An Automatic Method for Smoothing Two-Dimensional Data with a Piecewise Bicubic Polynomial

Fujiichi YOSHIMOTO*, Kozo ICHIDA** and Takeshi KIYONO**

Abstract

A method is described for smoothing two-dimensional data by the 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 good positions of knots is introduced. The number of knots is determined by the use of Akaike's information criterion (AIC).

1. Introduction

Two-dimensional data are treated in many fields. In general, measured data have errors and must be smoothed frequently. One of the methods for smoothing two-dimensional data is to fit an approximating function having two variables to the data by the method of least squares. This method is useful for smoothing or compressing data, for integrating or differentiating the underlying function of data and for drawing contours from data.

Some methods which use a single polynomial having two variables have been developed. However, these methods do not give a good approximation in cases where the underlying function of data has a disassociated nature. In these cases, piecewise polynomials are very useful. Hayes and Halliday [2] describe a method using bicubic spline functions. However, automatic determination of knots is not considered. This problem is very important, because goodness of fit with a spline (or a piecewise) function is considerably affected by the number of knots and their positions.

In this paper, we describe an automatic method for smoothing two-dimensional data with a piecewise bicubic polynomial which is continuous with its first derivative. The positions of knots are determined by the use of the second partial derivatives of the underlying function of data and the number of knots are determined by using Akaike's information criterion (AIC) [1]. Further details are given in the Japanese version of this paper (see footnote).

2. Representation of an Approximating Function

We assume that data are given on the mesh points in a rectangular domain $R = [a, b]\times[c, d]$ on the (x, y) plane and expressed by the equation

$$F_{tu} = f(x_t, y_u) + e_{tu}, \qquad (t = 1,2,\dots,K;\ u = 1,2,\dots,L). \qquad (1)$$

Here $f(x, y)$ denotes the underlying function of data and e_{tu}'s are mutually independent errors which follow the normal distribution with the mean zero and the unknown variance σ^2 $(< \infty)$. In this paper, $f(x, y)$ is called the "signal", (x_t, y_u)

This paper first appeared in Japanese in Joho－Shori (Journal of the Information Processing Society of Japan), Vol. 18, No. 2 (1977), pp. 128～134.

* Department of General Studies, Akashi Technical College
** Department of Information Science, Kyoto University

a sample point and F_{tu} a data value.

The domain R is subdivided into rectangular panels R_{mn} (m = 0,1,...,h; n = 0,1,...,k) and a piecewise bicubic polynomial which is continuous with its first derivative is used for an approximating function (Fig. 1). This piecewise bicubic polynomial can be expressed as

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

where $M_i(x)$ and $N_j(y)$ are normalized cubic B-splines, and γ_{ij}'s are parameters which are determined by the method of least squares [2]. The spline function (2) would, in general, have second derivative continuity: to allow discontinuities in second derivative, we must have two knots coinciding at each breakpoint of the piecewise bicubic polynomial. Let the knots for $M_i(x)$ ($i = 1,2,\ldots,2h+4$) be $\boldsymbol{\lambda} = (\lambda_{-3}, \lambda_{-2}, \ldots, \lambda_{2h+4})$, and let the knots for $N_j(y)$ ($j = 1,2,\ldots,2k+4$) be $\mu = (\mu_{-3}, \mu_{-2}, \ldots, \mu_{2k+4})$. In addition, let the breakpoints for x direction be $\boldsymbol{\xi} = (\xi_0, \xi_1, \ldots, \xi_{h+1})$ and let the breakpoints for y direction be $\boldsymbol{\zeta} = (\zeta_0, \zeta_1, \ldots, \zeta_{k+1})$, where $\xi_0 = a$, $\xi_{h+1} = b$, $\zeta_0 = c$ and $\zeta_{k+1} = d$.

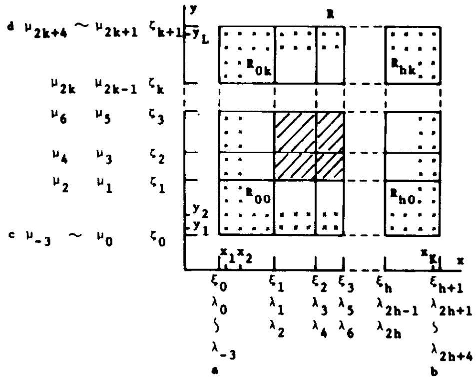

Fig. 1 The area of fitting, divided by breakpoints ξ_m, ζ_n into panels R_{mn}. Here λ_i and μ_j denote knots.

3. Good Positions of Knots

We first consider the choice of knot positions when the number of knots is given. A piecewise bicubic polynomial which is continuous with its first partial derivatives has the distinctive feature that its second partial derivatives are piecewise linear polynomials which are discontinuous at the breakpoints. Using this property, we determine good positions of knots automatically. If the second partial derivatives of the signal are sufficiently close to those of the approximating function, we can obtain a good approximation. let us consider this point.

Consider how to determine the knots $\boldsymbol{\lambda}$ to obtain a good approximation. If $y = y_u$, then data $F_u = (F_{1u}, F_{2u}, \ldots, F_{Ku})$ can be regarded as the data having one variable x. If we write as $H_m = [\xi_m, \xi_{m+1}]$ ($m = 0,1,\ldots,h$), then the square G_u of the L_2 norm of $\{f(x_t, y_u) - S(x_t, y_u)\}$ for $y = y_u$ and $a \le x_t \le b$ can be expressed as follows:

$$G_u = \sum_{m=0}^{h} \sum_{x_t \in H_m} \{f(x_t, y_u) - S(x_t, y_u)\}^2. \quad (3)$$

THEOREM.

Suppose that $f(\xi_m, y_u) = S(\xi_m, y_u)$, $\partial f(x, y_u)/\partial x|_{x=\xi_m} = \partial S(x, y_u)/\partial x|_{x=\xi_m}$ *and*

$$\left| \frac{\partial^2 f(x, y_u)}{\partial x^2} - \frac{\partial^2 S(x, y_u)}{\partial x^2} \right| \le \theta_{mu} \quad (4)$$

hold in each interval $\{H_m | m = 0,1,\dots,h\}$. *Then*

$$G_u \le \frac{1}{4} \sum_{m=0}^{h} (\theta_{mu})^2 \sum_{x_t \in H_m} (x_t - \xi_m)^4. \tag{5}$$

If the breakpoints ξ (hence the knots λ) are determined so that the right-hand side of (5) can be minimum, we may expect $S(x,y_u)$ to be a good approximating function to $f(x, y_u)$ from the theorem. However, the breakpoints ξ (hence the knots λ) minimizing (5) are functions of y_u $(u = 1,2,\dots,L)$. As is obvious from the descriptions in section 2, the knots λ must be independent of y. Therefore, we determine the positions of knots λ as follows.

First of all, we estimate the values of the second partial derivative $\partial^2 f(x, y_u)/\partial x^2|_{x=x_t}$ $(t = 1,2,\dots,K)$ from the data $\mathbf{F}_u$. Let these values be denoted by $\mathbf{g}_u = (g_{1u}, g_{2u}, \dots, g_{Ku})$. Secondly, we make an approximation to $\mathbf{g}_u$ with a piecewise linear polynomial $B_u(x)$ which is discontinuous at its breakpoints. Let the breakpoints of $B_u(x)$ be ψ_{mu} $(m = 0,1,\dots,h+1)$, where $\psi_{0u} = a$ and $\psi_{h+1,u} = b$. In the intervals $D_{mu} = [\psi_{mu}, \psi_{m+1,u})$ $(m = 0,1,\dots,h)$, suppose that

$$\max_{x_t \in D_{mu}} |g_{tu} - B_u(x_t)| = \delta_{mu}. \tag{6}$$

At this time, we make the approximation so that the values of

$$V_u = \sum_{m=0}^{h} (\delta_{mu})^2 \sum_{x_t \in D_{mu}} (x_t - \psi_{mu})^4 \equiv \sum_{m=0}^{h} \Delta_{mu} \tag{7}$$

can be almost minimum on the basis of the theorem. This approximation is computed easily, and the approximating segments can be increased one by one with an automatic manner [3].

Third, we determine the breakpoint ξ_m (hence the knots λ_{2m-1} and λ_{2m}) of the approximating function (2) using the weighted least squares approximation to ψ_{mu} $(u = 1,2,\dots,L)$ as follows. Suppose $m = 1$ (see Fig. 2). The weight w_{1u} for ψ_{1u} is determined to be the larger value between Δ_{0u} and Δ_{1u}; that is,

$$w_{1u} = \max(\Delta_{0u}, \Delta_{1u}), \quad (u = 1,2,\dots,L). \tag{8}$$

Then the breakpoint ξ_1 is taken at the value which minimizes the weighted sum

$$G_1 = \sum_{u=1}^{L} w_{1u}(\xi_1 - \psi_{1u})^2 \tag{9}$$

of the squares of the residuals $(\xi_1 - \psi_{1u})$, $u = 1,2,\dots,L$. Therefore

$$\xi_1 = \lambda_1 = \lambda_2 = \sum_{u=1}^{L} w_{1u}\psi_{1u} / \sum_{u=1}^{L} w_{1u}. \tag{10}$$

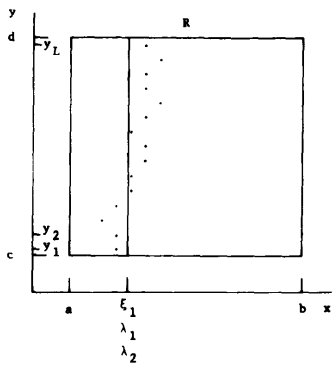

Fig. 2 Determination of the breakpoint ξ_1. The dots in the figure denote the breakpoints ψ_{1u} $(u = 1,2,\dots,L)$.

The breakpoints ξ_m (hence λ_{2m-1} and λ_{2m}) for $m = 2,3,\ldots,h$ are determined similarly.

The knots μ are determined by the same procedure.

4. Least Squares Fitting With a Piecewise Bicubic Polynomial

Equation (2) is fitted to the data (1) with the method of least squares. We get an approximating function for the various pairs of h and k using the positions of knots obtained by the method in section 3, and decide the adequate number of knots by use of AIC [1].

Let us consider the following regression model from (2):

$$F_{tu} = S(x_t, y_u) + e_{tu}, \qquad (t = 1,2,\ldots,K;\ u = 1,2,\ldots,L). \qquad (11)$$

If AIC is applied to this model as a criterion for fitting, then

$$AIC = KL \log_e Q + 2(2h+4)(2k+4), \qquad (12)$$

where Q is the sum of the squares of residuals and (2h+4)(2k+4) is the number of parameters of S(x, y) [3].

We continue the calculation of fitting by increasing the number of knots and choose the model which minimizes AIC. As the result of the calculation, a good approximating function is given by the minimum AIC model.

An outline of the algorithm of fitting is shown in Fig. 3. AIC is not always a unimodal function of the parameter. However, because of the fact that the best approximating function is of the minimum value of AIC, we can determine a satisfactory approximating function (in other words, good knots) automatically.

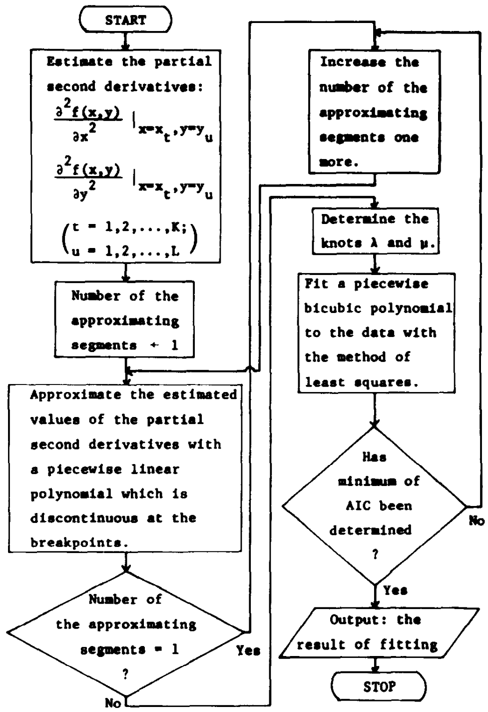

Fig. 3 A flow chart for the two-dimensional data fitting with a piecewise bicubic polynomial.

5. Numerical Example

We use the following data

$$F_{tu} = \frac{1}{0.01 + 2(x_t - 0.4)^2} + \frac{1}{0.02 + 2(y_u - 0.2)^2} + e_{tu},$$

$$(x_t\!: 0.0,\ 0.02,\ \ldots,\ 1.0;\ y_u\!: 0.0,\ 0.02,\ \ldots,\ 1.0), \qquad (13)$$

where e_{tu}'s are mutually independent errors which follow the normal distribution with the mean zero and the variance 1. The domain of fitting is determined as $[a, b]\times[c, d] = [0, 1]\times[0, 1]$.

Table 1 shows the computed values of AIC. In the case of the proposed method, AIC attains a minimum value at the point where the number of the parameters is equal to 196. The contour lines drawn from the result of fitting in this case is shown in

Fig. 4. In the case of the fitting with equal intervals of the breakpoints ξ and ζ, the values of AIC are larger than those of AIC of the proposed method. Fig. 5 shows the contour lines drawn from the result of fitting with the equal intervals to the same number of the parameters as Fig. 4. Fig. 6 shows the contour lines drawn from the original data. Fig. 7 shows the contour lines drawn from the signal. Fig. 4 is very similar to Fig. 7.

Table 1 The calculated values of AIC

The number of parameters (2h+4)(2k+4)	proposed method	equal intervals
36	32,943	32,810
64	28,830	30,989
100	23,350	30,529
144	21,392	29,175
196	20,738	26,848
256	20,837	22,596
324	20,827	21,632
400	20,850	22,784
484	20,952	22,074

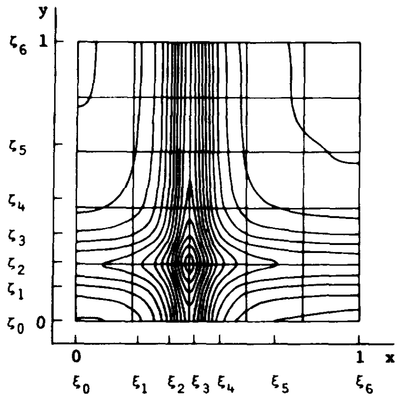

Fig. 4 The contour lines drawn from the result of fitting by use of the proposed method.

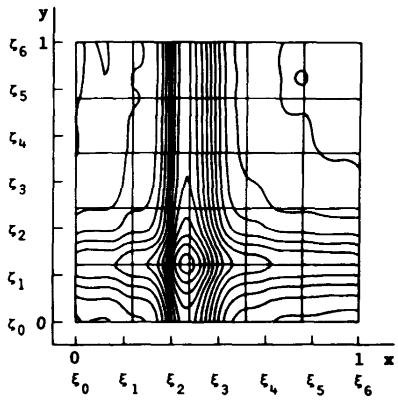

Fig. 5 The contour lines drawn from the result of fitting with equal intervals of the breakpoints.

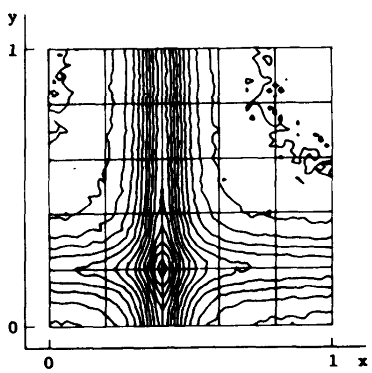

Fig. 6 The contour lines drawn from the original data.

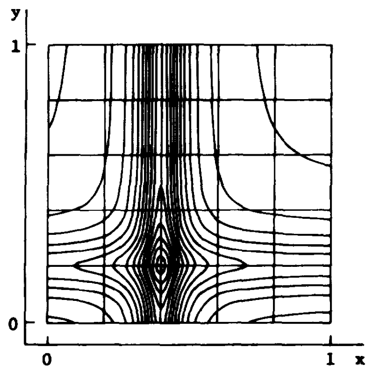

Fig. 7 The contour lines drawn from the signal.

References

[1] Akaike, H. A new look at the statistical model identification. IEEE Trans. Automat. Contr. AC-19 (1974), 716-723.

[2] Hayes, J.G., and Halliday, J. The least-squares fitting of cubic spline surfaces to general data sets. J. Inst. Math. Appl. 14 (1974), 89-103.

[3] Yoshimoto, F. Studies on data fitting with spline functions, Doctoral dissertation, Kyoto University, March, 1977.