

Nichtlineare Optimierung

1. Rechnerübungsblatt



TECHNISCHE
UNIVERSITÄT
DARMSTADT

Fachbereich Mathematik
Prof. Dr. Irwin Yousept
Hannes Meinlschmidt

WS 2012/13
15.11.2012

Rechnerübung

Aufgabe R1 (Gradientenverfahren)

Implementieren Sie das Gradientenverfahren aus der Vorlesung in MATLAB. Verwenden Sie zur Bestimmung der Schrittweite die Schrittweitenregel von Armijo. Implementieren Sie hierbei die Schrittweitsuche nach Armijo unabhängig vom Hauptalgorithmus des Gradientenverfahrens, so dass Sie die Schrittweitsuche später in anderen Verfahren wiederverwenden können. Beachten Sie, dass das Programm insbesondere auch für Zielfunktionen $f : \mathbb{R}^n \rightarrow \mathbb{R}$ mit $n > 1$ funktionieren soll. Verwenden Sie die folgende Anleitung:

- *Zum Gradientenverfahren:* Verwenden Sie den Funktionskopf

```
function [xn] = grad(x0,fg,tol,stepsize,maxit)
```

mit folgenden Parametern:

x0 Startpunkt (in der Vorlesung sind alle Vektoren Spaltenvektoren),

fg eine MATLAB-Funktion, die beim Aufruf $[f,g] = fg(x)$ den Funktionswert $f(x)$ der Zielfunktion und deren Gradient $g(x) = \nabla f(x)$ im Punkt x zurückgibt, beim Aufruf $[f] = fg(x)$ aber nur den Funktionswert $f(x)$ (siehe MATLAB-Funktion `nargout`)

tol Toleranz für die Abbruchbedingung: $\|g(x^k)\| \leq \text{tol} \cdot \min(1, \|g(x^0)\|)$,

stepsize definiert, welche Schrittweitenregel benutzt werden soll (z.B. `stepsize = 1` verwendet Armijo),

maxit gibt die maximale Anzahl durchzuführender Iterationen an.

- *Zur Armijo-Regel:* Verwenden Sie den Funktionskopf

```
function [sig] = armijo(xk,sk,stg,fg,fk,gamma)
```

mit folgenden Parametern:

xk aktueller Punkt,

sk aktuelle Suchrichtung,

stg aktuelle Richtungsableitung in Richtung s^k , also $\nabla f(x^k)^T s^k$,

fg Zielfunktion (s.o.),

fk aktueller Funktionswert,

gamma Parameter γ aus der Armijo-Regel.

Wählen Sie zur Berechnung der Schrittweite σ den Parameter $\beta = 0.5$ fest.

Testen Sie ihr Verfahren für folgende Funktionen:

- die quadratische Funktion $f_1(x_1, x_2) = x_1^2 + \alpha x_2^2$, mit verschiedenen $\alpha \geq 1$,
- die *Rosenbrock-Funktion*

$$f_2(x_1, x_2) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2,$$

globales Minimum ist $(1, 1)$,

- die *Dixon-Funktion:*

$$f_3(x_1, \dots, x_{10}) = (1 - x_1)^2 + (1 - x_{10})^2 + \sum_{i=1}^9 (x_i^2 - x_{i+1})^2,$$

globales Minimum ist der Vektor $(1, 1, \dots, 1)$.

Aufgabe R2 (Veranschaulichung des Gradientenverfahrens)

Ziel ist es nun, den Verlauf der Iterationspunkte zu visualisieren. Gehen Sie dazu wie folgt vor:

- Erweitern Sie zuerst ihr Programm zur Implementierung des Gradientenverfahrens: In jeder Iteration soll nun die Verbindungsgerade zwischen altem und neuem Iterationspunkt mit dem Befehl `plot` geplottet werden. In der k -ten Iteration entspricht dies genau der Richtung des steilsten Abstiegs im Punkt x^k mit der Länge der gewählten Schrittweite.
- Legen Sie nun eine weitere Matlab-Datei `runf1.m` an, in der Sie alle benötigten Schritte zum Aufruf und der Visualisierung Ihres Programms für die Funktion f_1 zusammenfassen (wählen Sie hier den Parameter $\alpha = 10$ fest):
 - Erzeugen Sie ein Höhenliniendiagramm zur Funktion f_1 mit dem Befehl `contour` (hierbei sollte sich auch `meshgrid` als nützlich erweisen).
 - Rufen Sie Ihr erweitertes Programm für das Gradientenverfahren auf, welches den Pfad der Iterationspunkte dort einzeichnet.

Hinweis: Der Befehl `hold on` könnte hilfreich sein.

Testen Sie Ihr erweitertes Programm für die Funktion f_1 und erstellen Sie dann analog eine Datei `runf2.m`, um die Funktion f_2 zu testen.

Hausübung

Aufgabe H1 (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:

- Sei $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ 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. $x_k = Ay_k + b$.
- Sei $f : \mathbb{R}^n \rightarrow \mathbb{R}$ 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. $x_k = Ay_k + b$.

Aufgabe H2 (Noch mehr Spaß mit Gradienten) (5 Punkte)

Wir erweitern die Definition eines Gradienten nochmals. Sei dazu $U \subseteq \mathbb{R}^n$ offen und $f : \mathbb{R}^n \supseteq U \rightarrow \mathbb{R}$ stetig differenzierbar. Sei zudem $\text{Inner}(\mathbb{R}^n)$ der (Vektor-) Raum aller Skalarprodukte auf \mathbb{R}^n .¹ Eine *Riemannsche Metrik* (auf der offenen Menge $U \subseteq \mathbb{R}^n$) ist nun definiert als eine stetige Funktion $g : U \rightarrow \text{Inner}(\mathbb{R}^n)$, d.h. $g(u)$ ist für jedes $u \in U$ ein Skalarprodukt auf \mathbb{R}^n , das wir mit $\langle \cdot, \cdot \rangle_{g(u)}$ notieren. Nun gibt es für jedes $u \in U$ ein eindeutiges Element $d(u) \in \mathbb{R}^n$, so dass

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

gilt. 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

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

Wir betrachten nun der Einfachheit halber $U = \mathbb{R}^n$. Sei die (sog. *Energiefunktion*) $\mathcal{E} : \mathbb{R}^n \rightarrow \mathbb{R}$ für stetig differenzierbares $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ gegeben durch

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

- Geben Sie zunächst die Riemannsche Metrik an, die das Skalarprodukt $\langle \cdot, \cdot \rangle_M$ aus Aufgabe H2 vom zweiten Übungsblatt erzeugt.
- Sei $F'(x)$ invertierbar für alle $x \in \mathbb{R}^n$. Das Gradientenverfahren bezüglich g erzeuge seine Iterierten durch

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

also mit Schrittweite 1 und Suchrichtung $\nabla_g \mathcal{E}$. 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 der Abbildung g nicht beweisen (man darf das aber natürlich trotzdem gerne tun).

¹ Der Name kommt vom englischen *inner product* für Skalarprodukt.

Bemerkung. Was man in (b) bemerken sollte, ist, dass man das Newtonverfahren für Nullstellen von F also als Gradientenverfahren für eine Energiefunktion von F bezüglich einer speziell auf F zugeschnittenen Metrik interpretieren kann.

Aufgabe H3 (Wiederholung: Newton-Verfahren für Gleichungssysteme) (5 Punkte)

Diese Aufgabe dient zum Wiederholen des Newton-Verfahrens. Gegeben sei die Funktion $F : \mathbb{R} \rightarrow \mathbb{R}$ 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?