演習問題 1.25

単一の目標変数 t の式 (1.87) の二乗損失関数のベクトル値 t で表される多変数の場合への、以下の一般化について考える。

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

… (1.151)

変分法によって、この期待損失を最小化する関数 $\mathbf{y}(\mathbf{x})$ が $\mathbf{y}(\mathbf{x}) = \mathbf{E}_{\mathrm{t}}[\ \mathbf{t}\,|\,\mathbf{x}\ ]$ で与えられることを示せ。単一の目標変数 $\mathbf{t}$ の場合は、この結果が式 (1.89) に帰着されることを確かめよ。

[期待損失]

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

… (1.87)

ここで、$\{\ y\,(\,\mathbf{x}\,) - t\ \}^2$ は損失関数である。⇒ $\mathrm{E}[\,L\,]$ を最小化する $y\,(\,\mathbf{x}\,)$ を選ぶ。

[回帰関数]

$$y\,(\,\mathbf{x}\,) = \frac{\int t\, p\,(\,\mathbf{x}, t\,)\, dt}{p\,(\,\mathbf{x}\,)} = \int t\, p\,(\,t\,|\,\mathbf{x}\,)\, dt = \mathrm{E}_t\,[\,t\,|\,\mathbf{x}\,]$$

… (1.89)

[解]

我々の目標は、$\mathbf{E}[\ L(\ \mathbf{t},\ \mathbf{y}(\mathbf{x})\ )\ ]$ を最小化する $\mathbf{y}(\mathbf{x})$ を選ぶことである。このために、式 (1.151) を関数 $\mathbf{y}(\mathbf{x})$ で微分すると、**⇒ 変分法**

$$\frac{\delta\, \mathbf{E}[\ L(\ \mathbf{t},\ \mathbf{y}(\mathbf{x})\ )\ ]}{\delta\, \mathbf{y}(\mathbf{x})} = \int 2\,(\ \mathbf{y}(\mathbf{x}) - \mathbf{t}\ )\, p(\mathbf{x},\ \mathbf{t})\, d\mathbf{t} = 0$$

より、これを $\mathbf{y}(\mathbf{x})$ について解くと、

$$\int 2\,(\ \mathbf{y}(\mathbf{x}) - \mathbf{t}\ )\, p(\mathbf{x},\ \mathbf{t})\, d\mathbf{t} = 0$$

$$\int (\ \mathbf{y}(\mathbf{x}) - \mathbf{t}\ )\, p(\mathbf{x},\ \mathbf{t})\, d\mathbf{t} = 0$$

$$\int \mathbf{y}(\mathbf{x})\, p(\mathbf{x},\ \mathbf{t})\, d\mathbf{t} - \int \mathbf{t}\, p(\mathbf{x},\ \mathbf{t})\, d\mathbf{t} = 0$$

$$\mathbf{y}(\mathbf{x}) \int p(\mathbf{x},\ \mathbf{t})\, d\mathbf{t} - \int \mathbf{t}\, p(\mathbf{x},\ \mathbf{t})\, d\mathbf{t} = 0$$

$$\mathbf{y}(\mathbf{x}) \int p(\mathbf{x},\ \mathbf{t})\, d\mathbf{t} = \int \mathbf{t}\, p(\mathbf{x},\ \mathbf{t})\, d\mathbf{t}$$

$$\mathbf{y}(\mathbf{x}) = \frac{\int \mathbf{t}\, p(\mathbf{x},\ \mathbf{t})\, d\mathbf{t}}{\int p(\mathbf{x},\ \mathbf{t})\, d\mathbf{t}}$$

… ※

と整理できる。ここで、確率の加法定理より、上記の式の分母は、

$$\int p(\mathbf{x},\ \mathbf{t})\, d\mathbf{t} = p(\mathbf{x})$$

であり、確率の乗法定理より、上記の式の分子は、

$$\int \mathbf{t}\, p(\mathbf{x},\ \mathbf{t})\, d\mathbf{t} = \int \mathbf{t}\, p(\mathbf{t} \mid \mathbf{x})\, p(\mathbf{x})\, d\mathbf{t}$$

となるから、最終的に式 ※ は、

$$= \frac{\int \mathbf{t}\, p(\mathbf{t} \mid \mathbf{x})\, p(\mathbf{x})\, d\mathbf{t}}{p(\mathbf{x})} = \int \mathbf{t}\, p(\mathbf{t} \mid \mathbf{x})\, d\mathbf{t} = \mathbf{E}_{\mathrm{t}}[\mathbf{t} \mid \mathbf{x}]$$

となる。よって、期待損失を最小化する関数 $\mathbf{y}(\mathbf{x})$ が $\mathbf{y}(\mathbf{x}) = \mathbf{E}_{\mathrm{t}}[\mathbf{t} \mid \mathbf{x}]$ で与えられることを示せた。また、単一の目標変数 $\mathbf{t}$ の場合は、この結果が式 (1.89) に帰着されることがわかる。

[変分法]

関数についての微分のことである。通常の微分は、変数に対して微分を行い、変数の値を求めるが、変分法では、関数に関して微分を行い、関数を求める。

[期待値]

ある関数 x の確率分布 $p(x)$ の下での平均値は、以下のように表され、

$$\mathrm{E}[x] = \int x\, p(x)\, dx$$

… (1.34)

のようになる。これを x の期待値と呼ぶ。