Exercise sheet 9

Exercise 1 [Criteria for admissibility of search directions and step sizes]

Let $f : \mathbb{R}^n \to \mathbb{R}$ be C^1 and let (x^k), (s^k) as well as (σ_k) be generated by Algorithm 2.8.1 from the lecture. Show that:

a) The sequence of search directions $(s^k)_K$ is admissible, if

$$-\nabla f(x^k)^\top s^k \geq \phi(\|\nabla f(x^k)\|)\|s^k\| \quad \forall\, k \in K,$$

where $\phi : [0, \infty) \to [0, \infty)$, $\phi(0) = 0$ is a strict monotonic and increasing function.

b) The sequence of step lengths $(\sigma_k)_K$ is admissable, if there is for every $\varepsilon > 0$ a $\delta(\varepsilon) > 0$ such that holds:

$$\frac{-\nabla f(x^k)^\top s^k}{\|s^k\|} \geq \varepsilon \quad \text{for infinitely many } k \in K$$
$$\Longrightarrow \; f(x^k) - f(x^k + \sigma_k s^k) \geq \delta(\varepsilon) \quad \text{for infinitely many } k \in K.$$

Exercise 2 [Admissibility of the Armijo step sizes]

Prove Lemma 2.79 from the lecture using Exercise 1.

Exercise 3 [Efficient step sizes]

Let $f \colon \mathbb{R}^n \to \mathbb{R}$ be C^1 and s^k be a descent direction of f in x^k. The step size $\sigma_k > 0$ is called efficient, if

$$f(x^k + \sigma_k s^k) \leq f(x^k) - \theta \left(\frac{\nabla f(x^k)^\top s^k}{\|s^k\|} \right)^2$$

holds, where $\theta > 0$ is a constant. Now let (x^k), (s^k) and (σ^k) be generated by Algorithm 2.8.1 and (40) from the lecture holds true. Show that:

a) Let $(\sigma_k)_K$ be a subsequence, for which all step sizes σ_k, $k \in K$ are efficient, then the subsequence of step sizes $(\sigma_k)_K$ is admissible.

b) Conclude that the Curry step size from Sheet 4, Exercise 3 is admissible.

Exercise 4 [Non admissible search directions]

If one chooses search directions, which are almost perpendicular to the gradient direction, it can happen that Algorithm 2.8.1 does not converge to the optimal solution. As an example we consider $f(x_1, x_2) = \frac{1}{2}(x_1^2 + x_2^2)$ with search directions

$$s^k = g^k - \frac{1}{2^{k+3}} \nabla f(x^k),$$

where g^k is chosen such that $\langle g^k, \nabla f(x^k) \rangle = 0$ and $\|s^k\|_2 = \|\nabla f(x^k)\|_2$. Show that Algorithm 2.8.1 with s^k and admissible step lengths does not converge for any initial guess $x^0 \in \mathbb{R}^n \setminus \{0\}$ to the optimal solution $x^* = 0$ and that x^* is not a accumulation point of (x^k).

Exercise 5 [Non admissible step sizes by the Armijo rule]

We consider Algorithm 2.8.1 with $f \colon \mathbb{R}^n \to \mathbb{R}$ in C^1 and the search directions $s^k = -2^{-k} \nabla f(x^k)$.

a) Proof that every subsequence of search directions $(s^k)_K$ is admissible.

b) Proof that the Armijo rule in general does not generate admissible step sizes σ_k. Use the example

$$f(x) = \frac{x^2}{8}, \quad \text{initial guess} \quad x^0 > 0$$

for your proof, i.e. show that there is no admissible subsequence $(\sigma_k)_K$.

c) What is the reason for non admissibility of the step sizes?