Nichtlineare Optimierung
4. Übungsblatt

TECHNISCHE UNIVERSITÄT DARMSTADT

Fachbereich Mathematik
Prof. Dr. Stefan Ulbrich
Hannes Meinlschmidt

WS 2011-2012
11.11.2011

Gruppenübung

Aufgabe G1 (Powell-Wolfe Algorithmus)
Beweisen Sie Satz 2.6.4 aus der Vorlesung:
Sei $f : \mathbb{R}^n \to \mathbb{R}$ stetig differenzierbar und seien $0 < \gamma < \theta < 1$ beliebige, aber feste Konstanten. Dann liefert Algorithmus 4 zur Berechnung der Powell-Wolfe-Schrittweite zu jedem nicht-stationären $x \in \mathbb{R}^n$ und zu jeder Abstiegsrichtung s von f in x, entlang der f nach unten beschränkt ist, eine Schrittweite $\sigma > 0$, die die Powell-Wolfe-Bedingungen erfüllt, also

$$f(x) - f(x + \sigma s) \geq -\gamma\sigma\nabla f(x)^T s \tag{PW1}$$

und

$$\nabla f(x + \sigma s)^T s \geq \theta\nabla f(x)^T s. \tag{PW2}$$

Bemerkung. Der Algorithmus könnte sich bei der nächsten Rechnerübung als relevant erweisen.

Aufgabe G2 (Energiefunktion und Newton-Richtung)
Sei $F : \mathbb{R}^n \to \mathbb{R}^n$ stetig differenzierbar. Betrachten Sie die Funktion $\mathscr{E} : \mathbb{R}^n \to \mathbb{R}$, gegeben durch

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

(a) Wie verhalten sich Extrema und Nullstellen von $\mathscr{E}$ und F zueinander? Welche Möglichkeiten ergeben sich, um Nullstellen der Funktion F zu finden?

(b) Nehmen Sie an, dass $F'(x)$ invertierbar ist und zeigen Sie, dass die Newton-Richtung für F in x eine Abstiegsrichtung für $\mathscr{E}$ ist.

Bemerkung. Funktionen wie $\mathscr{E}$, die entlang solcher spezieller Wege (wie hier der Newton-Schritt von F) abnehmen, sind in weitreichenden Kontexten interessant und werden überlicherweise *Energiefunktionen* bzw. manchmal auch *Lyapunov-Funktionen* für (in diesem Fall) F genannt, wegen ersterem auch der Buchstabe $\mathscr{E}$.

Aufgabe G3 (Newton-Verfahren für Gleichungssysteme)
Diese Aufgabe dient zum Wiederholen des Newton-Verfahrens. Gegeben sei die Funktion $F : \mathbb{R} \to \mathbb{R}$ mit $F(x) = x^3 - x$.

(a) Skizzieren Sie den Graphen der Funktion, z.B. im Intervall $[-2, 2]$.

(b) Führen Sie einige (vier genügen) Schritte des Newton-Verfahrens durch, beginnend mit dem Startpunkt $x_0 = 2$. Tragen Sie die einzelnen Schritte in die Skizze ein.

(c) Ist der Startpunkt $x_0 = 0.51$ geeignet, um die Nullstelle $x_N = 0$ mit dem Newton-Verfahren zu finden?

(d) Welche Startpunkte sind generell ungeeignet, um mit dem Newton-Verfahren eine Nullstelle zu finden?

Hausübung

Aufgabe H1 (Zulässigkeit der Powell-Wolfe Schrittweiten) (5 Punkte)
Beweisen Sie Satz 2.6.5 aus der Vorlesung:
Sei $f:\mathbb{R}^n \to \mathbb{R}$ stetig differenzierbar und nach unten beschränkt. Sei $x_0 \in \mathbb{R}^n$ ein beliebiger Startpunkt und sei $\nabla f(x_k)$ auf der Niveaumenge $N_f(x_0)$ gleichmäßig stetig. Dann ist die im allgemeinen Abstiegverfahren (Algorithmus 3) mit der Powell-Wolfe-Regel erzeugte Schrittweitenfolge (σ_k) zulässig.
Hinweis: Zeigen Sie die Zulässigkeitsbedingung durch Kontraposition; nehmen Sie also an, dass es eine Indexmenge $K \subset \mathbb{N}$ mit

$$-\frac{\nabla f(x_k)^T s_k}{\|s_k\|_2} \geq \epsilon > 0 \quad \text{für alle } k \in K$$

gibt. Folgern Sie daraus mithilfe der zweiten Powell-Wolfe-Bedingung, dass $\|x_{k+1} - x_k\| \geq \delta$ für ein $\delta > 0$ und alle $k \in K$ gilt, und zeigen Sie damit $f(x_k) - f(x_{k+1}) \not\to 0$.

Aufgabe H2 (Noch mehr Spaß mit Gradienten) (5 Punkte)
Wir erweiteren den, in G2 vom dritten Übungsblatt eingeführten, allgemeineren Gradienten nochmals. Sei $U \subseteq \mathbb{R}^n$ offen und $f:\mathbb{R}^n \supseteq U \to \mathbb{R}$ stetig differenzierbar und Inner($\mathbb{R}^n$) der (Vektor-) Raum aller Skalarprodukte auf $\mathbb{R}^n$ (der Name kommt vom englischen *inner product*). Eine *Riemannsche Metrik* auf einer offenen Menge $U \subseteq \mathbb{R}^n$ ist definiert als eine stetige Funktion $g: U \to \text{Inner}(\mathbb{R}^n)$. Wir bezeichnen das Skalarprodukt $g(u)$ im Punkt $u \in U$ mit $\langle \cdot,\cdot \rangle_{g(u)}$. Dann gibt es für jedes $u \in U$ ein eindeutiges Element $d(u) \in \mathbb{R}^n$, so dass

$$\langle d(u), v\rangle_{g(u)} = f'(u)v \quad \text{für alle } v \in \mathbb{R}^n.$$

Dieses Element nennen wir den *Gradienten bzgl. der Riemannschen Metrik* g im Punkt $u \in U$ und definieren damit $\nabla_g f(u) := d(u)$. Es gilt also

$$\left\langle \nabla_g f(u), v \right\rangle_{g(u)} = f'(u)v = \left\langle \nabla f(u), v \right\rangle \quad \text{für alle } v \in \mathbb{R}^n.$$

Es sei, wie schon in Aufgabe G3, die Energiefunktion $\mathscr{E}:\mathbb{R}^n \to \mathbb{R}$ (wir betrachten der Einfachheit halber $U = \mathbb{R}^n$) für stetig differenzierbares $F:\mathbb{R}^n \to \mathbb{R}^n$ gegeben durch

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

(a) Geben Sie die Riemannsche Metrik an, die das Skalarprodukt $\langle \cdot,\cdot \rangle_M$ aus Aufgabe G2 vom dritten Übungsblatt erzeugt.

(b) Sei $F'(x)$ invertierbar für alle $x \in \mathbb{R}^n$. Das Gradientenverfahren bezüglich g erzeuge seine Iterierten durch

$$x_{k+1} = x_k - \nabla_g \mathscr{E}(x_k),$$

also mit Schrittweite 1. Finden Sie eine Riemannsche Metrik g, so dass das angegebene Gradientenverfahren genau dem Newtonverfahren zur Bestimmung einer Nullstelle von F entspricht. Beweisen Sie, dass Ihre Riemannsche Metrik auch wirklich Skalarprodukte erzeugt. Sie müssen die Stetigkeit von g nicht beweisen.

Bemerkung. Man kann also das Newtonverfahren für F als Gradientenverfahren für eine Energiefunktion von F bezüglich einer speziell auf F zugeschnittenen Metrik interpretieren.

Aufgabe H3 (Invarianz des Newtonverfahrens unter Koordinaten-Transformationen) (5 Punkte)
Sei $A \in \mathbb{R}^{n,n}$ eine invertierbare Matrix und $b \in \mathbb{R}^n$. Zeigen Sie, dass das Newtonverfahren im folgenden Sinne invariant unter Koordinaten-Transformation der Form $x = Ay + b$ bzw. $y = A^{-1}(x - b)$ ist:

(a) Sei $F:\mathbb{R}^n \to \mathbb{R}^n$ eine stetig differenzierbare Funktion und seien x_k die Iterierten, die das Newton-Verfahren bei Anwendung auf die Funktion F mit Startpunkt x_0 erzeugt. Für die Iterierten y_k des Newton-Verfahrens, angewendet auf die Funktion $G(y) := A^T F(Ay + b)$ mit Startpunkt $y_0 = A^{-1}(x_0 - b)$, gilt $y_k = A^{-1}(x_k - b)$ bzw. $x_k = Ay_k + b$.

(b) Sei $f:\mathbb{R}^n \to \mathbb{R}$ eine zweimal stetig differenzierbare Funktion und seien x_k die Iterierten, die das Newton-Verfahren bei Anwendung auf die Funktion ∇f mit Startpunkt x_0 erzeugt. Für die Iterierten y_k des Newton-Verfahrens, angewendet auf die Funktion ∇g mit $g(y) := f(Ay + b)$ mit Startpunkt $y_0 = A^{-1}(x_0 - b)$, gilt $y_k = A^{-1}(x_k - b)$ bzw. $x_k = Ay_k + b$.