02 Minimierung quadratischer Polynome, Gradientenverfahren

Nichtlineare Optimierung

WS 2020/21

Quizfrage

A immer symmetrisch, sonst $A \rightsquigarrow \frac{1}{2}(A+A^T)$

Wie ist die Lösungsstruktur dieser Optimierungsaufgabe?

$$\text{Minimiere} \quad \frac{1}{2}x^\top \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}}_{A \text{ ist spd}} x - \begin{bmatrix} -7 \\ 3 \end{bmatrix}^\top x$$

gleichmäßig konvex

$\boxed{Ax = b}$

→ Umfrage

Lemma 4.1

8 A eine globale Lösung, keine weitere lokale

3 B mehrere globale Lösungen, keine weiteren lokalen

↯ strikte Konvexität

3 C eine globale Lösung und weitere lokale

↯ wegen Konvexität

0 D weder lokale noch globale Lösung

Quizfrage

$\lambda_1(A) = 1, \ \lambda_2(A) = 0$ konvex

quadr. Teil der Zielfunktion ist in Richtung $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ flach

Wie ist die Lösungsstruktur dieser Optimierungsaufgabe?

$$\text{Minimiere} \quad \frac{1}{2} x^\top \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}}_{A \text{ spsd}} x - \underbrace{\begin{bmatrix} -7 \\ 3 \end{bmatrix}}_{b}^\top x = \phi(x)$$

$\boxed{Ax = b \text{ nicht lösbar}}$

→ Umfrage

$$\psi(\alpha) = \phi\left(\alpha \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) = 0 - \alpha \begin{pmatrix} -7 \\ 3 \end{pmatrix}^\top \begin{pmatrix} 0 \\ 1 \end{pmatrix} = -3\alpha$$

unbeschränkt

5 A eine globale Lösung, keine weitere lokale

8 B mehrere globale Lösungen, keine weiteren lokalen

1 C eine globale Lösung und weitere lokale

0 D weder lokale noch globale Lösung

Quizfrage

$\lambda_1(A) = 1, \quad \lambda_2(A) = -2$

nicht konvexer Fall

in Richtung $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$:

Wie ist die Lösungsstruktur dieser Optimierungsaufgabe?

unbeschränkt

$$\text{Minimiere} \quad \frac{1}{2} x^\top \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix}}_{A} x - \begin{bmatrix} -7 \\ 3 \end{bmatrix}^\top x$$

$Ax=b$ lösbar (Sattelpunkt)

→ Umfrage

0 A eine globale Lösung, keine weitere lokale

0 B mehrere globale Lösungen, keine weiteren lokalen

1 C eine globale Lösung und weitere lokale

13 D weder lokale noch globale Lösung

Quizfrage

Wenn A spd ist, dann ist die Minimierung von $\frac{1}{2}x^\mathsf{T}A\,x - b^\mathsf{T}x + c$ äquivalent zur Lösung des LGS $\boxed{A\,x = b.}$
Warum sollte man überhaupt iterative Optimierungsverfahren dafür betrachten?

→ Umfrage

13 A für hochdimensionale Aufgaben

0 B wenn A nur als Matrix-Vektor-Operation vorliegt

1 C wenn Näherungslösungen ausreichen

0 D um z. B. fehlende positive Definitheit festzustellen

Gauß/Cholesky Aufwand $\sim n^3$

Residuum $r = Ax - b$

Oft sind quadr. Polynome Hilfsaufgaben bei der Lösung nichtlinearer Opt. aufgaben. Dann reicht es i.d.R. aus, inexakt zu lösen, also ein Residuum zu erlauben:

$\|r\|_{M^{-1}} \le \varepsilon$

$A = vv^T$
$Ax = v(v^T x)$

notfalls: $A = [(Ae_1), (Ae_2), \dots, (Ae_n)]$

weitere Beispiele: • fft • Mehrgitterverfahren

A liegt nur als Matrix-Vektor-Produkt vor:

$x \mapsto v$
linear

```
function v = matvec(x)
v(1) = 2*x(1) - 5*x(2);
v(2) = -x(1) + x(2);
```

$x = []$

$A = \begin{bmatrix} 2 & -5 \\ -1 & 1 \end{bmatrix}$

von A und M

Feststellen fehlender pos. Definitheit

im Gradientenverf.

Setze $k := 0$

Setze $r_0 := \boxed{A\, x_0} - b$

Setze $d_0 := -\boxed{M^{-1} r_0}$

Setze $\delta_0 := -r_0^T d_0$ // $\delta_0 = \|\nabla_M \phi(x_0)\|_M^2$

while Abbruchkriterium nicht erfüllt **do**

  Setze $q_k := \boxed{A\, d_k}$ $d_k^T q_k = d_k^T A d_k \le 0$ ↯

  Setze $\alpha_k := \delta_k / (d_k^T q_k)$

  Setze $x_{k+1} := x_k + \alpha_k\, d_k$

  Setze $r_{k+1} := r_k + \alpha_k\, q_k$

  Setze $d_{k+1} := -\boxed{M^{-1} r_{k+1}}$

  Setze $\delta_{k+1} := -r_{k+1}^T d_{k+1}$ // $\delta_{k+1} = \|\nabla_M \phi(x_{k+1})\|_M^2$

  Setze $k := k + 1$

end while

return x_k

$\delta_{k+1} = r_{k+1}^T M^{-1} r_{k+1} \le 0$ ↯

$= \|r_{k+1}\|^2_{M^{-1}}$

Quizfrage

Warum hängt der Gradient einer Funktion $f : \mathbb{R}^n \to \mathbb{R}$ vom verwendeten Skalarprodukt in $\mathbb{R}^n$ ab?

M-Skalarprodukt

Richtung des st. Anstiegs

$\to$ Umfrage

$$R(d) = \frac{f'(x)d}{\|d\|_M} \longrightarrow \max$$

0 A das stimmt gar nicht

8 C weil dazu die M- Richtungsableitung über die Einheitskugel minimiert wird und diese vom SP abhängt

0 B weil f vom Skalarprodukt abhängt

5 D weil die Ableitung von f vom Skalarprodukt abhängt

Quizfrage

$\nabla f(x) = $ Eukl. Gradient
$= f'(x)^T$

Wie bestimmt sich der Gradient $\nabla_M f(x)$ bzgl. des M-Skalarprodukts?

$\rightarrow$ Umfrage

LGS mit M lösen!

0 A $\nabla_M f(x) = \nabla f(x)$

11 B $\nabla_M f(x) = M^{-1} \nabla f(x)$

0 C $\nabla_M f(x) = M \nabla f(x)$

1 D $\nabla_M f(x) = M f'(x)$

- M heißt auch Vorkonditionierer
- „M anwenden" heißt Lösen eines LGS mit M als Koeffizientenmatrix $\Rightarrow M^{-1} v$ ausrechnen

Kantarovich-Ungleichung

Es sei

- $Q \in \mathbb{R}^{n \times n}$ spd
- $\alpha := \lambda_{\min}(Q)$
- $\beta := \lambda_{\max}(Q)$

$$\frac{x^\top Q x}{\|x\|^2} \le \beta \qquad \frac{y^\top Q^{-1} y}{\|y\|^2} \le \frac{1}{\alpha}$$

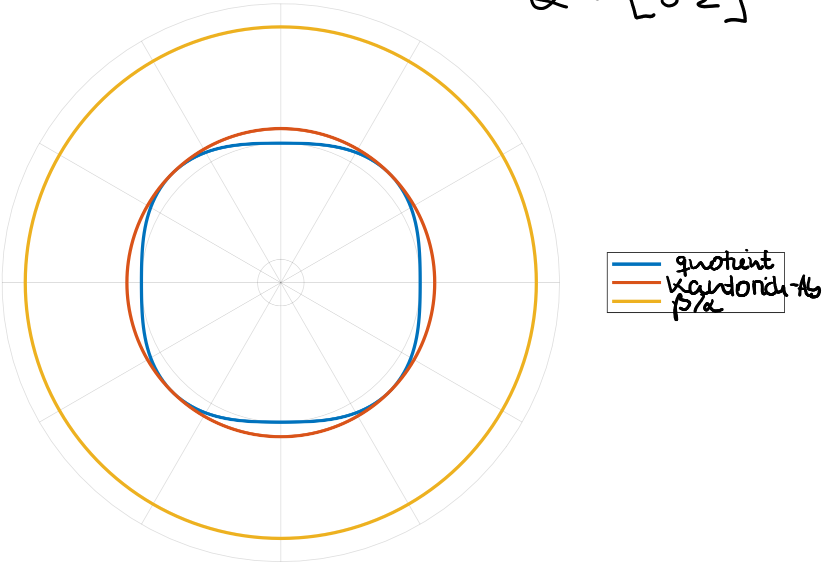
Dann gilt

$$1 \leqslant \boxed{\frac{(x^\top Q\, x)\,(x^\top Q^{-1} x)}{\|x\|^4}} \leqslant \boxed{\frac{(\alpha + \beta)^2}{4\,\alpha\,\beta}} \leqslant \frac{\beta}{\alpha}$$

für alle $x \in \mathbb{R}^n$, $x \neq 0$.

Visualisierung der Ungleichung

$$Q = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$$



Verallg. Kantarovich-Ungleichung

Es sei

- $Q \in \mathbb{R}^{n\times n}$ spd und $M \in \mathbb{R}^{n\times n}$ spd
- $\alpha := \lambda_{\min}(Q; M)$ $\qquad Qv = \lambda M v$
- $\beta := \lambda_{\max}(Q; M)$

Dann gilt

$$1 \leqslant \boxed{\frac{(x^\top Q\, x)\,(x^\top M Q^{-1} M\, x)}{\|x\|_M^4}} \leqslant \boxed{\frac{(\alpha+\beta)^2}{4\,\alpha\,\beta}} \leqslant \frac{\beta}{\alpha}$$

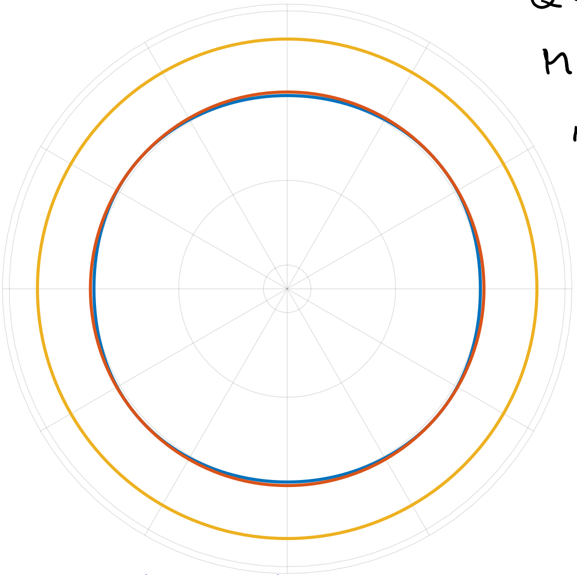
für alle $x \in \mathbb{R}^n$, $x \neq 0$.

Visualisierung der Ungleichung

$Q = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$

$M = \begin{bmatrix} 1 & 0 \\ 0 & 3/2 \end{bmatrix}$

„näher an Q als I"



Quizfrage

Durch welche Größe(n) kann man die Konvergenzgeschwindigkeit des Gradientenverfahrens bzgl. des M-Skalarprodukts für die Minimierung von $\frac{1}{2}x^\mathsf{T} A\, x - b^\mathsf{T} x + c$ abschätzen? (Satz 4.9)

$\rightarrow$ Umfrage

A $\lambda_{\max}(A)$

B $\lambda_{\min}(A; M)$

C $\lambda_{\min}(A)$

D $\lambda_{\max}(A; M)$

E $\lambda_{\max}(A; M)$ und $\lambda_{\min}(A; M)$

Konditionszahl

$\kappa = \dfrac{\lambda_{\max}(A;M)}{\lambda_{\min}(A;M)}$

„Kontrastverhältnis"

$\dfrac{\kappa-1}{\kappa+1}$ Kontraktionsfaktor

Demonstration Gradientenverfahren

`test_steepest_descent_quadratic.m`

im OPAL

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