Datasta Tietoon, Autumn 2007

EXERCISE PROBLEMS 3 [Nov 23rd 2007, Nov 28th 2007]

H3 / 1. (MLE-regression)

We have n pairs of observations $(y(i), x(i)),\ i = 1, .., n$ of some variables x, y which are believed to be linearly related by the model $y = \theta x$. However, the observations contain some errors: $y(i) = \theta x(i) + \epsilon(i)$ where $\epsilon(i)$ is the observation error ("noise") at the i:th point. Assume that the observation error $\epsilon(i)$ is gaussian distributed with mean value 0 and standard deviation σ.

Solve the parameter θ in the model using Maximum Likelihood estimation.

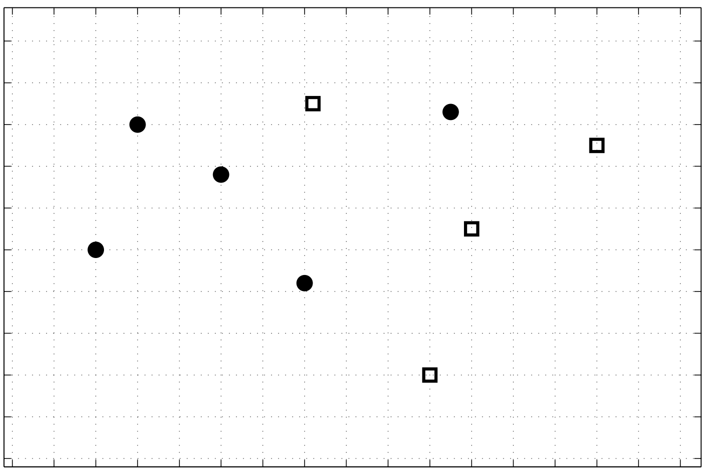
H3 / 2. (Bayes regression)

Let us add some prior knowledge to the previous problem. 1. Suppose that the parameter θ is roughly equal to 1. We model this uncertainty by assuming a gaussian prior density with mean value 1 and standard deviation 0.5.

2. Suppose that the regression line may not pass through the origin after all, but it has the form $y = \alpha + \theta x$. The relation between our observations is then $y(i) = \alpha + \theta x(i) + \epsilon(i)$. We model the uncertainty related to the new parameter α by assuming that is has a gaussian prior density with mean value 0 and standard deviation 0.1. Compute the Bayes estimates for the parameters α, θ.

H3 / 3. (Nearest neighbor classifier)

In the following picture there are 2 classes (circles and squares) in 2 dimensions. Plot the decision boundary (border between classes) of the NN classifier.



H3 / 4. (Bayes classifier)

Assume two classes for a scalar variable x. The class densities $p(x|\omega_1), p(x|\omega_2)$ are gaussian such that both have mean value 0 but different standard deviations σ_1, σ_2. The prior probabilities are $P(\omega_1), P(\omega_2)$. Plot the densities. Where would you place the decision boundaries? Then, derive the decision boundaries of the Bayes classifier.

H3 / 5. (Bayes classifier for binary data)

As shown in the lecture, we derive the Bayes classifier for binary vectors $\mathbf{x} = (x_1, ...x_d)^T$, with each x_i equal to either 0 or 1. There are two classes, and in class ω_1 the probability of having one is p and in class ω_2 it is $q < p$. We may assume that the elements of the vector are independent. The prior probabilities are $P(\omega_1), P(\omega_2)$.

a) Denote by N the number of ones in vector $\mathbf{x}$. Derive the probability of vector $\mathbf{x}$ in class 1 and class 2 in terms of N.

b) Derive the Bayes classifier giving decision regions for N.