演習 1.25

茂木 哲矢

2008 年 07 月 11 日

1. 25（**基本**）　$\boxed{\text{www}}$単一の目標変数 t の (1.87) の二乗損失関数のベクトル値 $\boldsymbol{t}$ 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．

$$\mathbb{E}[L] = \iint ||y(\boldsymbol{x}) - \boldsymbol{t}||^2 p(\boldsymbol{x}, \boldsymbol{t}) d\boldsymbol{x} d\boldsymbol{t} \tag{1.151}$$

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

$$\int tp(x|\boldsymbol{x})dt = \mathbb{E}_t[t|\boldsymbol{x}] \tag{1.89}$$

■Answer

$$F[y] = \int_{x_1}^{x_2} G(y(x), y'(x), x) dx. \tag{D.5}$$

の変分 $\frac{\delta F}{\delta y(x)}$ はオイラー-ラグランジュ方程式

$$\frac{\partial G}{\partial y} - \frac{d}{dx}\left(\frac{\partial G}{\partial y'}\right) = 0. \tag{D.8}$$

を解くことで得られる．

式 (1.151) を整理して，

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

式 (D.5) と比較すると，$G(y(x), y'(x), x)$ と $\int ||y(\boldsymbol{x}) - \boldsymbol{t}||^2 p(\boldsymbol{x}, \boldsymbol{t}) d\boldsymbol{t}$ が対応する．

よって

$$G(y(\boldsymbol{x}), y'(\boldsymbol{x}), \boldsymbol{x}) \equiv \int ||y(\boldsymbol{x}) - \boldsymbol{t}||^2 p(\boldsymbol{x}, \boldsymbol{t}) d\boldsymbol{t}$$

とすると

$$\frac{\partial G}{\partial y(\boldsymbol{x})} = 2 \int \{y(\boldsymbol{x}) - \boldsymbol{t}\} p(\boldsymbol{x}, \boldsymbol{t}) d\boldsymbol{t} \tag{1}$$

$$\frac{\partial G}{\partial y'(\boldsymbol{x})} = 0 \tag{2}$$

変分は

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

これを $y(\boldsymbol{x})$ について解く.

$$y(\boldsymbol{x}) = \frac{\int \boldsymbol{t} p(\boldsymbol{x}, \boldsymbol{t}) d\boldsymbol{t}}{\int p(\boldsymbol{x}, \boldsymbol{t}) d\boldsymbol{t}} \tag{3}$$

確率の加法定理・乗法定理より, 式 (3) は

$$y(\boldsymbol{x}) = \frac{\int \boldsymbol{t} p(\boldsymbol{x}, \boldsymbol{t}) d\boldsymbol{t}}{p(\boldsymbol{x})} = \int \boldsymbol{t} p(\boldsymbol{t}|\boldsymbol{x}) d\boldsymbol{t} = \mathbb{E}_t[\boldsymbol{t}|\boldsymbol{x}] \tag{4}$$

これは $\boldsymbol{x}$ が与えられた下での $\boldsymbol{t}$ の条件付き平均になっている.

また, t が単一変数の場合は,

$$y(\boldsymbol{x}) = \int t p(t|\boldsymbol{x}) dt = \mathbb{E}_t[t|\boldsymbol{x}]$$

とな式 (1.89) と等しくなる.