EECS 496-10 – Computational Learning Theory
Winter 2019
Problem Set 3

Lev Reyzin

Due: 3/15/19 by 9:30am

Instructions: Atop your problem set, please write your name and list your collaborators. You may consult outside references, but cite all the resources used (e.g. which resources on the internet you consulted). You should not, however, search for answers to these questions. All problems in this assignment require proof.

Problems

1. In lecture, the function of the slack variables we used in the optimization for "soft margin" hyperplanes has the form $\sum_i^m \xi_i$. Instead we could use $\sum_i^m \xi_i^2$. Derive the dual formulation of the optimization problem in this case.

2. Imagine that an online learning algorithm A that runs in T rounds and has an expected regret bound of $\epsilon + T/\epsilon$, where ϵ is set by the algorithm. Clearly the optimal setting is $\epsilon = \sqrt{T}$. The problem is that sometimes T is not known in advance. How do we fix this issue? We can have an algorithm A' that does the following: A' starts with a parameter ϵ_1 and runs A for T_1 rounds, then adjusts the parameter to ϵ_2 and runs A for T_2 rounds, and so on. Construct a schedule of (ϵ_i, T_i) that asymptotically achieves the $\sqrt{T}$ expected regret bound without knowing T in advance.

3. Suppose you have two coins, one perfectly fair, and one with bias toward H of $1/2+\epsilon$ for some $\epsilon > 0$. It is known that to tell which coin is biased (with probability $> 3/4$) one needs to perform at least c/ϵ^2 coin flips ($c > 0$ is some constant). Show that this implies that EXP3's (or any bandit or online learning algorithm's) asymptotic regret dependence of $T^{1/2}$ cannot be improved to $T^{1/2-\delta}$ for any constant $\delta > 0$.