Chemometrics and Intelligent Laboratory Systems xxx (2010) xxx–xxx

Contents lists available at ScienceDirect

Chemometrics and Intelligent Laboratory Systems

journal homepage: www.elsevier.com/locate/chemolab

Short Communication

Uncover the path from PCR to PLS via elastic component regression

Hong-Dong Li [a], Yi-Zeng Liang [a,*], Qing-Song Xu [b]

[a] Research Center of Modernization of Traditional Chinese Medicines, College of Chemistry and Chemical Engineering, Central South University, Changsha 410083, PR China
[b] School of Mathematic Sciences, Central South University, Changsha 410083, PR China

ARTICLE INFO

Article history:
Received 21 March 2010
Received in revised form 26 July 2010
Accepted 4 August 2010
Available online xxxx

Keywords:
Elastic component regression
Model path
Principal component regression
Partial least squares

ABSTRACT

This contribution introduces Elastic Component Regression (ECR) as an explorative data analysis method that utilizes a tuning parameter $\alpha \in [0,1]$ to supervise the **X**-matrix decomposition. It is demonstrated theoretically that the elastic component resulting from ECR coincides with principal components of PCA when $\alpha = 0$ and also coincides with PLS components when $\alpha = 1$. In this context, PCR and PLS occupy the two ends of ECR and $\alpha \in (0,1)$ will lead to an infinite number of transitional models which collectively uncover the model path from PCR to PLS. Therefore, the framework of ECR shows a natural progression from PCR to PLS and may help add some insight into their relationships in theory. The performance of ECR is investigated on a series of simulated datasets together with a real world near infrared dataset. (The source codes implementing ECR in MATLAB are freely available at http://code.google.com/p/ecr/.)

1. Introduction

With the development of modern analytical instrument, the routinely produced data from biological and environmental samples are tending to be characterized by a great number of high collinear variables measured on a few samples. Often the number of such variables, i.e. p, is much larger than that of the observations N. This usually refers to the "large p, small N" problem. As is known to us, ordinary least squares (OLS) cannot work in this situation. To tackle this problem, a number of methods have been developed, among which the latent variable-based multivariate regression technique has been regarded as a natural way for analyzing such data. Loosely speaking, this procedure works by computing a few latent variable, followed by regressing the response variable on the derived latent variables by means of OLS technique.

The most popular regression methods in the community of chemometrics are principal component regression (PCR) [1,2] and partial least squares (PLS) [3–6]. These methods are heavily promoted in chemometrics and have found very wide applications in a variety of fields, such as near infrared (NIR) [7,8], microarray data analysis [9] and metabolic profiles [10]. Other latent variable-based methods known to statisticians but rarely mentioned by chemometricians are canonical correlation regression (CCR), reduced rank regression (RRR) and some other variations [11].

It's very interesting to note that there exists some potential relationship between some of the above-mentioned methods. M. Stone and R. Brooks constructed a general objective function and introduced continuum regression (CR) which includes OLS and PCR as the two opposite ends of a continuum spectrum with PLS lying in between [12,13]. The established theories and algorithms for implementing CR are to some extent difficult for us chemists/chemometricians to understand. As the authors stated, the analysis of PCR model, in the context of CR, is rather technical and is left to be demonstrated only numerically. Following Stone and Brooks' work, Sijmen de Jong and Henk A.L. Kiers presented principal covariate regression (PCovR) by minimizing the proposed weighted least squares loss function [14]. In brief, the procedure of PCovR moves from ordinary least squares to principal component regression, which can be seen as a relatively simple alternative for continuum regression. John H. Kalivas et al. proposed Cyclic Subspace Regression (CSR) [15] as a new approach to the complex multivariate calibration problem. CSR was shown to produce solutions of PCR, PLS, LS and other related intermediate regressions with one algorithm [16–18]. Recently, Sijmen de Jong et al. developed the canonical partial least squares which offers a computationally efficient setting to carry out the continuum regression approach [19]. On the whole, CR found a variety of applications [20] in model calibration and has recently been discussed by Sundberg [21] and Geladi [22].

The goal of this paper is to develop a new theory for continuum regression, called Elastic Component Regression (ECR). What distinguishes ECR from the previous work [12,14,15,19,21,22] lies in that it specifically brings PCR and PLS under the same objective function thus providing a continuous progression from PCR to PLS. In this sense, the widely used algorithms, e.g. nonlinear iterative partial least squares (NIPALS), can be borrowed after a slight modification to efficiently compute the ECR model, which makes ECR easy to understand and implement. In brief, a supervising factor $\alpha \in [0,1]$ is defined for ECR. In the context of ECR, PCR and PLS

* Corresponding author.
E-mail address: yizeng_liang@263.net (Y.-Z. Liang).

0169-7439/$ – see front matter © 2010 Published by Elsevier B.V.
doi:10.1016/j.chemolab.2010.08.003

are special cases corresponding to $\alpha=0$ and 1, respectively. The ECR models when $\alpha\in(0,1)$ are therefore could be some kind of transitional models between PCR and PLS. It will be shown that the transitional models with increased α from 0 to 1 can collectively reveal how the PCR model evolves into the PLS model, leading to the intuitive model path from PCR to PLS. Finally, a series of simulated datasets and a NIR dataset are employed to explore the performance of ECR.

2. Theory and the algorithm for ECR

Assume first that the sample matrix $\mathbf{X}$ and the response vector $\mathbf{y}$ are both mean centered. From the optimization perspective, the first principal component in PCR is computed by maximizing the variance of the latent variable $\mathbf{Xw}_{pcr}$ which is a combination of the original variables in $\mathbf{X}$ with combination coefficients $\mathbf{w}_{pcr}$ without considering the effect of $\mathbf{y}$. It can be formulated into the following optimization problem

$$\begin{cases} \max\ \mathbf{w}_{pcr}^t\mathbf{X}^t\mathbf{X}\mathbf{w}_{pcr} \\ s.t.\ \mathbf{w}_{pcr}^t\mathbf{w}_{pcr}^t = 1. \end{cases} \tag{1}$$

By employing Lagrange multiplier method, it can be shown that the objective function is maximized when $\mathbf{w}_{pcr}$ is the eigenvector of the covariance matrix $\mathbf{X}^t\mathbf{X}$ corresponding to the largest eigenvalue. The first principal component can hence be calculated as

$$\mathbf{t}_{pcr} = \mathbf{X}\mathbf{w}_{pcr}. \tag{2}$$

Then the contribution of the first component is stripped, resulting in the residual sample matrix

$$\mathbf{X} = \mathbf{X} - \mathbf{t}_{pcr}\mathbf{w}_{pcr}^t. \tag{3}$$

Then the second, the third and etc. components can be sequentially computed by repeatedly solving the optimization problem in Eq. (1) based on the residual $\mathbf{X}$-matrix. For more detailed information on PCR, the readers are referred to the References [23,24].

While for SIMPLS [25], the objective function is to maximize the squared covariance between latent variables of $\mathbf{Xw}_{pls}$ and the original $\mathbf{y}$ vector. Mathematically, it can also be formulated as the optimization problem

$$\begin{cases} \max\ \mathbf{w}_{pls}^t\mathbf{X}^t\mathbf{y}^t\mathbf{y}\mathbf{X}\mathbf{w}_{pls} \\ s.t.\ \mathbf{w}_{pls}^t\mathbf{w}_{pls}^t = 1. \end{cases} \tag{4}$$

In the similar manner for computing principal components, the PLS component can also be computed sequentially. Due to the fact that the theory and algorithm of PLS are familiar to the chemometricians and there is a large amount of relevant literature [5,13,26], details of PLS are not presented here.

Note that the optimization problems in Eqs. (1) and (4) are of the same formulation and differ only in the criterion maximized. In fact, the two criteria of PCR and PLS can be combined into a general one which can be expressed as

$$\mathbf{H} = (1-\alpha)\mathbf{X}^t\mathbf{X} + \alpha\mathbf{X}^t\mathbf{y}^t\mathbf{y}\mathbf{X} \tag{5}$$

where $\mathbf{H}$ is easily seen as a symmetric matrix and $\alpha\in[0,1]$ is called supervising factor, which controls how much information of $\mathbf{y}$ is occupied in $\mathbf{H}$. Obviously, $\alpha=0$ indicates that $\mathbf{H}=\mathbf{X}^t\mathbf{X}$ coincides with the criterion of PCA, while $\alpha=1$ will result in that $\mathbf{H}=\mathbf{X}^t\mathbf{y}^t\mathbf{y}\mathbf{X}$ coincides with the criterion of PLS. The situations for $\alpha\in(0,1)$ may be considered as the compromise criteria. In this sense, the optimization problems in Eqs. (1) and (4) can be combined into a general one, which serves as the basis of elastic component regression:

$$\begin{cases} \max\ \mathbf{w}^t\mathbf{H}\mathbf{w} \\ s.t.\ \mathbf{w}^t\mathbf{w} = 1. \end{cases} \tag{6}$$

Also by utilizing Lagrange method or the Rayleigh–Ritz lemma, the objective function in Eq. (6) is maximized when $\mathbf{w}$ is the eigenvector of $\mathbf{H}$ corresponding to the largest eigenvalue. Then the first elastic component can be computed as

$$\mathbf{t} = \mathbf{H}\mathbf{w}. \tag{7}$$

In the similar manner of PLS decomposition, the first elastic loading of $\mathbf{X}$ can be computed using OLS as

$$\mathbf{p} = \mathbf{X}^t\mathbf{t}/\mathbf{t}^t\mathbf{t}. \tag{8}$$

The first loading of $\mathbf{y}$ can also be computed using OLS as

$$\mathbf{r} = \mathbf{y}^t\mathbf{t}/\mathbf{t}^t\mathbf{t}. \tag{9}$$

Then, the data $\mathbf{X}$ and $\mathbf{y}$ are deflated by removing the contribution of the first elastic component, resulting in the residual $\mathbf{X}$ and $\mathbf{y}$.

$$\mathbf{X} = \mathbf{X} - \mathbf{t}\mathbf{p}^t \tag{10}$$

$$\mathbf{y} = \mathbf{y} - \mathbf{t}\mathbf{r}^t. \tag{11}$$

Iteratively, the second, the third and etc. elastic components can be computed by repeatedly solving the optimization problem in Eq. (6) based on the residual $\mathbf{X}$ and $\mathbf{y}$.

At last, the algorithm for implementing ECR is given here. Assume that both $\mathbf{X}$ and $\mathbf{y}$ are mean centered and a supervising factor $\alpha\in[0,1]$ is given. The maximal number of elastic component is set to A.

1. Compute $\mathbf{H}=(1-\alpha)\mathbf{X}^t\mathbf{X}+\alpha\mathbf{X}^t\mathbf{y}^t\mathbf{y}\mathbf{X}$
2. Compute $\mathbf{w}$ as the eigenvector of $\mathbf{H}$ corresponding to the largest eigenvalue
3. Compute $\mathbf{X}$-scores $\mathbf{t}=\mathbf{H}\mathbf{w}$
4. Compute $\mathbf{X}$-loadings $\mathbf{p}=\mathbf{X}^t\mathbf{t}/\mathbf{t}^t\mathbf{t}$
5. Compute $\mathbf{y}$-loadings $\mathbf{r}=\mathbf{y}^t\mathbf{t}/\mathbf{t}^t\mathbf{t}$
6. Deflate $\mathbf{X}=\mathbf{X}-\mathbf{t}\mathbf{p}^t$ and $\mathbf{y}=\mathbf{y}-\mathbf{t}\mathbf{r}^t$.

Repeat this procedure until A elastic components are finished. Then the regression vector of ECR can be computed as $\mathbf{b}=\mathbf{W}(\mathbf{P}^t\mathbf{W})^{-1}\mathbf{R}^t$, where $\mathbf{W}$, $\mathbf{P}$ and $\mathbf{R}$ stand for the weight, $\mathbf{X}$-loading and $\mathbf{y}$-loading matrices, respectively.

3. Data and software

3.1. Simulated data

Five groups of datasets are simulated according to the method for producing **SIMUIN** data in Reference [27]. First, **SIM** data of size 25×100 are produced in three steps: (1) generating a matrix $\mathbf{X}$ of random numbers from 0 to 1 with dimension 25×100; (2) performing PCA on the mean-centered matrix $\mathbf{X}$, yielding scores and loadings and (3) reconstructing M using the first 5 principal components and the response vector $\mathbf{y}$ are defined as product of the first 5 components times a weight vector $[5, 4, 3, 2, 1]^t$. After generation of **SIM** data, a certain number, i.e. Q, of uninformative/noise variables are appended to the **SIM** data, leading to **SIMUI** data of size $25\times(100+Q)$. Finally random noises are added to the **SIMUI** data to produce the data called **SIMUIN**. The maximal value of noise is limited to 0.005 as the literature does. In this study, we considered 5 cases: $Q=20$, 50, 100, 200 and 500, respectively. For each case, 1000

Please cite this article as: H.-D. Li, et al., Uncover the path from PCR to PLS via elastic component regression, Chemometr. Intell. Lab. Syst. (2010), doi:10.1016/j.chemolab.2010.08.003

datasets are randomly produced, which results in 5×1000 datasets in all.

3.2. Corn NIR data

This benchmark dataset consists of NIR spectra of 80 corn samples, measured by mp5 NIR spectrometer. Each spectrum contains 700 data points measured in the wavelength range 2498–1100 nm at 2 nm intervals. The moisture value is selected as dependent variable **y**. The original spectra are shown in Fig. 1. This dataset can be freely available at the website: http://software.eigenvector.com/Data/index.html.

3.3. Software

All the calculations are performed using our in-house programs coded in MATLAB 7.6.0. The software for implementing ECR is freely available at: http://code.google.com/p/ecr/downloads/list.

4. Results and discussions

4.1. Simulated data

To build an ECR model, α and the number of latent variables (nLV) should be optimized first. In this work, two-way cross validation (CV) is performed to simultaneously determine the optimal value of α and nLV. Taking the **SIMUIN** data with $Q=20$ as an example, the RMSECV surface resulting from two-way cross validation is shown in Fig. 2 where the optimal α and nLV can be chosen from the lowest point of the RMSECV surface.

For each group of datasets, two-way CV is conducted on the 1000 randomly produced mean-centered datasets and the optimal α is recorded, resulting in 1000 optimal α. Then we count the number of the three cases: (1) $\alpha=0$, (2) $\alpha\in(0,1)$ and $\alpha=1$. Fig. 3 shows the frequency of each case for all the 5 groups of datasets. Note that the frequency of $\alpha=0$ is 0 for all the datasets and thus not shown in Fig. 3. It can be seen that the frequency of $\alpha\in(0,1)$ increases nearly in a linear fashion while increasing the number of noise variables. This observation may indicate that the ECR could to some extent resist the negative influence of noise variables.

Furthermore, in order to show how the PCR model evolves into the PLS model, we take four datasets with the number of noise variables equal to 20, 50, 100 and 500, as examples to illustrate the model path. For each of the four datasets, the number of latent variables is fixed to 5 and α uniformly increases from 0 to 1 with stepsize 0.1. Therefore, for each dataset we can compute altogether 11 ECR models, among which the first and the last coincides PCR and PLS models and the other 9 models in between are the transitional models. The 11 ECR models can hence collectively uncover the model path from PCR to PLS. In this work, PCA is performed on the 11 regression coefficient vectors and the first 2 principal components are extracted to display the model path in a two dimensional plane.

The 4 model paths for the 4 datasets are shown in Plots A, B, C and D of Fig. 4, respectively. Interestingly, the 9 transitional models become closer to the PLS model as the number of noise variable increases. An extreme case is shown in Plot D where all the transitional models are located very closely to PLS model. In our opinion, the phenomenon is worth further investigating. Also, the first two components of the four path plots capture 99.80, 99.93, 99.77 and 99.97% of the total variance of the corresponding regression coefficient matrix, respectively. This result may suggest that the models are highly correlated for a given data.

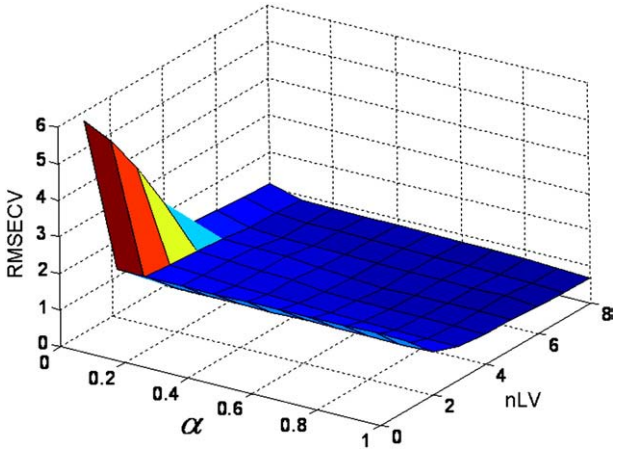

Fig. 2. Illustration of two-way cross validation for determining the optimal value of both α and nLV.

4.2. Corn NIR data

To comprehensively compare the predictive performance of PCR, PLS and ECR, 80% of the 80 corn samples are randomly selected into the calibration set, leaving the remaining 20% as test set. And we

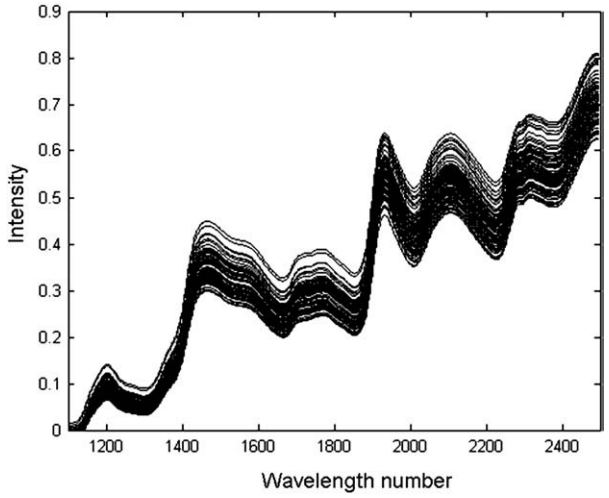

Fig. 1. The original NIR spectra of 80 corn samples.

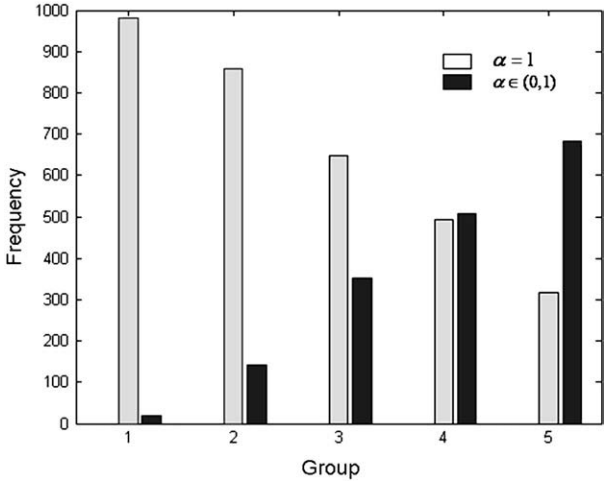

Fig. 3. The frequency of the two-way cross validation-based optimal α value for the two cases (1) $\alpha\in(0,1)$ and (2) $\alpha=1$.

Please cite this article as: H.-D. Li, et al., Uncover the path from PCR to PLS via elastic component regression, Chemometr. Intell. Lab. Syst. (2010), doi:10.1016/j.chemolab.2010.08.003

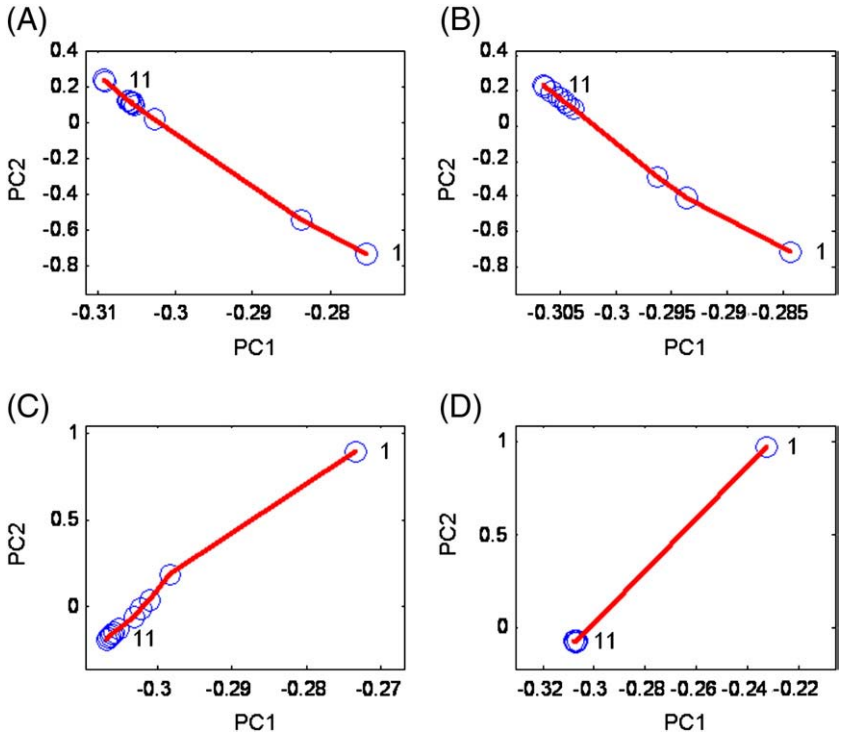

Fig. 4. Illustration of the model path from PCR to PLS when the number of noise variables is set to 20 (A), 50 (B), 100 (C) and 500 (D), respectively. The PCR and PLS models are marked by 1 and 11, respectively.

repeat this procedure 1000 times thus obtaining 1000 calibration sets and 1000 test sets. The rationale of using a large number of (1000) test sets is to avoid possibly unreliable comparison which is addressed in our previous work on model population analysis (MPA) [28]. For each pair of calibration/test set, the minimal root mean squared errors of CV (RMSECV) resulting from the aforementioned two-way cross validation is taken to determine the optimal values of α and nLV. The maximal nLV is set to 12 and the supervising factor α also uniformly increases from 0 to 1 with stepsize 0.1. This step leads to 1000 optimal α values, of which the frequency plot is shown in Fig. 5A. The count numbers for the three cases: (1) $\alpha=0$, (2) $\alpha\in(0,1)$ and $\alpha=1$ are 55, 329 and 616, respectively. This result shows that most of the optimal models lie between PCR and PLS.

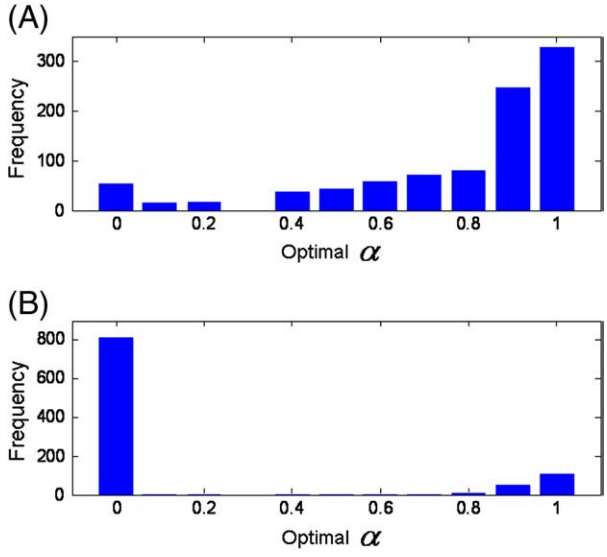

Fig. 5. The distribution of the optimal α values resulting from ECR before and after variable selection on the NIR data.

Then, we build on each calibration set a full-spectrum PCR, PLS and ECR model, respectively, followed by making predictions on the test set. Root mean squared errors of prediction (RMSEP) on the test set is employed to assess the predictive performance of PCR, PLS and ECR. The distribution of the RMSEP values for PCR, PLS and ECR is shown in the left panel of Fig. 6. The mean and standard deviations of the three distributions in the upper, middle and lower plots are 0.1455 ± 0.0253, 0.1365 ± 0.0220 and 0.1354 ± 0.0217, respectively. From this result, it can be found that ECR outperforms PCR and PLS in terms of lower prediction error and slightly smaller standard deviations, indicating that ECR may be a good alternative for multivariate calibration.

As is known, variable selection has great influence on model's predictive performance. The influence of variables may be larger than that of calibration methods. Therefore, the performances of PCR, PLS and ECR are also investigated on the data only including a subset of 19 variables/wavelengths selected by using competitive adaptive reweighted sampling (CARS) [8,29]. In the same way as aforementioned, 1000 calibration sets as well as 1000 test sets are also randomly generated. The distributions of the 1000 optimal α values resulting from two-way cross validation are shown in Fig. 5B. Compared to Fig. 5A, the distribution is dramatically changed. The optimal α values are dominated by PCR followed by PLS. The transitional models only occupy a minority of them. The distributions of RMSEP resulting from PCR, PLS and ECR are also presented in the right panel of Fig. 6. The mean and standard deviations of the three distributions in the upper, middle and lower plots are 0.1031 ± 0.0173, 0.1049 ± 0.0180 and 0.1038 ± 0.0177, respectively. It's very interesting to note that in this case study PCR performs best by possessing the lowest mean prediction error as well as smallest standard deviations. This study shows that the performances of PCR, PLS and ECR are data-dependent. One cannot say which one is the best. Besides, the fact that the three distributions of RMSEP overlap heavily with each other implies that unconvincing (even erroneous)

Please cite this article as: H.-D. Li, et al., Uncover the path from PCR to PLS via elastic component regression, Chemometr. Intell. Lab. Syst. (2010), doi:10.1016/j.chemolab.2010.08.003

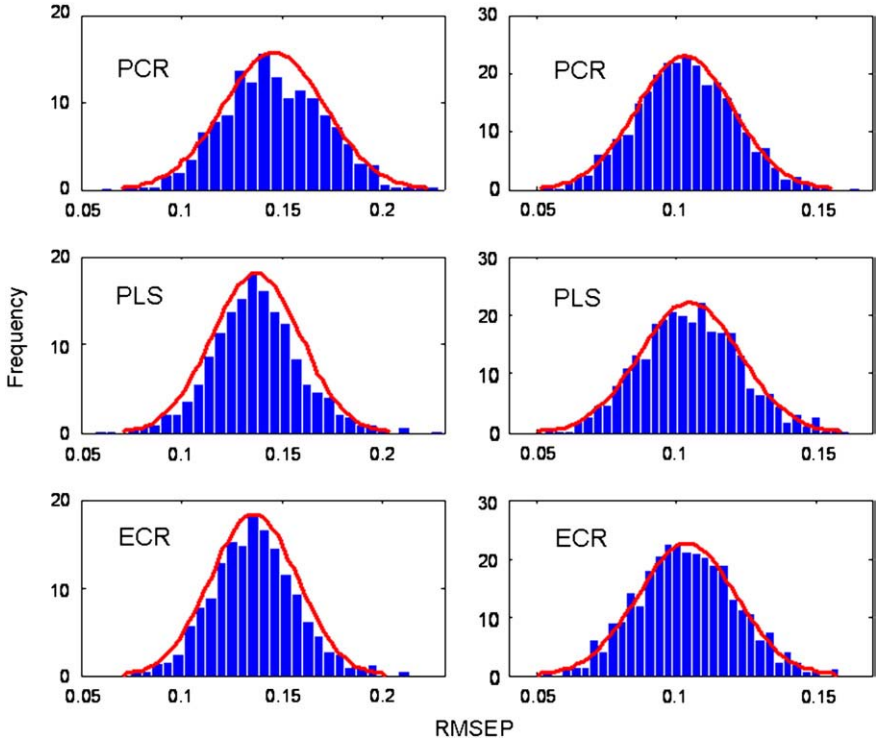

Fig. 6. The distribution of RMSEP values for PCR, PLS and ECR. Left: before variable selection. Right panel: after variable selection.

conclusions may be drawn if different methods are compared based on a single split of the data into a calibration and test set [28].

To show the model path from PCR to PLS, we take all the 80 samples to build 11 ECR models by increasing α from 0 to1 uniformly with stepsize 0.1. The number of latent variables is chosen to be 10 by 5-fold cross validation. This procedure results in 11 regression coefficient vectors, which are collected into a regression coefficient matrix. Then PCA is performed on the matrix and the first two principal components are utilized to display the model path from PCR to PLS, which is shown in Plot A of Fig. 7. Obviously, it can be found that ECR model gradually evolves from PCR into PLS model when α increases. In addition, the minimal RMSECV as a function of α is also shown in Plot B in Fig. 7. It is noteworthy that the RMSECV does not change much after α exceeds 0.6, which may be an indication that $\alpha = 0.6$ is an optimal choice in this case.

4.3. A brief comparison with continuum regression

To intuitively see whether the model path from PCR to PLS uncovered by ECR is the same as that by CR, the PCA-based model paths from both methods are computed by taking the NIR data as input. The number of latent variables is set to 10. For ECR, we compute 11 models with α uniformly increases from 0 to 1. For CR, the continuum power regression (CPR) algorithm is employed and the α values are chosen such that it covers the whole path from PCR to OLS with PLS lying in between. The two paths are shown in Fig. 8. From

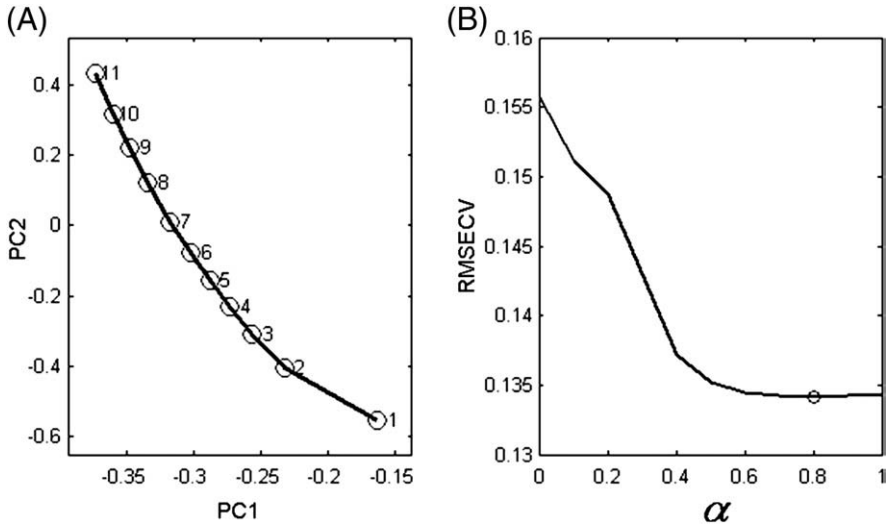

Fig. 7. Plot A shows the model path from PCR to PLS using all the 80 samples with nLV set to 10. The PCR and PLS models are marked by 1 and 11, respectively. Plot B shows the how RMSECV changes with increased α. The circle denotes the minimum.

Please cite this article as: H.-D. Li, et al., Uncover the path from PCR to PLS via elastic component regression, Chemometr. Intell. Lab. Syst. (2010), doi:10.1016/j.chemolab.2010.08.003

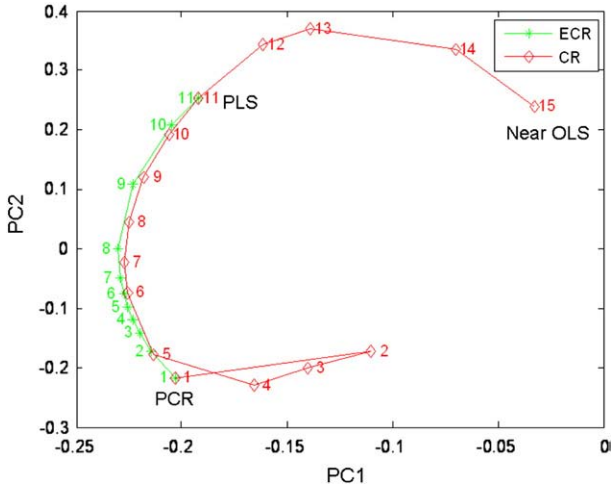

Fig. 8. The model paths computed by using ECR and CR, respectively.

this plot, it is clear that the model path provided by ECR directly evolves from PCR to PLS, while the path between PCR and PLS computed by CR is different. The different paths are in our opinion very interesting and thus deserve further investigating. Besides, the path from PLS to OLS from CR is also given. On the whole, it is shown that the model paths provided by ECR and CR are not the same, which will be further investigated in our future study.

5. Conclusions

Based on the theory of optimization, the elastic component regression is developed in this contribution for uncovering the path from PCR and PLS. ECR utilizes an adjustable supervising factor $\alpha \in [0,1]$ to supervise the **X**-matrix decomposition. In the context of ECR, PCR and PLS occupy the two opposite ends with an infinite number of transitional models that lie in between. The framework of ECR shows a natural progression from PCR to PLS and may help get some insight into their relationships in theory. The simulation study together with the application to the NIR data shows that in some cases the transitional models provided by ECR could perform better than PCR and PLS. Also indicated in this work is that it may reduce the risk of overfitting by controlling the value of the supervising factor α. Our future work will be focused on the model selection based on ECR. It is expected that ECR will find more applications in a variety of fields, e.g. genomics and metabolomics.

Acknowledgements

This work is financially supported by the National Nature Foundation Committee of P.R. China (Grant No. 20875104 and No. 10771217), and the International Cooperation Project on Traditional Chinese Medicines of Ministry of Science and Technology of China (Grant No. 2007DFA40680). The studies meet with the approval of the university's review board. The two potential reviewers are acknowledged for their enlightening comments.

References

[1] M.K. Hartnett, G. Lightbody, G.W. Irwin, Dynamic inferential estimation using principal components regression (PCR), Chemometr. Intell. Lab. 40 (1998) 215–224.
[2] S. Wold, K. Esbensen, P. Geladi, Principal component analysis, Chemometr. Intell. Lab. 2 (1987) 37–52.
[3] T. Rajalahti, R. Arneberg, F.S. Berven, K.M. Myhr, R.J. Ulvik, O.M. Kvalheim, Biomarker discovery in mass spectral profiles by means of selectivity ratio plot, Chemometr. Intell. Lab. 95 (2009) 35–48.
[4] M. Barker, W. Rayens, Partial least squares for discrimination, J. Chemometr. 17 (2003) 166–173.
[5] Q.-S. Xu, Y.-Z. Liang, H.-L. Shen, Generalized PLS regression, J. Chemometr. 15 (2001) 135–148.
[6] S. Wold, M. Sjöström, L. Eriksson, PLS-regression: a basic tool of chemometrics, Chemometr. Intell. Lab. 58 (2001) 109–130.
[7] G. Reich, Near-infrared spectroscopy and imaging: basic principles and pharmaceutical applications, Adv. Drug Deliver Rev. 57 (2005) 1109–1143.
[8] H.-D. Li, Y.-Z. Liang, Q.-S. Xu, D.-S. Cao, Key wavelengths screening using competitive adaptive reweighted sampling method for multivariate calibration, Anal. Chim. Acta 648 (2009) 77–84.
[9] I. Guyon, J. Weston, S. Barnhill, V. Vapnik, Gene selection for cancer classification using support vector machines, Mach. Learn. 46 (2002) 389–422.
[10] R. Madsen, T. Lundstedt, J. Trygg, Chemometrics in metabolomics—a review in human disease diagnosis, Anal. Chim. Acta 659 (2010) 23–33.
[11] A.J. Burnham, V. Roman, Frameworks for latent variable multivariate regression, J. Chemometr. 10 (1996) 31–45.
[12] M. Stone, R.J. Brooks, Continuum regression: cross-validated sequentially constructed prediction embracing ordinary least squares, partial least squares and principal components regression, J. R. Statist. Soc. B 52 (1990) 237–269.
[13] A. Bjorkstrom, R. Sundberg, A generalized view on continuum regression, Scand. J. Statist. 26 (1999) 17–30.
[14] S. De Jong, H.A.L. Kiers, Principal covariates regression part I, Theory. Chemometr. Intell. Lab. 14 (1992) 155–164.
[15] P.M. Lang, J.M. Brenchley, R.G. Nieves, J.H. Kalivas, Cyclic subspace regression, J. Multivariate Anal. 65 (1998) 58–70.
[16] J.H. Kalivas, Cyclic subspace regression with analysis of the hat matrix, Chemometr. Intell. Lab. 45 (1999) 215–224.
[17] G.A. Bakken, T.P. Houghton, J.H. Kalivas, Cyclic subspace regression with analysis of wavelength-selection criteria, Chemometr. Intell. Lab. 45 (1999) 225–239.
[18] P. Lang, A. Gironella, R. Venema, Properties of cyclic subspace regression, J. Multivariate Anal. 98 (2007) 625–637.
[19] S. de Jong, B.M. Wise, N.L. Ricker, Canonical partial least squares and continuum power regression, J. Chemometrics 15 (2001) 85–100.
[20] B.M. Wise, N.L. Ricker, Identification of finite impulse response models with continuum regression, J. Chemometr. 7 (1993) 1–14.
[21] R. Sundberg, Multivariate calibration—direct and indirect regression methodology, Scand. J. Statist. 26 (1999) 16–207.
[22] P. Geladi, Some recent trends in the calibration literature, Chemometr. Intell. Lab. 60 (2002) 211–224.
[23] P.J. Gemperline, A. Salt, Principal components regression for routine multicomponent UV determinations: a validation protocol, J. Chemometr. 3 (1989) 343–357.
[24] I.E. Frank, J.H. Friedman, A statistical view of some chemometrics regression tools, Technometrics 35 (1993) 109–148.
[25] S. De Jong, SIMPLS: an alternative approach to partial least squares regression, Chemometr. Intell. Lab. 18 (1993) 251–263.
[26] P. Geladi, B.R. Kowalski, Partial least-squares regression: a tutorial, Anal. Chim. Acta 185 (1986) 1–17.
[27] V. Centner, D.-L. Massart, O.E. de Noord, S. de Jong, B.M. Vandeginste, C. Sterna, Elimination of uninformative variables for multivariate calibration, Anal. Chem. 68 (1996) 3851–3858.
[28] H.-D. Li, Y.-Z. Liang, Q.-S. Xu, D.-S. Cao, Model population analysis for variable selection DOI: 10.1002/cem.1300 (2009).
[29] http://code.google.com/p/carspls/.

Please cite this article as: H.-D. Li, et al., Uncover the path from PCR to PLS via elastic component regression, Chemometr. Intell. Lab. Syst. (2010), doi:10.1016/j.chemolab.2010.08.003