125. Gegeben sei die Funktion $f : \mathbb{R} \mapsto \mathbb{R}$ mit $f(x) = -\frac{1}{4}x^4$. Zeigen Sie, dass das in der Vorlesung beschriebene lokale Newtonverfahren für jede Wahl des Startpunkts $x^{(0)} \in \mathbb{R}$ gegen das eindeutig bestimmte *Maximum* $x^* = 0$ der Funktion f konvergiert. Was läßt sich über das Konvergenzverhalten aussagen?

126. Betrachten Sie die Funktion $f(x) = |x| - \arctan|x|$. Zeigen Sie das lokale Newtonverfahren für keinen Startpunkt $x^{(0)}$ mit $|x^{(0)}| \geq 2$ gegen das (eindeutige) Minimum von f konvergiert.

127. Betrachten Sie das (lokale) Newtonverfahren zur Minimierung der Funktion $f : \mathbb{R} \mapsto \mathbb{R}$ mit $f(x) = |x|^p$ mit $p > 1$ und Startpunkt $x^{(0)} \neq 0$.

    (a) Bestimmen Sie geschlossene Ausdrücke für die benötigten Ableitungen.

    (b) Sei $p > 2$. Zeigen Sie dass das Newtonverfahren Q-linear gegen das globale Minimum konvergiert und bestimmen Sie die Konvergenzrate. Zeigen Sie dass die Konvergenz nicht Q-superlinear ist. Warum widerspricht dies nicht dem Konvergenzsatz aus der Vorlesung?

    (c) Wieso wurde $p = 2$ aus der obigen Betrachtung ausgeschlossen?

    (d) Sei nun $p \in (1, 2)$. Zeigen Sie, dass das Newtonverfahren auch für $p \in (\frac{3}{2}, 2)$ linear konvergiert. Was passiert für $p = \frac{3}{2}$?

128. Betrachten Sie die Funktion f mit

$$f(x_1, x_2) = 2x_1^2 + x_2^2 - 2x_1x_2 + 2x_1^3 + x_1^4$$

Bestimmen Sie die Newtonrichtung d im Punkt $x^{(0)} = (0, 1)^1$. Verwenden Sie eine Cholesky-Zerlegung (am besten von Hand zur Wiederholung wenn Ihnen die Zerlegungsmethode nicht mehr ausreichend geläufig ist) der Hessematrix von f im Punkt $x^{(0)}$ zur Bestimmung von d.

129. Sei $f \in \mathcal{C}^2(\mathbb{R}^n, \mathbb{R})$. Aus der Vorlesung ist bekannt, dass wenn in einem gegebenen Punkt $x \in \mathbb{R}^n$ die Matrix $\nabla^2 f(x)$ regulär ist, die Newtonrichtung d die

$$\nabla f(x) + \nabla^2 f(x) d = 0 \tag{1}$$

erfüllt existiert.

Wir nehmen im folgenden $\nabla f(x) \neq 0$ an.

    (a) Zeigen Sie durch Angabe eines Beispiels dass die Richtung aus (1) nicht notwendigerweise eine Abstiegsrichtung ist.

    (b) Zeigen Sie, dass wenn $\nabla^2 f(x)$ positiv definit ist, die Richtung aus (1) eine Abstiegsrichtung ist.

    (c) Bezeichne I die $n \times n$ Einheitsmatrix und sei μ ein reeller Parameter. Wir betrachten das Gleichungssystem

$$\nabla f(x) + (\nabla^2 f(x) + \mu I) d = 0 \tag{2}$$

    Bezeichne d_μ die Lösung des obigen Systems, wenn eine solche für den Wert μ existiert. Bezeichne mit $\lambda_{\min}$ den kleinsten Eigenwert von $\nabla^2 f(x)$.

    Zeigen Sie, dass das System (2) für alle $\mu > -\lambda_{\min}$ eine Lösung d_μ besitzt, die eine Abstiegsrichtung darstellt.

    (d) Zeigen Sie dass gilt:

$$\lim_{\mu \to \infty} d_\mu = 0 \qquad \text{und} \qquad \lim_{\mu \to \infty} \frac{d_\mu}{||d_\mu||} = -\frac{\nabla f(x)}{||\nabla f(x)||}$$

    (e) Gegeben sei das konkrete Beispiel mit $f(x_1, x_2) = x_1^2 - \frac{1}{2}x_2^2 + x_1 + x_2$ und $x = (0, 0)^t$. Berechnen Sie wenn möglich die Lösungen zu (1) und (2) und ermitteln Sie wann es sich um Abstiegsrichtungen handelt.

130. Beweisen Sie das folgende Lemma aus der Vorlesung: Sei $\tilde{x}$ ein isolierter Häufungspunkt der Folge $\{x^{(k)}\}$, $x^{(k)} \in \mathbb{R}^n$ für die $\{||x^{(k+1)} - x^{(k)}||\}_K \to 0$ für jede gegen $\tilde{x}$ konvergente Teilfolge $\{x^{(k)}\}_K$ gilt. Dann konvergiert die Folge $\{x^{(k)}\}$ gegen $\tilde{x}$.

131. Beweisen Sie das folgende Lemma aus der Vorlesung: $f : \mathbb{R}^n \to \mathbb{R}$ sei zweimal stetig differenzierbar und sei $\tilde{x} \in \mathbb{R}^n$ mit $\nabla^2 f(\tilde{x})$ positiv definit. Dann gibt es Konstanten $\delta_1 > 0$ und $\delta_2 > 0$ mit

$$\delta_2 ||d||^2 \leq d^t \nabla^2 f(x) d$$

für alle $x \in \mathbb{R}^n$ mit $||x - \tilde{x}|| \leq \delta_1$ und alle $d \in \mathbb{R}^n$.

132. Aufbauend auf der Beobachtung aus Aufgabe 129 läßt sich ein modifiziertes Newtonverfahren angeben, bei dem wenn die Hessematrix $\nabla^2 f(x^{(k)})$ nicht positiv definit ist, die modifizierte Newtonrichtung

$$d^{(k)} := -(\nabla^2 f(x^{(k)}) + \mu_k I)^{-1} \nabla f(x^{(k)})$$

verwendet wird. μ_k wird hier so gewählt, dass die Matrix $\nabla^2 f(x^{(k)}) + \mu_k I$ positiv definit ist, was sich durch die Wahl

$$\mu_k \geq \mu + \max\{0, -\lambda_{\min}(\nabla^2 f(x^{(k)})\}$$

erreichen lässt ($\mu > 0$). Die Schrittweite t_k wird wie mit der Armijo-Methode bestimmt.

(a) Wenden Sie dieses Verfahren auf die Funktion $f(x_1, x_2) = x_1^4 - 3x_1^2 + 2 + 2x_2^2$ an. Starten Sie im Punkt $x^{(0)} = (\frac{1}{2}, 1)^t$. Verwenden Sie $\mu = 1$ sowie für die Armijo-Regel $\sigma = \frac{1}{4}$ und $\beta = \frac{1}{2}$.

(b) Plotten/skizzieren Sie die Höhenlinien von f. Inkludieren Sie in die Graphik die klassische Newton-Richtung, die (normierte) steilste Abstiegsrichtung, sowie die modifizierte Newtonrichtung im Startpunkt $x^{(0)}$ ein.

133. Wir betrachten affin-lineare Variablentransformationen $\hat{x} = Ax + b$ wobei A eine invertierbare Matrix sei.

(a) Zeigen Sie, dass das (lokale) Newtonverfahren invariant gegenüber solchen Variablentransformationen ist.

(b) Veranschaulichen Sie dies an Hand des konkreten Beispiels mit

$$A = \begin{pmatrix} 3 & 1 \\ 4 & 1 \end{pmatrix} \qquad b = \begin{pmatrix} -1 \\ -2 \end{pmatrix}.$$

(c) Ist auch das steilste Abstiegsverfahren mit Schrittweitenwahl $t_k = 1$ für alle k invariant gegenüber den in dieser Aufgabe betrachteten Variablentransformationen?

134. Sei Q eine positiv definite Matrix. Zeigen Sie, daß das Newtonverfahren die Optimallösung von $\min \frac{1}{2} x^t Q x + c^t x + \gamma$ in einem Schritt bestimmt und zwar unabhängig von der Wahl der Startlösung.

135. Verwenden Sie das lokale Newtonverfahren um ein Minimum von $5x_1^4 + 6x_2^4 - 6x_1^2 + 2x_1x_2 + 5x_2^2 + 15x_1 - 7x_2 + 13$ ausgehend vom Startpunkt $(1, 1)^t$ zu bestimmen.

136. Gegeben sei die Funktion f mit

$$f(x_1, x_2) = \sqrt{x_1{}^2 + 1} + \sqrt{x_2{}^2 + 1}.$$

(a) Wenden Sie das lokale Newtonverfahren ausgehend von $x^{(0)} = (1, 1)^1$ an. Wieviele Schritte sind nötig um die Abbruchbedingung $||\nabla f(x^{(k)})|| \leq 10^{-8}$ zu erfüllen?

(b) Vergleichen Sie das Verhalten mit dem Verhalten der steilsten Abstiegsmethode.

(c) Wiederholen Sie das Experiment ausgehend von $x^{(0)} = (10, 10)^t$. Wie erklären Sie sich das beobachtete Verhalten der beiden Methoden im Vergleich zueinander für dieses Beispiel?

(d) Wenden Sie nun das globalisierte Newtonverfahren (wählen Sie $\rho = 10^{-8}$, $\beta = \frac{1}{2}$, $p = 2.1$, $\sigma = 10^{-4}$) für den Startpunkt $x^{(0)} = (10, 10)^t$ an.

137. Sei $F \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R}^n)$. Betrachten Sie die Funktion $E : \mathbb{R}^n \mapsto \mathbb{R}$ mit

$$E(x) = \frac{1}{2}||F(x)||^2.$$

(a) Wie verhalten sich Extrema und Nullstellen von E und F zueinander? Welche Möglichkeiten ergeben sich die Nullstellen von F zu bestimmen?

(b) Nehmen Sie an, dass $\nabla F(x)$ an der Stelle $x = x'$ regulär ist. Zeigen Sie, dass dann die Newtonrichtung für F in x' eine Abstiegsrichtung für E ist.

138. Unter welchen Bedingungen ist das steilste Abstiegsverfahren invariant bezüglich der Variablentransformation $\hat{x} = Ax$ mit A invertierbar?

139. Gauss-Newtonverfahren: Sei $F : \mathbb{R}^n \to \mathbb{R}^m$, $m \geq n$ zweimal stetig differenzierbar. Wir betrachten das Kleinste-Quadrate-Problem (KQ)

$$\min_{x \in \mathbb{R}^n} f(x) \qquad \text{mit} \qquad f(x) = \frac{1}{2}||F(x)||^2.$$

(KQ) kann mit der folgenden Variante des Newtonverfahrens gelöst werden:

- Wähle $x^{(o)} \in \mathbb{R}^n$.
- Für $k = 0, 1, 2, \ldots$ bestimme $s^{(k)}$ durch Lösen der Gauss-Newton-Gleichung (GN)

$$J_F(x^{(k)})^t J_F(x^{(k)}) s^{(k)} = -J_F(x^{(k)}) F(x^{(k)})$$

(J_F bezeichnet die Jacobimatrix von F) und setze $x^{(k+1)} = x^{(k)} + s^{(k)}$.

(a) Berechnen Sie $\nabla f(x)$ und $\nabla^2 f(x)$.

(b) Vergleichen Sie (GN) mit der Newtongleichung für das Problem (KQ). Worin besteht der Unterschied?

(c) Geben Sie eine konvexe quadratische Funktion $q_k(s)$ an für die (GN) äquivalent ist zu $\nabla q_k(s^{(k)}) = 0$ und interpretieren Sie $s^{(k)}$ als Lösung eines geeigneten quadratischen Optimierungsproblems.

(d) Sei $\bar{x} \in \mathbb{R}^n$ die Lösung von (KQ) und die Jacobimatrix $J_F(\bar{x})$ habe vollen Spaltenrang. Begründen Sie, dass dann das Gauss-Newton-verfahren in einer Umgebung von $\bar{x}$ wohldefiniert ist.

140. Fermat-Weber Problem und Weiszfeld Algorithmus: Das Fermat-Weber Problem, eines der grundlegenden geometrischen Standortprobleme, lässt sich wie folgt beschreiben: Gegeben seien m Punkte $a_1, \ldots . a_m$ im $\mathbb{R}^n$ sowie m positive Gewichte $w_1, \ldots, w_m$. Gesucht ist ein Punkt $x \in \mathbb{R}^n$, der die gewichtete Abstandssumme von x zu den gegebenen Punkten gegeben durch

$$\sum_{i=1}^{m} w_i ||x - a_i||$$

minimiert. Der folgende Algorithmus wurde 1937 von Weiszfeld angegeben:

- Wähle ein $x^{(0)} \in \mathbb{R}^n$ mit $x^{(0)} \neq a_1, \ldots, a_m$.
- Für $k = 0, 1, 2, \ldots$ berechne

$$x^{(k+1)} = T(x^{(k)}) = \frac{1}{\sum_{i=1}^{m} \frac{w_i}{||x^{(k)} - a_i||}} \sum_{i=1}^{m} \frac{w_i a_i}{||x^{(k)} - a_i||}$$

Dieser Algorithmus ist nur definiert wenn alle Iterierten $x^{(k)}$ verschieden von $a_1, \dots, a_m$ sind.

(a) Zeigen Sie, dass man den Weiszfeld Algorithmus als Gradientenverfahren mit geeigneter Schrittweitenwahl interpretieren kann.

(b) (Für Ambionierte) Zeigen Sie, dass die Folge $\{f(x^{(k)})\}$ monoton fallend ist (unter der oben genannten Voraussetzung die einen wohldefinierten Algorithmus garantiert) und sogar streng monoton fallend wenn man in einem stationären Punkt steckenbleibt.