

第 14 週目 追加問題

解答は答えだけでなく、途中経過も記述すること（プログラム作成の場合はこの限りでない）。解答用紙が足りない場合は別途解答用紙を印刷し、ホッチキス止めしたものを提出すること。その際は、それぞれの解答用紙に氏名、学籍番号を記載すること。

[問 1]

データの組 (x_i, y_i) ($i = 1, 2, \dots, N$) がある。これらのデータを最小二乗法を用いて 1 次式 $y = ax + b$ で近似する。このとき、

$$a = \frac{\text{Cov}(X, Y)}{\sigma_x^2}$$

$$b = \bar{y} - a\bar{x}$$

で表現されることを示せ。ここで、 $\text{Cov}(X, Y)$ はデータの共分散であり、

$$\text{Cov}(X, Y) = \overline{xy} - \bar{x} \cdot \bar{y}$$

と表現される。ここで、 \bar{x} , \bar{y} , \overline{xy} はそれぞれ x , y , xy の平均を表し、

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i, \quad \bar{y} = \frac{1}{N} \sum_{i=1}^N y_i, \quad \overline{xy} = \frac{1}{N} \sum_{i=1}^N x_i y_i$$

である。また、 σ_x^2 は x の分散であり、

$$\sigma_x^2 = \overline{x^2} - (\bar{x})^2$$

と表現される。ここで、 $\overline{x^2}$ は x^2 の平均であり、

$$\overline{x^2} = \frac{1}{N} \sum_{i=1}^N x_i^2$$

である。