Department of Mathematics and Computer Science
University of Southern Denmark, Odense

February 1, 2013
Marco Chiarandini

DM825 - Introduction to Machine Learning

Sheet 3, Spring 2013

Exercise 1
Redo exercise 1 from Sheet 1 using logistic regression (transform the response label -1 to 0). Alternatively use logistic regression on these data [classification.data]. Although, as we will see, logistic regression can be impleemnted in R via `glm`, you are asked here to implement the method by yourself. For the optimization you can reuse the gradient descent method developed in previous exercises or you can use `optim`.

Exercise 2
In exercise 3 of Sheet 2 use 1/2 of the data for training the models, 1/4 of the data to *select* the model (k-nearest neighbor or linear regression) and 1/4 to *assess* the performance of the best model selected.

Exercise 3 Bayesian prediction
In class we saw an example with binary variables. Often however we encounter discrete variables that can take on one of K possible mutually exclusive states. A way to handle this situation is to express such variables by a K-dimensional vector $\vec{x}$ in which one of the x_k elements equals to 1 and all remaining elements equal 0. Consider a sample described by m multinomial random variables $(X^1, X^2, \ldots, X^m)$, where $X^i \sim \mathrm{Mult}(\theta)$ for each m, and where the X^i are assumed conditionally independent given θ. Let $\theta \sim \mathrm{Dir}(\boldsymbol{\alpha})$. Now consider a random variable $X_{new} \sim \mathrm{Mult}(\theta)$ that is assumed conditionally independent of $(X^1, X^2, \ldots, X^m)$ given θ. Compute the predictive distribution:

$$p(x_{new}|x_1, x_2, \ldots, x_N, \alpha)$$

by integrating over θ.