Optimization in Machine Learning HWS 2024

Sheet 5

For the exercise class on the 21.11.2024. Hand in your solutions by 10:15 in the lecture on Tuesday 19.11.2024.

Exercise 1 (Conditional Expectation).

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, \mathcal{F} a subalgebra of \mathcal{A} and X, Y random vectors. Prove for \mathcal{F} -measurable $X \in \mathbb{R}^d$ that we have

 $\mathbb{E}[\langle X, Y \rangle \,|\, \mathcal{F}] = \langle X, \mathbb{E}[Y \,|\, \mathcal{F}] \rangle$

Let Z be a random variable. Let f(x) := f(x, Z) be a random function $(f(x, \omega) = f(x, Z(\omega))$ if you want) and its expectation

 $F(x) = \mathbb{E}[f(x)]$

Is
$$f$$
 almost surely convex if and only if F is convex? Prove or disprove both directions.

Exercise 3 (Convergence of SGD on Strongly Convex Functions). (4 Points) In the lecture we proved for *L*-smooth functions F and X_n generated by Algorithm 6 (SGD)

$$\|\nabla F(X_n)\|^2 \to 0$$
 a.s.

If we additionally have strong convexity of F, prove $||X_n - x_*|| \to 0$ almost surely.

Exercise 4 (Swap Integration with Differentiation).

(i) What formal requirements on $f: V \times \Omega \to \mathbb{R}$ with $V \subseteq \mathbb{R}$ and measure μ on Ω are needed, for the following argument using the fundamental theorem of calculus (FTC) to work?

$$\begin{split} \frac{\partial}{\partial t} \int_{\Omega} f(t_0, \omega) d\mu(\omega) &\stackrel{\text{linear}}{=} \lim_{\epsilon \to 0} \int \frac{f(t_0 + \epsilon, \omega) - f(t_0, \omega)}{\epsilon} d\mu(\omega) \\ &\stackrel{\text{FTC II}}{=} \lim_{\epsilon \to 0} \int \frac{1}{\epsilon} \int_{t_0}^{t_0 + \epsilon} \frac{\partial}{\partial t} f(t, \omega) dt d\mu(\omega) \\ &\stackrel{\text{Fubini}}{=} \lim_{\epsilon \to 0} \frac{1}{\epsilon} \int_{t_0}^{t_0 + \epsilon} \int \frac{\partial}{\partial t} f(t, \omega) d\mu(\omega) dt \\ &\stackrel{\text{def.+lin.}}{=} \frac{d}{dy} \int_{t_0}^{y} \int \frac{\partial}{\partial t} f(t, \omega) d\mu(\omega) dt \Big|_{y=t_0} \\ &\stackrel{\text{FTC I}}{=} \int \frac{\partial}{\partial t} f(t_0, \omega) d\mu(\omega). \end{split}$$

Formulate the corresponding theorem.

Universität Mannheim Prof. Simon Weißmann, Felix Benning

(4 Points)

(4 Points)

(12 Points)

(6 pts)

(ii) We want to find an example for a function, where you can not swap integration with differentiation. So for a function $f(t, \omega)$ we need some t_0 such that

$$\frac{\partial}{\partial t} \int_{\Omega} f(t_0, \omega) d\omega \neq \int_{\Omega} \frac{\partial}{\partial t} f(t_0, \omega) d\omega.$$

For this consider $f(t, \omega) = t^3 e^{-t^2 \omega}$. Prove the inequality at $t_0 = 0$ and $\Omega = [0, \infty)$. Why is this not a contradiction to (i)? (6 pts)

Hint. It is helpful to calculate the entire function

$$t \mapsto \int_0^\infty \frac{\partial}{\partial t} f(t,\omega) d\omega$$