Answer Key for Practice Quiz 2

CS 591Q/791V - Pattern Recognition

Posted on April 26, 2010

1. Now,

$$\hat{p}(x) = \frac{k_n/n}{V_n}.$$

Here, n = 13, h_n = 1 and V_n = 1. k_n can be computed by using the expression

$$k_n = \sum_{i=1}^{n} \phi\left(\frac{x - x_i}{h_n}\right),$$

where,

$$\phi(u) = \begin{cases} 1, & \text{if } |u| \leq 1/2 \\ 0, & \text{otherwise.} \end{cases}$$

Thus,

$\hat{p}(0) = 0$
$\hat{p}(1) = 2/13$
$\hat{p}(3) = 3/13$
$\hat{p}(5) = 1/13$
$\hat{p}(7) = 3/13$
$\hat{p}(9) = 0.$

2. (a) Condition 1: $\lim_{n\to\infty} \quad V_n = 0$
Condition 2: $\lim_{n\to\infty} \quad k_n = \infty$
Condition 3: $\lim_{n\to\infty} \quad k_n/n = 0$

(b) The window width, h_n, is the most critical parameter in the Parzen window approach. This parameter can be selected by cross-validation where a portion of the training set is used to form a validation set. The classifier is trained on the remaining patterns in the training set for different values of h_n. The h_n that results in the smallest error in the validation set is selected as the most "optimal" one.

3. Without any loss of generality, consider an arbitrary test sample in class ω_1. This sample will be correctly classified (as ω_1) if and only if more than $(k-1)/2$ samples in D belong to class ω_1 (since the two unit hyperspheres are separated by 10

units, the distance between pairs of samples within a class is always smaller than the distance between pairs of samples from different classes). Thus, $P_n(e)$ can be estimated by computing the probability that *at most* $(k-1)/2$ samples in D are from the same class. Note that samples in D can be labeled in 2^n different ways. So,

$$\begin{aligned} P_n(e) &= \frac{\binom{n}{0} + \binom{n}{1} + \ldots \binom{n}{(k-1)/2}}{2^n} \\ &= \frac{1}{2^n} \sum_{j=0}^{(k-1)/2} \binom{n}{j}. \end{aligned}$$
