



Nonlinear least squares with local polynomial interpolation for quantitative analysis of IR spectra

Yifan Wu^{a,b}, Silong Peng^{a,b,*}, Qiong Xie^a, Pengcheng Xu^c

^aInstitute of Automation, Chinese Academy of Sciences, Beijing 100190, China

^bUniversity of Chinese Academy of Sciences, Beijing 100190, China

^cAcademy of Mathematics and Systems Science, Chinese Academy of Sciences, 100190 Beijing, China

ARTICLE INFO

Article history:

Received 8 April 2018

Received in revised form 23 July 2018

Accepted 3 August 2018

Available online 4 August 2018

Keywords:

Spectroscopy

Quantitative analysis

Nonlinear least squares

Local polynomial interpolation

ABSTRACT

When using spectroscopic instrumentation for quantitative analysis of mixture, spectral intensity non-linearity and peak shift make it challenging for building calibration model. In this study, we investigated the performance of a nonlinear model, namely nonlinear least squares with local polynomial interpolation (NLSLPI). In NLSLPI, the parameters to be optimized are the concentrations of the components. Levenberg-Marquardt (L-M) method is used to solve the nonlinear-least-squares optimization problem and local polynomial interpolation is used to generate the nonlinear function for each component. We tested the robustness of NLSLPI on a computer-simulation dataset. We also compared NLSLPI, in terms of RMSEP, to partial least squares (PLS), classical least squares (CLS) and piecewise classical least squares (PCLS) on a real-world dataset. Experimental results demonstrate the effectiveness of the proposed method.

© 2018 Elsevier B.V. All rights reserved.

1. Introduction

Infrared (IR) spectroscopic analysis, whose advantages are no requirement for sample preparation, fast response and high accuracy, has been used for the quantitative analysis of mixture in wide applications [1–3]. In order to estimate the concentration of an interesting component from a mixture spectrum which usually contains hundreds of measurement values, chemometric algorithms are developed.

Usually, two situations will be met when dealing with IR spectroscopy. The first one is that only one or more components among the mixture are interested and most information (such as the reference spectrum of the interested component) are not obtainable. To deal with this case, a large amount of representative samples of mixture together with the concentrations of the interested component are needed. Inverse modeling methods such as principle component regression (PCR) [4] and partial least squares (PLS) [5] are usually applied to predict the concentrations of the future samples. The second case is that the mixture contains several major components and the reference spectra of these components are known. In this case, classical least squares (CLS) [6,7] is usually applied to simulate the spectrum of the mixture. Assuming that the mixture spectrum

is a linear combination of reference spectra, CLS works by finding the concentrations that minimizes difference between the measured spectrum and the simulated spectrum. When the linear assumption failed, the piecewise classical least squares (PCLS) was used [8,9]. In PCLS, there are multiple reference spectra for each target component and then multiple CLS sub-models are generated. The two sub-models that predict concentrations in the measured spectrum nearest to those of specific reference spectra included in the set of sub-models are selected. The prediction value of PCLS is a weighted sum of the values of the selected sub-models. Even though CLS is widely used in commercial chemometric software, the investigation of its nonlinear extension is rare.

In this study, we focus on the second case with nonlinear assumption. The nonlinearity is presented in two aspects [10,11]. The first is a nonlinear relationship between spectral intensity and concentration. The second is peak shift with increasing concentration. In our point of view, CLS can still be used after some mathematical transformation when only the first nonlinearity presents. When the second nonlinearity or a combination of them present, nonlinear least squares with local polynomial interpolation (NLSLPI) is introduced to deal with quantitative analysis. Nonlinear least squares (NLS) has been proved to be a powerful tool for handling nonlinear systems [12,13]. In this study, NLS was employed to estimate the concentrations of the interested components. The Levenberg-Marquardt (L-M) method was used to solve the NLS problem. When using L-M, a difficulty is to estimate the function value together with

* Corresponding author.

E-mail address: silong.peng@ia.ac.cn (S. Peng).

its derivative, which are unknown usually. In this study, local polynomial interpolation is used to generate the nonlinear function for each component.

In the following, we will discuss the nonlinear system and introduce NLSLPI in Section 2. Experimental data and results are presented in Sections 3 and 4 respectively. Finally, we make conclusions of NLSLPI in Section 5.

2. Methodology

2.1. Nonlinear Model of Quantitative Analysis

Given that the mixture is composed of J components, whose spectral responses are linear to the concentrations, CLS can be used to predict the concentration of each component from the spectrum of the mixture:

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \left\| \mathbf{s} - \sum_{j=1}^J x_j \mathbf{f}_j \right\|_2^2 \quad (1)$$

where the mixture spectrum \mathbf{s} is row vector of length p . x_j and \mathbf{f}_j represent the concentration and the reference spectrum for component j . $\mathbf{x} = [x_1, x_2, \dots, x_j]^T$ is the concentration vector.

If the linearity condition can not be guaranteed, for example, the concentration falls outside the linear range, Eq. (1) should be modified:

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \left\| \mathbf{s} - \sum_{j=1}^J \tilde{\mathbf{f}}_j(x_j) \right\|_2^2 \quad (2)$$

where the spectral response function $\tilde{\mathbf{f}}_j(x_j)$ can be any function of the concentration, not necessarily linear one. When spectral response function exhibits the first nonlinearity, we only need to use a nonlinear function $g_j(x_j)$ to represent the nonlinear relationship:

$$\tilde{\mathbf{f}}_j(x_j) = g_j(x_j) \mathbf{f}_j \quad (3)$$

In practice, $g_j(x_j)$ can be generated by interpolating to the reference spectra with various concentrations of component j . Therefore, instead of estimating true concentration directly, we estimate pseudo-concentration first:

$$\hat{\mathbf{g}} = \underset{\mathbf{g}}{\operatorname{argmin}} \left\| \mathbf{s} - \sum_{j=1}^J g_j \mathbf{f}_j \right\|_2^2 \quad (4)$$

where $\mathbf{g} = [g_1, g_2, \dots, g_j]^T$ is the pseudo-concentration vector. Then the true concentration can be obtained by:

$$\hat{x}_j = g_j^{-1}(\hat{g}_j), j = 1, 2, \dots, J \quad (5)$$

where g_j^{-1} is the inverse function of g_j .

When the mixture exhibits the combination of two nonlinearity, the above linear squares method fails. Then the nonlinear least squares method must be developed to predict the components of the mixture. In this study, Levenberg-Marquardt (L-M) method is employed.

2.2. Levenberg-Marquardt Method

L-M method is a classical method for solving nonlinear least squares, which is based on the trust-region framework [14]. In this

study, L-M method is used for minimizing the following objective function:

$$l(\mathbf{x}) = \|\mathbf{s} - \mathbf{F}(\mathbf{x})\|_2^2 \quad (6)$$

where $\mathbf{F}(\mathbf{x}) = \sum_{j=1}^J \tilde{\mathbf{f}}_j(x_j)$ is the sum of the pure component spectra.

L-M method performs optimization in a stepwise manner. Suppose that in iteration k , the estimation is \mathbf{x}^k , according to L-M method the solution for updating \mathbf{x}^k can be calculated as following:

$$\delta = (\mathbf{J}^{kT} \mathbf{J}^k + \lambda \mathbf{I})^{-1} \mathbf{J}^{kT} (\mathbf{s} - \mathbf{F}^k) \quad (7)$$

where δ is the solution for updating \mathbf{x}^k and $\mathbf{F}^k = \mathbf{F}(\mathbf{x}^k)$. $\mathbf{J}(\mathbf{x}) = \frac{\partial \mathbf{F}}{\partial \mathbf{x}}$ is first derivative matrix and $\mathbf{J}^k = \mathbf{J}(\mathbf{x}^k)$. λ is the regularization parameter and it is critical to the convergence speed of the optimization. When λ is large enough, $(\mathbf{J}^{kT} \mathbf{J}^k + \lambda \mathbf{I})^{-1} \approx 1/\lambda$ and δ can be regarded yielded by Steepest Decent method. In this case, δ guarantees the reduction of the objective function but with slow convergence speed. When $\lambda = 0$, δ can be regarded yielded by Quasi Newton method. In this case, $\mathbf{J}^{kT} \mathbf{J}^k$ is not necessarily invertible and does not guarantees the reduction of the objective function. However, when the reduction condition is fulfilled, Quasi Newton method converges to the minimum much faster than Steepest Decent method. In practice, λ usually decreases from a relatively large value and we check the reduction condition after each decrease. The smallest value which fulfills the reduction condition is selected as the final solution of λ .

After obtaining δ , we update \mathbf{x}^k by:

$$\mathbf{x}^{k+1} = \mathbf{x}^k + \delta \quad (8)$$

2.3. Local Polynomial Interpolation

In order to implement L-M method in our study, $\mathbf{F}(\mathbf{x})$ and $\mathbf{J}(\mathbf{x})$ are obtained by interpolating to reference spectra with various concentrations. Polynomial interpolation is a common method for interpolation [15]. Suppose that $(n + 1)$ reference spectra are available, then polynomial interpolation with degree of n can be obtained:

$$\tilde{\mathbf{f}}_j(x_j) = \mathbf{a}_{j0} + x_j \mathbf{a}_{j1} + x_j^2 \mathbf{a}_{j2} \dots + x_j^n \mathbf{a}_{jn}, \quad j = 1, 2, \dots, J \quad (9)$$

where $\mathbf{a}_{j0}, \mathbf{a}_{j1}, \dots, \mathbf{a}_{jn}$ are the polynomial coefficients for component j . The derivative of Eq. (9) is:

$$\frac{\partial \tilde{\mathbf{f}}_j}{\partial x_j} x_j = \mathbf{a}_{j1} + 2x_j \mathbf{a}_{j2} + \dots + nx_j^{n-1} \mathbf{a}_{jn}, \quad j = 1, 2, \dots, J \quad (10)$$

Therefore, we can obtain:

$$\begin{aligned} \mathbf{F}(\mathbf{x}) &= \sum_{j=1}^J \tilde{\mathbf{f}}_j(x_j) \\ \mathbf{J}(\mathbf{x}) &= \frac{\partial \mathbf{F}}{\partial \mathbf{x}} \\ &= \begin{bmatrix} \frac{\partial \tilde{\mathbf{f}}_1}{\partial x_1} & \frac{\partial \tilde{\mathbf{f}}_2}{\partial x_2} & \dots & \frac{\partial \tilde{\mathbf{f}}_J}{\partial x_J} \end{bmatrix} \end{aligned} \quad (11)$$

Local polynomial interpolation (LPI) aims to find the closest data points to perform interpolation with low degree polynomial. For example, when three is set to be the polynomial degree, four pure component samples, $(x_j^{(1)}, \tilde{\mathbf{f}}_j^{(1)})$, $(x_j^{(2)}, \tilde{\mathbf{f}}_j^{(2)})$, $(x_j^{(3)}, \tilde{\mathbf{f}}_j^{(3)})$, $(x_j^{(4)}, \tilde{\mathbf{f}}_j^{(4)})$, which satisfy the condition of $x_j^{(1)} \leq x_j^{(2)} \leq x_j \leq x_j^{(3)} \leq x_j^{(4)}$ are located, and then Eqs. (9) and (10) are employed. If x_j is close to the endpoints,

Download English Version:

<https://daneshyari.com/en/article/7667185>

Download Persian Version:

<https://daneshyari.com/article/7667185>

[Daneshyari.com](https://daneshyari.com)