Der Satz von Newton Kantorovich

Nils Bechtloff

28. Mai 2008

1 Einführung

Wir betrachten das gewöhnliche Newton-Verfahren

$$F'(x^k)\Delta x^k = -F(x^k), \qquad x^{k+1} = x^k + \Delta x^k, \qquad k = 0, 1, ... \tag{1}$$

Um das Newton Verfahren anwenden zu können, muss notwendigerweise $F'(x^k)^{-1}$ für jedes x^k existieren und beschränkt sein.

Der Satz sagt folgendes aus:
Gilt die (affin invariante) Lipschitzbedingung, die Beschränktheit von $F'(x)^{-1}$ und ist $h_0 = \left\|x^1 - x^0\right\| \overline{\omega}_0 < \frac{1}{2}$, dann folgt quadratische Konvergenz gegen die eindeutige Nullstelle. Außerdem bleiben alle Iterationspunkte in der Kugel um x^0 mit Radius $\rho_- := (1 - \sqrt{1 - 2h_0})/\overline{\omega}_0$.
Eindeutigkeit beweisen wir nicht.

Theorem 1.1. *Sei $F : D \to Y$ eine stetig Fréchet differenzierbare Abbildung, $D \subseteq X$ offen und konvex. Für einen Startwert $x^0 \in D$ sei $F'(x^0)$ invertierbar. Gilt*

$$\left\|F'(x^0)^{-1}F(x^0)\right\| \leq \alpha,$$

$$\left\|F'(x^0)^{-1}(F'(y) - F'(x))\right\| \leq \overline{\omega}_0 \left\|y - x\right\| \qquad x, y \in D, \tag{2}$$

$$h_0 := \alpha\overline{\omega}_0 \leq \frac{1}{2}, \tag{3}$$

$$\overline{S}(x^0, \rho_-) \subset D, \qquad \rho_- := (1 - \sqrt{1 - 2h_0})/\overline{\omega}_0,$$

dann existiert die Folge $(x^k)_{k\in\mathbb{N}}$ der Newton Iterierten, bleibt in $\overline{S}(x^0, \rho_-)$ und konvergiert gegen ein x^* mit $F(x^*) = 0$. Für $h_0 < \frac{1}{2}$ ist die Konvergenz quadratisch.

2 Beweis

Der Beweis wird mittels Induktion über k geführt.
Nach den Voraussetzungen ist $F'(x^0)$ invertierbar und beschränkt. Wir nehmen an $F'(x^k)$ ist invertierbar und beschränkt und die folgende Lipschitzbedingung gelte:

$$\left\| F'(x^k)^{-1}(F'(y) - F'(x)) \right\| \leq \overline{\omega}_k \left\| y - x \right\|.$$

Wir definieren die erste Majorante

$$\overline{\omega}_k \left\| y - x \right\| \leq h_k.$$

Um den Induktionsschritt auszuführen, definieren wir den Operator

$$B_{k+1} := F'(x^k)^{-1} F'(x^{k+1}),$$

sowie die zweite Majorante

$$\left\| B_{k+1}^{-1} \right\| \leq \beta_{k+1}.$$

Ist B_{k+1} invertierbar, dann ist (mit $F'(x^k)$ invertierbar) auch $F'(x^{k+1})$ invertierbar.
Als Folgerung aus der Lipschitzbedingung für $k-1$ und der Definition der zweiten Majorante erhalten wir zuerst:

$$\overline{\omega}_{k-1} \left\| y - x \right\| \geq \left\| F'(x^{k-1})^{-1}(F'(y) - F'(x)) \right\|$$

$$\beta_k \overline{\omega}_{k-1} \left\| y - x \right\| \geq \beta_k \left\| F'(x^{k-1})^{-1}(F'(y) - F'(x)) \right\|$$

$$\beta_k \overline{\omega}_{k-1} \left\| y - x \right\| \geq \left\| F'(x^k)^{-1} F'(x^{k-1}) \right\| \left\| F'(x^{k-1})^{-1}(F'(y) - F'(x)) \right\|$$

$$\beta_k \overline{\omega}_{k-1} \left\| y - x \right\| \geq \left\| F'(x^k)^{-1} F'(x^{k-1}) F'(x^{k-1})^{-1}(F'(y) - F'(x)) \right\|$$

$$\beta_k \overline{\omega}_{k-1} \left\| y - x \right\| \geq \left\| F'(x^k)^{-1}(F'(y) - F'(x)) \right\|.$$

Wir schätzen die rechte Seite mit der Lipschitzbedingung für k ab und erhalten:

$$\overline{\omega}_k \leq \beta_k \overline{\omega}_{k-1}. \tag{4}$$

Wir wenden jetzt das Störungslemma für Operatoren an, um zu zeigen, dass B_{k+1} invertierbar ist.

Lemma 2.1. *Seien A, C lineare Operatoren zwischen den Banachräumen X und Y. Sei A invertierbar mit $\left\|A^{-1}\right\| \leq \alpha$. Ist $\|A - C\| \leq \beta$ und $\alpha\beta < 1$, dann ist auch C invertierbar und es gilt*

$$\left\|C^{-1}\right\| \leq \frac{\alpha}{1-\alpha\beta}.$$

Wir wählen $A = I$, die Identität, $C = B_{k+1}$, und vergewissern uns, dass die Vorraussetzungen des Störungslemma erfüllt sind:
$\|I - B_{k+1}\| = \left\|F'(x^k)^{-1}(F'(x^k) - F'(x^{k+1}))\right\| \leq \overline{\omega}_k \left\|\Delta x^k\right\| \leq h_k \leq \beta$
$\|A\| = \|I\| = 1 \leq \alpha$
(Für $k = 0$ ist die Induktionsvorrausetzung mit $\alpha\beta \leq h_0 \leq \frac{1}{2} \leq 1$ erfüllt). Aus dem Störungslemma folgt, dass B_{k+1} invertierbar ist (also auch $F'(x^{k+1})$), mit $\left\|B_{k+1}^{-1}\right\| \leq \frac{1}{1-h_k}$ und wir definieren β_{k+1}:

$$\beta_{k+1} = \frac{1}{1-h_k}. \tag{5}$$

Damit der Induktionsschritt funktioniert benötigen wir, dass (h_k) als Funktion von h_{k-1} eine Kontraktion ist, denn dann folgt mit dem Banachschen Fixpunktsatz die Konvergenz.

Hierfür benötigen wir folgende Hilfsgleichung:

$$\left\|F'(x^k)^{-1}(F(y) - F(x) - F'(x)(y-x))\right\| \leq \frac{\overline{\omega}_k}{2} \|y - x\|^2 \tag{6}$$

Diese Gleichung erhalten wit, indem wir die Differenz der Funktionswerte über das Integral ausdrücken:

$$F(y) - F(x) = F(t)\Big|_x^y = \int_x^y F'(t)dt$$

durch Substitution mit $g(s) = x + s(y - x)$ erhalten wir:

$$F(y) - F(x) = \int_{s=0}^1 F'(x + s(y-x))(y-x)ds$$

und damit:

$$\begin{aligned} F(y) - F(x) - F'(x)(y-x) &= \int_{s=0}^1 (F'(x+s(y-x)) - F'(x))(y-x)ds \\ F'(x^k)^{-1}(F(y) - F(x) - F'(x)(y-x)) &= \int_{s=0}^1 F'(x^k)^{-1}(F'(x+s(y-x)) - F'(x))(y-x)ds. \end{aligned}$$

Wir schätzen die rechte Seite mit der Lipschitzbedingung ab:

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

$$\leq \int_{s=0}^{1} \underbrace{\left\| F'(x^k)^{-1}(F'(x+s(y-x)) - F'(x)) \right\|}_{Lipschitz-Bedingung} \|(y-x)\| \, ds$$

$$\leq \int_{s=0}^{1} \overline{\omega}_k \, \|x + s(y-x) - x\| \, \|(y-x)\| \, ds$$

$$= \int_{s=0}^{1} s\overline{\omega}_k \, \|(y-x)\|^2 \, ds = \frac{\overline{\omega}_k}{2} \, \|(y-x)\|^2 .$$

Nun können wir die Konvergenzgeschwindigkeit bestimmen. Da $F(x^*) = 0$ gilt

$$\begin{aligned} x^{k+1} - x^* &= x^k \quad \underbrace{-F'(x^k)^{-1}F(x^k)}_{=\Delta x^k \text{(Newton-Iteration)}} \quad -x^* \\ &= x^k - x^* - F'(x^k)^{-1}(F(x^k) - F(x^*)) \\ &= F'(x^k)^{-1}(F'(x^*) - F(x^k) - F'(x^k)(x^* - x^k)) \end{aligned}$$

und mit (6) folgt:

$$\left\| x^{k+1} - x^* \right\| \leq \frac{\overline{\omega}_k}{2} \left\| x^k - x^* \right\|^2 .$$

Kombination von zwei Newton-Schritten für k und $k-1$ liefert:

$$\begin{aligned} F'(x^k)\Delta x^k = -F(x^k) \qquad \text{und} \qquad F'(x^{k-1})\Delta x^{k-1} = -F(x^{k-1}) \\ F'(x^k)\Delta x^k = -F(x^k) + F'(x^{k-1})\Delta x^{k-1} + F(x^{k-1}) \\ \Delta x^k = -F'(x^k)^{-1}(F(x^k) - F'(x^{k-1})\Delta x^{k-1} - F(x^{k-1})) \end{aligned}$$

Vergleich mit der linken Seite von (6) mit $y = x^k$ und $x = x^{k-1}$ liefert:

$$\left\| \Delta x^k \right\| \leq \frac{\overline{\omega}_k}{2} \left\| \Delta x^{k-1} \right\|^2 .$$

Wir schätzen mit (4) und der Majoranten h_{k-1} ab:

$$\overline{\omega}_k \left\| \Delta x^k \right\| \leq \frac{\overline{\omega}_k}{2} \left\| \Delta x^{k-1} \right\|^2 \leq \frac{1}{2}\beta_k^2 \overline{\omega}_{k-1} \left\| \Delta x^{k-1} \right\|^2 \leq \frac{1}{2}\beta_k^2 h_{k-1}^2 =: h_k \tag{7}$$

Kombination von (7) und (5) liefert uns die Rekursion für h_k:

$$h_k = \frac{\frac{1}{2} h_{k-1}^2}{(1 - h_{k-1})^2}.$$

Wir erhalten die Kontraktionseigenschaft, falls $\frac{h_k}{h_{k-1}} < 1$.

Anmerkung: f ist eine Kontraktion, wenn: $\exists\lambda : \forall x, y \in M, (M,d)$ metrischer Raum $d(f(x)-f(y)) \leq \lambda d(x,y)$. Man sieht, dass für $f(x) < x, \quad x, f(x) > 0$ und o.B.d.A $x > y$ direkt $\|f(x) - f(y)\| < \|x - y\|$, also Kontraktion folgt.

Wir fordern also:

$$\frac{h_k}{h_{k-1}} = \frac{\frac{1}{2}h_{k-1}}{(1-h_{k-1})^2} < 1.$$

Für die Induktionsverankerung muss also

$$\frac{\frac{1}{2}h_0}{(1-h_0)^2} < 1$$

gelten, woraus wir $h_0 < \frac{1}{2}$ schließen.
Wir betrachten noch den Grenzfall $h_0 = \frac{1}{2}$:
Aus der Rekursionsformel folgt direkt $h_k = \frac{1}{2} \qquad k = 1, 2, ...$, mit (5) und (4) folgt dann

$$\beta_k = 2, \qquad \overline{\omega}_k \leq 2^k \overline{\omega}_0.$$

Eingesetzt in (7) erhalten wir:

$$\lim_{k\to\infty} \overline{\omega}_k \left\|\Delta x^k\right\| \leq \lim_{k\to\infty} 2^k \overline{\omega}_0 \left\|\Delta x^k\right\| \leq \frac{1}{2},$$

woraus die Konvergenz folgt:

$$\lim_{k\to\infty} \left\|\Delta x^k\right\| \leq \lim_{k\to\infty} \frac{1}{2^{k-1}\overline{\omega}_0} = 0.$$