Published in final edited form as:

Appl Spectrosc. 2017 August ; 71(8): 1894–1905. doi:10.1177/0003702816685336.

Least-Squares Moving-Window Spectral Analysis

Young Jong Lee

Biosystems & Biomaterials Division, National Institute of Standards and Technology,
Gaithersburg, MD 20899, USA

Abstract

Least-squares regression is proposed as a moving-windows method for analysis of a series of spectra acquired as a function of external perturbation. The least-squares moving-window (LS-MW) method can be considered an extended form of the Savitzky-Golay differentiation for non-uniform perturbation spacing. LS-MW is characterized in terms of moving-window size, perturbation spacing type, and intensity noise. Simulation results from LS-MW are compared with results from other numerical differentiation methods, such as single-interval differentiation, autocorrelation moving-window and perturbation-correlation moving-window methods. It is demonstrated that this simple LS-MW method can be useful for quantitative analysis of non-uniformly spaced spectral data with high frequency noise.

Keywords

moving-window; least-squares regression; Savitzky-Golay; numerical differentiation; two-dimensional correlation spectroscopy

1. Introduction

The generalized two-dimensional correlation spectroscopy (2DCOS) is widely used for characterization of synchronous and asynchronous correlations of multiple spectral bands acquired as a function of external perturbation.[1,2] The generalized 2DCOS is useful for determining the direction and sequence of spectral changes occurring over an entire perturbation range. However, when spectral changes occur in narrow perturbation regions, the generalized 2DCOS cannot locate the perturbation regions and associate them with specific spectral changes. Various moving-window (MW) approaches have been proposed to overcome the limitation of 2DCOS over an entire perturbation range by performing 2DCOS within a narrow perturbation window and repeating it after moving the window.[3] An autocorrelation spectrum represents the overall extent of spectral intensity variation within a perturbation window. A contour map of autocorrelation moving-window (AC-MW) intensity as a function of spectral variable and average perturbation variable is used for characterization of phase transitions and chemical reactions occurring in various molecular systems, including polymers,[4,5] hydrogels,[6] and proteins.[7,8] However, AC-MW cannot determine the direction of spectral change with respect to perturbation because the

Notes

Official contribution of the National Institute of Standards and Technology; not subject to copyright in the United States.

output of AC-MW is intensity variation, which is always positive. An alternative method, called perturbation-correlation moving-window (PC-MW), has been proposed to provide direction by calculating the correlation between the perturbation variable and the spectral intensity within each moving window.[9] The synchronous correlation intensity of PC-MW represents direction of spectral intensity change along the perturbation change. Due to this unique information, PC-MW has become a popular MW method for spectral analysis of various physical and chemical changes.[10–17]

Correlation intensities of the two MW methods are often used as a surrogate for the first order derivative with respect to perturbation,[9,18] with their linear proportionality being described analytically and numerically.[19] In cases where spectral change is cumulative along perturbation, the first order derivative can become very helpful for data interpretation because it eliminates constant background and visualizes both the peak and the width of transition perturbation. Among various methods to calculate the first order derivative of discrete spectral data, the simplest is single-interval differentiation, also known as gradient mapping[20] or simple numerical differentiation.[9] It calculates $\Delta I / \Delta p$ per each perturbation interval, where ΔI is intensity change and Δp is perturbation interval. However, single-interval differentiation is known to be highly susceptible to high frequency noise with respect to perturbation even with subsequent smoothing.[18]

In 1964, Savitzky and Golay proposed a method of data smoothing and differentiation based on local least-squares polynomial regression.[21] They fit a small section of a dataset with a polynomial function and repeated the fitting for the next section, which is similar to the moving-window approach to 2DCOS. Because the Savitzky-Golay smoothing reduces noise reasonably well while maintaining the shape of the original, their method is widely used for signal processing. In addition, the least-squares based fitting method produces a straightforward, closed-form solution and does not rely on any iterative optimization. The Savitzky-Golay smoothing and differentiation is generally applied only to uniformly spaced data points. Experimental data, however, are not always acquired with uniformly spaced perturbation variables. For example, actual temperatures measured while spectra are acquired can be different from their set temperatures due to various experimental reasons. Also, when time-spaced data are converted into energy-spaced data, the perturbation spacing has to become non-uniform. Uniform spacing simplifies the formulae of the Savitzky-Golay method and reduces its computation time, but it is not an intrinsic condition for least-squares regression. Basic formulae of general least-squares regression of non-uniformly spaced data can be found in many textbooks.[22]

In this paper, least-squares regression for moving windows is proposed as a numerical differentiation method, which is versatile for both uniformly and non-uniformly spaced datasets. This least-squares moving-window (LS-MW) method is compared with AC-MW and PC-MW while effects of window size, non-uniform perturbation increment, and intensity noise are all examined.

2. Theory

2.1. General Description of Least-Squares Moving-Window (LS-MW)

The basic concept of moving-window analysis is illustrated in Fig. 1a. A series of spectra are measured as an external perturbation varies. Spectral intensity $y(\nu, p)$ is expressed as a function of spectral variable ν and perturbation variable p . As shown in Fig. 1b, a moving window is defined as a subset of $(2m + 1)$ spectra within $p_{j-m} \leq p \leq p_{j+m}$, where p_j is the median perturbation variable. When MW analysis is performed independently for single ν , e.g., AC-MW and PC-MW, spectral intensity $y(\nu, p)$ can be simplified as $y_j = y(p_j)$. For each window, the dynamic or mean-centered spectrum is expressed as $\tilde{y}_j = y_j - y_{ave}$, and the mean-centered perturbation variable, $\tilde{p}_j = p_j - p_{ave}$, where y_{ave} and p_{ave} are the average spectral intensity and the average perturbation variable, respectively.

$$y_{ave} = \frac{1}{2m+1} \sum_{j=j-m}^{j+m} y_j, \quad (1)$$

$$p_{ave} = \frac{1}{2m+1} \sum_{j=j-m}^{j+m} p_j. \quad (2)$$

It is noted that the average perturbation variable p_{ave} becomes equal to the median perturbation variable p_j when the perturbation spacing is uniform.

Fig. 1c shows a plot of $\tilde{y}_j$ and $\tilde{p}_j$ in a moving window, where the mean-centered dataset is least-squares fitted with a linear model function. This analysis method is called *linear-function least-squares moving-window* (*L-LS-MW*) or *LS-MW* in short unless further specified. The basic idea of least-squares differentiation of moving windows was originally proposed by Savitzky and Golay in 1964.[21] Briefly, a linear model function is expressed as

$$\tilde{y}_j = \alpha + \beta \tilde{p}_j + e_j, \quad (3)$$

where α is the constant; β is the linear coefficient; and e_j is the error made in prediction of $\tilde{y}_j$. The unknown coefficients α and β are to be determined in order to minimize the sum of the squared errors (*SSE*), which is defined as

$$SSE \equiv \sum_{j=j-m}^{j+m} e_j^2 = \sum \left(\tilde{y}_j - \alpha - \beta \tilde{p}_j \right)^2. \quad (4)$$

For the minimum SSE with respect to the coefficients α and β , the following two associated equations are solved

$$0 \equiv \frac{\partial SSE}{\partial \alpha} = \frac{\partial}{\partial \alpha} \sum (\tilde{y}_j - \alpha - \beta \tilde{p}_j)^2 = -2 \sum (\tilde{y}_j - \alpha - \beta \tilde{p}_j), \quad (5)$$

$$0 \equiv \frac{\partial SSE}{\partial \beta} = \frac{\partial}{\partial \beta} \sum (\tilde{y}_j - \alpha - \beta \tilde{p}_j)^2 = -2 \sum \tilde{p}_j (\tilde{y}_j - \alpha - \beta \tilde{p}_j). \quad (6)$$

By applying the mean-centering conditions ($\sum \tilde{y}_j = 0$ and $\sum \tilde{p}_j = 0$), Eq. 5 yields $\alpha = 0$. Then, Eq. 6 leads to the solutions for the linear coefficient

$$\beta = \frac{\sum \tilde{p}_j \tilde{y}_j}{\sum \tilde{p}_j^2}. \quad (7)$$

It is noted that this expression is good whether the perturbation spacing is uniform or not. After this calculation is repeated for each moving window, the determined linear coefficients β are expressed as a function of the average perturbation variable p_{ave} . The relation between β and the first order derivative $y^{(1)}(p)$ can be found by inserting a linear truncated Taylor series $\tilde{y}_j = y^{(1)}(p_{ave}) \tilde{p}_j$ into Eq. 7

$$\beta = \frac{y^{(1)}(p_{ave}) \sum \tilde{p}_j^2}{\sum \tilde{p}_j^2} = y^{(1)}(p_{ave}). \quad (8)$$

This equivalency between β and $y^{(1)}(p)$ is effective as long as $\tilde{y}_j$ is close to a linear function within a moving window. That condition can be satisfied either the function itself is close to a linear function or the window size is sufficiently small.

For quantitative characterization of the window size effect on β from LS-MW, simulation data are generated by using a modified error function

$$y(p) = \frac{1}{2} \operatorname{erfc} \left[-2 \sqrt{\ln 2} \left(\frac{p - p_c}{\sigma} \right) \right] = \frac{2 \sqrt{\ln 2}}{\sqrt{\pi} \sigma} \int_{-\infty}^p e^{-4 (\ln 2) \left(\frac{x}{\sigma} \right)^2} dx, \quad (9)$$

which is illustrated in Fig. 1b. Its first order derivative, $y^{(1)}(p)$, is a Gaussian function with a center position at p_c and a full-width-half-maximum (FWHM) of σ , which represents the bandwidth of spectral intensity change.

Fig. 2a shows the window size effect on intensity and shape of β calculated with simulated spectral data. When window size is narrower than the bandwidth of spectral change ($\sigma = 2$), the β profile becomes very close to the first order derivative plot (the dotted line). By contrast, as window size becomes larger than the bandwidth, the β profile becomes broader in shape. Figs. 2b and 2c display its peak height and FWHM as a function of window size

p . For the purpose of using β as a metric of the first order derivative, I suggest that a reference point $p = \sigma/2 = 1$, when the peak height of β is smaller by 4% than that of $y^{(1)}(p)$ and the FWHM of β is broader by 4% than that of $y^{(1)}(p)$.

2.2. Comparison of LS-MW with smoothed single-interval differentiation

Another method to measure the first order derivative is the single-interval differentiation, which calculates intensity change by each perturbation increment, also known as the gradient mapping method and simple numerical derivative. In fact, the gradient mapping method can be considered as an extreme case of LS-MW, where the smallest number ($= 2$) of spectra are used for determination of the slope of the change. Previously, Morita et al. reported the noise effect on single-interval derivatives of simulated data followed by Savitzky-Golay smoothing.[18] They found that smoothed single-interval differentiation is more susceptible to random noise than autocorrelation moving-window performed by the corresponding window size. Similarly, here smoothed single-interval differentiation is compared with LS-MW by using noise added intensity data. The average intensity of the simulation data is based on Eq. 9, and its perturbation variables are uniformly spaced. The noise is generated as a random number within $\pm 5\%$ of the maximum intensity, shown in Fig. 3a. By using the simulation data, a single-interval derivative plot is calculated (Fig. 3b). The derivative plot exhibits a significant amount of high-frequency noise, which requires a following smoothing process for practical uses. Fig. 3c shows a derivative plot after the plot of Figure 3b is smoothed by simple adjacent averaging by 10 points. It is noted that the 10-point adjacent smoothing indicates that 11 spectra per window are used for a single smoothed single-interval derivative value because of a prior 2-point differentiation process. For comparison, a linear coefficient plot from LS-MW are calculated with the same number of spectra per window, shown in Fig. 3d. It is clear that a much greater amount of high frequency noise still remains in the plot of smoothed single-interval differentiation than that of LS-MW. Degree of noise of a plot is measured by the root-mean-square (RMS) of difference between derivatives determined from the noise-including data and derivatives from the noise-free data. In Fig. 3e, the RMS determined from each derivative plot is plotted as a function of $(2m + 1)$. When $(2m + 1)$ is small, smoothed single-interval differentiation and LS-MW show similar RMS values to each other. As $(2m + 1)$ increases, the RMS of both smoothed single interval differentiation and LS-MW decreases but the RMS of LS-MW decreases more rapidly than that of smoothed single-interval differentiation. For example, at $2m + 1 = 21$, which corresponds to the half of the FWHM of the first order derivative of Eq. 9, the RMS of smoothed single-interval differentiation is twice larger than that of LS-MW. This comparison indicates that LS-MW performs better in noise suppression than single-interval differentiation followed by smoothing.

2.3. Comparison of LS-MW with other MW methods

In order to measure the first order derivative of spectral intensity with respect to perturbation variable, other MW methods have been used. For example, s -PC-MW intensity Π_Φ and the square root of AC-MW intensity $\sqrt{\Phi}$ have been used as they are known to be proportional to the first order derivative.[19] The numerical expressions of the two MW output intensities can be described in the following simple forms for spectral data with uniform perturbation spacing[19]

$$\Pi_{\Phi}(p_{ave}) = \frac{1}{2m+1} \sum_{j=j-m}^{j+m} \tilde{p}_j \tilde{y}_j, \quad (10)$$

$$\sqrt{\Phi}(p_{ave}) = \sqrt{\frac{1}{2m+1} \sum_{j=j-m}^{j+m} \tilde{y}_j^2}. \quad (11)$$

Figs. 2d–2i show $\sqrt{\Phi}$ and Π_{Φ} calculated for different window sizes by using the identical spectral data. Their results are compared side by side with the LS-MW results. The FWHM of both $\sqrt{\Phi}$ and Π_{Φ} in Figs. 2e and 2h are close to the FWHM σ of the original $y^{(1)}(p)$ and increases slowly with window size, which is similar to β in Fig. 2b. By contrast, the peak height of both $\sqrt{\Phi}$ and Π_{Φ} in Figs. 2f and 2i are quite different from that of β in Fig. 2c. The peak height of β decreases only slightly from the original peak height of $y^{(1)}(p)$ as window size increases while the peak height of $\sqrt{\Phi}$ and Π_{Φ} increases beginning from zero. These results suggest that β from LS-MW can represent the first order derivative both in intensity and shape while $\sqrt{\Phi}$ and Π_{Φ} from AC-MW and *s*-PC-MW can represent the first order derivative only in shape.

2.4. Effects of Non-Uniform Perturbation Spacing on LS-MW

The correlation-based moving-window analysis can be expressed in simple numerical forms when the perturbation variables are uniformly spaced, as shown in Eqs. 10 and 11 for PC-MW and AC-MW, respectively. Real data, however, are not always acquired with uniformly spaced perturbation variables. For example, actual temperatures measured while spectra are acquired can be different from their set temperatures due to various experimental reasons. Here, the effect of non-uniform spacing of perturbation variables on LS-MW is examined and compared with the effect on PC-MW and AC-MW. For comparison, identical simulation datasets are generated with two types of non-uniform perturbation spacing: pseudo-random spacing and increasing spacing. The model function used for simulation dataset with non-uniform spacing is the sum of two shifted error functions with identical bandwidths (illustrated as the green line in Fig. 1a.)

$$y(p) = \frac{1}{2} \operatorname{erfc} \left[-2 \sqrt{\ln 2} \left(\frac{p-2}{2} \right) \right] + \frac{1}{2} \operatorname{erfc} \left[-2 \sqrt{\ln 2} \left(\frac{p+2}{2} \right) \right]. \quad (12)$$

This function helps to elucidate the effect of increment size, in particular, for a dataset with increasingly spaced perturbation.

The numerical expression for LS-MW in Eq. 7 is valid for both uniformly and non-uniformly spaced data. However, Eqs. 10 and 11 for PC-MW and AC-MW are derived only for uniformly spaced datasets. Depending on a numerical integration model, their expression can be different for non-uniformly spaced data. For example, here a trapezoidal rule is used for Π_{Φ} from PC-MW as

$$\Pi_{\Phi}(p_{\text{ave}}) = \frac{1}{2(\tilde{p}_{j+m+1} - \tilde{p}_{j-m})} \sum_{j=j-m}^{j+m} (\tilde{p}_{j+1} \tilde{y}_{j+1} + \tilde{p}_j \tilde{y}_j)(\tilde{p}_{j+1} - \tilde{p}_j). \quad (13)$$

Similarly, by the trapezoidal rule, $\sqrt{\Phi}$ from AC-MW is expressed as

$$\sqrt{\Phi}(p_{\text{ave}}) = \sqrt{\frac{1}{2(\tilde{p}_{j+m+1} - \tilde{p}_{j-m})} \sum_{j=j-m}^{j+m} (\tilde{y}_{j+1}^2 + \tilde{y}_j^2)(\tilde{p}_{j+1} - \tilde{p}_j)}. \quad (14)$$

The calculation results of $\sqrt{\Phi}$ and Π_{Φ} , from non-uniformly spaced data are compared with those of LS-MW, shown in Fig. 4.

In Fig. 4, it should be noted that the LS-MW results from non-uniformly spaced data are almost identical to those from uniformly spaced data, which is very close to the first order derivative of the intensity function. The robustness of LS-MW to non-uniform perturbation spacing is quite contrasted to the PC-MW and AC-MW results. Even with the modifications of numerical expressions in Eqs. 13 and 14 for non-uniform spacing, the $\sqrt{\Phi}$ and Π_{Φ} profiles still exhibit noisy fluctuation from the data with pseudo-random spacing and noticeable shape distortion from the data with increasing spacing. These unwanted intensity noise and shape distortion of PC-MW and AC-MW outputs from non-uniform perturbation spacing is due to variation in the window size of moving windows.[19] To circumvent the complication due to non-uniform spacing, new uniformly spaced data are generated by interpolating the original non-uniformly spaced data. However, this data manipulation can cause unwanted bias and inconsistency in analysis depending on an interpolation model and a weighting method. On the contrary, LS-MW can be performed in a straightforward manner without need of any numerical integration or interpolation models for non-uniform spacing. Also, the linear coefficient β from LS-MW is very close to the first order derivative both in shape and in intensity whether data spacing is uniform or not.

2.5. Quadratic-function Least-Squares Moving-Window (q-LS-MW)

Earlier, the linear coefficient determined from *linear*-function LS-MW (*l*-LS-MW) is characterized. In this section, the two meaningful coefficients from *quadratic*-function LS-MW (*q*-LS-MW) are examined. Similar to the formulations for *l*-LS-MW, a quadratic model function is used for least-squares regression as

$$\tilde{y}_j = (\alpha + \beta \tilde{p}_j + \gamma \tilde{p}_j^2) + e_j. \quad (15)$$

The detailed derivation for the least-squares regression with a quadratic model function is described in the Appendix A. The three coefficients obtained from *q*-LS-MW are

$$\alpha = \frac{-(\sum \tilde{p}_j^2) \{ \sum \tilde{p}_j^2 \sum \tilde{p}_j^2 \tilde{y}_j - \sum \tilde{p}_j^2 \sum \tilde{p}_j \tilde{y}_j \}}{(2m+1) \sum \tilde{p}_j^2 \sum \tilde{p}_j^4 - (2m+1) (\sum \tilde{p}_j^3)^2 - (\sum \tilde{p}_j^2)^3}, \quad (16)$$

$$\beta = \frac{(2m+1) \sum \tilde{p}_j^4 \sum \tilde{p}_j \tilde{y}_j - (\sum \tilde{p}_j^2)^2 \sum \tilde{p}_j \tilde{y}_j - (2m+1) \sum \tilde{p}_j^3 \sum \tilde{p}_j^2 \tilde{y}_j}{(2m+1) \sum \tilde{p}_j^2 \sum \tilde{p}_j^4 - (2m+1) (\sum \tilde{p}_j^3)^2 - (\sum \tilde{p}_j^2)^3}, \quad (17)$$

$$\gamma = \frac{(2m+1) \{ \sum \tilde{p}_j^2 \sum \tilde{p}_j^2 \tilde{y}_j - \sum \tilde{p}_j^3 \sum \tilde{p}_j \tilde{y}_j \}}{(2m+1) \sum \tilde{p}_j^2 \sum \tilde{p}_j^4 - (2m+1) (\sum \tilde{p}_j^3)^2 - (\sum \tilde{p}_j^2)^3}. \quad (18)$$

It is noted that the y -intercept α in Eq. 16 is negatively proportional to the quadratic coefficient γ in Eq. 18 as $\alpha = - \left[\sum \tilde{p}_j^2 / (2m+1) \right] \gamma$. Therefore, only the two independent parameters, β and γ , remain to be determined by q -LS-MW.

The expression of Eq. 17 for the linear coefficient β from q -LS-MW is more complicated than the form of Eq. 7 for the linear coefficient β from I -LS-MW. However, Fig. 5 shows that their actual values are very close to each other both for uniform and non-uniform spacing. In fact, the two linear coefficients from q -LS-MW and I -LS-MW are mathematically identical when the perturbation spacing is uniform, whose derivation is presented in the Appendix B. Both linear coefficients are very close to the first order derivative $y^{(1)}(p)$ with difference $< 1\%$ for the simulation datasets including the slowly varying double dip caused by finite moving window size. It is interesting that β from q -LS-MW is better to suppress noise generated from perturbation spacing fluctuation than β from I -LS-MW.

The quadratic coefficient γ from q -LS-MW can be easily related with the second order derivative if an intensity function is assumed to be a second order truncated Taylor series as

$$\tilde{y}_j = y(p_{ave}) + y^{(1)}(p_{ave}) \tilde{p}_j + \frac{y^{(2)}(p_{ave})}{2} \tilde{p}_j^2. \quad (19)$$

By inserting Eq. 19 into Eq. 18, the γ value determined from q -LS-MW can be expressed with the second order derivative as

$$\gamma = y^{(2)}(p_{ave})/2. \quad (20)$$

Fig. 6a shows γ profiles from q -LS-MW with various moving window sizes. Similar to the window size effect of the linear coefficient β from I -LS-MW in Fig. 2a, γ from q -LS-MW

performed with a moving window narrower than $\sigma/2$ ($= 1$) turns out to be close to the asymptotic value of $y^{(2)}(p_{\text{ave}})/2$ within 5 % of the maximum intensity.

The shape of the second order derivative can also be represented by the negative of Π_{Ψ} from asynchronous PC-MW (*as*-PC-MW).[9,19] Here, the quadratic coefficient γ profiles from *q*-LS-MW is compared with the negative of Π_{Ψ} from *as*-PC-MW. For uniformly-spaced perturbation datasets, Π_{Ψ} can be expressed in a simplified form as[19]

$$\Pi_{\Psi} = \frac{1}{2m+1} \sum_{J=j-m}^{j+m} \tilde{y}_J \sum_{K=j-m}^{j+m} M_{JK} \tilde{p}_K, \quad (21)$$

where

$$M_{JK} = \begin{cases} 0 & J=K \\ \frac{1}{\pi(K-J)} & J \neq K \end{cases} \quad (22)$$

is the discrete Hilbert-Noda transformation matrix.[23] For a dataset with non-uniform spacing, a trapezoidal rule is used to re-express Π_{Ψ} as

$$\Pi_{\Psi} = \frac{1}{2(\tilde{p}_{j+m+1} - \tilde{p}_{j-m})} \sum_J \left(\tilde{y}_{J+1} \sum_{K=j-m}^{j+m} M_{J+1,K} \tilde{p}_K + \tilde{y}_J \sum_{K=j-m}^{j+m} M_{JK} \tilde{p}_K \right) (\tilde{p}_{J+1} - \tilde{p}_J). \quad (23)$$

Fig. 5b shows Π_{Ψ} calculated by Eqs. 21 and 22 from a uniformly-spaced dataset with various moving window sizes, which shows a significant moving window size effect on the Π_{Ψ} intensity. This dependence is already discussed in the previous report, where the shape of Π_{Ψ} is close to that of $y^{(2)}(p)$ and much less sensitive to moving window size.[19]

Fig. 7 shows both γ and Π_{Ψ} calculated with datasets with non-uniform spacing, which are identical to the datasets used for Fig. 4. The γ profiles from *q*-LS-MW for non-uniform spacing are as close to $y^{(2)}(p)$ as that for uniform spacing. By contrast, the negative Π_{Ψ} calculated with Eq. 23 for the non-uniformly spaced datasets are quite different not only from $y^{(2)}(p)$ but also from the negative Π_{Ψ} for uniform spacing. These features of γ and Π_{Ψ} are overall similar to those of β and Π_{Φ} from *ALS*-MW and *s*-PC-MW, respectively. While *as*-PC-MW results are strongly affected by non-uniformity of perturbation variable, *q*-LS-MW results are not affected noticeably. While only the shape of Π_{Ψ} from *as*-PC-MW is similar to that of $y^{(2)}(p)$, both the shape and absolute intensity of γ from *q*-LS-MW are close to those of $y^{(2)}(p)$. This re-confirms that the quadratic coefficient from *q*-LS-MW can be used as a quantitative, robust metric for the second order derivative whether a dataset is uniformly spaced or not.

3. Conclusion

Least-squares moving-window analysis has been described as a numerical differentiation method for spectral analysis. The effects on moving-window sizes, different perturbation spacing types, and intensity noises have been discussed by using simulation data. Unlike the original Savitzky-Golay differentiation method, the LS-MW can be used for both uniform and non-uniform perturbation spacing. The linear and quadratic coefficients determined from the LS-MW method have been compared with outputs from other widely used numerical differentiation methods, including single-interval differentiation, autocorrelation moving-window and perturbation-correlation moving-window methods. Those comparisons suggest the superior performance of LS-MW in suppression of high frequency noise and in versatility for both uniformly and non-uniformly spaced datasets.

References

1. Noda, I., Ozaki, Y. Two-Dimensional Correlation Spectroscopy. Wiley & Sons; Chichester, West Sussex: 2004.
2. Noda I. Generalized two-dimensional correlation method applicable to infrared, Raman, and other types of spectroscopy. *Appl Spectrosc.* 1993; 47:1329–1336.
3. Thomas M, Richardson HH. Two-dimensional FT-IR correlation analysis of the phase transitions in a liquid crystal, 4'-n-octyl-4-cyanobiphenyl (8CB). *Vib Spectrosc.* 2000; 24:137–146.
4. Šaši S, Katsumoto Y, Sato H, Ozaki Y. Applications of Moving Window Two-Dimensional Correlation Spectroscopy to Analysis of Phase Transitions and Spectra Classification. *Anal Chem.* 2003; 75:4010–4018. [PubMed: 14632112]
5. Jin Y, Kotula AP, Hight Walker AR, Migler KB, Lee YJ. Phase-specific Raman analysis of n-alkane melting by moving-window two-dimensional correlation spectroscopy. *J Raman Spectrosc.* 2016; 47:1375–1384. [PubMed: 28070140]
6. Peng L, Zhou T, Huang Y, Jiang L, Dan Y. Microdynamics mechanism of thermal-induced hydrogel network destruction of poly(vinyl alcohol) in D2O studied by two-dimensional infrared correlation spectroscopy. *J Phys Chem B.* 2014; 118:9496–9506. [PubMed: 25046657]
7. Batista ANL, Batista JM, Ashton L, Bolzani VS, Furlan M, Blanch EW. Investigation of DMSO-Induced Conformational Transitions in Human Serum Albumin Using Two-Dimensional Raman Optical Activity Spectroscopy. *Chirality.* 2014; 26:497–501. [PubMed: 25042763]
8. Ashton L, Blanch EW. pH-induced conformational transitions in α -lactalbumin investigated with two-dimensional Raman correlation variance plots and moving windows. *J Mol Struct.* 2010; 974:132–138.
9. Morita S, Shinzawa H, Noda I, Ozaki Y. Perturbation-Correlation Moving-Window Two-Dimensional Correlation Spectroscopy. *Appl Spectrosc.* 2006; 60:398–406. [PubMed: 16613636]
10. Watanabe A, Morita S, Ozaki Y. Study on temperature-dependent changes in hydrogen bonds in cellulose I β by infrared spectroscopy with perturbation-correlation moving-window two-dimensional correlation spectroscopy. *Biomacromolecules.* 2006; 7:3164–3170. [PubMed: 17096547]
11. Gomez De La Cuesta R, Goodacre R, Ashton L. Monitoring antibody aggregation in early drug development using raman spectroscopy and perturbation-correlation moving windows. *Anal Chem.* 2014; 86:11133–11140. [PubMed: 25329604]
12. Nishii T, Morita S, Genkawa T, Watari M, Ishikawa D, Ozaki Y. Moving-window two-dimensional heterospectral (MW2DHetero) correlation analysis and its application for the process monitoring of alcoholic fermentation. *Appl Spectrosc.* 2015; 69:665–670. [PubMed: 25955516]
13. Ye Z, Li Y, An Z, Wu P. Exploration of Doubly Thermal Phase Transition Process of PDEGA-b-PDMA-b-PVCL in Water. *Langmuir.* 2016; 32:6691–6700. [PubMed: 27299984]
14. Li C, Liu J, Li J, Shen F, Huang Q, Xu H. Studies of 4,4'-diphenylmethane diisocyanate (MDI)/1,4-butanediol (BDO) based TPUs by in situ and moving-window two-dimensional correlation

infrared spectroscopy: Understanding of multiple DSC endotherms from intermolecular interactions and motions level. *Polymer*. 2012; 53:5423–5435.

15. Chen Y, Mu T. A new way to interpret perturbation-correlation moving-window two-dimensional correlation spectroscopy: probing the dynamic interaction of ionic liquid 1-ethyl-3-methylimidazolium acetate to absorb atmospheric water. *Appl Spectrosc*. 2015; 69:130–42. [PubMed: 25280060]
16. Unger M, Morita S, Sato H, Ozaki Y, Siesler HW. Variable-Temperature Fourier Transform Infrared Spectroscopic Investigations of Poly(3-Hydroxyalkanoates) and Perturbation-Correlation Moving-Window Two-Dimensional Correlation Analysis. Part I: Study of Non-Annealed and Annealed Poly(3-Hydroxybutyrate). *Appl Spectrosc*. 2009; 63:1027–1033. [PubMed: 19796485]
17. Ryu SR, Noda I, Lee CH, Lee PH, Hwang H, Jung YM. Two-dimensional correlation analysis and waterfall plots for detecting positional fluctuations of spectral changes. *Appl Spectrosc*. 2011; 65:359–368. [PubMed: 21396181]
18. Morita S, Shinzawa H, Tsenkova R, Noda I, Ozaki Y. Computational simulations and a practical application of moving-window two-dimensional correlation spectroscopy. *J Mol Struct*. 2006; 799:111–120.
19. Lee YJ. Analytical and Numerical Characterization of Autocorrelation and Perturbation-Correlation Moving-Window Methods. *Appl Spectrosc*. 2016 in press.
20. Jung YM, Shin HS, Kim SB, Noda I. Two-Dimensional Gradient Mapping Technique Useful for Detailed Spectral Analysis of Polymer Transition Temperatures. *J Phys Chem B*. 2008; 112:3611–3616. [PubMed: 18318520]
21. Savitzky A, Golay MJE. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Anal Chem*. 1964; 36:1627–1639.
22. Chapra, SC., Canale, RP. Numerical methods for engineers: with programming and software applications. 3. WCB/McGraw-Hill; Boston: 1998.
23. Noda I. Determination of two-dimensional correlation spectra using the Hilbert transform. *Appl Spectrosc*. 2000; 54:994–998.

Appendix A. Derivation of the linear and quadratic coefficients from a quadratic modeled least-squares regression

A quadratic model function for least-squares fitting is expressed as

$$\tilde{y}_j = (\alpha + \beta \tilde{p}_j + \gamma \tilde{p}_j^2) + e_j. \quad (\text{A1})$$

Then, the coefficients, α , β , and γ , can be determined by solving the equations so that all partial derivatives of SSE equal zero as

$$0 = \frac{\partial SSE}{\partial \alpha} = \sum \frac{\partial}{\partial \alpha} (\alpha + \beta \tilde{p}_j + \gamma \tilde{p}_j^2 - \tilde{y}_j)^2 = 2 \sum (\alpha + \beta \tilde{p}_j + \gamma \tilde{p}_j^2 - \tilde{y}_j), \quad (\text{A2})$$

$$0 = \frac{\partial SSE}{\partial \beta} = \sum \frac{\partial}{\partial \beta} (\alpha + \beta \tilde{p}_j + \gamma \tilde{p}_j^2 - \tilde{y}_j)^2 = 2 \sum (\alpha + \beta \tilde{p}_j + \gamma \tilde{p}_j^2 - \tilde{y}_j) \tilde{p}_j, \quad (\text{A3})$$

$$0 = \frac{\partial SSE}{\partial \gamma} = \sum \frac{\partial}{\partial \gamma} \left(\alpha + \beta \tilde{p}_J + \gamma \tilde{p}_J^2 - \tilde{y}_J \right)^2 = 2 \sum \left(\alpha + \beta \tilde{p}_J + \gamma \tilde{p}_J^2 - \tilde{y}_J \right) \tilde{p}_J^2. \quad (\text{A4})$$

After $\sum \tilde{p}_J = \sum \tilde{y}_J = 0$ are applied from the mean-centering process, these equations can be rearranged in the form of matrices.

$$\begin{pmatrix} 2m+1 & 0 & \sum \tilde{p}_J^2 \\ 0 & \sum \tilde{p}_J^2 & \sum \tilde{p}_J^3 \\ \sum \tilde{p}_J^2 & \sum \tilde{p}_J^3 & \sum \tilde{p}_J^4 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} 0 \\ \sum \tilde{p}_J \tilde{y}_J \\ \sum \tilde{p}_J^2 \tilde{y}_J \end{pmatrix}, \quad (\text{A5})$$

where $(2m+1)$ is the number of data of the window. Then, the unknown coefficients can be calculated as

$$\begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \mathbf{M}^{-1} \begin{pmatrix} 0 \\ \sum \tilde{p}_J \tilde{y}_J \\ \sum \tilde{p}_J^2 \tilde{y}_J \end{pmatrix}, \quad (\text{A6})$$

where $\mathbf{M}$ is the 3 x 3 matrix in Eq. A5. Then, the determinant of $\mathbf{M}$ is expressed as

$$\det(\mathbf{M}) = (2m+1) \sum \tilde{p}_J^2 \sum \tilde{p}_J^4 - (2m+1) \left(\sum \tilde{p}_J^3 \right)^2 - \left(\sum \tilde{p}_J^2 \right)^3. \quad (\text{A7})$$

And the individual coefficients are

$$\alpha = -\frac{\sum \tilde{p}_J^2}{\det(\mathbf{M})} \left\{ \sum \tilde{p}_J^2 \sum \tilde{p}_J^2 \tilde{y}_J - \sum \tilde{p}_J^3 \sum \tilde{p}_J \tilde{y}_J \right\}, \quad (\text{A8})$$

$$\beta = \frac{1}{\det(\mathbf{M})} \left\{ (2m+1) \sum \tilde{p}_J^4 \sum \tilde{p}_J \tilde{y}_J - \left(\sum \tilde{p}_J^2 \right)^2 \sum \tilde{p}_J \tilde{y}_J - (2m+1) \sum \tilde{p}_J^3 \sum \tilde{p}_J^2 \tilde{y}_J \right\}, \quad (\text{A9})$$

$$\gamma = \frac{2m+1}{\det(\mathbf{M})} \left\{ \sum \tilde{p}_J^2 \sum \tilde{p}_J^2 \tilde{y}_J - \sum \tilde{p}_J^3 \sum \tilde{p}_J \tilde{y}_J \right\}. \quad (\text{A10})$$

Appendix B. Equivalency of the linear coefficients β from a quadratic model and a linear model least-squares fitting for uniform perturbation spacing

The linear coefficient from q -LS-MW, β_q , of Eq. A9 is expressed in a much more complicated form than the linear coefficient from l -LS-MW, β_l , is from Eq. 7

$$\beta = \frac{\sum_{j=j-m}^{j+m} \tilde{p}_j \tilde{y}_j}{\sum_{j=j-m}^{j+m} \tilde{p}_j^2}. \quad (\text{A11})$$

However, when perturbation spacing is uniform, the complicated expression of Eq. A9 can be simplified. First, the mean-centered perturbation variable can be expressed as $\tilde{p}_j = dJ$, where d is the constant increment. Then,

$$\sum_{j=j-m}^{j+m} \tilde{p}_j^2 = d^2 \sum_{J=-m}^m J^2 = 2d^2 \sum_{J=1}^m J^2, \quad (\text{A12})$$

$$\sum_{j=j-m}^{j+m} \tilde{p}_j^3 = d^3 \sum_{J=-m}^m J^3 = 0, \quad (\text{A13})$$

$$\sum_{j=j-m}^{j+m} \tilde{p}_j^4 = d^4 \sum_{J=-m}^m J^4 = 2d^4 \sum_{J=1}^m J^4. \quad (\text{A14})$$

The sum of the squares of the first n natural numbers can be calculated by using a summation method as

$$\sum_{k=1}^n k^3 = \sum_{k=0}^n (k+1)^3 - (n+1)^3. \quad (\text{A15})$$

This leads to

$$\sum_{k=1}^n k^2 = \frac{1}{6} n(n+1)(2n+1). \quad (\text{A16})$$

Similarly, $\sum_{k=1}^n k^3$ can be calculated as

$$\sum_{k=1}^n k^3 = \frac{1}{30} n(n+1)(2n+1)(3n^2+3n-1). \quad (\text{A17})$$

By inserting Eqs. A16 and A17 into Eq. A11, the linear coefficient β from the *linear* model LS-MW is expressed as

$$\beta = \frac{\sum \tilde{p}_j \tilde{y}_j}{\sum \tilde{p}_j^2} = \frac{3}{d^2 m(m+1)(2m+1)} \sum \tilde{p}_j \tilde{y}_j. \quad (\text{A18})$$

For the linear coefficient β from the *quadratic* model LS-MW, Eqs. A16 and A17 are inserted into Eq. A7, then

$$\det(\mathbf{M}) = \frac{1}{135} m^2 (m+1)^2 (2m+1)^3 (2m-1)(2m+3), \quad (\text{A19})$$

and the numerator $\{ \}$ in Eq. A9 becomes

$$\{ \} = \frac{1}{45} m(m+1)(2m+1)^2 (2m-1)(2m+3) \sum \tilde{p}_j \tilde{y}_j. \quad (\text{A20})$$

By inserting Eqs. A19 and A20 into Eq. A9, β from the quadratic model LS-MW can be expressed as

$$\beta = \frac{3}{d^2 m(m+1)(2m+1)} \sum \tilde{p}_j \tilde{y}_j, \quad (\text{A21})$$

which is identical to Eq. A18, from the linear model LS-MW.

Fig. 1.

(a) Illustration of a moving-window analysis of a series of spectra. (b) The spectral intensity $y(p)$ at $\nu = 700$ is plotted as a function of perturbation variable p . (c) An example of linear-function least-squares moving window (*L*-LS-MW) analysis, where the mean-centered variables ($\tilde{y}_j$ vs. $\tilde{p}_j$) of a moving window are least-squares fitted with a linear model function.

Fig. 2.

(a) Linear coefficient, β , from LS-MW is plotted as a function of average perturbation variable, p_{ave} , for various window sizes, Δp . The first order derivative, $y^{(1)}(p)$, are plotted as a reference (the black dotted line). The model intensity function, $y(p)$, is based on Eq. 9 with uniform perturbation spacing, d , of 0.1. For the uniformly spaced data, p is calculated from $d(2m+1)$, where $(2m+1)$ is the number of spectra per window. (b) Full-width-half-maximum (FWHM) of the calculated β is plotted as a function of Δp (the red solid line). For comparison, FWHM of adjacent-averaged $y^{(1)}(p)$ by $(2m+1)$ is plotted as a function of Δp (the red dash-dotted line). The black dotted line at 2 indicates the constant FWHM of the original $y^{(1)}(p)$. (c) Peak height of the calculated β is plotted as a function of Δp (the blue solid line). For comparison, the peak height of the adjacent-averaged $y^{(1)}(p)$ is plotted as a function of Δp (the blue dash-dotted line). The black dotted line at $\sqrt{(\ln 2)/\pi}$ indicates the peak height of the original $y^{(1)}(p)$. (d)–(f) Synchronous perturbation-correlation moving-window (s -PC-MW) intensity, Π_{Φ} , is calculated with Eq. 10 from the identical dataset for (a). (g)–(i) The square root of autocorrelation moving-window (AC-MW) intensity, $\sqrt{\Phi}$, is calculated with Eq. 11 from the identical dataset for (a).

Fig. 3.

(a) A plot of simulated intensity based on Eq. 9 added with a random number generated between -0.05 and $+0.05$. The perturbation spacing is set to be uniform with the increment of 0.05 . The dotted line indicates noise-free intensity data for comparison. (b) A single-interval differentiation plot of the intensity data from (a). (c) The single-interval differentiation plot of (b) is smoothed by adjacent averaging by 10 points. (d) A linear coefficient plot from LS-MW with $2m + 1 = 11$. The black dotted lines in (b)–(d) indicates plots calculated from noise-free data intensity data of (a). A s -PC-MW intensity (Π_{Φ}) plot becomes overlapped with the linear coefficient plot from LS-MW when their peak height is normalized. (e) Root-mean-square (RMS) is calculated from the difference between the derivative intensities determined from the noise-added and the noise-free intensity data. The RMS for both smoothed single-interval differentiation and LS-MW are plotted as a function of $(2m + 1)$.

Fig. 4.

Comparison of the non-uniform spacing effect on LS-MW, PC-MW, and AC-MW. Two sets of perturbation variables are prepared for non-uniform perturbation spacing: pseudo-random spacing (the left column) and increasing spacing (the right column). Spectral intensity is generated based on Eq. 12 from each set of perturbation variables. (a) Perturbation spacing is determined by a random number generated in the range between 0.05 and 0.15. (b) Perturbation spacing increases linearly from 0.072 to 0.122 at $p = -6$ to 6, respectively. At $p = 0$, the increment is set to be 0.1. Each MW result from the non-uniform perturbation spacing (the red solid lines) is compared with the corresponding MW result from the uniform perturbation spacing data with a constant spacing of 0.1 (the black dashed lines). For all calculations, $(2m + 1)$ is fixed as 5.

Fig. 5.

Comparison between the linear coefficients β determined from q -LS-MW and l -LS-MW. The identical simulation datasets are used for calculation as in Fig. 4. The difference profiles are calculated by subtracting the first order derivative value from each β . For all calculations, the number of spectra per window $2m + 1$ is fixed as 5.

Fig. 6. Window size effect on (a) the quadratic coefficient γ from q -LS-MW and (b) Π_{Ψ} intensity from as -PC-MW. The uniformly-spaced intensity data used for Fig. 5a are used for calculation. The indicated window size p of 0.5, 1.1, and 2.1 are calculated from $(2m + 1)$ of 5, 11, and 21, respectively.

Fig. 7.

Comparison of the non-uniform spacing effect on the quadratic coefficient γ from q -LS-MW and Π_{Ψ} intensity from as -PC-MW. The identical intensity data used for Fig. 4 are used for calculation. Each MW result from the non-uniform perturbation spacing (the solid magenta lines) is compared with the MW result from the uniform perturbation spacing (the black dashed lines). For all calculations, $(2m + 1)$ is fixed as 5.