Theoretical Assignment 2

Intro to Machine Learning, TAU

November 29, 2024

Question 1

Show that the hyperbolic set $C_2 = \{x \in \mathbb{R}_+^2 | x_1 x_2 \geq 1\}$ is convex.
As a generalization, show that $C_n = \{x \in \mathbb{R}_+^n | \prod_{i=1}^n x_i \geq 1\}$ is convex.
Hint: If $a, b \geq 0$ and $0 \leq \theta \leq 1$, then $a^\theta b^{1-\theta} \leq \theta a + (1-\theta)b$.

Question 2

The logistic loss function is frequently used in machine learning for binary classification. It is defined as:

$$f(z) = \log\left(1 + e^z\right),$$

where $z \in \mathbb{R}$.
Prove that $f(z)$ is convex.

Question 3

The following definition is a generalization of the definition of Jensen's inequality that we've seen in class for two points, to m points drawn from a distribution p:
Jensen's inequality states that for a convex function f and a probability distribution $\{p_i\}$ over points $\{\mathbf{x}_i\}$, we have:

$$f\left(\sum_{i=1}^m p_i \mathbf{x}_i\right) \leq \sum_{i=1}^m p_i f(\mathbf{x}_i).$$

Prove that $f(\mathbf{x}) = \|\mathbf{x}\|^2$ is convex using the above definition.

Question 4

Log in to Google Collab with your university email. Run this notebook, and follow the discussion points mentioned there. Add your answers to the pdf file with the solutions to the following questions.

Clarification: this is still considered a theoretical assignment as your goal is to analyze the results without implementing anything.