PRML Homework Assignment II

(Due on Saturday, May 30, 2020.)

Please submit to our TA–Siwen Liu before 8:30 am.

For every 5 minutes beyond the deadline, 5 points will be deducted from your score.

Problem 1: Probability Theory (20 points)

Suppose we have two bags each containing black and white balls. Bag A contains 50 black balls, and twice as many white balls as black balls. Bag B contains 30 white balls, and 3 times as many black balls as white balls.

(a) (5 points) Suppose we give Bag A to a robot who is programmed to randomly select one ball from the bag, and then put the ball back to the bag after recording its color. After the robot performs the task 3 times, what is the probability that it records 2 black balls and 1 white ball?

(b) (5 points) Suppose we replace Bag A with Bag B, and give the bag to the same robot. After the robot performs the task 3 times, what is the probability that it records 2 black balls and 1 white ball?

(c) (10 points) Suppose we randomly select one bag, and give it to this robot. After the robot performs the task 3 times, it records 2 black balls and 1 white ball. What is the probability that these balls are from Bag A?

Problem 2: Bayesian Decision Theory (20 points)

Consider a two-category one-dimensional problem. Suppose the class-conditional densities and the priors of the two classes are known (i.e. $p(x|\omega_1)$, $p(x|\omega_2)$, $P(\omega_1)$ and $P(\omega_2)$).

(a) (5 points) Suppose the decision rule is: Decide ω_1 if $x > \theta$; otherwise decide ω_2. Suppose θ is not the optimal decision boundary, as shown in Figure 1. Explain that the average probability of error for this rule is given by

$$P(error) = P(\omega_1)\int_{-\infty}^{\theta} p(x|\omega_1)dx + P(\omega_2)\int_{\theta}^{\infty} p(x|\omega_2)dx.$$

(b) (5 points) If we change the decision rule to "maximizing a posterior (MAP)", then show that the average probability of error is given by

$$P(error) = 1 - \int P(\omega_{max}|x)p(x)dx,$$

where $P(\omega_{max}|x) \geq P(\omega_i|x)$ for all $i, i = 1, 2$.

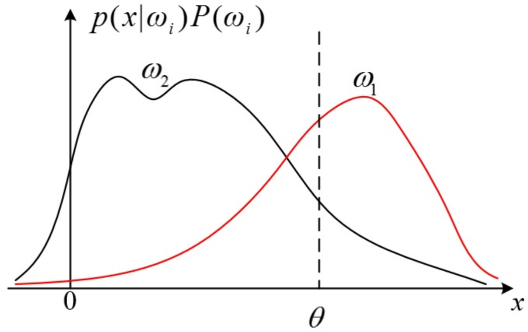

Figure 1: An example of two-category one-dimensional classification.

(c) (2 points) Describe a situation when the two decision rules are equivalent.

(d) (8 points) Suppose we extend this problem to a m-category one-dimensional case. Show that $P(\omega_{max}|x) \geq 1/m$ and $P(error) \leq (m-1)/m$.

Problem 3: Maximum-likelihood Estimation and Bayesian Parameter Estimation (35 points)

Maximum-likelihood estimation (MLE) and Bayesian parameter estimation (BPE) can be applied to estimate the prior probability as well. Let samples be drawn by successive, independent selections of a state of nature ω_i with unknown probability $P(\omega_i)$. Let $z_{ik} = 1$ if the state of nature for the kth sample is ω_i and $z_{ik} = 0$ otherwise.

(a) (2 points) We first reformulate this problem with the notations used in the lectures. For example, the sample set $\mathcal{D} = \{x_1, \ldots, x_n\} = \{z_{i1}, \ldots, z_{in}\}$. Please represent θ and $p(\mathcal{D}|\theta)$ with $\{z_{i1}, \ldots, z_{in}\}$ and $P(\omega_i)$.

(b) (6 points) For MLE approach, show that

$$p(\mathcal{D}|\theta) = \prod_{k=1}^{n} \theta^{x_k}(1-\theta)^{1-x_k}.$$

(Hint: $p(x|\theta)$ in (d)).

(c) (10 points) For MLE approach, show that the best estimate for θ is

$$\theta = \frac{1}{n}\sum_{k=1}^{n} x_k.$$

(d) (5 points) For BPE approach, we assume that

$$p(x|\theta) = \begin{cases} \theta & x = 1 \\ 1-\theta & x = 0 \end{cases} \Leftrightarrow p(x|\theta) = \theta^x (1-\theta)^{(1-x)} (x = 0, 1),$$

and $p(\theta)$ follows uniform distribution (i.e. $p(\theta) \sim U(0,1)$). Please derive $p(\theta|\mathcal{D})$. (Hint: $\int_0^1 \theta^m (1-\theta)^n d\theta = \frac{m!n!}{(m+n+1)!}$)

(e) (5 points) For BPE approach, show that

$$p(x|\mathcal{D}) = \frac{1}{n+2} \cdot \frac{(x + \sum_{k=1}^n x_k)!(n+1-x-\sum_{k=1}^n x_k)!}{(\sum_{k=1}^n x_k)!(n - \sum_{k=1}^n x_k)!}$$

(f) (5 points) What is the effective Bayesian estimate for θ? (Hint: What does $p(x=1)$ stand for?)

(g) (2 points) Compare the estimation results of MLE and BPE approaches and describe your findings.