

03 Freie Optimierung:
Verfahren hinter `fminunc`
Optimierung für Nichtmathematiker
WS 2020/21

Quizfrage

Was bedeutet es, wenn ein Optimierungsverfahren „iterativ“ arbeitet?

→ Umfrage *Iteration $x^{(k)}$*

A Die Lösung wird nur ungefähr berechnet.

B Manchmal wird keine Lösung gefunden.

C Die Lösung wird in mehreren Schritten ermittelt.

D Der Benutzer muss die Lösung überprüfen.

Quizfrage

Was passiert in einer Iteration von `fminunc` nicht?

→ Umfrage
↓ quadratisches

- g A Modell aufstellen und Modell minimieren
- 5 C Schrittlänge wählen

$x^{(k)} \rightarrow x^{(k+1)}$

- 2 B Iterierte aktualisieren
- g **D** Zielfunktion minimieren
stattdessen wird ein Modell der Zielfunktion minimiert

Quizfrage

quadratisch

linear (unbeschränkt)

Warum verwendet man gerade quadratische Polynome als Modelle in fminunc?

→ *Umfrage*

A lineare Modelle würden auch gehen

B quadratische Modelle sind die einfachsten, die sinnvoll sind

C quadratische Polynome sind in Matlab einfach

D quadratische Modelle sind oft exakt

genau dann, wenn Zielfunktion selbst quad. Polynom ist

Quizfrage

Warum verwendet man gerade quadratische Polynome als Modelle in `fminunc`?

→ Umfrage

- A lineare Modelle würden auch gehen
- B quadratische Modelle sind die einfachsten, die sinnvoll sind
- C quadratische Polynome sind in Matlab einfach
- D quadratische Modelle sind oft exakt

Quizfrage

Was wird mit dem quadratischen Modell in jeder Iteration gemacht?

→ Umfrage

- A es wird gezeichnet
 C es wird dem Benutzer mitgeteilt

leider nicht

die Ableitung

- B es wird nullgesetzt
 D es wird an Stelle der Zielfunktion minimiert



Quizfrage $\nabla f(x^{(k)}) = \begin{bmatrix} \\ \\ \end{bmatrix}$ $d = \begin{bmatrix} \\ \\ \end{bmatrix}$

Welche Gestalt hat das quadratische Modell für die Zielfunktion $f(x)$ an der Stelle $x^{(k)}$?

→ Umfrage

daß von $\nabla^2 f(x^{(k)})$ abweichen
aufwärts

- 18 A $f(x^{(k)}) + \nabla f(x^{(k)})^T d + \frac{1}{2} d^T B^{(k)} d$ 2 B $f(x^{(k)}) + \nabla f(x^{(k)})^T d$
3 C $f(x^{(k)}) + \nabla f(x^{(k)})^T d + \frac{1}{2} \|d\|^2$ 1 D $f(x^{(k)})$

$B^{(k)}$ über Quasi-Newton-Update
(BFGS)

Quizfrage

Welche Informationen über die Zielfunktion werden zum Aufstellen des quadratischen Modells an der Stelle $x^{(k)}$ benötigt?

→ Umfrage

A $f(x^{(k)})$

C $\nabla f(x^{(k)})$

B $f(x^{(k)})$ und $\nabla f(x^{(k)})$

D $f(x^{(k)})$, $\nabla f(x^{(k)})$ und $\nabla^2 f(x^{(k)})$

Quizfrage

$B^{(k)}$ klein (positiv)

$B^{(k)}$ groß (positiv)

Welche Bedeutung hat die Quasi-Newton-Matrix $B^{(k)}$ im Modell?

→ Umfrage

wenn $B^{(k)} = \nabla^2 f(x^{(k)})$,
dann Newton-Verf.

1 **A** bestimmt die Krümmung des Modells

3 **B** muss mit der Hessematrix $\nabla^2 f(x^{(k)})$ übereinstimmen

2 **C** ersetzt die echte Hessematrix $\nabla^2 f(x^{(k)})$

0 **D** definiert das Optimierungsproblem

Quizfrage

Was bedeutet es, dass in jedem Schritt von `fminunc` ein quadratisches Modell minimiert wird?

→ Umfrage

3 A `fminunc` ruft sich selbst auf

8 C Lösen eines linearen Gleichungssystems


6 B Liniensuche

8 D Funktionswerte durchsuchen

Quizfrage



$B^{(k)} > 0$



$B^{(k)} < 0$ kein Min.
da unbeschränkt

Warum wird die Quasi-Newton-Matrix $B^{(k)}$ in fminunc positiv definit gehalten?
alle Eigenwerte > 0

→ Umfrage

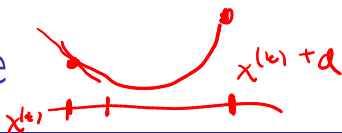
1 A weil die Werte der Zielfunktion oft auch positiv sind

0 C weil diese einfacher sind als negativ definite Matrizen

16 **B** ^{quadr.} damit das Modell ein eindeutiges Minimum besitzt

10 D damit man ein Minimum der Zielfunktion $f(x)$ findet

Quizfrage



Wann nennt man eine Richtung d eine Abstiegsrichtung für die Funktion $f(x)$ an der Stelle $x^{(k)}$?

→ Umfrage

14 A $f(x^{(k)} + d) < f(x^{(k)})$

0 C $d < 0$

$f(x^{(k)} + \alpha d) < f(x^{(k)})$
für kleine $\alpha > 0$

Abstiegsrichtungen

4 B $\nabla f(x^{(k)})^T d > 0$

9 D $\nabla f(x^{(k)})^T d < 0$

Innenprodukt z.B.
Grad. und Richtung
ist negativ

$$q^{(k)}(\underline{d}) = f(x^{(k)}) + \nabla f(x^{(k)})^T \underline{d} + \frac{1}{2} \underline{d}^T B^{(k)} \underline{d}$$

$$\nabla q^{(k)} = 0 + \nabla f(x^{(k)}) + B^{(k)} \underline{d} \stackrel{!}{=} 0$$

$$\Leftrightarrow B^{(k)} \underline{d} = -\nabla f(x^{(k)})$$

$\square \quad \square = \square$ lineares Gleichungssystem
 \uparrow Unbekannte

Falls $B^{(k)}$ positiv definit ist (wie bei BFGS), dann ist das LGS eindeutig lösbar und $\nabla f(x^{(k)})^T \underline{d} = -\underline{d}^T B^{(k)} \underline{d} < 0$ (Abstiegsrichtung!)

Quizfrage

$$x^{(k+1)} = x^{(k)} + \underbrace{\alpha^{(k)}}_{\text{bestimmt } d^{(k)}} d^{(k)}$$

Wozu dient die Liniensuche in fminunc?

→ Umfrage

- 10 A zur Bestimmung einer geeigneten Schrittlänge, mit der die Abstiegsrichtung **multipliziert** wird
- 4 B zum Testen, ob die Richtung eine Abstiegsrichtung ist
- 1 C zur Bestimmung einer geeigneten Schrittlänge, die zur Abstiegsrichtung addiert wird
- 0 D zum Testen, ob man die Lösung gefunden hat

Matlab-Demonstration

Hausbau.m

Quizfrage

Warum gehen die Funktionswerte der Zielfunktion $f(x)$ in jeder Iteration herunter?

→ Umfrage

- A das ist hier Zufall
- B weil f_{\min} ein Abstiegsverfahren ist
- C weil $f(x)$ minimiert werden soll
- D weil das Verfahren $\nabla f(x^{(k)})$ verwendet

Quizfrage

Was bedeutet die Ausgabe
First-order optimality?

→ Umfrage

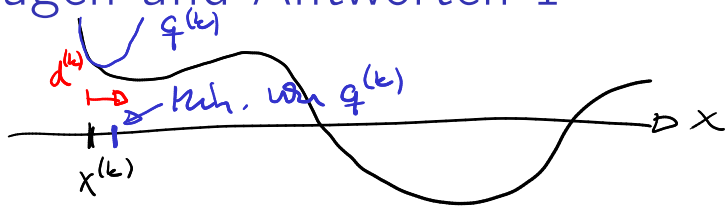
- A Wert der Zielfunktion $f(x^{(k)})$
- B Wert des Gradienten $\nabla f(x^{(k)})$
- C Norm des Gradienten $\|\nabla f(x^{(k)})\|$
- D $x^{(k)}$ ist ungefähr optimal

Zeit für Ihre Fragen

Was sind Ihre Fragen zu den Themen der Woche?

→ Benutzen Sie den **Chat**.

Fragen und Antworten 1



$d^{(k)}$ Richtung

$$x^{(k+1)} = x^{(k)} + \alpha^{(k)} d^{(k)}$$

↑ Schrittweite

Fragen und Antworten 2

- An jeder Iteration wird $f(x^{(k)})$ und $\nabla f(x^{(k)})$ benötigt!

↑ entweder als Benutzer programmiert
oder sonst: Finite-Differenzen-
Approximation (nä. Woche)

- $\boxed{a^T b}$

$= a \cdot b$

= Innenprod.

$= (a, b)$

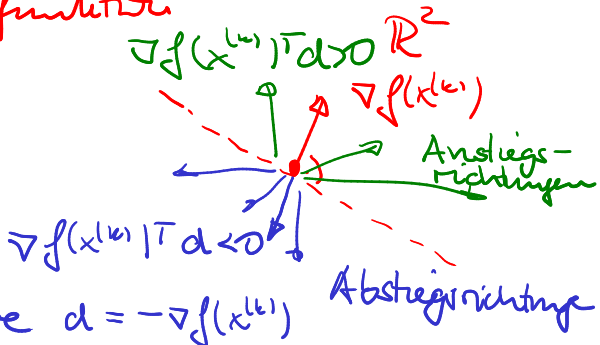
$a = \begin{bmatrix} \\ \\ \end{bmatrix} \quad b = \begin{bmatrix} \\ \\ \end{bmatrix} \text{ in } \mathbb{R}^n$

z.B. $\nabla f(x^{(k)})^T d$

$a^T b = \begin{bmatrix} \bullet & \bullet & \bullet & \bullet \end{bmatrix} \begin{bmatrix} \bullet \\ \bullet \\ \bullet \\ \bullet \end{bmatrix} \quad n \times 1$
 $\sum a_i b_i$

Fragen und Antworten 3

$\nabla f(x^{(k)})$ ist Richtung des steilsten Anstiegs der Zielfunktion



insbesondere $d = -\nabla f(x^{(k)})$