最小二乗法

現場透水試験結果の解析において，最小二乗近似法を用いて解析結果から人為的な誤差を排除する技法がとられることがある。最近では，データの整理に使われる表計算アプリケーション（spreadsheet）に当該機能が組み込まれていることが多く，個々の問題ごとにこの技法を振り返り見る機会が少なくなっている。

ここでは出来る限り汎用的な説明を示し，直線勾配法での適用を説明してみる。

1 次多元連立方程式を考え，データにより関係付けられる関係式（方程式）の式数が，求めようとする未知量（方程式の定数項）より多い場合（$N>M$）を考える。以下の方程式群が得られたとする。

$$
\begin{array}{l}
a_{11}x_1 + a_{12}x_2 + \cdots\cdots + a_{1M}x_M = b_1 \\
a_{21}x_1 + a_{22}x_2 + \cdots\cdots + a2_{1M}x_M = b_2 \\
\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\
\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\
a_{N1}x_1 + a_{12}x_2 + \cdots\cdots + a_{NM}x_M = b_N
\end{array}
\tag{1}
$$

ここで，総量 N：解析の用いる観測データ個数，総量 M：未知量個数，b_i：観測データ，x_j：未知量，a_{ij}：観測データ b_i に関係する条件値である。

これを行列式で表すと

$$[A]\{X\} = \{B\} \tag{2}$$

ここで，

$$[A] = \begin{bmatrix} a_{11} & a_{12} & \cdot & a_{1M} \\ a_{21} & a_{22} & \cdot & a_{2M} \\ \cdot & \cdot & \cdot & \cdot \\ a_{N1} & a_{N2} & \cdot & a_{NM} \end{bmatrix} \tag{3}$$

$$\{X\} = \begin{Bmatrix} x_1 \\ x_2 \\ \cdot \\ x_M \end{Bmatrix} \tag{4}$$

$$\{B\} = \begin{Bmatrix} b_1 \\ b_2 \\ \cdot \\ b_N \end{Bmatrix} \tag{5}$$

(1)式の連立方程式を最小二乗法で解く際には，以下の評価関数 Z を定義する。

$$Z = \sum_{i=1}^{N} \left[\left(a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{iM}x_M \right) - b_i \right]^2 \tag{6}$$

評価関数 Z が最小になるように，$\{X\}$を決定する。ここで，Z を x_i で個々偏微分したものが 0 となるとき，評価関数 Z は最小である。つまり以下の関係式となる。

$$\frac{\partial Z}{\partial x_1} = \sum_{i=1}^{N} 2a_{i1} \left[\left(a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{iM}x_M \right) - b_i \right] = 0 \tag{6.1}$$

$$\frac{\partial Z}{\partial x_2} = \sum_{i=1}^{N} 2a_{i2} \left[\left(a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{iM}x_M \right) - b_i \right] = 0 \tag{6.2}$$

・　・　・

$$\frac{\partial Z}{\partial x_M}=\sum_{i=1}^{N}2a_{iM}\left[\left(a_{i1}x_1+a_{i2}x_2+\cdots+a_{iM}x_M\right)-b_i\right]=0 \tag{6.3}$$

ここで，(6)式は以下のように整理できる。

$$\sum_{i=1}^{N}\left(a_{i1}a_{i1}\right)x_1+\sum_{i=1}^{N}\left(a_{i1}a_{i2}\right)x_2+\cdots+\sum_{i=1}^{N}\left(a_{i1}a_{iM}\right)x_M=\sum_{i=1}^{N}\left(a_{i1}b_i\right) \tag{7.1}$$

$$\sum_{i=1}^{N}\left(a_{i2}a_{i1}\right)x_1+\sum_{i=1}^{N}\left(a_{i2}a_{i2}\right)x_2+\cdots+\sum_{i=1}^{N}\left(a_{i2}a_{iM}\right)x_M=\sum_{i=1}^{N}\left(a_{i2}b_i\right) \tag{7.2}$$

・　・　・

$$\sum_{i=1}^{N}\left(a_{iM}a_{i1}\right)x_1+\sum_{i=1}^{N}\left(a_{iM}a_{i2}\right)x_2+\cdots+\sum_{i=1}^{N}\left(a_{iM}a_{iM}\right)x_M=\sum_{i=1}^{N}\left(a_{iM}b_i\right) \tag{7.3}$$

ここで，(1)(2)式と，(7)式を比較すると，M 個の未知量 x に対して，前者は N 個の方程式があったが，後者は M 個の方程式となっており，(7)式は正規方程式となっている。(7)式を行列式で表すと以下となる。

$$[A]^T[A]\{X\}=[A]^T\{B\} \tag{8}$$

ここで，$[A]^T$：転置行列（行と列の入れ替え）
よって，以下より最適解{X}が得られる。

$$\{X\}=\left([A]^T[A]\right)^{-1}[A]^T\{B\} \tag{9}$$

ここで，$\left([A]^T[A]\right)^{-1}$ ：$\left([A]^T[A]\right)$の逆行列である。