Analytica Chimica Acta 791 (2013) 25–35

Contents lists available at SciVerse ScienceDirect

Analytica Chimica Acta

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

Investigation of the equality constraint effect on the reduction of the rotational ambiguity in three-component system using a novel grid search method

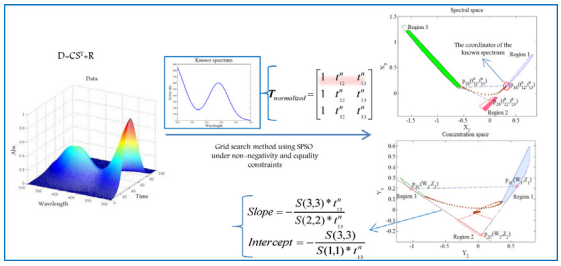
Samira Beyramysoltan[a,b], Róbert Rajkó[b], Hamid Abdollahi[a,*]

[a] Department of Chemistry, Institute for Advanced Studies in Basic Sciences, P.O. Box 45195-1159, Zanjan, Iran
[b] Institute of Process Engineering, Faculty of Engineering, University of Szeged, H-6725 Szeged, Moszkvai krt. 5-7, Hungary

HIGHLIGHTS

- Studying the effect of the equality constraint on the rotational ambiguity.
- Extending grid search method using SPSO to find solutions are corresponding.
- The equality constraint effect on complementary space shown by linear algebra.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 12 January 2013
Received in revised form 28 May 2013
Accepted 26 June 2013
Available online 3 July 2013

Keywords:
Equality constraint
Self-modeling curve resolution
Multivariate curve resolution
Rotational ambiguity
Species-based Particle Swarm Optimization

ABSTRACT

The obtained results by soft modeling multivariate curve resolution methods often are not unique and are questionable because of rotational ambiguity. It means a range of feasible solutions equally fit experimental data and fulfill the constraints. Regarding to chemometric literature, a survey of useful constraints for the reduction of the rotational ambiguity is a big challenge for chemometrician. It is worth to study the effects of applying constraints on the reduction of rotational ambiguity, since it can help us to choose the useful constraints in order to impose in multivariate curve resolution methods for analyzing data sets. In this work, we have investigated the effect of equality constraint on decreasing of the rotational ambiguity. For calculation of all feasible solutions corresponding with known spectrum, a novel systematic grid search method based on Species-based Particle Swarm Optimization is proposed in a three-component system.

1. Introduction

In the recent years, the progress of chemometric methods in different branches of science was impressive. Amongst, the self-modeling curve resolution (SMCR) methods owing to development of multivariate data sets – which cannot easily be analyzed by analytical chemistry methods – have found a significant character in chemistry. The recent analytical chemistry researches considered the SMCR methods for analyzing of data sets such as spectroscopic multi-wavelength measurements [1], hyphenated chromatography [2,3], and bio-chemical experiments [4–6]. However, the analysis of one chemical system using these methods is not typically well-determined due to rotational ambiguity. The rotational ambiguity is the property of bilinear data matrix and so inseparable part of soft-modeling methods. Based on the rotational ambiguity, a range of solutions can equivalently represent the measured data while there is one true chemical solution. Hence, the systematic errors can appear in quantitative measurements when soft methods are used to analyze data sets [7]. In order to decrease the feasible solution range, the useful knowledge of

* Corresponding author. Tel.: +98 241 4153122; fax: +98 241 4153232.
E-mail address: abd@iasbs.ac.ir (H. Abdollahi).

0003-2670/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.aca.2013.06.043

the studied chemical system can be applied as constraints. Non-negativity is the natural property of most chemical systems and sometimes it is the only information available about the systems (the concentration and most of the spectral intensities are always non-negative). This constraint constructs the basic idea of defining analytical curve resolution methods, Lawton–Sylvestre and Borgen methods [8,9]. The unimodality [10,11], closure, local rank and selectivity [12,13] are other useful constraints. However, the profiles can uniquely be determined when the specific information exists about the system [14]. The knowledge of underlying model in the some parts of the chemical systems, especially in kinetic and equilibrium reactions, can be considered as another important constraint in the soft modeling methods [15–18]. The effects of this constraint, hard modeling constraint, on reduction of rotational ambiguity were systematically investigated by Golshan et al. [19]. Recently, Sawall et al. investigated the reduction of rotational ambiguity in curve resolution techniques by introducing the partial knowledge of the factors [20]. They indicated the information of spectra leads to linear restrictions on the concentration profiles of the complementary species and vice versa. Though, the incorrect knowledge of the system causes to gain wrong results, and caution is required on imposing the constraint. Several methods were reported to calculate all feasible solutions in two and three component systems. The Lawton–Sylvestre method [8] for the two-component systems can be seen as the first step to develop these methods. Then, this method was extended by Borgen et al. [9] for three-component systems, and Rajko et al. [21] clarified the Borgen method using computational geometry tools. Gemperline [22] and Tauler [23] introduced a method, MCR-BANDs method, to estimate the band boundaries of feasible solutions based on fast maximization and minimization of the relative signal contribution of each component. MCR-BANDs method can also be used in case of data sets with more than three components. The method indicates the extent of rotational ambiguity for each component in the system, but does not provide explicitly full space of feasible solutions [24]. Meantime, the systematic grid search method was introduced in two-component systems, and subsequently extended for three-component systems [25,26].

In case of the three-component systems, the mentioned grid search method can calculate the area of the feasible regions for each component, but is not able to detect which solutions of these feasible regions are corresponding with each other. Therefore, it is valuable to modify pervious methods to develop one new method with this ability. The first section of this paper introduces a novel grid search method, which is able to obtain the area of feasible solutions associated to each component and correspondence relations among solutions simultaneously.

The second section includes a systematic study of effects of the equality constraint in three-component system. The results expose the efficiency of the equality constraint on the reduction of the rotational ambiguity. So, the probability of finding the real solution will extremely increase. The equality results show that the feasible regions of other components in complementary space reduced to a line. These observations explained based on linear algebra in this paper. This finding is an important step in outlining an analytical solution for the known profile effect on the rotational ambiguity.

2. Theoretical basis

2.1. Introduction to MCR methods

The second-order bilinear data set generally can be decomposed by multivariate curve resolution methods [27] according to Eq. (1):

$$\mathbf{D} = \mathbf{C}\mathbf{A}^{\mathrm{T}} + \mathbf{R} \tag{1}$$

In this equation $\mathbf{D}$ ($I \times J$) is a measured data matrix corresponding to a bilinear system with I different samples and J different variables. The matrix $\mathbf{C}$ ($I \times N$) includes the contributions of the N components in each sample, $\mathbf{A}$ ($J \times N$) is the pure response matrix of the components, and $\mathbf{R}$ ($I \times J$) is the residual error matrix. A range of $\mathbf{C}$ and $\mathbf{A}$ can construct matrix $\mathbf{D}$ and equally fulfill Eq. (1). There is no indication to specify which of the feasible solutions is true, unless the useful information is provided to the system.

$$\mathbf{D} = \mathbf{U}\mathbf{S}\mathbf{V}^{\mathrm{T}} + \mathbf{R} = \mathbf{X}\mathbf{V}^{\mathrm{T}} = \mathbf{U}\mathbf{Y} \tag{2}$$

According to Eq. (2), the bilinear matrix $\mathbf{D}$ can also be decomposed to the left eigenvectors and right eigenvectors by Singular Value Decomposition (SVD) [28] or Principal Component Analysis (PCA) [28]. $\mathbf{S}$ is a diagonal matrix including singular values and the columns of $\mathbf{U}$ and $\mathbf{V}$ are respectively the orthonormal basis for the column and row spaces of $\mathbf{D}$. The matrices $\mathbf{X}$ and $\mathbf{Y}$ involve coordinates of the projected row and column vectors of the matrix $\mathbf{D}$ in the row and column spaces, respectively. The PCA and SVD decomposition are unique due to the applied restrictions – orthogonality and maximum variance – in determination of $\mathbf{U}$ and $\mathbf{V}$ matrices. However, $\mathbf{U}$ and $\mathbf{V}$ are not physically meaningful and are related to physically interpretable $\mathbf{C}$ and $\mathbf{A}$ by a proper transformation matrix $\mathbf{T}$ as stated in Eq. (3),

$$\mathbf{A} = \mathbf{V}\mathbf{T}^{\mathrm{T}} \quad \mathbf{C} = \mathbf{U}\mathbf{S}\mathbf{T}^{-1} \tag{3}$$

In this equation, any invertible matrix $\mathbf{T}$ such as $\mathbf{T}\mathbf{T}^{-1} = \mathbf{I}$ can transform loading ($\mathbf{V}$) and score ($\mathbf{US}$), respectively, to a solution $\mathbf{A}$ and $\mathbf{C}$.

Thus, the calculation of the feasible matrices $\mathbf{C}$ and $\mathbf{A}$ is equivalent to finding the matrices $\mathbf{T}$. It is clear if a system does not have a unique feasible solution, there will be a set of matrices $\mathbf{T}$ satisfying Eq. (3). The results of MCR methods are complicated with two types of ambiguity, intensity and rotational ambiguity. The intensity ambiguity can be removed using normalization of profiles or matrix $\mathbf{T}$, but it cannot affect the rotational ambiguity. Based on the rotational ambiguity concept, all the feasible solutions have the same sum of the square of the residuals. So, the elements of matrix $\mathbf{T}$ can be treated as a set of unknown parameters that define correct $\mathbf{C}$ and $\mathbf{A}$ and result in minimum ssq.

2.2. Defining areas of feasible solutions by the grid search method

Amongst the methods for the calculation of feasible solutions, the systematic grid search methods for two [25] and three-component systems [26] are reformulated to find all feasible matrix $\mathbf{T}$ defining the correct $\mathbf{C}$ and $\mathbf{A}$. The heart of these methods is based on defining the ssq as a function of the matrix $\mathbf{T}$ elements. At first step, $\mathbf{C}$ and $\mathbf{A}$ are estimated by the normalized transformation matrix according to Eq. (3). Normalization of matrix $\mathbf{T}$ is useful as mentioned, because it generally removes the intensity ambiguity and also reduces the unknown elements of the matrix $\mathbf{T}$. In the next step all the physical constraints impose on the $\mathbf{C}$ and $\mathbf{A}$ profiles and result in matrices $\mathbf{C}'$ and $\mathbf{A}'$ according to Eq. (4). Finally, the sum of the square of the residual is computed by Eq. (5) as the difference between the reconstructed data (from $\mathbf{C}'$ and $\mathbf{A}'$) and the original data set ($\mathbf{D}$). This value will be the minimum for the accurate matrices $\mathbf{T}$.

$$\begin{aligned} &\mathbf{A}(\text{apply constraints}) \rightarrow \mathbf{A}' \\ &\mathbf{C}(\text{apply constraints}) \rightarrow \mathbf{C}' \end{aligned} \tag{4}$$

$$\mathbf{R} = \mathbf{D} - \mathbf{C}'\mathbf{A}' \quad \text{ssq} = \sum\sum \mathbf{R}_{i,j}^2 \tag{5}$$

In case of three-component systems, matrices $\mathbf{T}$ are of dimensions 3×3. Accordingly, the ssq is a function of nine elements of the