Labs
Optimization for Machine Learning
Spring 2024

Saarland University
CISPA Helmholtz Center for Information Security
Sebastian Stich
TAs: Yuan Gao & Xiaowen Jiang
https://cms.cispa.saarland/optml24/

Problem Set 5, May 14, 2024
(Newton and Adaptive gradient methods)

Exercise 1. *Prove Lemma 6.6 in the lecture notes: With the assumptions and terminology of Theorem 6.4, and if $\mathbf{x}_0 \in X$ satisfies*

$$\|\mathbf{x}_0 - \mathbf{x}^\star\| \leq \frac{\mu}{B},$$

then the Hessians in Newton's method satisfy the relative error bound

$$\frac{\left\|\nabla^2 f(\mathbf{x}_t) - \nabla^2 f(\mathbf{x}^\star)\right\|}{\|\nabla^2 f(\mathbf{x}^\star)\|} \leq \left(\frac{1}{2}\right)^{2^t - 1}, \quad t \geq 0.$$

Exercise 2. Prove Young's inequality

Let $\mathbf{a}, \mathbf{b} \in \mathbb{R}^d$ be arbitrary vectors. Prove that for any $\gamma > 0$:

$$\mathbf{a}^\top \mathbf{b} \leq \frac{\gamma^2}{2} \|\mathbf{a}\|^2 + \frac{1}{2\gamma^2} \|\mathbf{b}\|^2 .$$

Exercise 3. Prove Cauchy-Schwarz for random variables

Suppose $A, B \in \mathbb{R}$ are random variables. Then

$$\mathbb{E}[AB] \leq \sqrt{\mathbb{E}[A^2]\mathbb{E}[B^2]}\,.$$

Hint: Apply Young's inequality.