Wintersemester 2019/20
Fakultät für Mathematik und Informatik
Universität Leipzig

Dr. André Uschmajew
Dr. Max Pfeffer

4. Übungsblatt
Optimierung I

Die Abgabe der Hausaufgaben erfolgt in der Übung am 21.11.2019.

Übungsaufgaben

1. **Matrixwertige Funktionen.**

   In dieser Aufgabe betrachten wir Funktionen, die den Vektorraum der Matrizen $\mathbb{R}^{n\times n}$ in sich abbilden.

   (a) Gegeben seien differenzierbare Funktionen $F, U, V : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$ und es gelte $F(X) = U(X)V(X)$. Beweisen Sie die Produktregel
   $$F'(X)[H] = U'(X)[H] \cdot V(X) + U(X) \cdot V'(X)[H].$$

   (b) Sei $F : \mathrm{GL}(n) \to \mathbb{R}^{n\times n}$, $F(X) = X^{-1}$. Nehmen Sie an, dass F differenzierbar ist (dies folgt zum Beispiel aus der Formel $X^{-1} = \frac{1}{\det(X)} \operatorname{adj}(X)$, wobei $\operatorname{adj}(X)$ die Adjunkte von X ist) und finden Sie mit Hilfe von (a) eine Formel für die Ableitung $F'(X)[H]$.

   (c) Weisen Sie direkt nach, dass die gefundene Formel die Definition der Ableitung erfüllt.

2. **Banachscher Fixpunktsatz für eine matrixwertige Funktion.**

   Gegeben sei $A \in \mathbb{R}^{n\times n}$ sowie die matrixwertige Funktion $F : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$, $X \mapsto I + AX$.

   (a) Zeigen Sie, dass die zugehörige Fixpunktiteration für F genau dann für alle Startwerte konvergiert, falls $\|A\|_2 < 1$. Zeigen Sie außerdem, dass in diesem Fall der Fixpunkt X_* durch die Neumannsche Reihe gegeben ist
   $$X_* = (I - A)^{-1} = \sum_{i=0}^{\infty} A^i.$$

   (b) Zeigen Sie mit Hilfe von Aufgabenteil (a) einen Spezialfall des Lemmas von Banach: Falls $\|A\|_2 < 1$, so gilt
   $$\|(I - A)^{-1}\|_2 \le \frac{1}{1 - \|A\|_2}.$$

3. **Richtungen des steilsten Abstiegs.**

   Wir betrachten eine stetig differenzierbare Funktion $f : \mathbb{R}^n \to \mathbb{R}$ und einen Punkt $\hat{x} \in \mathbb{R}^n$ mit $\nabla f(\hat{x}) \neq 0$. Zu der symmetrischen, positiv definiten Matrix $A \in \mathbb{R}^{n\times n}$ definieren wir durch $\|x\|_A = \sqrt{x^\mathsf{T} A x}$ eine Norm auf $\mathbb{R}^n$. Bestimmen Sie die normierte Richtung des steilsten Abstiegs von f bezüglich dieser Norm, d.h. die Lösung des Problems
   $$\min_{\|d\|_A = 1} \nabla f(\hat{x})^\mathsf{T} d.$$

Hausaufgaben

1. **Newton-Verfahren für nichtlineare Gleichungssysteme.** (5 Punkte)

   Gegeben sei das nichtlineare Gleichungssystem
   $$x_1 = 0, \qquad \frac{10x_1}{x_1 + 0.1} + 2x_2^2 = 0,$$
   dessen eindeutig bestimmte Lösung $x^* = (x_1^*, x_2^*)^\mathsf{T} = (0, 0)^\mathsf{T}$ ist.

   (a) Führen Sie ausgehend von dem Startvektor $x^0 = (0, 1)^\mathsf{T}$ einen Newton-Schritt durch.

   (b) Konvergiert das Newton-Verfahren in diesem Beispiel für alle genügend nahe an der Lösung liegenden Startvektoren? Geben Sie die maximale Teilmenge aller Punkte in der offenen Kugel um den Ursprung mit dem Radius 1/10, d.h.
   $$B_{\frac{1}{10}}(0) = \left\{ x \in \mathbb{R}^2 : \|x\| < \frac{1}{10} \right\},$$
   an, sodass das Verfahren für alle Startwerte in dieser Menge gegen x^* konvergiert.

2. **Newton-Schulz-Verfahren.** (7 Punkte)

   Sei $A \in \mathrm{GL}(n) \subset \mathbb{R}^{n \times n}$ eine invertierbare Matrix. Gegeben sei die Funktion
   $$F : \mathrm{GL}(n) \to \mathbb{R}^{n \times n}, \qquad X \mapsto X^{-1} - A.$$

   (a) Finden Sie die Nullstelle(n) von f.

   (b) Wir wissen aus der Übung, dass
   $$F'(X)[H] = -X^{-1} H X^{-1}$$
   gilt. Berechnen Sie die inverse Abbildung $F'(X)^{-1}$.

   (c) Geben Sie das Newton-Verfahren zur Bestimmung einer Nullstelle von F an und geben Sie den Fixpunkt dieser Iterationsvorschrift an.

   (d) Begründen Sie, warum die Iteration lokal quadratisch konvergiert.

3. **Koordinaten- und Metrikwechsel.** (8 Punkte)

   Sei $f : \mathbb{R}^n \to \mathbb{R}$ zweimal stetig differenzierbar und $B \in \mathbb{R}^{n \times n}$ invertierbar.

   (a) Durch einen Koordinatenwechsel $y = Bx$ sei die Funktion $g(y) = f(x)$ gegeben. Führen Sie jeweils vom Punkt $x^0 \in \mathbb{R}^n$ resp. $y^0 = Bx^0$ einen Newton-Schritt (mit Schrittweite 1) durch und zeigen Sie, dass $x^1 = By^1$. Das Newton-Verfahren ist also gegenüber diesem Koordinatenwechsel invariant. Was passiert, wenn man stattdessen einen Gradientenschritt durchführt?

   (b) Sei B symmetrisch und positiv definit. Bestimmen Sie den Gradienten g und die Hesse-Matrix H bzgl. des Skalarprodukts $(x, y)_B \mapsto x^\mathsf{T} B y$, sodass also die Taylorformel
   $$f(x + h) = f(x) + (g, h)_B + \frac{1}{2}(h, Hh)_B + o(\|h\|_B^2)$$
   gilt. Führen Sie einen Newton-Schritt ausgehend von $x^0 \in \mathbb{R}^n$ sowohl bzgl. des üblichen Euklidischen Skalarprodukts, als auch mit g und H durch. Welche Invarianz lässt sich feststellen?

   (c) Sei f eine quadratische Funktion mit positiv definiter Hesse-Matrix $A \in \mathbb{R}^{n \times n}$. Führen Sie einen Metrikwechsel wie in Teilaufgabe (b) mit $B = A$ durch und betrachten Sie einen Gradientenschritt vom Startpunkt $x^0 \in \mathbb{R}^n$. Vergleichen Sie diesen mit beiden Newton-Schritten aus (b).