)$ folgt

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}^*} \mathbf{F}'(\mathbf{x}) = \begin{bmatrix} 0 & 1 \\ f(x^*) \frac{f''(x^*)}{2f'(x^*)} & f(x^*) \frac{f''(x^*)}{2f'(x^*)} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

Es wird für $\mathbf{X}(k) = (x(k), x(k-1))$ und $\mathbf{X}^* = (x^*, x^*)$ mit Matlab berechnet,

```
...
k = 3, ||X(k+1)-X*||/||X(k)-X*||^2 = 4.0e+00, ||X(k+1)-X*||/||X(k)-X*||^gs = 9.8e-01
k = 4, ||X(k+1)-X*||/||X(k)-X*||^2 = 6.5e+00, ||X(k+1)-X*||/||X(k)-X*||^gs = 6.7e-01
k = 5, ||X(k+1)-X*||/||X(k)-X*||^2 = 4.1e+01, ||X(k+1)-X*||/||X(k)-X*||^gs = 8.7e-01
```

Obwohl $\rho(\mathbf{F}'(\mathbf{x}^*)) = 0$ gilt, ist die Konvergenz nicht quadratisch. Für quadratische Konvergenz gibt es die Bedingung auf Seite 639 im Lehrbuch, dass $\mathbf{F}'(\mathbf{x}^*) = 0$ gelten soll, die in diesem Fall nicht erfüllt ist.

5. Sei $f(x) \in \mathcal{C}^3(\mathbb{R})$ eine Funktion mit genau einer Nullstelle $f(x^*) = 0$, an der $f^{(k)}(x^*) \neq 0$ gilt für $k = 1, 2, 3$. Sei die Funktion $g(x)$ implizit definiert durch

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

Die Fixpunktiteration $x_{k+1} = g(x_k)$ wird untersucht. Kreuzen Sie bei den wahren Behauptungen an.

- (a) Es gelten $g'(x^*) = g''(x^*) = 0$.
- (b) Die Fixpunktiteration erfüllt $|x^* - x_k|/|x^* - x_{k-1}|^3 \xrightarrow{k \rightarrow \infty} |f^{(3)}(x^*)/f^{(1)}(x^*)|$.
- (c) Für die Funktion $f(x) = x^3 - 1$ konvergiert die Fixpunktiteration mit $x_0 = \frac{1}{2}$ zur Nullstelle $x^* = 1$ mit doppelter Genauigkeit in genau 4 Iterationen, und zu 1 signifikanten Ziffern gelten $|x_k - x^*|/|x_{k-1} - x^*|^3 = 0.3$ für $k = 2, 3$.
- (d) Für die Funktion $f(x) = x - 1/x$ konvergiert die Fixpunktiteration mit $x_0 = \frac{1}{2}$ zur Nullstelle $x^* = 1$ mit doppelter Genauigkeit in genau 5 Iterationen, und zu 1 signifikanten Ziffern gelten $|x_k - x^*|/|x_{k-1} - x^*|^3 = 0.5$ für $k = 3, 4$.

Kommentare: Sehen Sie diesen [Code](#). Die Funktion $g(x)$ ist explizit gegeben durch

$$a(x) = f''(x)/2, \quad b(x) = f'(x) - xf''(x), \quad c(x) = f(x) - xf'(x) + x^2f''(x)/2$$

$$\begin{aligned} g(x) &= \frac{-b(x) + \text{sign}(f'(x))\sqrt{b^2(x) - 4a(x)c(x)}}{2a(x)} \\ &= x - \frac{f'(x)}{f''(x)} \left(1 - \sqrt{1 - 2\frac{f(x)f''(x)}{f'(x)^2}} \right) \end{aligned}$$

wobei $\text{sign}(f'(x))$ in der quadratischen Formel verwendet wird, damit $g(x^*) = x^*$ gilt. Die Ableitungen von g können aber bequem durch implizites Ableiten der für g definierenden Gleichung bestimmt werden:

$$\begin{aligned} 0 &= f'(x) + f''(x)(g(x) - x) + f'(x)(g'(x) - 1) + \dots \\ 0 &= f''(x) + f'''(x)(g(x) - x) + 2f''(x)(g'(x) - 1) + \dots \\ &\quad \text{usw.} \end{aligned}$$

Durch Auswertung dieser Ergebnisse an der Stelle $x = x^*$ ergibt sich

$$g'(x^*) = g''(x^*) = 0 \quad \text{und} \quad g^{(3)}(x^*) = f^{(3)}(x^*)/f^{(1)}(1)(x^*).$$

Durch eine Taylor Entwicklung, wie auf Seite 180 im Skriptum, ergibt sich für die Fixpunkt Iteration

$$|x^* - x_k|/|x^* - x_k|^3 \xrightarrow{k \rightarrow \infty} \frac{1}{3!}|g^{(3)}(x^*)| = \frac{1}{3!}|f^{(3)}(x^*)/f^{(1)}(1)(x^*)|.$$

Für Teil (c) wird daher erwartet,

$$|x^* - x_k|/|x^* - x_k|^3 \xrightarrow{k \rightarrow \infty} \frac{1}{3!}|f^{(3)}(1)/f^{(1)}(1)| = \frac{1}{3}$$

und mit Matlab werden berechnet,

```
k = 1, |x(k)-x*|/|x(k-1)-x*|^3 = 4.29e-01
k = 2, |x(k)-x*|/|x(k-1)-x*|^3 = 3.34e-01
k = 3, |x(k)-x*|/|x(k-1)-x*|^3 = 3.33e-01 <-
k = 4, |x(k)-x*|/|x(k-1)-x*|^3 = 0.00e+00
```

Für Teil (d) wird erwartet,

$$|x^* - x_k|/|x^* - x_k|^3 \xrightarrow{k \rightarrow \infty} \frac{1}{3!}|f^{(3)}(1)/f^{(1)}(1)| = \frac{1}{2}$$

und mit Matlab werden berechnet,

```
k = 1, |x(k)-x*|/|x(k-1)-x*|^3 = 1.50e+00
k = 2, |x(k)-x*|/|x(k-1)-x*|^3 = 1.03e+00
k = 3, |x(k)-x*|/|x(k-1)-x*|^3 = 4.90e-01 <-
k = 4, |x(k)-x*|/|x(k-1)-x*|^3 = 5.94e+04
k = 5, |x(k)-x*|/|x(k-1)-x*|^3 = 0.00e+00
```

Wie bei vorherigen Beispielen gesehen, ist die Berechnung solcher Quotienten nicht zuverlässig, sobald die Differenzen sich im Bereich der doppelten Genauigkeit befinden.

6. Basierend auf Beispiel 1 auf dem 4. Übungsblatt seien $n = 100$, $x_i = 2(i - \frac{1}{2})/n - 1$, $\mathbf{x} = \{x_i\}_{i=1}^n$ und $\mathbf{v} = \text{sign}(\mathbf{x})$ fixiert, und mit $\mathbf{u}(\alpha) = \tanh(\alpha \mathbf{x})$ soll die Funktion

$$F(\alpha) = \|\mathbf{u}(\alpha) - \mathbf{v}\|_2^2 + n\|\mathbf{Q}\mathbf{u}(\alpha)\|_2^2, \quad \mathbf{Q} = \begin{bmatrix} -1 & +1 & & 0 \\ & \ddots & \ddots & \\ 0 & & -1 & +1 \end{bmatrix} \in \mathbb{R}^{(n-1) \times n},$$

über $\alpha \in \mathbb{R}$ global minimiert werden. Sei $\alpha^* = \text{argmin}_{\alpha \in \mathbb{R}} F(\alpha)$. Für die folgenden Verfahren wird in den entsprechenden Pseudo-Codes im Skriptum das Abbruchskriterium mit $\text{tol} = 1.0e-6$ verwendet. Kreuzen Sie bei den wahren Behauptungen an.

(a) Wenn das Bisektionsverfahren

$$F'(a_k)F'(b_k) < 0 \quad \begin{cases} F'(c_k)F'(b_k) < 0 : & a_{k+1} = c_k, \quad b_{k+1} = b_k \\ c_k = (a_k + b_k)/2 & F'(a_k)F'(c_k) < 0 : \quad b_{k+1} = c_k, \quad a_{k+1} = a_k \end{cases} \quad k \in \mathbb{N}_0$$

mit dem Intervall $[a_0, b_0] = [\frac{1}{2}\alpha^*, \frac{3}{2}\alpha^*]$ gestartet wird, gilt $|c_{20} - \alpha^*| < 10^{-6}$ nach 23 Auswertungen von $F'(\alpha)$.

(b) Wenn das Newton-Verfahren

$$\alpha_k = \alpha_{k-1} - F'(\alpha_{k-1})/F''(\alpha_{k-1}), \quad k \in \mathbb{N}$$

mit $\alpha_0 = \frac{1}{2}\alpha^*$ gestartet wird, gilt $|\alpha_6 - \alpha^*| < 10^{-15}$ nach 6 Auswertungen jeweils von $F'(\alpha)$ und $F''(\alpha)$.

(c) Wenn das Sekant-Verfahren

$$\alpha_{k+1} = \alpha_k - F'(\alpha_k) \left/ \frac{F'(\alpha_k) - F'(\alpha_{k-1})}{\alpha_k - \alpha_{k-1}} \right., \quad k \in \mathbb{N}$$

mit $\alpha_0 = \frac{1}{2}\alpha^*$ und $\alpha_1 = \frac{3}{2}\alpha^*$ gestartet wird, gilt $|\alpha_{11} - \alpha^*| < 10^{-15}$ nach 11 Auswertungen von $F'(\alpha)$.

(d) Wenn das Abstieg-Verfahren

$$\alpha_k = \alpha_{k-1} - \omega F'(\alpha_{k-1}), \quad \omega = 1, \quad k \in \mathbb{N}$$

mit $\alpha_0 = \frac{1}{2}\alpha^*$ gestartet wird, gilt $|\alpha_{18} - \alpha^*| < 10^{-6}$ nach 18 Auswertungen von $F'(\alpha)$.

Kommentare: Sehen Sie diesen [Code](#). Die ersten zwei Ableitungen der Zielfunktion

$$F(\alpha) = \sum_{i=1}^n [\tanh(\alpha x_i) - \text{sign}(x_i)]^2 + n \sum_{i=1}^{n-1} [\tanh(\alpha x_{i+1}) - \tanh(\alpha x_i)]^2$$

sind

$$\begin{aligned} F'(\alpha) &= \sum_{i=1}^n 2x_i \operatorname{sech}^2(\alpha x_i) [\tanh(\alpha x_i) - \text{sign}(x_i)] \\ &+ \sum_{i=1}^{n-1} 2n(x_{i+1} \operatorname{sech}^2(\alpha x_{i+1}) - x_i \operatorname{sech}^2(\alpha x_i)) [\tanh(\alpha x_{i+1}) - \tanh(\alpha x_i)] \end{aligned}$$

und

$$\begin{aligned} F''(\alpha) &= \sum_{i=1}^n 2x_i^2 \operatorname{sech}^4(\alpha x_i) \\ &\quad - 4x_i^2 \operatorname{sech}^2(\alpha x_i) \tanh(\alpha x_i) [\tanh(\alpha x_i) - \text{sign}(x_i)] \\ &+ \sum_{i=1}^{n-1} 2n[x_{i+1} \operatorname{sech}^2(\alpha x_{i+1}) - x_i \operatorname{sech}^2(\alpha x_i)]^2 \\ &\quad - 2(\tanh(\alpha x_{i+1}) - \tanh(\alpha x_i)) \times \\ &\quad [2x_{i+1}^2 \operatorname{sech}^2(\alpha x_{i+1}) \tanh(\alpha x_{i+1}) - 2x_i^2 \operatorname{sech}^2(\alpha x_i) \tanh(\alpha x_i)] \end{aligned}$$

Diese können mit dem folgenden Matlab Code berechnet werden:

```

n = 100;
x = 2*((1:n)-0.5)/n-1; xp = x(2:n); xm = x(1:(n-1));
v = sign(x);
F0 = @(a) sum((tanh(a*x) - v).^2) + ...
        sum((tanh(a*xp) - tanh(a*xm)).^2)*n;
F1 = @(a) sum(2*x.*sech(a*x).^2.* (tanh(a*x) - v)) + ...
        sum((2*xp.*sech(a*xp).^2 - 2*xm.*sech(a*xm).^2).* ...
            (tanh(a*xp) - tanh(a*xm)))*n;
F2 = @(a) sum(2*x.^2.*sech(a*x).^4 - ...
        4*x.^2.*sech(a*x).^2.*tanh(a*x).*(tanh(a*x) - v)) ...
+sum(2*(xp.*sech(a*xp).^2 - xm.*sech(a*xm).^2).^2 - ...
        2*(tanh(a*xp) - tanh(a*xm)).* ...
        (2*xp.^2.*sech(a*xp).^2.*tanh(a*xp) - ...
        2*xm.^2.*sech(a*xm).^2.*tanh(a*xm)))*n;

```

Zu 2 signifikanten Ziffern gelten für Teil (a) $|c_{20} - \alpha^*| = 1.8\text{e-}06$, für Teil (b) $|\alpha_6 - \alpha^*| = 3.7\text{e-}12$, für Teil (c) $|\alpha_{11} - \alpha^*| = 4.4\text{e-}16$ und für Teil (d) $|\alpha_{18} - \alpha^*| = 4.5\text{e-}07$.

7. Für $N = 100$, $h = 1/N$, $x_i = (i - \frac{1}{2})h$, $\mathbf{x} = \{x_i\}_{i=1}^N$, $\mathbf{v} = \text{sign}(\mathbf{x} - \frac{1}{2})$, $\mu = h$ und $p = 1.1$ soll die Funktion

$$J_h(\mathbf{u}) = \frac{h}{2} \sum_{i=1}^N (u_i - v_i)^2 + \mu \frac{h}{p} \sum_{i=1}^{N-1} \left| \frac{u_{i+1} - u_i}{h} \right|^p$$

über $\mathbf{u} \in \mathbb{R}^N$ global minimiert werden. Kreuzen Sie bei den wahren Behauptungen an.

- (a) Der Gradient der Zielfunktion ist gegeben durch

$$\nabla J_h(\mathbf{u})/h = \mathbf{u} - \mathbf{v} + \frac{\mu}{h^2} Q^\top \Phi(\mathbf{u}) Q \mathbf{u}$$

wobei $Q \in \mathbb{R}^{(N-1) \times N}$ und $\Phi(\mathbf{u}) \in \mathbb{R}^{(N-1) \times (N-1)}$ gegeben sind durch

$$Q = \begin{bmatrix} -1 & +1 & & 0 \\ & \ddots & \ddots & \\ 0 & & -1 & +1 \end{bmatrix}, \quad \Phi(\mathbf{u}) = \text{diag} \left\{ \phi \left(\frac{u_{i+1} - u_i}{h} \right) \right\}_{i=1}^{N-1}.$$

und $\phi(t) = |t|^{p-2}$ für $t \neq 0$ und $\phi(0) = 0$.

- (b) Mit $B = \bar{B}_2(0, \|\mathbf{v}\|_2)$ gilt $\mathbf{G}_I(B) \subseteq B$ für die Funktion $\mathbf{G}_I(\mathbf{u})$ der impliziten Fixpunkt-Iteration $\mathbf{u}_{k+1} = \mathbf{G}_I(\mathbf{u}_k)$ definiert durch

$$\left[I + \frac{\mu}{h^2} Q^\top \Phi(\mathbf{u}_k) Q \right] \mathbf{u}_{k+1} = \mathbf{v}.$$

- (c) Wenn die implizite Fixpunkt Iteration der letzten Gleichung mit $\mathbf{u}_0 = \mathbf{v}$ durchgeführt wird, gilt $\|\mathbf{u}_k - \mathbf{u}_{k-1}\|_2 < 10^{-5} \|\mathbf{v}\|_2$ mit $k = 50$ Iterationen.
(d) Das Ergebnis der letzten Iteration erfüllt $\|\mathbf{u}_k - \mathbf{v}\|_2 < 10^{-2} \|\mathbf{v}\|_2$ mit $k = 50$ Iterationen.

Kommentare: Sehen Sie diesen [Code](#). Der Gradient der Zielfunktion ist gegeben durch

$$\begin{aligned}\frac{\partial J_h}{\partial u_i} &= h(u_i - v_i) + \mu h^{1-p} [|u_i - u_{i-1}|^{p-1} \text{sign}(u_i - u_{i-1})_{\delta_{i>1}} - |u_{i+1} - u_i|^{p-1} \text{sign}(u_{i+1} - u_i)_{\delta_{i<N}}] \\ &= h(u_i - v_i) + \mu \left[\phi\left(\frac{u_i - u_{i-1}}{h}\right) \frac{u_i - u_{i-1}}{h}_{\delta_{i>1}} - \phi\left(\frac{u_{i+1} - u_i}{h}\right) \frac{u_{i+1} - u_i}{h}_{\delta_{i<N}} \right]\end{aligned}$$

Das implizite Fixpunkt Iteration erfüllt

$$\|\mathbf{u}_{k+1}\|_2^2 \leq \mathbf{u}_{k+1}^\top \left[I + \frac{\mu}{h^2} Q^\top \Phi(\mathbf{u}_k) Q \right] \mathbf{u}_{k+1} = \mathbf{u}_{k+1}^\top \mathbf{v} \leq \|\mathbf{u}_{k+1}\|_2 \|\mathbf{v}\|_2$$

oder $\|\mathbf{u}_{k+1}\|_2 \leq \|\mathbf{v}\|_2$. Zu 2 signifikanten Ziffern gelten $\|\mathbf{u}_{50} - \mathbf{u}_{49}\|_2 = 2.9\text{e-}03$ für Teil (c) und $\|\mathbf{u}_{50} - \mathbf{v}\|_2 = 5.5\text{e-}02$ für Teil (d).

8. Für $N = 100$, $h = 1/N$, $x_i = (i - \frac{1}{2})h$, $\mathbf{x} = \{x_i\}_{i=1}^N$, $\mathbf{v} = \text{sign}(\mathbf{x} - \frac{1}{2})$, $\mu = h$ und $\epsilon = h$ soll die Funktion

$$J_h(\mathbf{u}) = \frac{h}{2} \sum_{i=1}^N (u_i - v_i)^2 + \mu h \sum_{i=1}^{N-1} \left[\left(\frac{u_{i+1} - u_i}{h} \right)^2 + \epsilon^2 \right]^{\frac{1}{2}}$$

global minimiert werden. Kreuzen Sie bei den wahren Behauptungen an.

- (a) Der Gradient der Zielfunktion ist gegeben durch

$$\nabla J_h(\mathbf{u})/h = \mathbf{u} - \mathbf{v} + \frac{\mu}{h^2} Q^\top \Phi(\mathbf{u}) Q \mathbf{u}$$

wobei $Q \in \mathbb{R}^{(N-1) \times N}$ und $\Phi(\mathbf{u}) \in \mathbb{R}^{(N-1) \times (N-1)}$ gegeben sind durch

$$Q = \begin{bmatrix} -1 & +1 & & 0 \\ & \ddots & \ddots & \\ 0 & & -1 & +1 \end{bmatrix}, \quad \Phi(\mathbf{u}) = \text{diag} \left\{ \phi\left(\frac{u_{i+1} - u_i}{h}\right) \right\}_{i=1}^{N-1}.$$

wobei $\phi(t) = 1/\sqrt{t^2 + \epsilon^2}$.

- (b) Sei die implizite Fixpunkt Iteration $\mathbf{u}_{k+1} = \mathbf{G}_I(\mathbf{u}_k)$ definiert durch

$$\left[I + \frac{\mu}{h^2} Q^\top \Phi(\mathbf{u}_k) Q \right] \mathbf{u}_{k+1} = \mathbf{v}.$$

Mit $\mathbf{u}_0 = \mathbf{v}$ gilt $\rho(\mathbf{G}'_I(\mathbf{u}_k)) < 0.9$ für $k = 1, \dots, 50$.

- (c) Das Ergebnis der letzten Iteration erfüllt $\|\mathbf{u}_k - \mathbf{u}_{k-1}\|_2 < 10^{-5} \|\mathbf{v}\|_2$ mit $k = 50$ Iterationen.
(d) Das Ergebnis der letzten Iteration erfüllt $\|\mathbf{u}_k - \mathbf{v}\|_2 < 10^{-2} \|\mathbf{v}\|_2$ mit $k = 50$ Iterationen.

Kommentare: Sehen Sie diesen [Code](#). Der Gradient der Zielfunktion

$$J_h(\mathbf{u}) = \frac{h}{2} \sum_{i=1}^N (u_i - v_i)^2 + \mu \sum_{i=1}^{N-1} \sqrt{(u_{i+1} - u_i)^2 + h^2 \epsilon^2}$$

ist gegeben durch

$$\begin{aligned}\frac{\partial J_h}{\partial u_i} &= h(u_i - v_i) + \mu \left[\frac{(u_i - u_{i-1})}{\sqrt{(u_i - u_{i-1})^2 + h^2 \epsilon^2}}_{\delta_{i>1}} - \frac{(u_{i+1} - u_i)}{\sqrt{(u_{i+1} - u_i)^2 + h^2 \epsilon^2}}_{\delta_{i>1}} \right] \\ &= h(u_i - v_i) + \frac{\mu}{h} \left[\phi \left(\frac{u_i - u_{i-1}}{h} \right) (u_i - u_{i-1})_{\delta_{i>1}} - \phi \left(\frac{u_{i+1} - u_i}{h} \right) (u_{i+1} - u_i)_{\delta_{i<N}} \right]\end{aligned}$$

mit $\phi(t) = 1/\sqrt{t^2 + \epsilon^2}$ oder

$$\nabla J_h(\mathbf{u})/h = \mathbf{u} - \mathbf{v} + \frac{\mu}{h^2} Q^\top \Phi(\mathbf{u}) Q \mathbf{u}$$

Die Ableitung $\mathbf{G}'_I(\mathbf{u})$ wird folgendermaßen bestimmt. Die Fixpunkt Gleichung

$$\left[I + \frac{\mu}{h^2} D(\mathbf{u}) \right] \mathbf{G}_I(\mathbf{u}) = \mathbf{v}$$

wird nach \mathbf{u} abgeleitet. Um die Notation zu erleichtern sei

$$\mathbf{y} = \mathbf{y}(\mathbf{u}) = \mathbf{G}_I(\mathbf{u}), \quad \mathbf{y} = \{y_i\}_{i=1}^N.$$

Für die Fixpunkt Gleichung gilt komponentenweise,

$$y_i + \frac{\mu}{h} \left[D_{i-1} \frac{y_i - y_{i-1}}{h}_{\delta_{i>1}} - D_i \frac{y_{i+1} - y_i}{h}_{\delta_{i<N}} \right] = v_i, \quad i = 1, \dots, N.$$

Die Ableitung dieser Gleichung nach u_{i-1} ist

$$\begin{aligned}& \frac{\mu}{h^2} \left[D_{i-1}^3 \left(\frac{u_i - u_{i-1}}{h} \right) \left(\frac{y_i - y_{i-1}}{h} \right) \right]_{\delta_{i>1}} \\ & + \frac{\partial y_i}{\partial u_{i-1}} + \frac{\mu}{h^2} \left[D_{i-1} \left(\frac{\partial y_i}{\partial u_{i-1}} - \frac{\partial y_{i-1}}{\partial u_{i-1}} \right) - D_i \left(\frac{\partial y_{i+1}}{\partial u_{i-1}} - \frac{\partial y_i}{\partial u_{i-1}} \right) \right] = 0.\end{aligned}$$

Die Ableitung nach u_{i+1} ist

$$\begin{aligned}& \frac{\mu}{h^2} \left[D_i^3 \left(\frac{u_{i+1} - u_i}{h} \right) \left(\frac{y_{i+1} - y_i}{h} \right) \right]_{\delta_{i<N}} \\ & + \frac{\partial y_i}{\partial u_{i+1}} + \frac{\mu}{h^2} \left[D_{i-1} \left(\frac{\partial y_i}{\partial u_{i+1}} - \frac{\partial y_{i-1}}{\partial u_{i+1}} \right) - D_i \left(\frac{\partial y_{i+1}}{\partial u_{i+1}} - \frac{\partial y_i}{\partial u_{i+1}} \right) \right] = 0.\end{aligned}$$

Die Ableitung nach u_i ist

$$\begin{aligned}& -\frac{\mu}{h^2} \left[D_{i-1}^3 \left(\frac{u_i - u_{i-1}}{h} \right) \left(\frac{y_i - y_{i-1}}{h} \right) \right]_{\delta_{i>1}} \\ & -\frac{\mu}{h^2} \left[D_i^3 \left(\frac{u_{i+1} - u_i}{h} \right) \left(\frac{y_{i+1} - y_i}{h} \right) \right]_{\delta_{i<N}} \\ & + \frac{\partial y_i}{\partial u_i} + \frac{\mu}{h^2} \left[D_{i-1} \left(\frac{\partial y_i}{\partial u_i} - \frac{\partial y_{i-1}}{\partial u_i} \right) - D_i \left(\frac{\partial y_{i+1}}{\partial u_i} - \frac{\partial y_i}{\partial u_i} \right) \right] = 0.\end{aligned}$$

Die Ableitung der Fixpunkt Gleichung nach u_j , $|i - j| > 1$, ist

$$\frac{\partial y_i}{\partial u_j} + \frac{\mu}{h^2} \left[D_{i-1} \left(\frac{\partial y_i}{\partial u_j} - \frac{\partial y_{i-1}}{\partial u_j} \right) - D_i \left(\frac{\partial y_{i+1}}{\partial u_j} - \frac{\partial y_i}{\partial u_j} \right) \right] = 0.$$

Zusammengefasst gilt

$$\left[I + \frac{\mu}{h^2} D(\mathbf{u}) \right] \mathbf{G}'_I(\mathbf{u}) = \frac{\mu}{h^4} Z$$

wobei mit $Z_i = D_i^3(u_{i+1} - u_i)(y_{i+1} - y_i)$ die Matrix Z gegeben ist durch

$$Z = \begin{bmatrix} Z_1 & -Z_1 & & & & 0 \\ -Z_1 & Z_1 + Z_2 & -Z_2 & & & \\ & \ddots & \ddots & \ddots & & \\ & & -Z_{N-2} & Z_{N-2} + Z_{N-1} & -Z_{N-1} & \\ 0 & & & -Z_{N-1} & Z_{N-1} & \end{bmatrix}.$$

Weiters sehen Sie die Lösung der Hausaufgabe auf Seite 187 im Skriptum.

Zu 2 signifikanten Ziffern gelten $\mathbf{G}_I(\mathbf{u}_{50}) = 9.5e-01$ für Teil (b), $\|\mathbf{u}_{50} - \mathbf{u}_{k-1}\|_2 = 1.8e-07$ für Teil (c) und $\|\mathbf{u}_5 - \mathbf{v}\|_2 = 2.0e-02$ für Teil (d).