最小2乗法の基礎

1　最小2乗法の原理

ある量を同じ条件の下で n 回測定をして，測定値 x_1, x_2, $\cdots$, x_n を得たとする．この量の真の値を M とすると，誤差 ε_1, ε_2, $\cdots$, ε_n は，それぞれ

$$\varepsilon_1 = x_1 - M,\ \varepsilon_2 = x_2 - M,\ \cdots,\ \varepsilon_n = x_n - M$$

で表される．測定は互いに独立であり，かつ，誤差はガウスの誤差の法則に従うとすると，これらの誤差が生じる確率は，それぞれ，

$$\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{\varepsilon_1^2}{2\sigma^2}}d\varepsilon,\ \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{\varepsilon_2^2}{2\sigma^2}}d\varepsilon,\ \cdots,\ \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{\varepsilon_n^2}{2\sigma^2}}d\varepsilon$$

である．これらの誤差が順に起こる確率は，上記のものを掛け合わせていくことになるので，以下の式のようになる．

$$\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n e^{-\frac{\varepsilon_1^2+\varepsilon_2^2+\ \cdots\ +\varepsilon_n^2}{2\sigma^2}}(d\varepsilon)^2 = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n e^{-\frac{(x_1-M)^2+(x_2-M)^2+\ \cdots\ +(x_n-M)^2}{2\sigma^2}}(d\varepsilon)^2$$

上式の確率が最大になる M の値が最も確からしい値と考えられる．そのためには，

$$(x_1-M)^2+(x_2-M)^2+\ \cdots\ +(x_n-M)^2$$

の部分が最小となる M の値を求めるということになる．そのためには上式を微分して極値を取る条件を探せばよいので，

$$\frac{d}{dM}\left[(x_1-M)^2+(x_2-M)^2+\ \cdots\ +(x_n-M)^2\right]=0$$

となればよいので，結局，

$$(x_1-M)+(x_2-M)+\ \cdots\ +(x_n-M)=0$$

より，

$$M=\frac{x_1+x_2+\ \cdots\ +x_n}{n}$$

となる．すなわち，真の値として最も確からしいものは測定値の算術平均（標本平均）ということである．

2　間接測定における最小2乗法

前節では単純に測定した値についての検討であったが，測定値が何かの変量の関数として与えられるような関係にある場合がある．例えば，鉛直方向につるしたバネに重りをつけてバネの長さを測る場合には，伸び自体は重りの質量に対して直線的に増加するはずであるが，バネ自体の自重により最初から少し伸びているので，その長さを知らなければ，力と伸びの関係を厳密に議論することはできない．そのような場合には，バネの伸びを l，重りの質量を m とすると，以下のように，

$$l = km + l_0$$

として1次関数で表現できるので，この式を求めることができれば自重による伸び l_0 を見積もることができる．このように一次式で表されるもっとも簡単な間接測定における最小2乗法について考えていこう．

2.1　未知の係数が 2 個の場合

一番単純な最小 2 乗法の例は以下の式で表されるような関係の場合である．

$$y = a + bx \tag{1}$$

このとき，測定を繰り返していくと，以下のような関係式が作られていく．

$$
\begin{aligned}
y_1 &= a + bx_1 && (2)\\
y_2 &= a + bx_2 && (3)\\
&\cdots \\
y_n &= a + bx_n && (4)
\end{aligned}
$$

上式においてそれぞれの式の誤差 ε_i は

$$\varepsilon_i = y_i - (a + bx_i) \tag{5}$$

であるので，2 乗誤差全体では，

$$\varepsilon^2 = \sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} [y_i - (a + bx_i)]^2 \tag{6}$$

を最小にすることが目的となる．表記が煩雑となるので，以後は添え字を省略して以下のように総和の記号を表記するものとする．

$$\varepsilon^2 = \sum [y_i - (a + bx_i)]^2 \tag{7}$$

x_i や y_i は誤差を含んでいるので，式 (2)〜(4) のような観測方程式における係数 a と b は式ごとにそれぞれ異なる結果となる．そこで，式 (6) より，この測定系の 2 乗誤差 ε^2 を最小にする係数 a と b を求めればよいことになる．よって，以下の連立方程式から未知の係数 a と b を求める．

$$
\begin{aligned}
\frac{\partial}{\partial a} \sum [y_i - (a + bx_i)]^2 &= 0 && (8)\\
\frac{\partial}{\partial b} \sum [y_i - (a + bx_i)]^2 &= 0 && (9)
\end{aligned}
$$

【演習】
式 (8) と (9) から変数 a と b を求めよ．

【解法】
式 (8) より

$$
\begin{aligned}
\frac{\partial}{\partial a} \sum (y_i^2 + a^2 + b^2 x_i^2 - 2ay_i - 2by_ix_i + 2abx_i) &= 0\\
\sum (2a - 2y_i + 2bx_i) &= 0\\
na + (\sum x_i)b &= \sum y_i && (10)
\end{aligned}
$$

同様に，式 (9) から

$$(\sum x_i)a + (\sum x_i^2)b = \sum x_i y_i \tag{11}$$

を得る．これらは単純な 2 元連立方程式であるので，簡単に解くことも可能であるが，ここでは行列表記で考えてみよう．すなわち，

$$\begin{pmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \sum y_i \\ \sum x_i y_i \end{pmatrix} \tag{12}$$

を解いて最も確からしい変数 a と b の値を決定する．線形代数で学習したように，これは左から逆行列をかける方法と Cramer の方法かのどちらかを使えばよいので，まずは逆行列を求める方法を見てみる．行列式の値を Δ とすると，逆行列は一般に以下の式で表される．

$$\frac{1}{\Delta} \begin{pmatrix} \Delta_{11} & \Delta_{12} \\ \Delta_{21} & \Delta_{22} \end{pmatrix} \tag{13}$$

ここで，Δ_{ij} は**余因子**と呼ばれるものであり，行列の j 行 i 列を抜いたものの行列式に $(-1)^{i+j}$ をかけたものである．式 (12) の場合には逆行列は式 (13) を用いて以下のように求められる．

$$\begin{pmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{pmatrix}^{-1} = \frac{1}{n \sum x_i^2 - \left(\sum x_i\right)^2} \begin{pmatrix} \sum x_i^2 & -\sum x_i \\ -\sum x_i & n \end{pmatrix} \tag{14}$$

よって，

$$\begin{aligned} \begin{pmatrix} a \\ b \end{pmatrix} &= \frac{1}{n \sum x_i^2 - \left(\sum x_i\right)^2} \begin{pmatrix} \sum x_i^2 & -\sum x_i \\ -\sum x_i & n \end{pmatrix} \begin{pmatrix} \sum y_i \\ \sum x_i y_i \end{pmatrix} \\ &= \frac{1}{n \sum x_i^2 - \left(\sum x_i\right)^2} \begin{pmatrix} \left(\sum x_i^2\right)\left(\sum y_i\right) - \left(\sum x_i\right)\left(\sum x_i y_i\right) \\ -\left(\sum x_i\right)\left(\sum y_i\right) + n \sum x_i y_i \end{pmatrix} \end{aligned} \tag{15}$$

と求められる．一方，Cramer の公式は $A\boldsymbol{x} = \boldsymbol{d}$ のような形の式があった場合，例えば 3 行 3 列であれば以下のようなものであることを思い出してほしい．

$$\begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix} \tag{16}$$

の場合には，それぞれの変数 x, y および z の値は A の行列式の値を $|A|$ とすると以下のように求められる．

$$x = \frac{1}{|A|} \begin{vmatrix} d_1 & A_{12} & A_{13} \\ d_2 & A_{22} & A_{23} \\ d_3 & A_{32} & A_{33} \end{vmatrix} \tag{17}$$

$$y = \frac{1}{|A|} \begin{vmatrix} A_{11} & d_1 & A_{13} \\ A_{21} & d_2 & A_{23} \\ A_{31} & d_3 & A_{33} \end{vmatrix} \tag{18}$$

$$z = \frac{1}{|A|} \begin{vmatrix} A_{11} & A_{12} & d_1 \\ A_{21} & A_{22} & d_2 \\ A_{31} & A_{32} & d_3 \end{vmatrix} \tag{19}$$

よって前述の a と b は，以下のように求められる．

$$a = \frac{1}{\Delta} \begin{vmatrix} \sum y_i & \sum x_i \\ \sum x_i y_i & \sum x_i^2 \end{vmatrix} \tag{20}$$

$$b = \frac{1}{\Delta} \begin{vmatrix} n & \sum y_i \\ \sum x_i & \sum x_i y_i \end{vmatrix} \tag{21}$$

2.2　一般的な線型方程式の最小 2 乗法

行列の解法を用いることで，前節の係数が 2 個のものを容易に拡張できる．以下のような線形関係にあると予想される測定量を考える．

$$y = ax_1 + bx_2 + \cdots + lx_m \tag{22}$$

観測方程式から前節と同様に誤差を求める．

$$\begin{aligned}
\varepsilon_1 &= y_1 - (ax_{11} + bx_{21} + \cdots + lx_m1) \\
\varepsilon_2 &= y_2 - (ax_{12} + bx_{22} + \cdots + lx_m2) \\
&\cdots \\
\varepsilon_n &= y_n - (ax_{1n} + bx_{2n} + \cdots + lx_mn)
\end{aligned}$$

2 乗誤差の和を最小にするために上式を変形していくが，経過は煩雑なので興味があれば自分で試してみて欲しい．結果としては，以下のような正規方程式の形に変形できる．

$$\begin{aligned}
(\sum x_{1i}^2)a + (\sum x_{1i}x_{2i})b + \cdots + (\sum x_{1i}x_{mi})l &= \sum x_{1i}y_i \\
(\sum x_{1i}x_{2i})a + (\sum x_{2i}^2)b + \cdots + (\sum x_{2i}x_{mi})l &= \sum x_{2i}y_i \\
\cdots \\
(\sum x_{1i}x_{mi})a + (\sum x_{2i}x_{mi})b + \cdots + (\sum x_{mi}^2)l &= \sum x_{mi}y_i
\end{aligned}$$

よって，以下の行列で表される連立方程式を解くことにより最も確からしい係数 $a, b, \cdots, l$ を求めることができる．

$$\begin{pmatrix}
\sum x_{1i}^2 & \sum x_{1i}x_{2i} & \cdots & \sum x_{1i}x_{mi} \\
\sum x_{1i}x_{2i} & \sum x_{2i}^2 & \cdots & \sum x_{1i}x_{mi} \\
 & & \cdots & \\
\sum x_{1i}x_{mi} & \sum x_{2i}x_{mi} & \cdots & \sum x_{mi}^2
\end{pmatrix}
\begin{pmatrix} a \\ b \\ \cdots \\ l \end{pmatrix}
=
\begin{pmatrix} \sum x_{1i}y_i \\ \sum x_{2i}y_i \\ \cdots \\ \sum x_{mi}y_i \end{pmatrix} \tag{23}$$

ただし，行列式の計算方法としてよく知られている要素の斜めの掛け合わせの和（Sarrus の式）は 3 行 3 列までしか使えないので，4 行以上もしくは 4 列以上の場合には，行列式の求め方を十分理解して行う必要がある．求め方については別途線形代数の教科書などを参考にすること．

また，実際に最小 2 乗法を適用するとしても，変数が多い場合には特殊な関数形でない限り意味のある値を得るのは難しい場合もあるので，なるべく簡潔な式の関係に帰結できるような場合に適用する方が無難と思われる．