最小二乗法（垂直法）その 2

2019.7.8　　鈴木　実

1　はじめに

回帰曲線への垂直距離の二乗和が最小になることが条件となる最小二乗法に関しては，2008.11.14 のエントリー [1] で述べたところである．この方法により最小二乗フィッティング曲線が得られるが，回帰曲線の種類によっては必ずしも最小値への収束が十分でない場合がある．極端な場合は，初期近似パラメータが最適値に比較的近い場合でも発散することもある．

これは，複雑な回帰関数の場合は複数の極値を持つことが考えられ，その場合はパラメータ初期値を十分適切に取らなければ最適値に収束しない．もう一つは，パラメータを計算する正規方程式が十分厳密な近似法に拠っていないという理由が考えられる．

本稿では，厳密に近似した場合の正規方程式を求める．垂直法最小二乗フィッティング計算が収束しない場合に収束が改善されるものと考えられる．ただし，数式の取り扱いが複雑になるのみで，ほとんど収束の改善が見られない場合もある．

2　これまでの正規方程式と問題点

回帰曲線を $F(x, y, a_k) = 0$ とすると，データ (x, y) と回帰曲線の間の垂直距離の二乗和は，

$$S = \sum_{i=1}^{n} \frac{F^2(x_i, y_i, a_k)}{\alpha_i^2 + \beta_i^2} \tag{1}$$

となる [1]．ここで，a_{k0} を a_k の初期値として F を微小項 $(a_k - a_{k0})$ で展開し 1 次の項までとると，

$$F(x_i, y_i, a_k) = F(x_i, y_i, a_{k_0}) + \sum_{k=1}^{n} F_{ak}(a_k - a_{k_0}) \tag{2}$$

$$F_{a_k} = \frac{\partial F}{\partial a_k}(x_i, y_i, a_{k_0}) \tag{3}$$

これを式 (1) に代入すると，

$$S = \sum_{i=1}^{n} \frac{1}{\alpha_i^2 + \beta_i^2} \left[F(x_i, y_i, a_{k_0}) + \sum_{k=1}^{m} F_{a_k}(a_k - a_{k_0}) \right]^2 \tag{4}$$

となる．この式において、S を極小とする条件 $\partial S/\partial a_k = 0 (k = 1, \cdots, m)$ より

$$\sum_{i=1}^{n} \frac{1}{\alpha_i^2 + \beta_i^2} \left[F(x_i, y_i, a_{k_0}) F_{a_k} + \sum_{j=1}^{m} F_{a_k} F_{a_j}(a_j - a_{j_0}) \right] = 0 \tag{5}$$

という $(a_k - a_{k0})$ に関する連立 1 次方程式が得られる．この式は，整理すると

$$\sum_{j=1}^{m} \left[\sum_{i=1}^{n} \frac{1}{\alpha_i^2 + \beta_i^2} F_{a_k} F_{a_j} \right] (a_j - a_{j_0}) = -\sum_{i=1}^{n} \frac{1}{\alpha_i^2 + \beta_i^2} F(x_i, y_i, a_{k_0}) F_{a_k} \tag{6}$$

となる．さらに，

$$\langle kj\rangle = \sum_{i=1}^{n}\frac{1}{\alpha_i^2+\beta_i^2}F_{a_k}F_{a_j} \tag{7}$$

$$\langle k0\rangle = -\sum_{i=1}^{n}\frac{1}{\alpha_i^2+\beta_i^2}F(x_i,y_i,a_{k_0})F_{a_k} \tag{8}$$

と書き換えると，

$$\left\{\begin{array}{lcrcl} \langle 11\rangle(a_1-a_{10})+ & \cdots & +\langle 1m\rangle(a_m-a_{m0}) & = & \langle 10\rangle \\ & \cdots & & & \\ \langle m1\rangle(a_1-a_{10})+ & \cdots & +\langle mm\rangle(a_m-a_{m0}) & = & \langle m0\rangle \end{array}\right. \tag{9}$$

これが以前のエントリーで示した正規方程式である．

しかし，この正規方程式は近似の仕方から $\alpha(x,y,a_k)=\partial F(x,y,a_k)/\partial x$，$\alpha(x,y,a_k)=\partial F(x,y,a_k)/\partial x$，と $\partial F(x,y,a_k)/\partial a_k$ が a_k の緩やかな関数の場合に成り立つ．そうでない場合は，式 (2) で展開近似するのではなく，式 (1) において展開し 2 次の項までとる必要がある．次節では，その場合の正規方程式を導く．

3 一般化

以下では，偏微分係数を以下のように略記することにする．

$$F = F(x_i,y_i,a_0,a_1,\cdots,a_n) \equiv F(x_i,y_i,a_k) \tag{10}$$

$$F_0 = F(x_i,y_i,a_0,a_1,\cdots,a_n) \equiv F(x_i,y_i,a_{k_0}) \tag{11}$$

$$F_k = \frac{\partial F}{\partial a_k}(x_i,y_i,a_{k_0}) \tag{12}$$

$$F_{kj} = \frac{\partial^2 F}{\partial a_k \partial a_j}(x_i,y_i,a_{k_0}) \tag{13}$$

$$\alpha = \frac{\partial F}{\partial x}(x_i,y_i,a_k) \tag{14}$$

$$\beta = \frac{\partial F}{\partial y}(x_i,y_i,a_k) \tag{15}$$

$$\alpha_0 = \frac{\partial F}{\partial x}(x_i,y_i,a_{k_0}) \tag{16}$$

$$\beta_0 = \frac{\partial F}{\partial y}(x_i,y_i,a_{k_0}) \tag{17}$$

$$\alpha_k = \frac{\partial \alpha}{\partial a_k}(x_i,y_i,a_{k_0}) \tag{18}$$

$$\beta_k = \frac{\partial \beta}{\partial a_k}(x_i,y_i,a_{k_0}) \tag{19}$$

$$\alpha_{kj} = \frac{\partial^2 \alpha}{\partial a_k \partial a_j}(x_i,y_i,a_{k_0}) \tag{20}$$

$$\beta_{kj} = \frac{\partial^2 \beta}{\partial a_k \partial a_j}(x_i,y_i,a_{k_0}) \tag{21}$$

F を a_{k_0} の周りに展開し，$a_k - a_{k_0}$ は十分小さいと考え、2 次の項までとる．

$$F = F_0 + \sum_k F_k(a_k-a_{k_0}) + \frac{1}{2!}\sum_{k,j}F_{kj}(a_k-a_{k_0})(a_j-a_{j_0}) \tag{22}$$

α，β についても同様に，

$$\alpha = \alpha_0 + \sum_k \alpha_k (ak - a_{k_0}) + \frac{1}{2!} \sum_{k,j} \alpha_{kj} (a_k - a_{k_0})(a_j - a_{j_0}) \tag{23}$$

$$\beta = \beta_0 + \sum_k \beta_k (ak - a_{k_0}) + \frac{1}{2!} \sum_{k,j} \beta_{kj} (a_k - a_{k_0})(a_j - a_{j_0}) \tag{24}$$

となる．これから，2 次の項までとると，

$$\begin{aligned} F^2 &= F_0^2 + 2\sum_k F_0 F_k (a_k - a_{k_0}) + \sum_{k,j} F_k F_j (a_k - a_{k_0})(a_j - a_{j_0}) + \sum_{k,j} F_0 F_{kj} (a_k - a_{k_0})(a_j - a_{j_0}) \\ &= F_0^2 + 2\sum_k F_0 F_k (a_k - a_{k_0}) + \sum_{k,j} (F_k F_j + F_0 F_{kj})(a_k - a_{k_0})(a_j - a_{j_0}) \end{aligned} \tag{25}$$

$$\begin{aligned} \alpha^2 &= \alpha_0^2 + 2\sum_k \alpha_0 \alpha_k (a_k - a_{k_0}) + \sum_{k,j} \alpha_k \alpha_j (a_k - a_{k_0})(a_j - a_{j_0}) + \sum_{k,j} \alpha_0 \alpha_{kj} (a_k - a_{k_0})(a_j - a_{j_0}) \\ &= \alpha_0^2 + 2\sum_k \alpha_0 \alpha_k (a_k - a_{k_0}) + \sum_{k,j} (\alpha_k \alpha_j + \alpha_0 \alpha_{kj})(a_k - a_{k_0})(a_j - a_{j_0}) \end{aligned} \tag{26}$$

$$\begin{aligned} \beta^2 &= \beta_0^2 + 2\sum_k \beta_0 \beta_k (a_k - a_{k_0}) + \sum_{k,j} \beta_k \beta_j (a_k - a_{k_0})(a_j - a_{j_0}) + \sum_{k,j} \beta_0 \beta_{kj} (a_k - a_{k_0})(a_j - a_{j_0}) \\ &= \beta_0^2 + 2\sum_k \beta_0 \beta_k (a_k - a_{k_0}) + \sum_{k,j} (\beta_k \beta_j + \beta_0 \beta_{kj})(a_k - a_{k_0})(a_j - a_{j_0}) \end{aligned} \tag{27}$$

とすることができる．したがって，

$$\alpha^2 + \beta^2 = \alpha_0^2 + \beta_0^2 + 2\sum_k (\alpha_0 \alpha_k + \beta_0 \beta_k)(a_k - a_{k_0}) + \sum_{k,j} (\alpha_k \alpha_j + \beta_k \beta_j + \alpha_0 \alpha_{kj} + \beta_0 \beta_{kj})(a_k - a_{k_0})(a_j - a_{j_0}) \tag{28}$$

となる．これを用いて $(\alpha^2 + \beta^2)^{-1}$ を展開し，2 次の項までをとると，

$$\begin{aligned} &\frac{\alpha_0^2 + \beta_0^2}{\alpha^2 + \beta^2} \\ &= \left[1 + \frac{2}{\alpha_0^2 + \beta_0^2} \sum_k (\alpha_0 \alpha_k + \beta_0 \beta_k)(a_k - a_{k_0}) + \frac{1}{\alpha_0^2 + \beta_0^2} \sum_{k,j} (\alpha_k \alpha_j + \beta_k \beta_j + \alpha_0 \alpha_{kj} + \beta_0 \beta_{kj})(a_k - a_{k_0})(a_j - a_{j_0}) \right]^{-1} \\ &= 1 - \frac{2}{\alpha_0^2 + \beta_0^2} \sum_k (\alpha_0 \alpha_k + \beta_0 \beta_k)(a_k - a_{k_0}) - \frac{1}{\alpha_0^2 + \beta_0^2} \sum_{k,j} (\alpha_k \alpha_j + \beta_k \beta_j + \alpha_0 \alpha_{kj} + \beta_0 \beta_{kj})(a_k - a_{k_0})(a_j - a_{j_0}) \\ &\qquad + \frac{1}{2!} \left[\frac{2}{\alpha_0^2 + \beta_0^2} \sum_k (\alpha_0 \alpha_k + \beta_0 \beta_k)(a_k - a_{k_0}) \right]^2 \\ &= 1 - \frac{2}{\alpha_0^2 + \beta_0^2} \sum_k (\alpha_0 \alpha_k + \beta_0 \beta_k)(a_k - a_{k_0}) - \frac{1}{\alpha_0^2 + \beta_0^2} \sum_{k,j} (\alpha_k \alpha_j + \beta_k \beta_j + \alpha_0 \alpha_{kj} + \beta_0 \beta_{kj})(a_k - a_{k_0})(a_j - a_{j_0}) \\ &\qquad + \frac{2}{(\alpha_0^2 + \beta_0^2)^2} \sum_{kj} (\alpha_0^2 \alpha_k \alpha_j + \alpha_0 \beta_0 \alpha_k \beta_j + \beta_0 \alpha_0 \beta_k \alpha_j + \beta_0^2 \beta_k \beta_j)(a_k - a_{k_0})(a_j - a_{j_0}) \end{aligned} \tag{29}$$

となる．式 (25) と式 (29) の積を展開し，2 次の項までをとると，

$$
\begin{aligned}
&\frac{F^2}{\alpha^2+\beta^2}(\alpha_0^2+\beta_0^2)\\
&=F_0^2+2\sum_k F_0F_k(a_k-a_{k_0})+\sum_{k,j}(F_kF_j+F_0F_{kj})(a_k-a_{k_0})(a_j-a_{j_0})\\
&\quad-\frac{2F_0^2}{\alpha_0^2+\beta_0^2}\sum_k(\alpha_0\alpha_k+\beta_0\beta_k)(a_k-a_{k_0})-\frac{F_0^2}{\alpha_0^2+\beta_0^2}\sum_{k,j}(\alpha_k\alpha_j+\beta_k\beta_j+\alpha_0\alpha_{kj}+\beta_0\beta_{kj})(a_k-a_{k_0})(a_j-a_{j_0})\\
&\quad+\frac{2F_0^2}{(\alpha_0^2+\beta_0^2)^2}\sum_{k,j}(\alpha_0^2\alpha_k\alpha_j+\alpha_0\beta_0\alpha_k\beta_j+\beta_0\alpha_0\beta_k\alpha_j+\beta_0^2\beta_k\beta_j)(a_k-a_{k_0})(a_j-a_{j_0})\\
&\quad-\frac{4}{\alpha_0^2+\beta_0^2}\sum_{k,j}F_0F_k(\alpha_0\alpha_j+\beta_0\beta_j)(a_k-a_{k_0})(a_j-a_{j_0}) \qquad (30)
\end{aligned}
$$

である．これから，

$$
\begin{aligned}
S&=\sum_i\frac{F^2}{\alpha^2+\beta^2}\\
&=\sum_i\frac{F_0^2}{\alpha_0^2+\beta_0^2}+2\sum_{i,k}\left[\frac{F_0F_k}{\alpha_0^2+\beta_0^2}-\frac{F_0^2}{(\alpha_0^2+\beta_0^2)^2}(\alpha_0\alpha_k+\beta_0\beta_k)\right](a_k-a_{k_0})\\
&\quad+\sum_{i,k,j}\left[\frac{F_kF_j+F_0F_{kj}}{\alpha_0^2+\beta_0^2}-\frac{F_0^2}{(\alpha_0^2+\beta_0^2)^2}(\alpha_k\alpha_j+\beta_k\beta_j+\alpha_0\alpha_{kj}+\beta_0\beta_{kj})\right.\\
&\quad+\frac{2F_0^2}{(\alpha_0^2+\beta_0^2)^3}(\alpha_0^2\alpha_k\alpha_j+\alpha_0\beta_0\alpha_k\beta_j+\beta_0\alpha_0\beta_k\alpha_j+\beta_0^2\beta_k\beta_j)\\
&\quad\left.-\frac{4F_0}{(\alpha_0^2+\beta_0^2)^2}F_k(\alpha_0\alpha_j+\beta_0\beta_j)\right](a_k-a_{k_0})(a_j-a_{j_0}) \qquad (31)
\end{aligned}
$$

となる．

最小点ではこれを a_k で偏微分して 0 になるから，

$$
\begin{aligned}
\frac{\partial S}{\partial a_k}&=\frac{\partial}{\partial a_k}\sum_i\frac{F^2}{\alpha^2+\beta^2}\\
&=2\sum_i\left[\frac{F_0F_k}{\alpha_0^2+\beta_0^2}-\frac{F_0^2}{(\alpha_0^2+\beta_0^2)^2}(\alpha_0\alpha_k+\beta_0\beta_k)\right]\\
&\quad+2\sum_{i,j}\left[\frac{F_kF_j+F_0F_{kj}}{\alpha_0^2+\beta_0^2}-\frac{F_0^2}{(\alpha_0^2+\beta_0^2)^2}(\alpha_k\alpha_j+\beta_k\beta_j+\alpha_0\alpha_{kj}+\beta_0\beta_{kj})\right.\\
&\quad+\frac{2F_0^2}{(\alpha_0^2+\beta_0^2)^3}(\alpha_0^2\alpha_k\alpha_j+\alpha_0\beta_0\alpha_k\beta_j+\beta_0\alpha_0\beta_k\alpha_j+\beta_0^2\beta_k\beta_j)\\
&\quad\left.-\frac{2F_0}{(\alpha_0^2+\beta_0^2)^2}(F_k\alpha_0\alpha_j+F_k\beta_0\beta_j+F_j\alpha_0\alpha_k+F_j\beta_0\beta_k)\right](a_j-a_{j_0})=0 \qquad (32)
\end{aligned}
$$

となり，

$$
\begin{aligned}
\langle\, kj\,\rangle&=\sum_i 2\left[\frac{F_kF_j+F_0F_{kj}}{\alpha_0^2+\beta_0^2}-\frac{F_0^2}{(\alpha_0^2+\beta_0^2)^2}(\alpha_k\alpha_j+\beta_k\beta_j+\alpha_0\alpha_{kj}+\beta_0\beta_{kj})\right.\\
&\quad+\frac{2F_0^2}{(\alpha_0^2+\beta_0^2)^3}(\alpha_0^2\alpha_k\alpha_j+\alpha_0\beta_0\alpha_k\beta_j+\beta_0\alpha_0\beta_k\alpha_j+\beta_0^2\beta_k\beta_j)\\
&\quad\left.-\frac{2F_0}{(\alpha_0^2+\beta_0^2)^2}(F_k\alpha_0\alpha_j+F_k\beta_0\beta_j+F_j\alpha_0\alpha_k+F_j\beta_0\beta_k)\right] \qquad (33)\\
\langle\, k0\,\rangle&=-\sum_i\left[\frac{F_0F_k}{\alpha_0^2+\beta_0^2}-\frac{F_0^2}{(\alpha_0^2+\beta_0^2)^2}(\alpha_0\alpha_k+\beta_0\beta_k)\right] \qquad (34)
\end{aligned}
$$

とおくことにより，

$$\left\{\begin{array}{lcll}\langle 11\rangle(a_1-a_{10})+ & \cdots & +\langle 1m\rangle(a_m-a_{m_0}) & = \langle 10\rangle \\ & \cdots & & \\ \langle m1\rangle(a_1-a_{10})+ & \cdots & +\langle mm\rangle(a_m-a_{m_0}) & = \langle m0\rangle \end{array}\right. \tag{35}$$

という正規方程式になる．これより $a_k-a_{k_0}$ が求められる．この結果から新しい a_{k_0} が得られるので、逐次法で近似の精度を高くすることができる。

参考文献

[1] 「最小二乗法（垂直法）」(2008/11/14 のエントリー)
`http://totoha.web.fc2.com/least_square_1.pdf`

(2019.10.12　修正)