

システム最適化演習問題解答 – その1-2

4.

平面上にデータ点 $(x_1, y_1), \dots, (x_N, y_N)$ が与えられたとする。次に入力 \bar{x} だけが得られたとき、 \bar{x} から最良の出力 \bar{y} を予測したい。予測式は一次式

$$y = ax + b$$

と仮定する。係数 a, b をどのように決めればよいか。

ガウスやルジャンドルによって、与えられたデータにもとづく予測値 $ax_i + b$ と実際の出力値 y_i の差の二乗和

$$F(a, b) = \frac{1}{2} \sum_{i=1}^N (y_i - ax_i - b)^2$$

を最小にする解 a, b を利用することが提案された。今日まで最も良く利用されている最小二乗法である。

問1a：最小二乗法による解 a, b は $\frac{\partial F}{\partial a} = 0, \frac{\partial F}{\partial b} = 0$ を解くことによって求められる。これらの計算を行い、解 a, b を求めるための連立方程式を導出せよ。

問1b：この連立方程式を解いて最小二乗法による解 a, b を求めよ。

問1c：5つのデータ $(-2, 0), (-1, 0), (0, -1), (1, -2), (2, -3)$ について式 $y = ax + b$ を求め、図示せよ。

(解答)

1a.

$$\frac{\partial F}{\partial a} = \sum_i x_i (ax_i + b - y_i) = 0$$

$$\frac{\partial F}{\partial b} = \sum_i (ax_i + b - y_i) = 0$$

より

$$\left(\sum_{i=1}^N x_i^2\right)a + \left(\sum_{i=1}^N x_i\right)b = \sum_{i=1}^N x_i y_i$$

$$\left(\sum_{i=1}^N x_i\right)a + Nb = \sum_{i=1}^N y_i$$

1b.

$S_{x^2} = \sum_{i=1}^N x_i^2, S_{xy} = \sum_{i=1}^N x_i y_i, T_x = \sum_{i=1}^N x_i, T_y = \sum_{i=1}^N y_i$ とおくと、

$$S_{x^2}a + T_x b = S_{xy}$$

$$T_x a + Nb = T_y$$

これを解いて $a = \frac{NS_{xy} - T_x T_y}{NS_{x^2} - T_x^2}, b = \frac{S_{x^2} T_y - S_{xy} T_x}{NS_{x^2} - T_x^2}$.

1c. $a = -0.8, b = -1.2$. グラフは省略するが、各自書いてみることに。