Exercise sheet 10

Let $f \in C^1$, $\gamma \in (0, 1/2)$ and $\eta \in (\gamma, 1)$. Then $\sigma > 0$ satisfies the Powell-Wolfe condition, if

$$f(x + \sigma s) - f(x) \leq \sigma\gamma\nabla f(x)^\top s, \tag{PW1}$$

$$\nabla f(x + \sigma s)^\top s \geq \eta\nabla f(x)^\top s. \tag{PW2}$$

Exercise 1 [Existence and admissibility of the Powell-Wolfe step size]

a) Visualize (PW1) and (PW2) with a sketch in dimension 1.

b) Prove Lemma 2.88 from the lecture.

c) Prove Lemma 2.90.

Implementation of the Powell-Wolfe step size rule:

1. If $\sigma = 1$ satisfies (PW1) go to step 3
2. Find the biggest number $\sigma_- \in \{2^{-1}, 2^{-2}, \dots\}$, such that $\sigma = \sigma_-$ satisfies (PW1). Set $\sigma_+ = 2\sigma_-$ and go to step 5.
3. If $\sigma = 1$ satisfies (PW2), Stop with $\sigma = 1$.
4. Find the smallest number $\sigma_+ \in \{2, 2^2, 2^3, \dots\}$, such that (PW1) for $\sigma = \sigma_+$ is not satisfied. Set $\sigma_- = \frac{\sigma_+}{2}$.
5. As long as $\sigma = \sigma_-$ does not satisfy (PW2) do the following:

   5.1 Set $\sigma = \frac{\sigma_- + \sigma_+}{2}$.

   5.2 If σ satisfies (PW1), set $\sigma_- = \sigma$, otherwise set $\sigma_+ = \sigma$.
6. Stop with $\sigma = \sigma_-$.

Exercise 2 [Powell-Wolfe step size algorithm]

Prove that under the assumptions of Lemma 2.88 the above algorithm terminates after finitely many steps with a step size $\sigma > 0$ satisfying (PW1) and (PW2).

Exercise 3 [Global Newton method for strongly convex f]

Let $f\colon \mathbb{R}^n \to \mathbb{R}$ be in C^2 and strongly convex. We apply the global Newton method for optimization problems Algorithm 2.9.1 to f. Though we do not use gradient steps, i.e. we always use the Newton step. We choose $\gamma \in (0, 1/2)$. Prove the following:

a) The described algorithm is well-defined.

b) The sequence (x^k) generated by the algorithm converges to the unique minimum x^* of f.

c) For all sufficiently large $k \in \mathbb{N}$ there holds $\sigma_k = 1$.

d) The sequence (x^k) converges q-superlinearly to x^*.

e) If $\nabla^2 f(x)$ is Lipschitz continuous in a neighbourhood of x^*, the convergence rate is q-quadratic.

Exercise 4 [Programming exercise: Globalized Newton and inverse BFGS-Method]

Repeat Exercise 4 a) from exercise sheet 4 using the globalized Newton method (Algorithm 2.9.1) for `tol` $= 10^{-3},\ 10^{-4}, \ldots,\ 10^{-12}$. Implement the following function

```
function[x] = gnewt(fct,gradfct,hessfct,x0,tol,maxsteps)
```

where `hessfct` is a function handle for the evaluation of $\nabla^2 f(x)$, $\alpha_1 = \alpha_2 = 10^{-6}$ and $p = 0.1$. Use your function for the calculation of the Armijo stepsize from Sheet 4. Moreover implement the globalized, inverse BFGS method (Algorithm 2.11.1) and apply it for solving Exercise 4 a) from exercise sheet 4 with `tol` $= 10^{-3},\ 10^{-4}, \ldots,\ 10^{-12}$. Implement the following function

```
function[x] = bfgs(fct,gradfct,x0,tol,maxsteps)
```

with $\eta = 0.9$. Moreover write a function which calculates the Powell-Wolfe step size similar to your function which calculates the Armijo step size. Create a table for each method which contains $\|x^k - (1,1)^\top\|_2$, $f(x^k)$, $\|\nabla f(x^k)\|_2$, σ_k and the number of function and gradient evaluations in the line search in step k for all k in the case `tol` $= 10^{-9}$. Describe your results. What convergence rate do you observe?
Hand in by email (philip.trautmann@uni-graz.at) until 07.01.2019, 23:59 o'clock.