Gleichungssysteme: Globale Newtonverfahren

Dirk Schröder

Residuum basierter Abstieg

Gedämpfte Newton Iteration:

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

Kontrahierendes Residuum:

$$||F(x^{k+1})|| < ||F(x^k)||$$

Programm des Vortrags

- Theorie
  - optimale Dämpfungsfaktoren
  - Globale Konvergenz in einem Bereich dieser Faktoren

Programm des Vortrags

- Theorie
  - optimale Dämpfungsfaktoren
  - Globale Konvergenz in einem Bereich dieser Faktoren
- Praxis
  Ziel: Residuum basierte Strategien für die Wahl der Dämpfungsfaktoren
  - exakte Newton Korrektur
  - inexakte Variante mit dem iterativen Löser von GMRES

Affine Contravarianz: Konvergenzanalyse

Ziel: Finde Schrittweite entlang der Newtonrichtung, so dass die Residuumreduktion in gewisser Weise optimal ist.

Lokaler Abstieg

Sei $F \in C^1(D)$ mit $D \subset \mathbb{R}^n$ offen konvex und $F'(x)$ nicht-singulär für alle $x \in D$. Es gelte weiterhin die spezielle affin, contravariante Lipschitzbedingung:

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

Lokaler Abstieg

Sei $F \in C^1(D)$ mit $D \subset \mathbb{R}^n$ offen konvex und $F'(x)$ nicht-singulär für alle $x \in D$. Es gelte weiterhin die spezielle affin, contravariante Lipschitzbedingung:

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

Dann gilt mit $h_k := \omega ||F(x^k)||$ und $\lambda \in [0, \min(1, 2/h_k)]$:

$$||F(x^k + \lambda \Delta x^k)||_2 \leq t_k(\lambda) ||F(x^k)||_2,$$

mit $t_k(\lambda) := 1 - \lambda + \frac{1}{2}\lambda^2 h_k$.

Lokaler Abstieg

Sei $F \in C^1(D)$ mit $D \subset \mathbb{R}^n$ offen konvex und $F'(x)$ nicht-singulär für alle $x \in D$. Es gelte weiterhin die spezielle affin, contravariante Lipschitzbedingung:

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

Dann gilt mit $h_k := \omega ||F(x^k)||$ und $\lambda \in [0, \min(1, 2/h_k)]$:

$$||F(x^k + \lambda \Delta x^k)||_2 \leq t_k(\lambda) ||F(x^k)||_2,$$

mit $t_k(\lambda) := 1 - \lambda + \frac{1}{2}\lambda^2 h_k$. Die optimale Wahl des Dämpfungsfaktor im Sinne dieser lokaler Abschätzung ist

$$\bar{\lambda}_k := \min(1, 1/h_k).$$

Globale Konvergenz

Zusätzlich zu den vorherigen Vorraussetzungen sei D_0 die wegzusammenhängende Komponente von $G(x^0) := \{y \in D | T(y) \leq T(x^0)\}$ in x^0 und sei $D_0 \subseteq D$ kompakt. Weiter sei die Jacobimatrix $F'(x)$ nicht-singulär für alle $x \in D_0$.

Globale Konvergenz

Zusätzlich zu den vorherigen Vorraussetzungen sei D_0 die wegzusammenhängende Komponente von $G(x^0) := \{y \in D | T(y) \leq T(x^0)\}$ in x^0 und sei $D_0 \subseteq D$ kompakt. Weiter sei die Jacobimatrix $F'(x)$ nicht-singulär für alle $x \in D_0$.
Dann konvergiert die gedämpfte Newtoniteration $(k = 0, 1, ...)$ mit Dämpfungsfaktoren im Bereich von

$$\lambda_k \in [\epsilon, 2\bar{\lambda}_k - \epsilon]$$

und von D_0 abhängenden genügend kleinen $\epsilon > 0$ gegen eine Lösung x^*.

Adaptive trust region strategy

Die h_k sind wegen der Lipschitzkonstante ω nicht verfügbar.

Adaptive trust region strategy

Die h_k sind wegen der Lipschitzkonstante ω nicht verfügbar.
Idee: Ersetze ω durch Schätzung $[\omega]$ und h_k durch
$[h_k] = [\omega]||F(x^k)||$.

Adaptive trust region strategy

Die h_k sind wegen der Lipschitzkonstante ω nicht verfügbar.
Idee: Ersetze ω durch Schätzung $[\omega]$ und h_k durch $[h_k] = [\omega]||F(x^k)||$.

$$\Rightarrow [\omega] \leq \omega, [h_k] \leq h_k$$

Adaptive trust region strategy

Die h_k sind wegen der Lipschitzkonstante ω nicht verfügbar.
Idee: Ersetze ω durch Schätzung $[\omega]$ und h_k durch $[h_k] = [\omega]||F(x^k)||$.

$$\Rightarrow [\omega] \leq \omega, [h_k] \leq h_k$$

Optimale Dämpfungsfaktoren:

$$[\bar{\lambda}_k] := \min(1, 1/[h_k])$$

Bit Counting Lemma

Angenommen das gedämpfte Newtonverfahren wurde mit Dämpfungsfaktoren $[\bar{\lambda}_k] := \min(1, 1/[h_k])$ durchgeführt. Weiterhin gelte für die Schätzungen:

$$0 \le h_k - [h_k] < \sigma \max(1, [h_k]) \text{ für ein } \sigma < 1.$$

Bit Counting Lemma

Angenommen das gedämpfte Newtonverfahren wurde mit Dämpfungsfaktoren $[\bar{\lambda}_k] := \min(1, 1/[h_k])$ durchgeführt. Weiterhin gelte für die Schätzungen:

$$0 \leq h_k - [h_k] < \sigma \max(1, [h_k]) \text{ für ein } \sigma < 1.$$

Dann ergibt der Residuum-Monotonietest:

$$||F(x^{k+1})|| \leq (1 - \frac{1}{2}(1 - \sigma)\lambda)||F(x^k)||.$$

Bit Counting Lemma

Angenommen das gedämpfte Newtonverfahren wurde mit Dämpfungsfaktoren $[\bar{\lambda}_k] := \min(1, 1/[h_k])$ durchgeführt. Weiterhin gelte für die Schätzungen:

$$0 \le h_k - [h_k] < \sigma \max(1, [h_k]) \text{ für ein } \sigma < 1.$$

Dann ergibt der Residuum-Monotonietest:

$$||F(x^{k+1})|| \le (1 - \frac{1}{2}(1-\sigma)\lambda)||F(x^k)||.$$

Es reicht also das führende Bit von λ_k mit $[\lambda_k]$ zu treffen, damit das Residuum monoton fällt.

geeignete Schätzer

Das gedämpfte Newtonverfahren kann als Abstieg des dazugehörigen Newtonpfades interpretiert werden.

geeignete Schätzer

Das gedämpfte Newtonverfahren kann als Abstieg des dazugehörigen Newtonpfades interpretiert werden.

$$\Rightarrow ||F(x^{k+1}) - (1-\lambda)F(x^k)|| \leq \frac{1}{2}\lambda^2\omega||F(x^k)||^2.$$

geeignete Schätzer

Das gedämpfte Newtonverfahren kann als Abstieg des dazugehörigen Newtonpfades interpretiert werden.

$$\Rightarrow ||F(x^{k+1}) - (1-\lambda)F(x^k)|| \leq \frac{1}{2}\lambda^2\omega||F(x^k)||^2.$$

$$\Rightarrow [h_k] := \frac{2||F(x^{k+1}) - (1-\lambda)F(x^k)||}{\lambda^2||F(x^k)||} \leq h_k.$$

Prädiktor- und Korrektorstrategie

Wir brauchen mindestens einen Versuchswert

$$x^{k+1} = x^k + \lambda_k^0 \Delta x^k.$$

Prädiktor- und Korrektorstrategie

Wir brauchen mindestens einen Versuchswert

$$x^{k+1} = x^k + \lambda_k^0 \Delta x^k.$$

Korrektorstrategie: $\lambda_k^{i+1} := \min(\frac{1}{2}\lambda_k^i, 1/[h_k^{i+1}])$.

Prädiktor- und Korrektorstrategie

Wir brauchen mindestens einen Versuchswert

$$x^{k+1} = x^k + \lambda_k^0 \Delta x^k.$$

Korrektorstrategie: $\lambda_k^{i+1} := \min(\frac{1}{2}\lambda_k^i, 1/[h_k^{i+1}])$.
Wir brauchen noch ein ein λ_k^0:

$$h_{k+1} = \frac{||F(x^{k+1})||}{||F(x^k)||} h_k \Rightarrow [h_{k+1}^0] = \frac{||F(x^{k+1})||}{||F(x^k)||}[h_k^{i_*}] < [h_k^{i_*}].$$

Prädiktor- und Korrektorstrategie

Wir brauchen mindestens einen Versuchswert

$$x^{k+1} = x^k + \lambda_k^0 \Delta x^k.$$

Korrektorstrategie: $\lambda_k^{i+1} := \min(\frac{1}{2}\lambda_k^i, 1/[h_k^{i+1}])$.
Wir brauchen noch ein ein λ_k^0:

$$h_{k+1} = \frac{||F(x^{k+1})||}{||F(x^k)||} h_k \Rightarrow [h_{k+1}^0] = \frac{||F(x^{k+1})||}{||F(x^k)||} [h_k^{i_*}] < [h_k^{i_*}].$$

Prädiktorstrategie: $\lambda_{k+1}^0 := \min(1, 1/[h_{k+1}^0])$.

Quasi-Newton-Zwischenschritte

Falls $\lambda_k = 1$ und der Residuum Monotonietest ergibt:

$$\Theta_= \frac{||F(x^{k+1})||}{||F(x^k)||} \leq \Theta_{\max} < 1$$

Dann ist das Residuumbasierte Quasi-Newtonverfahren anwendbar. Das heisst, Die Jacobimatrixauswertung ist duch ein Residuum rank-1 update ersetzt.

Inexaktes Newton-RES-Verfahren

Inexaktes globales Newton-Verfahren:

$$x^{k+1} = x^k + \lambda_k \delta x^k, \quad 0 < \lambda_k \leq 1$$

mit GMRES als iterativen linearen Gleichungslöser:

$$F'(x^k)\delta x^k = -F(x^k) + r^k.$$

Konvergenzanalyse

Vorraussetzungen sind wie bei der exakten Variante.

Konvergenzanalyse

Vorraussetzungen sind wie bei der exakten Variante.
Die inexakte Newton-GMRES Iteration erfüllt:

$$\Theta_k := \frac{||F(x^{k+1})||}{||F(x^k)||} \leq t_k(\lambda_k, \mu_k)$$

mit

$$t_k(\lambda, \mu) = 1 - (1-\mu)\lambda + \frac{1}{2}\lambda^2(1-\mu^2)h_k, \mu_k = \frac{||r^k||_2}{||F(x^k)||_2} < 1.$$

Die optimale Wahl des Dämpfungsfaktor ist:

$$\bar{\lambda_k} := \min(1, \frac{1}{(1+\mu_k)h_k}).$$

Monotonietest

Dämpfungsfaktoren:

$$[\bar{\lambda}_k] := \min(1, \frac{1}{(1+\mu_k)[h_k]})$$

Monotonietest

Dämpfungsfaktoren:

$$[\bar{\lambda_k}] := \min(1, \frac{1}{(1+\mu_k)[h_k]})$$

Beschränkter Residuum Monotonietest:

$$||F(x^{k+1})||_2 \leq (1 - \frac{1-\mu_k}{4}\lambda_k)||F(x^k)||_2$$

Adaptive trust region Methode

A-posteriori Schätzer:

$$[h_k](\lambda) := \frac{2||F(x^{k+1}(\lambda)) - (1-\lambda)F(x^k) - \lambda r^k||_2}{\lambda^2(1-\mu_k^2)||F(x^k)||_2} \leq h_k$$

Adaptive trust region Methode

A-posteriori Schätzer:

$$[h_k](\lambda) := \frac{2||F(x^{k+1}(\lambda)) - (1-\lambda)F(x^k) - \lambda r^k||_2}{\lambda^2(1-\mu_k^2)||F(x^k)||_2} \leq h_k$$

Korrekturstrategie (i=0,1,..,i_k^*):

$$\lambda_k^{i+1} = \min(\frac{1}{2}\lambda_k^i, \frac{1}{(1+\mu_k)[h_k^i]})$$

Adaptive trust region Methode

A-priori Schätzer:

$$[h_{k+1}^0] := \Theta_k [h_k^{i_*}] \leq h_{k+1}$$

Adaptive trust region Methode

A-priori Schätzer:

$$[h^0_{k+1}] := \Theta_k [h_k^{i_*}] \le h_{k+1}$$

Prädiktorstrategie:

$$\lambda^0_{k+1} := \min(1, \frac{1}{(1+\mu_{k+1})[h^0_{k+1}]})$$

Schluss

Vielen Dank für Ihre Aufmerksamkeit.