



Asymmetric least squares for multiple spectra baseline correction

Jiangtao Peng^{a,*}, Silong Peng^a, An Jiang^a, Jiping Wei^b, Changwen Li^b, Jie Tan^a

^a Institute of Automation, Chinese Academy of Sciences, Beijing 100190, PR China

^b Food Research Institute of Tianshili Group, Tianjin 300410, PR China

ARTICLE INFO

Article history:

Received 4 March 2010

Received in revised form 7 August 2010

Accepted 24 August 2010

Available online 15 October 2010

Keywords:

Asymmetric least squares

Multiple spectra baseline correction

Relaxation factor

ABSTRACT

In this paper, based on asymmetric least squares smoothing, a new algorithm for multiple spectra baseline correction is proposed. By means of the similarity among the multiple spectra, the algorithm estimates the baselines by penalizing the differences in the baseline corrected signals, which makes the algorithm possible to eliminate scatter effects on the spectra. In addition, a relaxation factor which measures the similarity of the baseline corrected spectra is incorporated into the optimization model and an alternate iteration strategy is used to solve the optimization problem. The proposed algorithm is fast and can output multiple baselines simultaneously. Experimental results on both simulated data and real data demonstrate the effectiveness and efficiency of the algorithm.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

Fourier transform infrared spectroscopy can be a valuable tool for both qualitative and quantitative analysis. However, it is the common problem that there are baselines for the collected spectra which consist of sharp features superimposed upon a continuous, slowly varying background. These baselines hamper the interpretation of spectra. In addition, the baseline varies greatly from spectrum to spectrum, even for otherwise similar samples. In quantitative analysis, these inconsistent baselines will reduce the simplicity and robustness of a calibration model that built on these spectra. Thus, it is necessary to remove the baseline drift in spectroscopic analysis.

Baseline correction is to pull the drifted baseline back zero absorbance. It can be implemented by either manual or automation means. Manual methods consist of selecting manually a series of representative points of the baseline from a spectrum. The baseline is then constructed by interpolating these points using linear, polynomial, or spline functions, etc. This approach is subjective, time-consuming and of limited reproducibility [1]. There have been numerous proposals for automatic baseline correction [2,3]. In the derivative method [4], first derivatives can remove constant offsets from the spectra. Linear baselines are eliminated by second derivatives. Frequency analysis based methods try to capture the differences in frequency content by separating the signal into a slow (the baseline) and a fast part (the signal). Fourier transform method yields the frequency components of the spectrum, then a high-pass

filter is used to suppress the spectral baseline. Wavelet transform method decomposes the original spectrum into an appropriate resolution level, and the baseline is constructed according to the smooth wavelet coefficients at this level. Morphological operators [5–7] can be used to extract geometry shape information of a signal by a structuring element. In particular, using morphological opening, peaks whose width is smaller than the structuring element will be eliminated, then the baseline is extracted. In curve fitting method, the drifted baseline is approximated with a mathematical function which fits some automatically selected “reference points”. Baseline is a slowly varying background, and background estimation based on Expectation-Maximization algorithm and a recursive histogram computation are proposed for DNA sequencing [8]. Another background estimation method based on Statistics-sensitive Nonlinear Iterative Peak-clipping algorithm is proposed in Refs. [9,10].

Recently, a baseline correction algorithm with asymmetric least squares smoothing is proposed [11,12]. It estimates a baseline \mathbf{z} based on the Whittaker smoother as follows:

$$\mathbf{z} = \arg \min_{\mathbf{z}} \left\{ \sum_i w_i (\mathbf{y}_i - \mathbf{z}_i)^2 + \mu \sum_i (\Delta^2 \mathbf{z}_i)^2 \right\}$$

where \mathbf{y} is a spectrum, Δ is a difference operator, weights w_i are chosen asymmetrically: $w_i = p$ if $\mathbf{y}_i > \mathbf{z}_i$ and $w_i = 1 - p$ otherwise, and μ is a regularization parameter. When $p = 0.5$, the algorithm is actually the Hodrick–Prescott filtering algorithm [13] which is widely used in trend estimation of macroeconomic time series.

Note that, all the above baseline (or trend) correction algorithms are performed on a single spectrum. In practice, the baseline of a spectrum usually is not known in advance. From a single spectrum,

* Corresponding author. Tel.: +86 10 62520293.

E-mail addresses: jiangtao.peng@ia.ac.cn, pengjt1982@yahoo.com.cn (J. Peng).

it is unclear what the true baseline is. When multiple spectra are available, we can infer the common characteristics in spectra and learn the slow varying baselines accordingly. Assume that multiple spectra are the spectra of a certain sample continuously collected several times, thus the multiple spectra have common characteristics and show small differences in baselines. We can estimate the baselines by penalizing the differences in the baseline corrected spectra. Motivated by the similarity among the baseline corrected spectra, we propose a multiple spectra baseline correction algorithm which will be introduced in the next section. In Section 3, experimental results and discussion are given. In Section 4, we give a brief conclusion.

2. The algorithm

The goal of multiple spectra baseline correction algorithm is to learn baselines that perform well on the corresponding spectra, and then “co-regularize” the choices by penalizing disagreements among the baseline corrected spectra based on asymmetric least squares. Assume that m spectra $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_m$ are column vectors of length n , and $\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_m$ are the corresponding baselines to be estimated. The multiple spectra regularization solves the following optimization problem for $k = 1, \dots, m$:

$$\mathbf{z}_k = \arg \min_{\mathbf{z}_k} \left\{ \sum_{u=1}^m \sum_{v=1, v>u}^m \|(\mathbf{y}_u - \mathbf{z}_u) - (\mathbf{y}_v - \mathbf{z}_v)\|^2 + \lambda \sum_{i=1}^m (\mathbf{y}_i - \mathbf{z}_i)^T Q_i (\mathbf{y}_i - \mathbf{z}_i) + \sum_{i=1}^m \mu_i \|D\mathbf{z}_i\|^2 \right\} \quad (1)$$

where D is the second order differential matrix, each Q_i is a diagonal matrix with diagonal elements $w_i(j) = p$ if $\mathbf{y}_i(j) > \mathbf{z}_i(j)$ and $w_i(j) = 1 - p$ otherwise, and λ, μ_i are regularization parameters.

The first term on the right side of Eq. (1) reflects the pairwise differences between the corrected spectra, while the second term measures the fitting error to the data as an asymmetric way and the last term is smoothness constraint of the baselines. When all regularization matrices Q_i are chosen as the identity matrix (or $p = 0.5$) in Eq. (1), the optimization framework (1) is actually a joint Hodrick–Prescott filtering algorithm which can extract the trends of multiple NIR spectra.

Up to a constant factor m , the optimization problem (1) is equivalent to the following problem (see the proof in Appendix):

$$\mathbf{z}_k = \arg \min_{\mathbf{z}_k} \left\{ \sum_{u=1}^m \left\| (\mathbf{y}_u - \mathbf{z}_u) - \frac{1}{m} \sum_{v=1}^m (\mathbf{y}_v - \mathbf{z}_v) \right\|^2 + \lambda \sum_{i=1}^m (\mathbf{y}_i - \mathbf{z}_i)^T Q_i (\mathbf{y}_i - \mathbf{z}_i) + \sum_{i=1}^m \mu_i \|D\mathbf{z}_i\|^2 \right\} \quad (2)$$

The above framework shows certain relations with multiplication scattering correction (MSC). From the first term on the right side of Eq. (2), it can be seen that each corrected spectrum must be close to the mean corrected spectrum. MSC also corrects each spectrum according to the mean spectrum. While MSC does scatter correction on original spectrum, and our algorithm can simultaneously correct scatter effects and eliminate baselines (or trends).

An underlying assumption in framework (1) or (2) is that the baseline corrected spectra are approximately the same. However, this assumption can be loosed by attaching a relaxation factor to each corrected spectrum and the framework (2) can be extended

to the following form:

$$(\mathbf{z}_k, a_k) = \arg \min_{\mathbf{z}_k, a_k} \left\{ \sum_{u=1}^m \left\| (\mathbf{y}_u - \mathbf{z}_u) - a_u \frac{1}{m} \sum_{v=1}^m (\mathbf{y}_v - \mathbf{z}_v) \right\|^2 + \lambda \sum_{i=1}^m (\mathbf{y}_i - \mathbf{z}_i)^T Q_i (\mathbf{y}_i - \mathbf{z}_i) + \sum_{i=1}^m \mu_i \|D\mathbf{z}_i\|^2 \right\} \quad (3)$$

where a_u is a relaxation factor, which reflects the similarity between the corresponding corrected spectrum and the mean corrected spectrum. Unlike Eq. (2), however, each corrected spectrum needs to be close to the mean corrected spectrum in the sense of a scale factor.

Denote $\gamma_k = a_k(2 - a_k)$ and $\theta = \frac{1}{m} \sum_{v=1}^m (\mathbf{y}_v - \mathbf{z}_v)$, the solution of Eq. (3) is as follows:

$$\begin{cases} \mathbf{z}_k = [(m - \gamma_k)E + \lambda^* Q_k + \mu_k^* D^T D]^{-1} \cdot \\ \quad \left[(m - \gamma_k) \mathbf{y}_k - \gamma_k \sum_{i=1, i \neq k}^m (\mathbf{y}_i - \mathbf{z}_i) + \lambda^* Q_k \mathbf{y}_k \right], \\ a_k = (\theta^T \theta)^{-1} \theta^T (\mathbf{y}_k - \mathbf{z}_k). \end{cases} \quad (4)$$

where E is the $n \times n$ identity matrix, and $\lambda^* = m\lambda$, $\mu_k^* = m\mu_k$.

Given initial factors $a_k^{(0)} = 1$, weight matrices $Q_k^{(0)} = E$, and initial baselines $\mathbf{z}_k^{(0)}$, the first baselines $\mathbf{z}_k^{(1)}$ of the multiple spectra can be estimated according to Eq. (4), and then the weight matrices $Q_k^{(1)}$ are updated as follows:

$$Q_k^{(1)}(i, j) = \begin{cases} p, & \text{if } y_k(i) > z_k^{(1)}(i) \text{ and } i = j; \\ 1 - p, & \text{if } y_k(i) \leq z_k^{(1)}(i) \text{ and } i = j; \\ 0, & \text{otherwise.} \end{cases} \quad (5)$$

The algorithm is insensitive to the initial baselines $\mathbf{z}_k^{(0)}$ which are usually set as constant spectra in the experiment, where the constant value of each initial baseline spectrum is chosen as the minimum of the corresponding spectrum. The weight parameter p controls the position of the baseline and is mainly related to the quality of the spectra. When the noise is low, p can be chosen as a small value. In cases of substantial noise, p should be chosen such that the estimated baseline runs in the middle of the noise band in areas without signal [12]. The regularization parameters μ_k can simply set to be the same value μ which is usually much larger than λ . The parameter pair (λ, μ) balances the bias of the similarity assumption against the error of not estimating the baselines well. The choice of large values for the parameter pair will weaken the similarity constraint, and the algorithm will more or less reduce to asymmetric least squares for the single spectrum separately.

The final baselines can be estimated by updating the baselines and relaxation factors alternately according to Eq. (4). Algorithm 1 describes the proposed multiple spectra baseline correction algorithm.

Algorithm 1. Multiple spectra baseline correction algorithm.

Given λ , μ_k , p , and spectra \mathbf{y}_k , $k = 1, \dots, m$.

Estimate the baselines $\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_k$

1: Initialize: $j = 0$, $Q_k^{(j)} = Q_k^{(0)}$, $\mathbf{z}_k^{(j)} = \mathbf{z}_k^{(0)}$, $a_k^{(j)} = a_k^{(0)} = 1$

2: Compute $\mathbf{z}_k^{(j+1)}$ according to Eq. (4) as follows:

$$\mathbf{z}_k^{(j+1)} = \left[(m - \gamma_k^{(j)})E + \lambda Q_k^{(j)} + \mu_k D^T D \right]^{-1} \cdot \left[(m - \gamma_k^{(j)})\mathbf{y}_k - \gamma_k^{(j)} \sum_{i=1, i \neq k}^m (\mathbf{y}_i - \mathbf{z}_i^{(j)}) + \lambda Q_k^{(j)} \mathbf{y}_k \right]$$

where $\gamma_k^{(j)} = a_k^{(j)}(2 - a_k^{(j)})$.

3: Update the diagonal matrix $Q_k^{(j+1)}$ according to Eq. (5), where the diagonal elements equal to $1 - p$ if $\mathbf{y}_k(i) \leq \mathbf{z}_k^{(j+1)}(i)$ and p otherwise.

4: Calculate the relaxation factor $a_k^{(j+1)}$ as follows:

$$a_k^{(j+1)} = (\theta^T \theta)^{-1} \theta^T (\mathbf{y}_k - \mathbf{z}_k^{(j+1)})$$

where $\theta = \frac{1}{m} \sum_{i=1}^m (\mathbf{y}_i - \mathbf{z}_i^{(j+1)})$.

5: Set $j = j + 1$ and return to step 2 until convergence.

3. Experimental results and discussion**3.1. Materials**

The programs are written in house in Matlab Version R2008a (The MathWorks, Inc.) and run in a personal computer with an

2.20 GHz Intel Core 2 processor, 2 GB RAM, and a Windows XP operating system. The Matlab codes and data used in this research are included as **Supplementary material**.

3.2. Dataset

Dataset 1 consists of 32 FTIR spectra which are the spectra of the same Chinese liquor Guotai sample collected 32 times continuously. The IR spectra in the region $4000\text{--}650\text{ cm}^{-1}$ have been recorded with a Perkin-Elmer Spectrum GX FTIR spectrometer, equipped with the Universal ATR Sampling Accessory (ZnSe cell). The spectral resolution is set at 4 cm^{-1} . Sixteen scans are added for each spectrum.

Dataset 2 consists of 80 NIR spectra of corn measured on spectrometers mp5 and mp6. The spectra are collected in 2 nm intervals within the spectral range $1100\text{--}2498\text{ nm}$. Reference values for moisture, oil, protein and starch are available for all the samples. The data set is available via internet <http://www.eigenvector.com/data/index.htm>.

3.3. Multiple spectra baseline correction algorithm

With a simulation, the performance of multiple spectra baseline correction algorithm in a situation where there are two spectra with known baselines is presented. Assume original spectrum be of 8 Gaussian peaks with different widths, and two synthetic baselines are added to it, yielding two drifted spectra shown in Fig. 1,

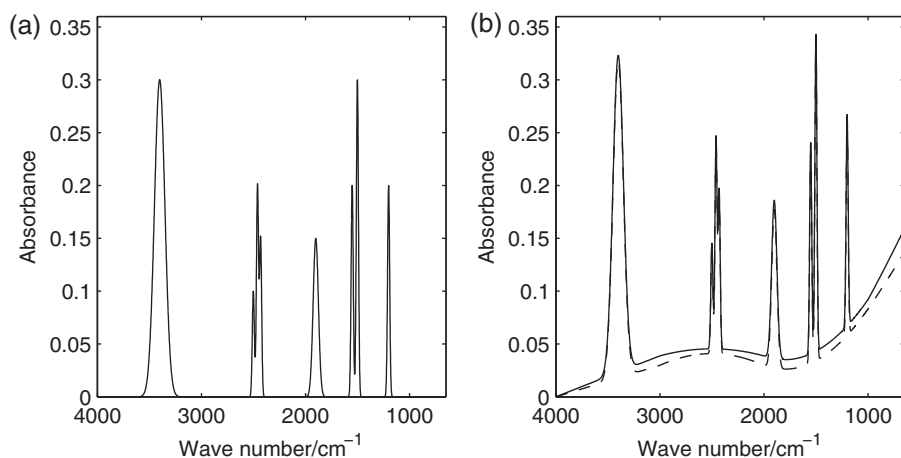


Fig. 1. (a) Original spectrum; (b) two synthetic spectra with different baselines.

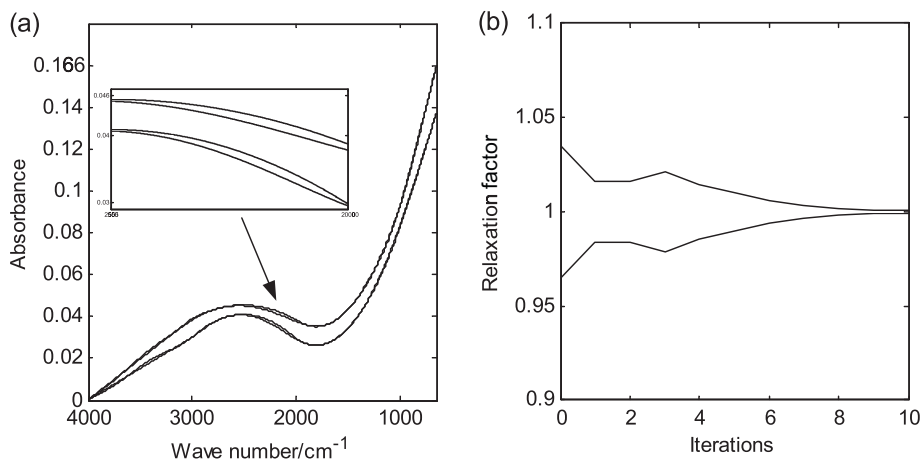


Fig. 2. (a) Two true baselines (solid lines) and estimated baselines (dotted lines); (b) two relaxation factors.

where the simulated baselines are piecewise functions with different combinations of linear and quadratic functions.

Setting the parameters $\lambda = 150$, $\mu_1 = 0.8 \times 10^8$, $\mu_2 = 10^8$, and $p = 0$, two baselines are estimated according to Algorithm 1. Fig. 2(a) gives the estimated baselines and the true baselines, which shows very small differences. Fig. 2(b) shows the relaxation factors corresponding to the two baseline corrected spectra. The two factors tend to 1 as the number of iterations increases. That is, after baseline correction, the two spectra are close to each other.

Next, a comparison between the multiple spectra baseline correction algorithm and asymmetric least squares (AsLS) baseline correction algorithm is presented. Note that, AsLS corrects the spectrum individually.

In order to find the optimal parameters p and μ in AsLS, a grid search method is used. First, p varies in $\{0.1, 0.01, 0.001, 0.0001\}$ and μ ranges from 10^1 to 10^{10} with exponent varying from 1 to 10. For each pair of parameters (p, μ) , a baseline is estimated and the error $RMSE = \sqrt{\sum_i (z(i) - b(i))^2}$ is computed, where z and b are the estimated baseline and true baseline respectively. After coarse tuning in the first step, parameters $(\bar{p}, \bar{\mu})$ corresponding to the minimum RMSE are chosen, then fine tuning is performed around \bar{p} and $\bar{\mu}$. Finally, parameters are chosen as follows: $p = 10^{-4}$, $\mu = 4 \times 10^5$ for spectrum y_1 , and $p = 10^{-4}$, $\mu = 9 \times 10^5$ for spectrum y_2 . The fitting errors are shown in Table 1, where RMSE1 (RMSE2) is the error between the estimated baseline z_1 (z_2) and true baseline b_1 (b_2), and RMSE is the error between two baseline corrected spectra. It

Table 1

The error between estimate baseline and true baseline.

Method	Parameters	RMSE1	RMSE2	RMSE
MSBC	$\lambda = 150, \mu_1 = 0.8 \times 10^8, \mu_2 = 10^8$	0.0314	0.0331	0.0119
AsLS	$p = 10^{-4}, \lambda = 4 \times 10^5$	0.0992	–	–
AsLS	$p = 10^{-4}, \lambda = 9 \times 10^5$	–	0.0538	0.0666

can be seen from the table that multiple spectra baseline correction algorithm outperforms AsLS for single spectrum, and makes the differences between corrected spectra smaller. Due to the joint regularization in multiple spectra baseline correction, the newly learned baseline may provide new information for the baseline to be estimated to improve its performance. Therefore, the performance of multiple spectra regularization is often significantly better than that of single spectrum approach.

Multiple spectra baseline correction algorithm on two real spectra selected from Dataset 1 is shown in Fig. 3, where parameters are selected as follows: $\lambda = 100$, $\mu_1 = 10^{10}$, $\mu_2 = 10^9$, and $p = 10^{-4}$. After baseline correction, two spectra are almost the same. It takes three seconds to estimate baselines of the 32 liquor spectra in Dataset 1, where each spectrum has length 3351. Fig. 4 shows the estimated baselines of 32 liquor spectra and baseline corrected spectra, where $p = 10^{-4}$, $\lambda = 500$, and all the regularization parameters μ_k are just set to be 10^{11} for simplicity. It can be seen from Figs. 3 and 4 that the baselines are nearly removed and the spectra scatterings around 3400 and 700 cm^{-1} are almost removed at the same time.

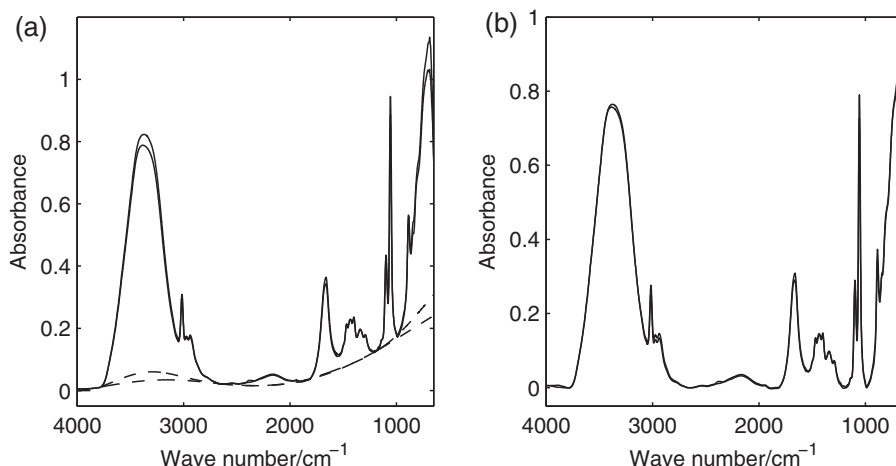


Fig. 3. (a) Two spectra (solid lines) and estimated baselines (dashed lines); (b) two corrected spectra.

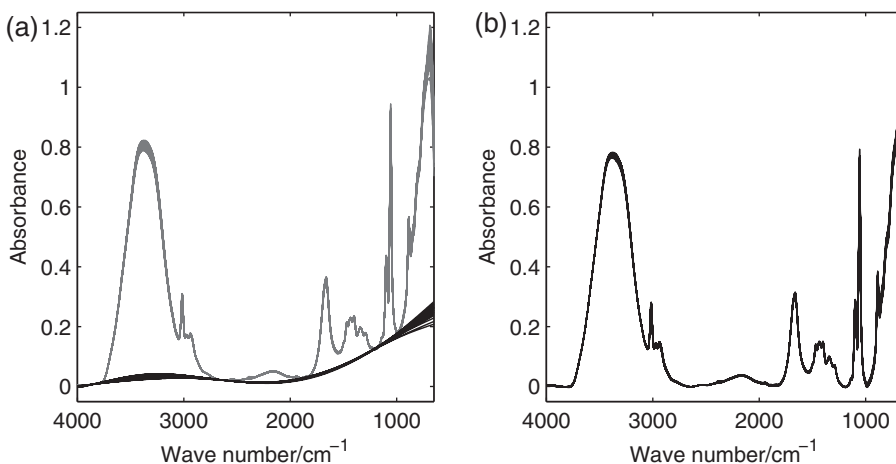


Fig. 4. (a) 32 spectra (gray lines) and estimated baselines (black lines); (b) corrected spectra.

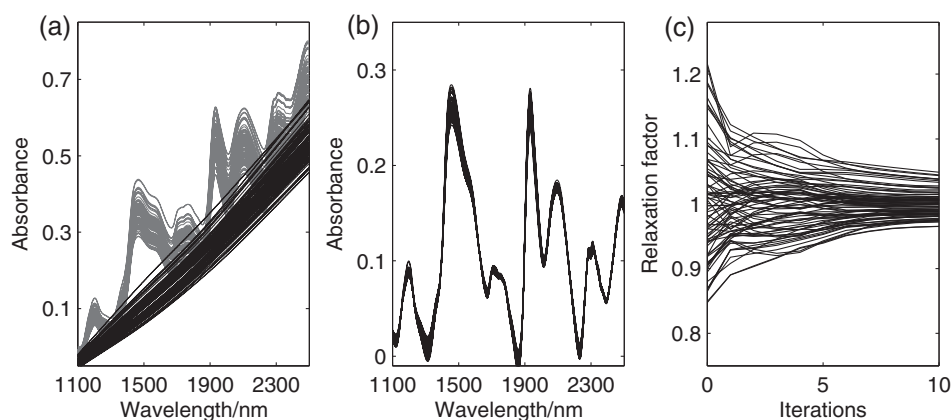


Fig. 5. (a) NIR spectra of corn samples (gray line) and estimated baselines (black line); (b) corrected spectra; (c) relaxation factors.

The multiple spectra baseline correction algorithm on NIR corn spectra scanned on mp6 is shown in Fig. 5, where $p=0$, $\lambda=10^3$ and $\mu_k=5 \times 10^9$ ($k=1, \dots, 80$).

It can be seen from Fig. 5(b) that the scattering effects are removed to a certain extent. Furthermore, the relaxation factors corresponding to the baseline corrected spectra in Fig. 5(c) also show the closeness of the corrected spectra, where the iteration 0 corresponds to the differences between the original spectra and mean spectra.

The scattering effects can dramatically and adversely impact the quantitative results and capabilities of the NIR calibrations, and they must be minimized. Many methods of standardization or correcting for these effects are used in NIR spectroscopy [16]. Multiplicative scatter correction [15] and standard normal variate transformations [14] are two widely used methods. In addition, detrending can be viewed as a baseline correction method [16], and it accounts for the variation in baseline shift and curvilinearity, with the use of a second order polynomial regression [14]. By minimizing scatter and baseline effects, these preprocessing methods are believed to benefit the calibration process. Next, calibration models on corn samples are built with the moisture, oil, protein and starch value as responses, respectively. The corn samples are first sorted by its corresponding response, and then the third ones of every fifth

sample are assigned for the prediction set, resulting in 16 prediction samples and 64 calibration samples. Calibration models are performed by partial least squares regression and the number of latent variables is determined by leave-one-out cross-validation. The major statistical determinations of calibration robustness and “goodness” are R^2 (correlation coefficient) and root mean square error of prediction (RMSEP).

The comparisons of the prediction abilities of the several preprocessing methods and the proposed multiple spectra baseline correction algorithm on mp5 and mp6 are performed in Tables 2 and 3, respectively.

For moisture, protein and starch responses, the multiple spectra baseline correction (MSBC) algorithm achieves better calibration results than all the other methods. For oil response, only detrending method improves the results, and the other three correction methods seem no use in this case. Nevertheless, the maximum difference in RMSEP (%) between the MSBC and detrending method is only 0.008. SNV almost has the same performance with MSC in all cases. In addition, MSBC yields much better results than SNV and MSC. On the whole, MSBC algorithm produces better performance. The lower RMSEP shows better predictive ability, and the higher correlation coefficients R^2 demonstrates the robust and universal of MSBC calibration.

Table 2
Calibration regression statistics for SNV, MSC, DT and MSBC transforms on mp5.

	Moisture		Oil		Protein		Starch	
	R^2	RMSEP	R^2	RMSEP	R^2	RMSEP	R^2	RMSEP
No	0.956	0.123(9)	0.882	0.087(7)	0.969	0.140(10)	0.890	0.406(9)
SNV	0.808	0.231(6)	0.815	0.106(5)	0.967	0.150(8)	0.919	0.342(8)
MSC	0.808	0.231(6)	0.815	0.106(5)	0.968	0.150(8)	0.915	0.354(8)
DT	0.955	0.131(8)	0.920	0.087(6)	0.971	0.121(13)	0.891	0.405(9)
MSBC	0.952	0.121(7)	0.896	0.091(7)	0.979	0.116(12)	0.907	0.338(8)

The values in parentheses refer to the number of latent variables. No = original spectra, DT = Detrending based on second order polynomial fitting, MSBC = multiple spectra baseline correction.

Table 3
Calibration regression statistics for SNV, MSC, DT and MSBC transforms on mp6.

	Moisture		Oil		Protein		Starch	
	R^2	RMSEP	R^2	RMSEP	R^2	RMSEP	R^2	RMSEP
No	0.918	0.167(8)	0.849	0.098(7)	0.979	0.134(9)	0.903	0.380(9)
SNV	0.787	0.246(6)	0.803	0.108(5)	0.966	0.162(8)	0.906	0.354(7)
MSC	0.786	0.246(6)	0.803	0.108(5)	0.966	0.162(8)	0.905	0.356(7)
DT	0.945	0.129(8)	0.859	0.091(7)	0.974	0.138(9)	0.866	0.414(7)
MSBC	0.954	0.115(8)	0.845	0.099(7)	0.981	0.130(11)	0.910	0.331(7)

The values in parentheses refer to the number of latent variables. No = original spectra, DT = Detrending based on second order polynomial fitting, MSBC = multiple spectra baseline correction.

4. Conclusions

The interpretation of spectroscopic data is largely hampered by the baseline or trend problem. In this paper, a novel multiple spectra baseline correction algorithm is presented. By exploiting the similarity between spectra, the algorithm removes the baselines effectively. Moreover, similar to SNV and MSC, the proposed algorithm also can eliminate the scatter effects on spectra. Simulated and real spectra are used for testing the efficiency of multiple spectra baseline correction algorithm. In particular, calibration regression statistics are preformed on NIR corn samples. The results on both simulated and real data show the superiority of the proposed algorithm.

Acknowledgements

The authors would like to thank the anonymous reviewers for their insightful comments and constructive suggestions. The research is supported in part by the National Natural Science Foundation of China (60972126), the Joint Funds of the National Natural Science Foundation of China (U0935002/L05), the Beijing Municipal Natural Science Foundation (4102060) and the State Key Program of National Natural Science of China (61032007).

Appendix A. Appendix A

Proof for the equivalence between Eqs. (1) and (2).

Denote the objective function in regularization framework (1) as

$$S_1 = \sum_{u,v=1, v>u}^m \|(\mathbf{y}_u - \mathbf{z}_u) - (\mathbf{y}_v - \mathbf{z}_v)\|^2 + \lambda \sum_{i=1}^m (\mathbf{y}_i - \mathbf{z}_i)^T Q_i (\mathbf{y}_i - \mathbf{z}_i) + \sum_{i=1}^m \mu_i \|D\mathbf{z}_i\|^2$$

Taking the partial derivative of S_1 with respect to \mathbf{z}_k ,

$$\frac{\partial S_1}{\partial \mathbf{z}_k} = -2 \sum_{i=1, i \neq k}^m [(\mathbf{y}_k - \mathbf{z}_k) - (\mathbf{y}_i - \mathbf{z}_i)] - 2\lambda Q_k (\mathbf{y}_k - \mathbf{z}_k) + 2\mu_k D^T D \mathbf{z}_k$$

Let $\frac{\partial S_1}{\partial \mathbf{z}_k} = 0$, it gets

$$\mathbf{z}_k = [(m-1)E + \lambda Q_k + \mu_k D^T D]^{-1} [(m-1)\mathbf{y}_k - \sum_{i=1, i \neq k}^m (\mathbf{y}_i - \mathbf{z}_i) + \lambda Q_k \mathbf{y}_k] \quad (6)$$

Denote $\theta = \frac{1}{m} \sum_{i=1}^m (\mathbf{y}_i - \mathbf{z}_i)$, the objective function in Eq. (2) can be written as:

$$S_2 = \sum_{i=1}^m \|(\mathbf{y}_i - \mathbf{z}_i) - \theta\|^2 + \lambda \sum_{i=1}^m (\mathbf{y}_i - \mathbf{z}_i)^T Q_i (\mathbf{y}_i - \mathbf{z}_i) + \sum_{i=1}^m \mu_i \|D\mathbf{z}_i\|^2$$

Taking the partial derivative of S_2 with respect to \mathbf{z}_k ,

$$\begin{aligned} \frac{\partial S_2}{\partial \mathbf{z}_k} &= (\mathbf{y}_k - \mathbf{z}_k - \theta) \left(-1 + \frac{1}{m} \right) + \frac{1}{m} \sum_{i=1, i \neq k}^m (\mathbf{y}_i - \mathbf{z}_i - \theta) \\ &\quad - \lambda Q_k (\mathbf{y}_k - \mathbf{z}_k) + \mu_k D^T D \mathbf{z}_k \\ &= (\mathbf{y}_k - \mathbf{z}_k - \theta) \left(-1 + \frac{1}{m} \right) + \frac{1}{m} \left[\sum_{i=1}^m (\mathbf{y}_i - \mathbf{z}_i) - (\mathbf{y}_k - \mathbf{z}_k) - (m-1)\theta \right] \\ &\quad - \lambda Q_k (\mathbf{y}_k - \mathbf{z}_k) + \mu_k D^T D \mathbf{z}_k \\ &= \mathbf{z}_k - \mathbf{y}_k + \theta - \lambda Q_k (\mathbf{y}_k - \mathbf{z}_k) + \mu_k D^T D \mathbf{z}_k \end{aligned}$$

Let $\frac{\partial S_2}{\partial \mathbf{z}_k} = 0$, it gets

$$\begin{aligned} \mathbf{z}_k &= \left[\left(\frac{1-1}{m} \right) E + \lambda Q_k + \mu_k D^T D \right]^{-1} \cdot \\ &\quad \left[\left(\frac{1-1}{m} \right) \mathbf{y}_k - \frac{1}{m} \sum_{i=1, i \neq k}^m (\mathbf{y}_i - \mathbf{z}_i) + \lambda Q_k \mathbf{y}_k \right] \\ &= [(m-1)E + \lambda^* Q_k + \mu_k^* D^T D]^{-1} \cdot \\ &\quad \left[(m-1)\mathbf{y}_k - \sum_{i=1, i \neq k}^m (\mathbf{y}_i - \mathbf{z}_i) + \lambda^* Q_k \mathbf{y}_k \right], \quad k = 1, \dots, m \end{aligned} \quad (7)$$

where $\lambda^* = m\lambda$, $\mu_k^* = m\mu_k$.

Up to a constant factor m in the regularization parameters λ and μ_k , the Eqs. (6) and (7) are equivalent. Thus, it proves the equivalence between the optimization framework (1) and (2).

Appendix B. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.aca.2010.08.033

References

- [1] A. Jirasek, G. Schulze, M.M.L. Yu, M.W. Blades, R.F.B. Turner, Appl. Spectrosc. 58 (2004) 1488–1499.
- [2] G. Schulze, A. Jirasek, M.M.L. Yu, A. Lim, R.F.B. Turner, M.W. Blades, Appl. Spectrosc. 59 (2005) 545–574.
- [3] D.A. Barkauskas, D.M. Rocke, Anal. Chim. Acta 657 (2010) 191–197.
- [4] M.N. Leger, A.G. Ryder, Appl. Spectrosc. 60 (2006) 182–193.
- [5] P. Maragos, R.W. Schafer, IEEE Trans. Acoust. Speech Signal Process 35 (1987) 1170–1184.
- [6] C.H. Chu, E.J. Delp, IEEE Trans. Biomed. Eng 36 (1989) 262–272.
- [7] J. Angulo, J. Serra, Bioinformatics 19 (2003) 553–562.
- [8] L. Andrade, E.S. Manolakos, J. VLSI Signal Process. 35 (2003) 229–243.
- [9] C.G. Ryan, E. Clayton, W.L. Griffin, S.H. Sie, D.R. Cousens, Nucl. Instr. and Meth. B 34 (1988) 396–402.
- [10] M. Morháč, Nucl. Instr. and Meth. A 600 (2009) 478–487.
- [11] P.H.C. Eilers, Anal. Chem. 75 (2003) 3631–3636.
- [12] P.H.C. Eilers, Anal. Chem. 76 (2004) 404–411.
- [13] R. Hodrick, E. Prescott, J. Money Credit Bank. 29 (1997) 1–16.
- [14] R.J. Barnes, M.S. Dhanoa, S.J. Lister, Appl. Spectrosc. 43 (1989) 772–777.
- [15] P. Geladi, D. MacDougall, H. Martens, Appl. Spectrosc. 3 (1985) 491–500.
- [16] Å. Rinnan, F.V.D. Berg, S.B. Engelsen, Trends Anal. Chem. 28 (2009) 1201–1222.