Solutions to
Pattern Recognition and Machine Learning

Yunwei Ren

Contents

1 Introduction

1.1 Example: Polynomial Curve Fitting

1.

Proof. Suppose that $E(w)$ attains it minimum at $w = w^*$, then the partials of E are 0 at w^*. Hence, it is a necessary condition that $D_i E(w) = 0$ for each $i = 0, \dots, m$, namely,

$$0 = D_i E(w) = \sum_{n=1}^{N} \left(\sum_{j=0}^{M} w_j x_n^j - t_n \right) x_n^i \quad \Leftrightarrow \quad \sum_{j=0}^{M} \left(\sum_{n=1}^{N} x_n^{i+j} \right) w_j = \sum_{n=1}^{N} t_n x_n^i,$$

which is just (1.122) and (1.123). Meanwhile, since E is a quadratic function in w and is bounded below, E has a unique minimum. Hence, the above condition is also sufficient. $\square$

2.

Solution. $(A - \lambda I_{m+1})w = T$, where I_{m+1} is the identity matrix. $\square$