Einführung in die Numerische Mathematik

Sommersemester 2013
Prof. Dr. Sven Beuchler
Daniel Wissel

universitätbonn

Übungsblatt 6.

Abgabe am **Dienstag, 28.05.2013**

Aufgabe 22. (Gradientenverfahren mit konstanter Schrittweite) (4 Punkte)

Betrachten Sie das durch die Vorschrift

$$x^{(k+1)} = x^{(k)} - \alpha \cdot \nabla f(x^{(k)})$$

gegebene Gradientenverfahren mit fester Schrittweite $\alpha > 0$.

a) Sei $f : \mathbb{R}^d \to \mathbb{R}$ gegeben durch $f(x) = \|x\|^{3/2}$. Zeigen Sie, dass ∇f nicht Lipschitz-stetig ist, d.h. dass kein $L > 0$ existiert, so dass

$$\|\nabla f(x) - \nabla f(y)\| \leq L\|x - y\|$$

für alle $x, y \in \mathbb{R}^d$ gilt. Beweisen Sie, dass dieses Verfahren entweder nach endlich vielen Schritten das Optimum $x^* = 0$ erreicht oder es nicht gegen x^* konvergiert.

b) Sei f nun definiert als $f(x) = \|x\|^{\beta+2}$ mit $\beta > 0$. Geben Sie Bedingungen für α und x_0 an, für die das Gradientenverfahren mit konstanter Schrittweite konvergiert bzw. divergiert.

Aufgabe 23. (Wahl einer effizienten Schrittweite) (6 Punkte)

Sei $f : \mathbb{R}^d \to \mathbb{R}$ die quadratische Funktion

$$f(x) = \frac{1}{2}x^T Ax + b^T x + c$$

mit einer symmetrisch positiv definiten Matrix $A \in \mathbb{R}^{d\times d}$, $b \in \mathbb{R}^d$ und $c \in \mathbb{R}$. Seien weiter $x \in \mathbb{R}^d$ und $q \in \mathbb{R}^d$ mit $\langle \nabla f(x), q\rangle < 0$ beliebig gegeben. Zeigen Sie folgende Aussagen:

a) Die Schrittweite

$$\alpha_{\min} := -\frac{\langle \nabla f(x), q\rangle}{q^T Aq}$$

liefert den stärksten Abstieg von f entlang der Richtung q, d.h. es ist

$$f(x + \alpha_{\min} q) \leq f(x + \alpha q)$$

für alle $\alpha \in \mathbb{R}$; für $\alpha \neq \alpha_{\min}$ gilt dabei sogar die strikte Ungleichung.

b) Es existiert eine von x und q unabhängige Konstante $\theta > 0$ mit

$$f(x + \alpha_{\min} q) \leq f(x) - \theta \left(\frac{\langle \nabla f(x), q\rangle}{\|q\|}\right)^2,$$

d.h. die Schrittweite $\alpha_{\min}$ ist effizient.

Aufgabe 24. (Drei Lemmata) (6 Punkte)

Beweisen Sie die folgenden Aussagen:

a) Für alle $x \in \mathbb{R}^d$ gilt $\|x\| = \max_{\|z\|=1} |x^T z|$.

b) Für alle $x, y \in \mathbb{R}^d$ gilt $\|xy^T\| = \|x\|\|y\|$.

c) Für eine beliebige Matrix $A \in \mathbb{R}^{d\times d}$ und eine Orthonormalbasis $\{v_1, \dots, v_d\}$ des $\mathbb{R}^d$ lässt sich die Frobeniusnorm berechnen über

$$\|A\|_F^2 = \sum_{i=1}^{d} \|Av_i\|_2^2.$$

Aufgabe 25. (Powell-symmetric-Broyden-Formel) (6 Punkte)

Für ein Quasi-Newton-Verfahren nennt man die Forderung (1.53) aus der Vorlesung

$$A^{(k+1)}q^{(k)} = y^{(k)} \tag{1}$$

auch *Quasi-Newton-Bedingung*. Um aus einer gegebenen Matrix $A := A^{(k)}$ die Matrix des nächsten Iterationsschritts $A_+ := A^{(k+1)}$ zu berechnen, soll nun die folgende Quasi-Newton-Aufdatierungsformel hergeleitet werden:

Seien $q \in \mathbb{R}^d$ mit $q \neq 0$, $y \in \mathbb{R}^d$ und eine symmetrische Matrix $A \in \mathbb{R}^{d\times d}$ gegeben. Dann ist die eindeutige Lösung des Problems

$$\min \|A_+ - A\|_F^2 \quad \text{u.d.N.} \quad A_+q = y,\ A_+^T = A_+ \tag{2}$$

gegeben durch die sog. Powell-symmetric-Broyden-Formel (PSB-Formel):

$$A_+^{\text{PSB}} := A + \frac{(y - Aq)q^T + q(y - Aq)^T}{q^T q} - \frac{(y - Aq)^T q}{(q^T q)^2} q q^T.$$

Hinweise:

a) Zeigen Sie zunächst, dass die Matrix A_+^{PSB} der Quasi-Newton-Bedingung (1) genügt und für das Problem (2) zulässig ist.

b) Sei nun $M \in \mathbb{R}^{d\times d}$ eine beliebige Matrix, welche die Nebenbedingungen des Problems (2) erfüllt, sowie $v \in \mathbb{R}^d$ ein beliebiger Vektor mit $q^T v = 0$. Zeigen Sie die Abschätzung $\|(A_+^{\text{PSB}} - A)v\| \leq \|(M - A)v\|$ mit Hilfe von Teil b) aus Aufgabe 24.

c) Betrachten Sie schließlich eine Orthonormalbasis des $\mathbb{R}^d$, welche den Vektor $v_1 := q/\|q\|$ enthält. Verwenden Sie Teil c) aus Aufgabe 24, um schließlich die Abschätzung $\|A_+^{\text{PSB}} - A\|_F^2 \leq \|M - A\|_F^2$ zu zeigen.

Programmieraufgabe 6. (Powell–symmetric–Broyden–Verfahren) (12 + 6 Punkte)

Wir betrachten den folgenden

Algorithmus 1. *Powell–symmetric–Broyden–Verfahren*

(1) Wähle $x^{(0)} \in \mathbb{R}^d$, $A^{(0)} \in \mathbb{R}^{d\times d}$ symmetrisch, $\epsilon \geq 0$, und setze $k := 0$.

(2) Ist $\|\nabla F(x^{(k)})\|_2 \leq \epsilon$: STOP.

(3) Bestimme $d^{(k)}$ durch $\quad d^{(k)} := -\left(A^{(k)}\right)^{-1} \nabla F(x^{(k)})$.

(4) Setze $x^{(k+1)} := x^{(k)} + d^{(k)}$, $q^{(k)} := x^{(k+1)} - x^{(k)}$, $y^{(k)} := \nabla F(x^{(k+1)}) - \nabla F(x^{(k)})$ und

$$\begin{aligned} A^{(k+1)} :=& A^{(k)} + \frac{(y^{(k)} - A^{(k)}q^{(k)})(q^{(k)})^T + q^{(k)}(y^{(k)} - A^{(k)}q^{(k)})^T}{(q^{(k)})^T q^{(k)}} \\ &- \frac{(q^{(k)})^T(y^{(k)} - A^{(k)}q^{(k)})q^{(k)}(q^{(k)})^T}{((q^{(k)})^T q^{(k)})^2}. \end{aligned}$$

(5) Setze $k := k + 1$ und gehe zu (2).

a) Implementieren Sie zunächst den obigen Algorithmus, wobei Sie für die Lösung des Gleichungssystems in Schritt (3) einen Löser aus LAPACK verwenden dürfen. Achten Sie auf Effizienz bei der Berechnung der neuen Matrix in Schritt (4). Wie viele Hilfsvektoren benötigen Sie?

b) Testen Sie Ihr Programm anhand der *Rosenbrock–Funktion* sowie der *Funktion von Himmelblau* (siehe letzte Programmieraufgabe) mit entsprechenden Werten. Als Startmatrix $A^{(0)}$ wählen Sie beispielsweise die Einheitsmatrix.

c) **Bonusaufgabe:** Modifizieren Sie Ihre Implementierung so, dass nicht in jedem Iterationsschritt erneut ein Gleichungssystem gelöst werden muss, sondern dass eine vorhandene Dreieckszerlegung von $A^{(k)}$ zu einer neuen Dreieckszerlegung von $A^{(k+1)}$ aufdatiert wird. Dabei können Sie mit einer symmetrisch positiv definiten Jacobi-Matrix sowie einer zugehörigen Cholesky-Zerlegung arbeiten.

Abgabe **Mo 27.05.** und **Di 28.05.** im **CIP-Pool** (`www.iam.uni-bonn.de/pcpool/`) in der Wegelerstraße. Ab Mo 13.05. hängt dort eine Terminliste für diese beiden Tage aus; bitte tragen Sie sich alleine oder in 2er Gruppen ein.

Gesamtpunktzahl: 22 + 18 Punkte