最小二乗法

大枝 真一

1　はじめに

複雑なデータや関数を簡単な関数の和で近似する代表的な手法が「最小二乗法」である．これはコンピュータによるデータ解析の最も重要な基礎である．本授業ではこれを学ぶと共に，ベクトルや行列による線形計算に慣れることを目的とする．

2　最小二乗法とは

N 個のデータ $(x_1, y_1), \ldots, (x_N, y_N)$ に直線を当てはめたいとする．当てはめたい直線を $y = ax + b$ と置く．a, b はこれから定める未知の定数である．

理想的には $y_\alpha = ax_\alpha + b, \alpha = 1, \ldots, N$，となることが望ましいが，データ点 (x_α, y_α) が厳密に同一直線上にあるとは限らないので，a, b をどう選んでも多くの α に対して $y_\alpha \neq ax_\alpha + b$ となる．そこで

$$y_\alpha \approx ax_\alpha + b, \qquad \alpha = 1, \ldots, N \tag{1}$$

となるように a, b を定める (図 1)．記号 $\approx$ は「ほぼ等しい」という意味である．これを次のように解釈する．ただし $\rightarrow$ はその左側の式を最小にすることを表す．

$$J(a, b) = \frac{1}{2} \sum_{\alpha=1}^{N} (y_\alpha - (ax_\alpha + b))^2 \rightarrow \min \tag{2}$$

これは食い違いの二乗の和を最小にする方法であることから，**最小二乗法**と呼ばれている．全体を 1/2 倍するのは後の計算を見やすくするためである．

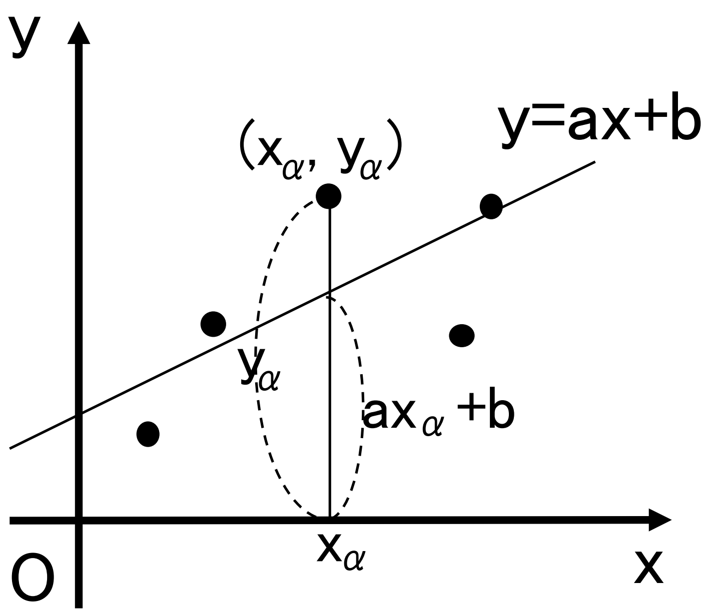

図 1　直線の当てはめ

2.1　1 次の最小二乗法の正規方程式

N 個のデータ $(x_1, y_1), \ldots, (x_N, y_N)$ に直線 $y = ax + b$ を当てはめよ.

(解)

式 (2) は a, b の関数である. 解析学で知られるように多変数の関数が最大値や最小値をとる点では，各変数に関する偏導関数が 0 でなければならない．したがって,

$$\frac{\partial J(a,b)}{\partial a} = 0, \qquad \frac{\partial J(a,b)}{\partial b} = 0 \tag{3}$$

を解いて a, b を定めればよい．式 (2) を a, b でそれぞれ偏微分すると次式を得る.

$$\frac{\partial J(a,b)}{\partial a} = \sum_{\alpha=1}^{N}(y_\alpha - ax_\alpha - b)(-x_\alpha) = a\sum_{\alpha=1}^{N} x_\alpha^2 + b\sum_{\alpha=1}^{N} x_\alpha - \sum_{\alpha=1}^{N} x_\alpha y_\alpha = 0$$

$$\frac{\partial J(a,b)}{\partial b} = \sum_{\alpha=1}^{N}(y_\alpha - ax_\alpha - b)(-1) = a\sum_{\alpha=1}^{N} x_\alpha + b\sum_{\alpha=1}^{N} 1 - \sum_{\alpha=1}^{N} y_\alpha = 0 \tag{4}$$

これから次の連立 1 次方程式を得る．

$$\begin{pmatrix} \sum_{\alpha=1}^{N} x_\alpha^2 & \sum_{\alpha=1}^{N} x_\alpha \\ \sum_{\alpha=1}^{N} x_\alpha & \sum_{\alpha=1}^{N} 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \sum_{\alpha=1}^{N} x_\alpha y_\alpha \\ \sum_{\alpha=1}^{N} y_\alpha \end{pmatrix} \tag{5}$$

これを**正規方程式**と呼ぶ．これを解いて a, b が定まる．

2.2　2 次の最小二乗法の正規方程式

N 個のデータ $(x_1, y_1), \ldots, (x_N, y_N)$ に 2 次式 $y = ax^2 + bx + c$ を当てはめよ．

(解) 当てはめる 2 次式を $y = ax^2 + bx + c$ とし，

$$y_\alpha \approx ax_\alpha^2 + bx + c, \qquad \alpha = 1, \ldots, N \tag{6}$$

となる a, b, c を最小二乗法

$$J(a,b,c) = \frac{1}{2}\sum_{\alpha=1}^{N}(y_\alpha - (ax_\alpha^2 + bx + c))^2 \to \min \tag{7}$$

によって定める．それには

$$\frac{\partial J(a,b,c)}{\partial a} = 0, \qquad \frac{\partial J(a,b,c)}{\partial b} = 0, \qquad \frac{\partial J(a,b,c)}{\partial c} = 0 \tag{8}$$

を解いて a, b, c を定めればよい．式 (7) を a, b, c でそれぞれ偏微分すると次式を得る．

$$\begin{aligned}
\frac{\partial J(a,b,c)}{\partial a} &= \sum_{\alpha=1}^{N}(y_\alpha - ax_\alpha^2 - bx - c)(-x_\alpha^2) = a\sum_{\alpha=1}^{N} x_\alpha^4 + b\sum_{\alpha=1}^{N} x_\alpha^3 + c\sum_{\alpha=1}^{N} x_\alpha^2 - \sum_{\alpha=1}^{N} x_\alpha^2 y_\alpha = 0 \\
\frac{\partial J(a,b,c)}{\partial b} &= \sum_{\alpha=1}^{N}(y_\alpha - ax_\alpha^2 - bx_\alpha - c)(-x\alpha) = a\sum_{\alpha=1}^{N} x_\alpha^3 + b\sum_{\alpha=1}^{N} x_\alpha^2 + c\sum_{\alpha=1}^{N} x_\alpha - \sum_{\alpha=1}^{N} x_\alpha y_\alpha = 0 \\
\frac{\partial J(a,b,c)}{\partial c} &= \sum_{\alpha=1}^{N}(y_\alpha - ax_\alpha^2 - bx_\alpha - c)(-1) = a\sum_{\alpha=1}^{N} x_\alpha^2 + b\sum_{\alpha=1}^{N} x_\alpha + c\sum_{\alpha=1}^{N} 1 - \sum_{\alpha=1}^{N} y_\alpha = 0
\end{aligned} \tag{9}$$

これから次の連立 1 次方程式を得る．

$$\begin{pmatrix} \sum_{\alpha=1}^N x_\alpha^4 & \sum_{\alpha=1}^N x_\alpha^3 & \sum_{\alpha=1}^N x_\alpha^2 \\ \sum_{\alpha=1}^N x_\alpha^3 & \sum_{\alpha=1}^N x_\alpha^2 & \sum_{\alpha=1}^N x_\alpha \\ \sum_{\alpha=1}^N x_\alpha^2 & \sum_{\alpha=1}^N x_\alpha & \sum_{\alpha=1}^N 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} \sum_{\alpha=1}^N x_\alpha^2 y_\alpha \\ \sum_{\alpha=1}^N x_\alpha y_\alpha \\ \sum_{\alpha=1}^N y_\alpha \end{pmatrix} \tag{10}$$

これを解いて，a, b, c が定まる．

2.3 n 次の最小二乗法の正規方程式

N 個のデータ $(x_1, y_1), \ldots, (x_N, y_N)$ に n 次式 $y = c_0 x^n + c_1 x^{n-1} + \ldots + c_n$ を当てはめよ．

(解) 当てはめる n 次式を $y = c_0 x^n + c_1 x^{n-1} + \ldots + c_n$ とし，

$$y_\alpha \approx c_0 x_\alpha^n + c_1 x_\alpha^{n-1} + \ldots + c_n, \qquad \alpha = 1, \ldots, N \tag{11}$$

となる $c_1, \ldots, c_n$ を最小二乗法

$$J(c_0, \cdots, c_n) = \frac{1}{2} \sum_{\alpha=1}^N (y_\alpha - (c_0 x_\alpha^n + c_1 x_\alpha^{n-1} + \ldots + c_n))^2 \to \min \tag{12}$$

によって定める．それには

$$\frac{\partial J(c_0, \cdots, c_n)}{\partial c_0} = 0, \qquad \frac{\partial J(c_0, \cdots, c_n)}{\partial c_1} = 0, \qquad \ldots, \qquad \frac{\partial J(c_0, \cdots, c_n)}{\partial c_n} = 0 \tag{13}$$

を解いて $c_1, \ldots, c_n$ を定めればよい．式 (12) を c_k で偏微分すると次式を得る．

$$\frac{\partial J}{\partial c_k} = \sum_{\alpha=1}^N (y_\alpha - c_0 x_\alpha^n - c_1 x_\alpha^{n-1} - \cdots - c_n)(-x_\alpha^{n-k})$$

$$= c_0 \sum_{\alpha=1}^N x_\alpha^{2n-k} + c_1 \sum_{\alpha=1}^N x_\alpha^{2n-k-1} \cdots + c_n \sum_{\alpha=1}^N x_\alpha^{n-k} - \sum_{\alpha=1}^N x_\alpha^{n-k} y_\alpha \tag{14}$$

これを 0 と置いて $k = 0, 1, \ldots, n$ に対する式を並べると次の正規方程式を得る．

$$\begin{pmatrix} \sum_{\alpha=1}^N x_\alpha^{2n} & \sum_{\alpha=1}^N x_\alpha^{2n-1} & \cdots & \sum_{\alpha=1}^N x_\alpha^n \\ \sum_{\alpha=1}^N x_\alpha^{2n-1} & \sum_{\alpha=1}^N x_\alpha^{2n-2} & \cdots & \sum_{\alpha=1}^N x_\alpha^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{\alpha=1}^N x_\alpha^n & \sum_{\alpha=1}^N x_\alpha^{n-1} & \cdots & \sum_{\alpha=1}^N 1 \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} \sum_{\alpha=1}^N x_\alpha^n y_\alpha \\ \sum_{\alpha=1}^N x_\alpha^{n-1} y_\alpha \\ \vdots \\ \sum_{\alpha=1}^N y_\alpha \end{pmatrix} \tag{15}$$

これを解いて，$c_1, \ldots, c_n$ が定まる．