Homework #4
CMSC 35401: The Interplay of Economics and Machine Learning (Winter'24)

Due Saturday 03/02, 9:00 pm

General Instructions The assignment is meant to practice your understanding of course materials, and some of them are challenging. You are allowed to discuss with fellow students, however please write up your solutions independently (e.g., start writing solutions after a few hours of any discussion) and, equally importantly, acknowledge everyone you discussed the homework with on your writeup. All course materials are available on the course webiste here https://www.haifeng-xu.com/cmsc35401win24. You may refer to any materials covered in our class. However, any attempt to consult outside sources, on the Internet or otherwise, for solutions to any of these homework problems is *not* allowed.

Whenever a question asks you to "show" or "prove" a claim, please provide a formal mathematical proof. These problems have been labeled based on their difficulties. `Short` problems are intended to take you 5-15 minutes each and `medium` problems are intended to take 15-30 minutes each. `Long` problems may take anywhere between 30 minutes to several hours depending on whether inspiration strikes. Note that, the total score is meant to *not* be normalized to 100 (for instance, this HW has 30 in total for regular students and additional 15 points for those who take it as elective).

Finally, please write your solutions in latex — **hand written solutions will not be accepted.** Hope you enjoy the homework!

Problem: Linear Regression from Strategic Data Sources

Suppose you are an insurance company who is trying to design a linear regression function $w \cdot x + b$ (with parameter $w \in \mathbb{R}^d, b \in \mathbb{R}$) to determine the insurance payment for any customer with feature vector $x \in \mathbb{R}^d$ (e.g., riskier customers need to pay more). Naturally, any customer with true feature vector x would have incentives to misreport his feature as some $z \in \mathbb{R}^d$ to hopefully induce a lower payment, i.e., smaller $w \cdot z + b$. However, such manipulation of pretending to be z comes with a cost $c(z; x)$. Therefore, there is a tradeoff between the benefit and cost of manipulating their feature x. Formally, we assume any customer with true feature x will try to pick

$$z^*(x) = \arg\min_z [w \cdot z + b + c(z; x)].$$

We say a *regressor* (w, b) is *incentive compatible* if $z^*(x) = x$ for any $x \in \mathbb{R}^d$. Note that a regressor with $w = 0$ is trivially incentive compatible as its payment does not depend on the feature x at all. We call such a regressor *trivial*. However, there might be non-trivial regressors.

- (`Medium, 11 points`) Suppose $c(z; x) = \sqrt{\sum_{i=1}^d (z_i - x_i)^2}$ is the standard Euclidean distance. Please characterize the set of all incentive compatible regressors.

- (`Medium, 11 points`) Suppose $c(z;x) = \sum_{i=1}^{d}(z_i - x_i)^2$ is the *squared* Euclidean distance. Please characterize the set of all incentive compatible regressors. How many non-trivial regressors are there in this set?

- (`Long, 18 points`) In practice, the regressor is learned from training data by running a linear regression. Suppose you have n *un-manipulated* data $(x_1, y_1), \cdots, (x_n, y_n)$, where $y_j \in \mathbb{R}$ is the payment of customer j. Now you want to use this "pure" data to learn a regressor that will perform well when facing strategic customers with manipulation cost $c(z;x) = \sum_{i=1}^{d}(z_i - x_i)^2$, i.e., the *squared* Euclidean distance.

  Formulate the problem of learning the regressor by minimizing the empiric risks but taking into account strategic behaviors of future customers. Prove that your formulated problem is a convex optimization problem.