

講 座

$$H = \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 2 & 2 \end{bmatrix} \quad W = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$H^T W H = \begin{bmatrix} 1 & 2 & 2 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 2 & 2 \end{bmatrix} = \begin{bmatrix} 9 & 7 \\ 7 & 6 \end{bmatrix}$$

$$(H^T W H)^{-1} = \frac{1}{5} \begin{bmatrix} 6-7 \\ -7 & 9 \end{bmatrix}$$

したがって、最適推定値は以下のようになる。

$$\begin{aligned} \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} &= (H^T W H)^{-1} H^T W z \\ &= \frac{1}{5} \begin{bmatrix} 6-7 \\ -7 & 9 \end{bmatrix} \begin{bmatrix} 1 & 2 & 2 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 5 \\ 7 \end{bmatrix} \\ &= \frac{1}{5} \begin{bmatrix} 9 \\ 7 \end{bmatrix} = \begin{bmatrix} 1.8 \\ 1.4 \end{bmatrix} \quad \dots \dots \dots \text{(el-4)} \end{aligned}$$

3.2.2 最尤法によるパラメーター推定

(1) 尤度関数の定義

対象としている系の応答の観測値を z (m 次元ベクトル) とし、その真値 z_0 が与えられている場合を考えると、観測誤差 ε (m 次元ベクトル) は次式で与えられる。

$$\varepsilon = z - z_0 \quad \dots \dots \dots \text{(6)}$$

ε の確率密度関数 $p(\varepsilon) = p(z; z_0)'$ が与えられているものとする。

ここで、実際に得られた観測値 z から真値 z_0 を推定することを考える。一つの観測値 z_1 が得られれば、真値 z_0 は z_1 の近傍にあると考えられる。そこで、観測値 z_1 を定数として、真値の推定値 \hat{z} に関する次式のような関数を定義する。

$$L(\hat{z}|z_1) = p(z_1; \hat{z}) \quad \dots \dots \dots \text{(7)}$$

$L(\hat{z}|z_1)$ は観測値 z_1 が与えられたときの推定値 \hat{z} に関する尤度と呼ばれている。 $p(z; z_0)$ は真値が z_0 であるとき、実際に得られた値 z が観測される確率を表しているので、 $L(\hat{z}|z_1)$ の値が大きい場合には観測値 z_1 に対し、真値 z_0 が \hat{z} の近傍にある可能性が高く、逆に $L(\hat{z}|z_1)$ の値が小さければ、真値 z_0 が \hat{z} の近傍にある可能性は低くなる。

さらに観測を繰り返して、 z_1, z_2, \dots, z_k が得られたとすると、これら k 個の観測ベクトルを総合して、真値の推定値（最尤推定値） \hat{z} の尤度関数 $L(\hat{z})$ を求めるためには、各々の観測ベクトルに対して定義された k 個の尤度関数の積をとればよい。

$$L(\hat{z}) = \prod_{i=1}^k L(\hat{z}|z_i) = \prod_{i=1}^k p(z_i; \hat{z}) \quad \dots \dots \dots \text{(8)}$$

尤度関数を概念的に理解するために、観測値がスカラーワークで与えられる場合を考えてみる。今、観測値が1個与えられたとし、それを z_1 とすれば、図-1に示すように尤度関数は確率分布関数の左右を、式(7)では z_1 が定数で \hat{z} が変数となるので逆転させ、中心を z_1 に置いたものになっている。観測値の数が多くなるにつれて、それらを総合して求めた尤度関数は、最尤推定値 \hat{z} が真値 z_0 に近づき、その分布幅も観測値の確率分布関数 P

$(z; z_0)$ の分布幅よりしだいに狭くなっていくことがわかる。

以下に、観測誤差ベクトル ε が多次元正規分布で表される場合について尤度関数を求めておく。この場合には観測ベクトル z が得られる確率は

$$p(z; z_0) = \frac{1}{(2\pi)^{\frac{m}{2}} |\mathbf{R}_1|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}(z - z_0)^T \mathbf{R}_1^{-1}(z - z_0)\right\}$$

と表現できる。ここに、 \mathbf{R}_1 は観測値に対する共分散行列である。したがって、観測ベクトル z_1 が与えられたとすれば、尤度関数は次式となる。

$$\begin{aligned} L(\hat{z}|z_1) &= p(z_1; \hat{z}) \\ &= \frac{1}{(2\pi)^{\frac{m}{2}} |\mathbf{R}_1|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}(z_1 - \hat{z})^T \mathbf{R}_1^{-1}(z_1 - \hat{z})\right\} \dots \dots \dots \text{(9)} \end{aligned}$$

ここに、 \mathbf{R}_1 は観測ベクトル z_1 に対する共分散行列である。

各観測ベクトル z_i について与えられる式(9)と同様な関数を式(8)に代入すれば、 k 個の観測ベクトルが与えられたときの尤度関数として次式を得る。

$$\begin{aligned} L(\hat{z}) &= \frac{1}{(2\pi)^{\frac{m}{2}} |\mathbf{R}_1||\mathbf{R}_2| \cdots |\mathbf{R}_k|} \\ &\exp\left\{-\frac{1}{2} \sum_{i=1}^k (z_i - \hat{z})^T \mathbf{R}_i^{-1}(z_i - \hat{z})\right\} \dots \dots \dots \text{(10)} \end{aligned}$$

尤度関数を最大にする \hat{z} が最尤推定値となるが、観測誤差が正規分布に従うときには、式(10)の \exp の中の値を最小にすればよい。すなわち、次式の評価関数を最小にするように \hat{z} を定めればよい。

$$J_1 = \frac{1}{2} \sum_{i=1}^k (z_i - \hat{z})^T \mathbf{R}_i^{-1}(z_i - \hat{z}) \dots \dots \dots \text{(11)}$$

すなわち、

$$dJ_1 = -d\hat{z} \sum_{i=1}^k \mathbf{R}_i^{-1}(z_i - \hat{z}) = 0 \dots \dots \dots \text{(12)}$$

これから \hat{z} が次式で与えられる。

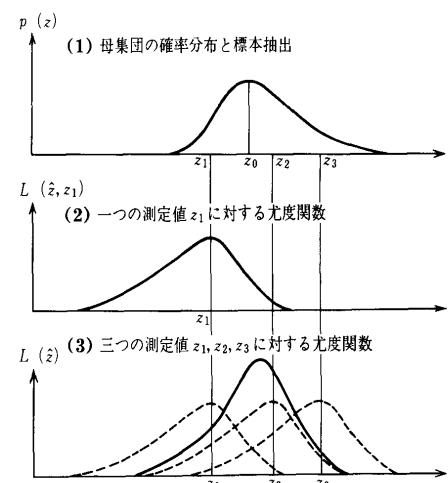
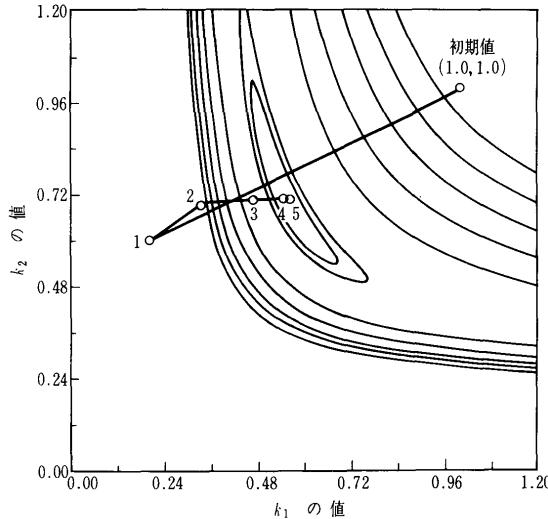


図-3.1 誤差の確率密度関数と尤度関数

- (1) 母集団の確率分布と標本抽出
- (2) 一つの測定値 z_1 に対する尤度関数
- (3) 三つの測定値 z_1, z_2, z_3 に対する尤度関数

講 座



図一例2 観測方程式が非線形の場合の評価関数 J_1 の等高線と収束過程 (ガウスニュートン法による)

$$L(\hat{z}) = \prod_{i=1}^k p(z_i; \hat{z}) = \prod_{i=1}^k p(z_i; \mathbf{h}(\hat{x})) = L(\hat{x}) \quad \dots \dots \dots \quad (25)$$

式(25)の対数をとった対数尤度関数

$$\log L(\hat{x}) = \sum_{i=1}^k \log p(z_i; h_i(\hat{x})) \quad \dots \dots \dots \quad (26)$$

を用いて、この最大値を与えるような \hat{x} を求めればよいことになる。

例題2

例題1では $x_1 = \frac{1}{k_1}$, $x_2 = \frac{1}{k_2}$ と置いたので、観測方程式が線形となつたが、

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} k_1 \\ k_2 \end{pmatrix}$$

とすると、観測方程式が非線形になるので、式(24)に基づいた繰返し計算が必要になる。

例題1の式(e1-2)の第2式を書き直すと、

$$u_2 = (f_1 + f_2) \frac{1}{x_1} + f_2 \frac{1}{x_2}$$

となるので、例題1の式(e1-1)の観測値を上式に代入すれば、観測方程式として次式を得る。

$$\begin{pmatrix} 2 \\ 5 \\ 7 \end{pmatrix} = \begin{pmatrix} \frac{1}{x_1} + \frac{1}{x_2} \\ \frac{2}{x_1} + \frac{1}{x_2} \\ \frac{2}{x_1} + \frac{2}{x_2} \end{pmatrix} \quad \dots \dots \dots \quad (e2-1)$$

式(e2-1)の右辺は式(19)の $\mathbf{h}(\mathbf{x})$ に相当するから次式を得る。

$$\mathbf{h}(\mathbf{x}) = \begin{pmatrix} h_1(x_1, x_2) \\ h_2(x_1, x_2) \\ h_3(x_1, x_2) \end{pmatrix} = \begin{pmatrix} \frac{1}{x_1} + \frac{1}{x_2} \\ \frac{2}{x_1} + \frac{1}{x_2} \\ \frac{2}{x_1} + \frac{2}{x_2} \end{pmatrix} \quad \dots \dots \dots \quad (e2-2)$$

式(24)に基づいた繰返し計算のためには、 \mathbf{H}_k と $\mathbf{h}(\mathbf{x}_k)$

の値が必要となるが、それらは k 回目の繰返し段階での \mathbf{x} の値を次式のように表現すれば、

$$\mathbf{x}_k = \begin{pmatrix} x_{1,k} \\ x_{2,k} \end{pmatrix}$$

以下のように与えられる。

$$\mathbf{H}_k = \begin{bmatrix} \frac{\partial h_1}{\partial x_1} & \frac{\partial h_1}{\partial x_2} \\ \frac{\partial h_2}{\partial x_1} & \frac{\partial h_2}{\partial x_2} \\ \frac{\partial h_3}{\partial x_1} & \frac{\partial h_3}{\partial x_2} \end{bmatrix}_{\mathbf{x}=\mathbf{x}_k} = \begin{bmatrix} -\frac{1}{x_{1,k}^2} & -\frac{1}{x_{2,k}^2} \\ -\frac{2}{x_{1,k}^2} & -\frac{1}{x_{2,k}^2} \\ -\frac{2}{x_{1,k}^2} & -\frac{2}{x_{2,k}^2} \end{bmatrix} \quad \dots \dots \dots \quad (e2-3)$$

$$\mathbf{h}(\mathbf{x}_k) = \begin{pmatrix} \frac{1}{x_{1,k}} + \frac{1}{x_{2,k}} \\ \frac{2}{x_{1,k}} + \frac{1}{x_{2,k}} \\ \frac{2}{x_{1,k}} + \frac{2}{x_{2,k}} \end{pmatrix} \quad \dots \dots \dots \quad (e2-4)$$

図一例2に \mathbf{x} の初期値を $x_1 = 1.0$ と $x_2 = 1.0$ としたときの収束過程を示した。図中の○印に添えられた数字は繰返し回数を意味している。また、図中の曲線は式(20)に定義された評価関数の等高線を描いたものである。5回の繰返しで評価関数の極小点へ到達していることが分かる。このときの \mathbf{x} の値は

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0.5555556 \\ 0.7142856 \end{pmatrix}$$

であり、例題1の答の逆数になっていることが分かる。例題1では簡単な線形代数の計算で問題が解けたが、同じ問題でもパラメーターの設定の仕方によって問題が非線形になり、かなり複雑な計算を行わないと求めようとする解が得られないばかりでなく、時には解の収束しないような事態が発生する。逆解析では観測方程式を書き下すときに、変数変換などを行って、できるだけ線形問題に帰着できるような努力をすることが、問題の見通しをよくする一つの方法である。なお、非線形問題では \mathbf{H}_k を解析的に求めることが困難な場合が多いので、数值微分によって計算するのが一般的である。

参考文献

- 1) 中川 徹・小柳義夫：最小二乗法による実験データ解析、東京大学出版会、1986.
- 2) Ang A. H-S. and Tang W. H. : 土木・建築のための確率統計の基礎 (伊藤・亀田訳), 丸善, 1977.
- 3) Jazwinski A. H. : Stochastic process and filtering theory, Academic Press Inc., 1970.
- 4) 加藤寛一郎：最適制御入門、東京大学出版会, 1987.
- 5) 有本 卓：カルマンフィルター、産業図書、1977.
- 6) 片山 徹：応用カルマンフィルタ、朝倉書店、1983.
- 7) 吉田郁政・黒瀬浩公・福井史朗：確率論に基づく逆解析手法の基礎的研究、土木学会論文集, No. 483/I-26, pp. 61~68, 1994.

(以下、次号に続く)