Technische Universität München
Zentrum Mathematik
Lehrstuhl für Mathematische Optimierung, M1
Prof. Dr. Michael Ulbrich
Dipl.-Math. Florian Lindemann

---

4. Nichtlineare Optimierung: Grundlagen (WS 2011/12)

Aufgabe 4.1 (Gradientenverfahren mit fester Schrittweite): (ca. 6 Punkte)
Sei $f: \mathbb{R}^n \to \mathbb{R}$ eine stetig differenzierbare Funktion, für die eine feste Konstante $L > 0$ existiert, so dass gilt:

$$\|\nabla f(x) - \nabla f(y)\| \leq L\|x - y\| \quad \forall x, y \in \mathbb{R}^n.$$

Wir betrachten nun ein Gradientenverfahren analog zu Algorithmus 2.3.5. Hierbei werden in Schritt 3 die σ_k nicht mit der Armijo-Regel bestimmt, sondern stattdessen $\sigma_k = \alpha_k$ gewählt mit

$$\varepsilon \leq \alpha_k \leq \frac{(1-\varepsilon)}{L}.$$

Dabei ist $0 < \varepsilon < 1/(1 + L)$ eine über den gesamten Algorithmus feste Konstante,

a) Zeigen Sie, dass für die durch den Algorithmus erzeugte Folge (x^k) gilt:

$$f(x^k) - f(x^{k+1}) \geq \varepsilon^2 ||\nabla f(x^k)||^2.$$

b) Beweisen Sie, dass der so modifizierte Algorithmus entweder mit einem stationären Punkt x^k abbricht oder eine unendliche Folge (x^k) erzeugt, deren Häufungspunkte stationäre Punkte von f sind.

Aufgabe 4.2 (Programmieraufgabe): (ca. 10 Punkte)
Implementieren Sie das Gradientenverfahren aus der Vorlesung (Algorithmus 2.3.5) mit der Armijo-Schrittweitenregel in Matlab. Teilen Sie dazu das Verfahren in zwei Funktionen auf:

1. Eine Funktion `Armijo`, die bei gegebenen Daten die Schrittweite berechnet und zurückgibt:

```
function [sigma] = Armijo(f, gfx, x, s, beta, gamma)
```

   mit Eingabeparameter

   - `f`: Ein Funktionshandle auf die Funktion f (vgl. hierzu `feval` in der Matlab-Hilfe)
   - `gfx`: Der Gradient ausgewertet an der Stelle `x`
   - `x`: Die aktuelle Iterierte
   - `s`: Die Suchrichtung
   - `beta`, `gamma`: Die Parameter der Armijo-Suche.

2. Eine Funktion `GradientenVerf`, welche die Funktion `Armijo` verwendet und folgendes Interface hat:

```
function [x,steps] = GradientenVerf(f, gf, x0, eps, maxsteps)
```

   Die Eingabeparameter sind:

- `f`, `gf`: Ein Funktionshandle auf die Funktion f bzw. den Gradienten von f
- `x0`: Der Startpunkt x^0
- `eps`: Die Abbruchtoleranz für die Norm des Gradienten: $\|\nabla f(x^k)\|_2 \leq \varepsilon$
- `maxsteps`: Die maximale Anzahl an Schritten, die durchgeführt werden sollen.

a) Testen Sie Ihre Implementierung am Beispiel der *Rosenbrockfunktion*

$$f(x) := 100(x_2 - x_1^2)^2 + (1 - x_1)^2$$

mit Startpunkt $x^0 = (-1.2, 1)^T$ und Parametern maxsteps $= 10\,000$, $\beta = 0.5$ und $\gamma = 10^{-4}$. Führen Sie das Verfahren für $\varepsilon = 10^{-1}$, 10^{-2}, 10^{-3}, 10^{-4}, 10^{-5} aus. Für $\varepsilon = 10^{-3}$ sollte das Verfahren ca. 5231 Iterationen benötigen und den Punkt $x = (0.9992, 0.9984)^T$ zurückgeben.

b) Lassen Sie die Höhenlinien der Rosenbrockfunktion zeichnen (Matlab-Befehl `contour`) und zeichnen Sie die Folge der durch Ihr Verfahren erzeugten Iterierten (x^k) ein. Können Sie erklären, warum das Gradientenverfahren so ineffizient ist?

c) Wir wollen noch eine andere Schrittweitenregel testen: In der Vorlesung wurde im Kapitel 2.3.4 die Minimierungsregel eingeführt. Wir möchten nun das Gradientenverfahren mit Minimierungsregel statt Armijoregel auf eine streng konvexe quadratische Funktion anwenden. Sei also $\hat{\sigma}_k$ die in der k-ten Iteration zur Suchrichtung s^k durch die Minimierungsregel bestimmte Schrittweite für eine solche Funktion.

   Wandeln Sie nun die von Ihnen programmierte Funktion `GradientenVerf` so ab, dass sie die Minimierungsregel statt der Armijoregel benutzt. Sie können davon ausgehen, dass `f` eine streng konvexe quadratische Funktion ist. Nutzen Sie dabei die explizite Formel für die Schrittweite $\hat{\sigma}_k$ (Vorlesung!). Testen Sie Ihre Implementierung für die Funktion

$$f(x) := x_1^2 + a x_2^2$$

   wobei $a = 10$ ist. Starten Sie dabei mit dem Punkt $x^0 = (a, 1)^T$. Experimentieren Sie etwas mit verschiedenen Werten von a, anderen Startpunkten und/oder der Armijo-Regel für diese Funktion. Visualisieren Sie sich die Ergebnisse wie in b).

Geben Sie bitte Ihren Quellcode und die Ausgabe ab. Hinweis: Sie können als Student einen aus Studienbeiträgen finanzierten Zugang zu Matlab erhalten. Genauere Informationen dazu finden Sie auf der Website dieser Veranstaltung und bei Ihrem Tutor!

Aufgabe 4.3 (Multiple Choice): **(ca. 3,5 Punkte)**

Bewertung: -0.5/0/0.5 Punkte für falsche/keine/richtige Antwort. In der Gesamtwertung dieser Aufgabe können keine negativen Punkte erzielt werden, d. h. bei einer negativen Gesamtpunktzahl wird diese Aufgabe mit null Punkten bewertet.

In dieser Aufgabe sei $f\colon \mathbb{R}^n \to \mathbb{R}$ eine stetig differenzierbare Funktion. Wenn in dieser Aufgabe vom Gradientenverfahren gesprochen wird, dann ist stets das Gradientenverfahren mit Armijo-Schrittweitenregel (Alg. 2.3.5) gemeint.

a) Für eine vom Gradientenverfahren erzeugte Folge (x^k) gilt
   $\|\nabla f(x^{k+1})\| < \|\nabla f(x^k)\| \quad \forall k \in \mathbb{N}$. **wahr** ☐ **falsch** ☐

b) Es seien $x \in \mathbb{R}^n$ und $s \in \mathbb{R}^n$, wobei s *keine* Abstiegsrichtung von f im Punkt x sei. Dann gilt: $\nexists \tau > 0$ mit $f(x + \sigma s) < f(x)\ \forall \sigma \in (0, \tau]$. **wahr** ☐ **falsch** ☐

c) Wir betrachten das Gradientenverfahren und wenden es auf f an. In der Vorlesung wurde die *globale Konvergenz* dieses Verfahrens bewiesen. Also gilt:
Das Verfahren konvergiert gegen das globale Minimum von f. **wahr** ☐ **falsch** ☐

d) Beim allgemeinen Abstiegsverfahren wird gefordert, dass man von x^k ausgehend als Suchrichtung s^k eine Abstiegsrichtung wählt.
Man muss also $s^k = -\lambda \nabla f(x^k)$ wählen mit $\lambda > 0$. **wahr** ☐ **falsch** ☐

e) Wendet man das Gradientenverfahren auf f an, so gilt: Das Verfahren terminiert endlich oder erzeugt eine Folge, die mindestens einen Häufungspunkt besitzt. **wahr** ☐ **falsch** ☐

f) Es sei $x^0 \in \mathbb{R}^n$ ein Punkt mit $\nabla f(x^0) \neq 0$. Weil der negative Gradient eine Abstiegsrichtung ist, gilt für $x^1 := x^0 - \nabla f(x^0)$: $f(x^1) < f(x^0)$. **wahr** ☐ **falsch** ☐

g) Wir betrachten ein allgemeines Abstiegsverfahren (Alg. 2.4.1), wobei $s^k = -\nabla f(x^k)$ gewählt wird und als σ_k irgendeine Schrittweite bestimmt wird, welche die Armijo-Bedingung (2.8) erfüllt. Dann gelten für dieses Verfahren die gleichen Aussagen wie in Satz 2.3.6.
wahr ☐ **falsch** ☐

Hinweis: Die Aufgabe 4.3 ist nicht relevant für den Notenbonus.

Abgabe: Bitte geben Sie bis **Mittwoch, den 07.12.2011, 11 Uhr** Ihre Lösungen im Briefkasten im Untergeschoss ab.