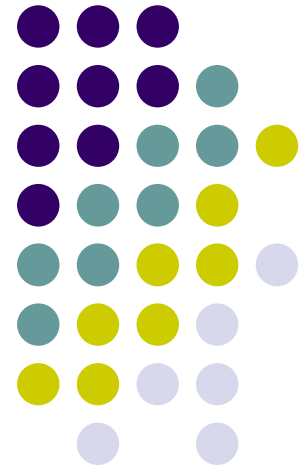
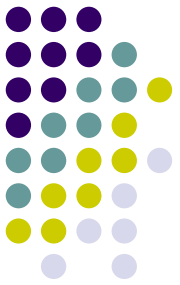


2.2 Residuum basierende Algorithmen

2.2.1 Gewöhnliche Newton Methode



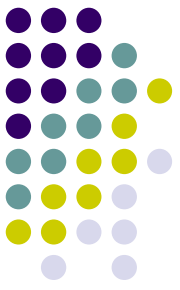


Zur Erinnerung:

$$F'(x^k) \Delta x^k = -F(x^k), \quad x^{k+1} = x^k + \Delta x^k, \quad k = 0, 1, \dots \quad (2.63)$$

Wir analysieren hier die Konvergenz der iterativen Residuen $\{F(x^k)\}$, die zur einer affinen kontravarianten Version vom Newton-Mysovskikh-Theorem führt.

Theorem 2.12:



Sei $F: D \rightarrow \mathbb{R}^n$ eine differenzierbare Funktion mit $D \subset \mathbb{R}^n$ offen und konvex und sei $F'(x)$ invertierbar $\forall x \in D$.

Weiter gelte die affine kontravariante Lipschitz-Bedingung:

$$\|(F'(y) - F'(x))(y - x)\| \leq \omega \|F'(x)(y - x)\|^2 \text{ für } x, y \in D.$$

Definieren die offene Menge:

$$L_\omega := \left\{ x \in D \mid \|F(x)\| < \frac{2}{\omega} \right\} \text{ und sei } \overline{L_\omega} \subset D \text{ beschränkt.}$$

Für einen gegebenen Anfangswert x^0 von einer unbekanntem Lösung x^* sei

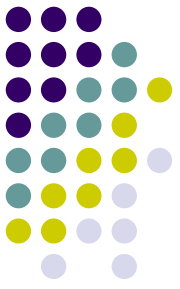
$$h_0 := \omega \cdot \|F(x^0)\| < 2, \text{ dh. } x^0 \in L_\omega \tag{2.64}$$

Dann bleiben die gewöhnlichen Newton Iterierten $\{x^k\}$ in L_ω und konvergieren gegen einen Lösungspunkt $x^* \in L_\omega$ mit $F(x^*) = 0$.

Die iterativen Residuen $\{F(x^k)\}$ konvergieren dann gegen 0 bei einem geschätzten Wert:

$$\|F(x^{k+1})\| \leq \frac{1}{2} \omega \cdot \|F(x^k)\|^2 \tag{2.65}$$

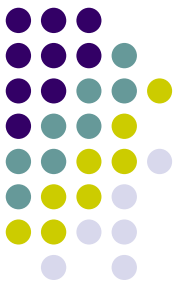
Beweis:



$$\begin{aligned} & \|F(x^k + \lambda \Delta x^k)\| = \|F(x^k) + F(x^k + \lambda \Delta x^k) - F(x^k)\| \\ & = \|F(x^k) + \int_{x^k}^{x^k + \lambda \Delta x^k} F'(u) du\| = \|F(x^k) + \int_0^\lambda F'(x^k + t \Delta x^k) \Delta x^k dt\| \\ & = \|\lambda F(x^k) + \int_0^\lambda F'(x^k + t \Delta x^k) \Delta x^k dt + (1 - \lambda) F(x^k)\| \\ & = \|\lambda F'(x^k) \Delta x^k + \int_0^\lambda F'(x^k + t \Delta x^k) \Delta x^k dt + (1 - \lambda) F(x^k)\| \\ & = \|\int_0^\lambda (F'(x^k + t \Delta x^k) - F'(x^k)) \Delta x^k dt + (1 - \lambda) F(x^k)\| \\ & \leq \int_0^\lambda \| (F'(x^k + t \Delta x^k) - F'(x^k)) \Delta x^k \| dt + (1 - \lambda) \|F(x^k)\| \\ & \leq \omega \int_0^\lambda \|F'(x^k) \Delta x^k\|^2 t dt + (1 - \lambda) \|F(x^k)\| \\ & = \omega \int_0^\lambda \|F(x^k)\|^2 t dt + (1 - \lambda) \|F(x^k)\| = \frac{1}{2} \omega \lambda^2 \|F(x^k)\|^2 + (1 - \lambda) \|F(x^k)\| \\ & = \left(1 - \lambda + \frac{1}{2} \omega \lambda^2 \|F(x^k)\|\right) \|F(x^k)\| \leq (1 - \lambda + \lambda^2) \cdot \|F(x^k)\| \end{aligned}$$

$$\Rightarrow \|F(x^k + \lambda \Delta x^k)\| < \frac{2}{\omega}$$

$$\Rightarrow x^k + t \Delta x^k \in L_\omega \text{ für } t \in [0, \lambda].$$



Nehmen nun an, dass $x^{k+1} \notin L_\omega$. Dann existiert ein minimales $\bar{\lambda} \in (0, 1)$ mit $x^k + \bar{\lambda} \Delta x^k \in \partial L_\omega$

und $\|F(x^k + \bar{\lambda} \Delta x^k)\| < (1 - \bar{\lambda} + \bar{\lambda}^2) \cdot \|F(x^k)\| < \frac{2}{\omega}$, Widerspruch!

Für $\lambda = 1$ bekommt man:

$$\|F(x^k + \Delta x^k)\| = \|F(x^{k+1})\| \leq \left(1 - 1 + \frac{1}{2} \omega \|F(x^k)\|\right) \cdot \|F(x^k)\| = \frac{1}{2} \omega \cdot \|F(x^k)\|^2$$

Um zu zeigen, dass die Newton Iterierten $\{x^k\}$ gegen ein $x^* \in L_\omega$ mit $F(x^*)=0$ konvergieren, führen wir die Kantorovich-Größen ein:

$$h_k := \omega \cdot \|F(x^k)\| \quad (2.66)$$

Mit ihnen erhält man die quadratische Rekursion:

$$h_{k+1} = \omega \cdot \|F(x^{k+1})\| < \frac{1}{2} \omega^2 \|F(x^k)\|^2 = \frac{1}{2} h_{k^2} = \left(\frac{1}{2} h_k\right) \cdot h_k \quad (2.67)$$

Wegen der Voraussetzung $h_0 < 2$ erhält man dann:

$$h_1 < \left(\frac{1}{2} h_0\right) \cdot h_0 < 2.$$

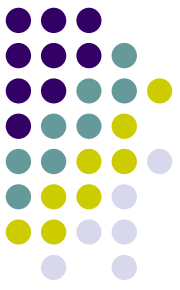
Durch wiederholte Induktion über k :

$$h_{k+1} < \left(\frac{1}{2} h_k\right) \cdot h_k < h_k < 2, \quad k=0,1,2,\dots \quad \Rightarrow \quad \lim_{k \rightarrow \infty} h_k = 0.$$

erhält man eine *monoton* fallende Folge, die gegen 0 konvergiert.

Setzen wir jetzt wieder die Residuen $\{F(x^k)\}$ ein, erhalten wir

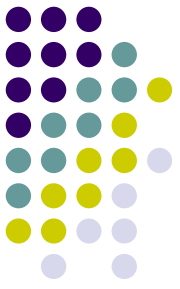
$$\|F(x^{k+1})\| < \|F(x^k)\| < \frac{2}{\omega} \quad \Rightarrow \quad \lim_{k \rightarrow \infty} \|F(x^k)\| = 0.$$



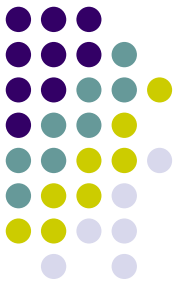
Bezüglich der Iterierten $\{x^k\}$ haben wir dann $\{x^k\} \subset L_\omega$.

Und da L_ω beschränkt ist, existiert ein Häufungspunkt x^* von $\{x^k\}$ mit $F(x^*)=0$.

D.h. x^* ist unser Lösungspunkt in L_ω , aber nicht notwendigerweise eindeutig.



Konvergenz-Monitor



Als nächstes wollen wir uns die Konvergenz des Theorems genauer betrachten. Dazu führen wir Kontraktionsfaktoren ein:

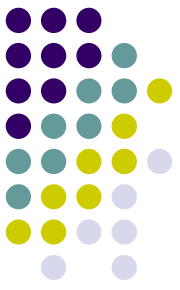
$$\theta_k := \frac{\|F(x^{k+1})\|}{\|F(x^k)\|}$$

und schreiben $h_{k+1} < \frac{1}{2} h_k^2$ um in

$$\theta_k = \frac{h_{k+1}}{h_k} < \frac{1}{2} h_k \quad (2.68)$$

Für $k=0$ und wegen $h_0 < 2$ erhalten wir die

$$\text{Residuums-Monotonie} \quad \theta_0 < 1 \quad (2.69)$$



Wann immer also $\theta_0 \geq 1$ ist, ist die Voraussetzung $h_0 < 2$ im Theorem 2.12 verletzt, was soviel heißt wie:

Der geschätzte Anfangswert x^0 ist nicht genügend dicht am Lösungspunkt x^* .

Wir nehmen nun an, dass $\theta_0 < 1$ ist.

Für die Konstruktion eines quadratischen Konvergenz-Monitors führen wir, für die unbekanntes theoretischen Größen h_k , rechenbetonte gültige

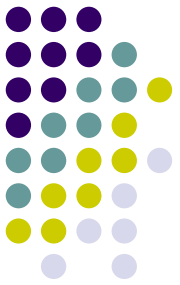
Abschätzungen $[h_k]$ ein.

Mit $\theta_k = \frac{h_{k+1}}{h_k} \leq \frac{1}{2} h_k$ definieren wir die rechenbetonte a-posteriori

Abschätzung: $[h_k]_1 := 2\theta_k \leq h_k$

und, wenn $h_{k+1} = \theta_k h_k$ ist, die a-priori Abschätzung

$$[h_{k+1}] = \theta_k [h_k]_1 = 2\theta_k^2 = \frac{h_{k+1}^2}{\frac{1}{2} h_k^2} \leq \frac{h_{k+1}^2}{h_{k+1}} = h_{k+1}.$$



Durch die Identifikation von $[h_{k+1}] \approx [h_{k+1}]_1$ bekommen wir die ungefähre Rekursion ($k=0,1,2,\dots$):

$$\theta_{k+1} \approx \theta_k^2 \leq \theta_0 < 1.$$

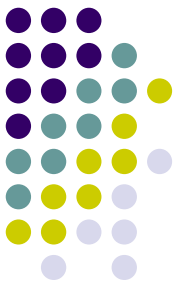
Die Verletzung dieser Rekursion im milderen Fall:

$$\theta_{k+1} > \theta_0$$

oder im strengeren Fall:

$$\theta_{k+1} \geq 2\theta_k^2$$

bricht die gewöhnliche Newton Iteration als *nicht konvergent* ab.

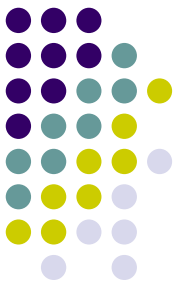


Abbruchs-Kriterium

Diese affine kontravariante Theorie stimmt mit dem Abbruchs-Kriterium der Form

$$\|F(\bar{x})\| \leq FTOL \quad (2.70)$$

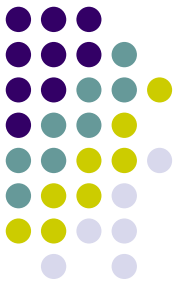
überein, wobei FTOL eine vom Benutzer vorgeschriebener Residuums-Fehler-Toleranz-Wert ist.



Rechenbetonter Aufwand

Eine kurze Berechnung zeigt, dass für einen gegebenen Startwert x^0 , die Anzahl q der Iterationen (so dass $\bar{x} = x^{q+1}$ die obige Abbruchs-Anforderung erreicht) erfüllt annähernd:

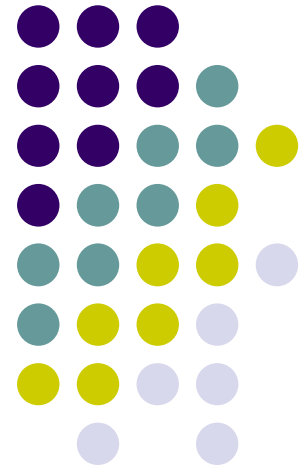
$$q \approx \text{ld} \frac{\log(\text{FTOL} / \|F(x^0)\|)}{\log \theta_0} \quad (2.71)$$

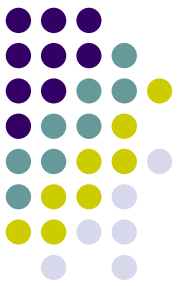


Mit anderen Worten, mit einem genügend gutem Startwert x^0 von einer vorliegenden Lösung x^* , ist der rechenbetonter Aufwand vom nicht-linearem Problem vergleichbar mit dem Aufwand vom linearisierten Problem.

Solche Probleme werden manchmal auch *einbisschen* $\{ \}$ # *nicht-linear* genannt.

2.2.2 Vereinfachte Newton Methode





Zur Erinnerung:

$$F'(x^0)\overline{\Delta x^k} = -F(x^k), \quad x^{k+1} = x^{k+1} + \overline{\Delta x^k}, \quad k=0,1,2,\dots \quad (2.72)$$

Wir analysieren hier die Konvergenz bezüglich der iterativen Residuen $\{F(x^k)\}$, um eine affine kontravariante Nebenform des Newton-Kontorovich-Theorems zu bekommen -ohne irgendwelche eindeutige Resultate.

Theorem 2.12:

Sei $F:D\rightarrow\mathbb{R}^n$ in $C^1(D)$ für $D\subset\mathbb{R}^n$ konvex. Außerdem sei $x^0\in D$ ein Anfangswert für die vereinfachte Newton Iteration.

Es gelte die folgende affine kontravariante Lipschitz-Bedingung:

$$\|F'(x)-F'(x^0)\|v\|\leq\omega\|F'(x^0)(x-x^0)\|\cdot\|F'(x^0)v\| \quad (2.73)$$

für $x,x^0\in D$, $v\in\mathbb{R}^n$ und $0\leq\omega<\infty$.

Definieren die Menge

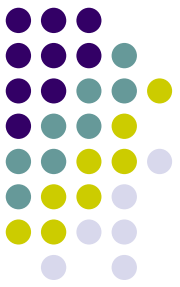
$$L_\omega:=\left\{x\in\mathbb{R}^n\mid\|F(x)\|\leq\frac{1}{2\omega}\right\} \quad \text{und sei } \overline{L}_\omega\subseteq D \text{ beschränkt.}$$

$$\text{Für } x^0\in L_\omega \text{ ist } h_0:=\omega\cdot\|F(x^0)\|\leq\frac{1}{2} \quad (2.74)$$

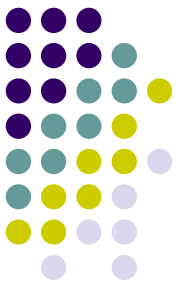
Dann bleiben die Iterierten $\{x^k\}$ in L_ω und konvergieren gegen einen Lösungspunkt x^* . Die iterativen Residuen $\{F(x^k)\}$ Normen konvergieren gegen 0 bei einem geschätzten Wert:

$$\frac{\|F(x^{k+1})\|}{\|F(x^k)\|}\leq\frac{1}{2}(t_k+t_{k+1})<1-\sqrt{1-2h_0},$$

$$\text{mit } t_0=0 \text{ und } t_{k+1}=h_0+\frac{1}{2}t_k^2, \quad k=0,1,2,\dots$$



Beweis:



$$\begin{aligned} \|F(x^{k+1})\| &= \|F(x^k + \overline{\Delta x^k}) + F(x^k) - F(x^k)\| \\ &= \left\| -F'(x^0)\overline{\Delta x^k} + \int_0^1 F'(x^k + t\overline{\Delta x^k})\overline{\Delta x^k} dt \right\| = \left\| \int_0^1 (F'(x^k + t\overline{\Delta x^k}) - F'(x^0))\overline{\Delta x^k} dt \right\| \\ &\leq \int_0^1 \| (F'(x^k + t\overline{\Delta x^k}) - F'(x^0))\overline{\Delta x^k} \| dt \\ &\leq \omega \cdot \|F'(x^0)\overline{\Delta x^k}\| \cdot \int_0^1 \|F'(x^0)(x^k - x^0 + t\overline{\Delta x^k})\| dt \\ &\leq \omega \|F(x^k)\| \cdot \left(\|F'(x^0)(x^k - x^0)\| + \frac{1}{2} \|F'(x^0)\overline{\Delta x^k}\| \right) \\ &= \omega \|F(x^k)\| \cdot \left(\|F'(x^0)(x^k - x^0)\| + \frac{1}{2} \|F(x^k)\| \right) \end{aligned} \tag{2.75}$$

Führen nun die Majoranten ein:

$$\omega \|F'(x^0)(x^k - x^0)\| \leq t_k$$

und
$$\omega \|F'(x^0)(x^{k+1} - x^k)\| = \omega \|F(x^k)\| \leq h_k$$

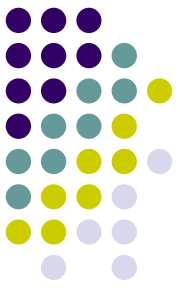
mit den Anfangswerten $t_0 = 0$ und $h_0 = \frac{1}{2}$.

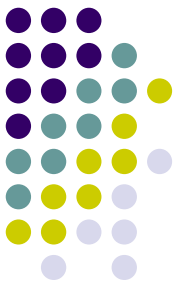
Mit
$$\begin{aligned} \|F'(x^0)(x^{k+1} - x^0)\| &= \|F'(x^0)(x^{k+1} - x^k + x^k - x^0)\| \\ &= \|F'(x^0)(x^{k+1} - x^k) + F'(x^0)(x^k - x^0)\| \\ &\leq \|F'(x^0)(x^{k+1} - x^k)\| + \|F'(x^0)(x^k - x^0)\| \end{aligned}$$

und mit $\|F(x^k)\| \leq \omega \|F(x^{k-1})\| \cdot \left(\|F'(x^0)(x^{k-1} - x^0)\| + \frac{1}{2} \|F(x^{k-1})\| \right)$ bekommen

wir dieselben Majoranten-Gleichungen wie in Kapitel 2.12:

$$t_{k+1} = h_k + t_k \quad \text{und} \quad h_k = h_{k-1} \left(t_{k-1} + \frac{1}{2} h_{k-1} \right)$$





und mit diesen beiden Gleichungen erhalten wir:

$$t_{k+1} - t_k = h_k = h_{k-1} \cdot \left(t_{k-1} + \frac{1}{2} h_{k-1} \right) = (t_k - t_{k-1}) \cdot \left(t_{k-1} + \frac{1}{2} (t_k - t_{k-1}) \right)$$

$$\dot{i} (t_k - t_{k-1}) \cdot \frac{1}{2} (t_k + t_{k-1}) = \frac{1}{2} (t_k^2 - t_{k-1}^2)$$

Die Umstellung dieser Gleichung erlaubt uns die Anwendung des Ortega-Tricks:

$$t_{k+1} - \frac{1}{2} t_k^2 = t_k - \frac{1}{2} t_{k-1}^2 = t_{k-1} - \frac{1}{2} t_{k-2}^2 = \dots = t_1 - \frac{1}{2} t_0^2 = t_1 = h_0$$

Mit der skalaren Gleichung:

$$g(t) = h_0 - t + \frac{1}{2} t^2 = 0$$

erhalten wir wieder

$$t_{k+1} - t_k = t_{k+1} - \frac{1}{2} t_k^2 - t_k + \frac{1}{2} t_k^2 = h_0 - t_k + \frac{1}{2} t_k^2 = -\frac{g(t_k)}{g'(t_0)} = g(t_k) = h_k$$

und somit

$$g(t_{k+1}) < g(t_k)$$

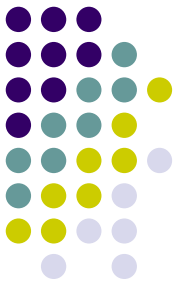
⇒

$$h_{k+1} < h_k$$

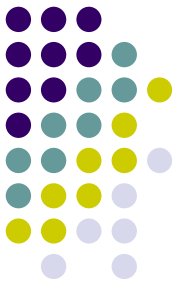
⇒

$$\|F(x^{k+1})\| < \|F(x^k)\| \leq \frac{1}{2\omega}$$

$$\omega \|F(x^k)\| \leq h_k$$



Damit ist gezeigt, dass die vereinfachten Newton Iterierten $\{x^k\}$ in $L_\omega \subset D$ bleiben. Für die Konvergenz gegen einen Lösungspunkt $x^* \in L_\omega \subset D$ können die gleichen Argumente angewandt werden, wie im Beweis davor.



Für die Konvergenzrate benutzen wir die Gleichung

$$\|F(x^{k+1})\| \leq \omega \cdot \|F(x^k)\| \cdot \left(\|F'(x^0)(x^k - x^0)\| + \frac{1}{2} \|F(x^k)\| \right)$$

und erhalten:

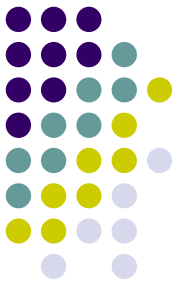
$$\frac{\|F(x^{k+1})\|}{\|F(x^k)\|} \leq \omega \cdot \|F'(x^0)(x^k - x^0)\| + \frac{1}{2} \omega \cdot \|F(x^k)\|$$

$$\leq t_k + \frac{1}{2} h_k$$

$$< t_* \cdot \sqrt{1 - 2h_0}$$



Konvergenz-Monitor



Für eine effektive Anwendung dieses Theorems definieren wir die Residuums-Kontraktionsfaktoren ($k=0,1,2,\dots$)

$$\Theta_k := \frac{\|F(x^{k+1})\|}{\|F(x^k)\|} \leq \frac{1}{2}(t_k + t_{k+1}).$$

Für $k=0$ ist der lokale Konvergenz-Bereich charakterisiert

mit
$$\Theta_0 \leq \frac{1}{2} t_1 = \frac{1}{2} h_0 \leq \frac{1}{4},$$

welches um einiges einschränkender ist als die Bedingung $\Theta_0 < 1$ für die gewöhnliche Newton Methode.