Application of non-linear optimization methods to the estimation of multivariate curve resolution solutions and of their feasible band boundaries in the investigation of two chemical and environmental simulated data sets

R. Tauler*

Department of Environmental Chemistry, IIQAB-CSIC, Jordi Girona 18-26, Barcelona 08034, Spain

Received 18 October 2006; received in revised form 20 December 2006; accepted 22 December 2006

Available online 10 January 2007

Abstract

Although alternating least squares algorithms have revealed extremely useful and flexible to solve multivariate curve resolution problems, other approaches based on non-linear optimization algorithms using non-linear constraints are possible. Once the subspaces defined by PCA solutions are identified, appropriate rotation and perturbation of these solutions can produce solutions fulfilling the constraints obeyed by the physical nature of the investigated systems. In order to perform such a rotation, an optimization algorithm based in the fulfilment of constraints and some examples of application in chemistry and environmental chemistry are given. It is shown that the solutions obtained either by alternating least squares or by the new proposed algorithm are rather similar and that they are both within the boundaries of the band of feasible solutions obtained by an algorithm previously developed to estimate them.

© 2007 Elsevier B.V. All rights reserved.

Keywords: Multivariate curve resolution (MCR); Rotation ambiguities; Feasible band boundaries; Non-linear constrained optimization

1. Introduction

Chemometric methods based on bilinear models have revealed extremely useful to investigate chemical and environmental systems. It was already early in the seventies that factor analysis derived methods [1] such as principal component analysis, PCA [2] and multivariate curve resolution, MCR [3–7] were proposed and satisfactorily used to solve chemical problems related to mixture analysis. Thus, in general, modern analytical instrumentation (i.e. in spectroscopy, electrochemistry, etc.) provides rich multivariate information describing the composition, state or evolution of a particular chemical system or process. This numeric information is generally stored in a rectangular data table or data matrix. This information however is not selective, i.e. it describes the composition, state or evolution of the different chemical components present in the system simultaneously and not separately. In analytical chemistry this lack of ‘selectivity’ of the instrumental measures is usually resolved by means of the previous physical separation of the chemical components in the mixtures using chromatography, capillary electrophoresis or flow injection analysis techniques, coupled to spectroscopic detection methods, among others.

Chemometric methods based on bilinear models attempt doing a similar task (resolving chemical components in mixtures) but without their physical separation, by means of the mathematical resolution of their signal contributions. The main assumption of bilinear models is that multivariate instrumental measurements are a linear sum of the pure individual contributions of the different chemical components in the system and that each one of this component contributions can be described by the product of two factors, one related with its nature (i.e. the pure spectrum of a component) and another one related with the scale or concentration of this contribution in the particular analyzed sample.

In many factor analysis methods, like in principal component analysis [2], the mathematical resolution of the contributions of the different components is performed under well-defined orthogonal constraints, normalization and maximum variance

* Tel.: +34 93 4006140.
E-mail address: rtaqam@iiqab.csic.es.
explained by successive resolved components. Using these constraints, PCA solutions are unique and there is no ambiguity in their estimation. However, since orthogonal constraints are not in general fulfilled by the contributions of the true underlying components in their mixtures (since in general these contributions do overlap and they are partially correlated), the component contributions deduced by PCA are completely different to the true ones. In fact, PCA solutions are also a particular linear combination of the true underlying component contributions including some embedded noise (see below). This is the reason why PCA is so useful to investigate the mathematical structure of the data and to interpret the sources of data variation produced by the true underlying components, but it is not so useful to estimate how the ‘true’ sources of data variation really are. Moreover, in many circumstances, interpretation of component contributions from PCA factor decompositions is rather cumbersome and not directly interpretable.

On the other hand, in multivariate curve resolution methods [3–7], although they also assume the same bilinear model as in PCA, the mathematical resolution of the components is performed using other type of more natural (with more physical sense) constraints like non-negativity, unimodality or mass balance-closure equations. For instance it has no sense to have a negative chemical concentration nor a negative absorption spectra. So, MCR methods attempt to recover the true underlying contributions of the components in the mixture from the mathematical resolution of the original data system. The difficulty or trade off here is that the constraints applied in MCR do not assure a unique solution like in PCA. The mathematical properties of the MCR decomposition are not so well-defined as in PCA and the resolved contributions are more ambiguous. The art of MCR in practice is to find and apply appropriate constraints to decrease and hopefully eliminate the ambiguity and obtain unique solutions. If this is achieved, it is then clearly advantageous using MCR compared to using PCA, because the solutions obtained by MCR will have direct physical meaning and describe the true nature of the component contributions, without the need of additional effort nor interpretation of them. They give directly the sought analytical information, like in chromatographic separation methods. A possible additional advantage of using MCR methods for mixture resolution is that the mathematical resolution in MCR methods can be achieved with less time, costs and chemical efforts than using expensive chromatographic methods. Among MCR methods, alternating least squares (MCR-ALS) algorithms [8–11] have revealed specially useful and simple to use.

In this work, two complementary methods are proposed to obtain MCR solutions and examine their reliability. In the first method, a new approach based on the use of unconstrained non-linear optimization, called here MCR-FMIN, is proposed and compared to the previously developed alternating least squares approach, MCR-ALS. And in the second method, an approach previously proposed to define the boundaries of the bands describing the set of feasible MCR solutions [13–15] and called here MCR-BANDS, is applied to compare the reliability of MCR solutions obtained either by MCR-ALS or by MCR-FMIN methods. The approach used by MCR-BANDS is based on a non-linear optimization method under non-linear constraints. This method is applied to investigate if both solutions, MCR-ALS and MCR-FMIN, are in fact equally feasible, i.e. the solutions obtained by both approaches are indistinguishable under the set of applied constraints.

Two data examples are used for illustration. A first example is the mathematical resolution of three coeluted peaks in their partial separation by liquid chromatography using multivariate diode array detection. The second example is the resolution of contamination sources present in multicomponent analysis of a set of samples. In both cases, the data have been mathematically simulated from the previous knowledge of the true experimental systems. The reason for using simulated data instead of experimental data in this work was to allow a better examination of the results obtained by the proposed methods from a profiles recovery point of view, i.e. to examine if the true (simulated) underlying factors were appropriately recovered by the proposed methods and to check for the presence of rotation ambiguities in the obtained solutions.

2. Theory and methods

Two similar bilinear models can be written for principal component analysis and multivariate curve resolution:

\[
\text{PCA: } X = X_{\text{PCA}} + E_{\text{PCA}} = UV^T + E_{\text{PCA}} \quad (1)
\]

\[
\text{MCR: } X = X_{\text{MCR}} + E_{\text{MCR}} = GF^T + E_{\text{MCR}} \quad (2)
\]

where \(X\) is the data matrix; \(X_{\text{PCA}}\) and \(X_{\text{MCR}}\) are the reproduced data matrices using, respectively, the PCA and MCR models; \(E_{\text{PCA}}\) and \(E_{\text{MCR}}\) are the residuals or non-modelled parts of \(X\) by PCA and MCR. \(U\) and \(G\) are the score (row factor) matrices obtained, respectively, by PCA and MCR and \(V^T\) and \(F^T\) are the loadings (column factor) matrices obtained, respectively, by PCA and MCR. Superscript ‘T’ in Eqs. (1) and (2) represent transpose matrices.

Matrix decomposition using PCA model in Eq. (1) is unique (no ambiguity) due to the very restrictive constraints applied during the determination of the component \(U\) and \(V^T\) (scores and loadings) matrices: orthogonal component matrices \(U\) and \(V^T\); normalization of \(V^T\) loadings and maximum explained variance of \(V^T\) loadings [1,2]. Matrix decomposition using MCR model in Eq. (2) is not unique in general because the constraints imposed during the determination of factor matrices are less restrictive and allow for an undetermined number of equivalent solutions (fitting equally well the data). Constraints usually applied in MCR are non-negativity, unimodality, closure, and some type of equality constraints based on local rank and on selectivity from previously known information [8]. As it was stated in Section 1 however, whereas MCR solutions have physical meaning because they are based on natural constraints fulfilled by the true sources of data variation, PCA solutions lack this physical meaning because they are based on mathematical properties which are not fulfilled in general by the true sources of data variation. So, the good mathematical properties of PCA decomposition are opposed to the physically sounder good resolution properties of MCR decompositions.
The question is then to look whether it is possible to go from good mathematically PCA solutions to the good physically MCR solutions. This may be stated by the following equation:

$$X_{\text{PCA}} = U V^T = U T T^{-1} V^T \Rightarrow G F^T = X_{\text{MCR}}$$

(3)

where $T$ would be an appropriate invertible matrix which rotates $U$ to $G$ and $V^T$ to $F^T$. This rotation should be performed taking into account that $G$ and $F^T$ should fulfill the constraints imposed to MCR solutions (see above).

In the general case however, and due to experimental noise and to the different ways on how matrix decompositions of Eqs. (1) and (2) are performed by PCA and MCR methods, $E_{\text{PCA}}$ will not be equal to $E_{\text{MCR}}$ and $X_{\text{PCA}}$ will be also different to $X_{\text{MCR}}$, and therefore going from PCA solutions to MCR solutions is not always an easy trivial task. The first problem is that straightforward application of PCA and MCR usually overfit the solutions, and therefore going from PCA solutions to MCR solutions may cause some problems due to the difficulty to make constraints to be fulfilled by these PCA rotated solutions. Or stated in another way, the subspace spanned by PCA solutions is not exactly equal to the subspaces spanned by MCR solutions, simply because the embedded noise in PCA solutions compels the total fulfilment of the required constraints. Therefore, direct rotation of PCA solutions by matrix $T$ might not give solutions fulfilling the desired constraints, like non-negativity and other. This may be stated in the following equations:

$$E_{\text{PCA}} \neq E_{\text{MCR}}$$

(4)

$$X_{\text{PCA}} \neq X_{\text{MCR}}$$

(5)

$$U V^T \neq G F^T$$

(6)

where the notation ‘#’ is used to indicate that both sides of the equation are not exactly equal. Taking into account Eqs. (4)–(6), it is obvious that deriving $G$ and $F^T$ from $U$ and $V^T$ rotation is not obvious (and in many cases not possible) in the presence of embedded noise in PCA solutions. This situation is often encountered for experimental data where noise contributions are important and traditional PCA algorithms do not account for them. It could probably not be the case when maximum likelihood [16,17] or total least squares algorithms [18] for PCA were used, but the general use of these more statistically sounder approaches is however hindered by the fact that they require a previous knowledge of the data noise structure, which is not the case in many practical situations.

In the case of alternating least squares approaches like MCR-ALS, the data are fitted directly from initial estimates of $G$ or $F^T$ already fulfilling most of the required constraints (for instance non-negativity) and PCA is only used to define the data subspace to approach and fit. ALS optimization implies a set of linear regression steps to solve local linear models where alternatively one of the two factor matrices, $G$ or $F^T$ are fixed. In any case, solution subspaces obtained by MCR-ALS solutions are found to approach PCA solutions subspaces, although they will not be exactly the same.

In order to evaluate how good the model (PCA or MCR) fits the data, the two following parameters and equations are used:

$$\text{Lack of fit: } \text{lof} = \frac{\sum_{i=1}^{I} \sum_{j=1}^{J} (x_{i,j} - \hat{x}_{i,j})^2}{\sum_{i=1}^{I} \sum_{j=1}^{J} x_{i,j}^2} \times 100$$

(7)

$$R^2: \quad R^2 = \left(1 - \frac{\sum_{i=1}^{I} \sum_{j=1}^{J} (x_{i,j} - \hat{x}_{i,j})^2}{\sum_{i=1}^{I} \sum_{j=1}^{J} x_{i,j}^2}\right) \times 100$$

(8)

In these equations $x_{i,j}$ and $\hat{x}_{i,j}$ are, respectively, the experimental and calculated data values using PCA or MCR. The sums are extended to $i=1, \ldots, I$ rows and $j=1, \ldots, J$ columns of the data matrix $X$. The closer the lack of fit and $R^2$ are, respectively, to zero and one, the best model fitting of the experimental data has been achieved. When the fit is good, lof provides more discrimination between similar models, as it is the case when noise is very low like in spectroscopic measures. In the case of environmental data tables however, data fitting is not so good because of larger model unexplained variances and $R^2$ is then preferred. Both lack of fit and $R^2$ values should be compared with expected levels of experimental noise.

In order to evaluate the correct recovery of component profiles in factor matrices $G$ and $F^T$, and test for the presence of rotational ambiguities, a similarity or correlation coefficient between the ‘true’ component profiles and the profiles obtained by MCR may be calculated. A similarity value close to one means a correct recovery of this particular profile and the absence of rotational ambiguity for this profile. A value differing considerably from one indicates a bad recovery of this component profile and the presence of significant rotation ambiguity. So, good values of lack of fit (close to zero) and $R^2$ (close to one) together with similarity values for all the component profiles close to one would mean that MCR achieved the right solution without ambiguities. When also good values of lack of fit and $R^2$ are obtained, but similarity values are poor, the model fits well the data but the solutions are not the true ones, but linear combination of them, and therefore rotation ambiguities are still present and they have not been solved by the imposed constraints. However, similarity evaluations can only be performed for cases where the model profiles are already known and this obviously is not the case in general, since their estimation is the goal of the study. This is the reason why simulations are useful to compare between different MCR methods. Since for simulated data the ‘true’ profiles are known in advance, the calculation of similarity values is possible. Below, another method (MCRBANDS) is proposed for the evaluation of rotation ambiguity effects, based in the calculation of the boundaries of the bands of feasible solutions caused by rotation ambiguities.

3. New algorithm for MCR based on non-linear optimization (MCR-FMIN)

A new MCR method is proposed based on the minimization of an objective function defined directly from the constraints non-fulfilment and being always in the subspace spanned by PCA solutions. This objective optimization function is
defined by:

\[
f(T) = c_{\text{norm}}(T) + c_{\text{non-neg}}(T) + c_{\text{unimod}}(T) + c_{\text{clos}}(T) + c_{\text{equi}}(T) + \cdots \tag{9}\]

where \( f(T) \) is the objective scalar function to minimize and \( c_{\text{norm}}(T), c_{\text{non-neg}}(T), c_{\text{unimod}}(T), c_{\text{clos}}(T), c_{\text{equi}}(T) \) are the also scalar functions (for normalization, non-negativity, unimodality, closure and equality constraints, respectively) of the constraints having large values when constraints are not fulfilled and having close to or zero values when the constraints are nearly or totally fulfilled. For instance in the case of non-negativity constraints these scalar functions may be easily defined by the following equations

\[
c_{\text{non-neg}}(T) = \sum (g_i < 0)^2 \tag{10}\]

\[
c_{\text{fnon-neg}}(T) = \sum (f_j < 0)^2 \tag{11}\]

\[
c_{\text{non-neg}}(T) = c_{\text{fnon-neg}}(T) + c_{\text{non-neg}}(T) \tag{12}\]

In Eqs. (10) and (11), the non-negativity constraints scalar functions are obtained from the sum of the squares of the negative values found in \( G \) and \( F^T \) components. For positive or zero value, these functions are equal to zero. In a similar way, appropriate scalar functions may be defined for the other constraints.

The goal of the optimization is then to find a \( T \) rotation matrix that minimizes \( f(T) \), which is not a least squares error loss function defined from the residuals \( E_{\text{PCA}} \) or \( E_{\text{MCR}} \) in Eqs. (1) and (2) like in resolving factor analysis, RFA [19], positive matrix factorization, PMF [20], or in alternating least squares methods, ALS [6–12], but from an objective function directly defined from the non-fulfillment of the constraints. In fact, MCR-FMIN solutions will always fit equally well the data, and their rotation will not change the fit, i.e. they will fit the same since PCA rotation matrix \( T \) does not change the fit and the objective function only reflects the fulfillment of constraints. In contrast, in the three methods before mentioned, RFA, PMF and ALS, the objective function to minimize is always defined as a least squares function of the residuals (error). In RFA, rotated PCA solutions are modified iteratively to fulfill the constraints and the perturbed solutions are then used to calculate the residuals of the least squares function to be minimized by a non-linear optimization procedure. This may be problematic when rotation of PCA solutions cannot converge to a solution fulfilling the constraints which minimizes at the same time the residuals least squares function. In the case of PMF and ALS, the solutions estimated during the optimization are not obtained from a direct rotation of PCA solutions, using \( T \) rotation matrix. In PMF, estimates of \( G \) and \( F^T \) matrices are obtained directly and iteratively from the non-linear minimization of a weighted residuals least squares function. The algorithm used in this case for such an optimization is similar to that used in RFA, but in this case the optimization task is even more involved due to the very large number of non-linear parameters to be simultaneously optimized, much larger than the number of \( T \) values in RFA or FMIN. Finally in the case of ALS, the situation is somewhat intermediate, since the parameters of the optimization are separated in two groups according to their bilinear structure (Eqs. (1) and (2)). If alternatively, the \( G \) and \( F^T \) matrices are considered constant, then the least squares optimization becomes linear in the parameters, and it runs rather easily and fast. See references [6–12,19,20] for more details and comparison among these different approaches.

Function \( f(T) \) in Eq. (9) can be minimized using different minimization general routines available in scientific software packages. In particular, the \textit{fminunc} general unconstrained minimization function from MATLAB Optimization Toolbox (version 3.0, 2004, MATLAB is a trademark from The Mathworks Inc., Natic, Ma, USA [23]) was selected since it is considered the state of the art for this type of optimization problems [21–23]. The algorithm in the \textit{fminunc} function uses a quasi-Newton non-linear optimization with a cubic interpolation line search and allowed for accurate numerical approximation of both, for the gradient and for the Hessian of the objective function to minimize. Alternatively to perform also the same task, a simplex optimization can also be proposed, although at much larger computational and time expenses. Such an optimization can be for instance performed using the MATLAB Optimization Toolbox \textit{fminsearch} function. The minimization can also be carried out by a combination of both approaches to assure that further improvement of the minimization function cannot be achieved after repeated attempts. Values finally achieved of \( f(T) \) at the optimum are checked to test for constraints fulfillment. The solutions obtained for \( T, G \) and \( F^T \) matrices, are then checked and the lack of fit and \( R^2 \) parameters evaluated. All the individual elements of the \( T \) matrix are optimized simultaneously. Scaling problems are avoided using either normalization or closure mass balance constraints (\( c_{\text{norm}}(T) \) and \( c_{\text{clos}}(T) \) in Eq. (9)) during the optimization. Rotation matrix \( T \) was initialized using the identity matrix of size equal to \( T \) matrix. In some cases, this initialization was difficult because of local minima and difficulty finding a feasible region where constraints are fulfilled at different \( T \) values. In these circumstances, initialization is performed using a different set of \( T \) values obtained either randomly or from the previous knowledge of the system. Starting with a simplex optimization instead of using the quasi-Newton optimization method can also help in these cases.

4. Calculation of the band boundaries of feasible solutions in MCR (MCR-BANDS)

As it was mentioned above, once a particular MCR solution (from MCR-ALS or from MCR-FMIN) is obtained, it is of great interest the evaluation of the possible rotation ambiguity associated with such a solution and to check whether the estimated solution is unique. If the MCR solution is unique, then the MCR solution should be the correct one (apart from noise effects). This means, that when the solutions are not unique due to the presence of rotation ambiguities, a linear combination of the MCR solutions fulfilling the constraints can have been obtained, and there is no way to be sure that the obtained one is the true one, unless more information is provided to the system. Linear combinations of the solutions may be simply obtained by appropriate rotation matrices \( T \) in Eq. (3). In a previous work [14], a method for the evaluation of the rotation ambiguity associated with MCR solu-
tions has been proposed. The method performs the calculation of the boundaries of the bands of the feasible solutions fitting equally well the experimental data and fulfilling the constraints of the system. These boundaries are obtained from a constrained non-linear optimization of an objective function which depends on rotation matrix $\mathbf{T}$ of Eq. (3) and which is defined in terms of the component profiles signal contribution to the whole signal contribution:

$$f_k(\mathbf{T}) = \sum_{i=1}^{f} \sum_{j=1}^{f}(g_{i,k}(\mathbf{T})f_{k,j}(\mathbf{T}))^2$$

$$= \frac{||g_i(\mathbf{T})f_k^T(\mathbf{T})||}{||\mathbf{G}^T||}$$

(13)

This function $f_k(\mathbf{T})$ is scaled between 0 and 1 and provides an estimation of the relative contribution of component $k$ to the whole signal expressed by the model in the component matrices $\mathbf{G}$ and $\mathbf{F}^T$. The method attempts the estimation of $\mathbf{T}$ matrices in Eq. (3) giving $g_i(\mathbf{T})$ and $f_i(\mathbf{T})$ component profiles making maximum and minimum $f/k(T)$ of Eq. (13). Only those $\mathbf{T}$ values producing $\mathbf{G}$ and $\mathbf{F}^T$ component matrices fulfilling the constraints of the system should be considered. So, the optimization of the objective function in Eq. (13) should be performed under constraints, which are defined by a vector of constraints, $c(\mathbf{T})$, in a similar way than previously in Eqs. (10)–(12) for non-negativity constraints.

The minimization and maximization of the optimization function defined in Eq. (13) is performed using the MATLAB Optimization Toolbox function $fmincon$. This function is based on a Sequential Quadratic Programming algorithm which uses a constrained quasi-Newton (Hessian approximated) method and a line search procedure to solve the Kuhn–Tucker equations obtained from necessary and sufficient conditions of the non-linear optimization with non-linear constraints problem [21–23]. See also previous works [14,15] for more details about the implementation and previous applications of the proposed method. In this work, this method has been called MCR-BANDS.

For each component $k$, MCR-BANDS calculates the profiles found at the maximum and minimum boundaries of the $f_k(\mathbf{T})$ function defined by Eq. (13). These profiles are plotted to show their difference. In case they are practically equal, the component profiles obtained by MCR are practically unique and no rotation ambiguities are associated with them. In case this happens for all component profiles, rotation ambiguities have been totally solved and MCR solutions are considered to be the true ones (apart from noise). In the more general case however, when the maximum and minimum boundary solutions differ for one or for several of the component profiles, the rotation ambiguity will be still present in spite of the constraints applied, and therefore, the obtained MCR solutions are only one among the possible ones and we do not have any proof that they correspond to the true one unless more information in the form of constraints are provided. Thus, MCR-BANDS method provides an easy way to check for rotation ambiguity of MCR solutions and also to see how different could be ‘true’ solutions from MCR solutions. Another use of MCR-BANDS employed in this work is the comparison between solutions obtained using different MCR methods, like MCR-ALS and MCR-FMIN, to see whether they are both within the boundaries of feasible band solutions.

5. Software

MCR-FMIN and MCR-BANDS methods have been implemented as m functions for the MATLAB computational environment. Source files containing them are available under request to the author of this paper (e-mail: rtaqam@iqab.csic.es). They require the use of $fminuncon$, $fminsearch$ and $fmincon$ functions from the version 3 or higher of the MATLAB Optimization Toolbox [23].
7. Results and discussion

7.1. Application of the MCR-FMIN method

In Fig. 1D results obtained using the new MCR-FMIN method when applied to the chromatographic coelution data (data example 1) are given. First in Fig. 1A, the plot of the data matrix, both column-wise (elution profiles at the different wavelengths) and row-wise (spectra at the different elution times) are given. Application of PCA gives the scores and loadings profiles plotted in Fig. 1C. These profiles should show shape features related with the true elution and spectra profiles, but are obviously different to them (they are an orthogonal linear combination of them). Rotation of these profiles using the algorithm MCR-FMIN here proposed, based on the minimization of the function defined through the fulfilment of constraints (non-negative for both elution and spectra profiles, unimodality for elution profiles and spectra normalization) are given in Fig. 1D. Lines with the ‘+’ symbols are those obtained using the MCR-FMIN method and continuous lines without any symbol are the profiles used for data simulation. Some differences exist between ‘true’ profiles and those obtained by MCR-FMIN. Similarly, MCR-ALS resolved profiles (in Fig. 1C, profiles with discontinuous broken lines) using the same constraints are not exactly equal to the ‘true’ ones nor to the MCR-FMIN ones, although the differences are small. In Table 1, these differences are quantified numerically. In the first two columns, the fitting values are given, both in terms of lack of fit and in terms of $R^2$. In the three cases, the fit values are practically the same, for the true, for the MCR-ALS and for the MCR-FMIN profiles. This indi-

<table>
<thead>
<tr>
<th>Item</th>
<th>Theoretical</th>
<th>MCR-ALS</th>
<th>MCR-FMIN</th>
</tr>
</thead>
<tbody>
<tr>
<td>lof (%)</td>
<td>0.385</td>
<td>0.416</td>
<td>0.354</td>
</tr>
<tr>
<td>$R^2$ (%)</td>
<td>99.998</td>
<td>99.998</td>
<td>99.999</td>
</tr>
<tr>
<td>$g_1$</td>
<td>–</td>
<td>0.999</td>
<td>0.993</td>
</tr>
<tr>
<td>$g_2$</td>
<td>–</td>
<td>0.977</td>
<td>0.998</td>
</tr>
<tr>
<td>$g_3$</td>
<td>–</td>
<td>0.995</td>
<td>0.990</td>
</tr>
<tr>
<td>$f_1$</td>
<td>–</td>
<td>0.979</td>
<td>0.982</td>
</tr>
<tr>
<td>$f_2$</td>
<td>–</td>
<td>0.995</td>
<td>0.969</td>
</tr>
<tr>
<td>$f_3$</td>
<td>–</td>
<td>0.996</td>
<td>0.972</td>
</tr>
</tbody>
</table>

lof and $R^2$ are defined by Eqs. (7) and (8); values in the table were obtained for theoretical, MCR-ALS and MCR-FMIN $G$ and $F^T$ profiles. $g_1$, $g_2$ and $g_3$ are the elution (column) profiles in factor matrix $G$; numerical values give the similarities (correlation coefficients) between theoretical and MCR-ALS or MCR-FMIN profiles. $f_1$, $f_2$ and $f_3$ are the spectra (row) profiles in factor matrix $F^T$; numerical values give the similarities (correlation coefficients) between theoretical and MCR-ALS or MCR-FMIN profiles.
cates, that since the three types of solutions fulfil the constraints used (non-negativity, unimodality and normalization), the differences obtained in every case should be related to the unresolved rotation ambiguities. The value achieved for the objective $f(T)$ function (Eq. (9)) for the non-fulfilment of constraints was of only $2.35E^{-05}$ after application of fminuncon and fminsearch MATLAB Optimization toolbox functions, which was considered to be very low and accepted. The fit values are in all cases (MCR-ALS and MCR-FMIN) in agreement with the noise level used in this case, which was around 0.4% (see data description).

In the next six lines of Table 1, similarity (correlation) values between true and MCR-FMIN or MCR-ALS profiles are given. As it can be seen from these similarity values in Table 1, recoveries of MCR-FMIN and MCR-ALS profiles are in general rather good, although they are not exactly equal to the true ones, as it was also clear from Fig. 1C and D. MCR-FMIN and MCR-ALS did not produce the same solutions because rotation ambiguity could not be totally solved in this case by the proposed constraints. This will be better evaluated below by the application of the MCR-BANDS method.

The second example of application is shown in Fig. 2, where the results obtained in the analysis of the environmental data set at the 1% noise level are given. In this case also, profiles obtained by PCA resemble little to true ones, although they are obviously related with them. Either MCR-FMIN solutions obtained by rotation of PCA solutions under constraints (non-negativity for profiles in the two modes and normalization of column- or loading profiles in matrix $F_T$) or MCR-ALS solutions using the same constraints produced component profiles similar but not equal to the true ones, in this case with less ambiguity than in the previous chromatographic case. In Table 2, a numerical evaluation of these results is given in more detail. It is important to realize first, that in this case both MCR-FMIN and MCR-ALS slightly overfit the data, since the lack of fit obtained by these two methods is lower (4.17 and 4.18%) than it should be (4.83%) according to the true component profiles used for the data simulation. This is also clearly reflected when the values of $R^2$ are compared. For MCR-FMIN and MCR-ALS $R^2$ values are larger (99.83 and 99.82%) than they should be (99.76%). The differences are however small and they will not be visible in the shape of the recovered profiles. In this case, the value achieved for the objective $f(T)$ function is $4.46E^{-04}$ (after repeated application of fminuncon and fminsearch MATLAB Optimization Toolbox functions [23]), which it is a little larger than before, but it is still considered very low and a measure of the embedded noise in PCA rotations compelling further constraint fulfilment. Looking
is clearly reflected in the differences between lack of fit and more noise have been analyzed as previously. Overfitting here 
ponent profiles as for Table 2 and Fig. 2) but with considerably 
gf

$R^2$ (%) 99.76 99.82 99.83
$f_1$ 0.996 0.995
$f_2$ 0.984 0.950
$f_3$ 0.970 0.893
$f_4$ 0.999 0.998

Rlof and $R^2$ are defined by Eqs. (7) and (8); values in the table were obtained for theoretical, MCR-ALS and MCR-FMIN G and $F^T$ profiles. $g_i, g_j, g_k$ and $g_l$ are the column component profiles in factor matrix $G$; numerical values give the similarities (correlation coefficients) between theoretical and MCR-ALS or MCR-FMIN profiles.

at the similarity values to evaluate the presence of rotation ambiguities, it is evident that solutions obtained either by MCR-ALS or by MCR-FMIN are very close to the true ones, since similarity correlation values are close to one in all cases. It is worth to note here that in this case the presence of rotation ambiguities was less significant due to the intrinsic nature of the environmental profiles, with less smooth and more distinct shape features (less overlap), together with the fact that environmental profiles had some zero or close to zero values. The same has been observed by other authors in the investigation of rotation ambiguities in source apportionment of environmental data [24].

Finally in Table 3, the same simulated data (the same component profiles as for Table 2 and Fig. 2) but with considerably more noise have been analyzed as previously. Overfitting here is clearly reflected in the differences between lack of fit and $R^2$ values (43.7 and 80.9%) obtained for true component profiles, compared to those obtained by MCR-FMIN (36.5 and 88.8%) and MCR-ALS (37.4 and 86.8%). This overfit is also reflected in the shapes of the profiles (not shown) and recoveries of component profiles. Indeed, the values of the similarities shown in Table 3 at the high noise level case are worse than those obtained at the low noise level case in Table 2. Moreover, the optimization was now more difficult, with a value finally achieved for the objective $f(T)$ function of only 0.2027, considerably larger than for the two previous cases after repeated application of fminunc and fminsearch, reflecting a more difficult optimization task in the constraints fulfilment when PCA solutions were rotated. The reason for this should be related with the amount of embedded noise in PCA solutions which compels the fulfilment of constraints simply by their rotation. Moreover, since a certain amount of rotation ambiguities were still present in this case, it is difficult to clearly distinguish what is causing what, either noise propagation or increase of rotation ambiguity (see [25,26] for recent discussions about this issue). For the moment the conclusion is that the degree of overfitting and quality of results depends obviously on the noise level (more overfitting at higher noise levels) and on the noise structure (random noise produces less overfitting than systematic and proportional noise). Further research is needed to investigate this and some work is pursued in this direction.

7.2. Calculation of the band boundaries of feasible solutions using the MCR-BANDS method

In Fig. 3, results of calculation of band boundaries of feasible solutions using the MCR-BANDS method [14] previously explained for example 1 LC-DAD data set are given. Thick continuous lines give the boundaries calculated for the three different components of the system. In Fig. 3A for the elution profiles (matrix $G$ in Eq. (2)) and in Fig. 3B for the spectra profiles (matrix $F^T$ in Eq. (2)). Theoretical, MCR-ALS and MCR-FMIN profiles are also given in discontinuous broken lines (see plot labels). In all cases, the theoretical values

| Table 2 | Comparison MCR-ALS vs. MCR-FMIN (environmental data, 1% low noise case) |
|---------|---------------------------|-----------------------------|
| Item    | Theoretical | MCR-ALS | MCR-FMIN |
| Rlof (%)| 4.83        | 4.18    | 4.17     |
| $R^2$ (%) | 99.76  | 99.82  | 99.83    |
| $g_1$  | 0.996     | 0.995   |
| $g_2$  | 0.984     | 0.950   |
| $g_3$  | 0.970     | 0.893   |
| $g_4$  | 0.999     | 0.998   |

| Table 3 | Comparison MCR-ALS vs. MCR-FMIN (environmental data, 10% high noise case) |
|---------|---------------------------|-----------------------------|
| Item    | Theoretical | MCR-ALS | MCR-FMIN |
| Rlof (%)| 43.7       | 37.4    | 36.6     |
| $R^2$ (%) | 80.9    | 86.0    | 86.8     |
| $g_1$  | 0.966     | 0.956   |
| $g_2$  | 0.814     | 0.785   |
| $g_3$  | 0.894     | 0.717   |
| $g_4$  | 0.987     | 0.935   |
| $f_1$  | 0.959     | 0.935   |
| $f_2$  | 0.860     | 0.854   |
| $f_3$  | 0.845     | 0.746   |
| $f_4$  | 0.980     | 0.922   |

lof and $R^2$ are defined by Eqs. (7) and (8); values in the table were obtained for theoretical, MCR-ALS and MCR-FMIN G and $F^T$ profiles. $g_i, g_j, g_k$ and $g_l$ are the column component profiles in factor matrix $G$; numerical values give the similarities (correlation coefficients) between theoretical and MCR-ALS or MCR-FMIN profiles. $f_1, f_2, f_3$ and $f_4$ are the row component profiles in factor matrix $F^T$; numerical values give the similarities (correlation coefficients) between theoretical and MCR-ALS or MCR-FMIN profiles.

Fig. 3. MCR-BANDS band boundaries of feasible solutions, and component profiles obtained by MCR-ALS, MCR-FMIN and from data simulation for the LC-DAD data set. (A) Plot of elution profiles. (B) Plot of spectra profiles.
and the MCR solutions are within the boundaries estimated by MCR-BANDS. This simply means that either MCR-ALS or MCR-FMIN solutions are equivalent to the true solutions (used in the data simulation) apart from noise, since all MCR solutions fit equally well the data and obey the same constraints of non-negativity, unimodality and spectra normalization. There is no way to say which one is better unless more information is provided to the system in the form of constraints. This proves therefore, that either MCR-ALS or MCR-FMIN provided good MCR solutions, equivalent to the true ones, apart from the unresolved rotation ambiguities still present. In order to get closer solutions to the true ones and still produce narrower feasible bands, more constraints should be applied. Two possibilities could be proposed at this stage. A first possibility to decrease the influence of rotational ambiguities in this case could be the use of additional constraints defining precisely the elution time windows (selectivity and local rank constraints [27]) of each of the coeluted components, information which could be obtained from local rank analysis or by evolving factor analysis methods [28,29], although this information could be not enough accurate in practice due to noise level perturbing effects. A second possibility would be the simultaneous analysis of multiple chromatographic runs of the same system at different experimental conditions, or even better, the simultaneous analysis of the data matrix of example 1 together with other data matrices obtained in similar chromatographic analysis of samples with only one of the components (i.e. samples of standards). Examples of difficult resolution of strongly overlapped profiles in chromatography using MCR-ALS have been solved in this way [30–32] and they are not further discussed here for brevity.

In Fig. 4, results of calculation of band boundaries of feasible solutions using the MCR-BANDS method [14] for the example 2 environmental data set are given. As before for Fig. 3, thick continuous lines are for the band boundaries and discontinuous broken lines are, respectively, for the true, MCR-ALS and MCR-FMIN solutions. In this case, the bands are so narrow, that they can be hardly distinguished, much narrower than in the case of the chromatographic data. This means that there is much less rotation ambiguity associated to the resolved environmental component profiles than for the chromatographic component profiles. The reason for this should be related with the different shape features of the component profiles in both cases. In the case of the environmental data set, composition ($F^T$) and contribution ($G$) profiles have shapes with very different features, much less smooth than those obtained for the chromatographic data. The number of possible linear combinations giving the same data fit under constraints in the MCR analysis of the environmental data set resulted to be much lower than in the MCR analysis of the chromatographic data set. One of the more decisive aspects for this drastically diminution of rotation ambiguity effects under non-negativity constraints is the presence of zero or close to zero values in the component profiles in one or both modes, as it happens in this case with the environmental data set. This has been already pointed out by different authors in previous works [8,24,27] and it may be proved mathematically. In fact, the presence of zero (or close to zero values) in some of the variables of the component profile should be related with the presence of selective wavelengths (channels) in spectroscopy or in the concentration profiles (time, pH, etc.), which is often recognized as the key feature to solve rotation ambiguities in curve resolution studies of chemical reactions and chromatographic systems [8,27].

8. Conclusions

The new approach for multivariate curve resolution (MCR-FMIN) proposed in this work has been shown to be a complementary tool to previous developed curve resolution methods using alternating least squares algorithms (MCR-ALS). This approach resembles a previous one developed to define the band boundaries of feasible solutions from MCR (MCR-BANDS) and it opens a new possibility to check for their reliability. The following conclusions have been derived from the present work: (1) state of art non-linear optimization algorithms (like those in MATLAB optimization Toolbox [23], fminunc and fmincon) are extremely powerful and useful to solve multiple problems in chemometrics; (2) good MCR solutions can be obtained using these non-linear optimization algorithms based on the rotation of PCA solutions and by minimization of a function defined from the non-fulfilment of constraints (new MCR-FMIN method); (3) solutions obtained using different multivariate curve resolution methods like MCR-ALS or MCR-FMIN are practically analogous and their differences are mostly due to the unresolved rotation ambiguities (they are within the band boundaries of feasible solutions evaluated by MCR-BANDS); (4) resolution of environmental profiles is less affected by rotational ambiguities than other type of smoother profiles, like spectroscopic, chromatographic, or chemical reaction (kinetic, equilibria) profiles. The smoother and featureless are the profiles (like in chromatography, equilibrium and kinetic reaction concentration profiles), the largest is the possibility of rotation ambiguity effects. Further work is planned to test the approaches presented here.

Fig. 4. MCR-BANDS band boundaries of feasible solutions, and component profiles obtained by MCR-ALS, MCR-FMIN and from data simulation for the environmental data set. (A) Plot of column component profiles. (B) Plot of row component profiles.
and to extend their use in the analysis of more complex data sets.

Acknowledgement

Research Project: Ministerio de Educación y Ciencia, Spain, BQU2003-00191.

References