EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

Mathematisch-Naturwissenschaftliche Fakultät

Fachbereich Mathematik

Prof. Dr. Andreas Prohl
Benita Tscheppe

Nichtlineare Optimierung

Sommersemester 25 **Tübingen, 01.05.2025**

Übungsaufgaben 3

Problem 1. Consider the quadratic function $f : \mathbb{R}^n \to \mathbb{R}$ of the form:

$$f(\mathbf{x}) := \frac{1}{2}\langle \mathbf{x}, \boldsymbol{Q}\mathbf{x}\rangle + \langle \mathbf{c}, \mathbf{x}\rangle + \delta\,,$$

where $\boldsymbol{Q} \in \mathbb{R}^{n\times n}_{\text{sym}}$, $\mathbf{c} \in \mathbb{R}^n$, and $\delta \in \mathbb{R}$. Now fix a position $\mathbf{x} \in \mathbb{R}^n$, as well as a *search direction* $\mathbf{r} \in \mathbb{R}^n$. It was detailed in the lecture that the *optimal* step size $\alpha_* > 0$ to go in direction $\mathbf{r} \in \mathbb{R}^n$ is

$$\alpha_* = -\frac{\langle \boldsymbol{Q}\mathbf{x} - \mathbf{c}, \mathbf{r}\rangle}{\langle \boldsymbol{Q}\mathbf{r}, \mathbf{r}\rangle}\,.$$

Show that this α_* satisfies the following condition in *Armeijo's rule*

$$f(\mathbf{x} + \alpha_*\mathbf{r}) - f(\mathbf{x}) \le \alpha_*\gamma\,\langle \nabla f(\mathbf{x}), \mathbf{r}\rangle \qquad \forall\,\gamma \in (0, \frac{1}{2}]\,,$$

while it does *not* hold for $\gamma > \frac{1}{2}$. What does this observation imply if you compare the gradient descent method with Armeijo's rule with the descent method with *explicit* step size selection — which is available for quadratic functions?

Hints: Use Taylor's formula and recall what it means that $\mathbf{r}$ is a 'search direction'.

Problem 2. In the lecture we did not complete yet the proof of a theorem which states convergence of a subsequence $\{\mathbf{x}^{(t')}\}_{t'\in\mathbb{N}_0} \subset \mathbb{R}^n$ towards a stationary point $\mathbf{x}^*$ of a coercive $f \in C^1(\mathbb{R}^n)$ — where the original sequence $\{\mathbf{x}^{(t)}\}_{t\in\mathbb{N}_0} \subset \mathbb{R}^n$ is generated by the gradient descent method with Armeijo's step size rule. What we know so far is that the *product* of two relevant sequences tends to zero:

$$\alpha^{(t')}\|\nabla f(\mathbf{x}^{(t')})\|^2 \to 0 \qquad (t' \uparrow \infty)\,,$$

which allows for two possibilities:

$$\text{(i)}\quad \lim_{t'\uparrow\infty} \alpha^{(t')} = 0 \qquad \text{or} \qquad \text{(ii)}\quad \nabla f(\mathbf{x}^*) = \mathbf{0}\,.$$

Here $\mathbf{x}^* = \lim_{t'\uparrow\infty} \mathbf{x}^{(t')}$, with $\nabla f(\mathbf{x}^*) = \lim_{t'\uparrow\infty} \nabla f(\mathbf{x}^{(t')})$ since $f \in C^1(\mathbb{R}^n)$. Conclude from it that $\mathbf{x}^*$ is stationary point of f.

Hints: **1.** Use Armeijo's step size criterion to verify the result if (i) applies. For this purpose recall that each $\alpha^{(t')}$ is the *largest* number out of $\{1, \beta, \beta^2, \cdots\}$ such that Armeijo's step size criterion applies. But this means that the *next largest* 'step size candidate' $\frac{\alpha^{(t')}}{\beta}$ still *violates* Armeijo's step size criterion, *i.e.*, a *strict* (opposite) inequality sign appears there for $\frac{\alpha^{(t')}}{\beta}$.

2. Use the mean value theorem to replace here the appearing difference of evalutions with f by a single evaluation of ∇f. Now complete the argument that settles that $\mathbf{x}^*$ is stationary point, by exploiting that $\gamma < 1$.

Abgabe: 08.05.2025.