

Reading Notes of Pattern Classification and Machine Learning

Tianyi Cui

August 12, 2012

1 Introduction

Different kinds of tasks of machine learning:

- supervised learning: known input and target vectors
- classification: output is one of a finite number of discrete categories
 - regression: output is one or more continuous variables
- unsupervised learning: no corresponding target values
 - clustering: discover groups of similar examples within the data
 - density estimation: determine the distribution of data within the input space
 - dimension reduction
- reinforcement learning: finding suitable actions to take in a given situation in order to maximize a reward

1.1 Example: Polynomial Curve Fitting

In regression problems, we can use a polynomial function

$$y(x, \mathbf{w}) = w_0 + w_1x + w_2x^2 + \dots + w_Mx^M = \sum_{j=0}^M w_jx^j \quad (1.1)$$

to fit the underlying function.

We need to minimize the error function

$$E(\mathbf{w}) = \frac{1}{2} \sum_{n=1}^N \{y(x_n, \mathbf{w}) - t_n\}^2 \quad (1.2)$$

in which unique solution \mathbf{w}^* can be found in closed form.

The root-mean-square (RMS) is error defined by

$$E_{RMS} = \sqrt{2E(\mathbf{w}^*)/N} \quad (1.3)$$

When M is large, over-fitting occurs, i.e. E_{RMS} against test data becomes large. One technique to control over-fitting is regularization, by adding a penalty term to the error function (1.2) in order to discourage the coefficients from reaching large values:

$$\tilde{E}(\mathbf{w}) = \frac{1}{2} \sum_{n=1}^N \{y(x_n, \mathbf{w}) - t_n\}^2 \quad (1.4)$$