Exercise Sheet 5

Due: 12.12.2018, 09:00

Exercise 5.1

Let y be a random variable which takes finitely many values (e.g. class labels) and x be a discrete random variable. Specify and **prove** Bayes formula for $P(y \mid x)$. Hint: How is the conditional probability defined?

Exercise 5.2

Consider the following 2-class classification problem: Class 0 occurs with probability p, class 1 occurs with probability $1-p$. For $i = 0{,}1$ let $x_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$ be a one-dimensional, normally distributed random variable with class label i. The Bayes classifier with zero-one loss assigns the class i to an $x \in \mathbb{R}$ with the maximum posterior probability $P(y = i|x)$.

a) Show that the decision boundary of the Bayes classifier can be described by the zero of a polynomial of maximum degree 2.
b) In which case is the degree 1? Specify the decision rule for this case. What is the influence of the priors here?
c) In which case is the degree 0? Specify the decision rule for this case. What is the influence of the priors here?

Exercise 5.3

The d-dimensional normal distribution with mean $\mu \in \mathbb{R}^d$ and covariance matrix $\Sigma \in \mathbb{R}^{d\times d}$ has the probability density

$$p(x) = \frac{1}{(2\pi)^{d/2} \, |\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right).$$

For $i = 0{,}1$ let $x_i \sim \mathrm{N}(\mu_i, \Sigma_i)$ be d-dimensional, normally distributed random variables with class label i. Consider the Bayes classifier with equal priors for both classes. Show that an $\Sigma_0 = \Sigma_1, \mu_0 \neq \mu_1$ implies a linear decision boundary. Describe the decision boundary for $\Sigma_0 = \Sigma_1, \mu_0 = \mu_1$.

Exercise 5.4

In this exercise, we illustrate decision boundaries of a Bayes classifier with normally distributed likelihoods. For $i = 0{,}1$ let $x_i \sim \mathrm{N}(\mu_i, \Sigma_i)$ 2-dimensional, normally distributed random variables with class label $y = i$ and prior $P(y_i)$. Plot data points (`numpy.random.multivariate_normal`) of the corresponding classes and the class regions in different colors for the Bayes classifier with zero-one loss for the following parameters. Each plot should contain 900 data points in total:

a) Two classes: $\mu_0 = (0,0), \Sigma_0 = \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}, \mu_1 = (4,2), \Sigma_1 = \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}, P(y_0) = P(y_1)$

b) Two classes: $\mu_0 = (0,0), \Sigma_0 = \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}, \mu_1 = (4,2), \Sigma_1 = \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}, P(y_0) = 2 \cdot P(y_1)$

c) Two classes: $\mu_0 = (0,0), \Sigma_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \mu_1 = (2,1), \Sigma_1 = 2 \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, P(y_0) = P(y_1)$

d) Two classes: $\mu_0 = (0,0), \Sigma_0 = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}, \mu_1 = (2,1), \Sigma_1 = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}, P(y_0) = 3 \cdot P(y_1)$

e) Two classes: $\mu_0 = (0,0), \Sigma_0 = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}, \mu_1 = (2,1), \Sigma_1 = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}, P(y_0) = P(y_1)$

f) Two classes: $\mu_0 = (0,0), \Sigma_0 = \begin{bmatrix} 4 & 0.5 \\ 0.5 & 1 \end{bmatrix}, \mu_1 = (0,0), \Sigma_1 = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 4 \end{bmatrix}, P(y_0) = P(y_1)$

g) Three classes:

$$\mu_0 = (0,0), \Sigma_0 = \begin{bmatrix} 2 & 0.5 \\ 0.5 & 1 \end{bmatrix}, \mu_1 = (-2,4), \Sigma_1 = \begin{bmatrix} 1 & 0 \\ 0 & 8 \end{bmatrix}, \mu_2 = (2,-2), \Sigma_2 = \begin{bmatrix} 1 & 0.25 \\ 0.25 & 2 \end{bmatrix}, P(y_0) = P(y_1) = P$$

Discuss the results.
Hint: To plot the class regions, one could create a grid of points and determine for each point in the grid the class membership and color the point accordingly.