Exercise Sheet: Gradient Decent Method

Exercises

Exercise 1. Let $f : \mathbb{R} \to \mathbb{R}$, $f(x) = e^x$. By which choice of γ_k can you ensure that $\big(f(x^{(k)})\big)_{k\in\mathbb{N}}$ with

$$x^{(k+1)} = x^{(k)} - \gamma_k f'(x^{(k)})$$

converges to $\inf_{x\in\mathbb{R}} f(x)$?

Exercise 2. Let $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$. Examine the sequence $(x^{(k)})_{k\in\mathbb{N}}$ with

$$x^{(k+1)} = x^{(k)} - 2^{-k} f'(x^{(k)})$$

for convergence. Are you surprised by the result?

Exercise 3. Let $f : \mathbb{R} \to \mathbb{R}$, $f(x) = 3x^4 + 4x^3 - 36x^2 + 42$, be given. Execute gradient descent with constant step size and show that the corresponding recursively defined sequence may converge to a local minimum point, which is not a (global) minimizer, if the initial value is chosen unfavorably.

Exercise 4. Show that the derivative of the following function decreases exponentially fast for $x \searrow 0$. Thus, the gradient method, once it is near zero, only progresses very slowly.

$$f : \mathbb{R} \to \mathbb{R}, \qquad f(x) = \begin{cases} e^{-1/x}, & \text{if } x \in (0, \frac{1}{4}], \\ \big(x + 8e^{-4} - \frac{1}{4}\big)^2 - 64e^{-8} + e^{-4}, & \text{if } x \in (\frac{1}{4}, \infty), \\ 0, & \text{otherwise.} \end{cases}$$

Do you believe that one can still get a sequence *(17.1)* with constant step size that converges to a minimizer regardless of the starting value?

Exercise 5. Show that $f : \mathbb{R}^n \to \mathbb{R}$,

$$f(x) = \tfrac{1}{2}\langle x, Ax\rangle + \langle b, x\rangle + c$$

with $A \in \mathbb{R}^{n\times n}$ positive definite, as well as $b \in \mathbb{R}^n$ and $c \in \mathbb{R}$ arbitrary, is μ-convex and L-smooth and determine for which μ and L this is the case.

Exercise 6. Let $f : \mathbb{R}^n \to \mathbb{R}$ be convex and L-smooth with $L > 0$. Let $x^* \in X^*$ and $x^{(0)}$ be arbitrary, let $(x^{(k)})_{k\in\mathbb{N}}$ be the sequence from *(17.1)* with constant step size $\gamma \in (0, 1/L)$. Show that then

$$f(x^{(k)}) - f(x^*) \leq \frac{2}{k\gamma}\,\|x^{(0)} - x^*\|^2$$

holds for $k \geq 0$.

Exercise 7. In Lemma 17.22, the auxiliary function φ is in fact even $\frac{L-\mu}{2}$-smooth, as we have "given away" a factor of $1/2$ in the very first estimate. Go through the proof again and find out how the estimate in Theorem 17.23 improves if you take the $1/2$ into account.