....
Note

Name	Vorname

Matrikelnummer	Studiengang (Hauptfach)	Fachrichtung (Nebenfach)

Obige Angaben sind richtig:

Unterschrift der Kandidatin/des Kandidaten

	I	II
1		
2		
3		

Σ		

I ....
Erstkorrektur

II ....
Zweitkorrektur

TECHNISCHE UNIVERSITÄT MÜNCHEN

Fakultät für Mathematik

Grundlagen der Nichtlinearen Optimierung

Klausur zur Vorlesung WS 2008/09

Termin: Donnerstag 19.02.2009 09:40 – 10:40 Uhr

Prof. Dr. M. Ulbrich

Hörsaal: .... Reihe: .... Platz: ....

Hinweise:

Überprüfen Sie die Angabe: Es sind **3 Aufgaben** auf den **Seiten 1 bis 7** !

Jede Aufgabe ist in dem unmittelbar anschliessenden eingerahmten Platz zu bearbeiten. Rückseiten möglichst nicht beschreiben!

Zum Bestehen sind mindestens 17 Punkte nötig!

Das letzte Blatt mit der Aufgabenübersicht kann zur Bearbeitung abgetrennt werden.

Bei eventueller vorzeitiger Beendigung der Prüfung ist dieses Blatt mit abzugeben.

Nur von der Aufsicht auszufüllen:

Hörsaal verlassen von .... bis ....

Vorzeitig abgegeben um ....

Besondere Bemerkungen:

Wir betrachten das Optimierungsproblem

$$\min_{x \in \mathbb{R}^n} q(x) := \frac{1}{2}\|Ax - b\|_2^2 + \frac{\alpha}{2}\|x\|_2^2 \tag{P}$$

mit der Matrix $A \in \mathbb{R}^{m \times n}$, dem Vektor $b \in \mathbb{R}^m$ und der positiven Zahl $\alpha > 0$.

(a) Zeigen Sie, dass es sich bei (P) um ein quadratisches Optimierungsproblem handelt, indem Sie q in die Standardform einer quadratischen Funktion, d.h.

$$q(x) = \frac{1}{2}x^T C x + c^T x + \gamma, \tag{*}$$

überführen, mit symmetrischer Matrix $C \in \mathbb{R}^{n \times n}$, $c \in \mathbb{R}^n$ und $\gamma \in \mathbb{R}$.

(b) Zeigen Sie, dass die von Ihnen ermittelte Matrix C positiv definit ist.

Hinweis: Sollten Sie Teil (a) nicht bearbeitet haben, können Sie ab Teil (c) die Form (*) für q benutzen.

(c) Begründen Sie, dass q genau ein lokales Minimum $x^* \in \mathbb{R}^n$ besitzt und dass dieses sogar ein globales Minimum ist. Berechnen Sie x^*.

(d) Es sei $x^0 \in \mathbb{R}^n$ gegeben. Wir betrachten einen Schritt des lokalen Newtonverfahrens zur Minimierung von q (Startpunkt x^0). Bestimmen Sie dazu zunächst s^0 und anschließend $x^1 = x^0 + s^0$.

(e) Wieviele Iterationen sind bei Verwendung des lokalen Newtonverfahrens zur Bestimmung des Minimums von q maximal erforderlich?

⟸ **Bei Bedarf dort weiterschreiben!**

Aufgabe 2 (11 Punkte: -1/0/1 Punkte für falsche/keine/richtige Antwort)

Bitte kreuzen Sie nur die Kästchen an – Begründungen werden in dieser Aufgabe nicht bewertet.
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 bezeichne stets $f : \mathbb{R}^n \to \mathbb{R}$ eine zweifach stetig differenzierbare Funktion.

(a) Wenn f konvex ist, dann ist $\nabla^2 f(x)$ positiv definit für alle $x \in \mathbb{R}^n$. wahr ☐ falsch ☐

(b) Ist $g : I \to \mathbb{R}$ konvex, $I \subset \mathbb{R}$ ein Intervall, so ist g^2 strikt konvex. wahr ☐ falsch ☐

(c) Es sei $x \in \mathbb{R}^n$ ein Punkt mit $\nabla f(x) \neq 0$. Dann gilt:
$-\nabla f(x)$ ist eine Abstiegsrichtung von f in x. wahr ☐ falsch ☐

(d) 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:
$\nexists \tau > 0 : f(x + \sigma s) < f(x) \;\; \forall \sigma \in (0, \tau]$. wahr ☐ falsch ☐

(e) Wendet man das Gradientenverfahren mit Armijoregel (Algorithmus 2.3.5) auf f an, so gilt:
Das Verfahren terminiert endlich oder erzeugt eine Folge, die mindestens einen Häufungspunkt besitzt.
wahr ☐ falsch ☐

(f) Es seien $\gamma \in (0, 1)$ und $x^* \in \mathbb{R}^n$. Ferner sei $(x^k) \subset \mathbb{R}^n$ eine Folge, die folgender Abschätzung genüge:

$$\|x^{k+1} - x^*\|_2 \leq \gamma \|x^k - x^*\|_2 \quad \forall k \geq 0.$$

Dann gilt:
Die Folge (x^k) ist konvergent. wahr ☐ falsch ☐

(g) Sei $x^* \in \mathbb{R}^n$ ein lokales Minimum von f, in dem die hinreichenden Optimalitätsbedingungen zweiter Ordnung erfüllt sind (insbesondere ist also $\nabla f(x^*) = 0$). Dann gilt:
Es gibt eine Umgebung von x^*, in der sich keine weitere Gradientennullstelle befindet.
wahr ☐ falsch ☐

(h) Sei $x^* \in \mathbb{R}^n$ ein lokales Minimum von f. Wir wenden das lokale Newtonverfahren für Optimierungsprobleme (Algorithmus 2.6.6) auf f an. Es erzeuge eine gegen x^* konvergente Folge. Dann gilt:
In x^* gelten die hinreichenden Optimalitätsbedingungen zweiter Ordnung. wahr ☐ falsch ☐

(i) Wir wenden das globalisierte Newtonverfahren (Algorithmus 2.6.10) auf f an. Es erzeuge die Folge $(x^k) \subset \mathbb{R}^n$. Dann gilt:
Jeder Häufungspunkt von (x^k) ist stationär. wahr ☐ falsch ☐

(j) Wir betrachten ein lokales newtonartiges Verfahren zur Minimierung von f. Das Verfahren erzeuge unter Verwendung der Matrizenfolge $(M_k) \subset \mathbb{R}^{n \times n}$ mit M_k positiv definit, symmetrisch für alle k, die Folge von Iterierten (x^k). Die Folge (x^k) konvergiere gegen ein lokales Minimum x^* von f, in dem die hinreichenden Optimalitätsbedingungen zweiter Ordnung gelten. Dann gilt:
Falls $M_k \to \nabla^2 f(x^*)$ für $k \to \infty$ gilt, so liegt q-superlineare Konvergenz von (x^k) gegen x^* vor.
wahr ☐ falsch ☐

(k) Quasi-Newton-Verfahren sind newtonartige Verfahren. wahr ☐ falsch ☐

⟸ **Bei Bedarf dort weiterschreiben!**

Es sei $f : \mathbb{R}^n \to \mathbb{R}$ eine stetig differenzierbare Funktion.
Wir betrachten das folgende Abstiegsverfahren:

0. Wähle $\beta \in (0,1), \gamma \in (0,1)$ und einen Startpunkt $x^0 \in \mathbb{R}^n$.

Für $k = 0, 1, 2, \ldots$:

1. Falls $\nabla f(x^k) = 0$: STOP.

2. Bestimme $s^k \in \mathbb{R}^n$ durch

$$s^k = -\frac{\partial f}{\partial x_j}(x^k) \cdot e_j,$$

wobei $e_j \in \mathbb{R}^n$ den j-ten kanonischen Einheitsvektor bezeichnet und $j \in \{1, 2, \ldots, n\}$ ein Index ist mit $\left|\frac{\partial f}{\partial x_j}(x^k)\right| = \max_{1 \le i \le n} \left|\frac{\partial f}{\partial x_i}(x^k)\right|$.

3. Ermittle die Schrittweite $\sigma_k \in (0,1]$ mit Hilfe der Armijoregel (Parameter β und γ).

4. Setze $x^{k+1} = x^k + \sigma_k s^k$.

(a) Zeigen Sie, dass Schritt 3 wohldefiniert ist.

(b) Zeigen Sie, dass jede Teilfolge von Suchrichtungen $(s^k)_K$ zulässig ist.

(c) Zeigen Sie, dass $f(x^{k+1}) < f(x^k)$ für alle k gilt.

(d) Zeigen Sie, dass für jede beschränkte Teilfolge von Iterierten $(x^k)_K$ die entsprechende Schrittweitenteilfolge $(\sigma_k)_K$ zulässig ist.

Tipp: Begründen Sie zunächst, dass der angegebene Algorithmus ein Spezialfall des allgemeinen Abstiegsverfahrens (Algorithmus 2.4.1) ist. Benutzen Sie dann den folgenden Satz aus der Vorlesung:

Satz 2.5.2. *Sei $f : \mathbb{R}^n \to \mathbb{R}$ stetig differenzierbar und die Teilfolge $(x^k)_K$ sei beschränkt [...]. Weiter gebe es eine streng monoton wachsende Funktion $\varphi : [0,\infty) \to [0,\infty)$, so dass die durch das allgemeine Abstiegsverfahren erzeugten Suchrichtungen folgender Bedingung genügen:*

$$\|s^k\|_2 \ge \varphi\left(\frac{-\nabla f(x^k)^T s^k}{\|s^k\|_2}\right) \quad \forall k \in K.$$

Dann ist die durch die Armijoregel erzeugte Schrittweitenteilfolge $(\sigma_k)_K$ zulässig.

(e) Es sei (x^k) eine durch den angegebenen Algorithmus erzeugte Folge von Iterierten. Begründen Sie, dass jeder Häufungspunkt von (x^k) ein stationärer Punkt von f ist.

⟸ **Bei Bedarf dort weiterschreiben!**

Grundlagen der Nichtlinearen Optimierung

Aufgabe 1 (ca. 3+2+4+3+1=13 Punkte)
Wir betrachten das Optimierungsproblem

$$\min_{x\in\mathbb{R}^n} q(x) := \frac{1}{2}\|Ax-b\|_2^2 + \frac{\alpha}{2}\|x\|_2^2 \tag{P}$$

mit der Matrix $A \in \mathbb{R}^{m\times n}$, dem Vektor $b \in \mathbb{R}^m$ und der positiven Zahl $\alpha > 0$.

(a) Zeigen Sie, dass es sich bei (P) um ein quadratisches Optimierungsproblem handelt, indem Sie q in die Standardform einer quadratischen Funktion, d.h.

$$q(x) = \frac{1}{2}x^T Cx + c^T x + \gamma, \tag{*}$$

überführen, mit symmetrischer Matrix $C \in \mathbb{R}^{n\times n}$, $c \in \mathbb{R}^n$ und $\gamma \in \mathbb{R}$.

(b) Zeigen Sie, dass die von Ihnen ermittelte Matrix C positiv definit ist.

Hinweis: Sollten Sie Teil (a) nicht bearbeitet haben, können Sie ab Teil (c) die Form (*) für q benutzen.

(c) Begründen Sie, dass q genau ein lokales Minimum $x^* \in \mathbb{R}^n$ besitzt und dass dieses sogar ein globales Minimum ist. Berechnen Sie x^*.

(d) Es sei $x^0 \in \mathbb{R}^n$ gegeben. Wir betrachten einen Schritt des lokalen Newtonverfahrens zur Minimierung von q (Startpunkt x^0). Bestimmen Sie dazu zunächst s^0 und anschließend $x^1 = x^0 + s^0$.

(e) Wieviele Iterationen sind bei Verwendung des lokalen Newtonverfahrens zur Bestimmung des Minimums von q maximal erforderlich?

Aufgabe 2 (11 Punkte: -1/0/1 Punkte für falsche/keine/richtige Antwort)
Bitte kreuzen Sie nur die Kästchen an – Begründungen werden in dieser Aufgabe nicht bewertet.
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 bezeichne stets $f : \mathbb{R}^n \to \mathbb{R}$ eine zweifach stetig differenzierbare Funktion.

(a) Wenn f konvex ist, dann ist $\nabla^2 f(x)$ positiv definit für alle $x \in \mathbb{R}^n$.

(b) Ist $g : I \to \mathbb{R}$ konvex, $I \subset \mathbb{R}$ ein Intervall, so ist g^2 strikt konvex.

(c) Es sei $x \in \mathbb{R}^n$ ein Punkt mit $\nabla f(x) \neq 0$. Dann gilt:
$-\nabla f(x)$ ist eine Abstiegsrichtung von f in x.

(d) 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:
$\nexists \tau > 0 : f(x+\sigma s) < f(x)\ \ \forall \sigma \in (0,\tau]$.

(e) Wendet man das Gradientenverfahren mit Armijoregel (Algorithmus 2.3.5) auf f an, so gilt:
Das Verfahren terminiert endlich oder erzeugt eine Folge, die mindestens einen Häufungspunkt besitzt.

(f) Es seien $\gamma \in (0,1)$ und $x^* \in \mathbb{R}^n$. Ferner sei $(x^k) \subset \mathbb{R}^n$ eine Folge, die folgender Abschätzung genüge:

$$\|x^{k+1} - x^*\|_2 \leq \gamma \|x^k - x^*\|_2 \quad \forall k \geq 0.$$

Dann gilt:
Die Folge (x^k) ist konvergent.

(g) Sei $x^* \in \mathbb{R}^n$ ein lokales Minimum von f, in dem die hinreichenden Optimalitätsbedingungen zweiter Ordnung erfüllt sind (insbesondere ist also $\nabla f(x^*) = 0$). Dann gilt:
Es gibt eine Umgebung von x^*, in der sich keine weitere Gradientennullstelle befindet.

(h) Sei $x^* \in \mathbb{R}^n$ ein lokales Minimum von f. Wir wenden das lokale Newtonverfahren für Optimierungsprobleme (Algorithmus 2.6.6) auf f an. Es erzeuge eine gegen x^* konvergente Folge. Dann gilt:
In x^* gelten die hinreichenden Optimalitätsbedingungen zweiter Ordnung.

(i) Wir wenden das globalisierte Newtonverfahren (Algorithmus 2.6.10) auf f an. Es erzeuge die Folge $(x^k) \subset \mathbb{R}^n$. Dann gilt:
Jeder Häufungspunkt von (x^k) ist stationär.

(j) Wir betrachten ein lokales newtonartiges Verfahren zur Minimierung von f. Das Verfahren erzeuge unter Verwendung der Matrizenfolge $(M_k) \subset \mathbb{R}^{n\times n}$ mit M_k positiv definit, symmetrisch für alle k, die Folge von Iterierten (x^k). Die Folge (x^k) konvergiere gegen ein lokales Minimum x^* von f, in dem die hinreichenden Optimalitätsbedingungen zweiter Ordnung gelten. Dann gilt:
Falls $M_k \to \nabla^2 f(x^*)$ für $k \to \infty$ gilt, so liegt q-superlineare Konvergenz von (x^k) gegen x^* vor.

(k) Quasi-Newton-Verfahren sind newtonartige Verfahren.

Aufgabe 3 (ca. 3+4+2+4+3=16 Punkte)
Es sei $f : \mathbb{R}^n \to \mathbb{R}$ eine stetig differenzierbare Funktion.
Wir betrachten das folgende Abstiegsverfahren:

0. Wähle $\beta \in (0,1), \gamma \in (0,1)$ und einen Startpunkt $x^0 \in \mathbb{R}^n$.
Für $k = 0,1,2,\ldots$:
1. Falls $\nabla f(x^k) = 0$: STOP.
2. Bestimme $s^k \in \mathbb{R}^n$ durch

$$s^k = -\frac{\partial f}{\partial x_j}(x^k) \cdot e_j,$$

wobei $e_j \in \mathbb{R}^n$ den j-ten kanonischen Einheitsvektor bezeichnet und $j \in \{1,2,\ldots,n\}$ ein Index ist mit $\left|\frac{\partial f}{\partial x_j}(x^k)\right| = \max_{1\le i\le n} \left|\frac{\partial f}{\partial x_i}(x^k)\right|$.
3. Ermittle die Schrittweite $\sigma_k \in (0,1]$ mit Hilfe der Armijoregel (Parameter β und γ).
4. Setze $x^{k+1} = x^k + \sigma_k s^k$.

(a) Zeigen Sie, dass Schritt 3 wohldefiniert ist.

(b) Zeigen Sie, dass jede Teilfolge von Suchrichtungen $(s^k)_K$ zulässig ist.

(c) Zeigen Sie, dass $f(x^{k+1}) < f(x^k)$ für alle k gilt.

(d) Zeigen Sie, dass für jede beschränkte Teilfolge von Iterierten $(x^k)_K$ die entsprechende Schrittweitenteilfolge $(\sigma_k)_K$ zulässig ist.

Tipp: Begründen Sie zunächst, dass der angegebene Algorithmus ein Spezialfall des allgemeinen Abstiegsverfahrens (Algorithmus 2.4.1) ist. Benutzen Sie dann den folgenden Satz aus der Vorlesung:

Satz 2.5.2. *Sei $f : \mathbb{R}^n \to \mathbb{R}$ stetig differenzierbar und die Teilfolge $(x^k)_K$ sei beschränkt [...]. Weiter gebe es eine streng monoton wachsende Funktion $\varphi : [0,\infty) \to [0,\infty)$, so dass die durch das allgemeine Abstiegsverfahren erzeugten Suchrichtungen folgender Bedingung genügen:*

$$\|s^k\|_2 \ge \varphi\left(\frac{-\nabla f(x^k)^T s^k}{\|s^k\|_2}\right) \quad \forall k \in K.$$

Dann ist die durch die Armijoregel erzeugte Schrittweitenteilfolge $(\sigma_k)_K$ zulässig.

(e) Es sei (x^k) eine durch den angegebenen Algorithmus erzeugte Folge von Iterierten. Begründen Sie, dass jeder Häufungspunkt von (x^k) ein stationärer Punkt von f ist.