PRML study note

Booil Jung

April 9th 2018

Chapter 1

Introduction

Supervised learning problems:
Applications in which the training data comprises examples of the input vectors along with their corresponding target vectors.

Unsupervised learning problems:
The training data consists of a set of input vectors $\mathbf{x}$ without any corresponding target values

may be to discover groups of similar examples within the data, where it is called clustering, or to determine the distribution of data within the input space, known as density estimation, or to project the data from a high-dimensional space down to two or three dimensions for the purpose of visualization.

1.1 Example: Polynomial Curve Fitting

Input vector $\mathbf{x}$:

$$\mathbf{x} = x_1, \cdots, x_D$$

$$x: \text{ input variable}$$

$$D: \text{ dimension}$$

Training set:

$$\{\mathbf{x}_1, \cdots \mathbf{x}_N\}$$

$$N: \text{ observations}$$

Machine learing function:

$$y(\mathbf{x})$$

Target vector $\mathbf{t}$:

$$\mathbf{t} = t_1, \cdots, t_D$$

$$t: \text{ target variable}$$

$$D: \text{ dimension}$$

1.1.1 Example: Polynomial Curve Fitting

A training set $\mathbf{x}$:

$$\mathbf{x} = (x_1 \quad \cdots \quad x_N)^T$$

A corresponding observations of target set t:

$$\mathbf{t} = (t_1 \quad \cdots \quad t_N)^T$$

Predictions of target value:

$$\hat{t}$$

Input value of $\hat{t}$

$$\hat{x}$$

A polynomial function:

$$y(x, \mathbf{w}) = w_0 + w_1 x + w_2 x^2 + \cdots + w_M x^M = \sum_{j=0}^{M} w_j x^j \tag{1.1}$$

$$\begin{aligned} x &: \text{ input value} \\ M &: \text{ order of the polynomial} \\ \mathbf{w} &: \text{ polynomial coefficients vector } (w_0 \quad \cdots \quad w_M)^T \end{aligned}$$

Error function:

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

Root-mean-square error:

$$E_{\text{RMS}} = \sqrt{2 \frac{E(\mathbf{w}^\star)}{N}} \tag{1.3}$$

$$\begin{aligned} \mathbf{w}^\star &: \text{ minimized error solution} \\ y(x, \mathbf{w}^\star) &: \text{ minimized polynomial error function} \end{aligned}$$

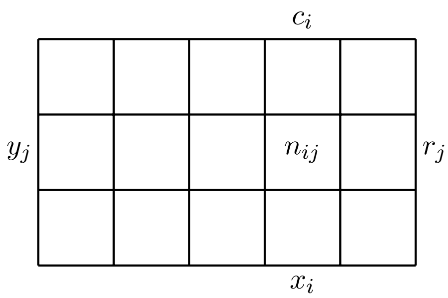
modified error function:

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

$$\|\mathbf{w}\|^2 = w_0^2 + w_1^2 + \cdots + w_N^2$$

$$\lambda : \text{ coefficient}$$

1.1.2 Probability Theory



joint probability of $X = x_i$ and $Y = y_j$:

$$p(X = x_i, Y = y_j) = \frac{n_{ij}}{N} \tag{1.5}$$

$X = x_i :$ probability X take value x_i

$Y = y_j :$ probability Y take value y_j

()

! !

! !