



## 論文

## 区分的3次関数を用いた2次元データの 平滑化の自動的方法\*

吉本 富士市\*\* 市田 浩三\*\*\* 清野 武\*\*\*

### Abstract

A method is described for smoothing two-dimensional data by a least squares fitting with a piecewise bicubic polynomial of class  $C^1$ . The number of knots and their positions are determined automatically. A condition for the good positions of knots is introduced. The condition is based on the fact that if the partial second derivatives of the underlying function of data are sufficiently close to those of the approximating function, we can obtain a good approximation. The values of the partial second derivatives are estimated from data. The estimated values of the partial second derivatives are approximated by a discontinuous piecewise linear function. The positions of knots are determined using the breakpoints of this function. The number of knots is decided by use of an information theoretical criterion (AIC).

### 1. ま え が き

2次元データは、地形とか画像のデータ処理など、多くの分野で扱われている。一般にデータには誤差があるので、しばしば得られたデータを平滑化する必要がある。

2次元データの平滑化を行う方法の1つに、最小2乗法により2変数の近似関数を2次元データへあてはめる方法がある。この方法は、単にデータを平滑化するだけでなく、データの圧縮とか信号(2.参照)の微分や積分にも有効である。また、データから等高線図を描く場合にも有用である。

これまでに開発されている方法には、まず2変数の多項式を用いるものがある<sup>1)~4)</sup>。しかし、これらは複雑な関数形をしたデータの場合には、よい近似関数を得られないことが多い。そこで、区分的な関数を用いることが考えられた<sup>5)</sup>。区分的な関数は、データが複雑な関数形をしていても、多くの場合に有効である。

しかし、それは節点の数と位置によって大きな影響を受けるので、よい節点をどのようにして決めるかという重要な問題がある。この節点の決定方法に関する研究は、1次元データの場合についてはすでにあるが<sup>6)</sup>、2次元データの場合についてはないようである。

本論文では、信号(2.参照)の2次微分を利用して節点のよい位置を求め、情報量基準 AIC<sup>7),8)</sup>を用いて節点の適切な数を決定する方法を述べる。これによれば、自動的に節点のよい位置を決定してデータ平滑化を行うことができる。

### 2. 近似関数の表現

データは、 $x-y$  平面内の矩形領域  $R=[a, b] \times [c, d]$  上の格子点で与えられており、

$$F_{i,u} = f(x_i, y_u) + e_{i,u} \quad (2.1)$$

$$\left( \begin{array}{l} i=1, 2, \dots, K; \\ u=1, 2, \dots, L \end{array} \right)$$

であると仮定する。ここに、 $f(x, y)$  は未知関数(ここでは信号と呼ぶ)で、 $e_{i,u}$  は平均値が0で未知分散が $\sigma^2$ の正規分布に従う誤差で互に独立であるとする。本論文では $(x_i, y_u)$ をサンプル点、 $F_{i,u}$ をデータ点ということにする。

領域  $R$  を格子状に小領域  $R_{m,n}$  ( $m=0, 1, \dots, h$ ;

\* An automatic method for smoothing two-dimensional data with a piecewise bicubic polynomial by Fujiichi YOSHIMOTO (Akashi Technical College), Kozo ICHIDA and Takeshi KIYONO (Faculty of Engineering, Kyoto University).

\*\* 明石工業高等専門学校

\*\*\* 京都大学工学部情報工学科

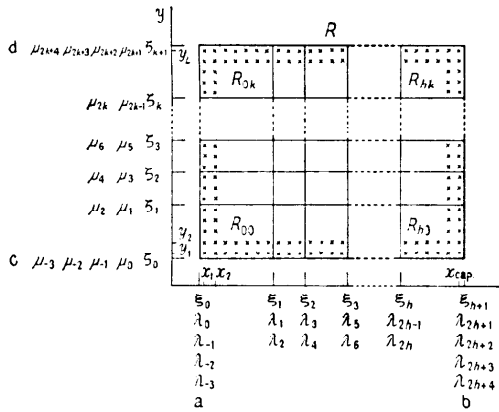


Fig. 1. The area of fitting, divided by breakpoints  $\xi_m, \zeta_n$  into panels  $R_{mn}$ . Here  $\lambda_i$  and  $\mu_j$  denote knots.

$n=0, 1, \dots, k$ ) に分けて, 各小領域で別々の近似関数を用いる (Fig. 1). ただし, 近似関数をなめらかにするために小領域の継ぎ目で関数と微分を連続にする. このとき近似関数は

$$S(x, y) = \sum_{i=1}^{2h+4} \sum_{j=1}^{2k+4} \gamma_{ij} M_i(x) N_j(y) \quad (2.2)$$

と表現できる. ここに,  $M_i(x)$  および  $N_j(y)$  は正規化された3次の B-スプラインであり,  $\gamma_{ij}$  は最小 2 乗法で決定するパラメータである<sup>5), 9)</sup>.  $M_i(x)$  ( $i=1, 2, \dots, 2h+4$ ) に対する節点を  $\lambda = (\lambda_{-3}, \lambda_{-2}, \dots, \lambda_{2h+4})$  とし,  $N_j(y)$  ( $j=1, 2, \dots, 2k+4$ ) に対する節点を  $\mu = (\mu_{-3}, \mu_{-2}, \dots, \mu_{2k+4})$  とする. また, 各小領域の継ぎ目を,  $x$  方向に対しては  $\xi = (\xi_0, \xi_1, \dots, \xi_{h+1})$ ,  $y$  方向に対しては  $\zeta = (\zeta_0, \zeta_1, \dots, \zeta_{k+1})$  とする. ただし  $\xi_0 = a, \xi_{h+1} = b, \zeta_0 = c, \zeta_{k+1} = d$  である. すると関数と微分の連続条件から,  $x$  方向に対しては

$$a < \lambda_1 = \lambda_2 = \xi_1 < \lambda_3 = \lambda_4 = \xi_2 < \dots < \lambda_{2h-1} = \lambda_{2h} = \xi_h < b \quad (2.3)$$

となる. 領域  $R$  の端では節点を 4 重にして

$$\left. \begin{aligned} \lambda_{-3} = \lambda_{-2} = \lambda_{-1} = \lambda_0 = \xi_0 = a, \\ \lambda_{2h+1} = \lambda_{2h+2} = \lambda_{2h+3} = \lambda_{2h+4} = \xi_{h+1} = b \end{aligned} \right\} \quad (2.4)$$

とする.  $y$  方向についても同様である. B-スプライン  $M_i(x), N_j(y)$  は Cox-de Boor アルゴリズムによって安定に計算できる<sup>9), 10)</sup>.

### 3. 節点のよい位置

2変数  $x, y$  に関して 1 次微分まで連続な区分的 3 次関数は,  $x$  または  $y$  に関するその 2 次微分が継ぎ目で不連続な折れ線になる特徴がある. この性質を利

用して, 節点のよい位置を自動的に決める. 信号の 2 次微分と近似関数の 2 次微分が十分近くなるように考慮すれば, よい近似関数を得ることができると考えられる. 以下, この点を検討しよう.

まず, 節点  $\lambda$  をどのように決めればよいかを考える. いま  $y = y_u$  とすると, データ  $F_u = (F_{1u}, F_{2u}, \dots, F_{ku})$  は  $x$  に関する 1 変数のデータと考えられ, 式 (2.2) は次のように書ける.

$$\begin{aligned} S(x, y_u) &= \sum_{i=1}^{2h+4} \sum_{j=1}^{2k+4} \gamma_{ij} M_i(x) N_j(y_u) \\ &\equiv \sum_{i=1}^{2h+4} \tau_{iu} M_i(x). \end{aligned} \quad (3.1)$$

このとき, 節点  $\lambda$  がどのようになっておれば, 式 (3.1) は信号  $f(x, y_u)$  のよい近似関数となるであろうか. ここで,  $H_m = [\xi_m, \xi_{m+1})$  ( $m=0, 1, \dots, h$ ) と書く. すると,  $y = y_u, a \leq x \leq b$  における  $f(x, y_u)$  と  $S(x, y_u)$  の 2 乗ノルム  $G_u$  は次式となる.

$$G_u = \sum_{m=0}^h \sum_{x \in H_m} \{f(x, y_u) - S(x, y_u)\}^2 \quad (3.2)$$

定理 各区間  $\{H_m | m=0, 1, \dots, h\}$  において  $f(\xi_m, y_u) = S(\xi_m, y_u), \partial f(x, y_u) / \partial x |_{x=\xi_m} = \partial S(x, y_u) / \partial x |_{x=\xi_m}$  および

$$\left| \frac{\partial^2 f(x, y_u)}{\partial x^2} - \frac{\partial^2 S(x, y_u)}{\partial x^2} \right| \leq \theta_{mu} \quad (3.3)$$

ならば

$$G_u \leq \frac{1}{4} \sum_{m=0}^h (\theta_{mu})^2 \sum_{x \in H_m} (x - \xi_m)^4. \quad (3.4)$$

証明 式 (3.3) は次のように書ける.

$$\frac{\partial^2 S(x, y_u)}{\partial x^2} - \theta_{mu} \leq \frac{\partial^2 f(x, y_u)}{\partial x^2} \leq \frac{\partial^2 S(x, y_u)}{\partial x^2} + \theta_{mu}. \quad (3.5)$$

そこで, 次の積分

$$\begin{aligned} \int_{\xi_m}^x \left\{ \frac{\partial^2 S(z, y_u)}{\partial z^2} - \theta_{mu} \right\} dz &\leq \int_{\xi_m}^x \frac{\partial^2 f(z, y_u)}{\partial z^2} dz \\ &\leq \int_{\xi_m}^x \left\{ \frac{\partial^2 S(z, y_u)}{\partial z^2} + \theta_{mu} \right\} dz \end{aligned} \quad (3.6)$$

を行い, 仮定  $\partial f(x, y_u) / \partial x |_{x=\xi_m} = \partial S(x, y_u) / \partial x |_{x=\xi_m}$  を用いると

$$\begin{aligned} \frac{\partial S(x, y_u)}{\partial x} - \theta_{mu}(x - \xi_m) &\leq \frac{\partial f(x, y_u)}{\partial x} \leq \frac{\partial S(x, y_u)}{\partial x} \\ &\quad + \theta_{mu}(x - \xi_m). \end{aligned} \quad (3.7)$$

この不等式に, 再び同様な積分を行って仮定

$f(\xi_m, y_u) = S(\xi_m, y_u)$  を用いれば

$$|f(x, y_u) - S(x, y_u)| \leq \frac{1}{2} \theta_{mu}(x - \xi_m)^2. \quad (3.8)$$

ゆえに式 (3.4) となる。(証明終)

この定理の仮定  $f(\xi_m, y_u) = S(\xi_m, y_u)$  および  $\partial f(x, y_u) / \partial x|_{x=\xi_m} = \partial S(x, y_u) / \partial x|_{x=\xi_m}$  は、節点不適切で近似が悪い場合には成り立たないが、よい節点が求まり満足できる近似関数となれば成り立つと考えられる。従って、式(3.4)の右辺の値をできるだけ小さくするように継ぎ目  $\xi$  (よって節点  $\lambda$ ) を決めれば、式(3.1)は信号  $f(x, y_u)$  のよい近似関数となる。しかし、このような継ぎ目  $\xi$  (ゆえに節点  $\lambda$ ) は  $y_u$  ( $u=1, 2, \dots, L$ ) の関数である。2. で述べたように、 $\lambda$  は  $y$  に関係しない必要がある。そこで、次のようにして  $\lambda$  のよい位置を求める。

まず信号の2次微分  $\partial^2 f(x, y_u) / \partial x^2|_{x=x_t}$  ( $t=1, 2, \dots, K$ ) を求める。これは未知であるから各サンプル点においてデータ  $F_u$  から推定する。この値を  $g_u = (g_{1u}, g_{2u}, \dots, g_{Ku})$  と書く。次に、この  $g_u$  を継ぎ目で不連続な折れ線  $B_u(x)$  を用いて最良近似する。この継ぎ目を  $\phi_{mu}$  ( $m=0, 1, \dots, h+1$ ) とする。ただし、 $\phi_{0u} = a, \phi_{h+1, u} = b$  である。また、各区間  $D_{mu} = [\phi_{mu}, \phi_{m+1, u}]$  ( $m=0, 1, \dots, h$ ) において

$$\max_{x \in D_{mu}} |g_{tu} - B_u(x)| = \delta_{mu} \quad (3.9)$$

とする。このとき、定理により

$$\begin{aligned} V_u &= \sum_{m=0}^h (\delta_{mu})^2 \sum_{x \in D_{mu}} (x - \phi_{mu})^4 \\ &\equiv \sum_{m=0}^h \Delta_{mu} \end{aligned} \quad (3.10)$$

を最小にするように近似する。

$y_u$  を変えて計算すれば、折れ線  $B_u(x)$  による近似の継ぎ目  $\phi_{mu}$  ( $m=0, 1, \dots, h+1; u=1, 2, \dots, L$ ) を得ることができる。次に、この  $\phi_{mu}$  ( $u=1, 2, \dots, L$ ) に対して次のような重みつき最小2乗法を行って、近似関数の式(2.2)の継ぎ目  $\xi_m$  を決定する。これから節点  $\lambda_{2m-1}$  および  $\lambda_{2m}$  は自動的に決まる。

いま、 $m=1$  としよう (Fig. 2)。  $u$  の各値に対して、継ぎ目  $\phi_{1u}$  の両側の  $\Delta_{0u}$  と  $\Delta_{1u}$  を比較して、大きい方を重み  $w_{1u}$  とする。すなわち、

$$w_{1u} = \max(\Delta_{0u}, \Delta_{1u}) \quad (u=1, 2, \dots, L) \quad (3.11)$$

継ぎ目  $\xi_1$  は、残差の2乗和

$$G_1 = \sum_{u=1}^L w_{1u} (\xi_1 - \phi_{1u})^2 \quad (3.12)$$

を最小にするように決定する。よって

$$\xi_1 = \lambda_1 = \lambda_2 = \frac{\sum_{u=1}^L w_{1u} \phi_{1u}}{\sum_{u=1}^L w_{1u}} \quad (3.13)$$

同様にして、 $m=2, 3, \dots, h$  に対しても継ぎ目  $\xi_m$

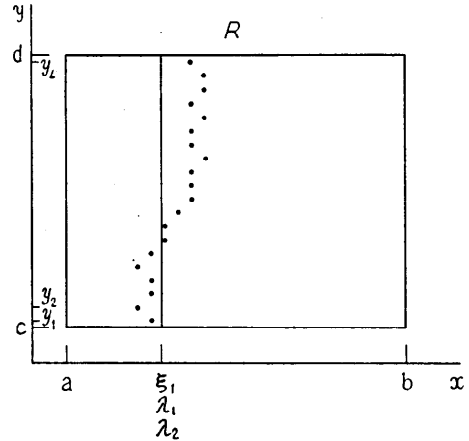


Fig. 2 Determination of the breakpoints  $\xi_1$ . The dots in the figure denote the breakpoints  $\phi_{1u}$ .

(従って節点  $\lambda_{2m-1}, \lambda_{2m}$ ) を決定できる。式(3.11)のように重みを決める理由は、2次微分の近似の誤差ノルム  $\Delta_{mu}$  の大きい継ぎ目を重視して節点を決めることにより、式(2.2)の2次微分が信号の2次微分に近くなるようにするということである。

節点  $\mu$  についても以上と同様である。

#### 4. 信号の2次微分の推定と近似

信号の  $x$  に関する2次微分の推定値  $g_{tu}$  ( $t=1, 2, \dots, K; u=1, 2, \dots, L$ ) は、各  $y_u$  においてデータ  $F_u$  に対して1変数の2次微分の推定アルゴリズム<sup>6)</sup>を適用して計算する。信号の  $y$  に関する2次微分も同様に推定する。

次に、こうして得られた2次微分の推定値を、継ぎ目で不連続な折れ線で近似する。ここでは信号の  $x$  に関する2次微分の推定値の近似についてだけ述べる ( $y$  に関しても同様である)。各  $u$  において  $g_u$  に対する近似を計算する。このとき、式(3.10)をできるだけ小さくするように近似する。そこで許容誤差  $\epsilon_u$  を与えて、まず次のような最良近似<sup>11)</sup>を行う。

近似線分は、 $g_u$  に対して左から右に向って次々と計算する。この近似の誤差ノルムとしては、式(3.10)の  $\Delta_{mu}$  を用いる。  $g_{1u}$  から近似を始めて、 $\Delta_{0u} \leq \epsilon_u$  であれば、2次微分の値を1つずつ増やしながら近似線分を更新する。こうして  $\Delta_{0u} \leq \epsilon_u$  となる近似の中で最も多くのデータを含んだ近似を第1近似線分とする。この線分に含まれる2次微分の値を  $g_{tu}$  ( $t=1, 2, \dots, J$ ) とすると、第2近似線分は  $g_{tu}$  ( $t=J+1, J+2, \dots$ ) に対して計算される。このプロセスを続けると、 $g_u$

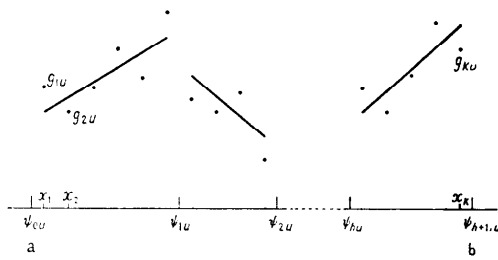


Fig. 3 Approximation of the estimated values of the partial second derivative of the signal.

はいくつかの線分によって近似され、各線分の誤差ノルムは  $\epsilon_u$  以下である。すなわち線分の数  $h+1$  とすると、 $\Delta_{mu} \leq \epsilon_u$  ( $m=0, 1, \dots, h$ ) である (Fig. 3)。また、この方法によって得られる線分数は、 $\Delta_{mu} \leq \epsilon_u$  ( $m=0, 1, \dots, h$ ) となる近似の中で、最も少ない。

次に、この方法で得られた線分数を増やさずに、 $\Delta_{mu}$  ( $m=0, 1, \dots, h$ ) の最大値を最小にする。これは文献 6) の 4. の“最適化”と同様な方法でできる。このアルゴリズムを“最大誤差ノルムの最小化”ということにする。この結果、式 (3.10) の値を相当に小さくするような近似となる。

この後で、次の簡単な最小化アルゴリズムを用いて式 (3.10) の値をできるだけ小さくするような近似を求める。この方法を“誤差ノルムの和の最小化”ということにする。

ステップ 1

$m=1$  に設定する。

ステップ 2

もしも  $\Delta_{m-1} \geq \Delta_m$  ならば

$$\psi_{mu}' = \psi_{mu} - \nu. \tag{4.1}$$

反対に  $\Delta_{m-1} < \Delta_m$  の場合は

$$\psi_{mu}' = \psi_{mu} + \nu. \tag{4.2}$$

ステップ 3

$\psi_{mu}'$  を用いて新しい  $\Delta_{m-1}'$  と  $\Delta_m'$  を計算する。

ステップ 4

もしも  $\Delta_{m-1} + \Delta_m > \Delta_{m-1}' + \Delta_m'$  ならば、 $\psi_{mu}$  を  $\psi_{mu}'$  でおきかえる。そうでなければ  $\psi_{mu}$  はそのままにしておく。

ステップ 5

$m \leftarrow m+1$  とする。もしも  $m \leq h$  ならばステップ 2 へもどる。そうでなければステップ 6 へいく。

ステップ 6

ステップ 1 ~ ステップ 5 で  $\psi_{mu}$  ( $m=1, 2, \dots, h$ ) が 1 つも動かなければ終了する。さもなければステップ

1 へもどる。

$\nu$  の値はサンプル点の間隔の大きさ程度にすればよい。 $\psi_{mu}$  ( $m=1, 2, \dots, h$ ) の初期値はその最適値に近いので、このアルゴリズムは速い。

以上のようにして、式 (3.10) の値をほとんど最小にする近似を自動的に得ることができる。従って、こうして得られた折れ線による近似の継ぎ目  $\psi_{mu}$  ( $m=1, 2, \dots, h; u=1, 2, \dots, L$ ) から、3. で述べたようにして継ぎ目  $\xi$  を決定し、節点  $\lambda$  を求められる。

近似線分の個数は、各  $\nu_u$  において  $\epsilon_u$  を次のように変えれば、自動的に 1 個ずつ増大する。まず

$$\epsilon_u \geq \left[ \frac{\max_{1 \leq i \leq K} (g_{iu}) - \min_{1 \leq i \leq K} (g_{iu})}{2} \right] \times \sum_{t=1}^K (x_t - \psi_{0u})^4 \tag{4.3}$$

とすれば、線分数 1 の近似となる。この誤差ノルムを  $\Delta_u^{(0)}$  と書く。次に、 $\epsilon_u = \rho \times \Delta_u^{(0)}$  とすれば線分数 2 の近似となる。ここに  $\rho$  は 1 を越えないが、しかし 1 に近いある定数である。この後、 $\epsilon_u = \rho \times \Delta_u^{(h-1)}$  とすれば、線分数  $h+1$  の近似を得られる。ここに、 $\Delta_u^{(h-1)}$  は線分数  $h$  において“最大誤差ノルムの最小化”を行った後の誤差ノルムの最大値である。従って、 $g_u$  に対して線分数 1, 2, ... の近似を連続的に得られる。ゆえに  $h=1, 2, \dots$  に対する継ぎ目  $\xi$ 、節点  $\lambda$  のよい位置を自動的に得られる。

ここで、 $\epsilon_u$  が非常に小さくなったときには、この  $g_u$  に対しては、それ以上に近似線分数を増大できない。従って、この  $g_u$  に対する近似はここまでとして式 (3.11) ~ (3.13) において、この  $\nu_u$  は考慮しないことにする。

5. 最小 2 乗法による近似関数のあてはめ

式 (2.1) のデータに対して、式 (2.2) を最小 2 乗法によりあてはめる。残差の 2 乗和  $Q$  は、

$$Q = \sum_{u=1}^L \sum_{t=1}^K \{S(x_t, y_u) - F_{tu}\}^2. \tag{5.1}$$

この式をパラメータ  $\gamma_{ij}$  ( $i=1, 2, \dots, 2h+4; j=1, 2, 2h+4$ ) で偏微分して、その結果を 0 とおくと正規方程式

$$A\gamma = b \tag{5.2}$$

を得る。ここに、

$$\gamma = (\gamma_{1,1}, \gamma_{2,1}, \dots, \gamma_{2h+4, 2h+4})^T \tag{5.3}$$

であり、 $T$  は転置行列を表わす。係数行列  $A$  は、 $B$ -スプラインの性質から帯行列となり、対角要素がかなり大きくて条件はよい。式 (5.2) の解を式 (2.2)

へ代入すれば近似関数を決定できる。

ここで、よい近似を得るために、3.および4.で述べたようにして節点のよい位置を求めるとともに、情報量基準 AIC<sup>7),8)</sup>を用いて節点の適切な数を決定する。AIC は一般的には

$AIC = (-2) \log_e(\text{最大尤度}) + 2(\text{パラメータ数})$  (5.4) と定義されている。ここに、パラメータ数はモデル内において自由に变化させられるもの数である。AIC の値が小さいほど、よいモデルとみなされる<sup>9)</sup>。

式 (2.2) から、次の回帰モデルを考えよう。

$$F_{iu} = S(x_i, y_u) + e_{iu}, \quad \begin{matrix} (i=1, 2, \dots, K) \\ (u=1, 2, \dots, L) \end{matrix} \quad (5.5)$$

あてはめの評価の基準として、このモデルに AIC を適用すれば、式 (5.4) より

$$AIC = KL \log_e Q + 2(2h+4)(2k+4) \quad (5.6)$$

とできる。従って、節点の数を増大しながらあてはめを計算して、AIC を最小にするモデルを選ぶ。このときよい近似関数となり、節点の数も適切なものと考えられる。ここに、AIC が有効であるためには  $KL \gg 2(2h+4) \times (2k+4)$  が必要である。

なお、式 (5.2) が一意的な解を持つためには、Schoenberg と Whitney の条件<sup>12)</sup>が成り立つ——すなわち各 B-スプラインが支持するデータ点が少なくとも1つある——必要がある。ここで述べた方法では、データは格子状にあると仮定したので、各区間  $[\xi_m, \xi_{m+1})$  ( $m=0, 1, \dots, h$ ) に  $x_i$  が2つ以上および各区間  $[\zeta_n, \zeta_{n+1})$  ( $n=0, 1, \dots, k$ ) に  $y_u$  が2つ以上入っているようにすれば、この条件は満足される。

### 6. あてはめのアルゴリズム

あてはめを行う場合に、 $x$  方向の節点  $\lambda$  と  $y$  方向の節点  $\mu$  の要素の数 (従って  $h$  と  $k$  の値) を、それぞれいくらにするかという問題がある。これは、2変数  $x$  および  $y$  の多項式による近似で、 $x$  と  $y$  を何次の項まで考慮するかというのと類似の問題である。理想的には、 $h$  と  $k$  の可能な組み合わせすべてについて、パラメータ数  $(2h+4) \times (2k+4)$  の値の小さい順にあてはめを計算し、AIC を最小にするものを選び出せばよい。しかし実際にそうすると、計算すべきあてはめの数は  $h$  および  $k$  の増大とともに急激に増えて、計算量が多くなりすぎる。そこで、 $h=k$  とした。こうしても、節点のよい位置を求めれば、多くの場合によいあてはめを得ることができる。

あてはめのアルゴリズムの概要は Fig. 4 のように

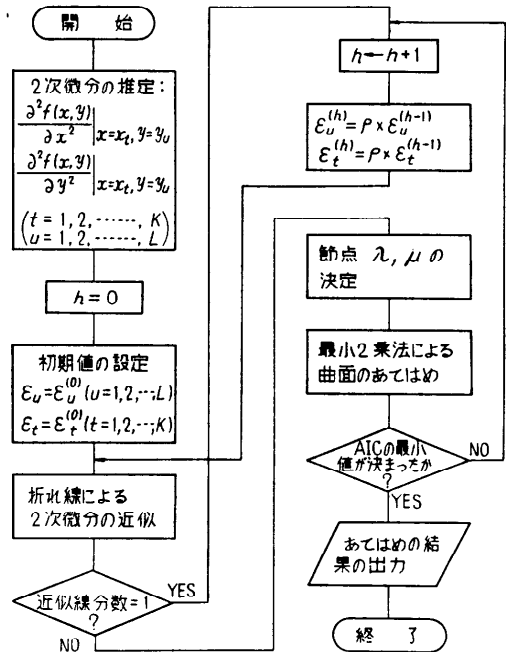


Fig. 4 The outline of the two-dimensional data fitting.

なる。AIC は、パラメータ数  $(2h+4) \times (2k+4)$  に対して必ずしも単峰性ではない。しかし、AIC が最小となるときのよい近似関数が得られるので、よい近似関数 (従ってよい節点) を自動的に決定できる。実際の計算では、 $h$  と  $k$  の値を同時に増やしながら計算していき、AIC の値がある値より続けて4つほど大きな値になったときに計算を終る。そして、このある値を AIC の最小値とする。

### 7. 計算例

次のデータを用いた。

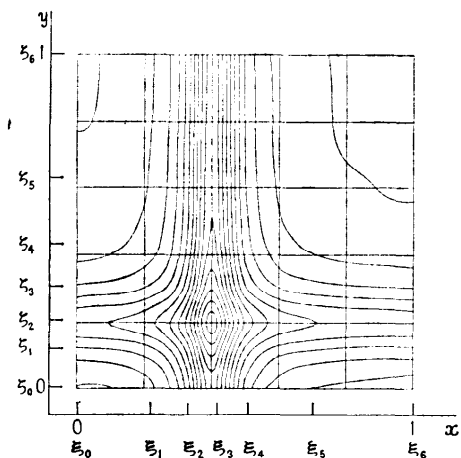
$$F_{iu} = \frac{1}{0.01 + 2(x_i - 0.4)^2} + \frac{1}{0.02 + 2(y_u - 0.2)^2} + e_{iu}, \quad \begin{matrix} (x_i : 0.0(0.02)1.0) \\ (y_u : 0.0(0.02)1.0) \end{matrix} \quad (7.1)$$

ここに、 $e_{iu}$  は互に独立な誤差で、正規分布をし、その平均値は0で分散は1である。あてはめの領域は、 $[a, b] \times [c, d] = [0, 1] \times [0, 1]$  とした。

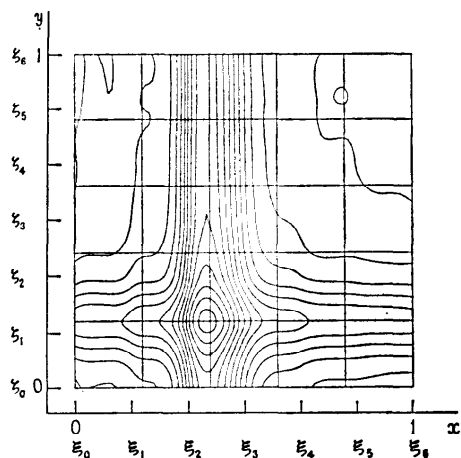
Table 1(次頁参照)に AIC の計算値を示す。ここで述べた方法では、パラメータ数  $(2h+4) \times (2k+4)$  が 196 のときに最小値になっている。この場合のあてはめの結果が Fig. 5(次頁参照)の等高線図であ

**Table 1** The calculated values of AIC and the unbiased estimator  $\delta$  of the variance of errors.

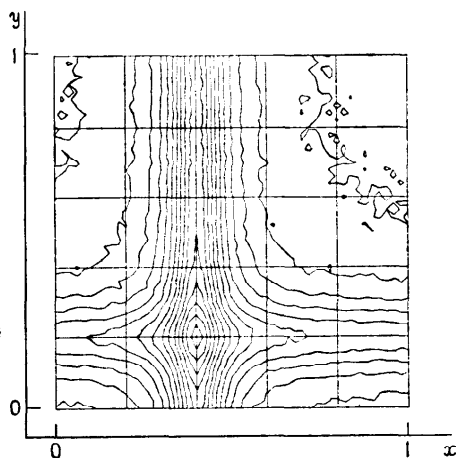
The number of parameters (2h+4) × (2k+4)	proposed method		equal intervals	
	AIC	$\delta$	AIC	$\delta$
36	32,943	120.078	32,810	114.081
64	28,830	24.442	30,989	56.046
100	23,350	2.933	30,529	46.339
144	21,392	1.360	29,175	27.093
196	20,738	1.038	26,848	10.870
256	20,837	1.056	22,596	2.077
324	20,827	1.028	21,632	1.401
400	20,850	1.012	22,784	2.129
484	20,952	1.026	22,074	1.578



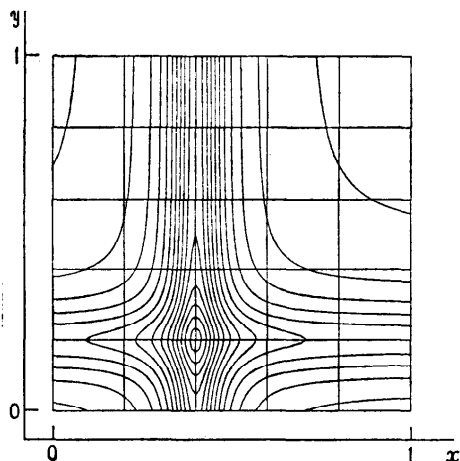
**Fig. 5** The result of fitting by use of the method described here.



**Fig. 6** The result of fitting with the equal intervals of breakpoints  $\xi_m$  and  $\xi_n$ .



**Fig. 7** The contour lines drawn from the original data.



**Fig. 8** The contour lines drawn from the signal.

る。継ぎ目  $\xi, \zeta$  を等間隔にした場合には、AIC の値は大きい。Fig. 6 はこの場合のあてはめの結果であり、パラメータ数は Fig. 5 と同じである。Fig. 7 はデータから直接に描いたものである。Fig. 8 は信号である。Fig. 5 と Fig. 8 はよく似ている。ここに、Fig. 5 ~ Fig. 8 の等高線は 5, 15, 25, ... と 10 おきの高さで描いている。Table 1 から、Fig. 5 のときの誤差分散の不偏推定量  $\delta$  はデータのその値 1 に近いことがわかる。

ここに、

$$\delta = Q / \{KL - (2h+4)(2k+4)\} \quad (7.2)$$

である。

## 8. む す び

2変数の区分的3次関数を用いた2次元データの平滑化について述べた。節点の数と位置は自動的に決定される。AICを用いる場合には、データの数パラメータの数に比べて十分多い必要がある。従って、信号が複雑な関数形をしているにもかかわらずデータ数が少ない場合(すなわち補間に近いあてはめ)に、あてはめの基準としてAICを適用するのは好ましくない。この場合については、さらに研究する必要がある。データが各点で異なる重みを持っている場合には、重み付きの最小2乗法を用いればよい。データが格子状でない不規則データの場合には、何らかの方法で適当な格子状上の2次微分の値を推定して、これにより節点の位置を決めることが考えられる。これについても、今後の研究が望まれる。

なお、2.で継ぎ目 $\xi$ は $y$ に無関係としたが、これを $y$ の関数にすれば、式(2.2)はいっそう柔軟な近似関数になると思われる。このとき、3.で $\xi$ を決める場合に、 $\psi_{mu}$  ( $u=1, 2, \dots, L$ )を $y$ の関数で近似することが考えられる。継ぎ目 $\xi$ についても同様である。これについては今後研究する予定である。

## 参 考 文 献

- 1) J.H. Cadwell: A least squares surface fitting program, *The Computer Journal*, Vol. 3, pp. 266~269 (1961).
- 2) J.H. Cadwell and D.E. Williams: Some or-

thogonal methods of curve and surface fitting, *The Computer Journal*, Vol. 4, pp. 260~264 (1962).

- 3) C.W. Clenshaw and J.G. Hayes: Curve and surface fitting, *J. Inst. Maths Applics*, Vol. 1, pp. 164~183 (1965).
- 4) J.G. Hayes: Fitting data in more than one variable, *Numerical Approximation to Functions and Data*, J. G. Hayes, ed., pp. 84~97, Athlone Press, London (1970).
- 5) J.G. Hayes and J. Halliday: The least-squares fitting of cubic spline surfaces to general data sets, *J. Inst. Maths Applics*, Vol. 14, pp. 89~103 (1974).
- 6) 吉本, 市田, 清野: 区分的3次関数を用いたデータ平滑化——節点の決定について——, *情報処理*, Vol. 17, pp. 200~206 (1976).
- 7) H. Akaike: A new look at the statistical model identification, *IEEE Trans. Automat. Contr.*, Vol. AC-19, pp. 716~723 (1974).
- 8) 赤池: 情報量基準 AIC とは何か, *数理科学*, No. 153, pp. 5~10 (1976).
- 9) C. de Boor: On calculating with B-splines, *J. Approx. Theory*, Vol. 6, pp. 50~62 (1972).
- 10) M.G. Cox: The numerical evaluation of B-splines, *J. Inst. Maths Applics*, Vol. 10, pp. 134~149 (1972).
- 11) 鈴木: 実験データの折線近似, *情報処理学会11回大会予稿集*, pp. 299~300 (1970).
- 12) I.J. Schoenberg and A. Whitney: On Pólya frequency functions III, *Trans. Amer. Math. Soc.*, Vol. 74, pp. 246~259 (1953).

(昭和51年6月30日受付)