Labs
Optimization for Machine Learning
Spring 2022

EPFL
School of Computer and Communication Sciences
Martin Jaggi & Nicolas Flammarion
github.com/epfml/OptML_course

Problem Set 5, March 24, 2022
(Recap on convexity and gradient descent algorithms.)

The goal of this exercise session is to consolidate your understanding of general convex theory and of the several gradient descent algorithms seen in class up until now.

Convex sets: Solve question 21 from exam 2018 (available on the github).

Gradient computation: Solve questions 7, 8 and 9 from exam 2018.

Convexity, smoothness: Solve questions 7 to 13 from exam 2019.

Gradient descent on a quadratic function. Consider the quadratic function $f(x) = \frac{1}{2}x^\top Ax + \langle b, x\rangle + c$, where A is a $d \times d$ symmetric matrix, $b \in \mathbb{R}^d$ and c in $\mathbb{R}$.

1. What are the minimal conditions on A, b and c that ensure that f is strictly convex ? For the rest of the exercise we assume that these conditions are fulfilled.
2. Is f strongly convex ?
3. Prove that f has a unique minimum x^* and give its closed form expression.
4. Show that f can be rewritten as $f(x) = \frac{1}{2}(x - x^*)^\top A(x - x^*) + f(x^*)$
5. From an initial point $x_0 \in \mathbb{R}^d$, assume we run gradient descent with step-size $\gamma > 0$ on the function f. Show that the n^{th} iterate x_n satisfies $x_n = x^* + (I_d - \gamma A)^n\ (x_0 - x^*)$, where I_d is the $d \times d$ identity matrix.
6. In which range must the step-size γ be so that the iterates converge towards x^* ?