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. □

2.

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

1.2 Probability Theory

7.

Proof. Make the change of variables $x = r\cos\theta$, $y = r\sin\theta$. Then

$$I^2 = \int_0^\infty \int_0^{2\pi} e^{-r^2/2\sigma^2} r \mathrm{d}r \mathrm{d}\theta = \pi \int_0^\infty e^{-r^2/2\sigma^2} \mathrm{d}r^2 = 2\pi\sigma^2.$$

Hence, $I = (2\pi\sigma^2)^{1/2}$ and therefore $\int \mathcal{N}(x|0, \sigma^2) = 1$. Since $\int_{-\infty}^{\infty}$ is translation invariant, this implies that the general Gaussian distribution is normalized. □

8.

Proof. Make the change of variable $y = x - u$. Then

$$\begin{aligned}
\mathbb{E}[x] &= \int_{-\infty}^{\infty} \mathcal{N}(x|\mu, \sigma^2) x \mathrm{d}x \\
&= \frac{1}{(2\pi\sigma^2)^{1/2}} \int_{-\infty}^{\infty} \exp\left\{ -\frac{1}{2\sigma^2} y^2 \right\} (y + \mu) \mathrm{d}y \\
&= \frac{1}{(2\pi\sigma^2)^{1/2}} \left\{ \int_{-\infty}^{\infty} \exp\left\{ -\frac{1}{2\sigma^2} y^2 \right\} y \mathrm{d}y + \mu \int_{-\infty}^{\infty} \exp\left\{ -\frac{1}{2\sigma^2} y^2 \right\} \mathrm{d}y \right\}.
\end{aligned}$$

Since $\exp\{-y^2/2\sigma^2\}y$ is odd, the first term in the braces equals 0. Then by the normalization condition, $\mathbb{E}[x] = \mu$.

Differentiating the both sides of (1.48) with respect to σ^2 yields

$$-\frac{1}{2\sigma^2} + \frac{1}{2\sigma^2}(\mathbb{E}[x^2] - 2\mu E[x] + \mu^2) = 0.$$

Namely, $\mathbb{E}[x^2] = \mu^2 + \sigma^2$. Thus, $\mathrm{var}[x] = \mathbb{E}[x^2] - \mathbb{E}[x]^2 = \sigma^2$. □