Solution Manual For Pattern Recognition and Machine Learning

Edited By

ZHENGQI GAO

Information Science and Technology School
Fudan University

Nov.2017

0.1 Introduction

Problem 1.1 Solution

We let the derivative of *error function* E with respect to vector $\mathbf{w}$ equals to $\mathbf{0}$, (i.e. $\frac{\partial E}{\partial \mathbf{w}} = 0$), and this will be the solution of $\mathbf{w} = \{w_i\}$ which minimizes *error function* E. To solve this problem, we will calculate the derivative of E with respect to every w_i , and let them equal to 0 instead. Based on (1.1) and (1.2) we can obtain :

=>

$$\frac{\partial E}{\partial w_i} = \sum_{n=1}^{N} \{y(x_n, \mathbf{w}) - t_n\} x_n^i = 0$$

=>

$$\sum_{n=1}^{N} y(x_n, \mathbf{w}) x_n^i = \sum_{n=1}^{N} x_n^i t_n$$

=>

$$\sum_{n=1}^{N} (\sum_{j=0}^{M} w_j x_n^j) x_n^i = \sum_{n=1}^{N} x_n^i t_n$$

=>

$$\sum_{n=1}^{N} \sum_{j=0}^{M} w_j x_n^{(j+i)} = \sum_{n=1}^{N} x_n^i t_n$$

=>

$$\sum_{j=0}^{M} \sum_{n=1}^{N} x_n^{(j+i)} w_j = \sum_{n=1}^{N} x_n^i t_n$$

If we denote $A_{ij} = \sum_{n=1}^{N} x_n^{i+j}$ and $T_i = \sum_{n=1}^{N} {x_n}^i t_n$, the equation above can be written exactly as (1.222), Therefore the problem is solved.

Problem 1.2 Solution

This problem is similar to Prob.1.1, and the only difference is the last term on the right side of (1.4), the penalty term. So we will do the same thing as in Prob.1.1 :

=>

$$\frac{\partial E}{\partial w_i} = \sum_{n=1}^{N} \{y(x_n, \mathbf{w}) - t_n\} x_n^i + \lambda w_i = 0$$

=>

$$\sum_{j=0}^{M} \sum_{n=1}^{N} x_n^{(j+i)} w_j + \lambda w_i = \sum_{n=1}^{N} x_n^i t_n$$

=>

$$\sum_{j=0}^{M} \{\sum_{n=1}^{N} x_n^{(j+i)} + \delta_{ji} \lambda\} w_j = \sum_{n=1}^{N} x_n^i t_n$$

where

$$\delta_{ji}\begin{cases}0 & \mathrm{j} \neq \mathrm{i}\\ 1 & \mathrm{j} = \mathrm{i}\end{cases}$$

Problem 1.3 Solution

This problem