International Symposium on Nonlinear Theory and its Applications
Xi'an, PRC, October 7–11, 2002

Endoharmonious Weighted Least-Square Polynomial Regression in the Two-Dimensional Space

Takamitsu ARAI† and Masao IRI‡

† Department of Information and System Engineering
Faculty of Science and Engineering, Chuo University
1-13-27 Kasuga, Bunkyo-ku, Tokyo, 112-8551, Japan
Phone: +81-3-3817-1680, FAX:+81-3-3817-1681
Email: taarai@iri-lab.ise.chuo-u.ac.jp

‡ Department of Information and System Engineering
Faculty of Science and Engineering, Chuo University
1-13-27 Kasuga, Bunkyo-ku, Tokyo, 112-8551, Japan
Phone:+81-3-3817-1690, Fax:+81-3-3817-1681
Email: iri@ise.chuo-u.ac.jp

1. Introduction

When we try to extract a reliable mathematical model from a series of noisy data, the method of least squares is the typical useful tool. Sometimes, we would like to focus our attention in a neighbourhood of a "*point of interest*" using an appropriate "*window or weight function*" to resort to the weighted least–square polynomial regression. However, the theory of the method of least squares does not in general guarantee that the coefficients in the resulting polynomials satisfy the differential relations which we would naively expect to hold with respect to the shift of the point of interest. If such relations hold, the method is said to be "*endoharmonious*". The concept of endoharmony is new and has been introduced by the second author. It has been fully developed and investigated for the one-dimensional case [3]. An exploratory extension for the two-dimensional case has also been made with DEM (digital elevation model) as an example [1]. In this paper, the authors will develop more concretely the new concept for the two-dimensional case, theoretically as well as practically-numerically, and illustrate the validity from a number of viewpoints.

Specifically we will take up a methematical model of the phenomenon concerned:

$$\begin{aligned} z(x,y) &= \frac{1}{2}ax'^2 + bx'y' + \frac{1}{2}cy'^2 + dx' + ey' + f \\ &= g(x,y;a,b,c,d,e,f) \\ &\qquad (x' = x - x_0,\ \ y' = y - y_0), \end{aligned} \tag{1}$$

where x and y are the independent variables, z is the dependent, a, b, c, d, e and f are parameters (to be determined) of the model, and (x_0, y_0) is the "point of interest" around which we focus attention. Then we consider to estimate the parameters a, b, c, d, e and f based on the actually observerd values z_i $(i = 1, \cdots, N)$ at (x_i, y_i) $(i = 1, \cdots, N)$. The method of weighted least squares, or more precisely, the weighted least-square regression, is to determine a, b, c, d, e and f in such a way that the *weighted sum of squares of residuals*:

$$\sum_{i=1}^{N} w({x_i}', {y_i}'){\varepsilon_i}^2 \tag{2}$$

may be minimum, where

$$\varepsilon_i \stackrel{\text{def}}{=} z_i - g(x_i, y_i; a, b, c, d, e, f) \tag{3}$$

and $w(\xi, \eta)$ is the "weight function" or "window function" (having nonnegative values). Since the parameters a, b, c, d, e and f thus determined are functions of the point of interest (x_0, y_0), i.e., $a = a(x_0, y_0)$, $\cdots$, $f = f(x_0, y_0)$ we had better denote the model (1) as

$$\begin{aligned} &z(x, y, x_0, y_0) \\ &\quad = \frac{1}{2}a(x_0,y_0)x'^2 + b(x_0,y_0)x'y' + \frac{1}{2}c(x_0,y_0)y'^2 \\ &\qquad + d(x_0,y_0)x' + e(x_0,y_0)y' + f(x_0,y_0) \end{aligned} \tag{4}$$

to explicitly express the dependence on (x_0, y_0).

Here arises a natural question. From the meaning of those parameters, we might well expect to have

$$\begin{aligned} \frac{\partial f}{\partial x_0}(x_0,y_0) = f_{x_0}(x_0,y_0) &= d(x_0,y_0), \\ f_{y_0}(x_0,y_0) &= e(x_0,y_0), \\ d_{x_0}(x_0,y_0) &= a(x_0,y_0), \\ e_{y_0}(x_0,y_0) &= c(x_0,y_0), \\ d_{y_0}(x_0,y_0) = e_{x_0}(x_0,y_0) &= b(x_0,y_0), \end{aligned} \tag{5}$$

but we would quickly recognize that there is no reason for the parameter values determined by the method of least squares to satisfy such a relation in general. Then, is it impossible to think of the case where (5) holds? And, if it is possible, under what conditions is it? We shall say, if (5) holds, that the method is *endoharmonious* at (x_0, y_0).

As in the one-dimensional case [3], we may take

$$w(\xi, \eta) = \exp\left(-\frac{\xi^2 + \eta^2}{2R^2}\right) \tag{6}$$

for the weight function with resolution paramater R. By so doing, we can realize the endoharmony at those points which are not very close (in comparison with R) to the boundary of the considered region provided the following conditions hold.

1. The observation points (x_i, y_i) are placed on a regular rectangular grid.

2. The space between the grid points is small enough in comparison with R, "small enough" here meaning "0.6~0.7 times R" or less.

Those theoretical arguments will be illustrated with digital elevation data of different areas of Japan as examples.

2. Formulation of the problem

$N \times 6$ data matrix X and the vector $\boldsymbol{z}$ are defined as

$$X = \begin{bmatrix} 1 & y_1' & x_1' & \frac{y_1'^2}{2} & x_1'y_1' & \frac{x_1'^2}{2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & y_N' & x_N' & \frac{y_N'^2}{2} & x_N'y_N' & \frac{x_N'^2}{2} \end{bmatrix}, \tag{7}$$

$$\boldsymbol{z} = \begin{bmatrix} z_1 \\ \vdots \\ z_N \end{bmatrix}. \tag{8}$$

W is the diagonal matrix representing weights:

$$W = \mathrm{diag}[w(x_1', y_1'), \cdots, w(x_N', y_N')]. \tag{9}$$

$\boldsymbol{a}$ is the vector with six unknown components:

$$\begin{aligned}\boldsymbol{a} = [&f(x_0, y_0), e(x_0, y_0), d(x_0, y_0), \\ &c(x_0, y_0), b(x_0, y_0), a(x_0, y_0)]^{\mathrm{T}}.\end{aligned} \tag{10}$$

Usually these six unknowns are determined by solving the so-called "normal equations":

$$X^{\mathrm{T}}WX\boldsymbol{a} = X^{\mathrm{T}}W\boldsymbol{a}. \tag{11}$$

From definition (3), the residuals may be denoted in the vector form:

$$\boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_N \end{bmatrix} = \boldsymbol{z} - X\boldsymbol{a}, \tag{12}$$

and normal equation (11) in the form:

$$X^{\mathrm{T}}W\boldsymbol{\varepsilon} = \boldsymbol{0}. \tag{13}$$

For the sake of simplicity of notation, we will denote the derivative with respect to x_0 by ∂_0. In the following if we operate ∂_0 on (13), we have

$$(\partial_0 X^{\mathrm{T}})W\boldsymbol{\varepsilon} + X^{\mathrm{T}}(\partial_0 W)\boldsymbol{\varepsilon} + X^{\mathrm{T}}W(\partial_0\boldsymbol{\varepsilon}) = \boldsymbol{0}, \tag{14}$$

$$\partial_0\boldsymbol{\varepsilon} = -(\partial_0 X)\boldsymbol{a} - X\partial_0\boldsymbol{a}. \tag{15}$$

Furthermore we introduce 6×6 matrix C to denote the differentiation relations among basis polynomials, i.e.,

$$\partial_0 X = -XC, \quad \partial_0 X^{\mathrm{T}} = -C^{\mathrm{T}}X^{\mathrm{T}}. \tag{16}$$

Them we have for the third term of (14)

$$\partial_0\boldsymbol{\varepsilon} = X(C\boldsymbol{a} - \partial_0\boldsymbol{a}), \tag{17}$$

$$C\boldsymbol{a} - \partial_0\boldsymbol{a} = \begin{bmatrix} d(x_0, y_0) - \partial_0 f(x_0, y_0) \\ b(x_0, y_0) - \partial_0 e(x_0, y_0) \\ a(x_0, y_0) - \partial_0 d(x_0, y_0) \\ -\partial_0 c(x_0, y_0) \\ -\partial_0 b(x_0, y_0) \\ -\partial_0 a(x_0, y_0) \end{bmatrix}, \tag{18}$$

and we see that the first term of (14) vanishes by virtue of (13)

$$(\partial_0 X^{\mathrm{T}})W\boldsymbol{\varepsilon} = -C^{\mathrm{T}}X^{\mathrm{T}}W\boldsymbol{\varepsilon} = \boldsymbol{0}. \tag{19}$$

In order to investigate the second term of (14), we will take up the Gaussian type weight function (6), with R as the physical parameter representing the width of our view called the "resolution parameter". Then we have

$$\frac{\partial}{\partial\xi}w(\xi, \eta) = -\frac{\xi}{R^2}\exp\left(-\frac{\xi^2 + \eta^2}{2R^2}\right), \tag{20}$$

or

$$\partial_0 W = \mathrm{diag}\left[\frac{x_i'}{R^2}w(x_i', y_i')\right]. \tag{21}$$

Here we may introduce a 3×6 matrix D and another $N \times 3$ matrix E, to denote the coefficients of $\boldsymbol{\varepsilon}$ in the second term of (14) as

$$X^{\mathrm{T}}\partial_0 W = \begin{bmatrix} DX^{\mathrm{T}} \\ E^{\mathrm{T}} \end{bmatrix} W, \tag{22}$$

so that we have from (13)

$$X^{\mathrm{T}}(\partial_0 W)\boldsymbol{\varepsilon} = \begin{bmatrix} DX^{\mathrm{T}} \\ E^{\mathrm{T}} \end{bmatrix} W\boldsymbol{\varepsilon} = \begin{bmatrix} 0 \\ E^{\mathrm{T}}W\boldsymbol{\varepsilon} \end{bmatrix}. \tag{23}$$

Then (14) is reduced to

$$X^{\mathrm{T}}WX \begin{bmatrix} d(x_0,y_0)-\partial_0 f(x_0,y_0) \\ b(x_0,y_0)-\partial_0 e(x_0,y_0) \\ a(x_0,y_0)-\partial_0 d(x_0,y_0) \\ -\partial_0 c(x_0,y_0) \\ -\partial_0 b(x_0,y_0) \\ -\partial_0 a(x_0,y_0) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \\ -E^{\mathrm{T}}W\varepsilon \\ \\ \end{bmatrix}. \tag{24}$$

If $X^{\mathrm{T}}WX$ is diagonal, part of (5) is seen to hold.

If we take ∂_0 for the differentiation with respect to y_0, similar arguments will lead us to the remaining part of the conditions of endoharmony.

We may adopt any set of polynomials which constitutes a basis of polynomial up to degree 2 in the model (4):

$$\begin{aligned} z(x,y) = {} & f'(x_0,y_0)\psi_{0,0}(x-x_0,y-y_0) \\ & +e'(x_0,y_0)\psi_{0,1}(x-x_0,y-y_0) \\ & +d'(x_0,y_0)\psi_{1,0}(x-x_0,y-y_0) \\ & +c'(x_0,y_0)\psi_{0,2}(x-x_0,y-y_0) \\ & +b'(x_0,y_0)\psi_{1,1}(x-x_0,y-y_0) \\ & +a'(x_0,y_0)\psi_{2,0}(x-x_0,y-y_0). \end{aligned} \tag{25}$$

In that case diagonality of $X^{\mathrm{T}}WX$ in (24) is equivalent to the collocation orthogonality of the basis polynomials with weight $w(\xi,\eta)$, i.e.,

$$\sum_{i,j=-\infty}^{\infty} w(ih,jh')\psi_{k,l}(ih,jh')\psi_{p,q}(ih,jh') = 0 \quad \text{for } k \neq p \text{ or } k \neq p, \tag{26}$$

where we take into account that the number N of observation points is large enough and that the weight function $w(\xi,\eta)$ vanishes rapidly as $\xi \to \infty$ or $\eta \to \infty$, so that the finite sum may approximately be replaced by the infinite sum except near the boundary of the considered region.

If we adopt the Hermite polynomials (see [3]) and assume that

(a) there are sufficiently many points (x_i,y_j)'s $(i,j=1,2,\cdots,N)$ located at a regular rectangular grid,
(b) weight function is of the Gaussian type (6),

then the diagonality of $X^{\mathrm{T}}WX$ is numerically materialized as is shown in [3].

3. An endoharmonious procedure in the two-dimensional space

We will illustrate a concrete "endoharmonious" procedure for the least-square polynomial regression of data on a rectangular regular grid (with spaces h and h'). We adopt weight function (6) and basis polynomials

$$\begin{aligned} \psi_{k,l}(\xi,\eta) &= \frac{R^{k+l}}{k!l!} H_k\left(\frac{\xi}{R}\right) H_l\left(\frac{\eta}{R}\right), \\ H_n(\tau) &= (-1)^n \exp\left(\frac{\tau^2}{2}\right) \frac{\mathrm{d}^n}{\mathrm{d}\tau^n} \exp\left(-\frac{\tau^2}{2}\right). \end{aligned} \tag{27}$$

Then we have

$$\begin{aligned} -\frac{\partial}{\partial\xi}\psi_{k,l}(\xi,\eta) &= \psi_{k-1,l}(\xi,\eta), \\ -\frac{\partial}{\partial\eta}\psi_{k,l}(\xi,\eta) &= \psi_{k,l-1}(\xi,\eta). \end{aligned} \tag{28}$$

It is also seen that the condition of endoharmony takes the simple form in this case:

$$\begin{aligned} &\frac{\partial}{\partial x_0}d'(x_0,y_0) = a'(x_0,y_0), \ \frac{\partial}{\partial x_0}e'(x_0,y_0) = b'(x_0,y_0), \\ &\frac{\partial}{\partial x_0}f'(x_0,y_0) = d'(x_0,y_0), \ \frac{\partial}{\partial y_0}d'(x_0,y_0) = b'(x_0,y_0), \\ &\frac{\partial}{\partial y_0}e'(x_0,y_0) = c'(x_0,y_0), \ \frac{\partial}{\partial y_0}f'(x_0,y_0) = e'(x_0,y_0). \end{aligned} \tag{29}$$

4. Examples

The foregoing arguments will be illustrated for the endoharmonious weighted least-square polynomial regression of 100×100 digital elevation data covering Taniguchi. (x,y) coordinates are taken along latitude and longitude both in km. We set the origin $(0,0)$ at (134°07′30″E, 33°50′00″N). Around the area considered $x_{i+1}-x_i \doteqdot 57.8$m and $y_{i+1}-y_i \doteqdot 46.2$m. We set $R=100$m and approximated the differentiations with respect to x_0 and y_0 by the divided differences such as

$$\frac{\partial}{\partial x_0}p(x_0,y_0) = \frac{p(x_0+h,y_0)-p(x_0-h,y_0)}{2h} \quad (p=a,b,c,d,e,f), \tag{30}$$

with $h=5$m. Similarly for y with h'=5m. We compared the values of the partial derivative (with respect to x_0) of $d(x_0,y_0)$ with $a(x_0,y_0)$ (Fig.1) and the values of the partial derivatives of $e(x_0,y_0)$ and $d(x_0,y_0)$ with $b(x_0,y_0)$ (Fig.2). As is seen form those figures, the Gaussian weight exhibits the endoharmony. It should be noted that, if we took a non-Gaumian weight such as a triangular weight, we could not realize the endoharmony.

As is shown in Fig.3, the condition of endoharmony $\frac{\partial}{\partial x_0}f(x_0,y_0) = d(x_0,y_0)$ is not satisfied even by the Gaussian weight, whereas, as is shown in Fig.4, $d(x_0, y_0)$ coincides — not with $\frac{\partial}{\partial x_0} f(x_0, y_0)$ but — with $\frac{\partial}{\partial x_0}(f(x_0, y_0) + \frac{R^2}{2} a(x_0, y_0))$. It may be interesting to note that the quantity $f(x_0, y_0) + \frac{R^2}{2} a(x_0, y_0)$ is nothing but the coefficient of the Hermite polynomial of degree 0 of x' which we should have if we made regression analysis in terms of the Hermite polynomials.

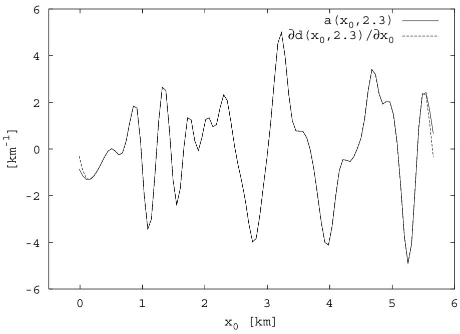

Fig. 1. Comparison of $a(x_0, 2.3)$ with $\partial d(x_0, 2.3)/\partial x_0$ (Gaussian weight, R =100m, h =5m)

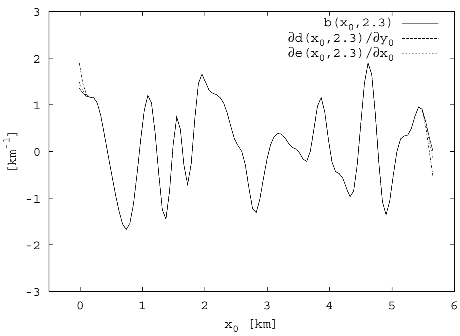

Fig. 2. Comparison of $b(x_0, 2.3)$ with $\partial d(x_0, 2.3)/\partial y_0$ and $\partial e(x_0, 2.3)/\partial x_0$ (Gaussian weight, R =100m, h =5m)

5. Concluding Remarks

We have developed more concretely the new concept of endoharmony for the two-dimensional case, theoretically as well as practically-numerically, and illustrated the validity from a number of viewpoints. The results which we could not include in this abstract because of the space limitation will be presented in the symposium.

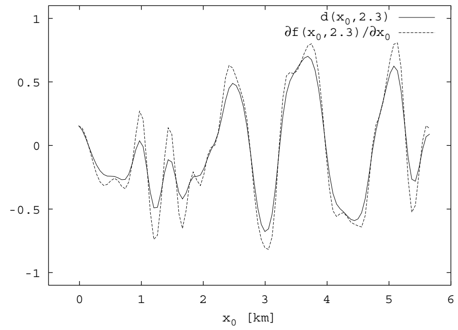

Fig. 3. Comparison of $d(x_0, 2.3)$ with $\partial f(x_0, 2.3)/\partial x_0$ (Gaussian weight, R =100m, h =5m)

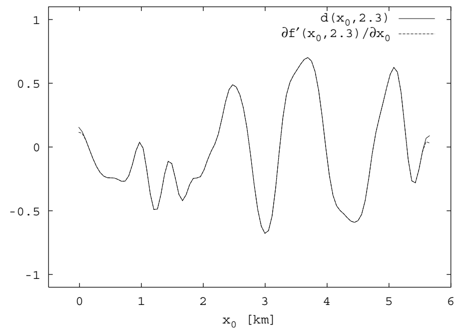

Fig. 4. Comparison of $d(x_0, 2.3)$ with $\partial f'(x_0, 2.3)/\partial x_0$ (Gaussian weight, R =100m, h =5m)

References

[1] M. Iri. Y. Shimakawa and T. Nagai: Extraction of invariants from digital elevation data with application to terrain topography, *Nonlinear Analysis: Series A – Theory and Methods*, vol. 47, issue 8 (2001), pp.5585–5598.

[2] M. Iri: How to harmonize the method of least squares with differentiation (in Japanese), *Papers and Proceedings of the Geographic Information Systems Association, Japan*, vol.9 (2000), pp.389–394.

[3] M. Iri: A new viewpoint on the method of least squares, *TRISE 01–04*, Department of Information and System Engineering, Faculty of Science and Engineering, Chuo University, July 2001.