Statistical learning

2. Classification

Kevin Zagalo

<kevin.zagalo@inria.fr>

January 29, 2020

Overview

1. Logistic Regression
2. Linear Discriminant Analysis
3. Classification And Regression Trees

Logistic regression

- Binary classification :

$$Y \in \{0, 1\}$$

- Logistic function :

$$p(x; \theta) = \mathbf{P}[Y = 1 | X = x] = \left[1 + e^{-\langle \theta, x \rangle}\right]^{-1}$$

- Cost function : likelihood of the associated probability

$$\ell(\theta; (x_i, y_i)_i) = \prod_i p(x_i; \theta)^{y_i} (1 - p(x_i; \theta))^{1 - y_i}$$

Maximum likelihood

We are looking for the **maximum likelihood estimator** of the parameter θ, which is the equivalent of maximizing the log-likelihood

$$L(\theta) = \log \ell(\theta; (x_i, y_i)_i)$$

Finally we are interested in the optimization problem

$$\hat{\theta} = \operatorname*{argmax}_{\theta} L(\theta) \Leftrightarrow \nabla L(\hat{\theta}) = 0$$

Then everything behaves as a regression.

Linear discriminant analysis (LDA)

youpi

Classification and regression trees

youpiyo