117. Sei $f \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R})$. Zeigen Sie, dass wenn im generischen Abstiegsverfahren eine gradientenähnliche Abstiegsrichtung d^k verwendet wird und die Schrittweite t_k mit der Armijo-Regel als

$$t_k = \max\{\beta^i \,|\, i \in \mathbb{N}_0 \text{ sodass } f(x^k + \beta^i d^k) \leq f(x^k) + \sigma \beta^i \nabla f(x^k)^t d^k\}$$

bestimmt wird, jeder Häufungspunkt der Folge $\{x^k\}$ ein stationärer Punkt von f ist.

118. Sei $f \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R})$ und H_k eine Folge von symmetrischen, positiv definiten Matrizen, die für Konstante $\mu_1 > 0$ und $\mu_2 > 0$

$$\mu_1 ||x||^2 \leq x^t H_k x \leq \mu_2 ||x||^2$$

für alle $x \in \mathbb{R}^n$ und alle $k \in \mathbb{N}$ erfüllen. Zeigen Sie, dass wenn im generischen Abstiegsverfahren die Abstiegsrichtungen d^k so gewählt werden, dass

$$H_k d^k = -\nabla f(x^k)$$

und die Schrittweite t_k mit der Armijo-Regel wie in Aufgabe 117 bestimmt wird, jeder Häufungspunkt der Folge $\{x^k\}$ ein stationärer Punkt von f ist.

119. Sei $f \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R})$ und koerziv, d.h., für jede Folge $\{z^k\}$ im $\mathbb{R}^n$ gelte

$$||z^k|| \to \infty \implies f(z^k) \to \infty \qquad \text{für } k \to \infty$$

Sei x^* ein globaler Minimierer von f. Sei $\{x^k\}$ eine vom generischen Abstiegsverfahren erzeugte Folge wobei als Abstiegsrichtung d^k eine gradientenähnliche Richtung gewählt wird und die Schrittweite t_k mit der Armijo-Regel wie in Aufgabe 117 bestimmt wird. Zeigen Sie dass dann folgendes gilt:

(a) Die Folge $\{x^k\}$ besitzt eine Teilfolge, die gegen einen stationären Punkt von f konvergiert.

(b) Wenn f konvex ist, dann gilt $f(x^k) \to f(x^*)$ für $k \to \infty$.

(c) Wenn f strikt konvex ist, dann gilt $x^k \to x^*$ für $k \to \infty$.

120. (a) Bestimmen Sie in Abhängigkeit von den Eigenwerten der positiv definiten symmetrischen Matrix Q eine untere und eine obere Schranke für den Quotienten

$$\frac{x^t Q x}{x^t x}.$$

(b) Der Punkt x^* sei der Punkt, an dem die Funktion $f(x) = \frac{1}{2} x^t Q x + c^t x$ ihr Minimum annimmt, und $x_0, x_1, \ldots$ seien die aufeinanderfolgenden Näherungspunkte bei der Methode des steilsten Abstiegs. Bestimmen Sie eine Schranke für $\|x_k - x^*\|$ in Abhängigkeit von k, $\|x_0 - x^*\|$, und den Eigenwerten von Q.

121. 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?

122. 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.

123. 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 $p = \frac{3}{2}$?

124. 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.

125. 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.

126. Aufbauend auf der Beobachtung aus Aufgabe 125 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.

127. 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?

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

129. 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.

130. 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.

131. 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.