8. Minimierungsprobleme ohne Nebenbedingungen

8.1 Motivation

(i) Seien $f : \mathbb{R}^n \to \mathbb{R}^n$ zweimal stetig differenzierbar und $Df(x)$ invertierbar für alle $x \in \mathbb{R}^n$.

Die folgenden Probleme sind äquivalent:

Problem 7: Löse $f(x) = 0$.

Problem 8: Bestimme $\min\limits_{x \in \mathbb{R}^n} h(x), \quad h(x) = \|f(x)\|_2^2.$

(ii) Seien $A \in \mathbb{R}^{n \times n}$ positiv definit und $b \in \mathbb{R}^n$.

Die folgenden Probleme sind äquivalent:

Problem 2: Löse $Ax = b$.

Problem 8: Bestimme $\min\limits_{x \in \mathbb{R}^n} h(x), \quad h(x) = \frac{1}{2} x^T A x - x^T b.$

Problem 8: Sei $h : \mathbb{R}^n \to \mathbb{R}$, $h \in C^2(\mathbb{R}^n)$.

$$\text{Bestimme} \qquad \min_{x \in \mathbb{R}^n} h(x)$$

Definiere den Gradienten von h

$$\nabla h(x) = \begin{pmatrix} \frac{\partial h}{\partial x_1} & \dots & \frac{\partial h}{\partial x_n} \end{pmatrix} (x)$$

und die Hesse-Matrix von h

$$H(x) = \begin{pmatrix} \frac{\partial^2 h}{\partial x_1^2} & \cdots & \frac{\partial^2 h}{\partial x_1 \partial x_n} \\ \vdots & & \vdots \\ \frac{\partial^2 h}{\partial x_n \partial x_1} & \dots & \frac{\partial^2 h}{\partial x_n^2} \end{pmatrix} (x), \quad x \in \mathbb{R}^n.$$

Das Problem 8 ist lösbar, falls $\nabla h(x^*) = 0$ und $H(x^*)$ positiv definit ist.

8.2 Minimierung nach Oren und Luenberger

Wähle beliebigen Startvektor $x^{(0)} \in \mathbb{R}^n$ und $TOL > 0$

Setze $H_0 = I$ oder wähle eine andere positiv definite Matrix H_0

$k = 0$

while $\|\nabla h(x^{(k)})\| \geq TOL$

1. Berechne $s_k = H_k \nabla h(x^{(k)})$
2. $j = 0$

   while $h(x^{(k)} - 2^{-j} s_k) > h(x^{(k)})$

   $\quad j = j + 1$

   end
3. Setze $x^{(k+1)} = x^{(k)} - 2^{-j} s_k$
4. Wähle $\gamma_k > 0$ und $\theta_k \geq 0$ (z.B. $\gamma_k = 1$ und $\theta_k = 0$)
5. Für $p_k = x^{(k+1)} - x^{(k)}$ und $q_k = \nabla h(x^{(k+1)}) - \nabla h(x^{(k)})$ setze

$$H_{k+1} = \Psi(\gamma_k, \theta_k, H_k, p_k, q_k)$$

end

Bemerkung:

(*i*)

$$\Psi(\gamma,\theta,H,p,q) = \gamma H \quad + \quad \left(1+\gamma\theta\frac{q^T H q}{p^T q}\right)\frac{pp^T}{p^T q} - \gamma\frac{1-\theta}{q^T H q} H q q^T H$$
$$- \quad \frac{\gamma\theta}{p^T q}(pq^T H + Hqp^T).$$

(*ii*) Die Newton-Richtung $[H(x^{(k)})]^{-1}\nabla h(x^{(k)})$ wird durch geeignete einfacher zu berechnende Abstiegsrichtung $s_k = H_k \nabla h(x^{(k)})$ ersezt.

(*iii*) Man spricht von einem Quasi-Newton-Verfahren, falls gilt

$$x^{(k+1)} = x^{(k)} + H_{k+1}(\nabla h(x^{(k+1)}) - \nabla h(x^{(k)})), \quad k \in \mathbb{N}_0.$$

Satz: Ein Minimierungsverfahren der Oren-Luenberger-Klasse hat folgende Eigenschaften:

(i) Falls für ein $k \in \mathbb{N}_0$ sowohl H_k positiv definit ist als auch $\nabla h(x^{(k)}) \neq 0$ gilt, so ist die Matrix $H_{k+1} = Y(\gamma_k, \theta_k, H_k, p_k, q_k)$ wohldefiniert und positiv definit. Ferner erfüllt H_{k+1} die Quasi-Newton- Gleichung $H_{k+1} q_k = p_k$.

(ii) Für $h(x) = \frac{1}{2} x^T A x + x^T b + c$ mit positiv definiter $n \times n$ Matrix A liefert das Verfahren in höchstens n Schritten das Minimum, soferne die exakte Liniensuche verwendet wird.

(iii) Ist $H(x^*)$ positiv definit, und ist $H(x)$ für x^* Lipschitz-stetig, d.h.

$$\exists L \geq 0 \ : \ \|H(x) - H(x^*)\| \leq L \|x - x^*\| \quad \text{für alle } x \text{ in einer Umgebung von } x^*,$$

so konvergiert die Folge $\{x^{(k)}\}_{k \in \mathbb{N}}$ lokal superlinear gegen x^* (in Spezialfällen sogar quadratisch).

Beweisidee (iii): siehe J.Stoer, Einführung in die numerische Mathematik I, S. 278-279.

Bildkompression und Optimierung

Originalbild $A \in \mathbb{R}^n$

JPEG:

$$A \mapsto \tilde{A} \in \mathbb{R}^m, \quad m << n$$

mit Hilfe

diskreter Kosinus-Transformation (DCT)

Huffman- oder arithmetischer Kodierung

Rekonstruiertes Bild (Digitalkamera)

Originalbild $A \in \mathbb{R}^n$

Alternative: nimm direkt

$$\tilde{A} \in \mathbb{R}^m, \quad m << n \text{ auf}$$

mit Hilfe

von Zufallsmatrizen und

linearer Programmierung

Rekonstruiertes Bild (One-Pixel-Camera)