数理解析研究所講究録
483巻 1983年 86-111

86

ベイズの方法によるデータのあてはめ

電通大　情報数理・統数研　田中　輝雄 (Teruo Tanaka)
統数研　田辺　國士 (Kunio Tanabe)

はじめに

測定，観測によって得られたデータから対象に関する情報を取り出すことは，自然科学をはじめとする各分野の基本的な仕事である。データは一般に誤差を含んでいる。つまり，

［データ］＝［構造］＋［誤差］

と考えられる。［構造］と［誤差］を分離し，［構造］を取り出すためにはあらかじめ［構造］あるいは［誤差］に対してのなんらかの仮定（事前情報）がなくてはならない。“［構造］を特定の形を持つ曲線（曲面）で近似する”と仮定すると，曲線（曲面）のあてはめ（回帰）問題に帰着する。

近似関数としては，ふつう多項式やスプライン関数などが用いられている（石黒，荒畑，市川，吉本，田辺［9，10，19，21］）。近似関数のパラメタの値にデータに対する近似の良さを測る尺度に，$\|\cdot\|_1$，$\|\cdot\|_2$，$\|\cdot\|_\infty$ などのノルムを採用することにより，最小二乗問題やＬＰ問題に帰着させてパラメタの値を求める事が多い。しかし，近似関数としてこのような多項式やスプライン関数などを用いる方法にはいくつかの問題がある。一つは，［構造］を反映しない近似関数の“くせ”が出てしまうことである。たとえば，近似関数として高次の多項式をあてはめると不必要な振動を起こしてしまうし，スプライン関数も節点が適当でないと良い結果が得られない。また，推定すべき［構造］が単調性，非負性などとあらかじめ分かっている場合，それらを制約条件として前出のような近似関数に組み込むことは容易ではない。特に制約条件が不等式である場合取り扱いが面倒である。

そこで，本稿では近似関数ｆの形状をあらかじめ定めることを最低限におさえ，事前情報を取り込み易い『離散スプライン』の形にｆを表現し，評価関数

（データと近似関数ｆとの距離）＋ $\alpha^2\times$（近似関数ｆのなめらかさの強さ）　（1）

を最小化するｆを最適近似関数とみる方法を述べる。このような方法においては，スカラーαの値の選択が重要である。

ここでは，赤池，石黒，荒畑，柏木，北川，中村らの研究([7，8，9，11，12，13，14，18])で有用性が示されている情報量規準ABICを用いてfをベイズモデルに定式化し，データに基づいてαの値を決定してfを推定する。本稿の目的は，

①『離散スプライン』の有用性

②スパースで構造のある行列を持つ最小二乗問題に対するGivens変換による方法の有効性

③情報量規準ABICの有用性

を示すことにある。第1節では，近似関数fをベイズモデルとして定式化し，最小二乗問題に帰着させる。第2節では，その最小二乗問題をGivens変換法によって効率よく解くアルゴリズムについて述べる。第3節では近似関数fが三次の『離散スプライン』となること示す。第4節でこの方法を適用した例を，一次元問題，二次元問題について示す。副産物として離散スプラインが二次元問題のデータ補間としても用いることができることについて触れる。第5節では，事前情報を組み込むために，問題を制約条件付き二次計画問題として定式化し解く方法について述べる。第6節では二相問題のモデル化について述べる。第7節では不等間隔データ問題のモデル化について述べる。

1．ベイズモデルと情報量規準ABIC

推定すべき関数fを，離散点 x_j 上の値 $f_j = f(x_j)$，$(1 \leq j \leq n)$ で表現する。つまり，

$$f = (f_1, f_2, \cdots, f_j, \cdots f_n)^t \quad (2)$$

とfをパラメトライズする（tは転置）。以下では関数とベクトルfを同一視する。x_j は等間隔にとりnは十分大きくとっておく。二次元問題では長方形領域を考え x_j を格子点 $(n_1 \times n_2)$ 上の点と考える $(n_1 \times n_2 = n)$。一次元問題の場合，図1.1 のようになる。

今，N個のデータ $\bar{y}_i (1 \leq i \leq N)$ が得られているとする。$\bar{y}$ を，

$$\bar{y} = (\bar{y}_1, \bar{y}_2, \cdots, \bar{y}_i, \cdots \bar{y}_N)^t \quad (3)$$

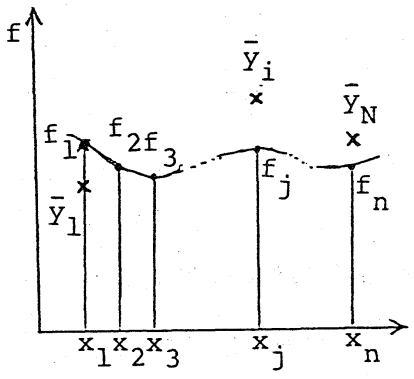

図1.1　モデル関数 f

と表す。nはNより十分大きくとる。fit-point x_j 上に常にデータ $\bar{y}_i$ がなくてもよい。

データ $\bar{y}_i$ の誤差 ε_i を

$$\varepsilon_i = \text{i.i.d.N}(0,\sigma^2) \tag{4}$$

とすれば y_i の分布は,

$$p_i(y_i|f_j,\sigma^2) = \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{1}{2}} \times \exp\left(-\frac{1}{2\sigma^2}(y_i-f_j)^2\right) \tag{5}$$

$$E = \begin{bmatrix} 1 & & & & & & & j \\ & 1 & 0 & 1 & 0 & 0 & 1 & \\ & & & & & & & i \\ & & & & & & 0 & 1 \end{bmatrix}$$

図1.2 Eの概形

となる。ただし, $\bar{y}_i$ は f_j 上のデータである。よって, yの分布は,

$$p(y|f,\sigma^2) = \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{N}{2}} \exp\left(-\frac{1}{2\sigma^2}\|y-Ef\|^2\right) \tag{6}$$

と表せる。ここで $\|\cdot\|$ は, ユークリッドノルムを表す。またEは, yとfの対応を表す行列である。

つぎに, fのなめらかさを二階差分の大きさで表現する。各点での二階差分の大きさを次のような量,

(一次元モデル) $|f_{i-1}-2f_i+f_{i+1}|$, $2 \leqq i \leqq n-1$, (7)

(二次元モデル) $|f_{i,j-1}+f_{i-1,j}-4f_{i,j}+f_{i,j+1}+f_{i+1,j}|$,

$2 \leqq i \leqq n_1-1$, $2 \leqq j \leqq n_2-1$, (8)

$|f_{i-1,j}-2f_{i,j}+f_{i+1,j}|$, $j=1,n_2$, $2 \leqq i \leqq n_1-1$, (9)

$|f_{i,j-1}-2f_{i,j}+f_{i,j+1}|$, $i=1,n_1$, $2 \leqq j \leqq n_2-1$ (10)

で測り, fのなめらかさの程度を $\|dDf\|^2$ で測る。Dは構造のある(帯行列)スパースな行列である。各行の順はどのように取ってもかまわないがGivens変換法による数値計算の効率を上げるために図1.3,図1.4のようにとる(2節参照)。

$$D = \begin{bmatrix} 1 & -2 & 1 & & & \\ & 1 & -2 & 1 & & \\ & & 1 & -2 & 1 & \\ & & & 1 & -2 & 1 \end{bmatrix}$$

図1.3 一次元モデル(n=7)

$$D = \begin{bmatrix}
1 & -2 & 1 & & & & & & & & & & & & & & & & & \\
 & 1 & -2 & 1 & & & & & & & & & & & & & & & & \\
 & & & & -2 & & & 1 & & & & & & & & & & & & \\
1 & & & & 1 & -4 & 1 & & & 1 & & & & & & & & & & \\
 & 1 & & & & 1 & -4 & 1 & & & 1 & & & & & & & & & \\
 & & 1 & & & & & -2 & & & & 1 & & & & & & & & \\
 & & & & 1 & & & & -2 & & & & 1 & & & & & & & \\
 & & & & & 1 & & & 1 & -4 & 1 & & & 1 & & & & & & \\
 & & & & & & 1 & & & 1 & -4 & 1 & & & 1 & & & & & \\
 & & & & & & & 1 & & & & -2 & & & & 1 & & & & \\
 & & & & & & & & 1 & & & & -2 & & & & 1 & & & \\
 & & & & & & & & & 1 & & & 1 & -4 & 1 & & & 1 & & \\
 & & & & & & & & & & 1 & & & 1 & -4 & 1 & & & 1 & \\
 & & & & & & & & & & & 1 & & & & -2 & & & & 1 \\
 & & & & & & & & & & & & & & & & 1 & -2 & 1 & \\
 & & & & & & & & & & & & & & & & & 1 & -2 & 1
\end{bmatrix}$$

図1.4 二次元モデル $(n_1 \times n_2) = 4 \times 5$

ここで，超パラメタ（パラメタの分布のパラメタ）$d \in R^+$ を導入し，f のimproperな事前分布としてfのなめらかさを

$$\pi(f|d) = \left(\frac{1}{2\pi}\right)^{\frac{\ell}{2}} d^{\ell} \psi \exp\left(-\frac{1}{2}\| dDf \|^2\right) \tag{11}$$

と表現する。一次元モデルの場合 $\ell = n-2$，二次元モデルの場合 $\ell = n-4$ で，ψ は D^tD の非零固有値の積である。ここではimproperな事前分布を用いているが，これはある種のproperな事前分布を用いる方法の近似と考えられる（補遺参照）。

よって，データ $\bar{y}$ が与えられた時の周辺尤度を $L(\sigma^2, d)$ とすれば，

$$L(\sigma^2, d) = \int p(\bar{y}|f, \sigma^2)\pi(f|d)df \tag{12}$$

となる。

赤池はベイズ型情報規準量ABIC（赤池［1，2，3］）

$$ABIC = -2\log L(\sigma^2, d) \tag{13}$$

を最小にする σ^2，d をfの推定に用いる方法を提案している。つまり，ABICを最小にするσ^2，d を $\hat{\sigma}^2$，$\hat{d}$（最尤推定値）とすると，ベイズの定理により，f の事後分布は，

$$p(f|\bar{y}) = \frac{p(\bar{y}|f, \hat{\sigma}^2)\pi(f|\hat{d})}{\int p(\bar{y}|f, \hat{\sigma}^2)\pi(f|\hat{d})df} \tag{14}$$

と表されるから，(14) 式を最大にするfを最適近似関数$\hat{f}$として選ぶことにするのである。

$\alpha = d\sigma$と超パラメタの変数変換を行うと，

$$L(\sigma^2,\alpha)=\left(\frac{1}{2\pi}\right)^{\frac{N+\ell-n}{2}}\left(\frac{1}{\sigma}\right)^{\ell}\alpha^{\ell}\psi\left|\det(Z_\alpha^t Z)\right|^{-\frac{1}{2}}\exp\left(-\frac{1}{2\sigma^2}\| b-Z_\alpha f_*\|^2\right), \tag{15}$$

$$Z_\alpha=\begin{bmatrix}E\\ \alpha D\end{bmatrix},\quad b=\begin{bmatrix}y\\ 0\end{bmatrix}。$$

ここで，$\|b-Z_\alpha f_*\|^2$ は最小二乗残差で f_* はそのときの最小二乗解である。つまり，$\hat{f}$ は $\alpha=\hat{\alpha}$，$\sigma^2=\hat{\sigma}^2$ のときの f_* に他ならない。よって，問題は (15) 式を最大にする σ^2，αを求めることになる。

尤度方程式 $\frac{\partial}{\partial\sigma^2}L(\sigma^2,\alpha)\big|_{\sigma^2=\hat{\sigma}^2}=0$ より，

$$\hat{\sigma}^2=\frac{1}{N+\ell-n}\| b-Z_\alpha f_*\|^2 \tag{16}$$

これを (15) 式に代入して，

$$L'(\alpha)=\left(\frac{1}{2\pi}\right)^{\frac{N+\ell-n}{2}}\psi\exp\left(-\frac{N+\ell-n}{2}\right)\alpha^{\ell}\left|\det(Z_\alpha^t Z_\alpha)\right|^{-\frac{1}{2}} \tag{17}$$

$$\times\left(\frac{1}{N+\ell-n}\| b-Z_\alpha f_*\|^2\right)^{-\frac{N+\ell-n}{2}}$$

となる。よって，αについて，その尤度方程式

$$\frac{\partial}{\partial\alpha}L'(\alpha)\big|_{\alpha=\hat{\alpha}}=0 \tag{18}$$

を解けばよいのであるが，(18) 式は非線形方程式で解析的に解くことが困難である。そこでαを数値的に求める。いろいろな値αに対してABICを用いて，

$$\mathrm{ABIC}(\alpha)=\log\left|\det(Z_\alpha^t Z_\alpha)\right| - 2(n-2)\log\alpha$$
$$+(N+\ell-n)\log\left(\| b-Z_\alpha f_*\|^2\right)+C \tag{19}$$

を計算し，このABIC最小になるようなαを求め，このαに対応するfを推定関数にする。(19) 式の第一項を行列式項，第三項を残差項と呼ぶことにする。

2．Givens変換法

(19) 式を解くのは，次の最小二乗問題

$$\min_{f} \| b - Z_\alpha f \|^2 \tag{20}$$

を解けばよい。各αに対して毎回 (20) 式を解くのでその解法は高速なことが必要である。最小二乗問題を解く場合，ふつうHouseholder変換法が用いられるが，スパースな大行列に対してはGivens変換法の方がより有効な解法である（田中 [21] ）。ここでもGivens変換法を用いることにより，行列 Z_α の特殊な構造に着目して fill-in（零要素が非零要素になること）をおさえ，Householder変換法より計算量を減らすことができることを示そう。

(20) 式を解くにはまず Z_α をＱＲと分解する。Ｑは直交行列で，Ｒは (r_{ij}) とすれば i ＞ j の部分がすべて零である行列である。$\| Q \| = \| Q^t \| = 1$ に注意すると，

$$\begin{aligned} &= \| b - Z_\alpha f \|^2 \\ &= \| Q^t b - Rf \|^2 \\ &= \left\| \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} - \begin{bmatrix} R_1 \\ 0 \end{bmatrix} [f] \right\|^2 \end{aligned} \tag{21}$$

であるから，

$$b_1 = R_1 f \tag{22}$$

となり，後退代入よりｆが求まる。そのとき残差は$\| b_2 \|^2$となり (19) 式の残差項が求まる。　また (19) 式の行列式項は，

$$\log |\det Z_\alpha^t Z_\alpha| = \log |\det R^t R| = \log \| r_{ii} \|^2 = 2\Sigma \log r_{ii} \tag{23}$$

となり求めることができる。よって，ABIC (α) の値は Z_α をＱＲと分解することによって求めることができるわけである。このとき，変換行列Ｑとして二次元の回転行列を用いるのが，Givens変換法である。

Z_{ik} を零にする変換行列 Q_{ik} を図2.1 のように構成する。ここで，

$$r = \sqrt{z_{kk}^2 + z_{ik}^2}, \quad c_{ik} = z_{kk}/r, \quad s_{ik} = z_{ik}/r \quad . \tag{24}$$

$$Q_{ik} = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & c_{ik} & & s_{ik} & \\ & & & \ddots & & \\ & & -s_{ik} & & c_{ik} & \\ & & & & & \ddots \\ & & & & & & 1 \end{bmatrix}$$

図2.1 Givens変換行列

Q_{ik}を Z_α の左から作用させると，k行とi行のみが，

$$z_{kj} = c_{ik}z_{kj} + s_{ik}z_{ij}, \quad k \leqq j \leqq n, \tag{25}$$

$$z_{ij} = -s_{ik}z_{kj} + c_{ik}z_{ij}$$

と変形される。 Z_αの構造からfill-in の起こる場所があらかじめ分り，なるべくfill-in を押さえるようにピボットを選択することができる（このことがDを図1.3,図 1.4とした理由である）。

fill-in を押さえるために $\|\bar{y} - Ef\|$ を $\|E^t\bar{y} - E^tEf\|$ と変形して考える。E^tE は対角要素が0か1（$E^t\bar{y}$の要素にデータがあるかないかに対応する）で，非対角要素がすべて0となるような行列になる。

そして，Dの要素を上から一行ごとに生成し（Dを記憶しておく必要がないことに注意），Givens変換によりその非零要素を消去する。このときDの非零要素の長さ（帯幅）以上にはfill-in は起こらない。その理由は Z_α の非零要素が E^tEの対角成分にしかなく，Dを図1.3,図 1.4のようにとっているからである。よって，QR分解に必要な記憶容量は最終形がRとなる図2.2 の斜線の部分のみである。

なお，Householder変換法を用いた場合は点線内の部分すべてにfill-in がおこる。

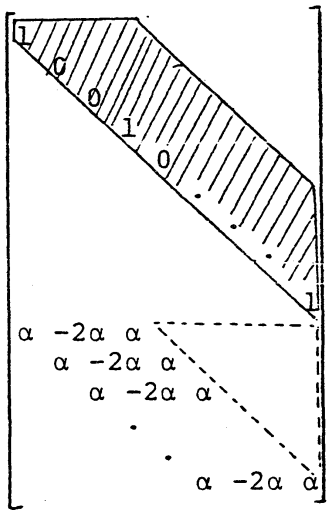

図2.2 QR分解時の Z_α

Givens変換を用いたQR分解のときに必要な計算量，記憶容量を表2.3 に掲げる。

表2.3 Givens分解法を用いたQR分解時に必要な計算量と記憶容量

	記憶容量	計算量	
		平方根計算	乗除算
一次元モデル	3n	3n	24n
二次元モデル	$2n_1^2n_2$	$2n_1^2n_2$	$8n_1^3n_2$

最大次数のみを示した

二次元モデルの場合

$n=n_1 \times n_2,\ n_1 \leqq n_2$

fを求めるアルゴリズムは次のようになる。

アルゴリズム1.

ステップ1：αを定める。

ステップ2：Z_αをGivens変換でQRと分解する。

ステップ3：ABIC最小ならステップ5へ。

ステップ4：αを変更してステップ2へもどる。

ステップ5：$R_1 f=b_1$　を計算してfを求める。

ただし，実際の計算上ではABIC（α）は非線形で凸性の保証がないので（実験的には，"ほとんど"凸である）大域的にいくつかの点を選んでABIC（α）を計算し，ある程度区間を狭めてから黄金分割による区間縮小法を用いる。ただし，αは必ずしも厳密に求める必要がないことを注意しておく。

3. 離散スプライン

この節では，f の持つ性質について述べる。簡単のためにまず一次元問題について考える。（この節ではデータのあるfit-point のことをdata-pointと書くことにする。）本稿のモデルは（1）式で言いたように，

$$\min_f \left(\|\bar{y} - Ef\|^2 + \alpha^2 \|Df\|^2 \right) \tag{26}$$

という形になっている。つまり，次のような連立方程式を解くことと同値である。

$$\alpha^2 D^t D f = E^t E f - E^t \bar{y}\,。 \tag{27}$$

(27) の各方程式は x_j がdata-pointの場合とfit-point の場合の二種類の方程式に分けることができる（ただし，両端二点ずつは除く）。

① (data-pointの場合)　$\nabla^4 f_j = \frac{1}{\alpha^2}(f_j - \bar{y}_i)$,　(28)

② (fit-point の場合)　$\nabla^4 f_j = 0$,　(29)

ここで

$$\nabla^4 f_j = f_{j-2} - 4f_{j-1} + 6f_j - 4f_{j+1} + f_{j+2}。 \tag{30}$$

②の場合，f_j においてfの四回差分が零ということは，f_{j-2}，f_{j-1}，f_j，f_{j+1}，f_{j+2} が同じ三次式に乗っていることを示している。また，$\nabla^4 f_j = 0$ ，$\nabla^4 f_{j+1} = 0$ であれば，二つの三次式は4点を共有しその4点で三次式が決定するので，f_{j-2} から f_{j+3} までが同一の三次式上の点となる。つまり，fは隣りあうdata-point間とそれらのdata-pointの外側のfit-point までの間がひとつの三次式にのっていることになる（図3のf_{j-1} からf_{j+4} までの点）。

data-point f_{j-4} と f_j 間とdata-point f_j とf_{j+3} 間の二つの三次式はf_{j-1}，f_j，f_{j+1} で交わっているが，その微係数は一致していない（ただし，f_j でのみ二階微係数が一致している）。data-pointの両端にあるfit-point がdata-pointに近ければ近い程 はなめらかな関数を表現しうることになる。これらのことからこのモデルfを三次の『離散スプライン』と呼ぶことにする。fは連続型モデル（三次スプライン関数）をサンプリングしたものではない。(28～30) 式はαの値にかかわりなく成り立つことに注意しておく。

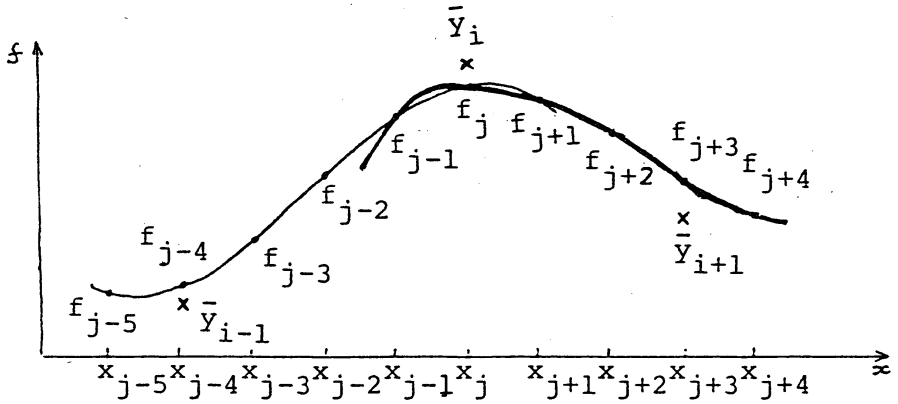

図3　離散スプラインの概念図

data-point間にfit-point が少なくとも一点あれば離散スプラインを表現できる。fit-point を増やした場合は，またデータ間で新しい三次式が生成される。これはfit-point を増やす前の三次式とは一致するとはかぎらない。

離散スプラインを用いる時の注意として，両端の影響をできるだけ少なく，かつデータのfに与える影響の仕方をすべて同等にするために最左端のdata-pointの左に二点のfit-point ，最右端のdata-pointの右に二点のfit-point をつねに取る必要がある。またfit-point 間の内挿が要求される場合にはdata-point間を三次式とすることが自然であろう。

離散スプラインが三次式になるのはDとして二階差分を用いているからである。Dとしてn階差分を用いれば離散スプラインは2n－1次式となることは (27) 式より容易にわかる。

二次元モデルの場合，一次元モデルの (30) 式に対応する式は，

$$\begin{aligned}\nabla^4 f_{i,j} = {} & f_{i-2,j}+2f_{i-1,j-1}-8f_{i-1,j}+2f_{i-1,j+1} \\ & +f_{i,j-2}-8f_{i,j-1}+20f_{i,j}-8f_{i,j+1}+f_{i,j+2} \\ & +2f_{i+1,j-1}-8f_{i+1,j}+2f_{i+1,j+1}+f_{i+2,j}\end{aligned} \tag{31}$$

となる。二次元モデルでは一次元モデルのときのように隣りの点を中心としたものと同じ曲面にはならない。また，一次元モデルのとき両端に二つずつfit-point をとったように二次元モデルでは二重のfit-point の枠をとる必要がある。

4．適用例

4．1　一次元問題

一次元の問題にアルゴリズム1を適用した例をみる。

図4.1aはビール箱の市場残存率を示す統計データに離散スプラインをあてはめた例である。fit-point は，データ数の10倍取ってある。図4.1b，図4.1cはそれぞれ図4.1aの一階差分，二階差分を表したものである。二階差分のグラフが折れ線グラフになっていることからも前節の議論の正当性が示されている。

図4.2 は同じデータに対してfit-point の数nを変化させてみたものである。ABICを用いるとfit-point の数nが多くなるに従ってαの値を大きくしてfのなめらかさを調整していることが観察される。

図4.3 はαの値を人為的に定めて，αによる離散スプラインの挙動を調べたものである。(26) 式からもわかるように，αが最適値$\hat{\alpha}$より大きいと近似関数 fのなめらかさを強くするのであるからfは直線に近づく。αを十分大きくとれば，線形回帰となる (図4.3a)。

逆にαの値を小さくしていくと，この場合は先程とは逆にfのなめらかさを弱くして離散スプラインはデータの動きを敏感に追うようになってくる。十分にαの値を小さくしていくとfはデータ補間に近づいていく (図4.3b,c,d)。

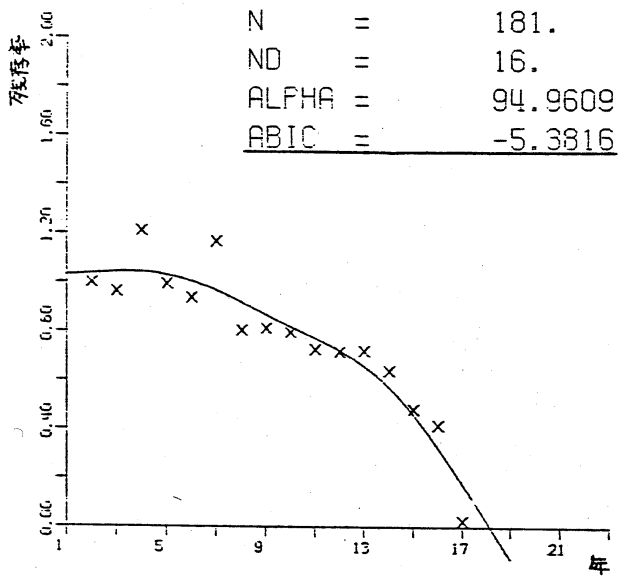

図4.1a

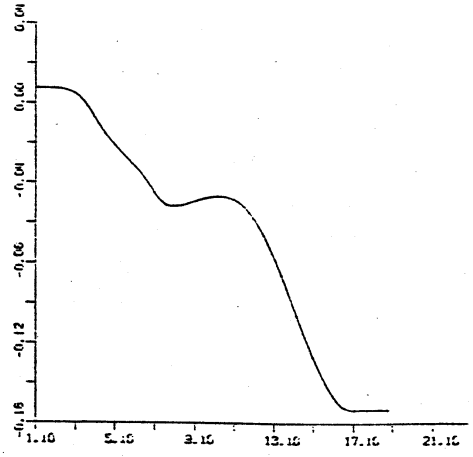

図4.1b　一階差分

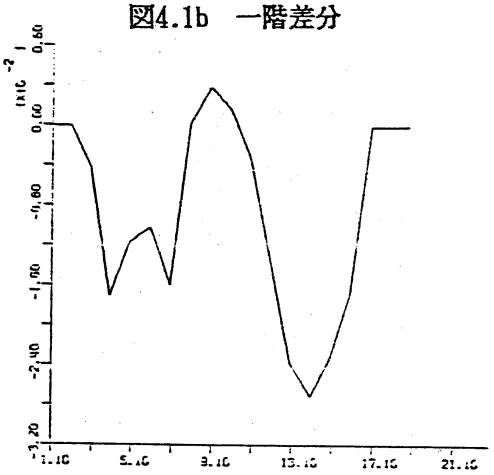

図4.1c　二階差分

図4.1 離散スプラインによるあてはめ

(ビール箱の市場残存率)

N　はfit-point 数

NDはdata-point数

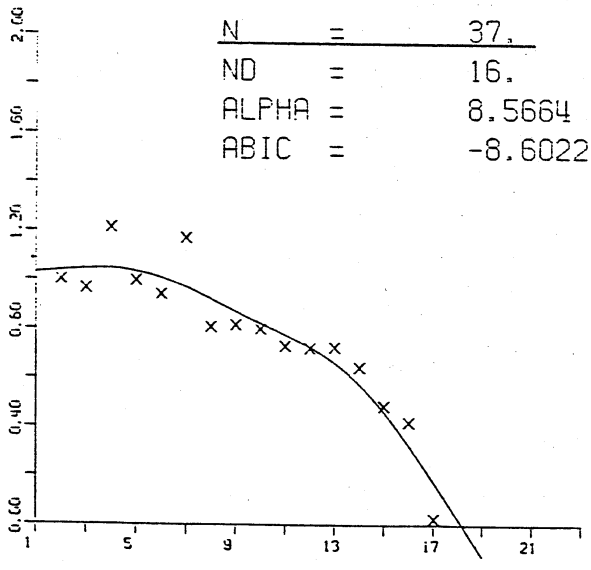

図4.2a　nを小さくした場合

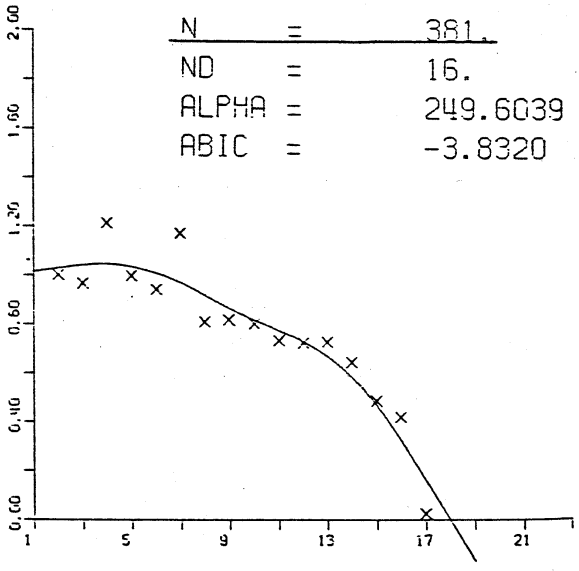

図4.2b　nを大きくした場合

図4.2　fit-point 数nによるαの変化

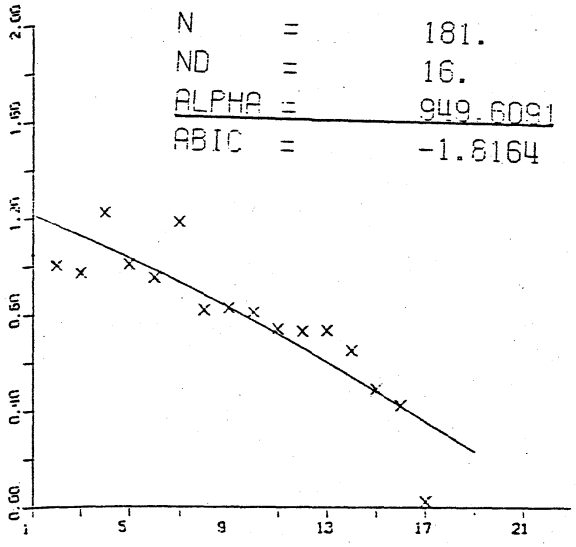

図4.3a

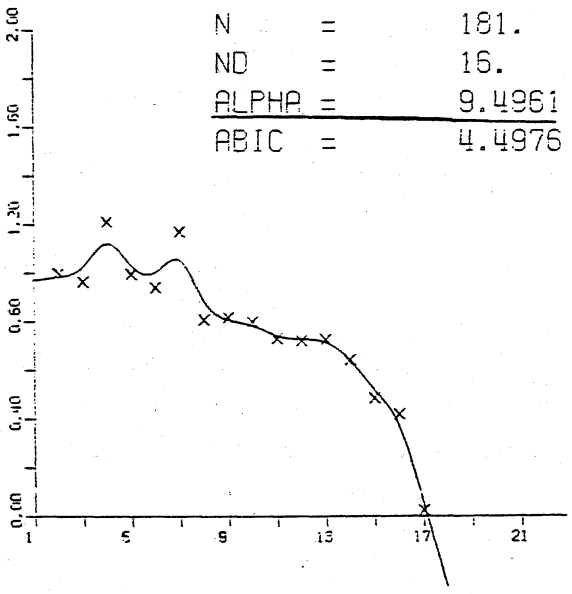

図4.3b

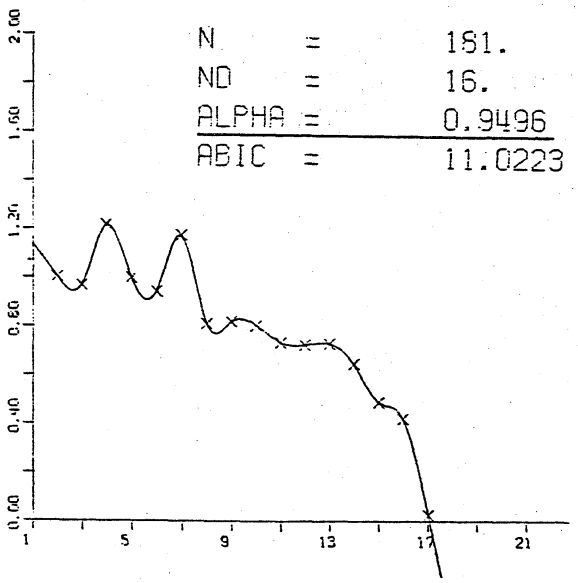

図4.3c

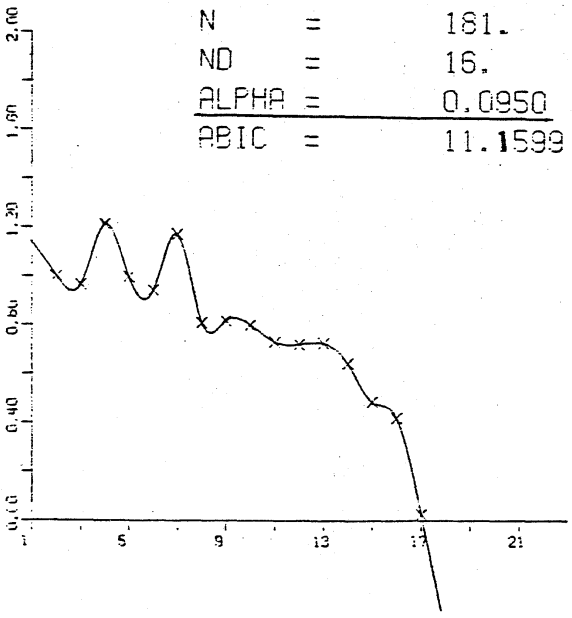

図4.3d

図4.3　αを変えた時の離散スプラインの挙動

4．2　二次元問題

一般に，一次元問題の場合とちがって二次元の問題に対してのあてはめや補間は近似関数が取りにくく難しい（Franke［6］）。しかし，本稿のモデルは一次元問題の場合とおなじアルゴリズムを用いることができる。

図4.4 に結果を示す。図4.4aは二つの山と一つの谷を持っているモデル関数である。このモデル関数を1089（33×33）の格子点上のfit-point であてはめてみる。図4.4bはモデル関数に正規誤差を加えたもので，これをデータとする。データは 289（17×17）の格子点上にある。このデータをあてはめた結果が図4.4cである。

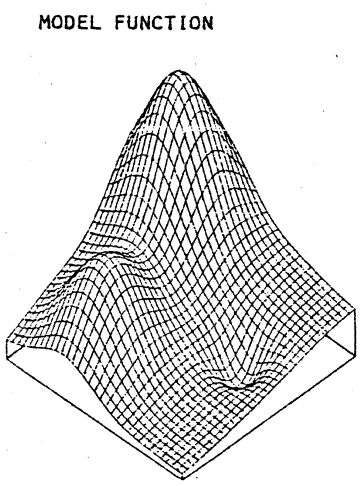

図4.4a　モデル関数

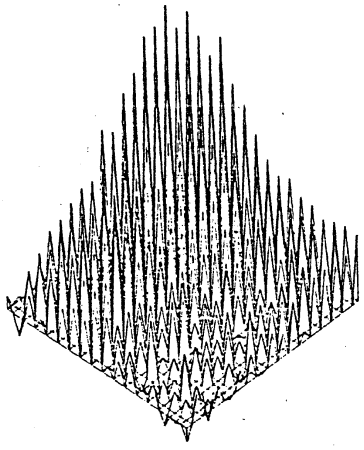

図4.4b　データ（データのない格子点上は0）

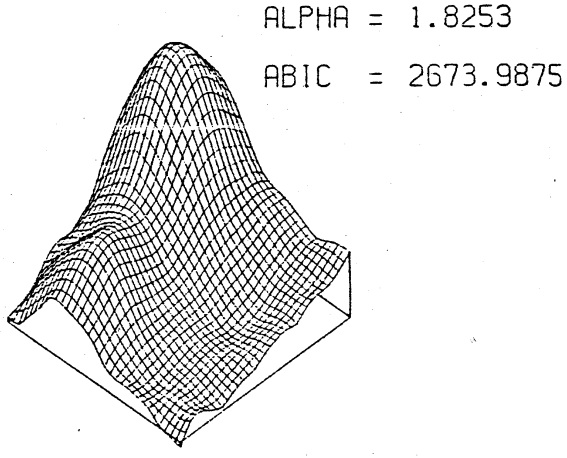

図4.4c　離散スプライン

図4.4　二次元データのあてはめ
fit-point 数1089（33×33）
data-point数 289（17×17）
すべて格子点上

次にαの値を変えてABICの値を比較した結果を図4.5 に示す。fit-point は前図と同じ格子点上ですべてのfit-point 上に誤差を乗せたデータがある。一次元の場合と同様にαが小さいと振動が激しく，αが大きくなると山や谷が消えて面が平になってしまう。 ABIC最小の時が（5つの中で），一番モデルに似ていることがわかる。

αを小さくするとデータ補間に近づく事を一次元問題でみたが，これを利用して二次元データ補間の問題として離散スプラインを用いることを考える。

図4.6 ，図4.7 はデータに誤差を乗せずにあてはめた例である。データは格子点81（9×9）の場合と25（5×5）の場合を示した。

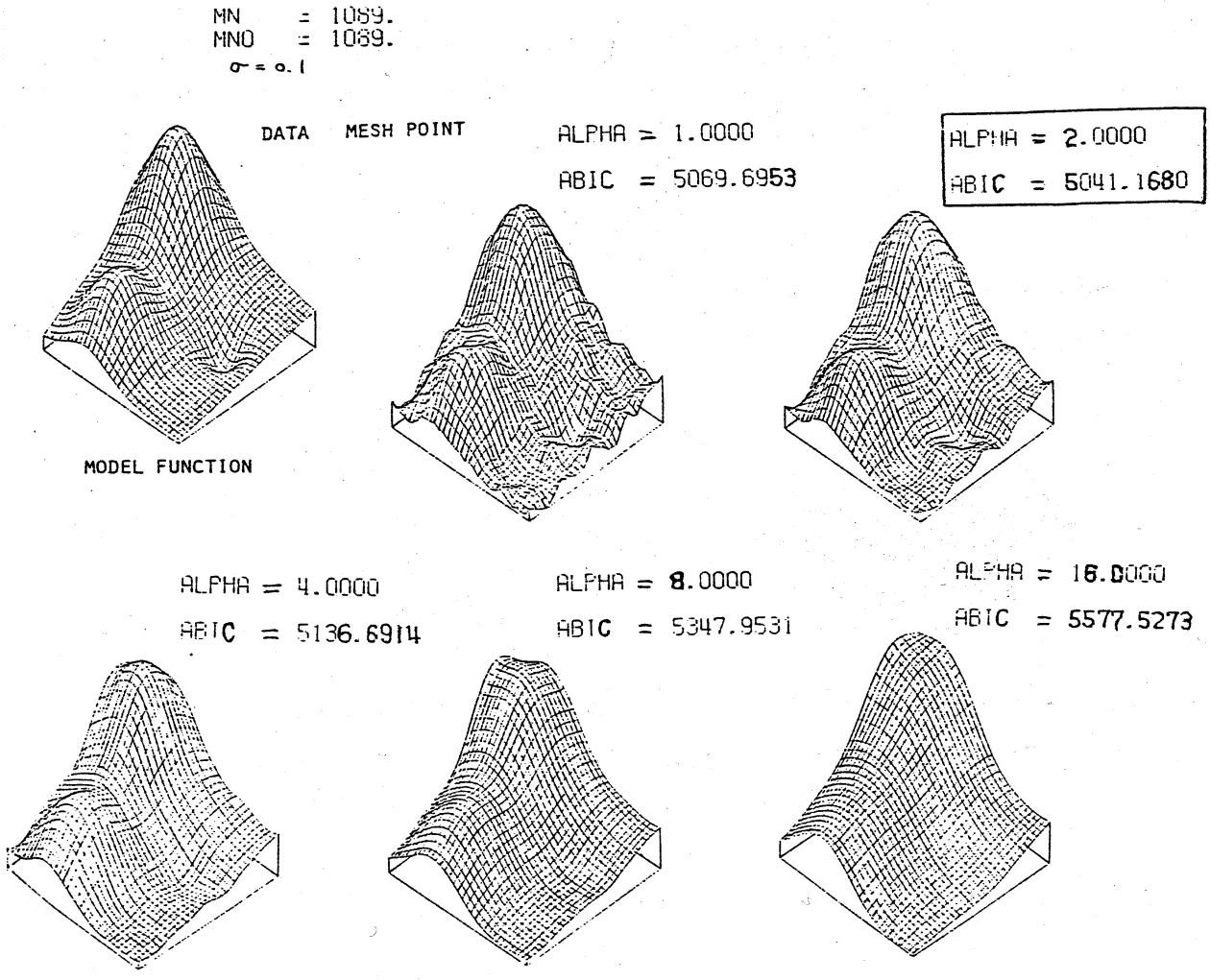

図4.5　αを変えた時の離散スプラインの挙動

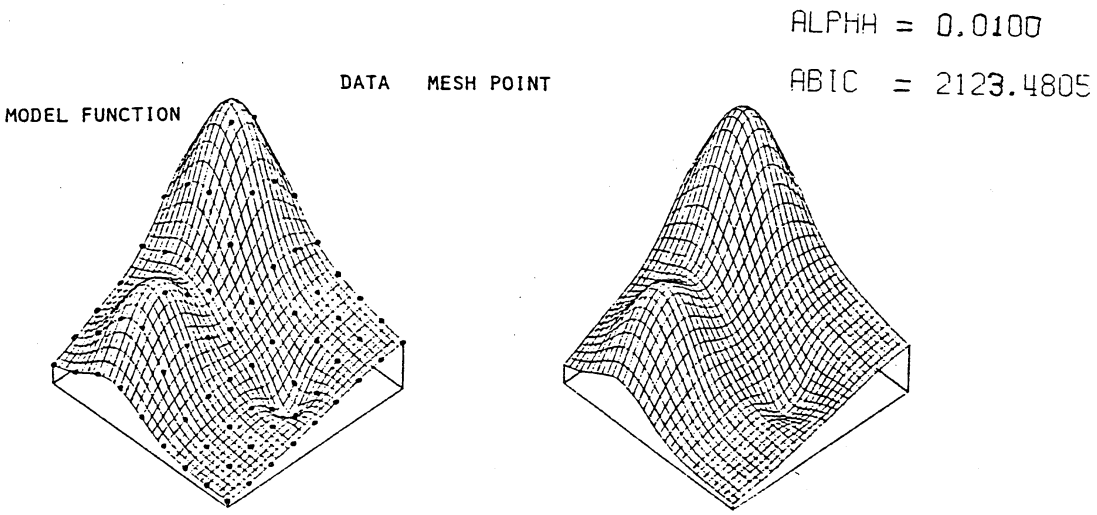

図4.6a データ（黒丸）　　図4.6b 離散スプライン

図4.6　データ補間としての離散スプライン（データ数81（9×9））

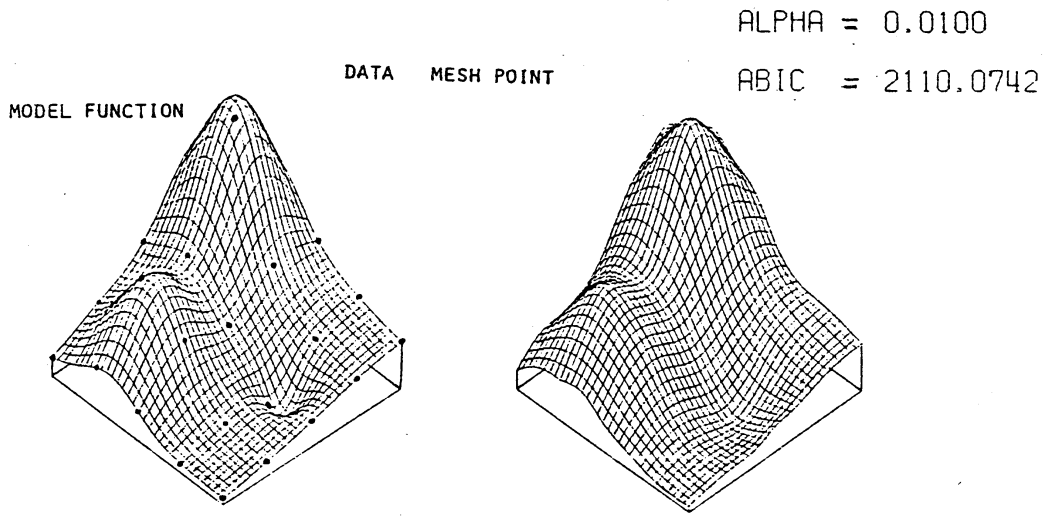

図4.7a データ（黒丸）　　図4.7b 離散スプライン

図4.7　データ補間としての離散スプライン（データ数25（5×5））

5. 制約条件付き問題

推定すべき〔構造〕に関して，非負性，単調性などがあらかじめ分かっている場合，その情報をfの制約条件として適用する。

ここでは，制約条件は一次式とする。線形の制約条件式としては，

正値性　$f_i \geq 0$,　$1 \leq i \leq n$,　(32)

単調性　$-f_i + f_{i+1} \geq 0$,　$1 \leq i \leq n-1$,　(33)

凸　性　$-f_{i-1} + 2f_i - f_{i+1} \geq 0$,　$2 \leq i \leq n-1$,　(34)

固定点　$f_i = c_i$　(35)

などが挙げられる。

一般形としては，$g_i \in R^n$, $1 \leq i \leq k$とすれば，線形制約条件付きの凸二次計画問題ＱＰ

$$(QP) \quad \left[\begin{array}{l} \min\limits_f \frac{1}{2} \| b - Z_\alpha f \|^2, \\ g_i^t f = c_i, \quad i \in I_1, \\ g_i^t f \geq c_i, \quad i \in I_2, \quad I_1 \cap I_2 = \phi, \ I_1 \cup I_2 = I \end{array} \right. \tag{36}$$

と定式化できる。これは，次のKuhn-Tucker 条件を解くことと同値である。

$$\left[\begin{array}{l} Z_\alpha^t Z_\alpha f - G^t \lambda = Z_\alpha^t b, \quad G \in R^{k \times n}, \ G = \begin{bmatrix} g_1^t \\ \vdots \\ g_k^t \end{bmatrix}, \quad (37) \\ \\ Gf \geq c, \quad (38) \\ \\ \lambda \geq 0, \quad \lambda^t (Gf - c) = 0 。 \quad (39) \end{array} \right.$$

これを解くために，有効制約戦略（active set strategy ）を用いる（今野，山下 [15] ）。

この解法の概略は次のようになる。

アルゴリズム２

ステップ1：制約条件なしの問題として，アルゴリズム1でfを求める。

ステップ2：次のような添字の集合 I_0 を考える。

$$I_0 = \{i \mid i \in I_1\} \cup \{i \mid g_i^t f \leqq c_i,\ i \in I_2\}。$$

ステップ3：k＝0とする。

ステップ4：Gの行ベクトルのうち I_k に含まれている添字のものからなる行列を$\tilde{G}$とする。

ステップ5：$\tilde{G}$に対応する λ，cを$\tilde{\lambda}$，$\tilde{c}$として，(37, 38) 式を連立させて，

$$\begin{bmatrix} Z_\alpha^t Z_\alpha & -\tilde{G}^t \\ -\tilde{G} & 0 \end{bmatrix} \begin{bmatrix} f \\ \tilde{\lambda} \end{bmatrix} = \begin{bmatrix} Z_\alpha^t b \\ -\tilde{c} \end{bmatrix} \tag{40}$$

を解く。このとき，f，λが (39) 式を満たせばfが解となる。さもなければ，ステップ6に行く。

ステップ6：$I_k = \{i \mid i \in I_1\} \cup \{i \mid g_i^t f \leqq c_i, i \in I_2\} \cup \{i \mid \lambda_i \geqq 0, i \in I_{k-1}\}$

k＝k＋1としてステップ4に戻る。

このアルゴリズムを用いた例をあげる。

(1)単調増加性

Akima [4] の問題にアルゴリズム1で離散スプラインをあてはめた図を図5.1 に示す。これに単調増加性の条件を付けてみたのが図5.2 である。図5.1bと図5.2bより，一回差分がすべて非負に変わったことがわかる。

(2)凹性

(1)と同様にLaFata [16] の問題にアルゴリズム1で離散スプラインをあてはめたものに（図5.3），凹性の条件を付けた場合を図5.4 に示す。図5.3cと図5.4cより，二回差分がすべて非正に変わったことがわかる。

(3)端点固定，単調減少性，凹性

4節で扱ったビール箱の問題に混合制約条件（端点固定，単調減少性，凹性）を付

N = 41.
ND = 10.
ALPHA = 5.4549
ABIC = 61.1163

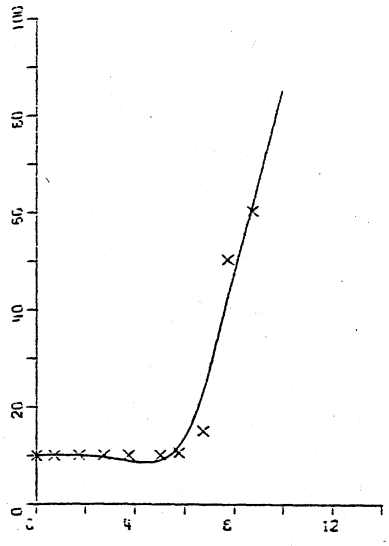

図5.1a

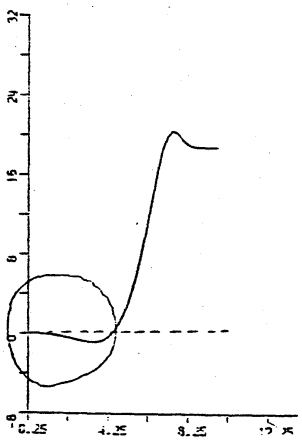

図5.1b 一階差分

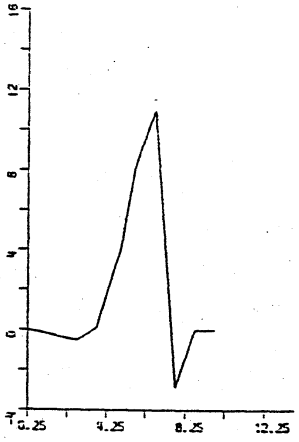

図5.1c 二階差分

図5.1 制約条件なしのあてはめ

MONOTONE

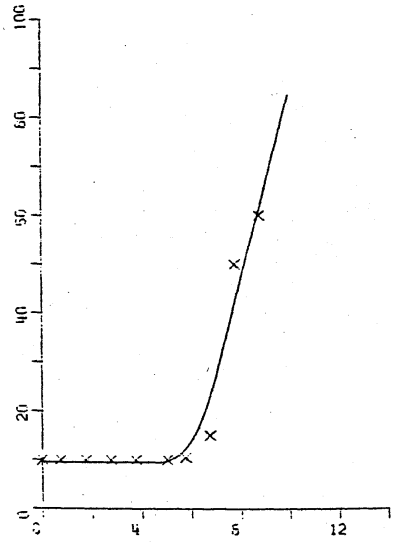

図5.2a

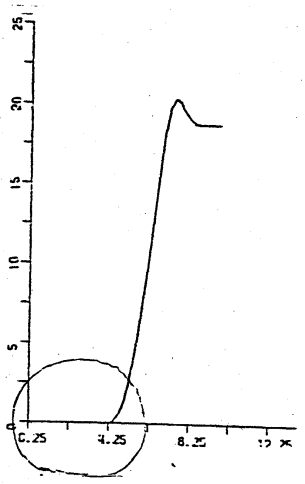

図5.2b 一階差分

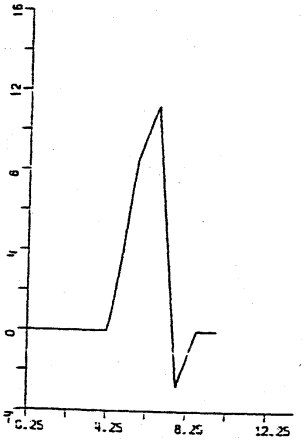

図5.2c 二階差分

図5.2 制約条件（単調増加性）付きのあてはめ

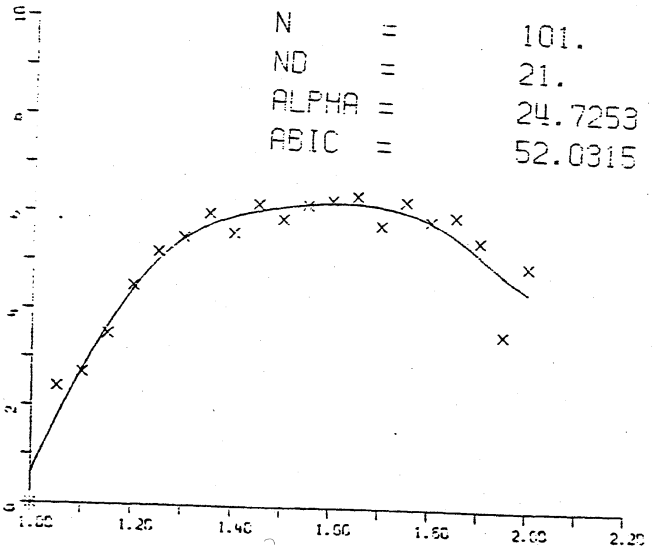

図5.3a

図5.3　制約条件なしのあてはめ

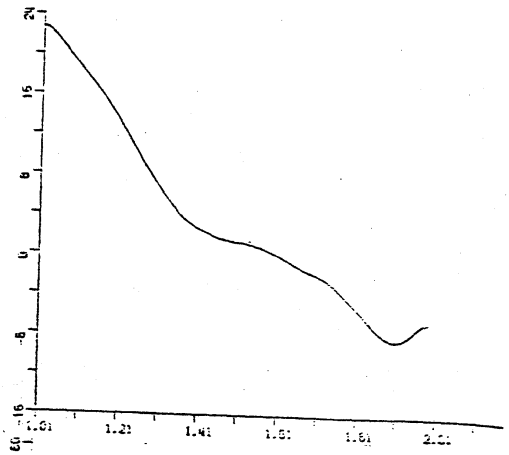

図5.3b　一階差分

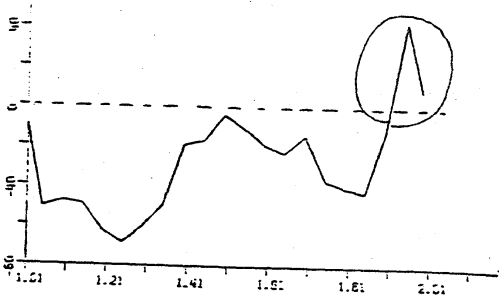

図5.3c　二階差分

CONCAVE

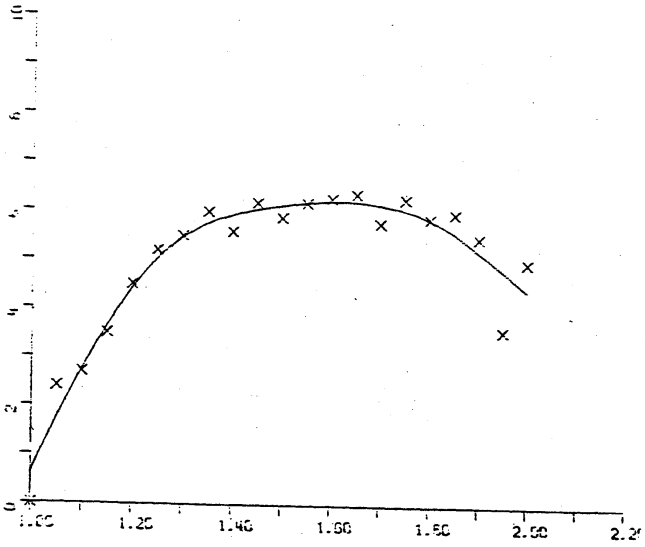

図5.4a

図5.4　制約条件（凹性）付きのあてはめ

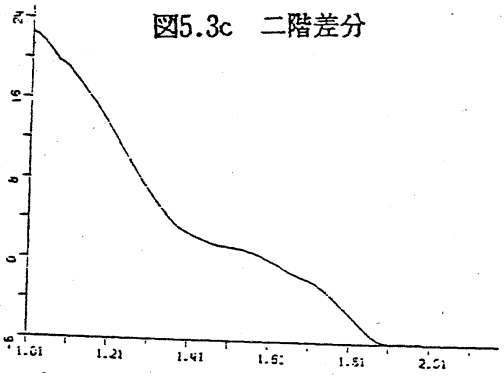

図5.4b　一階差分

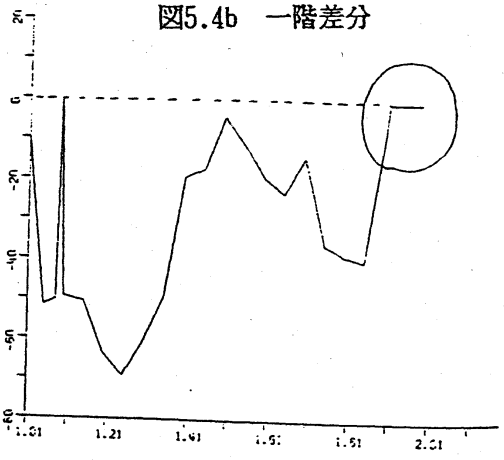

図5.4c　二階差分

けてみる。この場合の混合制約条件は次のようにすれば十分である。

$$\begin{cases} f_1 = 1, \\ f_1 - f_2 \geqq 0, \\ f_{i-1} - 2f_i + f_{i+1} \leqq 0, \quad 2 \leqq i \leqq n-1 \end{cases} \tag{41}$$

結果を図5.5 に示す。図4.1 と比べてみると(1), (2)と同様に一回差分，二階差分が条件を満たすように変化していることがわかる。

なお，これらは制約条件なしのモデルでABIC最小とするαを選び，そのαを固定してＱＰを解いたものである。よって，制約条件付きのモデルをABICで評価しているのではないことを注意しておく。

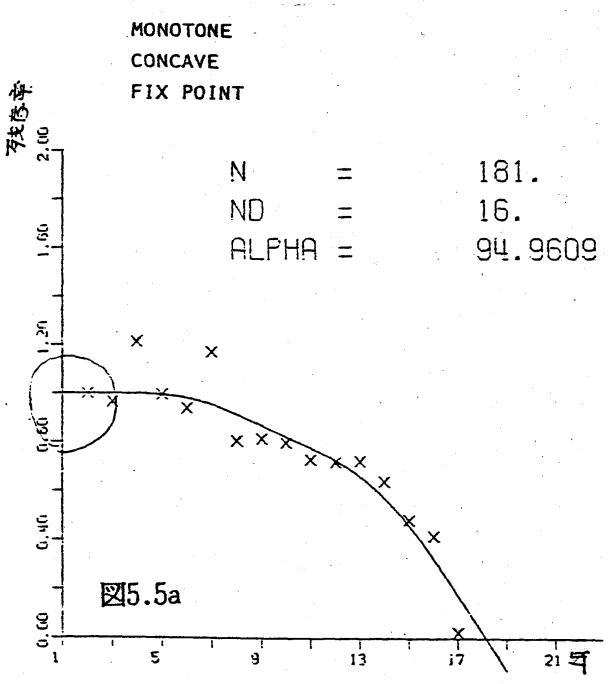

図5.5a

図5.5　制約条件付きのあてはめ

単調減少性

凹性

端点固定

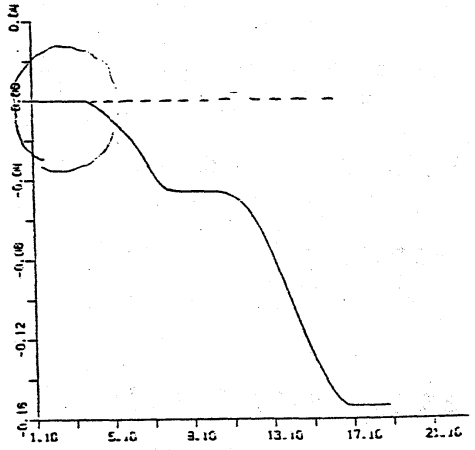

図5.5b　一階差分

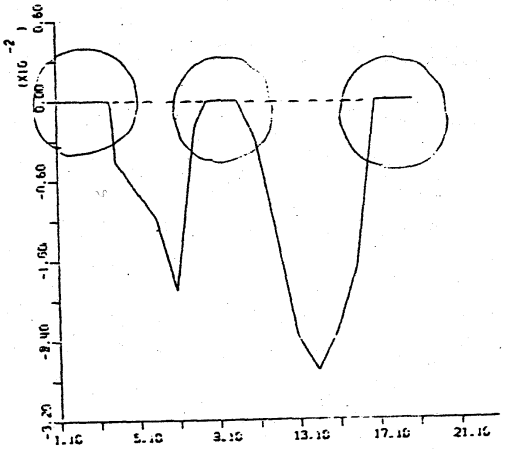

図5.5c　二階差分

6．二相問題のあてはめ（一次元問題）

あらかじめ［構造］がある点 x_p で二つの相に分けられることがわかっている問題について考察する（二相問題・two phase probrem）。ここでは二つの問題点について考える。一つは分岐点 x_p の両側をそれぞれ別の相としてあてはめること。もうひとつは x_p の位置を推定することである。

まず，前者の問題について考える。求める区間をｎ分割し，x_r で二相に分かれているとモデリングする。離散スプラインはｆのなめらかさを事前分布として扱っているが，この事前分布の仮定を次のように変更する。Ｄを三つの部分（第１行から第 r-1 行まで，第ｒ行，第 r-1 行から第ｎ－２行まで）に分離し，それぞれに α,β,γ という重みを超パラメタとして対応させる。１節での議論を繰り返すと（26）式に対応する式は，

$$\min_{f}\left(||\bar{y}-Ef||^2 + \alpha^2||D_1 f||^2 + \gamma^2||D_2 f||^2 + \beta^2||D_3 f||^2\right) \tag{42}$$

となる。このときのABICは，

$$\begin{aligned} ABIC(\alpha,\gamma,\beta,r) = (N-2)\log\left(||b-Z_{\alpha\beta\gamma}f_*||^2\right)+\log|\det(Z^t_{\alpha\beta\gamma}Z_{\alpha\beta\gamma})| \\ -2(r-1)\log\alpha - 2\log\gamma - 2(n-2-r)\log\beta + C \end{aligned} \tag{43}$$

となり，アルゴリズム１を用いてあてはめを行なうことができる。

x_p の位置の推定の問題に対してはｒをパラメタとして扱い，ｒを３からｎ－３まで動かし（端の近くを選ぶときは注意が必要），ABICで比較して最適なｒを求める。このとき x_r が分岐点として選ばれることになる。よってこの方法では x_p を厳密に求めるのでなく x_p に近い点 x_r を求めることになる。

この方法でのあてはめ例をみる。例題はポリアクリルニトリル中のｍ－ニトロアニリンの拡散係数ｄのArrhenius プロット図である。この問題はあらかじめ相が変わることがわかっている。まず，一相問題として離散スプラインであてはめたものが図6.1 である。図6.2 は二相問題としてｒを変えて各ｒに対してABIC最小の値を図示したものである。最適のｒの時の二相あてはめの結果を図6.3 に示す。

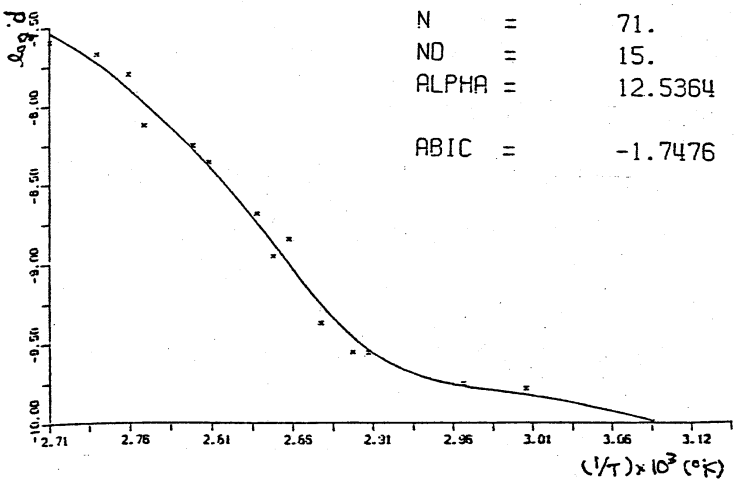

図6.1　離散スプラインよるあてはめ
(問題を一相として)

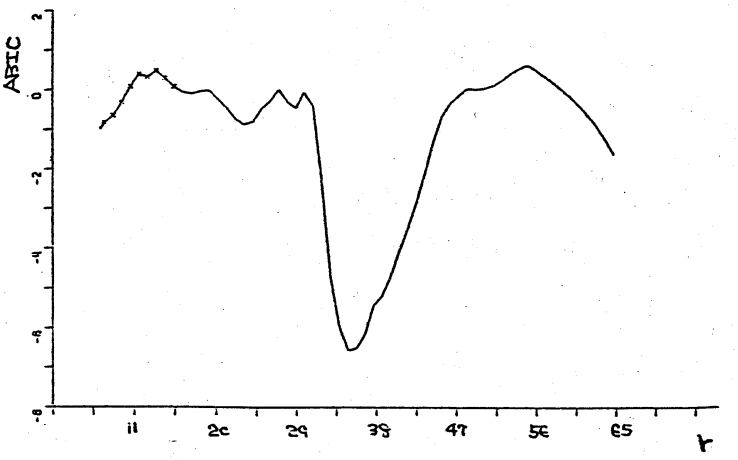

図6.2　rを変えた場合のABICの変化

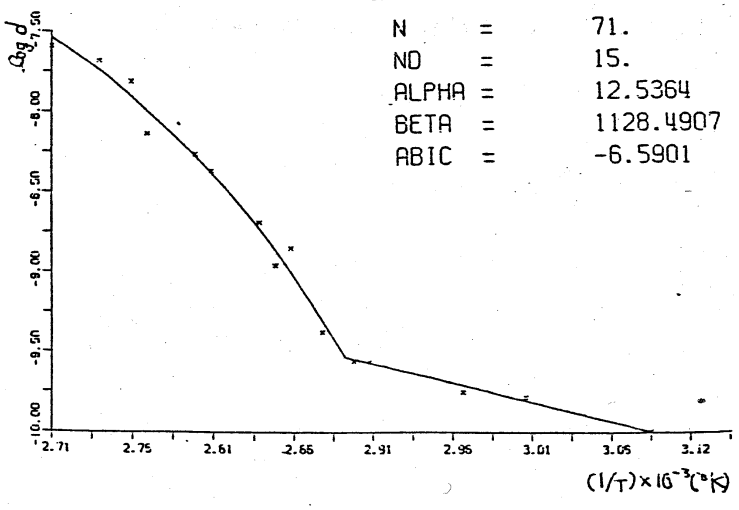

図6.3　最適のrで二相に
分けた場合のあてはめ

このモデルで x_r を決めて，α , β , γ を決定しfを求めることは，三次のスプライン関数の連続性を弱める時に節点を三重節点にし，その点で関数のみを連続にした場合と対応する。スプライン関数で節点の多重度を上げて不連続点を表現できるように，本稿の離散スプラインでも，第r行，第r+1 行を二つ続けて D_2 として (42) 式の第二項とすれば，その点で不連続な関数を表現できる。

離散スプラインの利点は相の分岐点をABICによって，評価できることである。

7. fit-point を不等間隔に取る場合（一次元問題）

データの出現がランダムに起こる問題（irreguraly spaced data probrem ）においては，fit-point を等間隔にとっては，nを十分大きくとるか，データを補正しなければ，データをfit-point 上に持ってくることができない。

そこで，fit-point x_j を等間隔にとるのではなくデータの間隔にあわせてfit-point をとる。つまり，データのある位置をfit-point にして，その間に同じ数だけのfit-point をとる。

そのためには二階差分をつぎのようにとればよい。

$$\left| -h_j f_{j-1} + (h_j+h_{j+1}) f_j - h_{j+1} f_{j+1} \right| / h_0 , \qquad (44)$$

$$h_0 = h_j h_{j+1} (h_j+h_{j+1}), \quad h_j = x_j - x_{j-1}, \quad h_{j+1} = x_{j+1} - x_j$$

(44) 式で $h_j = h_{j+1} = h$ とおき，$1/h^2$ をαに含めれば（7）式が得られる。

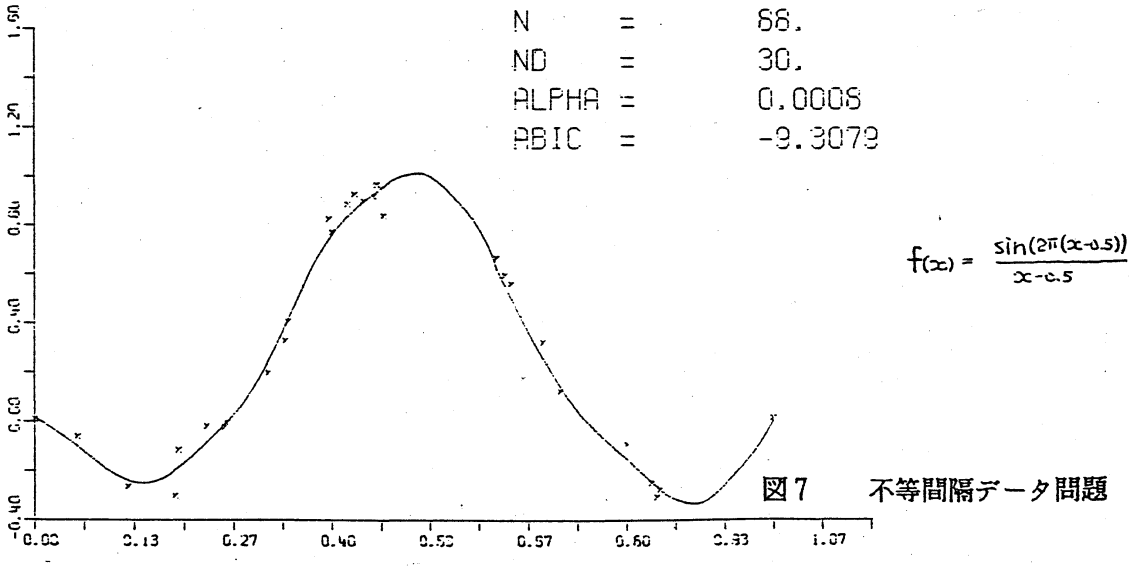

図7　不等間隔データ問題

ま と め

①三次の離散スプラインの有効性について

最大の特徴は一次元のアルゴリズムを二次元にそのまま拡張できることである。いままでのdata fittingにおいては二次元の近似関数をどのようにとるかが問題であったが，このことを解決する一つの手段となっている。離散スプラインは，ゆるやかにデータ間を結んでいるので柔軟性があり表現力が大きい。事前情報（特に不等式条件）が容易に導入できる。また，二次元問題のデータ補間として取り扱える。

②スパースで構造のある行列を持つ最小二乗問題に対するGivens法の有効性

3節において，Givens変換法がHouseholder変換法よりも計算量，記憶容量の両面で有効であることがわかった。また，このアルゴリズムがパラレル計算に向いていることは効率の面で注目すべきことである。

③情報量規準ABICの有用性

モデル選択の判断の規準としてABICがうまく働いていることがわかった。

補 遺

(11) 式のimproperな事前分布を次のようなproperな事前分布とみることができること示す（簡単のために一次元問題の場合について述べる）。

$$\tilde{\pi}(f|d,\varepsilon)=\left(\frac{1}{2\pi}\right)^{\frac{n}{2}} d^{n-2}\,\varepsilon^{2}\exp\left(-\frac{1}{2}\|dDf\|^{2}-\frac{1}{2}\|\varepsilon D' f\|^{2}\right) \tag{45}$$

このとき周辺尤度は，

$$\tilde{L}(\sigma^2,\alpha,\varepsilon)=\left(\frac{1}{2\pi}\right)^{\frac{N}{2}}\left(\frac{1}{\sigma}\right)^{n-2}\alpha^{n-2}\,\varepsilon^{2}\left|\det\left(\tilde{Z}^{t}_{\alpha\varepsilon\sigma}\tilde{Z}_{\alpha\varepsilon\sigma}\right)\right|^{-\frac{1}{2}}$$

$$\times\exp\left(-\frac{1}{2\sigma^2}\|b-\tilde{Z}_{\alpha\varepsilon\sigma}f_{*}\|^{2}\right)^{\frac{1}{2}}, \tag{46}$$

となる。

$$\tilde{Z}_{\alpha\varepsilon\sigma}=\begin{bmatrix} E \\ \alpha D \\ \varepsilon\sigma D \end{bmatrix}$$

ふつう σ^2 は有界と考えてよいから　$0<\sigma^2\leq M$　と仮定すると，

任意の $\delta_1,\delta_2>0$ に対して，一様に

$$\left|\left|\det(\tilde{Z}^{t}_{\alpha\varepsilon\sigma}\tilde{Z}_{\alpha\varepsilon\sigma})\right|^{-\frac{1}{2}} - \left|\det(Z^{t}_{\alpha}Z_{\alpha})\right|^{-\frac{1}{2}}\right| < \delta_1, \tag{47}$$

$$\left|\,\|\,b-\tilde{Z}_{\alpha\varepsilon\sigma}f_{*}\,\|^{\frac{1}{2}} - \|\,b-Z_{\alpha}f_{*}\,\|^{\frac{1}{2}}\right| < \delta_2 \tag{48}$$

となるように $\varepsilon>0$ をいくらでも小さくとることができる。よって (11) 式は (45) 式を近似したものと考え，(45) 式を用いた場合と同様に扱うことができる。

謝　辞

本文中で使用したデータは，統計数理研究所指導普及室長　野田　一雄氏，東京都公害研究所　伊藤　政志氏にいただいたものです。また，統計数理研究所第6研究部部長　赤池　弘次氏，同研究所研究員　北川　源四郎，柏木　宣久，岸野　洋久各氏には有益なコメントをいただきました。感謝致します。

参　考　文　献

[1] Akaike, H. (1980) : Likelihood and Bayes procedure, In Bayesian Statistics, J.M.Bernardo, M.H.DeGroot, D.V.Lindley and A.F.M.Smith, eds, University Press, Valencia, Spain, pp.143-166.

[2] 赤池　弘次 (1980) : 統計的推論のパラダイムの変遷について，統計数理研究所彙報，Vol.27, No.1, pp.5-12.

[3] 赤池　弘次 (1981) : モデルによってデータを測る，数理科学，No.218, pp.7-10.

[4] Akima,H. (1970) : A New Method of Interpolation and Smooth Curve Fitting Based on Local Procedures, J. Asso. Comp. Mach., vol.17, No.4, pp.589-602.

[5] Arahata,E. (1982) : A program package for drawing graphs with an X-Y PLOTTER, Computer Science Monographs, No.18, Inst. Stat. Math.

[6] Franke, R. (1979) : A critical comparison of some methods for interporation of scattered data, Naval Postgraduae School, Monterey, California, NPS-53-79-003.

[7] Ishiguro, M., Akaike, H. (1980) : Trend estimation with missing obsevations,

Ann. Inst. Statist. Math., vol.32, Part B, pp.481-488.

[8] 石黒　真木夫 (1981) ：ベイズ型季節調整モデル，数理科学，No.218，pp57-61 .

[9] 石黒　真木夫，荒畑　恵美子 (1982) ：ベイズ型スプライン回帰，統計数理研究所彙報，Vol.30，No.1.

[10] 市川　浩三，吉本　富士市 (1979) ：スプライン関数とその応用，新しい応用の数学20，教育出版.

[11] Kashiwagi,N. (1982) ：A Bayes estimation procedure for fertilities in field experiments, Research Memorandum, No.220, Inst. Statist. Math.

[12] 柏木　宣久 (1982) ：圃場試験に於ける地力の推定，統計数理研究所彙報，Vol.30，No.1，pp.1-10.

[13] 北川　源四郎 (1981) ：異状値解析ベイズモデル，数理科学，No.218，pp62-66 .

[14] Kitagawa,G., Akaike,H. (1982) ：A Quasi Baysian Approach to outlier detection, Ann. Inst. Statist. Math., vol.34, Part B, pp.389-398.

[15] 今野　浩，山下　浩 (1978) ：非線形計画法，日科技連.

[16] LaFata,P.,Rosen,J.B. (1970) ：An Interactive Display for Approximation by Linear Programming, Comm. ACM, vol.13, pp.651-659.

[17] 中川　撤，小柳　義夫 (1982) ：最小二乗法による実験データの解析　プログラムＳＡＬＳ，東京大学出版会.

[18] 中村　隆 (1982) ：ベイズ型コウホート・モデル -標準コウホート表への適用- ，統計数理研究所彙報，Vol.29，No.2，pp.77-98.

[19] 田辺　國士 (1975) ：統計データの誤差の処理，bit臨時増刊『数値計算における誤差』，pp.113-125.

[20] 田辺　國士 (1975) ：不適切問題の統計的および数値的取り扱いについて，日本数学会，応用数学分科会，講演予稿集.

[21] 田辺　國士 (1976) ：不適切問題への統計的アプローチ，数理科学，No.153，pp.1-5.

[22] 田中　輝雄 (1983) ：最小二乗問題におけるGivens法とHouseholder法について，統計数理研究所彙報，Vol.30，No.2，掲載予定.