UC Berkeley
Department of Electrical Engineering and Computer Sciences

EECS 126: Probability and Random Processes

Homework 11
Spring 2023

1. **Estimating Parameter of Random Graph Given Average Degree**

   Consider an Erdős–Rényi random graph on n vertices, in which each edge appears independently with probability p. Let D be the average degree of a vertex in the graph. Compute the maximum likelihood estimator of p given D. You may approximate $\text{Binomial}(k, p) \approx \text{Poisson}(kp)$.

2. **Community Detection Using MAP**

   It may be helpful to work on this problem in conjunction with the relevant lab. The *stochastic block model* (SBM) defines the random graph $\mathcal{G}(n, p, q)$ consisting of two communities of size $\frac{n}{2}$ each, such that the probability an edge exists between two nodes of the same community is p, and the probability an edge exists between two nodes in different communities is $q < p$. The goal of the problem is to exactly determine the two communities, given only the graph.

   Show that the MAP estimate of the two communities is equivalent to finding the *min-bisection* or *balanced min-cut* of the graph, the split of G into two groups of size $\frac{n}{2}$ that has the minimum edge weight across the partition. Assume that any assignment of the communities is a priori equally likely.

3. **MLE of Uniform Distribution**

   Find the MLE of θ given $X_1, \ldots, X_n \sim_{\text{i.i.d.}} \text{Uniform}([0, \theta])$.

4. **Linear Regression, MLE, and MAP**

   Suppose you draw n i.i.d. data points $(x_1, y_1), \ldots, (x_n, y_n)$, where the true relationship is given by $Y = WX + \varepsilon$ for $\varepsilon \sim \mathcal{N}(0, \sigma^2)$. In other words, Y has a linear dependence on X with additive Gaussian noise.

   a. Show that finding the MLE of W given the data points $\{(x_i, y_i)\}_{i=1}^n$ is equivalent to minimizing mean squared error, or minimizing the cost function

   $$J(w) = \sum_{i=1}^{n} (y_i - wx_i)^2.$$

   b. Now suppose that W has a *Laplace* prior distribution,

   $$f_W(w) = \frac{1}{2\beta} e^{-|w|/\beta}.$$

   Show that finding the MAP estimate of W given the data points $\{(x_i, y_i)\}_{i=1}^n$ is equivalent to minimizing the cost function

   $$J(w) = \sum_{i=1}^{n} (y_i - wx_i)^2 + \lambda |w|.$$

   (You should determine what λ is.) This is interpreted as a one-dimensional ℓ^1-regularized least-squares criterion, also known as LASSO.

5. **Poisson Process MAP**

   Customers arrive to a store according to a Poisson process with rate 1. The store manager learns of a rumor that one of the employees is sending every other customer to the rival store, so that *deterministically*, every odd-numbered customer $1, 3, 5, \ldots$ is sent away.

   Let $X = 1$ be the hypothesis that the rumor is true and $X = 0$ the rumor is false, assuming that both hypotheses are equally likely. Suppose a customer arrives to the store at time 0. After that, the manager observes $T_1, \ldots, T_n$, where T_i is the time of the ith subsequent sale, $i = 1, \ldots, n$. Derive the MAP rule to determine whether the rumor was true or not.

6. **Minimum-Error Property of MAP**

a. Let $X \in \{0,1\}$, and suppose we have the prior $\mathbb{P}(X=0)=\pi_0$ and $\mathbb{P}(X=1)=\pi_1$. Let $\hat{X}_{\text{MAP}}$ be the MAP estimate of X given the random variable Y, and let $\hat{X}$ be any other estimate of X given Y. Show that

$$\mathbb{P}(X \neq \hat{X}_{\text{MAP}}) \leq \mathbb{P}(X \neq \hat{X}).$$

b. Now, also suppose that type I errors (declaring $\hat{X}=1$ when $X=0$) incur a cost of $c_1 \geq 0$ and type II errors (declaring $\hat{X}=0$ when $X=1$) a cost of $c_2 \geq 0$. Derive the decision rule $\hat{X}$ that minimizes the total cost

$$c_1\,\mathbb{P}(\hat{X}=1, X=0) + c_2\,\mathbb{P}(\hat{X}=0, X=1).$$