Optimization - Quiz 1

Joris LIMONIER

January 18, 2022

1 Question 1

No, it is not possible to pursue directly this goal because we don't know the true distribution $\mathcal{D}$. This is a fundamental difference between Machine Learning and Statistics. We know however that our data was sampled from $\mathcal{D}$, and we know by the law of large numbers that our empirical loss will converge towards the expected value of the loss, as the number of samples increases.

2 Question 2

One way to learn a model is by performing a train-test split in order to verify that our function (that we train on the train set) performs well on a set that is never seen before (*i.e.* the test set). We need to find the right model with not too many parameters (otherwise we over-fit our training set), and not too few parameters (otherwise we under-fit and do not learn enough from data).
Other solutions, especially in case with small data sets, include K-fold cross-validation. One of its variations consists in disregarding a fold of the data set, while looking only at the $K-1$ other folds. Then repeat this step with the other folds.