Machine Learning **Spring 2017**

Solution : Homework 2

Lecturer: Yang Yang *Homework taker: Li Xu*

Due Time: April 9

Problem 1.

Solution:

$$\begin{aligned}
\arg\min_f \mathbb{E}\ell_{\alpha,\beta}(f(x),y) &= \arg\min_f \mathbb{E}_{X,Y}[\alpha\mathbf{1}\{f(X)=1,Y=0\}+\beta\mathbf{1}\{f(X)=0,Y=1\}] \\
&= \arg\min_f \mathbb{E}_X[\mathbb{E}_{Y|X}[\alpha\mathbf{1}\{f(X)=1,Y=0\}+\beta\mathbf{1}\{f(X)=0,Y=1\}]] \\
&= \arg\min_f \mathbb{E}_X[\int_y \alpha\mathbf{1}\{f(X)=1,y=0\}+\beta\mathbf{1}\{f(X)=0,y=1\}dP(y|x)] \\
&= \arg\min_f \int_x [\alpha\mathbf{1}\{f(x)=1\}P(y=0|x)+\beta\mathbf{1}\{f(x)=0\}P(y=1|x)]dP(x)
\end{aligned}$$

Thus, we can minimize the integrand by taking:

$$f(x) = \{ \begin{matrix} 1 & \beta P(y=1|x) \geq \alpha P(y=0|x) \\ 0 & \alpha P(y=0|x) > \beta P(y=1|x) \end{matrix}$$

■

Problem 2.

Solution:

Answer for problem (1):

Likelihood term:

$$\begin{aligned}
P(x_1,\cdots,x_N|\mu) &= \prod_{i=1}^N P(x_i|\mu) \\
&= \prod_{i=1}^N \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}
\end{aligned}$$

$$\log(P(x_1,\cdots,x_N|\mu)) = \sum_{i=1}^N -\frac{(x_i-\mu)^2}{2\sigma^2}\log(\frac{1}{\sqrt{2\pi\sigma^2}})$$

$$\frac{d\log(P(x_1,\cdots,x_N|\mu))}{d\mu} = \sum_{i=1}^N \frac{x_i-\mu}{\sigma^2}$$

If the left part of equation equals to 0:

$$\sum_{i=1}^{N} \frac{x_i - \mu}{\sigma^2} = 0$$

$$\sum_{i=1}^{N} (x_i - \mu) = 0$$

$$\sum_{i=1}^{N} \mu = \sum_{i=1}^{N} x_i$$

MLE estimator for the mean μ:

$$\hat{\mu} = \frac{\sum_{i=1}^{N} x_i}{N}$$

Answer for problem (2):

$$P(\mu|x_1, \cdots, x_N) = \frac{P(x_1, \cdots, x_N|\mu)P(\mu)}{P(x_1, \cdots, x_N)}$$

where

$$P(\mu) = \frac{1}{\sqrt{2\pi\beta^2}} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}$$

The problem is reduce to find the value of μ which maximizes:

$$P(\mu|x_1, \cdots, x_N) = \frac{(\prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}) \frac{1}{\sqrt{2\pi\beta^2}} e^{-\frac{(\mu-\upsilon)^2}{2\beta^2}}}{P(x_1, \cdots, x_N)}$$

$$\log P(\mu|x_1, \cdots, x_N) = (\sum_{i=1}^{N} \frac{(x_i - \mu)^2}{2\sigma^2} \log(\sqrt{2\pi\sigma^2}) + \frac{(\mu - \upsilon)^2}{2\beta^2} \log(\sqrt{2\pi\beta^2})$$

$$\frac{\partial \log P(\mu|x_1, \cdots, x_N)}{\partial \mu} = (\sum_{i=1}^{N} \frac{x_i - \mu}{\sigma^2}) - \frac{\mu - \upsilon}{\beta^2}$$

If the left part of equation equals to 0:

$$(\sum_{i=1}^{N} \frac{x_i - \mu}{\sigma^2}) - \frac{\mu - \upsilon}{\beta^2} = 0$$

$$(\sum_{i=1}^{N} \frac{x_i - \mu}{\sigma^2}) = \frac{\mu - \upsilon}{\beta^2}$$

$$\frac{\mu}{\beta^2} + \frac{N\mu}{\sigma^2} = \frac{\sum_{i=1}^{N} x_i}{\sigma^2} + \frac{\upsilon}{\beta^2}$$

$$\frac{(\sigma^2 + N\beta^2)\mu}{\sigma^2\beta^2} = \frac{\sigma^2\upsilon + \beta^2 \sum_{i=1}^{N} x_i}{\sigma^2\beta^2}$$

MLE estimator for the mean μ:

$$\hat{\mu} = \frac{\sigma^2\upsilon + \beta^2 \sum_{i=1}^{N} x_i}{\sigma^2 + N\beta^2}$$

Answer for problem (3):

$$N \to \infty$$
$$\frac{\sigma^2}{N\beta^2} \to 0$$
$$\frac{\sigma \upsilon}{N\beta^2 + \sigma^2} \to 0$$
$$\hat{\mu}_M AP \to \frac{\sum_{i=1}^{N} x_i}{N} = \hat{\mu}_M LE$$

■

Problem 3.

Solution:

Answer for problem (1):

$$P(Class = X) = \frac{2}{3}, P(Class = Y) = \frac{1}{3}$$
$$P(A1 = 0|Class = X) = \frac{1}{2}, P(A1 = 1|Class = X) = \frac{1}{4}, P(A1 = 2|Class = X) = \frac{1}{4}$$
$$P(A2 = 0|Class = X) = 0, P(A2 = 1|Class = X) = \frac{3}{4}, P(A2 = 2|Class = X) = \frac{1}{4}$$
$$P(A1 = 0|Class = Y) = 0, P(A1 = 1|Class = Y) = \frac{1}{2}, P(A1 = 2|Class = Y) = \frac{1}{2}$$
$$P(A2 = 0|Class = Y) = \frac{1}{2}, P(A2 = 1|Class = Y) = 0, P(A2 = 2|Class = Y) = \frac{1}{2}$$
$$P(Class = X|A1 = 2, A2 = 2) = P(Class = X)P(A1 = 2|Class = X)P(A2 = 2|Class = X) = \frac{1}{24}$$
$$P(Class = Y|A1 = 2, A2 = 2) = P(Class = Y)P(A1 = 2|Class = Y)P(A2 = 2|Class = Y) = \frac{1}{12}$$

So Class is predicted to Y.

Answer for problem (2):

$$P(Class = X) = \frac{5}{8}, P(Class = Y) = \frac{3}{8}$$
$$P(A1 = 0|Class = X) = \frac{3}{7}, P(A1 = 1|Class = X) = \frac{2}{7}, P(A1 = 2|Class = X) = \frac{2}{7}$$
$$P(A2 = 0|Class = X) = \frac{1}{7}, P(A2 = 1|Class = X) = \frac{4}{7}, P(A2 = 2|Class = X) = \frac{2}{7}$$
$$P(A1 = 0|Class = Y) = \frac{1}{5}, P(A1 = 1|Class = Y) = \frac{2}{5}, P(A1 = 2|Class = Y) = \frac{2}{5}$$
$$P(A2 = 0|Class = Y) = \frac{2}{5}, P(A2 = 1|Class = Y) = \frac{1}{5}, P(A2 = 2|Class = Y) = \frac{2}{5}$$
$$P(Class = X|A1 = 2, A2 = 2) = P(Class = X)P(A1 = 2|Class = X)P(A2 = 2|Class = X) = \frac{5}{98}$$
$$P(Class = Y|A1 = 2, A2 = 2) = P(Class = Y)P(A1 = 2|Class = Y)P(A2 = 2|Class = Y) = \frac{3}{50}$$

Result of previous question do not change. ■

Problem 4.

Solution:

The classification accuracy is 83.33%

Using col 7 - col 54 the accuracy raise to 85.00% ■