

最小二乗法

山内 仁喬

2021 年 7 月 17 日

1 線形最小二乗法

次のような n 組みのデータのあてはめ問題を考える:

$$\text{観測点: } t_1, t_2, \dots, t_n \quad (1)$$

$$\text{測定値: } f_1, f_2, \dots, f_n \quad (2)$$

$$\text{測定値の分散: } \sigma_1, \sigma_2, \dots, \sigma_n \quad (3)$$

測定の誤差 (分散) σ_i^2 は, 測定値 f_i の信頼性を表す尺度として考えることができる. このデータを, ある一次独立な関数系

$$\phi_1(t), \phi_2(t), \dots, \phi_m(t) \quad (4)$$

の一次結合

$$f(t) = \sum_j^m x_j \phi_j(t) \quad (5)$$

によってあてはめる. この形の $f(t)$ は $\{\phi_j(t)\}$ に関して線形であるから, 線形モデルと呼ばれる. 一次独立な関数系として, 単項式

$$\phi_j(t) = t^{j-1} \quad (6)$$

が最も広く採用されている.

最小二乗法は,

$$S(x_1, x_2, \dots, x_m) = \sum_{i=1}^n \frac{[f_i - f(t_i)]^2}{\sigma_i^2} \quad (7)$$

$$= \sum_{i=1}^n \frac{[f_i - \sum_{j=1}^m x_j \phi_j(t_i)]^2}{\sigma_i^2} \quad (8)$$

を最小にすることによって, 未知係数 x_1, \dots, x_m の組みを決定する方法である. S を最小にする条件は,

$$\frac{\partial S}{\partial x_k} = 0, \quad k = 1, 2, \dots, m \quad (9)$$

によって与えられるので,

$$\sum_{i=1}^n \frac{[f_i - \sum_{j=1}^m x_j \phi_j(t_i)] \phi_k(t_i)}{\sigma_i^2}, \quad k = 1, 2, \dots, m \quad (10)$$

とかける. ここで i, j 成分が

$$A_{ij} = \frac{\phi_j(t_i)}{\sigma_i}, \quad (i = 1, 2, \dots, n, j = 1, 2, \dots, m) \quad (11)$$

で定義される $n \times m$ 行列を導入する. この A_{ij} はヤコビアン行列あるいは計画行列と呼ばれている. このとき, 式 (10) は

$$\sum_{i=1}^n \sum_{j=1}^m \frac{\phi_j(t_i)}{\sigma_i} \frac{\phi_k(t_i)}{\sigma_i} = \sum_{i=1}^n \frac{\phi_k(t_i)}{\sigma_i} \frac{f_i}{\sigma_i} \quad (12)$$

$$\rightarrow \sum_{j=1}^m \left(\sum_{i=1}^n A_{ik} A_{ij} \right) x_j = \sum_{i=1}^n A_{ik} \frac{f_i}{\sigma_i}, \quad (k = 1, 2, \dots, m) \quad (13)$$

となる. さらに

$$b_i = \frac{f_i}{\sigma_i}, \quad (i = 1, 2, \dots, n) \quad (14)$$

を第 i 成分にもつベクトル \mathbf{b} , x_j を第 j 成分にもつベクトルを \mathbf{x} とおくと, 式 (13) は次のようにかける:

$$A^t A \mathbf{x} = A^t \mathbf{b} \quad (15)$$

ここで, A^t は行列 A の転置である. これを正規方程式という. この方程式を解けば, 未知係数 x_i を求めることができる.

1.1 単純な多項式 $\phi_j(t) = t^{j-1}$ の場合

単純な多項式 $\phi_j(t) = t^{j-1}$ を用いた, 最小二乗法を考える.

$$f(t) = \sum_{j=1}^m x_j t^{j-1} = x_1 + x_2 t + \dots + x_m t^{m-1}, \quad (16)$$

$$A_{ij} = \frac{t_i^{j-1}}{\sigma_i}, \quad (17)$$

$$A_{ik} = \frac{t_i^{k-1}}{\sigma_i}, \quad (18)$$

$$b_i = \frac{f_i}{\sigma_i} \quad (19)$$

であるので, 式 (13) に代入すると,

$$\sum_{j=1}^m \left(\sum_{i=1}^n \frac{t_i^{k-1}}{\sigma_i} \frac{t_i^{j-1}}{\sigma_i} \right) x_j = \sum_{i=1}^n \frac{t_i^{k-1}}{\sigma_i} \frac{f_i}{\sigma_i} \quad (20)$$

を得る. 行列形式で愚直に書くと,

$$\begin{bmatrix} \sum_{i=1}^n \frac{1}{\sigma_i^2} t_i^0 t_i^0 & \sum_{i=1}^n \frac{1}{\sigma_i^2} t_i^0 t_i^1 & \cdots & \sum_{i=1}^n \frac{1}{\sigma_i^2} t_i^0 t_i^{m-1} \\ \sum_{i=1}^n \frac{1}{\sigma_i^2} t_i^1 t_i^0 & \sum_{i=1}^n \frac{1}{\sigma_i^2} t_i^1 t_i^1 & \cdots & \sum_{i=1}^n \frac{1}{\sigma_i^2} t_i^1 t_i^{m-1} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^n \frac{1}{\sigma_i^2} t_i^m t_i^0 & \sum_{i=1}^n \frac{1}{\sigma_i^2} t_i^m t_i^1 & \cdots & \sum_{i=1}^n \frac{1}{\sigma_i^2} t_i^m t_i^{m-1} \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_m \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n \frac{1}{\sigma_i^2} f_i \\ \sum_{i=1}^n \frac{1}{\sigma_i^2} f_i t_i \\ \vdots \\ \sum_{i=1}^n \frac{1}{\sigma_i^2} f_i t_i^{m-1} \end{bmatrix} \quad (21)$$

となる。左辺の左の行列について、

$$D = \begin{bmatrix} \sum_{i=1}^n \frac{1}{\sigma_i^2} t_i^0 & \sum_{i=1}^n \frac{1}{\sigma_i^2} t_i^1 & \cdots & \sum_{i=1}^n \frac{1}{\sigma_i^2} t_i^{m-1} \\ \sum_{i=1}^n \frac{1}{\sigma_i^2} t_i^1 & \sum_{i=1}^n \frac{1}{\sigma_i^2} t_i^2 & \cdots & \sum_{i=1}^n \frac{1}{\sigma_i^2} t_i^m \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^n \frac{1}{\sigma_i^2} t_i^{m-1} & \sum_{i=1}^n \frac{1}{\sigma_i^2} t_i^m & \cdots & \sum_{i=1}^n \frac{1}{\sigma_i^2} t_i^{2m-1} \end{bmatrix} \quad (22)$$

と置くと、正規方程式は

$$\begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_m \end{bmatrix} = D^{-1} \begin{bmatrix} \sum_{i=1}^n \frac{1}{\sigma_i^2} f_i \\ \sum_{i=1}^n \frac{1}{\sigma_i^2} f_i t_i \\ \vdots \\ \sum_{i=1}^n \frac{1}{\sigma_i^2} f_i t_i^{m-1} \end{bmatrix} \quad (23)$$

$$= \begin{bmatrix} D_{11}^{-1} & D_{12}^{-1} & \cdots & D_{1m}^{-1} \\ D_{21}^{-1} & D_{22}^{-1} & \cdots & D_{2m}^{-1} \\ \vdots & \vdots & \ddots & \vdots \\ D_{m1}^{-1} & D_{m2}^{-1} & \cdots & D_{mm}^{-1} \end{bmatrix} \begin{bmatrix} \sum_{i=1}^n \frac{1}{\sigma_i^2} f_i \\ \sum_{i=1}^n \frac{1}{\sigma_i^2} f_i t_i \\ \vdots \\ \sum_{i=1}^n \frac{1}{\sigma_i^2} f_i t_i^{m-1} \end{bmatrix} \quad (24)$$

$$= \begin{bmatrix} D_{11}^{-1} \sum_{i=1}^n \frac{1}{\sigma_i^2} f_i + D_{12}^{-1} \sum_{i=1}^n \frac{1}{\sigma_i^2} f_i t_i + \cdots + D_{1m}^{-1} \sum_{i=1}^n \frac{1}{\sigma_i^2} f_i t_i^{m-1} \\ D_{21}^{-1} \sum_{i=1}^n \frac{1}{\sigma_i^2} f_i + D_{22}^{-1} \sum_{i=1}^n \frac{1}{\sigma_i^2} f_i t_i + \cdots + D_{2m}^{-1} \sum_{i=1}^n \frac{1}{\sigma_i^2} f_i t_i^{m-1} \\ \vdots \\ D_{m1}^{-1} \sum_{i=1}^n \frac{1}{\sigma_i^2} f_i + D_{m2}^{-1} \sum_{i=1}^n \frac{1}{\sigma_i^2} f_i t_i + \cdots + D_{mm}^{-1} \sum_{i=1}^n \frac{1}{\sigma_i^2} f_i t_i^{m-1} \end{bmatrix} \quad (25)$$

のように解くことができる。また、係数の誤差は、

$$\sigma_{xk} = \sqrt{\sum_{i=1}^n \left(\frac{\partial x_k}{\partial f_i} \sigma_i \right)^2} \quad (26)$$

であるので、式 (25) を用いると具体的に

$$\sigma_{xk} = \sqrt{\sum_{i=1}^n \left(\sum_{j=1}^m D_{kj}^{-1} \frac{1}{\sigma_i} t_i^{j-1} \right)^2} \quad (27)$$

と計算される。

1.2 具体例: 一次関数 $f(t) = x_0 t + x_1$ で最小二乗法

一次関数 $f(t) = x_0 t + x_1$ で最小二乗法を実行するとき、具体的な未知係数と誤差の表式を見ていく。行列 (22) を具体的に計算すると、

$$D = \begin{bmatrix} \sum_{i=1}^n \frac{1}{\sigma_i^2} & \sum_{i=1}^n \frac{1}{\sigma_i^2} t_i \\ \sum_{i=1}^n \frac{1}{\sigma_i^2} t_i & \sum_{i=1}^n \frac{1}{\sigma_i^2} t_i^2 \end{bmatrix} \equiv \begin{bmatrix} \sum_{i=1}^n w_i & \sum_{i=1}^n w_i t_i \\ \sum_{i=1}^n w_i t_i & \sum_{i=1}^n w_i t_i^2 \end{bmatrix} \quad (28)$$

なので、逆行列は

$$\mathbf{D}^{-1} = \frac{1}{\Delta} \begin{bmatrix} \sum_{i=1}^n w_i t_i^2 & \sum_{i=1}^n w_i t_i \\ \sum_{i=1}^n w_i t_i & \sum_{i=1}^n w_i t_i^2 \end{bmatrix} \quad (29)$$

$$\Delta \equiv \left(\sum_{i=1}^n w_i \right) \left(\sum_{i=1}^n w_i t_i^2 \right) - \left(\sum_{i=1}^n w_i t_i \right)^2 \quad (30)$$

と計算される。ここで、 $w_i = 1/\sigma_i^2$ とおいた。したがって、未知係数は

$$x_0 = \frac{(\sum_{i=1}^n w_i t_i^2) (\sum_{i=1}^n w_i f_i) - (\sum_{i=1}^n w_i t_i) (\sum_{i=1}^n w_i t_i f_i)}{\Delta} \quad (31)$$

$$x_1 = \frac{(\sum_{i=1}^n w_i) (\sum_{i=1}^n w_i t_i f_i) - (\sum_{i=1}^n w_i t_i) (\sum_{i=1}^n w_i f_i)}{\Delta} \quad (32)$$

と計算される。

続いて、係数の誤差を求めていく。 x_0 の誤差は、

$$\sigma_{x_0} = \sqrt{\sum_{i=1}^n \left(\frac{\partial x_0}{\partial f_i} \sigma_i \right)^2} \quad (33)$$

である。以下、具体的に計算をしていく：

$$\frac{\partial x_0}{\partial f_i} \sigma_i = \frac{1}{\Delta} \left[\left(\sum_{i=1}^n w_i t_i^2 \right) w_i - \left(\sum_{i=1}^n w_i t_i \right) w_i t_i \right] \sigma_i \quad (34)$$

$$\left(\frac{\partial x_0}{\partial f_i} \sigma_i \right)^2 = \frac{1}{\Delta^2} \left[\left(\sum_{i=1}^n w_i t_i^2 \right)^2 w_i^2 + \left(\sum_{i=1}^n w_i t_i \right)^2 w_i^2 t_i^2 - 2 \left(\sum_{i=1}^n w_i t_i^2 \right) \left(\sum_{i=1}^n w_i t_i \right) w_i^2 t_i \right] \sigma_i^2 \quad (35)$$

$$= \frac{1}{\Delta^2} \left[\left(\sum_{i=1}^n w_i t_i^2 \right)^2 w_i + \left(\sum_{i=1}^n w_i t_i \right)^2 w_i t_i^2 - 2 \left(\sum_{i=1}^n w_i t_i^2 \right) \left(\sum_{i=1}^n w_i t_i \right) w_i t_i \right] \quad (36)$$

最後の式変形には、 $w_i = 1/\sigma_i^2$ であることを用いた。さらに、 i について和をとると、

$$\begin{aligned} & \sum_{i=1}^n \left(\frac{\partial x_0}{\partial f_i} \sigma_i \right)^2 \\ &= \frac{1}{\Delta^2} \left[\left(\sum_{i=1}^n w_i t_i^2 \right)^2 \left(\sum_{i=1}^n w_i \right) + \left(\sum_{i=1}^n w_i t_i \right)^2 \left(\sum_{i=1}^n w_i t_i^2 \right) - 2 \left(\sum_{i=1}^n w_i t_i^2 \right) \left(\sum_{i=1}^n w_i t_i \right)^2 \right] \\ &= \frac{1}{\Delta^2} \left[\left(\sum_{i=1}^n w_i t_i^2 \right)^2 \left(\sum_{i=1}^n w_i \right) - \left(\sum_{i=1}^n w_i t_i \right)^2 \left(\sum_{i=1}^n w_i t_i^2 \right) \right] \\ &= \frac{\sum_{i=1}^n w_i t_i^2}{\Delta^2} \left[\left(\sum_{i=1}^n w_i t_i^2 \right) \left(\sum_{i=1}^n w_i \right) - \left(\sum_{i=1}^n w_i t_i \right)^2 \right] \\ &= \frac{\sum_{i=1}^n w_i t_i^2}{\Delta} \end{aligned} \quad (37)$$

したがって、 x_0 の誤差は、

$$\sigma_{x_0} = \sqrt{\frac{\sum_{i=1}^n w_i t_i^2}{\Delta^2}} \quad (38)$$

である。

同様にして, x_1 の誤差,

$$\sigma_{x_1} = \sqrt{\sum_{i=1}^n \left(\frac{\partial x_1}{\partial f_i} \sigma_i \right)^2} \quad (39)$$

を計算する.

$$\frac{\partial x_1}{\partial f_i} \sigma_i = \frac{1}{\Delta} \left[\left(\sum_{i=1}^n w_i \right) w_i t_i - \left(\sum_{i=1}^n w_i t_i \right) w_i \right] \sigma_i \quad (40)$$

$$\left(\frac{\partial x_1}{\partial f_i} \sigma_i \right)^2 = \frac{1}{\Delta^2} \left[\left(\sum_{i=1}^n w_i \right)^2 w_i^2 t_i^2 + \left(\sum_{i=1}^n w_i t_i \right)^2 w_i^2 - 2 \left(\sum_{i=1}^n w_i \right) \left(\sum_{i=1}^n w_i t_i \right) w_i^2 t_i \right] \sigma_i^2 \quad (41)$$

$$= \frac{1}{\Delta^2} \left[\left(\sum_{i=1}^n w_i \right)^2 w_i t_i^2 + \left(\sum_{i=1}^n w_i t_i \right)^2 w_i - 2 \left(\sum_{i=1}^n w_i \right) \left(\sum_{i=1}^n w_i t_i \right) w_i t_i \right] \quad (42)$$

最後の式変形には, $w_i = 1/\sigma_i^2$ であることを用いた. i について和をとると,

$$\begin{aligned} & \sum_{i=1}^n \left(\frac{\partial x_1}{\partial f_i} \sigma_i \right)^2 \\ &= \frac{1}{\Delta^2} \left[\left(\sum_{i=1}^n w_i \right)^2 \left(\sum_{i=1}^n w_i t_i^2 \right) + \left(\sum_{i=1}^n w_i t_i \right)^2 \left(\sum_{i=1}^n w_i \right) - 2 \left(\sum_{i=1}^n w_i \right) \left(\sum_{i=1}^n w_i t_i \right) \left(\sum_{i=1}^n w_i t_i \right) \right] \\ &= \frac{\sum_{i=1}^n w_i}{\Delta^2} \left[\left(\sum_{i=1}^n w_i \right) \left(\sum_{i=1}^n w_i t_i^2 \right) - \left(\sum_{i=1}^n w_i t_i \right)^2 \right] \\ &= \frac{\sum_{i=1}^n w_i}{\Delta} \end{aligned} \quad (43)$$

したがって, x_1 の誤差は,

$$\sigma_{x_1} = \sqrt{\frac{\sum_{i=1}^n w_i}{\Delta}} \quad (44)$$

である.