

【数物系科学】

【工学】

研究論文

空間補正機能を付与した線形最小二乗法による 非線形カーブフィッティング

金丸 英幸*¹

Non-linear Curve Fitting by Linear Least-squares Method with Space Compensation Ability

KANAMARU Hideyuki

Summary

Recently, new technologies such as Industry4.0 and/or Internet of Thing (IoT) are being studied as part of rapid advances in information and communication technology (ICT), where the importance of real-time processing is anticipated to increase.

The least-squares method (LSM) is one of the most popular ways to determine the characteristics of measured data. Iteration calculation is frequently used to obtain a high preciseness on the batch processing in LSM applied to non-linear curve fitting, but adaptive control equipment and other real-time computer systems need to reduce the steps of data processing for time-sharing.

In this paper, a new linear LSM with space compensation ability is presented to balance high preciseness and reduction of the amount of data processing even for non-linear curve fitting, and it is evaluated in comparison with results from linear LSM without space compensation ability and non-linear LSM with iteration.

Keyword : (LSM , space compensation , non-linear curve fitting , real-time processing)

1. はじめに

情報通信技術の飛躍的發展によって情報化社会がますます拡大しており、産業界においてはIoT (Internet of Thing) や第4次産業革命と言われる Industry4.0 への挑戦が続けられている。この分野ではデータ計測やその特性把握、データ解析に基づく制御等が時分割で行われるリアルタイム性が重要となり、処理の効率化が要求される。

各種計測データの特性解析によく使われている線形

最小二乗法を用いて非線形関数を当て嵌める場合には、前処理として独立変数と従属変数を各々非線形変換する方法がしばしば用いられる。しかしながら、この方法では求められたパラメータの精度が十分でなく、標本データの点列と当て嵌め曲線(多次元の場合は曲面)が乖離する場合もあることが知られている。

独立変数と従属変数を非線形変換すると各座標軸が伸縮されるが、その非一様性のために現空間上の直線(多次元の場合は平面)が写像空間上では曲線(多次元

*¹ 新技術創成研究所 客員研究員 産官学連携コーディネータ
2016年3月14日受付
2016年5月12日受理

の場合は曲面) となって写像空間に歪みが生じる。この歪んだ空間上で直線(多次元の場合は平面)を当て嵌めることが、原空間上で標本データと当て嵌め曲線(多次元の場合は曲面)の間に乖離が生じる原因と考えられる。

一方、Matlab^(注)等の技術計算ソフトウェアでは、反復計算によって当て嵌め曲線(2次元の場合は曲面)を収束させるバッチ処理が基本となっているため、精度は高いが以下の問題を抱えている。

- (1) 適応制御システム等、制御対象の特性解析・把握と、その結果に即応した制御系の自動調整がリアルタイムで行われるシステムでは、時分割処理のために計算時間(計算量)に制約がある。したがって、設計段階で演算量が予め計算でき、かつ収束が保証されることが重要で、それらが標本データに依存する反復計算法は適用し難い。
- (2) リアルタイム系のシステムでは、反復計算で解を収束させるための初期条件や各種制約条件を、人が介在して試行錯誤的に変更・調整することは許されない。
- (3) 一般に、Matlab 等の技術計算ソフトウェアは規模が大きく高価で、処理時間的にも記憶容量的にも効率化が要求されるリアルタイム系の組込み型システムへの実装は実用的でない。

本稿では、従来の最小二乗法を用いて非線形関数を当て嵌めるために荷重係数付最小二乗法を適用することを試みる。従来、荷重係数付最小二乗法では、標本データの当て嵌め精度を「部分的に」向上させることを目的(他部分は相対的に劣化)として、荷重の値が恣意的に決められ初期条件として与えられていた。ここでは、原空間と写像空間の間の歪みの原因である各座標軸の伸縮の度合いに応じて、標本点毎に当て嵌め誤差の評価の重みを調整する手法を用いる。特に、標本データから、各標本点の荷重係数を一意に決定する方法を示して、全体的な当て嵌め精度の改善と恣意性の排除を目指す。尚、空間補正を行わない関数の当て嵌めは、荷重係数を一様に「1」と置くことにより本手法に包含される。

本手法による非線形関数の当て嵌めは、反復計算による関数の当て嵌めに比べて対象とする関数が限定され

るが、解を収束させるための反復計算を伴わないため、リアルタイム処理に適しているという利点を持つ。

最後に、数値実験によって、空間補正の有無による当て嵌め精度の違いやMatlabによる解析結果との精度差を例示し、本手法の関数当て嵌め能力を評価する。

2. 荷重係数付最小二乗法

2.1 正規方程式

$(m+1)$ 次元の $x^k - y$ 空間 ($k = 1 \cdots m$) において標本データ $(x_i^k, y : i = 1 \cdots n)$ の分布に対して、関数

$$y = f(x^k; a^k, b | k = 1 \cdots m) \quad (1)$$

を当て嵌める時、

$$X^k = g_{x^k}(x^k) \quad (2)$$

$$Y = g_y(y) \quad (3)$$

$$A^k = h_{a^k}(a^l, b) \quad (l = 1 \cdots m) \quad (4)$$

$$B = h_b(a^l, b) \quad (l = 1 \cdots m) \quad (5)$$

で定義される変数とパラメータの写像を行い、その結果、 X^k と Y の間に平面を表す関係

$$Y = \sum_k A^k X^k + B \quad (6)$$

が成り立つと仮定する。ここに、添字「 k 」は独立変数の軸を表す。

(6)式は、(2)、(3)式による座標変換を受けた $(m+1)$ 次元空間上の平面を示す。一方、(2)、(3)式は $x^k - y$ 座標系で表現される原空間を各軸方向に歪めて $X^k - Y$ 座標系を構築していることを意味する。したがって、 $X^k - Y$ 空間上でパラメータ A^k, B を決定するために従来の線形最小二乗法を適用した場合、 A^k, B から a^k, b が求まったとしても原空間である $x^k - y$ 座標系において、(1)式が合理的に当て嵌まるとは限らない。

ここでは、 $x^k - y$ 空間、したがって $X^k - Y$ 空間 ($k = 1 \cdots m$) 上で n 組の標本データ

(注) Matlab は MathWorks 社の登録商標です

$$\{(X_1^k, Y_1), (X_2^k, Y_2), \dots, (X_n^k, Y_n)\}$$

に対して、荷重係数 p_i を付与した誤差関数

$$\varepsilon = \sum_i p_i \{Y_i - (\sum_k A^k X_i^k + B)\}^2 = \sum_i p_i \Delta Y_i^2 \quad (7)$$

を最小にする A^k と B を求める荷重係数付最小二乗法の適用について考察し、両空間の歪みを補正する係数として荷重係数を定式化することを試みる。

誤差関数(7)式最小化の必要条件

$$\frac{\partial \varepsilon}{\partial A^k} = 0 \quad (8)$$

$$\frac{\partial \varepsilon}{\partial B} = 0 \quad (9)$$

より、以下の A^k と B を変数とする正規方程式を得る。

$$\sum_{j=1}^m \{(\sum_i p_i X_i^j X_i^k) A^j\} + (\sum_i p_i X_i^k) B = \sum_i p_i X_i^k Y_i \quad (10)$$

$$\sum_{j=1}^m \{(\sum_i p_i X_i^j) A^j\} + (\sum_i p_i) B = \sum_i p_i Y_i \quad (11)$$

($m + 1$)元連立1次方程式(10)、(11)を A^k, B について解き、(4)、(5)式を連立させることによって、原空間上の曲線(1)のパラメータ a^k, b が決定される。

2.2 荷重係数の決定

前節で詳述したパラメータ同定は、原空間 ($x^k - y$ 空間) に対して x^k 軸($k = 1 \dots m$)、 y 軸の両方を(2),(3)式で示される関数で写像したことによる歪んだ空間 ($X^k - Y$ 空間) 上で行なわれる。したがって、写像空間上で原空間のパラメータ同定を正しく行なうためには、 X^k, Y 各軸に対して適切に空間の変形を補正しなければならない。

原データ(x_i^k, y_i)に対して、原空間 ($x^k - y$ 空間) 上で最小二乗法を用いて(1)式の関数を当て嵌める場合、誤差関数 ε は、次式で定義される。

$$\varepsilon = \sum_i \{\Delta y_i(x_i^k; a^k, b)\}^2 \quad (12)$$

$$\Delta y_i(x_i^k; a^k, b) = y_i - f(x_i^k; a^k, b) \quad | k = 1 \dots m \quad (13)$$

y_i したがって Δy_i が x_i^k とは独立な関係にあることを考慮すると、この誤差関数は δ 関数を用いて以下の積分形式で表すことができる。

$$\begin{aligned} \varepsilon &= \sum_i \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \delta(x^1 - x_i^1) \dots \delta(x^m - x_i^m) \\ &\quad \cdot \{\Delta y_i(x_i^k; a^k, b)\}^2 dx^1 \dots dx^m \\ &= \sum_i [\prod_k \int_{-\infty}^{\infty} \{\delta(x^k - x_i^k) dx^k\}] \Delta y_i^2 \quad (14) \end{aligned}$$

ここで、 δ 関数の定義式として以下の表現を用いる。

$$\delta(x^k - x_i^k) = \lim_{\Delta x^k \rightarrow 0} \begin{cases} 0 & \left(x^k < x_i^k - \frac{\Delta x^k}{2}\right) \\ \frac{1}{\Delta x^k} & \left(x_i^k - \frac{\Delta x^k}{2} \leq x^k \leq x_i^k + \frac{\Delta x^k}{2}\right) \\ 0 & \left(x_i^k + \frac{\Delta x^k}{2} < x^k\right) \end{cases} \quad (15)$$

(14)式に、(15)式を適用することにより、誤差関数は次式で与えられる。

$$\varepsilon = \sum_i \left[\prod_k \left\{ \lim_{\Delta x^k \rightarrow 0} \left(\frac{1}{\Delta x^k} \int_{x_i^k - \Delta x^k/2}^{x_i^k + \Delta x^k/2} dx^k \right) \right\} \Delta y_i^2 \right] \quad (16)$$

(2)、(3)式の逆関数

$$x^k = g_{x^k}^{-1}(X^k) \quad (17)$$

$$y = g_y^{-1}(Y) \quad (18)$$

を用いると

$$dx^k = \left(\frac{dg_{x^k}^{-1}}{dX^k} \right) dX^k \quad (19)$$

$$\Delta y = \left(\frac{dg_y^{-1}}{dY} \right) \Delta Y \quad (20)$$

したがって、(16)式は、

$$\varepsilon = \sum_i \left[\prod_k \lim_{\Delta x^k \rightarrow 0} \left\{ \frac{1}{\Delta x^k} \int_{g_{x^k}(x_i^k - \Delta x^k/2)}^{g_{x^k}(x_i^k + \Delta x^k/2)} \left(\frac{dg_{x^k}^{-1}}{dX^k} \right) dX^k \right\} \right. \\ \left. \cdot \left\{ \left(\frac{dg_y^{-1}}{dY} \right)_i \Delta Y_i \right\}^2 \right] \quad p_i = \left(\frac{dg_y^{-1}}{dY} \right)_i^2 \quad (28)$$

即ち、荷重係数 p_i は x^k 軸の変換 $g_{x^k}(x^k)$ には依存しない。 δ 関数の性質を考慮すると、この結果は(14)式から予想することもできる。

以降、(28)式の p_i を「空間補正係数」と称する。

さらに、

$$\lim_{\Delta x^k \rightarrow 0} \left(\frac{dg_{x^k}^{-1}}{dX^k} \right) = \left(\frac{dg_{x^k}^{-1}}{dX^k} \right)_i \quad (22)$$

$$dX^k = \lim_{\Delta x^k \rightarrow 0} \{ g_{x^k}(x^k + \Delta x^k/2) - g_{x^k}(x^k - \Delta x^k/2) \} \quad (23)$$

を用いて(21)式を加算形式に戻すと、

$$\varepsilon = \sum_i \left[\prod_k \lim_{\Delta x^k \rightarrow 0} \left\{ \frac{g_{x^k}(x_i^k + \Delta x^k/2) - g_{x^k}(x_i^k - \Delta x^k/2)}{\Delta x^k} \left(\frac{dg_{x^k}^{-1}}{dX^k} \right)_i \right\} \right. \\ \left. \cdot \left(\frac{dg_y^{-1}}{dY} \right)_i^2 \Delta Y_i^2 \right] \quad (24)$$

となり、原空間 ($x^k - y$ 空間) の誤差関数が写像空間 ($X^k - Y$ 空間) 上の誤差関数として表現される。任意の k に対して

$$\lim_{\Delta x^k \rightarrow 0} \frac{g_{x^k}(x_i^k + \Delta x^k/2) - g_{x^k}(x_i^k - \Delta x^k/2)}{\Delta x^k} = \left(\frac{dg_{x^k}}{dx^k} \right)_i \quad (25)$$

$$\left(\frac{dg_{x^k}}{dx^k} \right)_i \left(\frac{dg_{x^k}^{-1}}{dX^k} \right)_i = \left(\frac{dX^k}{dx^k} \right)_i \left(\frac{dx^k}{dX^k} \right)_i = 1 \quad (26)$$

が成り立つことを考慮し、(25)、(26)式を(24)式に適用すると、

$$\varepsilon = \sum_i \left(\frac{dg_y^{-1}}{dY} \right)_i^2 \Delta Y_i^2 \quad (27)$$

したがって、(7)式と(27)式より写像空間上の補正係数 p_i は次式で与えられる。

3. 2次元空間への適用

3.1 正規方程式

本節からは、より応用の機会が多い2次元の $x - y$ 空間 ($k = 1$) に限定して議論を進めることとし、次元を表現する右肩の添字「 k 」の表記を省略する。即ち、原データ (x_i, y_i) の分布が直線的でなく、関数

$$y = f(x; a, b) \quad (29)$$

を当て嵌める時、

$$X = g_x(x) \quad (30)$$

$$Y = g_y(y) \quad (31)$$

$$A = h_a(a, b) \quad (32)$$

$$B = h_b(a, b) \quad (33)$$

で定義される変数とパラメータの変換を行い、その結果、 X と Y の間にリニアな関係

$$Y = AX + B \quad (34)$$

が成り立つと仮定する。この時、原空間上の誤差関数(12)は、空間補正係数 p_i を用いて

$$\varepsilon = \sum_i p_i \{ Y_i - (AX_i + B) \}^2 = \sum_i p_i \Delta Y_i^2 \quad (35)$$

と表され、正規方程式(10)および(11)は、

$$X^2 A + X B = Z \quad (36)$$

$$X A + P B = Y^* \quad (37)$$

となる。ここに

$$X^* = \sum_i p_i X_i \quad (38)$$

$$Y^* = \sum_i p_i Y_i \quad (39)$$

$$X^{2*} = \sum_i p_i X_i^2 \quad (40)$$

$$Z^* = \sum_i p_i X_i Y_i \quad (41)$$

$$P^* = \sum_i p_i \quad (42)$$

したがって、(36), (37)式より、係数 A, B が、

$$A = \frac{P^* Z^* - X^* Y^*}{P^* X^{2*} - X^{*2}} \quad (43)$$

$$B = \frac{X^{2*} Y^* - X^* Z^*}{P^* X^{2*} - X^{*2}} \quad (44)$$

で求められ、原関数(29)のパラメータ a, b は(32), (33)式を連立して解くことにより決定される。

また、空間補正係数 p_i は x 軸の変換に依存しないため(28)式で与えられる。

$$p_i = \left(\frac{dg_y^{-1}}{dy} \right)_i^2 \quad (28)$$

以降、このケースを基本型と称する。

3.2 正規分布の確率密度関数

正規分布の確率密度関数

$$y = f(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (45)$$

に対しては、以下の手順を踏む。

3. 1 節では、原空間の 2 通りのパラメータ a, b の求解について論じたが、ここでは原データ (x_i, y_i) を用いて数値計算により期待値 μ を事前に計算する。

$$\mu = \oint_{-\infty}^{\infty} x f(x; \mu, \sigma) dx \quad (46)$$

その上で最小二乗法により、残されたパラメータである標準偏差 σ を求める。解法の詳細については次節の適用例の中で述べる。尚、 $\oint dx$ は数値計算であることを示す。

因みに、本ケースも基本型に属する。

3.3 適用例

- (1) 冪関数
冪関数

$$y = \left(\frac{x}{a} \right)^b \quad (47)$$

に対しては、以下の座標変換、空間補正係数の求解とパラメータ変換が行われる。

- 座標変換

$$X = g_x(x) = \ln x$$

$$Y = g_y(y) = \ln y$$

- パラメータ変換

$$A = h_a(a, b) = b$$

$$B = h_b(a, b) = -b \ln a$$

- 空間補正係数

$$p = e^{2Y}$$

- パラメータ逆変換

$$a = e^{\frac{A}{B}}$$

$$b = A$$

- (2) 正規分布 (確率密度関数)

正規分布の確率密度関数

$$y = f(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (45)$$

に対しては、以下の座標変換、空間補正係数の求解、およびパラメータ変換が行われる。

- 座標変換

$$X = g_x(x) = (x - \mu)^2$$

$$Y = g_y(y) = \ln y$$

・パラメータ変換

$$\sigma = \sigma_b = \frac{e^{-B}}{\sqrt{2\pi}}$$

$$A = h_a(\sigma) = -\frac{1}{2\sigma^2}$$

$$B = h_b(\sigma) = -\ln(\sqrt{2\pi}\sigma)$$

σ_a は、原データ (x_i, y_i) によっては虚数となる可能性があるので破棄する。したがって

・空間補正係数

$$\sigma = \sigma_b = \frac{e^{-B}}{\sqrt{2\pi}}$$

$$p = e^{2Y}$$

代表的な基本型関数に対して $x-y$ 空間を変換する写像関数と空間補正係数 p_i およびパラメータ a, b の計算方法を Table 1 に示す。

・パラメータ逆変換

$$\sigma = \sigma_a = \frac{1}{\sqrt{-2A}} \quad (\text{破棄})$$

Table 1 空間補正機能を備えた最小二乗法によるパラメータ求解方法 (基本型)

番号	当て嵌め関数		座標変換	荷重係数	パラメータ求解	備考
	関数名	関数形				
(0)	基本形 (直線形式に変換可能な関数)	$y = f(x; a, b)$ ↓ 変換可能 $Y = AX + B$	$A = h_a(a)$ $B = h_b(b)$ $X = g_x(x)$ $Y = g_y(y)$	$p = \left(\frac{\partial g_y^{-1}}{\partial Y}\right)^2$	連立方程式 $A = h_a(a)$ $B = h_b(b)$	一般形
(1)	直線	$y = ax + b$	$X = x$ $Y = y$	$p = 1$	$a = A$ $b = B$	直線回帰
(2)	冪関数 (単一項)	$y = \left(\frac{x}{a}\right)^b$	$X = \ln x$ $Y = \ln y$	$p = e^{2Y}$	$a = e^{-B/A}$ $b = A$	
(3)	指数関数 (一次式)	$y = ae^{bx}$	$X = e^x$ $Y = y$	$p = 1$	$a = A$ $b = B$	
(4)	指数関数 (冪型)	$y = ae^{ax}$	$X = x$ $Y = \ln y$	$p = e^{2Y}$	$a = e^B$ $b = A$	
(5)	対数関数 (一次式)	$y = a \ln x + b$	$X = \ln x$ $Y = y$	$p = 1$	$a = A$ $b = B$	
(6)	対数関数 (冪型)	$y = a(\ln x)^b$	$X = \ln(\ln x)$ $Y = \ln y$	$p = e^{2Y}$	$a = e^B$ $b = A$	
(7)	ワイブル分布 (累積)	$y = 1 - e^{-\left(\frac{x}{a}\right)^b}$	$X = \ln x$ $Y = \ln\{-\ln(1-y)\}$	$p = e^{2(Y-e^Y)}$	$a = e^{-B/A}$ $b = A$	
(8)	指数型成長曲線 (累積)	$y = \frac{1}{1 + be^{-at}}$	$X = t$ $Y = \ln\left(\frac{1-y}{y}\right)$	$p = \frac{1}{(e^{Y/2} + e^{-Y/2})^4}$	$a = -A$ $b = e^B$	
(9)	冪型成長曲線 (累積)	$y = \frac{1}{1 + bt^{-a}}$	$X = \ln t$ $Y = \ln\left(\frac{1-y}{y}\right)$	$p = \frac{1}{(e^{Y/2} + e^{-Y/2})^4}$	$a = -A$ $b = e^B$	
(10)	正規分布 (密度)	$y = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	$X = (x-\mu)^2$ $Y = \ln y$	$p = e^{2Y}$	$(\sigma = \frac{1}{\sqrt{-2A}}$ $\sigma = \frac{e^{-B}}{\sqrt{2\pi}}$	(注) 数値積分に台形則を使用した場合
			y の積分値: $Y_{\infty} = \sum \frac{y_i + y_{i-1}}{2} (x_i - x_{i-1})$ 期待値 μ : $\mu = \left\{ \frac{1}{4} \sum (x_i^2 - x_{i-1}^2)(y_i + y_{i-1}) \right\} / Y_{\infty}$			
(11)	Lorentz型関数	$y = \frac{a}{1 + b(x-x_0)^2}$	$X = (x-x_0)^2$ $Y = \frac{1}{y}$	$p = \frac{1}{Y^4}$	$a = \frac{1}{B}$ $b = \frac{A}{B}$	

4. 微分型関数の当て嵌め

4.1 2次元空間における応用

供試関数の積分関数に対して、 x 軸と y 軸の写像の結果(34)式の条件が成り立つ場合には、供試関数そのものではなく、その積分関数

$$F(x; a, b) = \int f(x; a, b) dx$$

に2節および3節で論じた解法を適用することができる。但し、積分区間外の積分値を補うために以下の処置を施す。

- (1) 区間 $[x_1, x_n]$ の積分値 z_i を計算する(点列に基づく数値積分)。

$$z_i = \oint_{x_1}^{x_i} f(x; a, b) dx$$

$$c_0 = z_n$$

- (2) 以下の場合に対応して、積分値 z_i を加工処理する。

- (a) 一般の微分型関数の場合

$$\hat{z}_i = z_i$$

- (b) 確率密度関数や成長率曲線など、積分関数が単調増加で、 $x \rightarrow \infty$ で有限値に収束する場合

$$\hat{z}_i = \frac{z_i}{c_0} \quad (i = 1 \cdots n)$$

- (3) (x_i, \hat{z}_i) に対して、3節の手順を用いて $F(x)$ を当て嵌め、パラメータ \hat{a} 、 \hat{b} を求める。

- (4) パラメータ \hat{a} 、 \hat{b} を用いて積分区間外の積分値を計算し、区間 $[x_1, x_n]$ の積分値 c_0 に加えて全区間の積分値を近似計算する。

- (a) $c = c_0 + F(x_1; \hat{a}, \hat{b})$
 (b) $c = c_0 + F(x_1; \hat{a}, \hat{b}) + \{1 - F(x_n; \hat{a}, \hat{b})\}$

- (5) 積分値 z_i を $F(x_1; \hat{a}, \hat{b})$ で補正し、積分値 \hat{z}_i を加工処理する。

$$(a) \hat{z}_i = z_i + F(x_1; \hat{a}, \hat{b}) \quad (i = 1 \cdots n)$$

$$(b) \hat{z}_i = \frac{z_i + F(x_1; \hat{a}, \hat{b})}{c} \quad (i = 1 \cdots n)$$

- (6) (x_i, \hat{z}_i) に対し、再び3節の手順を用いて $F(x)$ を当て嵌め、原空間のパラメータ a 、 b を決定する

- (7) 最後に、当て嵌め関数は次式で与えられる。

$$(a) y = f(x; a, b)$$

$$(b) y = c f(x; a, b)$$

この場合は、原関数ではなくその積分関数に最小二乗法を適用しているため、原空間における誤差 ϵ の最小化は保証されないことに注意を要する。また、上記の解法により、微分型関数への当て嵌めの場合には反復計算は1回のみ抑えられる。

以降、このケースを積分型と称する。

4.2 適用例

- (1) 微分型対数関数

微分型対数関数

$$y = \frac{ab(\ln x)^{b-1}}{x} \tag{48}$$

の場合は、その積分関数

$$y = F(x; a, b) = a(\ln x)^b$$

に対して、以下の座標変換、空間補正係数の計算およびパラメータ変換が行われる。

・座標変換

$$X = g_x(x) = \ln(\ln x)$$

$$Y = g_y(y) = \ln y$$

・パラメータ変換

$$A = h_a(a, b) = b$$

$$B = h_b(a, b) = \ln a$$

・空間補正係数

$$p = e^{2Y}$$

・パラメータ逆変換

$$a = e^B$$

$$b = A$$

およびパラメータ変換が行われる。

・座標変換

$$X = g_x(x) = \ln x$$

$$Y = g_y(y) = \ln\left(\frac{1-y}{y}\right)$$

・パラメータ変換

$$A = h_a(a, b) = -a$$

$$B = h_b(a, b) = \ln b$$

・空間補正係数

(2) 成長率曲線型 (冪型) 関数
成長率曲線型 (冪型) 関数

$$y = c \frac{abx^{-(a+1)}}{(1+bx^{-a})^2} \quad (a, b, c > 0) \quad (49)$$

の場合は、その積分関数

$$y = \frac{c}{1+bx^{-a}} \quad (a, b, c > 0)$$

$$p = \frac{1}{(e^{Y/2} + e^{-Y/2})^4}$$

・パラメータ逆変換

$$a = -A$$

$$b = e^B$$

代表的な微分型関数に対して $x-y$ 空間を変換する写像関数と空間補正係数 p_i およびパラメータ a, b の計算方法を Table 2 に示す。

に対して、以下の座標変換、空間補正係数の計算

Table 2 空間補正機能を備えた最小二乗法によるパラメータ求解方法 (積分型)

番号	当て嵌め関数		座標変換	荷重係数	パラメータ求解	備考
	関数名	関数形				
(0)	基本形 (直線形式に変換可能な関数)	$y = f(x; a, b)$ ↓ 変換可能 $Y = AX + B$	$A = h_a(a)$ $B = h_b(b)$ $X = g_x(x)$ $Y = g_y(y)$	$p = \left(\frac{\partial g_y^{-1}}{\partial Y}\right)^2$	連立方程式 $A = h_a(a)$ $B = h_b(b)$	一般形
(1)	微分型対数関数	$y = \frac{ab(\ln x)^{b-1}}{x}$	点列を数値積分して基本型関数(6)を適用。		$a = e^B$ $b = A$	4.1節
(2)	ワイブル分布型 (密度)	$y = c \left\{ \frac{b}{a} \left(\frac{x}{a}\right)^{b-1} e^{-\left(\frac{x}{a}\right)^b} \right\}$	点列を数値積分して基本型関数(7)を適用。		$a = e^{-B/A}$ $b = A$	4.1節 c : 積分値
(3)	成長率曲線型 (指数型)	$y = c \frac{abe^{-at}}{(1+be^{-at})^2}$	点列を数値積分して基本型関数(8)を適用。		$a = -A$ $b = e^B$	4.1節 c : 積分値
(4)	成長率曲線型 (冪型)	$y = c \frac{abt^{-a-1}}{(1+bt^{-a})^2}$	点列を数値積分して基本型関数(8)を適用。		$a = -A$ $b = e^B$	4.1節 c : 積分値

5. 直線の式への変形ができない場合への応用

5.1 2次元空間における応用

直線の形に変換することが難しく、線形最小二乗法をそのまま適用することができない場合には、3節の手順を基本として反復計算による収束を図る。一例として、原関数が

$$y = f(x; a, b, c) \tag{50}$$

と表され、パラメータ c が既知の場合には、(50)式を直線の式に変換する関数(30)および(31)が存在すると仮定する。

この場合は、 c と a, b を交互に定数扱いとし、1サイクルの中で空間補正機能を付与した最小二乗法の計算を2度適用することによって、パラメータ a, b, c を(準)最適値に収束させることができる。即ち

- (a) まず、パラメータ c に初期値 c_j ($j = 0$) を与える。
 - (b) c_j を定数と考え、3.1節の手順を用いて(43)、(44)式と(32)、(33)式より、 a と b の暫定値 a_j, b_j および(35)式の誤差 ϵ_j を求める。
 - (c) 次に、 ϵ_j の変化率が微小な既定値 δ より大きければ、パラメータ a_j, b_j を定数として c_j を計算し、その結果を用いて ϵ_j を再計算する。
 - (d) 上記(b)、(c)の計算を繰り返し、(b)で ϵ_j の変化率が既定値 δ 未満となった時点で反復計算を終了して a, b, c の推定値を確定する。
- 以降、このケースを反復型と称する。

5.2 適用例

- (1) 定数項付き対数関数 (冪型)
- 定数項付き対数関数 (冪型)

$$y = a(\ln x)^c + b \tag{51}$$

の場合は、上記の手順 (a)、(b) を第1ステップ、手順(c)を第2ステップとして処理する。

① 第1ステップ

- ・座標変換

$$X = g_x(x) = (\ln x)^c$$

$$Y = g_y(y) = y$$

- ・パラメータ変換

$$A = h_a(a, b) = a$$

$$B = h_b(a, b) = b$$

- ・空間補正係数

$$p = 1$$

- ・パラメータ逆変換

$$a = A$$

$$b = B$$

② 第2ステップ

- ・座標変換

$$X = g_x(x) = \ln(\ln x)$$

$$Y = g_y(y) = \ln(y - b)$$

- ・パラメータ変換

$$A = h_a(a, b) = c$$

$$B = h_b(a, b) \rightarrow \text{棄却}$$

- ・空間補正係数

$$p = e^{2Y}$$

- ・パラメータ逆変換

$$c = \lim_{(\sum p_i) \rightarrow \infty} A$$

代表的な反復型関数に対して $x - y$ 空間を変換する写像関数と空間補正係数 p_i およびパラメータ a, b の計算方法を Table 3 に示す。

6. 数値実験と評価

6.1 関数当て嵌め

空間補正無、空間補正有、Matlab に対し、以下の 5 ケース (a) ~ (e) について標本点に指定の関数を当て嵌めた数値実験を実施し、各々の結果を重畳表示して比較した。各図には、それぞれの原空間上の二乗平均誤差値

$$RMS = \sqrt{\frac{1}{n}\varepsilon} = \sqrt{\frac{1}{n}\sum_i\{\Delta y_i(x_i; a, b)\}^2} \quad (52)$$

の値と空間補正無の場合に対する比率を付記し、当て嵌め精度評価の基礎データとした。尚、Matlab の適用においては、種々の計算条件や収束条件としてデフォルト値を用い、手動による条件の変更・調整は行っていない。

(1) 基本関数

(a) 冪関数

$$y = \left(\frac{x}{a}\right)^b \quad (47)$$

(b) 正規分布の密度関数

$$y = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (45)$$

冪関数と正規分布の密度関数の当て嵌め結果をそれぞれ Fig. 1 および Fig. 2 に示す。

(2) 微分型関数

(c) 微分型対数関数

$$y = \frac{ab(\ln x)^{b-1}}{x} \quad (48)$$

(d) 成長率曲線型 (冪型) 関数

$$y = c \frac{abx^{-(a+1)}}{(1+bx^{-a})^2} \quad (a, b, c > 0) \quad (49)$$

微分型対数関数と成長率曲線型 (冪型) 関数の当て嵌め結果をそれぞれ Fig. 3 および Fig. 4 に示す。

Table 3 空間補正機能を備えた最小二乗法によるパラメータ求解方法 (反復型)

番号	当て嵌め関数		座標変換	荷重係数	パラメータ求解	備考
	関数名	関数形				
(0)	基本形 (直線形式に変換可能な関数)	$y = f(x; a, b)$ ↓ 変換可能 $Y = AX + B$	$A = h_a(a)$ $B = h_b(b)$ $X = g_x(x)$ $Y = g_y(y)$	$p = \left(\frac{\partial g_y^{-1}}{\partial Y}\right)^2$	連立方程式 $A = h_a(a)$ $B = h_b(b)$	一般形
(1)	定数項付冪型関数 【収束計算】	$y = ax^c + b$	$X = x^c$ $Y = y$	$p = 1$	$a = A$ $b = B$	第1ステップ (c : 固定)
			$X = \ln x$ $Y = \ln(y - b)$	$p = e^{2Y}$	$c = \lim_{(\sum p_i) \rightarrow \infty} A = \frac{\sum_i X_i Y_i}{\sum_i X_i^2}$	第2ステップ (b : 固定)
(2)	定数項付き指数関数 (冪型) 【収束計算】	$y = ae^{cx} + b$	$X = e^{cx}$ $Y = y$	$p = 1$	$a = A$ $b = B$	第1ステップ (c : 固定)
			$X = x$ $Y = \ln(y - b)$	$p = e^{2Y}$	$c = \lim_{(\sum p_i) \rightarrow \infty} A = \frac{\sum_i X_i Y_i}{\sum_i X_i^2}$	第2ステップ (b : 固定)
(3)	定数項付き対数関数 (冪型) 【収束計算】	$y = a(\ln x)^c + b$	$X = (\ln x)^c$ $Y = y$	$p = 1$	$a = A$ $b = B$	第1ステップ (c : 固定)
			$X = \ln(\ln x)$ $Y = \ln(y - b)$	$p = e^{2Y}$	$c = \lim_{(\sum p_i) \rightarrow \infty} A = \frac{\sum_i X_i Y_i}{\sum_i X_i^2}$	第2ステップ (b : 固定)

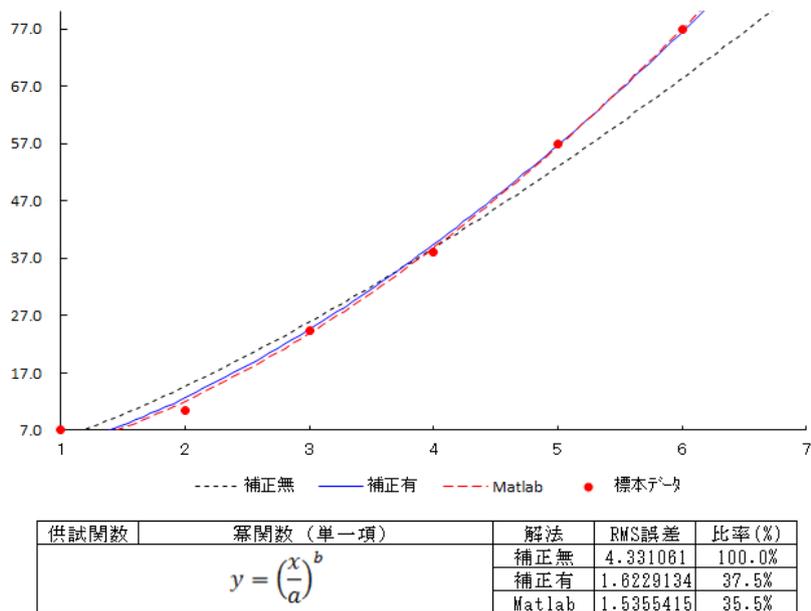


Fig.1 カーブフィッティング結果(幂関数: 単一項)

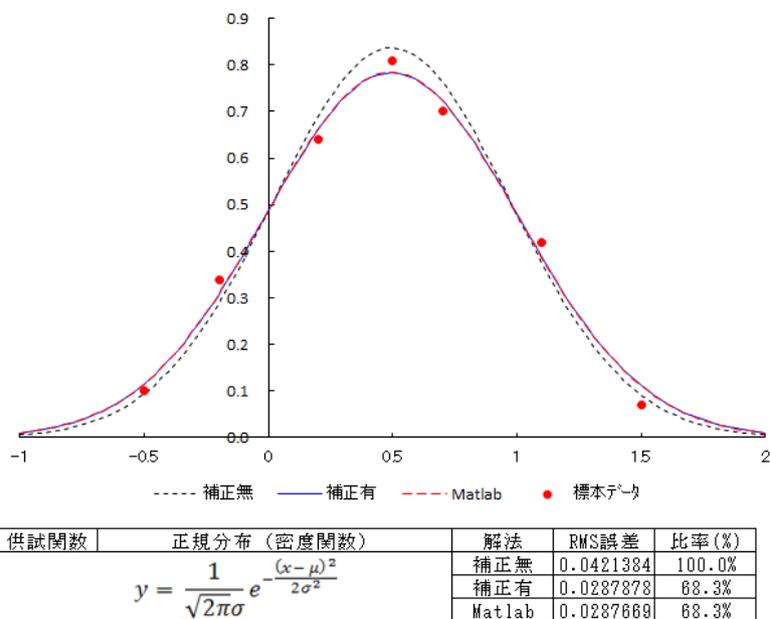
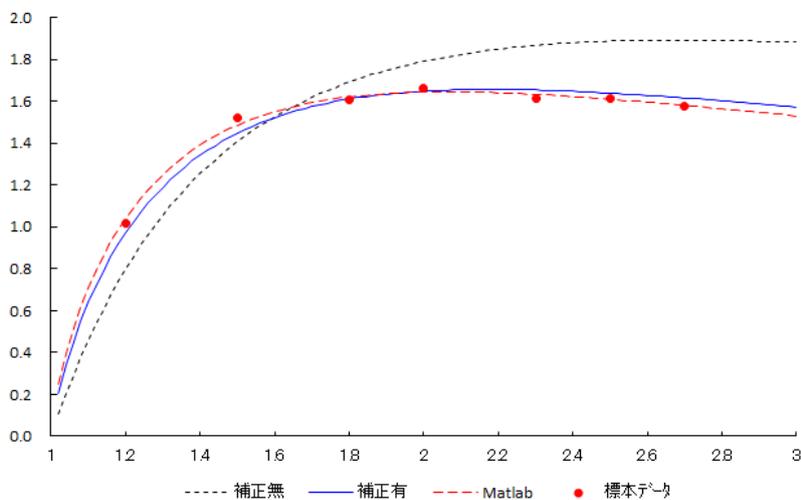
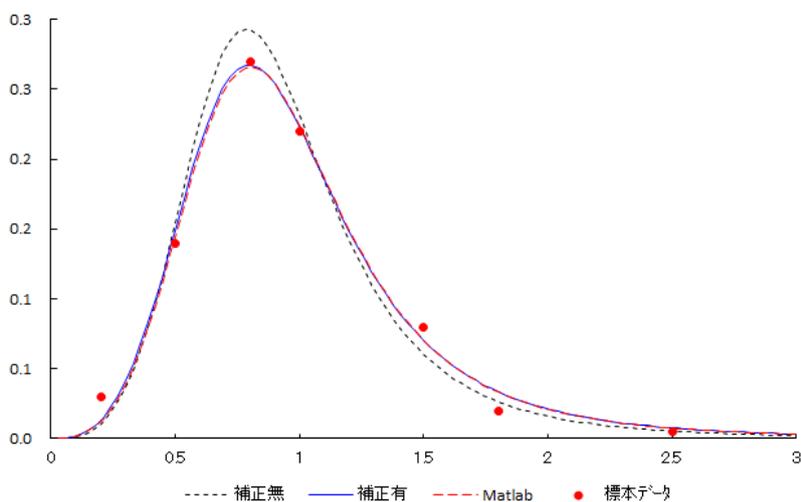


Fig.2 カーブフィッティング結果(正規分布: 密度関数)



供試関数	対数関数 (微分型)	解法	RMS誤差	比率(%)
$y = \frac{ab(\ln x)^{b-1}}{x}$		補正無	0.2156849	100.0%
		補正有	0.0406119	18.8%
		Matlab	0.0181434	8.4%

Fig.3 カーブフィッティング結果(微分型対数関数)



供試関数	成長率型曲線 (冪型)	解法	RMS誤差	比率(%)
$y = c \frac{abx^{-(a+1)}}{(1+bx^{-a})^2} \quad (a, b, c > 0)$		補正無	0.0153819	100.0%
		補正有	0.00969	63.0%
		Matlab	0.0094927	61.7%

Fig.4 カーブフィッティング結果(成長率型曲線:冪型)

(3) 反復型関数

(e) 定数項付き対数関数 (冪型)

$$y = a(\ln x)^c + b \tag{51}$$

定数項付き対数関数 (冪型) の当て嵌め結果を Fig. 5 に示す。

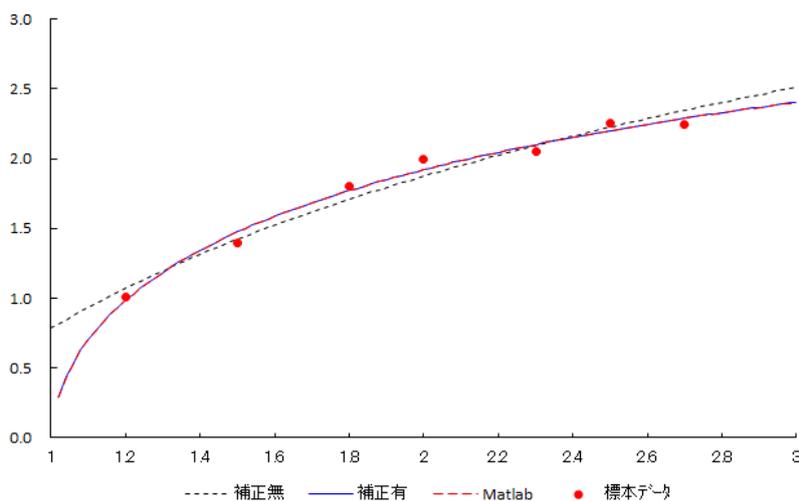
6.2 当て嵌め誤差の比較と評価

Table 1 ~ Table 3 の供試関数と標本データの組み合わせ 26 組に対して数値実験を行い、空間補正無、空間補正有および Matlab 使用時の RMS 誤差値を、空間補正無の誤差値を横軸にとって散布図で比較した。その結果を Fig. 6 に示す。

- (1) 数値実験の全てのケースにおいて、補正有の当て嵌めによって、補正無以上の精度が得られた。
- (2) 基本型、積分型、反復型を通して、一部の例外を除き、空間補正有の精度は、反復計算でパラメータを収束させることによって高い精度を求める Matlab よりは若干劣る傾向がある。しかしながら、その差は小さく、Fig. 1 ~ Fig. 5 に

おいて、空間補正有と Matlab の当て嵌め曲線は、ほぼ重なっている。

- (3) Fig. 6 においても、空間補正有の場合は、空間補正機能を伴わない線形最小二乗法に比して、高い精度で非線形関数が当て嵌められていることが分かる。また、殆どの数値実験結果で空間補正有の誤差は Matlab の誤差に重なっているか、接しており、両者の精度に大きな差がないことが明らかである。
- (4) 求解のための計算は、基本型関数の場合は一方向で反復処理を必要としない。また、微分型関数の場合も 1 回の反復のみに限定されるため、タイムシェアリングを基本とするリアルタイムシステムの基本設計段階で演算量を把握することができる。
- (5) 積分関数に対する誤差最小化を図った微分型関数の当て嵌め (積分型) においても、実施した数値実験の全てのケースで、補正無に比べて補正有の方が供試関数 (微分型関数) そのものの誤差も小さいという結果が得られた。



供試関数	対数関数 (冪関数 + 定数項)	解法	RMS誤差	比率 (%)
$y = a(\ln x)^c + b$		補正無	0.0772586	100.0%
		補正有	0.0534063	69.1%
		Matlab	0.0534057	69.1%

Fig.5 カーブフィッティング結果(定数項付き対数関数:冪型)

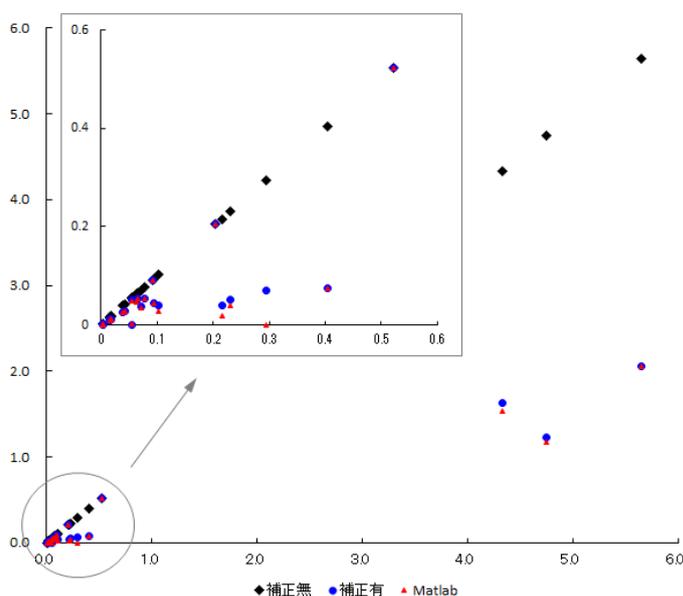


Fig.6 当て嵌め誤差比較

(6) 演算量を事前に把握することは難しいが、参考のために反復型関数についても数値実験を実施し、Matlab と同等の当て嵌め精度を得た。(Fig. 5) これは、求解の方法は異なるが、反復計算による収束という Matlab と同様の手法を用いた結果と思われる。

7. まとめ

研究開発が加速されている IoT や Industry4.0 では情報処理のリアルタイム性が重要となる。従来、データの特性解析に広く用いられてきた最小二乗法では、蓄積されたデータがオフラインでバッチ処理的に処理されるのが一般的で、リアルタイム処理への適用が難しかった。

本稿では、非線形カーブフィッティングにおいて適正な精度を確保しつつ、データ解析の即時性にとって重要となる求解の確実性と演算量の事前把握を保証するために、空間補正機能を付与した荷重係数付線形最小二乗法を提案し、数値実験による Matlab との比較で精度的にも十分実用性があることを示した。

(1) 線形最小二乗法において、標本点毎に適切な荷重係数 (空間補正係数) を与えることによって、十

分な空間補正能力を確保できることを示した。

- (2) この空間補正係数は、従属軸 (y 軸) 変換の逆関数の微係数を 2 乗した値で与えられる。一方、独立軸 (x^k 軸) の変換には依存しない。
- (3) 空間補正機能を付与した荷重係数付線形最小二乗法はリアルタイム処理向きではあるが、適用の対象は、変数変換によって多次元空間の平面 (2次元の場合は直線) に帰着できる関数に限定される。

情報処理分野においては、今後、大規模データベースやビッグデータを基盤とするオンラインシステムと即時性が求められるリアルタイムシステムの協調と融合が進むと考えられる。一方、多様な機器のインテリジェント化が進んで組み込み型システムが拡充され、計測データの特性解析においても精度と処理効率の両立が大きな課題となることが予想される。本稿のアプローチが、この問題を解決するための糸口の一つになれば幸いである。

参考文献

- (1) 朝香鉄一：品質管理のための統計数学，東京大学出版会 (1963)，pp.58-61