2.2 Residuum basierende Algorithmen

2.2.1 Gewöhnliche Newton Methode

Zur Erinnerung:

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

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 $F: D \to \Re^n$ eine differenzierbare Funktion mit $D \subset \Re^n$ offen und konvex und sei $F'(x)$ invertierbar $\forall x \in D$.

Weiter gelte die affine kontravariante Lipschitz-Bedingung:

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

Definieren die offene Menge:

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

Für einen gegebenen Anfangswert x^0 von einer unbekannten 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:

$$\left\|F\left(x^{k}+\lambda\Delta x^{k}\right)\right\|=\left\|F\left(x^{k}\right)+F\left(x^{k}+\lambda\Delta x^{k}\right)-F\left(x^{k}\right)\right\|$$

$$=\left\|F\left(x^{k}\right)+\int_{x^{k}}^{x^{k}+\lambda\Delta x^{k}} F'(u)\,du\right\| \quad =\left\|F\left(x^{k}\right)+\int_{0}^{\lambda} F'\left(x^{k}+t\Delta x^{k}\right)\Delta x^{k}\,dt\right\|$$

$$=\left\|\lambda F\left(x^{k}\right)+\int_{0}^{\lambda} F'\left(x^{k}+t\Delta x^{k}\right)\Delta x^{k}\,dt+(1-\lambda)F\left(x^{k}\right)\right\|$$

$$=\left\|-\lambda F'\left(x^{k}\right)\Delta x^{k}+\int_{0}^{\lambda} F'\left(x^{k}+t\Delta x^{k}\right)\Delta x^{k}\,dt+(1-\lambda)F\left(x^{k}\right)\right\|$$

$$=\left\|\int_{0}^{\lambda}\left(F'\left(x^{k}+t\Delta x^{k}\right)-F'\left(x^{k}\right)\right)\Delta x^{k}\,dt+(1-\lambda)F\left(x^{k}\right)\right\|$$

¿ $$\int_{0}^{\lambda}\left\|\left(F'\left(x^{k}+t\Delta x^{k}\right)-F'\left(x^{k}\right)\right)\Delta x^{k}\right\|dt+(1-\lambda)\left\|F\left(x^{k}\right)\right\|$$

¿ $$\omega\int_{0}^{\lambda}\left\|F'\left(x^{k}\right)\Delta x^{k}\right\|^{2} t\,dt+(1-\lambda)\left\|F\left(x^{k}\right)\right\|$$

$$=\omega\int_{0}^{\lambda}\left\|F\left(x^{k}\right)\right\|^{2} t\,dt+(1-\lambda)\left\|F\left(x^{k}\right)\right\| \quad =\frac{1}{2}\omega\lambda^{2}\left\|F\left(x^{k}\right)\right\|^{2}+(1-\lambda)\left\|F\left(x^{k}\right)\right\|$$

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

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

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

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:

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

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:

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

Mit ihnen erhält man die quadratische Rekursion:

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

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:

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

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

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

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

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

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

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

Residuums-Monotonie $\Theta_0 < 1$ $\qquad (2.69)$

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 unbekannten theoretischen Größen h_k, rechenbetonte gültige Abschätzungen $[h_k]$ ein.

Mit $\Theta_k = \dfrac{h_{k+1}}{h_k} \leq \dfrac{1}{2} h_k$ 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 $(k = 0,1,2,\ldots)$*:*

$$\Theta_{k+1} \approx \Theta_k^2 \leq \Theta_0 < 1.$$

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

$$\|F(\overline{x})\| \leq FTOL \qquad (2.70)$$

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:

$$q \approx \mathrm{ld} \frac{\log\left(\mathrm{FTOL}/\|F(x^0)\|\right)}{\log \Theta_0} \qquad (2.71)$$

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:

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

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 $F:D\to\Re^n$ in $C^1(D)$ für $D\subset\Re^n$ konvex. Außerdem sei $x^0\in D$ *ein* Anfangswert für die vereinfachte Newton Iteration.

Es gelte die folgende affine kontravariante Lipschitz-Bedingung:

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

für $x,x^0\in D$, $v\in\Re^n$ und $0\leq\omega<\infty$.

Definieren die Menge

$$L_\omega:=\left\{x\in\Re^n \,\middle|\, \|F(x)\|\leq\frac{1}{2\omega}\right\}$$ *und* sei $\overline{L_\omega}\subseteq D$ beschränkt.

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

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:

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

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

Beweis:

$$\left\|F(x^{k+1})\right\| = \left\|F(x^k + \overline{\Delta x^k}) + F(x^k) - F(x^k)\right\|$$

$$= \left\| -F'(x^0)\overline{\Delta x^k} + \int_0^1 F'(x^k + t\overline{\Delta x^k})\overline{\Delta x^k}\,dt \right\| = \left\| \int_0^1 \left(F'(x^k + t\overline{\Delta x^k}) - F'(x^0)\right)\overline{\Delta x^k}\,dt \right\|$$

¿ $$\int_0^1 \left\| \left(F'(x^k + t\overline{\Delta x^k}) - F'(x^0)\right)\overline{\Delta x^k} \right\| dt$$

¿ $$\omega \cdot \left\| F'(x^0)\overline{\Delta x^k} \right\| \cdot \int_0^1 \left\| F'(x^0)(x^k - x^0 + t\overline{\Delta x^k}) \right\| dt$$

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

$$= \omega \left\| F(x^k) \right\| \cdot \left(\left\| F'(x^0)(x^k - x^0) \right\| + \frac{1}{2} \left\| F(x^k) \right\| \right) \qquad (2.75)$$

Führen nun die Majoranten ein:

$$\omega\left\|F'(x^0)(x^k - x^0)\right\| \le t_k$$

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

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

Mit $\left\|F'(x^0)(x^{k+1} - x^0)\right\| = \left\|F'(x^0)(x^{k+1} - x^k + x^k - x^0)\right\|$

$$= \left\|F'(x^0)(x^{k+1} - x^k) + F'(x^0)(x^k - x^0)\right\|$$

$$\le \left\|F'(x^0)(x^{k+1} - x^k)\right\| + \left\|F'(x^0)(x^k - x^0)\right\|$$

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

wir dieselben Majoranten-Gleichungen wie in Kapitel 2.12:

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

und mit diesen beiden Gleichungen erhalten wir:

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

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

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

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

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

$\Rightarrow \quad h_{k+1} < h_k$

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

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 benützen 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:

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

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

$$< t * ¿ 1 - \sqrt{1 - 2h_0}$$

$\square$

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.