應用疊代式移動型小平方誤差法之波折解法研究具有跳躍斷點的數據特性

Wave Decomposition Across Discontinuity Using Iterative Moving Least Squares Methods

Yih Nen Jeng (鄭育能)
Professor , Aero. and Astro. Dept., National Cheng-Kung Univ., Taiwan
Z6208016@email.ncku.edu.tw
You-Chi Cheng (鄭又齊)
Engineer, Zu-Yuan Software Inc., Taipei, Taiwan
Tzung-Ming Yang (楊宗明)
Graduate Student, Aero. and Astro. Dept., National Cheng-Kung Univ., Taiwan

ABSTRACT

The iterative filter based on the iterative moving Gaussian smoothing method is employed to decompose wave from a data string with discontinuity. Around a discontinuous jump, the resulting smoothed data process the oscillatory Runge phenomenon whose wavelengths are in the same order of the smoothing factor σ of the Gaussian smoothing factor. On the other hand, the amplitude is only proportional to the magnitude of jump. For the discontinuity of derivatives, the induced oscillation is not significant. Similar result can be obtained for the iterative cubic moving least squares method which consists with the uniqueness theorem of the iterative filter. The limited filtering range method is employed to eliminate the Runge phenomenon. Numerical tests show that the proposed can effectively capture the original smoothed wave.

Keywords: iterative moving least squares method, discontinuous jump, limited filtering range.

INTRODUCTION

Because of the rapid development of computer hardware and software, especially after the nano-technology becomes a practical engineering work, how to automatically handle a tremendous amount of data within a reasonable short time period becomes more and more important. Today, people heavily reply on human inspection to remove failed data string and obtain available data for analysis. How to develop fundamental tools of the automatic inspection is a key technology to overcome the burry of human inspection. To the authors' opinion, a reliable wave decomposition technology which gives complete information of wave amplitude, frequency and phase velocity everywhere will give all necessary information for automatic inspection. For example, around data failure point, strange behavior will be reflected by a rapid variation of amplitude, frequency and/or phase.

So far there are several successful methods of wave decomposition: Fourier series expansion [1]; matrix pencil method [2,3]; and the empirical mode decomposition of Huang et. al. [4,5]. The result of the first method generally gives fixed spectrum over the expansion range. The second method is a zero-th order approximation. The last method is contaminated by the numerical modification of cubic spline interpolation. In other words, these method are not robust enough to take care of the data failure region.

In a paper presented in this conference [6], an iterative procedure to decompose different wave form from a data string composed of several waves via the moving least squares method is proposed. An new strategy is also proposed that one can employ a low order least squares smoothing method to achieve the performance of a high order method provided that the closest frequencies of two waveforms around the cut-off point are separable. Moreover, if these two frequencies close to each other, the required iteration steps increases exponentially. Although the limitation of lengthy iteration might restrict the application of this method, it is possible to replace the basic smoothing method (such as the diffusive Gaussian smoothing method) by a less diffusive smoothing method that their finally results are the same. Therefore, it is valuable to study the properties of the filtered result such as the behavior across a discontinuity and how to improve any possible undesirable phenomenon. In this study, the Runge phenomenon of the final result across a discontinuity and a remedy strategy will be studied. The discontinuity of a data string frequently relates to data failure or a deterioration of data quality.

THEORETICAL ANALYSIS

Previous Works

Consider a set of data $(x_i, y_i), i = 0,1,...,n$ to be approximated by a polynomial. The error targeting function of the weighted moving least squares method is defined at every data point x_k as [6-8]

$$I_k = \sum_j \exp[-\frac{(x_j - x_k)^2}{2\sigma^2}][y_j - \sum_{m=0}^{M} A_{m,k}(x_j - x_k)^m]^2 \qquad (1)$$

The least squares method requires that I_k is minimum with respect to $A_{0,k}, A_{1,k}, ..., A_{m,k}$. Thus, the following simultaneous algebraic equations are to be solved:

$$\partial I_k / \partial A_{0,k} = 0,\ \partial I_k / \partial A_{1,k} = 0,\ ...,\ \partial I_k / \partial A_{m,k} = 0, \qquad (2)$$

The solution of these equation yields the least-squares weighted value, $\bar{y}_k = A_{0,k}$, at $x = x_k$ [2-4]. If the polynomial contains only the zero degree term, the method becomes the well-known Gaussian smoothing method [9-11].

Assume that a data string y defined on uniform spacing $x_k = k\Delta x$, it is expanded in form of

$$y = \sum_{j=0}^{J} y_j(\lambda_j)\ , \qquad (3)$$

where $y_j(\lambda_j)$'s are wave components consist of sine or cosine functions. After employing the Gaussian smoothing method, the resulting data string can be written in the following form:

$$\bar{y}_i = S(y) = S\left(\sum_{j=0}^{J} y_j(\lambda_j)\right) = \sum_{j=-}^{J} S(y_j(\lambda_j)) = \sum_{j=0}^{J} a(\sigma/\lambda_j, 0) y_j(\lambda_j)\ , \qquad (4)$$

where $S(y)$ denotes the linear smoothing operator, $a(\sigma/\lambda_j, 0)$ is the attenuation factor reflecting the effect of S upon the j -th wave, and 0 denotes $M = 0$. In Ref.[6], the following theorems state the property of $a(\sigma/\lambda_j, M)$.

Theorem 1

The result of applying the Gaussian smoothing method to a data string described by Eq.(3) has the following dissipative property:

$$0 \le a(\sigma/\lambda_j, M) \le 1\ . \qquad (5)$$

Theorem 2

By applying the moving least squares method with $M = 1,2,...$ in Eq.(1) to a data string y , the dissipative property of Eq.(5) is also satisfied.

It can be shown that, for the Gaussian smoothing method, the resulting explicit formula for the data y is [9-11]

$$\bar{y}_k = \sum_{j=-\infty}^{\infty} \exp[-\frac{(x_j - x_k)^2}{2\sigma^2}] y_j \div \sum_{j=-\infty}^{\infty} \exp[-\frac{(x_j - x_k)^2}{2\sigma^2}] \qquad (6)$$

In Ref.[6], an iterative procedure was derived as follows.

1. Select a smoothing filter having the following property
$$\left|1 - a(\sigma/\lambda_j, M)\right| \le 1\ . \qquad (7)$$
2. If one uses the Gaussian smoothing technique, use equation (9) to estimate the critical values of σ and m . Perform m -th number of iterations as described in (3) and (4) using the critical σ and then go to (6). On the other hand, for smoothing filters in which the values of m and σ can not be identified, one may use trials and errors to select a σ such that steps 2 and 4 can converge.
3. Employ the filter to the original data string to separate both the smoothed and residual waveforms.
4. Apply the same smoothing method to the resulting residual waveform to yet obtain another smoothed and residual waveforms.
5. Repeat (3) and (4) until the increment smoothed waveform is negligible.
6. The desired high frequency waveform is the final residual waveform and the low frequency waveform is the difference of the high frequency waveform and the original data.

Subsequently, the resulting smoothed data y is

$$\bar{y}(m) = \sum_{j=0}^{J} \{1 - [1 - a(\sigma/\lambda_j, M)]^m\} y_j(\lambda_j) = \sum_{j=0}^{J} b(\sigma/\lambda_j, m) y_j(\lambda_j) \qquad (8)$$

If the original data are composed of smooth waves and the represented wavelength around the cut-off point in the frequency domain are λ_{c1} and λ_{c2} , respectively, the necessary factor σ and iteration step m for the Gaussian smoothing and the linear moving least squares methods can be solved from the following equation

$$\begin{aligned} b(\sigma/\lambda_{c1}, m) &= 1 - [1 - \exp\{-\frac{2\pi^2\sigma^2}{\lambda_{c1}^2}\}]^m = b_1 \\ b(\sigma/\lambda_{c2}, m) &= 1 - [1 = \exp\{-\frac{2\pi^2 o^2}{\lambda_{c2}^2}\}]^m = b_2 \end{aligned} \qquad (9)$$

provided that parameters b_1 and b_2 are given as 0.001 and 0.999, respectively. For the quadratic and cubic moving least squares methods, a similar set of equation can be derived too [6].

Using the iterative scheme and above mentioned theorems, the following theorems are also derived in Ref.[6].

Theorem 3 (Uniqueness Theorem)

If the lower and upper limits of the transition zone, b_1 and b_2 , are identical and the data string contain two adjacent waveforms having a difference in wavelength larger than two limits of the transition zone , the resulting low and high frequency parts can be uniquely defined provided that the filters employed have the property of Eq.(5).

Corollary 3.1

If a filter has the non-negative property

$$0 \le a(\sigma, \lambda_j, M) < \infty\ , \qquad (10),$$

an iterative filter can be constructed to decompose two waveforms by applying under-relaxation factors to all the smoothed waveforms.

$$\begin{aligned}&\bar{y}=\omega_1\bar{y}_1+\omega_2\bar{y}_2+\omega_3\bar{y}_3+...+\omega_m\bar{y}_m\\&y'=y-\bar{y}\end{aligned}, \tag{11}$$

$$0<\omega_1,\omega_2,...,\omega_m<1\,, \tag{12}$$

provided that the waveforms of the data string are separable.

Limited Filtering Range Method

If a data string involves a discontinuity, a Fourier series approximation across there always introduces the Gibbs phenomenon, say spurious oscillation. If the approximation is restricted to continuous region, the oscillation can be essentially avoid which is similar to that monotonic interpolation scheme of Ref.[11]. A convenient method to identify boundaries of these piecewise continuous segments are the location where the first cycle's high frequency part change sign and magnitudes at two sides are significant. For example, in Fig.1 the original data has a jump at $x=0.5$, the first cycle's high frequency part of Gaussian smoothing method satisfies the following properties.

$$\begin{aligned}&y_1^{'}(x_i)\times y_1^{'}(x_{i+1})<0\\&\left|y_1^{'}(x_{i-p})\right|_{local\ \max},\left|y_1^{'}(x_{i+1+q})\right|_{local\ \max}>K\end{aligned} \tag{13}$$

where $K\geq\alpha\left\|y_1^{'}\right\|_2$, α is an user specified parameter, and $y_1^{'}$ is the high frequency residual part of the first cycle. In many practical problems, the index number p and q are greater than unity. Moreover, small oscillations also exist across a discontinuity of derivatives dy/dx, d^2y/dx^2, and high order derivatives, respectively. Therefore, it seems reasonable to put a segment boundary point at

$$\left|y_1^{'}\right|_{local\ \max}\geq K \tag{14}$$

Although such a discriminator will embed a small segment between local maximum and minimum points of Eq.(14) which results in a slightly smeared jump resolution, it is rather simple and easy to implement. Since within the segment there is only a few points that may introduce oscillatory result, it is recommended to enlarge the segment boundaries at two ends.

For a practical implementation, the high frequency residual part of the first cycle, $y_1^{'}$ should be smoothed by the Gaussian smoothing method with a proper value of smoothing factor σ. As shown in Fig.1, the original data involves a discontinuous jump at $x=5$ and random noise with a large variation so that the resulting $y_1^{'}$ has a large local variation. After smoothing with $\sigma=0.1$, the resulting data has four local maximum points around $x=5$. Among these local maximum points, two points are not the desired points and can be excluded by adding the following constraint, say

$$\left|y_i-y_{i-1}\right|>\frac{\beta}{n}\sum_j\left|y_j-y_{j-1}\right| \tag{15}$$

where $n+1$ is total point number and $\beta>3$. After applying the restriction of Eqs.(14) and (15) to the smoothed high frequency residual part, shown as the heavy solid line in the figure, the interior segment boundaries are chosen at $x=4.96$ and 5.06. Although this restriction gives a relatively wide transition zone from $y\approx0$ to $y\approx1$, the proposed strategy is a robust for many tests.

RESULTS AND DISCUSSIONS

The first test case is the unit step function

$$\begin{aligned}y=u(5)&=0,\quad x<5\\&=1,\quad x\geq5\end{aligned} \tag{16}$$

The case is different from those discussed in Ref.[6] which are all of smoothed waves. Now the optimal factor σ and iteration step m do not have any meaning. There is a limit for the smoothed part as the iteration step m increases indefinitely. Figures 2 and 3 show the resulting smoothed value by using the iterative Gaussian smoothing method with $\sigma=0.25$ and 0.5, respectively. There exist oscillations (often termed as the Runge phenomenon) across the jump in both figure. Their oscillation amplitude are slowly die out within a distance away from the jump and their magnitudes of amplitude are about the same (both first peaks are at 1.1 for $x>5$). Their first oscillatory wave lengths are 0.35 and 0.71 for $\sigma=0.25$ and 0.5, respectively. In other words the wavelength of the oscillatory wave length is about $\sqrt{2}$ times the smoothing factor σ. If the jump is reduced to be 0.1, the result is shown in Fig.4. The wavelength changes slightly and the ratio between amplitude to jump is the same as that of Figs.2 and 3. In other words, the Runge phenomenon always exists. Figure 5 shows the result of employing the iterative moving least squares method, except that the iteration number is different, the final result is the same as that of Fig.3 that consists with the uniqueness theorem.

Next, the random number is added to the original wave so that

$$\begin{aligned}y&=0.1*(r-0.5),\qquad x<5\\&=1+0.1*(r-0.5),\quad x\geq5\end{aligned} \tag{17}$$

where r is the random number generated by the RANDOM subroutine of the Microsoft F-77 software. After applying the iterative Gaussian smoothing method, results are shown in Figs.6 and 7 for $\sigma=0.25$ and 0.5, respectively. A careful comparison between these figures to Figs.2 and 3, respectively, reveals that the Runge phenomenon is not affected by the imposed random noise. Around the jumping point, the corresponding amplitude and wavelength distributions of each iteration shown in figure are nearly the same with each other. This result confirms the fact that the wave with length shorter than 1.6 σ will be removed.

For a continuous data string with slope discontinuity, the smoothed part generated by the proposed iterative scheme generate a much more smaller Runge phenomenon than that discussed above. The original data in Fig.8 has a first order discontinuity is found at $x=0.3,0.5,$ and 0.7. The convergent smoothed part generated by the present iterative Gaussian smoothing method with $\sigma=0.05$ has small error around the slope discontinuous point as comparing with that of Figs.2 through 7. For a factor of $\sigma = 0.02$, both the amplitude and wavelength of the deviation from original data are less than that with $\sigma = 0.05$ as shown in Fig.9. For the discontinuity of the second order derivative, the error is smaller than that shown in Figs.8 and 9 and is not shown here. Generally, a practical data string is frequently contaminated by noise or short waves. So far it is nearly impossible to correctly capture the Runge phenomenon due to jumps of the first and high order derivatives. It is thus recommended to employ a factor σ as small as possible to ensure that the noise is eliminated and high frequency waves are captured.

The above discussions reflect a fact that, the convergent smooth part of the proposed smoothing method for a data string with a discontinuous jump will produce the undesirable Runge phenomenon and should be removed.

Figure 10 is the result of employing the iterative Gaussian smoothing method with $\sigma = 0.05$ for 2 cycle (estimated by Eq.(9)). Since the random number added to the original data is only but a pseudo-random, in addition to the main wavelength of $\lambda \approx 2\Delta x$ it has an additional characteristic wave with $\lambda_{eff} = 0.08$, if $\sigma < 0.04$ will reserve it to the smoothed part. It is clear that the present limited range method works very well.

Figure 11 shows a test case of block function added by random number. Since the employed factor $\sigma = 0.5$ is much larger than the wave length of the random number, only one cycle of the Gaussian smoothing method together with the limiting range strategy gives a satisfactory result. That shown in Fig.12 employs two waves plus the random number as

$$y = 0.2[\sin \pi x + \sin 2\pi x] + 0.07 * (r - 0.5) \quad (18)$$

where r is the random number in the range of $0 \le r \le 1$ generated by the RANDOM subroutine. Since the smooth composed wave has an effective wavelength of $\lambda = 0.8$ and the random number has an wavelength of $\lambda_{eff} = 0.08$, Eq.(9) gives parameters $\sigma = 0.05$ and $m = 1$. Again, the result of applying the limited range strategy and iterative Gaussian smoothing method gives a satisfactory filtered smoothed wave. A careful inspection of Figs.10 through 12 shows that the proposed method effectively eliminates the Runge phenomenon across the discontinuity.

CONCLUSIONS

The iterative moving least squares method is employed to study the properties of the oscillatory Runge phenomenon around a discontinuous jump. A strategy of limited filtering range is successively proposed to eliminate the oscillation.

ACKNOWLEDGEMENT

This work is supported by the National Science Council of Taiwan under the grant number NSC-92- 2212-E006-077.

REFERENCE

1. P. M. Morse and H. Feshbach, “Methods of Theoretical Physics,” New York, McGraw-Hill , vol.2, 1953.
2. Hua, R. and Sarkar, T. K., “Generalized Pencil-of-Function Method for Extracting Poles of an EM System from Its Transient Response,” IEEE Trans. Antennas and Propagation, vol.37, no.2, Feb. 1989, pp.229-234.
3. Ruscio, D. D., “On Subspace Identification of Extended Observability Matrix,” Proc. Of 36th Conf. on Decision & Control., San Diego, Cal. U.S.A., Dec. 1997, Paper no.TA05, pp.1841-1847.
4. N. E. Huang, Z. Shen, and S. R. Long, "A New View of Nonlinear Water Waves: the Hilbert Spectrum," Annual Reviews of Fluid Mechanics, 1999, vol.51, pp.417-457.
5. Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N. C., Tung, C. C. and Liu, H. H., “The Empirical Mode Decomposition and the Hilbert Spectrum for Nonlinear and Non-stationary Time Series Analysis,” Proc. R. Soc. Lond. A., vol. 454, pp.903-995, 1998.
6. Jeng, Y. N., Huang, G. P. G. and H. Chen, “Wave Decomposition in Physical Space Using Iterative Moving Least Squares Methods,” also presented in this conference.
7. Y. N. Jeng, “The Moving Least Squares and Least p-Power Methods for Random Data,” The 7-th National Computational Fluid Dynamics Conference, P-9 to P-14, Aug. 2000.
8. A. K. Mackworth, and F. Mokhtarian, “Scale-Based Description of Planar Curves,” Proc. 5th Canadian Soc. Comp. Studies Intell. London, Ontario pp.114- 119, May 1984.
9. F. Mokhtarian and A. Mackworth, “Scale-Based Description and Recognition of Planar Curves and Two-Dimensional Shapes,” IEEE Trans. Pattern Anal. Machine Intell., vol. PAMI-8, no. 1, pp. 34-43, Jan. 1986.
10. D. G. Lowe, “Organization of Smooth Image Curves at Multiple Scales,” Int. J. Computer Vision, vol.3, pp.119-130, 1989.
11. Huynh, H. T “Accurate Monotone Cubic Interpolation,” SIAM J. Number. Anal. vol.30, no.1, pp57-100, Feb.1993.

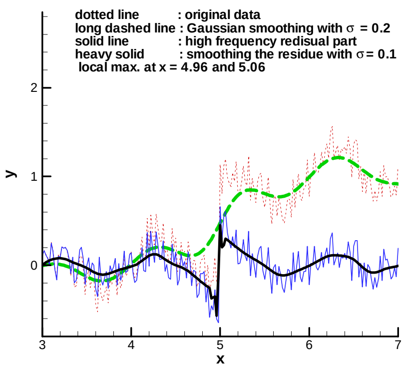

Fig.1 The high and low frequency parts of employing the Gaussian smoothing method one cycle across a jump with $\sigma = 0.2$, the heavy solid line is the smoothed high frequency residue with $\sigma = 0.1$.

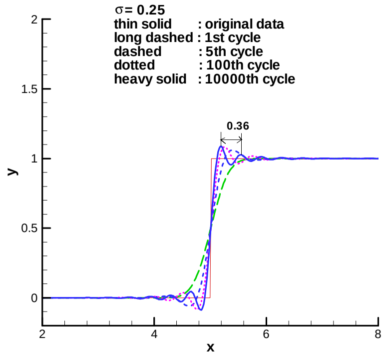

Fig.2 The results of employing the iterative Gaussian smoothing method with $\sigma = 0.25$: thin solid line is the original data; long dashed line is the first cycle result; dashed line is the 5th cycle result; dotted line is the 100th cycle result; and heavy solid line is the 10000th cycle result.

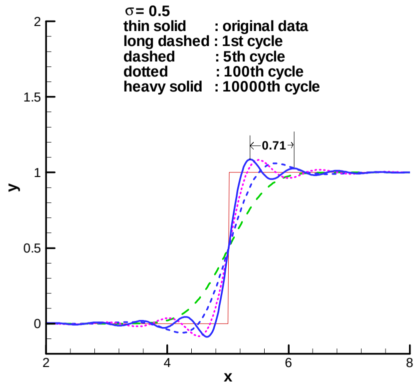

Fig.3 The results of employing the iterative Gaussian smoothing method with $\sigma = 0.5$: thin solid line is the original data; long dashed line is the 1st cycle result; dashed line is the 5th cycle result; dotted line is the 100th cycle result; and heavy solid line is the 10000th cycle result.

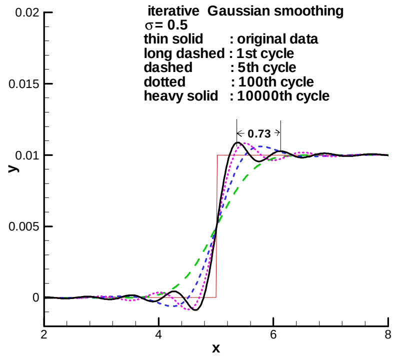

Fig.4 The results of employing the iterative Gaussian smoothing method with $\sigma = 0.5$, the jump is reduced to be 0.1: thin solid line is the original data; long dashed line is the first cycle result; dashed line is the 5th cycle result; dotted line is the 100th cycle result; and heavy solid line is the 10000th cycle result.

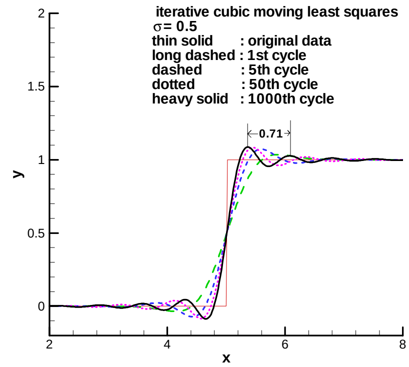

Fig.5 The results of employing the iterative cubic moving least squares method with $\sigma = 0.5$: thin solid line is the original data; long dashed line is the first cycle result; dashed line is the 5th cycle result; dotted line is the 100th cycle result; and heavy solid line is the 10000th cycle result.

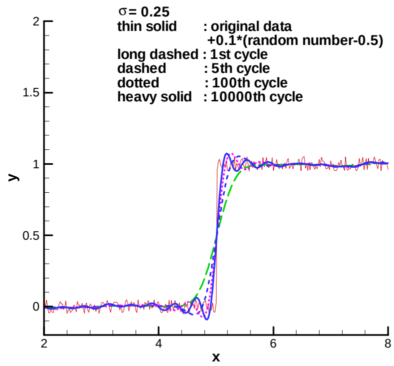

Fig.6 The results of employing the iterative Gaussian smoothing method with $\sigma = 0.25$: the zigzag thin solid line is the original data; long dashed line is the first cycle result; dashed line is the 5th cycle result; dotted line is the 100th cycle result; and heavy solid line is the 10000th cycle result.

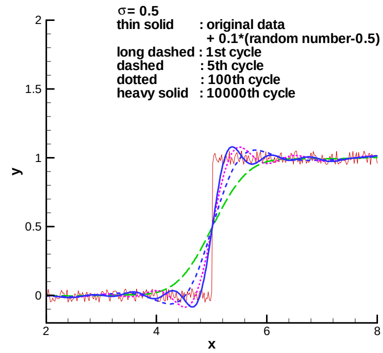

Fig.7 The results of employing the iterative Gaussian smoothing method with $\sigma = 0.5$: the zigzag thin solid line is the original data; long dashed line is the first cycle result; dashed line is the 5th cycle result; dotted line is the 100th cycle result; and heavy solid line is the 10000th cycle result.

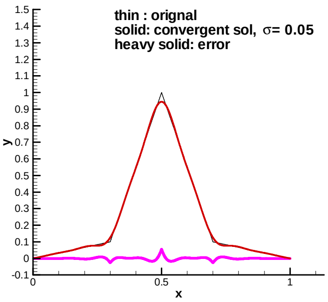

Fig.8 Result of slope discontinuity: thin solid line is the original data; upper solid line is the convergent smoothed part with $\sigma = 0.05$; and lower solid line is the error of the smoothed part.

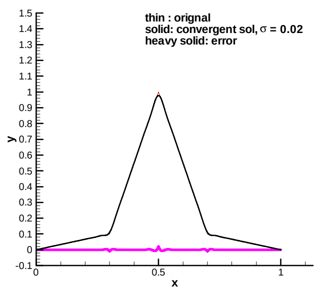

Fig.9 Result of slope discontinuity: thin solid line is the original data; upper solid line is the convergent smoothed part with $\sigma = 0.02$; and lower solid line is the error of the smoothed part.

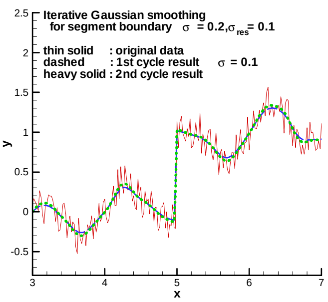

Fig.10 Result of employing the limited range method and iterative Gaussian smoothing method: zigzag line is the original data, dashed line is the first cycle result and heavy solid line is that of 2nd cycle, with $\sigma = 0.2$ and $\sigma_{res} = 0.1$ to judge segment boundaries and . $\sigma = 0.1$ for main iteration.

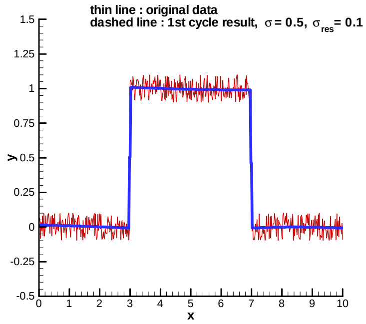

Fig.11 Result of employing the limited range method and iterative Gaussian smoothing method: zigzag line is the original data, the heavy line is the first cycle result, with $\sigma = 0.2$ and $\sigma_{res} = 0.1$.

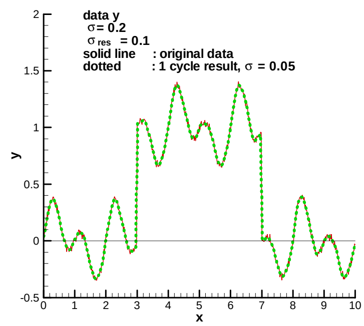

Fig.12 The result of recovering smoothed data with discontinuous jump by filtering out the random number whose effective wavelength is about $\lambda_{effect} \approx \Delta x = 0.02$: the original data is shown as thin solid line; 1st cycle result is shown as heavy solid line; the 10th cycle result is shown as dotted line; $\sigma = 0.2,\ \sigma_{res} = 0.1$.

應用疊代式移動型小平方誤差法之波拆解法研究具有跳躍斷點的數據特性

鄭育能 成大航太系教授
鄭又齊 儒園科技軟體公司工程師
楊宗明 成大航太系研究生

摘　要

本文應用疊代式移動型最小平方誤差之波拆解法研究據有斷點的數據串，經平滑後的 RUNGE 震盪現象。震盪波從斷點位置起向兩邊衰減，第一個震盪波的波長只與高斯函數之平滑參數 σ 成正比率，但震幅則只與斷點的高低大小有關。斜率及高次導函數的不連續性所產生的震盪誤差不大且其震幅和波長都和平滑參數 σ 成比率，因此若使用合理小的 σ ，可以不需特意處理這類的斷點誤差。本文提出一種簡易的數據串分段濾波法，對各斷分別做平滑，數值證明此種簡易法可以消除震盪性誤差。

關鍵詞：疊代式波拆解法，斷點，RUNGE 震盪現象，數據串分段濾波法。