Exercise sheet 3

Exercise 1 [Goldstein's rule]

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a differentiable function. We assume that its gradient is Lipschitz-continuous with modulus $L > 0$. Let $x \in \mathbb{R}^n$, let $p \in \mathbb{R}^n$ be a descent direction, let $c \in (0, 1/2)$. We set $\phi(\alpha) = f(x + \alpha p)$ for $\alpha \geq 0$. We consider the following rules:

$$\begin{aligned} A(\alpha) &\iff \phi(\alpha) \leq \phi(0) + c\phi'(0)\alpha \\ B(\alpha) &\iff \phi(\alpha) > \phi(0) + (1-c)\phi'(0)\alpha. \end{aligned}$$

a) Show the existence of $\bar{\alpha} > 0$ such that for all $\alpha \in [0, \bar{\alpha}]$, $B(\alpha)$ does not hold.

b) Prove that $\phi(\alpha) < \phi(0)$ is satisfied for all α satisfying both rules.

c) Propose an algorithm for finding a steplength satisfying both rules and justify that it terminates.

Exercise 2 [A simple gradient-descent method]

In this exercise, $\| \cdot \|$ denotes the Euclidean norm. We recall first two results, which will be used for this exercise.

- Let D be a closed subset of $\mathbb{R}^n$. Let $F : D \to D$ be a contraction, that is to say, a mapping such that there exists $\lambda \in [0, 1)$ such that for all x and $y \in D$,

  $$\|F(y) - F(x)\| \leq \lambda \|y - x\|.$$

  Then, there exists a unique $\bar{x} \in D$ such that $F(\bar{x}) = \bar{x}$. Moreover, all sequences $(x_k)_{k \in \mathbb{N}}$ satisfying $x_{k+1} = F(x_k)$ (for all $k \in \mathbb{N}$) converge to $\bar{x}$.

- Let S_n denote the set of symmetric matrices of size n. For all $M \in S_n$, denote by $\sigma^-(M)$ its smallest eigenvalue and by $\sigma^+(M)$ its largest eigenvalue. The operator norm of a symmetric matrix M (associated with the Euclidean norm) is given by:

  $$\|M\| = \max\big(|\sigma^-(M)|, |\sigma^+(M)|\big).$$

We consider now a twice continuously differentiable function f. We assume that there exists $\bar{x}$ such that $Df(\bar{x}) = 0$ and such that $D^2 f(\bar{x})$ is positive definite. We aim at proving the following result: there exist $\delta > 0$ and $\rho > 0$ such that for all $x \in \mathbb{R}^n$ with $\|x - \bar{x}\| \leq \delta$, the sequence $(x_k)_{k \in \mathbb{N}}$ defined below converges to $\bar{x}$:

$$x_0 = x, \quad x_{k+1} = x_k - \rho \nabla f(x_k), \ \forall x \in \mathbb{N}.$$

a) It is possible to show that there exist $\delta > 0$, $a > 0$ and $A \geq a$ such that for all $x \in \mathbb{R}^n$ with $\|x - \bar{x}\| \leq \delta$,

$$a \leq \sigma^-(D^2 f(x)) \leq \sigma^+(D^2 f(x)) \leq A.$$

Prove the existence of $\rho > 0$ and $\lambda \in [0, 1)$ such that for all $x \in \mathbb{R}^n$ with $\|x - \bar{x}\| \leq \delta$, we have $\|\mathrm{Id} - \rho D^2 f(x)\| \leq \lambda$.

b) Prove the main result of the exercise, by applying the fixed-point theorem.