Nichtlineare Optimierung
2. Übungsblatt

TECHNISCHE UNIVERSITÄT DARMSTADT

Fachbereich Mathematik
Prof. Dr. Irwin Yousept
Hannes Meinlschmidt

WS 2012/13
1.11.2012

Gruppenübung

Aufgabe G1 (Exakte Schrittweitensuche und quadratische Zielfunktion) (5 Punkte)
Wir untersuchen $f(x) = \frac{1}{2}x^TQx + b^Tx + c$, mit $c \in \mathbb{R}$, $b \in \mathbb{R}^n$ und $Q \in \mathbb{R}^{n\times n}$ symmetrisch positiv definit.

(a) Sei $x^k \in \mathbb{R}^n$ und sei s^k eine zulässige Suchrichtung von f in x^k. Bestimmen Sie die Lösung der exakten Schrittweitensuche, d.h. berechnen Sie

$$\sigma_k = \arg\min_{\sigma>0} f(x^k + \sigma s^k).$$

(b) Zeigen Sie, dass das Gradientenverfahren mit exakter Schrittweite genau dann in einem Schritt das globale Optimum $\bar{x} = -Q^{-1}b$ erreicht, wenn der Startpunkt x_0 so gewählt wird, dass $\nabla f(x_0)$ ein Eigenvektor von Q ist.

Aufgabe G2 (Ergänzung zu: Globale Konvergenz des Gradientenverfahrens)
Sei $f : \mathbb{R}^n \to \mathbb{R}$ stetig differenzierbar, $x_0 \in \mathbb{R}^n$ und die Niveaumenge $N_f(x_0) := \{x \in \mathbb{R}^n : f(x) \le f(x_0)\}$ sei kompakt. Zeigen Sie: Falls das Gradientenverfahren mit Startpunkt x_0 nicht endlich terminiert, so erzeugt es eine Folge (x^k), für die folgendes gilt:

(a) Die Folge (x^k) besitzt mindestens einen Häufungspunkt in $N_f(x_0)$.

(b) Es gilt $\lim_{k\to\infty} \nabla f(x^k) = 0$.

Aufgabe G3 (Negative Krümmung)
Sei $f : \mathbb{R}^n \to \mathbb{R}$ zweimal stetig differenzierbar, und sei $x \in \mathbb{R}^n$. Beweisen oder widerlegen Sie: Hat $\nabla^2 f(x)$ einen negativen Eigenwert, so existiert eine Abstiegsrichtung $s \in \mathbb{R}^n$ von f in x oder eine Richtung $d \in \mathbb{R}^n$ mit der Eigenschaft $f(x) > f(x + \alpha d)$ für ein $\alpha > 0$.

Aufgabe G4 (Zum Üben/Erinnern: Matrixnorm und Eigenwerte)
Sei $G \in \mathbb{R}^{n\times n}$ symmetrisch mit Eigenwerten $\lambda_1 \le \lambda_2 \le \cdots \le \lambda_n$. Bezeichne $\|\cdot\|_2$ die euklidische Norm auf $\mathbb{R}^n$ und

$$\|G\|_2 := \sup_{x\neq 0} \frac{\|Gx\|_2}{\|x\|_2}$$

die induzierte Matrixnorm (Spektralnorm). Zudem gelte mit $0 < \mu \le \eta$

$$\mu\|x\|_2^2 \le x^TGx \le \eta\|x\|_2^2 \quad \text{für alle } x \in \mathbb{R}^n.$$

Weisen Sie für alle $x \in \mathbb{R}^n$ die folgenden Ungleichungen bzw. Behauptungen nach:

(a) G ist positiv definit und es gilt $\mu \le \lambda_1 \le \cdots \le \lambda_n \le \eta$,

(b) die möglichen Werte von x^TGx sind noch genauer eingegrenzt durch

$$\mu\|x\|_2^2 \le \lambda_1\|x\|_2^2 \le x^TGx \le \lambda_n\|x\|_2^2 \le \eta\|x\|_2^2,$$

(c) die Bezeichnung *Spektralnorm* ist sinnvoll, denn es gilt $\|G\|_2 = \lambda_n \le \eta$,

(d) G ist invertierbar mit $\|G^{-1}\|_2 = \frac{1}{\lambda_1} \le \frac{1}{\mu}$, zudem gilt

$$\frac{1}{\eta}\|x\|^2 \le x^TG^{-1}x \le \frac{1}{\mu}\|x\|^2 \quad \text{und} \quad \frac{1}{\eta}\|x\|_2^2 \le \frac{1}{\lambda_n}\|x\|_2^2 \le x^TG^{-1}x \le \frac{1}{\lambda_1}\|x\|_2^2 \le \frac{1}{\mu}\|x\|_2^2.$$

Insbesondere zeigt (d), dass man die obigen Forderungen an G auch äquivalenterweise an G^{-1} stellen kann.

Hausübung

Aufgabe H1 (Variationsungleichung) (4 Punkte)
Sei $X \subseteq \mathbb{R}^n$ konvex und $f : \mathbb{R}^n \to \mathbb{R}$ stetig differenzierbar auf einer offenen Umgebung von X, sowie konvex auf X. Zeigen Sie: Der Punkt $\bar{x} \in X$ ist genau dann Minimum von f über X, wenn die *Variationsungleichung*

$$f'(\bar{x})(x - \bar{x}) \geq 0 \quad \text{für alle } x \in X$$

gilt. Welche Implikationen gelten auch ohne Konvexitätsannahmen an f oder X?

Aufgabe H2 (Spaß mit Gradienten) (5 Punkte)
In dieser Aufgabe soll die handelsübliche Definition des Gradienten einer stetig differenzierbaren Funktion $f : \mathbb{R}^n \to \mathbb{R}$, nämlich

$$\nabla f(x) = \left(\frac{\partial f}{\partial x_1}(x), \dots, \frac{\partial f}{\partial x_n}(x) \right)^T,$$

ein wenig aufgeweicht, bzw. verallgemeinert werden, und zwar zu folgender: Der Gradient von f im Punkt $x \in \mathbb{R}^n$ ist eindeutig gegeben als der Vektor $\nabla f(x) \in \mathbb{R}^n$, für den

$$\langle \nabla f(x), u \rangle = f'(x)u \quad \text{für alle } u \in \mathbb{R}^n$$

gilt, wobei $f'(x)$ die Ableitung bzw. Jacobimatrix von f bezeichne. Klarerweise sind die beiden Definitionen äquivalent (wieso?), allerdings erlaubt die letztere folgende Verallgemeinerung:
Sei $M \in \mathbb{R}^{n \times n}$ symmetrisch und positiv definit, und sei $\langle x, y \rangle_M = x^T M y$ das durch M induzierte Skalarprodukt auf $\mathbb{R}^n$. Dann definieren wir den Gradienten von f im Punkt $x \in \mathbb{R}^n$ bzgl. $\langle \cdot, \cdot \rangle_M$ als den eindeutigen Vektor $\nabla_M f(x)$, für den

$$\langle \nabla_M f(x), u \rangle_M = f'(x)u \quad \text{für alle } u \in \mathbb{R}^n$$

gilt. Zeigen Sie:

(a) Der Gradient $\nabla_M f(x)$ ist gegeben durch $M^{-1} \nabla f(x)$.

(b) Die anschauliche Vorstellung des steilsten An- bzw. Abstiegs ist intakt, denn das Problem

$$\min_{\substack{d \in \mathbb{R}^n \\ \|d\|_M = 1}} \langle \nabla f(x), d \rangle$$

hat die eindeutige Lösung

$$d^* = -\frac{\nabla_M f(x)}{\left\| \nabla_M f(x) \right\|_M} = -\frac{M^{-1} \nabla f(x)}{\left\| M^{-1} \nabla f(x) \right\|_M}.$$

Hier ist $\| \cdot \|_M$ die durch

$$\|x\|_M = \sqrt{\langle x, x \rangle_M} = \sqrt{x^T M x}$$

definierte Norm auf $\mathbb{R}^n$.

Aufgabe H3 (Konvergenzrate des Gradientenverfahrens) (6 Punkte)
Für eine quadratische Funktion $f : \mathbb{R}^n \to \mathbb{R}$, $f(x) = \frac{1}{2} x^T Q x + b^T x + c$ mit $c \in \mathbb{R}$, $b \in \mathbb{R}^n$ und symmetrisch positiv definiter Matrix $Q \in \mathbb{R}^{n \times n}$ wurde in der Vorlesung folgende Abschätzung für die Konvergenzrate des Gradientenverfahrens bei exakter Schrittweitensuche bewiesen:

$$f(x^{k+1}) - f(\bar{x}) \leq \left(\frac{\lambda_{\max}(Q) - \lambda_{\min}(Q)}{\lambda_{\max}(Q) + \lambda_{\min}(Q)} \right)^2 (f(x^k) - f(\bar{x})).$$

Zeigen Sie, dass diese Abschätzung scharf ist, indem Sie die Funktion $f(x_1, x_2) = \frac{1}{2}(x_1^2 + \kappa x_2^2)$ für $\kappa \geq 1$ mit Startpunkt $x_0 = (1, \frac{1}{\kappa})^T$ betrachten.