

Problem Set 2:

Sean Smith

February 29, 2016

Problem 1

- (a) Minimize the function

$$E_D(w) = \frac{1}{2} \sum_{n=1}^N r_n \{t_n - w^T \phi(x_n)\}^2$$

- (b)
(c)
(d)
(e)

Problem 2

- a) I used python and used the math function Pearson Correlation Coefficient:

$$r = r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}$$

- b) i I created a script `process` that takes the two files `crx.data.testing` and `crx.data.training` and imputes the missing values denoted by a question mark in column 1. Since the mean value of a's and b's is b I just filled in each question mark with a b.
To run:
`./processed crx.data.training crx.data.testing`