Technische Universität München

Zentrum Mathematik

Prof. Dr. P. Gritzmann, Dipl.-Inf. Dipl.-Math. S. Borgwardt

Optimierung 3, SS 2009

Übungsblatt 11

Aufgabe 11.1

Seien $\rho \in (0,1)$ und $\kappa \in (\rho, 1)$.

a) Zeigen Sie: Für jede stetig differenzierbare Funktion $\varphi : \mathbb{R}^n \to \mathbb{R}$ und jeden Startpunkt $x_0 \in \mathbb{R}^n$ erfüllt die Schrittweite $\lambda^{(k)}$, die beim Gradientenverfahren mit exakter line search gewählt wird, die zweite Wolfe-Powell Bedingung $\left(\varphi'(x^{(k)})\right)^T s^{(k)} \geq \kappa \left(\varphi'(x^{(k-1)})\right)^T s^{(k)}$.

b) Zeigen Sie anhand eines Beispiels, dass das Gradientenverfahren mit exakter line search die erste Wolfe-Powell Bedingung $\varphi(x^{(k)}) \leq \varphi(x^{(k-1)}) + \lambda^{(k)} \rho \left(\varphi'(x^{(k-1)})\right)^T s^{(k)}$ verletzen kann.

Lösung zu Aufgabe 11.1

a) Beim Gradientenverfahren ist $s^{(k)} = -\varphi'(x^{(k-1)})$, also ist die rechte Seite der zweiten Wolfe-Powell Bedingung

$$\kappa \left(\varphi'(x^{(k-1)})\right)^T s^{(k)} = -\kappa \|\varphi'(x^{(k-1)})\|_2^2 \leq 0.$$

Da andererseits $\left(s^{(k)}\right)^T s^{(k+1)} = 0$ ist, und $s^{(k+1)} = -\varphi'(x^{(k)})$ gilt, ist die linke Seite immer gleich 0. Damit ist die Ungleichung auf jeden Fall erfüllt.

b) Wir betrachten die Funktion $\varphi : \mathbb{R} \to \mathbb{R}$, $\varphi(x) := x^{2k}$ für $k \in \mathbb{N}$ hinreichend groß und den Startpunkt $x_0 = -2$. Damit ergibt sich die Abstiegsrichtung $s^{(1)} = -2k \cdot (-2)^{2k-1} = 2^{2k}k$, also ist $\phi(\lambda) = (-2 + \lambda \cdot 2^{2k}k)^{2k}$. Mit $\phi'(\lambda) = 2k(-2 + \lambda \cdot 2^{2k}k)2^{2k}k = 2^{2k+1}k^2(-2 + \lambda \cdot 2^{2k}k)$ ergibt sich als Schrittweite $\lambda^{(1)} = \frac{2}{2^{2k}k} = \frac{1}{2^{2k-1}k}$ und damit $x^{(1)} = -2 + \frac{1}{2^{2k-1}k} \cdot 2^{2k \frac{1}{2^{2k-1}k} \rho \cdot 2k \cdot 2^{2k-1} \cdot 2^{2k} k} k = -2 + 2 = 0$.

Jetzt setzen wir das in die erste Wolfe-Powell Bedingung ein:

$$\begin{aligned}
\varphi(x^{(1)}) &\leq \varphi(x^{(0)}) + \lambda^{(1)}\rho\left(\varphi'(x^{(0)})\right)^T s^{(1)} \\
&\Leftrightarrow 0 \leq (-2)^{2k} + \frac{1}{2^{2k-1}k}\rho \cdot 2k \cdot (-2)^{2k-1} \cdot 2^{2k}k \\
&\Leftrightarrow 0 \leq 2^{2k} - 2^{2k+1}k\rho \\
&\Leftrightarrow 0 \leq 2^{2k}(1-\rho 2k) \\
&\Leftrightarrow \rho \leq \frac{1}{2k}
\end{aligned}$$

Für jede Wahl von ρ gibt es also ein hinreichend großes k, so dass die verwendete Schrittweite nicht der ersten Wolfe-Powell Bedingung genügt.

Aufgabe 11.2

Konstruieren Sie ein Paar (φ, x_0) einer Funktion $\varphi : \mathbb{R}^n \to \mathbb{R}$ $(n \geq 2)$ und eines Punktes $x_0 \in \mathbb{R}^n$ so, dass das Gradientenverfahren für φ mit exakter Schrittweitenbestimmung und nichtstationärem Startpunkt x_0 einen stationären Punkt von φ liefert, der kein lokales Minimum ist, obwohl ein solches existiert.

Lösung zu Aufgabe 11.2

Wir betrachten die Funktion $\varphi : \mathbb{R}^2 \to \mathbb{R}$ definiert für $x = (\xi_1, \xi_2)^T \in \mathbb{R}^2$ durch

$$\varphi(x) := \xi_1^4 - \xi_1^2 + \xi_2^2$$

Die lokalen (und auch globalen) Minimalpunkte von φ sind offensichtlich $x^* = (\pm\frac{1}{2}\sqrt{2}, 0)^T$, es gilt $\varphi(x^*) = -\frac{1}{4}$.

Wir wählen als Startpunkt $x_0 = (0,1)^T$. Mit $\varphi'(x) = (4\xi_1^3 - 2\xi_1, 2\xi_2)^T$ ergibt sich $\varphi'(x_0) = (0,1)^T$. Wir benutzen also einen Schritt $x_0 + \lambda(0,-1)^T$ mit exakt bestimmter Schrittweite $\lambda \in [0,\infty[$. φ nimmt auf dem Strahl $(0,1)^T + \lambda(0,-1)^T$ ein globales Minimum im Punkt $(0,0)^T$ an, wir wählen daher $\lambda = 1$ und erhalten $x_1 = (0,0)^T$. Das Gradientenverfahren läuft aus dem Sattelpunkt x_1 nicht mehr weg, da $\varphi'(x_1) = (0,0)^T$.

Die konstruierte Funktion lässt sich in der Form $\varphi : \mathbb{R}^n \to \mathbb{R}$ mit $\varphi(x) := \xi_1^4 - \xi_1^2 + \xi_2^2$ für $x = (\xi_1, \ldots, \xi_n)^T$ direkt auf den $\mathbb{R}^n$ übertragen.