Optimierung 1 SS 2017

11. Übungsblatt

74. Bestimmen Sie die Konvergenzordnung folgender Folgen $\{x_k\}$, $k \in \mathbb{N}$.

 (a) $x_k = \dfrac{1}{k}$

 (b) $x_k = \left(\dfrac{1}{k}\right)^k$

75. Betrachten Sie die rekursiv definierte Folge x_k mit $x_{k+1} = \frac{1}{2}(x_+ + \frac{a}{x_k})$ für $a > 0$. Nehmen Sie an, dass die Folge konvergiert und bestimmen Sie unter diese Annahme die Konvergenzordnung.

76. Bestimmen Sie die Konvergenzordnung und die durchschnittliche Konvergenzordnung folgender Folgen x_k, $k \in \mathbb{N}$.

 (a) $x_k = ca^k$ wobei $a \in (0,1)$ und $c \in \mathbb{R} \setminus \{0\}$.

 (b) $x_k = a^{2^k}$ mit $a \in (0,1)$.

77. Zeigen Sie: Falls $\lim_{k\to\infty} r_k = r_*$ linear mit Konvergenzrate β, dann ist auch die durchschnittliche Konvergenz linear mit Konvergenzrate höchstens β.

 Erinnerung: Die Konvergenz ist *linear mit Rate* β falls $\lim_{k\to\infty} |r_{k+1} - r_*|/|r_k - r_*| = \beta \in (0,1)$. Wenn in der obigen Gleichung $\beta = 0$ gilt, dann heißt die Konvergenz *superlinear*. Die durchschnittliche Konvergenzordnung wird als $\inf\{p > 1 : \limsup_{k\to\infty} |r_k - r_*|^{1/p^k} = 1\}$ definiert. Wenn die durchschnittliche Konvergenzordnung 1 ist, dann heißt $\beta' := \limsup_{k\to\infty} |r_k - r_*|^{1/k}$ die Konvergenzrate. Fall $\beta' \in (0,1)$ spricht man von durchschnittlicher linearer Konvergenz, falls $\beta' = 0$ von durchschnittlicher superlinearer Konvergenz .

78. Sei $f \in \mathcal{C}^0(\mathbb{R}^n, \mathbb{R})$. Betrachten Sie ein Verfahren $\bar{S}\colon \mathbb{R}^n \to \mathbb{R}$ der Liniensuche, das das kleinste lokale Minimum der Funktion $f(x + \alpha d)$, $\alpha \geq 0$, bestimmt, wobei d ein Richtungsvektor $d \in \mathbb{R}^n \setminus \{\tilde{0}\}$ ist und $x \in \mathbb{R}^n$. Ist $\bar{S}$ abgeschlossen?

79. Betrachten Sie das Problem $\min\{f(x)\colon x \in \mathbb{R}^2\}$ mit $f(x,y) := 5x^2 + 5y^2 - xy - 11x + 11y + 11$.

 (a) Finden Sie einen Punkt, der die notwendigen Bedingungen ersten Grades für ein Minimum erfüllt.

 (b) Zeigen Sie, dass dieser Punkt auch ein globales Minimum darstellt.

 (c) Geben Sie für dieses Problem die Konvergenzrate der Methode des steilsten Abstiegs an.

 (d) Sei $(0,0)$ der Startpunkt der Methode des steilsten Abstiegs. Geben Sie eine obere Schranke für die Anzahl von Iterationen des steilsten Abstiegsverfahrens, die notwendig sind, um den Zielfunktionswert 10^{-11} zu erreichen.

80. Angenommen die Methode des steilsten Abstiegs wird verwendet um die quadratische Funktion $(x - x^*)^t Q(x - x^*)$ zu minimieren, wobei eine Toleranz von $\pm\delta\bar{\alpha}_k$ in der Liniensuche erlaubt wird ($\delta > 0$), d.h.

$$x_{k+1} = x_k - \alpha_k g_k \text{ mit } (1-\delta)\bar{\alpha}_k \leq \alpha_k \leq (1+\delta)\bar{\alpha}_k,$$

 mit dem Minimierer $\bar{\alpha}_k$ der Funktion $f(x_k - \alpha g_k)$, für $\alpha \geq 0$.

 (a) Bestimmen Sie die Konvergenzrate des Algorithmus in Abhängigket vom a, A, und δ, wobei a und A der kleinste bzw. größte Eigenwert von Q sind. Hinweis: Betrachten Sie den extremen Fall $\alpha_k = (1+\delta)\bar{\alpha}_k$.

(b) Für welche Werte von δ kann die Konvergenz des Verfahrens garantiert werden? Können Sie dieses Ergebnis geometrisch interpretieren?

81. Betrachten Sie das Problem $\min\{f(x) \colon x \in \mathbb{R}^n_+\}$ mit $f \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R})$. In einem gradientenbasierten Verfahren für dieses Problem, das auch die Nebenbedingungen berücksichtigt, sei der Richtungsvektor $d = (d_1, \ldots, d_n)^t$ folgendermaßen definiert:

$$d_i = \begin{cases} -g_i & \text{falls } x_i > 0 \text{ oder } g_i < 0 \\ 0 & \text{falls } x_i = 0 \text{ und } g_i \geq 0 \end{cases},$$

wobei $g := (g_1, g_2, \ldots, g_n) = \nabla f(x)^t$. Der Vektor d wird dann als Richtung der Liniensuche verwendet, wie in Abstiegsverfahren üblich.

(a) Wie lauten in diesem Problem die notwendigen Bedingungen erster Ordnung für ein Minimum?

(b) Zeigen Sie, dass der wie oben definierter Vektor d nur and den Punkten, die die notwendigen Bedingungen erster Ordnung erfüllen, gleich $\vec{0}$ ist.

(c) Zeigen Sie, dass entlang der Halbgeraden $x + \alpha d$ Punkte z mit $f(z) < f(x)$ gibt, falls $d \neq \vec{0}$.

(d) Lässt sich der globale Konvergenzsatz anwenden falls der zulässige Bereich eine kompakte Teilmenge von $\mathbb{R}^n_+$ bildet?

82. 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 linear gegen das globale Minimum konvergiert und bestimmen Sie die Konvergenzrate. Zeigen Sie dass die Konvergenz nicht 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 $p = \frac{3}{2}$?

83. 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.

84. 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?