演習問題 1.26

式 (1.151) の 2 乗を展開し、式 (1.90) に類似の結果を導き、目標変数ベクトル $\mathbf{t}$ の場合に、期待二乗損失を最小にする関数 $\mathbf{y}(\mathbf{x})$ が、やはり $\mathbf{t}$ の条件付き期待値で与えられることを示せ。

$$\mathbf{E}[\,L(\,\mathbf{t},\ \mathbf{y}(\mathbf{x})\,)\,] = \int\int \|\,\mathbf{y}(\mathbf{x}) - \mathbf{t}\,\|^2\, p(\mathbf{x},\ \mathbf{t})\, d\mathbf{x}\, d\mathbf{t}$$

⋯ (1.151)

$$\mathrm{E}[L] = \int\int \{y(\mathbf{x}) - t\}^2\, p(\mathbf{x}, t)\, d\mathbf{x}\, dt$$

$$= \int \{y(\mathbf{x}) - \mathrm{E}[t\,|\,\mathbf{x}]\}^2\, p(\mathbf{x})\, d\mathbf{x} + \int \mathrm{var}[t\,|\,\mathbf{x}]\, p(\mathbf{x})\, d\mathbf{x}$$

⋯ (1.90)

[解]

式 (1.151) の 2 乗部分を展開すると、

$$\|\,\mathbf{y}(\mathbf{x}) - \mathbf{t}\,\|^2 = \|\,\mathbf{y}(\mathbf{x}) - \mathrm{E}[\mathbf{t}\,|\,\mathbf{x}] + \mathrm{E}[\mathbf{t}\,|\,\mathbf{x}] - \mathbf{t}\,\|^2$$

$$= \|\,\mathbf{y}(\mathbf{x}) - \mathrm{E}[\mathbf{t}\,|\,\mathbf{x}]\,\|^2 + 2\,\|\,\mathbf{y}(\mathbf{x}) - \mathrm{E}[\mathbf{t}\,|\,\mathbf{x}]\,\|\,\|\,\mathrm{E}[\mathbf{t}\,|\,\mathbf{x}] - \mathbf{t}\,\| + \|\,\mathrm{E}[\mathbf{t}\,|\,\mathbf{x}] - \mathbf{t}\,\|^2$$

となる。ここで、この 2 乗部分を式 (1.151) に代入し直し、$\mathbf{t}$ について積分を行うと、

$$\mathbf{E}[\,L(\,\mathbf{t},\ \mathbf{y}(\mathbf{x})\,)\,] = \int\int \|\,\mathbf{y}(\mathbf{x}) - \mathbf{t}\,\|^2\, p(\mathbf{x},\ \mathbf{t})\, d\mathbf{x}\, d\mathbf{t}$$

$$= \int\int \|\,\mathbf{y}(\mathbf{x}) - \mathrm{E}[\mathbf{t}\,|\,\mathbf{x}]\,\|^2\, p(\mathbf{x},\ \mathbf{t})\, d\mathbf{x}\, d\mathbf{t}$$

$$+ 2\int\int \|\,\mathbf{y}(\mathbf{x}) - \mathrm{E}[\mathbf{t}\,|\,\mathbf{x}]\,\|\,\|\,\mathrm{E}[\mathbf{t}\,|\,\mathbf{x}] - \mathbf{t}\,\|\, p(\mathbf{x},\ \mathbf{t})\, d\mathbf{x}\, d\mathbf{t}$$

$$+ \int\int \|\,\mathrm{E}[\mathbf{t}\,|\,\mathbf{x}] - \mathbf{t}\,\|^2\, p(\mathbf{x},\ \mathbf{t})\, d\mathbf{x}\, d\mathbf{t}$$

となる。上記の式の第二項は消えるので、

$$= \int \|\,y(\mathbf{x}) - \mathrm{E}[\mathbf{t}\,|\,\mathbf{x}]\,\|^2\, p(\mathbf{x})\, d\mathbf{x} + \int \mathrm{var}[\mathbf{t}\,|\,\mathbf{x}]\, p(\mathbf{x})\, d\mathbf{x}$$

となり、目標変数ベクトル $\mathbf{t}$ の場合に、期待二乗損失を最小にする関数 $\mathbf{y}(\mathbf{x})$ が、やはり $\mathbf{t}$ の条件付き期待値で与えられることを示せた。