Einführung in die Grundlagen der Numerik

Wintersemester 2018/19
Prof. Dr. Ira Neitzel
Fabian Hoppe

UNIVERSITÄT BONN

Übungsblatt 4.

Abgabe am **8. November vor der Vorlesung**.

Aufgabe 1. (Pseudoinverse)

Sei $A \in \mathbb{R}^{m\times n}$ und $b \in \mathbb{R}^m$. Über den Rang von A wird keine Annahme gemacht. In der Vorlesung wurde die Pseudoinverse A^+ von A durch folgende vier Axiome eingeführt:

(i) $(A^+A)^T = A^+A$

(ii) $(AA^+)^T = AA^+$

(iii) $A^+AA^+ = A^+$

(iv) $AA^+A = A$.

Insbesondere wurde gezeigt, dass A^+ durch die Axiome (i)-(iv) bereits eindeutig bestimmt ist und dass $A^+ = \hat{U}\hat{\Sigma}^{-1}\hat{V}^T$ gilt. Hierbeit ist $A = \hat{U}\hat{\Sigma}\hat{V}^T$ die reduzierte Singulärwertzerlegung aus der Vorlesung, d.h. $\hat{U} \in \mathbb{R}^{m\times r}, \hat{V} \in \mathbb{R}^{n\times r}$ haben orthogonale Spalten und $\hat{\Sigma} \in \mathbb{R}^{r\times r}$, $r = \operatorname{rank}(A)$, ist eine Diagonalmatrix mit den Singulärwerten von A als Diagonalelemente.
Weiterhin ist bereits bekannt, dass

$$A^+b = \underset{x\in S}{\operatorname{argmin}} \|x\|_2^2 \qquad \text{mit } S := \{x \in \mathbb{R}^n : A^TAx = A^Tb\}$$

gilt, d.h. die Pseudoinverse bestimmt diejenige Lösung der Normalengleichung mit minimaler euklidischer Norm. In dieser Aufgabe lernen wir eine weitere Interpretation der Pseudoinversen kennen. Dazu definieren wir für $\lambda \geq 0$

$$J_\lambda(x) := \frac{1}{2}\|Ax - b\|_2^2 + \frac{\lambda}{2}\|x\|_2^2.$$

a. Zeigen Sie, dass für $\lambda > 0$ durch $x_\lambda = (A^TA + \lambda I)^{-1}A^Tb$ das eindeutige, strikte globale Minimum von J_λ auf $\mathbb{R}^n$ gegeben ist.

b. Zeigen Sie, dass gilt:
$$x_\lambda \to A^+b \qquad \text{für } \lambda \searrow 0.$$

c. Sei $x_0 \in \mathbb{R}^n$ gegeben. Geben Sie die erste Iterierte des Gradientenverfahrens zur Minimiernung von J_λ mit Startwert x_0 und exakter Schrittweite an.

(3+3+2 = 8 Punkte)

Aufgabe 2. (Nichtlineare Ausgleichsprobleme)

Sei $f : \mathbb{R}^n \to \mathbb{R}^m$ zweimal stetig differenzierbar mit $\operatorname{rank}(f'(x)) = n \leq m$ für alle $x \in \mathbb{R}^n$. Für $y \in \mathbb{R}^m$ wollen wir das nichtlineare Ausgleichsproblem

$$\min_{x \in \mathbb{R}^n} r(x), \qquad r(x) := \|y - f(x)\|_2^2$$

betrachten.

a. Bestimmen Sie den Gradienten ∇r und die Hesse-Matrix $\nabla^2 r$. Geben Sie die Newton-Iteration für die Lösung der Gleichung $\nabla r = 0$ an.

Beim sog. *Gauss-Newton-Verfahren* wird ausgehend von einer Startnäherung $x_0 \in \mathbb{R}^n$ mit der Vorschrift

$$x_{k+1} := \operatorname*{argmin}_{x \in \mathbb{R}^n} \|y - f(x_k) - f'(x_k)(x - x_k)\|_2^2 \tag{1}$$

iteriert. Unter gewissen geeigneten Bedingungen kann lokal mindestens lineare Konvergenz dieses Verfahrens gezeigt werden.

b. Geben Sie die Normalengleichung für (1) an. Wo liegt der Vorteil des Gauss-Newton-Verfahrens gegenüber dem Newton-Verfahren?

Der Zusammenhang zweier physikalischer Größen p und q lasse sich aufgrund theoretischer Überlegungen annähern durch

$$q = q(p) = \alpha \exp(\beta p)$$

mit noch zu bestimmenden Parametern α, β. Durch ein Experiment hat man die folgenden drei $(p, q)-$Wertepaare ermittelt:

$$(0, 2), (1, 5), (3, 50).$$

c. Stellen Sie ein geeignetes nichtlineares Ausgleichsproblem zur Bestimmung der Parameter α, β auf und führen Sie von den Startwerten $\alpha_0 = 2$, $\beta_0 = 1$ ausgehend einen Schritt des Gauss-Newton-Verfahrens durch.

   Hinweis: Hier darf ein Taschenrechner benutzt und sinnvoll gerundet werden.

(3+2+3 = 8 Punkte)

Aufgabe 3. (Versagen konstanter Schrittweite)

Es sei die stetig differenzierbare Funktion

$$f : \mathbb{R}^n \to \mathbb{R}, \qquad x \mapsto \frac{2}{3}\|x\|_2^{3/2},$$

gegeben. Das *Gradientenverfahren mit fester Schrittweite* $\alpha > 0$ ist durch die Vorschrift

$$x_{k+1} = x_k - \alpha \nabla f(x_k)$$

definiert. Beweisen Sie, dass das Gradientenverfahren mit fester Schrittweite $\alpha < 2$ für f entweder nach endlich vielen Schritten das Minimum $x^* = 0$ erreicht oder nicht gegen x^* konvergiert.

Hinweis: Betrachten Sie das Verhalten von $\phi(x) := x - \alpha \nabla f(x)$.

(4 Punkte)

Aufgabe 4. (Existenz effizienter Schrittweiten)

Sei $f : \mathbb{R}^n \to \mathbb{R}$ gegeben und $x_0 \in \mathbb{R}^n$, derart, dass

- $N = N(f, f(x_0)) := \{x \in \mathbb{R}^n : f(x) \leq f(x_0)\}$ kompakt ist,
- f stetig differenzierbar ist auf einer offenen Menge $U \subset \mathbb{R}^n$ mit $N \subset U$ und
- f' Lipschitz stetig auf N ist, d.h.

$$\|\nabla f(x) - \nabla f(y)\|_2 \leq L\|x - y\|_2, \qquad \forall x, y \in N,$$

  mit einem $L > 0$ gilt.

Beweisen Sie unter diesen Voraussetzungen folgendes Lemma aus der Vorlesung:

Lemma. *Sei $x \in N$, $d \in \mathbb{R}^n$ mit $\nabla f(x)^T d < 0$ sowie $\delta \in (0, 1)$ gegeben. Dann gibt es $\tau = \tau(x, d, \delta)$ mit folgenden Eigenschaften:*

(i) Für alle $\sigma \in (0, \tau)$ gilt: $f(x + \sigma d) < f(x) + \delta\sigma\nabla f(x)^T d$,

(ii) $f(x + \tau d) = f(x) + \delta\tau\nabla f(x)^T d$,

(iii) $\tau \geq \rho := -\frac{2(1-\delta)}{L}\frac{\nabla f(x)^T d}{\|d\|_2^2}$,

(iv) Für $\sigma \in [0, \rho/2]$ gilt: $\frac{\partial}{\partial\sigma} f(x + \sigma d) = \nabla f(x + \sigma d)^T d < \delta\nabla f(x)^T d$

Hinweis: Zeigen Sie erst, dass die Menge T aller τ, die (i) erfüllen nichtleer ist. Machen Sie sich dann klar, dass $\tau := \sup T$ endlich ist und alle Anforderungen erfüllt.

(6 Punkte)