Introduction to Machine Learing: Homework IV

Due on Dec 7th, 2022 at 11:59pm

Professor Ziping Zhao

Bingnan Li
2020533092

1. [*Clustering and Mixture Models*]
   (a) K-means algorithm.
   (b) Cluster the samples into 2 clusters.
2. [*Clustering and Mixture Models*]
   (a) Advantages of GMM and Why it can be used for clustering.
   (b) Estimate the parameters of the GMM.
3. [*Nonparametric Density Estimation*]
   (a) Expression of $\hat{p}(x)$.
   (b) Expression of $L'(h)$ based on the histogram estimator $\hat{p}(x)$.
   (c) h that minimizes $L'(h)$.
4. [Nonparametric Regression]
   (a) Estimated output $\hat{y}$ and is linear regression a linear smoother?
   (b) In kernel regression, if we use kernel $K(x_i, x) = exp\left\{\frac{-||x_i - x||^2}{2\sigma^2}\right\}$, given an input x, please derive the estimated output $\hat{y}$. Furthermore, is this kernel regression a linear smoother?