06 Newtonartige Verfahren, inexaktes Newton-Verfahren

Nichtlineare Optimierung

WS 2020/21

Quizfrage

$\nabla^2 f(x_k)\, d_k = -\nabla f(x_k)$
Newton-Gleichung / -Richtung

Was ist ein Nachteil des Newton-Verfahrens?

→ Umfrage
zumindest Matrix-Vektor-Produkte

5 **A** erfordert Bestimmung der Hessematrix $\nabla^2 f(x_k)$

0 **B** langsame Konvergenz

2 **C** Aufwand bei der Lösung von $\nabla^2 f(x_k)\, d_k = -\nabla f(x_k)$

0 **D** d_k evtl. keine Abstiegsrichtung

Falls $\nabla^2 f(x_k) > 0$, dann ist d_k Abstiegsrichtung.

Quizfrage

$$H_k d_k = -\nabla f(x_k) + r_k$$

Residuum

Welche Freiheiten hat man bei newtonartigen Verfahren?

→ Umfrage

1 A Wahl der Modell-Hesssematrix H_k

0 B Wahl der Genauigkeit der Lösung des LGS

7 C Wahl der Modell-Hessematrix und der Genauigkeit der Lösung des LGS

→ Residuum r_k

0 D Wahl der Dimension der Optimierungsvariablen

(Multilevel-Optimierungsverfahren)

Quizfrage

Wann spricht man von einem newtonartigen Verfahren?

→ Umfrage

4 A wenn man q-superlineare Konvergenz erreichen kann

4 C wenn d_k ungefähr der Newton-Schritt ist

1 B wenn es weniger Iterationen benötigt als das Gradientenverfahren

0 D wenn jeder zweite Schritt ein Newton-Schritt ist

Quizfrage

$H_k d_k = -\nabla f(x_k) + \zeta_k$

Auf die Norm welcher Größe kommt es bei der Untersuchung der q-superlinearen Konvergenz bei der Bestimmung der Suchrichtungen d_k an?

$\|x_{k+1} - x^*\| \le$ Nullfolge $\|x_k - x^*\|$

→ Umfrage

1 A $\nabla^2 f(x_k)\, d_k + \nabla f(x_k)$

0 B $\nabla f(x_k)$

1 C d_k

8 D $[\nabla^2 f(x_k) - H_k]\, d_k + \zeta_k$

$\nabla f(x_k) + \nabla^2 f(x_k) d_k$ ← Wie gut erfüllt d_k die echte Newton-Gleichung

$= \nabla f(x_k) + \nabla^2 f(x_k) d_k - \nabla f(x_k) - H_k d_k + \zeta_k$

$= [\nabla^2 f(x_k) - H_k] d_k + \zeta_k$

$[\nabla^2 f(x_k) - H_k] d_k + \zeta_k$

Quizfrage

Wann spricht man von einem inexakten Newton-Verfahren?

≈ Verwendung von $H_k = \nabla^2 f(x_k)$

→ Umfrage

6 A $H_k = \nabla^2 f(x_k)$, aber Residuum ζ_k erlaubt

1 B $H_k \approx \nabla^2 f(x_k)$, kein Residuum ζ_k erlaubt

0 C $H_k \equiv M$, kein Residuum ζ_k erlaubt

4 D $H_k \approx \nabla^2 f(x_k)$ und Residuum ζ_k erlaubt

Gradienten ≠ Newton ≠ Quasi-Newton

$H_k = M$ $\quad H_k = \nabla^2 f(x_k)$ $\quad H_k$ aufdatiert

Quizfrage

Beim inexakten Newton-Verfahren löst man das Newton-System mit relativer Genauigkeit η_k:

$$\|\underbrace{\nabla^2 f(x_k)\, d_k + \nabla f(x_k)}_{=\zeta_k}\|_{M^{-1}} \leqslant \eta_k \, \|\nabla f(x_k)\|_{M^{-1}}$$

Welche Konvergenzrate kann man bei $\eta_k \leqslant \overline{\eta} < 1$ erwarten?

(Satz 5.30)

→ Umfrage

wie Gradientenverfahren

0 A keine Konvergenz

6 B q-linear

4 C q-superlinear

0 D q-quadratisch

Quizfrage

forcing sequence

Beim inexakten Newton-Verfahren löst man das Newton-System mit relativer Genauigkeit η_k:

$$\|\underbrace{\nabla^2 f(x_k)\, d_k + \nabla f(x_k)}_{=\zeta_k}\|_{M^{-1}} \leqslant \eta_k \, \|\nabla f(x_k)\|_{M^{-1}}$$

Welche Konvergenzrate kann man bei $\eta_k \searrow 0$ erwarten?

$\rightarrow$ Umfrage

0 A keine Konvergenz

0 B q-linear

10 C q-superlinear

0 D q-quadratisch

Quizfrage

Praxis:
$\eta_k = \min\{\underbrace{\bar{\eta}}_{\text{z.B. } 1/2}, \|\nabla f(x_k)\|_{M^{-1}}^{1/2}\}$ Rate: $1+\frac{1}{2}$

Beim inexakten Newton-Verfahren löst man das Newton-System mit relativer Genauigkeit η_k:

$$\|\underbrace{\nabla^2 f(x_k)\, d_k + \nabla f(x_k)}_{=\zeta_k}\|_{M^{-1}} \leqslant \eta_k \,\|\nabla f(x_k)\|_{M^{-1}}$$

Welche Konvergenzrate kann man bei $\eta_k \leqslant c\,\|\nabla f(x_k)\|_{M^{-1}}$ erwarten?

$\rightarrow$ Umfrage

0 A keine Konvergenz

0 B q-linear

0 C q-superlinear

10 D q-quadratisch

Quizfrage

Falls $\nabla^2 f(x_k) > 0$ ist, dann ist das Lösen des LGS äquivalent zu
Minimiere $\frac{1}{2} d^T \nabla^2 f(x_k) d + \nabla f(x_k)^T d$

Mit welchem Lösungsverfahren kann man die erlaubte Inexaktheit bei der Lösung des Gleichungssystems $\nabla^2 f(x_k)\, d_k = -\nabla f(x_k)$ gut ausnutzen?

$\rightarrow$ Umfrage

0 **A** mit dem Gauss-Algorithmus

1 **B** mit dem cg-Verfahren

0 **C** mit dem Gradientenverfahren

9 **D** mit einem modifizierten cg-Verfahren

„abgeschnitten"

Das truncated-cg-Verfahren

$\nabla^2 f(x_k)$ $\qquad -\nabla f(x_k)$

für $A d = b$

mit A symmetrisch

Setze $\ell := 0$
Setze $d_0 := 0$
Setze $r_0 := -b$ $\qquad r_0 = \nabla f(x_k)$
Setze $p_0 := -M^{-1} r_0$
Setze $\delta_0 := -r_0^\top p_0$ $\qquad // \ \delta_0 = \|r_0\|^2_{M^{-1}}$
while Abbruchkriterium nicht erfüllt **do**
  Setze $q_\ell := A\, p_\ell$
  Setze $\kappa_\ell := p_\ell^\top q_\ell$
  if $\kappa_\ell > 0$ **then** $\qquad \kappa_\ell = p_\ell^T A p_\ell$
    Setze $\alpha_\ell := \delta_\ell / \kappa_\ell$
    Setze $d_{\ell+1} := d_\ell + \alpha_\ell\, p_\ell$
    Setze $r_{\ell+1} := r_\ell + \alpha_\ell\, q_\ell$
    Setze $p_{\ell+1} := -M^{-1} r_{\ell+1}$
    Setze $\delta_{\ell+1} := -r_{\ell+1}^\top p_{\ell+1}$ $\qquad // \ \delta_{\ell+1} = \|r_{\ell+1}\|^2_{M^{-1}}$
    Setze $\beta_{\ell+1} := \delta_{\ell+1} / \delta_\ell$
    Setze $p_{\ell+1} := p_{\ell+1} + \beta_{\ell+1}\, p_\ell$
    Setze $\ell := \ell + 1$
  else
    Abbruch der **while**-Schleife
  end if
end while
return d_ℓ

Alg. 4.2 b

Iterierte d_ℓ (statt x_k)

Richtungen p_ℓ (statt d_k)

Residuen r_ℓ (statt r_k)

$d_1 = \cancel{d_0}^{\,0} + \alpha_0 p_0$

$= -\alpha_0 M^{-1} \nabla f(x_k)$

$\sim -M^{-1} \nabla f(x_k)$

Zeit für Ihre Fragen

Was sind Ihre Fragen zu den Themen der Woche?

$\rightarrow$ Benutzen Sie den Chat.

Fragen und Antworten 1

Fragen und Antworten 2