# 1. Question: Convex Functions *(elementary)*

**1.1.** Which of the following functions are convex? (Hint: draw a picture)

(i) $f : \mathbb{R} \longrightarrow \mathbb{R},\ x \mapsto |x|$
(ii) $f : \mathbb{R} \longrightarrow \mathbb{R},\ x \mapsto \cos(x)$
(iii) $f : \mathbb{R} \longrightarrow \mathbb{R},\ x \mapsto x^2$

**1.2.** Prove that the following functions are convex.

(i) affine linear functions, i.e. $f : \mathbb{R}^d \longrightarrow \mathbb{R},\ x \mapsto a^T x + c$ for $a \in \mathbb{R}^d, c \in \mathbb{R}$,
(ii) norms, i.e. $x \mapsto \|x\|$,
(iii) sums of convex functions $f_k$, i.e. $f(x) = \sum_{k=1}^n f_k(x)$,
(iv) $F(x) := \sup_{f \in \mathcal{F}} f(x)$ for a set of convex functions $\mathcal{F}$.

# 2. Question: Lipschitz Continuous Functions *(elementary)*

**2.1.** Which of the following functions are Lipschitz

(i) $f : [1, 2] \to \mathbb{R},\ x \mapsto x^3$
(ii) $f : \mathbb{R} \longrightarrow \mathbb{R},\ x \mapsto x^2$

**2.2.** Prove the following for Lipschitz functions $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$.

(i) The composition $f \circ g : \mathbb{R} \to \mathbb{R}$ is Lipschitz.
(ii) The sum $f + g : \mathbb{R} \to \mathbb{R}$ defined by $(f + g)(x) = f(x) + g(x)$ is Lipschitz.

**2.3.** Show that any Lipschitz function $f : [a, b] \to \mathbb{R}$ defined on an interval of the form $[a, b]$ is a bounded function.

**2.4.** Show that $h : [0, 1] \to \mathbb{R}$ given by $h(x) = \sqrt{x}$, is bounded, but not Lipschitz.

# 3. Question: Optimization *(elementary)*

Consider an optimization problem

$$\begin{aligned} &\min f(x) \\ &\text{s.t. } x \in \Omega. \end{aligned} \tag{$\star$}$$

**3.1.** Prove that if $\Omega = \mathbb{R}^n$ and $f : \mathbb{R}^n \to \mathbb{R}$ is convex and differentiable, any point $\bar{x}$ that satisfies $\nabla f(\bar{x}) = 0$ is a global minimum.

**3.2.** Prove that if $f : \mathbb{R}^n \to \mathbb{R}$ is strictly convex on $\Omega$ and $\Omega$ is a convex set, the optimal solution (assuming it exists) must be unique.

**3.3.** Consider the optimization problem of ($\star$) under the additional constraint that $Ax = b,\ A \in \mathbb{R}^{m \times n}$. Prove that if $f$ is a convex function, a point $x \in \mathbb{R}^n$ is optimal to this constrained optimization problem if and only if it is feasible and $\exists \mu \in \mathbb{R}^m$ s.t.

$$\nabla f(x) = A^T \mu.$$

*Hint: Start with what the first order condition for convexity tells us about the term $\nabla f^T(x)(y - x)$ for $y :\ Ay = b$ and use the fact that $y$ with $Ay = b$ can be written as $y = x + v$, for $v \in \mathrm{Nul}(A)$.*

## 4. Question: Bregman Divergence
### (*advanced, to see what Lipschitz continuity can be used for*)

The Bregman Divergence $D_f^{(B)}$ of a continuously differentiable function $f : \mathbb{R}^d \to \mathbb{R}$ is defined as the error of the linear approximation and is related to $\mu$-strong convexity and Lipschitz continuous gradients as follows

$$\frac{\mu}{2}\|x-x_0\|^2 \overset{\substack{\mu\text{-strongly convex}\\ \text{(definition)}}}{\leq} \underbrace{f(x) - \overbrace{f(x_0) - \langle \nabla f(x_0), x - x_0\rangle}^{\text{linear approximation}}}_{=:D_f^{(B)}(x,x_0)} \overset{\substack{\text{L-Lipschitz gradient}\\ \text{(Descent Lemma)}}}{\leq} \frac{L}{2}\|x-x_0\|^2$$

For $\mu = 0$ this is simply the convexity condition. So non-negativity of the Bregman divergence implies convexity. The $L$-Lipschitz gradients provide us with an upper bound on the Bregman divergence on the other hand which immediately results in an upper bound on $f$

$$f(x) = f(x_0) + \langle \nabla f(x_0), x - x_0\rangle + \underbrace{D_f^{(B)}(x, x_0)}_{\leq \frac{L}{2}\|x-x_0\|^2}. \tag{UB}$$

Prove for functions $f$ with $L$-Lipschitz gradients, we have for all $x_0$

$$\min_x f(x) \leq f(x_0) - \frac{1}{2L}\|\nabla f(x_0)\|^2$$

by minimizing the upper bound (UB). What is the minimizer of the upper bound?
*Hint: Try minimizing first w.r.t $x : \|x - x_0\| = r$ and then $r$. Additionally you will need the Cauchy-Schwartz inequality, whereby, for vectors $\mathbf{u}$ and $\mathbf{v}$: $|\langle \mathbf{u}, \mathbf{v}\rangle| \leq \|\mathbf{u}\|\|\mathbf{v}\|$.*

---

If you have any questions or feedback, please feel free to contact me via E-mail at hannah.kuempel@stat.uni-muenchen.de!!

Also, thank you to Felix Benning & Prof. Dr. Simon Weißmann, Andy Hammerlindl, Kevin Jamieson & Anna Karlin, and A.A. Ahmadi whose exercises this sheet was inspired by.