PS 5 Solutions

Keertana V. Chidambaram.

Problem 1

$$R(h_r; q) = \int_x \sum_{c=1}^{C} \sum_{c'=1}^{C} L_{0/1}(c', c)\, q(r = c' | x)\, p(x, y = c)\, dx.$$

$$h^*(x) = \underset{c}{\operatorname{argmax}}\ p(c|x).$$

$$R(h^*(x)) = \int_x \sum_{c=1}^{C} \sum_{c'=1}^{C} L_{0/1}(c', c)\, I[h^*(x) = c']\, p(x, y = c)\, dx.$$

To prove $R(h_r; q) \geq R(h^*)$,

it is enough to prove $R(h_r; q | x) \geq R(h^* | x)$.

can write $p(x, y = c) = p(y = c | x)\, p(x)$

$$R(h_r; q | x) = \sum_{c=1}^{C} \sum_{c'=1}^{C} L_{0/1}(c', c)\, q(r = c' | x)\, p(y = c | x)$$

$$R(h^*(x) | x) = \sum_{c=1}^{C} \sum_{c'=1}^{C} L_{0/1}(c', c)\, I[h^*(x) = c']\, p(y = c | x)$$

Since we know $L(c', c) = 1$ if $c' \neq c$ and
$L(c', c) = 0$ if $c' = c$,

$$R(h_r, q|x) = \sum_c \sum_{c' \neq c} q(r = c'|x)\, p(y = c|x)$$

$$= \sum_c p(y=c|x)\,[1 - q(c|x)]$$

$$= 1 - \sum_c p(y=c|x)\, q(c|x) \quad -(1)$$

$$R(h^*(x)|x) = \sum_c \sum_{c' \neq c} I[h^*(x) = c']\, p(y=c|x)$$

$$\sum_{c' \neq c} I[h^*(x) = c'] = \begin{cases} 0 & \text{if } h^*(x) = c \\ 1 & \text{if } h^*(x) \neq c. \end{cases}$$

$$R(h^*(x)|x) = \sum_c [1 - I[h^*(x) = c]]\, p(y=c|x).$$

$$= 1 - \sum_c p(y=c|x)\, I[h^*(x) = c] \quad -(2).$$

$$R(h_r; q|x) = 1 - \sum_c q(c|x)\, p(c|x)$$

$$R(h^*(x)|x) = 1 - \sum_c I[h^*(x) = c]\, p(c|x).$$

Since we know that $h^*(x) = \arg\max_c p(c|x)$.

$$\sum_c I[h^*(x) = c]\, p(c|x) = \max_c p(c|x).$$

since $q(c|x)$ is a probability distribution,

$$\max_c p(c|x) \geq \sum_c q(c|x)\, p(c|x).$$

$$\therefore \quad R(h^*(x)|x) \leq R(h_r; q|x)$$

$$\Rightarrow \underline{\underline{R(h^*) \leq R(h_r; q).}} \quad \square$$

Problem 2

Given the parameter assumptions,

$$\delta_+(x) = \log p(x|y=1) + \log p(y=1) \text{ and}$$

$$\delta_-(x) = \log \hat{p}(x|y=0) + \log p(y=0)$$

The decision boundary is given as

$$\delta_+(x) - \delta_-(x) = wx + w_0 \quad \{w_0 = \text{intercept}\}$$

$$\Rightarrow \log p(x|y=1) + \log p(y=1) - \log p(x|y=0) - \log p(y=0) = wx + w_0$$

$$\Rightarrow \log \frac{p(x|y=1)\, p(y=1)}{p(x|y=0)\cdot p(y=0)} = wx + w_0$$

$$\Rightarrow \log \frac{p(y=1|x)p(x)}{p(y=0|x)\, p(x)} = \log \left[\frac{p(y=1|x)}{p(y=0|x)}\right] = wx + w_0$$

$$\Rightarrow \log \left[\frac{p(y=1|x)}{1-p(y=1|x)}\right] = wx + w_0$$

$$\Rightarrow \quad p(y=1|x) = \frac{1}{1+e^{-wx-w_0}}.$$

Problem 3

The two models will not produce the same classifier when estimated from a training set. This is because the logistic regression makes no assumption about the underlying data structure but the Gaussian classifier (as given in the problem) makes multiple assumptions (identical isotropic Σ). In this case, if the actual data doesn't have the same structure, both the models will predict different parameters.