

平面曲線の平滑化*

市田 浩三** 吉本富士市*** 清野 武**

Abstract

A method is described for fitting a plane curve given as discrete set of points. As an approximating function, we use a parametric piecewise cubic polynomial which is continuous with its first derivative. The fitting is repeated by inserting a knot one by one until a criterion of fitting is satisfied. As the criterions of fitting, AIC and the trend are used and compared.

1. まえがき

最小2乗法はデータの平滑化によく用いられるが、普通、独立変数には誤差がなく、従属変数のみに誤差が入っている場合が取り扱われている。しかし、変数のどちらにも誤差が入っていると考えられる場合もある。例えば、画像を線図形に変換し、雑音を除いて滑らかな線図形を出力したい場合などがそうである。このような平面曲線の平滑化に対しては、多項式を用いる方法^{1),2)}、つなぎ目で2次微分まで連続な3次のスプライン関数を用いる方法³⁾などが知られている。曲線が簡単な場合を除いては、区分的多項式であるスプライン関数を用いる方が柔軟性の点で優れていると思われる。ただし、区分的3次関数を用いるときは、1次微分までの連続性を仮定すれば十分であると思われる⁴⁾。区分的多項式を用いるときはそのつなぎ目(節点)の位置をどこにとるかが重要な問題である。文献3)では人間が(前もって、あるいは対話的に)適当に選ぶとしているが、計算機が自動的に節点をうまく決めることができれば処理量は飛躍的に増大する。

ここでは、1次微分まで連続な区分的3次関数を用いた平面曲線の平滑化について考察した。節点の位置については、データから微係数を推定して求めることもできるが⁴⁾、順次あてはめの悪い区間に節点を追加

していく方法を用いた。この方法は簡単であり、短時間でかなりよい結果が得られる。節点数の決定については赤池のAIC⁵⁾とPowellのtrend⁶⁾を用いる方法について実験した。これによって自動的に平面曲線の平滑化を行うことができる。

2. 区分的3次関数による平滑化

いま、 $x-y$ 平面上に順序づけされたデータ (x_j, y_j) ($j=1, 2, \dots, N$)が与えられているとする。平面曲線は一般に1価の関数 $y=f(x)$ (または $x=g(y)$)で表わせないで、次のようなパラメータ t を用いて表現する。

$$\left. \begin{aligned} t_1 &= 0, \\ t_j &= t_{j-1} + [(x_j - x_{j-1})^2 + (y_j - y_{j-1})^2]^{1/2}, \end{aligned} \right\} \quad (1)$$

$(j=2, 3, \dots, N)$.

即ち区間 $[t^{(i)}, t^{(i+1)}]$ において x と y を t の3次式

$$\left. \begin{aligned} x &= S_i(t) = m_i a_i(t) + m_{i+1} b_i(t) \\ &\quad + x^{(i)} c_i(t) + x^{(i+1)} d_i(t), \\ y &= R_i(t) = n_i a_i(t) + n_{i+1} b_i(t) \\ &\quad + y^{(i)} c_i(t) + y^{(i+1)} d_i(t) \end{aligned} \right\} \quad (2)$$

で表わす。ここで $x^{(i)} = S_i(t^{(i)})$, $y^{(i)} = R_i(t^{(i)})$, $m_i = dS_i(t)/dt|_{t=t^{(i)}}$, $n_i = dR_i(t)/dt|_{t=t^{(i)}}$,

$$\left. \begin{aligned} a_i(t) &= (t^{(i+1)} - t)^2 (t - t^{(i)}) / h_i^2, \\ b_i(t) &= -(t - t^{(i)})^2 (t^{(i+1)} - t) / h_i^2, \\ c_i(t) &= (t^{(i+1)} - t)^2 \{2(t - t^{(i)}) + h_i\} / h_i^3, \\ d_i(t) &= (t - t^{(i)})^2 \{2(t^{(i+1)} - t) + h_i\} / h_i^3 \end{aligned} \right\} \quad (3)$$

であり、 $t^{(i)}$ ($i=1, 2, \dots$)は節点で $h_i = t^{(i+1)} - t^{(i)}$ である。

平滑化のアルゴリズムは次の通りである。

* Data Fitting of Plane Curves by Kozo ICHIDA, Takeshi KIYONO (Department of Information Science, Kyoto University) and Fujitichi YOSHIMOTO (Akashi Technical College).

** 京都大学工学部情報工学科

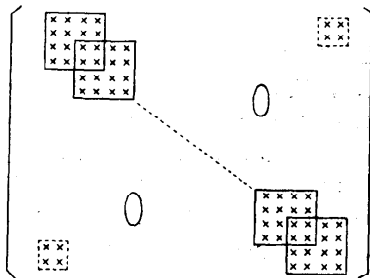
*** 明石工業高等専門学校

- ステップ 1: 最初 2 区間について最小 2 乗法の計算を行う。正規方程式を求めてステップ 4 へ行く。
- ステップ 2: 分割する区間を決定し、新しい節点を追加する。
- ステップ 3: 正規方程式を求める。
- ステップ 4: 正規方程式を解く。
- ステップ 5: 平滑化の結果をテストし、満足すれば終了。そうでなければステップ 2 へ行く。

最初の 3 つの節点は $t^{(1)}=t_1, t^{(3)}=t_N$ として $t^{(2)}$ は $[t^{(1)}, t^{(2)}]$ と $[t^{(2)}, t^{(3)}]$ ではほぼデータ数が等しいように選ぶ、以後、節点を順次追加して、節点数が K のときの正規方程式について考える。残差の 2 乗和 J は

$$\begin{aligned}
 J &= \sum_{i=1}^{K-1} \sum_{k=p_i}^{q_i} [(S_i(t_k) - x_k)^2 + (R_i(t_k) - y_k)^2] \\
 &= \sum_{i=1}^{K-1} \sum_{k=p_i}^{q_i} (S_i(t_k) - x_k)^2 + \sum_{i=1}^{K-1} \sum_{k=p_i}^{q_i} (R_i(t_k) - y_k)^2 \\
 &\equiv J_x + J_y \quad (4)
 \end{aligned}$$

と表わせる。ここで $p_i(q_i)$ は区間 $[t^{(i)}, t^{(i+1)}]$ に含まれる最初(最後)のデータの添字である。(4)式の最小値を求めるために、パラメータ $x^{(i)}, m_i, y^{(i)}, n_i$ ($i=1, 2, \dots, K$) について微分して 0 とおくと正規方程式が得られるが、この方程式では x に関する式と y に関する式は独立であるから別々に解くことができる。以後 x に対する平滑化を考えるが y についても同様である。正規方程式の係数行列は次の形をもつ。



ここで点線で囲んだ要素は平面曲線が閉曲線のときのみ表われる。新しい節点が追加されたとき、分割された区間に対する係数行列の要素のみを新たに計算すればよく、係数行列は簡単に更新できる。正規方程式の定数項についても同様である。

3. 平滑化の判定

節点を追加していくとき、いくつの節点に対する平滑化が最もよいかを判定する必要がある。ここでは次

の 2 つの判定法を用いた。

- 1) AIC^{4),5)} 各区間において残差の 2 乗和をその区間内のデータ数で割ったものが最大の区間に節点を順次追加してゆき、統計量

$$AIC = N \log_e J_x + 4K \quad (5)$$

が最小になるときの平滑化が最もよいと判定する。

- 2) trend⁶⁾ 区間 $[t^{(i)}, t^{(i+1)}]$ における残差を $r_k = S_i(t_k) - x_k$ ($k=p_i, \dots, q_i$) としたとき

$$T_i \equiv \sum_{k=p_i+1}^{q_i} r_{k-1} r_k < \sum_{k=p_i}^{q_i} r_k^2 / \sqrt{q_i - p_i} \equiv U_i \quad (6)$$

であればこの区間における x に対する平滑化は適当であると判定する。そしてすべての区間について(6)式が成立すれば終了し、そのときの節点に対する平滑化を求めるものとする。そうでなければ $T_i - U_i$ の値が最大である区間に節点を追加する。判定条件として trend を用いたときの流れ図を Fig. 1 に示す。AIC を用いたときも同様に考えられる。

4. 計算例

例 1 次の関数

$$\begin{aligned}
 x_j &= 5s_j - 11.5s_j^2 + 7.5s_j^3 + \varepsilon_j, \\
 y_j &= 2s_j - s_j^2 - 0.67s_j^3 + \eta_j, \\
 &(j=1, 2, \dots, 50)
 \end{aligned} \quad (7)$$

を用いて計算した。ここで s_j は 0.02 から 0.02 おきに 1 までとった。また ε_j, η_j は平均値 0, 分散 0.001

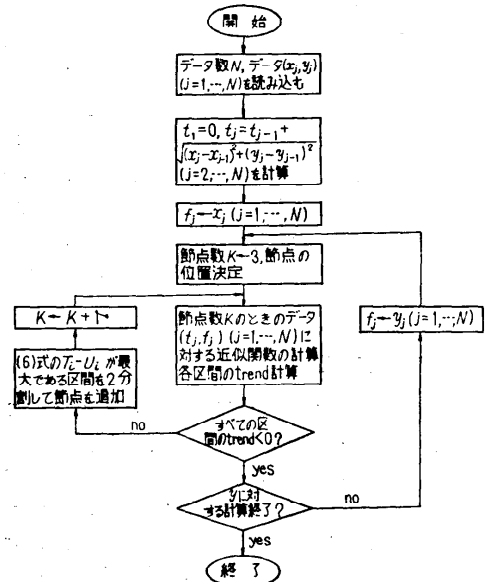


Fig. 1 Scheme of fitting using trend criterion

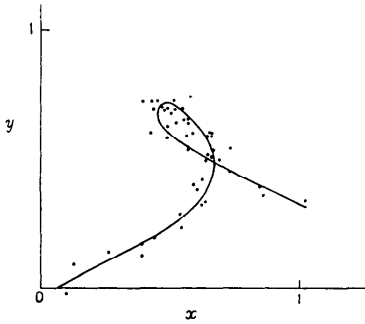


Fig. 2 Data points and the result of fitting of example 1

の正規分布をする誤差である。データ点と平滑化の結果を Fig. 2 に示す。Fig. 2 において節点数は x, y に対してともに 3 で、このとき AIC は最小である。平滑化した結果はもとの関数 ((7) 式で $\varepsilon_j = \eta_j = 0$ ($j = 1, 2, \dots, 50$) とおいたもの) とほとんど一致している、図で表わしたとき区別できなかった。

例 2 Fig. 3 にデータ (人の顔の輪郭) および平滑化した結果を示す。節点数は x, y とも 11 で、このときすべての区間で (6) 式が満足された。このときの計算時間は 37 秒であった。比較のため Forsythe⁷⁾ の直交多項式による方法で計算を行ったが (trend 判定法利用)、多くのパラメータ (x, y 合わせて 56) が必要であり、しかも結果は Fig. 3 より悪かった。なお、計算時間は 1 分 1 秒であった。

例 3 Fig. 4 にデータ (染色体) および平滑化した結果を示す。節点数は x, y とも 13 で (6) 式によって判定した。このときの計算時間は 48 秒であった。この例は t によってパラメータ表示したとき t の周期関数であるから、比較のため三角関数を用いて計算を行ったところ、ほとんど同じパラメータ数で同じ程度の結果が得られた。ただし、計算時間は 3 分 53 秒であった。これらの計算はすべて HITAC 8350 を利用して行った。

5. むすび

区分的 3 次関数を用いた平面曲線の平滑化について述べた。パラメータ表現を用いることにより開曲線、閉曲線、多価関数などを同じように取り扱うことができる。ここで用いた区分的多項式は、一般的多項式や三角関数に比べて柔軟性が優っており、正規方程式は条件の良い帯構造をしていて計算時間は短い。平滑化の判定規準としては AIC と trend を用いた。(5) 式

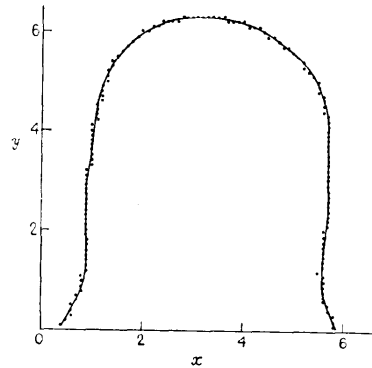


Fig. 3 Data points (outline of a human face) and the result of fitting of example 2

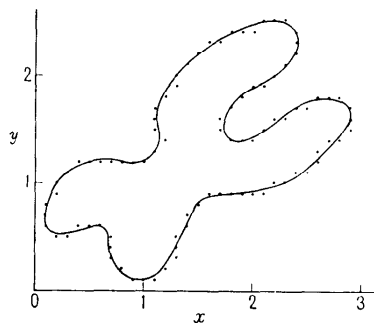


Fig. 4 Data points (chromosome) and the result of fitting of example 3

で表現された AIC を用いるときは、誤差が独立で正規分布をする、データがパラメータに比べて十分多い、などの条件が必要であり、画像から得られた線図形 (Fig. 3, Fig. 4) ではこの条件が満たされず、AIC では判定できなかった。また AIC ではかなり余分の計算をする必要がある (AIC が最小となる場合を見つけるためには多くの場合について計算する必要がある)。trend を用いれば計算量はかなり少ない。Fig. 3, Fig. 4 のようなデータのときには trend で判定する方がよいと思われるが、さらに実験が必要である。また、節点で関数や微係数が不連続な場合も考えられるので、節点での連続性を自動的に判定して平滑化を行う方法を開発することも今後の課題である。さらに、ここではデータ点は順序づけられていると仮定したが、データ点の順序づけも重要な問題である。

参考文献

- 1) M. O'Neill, I. G. Sinclair & F. J. Smith: Polynomial Curve Fitting When Abscissas and

- Ordinates Are both Subject to Error, *Comp. J.*, Vol. 12, No. 1, pp. 52~56 (1969).
- 2) M. Grossman: Parametric Curve Fitting, *Comp. J.*, Vol. 14, No. 2, pp. 169~172 (1971).
 - 3) R. E. Smith, Jr., J. M. Price & L. M. Howser: A Smoothing Algorithm Using Cubic Spline Functions, NASA Technical Note, TN D-7397 (1974).
 - 4) 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, No. 3, pp. 200~206 (1976).
 - 5) H. Akaike: A New Look at the Statistical Model Identification, *IEEE Trans. Automat. Contr.*, Vol. AC-19, No. 6, pp. 716~723 (1974).
 - 6) M. J. D. Powell: Curve Fitting by Splines in One Variable, Numerical Approximation to Functions and Data, J. G. Hayes, ed., pp. 65~83, Athlone Press, London (1970).
 - 7) G. E. Forsythe: Generation and Use of Orthogonal Polynomials for Data-fitting with a Digital Computer, *J. Soc. Indust. Appl. Math.*, Vol. 5, pp. 74~88 (1957).
- (昭和51年3月31日受付)
(昭和51年10月15日再受付)