

Quick Review on Non-Parametric Estimation | methods

11/4/2019

(1)

Types:

1. Estimation of density function $P(x|w_i)$ using samples patterns
2. Estimation of a posterior probability $P(w_i/x)$ directly based on sample patterns.

We will consider two approaches for both types of non-parametric methods above.

1. parzen-window-based
2. k-nearest neighbour-based,

Density Estimation:

Estimating the PDF $P(x)$ based on a given set of training samples $D = \{x_1, \dots, x_n\}$

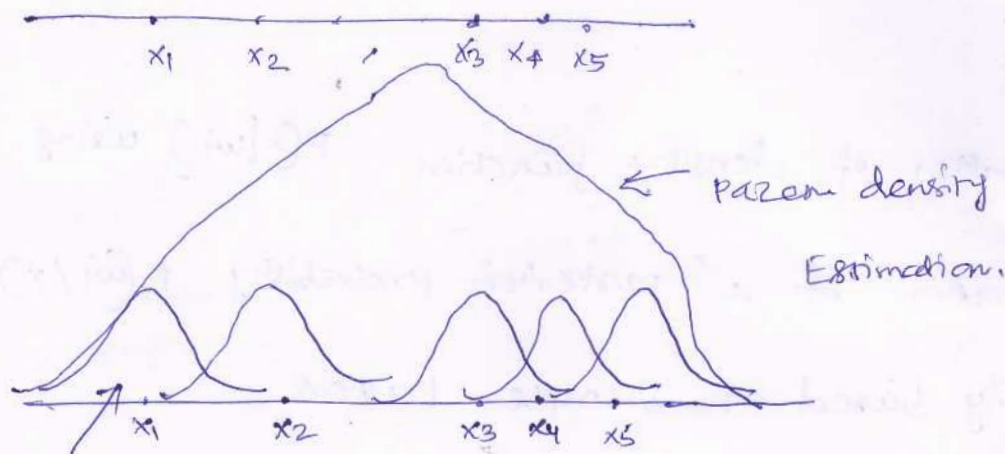
Histograms are the simplest way to density Estimation.

$$p(x) = \frac{n_i}{V_n} =$$

no. of training samples
volume of the cell/bin

$n_i, i=1, 2, \dots, n$

This method is not good for many features



Window

function for x_i



PNN - Implement the classification using Parzen windows

KNN - for density estimation.

An estimate for the joint probability is

$$P_n(x, w_i) = \frac{k_i}{n \cdot v_i}$$

\leftarrow captures k_i samples x_i of which turns out to be labeled w_i
 \uparrow volume around x

Then, we can estimate $P(w_i | x)$ by

$$P(w_i | x) = \frac{P_n(x, w_i)}{\sum P_n(x, w_i)} = \frac{k_i}{k}$$

5/4/2019

(2)

Linear Discriminant function:

* we have also a third class of the methods for designing classifiers. They are based on the discriminant functions.

* Discriminant functions (DFs) are a useful tool for representing classifiers in a simpler way that could be done with the posterior probabilities.

* The decision rule given DFs g_1, \dots, g_c was classify x to the class w_i if $g_i(x) \geq g_j(x)$ for all $i \neq j$.

① with the Bayes classifier, our classifier design procedure was the following:

i) Assume that class conditional probability (ccpdf) have a specific parametric form and prior probabilities are all known

ii) Then using the training data estimate the optimum parameter vector for each ccpdf based on the MLE / Bayesian techniques.

④ The third class of designing the classifier is based on the Non-Parametric form./ methods.

- nearest-neighbour rule
- Parzen window

} In these cases, we estimated the posterior probabilities directly without any assumptions on the form of ccpdfs.

* LDF:

Linear Discriminant functions, is one way of non-parametric form.

- Instead of estimating parameters for cpdfs, we can parameterize the DFs and estimate parameters for them [like: weight vectors and bias].
- The estimation will be formulated as a problem of minimizing a criterion functions
- the criterion function can be training error - the avg. classification error in classifying the training samples.