

COMPUTATIONAL STATISTICS

LAB IV - LOGISTIC REGRESSION

Luca Bortolussi

Department of Mathematics and Geosciences
University of Trieste

Office 238, third floor, H2bis
`luca@dmi.units.it`

Trieste, Winter Semester 2015/2016

OUTLINE

1 PERCEPTRON

2 LOGISTIC REGRESSION

TASK 1

- Download data from moodle2.
- Implement the perceptron algorithm for a generic set of basis function in any dimension, which has to be defined in an external function, taking as input the vector of input points, and returning the vector of basis functions as output.
- Apply the algorithm to the linear and non-linear datasets without overlap. They are all in dimension 2.

OUTLINE

1 PERCEPTRON

2 LOGISTIC REGRESSION

TASK 2

- Implement the (regularised) logistic regression (same mechanisms to deal with basis functions), using Gradient Descent and/or Newton-Rapson as an optimisation routine (to be implemented, too).
- Apply the non-regularised logistic regression the nonlinear dataset with overlap (use e.g. polynomials of degree 2)
- Apply the regularised logistic regression to the nonlinear dataset with overlap, set the regularisation coefficient by cross-validation.

TASK 3

- Implement the (regularised) logistic regression in the dual formulation, using the Gaussian kernel

$$k(\mathbf{x}, \mathbf{y}) = \exp \left[\frac{\|\mathbf{x} - \mathbf{y}\|^2}{\lambda^2} \right]$$

- Apply the method on max 30 input points. Set λ by using a validation dataset.