Convex Optimization and Modeling	Jun.-Prof. Matthias Hein
Exercise Sheet 9 - 09.06.2010	due: 16.06.2010

Exercise 19 - Subgradient Steepest Descent

a. **(3 Points)** Consider the minimization of the two-dimensional function

$$f(x_1, x_2) = \begin{cases} 5\sqrt{9x_1^2 + 16x_2^2}, & \text{if } x_1 > |x_2|, \\ 9x_1 + 16|x_2|, & \text{if } x_1 \leq |x_2| \end{cases}$$

using the steepest descent method, which moves from the current point in the opposite direction of the minimum norm subgradient with exact line search. Suppose that the algorithm starts anywhere within the set

$$\{(x_1, x_2) \,|\, x_1 > |x_2| > (9/16)^2 \,|x_1|\}.$$

Verify computationally that it converges to the nonoptimal point $(0, 0)$. What happens if a subgradient method with a constant stepsize is used instead? Check computationally. Provide for both cases the code you have used.

Exercise 20 - L_1-Minimization

We want to find the solution of a problem of form,

$$\min_{x \in \mathbb{R}^d} F(x) := f(x) + \lambda \|x\|_1 ,$$

where $\lambda \geq 0$ and f is a convex, continuously twice-differentiable function which has Lipschitz-continuous gradient, that is,

$$\|\nabla f(x) - \nabla f(y)\| \leq L \, \|x - y\| ,$$

where L is the Lipschitz-constant. Problems of this type arise in machine learning (loss function + l1-regularization e.g. the lasso), signal processing and image processing.

a. **(3 Points)** Show that the Lipschitz-condition implies that,

$$f(y) \leq f(x) + \langle \nabla f(x), y - x \rangle + \frac{L}{2} \|y - x\|^2 .$$

b. **(3 Points)** Using the result of the previous part, we derive the following upper bound on the objective $F(x)$. For any $z \in \mathbb{R}^d$,

$$F(x) \leq f(z) + \langle \nabla f(z), x - z \rangle + \frac{L}{2} \|z - x\|^2 + \lambda \|x\|_1 .$$

Derive the closed-form solution of the following problem, which corresponds to the minimization of the upper bound for a particular point z,

$$\arg\min_x \; \langle \nabla f(z), x - z \rangle + \frac{L}{2} \|z - x\|^2 + \lambda \|x\|_1$$

c. **(3 Points)** Implement the following iterative method,

$$x^{(k+1)} = \arg\min_x \; \left\langle \nabla f(x^k), x - x^k \right\rangle + \frac{L}{2} \left\|x^k - x\right\|^2 + \lambda \|x\|_1 ,$$

which minimizes the upper bound on the objective using the point from the previous iteration for the function. This method is called: **Iterative Shrinkage Thresholding Algorithm (ISTA)**. Use the particular quadratic function f,

$$f(x) = \|Ax - Y\|^2 \,,$$

This corresponds to the Lasso-Problem. Data and dummy code will be provided on the webpage.