

CMPS 242, HW1

Steven Reeves, Batu Aytemiz, Xi Hu

September 29, 2017

1 Introduction

Traditionally, a least-squares regression technique is used to fit a predictor curve to data, in order to match response variables. This however, may not reveal the best results. For example, if one has 9 data points, one can construct a 9th degree polynomial and perfectly fit all data points. However, there is no guarantee that additional sample points from the underlying distribution will be best fit by this polynomial. Additionally, when constructing a "perfect" fit polynomial, the coefficients can be large and erratic. To remedy this issue, we change the error function from the L_2 error to

$$E(\mathbf{w}) = \|\mathbf{X}^T \mathbf{w} - \mathbf{t}\|^2 + \lambda \|\mathbf{w}\|^2.$$

This penalised error function can now be minimized using classic optimization techniques. That is the optimal weights are ones that satisfy

$$\mathbf{w}^* = (\mathbf{X}\mathbf{X}^T + \lambda \mathbf{I})^{-1} \mathbf{X}\mathbf{t}$$

where \mathbf{w}^* are the optimal weights, \mathbf{X} is the coefficient matrix generated by the predictor data, \mathbf{I} is the identity matrix, \mathbf{t} is the response data, and λ is the penalization parameter.

In this report we present an optimal polynomial model for predicting the test data along with the best value of λ when evaluating the model against the test data. To do this we discretize λ using a ln scale. The exponential form of the penalization parameter has been shown to be an effective format, see CMPS 242 Lecture Notes 1.

2 Cross validation Method