

Anwendung affin invarianter Konvergenzsätze auf newtoneske Verfahren

im Rahmen des Seminars „NEWTON- und NEWTON-ähnliche Verfahren“
unter Leitung von Dr. Ekaterina Kostina und Dr. Moritz Diehl,
Wintersemester 2004/05, Universität Heidelberg

Hans Joachim Ferreau
ferreau@urz.uni-heidelberg.de

03.11.2004

Inhaltsverzeichnis

1	Residuen-basierte Algorithmen	2
1.1	Einleitung	2
1.2	Gewöhnliches NEWTON-Verfahren	2
1.3	Vereinfachtes NEWTON-Verfahren	4
1.4	Quasi-NEWTON-Verfahren	5
1.5	Inexaktes NEWTON-Verfahren	8
1.6	Numerische Tests	10
2	Konvexe Optimierung	13
2.1	Einleitung	13
2.2	Gewöhnliches NEWTON-Verfahren	13
2.3	Vereinfachtes NEWTON-Verfahren	15
2.4	Inexaktes NEWTON-PCG-Verfahren	17
2.5	Numerische Tests	20
	Schlussbemerkungen	23
	Literatur	24

1 Residuen-basierte Algorithmen

1.1 Einleitung

Eine Gleichung der Form

$$F(x) = 0$$

kann mit dem gewöhnlichen NEWTON-Verfahren

$$F'(x^k)\Delta x^k = -F(x^k), \quad x^{k+1} = x^k + \Delta x^k \quad \forall k = 0, 1, \dots \quad (1.1)$$

iterativ gelöst werden. Für ein Problem der Form

$$G(y) := F(By) = 0, \quad x = By, \quad B \in \text{GL}(n, \mathbb{R})$$

gilt

$$G'(y) = F'(By)B.$$

Damit transformiert sich (1.1) wie folgt:

$$G'(y^k)\Delta y^k = -G(y^k) \iff F'(By^k)B\Delta y^k = -F(By^k) \iff F'(x^k)\underbrace{B\Delta y^k}_{=\Delta x^k} = -F(x^k)$$

Man sieht, dass die Korrekturen exakt so wie der gesamte Definitionsbereich transformiert werden. Dies gilt auch für die Iterierten selbst, falls man entsprechend den Startwert $y^0 := B^{-1}x^0$ verwendet. Diese Eigenschaft bezeichnet man als *affine Kontravarianz*.

Dagegen ändern sich die Residuen $F(x^k)$ bzw. ihre Norm $\|F(x^k)\|$ wegen

$$G(y^k) := F(By^k) = F(x^k)$$

nicht. Ziel dieses Kapitels ist daher die Herleitung von Residuen-basierten Konvergenzkriterien für verschiedene newtoneske^a Verfahren, da diese unter affin kontravarianten Transformationen invariant bleiben.

1.2 Gewöhnliches Newton-Verfahren

Um die angestrebten Konvergenzkriterien zu erhalten, nutzen wir folgende affin kontravariante Version des Satzes von NEWTON-MYSOVSKIKH:

Satz 1.1 (Newton-Mysovskikh, affin kontravariante Version)

Sei $F : \mathcal{D} \rightarrow \mathbb{R}^n$ eine differenzierbare Abbildung, $\mathcal{D} \subseteq \mathbb{R}^n$ offen und konvex, und gelten folgende Bedingungen:

1. $F'(x)$ ist für alle $x \in \mathcal{D}$ invertierbar.
2. Es gilt die affin kontravariante LIPSCHITZ-Bedingung:

$$\exists \omega, 0 < \omega < \infty : \quad \|(F'(y) - F'(x))(y - x)\| \leq \omega \|F'(x)(y - x)\|^2 \quad \forall x, y \in \mathcal{D}.$$

3. Durch

$$\mathcal{L}_\omega := \left\{ x \in \mathcal{D} \mid \|F(x)\| < \frac{2}{\omega} \right\}$$

ist eine offene Menge definiert und $\overline{\mathcal{L}_\omega} \subset \mathcal{D}$ ist beschränkt.

4. Für den Startwert x^0 gilt

$$h_0 := \omega \|F(x^0)\| < 2, \quad (1.2)$$

d. h. $x^0 \in \mathcal{L}_\omega$.

^aSammelbegriff für NEWTON- und NEWTON-ähnliche Verfahren

Dann gilt:

1. Die Iterierten $\{x_k\}$ des gewöhnlichen NEWTON-Verfahrens (1.1) bleiben in \mathcal{L}_ω und konvergieren gegen eine Lösung $x^* \in \mathcal{L}_\omega$ mit $F(x^*) = 0$.
2. Die Residuen $\{F(x^k)\}$ konvergieren gegen 0 mit der Konvergenzrate

$$\|F(x^{k+1})\| \leq \frac{\omega}{2} \|F(x^k)\|^2. \quad (1.3)$$

Bemerkung 1.2

Eine Variante von Satz 1.1 mit schwächeren Differenzierbarkeits-Bedingungen findet man in [8].

Wir definieren nun die Kontraktionsfaktoren

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

und die sog. KANTOROWITSCH-Werte

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

Dann besagt (1.3)

$$h_{k+1} \leq \frac{1}{2} h_k^2 = \left(\frac{1}{2} h_k\right) h_k,$$

woraus man rekursiv die Ungleichungen

$$h_{k+1} < h_k < \dots < h_0 < 2$$

erhält. Dies führt zu

$$\Theta_k = \frac{h_{k+1}}{h_k} \leq \frac{1}{2} h_k < 1$$

und liefert speziell für $k = 0$:

$$\Theta_0 \leq \frac{1}{2} h_0 < \frac{1}{2} \cdot 2 = 1.$$

Damit erhalten wir das

Startwert-Kriterium: Falls

$$\Theta_0 \geq 1$$

gilt, weisen wir den Startwert x^0 als *nicht hinreichend nahe* zurück.

Da wir ω nicht kennen, können die KANTOROWITSCH-Werte nicht exakt berechnet werden. Wir bedienen uns daher der verfügbaren a-posteriori-Abschätzung^b

$$[h_{k+1}]_1 := 2\Theta_{k+1} \leq h_{k+1}.$$

$h_{k+1} = \Theta_k h_k$ liefert außerdem die a-priori-Abschätzung

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

Nimmt man nun näherungsweise $[h_{k+1}]_1 \approx [h_{k+1}]$ an, erhält man

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

^bIm Folgenden werden a-posteriori-Abschätzung einer Größe α mit $[\alpha]_1$ bzw. $[\alpha]_2$ bezeichnet. A-priori-Abschätzungen mit $[\alpha]$.

Dies liefert uns

Divergenz-Kriterium 1: Wir brechen die gewöhnliche NEWTON-Iteration als *divergent* ab, falls

$$\Theta_{k+1} > \Theta_0.$$

Die zusätzliche Annahme

$$0 \leq \frac{h_k - [h_k]}{[h_k]} < 1$$

führt zu

$$\Theta_{k+1} \leq \frac{1}{2}h_{k+1} < \frac{1}{2} \cdot 2[h_{k+1}] = 2\Theta_k^2$$

und liefert das strikere

Divergenz-Kriterium 2: Wir brechen die gewöhnliche NEWTON-Iteration als *divergent* ab, falls

$$\Theta_{k+1} \geq 2\Theta_k^2.$$

Sei *FTOL* die vom Benutzer definierte *Residuum-Fehlertoleranz*.

Konvergenz-Kriterium: Wir beenden die gewöhnliche NEWTON-Iteration, falls

$$\|F(x^k)\| \leq FTOL$$

erfüllt ist und geben x^k als Lösung aus.

1.3 Vereinfachtes Newton-Verfahren

Beim vereinfachten NEWTON-Verfahren wird die JACOBI-Matrix im Startpunkt x^0 die gesamte Iteration hindurch als Näherung für $F'(x^k)$ benutzt:

$$F'(x^0)\overline{\Delta x}^k = -F(x^k), \quad x^{k+1} = x^k + \overline{\Delta x}^k \quad \forall k = 0, 1, \dots \quad (1.4)$$

Es gilt folgender

Satz 1.3 (Newton-Kantorowitsch, affin kontravariante Version)

Sei $F : \mathcal{D} \rightarrow \mathbb{R}^n$ eine stetig-differenzierbare Abbildung, $\mathcal{D} \subseteq \mathbb{R}^n$ konvex, und gelten folgende Bedingungen:

1. Es existiert ein $0 \leq \omega < \infty$ so, dass

$$\|(F'(x) - F'(x^0))v\| \leq \omega \|F'(x^0)(x - x^0)\| \|F'(x^0)v\| \quad \forall x \in \mathcal{D}, v \in \mathbb{R}^n.$$

2. Der Abschluss $\overline{\mathcal{L}_\omega} \subseteq \mathcal{D}$ von

$$\mathcal{L}_\omega := \left\{ x \in \mathbb{R}^n \mid \|F(x)\| \leq \frac{1}{2\omega} \right\}$$

ist beschränkt.

3. Für den Startwert x^0 gilt

$$h_0 := \omega \|F(x^0)\| \leq \frac{1}{2},$$

d. h. $x^0 \in \mathcal{L}_\omega$.

Dann gilt:

1. Die Iterierten $\{x_k\}$ des vereinfachten NEWTON-Verfahrens (1.4) bleiben in \mathcal{L}_ω und konvergieren gegen eine Lösung $x^* \in \mathcal{L}_\omega$ mit $F(x^*) = 0$.

2. Die Norm der Residuen konvergiert mit der Rate

$$\frac{\|F(x^{k+1})\|}{\|F(x^k)\|} \leq \frac{1}{2}(t_k + t_{k+1}) < 1 - \sqrt{1 - 2h_0} \quad (1.5)$$

gegen 0, wobei die $\{t_k\}$ durch

$$t_0 := 0, \quad t_{k+1} = h_0 + \frac{1}{2}t_k^2$$

definiert sind.

Man kann also i. Allg. nur noch *lineare* Konvergenz erwarten. Drückt man (1.5) durch die oben definierten Kontraktionsfaktoren Θ_k aus, erhält man im Falle $k = 0$:

$$\Theta_0 \leq \frac{1}{2}(t_0 + t_1) = \frac{1}{2}(0 + h_0) \leq \frac{1}{4}.$$

Dies führt zu folgendem

Startwert-Kriterium: Falls

$$\Theta_0 > \frac{1}{4}$$

gilt, weisen wir den Startwert x^0 als *nicht hinreichend nahe* zurück.

Diese Bedingung ist – wie zu erwarten – deutlich stärker als das Startwert-Kriterium des gewöhnlichen NEWTON-Verfahrens. Den geringeren Rechenaufwand innerhalb *einer* Iteration muss man sich beim vereinfachten NEWTON-Verfahren also u. a. durch einen besseren Startwert erkaufen.

1.4 Quasi-Newton-Verfahren

BROYDEN stellte 1965 verschiedene Quasi-NEWTON-Methoden vor (siehe [1]):

$$J_k \delta x^k = -F(x^k), \quad x^{k+1} = x^k + \delta x^k \quad \forall k = 0, 1, \dots \quad (1.6)$$

Dabei wird die JACOBI-Matrix im Punkt x_k durch eine Näherung J_k approximiert. Dabei wird die jeweils nachfolgende Näherung J_{k+1} mit Hilfe eines Rang-1-Updates aus J_k berechnet. Eine spezielle Variante, die BROYDEN auf Grund praktischer Tests als „schlecht“ bezeichnete, spiegelt affine Kontravarianz wider, so dass wir sie nun genauer untersuchen wollen. Ihr Verhalten innerhalb einer Iteration beschreibt

Satz 1.4

Sei

$$J_{k+1}^{-1} = J_k^{-1} \left(Id - \frac{F(x^{k+1})(\delta F_{k+1})^T}{\|\delta F_{k+1}\|^2} \right) \quad (1.7)$$

das affin kontravariante „schlechte“ BROYDEN-Rang-1-Update und gelte die Residuen-Kontraktion

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

Dann gilt:

1. Die Update-Matrix J_{k+1} erfüllt

$$\begin{aligned} \|E_k(J_{k+1})\| &\leq \|E_k(J)\| \quad \forall J : E_k(J)\delta F_{k+1} = F_{k+1} \quad \text{und} \\ \|E_k(J_{k+1})\| &\leq \frac{\Theta_k}{1 - \Theta_k}, \end{aligned}$$

wobei

$$E_k(J) := Id - J_k J^{-1}.$$

2. Die Update-Matrix J_{k+1} ist regulär, wenn J_k regulär ist und besitzt die Darstellung

$$J_{k+1} = \left(Id - \frac{F(x^{x+1})(\delta F_{k+1})^T}{(\delta F_{k+1})^T F(x^k)} \right) J_k.$$

3. Die nächste Quasi-NEWTON-Korrektur ist

$$\delta x^{k+1} = -J_{k+1}^{-1} F(x^{k+1}) = \left(1 - \frac{(\delta F_{k+1})^T F(x^{x+1})}{\|\delta F_{k+1}\|^2} \right) (-J_k^{-1} F(x^{k+1})).$$

Mit Hilfe dieses Satzes können wir das Verhalten des „schlechten“ BROYDEN-Rang-1-Updates im Laufe der gesamten Iteration beschreiben:

Satz 1.5

Sei $F : \mathcal{D} \rightarrow \mathbb{R}^n$ eine stetig-differenzierbare Abbildung, $\mathcal{D} \subseteq \mathbb{R}^n$ konvex, und x^* der eindeutige Lösungspunkt von F mit $F'(x^*)$ regulär. Gelten ferner folgende Bedingungen:

1. Es existiert ein $\omega < \infty$ so, dass

$$\|(F'(x) - F'(x^*))(y - x)\| \leq \omega \|F'(x^*)(x - x^*)\| \|F'(x^*)(y - x)\| \quad \forall x, y \in \mathcal{D}.$$

2. Für die Quasi-NEWTON-Iteration aus Satz 1.4 gibt es ein $\bar{\Theta}$ mit $0 < \bar{\Theta} < 1$ so, dass

(a) die anfängliche Näherung J_0 der JACOBI-Matrix der Ungleichung

$$\bar{\eta}_0 := \|E_0\| < \bar{\Theta}$$

genügt (wobei $E_k := Id - F'(x^*)J_k^{-1}$) und

(b) der Startwert x^0 die Ungleichung

$$t_0 := \omega \|F'(x^*)(x^0 - x^*)\| \leq \frac{\bar{\Theta} - \bar{\eta}_0}{1 + \bar{\eta}_0 + \frac{4}{3}(1 - \bar{\Theta})^{-1}}$$

erfüllt.

Dann gilt:

1. Die Quasi-NEWTON-Iterierten $\{x^k\}$ konvergieren gegen x^* und zwar

(a) hinsichtlich des Fehlers mit

$$\|F'(x^*)(x^{k+1} - x^*)\| \leq \bar{\Theta} \|F'(x^*)(x^k - x^*)\|$$

(b) und bezüglich der Residuen mit

$$\|F(x^{k+1})\| \leq \bar{\Theta} \|F(x^k)\|.$$

2. Die Konvergenz ist superlinear mit

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

3. Die Abweichung der Näherungen der JACOBI-Matrix ist beschränkt durch

$$\|E_k\| \leq \bar{\eta}_0 + \frac{t_0}{(1 - t_0)(1 - \bar{\Theta})} \leq \bar{\Theta}$$

und es gilt die asymptotische Gleichung

$$\lim_{k \rightarrow \infty} \frac{\|E_k \delta F_{k+1}\|}{\|\delta F_{k+1}\|} = 0.$$

Die beiden vorangegangenen Sätze liefern die theoretische Grundlage für den QNRES-Algorithmus (**residual based quasi-Newton algorithm**). Wir geben ihn als C++-Fragment an. Dabei wurde auf verschiedene Optimierungen verzichtet, um den Algorithmus klar hervortreten zu lassen. Eine lauffähige Version steht zum Download bereit (siehe Seite 23).

```

// QNRES (Codefragment)

F[0] = func(x);           //  $F_0 := F(x^0) = func(x^0)$ 
sigma[0] = F[0]*F[0];    //  $\sigma_0 := \|F_0\|^2$ 

J0 = GetJacobiMatrix(x); //  $J_0 = F'(x_0)$ 
J0.Solve(dx,-F[0]);     // Löse  $J_0\delta x^0 = -F_0$ 

double kappa = 1;       //  $\kappa := 1$ 
int k = 0;

while (k < kmax);
{
  ++k;

  x = x + dx;           //  $x^{k+1} := x^k + \delta x^k$ 
  F[k+1] = func(x);    //  $F_{k+1} := F(x^{k+1})$ 
  dF[k+1] = F[k+1] - F[k]; //  $\delta F_{k+1} := F_{k+1} - F_k$ 

  sigma[k+1] = F[k+1]*F[k+1]; //  $\sigma_{k+1} := \|F_{k+1}\|^2$ 
  if (sigma[k+1] <= FTOL*FTOL) //  $\sigma_{k+1} \leq FTOL^2?$  (Konvergenz-Krit.)
    { status = solutionfound; break; } // Lösung gefunden!

  theta[k] = sqrt(sigma[k+1]/sigma[k]); //  $\Theta_k := \sqrt{\sigma_{k+1}/\sigma_k}$ 
  if (theta[k] >= thetamax) //  $\Theta_k \geq \Theta_{max}?$  (Divergenz-Kriterium)
    { status = noconvergence; break; } // Keine Konvergenz!

  kappa = kappa/(1-2*theta[k]); //  $\kappa := \kappa/(1 - 2\Theta_k)$ 
  if (kappa >= kappamax) //  $\kappa \geq \kappa_{max}?$ 
    { status = illconditioned; break; } // Schlecht konditioniertes Update!

  gamma[k] = dF[k+1]*dF[k+1]; //  $\gamma_k := \|\delta F_{k+1}\|^2$ 

  v = F[k+1]*(1 - dF[k+1]*F[k+1]/gamma[k]); //  $v := F_{k+1}(1 - \langle \delta F_{k+1}, F_{k+1} \rangle / \gamma_k)$ 

  double beta = 0;
  for (int j=k-1; j>=0; --j)
  {
    beta = dF[j+1]*v/gamma[j]; //  $\beta := \langle \delta F_{j+1}, v \rangle / \gamma_j$ 
    v = v - F[j+1]*beta; //  $v := v - \beta F_{j+1}$ 
  }

  J0.Solve(dx,-v); // Löse  $J_0\delta x^{k+1} = -v$ 
}

```

Man beachte, dass der Algorithmus im Wesentlichen nur die Residuen F_0, \dots, F_{k+1} und die Residuen-Differenzen $\delta F_1, \dots, \delta F_{k+1}$ speichern muss. Letztere können natürlich auch jedes Mal neu berechnet werden. Für Θ_{max} lässt sich ein sinnvoller Wert von $\approx \frac{1}{4}$ ableiten.

1.5 Inexaktes Newton-Verfahren

Beim gewöhnlichen NEWTON-Verfahren muss in jedem Schritt ein lineares Gleichungssystem gelöst werden. Wenn die Zahl der Unbekannten sehr groß ist, wächst der Rechenaufwand für das direkte Lösen schnell über ein sinnvolles Maß. Auch wenn die Iterierten noch relativ weit von der Lösung entfernt sind, ist das exakte Lösen des Gleichungssystems ggf. nicht gerechtfertigt. Dies führt zur Anwendung von inexakten NEWTON-Verfahren:

$$F'(x^k)\delta x^k = -F(x^k) + r^k, \text{ mit } \frac{\|r^k\|}{\|F(x^k)\|} \leq \eta_k, \quad x^{k+1} = x^k + \delta x^k \quad \forall k = 0, 1, \dots \quad (1.8)$$

Für $\eta_k \equiv 0$ erhält man das gewöhnliche NEWTON-Verfahren. Für $\eta_k \geq 1 \forall k$ wird diese Gleichung bereits durch $\delta_k \equiv 0$ erfüllt. Daher ist $\eta_k < 1 \forall k$ eine sinnvolle Voraussetzung und bereits 1982 wurde in [2] gezeigt, dass das inexakte NEWTON-Verfahren in diesem Falle lokal konvergiert.

Wir werden den iterativen Gleichungssystem-Löser GMRES (vgl. [3, S. 28] oder [11]) verwenden. Kritisch ist hierbei die Wahl des Parameters η_k : Sind die (äußeren) NEWTON-Iterierten noch weit von der Lösung entfernt, so ist es unnötig, das innere Gleichungssystem sehr genau zu lösen. Andererseits kann durch ein zu ungenaues Lösen des inneren Gleichungssystems die Konvergenz des (äußeren) NEWTON-Verfahrens verlangsamt werden. Um η_k so zu steuern, dass eine möglichst schnelle Konvergenz erreicht wird, benötigen wir folgenden

Satz 1.6

Sei $F : \mathcal{D} \rightarrow \mathbb{R}^n$ eine stetig-differenzierbare Abbildung, $\mathcal{D} \subseteq \mathbb{R}^n$ konvex, und gelten folgende Bedingungen:

1. Es existiert ein $0 \leq \omega < \infty$ so, dass

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

2. Die Menge

$$\mathcal{L}_0 := \left\{ x \in \mathbb{R}^n \mid \|F(x)\| \leq \|F(x^0)\| \right\} \subseteq \mathcal{D}$$

ist kompakt.

Dann gilt:

Für jede wohldefinierte Iterierte $x^k \in \mathcal{D}$ des inexakten NEWTON-Verfahrens (1.8) ist die (äußere) Residuen-Norm durch

$$\|F(x^{k+1})\| \leq \left(\eta_k + \frac{1}{2}(1 - \eta_k^2)h_k \right) \|F(x^k)\|$$

beschränkt, wobei $h_k := \omega \|F(x^k)\|$ und $\eta_k := \frac{\|r^k\|}{\|F(x^k)\|}$ definiert wird.

Die Konvergenzrate kann wie folgt abgeschätzt werden:

1. Linearer Konvergenzmodus:

Angenommen der Startwert erfüllt $h_0 < 2$, dann kann $\bar{\Theta}$ aus dem Bereich $\frac{h_0}{2} < \bar{\Theta} < 1$ gewählt werden. Wird dann die innere GMRES-Iteration durch

$$\eta_k \leq \bar{\Theta} - \frac{1}{2}h_k \quad (1.9)$$

kontrolliert, dann konvergieren die NEWTON-GMRES Iterierten $\{x^k\}$ mindestens linear gegen eine Lösung $x^* \in \mathcal{L}_0$ mit der Rate

$$\|F(x^{k+1})\| \leq \bar{\Theta} \|F(x^k)\|.$$

2. Quadratischer Konvergenzmodus:

Falls für ein $\rho > 0$ der Startwert

$$h_0 < \frac{2}{1 + \rho}$$

erfüllt und die innere GMRES-Iteration durch

$$\frac{\eta_k}{1 - \eta_k^2} \leq \frac{1}{2} \rho h_k \quad (1.10)$$

kontrolliert wird, dann ist die Konvergenz quadratisch mit der Rate

$$\|F(x^{k+1})\| \leq \frac{\omega}{2} (1 + \rho) (1 - \eta_k^2) \|F(x^k)\|^2. \quad (1.11)$$

Entsprechen Satz 1.6 werden wir während der Iteration

$$\Theta_k := \frac{\|F(x^{k+1})\|}{\|F(x^k)\|} \leq \bar{\Theta} < 1$$

verlangen:

Divergenz-Kriterium: Wir brechen die Quasi-NEWTON-Iteration als *divergent* ab, falls

$$\Theta_{k+1} > \bar{\Theta}.$$

Als nächsten wollen wir für den *quadratischen Konvergenzmodus* Schätzungen für die h_k einführen, da ω nicht bekannt ist. Aus der Ungleichung (1.11) erhält man

$$\|F(x^{k+1})\| \leq \frac{1}{2} (1 + \rho) (1 - \eta_k^2) \|F(x^k)\| h_k \iff [h_k]_2 := \frac{2\Theta_k}{(1 + \rho)(1 - \eta_k^2)} \leq h_k$$

und wegen $h_{k+1} = \Theta_k h_k$ auch die a-priori-Abschätzung

$$[h_{k+1}] := \Theta_k [h_k]_2 \leq h_{k+1}.$$

Bedingung (1.10) liefert dann das

Iterations-Kriterium für den quadratischen Konvergenzmodus: Um quadratische Konvergenz zu erreichen, muss die GMRES-Iteration solange fortgeführt werden bis

$$\frac{\eta_k}{1 - \eta_k^2} \leq \frac{1}{2} \rho [h_k]$$

erfüllt ist. Für $\rho \approx 1$ führt dies asymptotisch zu $\eta_k \rightarrow 0$.

Wenden wir uns nun dem *linearen Konvergenzmodus* zu. Dem Beweis von Satz 1.6 entnimmt man die Ungleichung

$$\|F(x^{k+1}) - r^k\| \leq \frac{1}{2} (1 - \eta_k^2) \|F(x^k)\| h_k$$

und erhält die a-posteriori-Abschätzung

$$[h_k]_1 := \frac{2 \|F(x^{k+1}) - r^k\|}{(1 - \eta_k^2) \|F(x^k)\|} \leq h_k, \text{ mit } r^k = F'(x^k) \delta x^k + F(x^k).$$

Wegen $h_{k+1} = \Theta_k h_k$ gilt die a-priori-Abschätzung

$$[h_{k+1}] := \Theta_k [h_k]_1 \leq h_{k+1}.$$

Bedingung (1.9) liefert dann das

Iterations-Kriterium für den linearen Konvergenzmodus: Um lineare Konvergenz zu erreichen, muss die GMRES-Iteration solange fortgeführt werden bis

$$\eta_k \leq \bar{\Theta} - \frac{1}{2} [h_k] =: \bar{\eta}_k$$

erfüllt ist. Ist $\bar{\eta}_k$ kleiner als der Wert, den man im quadratischen Konvergenzmodus fordert, so sollte man in diesen wechseln.

1.6 Numerische Tests

In diesem Abschnitt sollen das exakte und das vereinfachte NEWTON-Verfahren (NV) sowie der QNRES-Algorithmus praktischen Tests unterzogen werden. Dazu wurden diese Algorithmen basierend auf einer einfachen Matrix-Vektor-Klasse in C++ implementiert (auf Seite 23 ist die Download-Adresse angegeben). Tabellarisch sind jeweils die Startwerte und die Erfüllung der oben hergeleiteten Startwert- (*SWK*), Divergenz- (*DivK*) und Konvergenz-Kriterien (*KonK*) aufgelistet^c. Dabei wird im Falle der Konvergenz die Anzahl der benötigten Iterationen in Klammern angegeben.

Auch wenn die Aussagekraft dieser kleinen Testserie sicherlich begrenzt ist, so gibt sie doch einen ersten Eindruck von der Anwendbarkeit dieser Kriterien. Auf eine Implementierung des inexakten NEWTON-Verfahrens wurde insbesondere deshalb verzichtet, da dieser Algorithmus nur für (sehr) große Systeme verwendet wird und derartige Beispiele wenig anschaulich sind.

Test 1: Wir beginnen mit der eindimensionalen Funktion

$$f(x) := \exp(-x) + \frac{1}{x+2}.$$

Sie hat eine Nullstelle bei

$$x^* \approx 1.146.$$

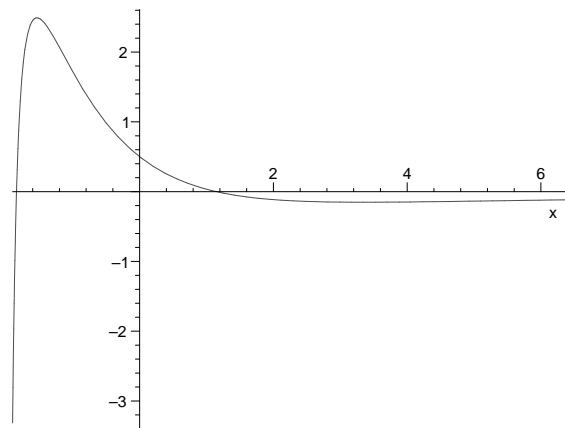


Abbildung 1: $f(x) = \exp(-x) + \frac{1}{x+2}$.

Mit $FTOL = 10^{-8}$ erhalten wir folgende Ergebnisse:

Verfahren:	Startwert:	SWK:	DivK1:	DivK2:	KonK1:
exaktes NV	0.0	nein	nein	ja	ja (5)
	0.5	nein	nein	nein	ja (4)
	1.7	nein	nein	nein	ja (5)
	1.8	ja	nein	nein	ja (5)
	2.3	ja	nein	ja	ja (7)
	2.5	ja	nein	ja	nein ^d
vereinfachtes NV	0.0	ja	–	–	ja (48)
	0.3	nein	–	–	ja (32)
	1.4	nein	–	–	ja (15)
	1.7	ja	–	–	ja (657)
	1.8	ja	–	–	nein

^cWenn ein Kriterium in der theoretischen Herleitung keine Ziffer erhalten hat, so wird es als Kriterium 1 geführt.

^dKonvergenz gegen eine andere Nullstelle.

Verfahren:	Startwert:	SWK:	DivK1:	DivK2:	KonK1:
QNRES	0.0	–	nein	–	ja (7)
	0.5	–	nein	–	ja (6)
	1.6	–	nein	–	ja (6)
	1.7	–	nein	–	ja (6)
	3.0	–	nein	–	nein

Das *exakten* NEWTON-Verfahren konvergiert mit einer Ausnahme für alle von uns geprüften Startwerte, entsprechend zeigt das mildere Divergenz-Kriterium 1 keine Divergenz an. Entfernt man sich von der Lösung, so wird das strikere Divergenz-Kriterium 2 und für $x > x^*$ auch das Startwert-Kriterium erfüllt.

Beim *vereinfachten* NEWTON-Verfahren fängt das Startwert-Kriterium Divergenz bzw. langsame Konvergenz ab.

Das Divergenz-Kriterium des *QNRES-Verfahren* wird – mit einer Ausnahme zurecht – nie erfüllt.

Test 2: Unsere zweite Testfunktion hat folgende Form:

$$\begin{aligned} f_1(x_1, x_2) &:= (x_1 + 3)(x_2^3 - 7) + 18, \\ f_2(x_1, x_2) &:= \sin(x_2) \exp(x_1) - 1. \end{aligned}$$

Sie besitzt die Nullstelle

$$x^* \approx (0.128, 1.076).$$

Mit $FTOL = 10^{-8}$ erhalten wir folgende Ergebnisse:

Verfahren:	Startwert:	SWK:	DivK1:	DivK2:	KonK1:
exaktes NV	(0, 0)	nein	nein	nein	ja (6)
	(0, 1)	nein	nein	nein	ja (4)
	(0, 2.2)	nein	nein	ja	ja (5)
	(0, 3.2)	nein	ja	ja	ja (6)
	(0, 3.5)	ja	ja	ja	nein
vereinfachtes NV	(0, 0)	ja	–	–	nein
	(0, 0.4)	ja	–	–	ja (314)
	(0, 1.1)	nein	–	–	ja (8)
	(0, 2)	ja	–	–	ja (22)
	(0, 3)	ja	–	–	ja (134)
	(0, 3.2)	ja	–	–	nein
QNRES	(0, 0)	–	ja	–	ja (14)
	(0, 1)	–	ja	–	ja (7)
	(0, 1.1)	–	nein	–	ja (6)
	(0, 2)	–	ja	–	ja (9)
	(0, 3)	–	ja	–	ja (12)
	(0, 3.2)	–	ja	–	nein

Beim *exakten* NEWTON-Verfahren funktionieren die Startwert- und Divergenz-Kriterien tadellos.

Das Startwert-Kriterium des *vereinfachten* NEWTON-Verfahrens verlangt eine sehr genaue Schätzung der Lösung. Langsame Konvergenz – aber auch Divergenz – würde dadurch verhindert.

Beim *QNRES-Algorithmus* zeigt sich ein ähnliches Bild wie beim vereinfachten NEWTON-Verfahren. Hier achtet jedoch das Divergenz-Kriterium auf hinreichende Nähe zur Lösung.

Test 3: Als letztes testen wir ein etwas größeres System ($n = 20$). Dazu benutzen wir folgende Gleichungen, die [1] entnommen sind:

$$\begin{aligned} f_1(x_1, \dots, x_n) &:= -(3 + \alpha x_1)x_1 + 2x_2 - \beta, \\ f_i(x_1, \dots, x_n) &:= x_{i-1} - (3 + \alpha x_i)x_i + 2x_{i+1} - \beta, \quad \forall i = 2, 3, \dots, n-1, \\ f_n(x_1, \dots, x_n) &:= x_{n-1} - (3 + \alpha x_n)x_n - \beta. \end{aligned}$$

Für $\alpha := -0.5$ und $\beta := 1.0$ bekommt man z. B. die Lösung

$$x^* \approx (-1.032, -1.315, -1.389, -1.408, -1.412, -1.414, -1.414, -1.414, -1.414, -1.413, \\ -1.412, -1.410, -1.406, -1.397, -1.381, -1.350, -1.291, -1.178, -0.968, -0.597).$$

Mit $FTOL = 10^{-8}$ erhalten wir folgende Ergebnisse:

Verfahren:	Startwert:	SWK:	DivK1:	DivK2:	KonK1:
exaktes NV	$(0, \dots, 0)$	ja	nein	ja	ja (8)
	$(-0.7, \dots, -0.7)$	nein	nein	ja	ja (5)
	$(-0.81, \dots, -0.81)$	nein	nein	nein	ja (4)
	$(-1, \dots, -1)$	nein	nein	nein	ja (4)
	$(-1.2, \dots, -1.2)$	nein	nein	ja	ja (4)
	$(-100, \dots, -100)$	nein	nein	ja	ja (10)
vereinfachtes NV	$(0, \dots, 0)$	ja	–	–	nein
	$(-0.6, \dots, -0.6)$	ja	–	–	ja (4535)
	$(-0.93, \dots, -0.93)$	nein	–	–	ja (24)
	$(-1000, \dots, -1000)$	nein	–	–	ja (14766)
QNRES	$(0, \dots, 0)$	–	ja	–	nein
	$(-0.19, \dots, -0.19)$	–	ja	–	ja (26)
	$(-0.82, \dots, -0.82)$	–	nein	–	ja (12)
	$(-3, \dots, -3)$	–	nein	–	ja (16)
	$(-3.7, \dots, -3.7)$	–	ja	–	ja (20)
	$(-100, \dots, -100)$	–	ja	–	ja (72)

Das *exakten* NEWTON-Verfahren konvergiert für alle von uns geprüften Startwerte. Entsprechend zeigt das mildere Divergenz-Kriterium 1 keine Divergenz an, der Startwert wird mit einer Ausnahme akzeptiert. Das striktere Divergenz-Kriterium 2 lässt jedoch nur einen sehr kleinen Startwertebereich zu.

Das Startwert-Kriterium des *vereinfachten* NEWTON-Verfahrens unterbindet die Divergenz bzw. sehr langsame Konvergenz für $x_i > -0.75$. Die extrem langsame Konvergenz im Bereich $x_i \ll -10$ wird nicht abgefangen.

Für *QNRES* zeigt sich im Wesentlichen die gleiche Situation wie beim exakten NEWTON-Verfahren.

2 Konvexe Optimierung

2.1 Einleitung

In diesem Kapitel geht es um das Problem, ein strikt konvexes, zweimal stetig-differenzierbares Funktional

$$f : \mathcal{D} \longrightarrow \mathbb{R}, \quad \mathcal{D} \subseteq \mathbb{R}^n \text{ konvex},$$

zu minimieren. Dann ist das Auffinden des Minimums äquivalent mit dem Lösen der Gleichung

$$F(x) := \text{grad } f(x) = 0, \quad x \in \mathcal{D}.$$

Die Ableitung $F'(x) = f''(x)$ der Gradientenabbildung ist in diesem Falle symmetrisch und (strikt) positiv definit. Daher kann man durch

$$\langle \cdot, \cdot \rangle_{F'(x)} : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}, \quad \langle u, v \rangle_{F'(x)} \longmapsto \langle u, F'(x)v \rangle \quad \text{bzw.}$$

$$\| \cdot \|_{F'(x)} : \mathbb{R}^n \longrightarrow \mathbb{R}, \quad \|u\|_{F'(x)} \longmapsto \|F'(x)^{1/2}u\| = \sqrt{\langle u, u \rangle_{F'(x)}}$$

ein Energie-Skalarprodukt bzw. eine Energienorm definieren.

Mit $B \in \text{GL}(n, \mathbb{R})$ können wir das transformierte Minimierungsproblem

$$g(y) := f(By) \longrightarrow \min, \quad x = By$$

betrachten. Dann folgt (mit $G(y) := \text{grad } g(y)$)

$$G(y) = B^T F(By), \quad x = By \quad \implies \quad G'(y) = B^T F'(x)B, \quad x = By. \quad (2.1)$$

Die Matrix $G'(y)$ ist konjugiert zur Matrix $F'(x)$. Daher spricht man in diesem Zusammenhang von *affiner Konjugiertheit*. Der Trägheitssatz von SYLVESTER garantiert ferner, dass $G'(y)$ neben der Symmetrie auch die (strikte) positive Definitheit von $F'(x)$ erbt.

Mit beliebigen $u = B\bar{u}$, $v = B\bar{v}$, $x = By$ gilt wegen (2.1)

$$\langle u, v \rangle_{F'(x)} = u^T F'(x)v = \bar{u}^T B^T F'(x)B\bar{v} = \bar{u}^T G'(y)\bar{v} = \langle \bar{u}, \bar{v} \rangle_{G'(y)}.$$

Affin konjugierte Konvergenzsätze beinhalten daher Aussagen über die Energienorm der Korrekturen $\|F'(x)^{1/2}\Delta x\|$ und die Energienorm des Fehlers $\|F'(x)^{1/2}(x^k - x^*)\|$. Da außerdem wegen $g(y^k) := f(By^k) = f(x^k)$ die Funktionswerte $f(x^k)$ invariant bleiben, können auch sie zur Beschreibung herangezogen werden. Dies werden wir wieder zuerst am Beispiel des gewöhnlichen NEWTON-Verfahrens zeigen.

2.2 Gewöhnliches Newton-Verfahren

Zunächst formulieren eine affin konjugierte Variante des Satzes von NEWTON-MYSOVSKIKH:

Satz 2.1 (Newton-Mysovskikh, affin konjugierte Version)

Sei $f : \mathcal{D} \rightarrow \mathbb{R}$ ein strikt-konvexes zweimal stetig-differenzierbares Funktional über einer offenen und konvexen Menge $\mathcal{D} \subseteq \mathbb{R}^n$. Sei $F(x) := f'(x)^T$ und $F'(x) = f''(x)$ eine symmetrische und (strikt) positiv definite Matrix und gelten folgende Bedingungen:

1. Es existiert ein $0 \leq \omega < \infty$ so, dass

$$\left\| F'(z)^{-1/2} (F'(y) - F'(x)) (y - x) \right\| \leq \omega \left\| F'(x)^{1/2} (y - x) \right\|^2 \quad \forall \text{ kollineare}^e x, y, z \in \mathcal{D}.$$

^ed. h. auf derselben Gerade liegende

2. Die Menge

$$\mathcal{L}_0 := \left\{ x \in \mathcal{D} \mid f(x) \leq f(x^0) \right\}$$

ist kompakt.

3. Für den Startwert x^0 gilt

$$h_0 := \omega \left\| F'(x^0)^{1/2} \Delta x^0 \right\| < 2.$$

Dann gilt:

1. Die Iterationen des gewöhnlichen NEWTON-Verfahrens (1.1) bleiben in \mathcal{L}_0 und konvergieren gegen das Minimum x^* von f mit der Rate

$$\left\| F'(x^{k+1})^{1/2} \Delta x^{k+1} \right\| \leq \frac{\omega}{2} \left\| F'(x^k)^{1/2} \Delta x^k \right\|^2. \quad (2.2)$$

Oder, mit den Definitionen

$$\varepsilon_k := \left\| F'(x^k)^{1/2} \Delta x^k \right\|^2, \quad h_k := \omega \left\| F'(x^k)^{1/2} \Delta x^k \right\|,$$

mit den Abschätzungen

$$-\frac{1}{6} h_k \varepsilon_k \leq f(x^k) - f(x^{k+1}) - \frac{1}{2} \varepsilon_k \leq \frac{1}{6} h_k \varepsilon_k, \quad (2.3)$$

$$\frac{1}{6} \varepsilon_k \leq f(x^k) - f(x^{k+1}) \leq \frac{5}{6} \varepsilon_k. \quad (2.4)$$

2. Der Abstand zum Minimum ist durch

$$f(x^0) - f(x^k) \leq \frac{\frac{5}{6} \varepsilon_0}{1 - \frac{h_0}{2}}$$

beschränkt.

Bemerkung 2.2

In [5] wird Satz 2.1 vom \mathbb{R}^n auf einen beliebigen HILBERT-Raum verallgemeinert.

Seien ε_k und h_k wie in Satz 2.1 definiert und außerdem

$$\Theta_k := \sqrt{\frac{\varepsilon_{k+1}}{\varepsilon_k}} = \frac{\left\| F'(x^{k+1})^{1/2} \Delta x^{k+1} \right\|}{\left\| F'(x^k)^{1/2} \Delta x^k \right\|}.$$

Dann sind die wichtigsten Konvergenzaussagen der affin konjugierten Version des Satzes von NEWTON-MYSOVSKIKH zum einen (gemäß (2.2))

$$\Theta_k = \frac{h_{k+1}}{h_k} \leq \frac{1}{2} h_k < 1$$

und (gemäß (2.4))

$$f(x^{k+1}) - f(x^k) \leq -\frac{1}{6} \varepsilon_k$$

Im Falle von $k = 0$ liefert dies das

Startwert-Kriterium: Falls

$$\Theta_0 \geq 1$$

gilt, weisen wir den Startwert x^0 als *nicht hinreichend nahe* zurück.

Und weiter:

Divergenz-Kriterium 1: Wir brechen die gewöhnliche NEWTON-Iteration als *divergent* ab, falls

$$f(x^{k+1}) - f(x^k) > -\frac{1}{6}\varepsilon_k.$$

Analog zum residuen-basierten, gewöhnlichen NEWTON-Verfahren (siehe Seite 3) ergibt sich

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

und damit:

Divergenz-Kriterium 2: Wir brechen die gewöhnliche NEWTON-Iteration als *divergent* ab, falls

$$\Theta_k \geq \Theta_0 \text{ für } k > 0.$$

Divergenz-Kriterium 3: Wir brechen die gewöhnliche NEWTON-Iteration als *divergent* ab, falls

$$\Theta_{k+1} \geq \frac{\Theta_k^2}{\Theta_0}.$$

Sei *ETOL* die vom Benutzer definierte *Energienorm-Fehlertoleranz*. Wegen (2.3) gilt asymptotisch

$$f(x^k) - f(x^{k+1}) \longrightarrow \frac{1}{2}\varepsilon_k \text{ für } k \longrightarrow \infty.$$

Diese Beziehung kann als Konvergenz-Kriterium gelten:

Konvergenz-Kriterium 1: Wir beenden die gewöhnliche NEWTON-Iteration, falls

$$f(x^k) - f(x^{k+1}) \leq \frac{1}{2}ETOL^2$$

erfüllt ist und geben x^{k+1} als Lösung aus.

Konvergenz-Kriterium 2: Wir beenden die gewöhnliche NEWTON-Iteration, falls

$$\varepsilon_k \leq ETOL^2$$

erfüllt ist und geben x^k als Lösung aus.

2.3 Vereinfachtes Newton-Verfahren

Wir untersuchen nun das vereinfachte NEWTON-Verfahren (1.4) im Falle konvexer Optimierung:

Satz 2.3

Sei $f : \mathcal{D} \rightarrow \mathbb{R}$ ein strikt konvexes, zweimal stetig-differenzierbares Funktional über einer konvexen Menge $\mathcal{D} \subseteq \mathbb{R}^n$. Sei $F(x) := f'(x)^T$ und $F'(x) = f''(x)$ eine symmetrische und (strikt) positiv definite Matrix und gelten folgende Bedingungen:

1. Es existiert ein $0 \leq \omega < \infty$ so, dass

$$\left\| F'(x^0)^{-1/2} (F'(z) - F'(x^0)) v \right\| \leq \omega \left\| F'(x^0)^{1/2} (z - x^0) \right\| \left\| F'(x^0)^{1/2} v \right\| \quad \forall z \in \mathcal{D}, v \in \mathbb{R}^n.$$

2. Für den Startwert x^0 gilt

$$h_0 := \omega \left\| F'(x^0)^{1/2} \overline{\Delta x}^0 \right\| \leq \frac{1}{2}. \quad (2.5)$$

Sei ferner $t^* := 1 - \sqrt{1 - 2h_0}$.

Dann gilt:

1. Die Iterationen des vereinfachten NEWTON-Verfahrens (1.4) konvergieren gegen ein x^* mit

$$\omega \|x^* - x^0\| \leq t^* .$$

2. Die Konvergenzrate kann – mit $\varepsilon_k := \left\| F'(x^0)^{1/2} \overline{\Delta x}^k \right\|^2$ – durch

$$-\frac{1}{6}\varepsilon_k(t_{k+1} + 2t_k) \leq f(x^k) - f(x^{k+1}) - \frac{1}{2}\varepsilon_k \leq \frac{1}{6}\varepsilon_k(t_{k+1} + 2t_k) \quad (2.6)$$

abgeschätzt werden. Oder bezüglich der Energienorm durch

$$\Theta_k = \left(\frac{\varepsilon_{k+1}}{\varepsilon_k} \right)^{1/2} \leq \frac{1}{2}(t_{k+1} + t_k) ,$$

wobei die $\{t_k\}$ durch

$$t_0 := 0, \quad t_{k+1} = h_0 + \frac{1}{2}t_k^2 < t^*$$

definiert sind.

Analog zum affin kontravarianten Falle erfordert (2.5) das

Startwert-Kriterium: Falls

$$\Theta_0 > \frac{1}{4}$$

gilt, weisen wir den Startwert x^0 als *nicht hinreichend nahe* zurück.

Wegen (2.6) trifft die asymptotische Aussage $f(x^k) - f(x^{k+1}) \rightarrow \frac{1}{2}\varepsilon_k$ *nicht* auf das vereinfachte NEWTON-Verfahren zu. Wir können jedoch die Iterations-Kriterien im *affin kovarianten* Falle^e übertragen, indem wir dort die Norm durch die Energienorm ersetzen. Mit

$$\Theta_k := \frac{\left\| F'(x^0)^{1/2} \overline{\Delta x}^{k+1} \right\|}{\left\| F'(x^0)^{1/2} \overline{\Delta x}^k \right\|}$$

übernehmen wir

Divergenz-Kriterium 1: Wir brechen die vereinfachte NEWTON-Iteration als *divergent* ab, falls

$$\Theta_k \geq 1 .$$

Divergenz-Kriterium 2: Wir brechen die vereinfachte NEWTON-Iteration als *divergent* ab, falls

$$\Theta_k \geq [t^*] := 1 - \sqrt{1 - 4\Theta_0} .$$

Konvergenz-Kriterium: Wir beenden die vereinfachte NEWTON-Iteration, falls

$$\varepsilon_{k-1} = \left\| F'(x^0)^{1/2} \overline{\Delta x}^{k-1} \right\|^2 \leq ETOL^2$$

erfüllt ist und geben x^k als Lösung aus.

^eFür Details sei auf [3, S. 55] bzw. [9] verwiesen.

2.4 Inexaktes Newton-PCG-Verfahren

Auch für den affin konjugierten Fall wollen wir nun ein inexaktes NEWTON-Verfahren erarbeiten. Da die JACOBI-Matrix der Gradientenabbildung in diesem Falle symmetrisch und positiv definit ist, bietet sich ein (vorkonditioniertes) konjugiertes Gradientenverfahren (preconditioned conjugate gradient $\hat{=}$ PCG) an:

$$F'(x^k)(\delta x^k - \Delta x^k) = r^k, \text{ mit } \frac{\|r^k\|_{F'(x^k)}}{\|\delta x^k\|_{F'(x^k)}} \leq \delta_k, \quad x^{k+1} = x^k + \delta x^k \quad \forall k = 0, 1, \dots \quad (2.7)$$

Hierbei bezeichnet Δx^k die Korrektur des *exakten* NEWTON-Verfahrens.

Das PCG-Verfahren erfüllt darüber hinaus die Orthogonalitäts-Bedingung (GALERKIN-Bedingung)

$$\langle \delta x^k, F'(x^k)(\delta x^k - \Delta x^k) \rangle = \langle \delta x^k, r^k \rangle = 0, \quad (2.8)$$

welche Voraussetzung für unsere theoretische Betrachtung ist. ([5] entnimmt man, dass die GALERKIN-Bedingung im Raum H^1 mit dem L_2 -Skalarprodukt von jeder beliebigen Finite-Elemente-Methode erfüllt wird.)

Satz 2.4

Sei $f : \mathcal{D} \rightarrow \mathbb{R}$ ein strikt konvexes, zweimal stetig-differenzierbares Funktional über einer offenen und konvexen Menge $\mathcal{D} \subseteq \mathbb{R}^n$. Sei $F'(x) := f''(x)$ eine symmetrische und (strikt) positiv definite Matrix und gelten folgende Bedingungen:

1. Es existiert ein $0 \leq \omega < \infty$ so, dass für alle $v \in \mathbb{R}^n$

$$\left\| F'(z)^{-1/2} (F'(y) - F'(x)) v \right\| \leq \omega \left\| F'(x)^{1/2} (y - x) \right\| \left\| F'(x)^{1/2} v \right\| \quad \forall \text{ kollineare } x, y, z \in \mathcal{D}.$$

2. Die Menge

$$\mathcal{L}_0 := \left\{ x \in \mathcal{D} \mid f(x) \leq f(x^0) \right\}$$

ist kompakt.

An jeder wohldefinierten Interierten x^k definieren wir die Größen des exakten NEWTON-Verfahrens

$$\begin{aligned} \varepsilon_k &:= \left\| F'(x^k)^{1/2} \Delta x^k \right\|^2, \\ h_k &:= \omega \left\| F'(x^k)^{1/2} \Delta x^k \right\|. \end{aligned}$$

Ferner sei für die Abschätzung des Fehlers der inneren PCG-Iteration

$$\begin{aligned} \delta_k &:= \frac{\left\| F'(x^k)^{1/2} (\delta x^k - \Delta x^k) \right\|}{\left\| F'(x^k)^{1/2} \delta x^k \right\|}, \\ \varepsilon_k^\delta &:= \left\| F'(x^k)^{1/2} \delta x^k \right\|^2 = \frac{\varepsilon_k}{1 + \delta_k^2}, \\ h_k^\delta &:= \omega \left\| F'(x^k)^{1/2} \delta x^k \right\| = \frac{h_k}{\sqrt{1 + \delta_k^2}}. \end{aligned}$$

Dann gilt für das inexakte NEWTON-PCG-Verfahren (2.7), welches die GALERKIN-Bedingung (2.8) erfüllt, und dessen innere PCG-Iteration jeweils mit $\delta x_0^k = 0$ gestartet wird:

1. Linearer Konvergenzmodus:

Angenommen der Startwert erfüllt

$$\frac{h_0}{2} < \bar{\Theta} < 1$$

und es gilt während der gesamten Iteration $\delta_{k+1} \geq \delta_k$. Wird dann die innere PCG-Iteration so kontrolliert, dass

$$\vartheta(h_k^\delta, \delta_k) := \frac{h_k^\delta + \delta_k \left(h_k^\delta + \sqrt{4 + (h_k^\delta)^2} \right)}{2\sqrt{1 + \delta_k^2}} \leq \bar{\Theta} \quad (2.9)$$

gilt, was

$$\delta_k \leq \frac{\bar{\Theta}}{\sqrt{1 - \bar{\Theta}^2}} \quad (2.10)$$

impliziert.

Dann bleiben die NEWTON-PCG-Iterierten $\{x^k\}$ in \mathcal{L}_0 und konvergieren mindestens linear gegen das Minimum $x^* \in \mathcal{L}_0$ und zwar mit

$$\begin{aligned} \left\| F'(x^{k+1})^{1/2} \delta x^{k+1} \right\| &\leq \bar{\Theta} \left\| F'(x^k)^{1/2} \delta x^k \right\|, \\ \left\| F'(x^{k+1})^{1/2} \Delta x^{k+1} \right\| &\leq \bar{\Theta} \left\| F'(x^k)^{1/2} \Delta x^k \right\|. \end{aligned} \quad (2.11)$$

2. Quadratischer Konvergenzmodus:

Falls der Startwert x^0 für ein $\rho > 0$

$$h_0^\delta < \frac{2}{1 + \rho}$$

erfüllt und die innere PCG-Iteration so kontrolliert wird, dass

$$\delta_k \leq \frac{\rho h_k^\delta}{h_k^\delta + \sqrt{4 + (h_k^\delta)^2}} \quad (2.12)$$

gilt, was zumindest

$$\delta_0 < \frac{\rho}{1 + \sqrt{1 + (1 + \rho)^2}} \quad (2.13)$$

erfordert.

Dann bleiben die NEWTON-PCG-Iterierten $\{x^k\}$ in \mathcal{L}_0 und konvergieren quadratisch gegen das Minimum $x^* \in \mathcal{L}_0$ und zwar mit

$$\begin{aligned} \left\| F'(x^{k+1})^{1/2} \delta x^{k+1} \right\| &\leq \frac{\omega}{2}(1 + \rho) \left\| F'(x^k)^{1/2} \delta x^k \right\|^2, \\ \left\| F'(x^{k+1})^{1/2} \Delta x^{k+1} \right\| &\leq \frac{\omega}{2}(1 + \rho) \left\| F'(x^k)^{1/2} \Delta x^k \right\|^2. \end{aligned}$$

3. Die Konvergenz kann wie folgt durch Funktionswerte abgeschätzt werden:

$$\left| f(x^k) - f(x^{k+1}) - \frac{1}{2} \varepsilon_k^\delta \right| \leq \frac{1}{6} h_k^\delta \varepsilon_k^\delta \quad (2.14)$$

Wir gehen im Folgenden davon aus, dass uns eine hinreichend gute Schätzung des Energienorm-Fehlers δ_k der inneren PCG-Iteration zur Verfügung steht (siehe dazu z. B. [4]). Dann ergibt sich

$$\varepsilon_k = (1 + \delta_k^2) \varepsilon_k^\delta = (1 + \delta_k^2) \left\| F'(x^k)^{1/2} \delta x^k \right\|^2.$$

Die Ungleichungen (2.11) und (2.14) liefern uns

Divergenz-Kriterium 1: Wir brechen die inexakte NEWTON-PCG-Iteration als *divergent* ab, falls

$$f(x^{k+1}) - f(x^k) > -\frac{1}{6}\varepsilon_k.$$

Divergenz-Kriterium 2: Wir brechen die inexakte NEWTON-PCG-Iteration als *divergent* ab, falls

$$\Theta_k := \sqrt{\frac{\varepsilon_{k+1}}{\varepsilon_k}} > \bar{\Theta}.$$

Um ein Iterations-Kriterium für den *quadratischen Konvergenzmodus* herzuleiten, nehmen wir an, dass $[h_0] < \frac{2}{1+\rho}$ für $\rho = 1$ erfüllt ist. In Übereinstimmung mit (2.13) sei ferner $\delta_0 \approx \frac{1}{4}$. Um nun Bedingung (2.12) kontrollieren zu können, brauchen wir eine Abschätzung von h_k^δ . (2.14) liefert die a-posteriori-Schätzung

$$[h_k^\delta]_2 := \frac{6}{\varepsilon_k^\delta} \left| f(x^k) - f(x^{k+1}) - \frac{1}{2}\varepsilon_k^\delta \right|.$$

Damit erhalten wir die a-priori-Schätzung

$$[h_k^\delta] := \Theta_{k-1} [h_{k-1}]_2 = \frac{6\Theta_{k-1}}{\varepsilon_{k-1}^\delta} \left| f(x^{k-1}) - f(x^k) - \frac{1}{2}\varepsilon_{k-1}^\delta \right|. \quad (2.15)$$

Dies ergibt das

Iterations-Kriterium für den quadratischen Konvergenzmodus: Um quadratische Konvergenz zu erreichen, muss die PCG-Iteration solange fortgeführt werden bis

$$\delta_k \leq \frac{\rho[h_k^\delta]}{[h_k^\delta] + \sqrt{4 + [h_k^\delta]^2}} \quad \text{mit} \quad [h_k^\delta] := \frac{6\Theta_{k-1}}{\varepsilon_{k-1}^\delta} \left| f(x^{k-1}) - f(x^k) - \frac{1}{2}\varepsilon_{k-1}^\delta \right|$$

erfüllt ist.

Man kann zeigen, dass

$$\delta_k \longrightarrow 0 \quad \text{für} \quad k \longrightarrow \infty$$

gilt. Das bedeutet, dass *im quadratischen Konvergenzmodus immer mehr innere PCG-Iterationen durchgeführt werden müssen, je näher die Iterierten dem Minimum kommen!*

Daher sollte zum *linearen Konvergenzmodus* gewechselt werden, falls Θ_k deutlich unter $\bar{\Theta}$ fällt. Hier muss Bedingung (2.9) sichergestellt werden. Dazu nähern wir $\vartheta(h_k^\delta, \delta_k)$ durch

$$[\vartheta(h_k^\delta, \delta_k)] := \vartheta([h_k^\delta], \delta_k) \leq \vartheta(h_k^\delta, \delta_k)$$

an, wobei wir die a-priori-Abschätzung (2.15) benutzen können. Dann liefert (2.9) das

Iterations-Kriterium für den linearen Konvergenzmodus: Um lineare Konvergenz zu erreichen, muss die PCG-Iteration solange fortgeführt werden bis

$$\vartheta([h_k^\delta], \delta_k) := \frac{[h_k^\delta] + \delta_k \left([h_k^\delta] + \sqrt{4 + [h_k^\delta]^2} \right)}{2\sqrt{1 + \delta_k^2}} \leq \bar{\Theta}$$

erfüllt ist.

Zusammen mit der Voraussetzung $\delta_{k+1} \geq \delta_k$ ergibt (2.10)

$$\delta_k \longrightarrow \frac{\bar{\Theta}}{\sqrt{1 - \bar{\Theta}^2}} \quad \text{für} \quad k \longrightarrow \infty.$$

Das bedeutet, dass *im linearen Konvergenzmodus immer weniger innere PCG-Iterationen durchgeführt werden müssen, je näher die Iterierten dem Minimum kommen!*

Schließlich erhalten wir wegen (2.14) analog zum gewöhnlichen NEWTON-Verfahren

Konvergenz-Kriterium 1: Wir beenden die inexakte NEWTON-PCG-Iteration, falls

$$f(x^k) - f(x^{k+1}) \leq \frac{1}{2} ETOL^2$$

erfüllt ist und geben x^{k+1} als Lösung aus.

Konvergenz-Kriterium 2: Wir beenden die inexakte NEWTON-PCG-Iteration, falls

$$\varepsilon_k \leq ETOL^2$$

erfüllt ist und geben x^k als Lösung aus.

2.5 Numerische Tests

Zum Abschluss soll das exakte und das vereinfachte NEWTON-Verfahren für konvexe Optimierung getestet werden. Die Ergebnisse sind wie auf Seite 10 beschrieben dargestellt. Wir werden insbesondere auch zwei *nicht* (strikt) konvexe Probleme untersuchen und dabei die Auswirkungen fehlender Konvexitätseigenschaften beobachten.

Test 1: Als erstes testen wir eine sehr einfache, auf ganz \mathbb{R}^2 strikt konvexe Funktion

$$f(x_1, x_2) := x_1^2 + x_1 + x_2^2$$

mit einem globalen Minimum im Punkt

$$x^* = (-0.5, 0).$$

Mit $ETOL = 10^{-6}$ erhalten wir folgende Ergebnisse:

Verfahren:	Startwert:	SWK:	DivK1:	DivK2:	DivK3:	KonK1:	KonK2:
exaktes NV	(0, 0)	nein	nein	nein	nein	ja (2)	ja (1)
	(10, -10)	nein	nein	nein	nein	ja (3)	ja (2)
	(-10, 100)	nein	nein	nein	nein	ja (3)	ja (2)
	(1000, 10)	nein	nein	nein	nein	ja (5)	ja (4)
vereinf. NV	(0, 0)	nein	nein	nein	-	ja (1)	-
	(10, -10)	nein	nein	nein	-	ja (2)	-
	(-10, 100)	nein	nein	ja	-	ja (3)	-
	(100, 10)	nein	nein	ja	-	ja (4)	-
	(700, 10)	ja	nein	nein	-	ja (64)	-
	(700, 700)	ja	nein	nein	-	ja (41)	-

Das *exakten* NEWTON-Verfahren konvergiert wie erwartet für alle von uns geprüften Startwerte. Keines der Divergenz-Kriterien wird erfüllt^f.

Das *vereinfachte* NEWTON-Verfahren konvergiert ebenfalls in allen Fällen. Entfernt man sich zu weit von der Lösung, steigt die Iterationszahl und das Startwert-Kriterium wird erfüllt.

^fEs kann passieren, dass beim allerletzten, winzigen Schritt durch Rundungsfehler eines der Kriterien scheinbar erfüllt wird.

Test 2: Unsere zweite Testfunktion (Abb. 2) ist nur lokal strikt konvex:

$$f(x_1, x_2) := x_1^3 - 20x_1 + x_2^3 - 10x_2.$$

Ein lokales Minimum nimmt sie im Punkt

$$x^* \approx (2.582, 1.826)$$

an. Mit $ETOL = 10^{-6}$ bzw. $ETOL = 10^{-5}$ erhalten wir folgende Ergebnisse:

Verfahren:	Startwert:	SWK:	DivK1:	DivK2:	DivK3:	KonK1:	KonK2:
exaktes NV	(2, 2)	nein	nein	nein	nein	ja (3)	ja (3)
	(3, 2)	nein	nein	nein	nein	ja (4)	ja (3)
	(10, 2)	nein	nein	nein	nein	ja (6)	ja (5)
	(2, 1)	ja	nein	nein	ja	ja (4)	ja (4)
	(2, 0.1)	ja	nein	nein	ja	ja (6)	ja (8)
vereinf. NV	(2, 2)	nein	nein	nein	–	ja (10)	–
	(3, 2)	nein	nein	ja	–	ja (6)	–
	(10, 2)	nein	nein	ja	–	ja (44)	–
	(2, 1)	ja	nein	nein	–	ja (64)	–
	(2, 0.1)	ja	ja	nein	–	nein	–

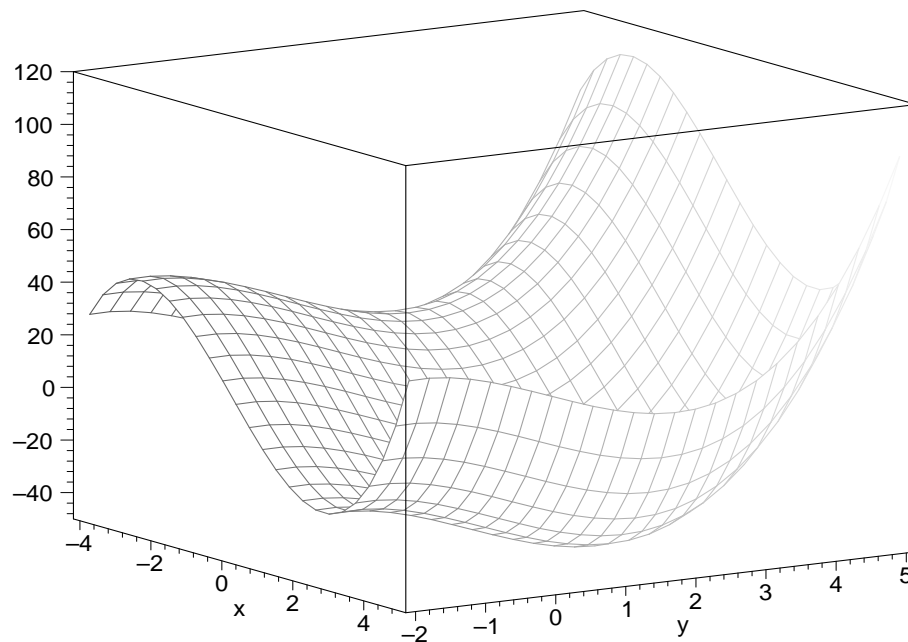


Abbildung 2: $f(x_1, x_2) = x_1^3 - 20x_1 + x_2^3 - 10x_2$ im Bereich eines lokalen Minimums.

Das *exakten* NEWTON-Verfahren konvergiert problemlos, wenn man den Startwert in x -Richtung die Steigung hinauf verlegt. Nähert man sich in y -Richtung dem Sattelpunkt, konvergiert das Verfahren noch, aber das Startwert- und ein Divergenz-Kriterium werden erfüllt.

Ähnlich ist auch die Situation beim *vereinfachten* NEWTON-Verfahren, solange man sich in der Nähe der Lösung befindet. Andernfalls nimmt die Iterationszahl deutlich zu oder das Verfahren konvergiert überhaupt nicht – das Startwert-Kriterium warnt in diesem Falle.

Test 3: Wir wollen nun eine Funktion betrachten, die zwar konvex, aber nicht strikt konvex ist (sie wurde aus [6] entnommen):

$$f(x_1, x_2, x_3, x_4) := (x_1 + 10x_2)^2 + 5(x_3 - x_4)^2 + (x_2 - 2x_3)^4 + 10(x_1 - x_4)^4.$$

Man rechnet leicht nach, dass ihre HESSE-Matrix im Minimum bei

$$x^* = (0, 0, 0, 0)$$

singulär ist. Mit $ETOL = 10^{-6}$ bzw. $ETOL = 10^{-5}$ erhalten wir folgende Ergebnisse:

Verfahren:	Startwert:	SWK:	DivK1:	DivK2:	DivK3:	KonK1:	KonK2:
exaktes NV	(3, -1, 0, 1)	nein	ja	ja	ja	ja (23)	ja (21)
	(1, -1, 0, 1)	ja	ja	nein	ja	ja (26)	ja (24)
	(0, 0, 0, 1)	ja	ja	nein	ja	ja (32)	ja (31)
vereinf. NV	(3, -1, 0, 1)	nein	nein	ja	-	ja (19340)	-
	(1, -1, 0, 1)	ja	ja	nein	-	nein	-
	(0, 0, 0, 1)	ja	nein	nein	-	nein	-

Wir sehen, dass die Divergenz-Kriterien unbrauchbarer werden, was vermutlich mit der fehlenden Konvexität im Lösungspunkt zusammenhängt. Das *vereinfachte* NEWTON-Verfahren ist für das Problem gänzlich ungeeignet.

Test 4: Als letztes Beispiel betrachten wir die nur teilweise konvexe ROSENBROCK-Funktion:

$$f(x_1, x_2) := 100(x_2 - x_1^2)^2 + (1 - x_1)^2.$$

Ihr globales Minimum nimmt sie im Punkt

$$x^* = (1, 1)$$

an und ist dort strikt konvex.

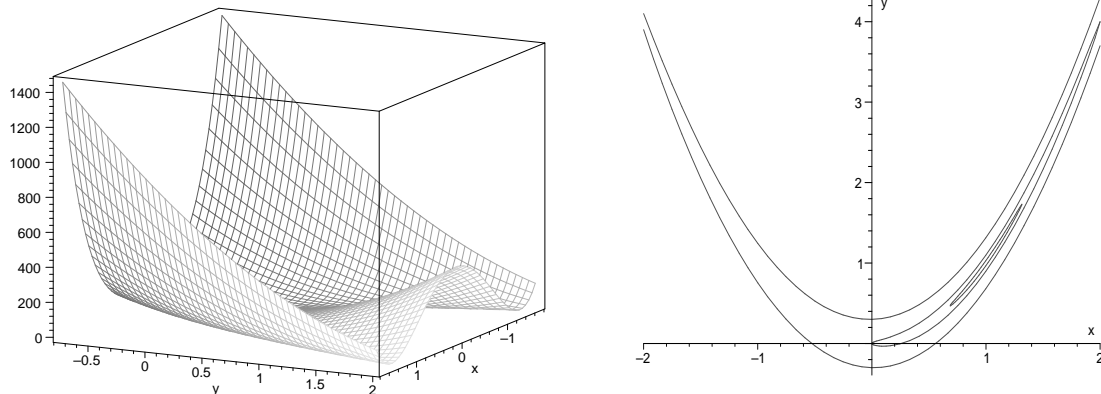


Abbildung 3: ROSENBROCKs Funktion mit parabelförmigem Tal (siehe [12] und [7]). Links: Dreidimensionale Darstellung. Rechts: Niveaulinien für die Werte 10, 1 und $\frac{1}{10}$ (von außen nach innen).

Mit $ETOL = 10^{-6}$ erhalten wir folgende Ergebnisse:

Verfahren:	Startwert:	SWK:	DivK1:	DivK2:	DivK3:	KonK1:	KonK2:
exaktes NV	(-1.2, 1)	nein	ja	ja	ja	ja (9)	ja (7)
	(0, 0)	ja	nein	nein	ja	ja (3)	ja (3)
	(0.8, 0.8)	nein	ja	ja	ja	ja (7)	ja (5)
	(2, 2)	nein	ja	ja	ja	ja (7)	ja (6)
vereinf. NV	(-1.2, 1)	nein	ja	ja	–	nein	–
	(0, 0)	ja	ja	nein	–	nein	–
	(0.8, 0.8)	nein	ja	nein	–	nein	–
	(2, 2)	nein	nein	ja	–	nein	–

Wie zu erwarten ist das Ergebnis bei einer teilweise nichtkonvexen Funktion verheerend: Beim *exakten* NEWTON-Verfahren versagen die Divergenz-Kriterien und das *vereinfachte* NEWTON-Verfahren konvergiert überhaupt nicht.

Schlussbemerkungen

Dieser Vortrag baut im Wesentlichen auf die Abschnitte 2.2 und 2.3 des Buchs *Newton Methods for Nonlinear Problems* von Peter Deuffhard ([3]) auf; die darin enthaltenen affin invarianten Konvergenzsätze wurden übernommen. Dort findet der interessierte Leser auch die zugehörigen – teilweise recht umfangreichen – Beweise, auf deren Abdruck hier verzichtet wurde.

Das Handout, die Präsentationsfolien und die Quelltexte (in C++) stehen unter

<http://www.mathi.uni-heidelberg.de/~ferreau/newton/>

zum Download bereit.

Literatur

- [1] BROYDEN, C. G.: *A Class of Methods for Solving Nonlinear Simultaneous Equations*. Mathematics of Computation, Vol. 19, No. 92, 577-583 (1965)
- [2] DEMBO, R. S.; EISENSTAT, S. C.; STEihaug, T.: *Inexact Newton Methods*. SIAM Journal on Numerical Analysis, Vol. 19, No. 2, 400-408 (1982)
- [3] DEUFLHARD, P.: *Newton Methods for Nonlinear Problems*. Springer, New York (2004).
- [4] DEUFLHARD, P.: *Cascadic Conjugate Gradient Methods for Elliptic Partial Differential Equations I. Algorithm and Numerical Results*. Konrad-Zuse-Zentrum für Informationstechnik Berlin, Preprint SC 93-23 (1993)
- [5] DEUFLHARD, P.; WEISER, M.: *Local Inexact Newton Multilevel FEM for Nonlinear Elliptic Problems*. Konrad-Zuse-Zentrum für Informationstechnik Berlin, Preprint SC 96-29 (1996)
- [6] FLETCHER, R.; POWELL, M. J. D.: *A rapidly convergent descent method for minimization*. The Computer Journal, Vol. 6, Issue 2, 163-168 (1963)
- [7] FLETCHER, R.; REEVES, C. M.: *Function minimization by conjugate gradients*. The Computer Journal, Vol. 7, Issue 2, 149-154 (1964)
- [8] HOHMANN, A.: *Inexact Gauss Newton Methods for Parameter Dependent Nonlinear Problems*. Dissertation, Freie Universität Berlin (1994)
- [9] KIRCHES, C.: *Fehlergesteuerte Algorithmen*. Seminarvortrag, Universität Heidelberg (2004).
- [10] NOCEDAL, J.; WRIGHT, S. J.: *Numerical Optimization*. Springer, New York (1999).
- [11] PONTA, O.: *P. Deufhard, Newton Methods for Nonlinear Problems, Kapitel 1*. Seminarvortrag, Universität Heidelberg (2004).
- [12] ROSENBROCK, H. H.: *An automatic method for finding the greatest or the least value of a function*. The Computer Journal, Vol. 3, Issue 3, 175-184 (1960)