Technische Universität München
Zentrum Mathematik, M1
Prof. Dr. Bastian von Harrach
Dipl.-Math. Boris von Loesch

6. Übung zu *Nichtlineare Optimierung: Grundlagen* (WS 2010/2011)

Aufgabe 6.1 (Vereinfachtes Newtonverfahren): (ca. 6 Punkte)
Es sei $F\colon \mathbb{R}^n \to \mathbb{R}^n$ eine stetig differenzierbare Abbildung. Das *vereinfachte Newtonverfahren* verläuft wie das Newtonverfahren, (Algorithmus 2.6.1) nur wird zur Berechnung des Schrittes $s^k \in \mathbb{R}^n$ anstelle der Newtongleichung die folgende vereinfachte Newtongleichung gelöst:

$$F'(x^0)s^k = -F(x^k).$$

Die neue Iterierte ergibt sich dann wie beim Newtonverfahren als $x^{k+1} = x^k + s^k$. Es sei $\bar{x} \in \mathbb{R}^n$ eine Nullstelle von F, in der $F'(\bar{x})$ invertierbar ist.

Ziel dieser Aufgabe ist es nachzuweisen, dass das vereinfachte Newtonverfahren für jeden Startpunkt x^0 aus einer geeigneten Umgebung $B_\varepsilon(\bar{x})$ von $\bar{x}$ q-linear gegen $\bar{x}$ konvergiert.

Hinweis: Schauen Sie sich die entsprechenden Umformungen im Beweis des lokalen Konvergenzsatzes für das Newtonverfahren, Satz 2.6.5, an.

a) Es sei $C := \|F'(\bar{x})^{-1}\| > 0$. Zeigen Sie für hinreichen kleines $\varepsilon > 0$ die Abschätzungen

$$\|F'(x)^{-1}\| \le 2C \quad \forall x \in B_\varepsilon(\bar{x}) \quad \text{und} \quad \sup_{x,y \in B_\varepsilon(\bar{x})} \|F'(x) - F'(y)\| \le \frac{1}{3C}.$$

b) Zeigen Sie nun, dass folgende Identität besteht:

$$x^{k+1} - \bar{x} = F'(x^0)^{-1} \int_0^1 \Big(F'(x^0) - F'(\bar{x} + t(x^k - \bar{x}))\Big)(x^k - \bar{x})\,dt.$$

c) Folgern Sie mit a) aus b) per Induktion über k, dass für alle k gilt: $\|x^{k+1} - \bar{x}\| \le 2/3\|x^k - \bar{x}\|$.

d) Begründen Sie, warum dieses Verfahren bei geeigneter Implementierung, im Allgemeinen deutlich weniger Rechenzeit pro Iteration benötigt als Algorithmus 2.6.1.

Aufgabe 6.2 (Affine Invarianz des Newtonverfahrens): (ca. 6 Punkte)
Gegeben sind die C^2-Funktion $f\colon \mathbb{R}^n \to \mathbb{R}$ und die Variablentransformation

$$T\colon y \in \mathbb{R}^n \mapsto Ay + b \in \mathbb{R}^n$$

mit $A \in \mathbb{R}^{n\times n}$ invertierbar und $b \in \mathbb{R}^n$.

Bezeichne $\hat{f}(y) = f(T(y))$ die auf y-Koordinaten transformierte Funktion f. Sei $x \in \mathbb{R}^n$ ein Punkt mit $\nabla f(x) \neq 0$ und invertierbarer Hesse-Matrix $\nabla^2 f(x)$. Weiter bezeichne x_g bzw. x_n das Ergebnis eines Gradienten- bzw. Newton-Schrittes zur Lösung des Problems $\min_{z\in\mathbb{R}^n} f(z)$ ausgehend von einem Punkt $x \in \mathbb{R}^n$:

$$x_g = x - \nabla f(x), \qquad x_n = x - \nabla^2 f(x)^{-1}\nabla f(x).$$

Entsprechend sei y_g bzw. y_n das Ergebnis eines Gradienten- bzw. Newton-Schrittes für $\hat{f}$ ausgehend von $y = T^{-1}(x)$:

$$y_g = y - \nabla \hat{f}(y), \qquad y_n = y - \nabla^2 \hat{f}(y)^{-1}\nabla \hat{f}(y).$$

a) Drücken Sie $\nabla \hat{f}(y)$ und $\nabla^2 \hat{f}(y)$ mit Hilfe von $\nabla f(x)$, $\nabla^2 f(x)$, A und b aus.

b) Zeigen Sie, dass das Newton-Verfahren invariant unter der Transformation T ist, d. h. $T(y_n) = x_n$.

c) Für welche Klasse von Matrizen A ist das Gradientenverfahren invariant unter der Transformation T, d. h. $T(y_g) = x_g$? Für welche nicht?

Aufgabe 6.3 (Levenberg-Marquard-Regularisierung): (ca. 8 Punkte)
Wir betrachten die *Levenberg-Marquardt-Regularisierung* des Newton-Verfahrens für konvexe Funktionen:

0. Wähle $x^0 \in \mathbb{R}^n,\ \beta \in (0,1),\ \gamma \in (0,1/2)$.

Für $k = 0, \dots$:

1. Falls $\nabla f(x^k) = 0$, STOP.
2. Berechne s^k durch Lösen des linearen Gleichungssystems

$$(\nabla^2 f(x^k) + \lambda_k I)s^k = -\nabla f(x^k)$$

   mit $\lambda_k \geq 0$.
3. Bestimme die Schrittweite $\sigma_k > 0$ mit Hilfe der Armijo-Regel (2.7).
4. Setze $x^{k+1} = x^k + \sigma_k s^k$.

Sei $f \in C^2$ und konvex. Wir wählen im folgenden $\lambda_k = \rho(\|\nabla f(x^k)\|)$ mit einer stetigen, monoton wachsenden Funktion $\rho\colon \mathbb{R} \to \mathbb{R}$ mit $\rho(0) = 0$ und $\rho(\tau) > 0$ für $\tau > 0$.

a) Zeigen Sie, dass die durch den Algorithmus erzeugten Suchrichtungen s^k Abstiegsrichtungen sind.

b) Sei $(x^k)_K$ eine konvergente Teilfolge. Zeigen Sie, dass die Suchrichtungs– und Schrittweiten–Teilfolgen $(s^k)_K$ und $(\sigma_k)_K$ zulässig sind.

c) Folgern Sie: Der Algorithmus terminiert entweder mit $\nabla f(x^k) = 0$ oder er erzeugt eine unendliche Folge (x^k), deren Häufungspunkte globale Minima von f sind.

d) Sei x^* ein Minimum von f in dem die Hesse-Matrix $\nabla^2 f(x^*)$ invertierbar ist. Zeigen Sie: Es existiert ein $\epsilon > 0$, so dass für alle $x \in B_\epsilon(x^*)$ und alle $\sigma \in (0,1]$ die Armijo-Bedingung

$$f(x + \sigma s) - f(x) \leq \gamma\sigma\nabla f(x)^T s$$

erfüllt ist.

Abgabe:

- Bitte geben Sie bis **Mittwoch, den 02.02.2010, 11 Uhr** Ihre Lösungen im Briefkasten im Untergeschoss ab.
- Die Aufgaben werden in den Übungsgruppen vom 02.02.2010 - 10.02.2010 besprochen.