

Seminar

Newton-Methoden

Universität Heidelberg
Wintersemester 2004/05
unter der Leitung von Dr. Ekaterina Kostina

Globalisierung der Konvergenz bei
verallgemeinerten Gauß-Newtonverfahren
und
Fortsetzungsverfahren für beschränkte
Ausgleichsprobleme

Vortragender

Jens Schöbel

am

16. Februar 2005

Inhaltsverzeichnis

1	Vorwort	2
2	Globalisierung der Konvergenz	3
2.1	Niveaufunktion im Bildraum	3
2.2	Natürliche Niveaufunktion	6
2.3	Schrittweitenstrategien für natürliche Niveaufunktionen	7
2.4	Praktische Realisierung	8
2.5	Ein Algorithmus zur Schrittweitensteuerung	9
3	Ein Fortsetzungsverfahren für beschränkte Ausgleichsprobleme	10
3.1	Wahl des Iterationsverfahrens	10
3.2	Mögliche Einbettungen	11
3.3	Wahl des Prädiktors	11
3.4	Wahl der Schrittweite	12
3.5	Automatische Parametrisierung	14

1 Vorwort

Dieser Vortrag stützt sich auf Beweise und Erkenntnisse aus dem Vortrag meines Vordrners. Zum besseren Verständnis sind beide Vorträge als Einheit zu betrachten.

Im vorliegenden Handout steht die Abkürzung VGN für *Verallgemeinertes Gauß-Newton-Verfahren*.

Der erste Abschnitt behandelt die Grundlagen, um eine globale Konvergenz bei VGN zu erhalten. Am Ende dieses Kapitels wird ein Algorithmus zur Schrittweitensteuerung vorgestellt.

Im zweiten Abschnitt wird ein Fortsetzungsverfahren für beschränkte Ausgleichsprobleme vorgestellt. Es wird das Problem definiert und auf Verbesserungen von gewissen Strategien eingegangen werden.

2 Globalisierung der Konvergenz

Um mit einem Verfahren eine optimale Lösung zu erhalten, ist es von entscheidender Bedeutung, dass die lokalen Konvergenzbedingungen bezüglich eines Startwertes x_0 erfüllt sind. Im Allgemeinen kann man diese Voraussetzung nicht treffen. Es besteht die Möglichkeit einen Startwert zu wählen, der *zu weit weg* vom Optimalwert x^* liegt. Es liegt nahe zu versuchen den Konvergenzbereich zu vergrößern.

Die Iteration

$$y_{k+1} = x_k + \Delta x_k \quad (1)$$

wird daher verallgemeinert zu:

$$y_{k+1} = x_k + t_k \Delta x_k \quad (2)$$

Man spricht in diesem Fall von einer **Dämpfung** oder **Relaxion mit Relaxationsfaktor** t_k .

Als Maß für die Annäherung dient hier eine **Niveaufunktion** $T(x)$. Diese Funktion muss streng monoton fallend sein. Es gilt also:

$$T(x_k + t_k \Delta x_k) < T(x_k) \quad (3)$$

Eine solche Niveaufunktion muss natürlich in Richtung des Inkrementes von Δx fallen. Es muss also zusätzlich folgende Bedingung gelten:

$$\Delta x \neq 0 \Rightarrow \left. \frac{\partial}{\partial \varepsilon} T(x + \varepsilon \Delta x) \right|_{\varepsilon > 0} < 0 \quad (4)$$

2.1 Niveaufunktion im Bildraum

Im folgenden wird die globale Konvergenz der Niveaufunktion

$$T_1(x) = \frac{1}{2} \|F_1(x)\|_2^2 + \sum_{i=1}^{n_2} \beta_i |F_{2i}(x)| + \sum_{i=1}^{n_3} \gamma_i |\min(0, F_{3i}(x))| \quad (5)$$

gezeigt. Für die Bestimmung der Schrittweite t_k wird dabei folgender Algorithmus zu Grunde gelegt.

Algorithmus: Bestimmung Schrittweite t_k (i) löse $[PL(x_k)]$ (ii) bestimme t_k so, dass

$$T_1(x_k + t_k \Delta x_k) = \min_{t \in [0,1]} T_1(x_k + t \Delta x_k)$$

(iii) setze $y_{k+1} = x_k + t_k \Delta x_k$ und gehe zu (i)

Dabei lässt sich die Bestimmung des Minimums in (ii) durch eine Näherungslösung ersetzen.

Lemma 2.1.1 (Verträglichkeit von T_1 bzgl. VGN). *Sei $(\Delta x, \lambda, \mu)$ ein Kuhn-Tucker-Punkt des linearisierten Problems $[PL(x)]$, $x \in D$ und es gelte $|\lambda| < \beta_i$, $|\mu| < \gamma_i$. Dann ist (x, λ, μ) ein Kuhn-Tucker-Punkt von $[PN]$, oder es gilt:*

$$\frac{\partial}{\partial \varepsilon} T(x + \varepsilon \Delta x) \Big|_{\varepsilon > 0} < 0$$

Aus dieser Aussage lässt sich nun der folgende Satz herleiten.

Satz 2.1.1 (Globale Konvergenz des VGN - Verfahrens). *Sei $x_0 \in D$ beliebig, und die zugehörige Niveaumenge D_0 sei kompakt mit*

$$D_0 = \{x | T_1(x) \leq T_1(x_0)\} \subset D$$

Für alle $x \in D_0$ gelte für den Kuhn-Tucker-Punkt $(\Delta x, \lambda, \mu)$ des linearisierten Problems $[PL]$

$$|\lambda| < \beta_i \text{ und } |\mu| < \gamma_i.$$

Dann erzeugt obiger Algorithmus eine Folge von Dämpfungsfaktoren t_k , so dass die Iteration

$$y_{k+1} = x_k + t_k \Delta x_k \text{ mit Startwert } x_0$$

einen Kuhn-Tucker-Punkt des nichtlinearen Problems $[PN]$ als Häufungspunkt besitzt.

Bemerkung: *Schränkt man obige Aussage auf Häufungspunkte der Folge (t_k) ein, kann sogar auf Kompaktheit verzichtet werden. Numerisch gesehen ist dies nicht förderlich. Besser ist es die Kompaktheit und Existenz entsprechender Schranken β_i und γ_i sicherzustellen.*

Wie allerdings schon bei einem guten Startwert, ist auch hier das a priori Schätzen von guten Schranken nur bedingt möglich. Abhilfe kann hier eine Schätzung dieser Werte aus den adjungierten Variablen der Iteration schaffen.

Mit Hilfe des Satzes (2.1.1) wurde nachgewiesen, dass die Niveaufunktion T_1 aus 5 die theoretische Eigenschaft eines gedämpften VGN-Verfahrens erfüllt. Leider aber ist diese Niveaufunktion niedriger Ordnung als das VGN-Verfahren selber. Dadurch kann bei schwierigen Problemen die Konvergenzeigenschaft sogar verschlechtert werden.

Die Konvergenz des relaxierten Verfahrens kann in manchen Fällen erheblich verlangsamt werden. Um diesem Verhalten entgegenzuwirken, ist es möglich, die Suchrichtung zu verbessern oder eine andere Niveaufunktion zu wählen.

2.2 Natürliche Niveaufunktion

Auf der Suche nach Niveaufunktionen ist man natürlich bestrebt jene von höherer Ordnung zu finden, damit oben angesprochene Nachteile vermieden werden können. Man betrachte daher Funktionen, die den Abstand von einem Lösungspunkt x_0 zur Iterierten messen.

Lemma 2.2.1 ("Ideale Niveaufunktion"). *Sei $\Omega(x)$ definiert durch*

$$\Omega(x) := \|J(x^*)^+ F_1(x)\|_2^2 \quad (6)$$

so gilt

$$\Omega(x) = \|x - x^*\|_2^2 + O(\|x - x^*\|_2^3) \quad (7)$$

Weiterhin lässt sich zeigen, dass für ein $\kappa(x^*) < 1$ (bekannt aus Folgerungen aus dem Vorredner-Vortrag) Ω in einer Umgebung von x^* die Niveaufunktion die Verträglichkeitsbedingung für ein VGN-Verfahren erfüllt. Das heißt, dass diese Niveaufunktion damit also die Konvergenzbedingungen erfüllt.

Etwas genauer bedeutet dies, folgende Aussagen können verifiziert werden:

$$\begin{aligned} J(x)^+ F(x) &= x - x^* + B^{-1} P E(x - x^*) + O(\|x - x^*\|^2) \\ J(x^*)^+ F(x) &= x - x^* + O(\|x - x^*\|^2) \end{aligned}$$

und damit auch

$$\begin{aligned} -\frac{1}{2} \nabla \Omega(x) J(x)^+ F(x) &= -\|x - x^*\|^2 - (x - x^*)^t B^{-1} P E(x - x^*) + O(\|x - x^*\|^3) \\ &\leq -(1 - \kappa) \|x - x^*\|^2 + O(\|x - x^*\|^3) \end{aligned}$$

Auch hier zeigt sich wieder der theoretisch gute Ansatz. Aber wie schon in den vorangegangenen Problemen ist x^* a priori nicht bekannt. Abhilfe kann hier die Einführung von iterativ gewichteten Niveaufunktionen schaffen.

$$T^k(y) := \|J(x_k)^+ F(y)\|_2^2 \quad (8)$$

Bei auf diese Weise definierten Niveaufunktionen spricht man von **lokal definierten Niveaufunktionen**. Um die Eigenschaften solcher Niveaufunktionen nachzuweisen, betrachtet man zunächst die Abstiegseigenschaften allgemein gewichteter Niveaufunktionen des Typs

$$T^A(y) := \|AF(y)\|_2^2 \quad (9)$$

Lemma 2.2.2 (Abstiegslemma). *Sei $\Delta x \neq 0$ und $AF(x) \neq 0$. Sei weiterhin*

$$w(x, A, t) := \sup_{s \in (0, t]} \frac{\|A(J(x + s\Delta x) - J(x))\Delta x\|}{s \|AF(x)\|^2} \leq \hat{w}(A) < \infty \quad (10)$$

und

$$\nu(x, A) := \frac{\|AR(x)\|}{\|AF(x)\|} \geq 0 \quad (11)$$

so gilt die Abstiegsrichtung

$$T^A(x + t\Delta x) \leq \left\{ 1 - t(1 - \nu(x, A)) + \frac{t^2}{2} w(x, A, t) \frac{\|\Delta x\|^2}{\|AF(x)\|} \right\}^2 T^A(x) \quad (12)$$

für alle $x \in D$, $t \in [0, 1]$ mit $x + s\Delta x \in D$ und $s \in [0, t]$.

Bemerkung: *Es lässt sich außerdem leicht zeigen, dass obige Niveaufunktionen für $\nu(x, A) < 1$ die Verträglichkeitsbedingungen erfüllen.*

2.3 Schrittweisenstrategien für natürliche Niveaufunktionen

Von den Niveaufunktionen, welche obige Eigenschaften erfüllen, sind solche der Form $T^{J(x)^+}$ diejenigen mit dem besten linearen Abstieg.

Lemma 2.3.1 (Minimaleigenschaft natürlicher Niveaufunktionen). *Es gelten folgende Eigenschaften*

$$\min_A \nu(x, A) = \nu(x, J(x)^+) = 0 \quad (13)$$

$$T^{J(x)^+}(x + \Delta x) \leq \left(1 - t + \frac{t^2}{2} w(x, J(x)^+, t) \|\Delta x\| \right)^2 T^{J(x)^+} \quad (14)$$

$$\left. \frac{\partial}{\partial \varepsilon} T^{J(x)^+}(x + \varepsilon \Delta x) \right|_{\varepsilon > 0} = -2 \|J(x)^+ F(x)\|_2^2 \quad (15)$$

Mit Hilfe des Lemmas hat man also die Minimaleigenschaft von Niveaufunktionen der Form $T^{J(x)^+}$ gezeigt. Somit lassen sich nun auch Schrittweisenstrategien für natürliche Niveaufunktionen herleiten, welche den Aufwand der eindimensionalen Minimumsuche umgehen.

Sei x_k eine Iterierte und $D_k = \{x | T^k(x) \leq T^k(x_k)\} \subset D$, für w aus Lemma (2.2.2) gilt dann auch

$$h_k(t) := w(x_k, J_k^+, t) \|\Delta x_k\|$$

Aus weiteren Abschätzungen erhält man schließlich $x_k + t\Delta x_k \in D$. Dabei existieren Grenzschriftweiten

$$t_k^*(\mu) = \min\left(1, \frac{\mu}{h_k(t_k^*(\mu))}\right) \leq 1$$

mit $0 < \mu < 2$. Als *optimale Relaxationsstrategie* wird dabei die Wahl einer Schrittweite $t^*(1)$ bezeichnet. Wählt man nun t_k derart, dass

$$t_k \in I_k := [t_k^*(\mu_1), t_k^*(\mu_2)]$$

mit $0 < \mu_1 \leq \mu_2 < 2$. So geht das gedämpfte Verfahren in ein Vollschriftverfahren über.

2.4 Praktische Realisierung

In diesem Kapitel wird am konkreten Fall eine Strategie mit Vorausschätzungen und Korrekturschritten betrachtet. Damit eine solche Strategie auch von Vorteil ist, sollte eine gute Schätzung von w aus Lemma (2.2.2) möglich sein. Hierfür wird eine *a priori Schätzformel* eingeführt.

$$\bar{w}(x_k, A, t) = 2 \frac{\|A(F(x_k + t\Delta x_k) - F(x_k) - tJ_k\Delta x_k)\|}{\|tJ_k^+ F(x_k)\|^2} \quad (16)$$

$$\bar{w}(x_k, J_k^+, t) = 2 \frac{\|\overline{\Delta x_{k+1}}(t) - (1-t)\Delta x_k\|}{\|t\Delta x_k\|^2} \quad (17)$$

wobei

$$\overline{\Delta x_{k+1}}(t) := -J_k^+ F(x_k + t\Delta x_k)$$

Für diese Schätzformel lässt sich die *asymptotische Korrektheit* zeigen.

Lemma 2.4.1 (Asymptotische Korrektheit).

$$\bar{w}(x_k, A, t) = w(x_k, A, t) + O(t) \leq w(x_k, A, t) \quad (18)$$

Lemma 2.4.2 (Verschärfte Abstiegsbedingung). Sei $0 < t \leq 1$ und $t \leq \frac{\mu}{\bar{w}(x_k, J_k^+, t)\|\Delta x\|}$ dann gilt

$$\bar{w}(x_k, A, t) = w(x_k, A, t) + O(t) \leq w(x_k, A, t) \quad (19)$$

2.5 Ein Algorithmus zur Schrittweitensteuerung

Aus den bisher sehr theoretischen Erkenntnissen lässt sich nun ein Algorithmus ableiten, mit dem wir unsere Schrittweite steuern können.

(1) Bestimme Vorschätzung (*Prädiktor*)

$$\mu_k^{(0)} := \min\left(1, \frac{\mu_0}{\overline{w}(x_{k-1}, J_{k-1}^+, t_{k-1}) \|\Delta x\|}\right)$$

(2) $j = 0$

$$(3) \quad t_k^{(j)} := \begin{cases} 1 & \tau < \mu_k^j \\ \mu_k^j & \tau_{min} \leq \mu_k^j \leq \tau \\ \tau_{min} & \mu_k^j < \tau_{min} \end{cases}$$

wobei $0 < \tau_{min}$ und $\tau \in [0.5, 1]$ beliebig

(4.1) wenn gilt $t_k^j \overline{w}(x_k, J_k^+, t_k^j) \|\Delta x\| \leq \mu_2$ setze

$t_k = t_k^j$ und $k = k + 1$ und gehe zu (1)

(4.2) sonst $j = j + 1$ und bestimme einen a posteriori Schätzer (*Korrektor*)

$$\mu_k^j = \frac{\mu_0}{\overline{w}(x_k, J_k^+, t_k^{j-1}) \|\Delta x\|}$$

und gehe zu (3)

Zu beachten ist an der Stelle, dass für den ersten Iterationsschritt kein μ_0^0 existiert, da wir noch keine aus einer vorigen Iteration bekannten Werte x_{-1} , J_{-1}^+ oder t_{-1} kennen. In diesem Fall setze man

$$\mu_0^0 = \min\left(1, \frac{\mu_0}{\overline{w}_0 \|\Delta x_0\|}\right)$$

mit z.B. $\overline{w}_0 = 100$.

Bemerkung: Aus Gründen der numerischen Verfahren muss hier eine Unterschranke τ_{min} gewählt werden, die jedoch recht klein sein sollte (z.B. $\tau_{min} = 0.001$). Wird selbst für solche kleinen Werte kein Abstieg erreicht, dann wird das Problem i.a. als singulär betrachtet.

3 Ein Fortsetzungsverfahren für beschränkte Ausgleichsprobleme

Wie im vorangegangenen Kapitel gezeigt, stellt eine Schrittweisensteuerung mit Schätzverfahren eine gute Möglichkeit dar, globale Konvergenz zu erreichen. Trotz dieser guten Eigenschaften ist es bei manchen schwierigen Problemen sinnvoll ein *Fortsetzungs-* oder *Homotopieverfahren* zu verwenden.

Definition 3.0.1. Ein *Eingebettetes beschränktes Ausgleichsproblem* $PN(t)$ ist wie folgt definiert:

$$\begin{aligned} \min_x \quad & \|H_1(x, t)\|_2^2 \quad t \in [0, 1] \\ \text{s.t.} \quad & H_2(x, t) = 0 \quad H_i \in C^3(\bar{D}), \bar{D} \subset \mathbb{R}^{m+1} \\ & H_3(x, t) \geq 0 \end{aligned}$$

Setzen wir $t = 1$, so soll dieses Problem gerade unserem Ausgangsproblem entsprechen. Für $t = 0$ sei eine Lösung x_0 bekannt.

Existiert ein *Homotopieweg*

$$x : I \rightarrow \mathbb{R}^m \quad x(0) = x_0$$

$$x(t) \quad \text{Lösung von } [PN(t)]$$

so zerlege man das Intervall I geeignet in Teilintervalle $[t_j, t_{j+1}]$ mit

$$\begin{aligned} 0 = t_1 < t_2 < \dots < t_l = 1 \quad \Delta t_j = t_{j+1} - t_j \\ j = 1 \dots l - 1 \end{aligned}$$

und löse sukzessiv die Teilprobleme $[PN(t_j)]$.

3.1 Wahl des Iterationsverfahrens

Die Durchführbarkeit und die Effizienz eines oben beschriebenen Verfahrens hängt u.a. von der richtigen Wahl des Iterationsverfahrens ab. Interessant ist dabei eine Variante VGN-Verfahren, bei der die aktiven Ungleichungsbeschränkungen über eine Funktion, welche vom Homotopieparameter t abhängt, bestimmt werden.

Auf diese Weise reduziert sich das Fortsetzungsverfahren auf Gleichungsbeschränkung.

$$\begin{aligned} \min_x \quad & \|H_1(x, t)\|_2^2 \quad t \in I = [t^1, t^2] \\ \text{s.t.} \quad & H_C(x, t) = 0 \end{aligned}$$

Dabei sind t^1 und t^2 Grenzpunkte, auf deren Bestimmung hier nicht weiter eingegangen wird.

3.2 Mögliche Einbettungen

Im Folgenden wird angenommen, dass ein zuverlässiger Parameter x_0 bekannt sei. Zuverlässig heißt in dem Zusammenhang, dass das Residuum im Punkt x_0 verschwindet. Weiter wird angenommen, der Wert x_0 sei keine Lösung des Problems $[PN]$. Nun gibt es zwei Alternativen für Einbettungen:

$$(1) \quad H_1(x, t) = \begin{pmatrix} tF_1(x) \\ (1-t)A(x-x_0) \end{pmatrix} \quad t \in [0, 1], \quad A \in L(m, m)$$

$$(2) \quad H_1(x, t) = F_1(x) - (1-t)F_1(x_0) \quad t \in [0, 1]$$

Der Ansatz (1) empfiehlt sich bei Ausgleichsproblemen, wohingegen der Ansatz (2) sich auch bei in x_0 verletzten Beschränkungen anbietet.

Bemerkung: Einen solchen Parameter x_0 könnte z.B. schon aus früheren Versuchen bekannt sein.

3.3 Wahl des Prädiktors

Definition 3.3.1. Eine Abbildung $P : I \rightarrow \mathbb{R}^m$ heißt **Prädiktor mit Wachstumsfunktion** W , wenn für $\mu > 0$ gilt:

$$\|P(\hat{t} + h) - x(\hat{t} + h)\| \leq \mu W(h) \quad , \hat{t} \leq \hat{t} + h \in I$$

Als *Wachstumsfunktion* kann hier jede beliebige monoton wachsende Funktion $W : \mathbb{R} \rightarrow \mathbb{R}$ mit $W(0) = 0$ genommen werden.

Definition 3.3.2. Sei x in $C^{k+1}(I)$, $j \geq k+1$ und für $t_1 < t_2 < \dots < t_j$ sei die Lösung des eingebetteten Problems bereits berechnet, dann ist das **Interpolationspolynom** p^k zu $\{(t_i, x(t_i)), j-k \leq i \leq j\}$ ein Prädiktor mit der Wachstumsfunktion

$$W^k(h) = \prod_{i=j-k}^j (h + \sum_{l=1}^{j-1} \Delta t_l)$$

Mit solchen Polynomen kann eine große Genauigkeit der Startwerte erreicht werden. Dies ist von Vorteil, da dadurch eine wesentliche Steigerung der Effizienz des Homotopieverfahrens erreicht werden kann.

Bemerkung: Anzumerken ist an dieser Stelle, dass die zu den Interpolationspolynomen p^k gehörige Wachstumsfunktionen $W^k(h)$ folgendes Ordnungskriterium erfüllen:

$$W^k(h) = O(\bar{h}^{k+1}) \quad \bar{h} = \max(h, \Delta t_{j-1}, \dots, \Delta t_{j-k})$$

Wie man sieht ist also die Wahl der oben definierten Prädiktoren auch in Bezug auf dessen Ordnung vorteilhaft.

3.4 Wahl der Schrittweite

In diesem Abschnitt wird darauf eingegangen, wie die maximalen Schrittweiten aussehen, bei denen noch lokale Konvergenz gesichert werden kann. Dieser Abschnitt ist eher theoretischer Natur, da solche Schrittweiten nur die Konvergenz sichern, aber keine konstruktiv guten Schrittweiten erzeugen.

Für den folgenden Satz werden verschiedene Voraussetzungen benötigt.

- (1) Für ein Gebiet \bar{D} gelte

$$H = \begin{pmatrix} H_1 \\ H_C \end{pmatrix} \in C^2(\bar{D})$$

mit $\{(x, t), t \in I, x \in D(t)\} \subset \bar{D}$

- (2) $x(t)$ sei eindeutige Lösung von $[PN(t)]$

- (3) Für alle $(x, t) \in \bar{D}$ besitze die Ableitung

$$J(x, t) = \frac{\partial}{\partial x} H(x, t)$$

eine verallgemeinerte Inverse $J(x, t)^+$

Satz 3.4.1 (Schrittweitensteuerung für ein VGN-Fortsetzungsverfahren).

(1) Für alle $t \in I$, $s \in [0, 1]$, $x, y, z \in D(t)$ sei (a) und (b) erfüllt

(a) $\|J(y, t)^+(J(x + s(y - x), t) - J(x, t))(y - x)\| \leq \omega(t)s \|y - x\|^2$

mit $\omega(t) \leq \omega < \infty$

(b) $\|(J(z, t)^+J(x, t)^+)R(x, t)\| \leq \kappa(t)s \|z - x\|^2$

mit $\kappa(t) \leq \kappa < 1$ und $R(x, t) := (I - J(x, t)J(x, t)^+)H(x, t)$

(2) Für beliebiges $t \in I$ sei ein Prädiktor P_t mit Wachstumsfunktion W in t und $\mu > 0$ gegeben.

(3) Für $\kappa < \delta < 1$, $\gamma := \kappa + 1$ sei

$$h_{max} := \min\left(W^{-1} \frac{\gamma(\sqrt{1 + \frac{4(\delta - \kappa)}{\gamma^2}} - 1)}{\omega\mu}, 1 - t\right)$$

(4) Für alle $h \in [0, h_{max}]$, $\hat{t} = t + h$, $\hat{x} = P_t(\hat{t})$ gelte

$$\hat{\kappa} := \bar{\kappa}(\hat{x}, \frac{\hat{a}}{1 - \delta}) \subset D(\hat{t})$$

mit $\hat{a} := \|J(\hat{x}, \hat{t})^+H(\hat{x}, \hat{t})\|$

Dann gilt für alle $\hat{t} \in [t, t + h_{max}]$

(1) Das Vollschrirt-VGN-Verfahren mit Startwert $x_0 = P_t(\hat{t})$ ist

(a) wohldefiniert

(b) bleibt in $\hat{\kappa}$ und

(c) konvergiert gegen $x(\hat{t})$

(2) Außerdem gelten folgende Abschätzungen

(a) $\|x_0 - x(\hat{t})\| \leq \frac{\gamma(\sqrt{1 + \frac{4(\delta - \kappa)}{\gamma^2}} - 1)}{\omega}$

(b) $\alpha_0 = \|J(x_0, \hat{t})^+H(x_0, \hat{t})\| \leq 2 \frac{(\delta - \kappa)}{\omega}$

(c) $\|x_k - x(\hat{t})\| \leq \alpha_0 \frac{\delta^k}{1 - \delta}$

3.5 Automatische Parametrisierung

Je nach Art des Problems kann es manchmal von Vorteil sein zwischen verschiedenen Iterationsschritten einen Wechsel der Parametrisierung vorzunehmen.

Sei eine beliebige Parametrisierung folgenden Pfades gegeben:

$$(x, t) : [0, 1] \rightarrow \mathbb{R}^{m+1} \quad t(0) = 0, t(1) = 1$$

$$z(s) := (x(s), t(s)) \text{ löse } [PN(t(s))]$$

dann empfiehlt sich als lokaler Iterator $z_{k+1} = z_k + \Delta z_k$, bei dem Δz_k folgendermaßen berechnet werden kann:

$$\min_{(\Delta x, \Delta t)} \|A^k \Delta x + a^k \Delta t + b^k\|_2 \quad (A^k a^k) \in L(m+1, n_a)$$

$$s.t. \quad (\Delta x, \Delta t) \text{ ist Lösung von}$$

$$\min_{\Delta x} \|J_1^k \Delta x + j_1^k \Delta t + H_1^k\|_2$$

$$s.t. \quad J_c^k \Delta x + j_c^k \Delta t + H_c^k = 0$$

wobei

$$(J_1^k, j_1^k) := \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial t}\right)H_1, \text{ mit } H_1^k := H_1(z_k)$$

$$A^k := A(z_k, s_j), a^k := a(z_k, s_j), b^k := b((z_k, s_j)),$$

Die Lösung dieses mehrstufigen Minimierungsproblems ist dann das gesuchte Δz_k .

Literatur

- [De] P. DEUFELHARD; *Newton Methods for Nonlinear Problems*; Springer; 2004
- [DB] P. DEUFELHARD, F. BORNEMANN; *Numerische Mathematik II, Gewöhnlicher Differentialgleichungen*; de Gruyter Lehrbuch; 2. Auflage; 2002
- [DBa] P. DEUFELHARD, F. BORNEMANN; *Numerische Mathematik II, Gewöhnlicher Differentialgleichungen*; de Gruyter Lehrbuch; 1. Auflage
- [1] R. RANNACHER; *Numerische Methoden für Gewöhnliche Differentialgleichungen, Numerische Mathematik 1*; Vorlesungsskriptum; Universität Heidelberg; WS 01/02
- [St] J. STARKE; *Gewöhnliche Differentialgleichungen und Dynamische Systeme*; Vorlesungsskriptum; Universität Heidelberg; WS 04/05
- [La] C. LAUX; *Gewöhnliche Differentialgleichungen, Randwertaufgaben*; Seminarvortrag; Universität Heidelberg; 2005
- [LaT] T. LAGEMANN; *Parameter-abhängige Systeme: Fortsetzung von Homotopiepfaden Globalisierung lokaler Newton Methoden*; Seminarvortrag; Universität Heidelberg; 2005
- [2] I.N.BRONSTEIN, K.A.SEMENDJAJEW, G.MUSIOL, H.MÜHLIG; *Taschenbuch der Mathematik*; Harri Deutsch; 5.Auflage; 2001
- [LM] Lexikon der Mathematik in 6 Bänden; Spektrum Akademischer Verlag; Berlin 2001