Exercise sheet 1

You can return your solutions to questions 6 and 7 to get them marked. If so, please upload them on Moodle before Monday 7/2 at 12noon.

1. Show that if $f : \mathbb{R}^n \to \mathbb{R}$ is convex, then the sublevel sets $S_t = \{x \in \mathbb{R}^n : f(x) \leq t\}$ are convex, for all t. Is the converse true? Prove or give a counterexample.

2. Show that if $f : \mathbb{R}^n \to \bar{\mathbb{R}}$ is convex, then $g(x,t) = tf(x/t)$ is convex for $t > 0$. What is the domain of g?

3. (a) Show that if $g : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ is convex, then $f(x) = \inf_{y\in\mathbb{R}^m} g(x,y)$ is convex. (b) Assuming g is a convex quadratic, i.e., $g(x,y) = \langle x, Ax\rangle + \langle y, Cy\rangle + 2\langle x, By\rangle$, where $\left[\begin{smallmatrix} A & B \\ B^T & C \end{smallmatrix}\right] \succ 0$, give an explicit expression for $f(x)$.

4. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function. (a) Show that f is m-strongly convex with respect to the Euclidean norm iff $f - (m/2)\|x\|_2^2$ is convex. (b) Show that ∇f is L-Lipschitz with respect to the Euclidean norm iff $(L/2)\|x\|_2^2 - f$ is convex.

5. Show that if $f : \mathbb{R}^n \to \mathbb{R}$ is convex and L-smooth, and $x^* \in \mathbf{int}\,\mathbf{dom}(f)$ is a minimizer of f, then for any $y \in \mathbf{dom}(f)$
$$f(y) - f(x^*) \leq \frac{L}{2}\|y - x^*\|^2.$$
Show further, that if $\mathbf{dom}(f) = \mathbb{R}^n$, then for all $y \in \mathbb{R}^n$
$$\frac{1}{2L}\|\nabla f(y)\|_*^2 \leq f(y) - f(x^*).$$

6. (*) Show that if $f : \mathbb{R}^n \to \mathbb{R}$ is m-strongly convex, and $x^* \in \mathbf{int}\,\mathbf{dom}(f)$ is the minimizer of f, then for any $y \in \mathbf{dom}(f)$
$$\frac{m}{2}\|y - x^*\|^2 \leq f(y) - f(x^*) \leq \frac{1}{2m}\|\nabla f(y)\|_*^2.$$

7. (*) Prove that the following functions are convex on their domain:
   (a) $f(x) = \|Ax - b\|_2^2$ where $x \in \mathbb{R}^n$
   (b) $f(x) = \log(\sum_{i=1}^n e^{x_i})$ where $x \in \mathbb{R}^n$
   (c) $f(x)$ = sum of k largest components of x, where $x \in \mathbb{R}^n$ and $k \in \{1,\dots,n\}$. (for example, $f(x) = \max_{i=1,\dots,n} x_i$ when $k = 1$, and $f(x) = x_1 + \dots + x_n$ when $k = n$.)
   (d) $f(X)$ = largest eigenvalue of X (X real symmetric $n \times n$ matrix)
   (e) $f(X) = -\log\det X$ where X is a symmetric positive definite matrix
   (f) $f(x,y) = \sum_{i=1}^n x_i \log(x_i/y_i)$ where $x, y \in \mathbb{R}^n_+$

   Also specify which functions are smooth, in which case provide an expression for the gradient and Hessian (if applicable).

8. Prove that the gradient method, with the following backtracking line search, converges at the rate $O(1/k)$: at each iteration k, initialize t_k to 1 and keep updating $t_k \leftarrow \beta t_k$ (where $\beta \in (0,1)$) until $f(x_k - t_k \nabla f(x_k)) \leq f(x_k) - (1/2)t_k \|\nabla f(x_k)\|_2^2$.

9. Consider the problem of minimizing a convex function $f(x)$ on a closed convex set C, i.e., we want to compute $\min_{x \in C} f(x)$. The *projected gradient method* works as follows: starting from $x_0 \in C$, let $x_{k+1} = P_C(x_k - t_k \nabla f(x_k))$ where P_C is the Euclidean projection on C defined by

$$P_C(x) = \operatorname*{argmin}_{y \in C} \|y - x\|_2^2.$$

   By adapting the convergence proof of the gradient method seen in lecture, show that the projected gradient method converges with a rate $O(1/k)$ when ∇f is assumed L-Lipschitz, and the step size t_k is fixed $t_k = t \in (0, 1/L]$.

10. Implement the gradient method and fast gradient method to minimize the following convex function (logistic regression loss)

$$f(x) = \sum_{i=1}^{N} \log \left[1 + \exp(y_i a_i^T x)\right]$$

   where $a_1, \ldots, a_N \in \mathbb{R}^n$ and $y_1, \ldots, y_N \in \{-1, +1\}$ are randomly generated. Take $N = 50$ and $n = 30$. Plot $f(x_k) - f^*$ as a function of k. Comment.