Machine Learning and Computational Statistics

Theodoros Georgiopoulos

Homework 2

Exercises for Unit 2: Classification and Statistics concepts

Solution for exercise 3

Consider the following non linear model:

$$y = 3x_1^2 + 4x_2^2 + 5x_3^2 + 7x_1x_2 + x_1x_3 + 4x_2x_3 - 2x_1 - 3x_2 - 5x_3 + \eta. \tag{1}$$

The suitable function ϕ that transforms the problem to a space where the problem of estimating the model becomes linear is

$$\boldsymbol{\phi}(x) = \begin{bmatrix} x_1^2 \\ x_2^2 \\ x_3^2 \\ x_1x_2 \\ x_1x_3 \\ x_2x_3 \\ x_1 \\ x_2 \\ x_3 \end{bmatrix}. \tag{2}$$

With this transformation the 2-dimensional space becomes 9-dimensional space and the problem is now linear.

Solution for exercise 4

For $\boldsymbol{x} = [x_1, x_2, x_3]^T$, consider the following non linear classification task:

$$x_1^2 + 3x_2^2 + 6x_3^2 + x_1x_2 + x_2x_3 > (<) 3 \rightarrow \boldsymbol{x} \in \omega_1(\omega_2). \tag{3}$$

The suitable function ϕ that transforms the problem to a space where the problem of estimating the border of the two classes becomes linear is

$$\boldsymbol{\phi}(x) = \begin{bmatrix} x_1^2 \\ x_2^2 \\ x_3^2 \\ x_1x_2 \\ x_2x_3 \end{bmatrix}. \tag{4}$$

The dimension of the original space is 2 and the dimension of the transformed space in which the problem is linear, is 5.

Solution for exercise 6

Let x, y be two random variables with sample spaces $X = \{x_1, x_2, \ldots, x_{n_x}\}$ and $Y = \{y_1, y_2, \ldots, y_{n_y}\}$ respectively. Moreover, we denote n the total number of experiments and as $\{n_1^X, n_2^X, \ldots, n_{n_x}^X\}$ the probabilities of occurence of $x_1, x_2, \ldots, x_{n_x}$. Similarly, we denote as $\{n_1^Y, n_2^Y, \ldots, n_{n_y}^Y\}$ the probabilities of occurence of $y_1, y_2, \ldots, y_{n_y}$. Finally, we denote as n_{ij} the number of the times that the value x_i and y_j occured simultaneously. The joint probability can be approximated as $P(x, y) \approx \frac{n_{ij}}{n}$. The total number n_i^X that value x_i has occured is $n_i^X = \sum_{i=1}^{n_y} n_{ij}$. Dividing both sides by n gives

$$P(x) = \sum_{y \in Y} P(x, y). \tag{5}$$

The conditional probability $P(A|B)$ is defined as

$$P(A|B) = \frac{n_{AB}}{n} \frac{n}{n_B} = \frac{n_{AB}}{n_B}. \tag{6}$$

So, for the product rule, we can write

$$P(A|B)p(B) = \frac{n_{AB}}{n_B} \frac{n_B}{n} = \frac{n_{AB}}{n} = P(A, B). \tag{7}$$

Using the product rule, we can verify the Bayes rule

$$P(A|B)p(B) = \frac{n_{AB}}{n_B} \frac{n_B}{n} = \frac{n_{AB}}{n_A} \frac{n_A}{n} = P(B|A)p(A). \tag{8}$$

So,

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)}. \tag{9}$$